2009 International ITG Workshop on Smart Antennas

Berlin, February 16-18

Fraunhofer Institute for Telecommunications
Heinrich-Hertz-Institut, Berlin, Germany

Technical support:
Fraunhofer Institute for Telecommunications
Heinrich-Hertz-Institut
Einsteinufer 37, 10587 Berlin
wsa2009@mk.tu-berlin.de
http://www.hhi.fraunhofer.de

ISBN: 3-923613-41

Preface

It is our pleasure to welcome you in Berlin to the 2009 International ITG Workshop on Smart Antennas. The workshop covers new developments in the design, implementation and application of smart antennas in wireless communication networks. It continues to offer a forum for discussion and presentation of recent research on smart antennas. As in the previous years, the objective of the workshop is to continue, accelerate, and broaden the momentum already gained with a series of ITG Workshops held in Munich and Zurich 1996, in Vienna and Kaiserslautern 1997, in Karlsruhe 1998, in Stuttgart 1999, in Ilmenau 2001, in Munich 2004, in Duisburg 2005, in Ulm 2006, in Vienna 2007 and in Darmstadt 2008.

From a large number of submissions, the program committee has selected 33 papers for oral presentation and 18 posters. The conference is organized in 10 sessions, two of them are poster sessions. In addition, there are two industry application talks and two keynote lectures given by distinguished experts.

The papers for both poster and oral presentations are published in the conference proceedings saved on USB memory sticks. In addition, the proceedings will be made available in the EURASIP Open Library.

The workshop is supported and organized by the Fraunhofer German-Sino Lab for Mobile Communications, the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut, and the Berlin University of Technology. We would like to acknowledge the assistance by the ITG and EURASIP, as well as thank all sponsors for their financial support.

The Fraunhofer Heinrich-Hertz-Institut is a perfect place for the workshop. In the nineteenth century, Heinrich R. Hertz, who obtained his PhD from the University of Berlin, demonstrated the existence of electromagnetic waves, which had been previously predicted by James Clerk Maxwell. Heinrich R. Hertz was the first to broadcast and receive radio waves. His experiments triggered the invention of the wireless telegraph and radio by Marconi and others. The unit of frequency of a radio wave – one cycle per second – is named the hertz, in honor of Heinrich R. Hertz.

We wish you a successful participation in the workshop.

Workshop chairs:

Volker Jungnickel
Martin Schubert
Slawomir Stanczak
Gerhard Wunder

Program Committee

Co-Chairs

Dr.rer.nat. Volker Jungnickel
Dr.-Ing. Martin Schubert
Dr.-Ing. Slawomir Stanczak
Dr.-Ing. Gerhard Wunder

Tutorial Chair

Prof. Dr-Ing. Dr.rer.nat. Holger Boche

Fraunhofer German-Sino Lab
for Mobile Communications
Berlin, Germany

Fraunhofer Institute
for Telecommunications
Heinrich-Hertz-Institut
Berlin, Germany

Heinrich-Hertz-Chair
for Mobile Communications
Technical University of
Berlin, Germany

Technical Committee Members

C. Anton-Haro, CTTC, Catalonia, Spain
G. Bauch, DoCoMo Euro-Labs, Germany
C. Anton-Haro, CTTC, Catalonia, Spain
G. Bauch, DoCoMo Euro-Labs, Germany
E. Biglieri, Universitat Pompeu Fabra, Barcelona, Spain
H. Boche, TU Berlin, Germany
H. Boelcskei, ETH Zurich, Switzerland
E. Bonek, Vienna University of Technology, Austria
A. Czylwik, University of Duisburg-Essen, Germany
D. Dahlhaus, University of Kassel, Germany
A. Dekorsy, Qualcomm, Munich, Germany
G. Fettweis, TU Dresden, Germany
R. Fischer, University Erlangen-Nuremberg, Germany
B. Fleury, Aalborg University, Denmark
J. Fonollosa, UPC, Catalonia, Spain
A. Gershman, Darmstadt University of Technology, Germany
D. Gesbert, Eurecom Institute, France
M. Haardt, Ilmenau University of Technology, Germany
D. Heberling, RWTH Aachen University, Germany
F. Hlawatsch, Vienna University of Technology, Austria
A. Hottinen, Nokia Research Center, Finland
V. Jungnickel, FhG-MCI, Berlin, Germany
A. Klein, TU Darmstadt, Germany
V. Kuehn, University of Rostock, Germany
M. A. Lagunas, CTTC, Catalonia, Spain
G. Leus, Delft University of Technology, The Netherlands
J. Lindner, Uni Ulm, Germany
U. Martin, Hochschule Mannheim, Germany
T. Matsumoto, CWC-Oulu, Finland
T. Matsumoto, CWC-Oulu, Finland
C. Mecklenbräuker, Vienna University of Technology, Austria
M. Meurer, University of Kaiserslautern, Germany
R. Mueller, NTNU, Trondheim, Norway
J. Nossek, TU Munich, Germany
B. Ottersten, Royal Institute of Technology, Sweden
S. Paul, University Bremen, Germany
M. Rupp, Vienna University of Technology, Austria
A. Scholtz, Vienna University of Technology, Austria
M. Schubert, FhG-MCI, Berlin, Germany
E. Schulz, Nokia Siemens Networks, Germany
K. Solbach, University of Duisburg-Essen, Germany
S. Stanczak, FhG-MCI, Berlin, Germany
G. Taricco, Politecnico di Torino, Italy
W. Teich, Ulm University, Germany
J. Thompson, University of Edinburgh, United Kingdom
W. Utschick, TU Munich, Germany
E. Viterbo, University of Calabria, Italy
T. Weber, University of Rostock, Germany
J. Wehinger, Infineon Technologies AG, Germany
W. Wiesbeck, University of Karlsruhe, Germany
G. Wunder, FhG-MCI, Berlin, Germany
T. Zemen, ftw., Vienna, Austria
A. Zoubir, Darmstadt University of Technology, Germany

Workshop Organizers

Fraunhofer German-Sino Lab Mobile Communications

Fraunhofer Institut Nachrichtentechnik Heinrich-Hertz-Institut

Sponsors

Program – Monday, February 16

09:00 AM-16:30 – Tutorials

- **IST MASCOT Tutorial**, **2nd Multiuser MIMO Industry Course**
  http://www.ist-mascot.org

Program – Tuesday, February 17

08:30-09:10 – Industry Application Talk 1

- **Multiple Antenna Systems for LTE Advanced**

  Hans Peter Mayer (Alcatel-Lucent Stuttgart, DE)

  Going from LTE to LTE advanced the improvement of system performance will by far and large be realized by the massive use of multi-antenna technologies as well as the introduction of advanced interfernce avoidance techniques. The talk gives an overview on the approaches for MIMO and beamforming in single cell operation as well as on the different options for Cooperative Multipoint transmission.

09:10-10:50 – Channel Measurements and Antenna Design (S1)

Chair: Tobias Weber (University of Rostock, DE)

- **MIMO Efficiency of a LTE Terminal Considering Realistic Antenna Models** p. 14

  Alexander Geißler, Volker Wienstroer, Rainer Kronberger (Cologne University of Applied Sciences, DE), Christian Drewes (Infineon Technologies, DE), Frank Dietrich (Rohde & Schwarz GmbH & Co KG, Munich, DE)

- **Modeling of 3D Field Patterns of Downtilted Antennas and Their Impact on Cellular Systems** .... p. 19

  Lars Thiele, Thomas Wirth, Kai Börner, Michael Olbrich, Volker Jungnickel (Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut, Berlin, DE), Jürgen Rumold, Stefan Fritze (KATHREIN-Werke KG, DE)

- **Multi-User MIMO Channel Reference Data for Channel Modelling and System Evaluation from Measurements** .... p. 24

  Christian Schneider, Gerd Sommerkorn, Milan Narandzic, Martin Käske, Aihua Hong, Vadim Algeier, Wim Kotterman, Reiner Thomä (Ilmenau University of Technology, DE), Carsten Jandura (Technical University of Dresden, DE)

- **Arrays of Isotropic Radiators – A Field-theoretic Justification** .... p. 32

  Yordanov Hristomir, Michel Ivrlac, Peter Russer, Josef Nossek (Munich University of Technology, DE)

- **Differential Feedback in MIMO Communications: Performance with Delay and Real Channel Measurements** .... p. 36

  Daniel Sacristan Murga (Centre Tecnologic de Telecomunicacions de Catalunya (CTTC), ES), Florian Kaltenberger (Institut Eurecom, FR), Antonio Pascual Iserte (Universitat Politecnica de Catalunya, Barcelona, ES), Ana Perez-Neira (CTTC, ES)

Program – Tuesday, February 17

11:25-12:45 – Beamforming (S2)

Chair: Anja Klein (TU Darmstadt, DE)

13:45-14:25 – Keynote Speech

- **Conventional, less conventional, and optimal modulation**

  Rudolf Mathar (RWTH Aachen, DE)

14:30- 15:50 – Interference channel (S3)

Chair: Holger Boche (TU Berlin, DE)

- **Limited Transmitter Cooperation in Adjacent Broadcast Channels** ...... p. 82
  David Schmidt, Wolfgang Utschick (Munich University of Technology, DE)

- **Optimized Beamforming for the Two Stream MIMO Interference Channel at High SNR** ...... p. 88
  Christoph Hellings, David Schmidt, Wolfgang Utschick (Munich University of Technology, DE)

16:15- 17:15 – Networks (S4)

Chair: Peter Grant (University of Edinburgh, UK)

- **Analysis of Fusion and Combining for Wireless Source Detection** ...... p. 96
  Rajesh Sharma, Jon Wallace (Jacobs University Bremen, DE)

- **Device collaboration in ad-hoc MIMO networks** ...... p. 104
  Ari Hottinen (Nokia Research Center, FI), Tiina Heikkinen (University of Helsinki, FI), Emanuele Viterbo (University of Calabria, Rende (CS), IT)

- **Advanced Physical Layer Techniques for Wireless Mesh Networks with Network Coding** ...... p. 108
  Andreas Dotzler, Johannes Brehmer, Wolfgang Utschick (Munich University of Technology, DE)

10:00-18:00 – Poster Session (P1)

- **Performance Investigation on SCan-On-REceive and Adaptive Digital Beam-Forming for High-Resolution Wide-Swath Synthetic Aperture Radar** . p. 114
  Federica Bordoni, Marwan Younis, Eduardo Makhoul, Nicolas Gebert, Gerhard Krieger (German Aerospace Center (DLR), Wessling, DE)

- **Experimental study on the impact of the base station height on the channel parameters** ...... p. 122
  Aihua Hong, Reiner Thomä (Ilmenau University of Technology, DE)

- **Performance of Printed Dipoles on Dual Band High-Impedance Surface** p. 127
  Daniel Kornek, Sven Hampel, Isabell Kiral, Ilona Rolfes (Leibniz University of Hannover, DE)

- **Performance and Modeling of LTE H-ARQ** ...... p. 130
  Josep Colom Ikuno, Martin Wrulich, Markus Rupp (Vienna University of Technology, AT)

Program – Wednesday, February 18

08:30-09:10 – Industry Application Talk 2

- **Technical Challenges in Future Radiolocation Systems**

  Dirk Czepluch (Rohde&Schwarz Munich, DE)

  Technical Challenges in Future Radiolocation Systems The rapid development of radio communications has resulted in an increased in the importance of direction finding, while at the same time significantly boosting the associated requirements and the complexity of direction-finding systems. However, there is a spread between theoretical investigations and practical challenges. The antenna array design, which is necessary for direction finding systems, is a function of antenna pattern, number of antenna elements and the antenna array diameter. We will show theoretical antenna pattern and their practical implementations. Another topic is the synchronization of a multichannel receiver, to gain a sufficient phase- and amplitude balance. We will finish our presentation with the implementation aspects of high resolution direction finding methods such as MuSiC.

09:10-10:50 – Resource allocation (S5)

Chair: Eduard Jorswieck (TU Dresden, DE)

- **A low complexity space-frequency multiuser resource allocation algorithm** p. 162

  Ana Perez-Neira (CTTC, ES), Pol Henarejos (Universitat Politecnica de Catalunya, ES), Velio Tralli (University of Ferrara, IT), Miguel Angel Lagunas (Telecommunications Technological Center of Catalonia, Castelldefels, ES)

- **Resource Allocation in Multiantenna Systems under General Power Constraints** .... p. 170

  Angela Feistel (Technical University of Berlin, DE), Slawomir Stanczak, Michal Kaliszan (Fraunhofer German-Sino Lab for Mobile Comm., Berlin, DE)

- **On Proportional Fairness in Nonconvex Wireless Systems** .... p. 178

  Johannes Brehmer, Wolfgang Utschick (Munich University of Technology, DE)

- **DFT-based vs. Cooperative MET-based MU-MIMO in the Downlink of Cellular OFDM Systems** .... p. 182

  Lars Thiele, Malte Schellmann, Thomas Wirth, Volker Jungnickel (Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut, Berlin, DE), Federico Boccardi (Alcatel-Lucent, UK), Howard Huang (Alcatel-Lucent, US)

- **Channel Adaptive OFDM Systems with Packet Error Ratio Adaptation** p. 187

  Dagmar Bosanska, Christian Mehlführer, Markus Rupp (Vienna University of Technology, AT)

11:25-12:45 – Signal processing and estimation (S6)

Chair: Abdelhak M. Zoubir (TU Darmstadt, DE)

13:45-14:25 – Keynote Speech

- **Optimization-Boosted Beamforming: From Receive and Transmit Methods to Cooperative Relay Techniques**

  Alex Gershman (TU Darmstadt, DE)

  The aim of this talk is to present an overview of recent advances for receive, transmit, and distributed (relay) beamforming based on the theory of convex optimization. Our particular focus will be on robust beamforming techniques for data streaming applications. It will be demonstrated that convex optimization is a very efficient toolbox for advanced beamformer designs.

14:30- 15:50 – Coded modulation and capacity (S7)

Chair: Volker Kühn (University of Rostock, DE)

Mark Petermann, Carsten Bockelmann, Dirk Wübben, Karl-Dirk Kammeyer (University of Bremen, DE)

16:10- 17:30 – Multiuser MIMO (S8)

Chair: Karl-Dirk Kammeyer (University of Bremen, DE)

10:00-18:00 – Poster Session (P2)

MIMO Efficiency of a LTE Terminal Considering Realistic Antenna Models

*Alexander Geißler *, Volker Wienstroer *, Rainer Kronberger**

* High Frequency Laboratory,
Cologne University of Applied Sciences
Betzdorfer Str. 2, D-50679 Cologne, Germany
rainer.kronberger@fh-koeln.de

Frank Dietrich, Christian Drewes [+],

[+] Infineon Technologies AG
Am Campeon 1-12, 85579 Neubiberg, Germany
christian.drewes@infineon.com

ABSTRACT

This paper shows the influence of small coupled antenna arrays on the MIMO efficiency. Two different antenna models with varying antenna distances are investigated: dipoles and realistic terminal antennas. Two solutions to decouple the antennas are presented: 1) A parasitic element; 2) A hardware design of a matching network on a substrate which is the first implementation of a recent proposal in the literature. The resulting RF antenna properties are incorporated in the SCME channel model. The benefit in throughput is evaluated based on the information theoretic capacity and the throughput of a full-scale (coded) LTE link-level simulator.

1. INTRODUCTION

The performance boost of MIMO (Multiple Input Multiple Output) systems is based on uncorrelated signal paths. This is usually the case for non line of sight connections with a rich scattering environment between the transmitter and receiver (e.g. urban areas). Furthermore, the antennas have to receive/radiate uncorrelated signals. However, for small terminals like cell phones this scenario is not realistic. Due to shortage of space in a market-ready 3GPP LTE (Long Term Evolution) user equipment we are going to analyse the influence of the antenna coupling for varying antenna distances on the MIMO capacity. Most of the studies about MIMO systems are restricted to the channel model, antennas, or other isolated components. The extended analysis presented in this paper includes antennas, the RF frontend, the radio channel, and the baseband components (see Fig. 7). We show how parasitic elements and matching and decoupling networks affect the throughput and channel capacity in 2 x 2 MIMO systems for the downlink case.

2. THEORY OF ANTENNA MATCHING/DECOUPLING METHODS

Mutual coupling, caused by the size restriction between MIMO terminal antennas, leads to impedance mismatch and power absorption by the coupled antennas and therefore degrades the overall system efficiency. Antenna coupling modifies the current distribution along the antennas. Thus the radiation characteristic of the antenna system is influenced and the far field pattern changed [1]. Due to the impact on the signal correlation, the mutual coupling can have a negative as well as a positive influence on the channel capacity [5, 6, 7].

The challenge of our investigation was to minimize the signal correlation ρ_e between two antennas. As the correlation ρ_e of two identical antennas can be calculated with

$$\rho_e = \frac{|2Re\{S_{11}S_{12}^*\}|^2}{(1-|S_{11}|^2-|S_{12}|^2)^2} \quad ,$$

the correlation can be minimized by decreasing S_{11} (antenna reflection factor) and S_{12} (antenna coupling), respectively [4]. For this goal two different methods for matching/decoupling have been investigated. In Fig. 2 and 5 these methods are shown with practical examples. In the first method, a parasitic antenna element is used which is placed between the antennas. The other method requires an optimal matching/decoupling network.

Parasitic Element

A parasitic element can be a simple metal wire or a complex structure with an arbitrary termination. In our investigations, a wire that works like a reflector of a yagi-uda antenna has been chosen. The length of this element mainly affects the coupling between the antennas.

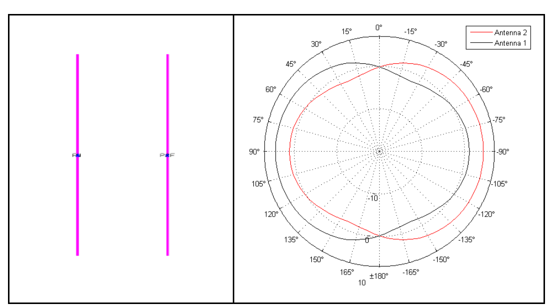


Fig. 1. Left: Two dipole antennas for 2 GHz, antenna separation 0.3 λ. Right: Far field diagram of the dipoles (Horizontal-Cut, dBi).

see the program

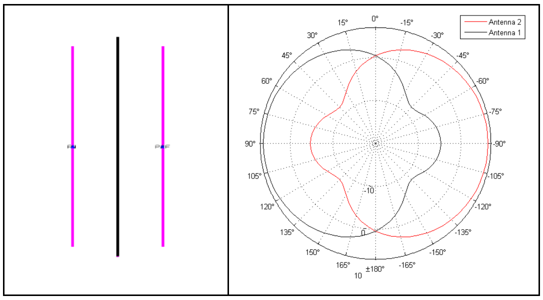


Fig. 2. Left: Two dipole antennas for 2 GHz with decoupling element (black stick), antenna separation 0.3 λ. Right: Far field diagram of the dipoles (Horizontal-Cut, dBi).

Figure 1 and 2 show the antenna simulation results for two dipoles with and without parasitic elements, respectively, with their associated azimuth far field pattern. Fig. 3 shows the associated S-parameters with and without parasitic element. With this simple method at antenna distance of 0.3 λ the coupling can be reduced from -10 dB to -26 dB and the correlation from 0.0656 to 0.0126. Additionally, the antenna matching is slightly improved. The results of the far field pattern in Figures 1 and 2 (right) and the computation result in Table 1 for the correlation ρe (Dipoles: 0.3 λ) shows the decorrelating effect of this decoupling method.

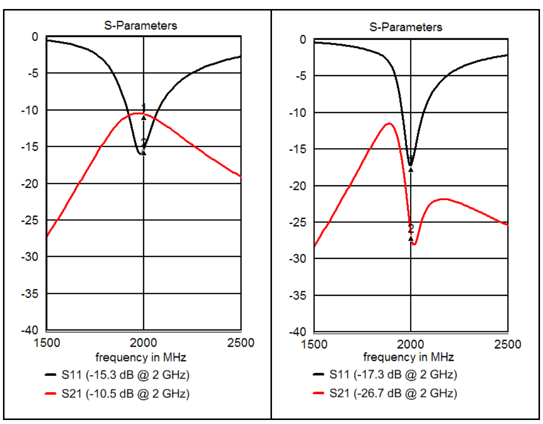


Fig. 3. S-parameters (in dB) of two dipoles with (right) and without (left) decoupling, antenna separation 0.3 λ.

Matching Network

The second method of reducing the antenna coupling is to insert a matching/decoupling network. In [4] it has been shown that such a network decreases the signal correlation between two identical antennas, but no hardware design has been reported before in the literature. Fig. 4 presents two short wideband monopoles on a substrate (MonSub). Based on this antenna configuration a matching/decoupling network has been designed. Fig. 5 shows the short wideband monopole antenna combined with this network. The network, consisting of microstrip lines, has been developed with Agilent's Advanced Design System© (ADS). Afterwards, it has been simulated in combination with the antennas in IMST's Empire XCcel©.

The S-Parameters in Fig. 4, 5 (right) show that at the antenna configuration with network both matching and coupling has been improved. The functionality of the network is described in [4].

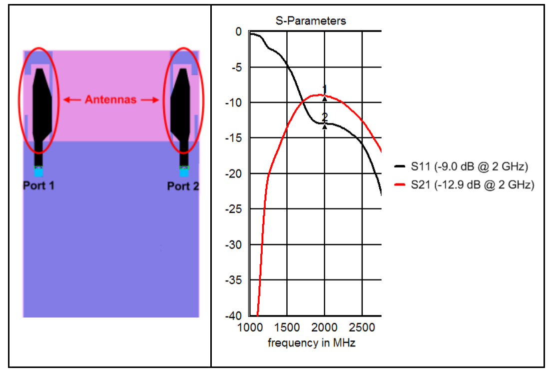


Fig. 4. Two element MIMO broadband antennas and their S-Parameters, antenna separation 0.3 λ. Bottom layer (ground): blue, top layer (antenna): black, substrate: pink.

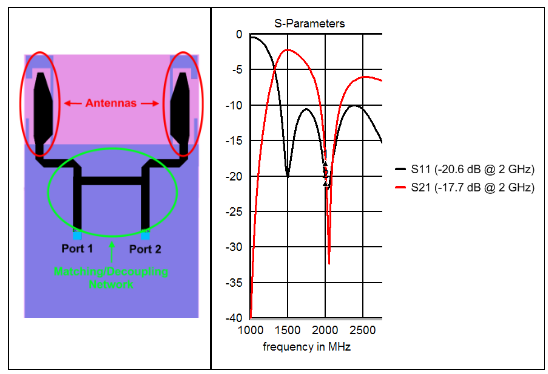


Fig. 5. Two element MIMO broadband antennas with decoupling network and their S-Parameters, antenna separation 0.3 λ. Bottom layer (ground): blue, top layer (antenna/network): black, substrate: pink.

The antenna efficiency and correlation values for all selected antenna distances in Table 1 show that the correlation can be decreased by optimal decoupling methods. Additionally the efficiency of the antennas and hence the antenna gain has been increased. This leads to a further capacity improvement. A theoretical analysis of the benefits of matching networks is shown in [6] and [7].

	Efficiency		**Correlation ρ_e**	
Antennas	**Z0-Load**	**Parasitic Elem./Match Network**	**Z0-Load**	**Parasitic Elem./Match Network**
Dipoles: 0.1 λ	56,9%	96,2%	0.3712	0.0091
Dipoles: 0.2 λ	80,7%	83,2%	0.0325	0.0258
Dipoles: 0.3 λ	88,1%	98,1%	0.0656	0.0126
Dipoles: 0.5 λ	90,9%	95,9%	0.0253	0.0242
MonSub: 0.1 λ	54,5%	98,2%	0,5709	0,1597
MonSub: 0.3 λ	82%	97,5%	0.0341	0.0217

Table 1. Efficiency and correlation for different antenna models, distances and decoupling methods

see the program

Both methods described above minimize losses in and between the antennas and change the radiation characteristics of the antennas which explain the decorrelation. Unfortunately, such methods in Figure 2 and 5 are only suitable for a small frequency band and are therefore only applicable for single band antennas. For practical use, adaptive networks are necessary.

3. SIMULATION MODELS

In this section the used radio channel, antenna, and LTE simulation models are described.

Antenna Simulation Models

Two different antenna models were designed: Dipoles and printed wideband monopoles. Dipoles are used because of the simple structure and their well known antenna characteristic. The wideband monopoles have been developed especially for terminal devices with the size of a typical PDA. All electromagnetic simulations have been performed in a free space environment with the software Empire XCcel©. The simulations include matching and material losses; therefore they are comparable with real antennas.

All antennas were simulated with distances between 0.01 λ and 0.6 λ (wideband monopoles from 0.06 λ to 0.6 λ because of their dimensions) in steps of 0.01 λ. For each antenna distance the obtained horizontal antenna pattern (H-Cut) with accuracy of [1:1:360] degree steps were transferred to the channel simulator.

If no matching network is used, all simulated antennas are terminated with the characteristic impedance $Z_0 = 50\ \Omega$.

Channel Simulation Model

The extended spatial channel model (SCME) [3], which is related to the 3GPP TR 25.996 [2], is used to simulate the behavior of the channel capacity regarding the channel environment and antenna behavior.

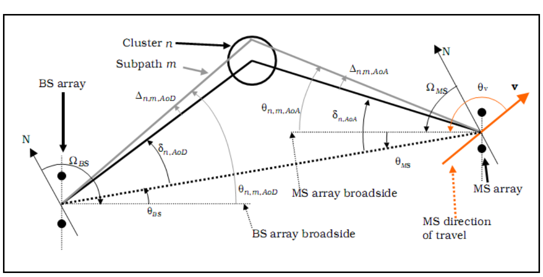


Fig. 6. BS and MS angular parameters in SCM specification [2]

The capacity simulations have been performed for a bandwidth of 20 MHz and for a center frequency of 2 GHz. The Urban Microcell scenario with one MS (Mobile Station) and one BS (Base Station) for the non line of sight case has been selected. We have varied the antenna distance at the MS array (see Fig. 6 MS array) as described above and have included the corresponding horizontal antenna far field pattern from the antenna simulations into the channel simulator. The current version of the SCME doesn't support the elevation angle and therefore only the azimuth angle is considered. The antennas at the BS array are separated at a distance of 0.5 λ and are assumed to be omni-directional radiators. For further SCME simulation parameters see [3].

Channel Scenario	Urban Micro
Number of paths (*N*)	6
Number of sub-paths (*M*) per-path	20
Mean AS at BS	NLOS: 19^0
Per-path AS at BS (Fixed)	5 deg (LOS and NLOS)
BS per-path AoD Distribution standard distribution	U(-40deg, 40deg)
Mean AS at MS	$E(\sigma_{AS,\ MS})=68^0$
Per-path AS at MS (fixed)	35^0
MS Per-path AoA Distribution	$\eta(0, \sigma^2_{AoA}(\mathrm{Pr}))$
Mean total RMS Delay Spread	$E(\sigma_{DS})=0.251\mu s$
Distribution for path delays	*U*(0, 1.2μs)

Table 2. Channel parameter for the Urban Microcell scenario [2]

In Figure 6 the angular parameters of the SCM specification are shown and Table 2 contains the main channel parameters. U(a,b) is a uniform distribution over (a,b) and $\eta(0,\sigma^2)$ a zero-mean Gaussian distribution with variance σ^2.

LTE-Simulator

To obtain simulation results like the throughput of the more complex LTE system, in the next step the correlation matrices and power delay profiles of the Matlab® channel simulations containing the channel and antennas properties are included into the LTE link level simulator (Fig. 7) based on Synopsys SystemStudio.

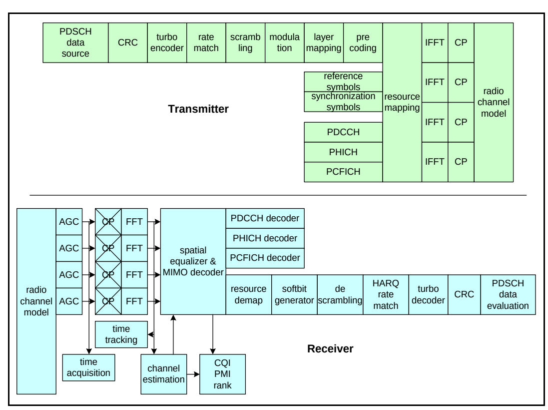


Fig. 7. Components of the MIMO LTE simulator

That simulator contains a complete LTE baseband model including the control channels (PDCCH, PHICH, PCFICH), reference and synchronization symbols, and the main data bearing channel, the physical downlink shared channel (PDSCH). The simulator includes functionality for the outer modem (hybrid ARQ (HARQ), cyclic redundancy check

see the program

(CRC), turbo coding, rate matching, bit scrambling), for MIMO processing (layer mapping, precoding) for up to four antennas, and for the OFDMA air interface (FFT, cyclic prefix (CP)). On the receiver side, modules for synchronization, channel estimation, and equalization are included for the inner receiver, as well as the corresponding outer receivers for the single physical channels. For closed loop MIMO operation channel quality indicators (CQI), precoding matrix indicators (PMI) and the channel rank is calculated.

4. SIMULATIONS

The channel matrix, which has been computed in the Matlab® SCME, can be used to obtain the channel capacity over a bandwidth B or to compute the corresponding correlation matrix and power delay profile for the LTE simulator.

Matlab SCME Simulations

The mean channel capacity is computed with

$$C = \frac{1}{B}\int_{B} log_2\left\{det\left(I + \frac{\rho}{N}H(f)H^H(f)\right)\right\}df,$$

where ρ is the SNR, N is the number of antennas on the MS side and $(.)^H$ denotes conjugate transpose. For every antenna distance the mean channel capacity is obtained with 5000 channel realizations. A SNR of 20 dB on each receive antenna and equal power distributions for the antennas on the BS side is assumed. Two dipoles without mutual coupling, i.e. they radiate and receive like ideal dipoles (2,14 dBi) and do not show any mutual influence on each other, are used as reference case (Ideal Dipoles). Fig. 8 shows the channel capacity over the antenna distance for different antenna configurations and furthermore the influence of the antenna matching/decoupling methods for selected antenna distances.

The mean channel capacity behavior of real dipoles with mutual coupling shows the red dashed curve (Dipoles, Z_0-Load). In comparison with the ideal dipoles for antenna separations smaller 0.18 λ mutual coupling increases the capacity. The same effect was also observed in [5], [6], and [7]. To improve the channel capacity with the dipoles for selected antenna distances 0.1 λ, 0.2 λ, 0.3 λ and 0.5 λ a parasitic element with a certain length has been placed between the antennas. By means of such a simple decoupling method the channel capacity could be increased and is much better than for ideal dipoles (see Fig. 8 circles at 0.1 λ, 0.2 λ, 0.3 λ and 0.5 λ).

For the case of the realistic terminal antennas short wideband monopoles on a substrate (MonSub) were simulated. Similar to the dipole case the mean channel capacity behavior of the MonSub with mutual coupling has been simulated (MonSub, Z_0-Load). The capacity decrease at the antenna distance 0.4 λ - 0.6 λ is due to the design of the MonSub; these have been optimized for small antenna distances. To investigate the influence of correlation and coupling at distances of 0,1 λ and 0.3 λ a matching/decoupling network has been designed. Similar to the dipoles, the capacity could be increased again (see Fig. 8 diamonds at 0.1 λ and 0.3 λ). It can be seen that at antenna distances larger than 0.3 λ decoupling has only small benefits. This is due to the already low mutual coupling and correlation at higher distances. Hence, the improvement in channel capacity with matching/decoupling networks is more effective at antenna distances ≤ 0.3 λ.

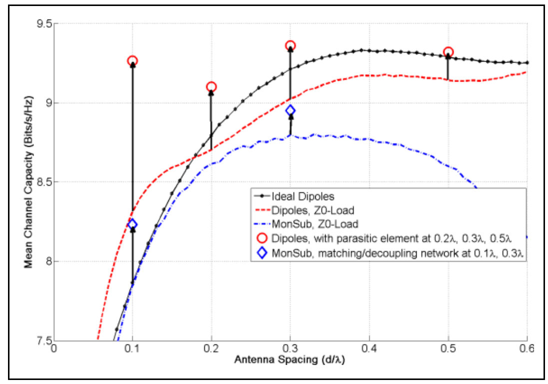


Fig. 8 2x2 MIMO-Channel capacity for antennas with matching networks and parasitic elements for selected antenna distances on the receiver side

In summery with the used matching/decoupling methods the channel capacity could be increased up to 5% for the wideband monopoles and approximately 2% - 10% for the dipoles. Moreover it could be seen that decoupled antennas have always a higher capacity than coupled antennas. Currently, with all capacity simulations only the azimuth angle has been included in the spatial channel model. For further detailed calculations, it is necessary to consider the full 3-dimensional antenna characteristics, which will be done in future work.

LTE Throughput Simulations

Due to the complexity of the link-level simulator it was not possible to incorporate the SCME model directly. Instead we consider one spatial scenario obtained from the SCME model and extract the power delay profile and the spatial correlations. In order to obtain a performance independent of the mobile's spatial orientation, 8 different azimuth orientations (in 45° steps) are considered. The throughput in Fig. 9 is the average of all 8 mobile positions based on the SCME model and the RF-model of the antenna coupling. Since the radio channel of the LTE simulator is also based on the 3GPP TR 25.996 SCM (Urban Micro) in the extended investigation the correlation matrices and the power delay profiles from the SCME Matlab® simulations were included in the LTE simulator. The parameters have been computed with the short wideband monopoles on substrate and an antenna separation of 0.1 λ. We compare the throughput performance for the antennas with and without matching/decoupling network.

For six SNR points we considered one LTE parameter set with the main parameters for a transmission bandwidth of 20 MHz, a center frequency of 2 GHz, FFT size of 2048, modulation order 16-QAM, Code-Rate 0.589 and 8 decoding iterations for the turbo decoder.

Figure 9 directly shows the throughput curves of the two antenna configurations and Figure 10 displays the computed percentage increase of the throughput with the short wideband monopoles against the configuration with the matching/decoupling network.

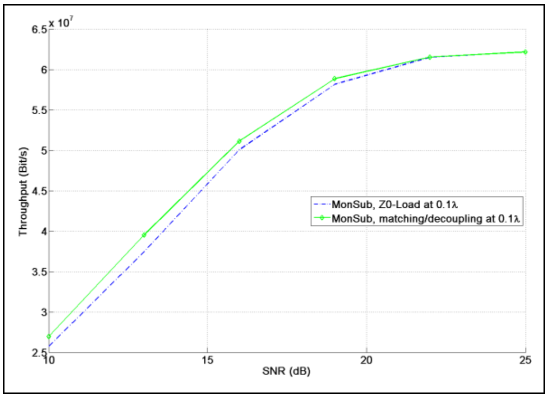


Fig. 9 LTE throughput for a 2x2 MIMO system

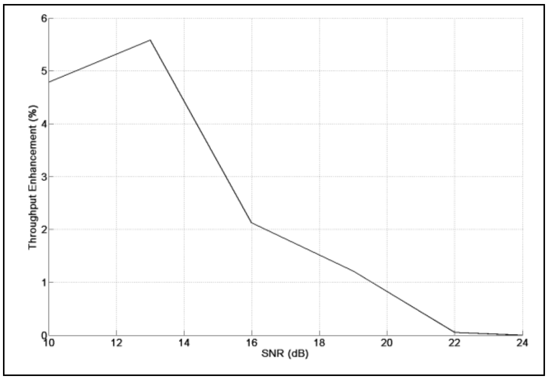


Fig. 10 LTE throughput enhancement of the MonSub (matching/decoupling) to the MonSub (Z0-Load) in percent

We observe that the performance increase of 5.5 % for lower SNR points is approximately the same as in the above capacity simulations. However, this small performance improvement can already minimize interferences or reduce transmitting power in noisy or weak signal environments.

For higher SNR values both curves converge to the maximum throughput of the system (with the selected simulation parameters) and antenna coupling plays no role anymore.

5. CONCLUSION

All important antenna parameters, like radiation, coupling and RF-network influences, in combination with the spatial channel model were included in a full-scale LTE link level simulation. We employed a simple dipole configuration and a realistic model of a terminal antenna design for several antenna separations. Two realistic matching/decoupling solutions were designed and we have shown that a parasitic element and a hardware design of an optimal matching/decoupling can improve the channel capacity and the throughput of the overall LTE system.

6. ACKNOWLEDGMENT

The authors wish to gratefully acknowledge the support of their colleagues at the Cologne University of Applied Science and at Infineon Technologies. Parts of this work are carried out within the scope of the EUREKA MEDEA+ project MIMOWA, which is partly funded by the German Federal Ministry of Education and Research.

7. REFERENCES

[1] R.W. Schlub, Practical Realization of Switched and Adaptive Parasitic Monopole Radiating Structures, PH.D. Dissertation, Griffith University, 2004.

[2] 3GPP TR 25.996, Spatial channel model for multiple input multiple output (MIMO) simulations, v8.0.0, December 2008.

[3] D. S. Baum, J. Salo, M. Milojevic, P. Kyösti, and J. Hansen, MATLAB implementation of the interim channel model for beyond-3G systems (SCME), May 2005. [Online]. Available: http://www.tkk.fi/Units/Radio/scm/

[4] S. Dossche, S. Blanch, J. Romeu, Optimum antenna matching to minimise signal correlation on a two-port antenna diversity system, Electronics Letters, Vol. 40, no. 19, pp. 1164-1165, September 2004.

[5] T. Svantesson, Mutual Coupling Effects on the Capacity of Multi Element Antenna Systems, Proc. IEEE ICASSP 01, Vol. 4, pp. 2485-2488, Salt Lake City, Utah, USA, May 2001.

[6] B.K. Lau, S. Ow, G. Kristensson, and A. F. Molisch, Capacity Analysis for Compact MIMO Systems, IEEE Vehicular Technology Conference, Vol.1, pp. 165-170, May 2005.

[7] J.W. Wallace and M.A. Jensen, Mutual coupling in MIMO wireless systems: A rigorous network theory analysis, IEEE Trans. Wireless Communication, Vol. 3, pp. 1317-1325, July 2004.

see the program

Modeling of 3D Field Patterns of Downtilted Antennas and Their Impact on Cellular Systems

L. Thiele, T. Wirth, K. Börner, M. Olbrich and V. Jungnickel
Fraunhofer Institute for Telecommunications
Heinrich-Hertz-Institut
Einsteinufer 37, 10587 Berlin, Germany
{thiele, thomas.wirth, jungnickel}@hhi.fraunhofer.de

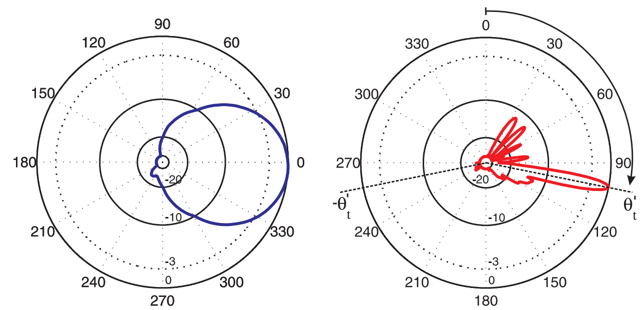
J. Rumold and S. Fritze
KATHREIN-Werke KG,
Anton-Kathrein-Strasse 1-3, 83004 Rosenheim, Germany
{juergen.rumold, stefan.fritze}@kathrein.de

***Abstract*—Advanced multi-antenna techniques, such as multi-user MIMO (MU-MIMO) and cooperative transmission are known to increase system performance in cellular deployments. However, it is well known that cellular systems suffer from multi-cell interference. Antenna downtilt is a common method used to adjust interference conditions especially in urban scenarios with a high base station density. Performance evaluation is generally based on multi-cell simulations using 2D models, neglecting the elevation component of the base station antennas. In this work we concentrate on 3D antenna models based on real world antennas with high directivity, their approximation and impact on cellular systems.**

I. Introduction

Recent advances valid for an isolated cell indicate huge performance gains obtained from multiple-input multiple-output (MIMO) communications [1], [2]. However, cellular systems still suffer from multi-cell interference. In order to develop advanced multi-antenna techniques, such as multi-user MIMO (MU-MIMO) [3] or cooperative transmission [4], we have to ensure a realistic modeling of multi-cell interference. Thus, we are able to investigate their performance more realistically. Performance evaluation is commonly based on multi-cell simulations using 2D models as e.g. 3GPP's extended spatial channel model (SCME) or WINNER Phase I model (WIM1). The WINNER Phase II model (WIM2), which was released recently [5], is capable of using 3D antenna geometries and field patterns.

In this work we concentrate on 3D antenna models based on real world antennas from KATHREIN, their approximation and impact on cellular systems. The goal is to provide appropriate antenna approximations, which can easily be included in channel models like the SCME used for performance evaluation of cellular MIMO communications. In general, 2D field patterns for the azimuth (ϕ) and elevation (θ) dimensions are available for various antenna types. Fig. 1 depicts these field patterns for the KATHREIN 80010541 antenna, which is one of the standard antennas used for future 3G Long Term Evolution (3G-LTE) sectorized cellular urban deployments. This antenna has an azimuth pattern with full width at half maximum (FWHM) of $\phi_{\text{FWHM}} = 60°$ and an elevation pattern with $\theta_{\text{FWHM}} = 6.1°$. The electrical downtilt angle $\alpha = \theta'_t - 90°$ is adjustable. For 3D approximation, we will use these 2D



(a) Azimuth; $\phi_{\text{FWHM}} = 60°$. (b) Elevation; $\theta_{\text{FWHM}} = 6.1°, \alpha = 7°$

Fig. 1. Radiation patterns from KATHREIN 80010541 antenna and 2.6 GHz carrier frequency

radiation patterns. In the following we compare two simple approximation techniques, their approximation errors and impact on cellular systems with respect to user geometries obtained from system level simulations.

II. Approximation of 3D Antenna Radiation Patterns

Common 3D antenna approximation approaches based on 2D field patterns are known from literature [6]–[8]. In this work we focus on the two most promising approaches, suitable for the approximation of highly directive antennas: a conventional method and a novel technique described in [8]. Fig. 2(a) depicts a 3D measured antenna diagram from KATHREIN 80010541 antenna in the phi-theta plane with a downtilt angle $\alpha = 10°$.

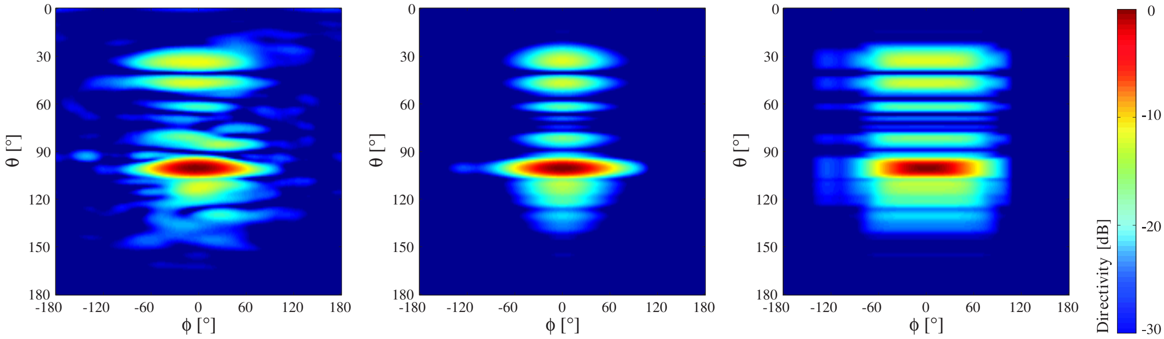
A. Conventional method

A simple way to create a quasi 3D pattern is to combine azimuth and elevation field patterns by adding their gains in both directions with equal weights.

$$G(\phi, \theta) = G_H(\phi) + G_V(\theta) \quad (1)$$

Using this method we obtain a 3D pattern which is symmetrical in ϕ direction. This method lacks in appropriate modeling at the back side of the antenna, i.e. between $-90° \leq \phi \leq 90°$. These directional gains at the backside of the antenna would be very small unlike those in a real antenna. Fig. 2(b) depicts the

see the program



(a) measured pattern (b) conventional method (c) novel technique

Fig. 2. Antenna patterns shown in the phi-theta plane. The attenuations are limited to 30 dB. The tilt angle is fixed to $\alpha = 10°$

resulting approximation in phi-theta-plane, where the overall dynamics are limited to 30 dB. It may be observed that the main lobe's shape is pretty close to the one obtained from the measured antenna radiation pattern. The components for $-60° \leq \phi \leq 60°$, which lie outside the 120°-sector, are suppressed for all side lobes in vertical dimension. However, their gain is 20 dB below the maximum. Thus, we expect these components to be less important for antennas with high directivity.

B. Novel technique

Another approximation method we consider in this work was proposed by [8], which we will refer to as *novel technique*. The difference to the conventional method is that the elevation and azimuth gains are both weighted to extrapolate the spatial gain

$$G(\phi,\theta) = \underbrace{\frac{\omega_1}{\sqrt[k]{\omega_1^k + \omega_2^k}}}_{A_1} G_H(\phi) + \underbrace{\frac{\omega_2}{\sqrt[k]{\omega_1^k + \omega_2^k}}}_{A_2} G_V(\theta) \quad (2)$$

$$\begin{aligned} \omega_1(\phi,\theta) &= vert(\theta) \cdot [1 - hor(\phi)] \\ \omega_2(\phi,\theta) &= hor(\phi) \cdot [1 - vert(\theta)] \end{aligned}$$

The weighting functions ω_1 and ω_2 are based on linear values $hor(\phi)$ and $vert(\theta)$ of the antenna gains in azimuth and elevation direction, respectively. k is a normalization-related parameter which referring to [8] provides best results for $k = 2$. Both A_1 and A_2 are smaller than one since $vert(\theta)$ and $hor(\phi)$ vary between zero and one.

The enumerators ω_n of A_n control the weight of elevation and azimuth pattern, respectively. For small gains in the azimuth direction, the antenna gains in the elevation pattern are reduced in their magnitude and vice versa. Thus, gain factors are more restricted in their range compared to (1). Further, if we choose values for θ where the elevation pattern has its maximum, i.e. $vert(\theta) = 1$, results in $\omega_2 = 0$. Hence, only the azimuth pattern is weighted according to $(1 - hor(\phi))$. In general, the lower the gain, the smaller the weight. Fig. 2(c) depicts the resulting approximation based on [8] in the phi-theta-plane. Comparing the resulting radiation pattern with the one in Fig. 2(b) shows a higher spread in power distribution among the phi-theta-plane. This property is closer to the measured pattern. However, the deformation in the main lobe seems to be more significant, which is not comparable with the measured pattern. Thus, we expect results obtained from the conventional approximation to be closer to reality. Due to that reason we mainly limit our investigations to the conventional method.

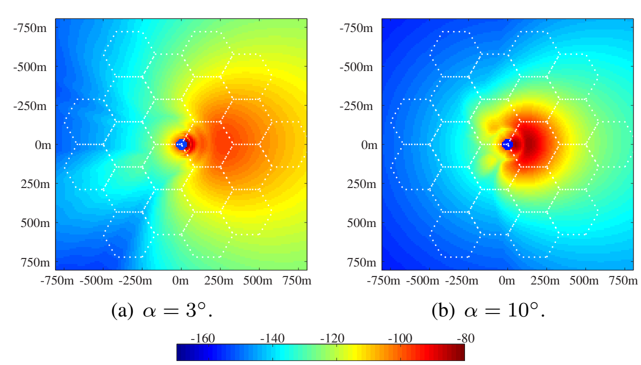
III. Effects in an Isolated Cell

As a starting point for our investigations we focus on the effects, which may be observed in an isolated cell. Therefore, we consider a channel with non line of sight (NLOS) propagation conditions in an urban-macro scenario. Thus, the path loss equation according to [9] is given by

$$\begin{aligned} \mathrm{PL_{dB}} &= 40(1 - 4 \cdot 10^{-3} \Delta h_{BS})\, log_{10}(d) \\ &\quad -18\, log_{10}(\Delta h_{BS}) + 21\, log_{10}(f_c) + 80, \end{aligned} \quad (3)$$

where d [km] is the distance between base station (BS) and mobile terminal (MT); f_c [MHz] is the carrier frequency and Δh_{BS} [m] is the BS height measured from the average rooftop level. Setting $f_c = 2.6$ GHz and $\Delta h_{BS} = 15$ m yields

$$\mathrm{PL_{dB}} = 130.5 + 37.6 \log_{10}(d) \quad (4)$$



(a) $\alpha = 3°$. (b) $\alpha = 10°$.

Fig. 3. Gains for the directive antenna in addition to an urban path loss for different downtilt angles

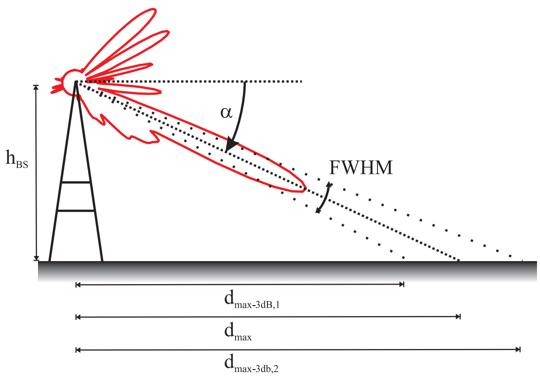
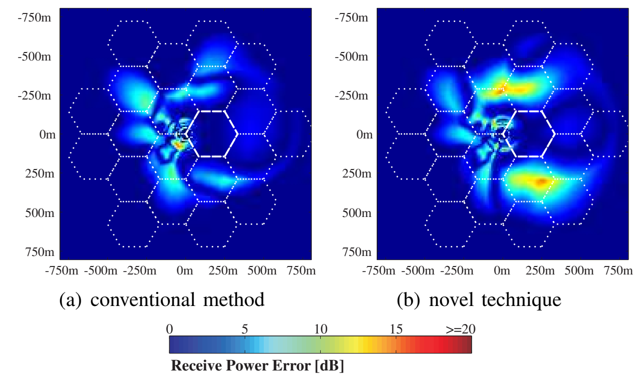
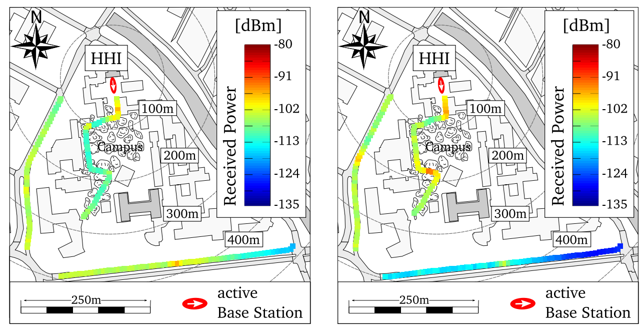


Fig. 4. Base station setup with downtilted antenna



(a) conventional method (b) novel technique

Fig. 5. Path loss plus directional gain errors in relation to the measured pattern plus path loss at 2.6 GHz



(a) 3° downtilt (b) 10° downtilt

Fig. 6. Measured path loss at $1W$ transmission power and $h_{BS} = 32$ m

In combination with the directive antenna gains obtained from a given 3D antenna diagram, we estimate the received power as a function of the distance between BS and position of the MT. Fig. 3 depicts the antenna gains for the directive antenna in addition to an urban path loss from (3), both seen at height of the terminal antenna and for different tilt angles. Hexagonal sectors are shown as white dotted lines with an inter-site distance (ISD) of 500 m. The BS antenna is located at the point [0,0] at $h_{BS} = 32$ m above ground level, i.e. the average rooftop height is assumed to be 17 m. It may be observed that in case of a downtilt angle $\alpha = 3°$ the sector antenna is serving up to 3 neighboring sectors with equivalent power as available in its own sector. In this application $\alpha = 10°$ seems to meet requirements of a cellular system with an ISD of 500 m best: high gain level in own sectors and low values for neighboring cells. This is also verified by (5) indicating the effective cell radius, i.e. the distance where main lobe and ground level intersect, refer to Fig. 4. The distance range covering the FWHM area can be determined to 129 m $\leq d \leq$ 244 m, assuming $\alpha = 10°$.

$$d_{max} = \frac{h_{BS}}{\tan\alpha} \qquad (5)$$

$$d_{max-3\text{ dB}} = \frac{h_{BS}}{\tan\left(\alpha \pm 0.5\,\theta_{\text{FWHM}}\right)} \qquad (6)$$

In the following, we focus on the quality of both approximation methods from (1) and (2). For comparison we use the distance dependent received power based on the measured radiation pattern combined with the urban path loss from (4). Fig. 5 shows the differential received power maps for the conventional and novel approximation for $\alpha = 10°$, respectively. The conventional method, depicted in Fig. 5(a), shows superior precision over the approximation based on [8], depicted in Fig. 5(b), both described in sections II-A and II-B.

To substantiate simulation results obtained for an isolated cell, we include outdoor measurement results from the city area of Berlin. These measurements were carried out in the campus area of Technical University Berlin (TUB) with an average rooftop height of approx. 30m. The BS antenna (KATHREIN 80010541) was fixed at $h_{BS} = 32$ m with a variable tilt angle, which was set to $\alpha = \{3°, 10°\}$. Figs. 6(a) and 6(b) show the received power for the given tilt angle. For $\alpha = 3°$ the BS antenna serves the whole area with approx. -100 dBm. Otherwise, for $\alpha = 10°$ the campus area and the outer region are both served with -90 dBm and -110 dBm, respectively. In particular, close to the BS there is gain of 10 dB. In general, we observe equivalent behavior as already found in Figs. 3(a) and 3(b): the smaller the tilt angle α the larger the area, which is served with equivalent, but lower received power. On the other hand with higher tilt angles, the BS focuses its transmit power to a smaller area.

IV. Effects in a cellular system

In the next section we turn our focus to the downtilted antenna and its effects observable in a cellular environment, i.e. effects on the neighboring cells. Therefore we consider a center cell surrounded by one tier of triple-sectorized cells. Each sector is served by single antenna with $h_{BS} = 32$ m, where all tilt angles are set to identical values. Thus, all BSs are assumed to cover a region of equivalent size. Fig. 7 indicates achievable signal to interference ratio (SIR) conditions in such a setup under the assumption of $\alpha = \{3°, 10°\}$. Again we employ the urban path loss model from (4) in combination with the radiation pattern obtained from the conventionally approximated antenna pattern (1). Since the evaluation scenario is limited to 7 cells, only the highlighted

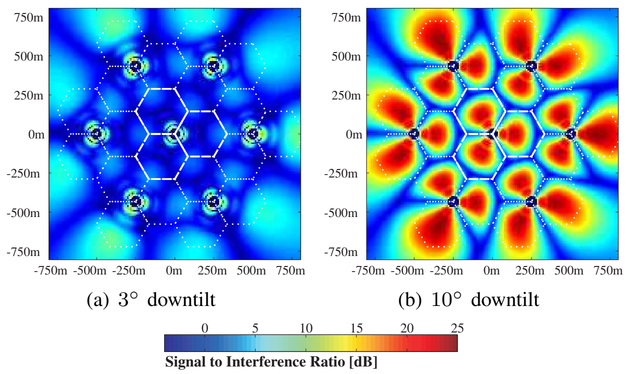


Fig. 7. SIR maps obtained from simplified multi-cell simulations

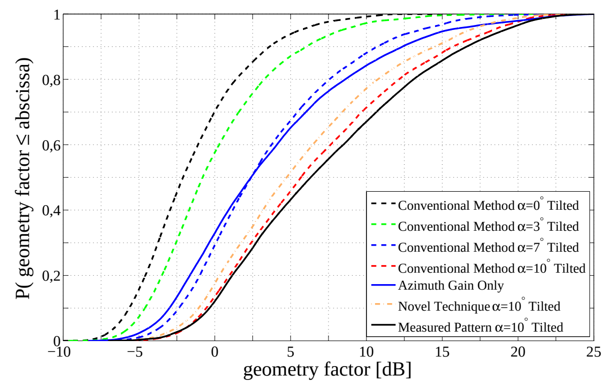


Fig. 8. User geometries obtained from multi-cell SCME simulations

TABLE I
SIMULATION ASSUMPTIONS

parameter	value
channel model	3GPP SCME
scenario	urban-macro
f_c	2.6 GHz
frequency reuse	1
signal bandwidth	18 MHz, 100 RBs
intersite distance	500m
transmit power	46 dBm
sectorization	triple, with FWHM of 68°
elevation pattern	with FWHM of 6.1°
BS height h_{BS}	32m
MT height	2m

center cell reflect reasonable SIR values. For small α, SIRs inside these 3 sectors are limited to an average value of 0 dB, refer to Fig. 7(a). In contrast, for $\alpha = 10°$ the achievable SIRs cover a range from 0 dB at the cell edge and up to 25 dB in the cell center, refer to Fig. 7(b). This result indicates that a tilt angle in the order of $\alpha = 10°$ is favorable for a generic cellular system with an ISD of 500 m and $h_{BS} = 32$ m. However, in real scenarios where BSs are not placed in a symmetric grid, α would be chosen according to the desired coverage area, i.e. cell size.

V. SYSTEM LEVEL SIMULATIONS USING SCME

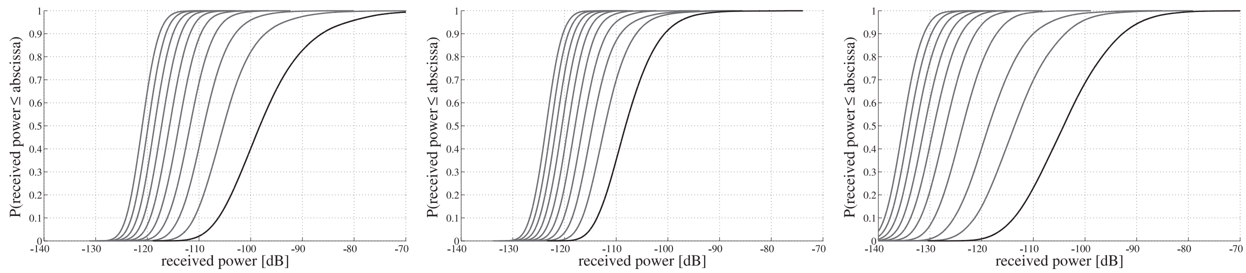
In the following we investigate the effects from modeling 3D antenna radiation patterns in a triple-sectored hexagonal cellular network with 19 BSs in total. This refers to the commonly used simulation assumption [10], i.e. a center cell surrounded by two tiers of interfering cells. The MTs are placed in the center cell and are always served by the BS whose signal is received with highest average power over the entire frequency band. In this way, BS signals transmitted from 1st and 2nd tier model the inter-cell interference. Simulation parameters are given in Table I. The SCME [11] with urban macro scenario parameters is used, yielding an equivalent user's geometry as reported in [12].

User geometries: Fig. 8 compares the resulting user geometries obtained from both approximation methods (1) and (2), while assuming $\alpha = 10°$. For validation we include the geometry factor distribution obtained from simulations using the measured radiation pattern from Fig. 2(a), where $\alpha = 10°$. The geometry from the conventional method is close to the cumulative distribution function (CDF) based on the measured pattern, while the CDF which is based on the novel approximation technique shows a significant gap. Hence, the choice of the approximation method influences simulation results considerably. Further, we show changes on the geometry factor due to the downtilt angle, which is selected from $\alpha = \{0°, 3°, 7°, 10°\}$ based on the conventional antenna approximation (dashed lines). Comparing these results with the user geometries obtained from simulations which consider 2D antenna modeling only, shows equivalent values for $\alpha = 7°$. For smaller downtilt angles we observe user geometries which are significantly below the well known values for the 2D case.

Top-N power distribution: Consider the application of a cellular radio system consisting of $\mathcal{K}$ BSs operating in the downlink direction. It is reasonable to assume that a MT located in a specific cell of that network is able to detect a subset of $N = |\mathcal{N}|$ strongest BS signals, i.e. a set of BSs $\mathcal{N} \subset \mathcal{K}$ of all BSs within the deployment. Based on the user-specific channels to all BSs, a so-called Top-N power distribution is generated by instantaneously sorting the estimated power distributions. These sorted received powers are put into one overall statistic, enabling us to observe the power distributions for all channels seen by a MT. At two given sample points, the strongest signals may be related to different sectors or sites and are included in the same CDF, referred to as top-1. The power distributions of the 1st to the 10th strongest channels are given in Fig. 9. We observe that power distributions are broadened due to large downtilt angles α. Intuitively spoken, cells become more separated, i.e. signal conditions with strong interference are mitigated directly at the physical layer (PHY layer).

Finally, we determine the source of the four strongest signals received at the MT. Fig. 10 depicts the histogram showing the probabilities for the source, i.e. center cell, 1st tier and 2nd tier, of the four strongest signals in the cellular scenario. These results are obtained from simulations using the azimuth

see the program



(a) azimuth pattern only (2D) (b) conventional method $\alpha = 3°$ (c) conventional method $\alpha = 10°$

Fig. 9. Power distributions for the 10 strongest signals

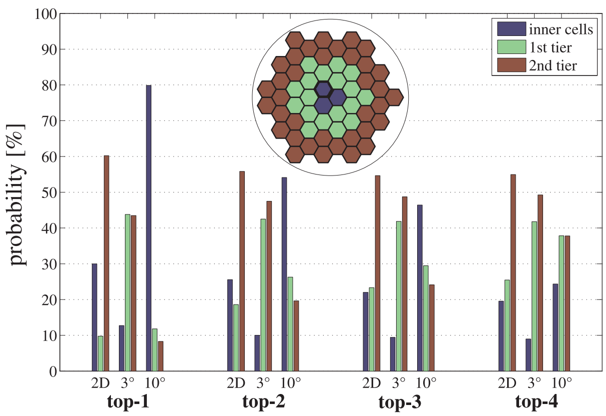


Fig. 10. Source of the four strongest signals, obtained from simulations considering azimuth pattern (2D) only, as well as 3D conventionally approximated antennas with $\alpha = \{3°, 10°\}$

pattern only, i.e the standard 2D assumption. Further, results are compared with probabilities using the conventional antenna approximation with downtilt angles $\alpha = \{3°, 10°\}$. Concentrating on the strongest signal, i.e. top-1, and comparing these probabilities, we can observe two main differences: for the 2D simulation the origin for top-1 signal lies with 60% in the 2nd tier and with 30% in the inner cell. With $\alpha = 10°$ the situation is changed. The top-1 signal has its origin with 80% in the inner cell and with less than 10% in the 2nd tier. For completeness note that for $\alpha = 10°$ signals are more likely to have their origin close to the terminal position.

VI. Conclusion

In this work we compared two simple 3D antenna approximation methods and focused on antenna types, typically used in 3G-LTE sectorized urban deployments. These methods use 2D radiation patterns for approximation, which are generally provided by antenna manufacturers. The simple conventional method for 3D antenna approximation provides results close to those obtained by using a 3D measured radiation pattern. Simulation and measurement results further showed significant SIR gains from downtilted BS antennas in cellular deployments. Finally, the results point out that full 3D antenna modeling is necessary to evaluate advanced multi-antenna techniques, like MU-MIMO and cooperative transmission. Especially, the evaluation of joint downlink transmission comparing dynamic versus fixed BS clustering will benefit from this work.

Acknowledgements

The authors are grateful for financial support from the German Ministry of Education and Research (BMBF) in the national collaborative project EASY-C under contract No. 01BU0631.

References

[1] G. Foschini and M. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, no. 3, pp. 311–335, 1998.
[2] L. Zheng and D. Tse, “Diversity and multiplexing: A fundamental tradeoff between in multiple antenna channels,” *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
[3] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, “Shifting the MIMO paradigm,” *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 36–46, Sept. 2007.
[4] F. Boccardi and H. Huang, “A near-optimum technique using linear precoding for the MIMO broadcast channel,” *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, pp. III–17–III–20, April 2007.
[5] L. Hentil, P. Kysti, M. Kske, M. Narandzic, and M. Alatossava, “MATLAB implementation of the WINNER Phase II Channel Model ver1.1,” Tech. Rep., Dec. 2007. [Online]. Available: https://www.ist-winner.org/phase_2_model.html
[6] W. Araujo Lopes, G. Glionna, and M. de Alencar, “Generation of 3d radiation patterns: a geometrical approach,” in *Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th*, vol. 2, 2002, pp. 741–744 vol.2.
[7] F. Gil, A. Claro, J. Ferreira, C. Pardelinha, and L. Correia, “A 3d interpolation method for base-station-antenna radiation patterns,” *Antennas and Propagation Magazine, IEEE*, vol. 43, no. 2, pp. 132–137, Apr 2001.
[8] T. Vasiliadis, A. Dimitriou, and G. Sergiadis, “A novel technique for the approximation of 3-d antenna radiation patterns,” *Antennas and Propagation, IEEE Transactions on*, vol. 53, no. 7, pp. 2212–2219, 2005.
[9] TR 101 112 v3.2.0, “Universal Mobile Telecommunications System (UMTS); Selection procedures for the choice of radio transmission technologies of the UMTS,” Apr. 1998.
[10] L. Thiele, M. Schellmann, T. Wirth, and V. Jungnickel, “Cooperative multi-user MIMO based on reduced feedback in downlink OFDM systems,” in *42nd Asilomar Conference on Signals, Systems and Computers*. Monterey, USA: IEEE, Nov. 2008.
[11] 3GPP TR 25.996 V7.0.0, “Spatial channel model for multiple input multiple output (MIMO) simulations (release 7),” July 2007. [Online]. Available: http://www.tkk.fi/Units/Radio/scm/
[12] H. Huang, S. Venkatesan, A. Kogiantis, and N. Sharma, “Increasing the peak data rate of 3G downlink packet data systems using multiple antennas,” vol. 1, april 2003, pp. 311–315 vol.1.

see the program

Multi-User MIMO Channel Reference Data for Channel Modelling and System Evaluation from Measurements

Christian Schneider, Gerd Sommerkorn,
Milan Narandžić, Martin Käske, Aihua Hong,
Vadim Algeier, W.A.Th. Kotterman, Reiner S. Thomä
Institute of Information Technology,
Ilmenau University of Technology, Germany
{christian.schneider}@tu-ilmenau.de

Carsten Jandura
Vodafone Chair Mobile Communications Systems
Technical University of Dresden, Germany
carsten.jandura@ifn.et.tu-dresden.de

Abstract—Novel channel models conceived in IST-WINNER, COST273 or standardisation bodies introduce *reference scenarios* to evaluate the link- and system-level performance of multi-antenna concepts. Although, parameter of those listed models are publicly available underlying measurement data are only rarely accessible. This paper presents an extensive multi-user and multi-base station MIMO measurement campaign in an urban macro cell scenario. The measurement setup allows high-resolution path parameter estimation through the use of dedicated measurement antennas. Furthermore, a framework for the extended use of channel sounding data is presented. Parts of the measurement data and estimated multipath parameters will be made publicly available.

***Index Terms*—channel sounding, measurement data, high resolution multipath parameter estimation, spatial channel modelling, reference scenario**

I. Introduction

In order to allow comparison of performance evaluations, by simulation, of different multi-antenna communication systems, the definition of reference scenarios is required. Channel models developed within COST [1][2], WINNER [4], or standardisation bodies[5] rely on scenario classifications like macro and micro cell, indoor and outdoor. Furthermore, based on system-deployment schemes, such scenarios are grouped into e.g. wide area, metropolitan area and local area concepts [4]. The parameter tables for these channel models for simulations on system or link level are based on the analysis of, often dedicated, radio channel sounding data. While the parameter tables are usually freely accessible, the measured channels are not. Recently, within the COST project various reference scenario sets have been presented and are made accessible via the web site [1]. However, what is lacking is a consistent view on the chain from channel sounding to processing for various applications, e.g. parameters for channel models or simulations including access to measurement data. Under this premise the presented work flow and measurement data in this paper can be understood as reference data for various application fields. In particular the used measurement configuration targets many system aspects that are of interest for the current standardization process, e.g. configurations as multiple base station and multiple users, relaying as well as frequency and bandwidth. Furthermore the presented data offers huge potential for scientific research, because of the considered system setup and quality of the data.

In this paper, we present an extensive multi-user and multi-base station MIMO measurement campaign in Section II. Furthermore in this section we discuss the outcome of such a campaign, including the data validation and notes how to share parts of the data. In Section III various fields of usage of realistic channel sounding data are discussed. First results based on the proposed data and their analysis is shown in Section IV. The paper ends with the summary in Section V.

II. Multi-User Multi-Base Station MIMO Channel Measurement Campaign

A. Scenario and Measurement System

Our measurement campaign focused on gathering realistic channel data in an urban macro cell environment in the 3GPP Long Term Evolution band (LTE [5]). In Table 1 and Table 2, the measurement and antenna setups are highlighted. Channel sounding is performed at 2.53 GHz in a band of 2x40 MHz.

Figure 1 Radio environment; BS2@25m, and reference points: 21 (upper right) and 13a (lower right).

see the program

Measurement Campaign		
Scenario		Urban macro cell
Location		City centre, Ilmenau, Germany
Measurement setup		3 Base stations, 1 relay station, 22 tracks, full MIMO
Channel Sounder		
Type		RUSK TUI-FAU Medav, GmbH
Transmit power @ PA	[dBm]	46
Centre frequency	[GHz]	2.53
Bandwidth	[MHz]	2x40 MHz
CIR length	[µs]	6.4
CIR sampling	[samples]	641
Snapshot rate	[Hz]	>75
MIMO sub-channels	[#]	944 (16×58 eff.)
AGC switching		in MIMO sub-channels
Positioning		Odometer and GPS

Table 1 Scenario and measurement setup

To allow high resolution path parameter estimations, dedicated antenna arrays at Tx and Rx side are used. On the Tx side (base station), an uniform linear array is used with 8 dual-polarised (H/V) elements, each of which consists of a stack of 4 patches in order to form a narrow transmit beam in elevation. At the mobile (passenger car), a circular array with 2 rings of 12 patches with H/V polarisations is used. Additionally, a MIMO cube is placed on top, see Figure 3. The mobile acts as Rx. In Figure 2, a map of the city centre of Ilmenau is shown including the position of the base stations, the relay station, and the routes. For each of the tracks and for each measured snapshot, geo-data information based on GPS, odometer and separated distance measurements via laser are available. The accuracy around the start and end points for each track is approximately 0.1 m and along the route 1m. During the measurement the position in height is only traced via GPS with low accuracy, therefore an interpolation is necessary and decreases the accuracy in the cases where the height along the track is changing. The information about the building/street structure and the landscape profile (height) has been acquired from the land registry office and 3D laser scans, respectively.

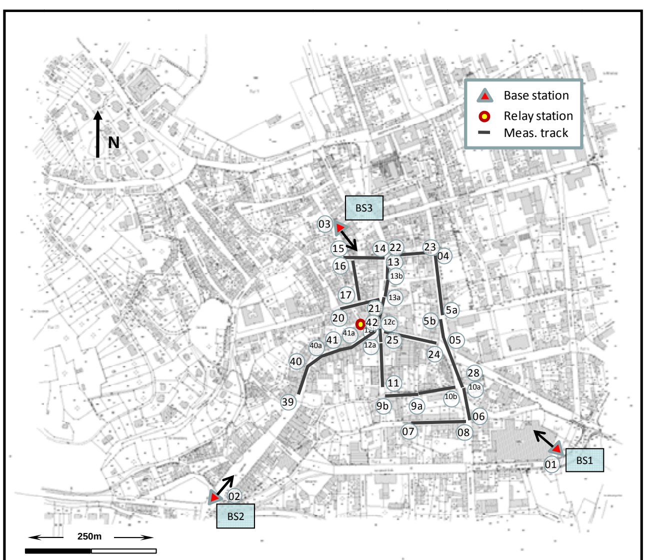


Figure 2 Sketch of the scenario, routes, and base stations

In total the measurement campaign covers 3 base station

		Tx Array	Rx Array
Name (type)		PULPA8	SPUCPA 2×12 + MIMO-Cube
Height	[m]	25, 15 and 3.5	1.9
Beamwidth (3dB) : azimuth	[°]	100	Omni
elevation	[°]	24	80
Tilt (down)	[deg.]	5	0
Mobility	[m/s]	0	3-5

Table 2 Antenna setup

positions with 25m and 15m height and additionally a relay point in the middle of the scenario. The intersite distance between the base stations is found to be for BS1-BS2 = 680m, BS2-BS3 = 580m and BS3-BS1=640m. More than 20 individual tracks with more than 120 measurement runs were performed.

B. *Campaign Outcome*

The result of the measurement campaign is composed of not only the *raw data* acquired by the channel sounder itself – rather additional information about the measurement system settings including the antenna arrays and accompanying data are needed for a full description, see Figure 4. The latter one

Tx: PULPA8, h/v-polarized **Rx: SPUCPA 2x12 with MIMO-Cube on top, h/v- polarized**

Figure 3 Pictures of Tx and Rx antenna arrays

we introduce as *meta data* holding the knowledge about the geographical positions, heights and orientations of the used antenna arrays as well as descriptions of the surrounding environment represented by media data.

The reliability of measurement raw data is always a crucial point, since a large number of errors might be occurred during the measurements. At one hand the measurement technique itself can be affected by hardware defects and at the other hand the operator might have forgot to adjust all settings in the

see the program

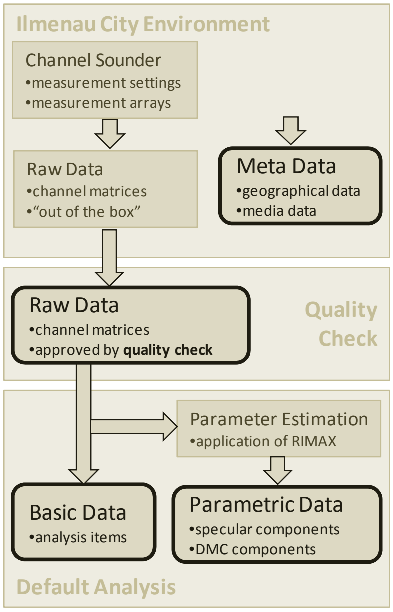


Figure 4 Schematic outcome of the reference campaign

correct way – both situations will lead to more or less useless measurement data. Additionally external influences like disturbing wireless systems in the same frequency range can be responsible for useless measurement data. That's why a so called *quality check* validating the raw data stored by the channel sounder should always be performed.

Dealing with a huge amount of data and for better accessing them, there's a need for a rough characterization of the raw data for further processing. For that simple non-directional analysis items have been introduced called *basic data*.

Since the measurements have been performed with customized antenna arrays an alternative representation of the mobile radio channel by RIMAX based parametric data sets [8][9][10] is given with a default analysis step. These *parametric data* contains an antenna independent set of parameters for the

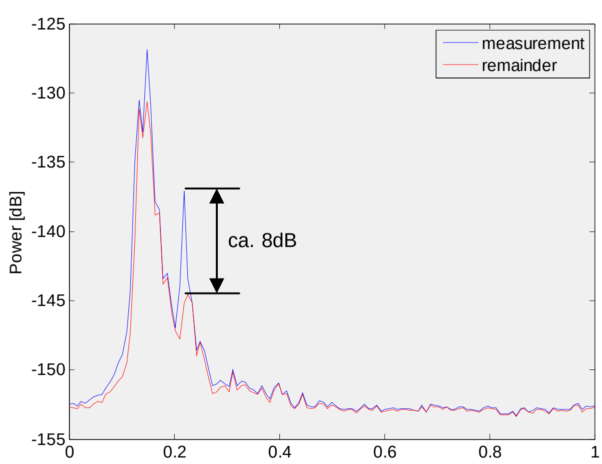


Figure 5 PDP of measurement and remainder after subtraction of estimated paths

specular and dense multipath components. With that additional feature the flexibility for the usage of the reference campaign is significant increased due to the chance of reconstructing the wave field (and therefore the channel matrices) for arbitrary application antenna arrays at both link ends.

C. *Data Validation*

An important and unfortunately often neglected part of a measurement campaign is the validation of the obtained data. The measurement equipment may be affected by errors that did not occur during the calibration of the system. This means although the system functions correctly at the beginning of the measurement one should not assume that this is the case throughout the whole campaign. In general the validation of measurement data is a challenging task. This is due to the complexity of the measurement system, errors can occur either by operating errors or malfunctioning of the hardware. Therefore the data should be validated on measurement side by on-line processing and later more intensive by post-processing procedures.

Since a MIMO channel sounder examines the structure of the channel both in the time and the spatial domain a proper data validation scheme needs to incorporate those domains as well. Therefore, the method presented in this section uses the high resolution parameter estimation framework RIMAX.

RIMAX uses a data model that describes the observation of a MIMO channel as the superposition of specular propagation paths (SC) and dense multipath components (DMC). The presence of DMC can be explained by diffuse scattering or otherwise unresolvable specular paths.

$$\mathbf{x} = \mathbf{s}(\theta_{sc}) + \mathbf{d}(\theta_{dmc}) \in \mathbb{C}^{M_T M_R M_f \times 1} \tag{1}$$

For the validation of measurements the physical meaning of DMC is of lower importance, it is in fact used as a method to improve the estimation of the SC.

The parameter vector θ_{sc}contains the parameters direction of arrival (DoA) and departure (DoD) as well as time-delay of each specular path. The contribution of all paths can be calculated using the following equation:

$$\mathbf{s}(\theta_{sc}) = \sum_{l=1}^{L} \mathbf{B}_{\mathrm{T}}(\varphi_{\mathrm{T,l}}, \vartheta_{\mathrm{T,l}}) \otimes \mathbf{B}_{\mathrm{R}}(\varphi_{\mathrm{R,l}}, \vartheta_{\mathrm{R,l}}) \otimes \mathbf{B}_f(\tau_l) \cdot \gamma_l \tag{2}$$

with γ_l being the complex path weight and $\mathbf{B}_f(\tau_l)$ the vector of complex exponentials in the frequency domain that are induced by the time-delay of the l'th path.

$$\mathbf{B}_f(\tau_l) = \left[e^{-j2\pi\left(-\frac{M_f-1}{2}\right) f_0 \tau_l} \quad \dots \quad e^{-j2\pi\left(+\frac{M_f-1}{2}\right) f_0 \tau_l} \right]^{\mathrm{T}} \tag{3}$$

$\mathbf{B}_R(\varphi_{R,l}, \vartheta_{R,l})$ and $\mathbf{B}_T(\varphi_{T,l}, \vartheta_{T,l})$ denote the steering vectors of the receive and transmit arrays respectively for the DoA and DoD of the l'th path.

see the program

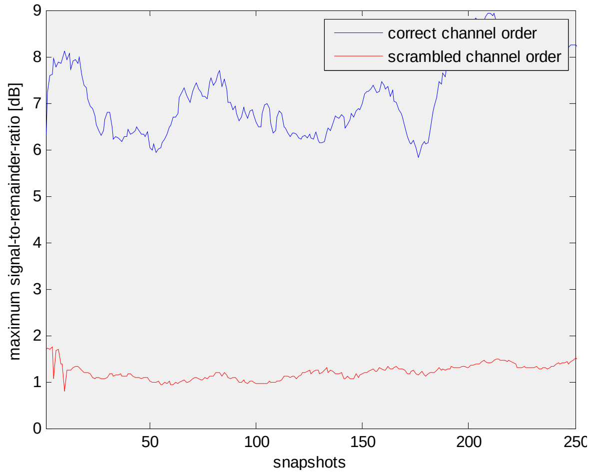


Figure 6 Developing of the maximum SRR of measurements with correct and scrambled channel order

$$\mathbf{B}_R(\varphi_{R,l},\vartheta_{R,l}) = [b_{R,1}(\varphi_{R,l},\vartheta_{R,l}) \quad \cdots \quad b_{R,M_R}(\varphi_{R,l},\vartheta_{R,l})]^{\mathrm{T}} \quad (4)$$

$$\mathbf{B}_T(\varphi_{T,l},\vartheta_{T,l}) = [b_{T,1}(\varphi_{T,l},\vartheta_{T,l}) \quad \cdots \quad b_{T,M_T}(\varphi_{T,l},\vartheta_{T,l})]^{\mathrm{T}} \quad (5)$$

For simplicity this equation applies only for the non-polarimetric case, however all of the following conclusions are also valid in the case of full-polarimetric array beampattern and path weights.

It is intuitive that the data model of RIMAX heavily depends on the order of the array elements of the Rx and Tx respectively. For the measurement this order is defined by the multiplexer switching sequence at the Rx and Tx. Therefore, it is mandatory that the same order is used for parameter estimation that was used during the measurement. In case of multiplexer malfunctioning it cannot be guaranteed that both channel orderings are the same. A typical failure of the multiplexers is that they are not switching at all because of hardware error or that they are switching in a different order because the wrong multiplexing table was loaded into the channel sounder software. A further cause of mismatching channel ordering is the loss of the MIMO-synchronisation. This synchronisation ensures that both Rx and Tx know which array element is currently enabled on either side. This is necessary in order to make sure that the transmitter is only switching to the next element after the receiver switched all of its elements.

If for any reason the channel order is scrambled, the parameters estimator will not be able to determine the path parameters correctly. Since RIMAX is a maximum-likelihood estimator, it will still estimate some specular paths, but when they are subtracted from the measured data they will not significantly reduce the power – which is to be expected when the multiplexer switching and parameters are correct. In order to decide whether the path parameters are valid or not a suitable metric must be introduced. Based upon experience with visual inspection of estimator results, the maximum ratio of the power-delay-profile (PDP) of the measured channel impulse response and the remainder after subtraction of the specular paths is used (maximum signal-to-remainder-ratio SRR). The relevant PDPs are thereby calculated as the mean PDPs for all channels.

High resolution parameter estimation is in general a computational complex task that requires both a lot of memory and computation time. However, for the purpose of data validation it is not necessary to estimate all propagation paths but only a few to demonstrate that the estimation procedure will work. Therefore, RIMAX is configured to estimate a reduced number of paths and to use limited bandwidth. Depending on the propagation environment the estimation of 5 to 10 specular paths seems to be sufficient, as long as individual paths exist that are distinct in all parameter dimensions. To further speed up the estimation RIMAX is only used to estimate the paths for the first 10 snapshot of a measurement file, after that an extended Kalman filter is used to track the parameters of the paths [15].

Figure 5 depicts an example of the PDP of the measured CIR of a snapshot and the PDP of the remainder after the subtraction of the estimated paths. In this case a total number of 5 paths were estimated. Even with this low number of paths it is possible to subtract approximately 8dB of power at a normalized delay of ca. 0.2. This is a strong indication that the multiplexing sequence is correct at both the Rx and Tx.

Figure 6 illustrates the development of the maximum SRR for a part of a measurement route. Here, two cases are distinguished. The first is the SRR using the correct channel order (blue line). It can be seen that it is always in the range of 6dB to 8dB, indicating correct multiplexing. The second case uses a scrambled channel order (red line), more precisely the channel order does not start with the first element of the transmitter and the transmitter switched in the middle of a receiver switching sequence. This results in the SRR not exceeding 2dB. Therefore, it can be concluded that a certain value of the SRR can be used to decide whether the measured data is valid or not. During the validation of all measurement files of the underlying measurement campaign in this paper a value of 3dB turns out to be a proper threshold.

Finally, it can be said that the use of high resolution parameter estimation – with proper limitation of the number of estimated paths - is a promising approach to validate measurement data before it is used for further processing.

D. Sharing the Campaign

Parts of the above described measurement campaign will be accessible for any researcher and institution. This will include the above described outcomes, including raw and meta data, as well as basic and parametric data from RIMAX. All data sets are collected and processed with high quality by considering current standards and available tools. Nevertheless no warranty can be given for improbable but possible mistakes or ambiguities, because of the experimental character of a channel sounding measurement campaign.

see the program

The aim in sharing the campaign to many researchers is to offer the possibility to apply realistic data with high quality and high resolution multipath estimation results to their research work in channel modelling and in particular in system and algorithm design evaluation. Based on this already in an early stage of the research the evaluation can be conducted under realistic conditions. This is important especially for advanced mobile communication concepts considering multiple access of multiple users to multiple base stations equipped with multiple antenna elements. Under these premises the spatial-temporal channel characteristic plays a significantly role, hence realistic data has to be considered for reliable system design and evaluation. Furthermore within the COST framework this measurement data will be placed as a reference scenario [2].
Based on the extensive measurement campaign 9 files describing the links from 3 base stations at the same height of 25m to 3 tracks in the city of Ilmenau are selected. The complete data set for the 9 files as described above is publicly available by accepting the “Terms of Use” at [16], whereby the raw data import functions from the RUSK channel sounder [17] is free accessible by a license agreement.

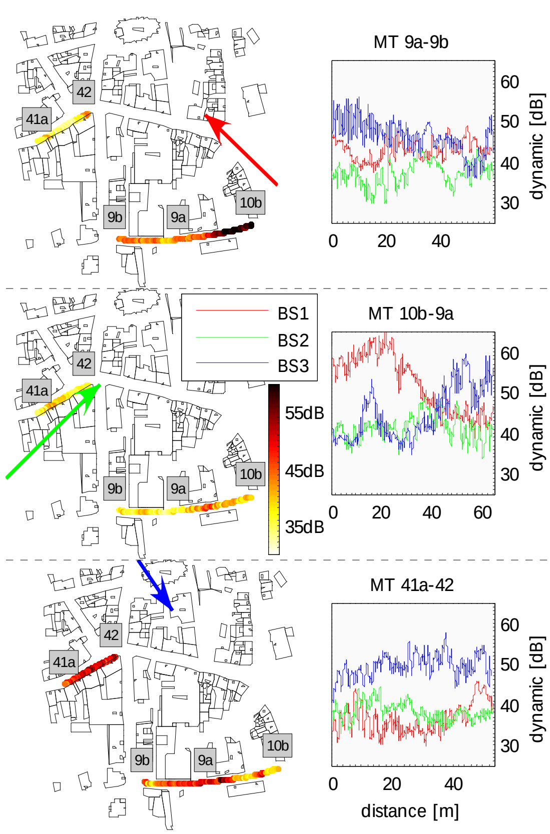


Figure 7 Dynamic along 9 measurement tracks

III. Data for Channel Modelling and System Evaluation

Within a consistent channel sounding work-flow various application fields for the measurement data can be identified. Those applications are basically grouped into analysis and extraction of individual channel parameters, into channel modelling and simulations for performance studies of various algorithms or complete systems. Furthermore the applied methods in these groups depend on the underlying data source: on the one hand the measurement data Figure 4 and on the other hand the results of the high resolution parameter estimation [8][9][10]. The later offers, compared to the pure measurement data, an antenna independent usage of the data, in particular access to the spatial domain of the multipath channel. However this method is only available if the specific antennas are used during the measurement.

Studies for *channel and propagation analysis* will guide to fundamental understanding of the mobile multipath channel. Well-known characterization as described in [13] can be applied for any type of antenna structures on both sides. Whereby receive power, delay spread and geometry factor are only few important examples. In the case of MIMO and/or polarimetric antenna elements the characterization should be extended to MIMO capacity and eigenvalue analysis. The aforementioned studies consider the direct use of the measurement data. By performing measurements with specific antenna arrays, high resolution path parameter estimation can be applied. This allows further studies of directional parameters as DoA and/or DOD and clustering of multipath parameters in various dimensions. The latter step offers a high potential for a closed description and characterisation in the spatio-temporal domain.

Reliable *Channel Models* are essential for network planning and link design. For the design of Next Generation Wireless Networks specified in WINNER or 3GPP LTE the SCM [5], SCME [6] and more advanced WINNER channel models were defined [4][7]**.** Using the measured and processed impulse responses a statistical analysis of parameters corresponding to the requirements, e.g. model structure, small and large scale parameters can be carried out. This leads to a more generalized approach to the measured channel data without neglecting the real scenario. Furthermore an extension of these models towards cooperative MIMO schemes and the influence of the dense multipath components (DMC) is possible.

Another application for the use of channel sounding based measurement data is the calibration and tuning of existing propagation simulation tools. On the one hand one can tune path loss and fading characteristics and on the other hand it is possible to calibrate the results of deterministic ray tracing or ~launching simulators for the actual scenario.

see the program

Link and system level simulations are a basic part in the development and evaluation of new algorithms and transmission strategies. Those simulations can be performed based on direct use of measurement data, on channel models and channel reconstruction. In all cases they will benefit from the use of real world data [11]. This is in particular important if the system under test aims to exploit the spatial domain (MIMO configurations).
The implementation of the raw data as part of the channel module in the software or hardware channel emulator is mapping the properties of the realistic measurement drive on the simulation machine. It acts like a drive or device test in the box, where even complete hardware demonstrators can be tested. One can compare different algorithms or scheme for the same channel realisations under the constraints of measurement antennas and side specific tracks.
Furthermore it is possible to reconstruct the pattern of the electromagnetic waves impinging the measurement antenna. Based on the estimation results of the RIMAX algorithm [9][10] the behaviour of other antennas can be investigated under realistic conditions. With this antenna embedding new antenna designs can be evaluated e.g. in terms of spatial multiplexing capacity.

IV. RESULTS ON DATA ANALYSIS

The following results are derived from the 9 raw data files, which will be publicly available. For these files the dynamic range along the measurement track is shown in Figure 7. On the left side the coverage of each base station to the 3 tracks is highlighted separately (orientation of base stations is depicted by colored arrows). Each base station can cover all 3 tracks, whereby the dynamic (maximum peak to off-peak ration in CIR) is typically larger than 30dB resulting in a good raw data quality for further processing. On the right side in Figure 7 the dynamic ranges are shown from the perspective of the mobile

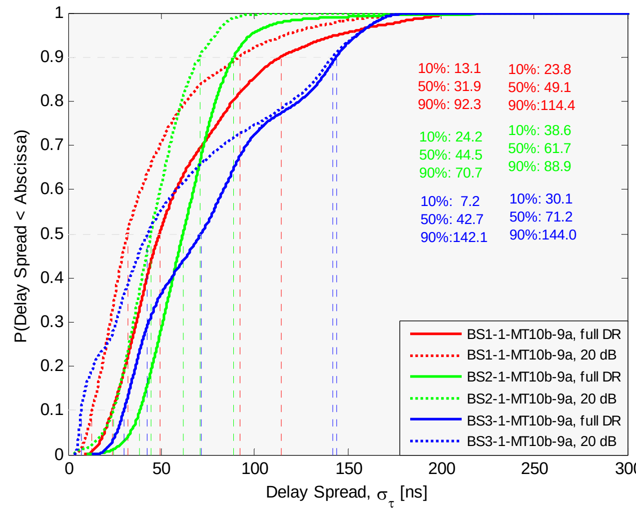


Figure 9 CDFs of per-snapshot delay spread (DS), calculated for same route of the mobile terminal toward different base stations.

terminal to the 3 base stations. Different colors identify the base stations in the same manner as in the left column of the figure. In order to acquire better observation of the channel, the RUSK TUI-FAU sounder uses dynamic AGC, meaning that each individual measurement channel, defined as a combination of the one Tx and one Rx antenna element in antenna-switched concept, has independent gain control. This feature requires proper interpretation during the channel characterization: in Figure 8 it is showed that cross-polar losses are partially compensated and therefore the observed dynamic ranges are similar for co-polarized (VV and HH) and cross-polarized combinations of Tx and Rx antenna elements.
The measurement data are taken sequentially over the same routes for the mobile terminal toward different base stations.

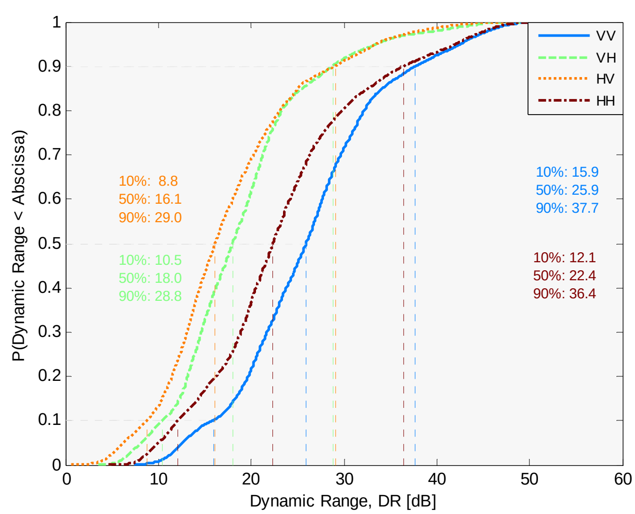


Figure 8 CDFs of per-snapshot dynamic ranges, for different combinations of Tx and Rx antenna polarizations.

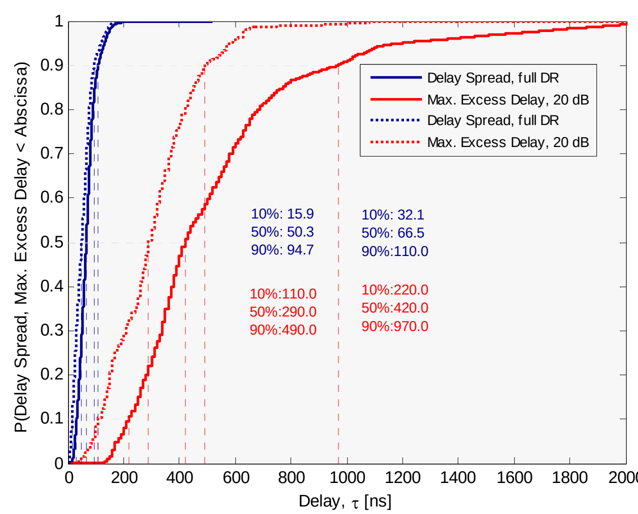


Figure 10 CDFs of Delay Spread and Max. Excess Delay for 9 selected measurement routes.

Figure 9 shows Delay-Spread (DS) distribution over the route 10b-9a for 3 used base stations. As noted in [14] parameters determined from Power-Delay-Profile (PDP) show significant dependence on available Dynamic-Range (DR). This is also illustrated on Figure 9: lower DS values are obtained if DR of PDP is limited to 20 dB. The extraction procedure and the underlying processing assumptions have significant influence on the derived results. Hence careful processing of the data should be performed to acquire results matching the goals for modeling, analysis or simulation.
When all values of DS and Maximum-Excess-Delay (MED, from every measured route of the mobile terminal toward different base stations) are used to generate probability distribution (Figure 10), we get the stochastic equivalent of measured radio-environment (scenario). When DR was set to 20 dB, DS values corresponding to different space-time samples become comparable. This complies with WINNER-modeling approach, and therefore obtained results should be comparable with previously reported parameters for C2 scenario (typical urban macro cell) [4]. Since WINNER model provides median DS parameters separately for LoS (41 ns) and NLoS (234 ns), we can confirm the matching with the computed median of 50.3 ns under the assumption that selected routes mainly exhibit the LoS condition. Calculated values show that in 99% DS is lower than 166 (152) ns, and MED is lower than 1950 (870) ns, depending weather full or limited (to 20 dB) DR is used.

Furthermore high resolution multipath parameter estimation based on RIMAX [9] [10] has been applied. Preliminary results are shown for the 9 measurement files along the tracks. Because of the computational complexity unfortunately not all snapshots in each of the files could be performed in time. Figure 11 shows the DoA spreads at the mobile. On the left hand side the results are shown in the map along the tracks. Each map indicates the results form base stations view jointly for 3 selected tracks. It is interestingly to note, that one particular base station position (BS2) offers for this 3 measurement tracks always very high azimuth spreads. Not shown here, but it can be expected that the MIMO capabilities for these links will be very good. On the right side of Figure 11 the results are shown jointly for the 3 base stations seen from the mobile along the individual tracks. Here the inter-site channel situation for the mobile routes can be studied. In general the characteristics of the azimuth spreads at the mobile are changing quite frequently, which indicates changing situation within the propagation channel in terms of the multipath scattering and richness. Furthermore in the lower map for the track MT 41a-42 a considerable correlation between the developing of the azimuth spreads for the links from BS1 and B2 is found. Further studies are required here.
Besides the azimuth characteristics at the mobile also the elevation is of interest. These results are highlighted in Figure 12. For the track MT 9a-9b lower elevation spreads with not to many changes in the developing compared to the other two

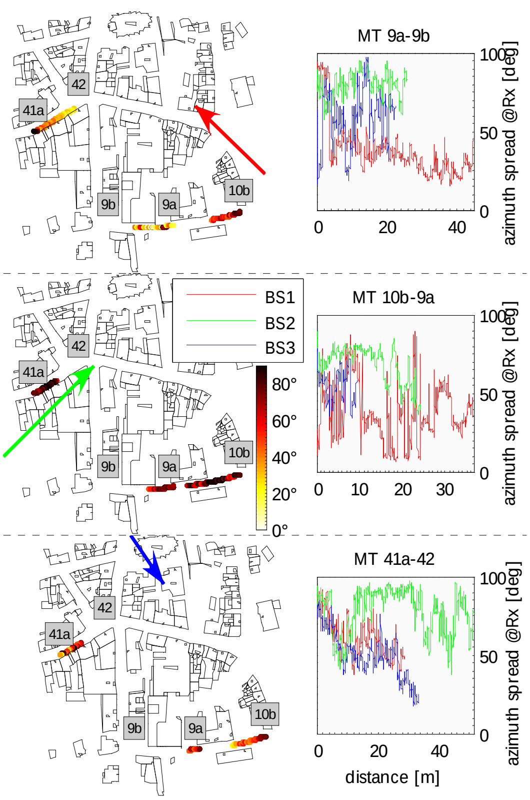


Figure 11 DoA azimuth spreads at the mobile along the measurement tracks

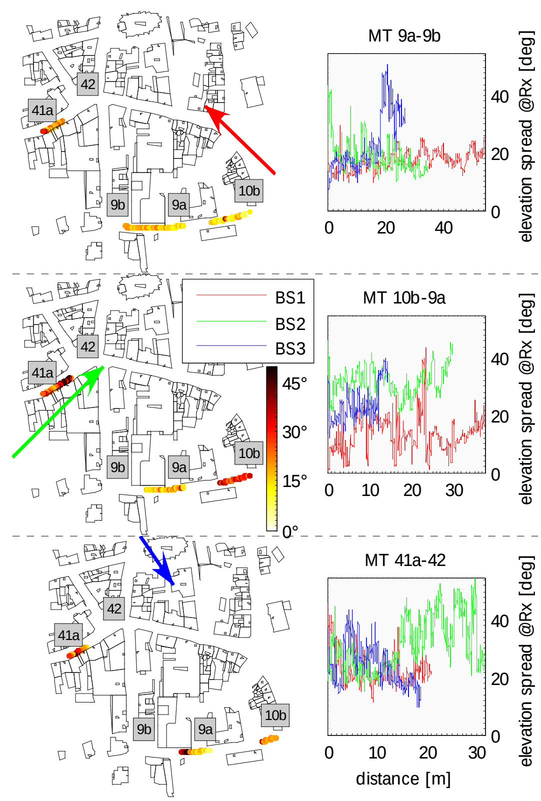


Figure 12 DoA elevation spread at mobile along the measurement tracks

see the program

tracks are found. In general the elevation spreads seems to oscillate between 20°-30°, but also easily reaches more than 45°. This underlines the importance especially in MIMO applications of the elevation in characterising and modelling of mobile channels.

V. Summary

In the baseline of the paper an extensive channel sounding campaign has been presented. The MIMO measurements were focused on configurations for multi user and multi base station applications in an urban macro cell scenario. The frequency range covers the deployment for LTE. Furthermore a sketch of the channel sounding work-flow is discussed including an approach for data validation. The data can be used as base for various applications, e.g. detailed channel analysis and channel modelling as well as on soft- and hardware simulators for link- and system-level evaluations of advanced multi-antenna concepts. First measurement campaign outcomes are shown. The results indicated a very good measurement data quality in terms of dynamic range coverage. A delay spread and maximum excess delay analysis support the findings from the WINNER project for the C2 scenario. High resolution path parameter estimation are applied and offered interesting insight to the spatial characteristics of the gathered measurement data. The measurement campaign will be introduced as reference scenario within the COST 2100 frame work. Parts of the measurement data including raw data, meta data and estimated multipath parameters will be publicly available.

VI. Acknowledgement

This work was supported in part by the research excellence cluster UMIC at RWTH Aachen. The authors would like to thank Volker Jungnickel from FhG-HHI, Berlin, for the support with the antenna arrays. Furthermore Carsten Jandura from Vodafone AG for the support with the measurement car, and Medav GmbH for providing parts of the sounder equipment.

References

[1] L. Correia, Ed., “Mobile Broadband Multimedia Networks,” Academic Press, 2006.
[2] http://grow.lx.it.pt/web/cost2100/
[3] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti,H. Hofstetter, P. Kyosti, D.Laurenson, G. Matz, A. Molisch, C. Oestges,and H. Ozcelik, “Survey of channel and radio propagation models for wireless MIMO systems,” in EURASIP Journal on Wireless Communications and Networking, 2007.
[4] https://www.ist-winner.org
[5] http://www.3gpp.org/
[6] D. S. Baum, J. Salo, G. Del Galdo, M. Milojevic, P. Kyösti, and J. Hansen, “An interim channel model for beyond-3G systems,” in Proc. IEEE VTC’05, Stockholm, Sweden, May 2005.
[7] Narandzic,Milan, Schneider,Christian, Thomä,Reiner S, Jämsä,T., Kyösti,P., Zhao,X., “Comparison of SCM, SCME, and WINNER Channel Models,” IEEE VTC2007-Spring, Dublin, Ireland, April 2007.
[8] R.S. Thoma, M. Landmann, A. Richter, et. al., Multidimensional High-Resolution Channel Sounding, in T. Kaiser et. al. (Ed.), Smart Antennas in Europe - State-of-the-Art, EURASIP Book Series on SP&C, Hindawi Publishing Corporation, Vol. 3.
[9] M. Landmann, W. Kotterman, R.S. Thomä, „Estimated Angular Distributions in Channel Characterisation,” in EUCAP 2007, Edinburgh, UK.
[10] Richter A., “On the Estimation of Radio Channel Parameters: Models and Algorithms (RIMAX),” Ph.D. dissertation, TU-Ilmenau, Ilmenau, Germany, May 2005.
[11] U. Trautwein, C. Schneider, R.S. Thomä, „Measurement Based Performance Evaluation of Advanced MIMO Transceiver Designs”, EURASIP Journal on Applied Signal Processing 2005, No.11, pp.1712-1724.
[12] C. Schneider, U. Trautwein, W. Wirnitzer, R.S. Thomä; “Performance Verification of MIMO Concepts using Multi-Dimensional Channel Sounding,” European Signal Processing Conference, EUSIPCO 2006, Florence, Italy, September 2006.
[13] T. S. Rappaport, “Wireless Communications: Principles and Practice”, 2nd Edition, Prentice Hall, 2002.
[14] M. Narandžić, M. Landmann, C. Schneider, and R. S. Thomä: "Influence of Extraction Procedures on Estimated Wideband MIMO Channel Parameters", Proc. of IST Mobile & Wireless Communication Summit, Budapest, Hungary, July 1-5, 2007.
[15] Salmi, J.; Richter, A.; Enescu, M.; Vainikainen, P. & Koivunen, V. “Propagation Parameter Tracking using Variable State Dimension Kalman Filter” Proc. VTC 2006-Spring Vehicular Technology Conference IEEE 63rd, 2006, 6, 2757-2761
[16] www-emt.tu-ilmenau.de
[17] www.channelsounder.de

see the program

ARRAYS OF ISOTROPIC RADIATORS — A FIELD-THEORETIC JUSTIFICATION

Hristomir Yordanov, Peter Russer

Institute for High Frequency Engineering
Technische Universität München
yordanov@tum.de

Michel T. Ivrlač, Josef A. Nossek

Institute for Circuit Theory and Signal Processing
Technische Universität München
ivrlac@tum.de

ABSTRACT

Because of their conceptual simplicity and mathematical convenience, isotropic radiators are all too popular in the analytical treatment of antenna arrays. However, because an isotropic electro-magnetic field is physically impossible, it is legitimate to ask on which ground the assumption of having isotropic radiators in the theoretical analysis of radio communications is justified. In this paper, we provide a field-theoretic answer to this question. To this end, we show that a uniform linear array (ULA) of isotropic radiators leads qualitatively to the same relative antenna coupling as if the ULA was composed of Hertzian dipoles. This implies that the transmit array gain – that is, the reduction of necessary transmit power to generate the same electric field strength at a given point of interest, compared to using a single antenna of the array – of a ULA of isotropic radiators, is *qualitatively the same* as of a ULA of Hertzian dipoles. This shows that, albeit isotropic antennas do not exist, their application in theoretical investigation of antenna arrays is legitimate, because they lead to physically reasonable antenna coupling which allows a qualitatively correct analysis of important performance parameters, like array gain.

1. INTRODUCTION

Antenna arrays can greatly enhance the performance of radio communication systems by increasing link reliability, capacity and coverage [1]. Therefore, antenna arrays are the object of study in a variety of different disciplines, like electro-magnetic field theory, signal and system theory, coding theory, and information theory. Of course, the methods and notions used are different in each discipline. In the signal processing and information theory literature, it is common to exercise a rather simplistic view on antenna arrays: the individual antennas of the antenna array are usually assumed to be *isotropic radiators* (see, e.g. [2], chapter 7). However, the appealing conceptual simplicity and the mathematical convenience of such an approach are overshadowed by the very fact that antennas which excite an isotropic vector-field do not exist [3]. Therefore, it is an interesting question, on which ground one can justify isotropic radiators in the analytical treatment of antenna arrays.

In this paper, we provide a field-theoretic justification of the legitimacy of the use of isotropic radiators in the theoretical treatment of antenna arrays. We show that a uniform linear array (ULA) of isotropic radiators leads to qualitatively the same relative antenna mutual coupling as if the ULA was composed of Hertzian dipoles. Because the transmit power is directly related with antenna coupling [4], it follows that the transmit array gain[1] of a ULA composed of isotropic radiators is qualitatively the same as the transmit array gain of a ULA of Hertzian dipoles. Albeit isotropic antennas do not exist, their application in theory is legitimate, because they lead to an antenna coupling, which is physically reasonable. This allows qualitatively correct analysis of important performance parameters, like array gain, without having to give up simplicity and convenience which isotropic radiators offer in system modeling.

2. UNIFORM LINEAR ARRAY OF ISOTROPS

A ULA of N isotropic radiators can be seen as a linear N-port. Assuming lossless radiators, the radiated power is equal to the electric input power:

$$P_{\mathrm{Tx}} = \mathrm{Re}\left\{\underline{\boldsymbol{V}}^{\mathrm{H}}\underline{\boldsymbol{I}}\right\}, \tag{1}$$

where $\underline{\boldsymbol{V}} \in \mathbb{C}^{N\times 1}\cdot\mathrm{V}$, and $\underline{\boldsymbol{I}} \in \mathbb{C}^{N\times 1}\cdot\mathrm{A}$, are the complex voltage and current envelopes, respectively, while $\mathrm{Re}\{.\}$, and $(.)^{\mathrm{H}}$ are the real-part, and the Hermitian operations, respectively. Because of linearity, we have that $\underline{\boldsymbol{V}} = \boldsymbol{Z}\underline{\boldsymbol{I}}$, where $\boldsymbol{Z} \in \mathbb{C}^{N\times N}\cdot\Omega$ is the impedance matrix of the array. Because of reciprocity, there is $\boldsymbol{Z} = \boldsymbol{Z}^{\mathrm{T}}$, and hence, the transmit power can be written:

$$P_{\mathrm{Tx}} = \underline{\boldsymbol{I}}^{\mathrm{H}}\mathrm{Re}\{\boldsymbol{Z}\}\underline{\boldsymbol{I}}. \tag{2}$$

Let us now compute $\mathrm{Re}\{\boldsymbol{Z}\}$. The strength of the electric far-field in a distance r and elevation angle θ, excited by the ULA located in the origin and aligned with the z-axis, can be written as (see e.g., [3] on pp 250 and 258):

$$\underline{E} = \alpha\,\frac{\mathrm{e}^{-\mathrm{j}kr}}{r}\,\boldsymbol{a}^{\mathrm{H}}\underline{\boldsymbol{I}}. \tag{3}$$

[1] The transmit array gain is the factor by which the transmit power can be reduced compared to a single antenna of the array, while the same strength of the electric field is excited at a position of interest.

see the program

Herein, $\boldsymbol{a} = \begin{pmatrix} 1 & \mathrm{e}^{\mathrm{j}\nu} & \cdots & \mathrm{e}^{\mathrm{j}(N-1)\nu} \end{pmatrix}^{\mathrm{T}}$, with $\nu = kd\cos\theta$, where d is the distance between the radiators, $k = 2\pi/\lambda$, with λ denoting the wavelength, and α is a constant. The power density is then given by (see e.g., [5], eq. (8) on page 571):

$$S = \alpha' \left|\underline{E}\right|^2 = \frac{\alpha' |\alpha|^2}{r^2} \underline{\boldsymbol{I}}^{\mathrm{H}} \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} \underline{\boldsymbol{I}}, \tag{4}$$

where $\alpha' \in \mathbb{R}_+ \cdot \Omega^{-1}$ is a constant. We can obtain the transmit power by integrating the power density over a sphere around the array:

$$P_{\mathrm{Tx}} = \iint\limits_{\text{sphere}} S \,\mathrm{d}A = 2\pi\alpha' |\alpha|^2 \underline{\boldsymbol{I}}^{\mathrm{H}} \left(\int_0^{\pi} \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} \sin(\theta)\,\mathrm{d}\theta \right) \underline{\boldsymbol{I}}. \tag{5}$$

Since $P_{\mathrm{Tx}} = \underline{\boldsymbol{I}}^{\mathrm{H}} \mathrm{Re}\{\boldsymbol{Z}\} \underline{\boldsymbol{I}} = \gamma \cdot \underline{\boldsymbol{I}}^{\mathrm{H}} \boldsymbol{C} \underline{\boldsymbol{I}}$, it follows from (5) that

$$(\mathrm{Re}\{\boldsymbol{Z}\})_{m,n} = \gamma \cdot (\boldsymbol{C})_{m,n} = \gamma \cdot j_0(kd|m-n|), \tag{6}$$

where $\gamma = 4\pi\alpha' |\alpha|^2$, is a constant, and

$$j_0(x) = \frac{\sin(x)}{x}. \tag{7}$$

Notice that the mutual coupling matrix $\boldsymbol{C} \in \mathbb{C}^{N\times N}$, is a dimensionless, real-valued Toeplitz matrix, which has got unity entries all along its main diagonal:

$$\boldsymbol{C} = \begin{bmatrix} 1 & j_0(kd) & j_0(2kd) & j_0(3kd) & \cdots \\ j_0(kd) & 1 & j_0(kd) & j_0(2kd) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{8}$$

The real-part of the impedance matrix can therefor be written compactly as:

$$\mathrm{Re}\{\boldsymbol{Z}\} = \gamma \cdot \boldsymbol{C}. \tag{9}$$

It is interesting to note that the computation of the mutual coupling matrix for the ULA of lossless isotropic radiators, is based solely on the superposition principle for the electric field, as expressed in (3), and the law of conservation of energy, from which (2) follows. Only for specific antenna separation the coupling matrix is the identity, namely only for $kd/\pi = 2d/\lambda \in \mathbb{N}$. For all other antenna separations, the very law of conservation of energy enforces that there *is* mutual coupling!

3. UNIFORM LINEAR ARRAY OF HERTZIAN DIPOLES

Let us first consider an array of two collinear Hertzian dipoles, located at a distance d from each other and aligned with the z axis of a rectangular coordinate system. The dipoles are situated on the x axis, as shown in Figure 1. We can consider the system of two dipoles as a two-port, which can be described by its impedance matrix. In particular:

$$\begin{aligned} \underline{V}_1 &= Z_{1,1}\underline{I}_1 + Z_{1,2}\underline{I}_2 \\ \underline{V}_2 &= Z_{2,1}\underline{I}_1 + Z_{2,2}\underline{I}_2 \end{aligned}, \tag{10}$$

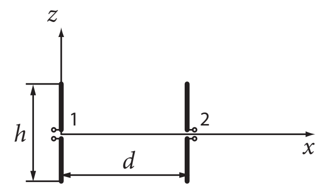


Fig. 1. An array of two Hertzian dipoles

where $\underline{I}_1$, and $\underline{I}_2$ are the current phasors which are impressed on antenna 1 and 2, respectively, while $\underline{V}_1$, and $\underline{V}_2$ are the corresponding induced voltage phasors. The Z_{11} and the Z_{22} elements of this matrix are defined as

$$Z_{1,1} = \left.\frac{\underline{V}_1}{\underline{I}_1}\right|_{\underline{I}_2=0}, \qquad Z_{2,2} = \left.\frac{\underline{V}_2}{\underline{I}_2}\right|_{\underline{I}_1=0}. \tag{11}$$

The conditions $\underline{I}_2 = 0$ and $\underline{I}_1 = 0$ mean that the matrix element $Z_{i,i}$ is simply the input impedance of the i-th antenna. The real part of the input impedance of a Hertzian dipole is given by [4]

$$R_{\mathrm{r}} = \mathrm{Re}\{Z_{1,1}\} = \mathrm{Re}\{Z_{2,2}\} = \frac{2}{3}\pi Z_{\mathrm{F0}} \left(\frac{h}{\lambda}\right)^2, \tag{12}$$

where $Z_{\mathrm{F0}} \approx 377\,\Omega$, is the free-space wave impedance, and λ is the free-space length of the radiated wave. Because of reciprocity, we have $Z_{1,2} = Z_{2,1}$. As of (10), the mutual impedance $Z_{2,1}$ is defined as

$$Z_{2,1} = \left.\frac{\underline{V}_2}{\underline{I}_1}\right|_{\underline{I}_2=0}, \tag{13}$$

that is, as the ratio between the induced voltage on the second antenna due to the impressed current on the first antenna, provided that the current on the second antenna vanishes. We can easily compute the induced voltage on the second antenna by integrating the tangential component of the electric field of the first antenna over the length of the second antenna [6]. The electric field of a Hertzian dipole, located in the origin of the coordinate system due to an impressed current $\underline{I}_1$ is given in spherical coordinates as:

$$\underline{E}_\theta = \frac{h\underline{I}_1}{4\pi\mathrm{j}\omega\varepsilon_0}\left(-\frac{k^2}{r} + \frac{\mathrm{j}k}{r^2} + \frac{1}{r^3}\right)\mathrm{e}^{-\mathrm{j}kr}\sin\theta, \tag{14}$$

$$\underline{E}_r = \frac{h\underline{I}_1}{2\pi\mathrm{j}\omega\varepsilon_0}\left(\frac{\mathrm{j}k}{r^2} + \frac{1}{r^3}\right)\mathrm{e}^{-\mathrm{j}kr}\cos\theta, \tag{14a}$$

where $k = 2\pi/\lambda = \omega\sqrt{\varepsilon_0\mu_0}$, and

$$r = \sqrt{x^2 + y^2 + z^2}, \tag{15}$$

$$\theta = \arctan\frac{\sqrt{x^2+y^2}}{z} \tag{15a}$$

The voltage which can be measured at the terminals of the receiving dipole antenna is given by the following expression:

$$\underline{V} = -\int\limits_{\text{Antenna}} \vec{\underline{E}} \cdot \mathrm{d}\vec{l}, \tag{16}$$

where $\vec{\underline{E}}$ is the electric field which is excited by the antenna, and the integration is performed over the antenna length. For the configuration, shown in Figure 1, we have:

$$\vec{\underline{E}} \cdot \mathrm{d}\vec{l} = \underline{E}_z \,\mathrm{d}z. \tag{17}$$

Since the Hertzian dipole is electrically small, i.e. $kh \ll 1$, we can approximate the z-component of the electric field on the antenna by the θ component

$$\underline{E}_z \approx \underline{E}_\theta(\theta = \pi/2) = \frac{h\underline{I}_1}{4\pi \mathrm{j}\omega\varepsilon_0}\left(-\frac{k^2}{d} + \frac{\mathrm{j}k}{d^2} + \frac{1}{d^3}\right)\mathrm{e}^{-\mathrm{j}kd}. \tag{18}$$

The voltage between the terminals of the second antenna is:

$$\underline{V}_2 = -\int_{h/2}^{-h/2} \underline{E}_\theta \,\mathrm{d}z = \int_{-h/2}^{h/2} \underline{E}_\theta \,\mathrm{d}z = \underline{E}_\theta h = $$
$$= \frac{h^2\underline{I}_1}{4\pi \mathrm{j}\omega\varepsilon_0}\left(-\frac{k^2}{d} + \frac{\mathrm{j}k}{d^2} + \frac{1}{d^3}\right)\mathrm{e}^{-\mathrm{j}kd}. \tag{19}$$

Therefore, the mutual impedance is given by

$$Z_{2,1} = \frac{\underline{V}_2}{\underline{I}_1} = \frac{h^2}{4\pi \mathrm{j}\omega\varepsilon_0}\left(-\frac{k^2}{d} + \frac{\mathrm{j}k}{d^2} + \frac{1}{d^3}\right)\mathrm{e}^{-\mathrm{j}kd}. \tag{20}$$

Using $k = 2\pi/\lambda = \omega\sqrt{\varepsilon_0\mu_0}$, and $Z_{\mathrm{F0}} = \sqrt{\mu_0/\varepsilon_0}$, as well as (12), the real part of this expression can be rewritten as

$$\mathrm{Re}\{Z_{2,1}\} = R_\mathrm{r} \cdot \Psi(kd), \tag{21}$$

where

$$\Psi(x) = \frac{3}{2}\left(\frac{\sin x}{x} + \frac{\cos x}{x^2} - \frac{\sin x}{x^3}\right). \tag{22}$$

This result can easily be generalized to the case of more than two Hertzian dipoles. For a ULA of N Hertzian dipoles, we obtain the following real-part of the impedance matrix:

$$\mathrm{Re}\{\boldsymbol{Z}_\mathrm{H}\} = R_\mathrm{r} \cdot \boldsymbol{C}_\mathrm{H}, \tag{23}$$

where

$$\boldsymbol{C}_\mathrm{H} = \begin{bmatrix} 1 & \Psi(kd) & \Psi(2kd) & \Psi(3kd) & \cdots \\ \Psi(kd) & 1 & \Psi(kd) & \Psi(2kd) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{24}$$

What is the major difference of the real-part of the impedance matrix of the ULA of Hertzian dipoles, when we compare it to the case of isotropic radiators? To answer this question, we just have to look at the pair of equations (24) and (23) for the case of Hertzian dipoles, and compare it with the corresponding pair of equations (8) and (9) for the case of isotropic radiators. Obviously, the main difference is that the relative mutual coupling is described by the function $j_0(x)$ from (7) for the case of isotropic radiators, while it is described by the function $\Psi(x)$ from (22), for the case of Hertzian dipoles. At first glance,

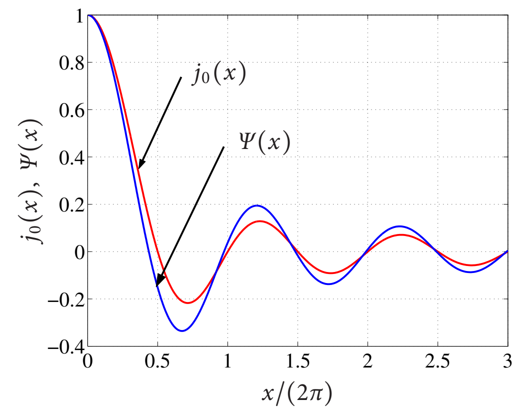


Fig. 2. The functions $j_0(x)$ from (7), and $\Psi(x)$ from (22).

$j_0(x)$ and $\Psi(x)$ appear to be rather different. However, when we look at Figure 2, we observe that there actually is a striking similarity. Both functions are decreasing at first, when x is increased from zero, and then display an oscillatory behavior, as x is further increased. While $j_0(x)$ has equidistant roots (integer multiples of π), the function Ψ has *almost* equidistant roots at *almost* the same positions as $j_0(x)$. From what we see in Figure 2, we may therefore conclude, that $j_0(x)$, and $\Psi(x)$ are *qualitatively the same*. This justifies the use of isotropic radiators in the theoretical treatment of antenna arrays, because they impose qualitatively the same relative mutual coupling.

4. TRANSMIT ARRAY GAIN

Array Gain is an important performance measure of antenna arrays. Because array gain critically depends on antenna mutual coupling, we expect – from what was said before – to find *qualitatively the same* results no matter if the array is equipped by Hertzian dipoles, or isotropic radiators. Let us check this. From (3) and (14), the electric far-field strength excited by a ULA of Hertzian dipoles can be written as:

$$\underline{E} = \xi \cdot \boldsymbol{a}^\mathrm{H}\underline{\boldsymbol{I}} \cdot \sin\theta, \tag{25}$$

where ξ is a (distance-dependent) constant, and $\boldsymbol{a} \in \mathbb{C}^{N\times 1}$ is the (direction dependent) Vandermonde array steering vector. Hence,

$$|\underline{E}|^2 = |\xi|^2 \cdot \underline{\boldsymbol{I}}^\mathrm{H}\boldsymbol{a}\boldsymbol{a}^\mathrm{H}\underline{\boldsymbol{I}} \cdot \sin^2\theta. \tag{26}$$

The power P_Rx, that can be extracted from the field by an antenna is then proportional to (26), such that:

$$P_\mathrm{Rx} = \zeta \cdot \underline{\boldsymbol{I}}^\mathrm{H}\boldsymbol{a}\boldsymbol{a}^\mathrm{H}\underline{\boldsymbol{I}}, \tag{27}$$

where $\zeta \in \mathbb{R}_+ \cdot \Omega$ is proportional to $|\xi|^2 \sin^2\theta$. The optimum excitation current, for a given steering vector $\boldsymbol{a}$, maximizes the received signal power for a given transmit power P_Tx:

$$\underline{\boldsymbol{I}}_\mathrm{opt} = \arg\max_{\underline{\boldsymbol{I}}} \underline{\boldsymbol{I}}^\mathrm{H}\boldsymbol{a}\boldsymbol{a}^\mathrm{H}\underline{\boldsymbol{I}}, \quad \text{s.t.} \quad R_\mathrm{r} \cdot \underline{\boldsymbol{I}}^\mathrm{H}\boldsymbol{C}_\mathrm{H}\underline{\boldsymbol{I}} \le P_\mathrm{Tx}. \tag{28}$$

see the program

The transmit power constraint follows from substituting (23) for $\mathrm{Re}\{\boldsymbol{Z}\}$ in (2). It can be shown that the receive power under optimum beamforming is given by [7]:

$$P_{\mathrm{Rx}}^{\max} = \frac{P_{\mathrm{Tx}}\,\zeta}{R_{\mathrm{r}}}\,\boldsymbol{a}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{H}}^{-1}\boldsymbol{a}. \tag{29}$$

If only *one* transmit antenna, say the n-th one, is excited by a non-zero electric current, the receive power becomes:

$$P_{\mathrm{Rx},n} = \frac{P_{\mathrm{Tx}}\,\zeta}{R_{\mathrm{r}}}\,|a_n|^2,$$

where a_n is the n-th component of the steering vector $\boldsymbol{a}$. The average received power over all transmit antennas becomes

$$P_{\mathrm{Rx,avr}}^{\mathrm{single}} = \frac{P_{\mathrm{Tx}}\,\zeta}{N\,R_{\mathrm{r}}}\sum_{n=1}^{N}|a_n|^2, \tag{30}$$

such that the *transmit array gain* is defined as:

$$A_{\mathrm{Tx}} = \frac{P_{\mathrm{Rx}}^{\max}}{P_{\mathrm{Rx,avr}}^{\mathrm{single}}}. \tag{31}$$

When we substitute (29), and (30) into (31), it follows that

$$A_{\mathrm{Tx}} = N\,\frac{\boldsymbol{a}^{\mathrm{H}}\boldsymbol{C}_{\mathrm{H}}^{-1}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\boldsymbol{a}}. \tag{32}$$

Note that A_{Tx} does not depend on R_{r}, or ζ, but solely on the array steering vector, and the relative antenna mutual coupling matrix. We can obtain the transmit array gain for the case of a ULA of isotropic radiators by simply replacing in (32), the matrix $\boldsymbol{C}_{\mathrm{H}}$ from (24), by the matrix $\boldsymbol{C}$ from (8). Notice that A_{Tx} quantifies how much more receive power we obtain by optimally using all the antennas of the array, compared to a single transmit antenna of the same type as used in the array, when the radiated power is the same in both cases.

For both a ULA of Hertzian dipoles and isotropic radiators, Figure 3 shows the array gain in the direction of the array axis, the so-called »end-fire« direction, for different antenna number N, as a function of the antenna spacing. We can observe that the array gain raises quickly as we reduce the antenna spacing below $\lambda/2$. The maximum array gain happens to grow proportional to N^2. These observations are true for both arrays of Hertzian dipoles and isotropic radiators. The most noticeable difference happens to be that the maximum array gain for the case of Hertzian dipoles is a little bit below the one achieved by isotropic radiators. For antenna separation larger than half the wavelength, the array gain is much lower than with closely spaced antennas, and displays an oscillatory nature. Again this is true for both Hertzian dipoles and isotropic radiators. From what we see in Figure 3, we can conclude that the transmit array gain of a ULA of Hertzian dipoles is *qualitatively the same* as if the array was equipped with isotropic radiators. This property clearly justifies the use of – mathematically and conceptually much simpler – isotropic radiators for theoretical treatment of antenna arrays.

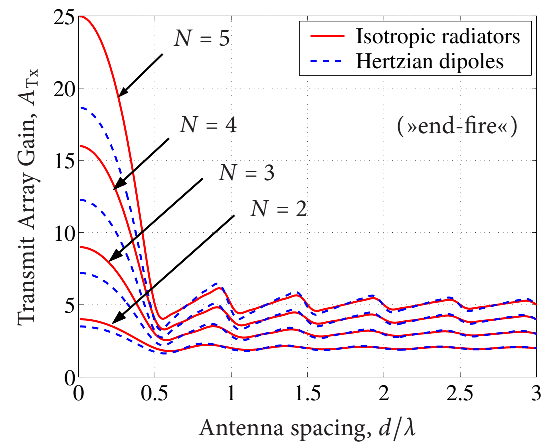


Fig. 3. Transmit array gain in »end-fire« direction, as function of antenna spacing for different antenna numbers. Results for Hertzian dipoles and isotropic radiators are shown.

5. CONCLUSION

A field-theoretic justification for the applicability of isotropic radiators in antenna array modeling is given. A uniform linear array of isotropic radiators leads qualitatively to the same relative antenna coupling as if Hertzian dipoles were used as the array elements. Albeit isotropic antennas do not exist, their application in theory is legitimate, because they lead to a mutual antenna coupling, which is reasonable from a physics point of view. This allows for qualitatively correct analysis of important performance parameters, like array gain.

6. REFERENCES

[1] J. C. Liberti and T. S. Rappaport, *Smart Antennas for Wireless Communications*. New Jersey: Prentice Hall, 1999.

[2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[3] A. Balanis, *Antenna Theory*. Second Edition, John Wiley & Sons, 1997.

[4] P. Russer, *Electromagnetics, Microwave Circuit and Antenna Design for Communications Engineering*, 2nd ed. Nordwood, MA: Artec House, 2006.

[5] D. Kraus, *Electromagnetics*. Fourth Edition, McGraw-Hill, 1992.

[6] S. A. Schelkunoff and H. T. Friis, *Antennas. Theory and Practice*. New York, NY: Wiley, 1952.

[7] M. T. Ivrlač and J. A. Nossek, "The Maximum Achievable Array Gain under Physical Transmit Power Constraint," in *Proc. IEEE International Symposium on Information Theory and its Applications*, Dec. 2008, pp. 1338–1343.

see the program

DIFFERENTIAL FEEDBACK IN MIMO COMMUNICATIONS: PERFORMANCE WITH DELAY AND REAL CHANNEL MEASUREMENTS

Daniel Sacristán-Murga[1], *Florian Kaltenberger*[2], *Antonio Pascual-Iserte*[1,3], *Ana I. Pérez-Neira*[1,3]

[1]Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) - Spain
[2]Mobile Communications - Institut Eurecom - France
[3]Dept. of Signal Theory and Communications - Universitat Politècnica de Catalunya (UPC) - Spain
email: daniel.sacristan@cttc.es, florian.kaltenberger@eurecom.fr, antonio.pascual@upc.edu, ana.isabel.perez@upc.edu

ABSTRACT

This work studies the performance of our recently proposed differential feedback scheme for multi-input-multi-output (MIMO) communication systems using real channel measurement data. The algorithm is applied to the channel correlation matrix exploiting geodesic curves and the intrinsic geometry of positive definite Hermitian matrices. The performance of this and a conventional non-differential feedback scheme are evaluated using real data and channel measurements obtained with the Eurecom MIMO OpenAir Sounder (EMOS). Additionally, the impact of having a delay in the feedback link is also studied in terms of a loss of performance in the communication through several simulations.

The results show that the differential feedback strategy performs much better than the non-differential strategies in low mobility channels, while in high mobility channels the performance is similar. A delay in the feedback channel affects specially high mobility channels while having a negligible impact in the slow-varying cases.

Topics: Precoding and limited feedback, Multi-antenna channel measurements, MIMO systems.

1. INTRODUCTION

Multi-input-multi-output (MIMO) communication systems are shown to provide improved performance when compared to single-antenna configurations, specially when both the transmitter and the receiver have some kind of channel state information (CSI). A possibility to obtain CSI at the transmitter consists in the exploitation of a low rate feedback channel from the receiver to the transmitter. A feedback channel is mandatory in frequency-division duplexing channels, where channel reciprocity does not hold.

In the literature, several feedback schemes have been proposed in order to provide CSI to the transmitter side. For time-varying channels, where the coherence time is higher than the time difference between consecutive feedback instants, a good approach consists in quantizing the channel response in a differential way. This lowers the required feedback load or improves the quality of the quantization for a fixed capacity of the feedback link. Taking this philosophy, there are several techniques, such as the direct scalar quantization of the entries of the channel variation matrix, or more sophisticated approaches, such as those based on geodesic curves over Grassmannian manifolds or correlation-type matrices [1–3].

All these real time feedback schemes suffer from the delay inherent to the feedback channel. This delay causes a mismatch between the true channel and the available CSI and, consequently, between the actual design of the transmitter and the optimum one, which results in a degradation of the performance. The effect of the delay can be alleviated using a channel predictor.

The main objective of this paper is to evaluate experimentally in a real environment the performance of different feedback strategies and the impact of feedback delay. This will be done taking as example the differential technique presented in [3] for feedback and channel quantization (this technique will be summarized in the subsequent sections) and a non-differential feedback strategy from [4]. Both of them will be applied to real channel measurements obtained with the Eurecom's MIMO OpenAir Sounder (EMOS) [5, 6]. In order to alleviate the effect of feedback delay a technique based on channel prediction will be studied. In particular a linear Wiener predictor will be considered.

The paper is organized as follows. The system and signal models are given in section 2. Section 3 summarizes the differential quantization technique used in the feedback link, and section 4 describes the EMOS and the channel measurements. The performance of differential versus non-differential feedback strategies applied to the measured channel, including an analysis of the effect of feedback delay, is shown in section 5. A solution based on prediction is presented in 6. Finally, section 7 concludes the paper.

This work was partially supported by the Catalan Government under grant 2005SGR-00996; by the Spanish Government under projects TEC2008-06327-C03-01 (FBMC-MULAN) and 2A103 (MIMOWA) from MEDEA+ program (AVANZA I+D TSI-020400-2008-150), by the European Commission under projects NEWCOM++ (216715) and CHORIST; and by Eurecom.

2. SYSTEM AND SIGNAL MODELS

This section and the next one summarize some of the ideas presented in [3] concerning the differential feedback technique that will be used in this paper to evaluate a realistic system performance according to real channel measurements.

We consider the transmission through a MIMO channel with n_T and n_R transmit and receive antennas represented at time instant n by matrix $\mathbf{H}(n) \in \mathbb{C}^{n_R \times n_T}$. The n_R received signals at the same time instant, assuming a linear transmitter, can be expressed as

$$\mathbf{y}(n) = \mathbf{H}(n)\mathbf{B}\big(\widehat{\mathbf{R}}_H(n)\big)\mathbf{x}(n) + \mathbf{w}(n) \in \mathbb{C}^{n_R}, \qquad (1)$$

where $\mathbf{x}(n) \in \mathbb{C}^{n_S}$ represents the n_S streams of signals to be transmitted with $\mathbb{E}\big[\mathbf{x}(n)\mathbf{x}^H(n)\big] = \mathbf{I}$, and $\mathbf{B} \in \mathbb{C}^{n_T \times n_S}$ is the linear transmitter matrix. Note that we explicitly indicate that the transmitter depends on the available estimate of the channel correlation matrix $\widehat{\mathbf{R}}_H(n)$, where the exact channel correlation matrix is $\mathbf{R}_H(n) = \mathbf{H}^H(n)\mathbf{H}(n)$. The additive white Gaussian noise (AWGN) at the receiver is represented by $\mathbf{w}(n) \in \mathbb{C}^{n_R}$ with $\mathbb{E}\big[\mathbf{w}(n)\mathbf{w}^H(n)\big] = \sigma_w^2\mathbf{I}$.

In the system setup, it will be considered that the receiver knows perfectly the current channel matrix $\mathbf{H}(n)$ and that the transmitter designs $\mathbf{B}$ assuming that the available CSI at its side represented by $\widehat{\mathbf{R}}_H(n)$ is also perfect. In reality, the CSI at the transmitter is not perfect because it is a quantized version of the perfect CSI obtained at the receiver. The transmitter design can be done according to different criteria, such as the maximization of the mutual information or signal to noise ratio (SNR), or the minimization of the mean square error (MSE) or the bit error rate (BER), among others. In all the cases, the optimum transmitter has been shown to depend only on the channel correlation matrix $\mathbf{R}_H(n)$ [7]. For each of them a cost function $d(\widehat{\mathbf{R}}_H(n), \mathbf{H}(n))$ can be defined, where the design objective is its minimization. A couple of examples of cost functions are given in the following paragraphs, although any criterion can be applied (we drop the dependency with respect to the time index n for the sake of clarity in the notation).

The cost function that follows a maximum SRN with single beamforming (number of streams $n_S = 1$) criteria can be expressed as

$$d(\widehat{\mathbf{R}}_H(n), \mathbf{H}(n)) = -\frac{1}{\sigma_w^2}\left\|\mathbf{HB}\right\|_F^2, \qquad (2)$$

where the transmission matrix $\mathbf{B} \in \mathbb{C}^{n_T \times 1}$ is defined as

$$\mathbf{B}\big(\widehat{\mathbf{R}}_H(n)\big) = \sqrt{P_T}\,\mathbf{u}_{\max}\big(\widehat{\mathbf{R}}_H(n)\big), \qquad (3)$$

and $\mathbf{u}_{\max}(\cdot)$ stands for the unit-norm eigenvector of maximum associated eigenvalue. P_T represents the maximum transmission power, i.e., $\|\mathbf{B}\|_F^2 \le P_T$, where subindex F stands for the Frobenius norm.

The cost function that maximizes the mutual information can be expressed as

$$d(\widehat{\mathbf{R}}_H(n), \mathbf{H}(n)) = -\log_2\left|\mathbf{I} + \frac{1}{\sigma_w^2}\mathbf{B}\mathbf{B}^H\mathbf{H}^H\mathbf{H}\right|, \qquad (4)$$

where the transmission matrix $\mathbf{B} \in \mathbb{C}^{n_T \times n_S}$ is defined as

$$\mathbf{B}\big(\widehat{\mathbf{R}}_H(n)\big) = \widetilde{\mathbf{U}}(n)\mathbf{P}^{1/2}(n), \qquad (5)$$
$$\mathbf{P}(n) = \operatorname{diag}(p_1, \ldots, p_{n_S}), \qquad (6)$$

and $\widetilde{\mathbf{U}}(n)$ consists of n_S columns that are the n_S unit-norm eigenvectors of $\widehat{\mathbf{R}}_H(n)$ associated to its n_S maximum eigenvalues $\{\lambda_i\}_{i=1}^{n_S}$. The power $\mathbf{P}(n)$ is allocated according to the waterfilling solution ($p_i = \max\{0, \mu - 1/\lambda_i\}$ where μ is a constant such that $\sum_{i=1}^{n_S} p_i = P_T$) [7].

The next section is devoted to summarize algorithm [3] for quantizing the actual correlation matrix $\mathbf{R}_H$ (instead of $\mathbf{H}$) from the receiver to the transmitter in a differential way. Since $\mathbf{R}_H$ belongs to the set of Hermitian positive definite matrices,[1] exploiting its inherent geometry will improve the performance of the quantization.

3. ALGORITHM DESCRIPTION FOR QUANTIZATION IN FEEDBACK LINK

In this section first we will give some comments on the concept of geodesic curves on the set of positive definite Hermitian matrices and then we will summarize the basic ideas concerning the algorithm presented in [3] for differential quantization.

3.1. Geodesic curves

As shown in [2] the set of Hermitian positive definite matrices $\mathcal{S} = \{\mathbf{R} \in \mathbb{C}^{n_T \times n_T} : \mathbf{R}^H = \mathbf{R}, \mathbf{R} \succ \mathbf{0}\}$ is a convex cone[2], i.e., $\forall \mathbf{R}_1, \mathbf{R}_2 \in \mathcal{S}, \forall s \ge 0,\ \mathbf{R}_1 + s\mathbf{R}_2 \in \mathcal{S}$ [8]. This set is characterized properly by means of differential geometry, which states a set of definitions for the distance, scalar products and routes within this set:

- *Scalar product:* At any point in this set $\mathcal{S}$ given by $\mathbf{R}$ (also named as *base point*), the scalar product between two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$ is defined as $\langle \mathbf{A}, \mathbf{B}\rangle_{\mathbf{R}} = \operatorname{Tr}(\mathbf{R}^{-1}\mathbf{A}\mathbf{R}^{-1}\mathbf{B})$. This definition implies that the norm is defined as $\|\mathbf{A}\|_{\mathbf{R}} = \sqrt{\operatorname{Tr}(\mathbf{R}^{-1}\mathbf{A}\mathbf{R}^{-1}\mathbf{A})}$.

[1] In the following, it will be assumed that the channel correlation matrix is strictly positive definite. If this cannot be guaranteed because, for example, if $n_R < n_T$, it is possible to work with extended correlation matrices defined as $\widetilde{\mathbf{R}}_H = \mathbf{H}^H\mathbf{H} + \epsilon\mathbf{I}$, $\epsilon > 0$, which are positive definite by construction.

[2] Actually, reference [2] is devoted to the case of real matrices, although the results and conclusions can be extended directly to the complex case.

see the program

- *Geodesic curve:* Given two points $\mathbf{R}_1$and $\mathbf{R}_2$ in the set $\mathcal{S}$, the geodesic curve, which is the curve connecting these points with minimum distance and with all its points belonging to $\mathcal{S}$, is given by

$$\Gamma(t) = \mathbf{R}_1^{1/2} \exp\left(t\mathbf{C}\right)\mathbf{R}_1^{1/2}, \tag{7}$$

where $\mathbf{C} = \log\left(\mathbf{R}_1^{-1/2}\mathbf{R}_2\mathbf{R}_1^{-1/2}\right)$, $\Gamma(0) = \mathbf{R}_1$, and $\Gamma(1) = \mathbf{R}_2$. The derivative of the geodesic curve at $t = 0$, which is in fact the *direction* of such curve at $t = 0$, is given by the Hermitian matrix $\Gamma'(0) = \mathbf{R}_1^{1/2}\mathbf{C}\mathbf{R}_1^{1/2}$.

- *Distance:* The geodesic distance between any two points in $\mathcal{S}$ is given by the length of the geodesic curve that connects them. According to the previous notation, it can be shown that this distance is given by

$$d_g(\Gamma(0), \Gamma(t)) = |t|\|\mathbf{C}\|_F, \ \Rightarrow \ d_g(\mathbf{R}_1, \mathbf{R}_2) = \|\mathbf{C}\|_F. \tag{8}$$

or, using an equivalent expression,

$$d_g(\mathbf{R}_1, \mathbf{R}_2) = \Big(\sum_i |\log\lambda_i|^2\Big)^{1/2}, \tag{9}$$

where $\{\lambda_i\}$ are the eigenvalues of matrix $\mathbf{R}_1^{-1/2}\mathbf{R}_2\mathbf{R}_1^{-1/2}$.

3.2. Differential quantization

In general terms, differential quantization is based on a quantization of the difference between the CSI at consecutive feedback intervals, instead of quantizing the complete CSI every time [4]. Depending on the design criterion and the allowed computational complexity, different strategies arise.

Some techniques can be based on the quantization of the variations of the MIMO channel matrix $\mathbf{H}(n)$ itself or even, on the differential quantization of the strongest right eigenspaces spanned by such matrices [1]. The technique that will used in this paper to evaluate experimentally the performance of the communication setup corresponds to reference [3]. It relies on the fact that in general, all the joint transmitter-receiver designs for MIMO channels and different quality criteria (SNR, MSE, mutual information, etc.) depend on the channel response matrix $\mathbf{H}(n)$ only through the channel correlation matrix defined as $\mathbf{R}_H(n) = \mathbf{H}^H(n)\mathbf{H}(n)$ [7]. Taking this into account, a possible strategy consists in applying a differential quantization exploiting the intrinsic geometry of the set of positive definite Hermitian matrices by means of the use of geodesic curves, as suggested in [2].

The fundamentals of the algorithm proposed in [3], which are summarized here, are based on a differential quantization of the channel correlation matrix $\mathbf{R}_H(n)$. The objective is to minimize the cost function as presented in section 2, which can be related to the quality measure of the system and, therefore, the receiver has to know which kind of design will be applied by the transmitter. If a more general setup is to be applied so that the feedback can be used for any transmitter design, another cost function could be added which is simply the geodesic distance between the actual channel correlation matrix and its fed back estimate, i.e., $d(\widehat{\mathbf{R}}_H(n), \mathbf{H}(n)) = d_g(\widehat{\mathbf{R}}_H(n), \mathbf{H}^H(n)\mathbf{H}(n))$.

The differential quantization algorithm for the feedback of the channel correlation matrix is an iterative procedure. At each iteration n the initial situation is described as follows: the receiver has a perfect knowledge of the current channel matrix $\mathbf{H}(n)$ and both the transmitter and the receiver know which is the last estimate of the channel correlation matrix sent through the feedback channel $\widehat{\mathbf{R}}_H(n-1)$. A possible initialization of the algorithm would correspond to starting the run of the algorithm from the cone vertex before the first iteration: $\widehat{\mathbf{R}}_H(0) = \mathbf{I}$.

At each iteration n, the following steps are followed (all these steps are represented conceptually in Fig. 1):

- *STEP 1:* Both the receiver and the transmitter generate a common set of Q random Hermitian matrices using the same pseudo-random generator and the same seed. Then, these matrices are orthonormalized using the Gram-Schmidt procedure [9] according to the definition of scalar product presented in section 3, producing the set $\{\mathbf{A}_i\}_{i=1}^{Q}$. Finally, each matrix $\mathbf{A}_i$ is rescaled individually so that $\mathbf{C}_i = \mathbf{R}^{-1/2}\mathbf{A}_i\mathbf{R}^{-1/2}$ has a norm equal to Δ ($\|\mathbf{C}_i\|_F = \Delta$) which is, in fact, the quantization step.

- *STEP 2:* Both the receiver and the transmitter use the previous matrices to generate a set of Q geodesic curves $\{\Gamma_i(t)\}_{i=1}^{Q}$, all of them having the same initial point $\mathbf{R} = \widehat{\mathbf{R}}_H(n-1)$ and with orthogonal directions: $\Gamma_i(t) = \widehat{\mathbf{R}}_H^{1/2}(n-1)\exp\left(t\mathbf{C}_i\right)\widehat{\mathbf{R}}_H^{1/2}(n-1)$.

- *STEP 3:* Each of these geodesic curves is used to generate two candidates for the feedback in the next iteration $\widehat{\mathbf{R}}_H(n)$ corresponding to $\Gamma_i(-1)$ and $\Gamma_i(1)$.

- *STEP 4:* The receiver evaluates the cost function for each of the candidates (there are $2Q$ candidates), and sends the selected index i_{FB} through the feedback channel to the transmitter. This index is the one for which the corresponding candidate minimizes the cost function. According to this, the number of feedback bits per iteration has to be higher than or equal to $\log_2(2Q)$. The matrix corresponding to the selected candidate will be used for the transmitter design and as the starting point in the next iteration.

All the previous steps are represented graphically in Fig. 1 for the case of a feedback using 2 bits and taking as optimization criterion the minimization of the geodesic distance to the actual channel correlation matrix $\mathbf{R}_H(n)$. Starting from $\widehat{\mathbf{R}}_H(n-1)$, the algorithm generates 2 orthogonal geodesic

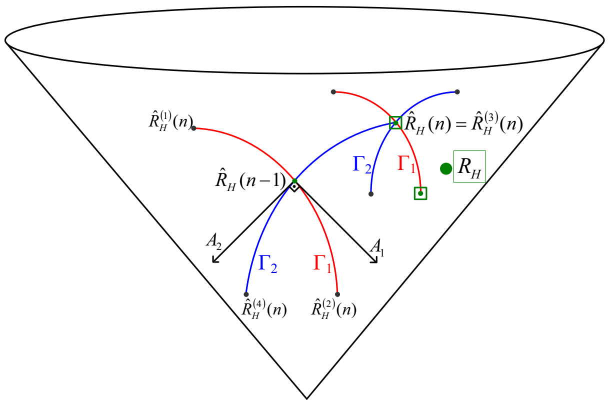


Fig. 1. 2-bit differential quantization in the space of channel correlation matrices.

routes $\Gamma_1(t)$ and $\Gamma_2(t)$ with velocity matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, producing four quantization candidates, all of them at distance Δ from the initial point. At the receiver, each candidate is compared to the actual $\mathbf{R}_H$ and the one with smallest distance (in this example candidate 3) is chosen. That is, its index $i_{FB} = 3$ is sent to the transmitter through the feedback channel and $\widehat{\mathbf{R}}_H(n) = \widehat{\mathbf{R}}_H^{(3)}(n)$. The next iteration starts from this point, generates 2 orthogonal routes and 4 quantization candidates, selects the closest candidate to $\mathbf{R}_H$, and so on.

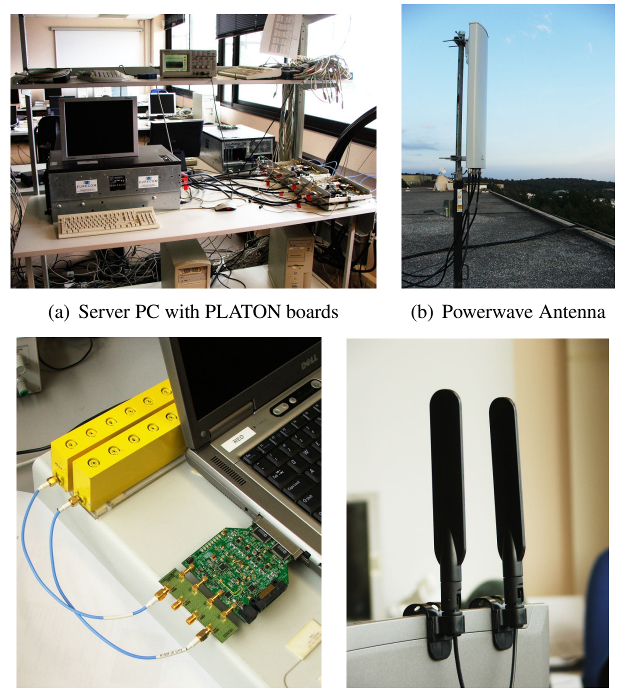
4. REAL CHANNEL MEASUREMENTS

Realistic MIMO channel measurements have been obtained using Eurecom's MIMO Openair Sounder (EMOS) [5, 6]. In this section we first describe the hardware of the EMOS platform and the channel sounding procedure and then the measurement campaign that was carried out for this paper. The obtained measurements are used in the next section to evaluate the previous feedback quantization technique from a realistic point of view.

4.1. Platform description

The EMOS is based on the OpenAirInterface hardware/software development platform at Eurecom.[3] It operates at 1.900-1.920 GHz with 5 MHz channels and can perform real-time channel measurements between a base station and multiple users synchronously. For the base station (BS), a workstation with four PLATON data acquisition cards (see Fig. 2(a)) is employed along with a Powerwave 3G broadband antenna (part no. 7760.00) composed of four elements which are arranged in two cross-polarized pairs (see Fig. 2(b)). The user equipment (UE) consists of a laptop computer with Eurecom's dual-RF CardBus/PCMCIA data acquisition card (see Fig. 2(c)) and two clip-on 3G Panorama Antennas (part no. TCLIP-DE3G, see Fig. 2(d)). The platform is designed for a full software-radio implementation, in the sense that all protocol layers run on the host PCs under the control of a Linux real time operation system.

[3]http://www.openairinterface.org



(a) Server PC with PLATON boards (b) Powerwave Antenna

(c) Dual-RF CardBus/PCMCIA Card (d) Panorama Antennas

Fig. 2. EMOS base-station and user equipment [10].

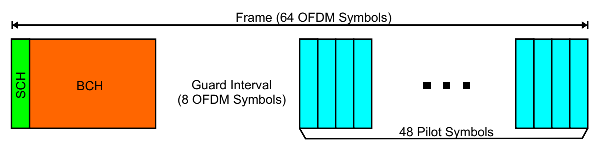


Fig. 3. Frame structure of the OFDM Sounding Sequence. The frame consists of a synchronization channel (SCH), a broadcast channel (BCH), and several pilot symbols used for channel estimation.

Sounding Signal. The EMOS uses an OFDM modulated sounding sequence with 256 subcarriers (out of which 160 are non-zero) and a cyclic prefix length of 64. One transmit frame is 64 OFDM symbols (2.667 ms) long and consists of a synchronization symbol (SCH), a broadcast data channel (BCH) comprising 7 OFDM symbols, a guard interval, and 48 pilot symbols used for channel estimation (see Fig. 3). The pilot symbols are taken from a pseudo-random QPSK sequence defined in the frequency domain. The subcarriers of the pilot symbols are multiplexed over the transmit antennas to ensure orthogonality in the spatial domain. We can therefore obtain

see the program

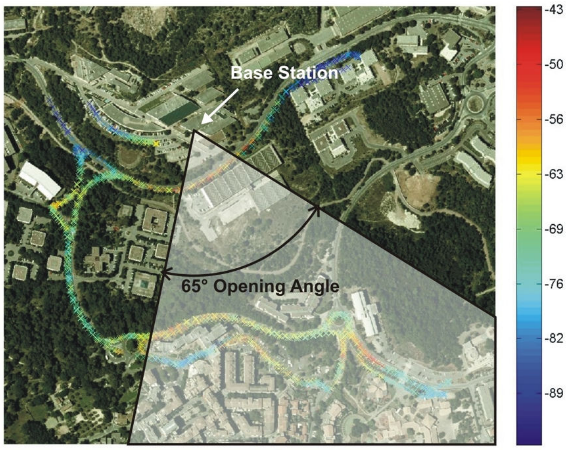


Fig. 4. Map of the measurement scenario. The position and the opening angle of the BS antenna are also indicated. The users were driving in cars along the indicated routes (the colors show the received signal strength in dBm along the routes).

Parameter	Value
Center Frequency	1917.6 MHz
Bandwidth	4.8 MHz
BS Transmit Power	30 dBm
Number of Antennas at BS	4 (2 cross polarized)
Number of UE	1
Number of Antennas at UE	2

Table 1. EMOS parameters.

one full MIMO channel estimate for one group of a number of subcarriers equal to the number of transmitter antennas. The BCH contains the frame number of the transmitted frame that is used for synchronization among the UEs.

Channel Estimation Procedure. Each UE first synchronizes to the BS using the SCH. It then tries to decode the data in the BCH. If the BCH can be decoded successfully, i.e., the cyclic redundancy check (CRC) is positive, then the channel estimation procedure is started. The channel estimation procedure consists of two steps. Firstly, the pilot symbols are derotated with respect to the first pilot symbol to reduce the phase-shift noise generated by the dual-RF CardBus/PCMCIA card. Secondly, the pilot symbols are averaged to increase the measurement SNR. The estimated MIMO channel is finally stored to disk. For a more detailed description of the synchronization and channel estimation procedure see [10, 11].

4.2. Measurements

The measurements were conducted outdoors in the vicinity of Eurecom in Sophia Antipolis, France[4]. The scenario is characterized by a semi-urban hilly terrain, composed by short buildings and vegetation with a predominantly present LOS. Fig. 4 shows a map of the environment. The BS is located at the roof of Eurecom's southmost building. The antenna is directed towards Garbejaire, a small nearby village. The measurement parameters are summarized in Table 1.

In this paper we use two different sets of measurements. In measurement 1, the UE was placed inside a standard passenger car which was being driven with an average speed of 50km/h along the routes shown in Fig. 4. The channel conditions are changing between line of sight (LOS) and non-LOS (NLOS). In measurement 2, the UE is more or less stationary on the parking lot in the bottom right corner of Fig. 4. This scenario is LOS.

5. REAL CHANNEL PERFORMANCE

In the simulations, we consider the particular real channel measured as commented in section 4 with 4 transmit and 2 receive antennas. Note that for the evaluations in this paper we have selected only one subcarrier to mimic a narrowband system. We show results for three cases: perfect CSI at the transmitter, non-differential Grassmannian packaging [4], and differential quantization of the channel correlation matrices $\mathbf{R}_H(n)$ using geodesic curves [3]. In all the cases, simulations were performed using the optimum strategies to maximize the mutual information and the SNR. The strategy that maximizes the mutual information corresponds to a waterfilling distribution of power over the eigenmodes of the channel, and the strategy that maximizes the SNR uses only the strongest eigenmode of the available channel response.

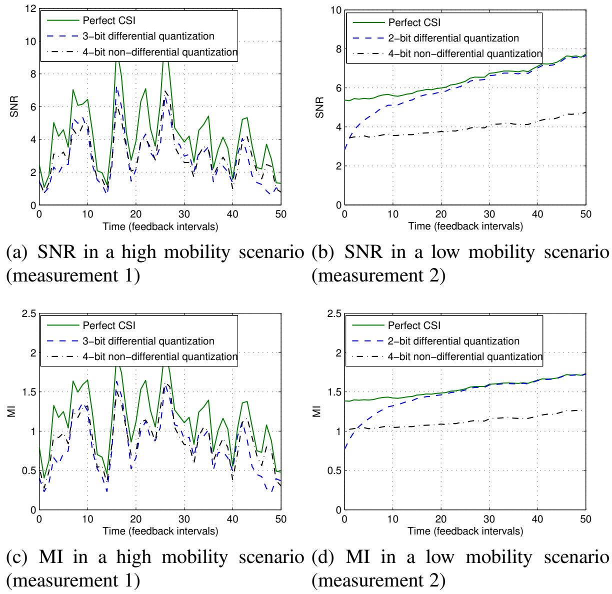
We considered two cases for the feedback. In the first case the quantized CSI is transmitted instantaneously from the receiver to the transmitter. That is, the transmitter had knowledge of the quantized version of the current channel matrix. In a real situation, however, the transmission delay through the feedback channel is not zero and this affects the performance of the system. Therefore we also studied the case where we introduce delay in the feedback channel.

5.1. Feedback with no delay

As shown in Fig. 5, the differential strategy exploits the time-correlation of the channel and converges to perfect CSIT case, while the performance using the non-differential quantization is lower, even when using more feedback bits. Also note that the differential quantization works better in more slow-varying channels and worse in the scenarios of high mobility

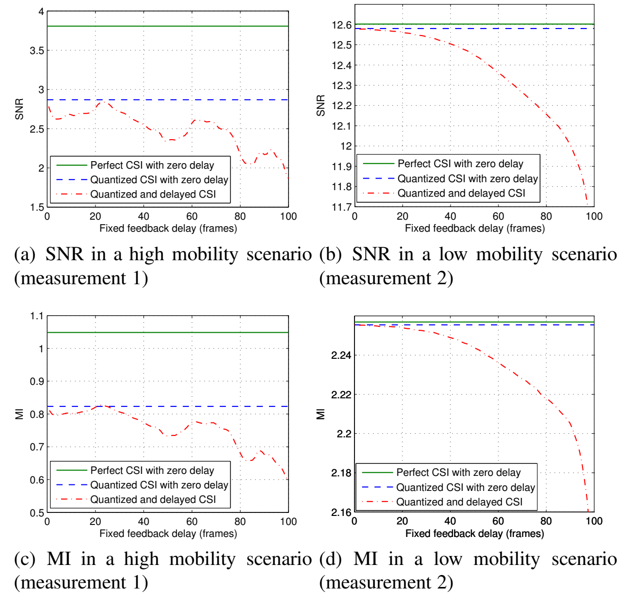
[4]Eurecom has a frequency allocation for experimentation around its premises.

see the program



(a) SNR in a high mobility scenario (measurement 1)
(b) SNR in a low mobility scenario (measurement 2)
(c) MI in a high mobility scenario (measurement 1)
(d) MI in a low mobility scenario (measurement 2)

Fig. 5. Different feedback techniques in realistic channels.



(a) SNR in a high mobility scenario (measurement 1)
(b) SNR in a low mobility scenario (measurement 2)
(c) MI in a high mobility scenario (measurement 1)
(d) MI in a low mobility scenario (measurement 2)

Fig. 6. Effects of feedback delay in realistic channels.

where the channel is fast-fading.

5.2. Delay in the feedback channel

The simulations corresponding to Fig. 6 analyze the impact of the feedback delay on the performance of the system. The plot shows the averaged SNR and mutual information (MI) for the high mobility and low mobility scenarios described in section 4 versus the delay measured in frames (e.g., a delay equal to 10 means that the delay is equal to 10 frames). For the simulations a window containing frames from 500 to 520 was used to calculate the average SNR and mutual information. Three situations are compared: perfect CSI at the transmitter, differential feedback with no delay, and differential feedback with different values for the delay in the feedback link. The main conclusion is that the performance rapidly decreases when the delay exceeds a threshold.

6. CHANNEL PREDICTION

In order to reduce or compensate the effect of feedback delay, channel prediction strategies can be used. If the receiver can predict the behavior of the channel response matrix and knows the value of the delay in the feedback channel, it is possible to send through the feedback link a quantized version of the prediction of the CSI. This way, the transmitter will receive the prediction of the current channel instead of the delayed CSI.

6.1. Algorithm description

The algorithm described in this section predicts each component of the channel response matrix $\mathbf{H}(n)$ separately using a Wiener predictor; this means that $n_R n_T$ predictors will be used. The objective of the Wiener predictor [12] is to minimize the quadratic error. The N predictor coefficients $\mathbf{g}_{ij}$ for the $_{ij}$ component of $\mathbf{H}(n)$ are defined as:

$$\mathbf{g}_{ij} = \arg\min_{\mathbf{g}_{ij}} \mathbb{E}[|e_{ij}(n)|^2] \in \mathbb{C}^N. \tag{10}$$

For a filter order N, and a prediction of L delay intervals, the error at time instant n is defined as

$$e_{ij}(n) = x_{ij}(n+L) - \mathbf{g}_{ij}^H \mathbf{x}_{ij}(n), \tag{11}$$

where the vector of samples $\mathbf{x}_{ij}(n)$ is

$$\mathbf{x}_{ij}(n) = \begin{bmatrix} x_{ij}(n) \\ x_{ij}(n-1) \\ \vdots \\ x_{ij}(n-N) \end{bmatrix} \in \mathbb{C}^N \text{and} \quad x_{ij}(n) = [\mathbf{H}(n)]_{ij}.$$

Then the quadratic error can be written as

$$\begin{aligned} |e_{ij}(n)|^2 = &\mathbf{g}_{ij}^H \mathbf{x}_{ij}(n) \mathbf{x}_{ij}^H(n) \mathbf{g}_{ij} + |x_{ij}(n+L)|^2 \\ &- \mathbf{g}_{ij}^H \mathbf{x}_{ij}(n) x_{ij}^*(n+L) - \mathbf{x}_{ij}^H(n) \mathbf{g}_{ij} x_{ij}(n+L). \end{aligned} \tag{12}$$

see the program

The predictor vector $\mathbf{g}_{ij}$ that minimizes the average of expression (12) can be easily proven to be [12]:

$$\mathbf{g}_{ij} = \left(\sum \mathbf{x}_{ij}(n)\mathbf{x}_{ij}^H(n)\right)^{-1}\left(\sum \mathbf{x}_{ij}(n)x_{ij}^*(n+L)\right), \tag{13}$$

where the sum is applied to the set of data that is available for the calculation of the predictor. This way, it is possible to predict the CSI with a delay of L time instants and design the transmitter based on the quantized version of the predicted CSI.

6.2. Simulations in real channel

The impact of channel prediction to compensate for the feedback delay will be analyzed in this section, using the realistic channel measured with EMOS. The same realistic channel studied in section 5 is used for this analysis. The simulations corresponding to Fig. 7 show the impact of prediction in a system with feedback delay. The plot shows the averaged SNR and mutual information for the high mobility and low mobility scenarios versus the delay measured in frames. For the simulations a window containing frames from 500 to 520 was used to calculate the average SNR and mutual information. Three situations are compared: perfect CSI at the transmitter, differential feedback with no delay, and differential feedback with different values for the delay in the feedback link using a Wiener predictor of order 10.

It can be seen from Fig. 7 that there is no substantial improvement when compared to the simulations without prediction from section 5.2. It would be interesting to observe the performance of the predictor in a channel that can be predicted more accurately, for example a channel that follows an autoregressive (AR) model. This case will be studied in the following section.

6.3. Simulations in an AR channel

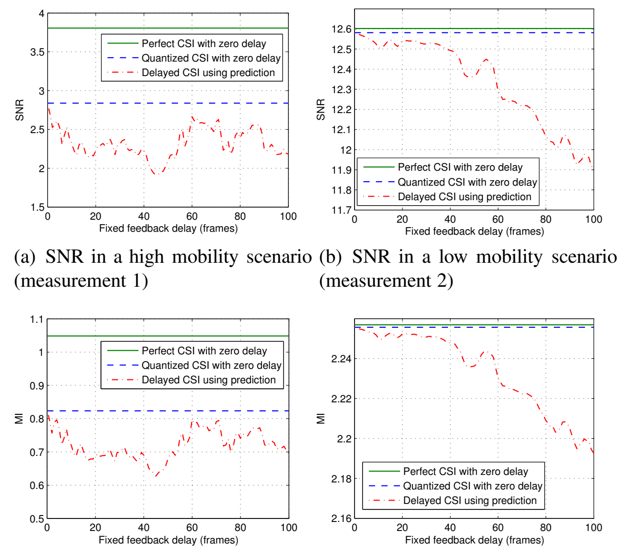
The predictor will now be tested on a synthetic channel generated using an autoregressive (AR) model of order Q. An AR model is described by the following equation [13]:

$$\mathbf{H}(n) = \sum_{q=1}^{Q} a_q \mathbf{H}(n-q) + \mathbf{W}(n), \tag{14}$$

where a_q are the autoregressive coefficients and the components of matrix $\mathbf{W}(n)$ are Gaussian, independent and with variance such that $\mathbb{E}\left[|[\mathbf{H}(n)]_{ij}|^2\right] = 1$.

Note that the predictor equals the coefficients of the AR model if the delay is 1 frame, but this is no longer the case for larger delays. In this last situation the coefficients should be computed as described in section 6.1.

In Fig. 8 we show the results of the differential feedback scheme applied to a channel generated using an AR model of order 10 with following AR coefficients: $a_1 = 1.4351 + 2.8308i$, $a_2 = 1.7196 - 2.9843i$, $a_3 = -0.9942 + 0.4452i$, $a_4 = -0.3001 - 0.6810i$, $a_5 = 0.1308 + 0.3598i$, $a_6 = -0.4145 + 0.0475i$, $a_7 = 0.1692 - 0.0565i$, $a_8 = -0.0661 + 0.3097i$, $a_9 = 0.1428 - 0.1975i$ and $a_{10} = -0.0702 + 0.0268i$. The simulations show that the predictor works best up to a delay of 10 frames, which corresponds to the order of the AR model. After that point, the gain decreases slowly.



(a) SNR in a high mobility scenario (measurement 1)

(b) SNR in a low mobility scenario (measurement 2)

(c) MI in a high mobility scenario (measurement 1)

(d) MI in a low mobility scenario (measurement 2)

Fig. 7. CSI prediction used to reduce the effect of feedback delay in realistic channels.

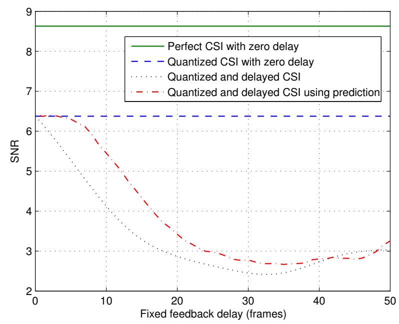
7. CONCLUSIONS

This paper has presented an evaluation of differential and non-differential feedback strategies in realistic MIMO systems. The main objective has been the study of the impact of such techniques using real channel measurements performed with the EMOS for high and low mobility scenarios and under different situations of delay in the feedback link.

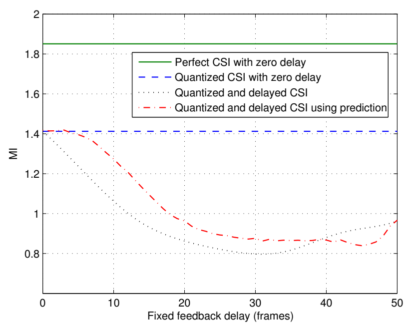
The differential feedback strategy performed much better than the non-differential strategies like Grassmannian packaging in low mobility channels, while in high mobility channels the performance was similar. Simulations using realistic channel data showed that a delay in the feedback channel affects specially high mobility channels because they vary faster. For small amounts of delay (less than 20 frames) the performance loss was around 10% in high mobility channels and less in the slow-varying cases.

The proposed technique to reduce the effect of feedback delay based on channel prediction performed well using a synthetic channel model. However, it was not able to compensate the delay in the measured channels. This showed that

see the program



(a) SNR performance



(b) Mutual Information performance

Fig. 8. CSI prediction used to reduce the effect of feedback delay in an AR(10) channel.

channel models are often too simplistic and do not provide realistic performance results.

8. REFERENCES

[1] J. Yang and D. Williams, "Transmission Subspace Tracking for MIMO Systems With Low-Rate Feedback," *IEEE Trans. on Communications*, vol. 55, no. 8, pp. 1629–1639, Aug. 2007.

[2] M. Talih, "Geodesic Markov Chains on Covariance Matrices," Statistical and Applied Mathematical Sciences Institute, Tech. Rep., March 2007. [Online]. Available: http://www.samsi.info/TR/tr2007-04.pdf

[3] D. Sacristán-Murga and A. Pascual-Iserte, "Differential Feedback of MIMO Correlation Matrices based on Geodesic Curves," in *Proc. IEEE International Conference on Accoustics, Speech and Signal Processing (ICASSP 2009)*, Taipei, Taiwan, April 2009, accepted.

[4] D. Love, R. Heath, and T. Strohmer, "Grassmannian Beamforming for Multiple-Input Multiple-Output Wireless Systems," *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[5] F. Kaltenberger, M. Kountouris, D. Gesbert, and R. Knopp, "Correlation and Capacity of Measured Multi-user MIMO Channels," in *Proc. IEEE Intl. Symp. on Personal, Indoor and Mobile Radio Comm. (PIMRC)*, Cannes, France, Sept. 2008.

[6] ——, "Performance of Multi-user MIMO Precoding with Limited Feedback over Measured Channels," in *Proc. IEEE Global Communications Conference (IEEE GLOBECOM 2008)*, New Orleans, USA, Nov.–Dec. 2008.

[7] D. Palomar, J. Cioffi, and M. Lagunas Hernández, "Joint Tx-Rx Beamforming Design for Multicarrier MIMO Channels: a Unified Framework for Convex Optimization," *IEEE Trans. on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sept. 2003.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[9] G. Golub and C. Van Loan, *Matrix Computations*. The Johns Hopkins University Press, Oct. 1996.

[10] R. de Lacerda, L. S. Cardoso, R. Knopp, M. Debbah, and D. Gesbert, "EMOS platform: real-time capacity estimation of MIMO channels in the UMTS-TDD band," in *Proc. International Symposium on Wireless Communication Systems (IWCS)*, Trondheim, Norway, Oct. 2007.

[11] F. Kaltenberger, L. Bernardo, and T. Zemen, "Characterization of measured multi-user mimo channels using the spectral divergence measure," COST 2100, Lille, France, Tech. Rep. TD(08) 640, Nov. 2008.

[12] N. Wiener, *The extrapolation, interpolation and smoothing of stationary time series with engineering applications*. Wiley, 1949.

[13] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Prentice Hall, 1997.

see the program

FILTER-AND-FORWARD DISTRIBUTED BEAMFORMING WITH INDIVIDUAL RELAY POWER CONSTRAINTS

Haihua Chen *Alex B. Gershman*

haihua.chen@nt.tu-darmstadt.de
gershman@nt.tu-darmstadt.de
Tel: +49-6151-162813
Communication Systems Group
Technische Universität Darmstadt
Merckstr. 25, Darmstadt, 64283 Germany

Shahram Shahbazpanahi

shahram.shahbazpanahi@uoit.ca
Tel: +1-905-7218668
Faculty of Eng. and Applied Science
Univ. of Ontario Institute of Technology
Oshawa, Ontario
LIH 7K4 Canada

ABSTRACT

A distributed beamforming technique for relay networks with frequency-selective fading channels is proposed. The network relays use the filter-and-forward (FF) strategy to compensate for the transmitter-to-relay and relay-to-destination channels using a finite impulse response (FIR) filter in each relay node. With the channel state information (CSI) known to the receiver, the destination quality-of-service (QoS) is maximized subject to individual relay transmitted power constraints. The resulting distributed beamforming problem is converted to a second-order cone programming (SOCP) form and, therefore, can be efficiently solved with polynomial time complexity. Simulation results demonstrate that in the frequency-selective fading case, the proposed beamforming technique provides substantial performance improvements in terms of the destination QoS as compared to the conventional amplify-and-forward (AF) distributed beamforming techniques.

***Index Terms*—** Cooperative communications, distributed beamforming, filter-and-forward relaying, relay networks

1. INTRODUCTION

Recently, cooperative wireless communication techniques have attracted much interest because they enable to exploit cooperative diversity without any need of having multiple antennas at each user [1]-[5]. In such cooperative schemes, different users share their resources to assist each other in conveying the information data through the network.

Different relaying strategies have been developed in the context of relay networks. The most popular relaying approaches are the amplify-and-forward (AF) and decode-and-forward (DF) techniques. Due to its simplicity, the AF technique is often preferred in distributed beamforming problems [3], [4].

This work was supported by the European Research Council under Advanced Investigator Grant Scheme.

In the case of frequency-selective transmitter-to-relay and relay-to-destination fading channels, there is a significant amount of inter-symbol interference (ISI). In such case, the AF distributed beamforming techniques developed in [3] and [4] are inefficient in suppressing ISI. In [6], a filter-and-forward (FF) strategy has been proposed to extend the AF approach to the frequency-selective fading case. The idea of the FF strategy is to use finite impulse response (FIR) filters in each relay to compensate for the effects of transmitter-to-relay and relay-to-receiver FIR channels. The relay beamforming problem considered in [6] is equivalent to minimizing the total relay transmitted power subject to the quality-of-service (QoS) constraint at the destination. However, because of a limited battery life of each relay, individual power criteria and/or constraints appear to be more meaningful than that operating with the total (aggregate) relay power. Moreover, the beamforming problem developed in [6] may be infeasible. In the present paper, we propose a FF relay beamforming approach alternative to [6] that maximizes the destination QoS subject to individual relay power constraints. The proposed approach can be shown to always correspond to a feasible problem.

2. SIGNAL MODEL

Let us consider a half-duplex relay network with one single-antenna transmitting source, one single-antenna destination and R single-antenna FF relay nodes, as shown in Fig. 1. We assume that there is no direct transmitter-destination link and each transmission consists of two stages. In the first stage, the transmitter sends its data to all the relay nodes. In the second stage, the received relay signals are passed through the relay FIR filters and sent to the destination node. We assume that the destination node has the full channel state information (CSI) which is used to obtain the filter coefficients of each relay according to a certain criterion. It is also assumed that there is a low-rate feedback link from the destination to each relay that is used to inform the relays about their optimal

see the program

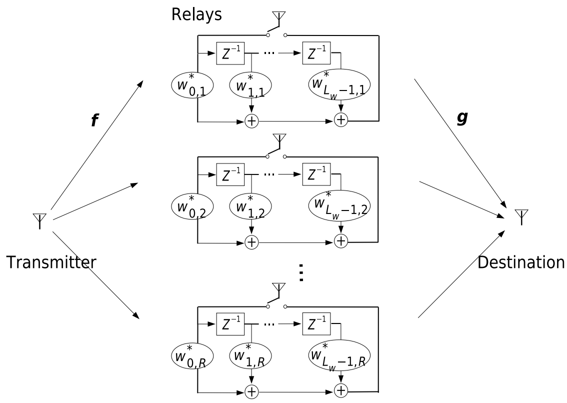


Fig. 1. System model for filter-and-forward relay network.

weight coefficients.

The transmitter-to-relay and relay-to-destination channels can be modeled as linear FIR filters [6]

$$\mathbf{f}(\omega) = \sum_{l=0}^{L_f-1} \mathbf{f}_l e^{-j\omega l}, \qquad \mathbf{g}(\omega) = \sum_{l=0}^{L_g-1} \mathbf{g}_l e^{-j\omega l} \tag{1}$$

respectively, where $\mathbf{f}_l$ and $\mathbf{g}_l$ are the $R \times 1$ channel impulse response vectors corresponding to the lth effective tap of the transmitter-to-relay and relay-to-destination channels, $\mathbf{f}(\omega)$ and $\mathbf{g}(\omega)$ are the $R \times 1$ vectors of channel frequency responses, and L_f and L_g are the channel lengths. The $R \times 1$ vector

$$\mathbf{r}(n) = [r_1(n), \cdots, r_R(n)]^T$$

of the received relay signals at time n can be written as

$$\mathbf{r}(n) = \sum_{l=0}^{L_f-1} \mathbf{f}_l s(n-l) + \boldsymbol{\eta}(n) \tag{2}$$

where $s(n)$ is the signal transmitted from the source,

$$\boldsymbol{\eta}(n) = [\eta_1(n), \cdots, \eta_R(n)]^T$$

is the $R \times 1$ vector of the relay noise, and $(\cdot)^T$ stands for the transpose. Let us introduce

$$\begin{aligned} \mathbf{F} &\triangleq [\mathbf{f}_0, \cdots, \mathbf{f}_{L_f-1}] \\ \mathbf{s}(n) &\triangleq [s(n), s(n-1), \cdots, s(n-L_f+1)]^T \end{aligned}$$

and rewrite (2) as [6]

$$\mathbf{r}(n) = \mathbf{F}\mathbf{s}(n) + \boldsymbol{\eta}(n). \tag{3}$$

The signal

$$\mathbf{t}(n) = [t_1(n), \cdots, t_R(n)]^T$$

transmitted by the relay nodes to the destination node can be expressed as [6]

$$\mathbf{t}(n) = \sum_{l=0}^{L_w-1} \mathbf{W}_l^H \mathbf{r}(n-l) \tag{4}$$

where

$$\mathbf{W}_l = \mathrm{diag}\{w_{l,1}, \cdots, w_{l,R}\}$$

is the diagonal matrix of relay filter impulse responses corresponding to the lth effective filter taps of the relays, L_w is the length of the relay FIR filters, and $(\cdot)^H$ is the Hermitian transpose. In the sequel, for any vector $\boldsymbol{x}$, $\mathrm{diag}\{\boldsymbol{x}\}$ denotes the diagonal matrix containing the entries of $\boldsymbol{x}$ on its main diagonal, and for any square matrix $\boldsymbol{X}$, $\mathrm{diag}\{\boldsymbol{X}\}$ denotes the vector built from the diagonal entries of $\boldsymbol{X}$. Let us introduce

$$\begin{aligned} \mathbf{W} &\triangleq [\mathbf{W}_0, \cdots, \mathbf{W}_{L_w-1}]^T \\ \boldsymbol{\mathcal{F}} &\triangleq \begin{bmatrix} \mathbf{f}_0 & \mathbf{f}_1 & \cdots & \mathbf{f}_{L_f-1} & \overbrace{\mathbf{0}_{R\times 1} \quad \cdots \quad \mathbf{0}_{R\times 1}}^{L_w-1 \text{ columns}} \\ \mathbf{0}_{R\times 1} & \mathbf{f}_0 & \mathbf{f}_1 & \cdots & \mathbf{f}_{L_f-1} \quad \cdots \quad \mathbf{0}_{R\times 1} \\ & & & \ddots & \\ \mathbf{0}_{R\times 1} & \mathbf{0}_{R\times 1} & \cdots & \mathbf{f}_0 & \mathbf{f}_1 \quad \cdots \quad \mathbf{f}_{L_f-1} \end{bmatrix} \\ \tilde{\mathbf{s}}(n) &\triangleq [s(n), s(n-1), \cdots, s(n-L_f-L_w+2)]^T \\ \tilde{\boldsymbol{\eta}}(n) &\triangleq [\boldsymbol{\eta}^T(n), \boldsymbol{\eta}^T(n-1), \cdots, \boldsymbol{\eta}^T(n-L_w+1)]^T \end{aligned}$$

where $\mathbf{0}_{N\times M}$ is the $N \times M$ matrix of zeros. Using these notations, we can express (4) as [6]

$$\mathbf{t}(n) = \mathbf{W}^H \boldsymbol{\mathcal{F}} \tilde{\mathbf{s}}(n) + \mathbf{W}^H \tilde{\boldsymbol{\eta}}(n). \tag{5}$$

The received signal at the destination can be written as

$$y(n) = \sum_{l=0}^{L_g-1} \mathbf{g}_l^T \mathbf{t}(n-l) + v(n) \tag{6}$$

where $v(n)$ is the receiver noise. Using (5) along with the fact that each $\mathbf{W}_l$ is diagonal, Equation (6) can be rewritten as [6]

$$\begin{aligned} y(n) = &\sum_{l=0}^{L_g-1} \mathbf{w}^H (\mathbf{I}_{L_w} \otimes \mathbf{G}_l) \boldsymbol{\mathcal{F}} \tilde{\mathbf{s}}(n-l) \\ &+ \sum_{l=0}^{L_g-1} \mathbf{w}^H (\mathbf{I}_{L_w} \otimes \mathbf{G}_l) \tilde{\boldsymbol{\eta}}(n-l) + v(n) \end{aligned} \tag{7}$$

where

$$\begin{aligned} \mathbf{w} &\triangleq [\mathbf{w}_0^T, \cdots, \mathbf{w}_{L_w-1}^T]^T \\ \mathbf{w}_l &\triangleq \mathrm{diag}\{\mathbf{W}_l\} \\ \mathbf{G}_l &\triangleq \mathrm{diag}\{\mathbf{g}_l\} \end{aligned}$$

$\mathbf{I}_N$ is the $N \times N$ identity matrix, and $\otimes$ denotes the Kronecker product. To rewrite (7) in a more compact form, let us

see the program

introduce

$$\begin{aligned}
&\boldsymbol{\mathcal{G}} \triangleq [\mathbf{I}_{L_w} \otimes \mathbf{G}_0, \cdots, \mathbf{I}_{L_w} \otimes \mathbf{G}_{L_g-1}] \\
&\breve{\mathbf{F}} \triangleq [\boldsymbol{\mathcal{F}}_0^T, \cdots, \boldsymbol{\mathcal{F}}_{L_g-1}^T]^T \\
&\boldsymbol{\mathcal{F}}_l \triangleq [\overbrace{\mathbf{0}_{RL_w\times 1}, \cdots, \mathbf{0}_{RL_w\times 1}}^{l \text{ columns}}, \boldsymbol{\mathcal{F}}, \overbrace{\mathbf{0}_{RL_w\times 1}, \cdots, \mathbf{0}_{RL_w\times 1}}^{(L_g-1-l) \text{ columns}}], \\
&\quad l = 0, \cdots, L_g - 1 \\
&\tilde{\mathbf{I}} \triangleq [\breve{\mathbf{I}}_0^T, \cdots, \breve{\mathbf{I}}_{L_g-1}^T]^T \\
&\breve{\mathbf{I}}_l \triangleq [\overbrace{\mathbf{0}_{RL_w\times R}, \cdots, \mathbf{0}_{RL_w\times R}}^{l \text{ blocks}}, \mathbf{I}_{RL_w}, \overbrace{\mathbf{0}_{RL_w\times R}, \cdots, \mathbf{0}_{RL_w\times R}}^{(L_g-1-l) \text{ blocks}}], \\
&\quad l = 0, \cdots, L_g - 1 \\
&\breve{\mathbf{s}}(n) \triangleq [s(n), s(n-1), \cdots, s(n-L_f-L_w-L_g+3)]^T \\
&\breve{\boldsymbol{\eta}}(n) \triangleq [\boldsymbol{\eta}^T(n), \boldsymbol{\eta}^T(n-1), \cdots, \boldsymbol{\eta}^T(n-L_w-L_g+2)]^T.
\end{aligned}$$

Using these notations, (7) can be expressed as [6]

$$y(n) = \mathbf{w}^H \boldsymbol{\mathcal{G}} \breve{\mathbf{F}} \breve{\mathbf{s}}(n) + \mathbf{w}^H \boldsymbol{\mathcal{G}} \tilde{\mathbf{I}} \breve{\boldsymbol{\eta}}(n) + \upsilon(n). \tag{8}$$

Let $\bar{\mathbf{f}}$ and $\bar{\mathbf{F}}$ denote the first column and the residue of $\breve{\mathbf{F}}$, respectively, so that $\breve{\mathbf{F}} = [\bar{\mathbf{f}}, \bar{\mathbf{F}}]$. Then, (8) can be written as

$$\begin{aligned}
y(n) &= \mathbf{w}^H \boldsymbol{\mathcal{G}} [\bar{\mathbf{f}}, \bar{\mathbf{F}}] \begin{bmatrix} s(n) \\ \bar{\mathbf{s}}(n) \end{bmatrix} + \mathbf{w}^H \boldsymbol{\mathcal{G}} \tilde{\mathbf{I}} \breve{\boldsymbol{\eta}}(n) + \upsilon(n) \\
&= \underbrace{\mathbf{w}^H \boldsymbol{\mathcal{G}} \bar{\mathbf{f}} s(n)}_{\text{signal}} + \underbrace{\mathbf{w}^H \boldsymbol{\mathcal{G}} \bar{\mathbf{F}} \bar{\mathbf{s}}(n)}_{\text{ISI}} + \underbrace{\mathbf{w}^H \boldsymbol{\mathcal{G}} \tilde{\mathbf{I}} \breve{\boldsymbol{\eta}}(n) + \upsilon(n)}_{\text{noise}} \quad (9)
\end{aligned}$$

where

$$\bar{\mathbf{s}}(n) \triangleq [s(n-1), \cdots, s(n-L_f-L_w-L_g+3)]^T.$$

In Equation (9),

$$\begin{aligned}
y_s(n) &= \mathbf{w}^H \boldsymbol{\mathcal{G}} \bar{\mathbf{f}} s(n) && (10) \\
y_i(n) &= \mathbf{w}^H \boldsymbol{\mathcal{G}} \bar{\mathbf{F}} \bar{\mathbf{s}}(n) && (11) \\
y_n(n) &= \mathbf{w}^H \boldsymbol{\mathcal{G}} \tilde{\mathbf{I}} \breve{\boldsymbol{\eta}}(n) + \upsilon(n) && (12)
\end{aligned}$$

are the signal, the ISI, and the noise components at the destination, respectively. Note that for the sake of simplicity of our algorithm developed in the next section, similarly to [6], we do not consider here *block processing* to coherently combine all the delayed copies of the signal.

The signal component in (10) can be further rewritten as [6]

$$\begin{aligned}
y_s(n) &= \mathbf{w}_0^H \mathbf{G}_0 \mathbf{f}_0 s(n) \\
&= \mathbf{w}_0^H (\mathbf{g}_0 \odot \mathbf{f}_0) s(n) \\
&= \mathbf{w}_0^H \mathbf{h}_0 s(n) && (13)
\end{aligned}$$

where

$$\mathbf{h}_0 \triangleq \mathbf{g}_0 \odot \mathbf{f}_0$$

and $\odot$ denotes the Schur-Hadamard matrix product.

3. FILTER-AND-FORWARD RELAY BEAMFORMING WITH INDIVIDUAL RELAY POWER CONSTRAINTS

The proposed distributed beamforming problem amounts to maximizing the destination QoS (measured in terms of signal-to-interference-plus-noise ratio (SINR)) subject to the individual relay transmitted power constraints:

$$\max_{\mathbf{w}} \text{SINR} \quad \text{s.t.} \quad p_m \le p_{m,\max}, \quad m = 1, \cdots, R \tag{14}$$

where p_m and $p_{m,\max}$ denote the transmitted power of the mth relay and its maximal allowed value, respectively. Using (5), the transmitted power of the mth relay can be written as

$$\begin{aligned}
p_m &= \mathrm{E}\{|t_m(n)|^2\} \\
&= \mathrm{E}\{\mathbf{e}_m^T \mathbf{W}^H \boldsymbol{\mathcal{F}} \tilde{\mathbf{s}}(n) \tilde{\mathbf{s}}^H(n) \boldsymbol{\mathcal{F}}^H \mathbf{W} \mathbf{e}_m\} \\
&\quad + \mathrm{E}\{\mathbf{e}_m^T \mathbf{W}^H \tilde{\boldsymbol{\eta}}(n) \tilde{\boldsymbol{\eta}}^H(n) \mathbf{W} \mathbf{e}_m\} \\
&= P_s \mathbf{e}_m^T \mathbf{W}^H \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{F}}^H \mathbf{W} \mathbf{e}_m + \sigma_\eta^2 \mathbf{e}_m^T \mathbf{W}^H \mathbf{W} \mathbf{e}_m \quad (15)
\end{aligned}$$

where $\mathbf{e}_m$ is the mth column of the identity matrix and $\mathrm{E}\{\cdot\}$ denotes the statistical expectation. In (15), it is assumed that the transmitted signal and the relay noise are statistically independent and

$$\begin{aligned}
\mathrm{E}\{\tilde{\mathbf{s}}(n) \tilde{\mathbf{s}}^H(n)\} &= P_s \mathbf{I}_{L_f+L_w-1} \\
\mathrm{E}\{\tilde{\boldsymbol{\eta}}(n) \tilde{\boldsymbol{\eta}}^H(n)\} &= \sigma_\eta^2 \mathbf{I}_{RL_w}
\end{aligned}$$

where P_s is the source transmitted power and σ_η^2 is the relay noise variance.

Using $\mathbf{E}_m \triangleq \mathrm{diag}\{\mathbf{e}_m\}$, Equation (15) yields

$$\begin{aligned}
p_m &= P_s \mathbf{w}^H (\mathbf{I}_{L_w} \otimes \mathbf{E}_m) \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{F}}^H (\mathbf{I}_{L_w} \otimes \mathbf{E}_m)^H \mathbf{w} \\
&\quad + \sigma_\eta^2 \mathbf{w}^H (\mathbf{I}_{L_w} \otimes \mathbf{E}_m) (\mathbf{I}_{L_w} \otimes \mathbf{E}_m)^H \mathbf{w} \\
&= \mathbf{w}^H \mathbf{D}_m \mathbf{w} && (16)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{D}_m \triangleq\ & P_s (\mathbf{I}_{L_w} \otimes \mathbf{E}_m) \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{F}}^H (\mathbf{I}_{L_w} \otimes \mathbf{E}_m)^H \\
& + \sigma_\eta^2 (\mathbf{I}_{L_w} \otimes \mathbf{E}_m) (\mathbf{I}_{L_w} \otimes \mathbf{E}_m)^H.
\end{aligned}$$

The SINR at the destination can be written as

$$\text{SINR} = \frac{\mathrm{E}\{|y_s(n)|^2\}}{\mathrm{E}\{|y_i(n)|^2\} + \mathrm{E}\{|y_n(n)|^2\}}. \tag{17}$$

Using (13), we obtain that

$$\begin{aligned}
\mathrm{E}\{|y_s(n)|^2\} &= \mathrm{E}\{|\mathbf{w}_0^H \mathbf{h}_0 s(n)|^2\} \\
&= P_s \mathbf{w}_0^H \mathbf{h}_0 \mathbf{h}_0^H \mathbf{w}_0 \\
&= P_s \mathbf{w}^H \mathbf{A}^H \mathbf{h}_0 \mathbf{h}_0^H \mathbf{A} \mathbf{w} \\
&= \mathbf{w}^H \mathbf{Q}_s \mathbf{w} && (18)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{A} &\triangleq [\mathbf{I}_R, \mathbf{0}_{R\times(L_w-1)R}] \\
\mathbf{Q}_s &\triangleq P_s \mathbf{A}^H \mathbf{h}_0 \mathbf{h}_0^H \mathbf{A}.
\end{aligned}$$

see the program

Using (11), we have

$$\begin{aligned}\mathrm{E}\{|y_i(n)|^2\} &= \mathrm{E}\{\mathbf{w}^H \boldsymbol{\mathcal{G}} \bar{\mathbf{F}} \bar{\mathbf{s}}(n) \bar{\mathbf{s}}^H(n) \bar{\mathbf{F}}^H \boldsymbol{\mathcal{G}}^H \mathbf{w}\} \\ &= P_s \mathbf{w}^H \boldsymbol{\mathcal{G}} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\mathcal{G}}^H \mathbf{w} \\ &= \mathbf{w}^H \mathbf{Q}_i \mathbf{w} \end{aligned} \tag{19}$$

where

$$\mathbf{Q}_i \triangleq P_s \boldsymbol{\mathcal{G}} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \boldsymbol{\mathcal{G}}^H.$$

Making use of (12), we also obtain

$$\begin{aligned}\mathrm{E}\{|y_n(n)|^2\} &= \mathrm{E}\{\mathbf{w}^H \boldsymbol{\mathcal{G}} \tilde{\mathbf{I}} \tilde{\boldsymbol{\eta}}(n) \tilde{\boldsymbol{\eta}}^H(n) \tilde{\mathbf{I}}^H \boldsymbol{\mathcal{G}}^H \mathbf{w}\} + \sigma_v^2 \\ &= \sigma_\eta^2 \mathbf{w}^H \boldsymbol{\mathcal{G}} \tilde{\mathbf{I}} \tilde{\mathbf{I}}^H \boldsymbol{\mathcal{G}}^H \mathbf{w} + \sigma_v^2 \\ &= \mathbf{w}^H \mathbf{Q}_n \mathbf{w} + \sigma_v^2 \end{aligned} \tag{20}$$

where

$$\mathbf{Q}_n \triangleq \sigma_\eta^2 \boldsymbol{\mathcal{G}} \tilde{\mathbf{I}} \tilde{\mathbf{I}}^H \boldsymbol{\mathcal{G}}^H.$$

Using (16) along with (18)-(20) and introducing an auxiliary variable $\tau > 0$, the problem in (14) can be expressed as

$$\begin{aligned} \max_{\mathbf{w},\tau} \quad & \tau \\ \text{s.t.} \quad & \frac{\mathbf{w}^H \mathbf{Q}_s \mathbf{w}}{\mathbf{w}^H \mathbf{Q}_i \mathbf{w} + \mathbf{w}^H \mathbf{Q}_n \mathbf{w} + \sigma_v^2} \geq \tau^2 \\ & \mathbf{w}^H \mathbf{D}_m \mathbf{w} \leq p_{m,\max}, \quad m = 1, \cdots, R. \end{aligned} \tag{21}$$

The first constraint in (21) is equivalent to

$$\sqrt{P_s} |\mathbf{w}^H \mathbf{h}| \geq \tau \sqrt{\mathbf{w}^H \mathbf{Q}_i \mathbf{w} + \mathbf{w}^H \mathbf{Q}_n \mathbf{w} + \sigma_v^2} \tag{22}$$

where $\mathbf{h} \triangleq \mathbf{A}^H \mathbf{h}_0$. We observe that any arbitrary phase rotation of $\mathbf{w}$ does not change the value of the objective function in (21). Therefore, using a proper rotation, (22) can be written as

$$\sqrt{P_s} \mathbf{w}^H \mathbf{h} \geq \tau \sqrt{\mathbf{w}^H \mathbf{Q}_i \mathbf{w} + \mathbf{w}^H \mathbf{Q}_n \mathbf{w} + \sigma_v^2}. \tag{23}$$

Let

$$\begin{aligned} \mathbf{B} &\triangleq \begin{bmatrix} \sigma_v^2 & \mathbf{0}_{RL_w \times 1}^T \\ \mathbf{0}_{RL_w \times 1} & \mathbf{Q}_i + \mathbf{Q}_n \end{bmatrix} = \mathbf{U}^H \mathbf{U} \\ \mathbf{D}_m &= \mathbf{V}_m^H \mathbf{V}_m, \quad m = 1, \cdots, R \end{aligned}$$

be the Cholesky factorizations of the matrices $\mathbf{B}$ and $\mathbf{D}_m$, respectively. Introducing new notations

$$\breve{\mathbf{w}} \triangleq [1, \mathbf{w}^T]^T, \quad \breve{\mathbf{V}}_m \triangleq [\mathbf{0}_{RL_w \times 1}, \mathbf{V}_m], \quad \breve{\mathbf{h}} \triangleq [0, \mathbf{h}^T]^T$$

we can rewrite (21) as

$$\begin{aligned} \max_{\breve{\mathbf{w}},\tau} \quad & \tau \\ \text{s.t.} \quad & \sqrt{P_s} \breve{\mathbf{w}}^H \breve{\mathbf{h}} \geq \tau \|\mathbf{U} \breve{\mathbf{w}}\| \\ & \|\breve{\mathbf{V}}_m \breve{\mathbf{w}}\| \leq \sqrt{p_{m,\max}}, \quad m = 1, \cdots, R \\ & \breve{\mathbf{w}}^H \mathbf{e}_1 = 1 \end{aligned} \tag{24}$$

where $\mathbf{e}_1 \triangleq [1, 0, 0, \cdots, 0]^T$

In contrast to the distributed beamforming problem considered in [6], problem (24) is always feasible. The latter problem is also quasi-convex [7], because for any value of τ it reduces to the following feasibility problem:

$$\begin{aligned} \text{find} \quad & \breve{\mathbf{w}} \\ \text{s.t.} \quad & \sqrt{P_s} \breve{\mathbf{w}}^H \tilde{\mathbf{h}} \geq \tau \|\mathbf{U} \breve{\mathbf{w}}\| \\ & \|\breve{\mathbf{V}}_m \breve{\mathbf{w}}\| \leq \sqrt{p_{m,\max}}, \quad m = 1, \cdots, R \\ & \breve{\mathbf{w}}^H \mathbf{e}_1 = 1. \end{aligned} \tag{25}$$

The latter feasibility problem can be identified as a second-order cone programming (SOCP) problem.

Let τ_* be the optimal objective value of (24). For any $\tau > \tau_*$, problem (25) is infeasible. Otherwise, it contradicts to the fact that τ_* is the optimal objective value. On the contrary, if (25) is feasible, then we have $\tau \leq \tau_*$. Hence, the maximum τ_* and the optimal weight vector $\breve{\mathbf{w}}$ can be found using bisection search technique discussed in [4]. Assuming that the maximum value τ_* lies in $[\tau_l, \tau_u]$, the bisection search procedure for solving (24) can be summarized as the following sequence of steps:

1. $\tau := (\tau_l + \tau_u)/2$.
2. Solve the convex feasibility problem (25). If (25) is feasible, then $\tau_l := \tau$, otherwise $\tau_u := \tau$.
3. If $(\tau_u - \tau_l) < \varepsilon$ then stop. Otherwise, go to step 1.

Here, ε is the error tolerance value.

The feasibility problem (25) is a standard SOCP problem which can be solved in polynomial time using modern interior point methods with the worst-case complexity of $\mathcal{O}((RL_w)^{3.5})$ [8].

4. SIMULATIONS

In our simulations, we consider a relay network with $R = 10$ relays and transmitter-to-relay and relay-to-destination channels of the length five. The channel coefficients are modeled as zero-mean complex Gaussian random values with exponential power delay profile. The relay and destination noises are assumed to have the same powers, and the source transmitted power is assumed to be 10 dB higher than the noise power. The maximal allowed relay transmitted powers are assumed to be equal to each other, that is, $p_{m,\max} = p_{\max}$ for all $m = 1, \ldots, R$.

Fig. 2 displays the destination SINR versus the maximal allowed relay transmitted power $p_{\max}$ for different lengths of the relay filters. Fig. 3 plots the destination SINR versus the length of the relay filter for different values of the maximal allowed relay transmitted power. From both these figures, it can be observed that the destination QoS improves significantly as the length of the relay filters increases. In the trivial case

see the program

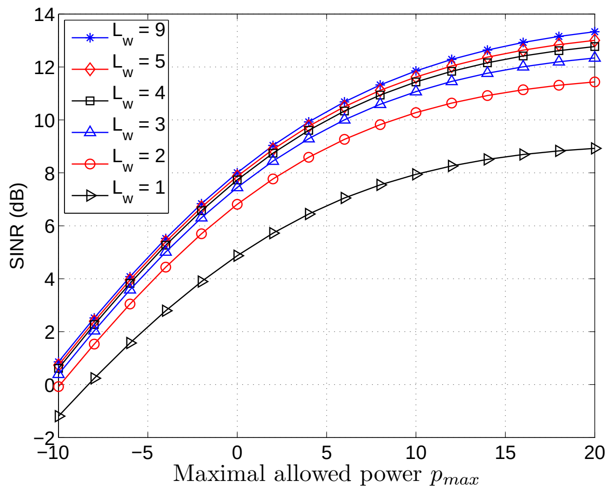


Fig. 2. Destination SINR versus the maximal allowed relay transmitted power.

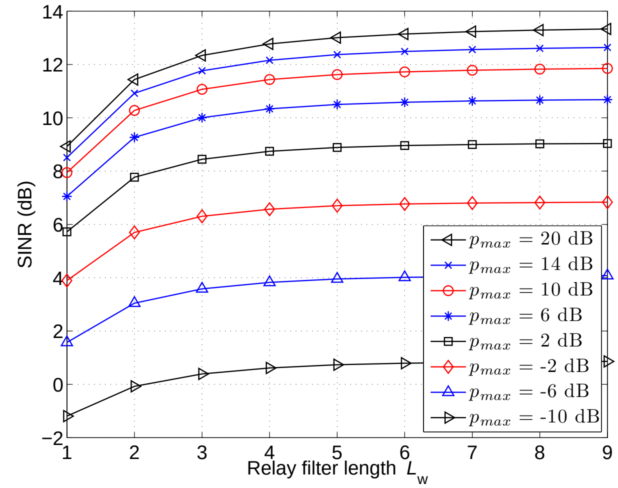


Fig. 3. Destination SINR versus the length of relay filter.

of $L_w = 1$, the FF strategy reduces to the conventional AF approach and, correspondingly, the proposed method boils down to that of [4]. It can be clearly seen from Figs. 2 and 3 that using the FF relaying strategy in the frequency-selective fading case, the performance of distributed beamforming can be substantially improved as compared to the conventional AF approach of [4].

5. CONCLUSIONS

In this paper, we have proposed a novel distributed beamforming approach for cooperative relay networks with frequency-selective fading transmitter-to-relay and relay-to-destination channels. To compensate for the effects of the channels, we have employed the filter-and-forward relaying strategy. In our approach, the destination quality-of-service is optimized subject to individual relay power constraints. It has been shown that this distributed beamforming problem can be solved using second-order cone programming and bisection search based techniques. Our simulations confirm that the proposed approach has substantial performance improvements over the existing amplify-and-forward relay beamforming techniques.

6. REFERENCES

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062-3080, Dec. 2004.

[2] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3037-3063, Sept. 2005.

[3] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *Proc. ICASSP'07*, vol. 3, pp. 473-476, Honolulu, HI, Apr. 2007.

[4] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Processing*, vol. 56, pp. 4306-4316, Sept. 2008.

[5] S. Fazeli-Dehkordy, S. Gazor, and S. Shahbazpanahi, "Distributed peer-to-peer multiplexing using ad hoc relay networks," *Proc. ICASSP'08*, vol. 2, pp. 2373-2376, Las Vegas, NV, Apr. 2008.

[6] H. H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming for relay networks in frequency selective fading channels," *Proc. ICASSP'09*, Taipei, April 2009, to appear.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[8] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones,"*Optimization Methods and Software*, Special issue on Interior Point Methods, vol. 11/12, pp. 625-653, 1999.

see the program

Handover Sequences for Interference-Aware Transmission in Multicell MIMO Networks

Volker Jungnickel, Konstantinos Manolakis, Lars Thiele, Thomas Wirth and Thomas Haustein
Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut
Einsteinufer 37, 10587 Berlin, Germany
jungnickel@hhi.de

***Abstract*—Providing cell-specific reference signals is a basic requirement for advanced transmission techniques reducing the inter-cell interference in the mobile radio network. In this paper, we introduce so-called handover sequences. They consist of a comb cyclically shifted in the frequency domain to identify the cells. Orthogonal sequences in time domain are used to identify the antennas within a cell. The scheme occupies few OFDM symbols per coherence interval. Sequence assignment to the cells follows a classical frequency-reuse scheme. Detection is easily implementable. With these sequences, the frequency-selective multi-cell channel can be identified with high precision also at the cell edge. We demonstrate that handover decisions are more reliable and channel estimation errors are reduced by more than a decade compared to the reference signals provided in 3GPP LTE Release 8.**

I. Introduction

Interference-aware transmission covers a new class of multi-antenna techniques reducing the inter-cell interference and further enhancing the spectral efficiency in the mobile radio access network. These new techniques can be classified into independent transmission at each base station and cooperative transmission of multiple base stations.

The independent approach has been thoroughly investigated in the multi-cell down-link scenario, refer to [1]. Explicit knowledge of the channel to the best server and to the strongest other base stations is assumed and exploited twice. At the physical layer, optimum combining is used [2]. At the medium access control layer (MAC) layer, knowing the multi-cell channel enables a precise, i.e. frequency-selective, estimation of the signal to interference and noise ratio (SINR). SINR versus frequency vectors after fixed beam forming, wireless channel and optimal combining for single and multiple streams in the cell are quantized and fed back to the base station. Knowledge about the interference improves the performance of fair scheduling algorithms [3], [4]. A striking advantage is the frequent use of multi-stream transmission even at the cell edge.

The cooperative mode receives more attention recently [5], [6]. There is a clear trend to organize it in a distributed manner [7]. Base station cooperation requires instantaneous exchange of channel state information obtained via feedback from the terminals and transferred between the base stations over a meshed backbone network with low latency. In this way, each base station gets global channel knowledge. Moreover, data are synchronously distributed to all cooperative base stations and the joint beam-forming is steered locally [8].

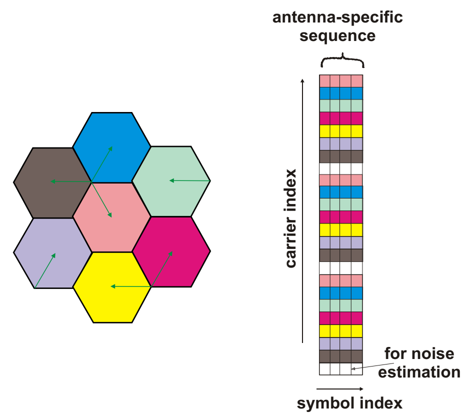


Fig. 1. **Left: Elementary interference scenario. Right: Cells are identified using a comb in frequency domain individually shifted in each cell. Antennas within each cell are identified using time-domain sequences.**

In this paper, we focus on the identification of multiple cells using specific reference signals. We illustrate the advantages of the proposed sequence design for handover decisions and multi-cell channel estimation.

II. Requirements and previous work

Interference-aware transmission techniques have two requirements in common: 1. Synchronized base stations. 2. Multi-cell reference signals. While synchronization can be realized with moderate effort [9], here we focus on the reference signals.

Release 8 of 3GPP LTE has not paid much attention to these requirements. LTE base stations are not necessarily synchronized. Cell-specific reference signals are defined (primary and secondary synchronization sequence, PSS and SSS) but they cover only part of the entire system bandwidth. Scrambling of channel estimation pilots is cell-specific along the frequency axis and it allows an estimation of the wide-band-averaged SINR in a cell. However, there are no explicit reference signals for estimating the multi-cell channel in a frequency-selective manner [10].

A first proposal of multi-cell pilots is reported in [11]. All cells use pilots at the same positions in the time-frequency grid. Cells are identified by scrambling in the time domain, i.e. there is no additional overhead. Proposed sequences have partial correlation properties, where certain subsets of sequences

see the program

are orthogonal already over a shorter correlation window. The idea is to assign sequences with shorter correlation windows to closer base stations and those with longer windows to more distant ones. From implementation point of view a large number of potential cells must be tracked simultaneously, which can be rather complex. Sequence assignment in the cellular network is not trivial but manageable. We feel that these virtual pilots are well suited for a rapid adaptation to the interference channel at the receiver side, using techniques such as optimum combining. Note that only a few interfering base stations must be actually tracked to realize most of the gain at the terminal side [12].

III. HANDOVER SEQUENCE DESIGN

In this paper we propose a second set of sequences allowing simpler implementation with almost instantaneous results. These sequences are not transmitted continuously but at a certain period to get a full overview of the multi-cell channel and to identify the most relevant interfering signals in a single shot. Such a design costs additional overhead and it may be suitable for handover decisions, evaluation of the SINR and feedback of the channel state information to the base station. The simple idea is to make cells orthogonal in the frequency domain and to exploit therefore the correlations in the frequency-selective channel. Classical frequency-reuse schemes are then applied to assign sequences to cells.

In a single cell, a handover sequence shall not interfere with another such sequence after passing through the multi-path channel. Definition in the frequency domain is then appropriate. We use a regularly spaced grid of active sub-carriers spanning the entire system bandwidth. For scrambling in the frequency domain, we use a Zadoff-Chu sequence with the same root index as defined in [10] but having a period adjusted to the total number of active sub-carriers in the sequence. Adjacent cells are identified by applying a cell-specific frequency shift of integer multiples of the sub-carrier spacing, see Fig. 1.

Further use of these sequences for interference-aware transmission is intended. Therefore we have to satisfy the fundamental sampling theorem of the channel. Interpolation in the frequency domain is possible if the total number of sub-carriers in the grid is not smaller than the number of resolved multi-path components L in the channel. In OFDM systems, an upper bound for L is given by the length of the cyclic prefix L_{CP}. In case of an 3GPP LTE system with 20 MHz bandwidth, more than $L_{CP} = 144$ pilot tones, scattered over 1.200 sub-carriers, are needed. Accordingly, each 8^{th} sub-carrier is active in the comb.

Next we apply a cyclic shift of the entire comb in the frequency domain by N_{shift} sub-carriers where the shift identifies the cell. With a grid spacing of 8, up to 8 cells can be identified. Consider the elementary interference scenario and the sequence assignment in Fig. 1. Each of the 7 cells is identified by a certain shift $N_{shift} = 0 \ldots 6$. Note that 7 is a magic number in the hexagonal deployment. It allows the reuse of the sequence in the second one after the next cell where the interference is significantly reduced due to path-loss and down-tilt of the antennas. Reuse of sequences and the resulting interference scenario for the pilots can be characterized by the frequency reuse factor $F = 7$. Sub-carriers at $N_{shift} = 7$ are not used at all. Such empty resources allow a precise noise estimation.

In modern cellular systems, we need to identify multiple antennas in each cell in addition. Therefore we concatenate a few OFDM symbols having the same comb structure and the same cyclic shift in one cell. On a given sub-carrier in the comb, an antenna-specific reference sequence is transmitted along the time domain over as many OFDM symbols as there are antennas in the cell. Estimation is based on a cross-correlation along the time axis on each sub-carrier in the comb. This technique is widely used for the high-throughput long training field (HT-LTF) in the IEEE 802.11n wireless LAN standard.

IV. PERFORMANCE EVALUATION

We have evaluated the performance in two ways. Firstly, we consider the handover fail rate, i.e. the probability that a false cell is detected. Secondly, we examine the channel estimation error in the presence of interference. Our evaluation is based on the 3GPP LTE Release 8 specification. At the receiver, the terminal applies OFDM symbol synchronization using the cyclic prefix. The radio frame start is then identified using the primary synchronization sequence (PSS). Next the terminal applies an FFT of the incoming sequences from multiple base stations.

A. Handover fail rate

The narrow-band PSS in LTE Release 8 is jointly transmitted with data in other parts of the spectrum while the comb is exclusively transmitted over the entire spectrum. For handover detection, autocorrelation with the cell-specific Zhadoff-Chu sequences is applied in the frequency domain and the peak magnitudes are compared. As a reference indicating the strongest cell, in the simulation we have separately measured the power received from each base station (which is not possible in practice, of course). From the autocorrelation peaks, the decision is derived in the presence of interference and compared to the reference. The fail rate is given by the percentage of wrong decisions. It is a monotonous function of the signal-to-interference ratio (SIR). The higher the SIR is the more likely handover decisions are correct.

B. Channel estimation error

For coarse channel estimation, cross-correlations along the time axis in Fig. 1 are applied with all antenna-specific sequences and at each receive antenna. Next the channels are interpolated in the frequency domain using a linear MMSE estimator. Interpolation in the time domain is not included in results. As an upper bound, perfect knowledge of both the signal to noise ratio (SNR) and the power delay profile is assumed for building the interpolation matrix.

see the program

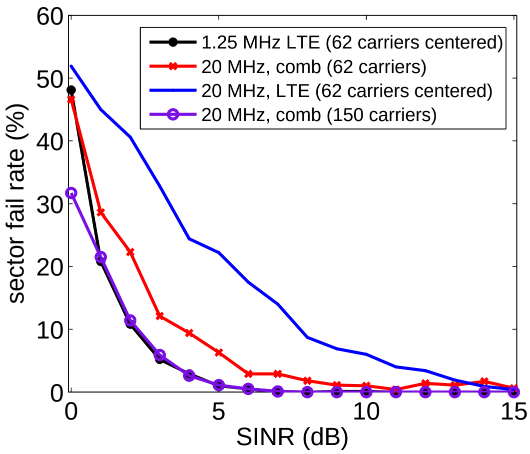


Fig. 2. **Reliability of handover decisions.**

We are most interested in comparing the estimation performance with LTE Release 8 where other cells transmit data on pilot positions in the desired cell. This is equivalent to having a frequency reuse factor of $F = 1$ during the channel estimation. Estimating the channel from an adjacent cell is interfered by the stronger data signals in the serving cell. The estimation error for negative SIR thus gives an impression of the corresponding performance degradation. For the comb sequences, there is no interference from data signals in other cells. Interference comes from the pilots of rather distant cells using the same cyclic shift for their comb. For evaluating the estimation performance, a frequency reuse factor of $F = 7$ is appropriate.

V. Results

A. Handover fail rate

We have tested the reliability of handover decisions in 1.25 and 20 MHz LTE systems using 3 cells at one site. A terminal is moved across the border between two sectors and the third sector points away. Random multi-path channels are used and instantaneous powers are considered.

At first we have considered the 62 sub-carriers wide PSS in [10], where other resource blocks (RBs) are filled with data. The second prototype is a comb with 62 sub-carriers in 20 MHz with data on intermediate sub-carriers. The original Zadoff-Chu sequence [10] is spread over the comb. Thirdly, the comb with 8 sub-carriers spacing is used where the presence of other base stations is emulated by random data signals on intermediate sub-carriers. Results for the handover fail rate are given in Fig. 2.

The fail rates r can be empirically modeled as

$$r = a * exp\left(-\frac{SIR}{SIR_0}\right) \tag{1}$$

with the parameters

System	a [%]	$SIR_0[dB]$
1.25 MHz LTE	50	1.8
20 MHz LTE	52	5
20 MHz 62 carriers	45	2.5
20 MHz 150 carriers	32	1.8

Good performance is achieved in the 1.25 MHz system where almost the full system bandwidth is also covered by the PSS. But if the system bandwidth is larger than the fixed PSS bandwidth [10], e.g. in the 20 MHz system, the performance gets worse. In 35 % of cases, a false server may be identified although the best cell has 3 dB more power. This is attributed to the frequency-selective fading temprorarily causing a lower-than-average power level in the limited bandwidth tested by the PSS. A good handover sequence should therefore probe the entire signal bandwidth. This is fulfilled for the comb. With 20 MHz and 62 subcarriers in the comb, the performance is already close to the 1.25 MHz case. With 20 MHz and 150 subcarriers in the comb, reliablility is further improved. Only in 5 % of cases the false cell is identified as the server if the best cell has 3 dB more power.

B. Channel estimation error

The channel estimation performance has been quantified for a 20 MHz LTE system where a single OFDM symbol contains 200 pilots. Note that the mean-square error (MSE) depends on the channel length. We can approximate the MSE as

$$MSE = \frac{G}{SIR} \tag{2}$$

where G denotes the estimator gain. In a multi-path Rayleigh fading channel with L independently and identically distributed taps, G is given by

$$G = \frac{N_{sc}}{d * L} \tag{3}$$

where N_{sc} is the total number of subcarriers and d the pilot spacing. The actual value of G can be read off from graphs like in Fig. 3 at $SNR = 0\ dB$ also for more realistic channel models. We have considered an exponential power decay profile over the entire cyclic prefix with a decay time of 2.3 μs to model a strongly shadowed NLOS channel which is a rather likely at the cell edge. Note that the kinks in the graphs come from using interpolation matrices in coarse SNR steps of 5 dB only. The resulting estimator gain G is obtained at 5 dB and 6 dB for LTE and the comb, respectively, due to the slightly smaller pilot density. The estimator gain is of course enhanced for shorter channels.

In the multi-cell environment, the channel estimation error follows the SIR statistics in general. The user experiences different SIR at different positions in the cell. Accordingly, the estimation error statistics can be modeled as

$$\rho_1(MSE) = \rho\,(SIR) * \left|\frac{dSIR}{dMSE}\right|. \tag{4}$$

Using the model for the MSE from above we obtain

see the program

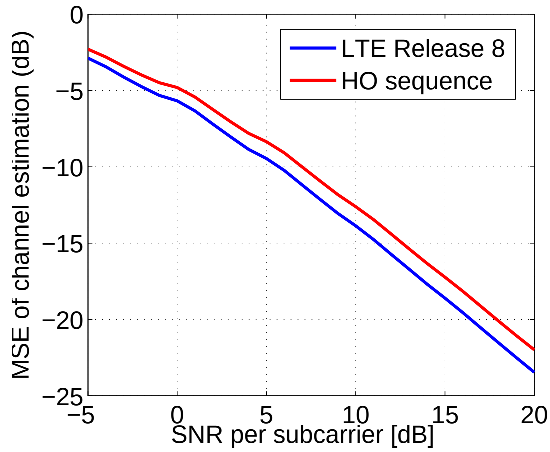


Fig. 3. **Channel estimation performance**

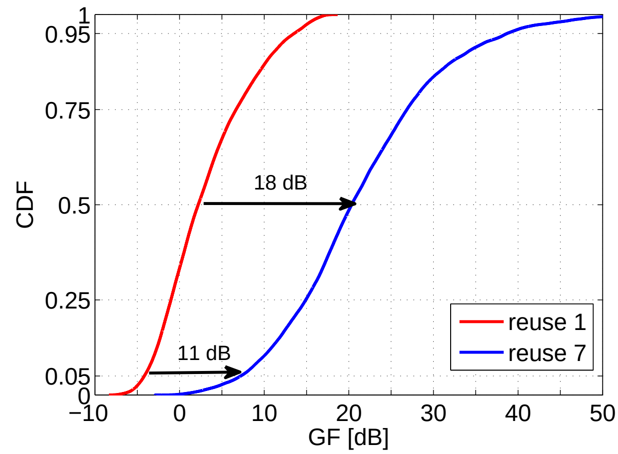


Fig. 4. **Geometry factor for frequency reuse 1 and 7.**

$$\rho_1(MSE) = \rho\left(\frac{G}{MSE}\right) * \frac{G}{MSE^2}. \tag{5}$$

The statistics depends significantly on the frequency reuse factor F. This causes a striking difference between the LTE pilot structure and the comb. A frequently used measure for the carrier-to-interference situation within a cell is the geometry factor

$$GF = \frac{\mathbb{E}(P_{signal})}{\mathbb{E}(P_{interference})}. \tag{6}$$

In Fig. 4, the geometry factor is shown for $F = 1$ and $F = 7$ in a hexagonal deployment with 57 sectors where each base station has three sectors. Obviously, the SIR is much higher with frequency-reuse, and thus the channel estimation is enhanced. From Fig. 4, we may read off the 5 and 50 percentiles of the cumulative geometry statistics to estimate the MSE at the cell edge and on average in the cell, respectively. Corresponding values are -2 dB and -8 dB for LTE Release 8 using $F = 1$ and -12 dB and -25 dB, respectively if the proposed comb is used with $F = 7$.

VI. Conclusions

We have proposed cell-specific reference signals which can be used for both handover decisions and multi-cell channel estimation. The proposed sequences consist of a comb in the frequency domain. Shifting the comb by integer multiples of the sub-carrier spacing identifies the cell. The frequency response of the channel can be estimated below a maximum spacing in the comb depending on OFDM symbol duration and cyclic prefix length. For sequence assignment in a full-coverage cellular network, classical frequency-reuse concepts can be used up to a reuse factor equal to the spacing. The handover fail rate is reduced in this way and the precision of the channel estimation is enhanced by 10 dB and 17 dB at the cell edge and on average in the cell, respectively, using a comb spacing of 8 and a frequency reuse factor of 7.

VII. Acknowledgments

This authors are grateful for financial support by the German Ministry of Education and Research (BMBF) in the national research project EASY-C under contract 01BU0631.

References

[1] L. Thiele, M. Schellmann, W. Zirwas, and V. Jungnickel, “Capacity scaling of multi-user MIMO with limited feedback in a multi-cell environment,” in *41st Asilomar Conference on Signals, Systems and Computers*. Monterey, USA: IEEE, Nov. 2007, invited.

[2] J. Winters, “Optimum combining in digital mobile radio with cochannel interference,” *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 4, pp. 528–539, 1984.

[3] L. Thiele, M. Schellmann, T. Wirth, and V. Jungnickel, “Interference-aware scheduling in the synchronous cellular multi-antenna downlink,” in *69th IEEE Vehicular Technology Conference, 2009. VTC 2009-Spring.*, May 2009, invited.

[4] V. Jungnickel, M. Schellmann, L. Thiele, T. Wirth, T. Haustein, O. Koch, W. Zirwas, and E. Schulz, “Interference aware scheduling in the multiuser mimo-ofdm downlink,” *IEEE Communications Magazine*, June 2009.

[5] P. Baier, M. Meurer, T. Weber, and H. Troger, “Joint transmission (jt), an alternative rationale for the downlink of time division cdma using multi-element transmit antennas,” in *Spread Spectrum Techniques and Applications, 2000 IEEE Sixth International Symposium on*, vol. 1, 2000, pp. 1–5 vol.1.

[6] M. Karakayali, G. Foschini, and R. Valenzuela, “Network coordination for spectrally efficient communications in cellular systems,” *Wireless Communications, IEEE*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[7] W. Zirwas, E. Schulz, J. H. Kim, V. Jungnickel, and M. Schubert, “Distributed organization of cooperative antenna systems,” in *European Wireless, Athens, Greece*, Apr. 2006.

[8] V. Jungnickel, L. Thiele, M. Schellmann, T. Wirth, A. Forck, W. Zirwas, T. Haustein, and E. Schulz, “Implementation concepts for distributed cooperative transmission,” *42nd Asilomar Conference on Signals, Systems and Computers*, Oct. 2008.

[9] V. Jungnickel, T. Wirth, M. Schellmann, T. Haustein, and W. Zirwas, “Synchronization of cooperative base stations,” in *IEEE International Symposium on Wireless Communication Systems 2008 (ISWCS08)*, Reykjavik, Iceland, Oct. 2008.

[10] 3GPP TS 36.211 V8.0.0, “E-UTRA - physical channels and modulation (release 8),” Sept. 2008.

[11] L. Thiele, M. Schellmann, S. Schiffermller, and V. Jungnickel, “Multi-cell channel estimation using virtual pilots,” in *IEEE 67th Vehicular Technology Conference VTC2008-Spring*, Singapore, May 2008.

[12] L. Thiele and V. Jungnickel, “Adaptive transmission in a realistic multicell scenario,” in *IEEE Radio and Wireless Symposium*, Jan. 2008.

see the program

MULTI-GROUP MULTICAST BEAMFORMING FOR MULTI-USER TWO-WAY RELAYING

Aditya Umbu Tana Amah[1], *Anja Klein*[1], *Yuri C. B. Silva*[2], *Andreas Ferneke߁*[1]

[1]Communications Engineering Lab, Technische Universität Darmstadt, Darmstadt, Germany
{a.amah, a.klein, a.fernekess}@nt.tu-darmstadt.de
[2]Wireless Telecommunications Research Group - GTEL, Federal University of Ceará, Fortaleza, Brazil
yuri@gtel.ufc.br

ABSTRACT

In this work, we consider a multi-user two-way relaying protocol. A multi-antenna relay station (RS) serves $2K$ nodes where K pairs of nodes would like to perform bidirectional communication. In the first phase, the $2K$ nodes transmit simultaneously and the RS spatially separates and decodes all the $2K$ bit sequences. In the second phase, after performing bit-wise XOR network coding for each bidirectional pair, which results in only K bit sequences, the RS sends all the K XOR-ed bit sequences simultaneously. At each node, having received its intended XOR-ed bit sequence, it cancels the self-interference by XOR-ing its received bit sequence with its transmitted one to obtain the bit sequence sent by its partner. Considering the second phase, as each node in a specific pair expects the same XOR-ed bit sequence from the RS, while seeing other pairs' XOR-ed bit sequences as interference, the RS has to separate the different pairs' XOR-ed bit sequences spatially. Thus, in the second phase, multi-group multicast beamforming can be applied with the node pairs being the multicast group. With this new perspective on the second phase, we propose to apply multi-group multicast beamforming algorithms which provide fairness to all pairs and have low computational complexity. We investigate several multi-group multicast beamforming algorithms, namely Zero Forcing (ZF), Multicast Aware ZF (MAZF) and SINR Balancing with Bisection Search (SINRB-Bisec), for the second phase transmission of the multi-user two-way relaying. We consider also two different XOR network coding approaches, namely with and without zero padding (ZP). The overall two-phase sum rate analysis is given which shows that both MAZF and SINRB-Bisec have similar performance and outperform ZF. Nevertheless, MAZF has lower computational complexity compared to SINRB-Bisec.

The work of Aditya U. T. Amah is supported by the 'Excellence Initiative' of the German Federal and State Governments and the Graduate School of Computational Engineering, Technische Universität Darmstadt

1. INTRODUCTION

In two-hop communication, a relay station (RS) assists the communication between a source node and a destination node. Due to the half-duplex constraint, two orthogonal resources are needed, one resource for the transmission from the source node to the RS and another one for the transmission from the RS to the destination node. This leads to a loss in capacity by a factor of 2 compared to single-hop communication [1, 2].

One way to mitigate this capacity loss is by allowing the RS to assist bidirectional communication between two nodes with the so-called two-way relaying protocol [2]. The communication in two-way relaying is performed in two phases [2]. In the first phase, the multiple access (MAC) phase, both nodes transmit at the same time to the RS. In the second phase, the broadcast (BC) phase, the RS forwards the superimposed signal to both nodes. Since each node a priori knows its own transmitted signal, it can subtract it from the received signal to obtain the signal from the other node.

At the RS, different types of signal processing can be performed. In Amplify and Forward (AF), the RS forwards the received signal after amplifying it [2–5]. In Decode and Forward (DF), the RS decodes the information of both nodes and forwards the re-encoded information [2, 6, 7].

Two-way relaying with two multi-antenna nodes and a multi-antenna RS has been treated in [8, 9] for AF and in [10] for DF. DF two-way relaying has the advantage that the noise at the RS is not propagated to the receive nodes. Its performance depends on the coding which is used to re-encode the information in the second phase. In [10, 11], for the DF multi-antenna two-way relaying case, it has been shown that for the re-encoding of the decoded information of both nodes, bit-wise XOR coding outperforms superposition coding.

The aforementioned contributions considered a scenario where an RS assists two nodes (single pair) to perform bidirectional communication with each other. Recently, two-way relaying for multi-user scenarios has attracted more attention and has been considered in [12–14]. In [12], a multi-antenna RS assists a Base Station to perform bidirectional communication with multiple nodes. A different multi-user scenario is treated in [13, 14] where two sets of nodes perform two-

see the program

way relaying communication with the assistance of an RS. Here, each node in each set communicates bidirectionally to only one specific node in the other set. While [13] assumes a single antenna RS which separates the users using Code Division Multiple Access, a multi-antenna RS is assumed in [14] to separate the users spatially. The work in [14] considers a DF protocol using XOR network coding where an optimization algorithm for the precoding matrix to maximize the overall two-phase sum rate is proposed and possible extensions to max-min fairness optimization are addressed.

In this work, we also consider multi-user two-way relaying where the multi-antenna DF RS serves multiple pairs of nodes using spatial separation and applies XOR network coding as in [14]. In the first phase, all nodes transmit simultaneously to the RS and the RS spatially separates and decodes the bit sequences of all nodes. In the second phase, the RS performs XOR network coding to each pair's bit sequences, which halves the number of bit sequences which are transmitted. We consider two different XOR network coding approaches, namely with and without zero padding (ZP). The use of XOR network coding at the RS needs self-interference cancellation at each node. Each node obtains its partner's bit sequence by XOR-ing its received bit sequence with its transmitted one.

In the second phase, each node in a specific pair, due to the XOR-ing of two bit sequences of the two nodes from the first phase at the RS, expects the same XOR-ed bit sequence from the RS. Thus, the RS multicasts the same bit sequence to this pair. However, the RS has to send all bit sequences to all pairs simultaneously and, consequently, each pair sees other pairs' bit sequences as interference. Since there are multiple bit sequences for multiple pairs, the bit sequences transmission for the second phase can be performed using multi-group multicast beamforming. With this new perspective on the second phase, we can apply multi-group multicast beamforming solutions known from single-hop transmission in the two-way relaying case.

Optimizing the precoding matrix to maximize the sum rate of multi-user two-way relaying as in [14] may lead to unfairness as the RS may distribute more power to pairs with good channel conditions and, consequently, cause very low data rate to other pairs. Thus, we propose to apply multi-group multicast beamforming that provides fairness to the pairs. For the second phase we investigate Multicast Aware Zero Forcing (MAZF) which was introduced in [15] for single-hop communication and consider as well the SINR balancing multi-group multicast beamforming with Bisection search (SINRB-Bisec) as introduced in [16] for single-hop. As computational complexity is a practical issue, we investigate as well the Zero Forcing (ZF) multi-group multicast beamforming given in [15] for single-hop.

This paper is organized as follows. Section 2 provides the system and signal model of the multi-user two-way relaying. The multi-group multicast beamforming algorithms under consideration are explained in Section 3. Section 4 gives the achievable overall two-phase sum rate. The performance analysis is given in Section 5. Finally, Section 6 provides the conclusion.

Throughout this paper, boldface lower case and upper case letters denote vectors and matrices, respectively, while normal letters denote scalar values. The superscripts $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathrm{H}}$ stand for matrix or vector transpose, complex conjugate, and complex conjugate transpose, respectively. The operators $\mathrm{E}\{\mathbf{X}\}$ and $\mathrm{tr}\{\mathbf{X}\}$ denote the expectation and the trace of $\mathbf{X}$, respectively, and $\mathcal{CN}(0, \sigma^2)$ denotes the zero-mean complex normal distribution with variance σ^2.

2. SYSTEM AND SIGNAL MODEL

We consider a multi-user two-way relaying scenario where a multi-antenna RS supports multiple single antenna node pairs. Each pair of nodes would like to perform an exclusive bidirectional communication. It is assumed that there exists no direct link between the two nodes in each pair, such that the bidirectional communication may happen only through the assistance of an RS. The K pairs, which consist of $N = 2K$ single antenna nodes, will be served by a multi-antenna RS which has $M \geq N$ antennas. A number $M \geq N$ of antennas at the RS is required as we need to decode N bit sequences in the first phase and we are using multi-group multicast beamforming addressed to N users for the second phase.

The RS applies a DF protocol. In the first phase, all N nodes transmit simultaneously to the RS and the RS decodes the bit sequences of all nodes. It then performs bit-wise XOR network coding to each pair's bit sequences. Thus, at the RS, there exist K XOR-ed bit sequences which need to be transmitted to K different pairs. In the second phase, to separate the pairs spatially from one another, the RS uses multi-group multicast beamforming and forwards each of the K bit sequences to its intended pair. Each node performs self-interference cancellation by XOR-ing its received bit sequence with its transmitted one to obtain the bit sequence from its bidirectional communication partner.

In the following, let us put each node S_n, $n \in \{1, \cdots, N\}$, into one of two sets of nodes, namely the set of odd numbered nodes $\mathcal{S}_{\mathrm{odd}}$ and the set of even numbered nodes $\mathcal{S}_{\mathrm{even}}$. Each node in $\mathcal{S}_{\mathrm{odd}}$ would like to perform bidirectional communication using the two-way relaying protocol with a specific node in $\mathcal{S}_{\mathrm{even}}$, where each node in each set can only have one bidirectional partner node in the other set. Let $G^{(k)}$, $k \in \{1, \cdots, K\}$ denote the k-th pair of nodes consisting of nodes $\mathrm{S}_{a_k} \in \mathcal{S}_{\mathrm{odd}}$ and $\mathrm{S}_{b_k} \in \mathcal{S}_{\mathrm{even}}$, where $\{a_k \in \mathbb{Z}_+ | a_k = 2k - 1,\ k \in \{1, \cdots, K\}\}$ and $\{b_k \in \mathbb{Z}_+ | b_k = 2k,\ k \in \{1, \cdots, K\}\}$.

Each node S_n wants to send to its partner a bit sequence denoted as $\mathbf{x}_n$. In the first phase, all N nodes transmit simultaneously, i.e., S_1 sends $\mathbf{x}_1$, S_2 sends $\mathbf{x}_2$, and so on. The RS decodes all the N nodes' bit sequences. In the second phase,

see the program

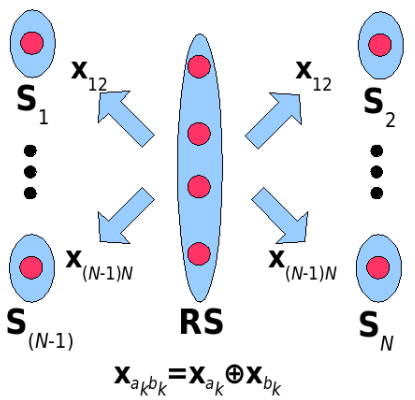


Fig. 1. Example of multi-group multicast beamforming and the XOR Network Coding for multi-user two-way relaying.

the RS performs bit-wise XOR operation to the bit sequences of each pair. The XOR operation of pair $G^{(k)}$ is defined by $\mathbf{x}_{a_k b_k} = \mathbf{x}_{a_k} \oplus \mathbf{x}_{b_k}$, for example $\mathbf{x}_{12} = \mathbf{x}_1 \oplus \mathbf{x}_2$. All the K XOR-ed bit sequences of all pairs are transmitted simultaneously by the RS using multi-group multicast beamforming. All nodes perform self-interference cancellation by XOR-ing the received information with their transmitted information to obtain the information being sent by their partner, for example S_1 performs $\mathbf{x}_{12} \oplus \mathbf{x}_1$ to receive $\mathbf{x}_2$ from node S_2. Figure 1 shows an example how the N nodes are divided into two sets, $\mathcal{S}_{\text{odd}}$ on the left and $\mathcal{S}_{\text{even}}$ on the right, and how the RS forwards the K bit sequences by spatially separating each pair using multi-group multicast beamforming.

In the following, we assume a frequency-flat block fading channel between all nodes and the RS where $\mathbf{h}_n \in \mathbb{C}^{M\times 1}$ denotes the channel vector from node S_n to the RS with identically and independently distributed $\mathcal{CN}(0, \sigma^2_{\mathbf{h}_n})$ entries. Assuming reciprocity and stationarity of the channel for both phases, the channel vector from the RS to node S_n is given by $\mathbf{h}_n^{\mathrm{T}}$.

2.1. First Phase - Multiple Access Phase

In the first phase, all nodes transmit simultaneously to the RS. The mapping process of the bit sequence of node S_n onto its corresponding transmitted symbol is defined by $\mathbf{x}_n \rightarrow t_n \in \mathbb{C}$. Each node S_n has power constraint $\mathrm{E}\{|t_n|^2\} \leq P_n, \forall n, n \in \{1, \cdots, N\}$. The received vector at the RS is given by

$$\mathbf{r}_{\mathrm{RS}} = \sum_{k=1}^{K} (\mathbf{h}_{a_k} \cdot t_{a_k} + \mathbf{h}_{b_k} \cdot t_{b_k}) + \mathbf{n}_{\mathrm{RS}}, \tag{1}$$

which can be rewritten as

$$\mathbf{r}_{\mathrm{RS}} = \mathbf{H} \cdot \mathbf{t} + \mathbf{n}_{\mathrm{RS}}, \tag{2}$$

where the channel matrix $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_N] \in \mathbb{C}^{M\times N}$ and the transmit vector $\mathbf{t} = [t_1, t_2, \cdots, t_N]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$ are the matrix consisting of all nodes' channel vectors and the vector consisting of all nodes' transmit symbols, respectively, and $\mathbf{n}_{\mathrm{RS}} \in \mathbb{C}^{M\times 1} \sim \mathcal{CN}(0, \sigma^2_{\mathrm{RS}}\mathbf{I}_M)$ is the complex additive white Gaussian noise (AWGN) vector at the RS, with $\mathbf{I}_M$ denoting an identity matrix of size $M \times M$.

Assuming that the RS has the ability to perform perfect decoding, there exist N bit sequences of all N nodes at the RS, i.e., $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N$.

2.2. Second Phase - Broadcast Phase

Having $N = 2K$ bit sequences available, the RS performs XOR operation of the bit sequences $\mathbf{x}_{a_k}$ and $\mathbf{x}_{b_k}$ of pair $G^{(k)}$, $\forall k, k \in \{1, \cdots, K\}$ resulting in $\mathbf{x}_{a_k b_k} = \mathbf{x}_{a_k} \oplus \mathbf{x}_{b_k}$. This XOR operation reduces the number of bit sequences to be transmitted from $2K$ to K, which means halving the number of interfering bit sequences for the BC phase. The next task for the RS is to forward these K bit sequences to the K pairs such that each pair can receive its intended bit sequence while receiving only small interference from unintended bit sequences. Multi-group multicast beamforming algorithms are used to spatially separate the pairs (and consequently the bit sequences), which will be explained in Section 3.

Assuming reciprocal and stationary channels for both phases and having transmit symbol $t_{a_k b_k} \in \mathbb{C}$ as the result of the mapping process $\mathbf{x}_{a_k b_k} \rightarrow t_{a_k b_k}$, the received vector of nodes S_{a_k} and S_{b_k} in pair $G^{(k)}$ can be written as

$$\begin{aligned} r_{a_k} &= \mathbf{h}_{a_k}^{\mathrm{T}} \cdot \mathbf{m}_k \cdot t_{a_k b_k} + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{h}_{a_k}^{\mathrm{T}} \cdot \mathbf{m}_l \cdot t_{a_l b_l} + n_{a_k} \\ &\text{and} \\ r_{b_k} &= \mathbf{h}_{b_k}^{\mathrm{T}} \cdot \mathbf{m}_k \cdot t_{a_k b_k} + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{h}_{b_k}^{\mathrm{T}} \cdot \mathbf{m}_l \cdot t_{a_l b_l} + n_{b_k}, \end{aligned} \tag{3}$$

where $\mathbf{m}_k \in \mathbb{C}^{M\times 1}$ and $\mathbf{m}_l \in \mathbb{C}^{M\times 1}$ are the precoding vectors for pairs $G^{(k)}$ and $G^{(l)}, \forall l, l \in \{1, \cdots, K\}, l \neq k$, respectively, and $n_n \sim \mathcal{CN}(0, \sigma_n^2)$ is the AWGN at node n. In matrix formulation, the received vector is given by

$$\mathbf{r}_{\text{nodes}} = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{M}_{\text{nodes}} \cdot \mathbf{t}_{\mathrm{RS}} + \mathbf{n}_{\text{nodes}}, \tag{4}$$

where $\mathbf{t}_{\mathrm{RS}} = [t_{12}, t_{34}, \cdots, t_{(N-1)N}]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ is the transmitted symbols vector at the RS, $\mathbf{n}_{\text{nodes}} \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_N)$ is the stacked vector of the complex AWGN at all nodes and $\mathbf{M}_{\text{nodes}} = [\mathbf{m}_1, \cdots, \mathbf{m}_K] \in \mathbb{C}^{M\times K}$ is the precoding matrix according to the multi-group multicast beamforming algorithm under consideration. It is assumed that the transmitted symbols are identically and independently distributed with $\mathrm{E}\{\mathbf{t}_{\mathrm{RS}}\mathbf{t}_{\mathrm{RS}}^{\mathrm{H}}\} = \sigma^2_{\mathbf{t}_{\mathrm{RS}}}\mathbf{I}_K$ and the RS has a power constraint defined by $\sigma^2_{\mathbf{t}_{\mathrm{RS}}}\mathrm{tr}(\mathbf{M}_{\text{nodes}}\mathbf{M}_{\text{nodes}}^{\mathrm{H}}) \leq P_{\mathrm{RS}}$.

see the program

After each node decodes its intended bit sequence, it performs self-interference cancelation by applying XOR operation of its received bit sequence and its a priori transmitted bit sequence to obtain its partner's bit sequence, such that at node S_{a_k}, $\mathbf{x}_{b_k} = \mathbf{x}_{a_k b_k} \oplus \mathbf{x}_{a_k}$ and at node S_{b_k}, $\mathbf{x}_{a_k} = \mathbf{x}_{a_k b_k} \oplus \mathbf{x}_{b_k}$ is obtained.

3. MULTI-GROUP MULTICAST BEAMFORMING

The capacity region for the MAC phase has been well investigated and may be found in the literature such as in [22]. In this work, we concentrate on the BC phase of multi-user two-way relaying, for which we propose the use of multi-group multicast beamforming for the second phase of multi-user two-way relaying. As we have multiple pairs, multi-group multicast beamforming algorithms which provide fairness to all pairs are desirable. An optimization objective which promotes fairness and is adequate for multicast beamforming is the maximization of the minimum SINR among the nodes [15].

In this work, we extend several single-hop multi-group multicast beamforming algorithms to be applied to the multi-user two-way relaying scenario. As computational complexity is a practical issue, we investigate ZF multi-group multicast beamforming, which was introduced in [15] for single-hop communication. For fairness consideration, we investigate the so-called MAZF, which provides fairness by performing fair power loading to all nodes and was proposed in [15] for single-hop communication and the SINRB-Bisec, which tries to solve the aforementioned optimization objective and was proposed in [16, 17] for single-hop communication.

In the following, we will only briefly explain these three algorithms, starting with ZF followed by MAZF and SINRB-Bisec. For more details, the interested reader is refered to the references directly. The corresponding precoding matrix $\mathbf{M}_{\text{nodes}}$, which is needed in (4), is derived for the three algorithms in the subsequent subsections.

3.1. Zero Forcing

The ZF optimization problem for multi-group multicast beamforming [15] can be rewritten for multi-user two-way relaying scenario as

$$\mathbf{M}_{\text{ZF}} = \underset{\mathbf{M}}{\operatorname{argmin}}\, \mathrm{E}\{\|\mathbf{r}_{\text{nodes}} - \mathbf{U}^{+}\mathbf{t}_{\text{RS}}\|^2\},$$
$$\text{subject to}: \quad \mathrm{E}\{\|\mathbf{M}\mathbf{t}_{\text{RS}}\|^2\} \leq P_{\text{RS}} \quad , \qquad (5)$$
$$\mathbf{r}_{\text{nodes}}|_{\mathbf{n}_{\text{nodes}}=\mathbf{0}} = \mathbf{U}^{+}\mathbf{t}_{\text{RS}}$$

where the second constraint corresponds to the ZF constraint and $\mathbf{U}^{+} \in \mathbb{R}^{N\times K}$ has its a_k-th and b_k-th rows given by $\mathbf{u}^{+}_{a_k} = \mathbf{u}^{+}_{b_k} = \mathbf{e}_k$, with $\mathbf{e}_k$ corresponding to the k-th column of an identity matrix of dimension K. For example, the $\mathbf{U}^{+}$ matrix for two pairs, $G^{(1)}$ and $G^{(2)}$, which consist of four nodes, S_1, S_2, S_3 and S_4, is given by $\mathbf{U}^{+} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$.

Using the optimization procedure in [18], the solution of (5) is given by

$$\mathbf{M}_{\text{ZF}} = \beta \mathbf{H}^{*}(\mathbf{H}^{\mathrm{T}}\mathbf{H}^{*})^{-1}\mathbf{U}^{+}, \qquad (6)$$

where the scalar factor

$$\beta = \sqrt{\frac{P_{\text{RS}}}{\sigma_{\mathbf{t}_{\text{RS}}} \mathrm{tr}((\mathbf{H}^{\mathrm{T}}\mathbf{H}^{*})^{-1}\mathbf{U}^{+}\mathbf{U}^{+\mathrm{T}})}} \qquad (7)$$

is needed to fulfill the power constraint in (5).

In ZF multi-group multicast beamforming, the interference is cancelled by inverting the gram matrix of the channel. Thus, energy is unnecessarily spent on cancelling interference of the own pair's partner.

3.2. Multicast Aware Zero Forcing

The idea of MAZF is to avoid unnecessarily energy spending to cancel the interference of the own group's partner and to provide fairness to the nodes [15]. The MAZF uses Block Diagonalization [19,20] to suppress only the interference among data streams of different pairs and then applies a fair power loading to all nodes. By having a Block Diagonalization, only the interference of other pairs' channels is cancelled and not of the pair partner's channel. Thus, the MAZF provides higher power when transmitting to the intended node compared to the aforementioned ZF. In the following, the extension to the multi-user two-way relaying scenario is given.

Let $\mathbf{H}_k^{\mathrm{T}} \in \mathbb{C}^{2\times M}$ and $\tilde{\mathbf{H}}_k^{\mathrm{T}} \in \mathbb{C}^{(N-2)\times M}$ denote the channel matrix of pair $G^{(k)}$ and the channel matrix of all other pairs $G^{(l)}$, $\forall l,\ l \in \{1, \cdots, K\},\ l \neq k$, respectively. The latter can be written as

$$\tilde{\mathbf{H}}_k^{\mathrm{T}} = [\mathbf{H}_1^{\mathrm{T}}, \cdots, \mathbf{H}_{k-1}^{\mathrm{T}}, \mathbf{H}_{k+1}^{\mathrm{T}}, \cdots, \mathbf{H}_K^{\mathrm{T}}]^{\mathrm{T}}. \qquad (8)$$

The channel matrix $\tilde{\mathbf{H}}_k^{\mathrm{T}}$ can be decomposed using Singular Value Decomposition as follows:

$$\tilde{\mathbf{H}}_k^{\mathrm{T}} = \tilde{\mathbf{U}}_k \tilde{\mathbf{S}}_k [\tilde{\mathbf{V}}_k^{(1)}, \tilde{\mathbf{V}}_k^{(0)}], \qquad (9)$$

where $\tilde{\mathbf{S}}_k \in \mathbb{R}^{(N-2)\times M}$ is a diagonal matrix, a unitary matrix given by $\tilde{\mathbf{U}}_k \in \mathbb{C}^{(N-2)\times(N-2)}$, $\tilde{\mathbf{V}}_k^{(1)} \in \mathbb{C}^{M\times \tilde{r}_k}$ and $\tilde{\mathbf{V}}_k^{(0)} \in \mathbb{C}^{M\times(N-\tilde{r}_k)}$ contain the right singular vectors of $\tilde{\mathbf{H}}_k^{\mathrm{T}}$, with $\tilde{r}_k$ denoting the rank of matrix $\tilde{\mathbf{H}}_k^{\mathrm{T}}$. The matrix $\tilde{\mathbf{V}}_k^{(0)}$ can be used to specify a beamforming vector that cancels the interference from the other pairs.

Given $\mathbf{H}_k^{\mathrm{T(eq)}} = \mathbf{H}_k^{\mathrm{T}}\tilde{\mathbf{V}}_k^{(0)} \in \mathbb{C}^{2\times(N-\tilde{r}_k)}$ as the equivalent channel matrix of pair $G^{(k)}$, one can assure that other pairs' interferences are suppressed. Based on this channel we can derive a single-group multicast beamforming vector $\mathbf{m}_k^{(\text{eq})} \in \mathbb{C}^{(N-\tilde{r}_k)\times 1}$ for pair $G^{(k)}$, which can be written as

$$\mathbf{m}_k^{(\text{eq})} = \mathbf{H}_k^{\mathrm{T(eq)}^{\mathrm{H}}}(\mathbf{H}_k^{\mathrm{T(eq)}}\mathbf{H}_k^{\mathrm{T(eq)}^{\mathrm{H}}})^{-1}\mathbf{1}, \qquad (10)$$

see the program

where $\mathbf{1} = [1, 1]^{\mathrm{T}}$.

The resulting beamforming vector for pair $G^{(k)}$ is then given by

$$\mathbf{m}_k = \tilde{\mathbf{V}}_k^{(0)} \mathbf{m}_k^{(\mathrm{eq})}, \tag{11}$$

which leads to the resulting beamforming matrix

$$\mathbf{M} = [\tilde{\mathbf{V}}_1^{(0)} \mathbf{m}_1^{(\mathrm{eq})}, \cdots, \tilde{\mathbf{V}}_K^{(0)} \mathbf{m}_K^{(\mathrm{eq})}]. \tag{12}$$

In MAZF, different to ZF, the received power is not balanced among the nodes. For this reason, a fair power loading which balances the received power among the nodes is considered in [15]. The power loading matrix $\mathbf{\Gamma} \in \mathbb{R}^{K \times K}$ is given by

$$\mathbf{\Gamma} = \mathrm{diag}(\min(|\mathbf{H}_1^{\mathrm{T}} \mathbf{m}_1|), \cdots, \min(|\mathbf{H}_K^{\mathrm{T}} \mathbf{m}_K|))^{-1}, \tag{13}$$

where the modulus operator $|\cdot|$ is assumed to be applied element-wise and the $\min$ function in this case returns the minimum element of a vector. This power loading ensures that the same amount of power is given to the worst node in each multicast pair. In order to satisfy the transmit power constraint, a normalization factor $\beta \in \mathbb{R}$ is needed and is given by

$$\beta = \sqrt{\frac{P_{\mathrm{RS}}}{\sigma_{\mathbf{t}_{\mathrm{RS}}} \mathrm{tr}(\mathbf{M}^{\mathrm{H}} \mathbf{M} \mathbf{\Gamma}^2)}}. \tag{14}$$

The beamforming solution for MAZF is given by

$$\mathbf{M}_{\mathrm{MAZF}} = \beta \mathbf{M} \mathbf{\Gamma}. \tag{15}$$

3.3. SINRB-Bisec

The optimization problem for obtaining the optimum precoding matrix which maximizes the worst node SINR can be written as

$$\begin{aligned} \mathbf{M}_{\mathrm{opt}} &= \underset{\mathbf{M}}{\mathrm{argmax}} \min_{n} \gamma_n, \quad n \in \{1, \cdots, N\} \\ &\text{subject to}: \sigma_{\mathbf{t}_{\mathrm{RS}}}^2 \, \mathrm{tr}(\mathbf{M}^{\mathrm{H}} \mathbf{M}) \leq P_{\mathrm{RS}}, \end{aligned} \tag{16}$$

with

$$\gamma_n = \frac{\sigma_{\mathbf{t}_{\mathrm{RS}}}^2 |\mathbf{h}_n^{\mathrm{T}} \mathbf{m}_k|^2}{\sum_{\substack{l=1 \\ l \neq k}}^{K} \sigma_{\mathbf{t}_{\mathrm{RS}}}^2 |\mathbf{h}_n^{\mathrm{T}} \mathbf{m}_l|^2 + \sigma_n^2}, \quad n \in \{1, \cdots, N\}, \tag{17}$$

being the SINR of node S_n. Note that the pair index k associated to node n can be obtained as $k = \lfloor (n+1)/2 \rfloor$. SINRB-Bisec is not adequate to maximize the sum rate but it provides equal SINR to all nodes such that it maximizes the fairness among nodes. Therefore, the performance comparison of SINRB-Bisec with MAZF and ZF is not that fair.

The contributions of [16, 17] have shown that (16) can be solved using bisection search method. First we need to specify the SINR interval where the optimal solution must lie and then determine the solution of (16) when considering the middle point of the interval as the target SINR. The interval is then successively bisected based on whether the required amount of power P_{req} exceeds the transmit power constraint P_{RS} or not. For each interval middle point, the problem (16) is solved. The bisection proceeds until a desired precision is reached with regard to $|P_{\mathrm{req}} - P_{\mathrm{RS}}|$.

4. ACHIEVABLE SUM RATE

In this section, the rate expressions for each phase and the overall two-phase sum rate will be given. It is assumed that $\log_2$ denotes the logarithm of base 2.

4.1. MAC Phase

In this subsection, the MAC phase rate is given. Having received vector $\mathbf{r}_{\mathrm{RS}}$, it is assumed that the RS is able to perform perfect decoding. Assuming Gaussian codebooks, the MAC phase capacity region is defined by

$$\sum_{n=1}^{N} R_n^{\mathrm{MAC}} \leq R^{\mathrm{MAC}} = \log_2 \det \left(\mathbf{I}_M + \frac{P_n}{\sigma_{RS}^2} \sum_{n=1}^{N} \mathbf{h}_n \mathbf{h}_n^{\mathrm{H}} \right) \tag{18}$$

[22], where

$$R_n^{\mathrm{MAC}} \leq R_n = \log_2(1 + \gamma(\|\mathbf{h}_n\|^2)), \forall n, n \in \{1, \cdots, N\} \tag{19}$$

means that each node's rate cannot be larger than when it is the only node transmitting.

4.2. BC Phase

In this subsection, the rate of the BC phase followed by the overall two-phase sum rate are given. The rate of node n is given by

$$R_n^* = \log_2(1 + \gamma_n), \ \forall n, n \in \{1, \cdots, N\} \tag{20}$$

[22], where γ_n is given in (17) due to having multiple transmit antennas at the RS and using multi-group multicast beamforming as explained in Section 3. Since the RS sends the same bit sequence to each node in pair $G^{(k)}$, the minimum rate between both nodes in $G^{(k)}$ will define the rate that can be transmitted in the second phase. Thus, the achievable rate for the second phase at both nodes S_{a_k} and S_{b_k} in $G^{(k)}$ is defined by

$$R_k^{2^{\mathrm{nd}}} = \min(R_{a_k}^*, R_{b_k}^*), \ \forall k, k \in \{1, \cdots, K\}. \tag{21}$$

The rate given in (21) considers only the second phase transmission when the RS multicasts the same bit sequences to both nodes. However, the information rate that can be received at each node depends on the available information rate of its partner in the first phase. Thus, $R_{a_k b_k}$ and $R_{b_k a_k}$, which

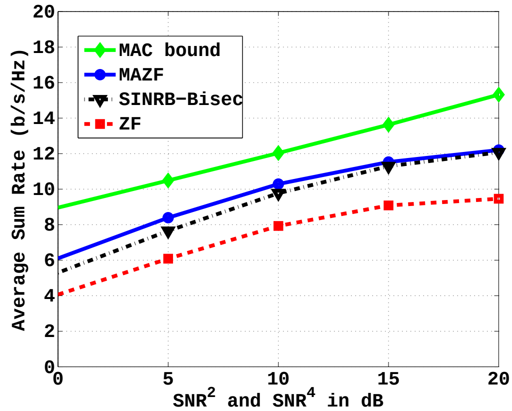


Fig. 2. Sum Rate performace of multi-group multicast beamforming: SNR^1 and SNR^3 fixed at 10 dB.

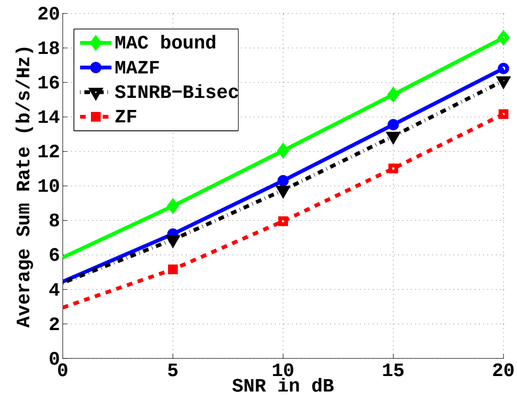


Fig. 3. Sum Rate performace of multi-group multicast beamforming: equal SNR for all nodes.

are the achievable rates for transmission from node S_{a_k} to node S_{b_k} and vice versa, respectively, are given by

$$R_{a_kb_k} = \min(R_{a_k}, R_k^{2\text{nd}}) \quad \text{and} \quad R_{b_ka_k} = \min(R_{b_k}, R_k^{2\text{nd}}). \tag{22}$$

The achievable BC rate for each pair is given by

$$R_k^{\text{BC}} = R_{a_kb_k} + R_{b_ka_k} \tag{23}$$

and the overall two-phase sum-rate is, thus, given by

$$R_{sum} = \frac{1}{2}\min\left(\sum_{k=1}^{K} R_k^{\text{BC}}, R^{\text{MAC}}\right), \tag{24}$$

where the pre-log factor $1/2$ is due to the use of two orthogonal resources. In [14], it is proposed to perform ZP such that a higher rate can be transmitted in the second phase. As an example, it is assumed that the RS is able to decode correctly the bit sequence $\mathbf{x}_{a_k}$ from S_{a_k} and the bit sequence $\mathbf{x}_{b_k}$ from S_{b_k}, which have lengths L_{a_k} and L_{b_k}, respectively. Assuming that the transmission rate from the first phase is sufficient to support the second phase, if for the second phase we have $R^*_{a_k} > R^*_{b_k}$ then the corresponding length of bit sequences that can be transmitted in the second phase is $L_{b_k} > L_{a_k}$, as node S_{a_k} expects to receive $\mathbf{x}_{b_k}$, which has the length of L_{b_k}, and vice versa. Using ZP, we append $L_{b_k} - L_{a_k}$ zeros to the bit sequence $\mathbf{x}_{a_k}$. At the second phase, the RS sends $\mathbf{x}_{a_kb_k}$, which is the XOR applied to the zero appended $\mathbf{x}_{a_k}$ and $\mathbf{x}_{b_k}$ and which is encoded with a codebook of rate $\max(R^*_{a_k}, R^*_{b_k}) = R^*_{a_k}$.

For the decoding process, node S_{a_k} employs a codebook with rate $R^*_{a_k}$ and obtains $\mathbf{x}_{a_kb_k}$ perfectly. On the other hand, node S_{b_k} knows a priori that there is ZP being used, so that it decodes only the beginning part of $\mathbf{x}_{a_kb_k}$ according to the rate $R^*_{b_k}$. Thus, as stated in [14], with ZP and a priori knowledge at the corresponding nodes, different rates are supported and the broadcast capacity as given in [21] is achieved.

Using the ZP approach, (22) needs to be adjusted according to

$$\begin{aligned} R_{a_kb_k} &= \min(R_{a_k}, R^*_{b_k}) \\ R_{b_ka_k} &= \min(R_{b_k}, R^*_{a_k}). \end{aligned} \tag{25}$$

The overall two-phase sum rate with ZP remains the same as in (24) when exchanging (22) by (25).

5. PERFORMANCE ANALYSIS

As in [14], the average signal to noise ratio (SNR) is defined as SNR $= P_s/\sigma_r^2$, where P_s is the transmit signal power at the transmitter and σ_r^2 is the noise variance at the receiver. In the first phase, the nodes are the transmitters and the RS is the receiver, and for the second phase it is vice versa. In the following, the RS has $N = 4$ antennas , so that it can support four nodes, namely S_1, S_2, S_3 and S_4. Two bidirectional pairs are considered, namely $G^{(1)}$ consisting of S_1 and S_2 and $G^{(2)}$ consisting of S_3 and S_4, respectively. All nodes and the RS are assumed to have the same noise variance and the RS is assumed to have the sum transmit power of all nodes, i.e., $P_{\text{RS}} = NP_n$. We assume reciprocal flat fading channels with unit average gain for both phases. The notation SNR^x denotes the SNR of the link between node x and the RS.

Figures 2 and 3 show the overall two-phase sum rate performance and the MAC bound when multi-group multicast beamforming is applied for the second phase and the XOR operation is performed without ZP. In Figure 2, SNR^1 and SNR^3 are fixed at 10 dB while SNR^2 and SNR^4 are varied. Figure 3 shows the average sum rate when all SNR values are equal. The MAZF outperforms the non-multicast aware ZF multi-group multicast beamforming as it does not perform null-space projection to the channel of the node within the

see the program

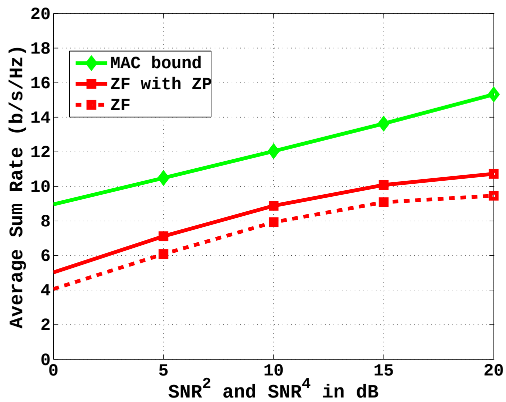


Fig. 4. Performance comparison of different XOR approaches for ZF: SNR^1 and SNR^3 fixed at 10 dB.

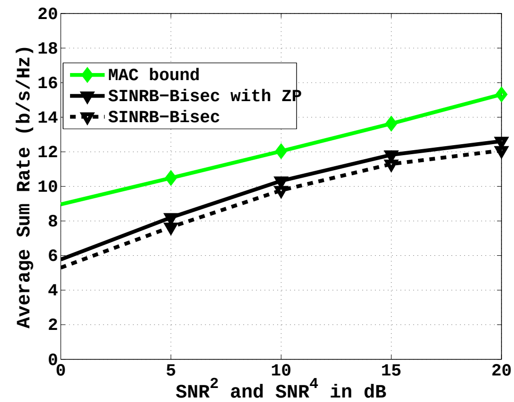


Fig. 6. Performance comparison of different XOR approaches for SINRB-Bisec: SNR^1 and SNR^3 fixed at 10 dB.

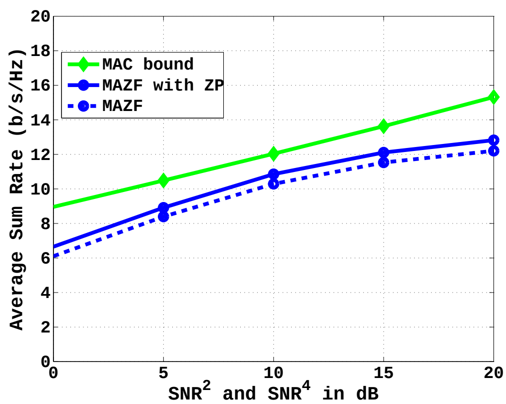


Fig. 5. Performance comparison of different XOR approaches for MAZF: SNR^1 and SNR^3 fixed at 10 dB.

same pair. The performance improvement of MAZF comes at the price of a higher computational complexity compared to the multi-group multicast ZF beamforming. The MAZF performs slightly better than the SINRB-Bisec. This happens since SINRB-Bisec provides higher fairness by guaranteeing equal SINR to all nodes, which is not the case when using MAZF. MAZF provides fairness by balancing the received power among the nodes, which cannot assure that the SINR is equal among all nodes. Furthermore, in [15] it is shown that MAZF has a lower complexity compared to SINRB-Bisec. It can be seen that the average sum rate will saturate when the SNR of one node in a pair is fixed while the other node's SNR is increased. This is due to the fact that the lower SNR limits the overall two-phase sum rate.

Figures 4, 5 and 6 show the performance improvement of all three multi-group multicast beamforming algorithms when using ZP. The usage of ZP is beneficial in the case of unbalanced SINR among two nodes in a pair. We measure the performance of multi-user two-way relaying when SNR^1 and SNR^3 are fixed at 10 dB while SNR^2 and SNR^4 are varied. The use of ZP at the RS when performing the XOR operation for each pair helps to achieve the two-way relaying broadcast capacity of each pair, resulting in a higher sum rate compared to the case without ZP, which is only able to transmit the minimum broadcast rate between the two nodes in a pair. From these three figures, it can be seen that ZF obtains the highest gain compared to the other multi-group multicast beamforming algorithms when the ZP approach is used. Without ZP, the rate is defined by the worst node in each pair. Both MAZF and SINRB-Bisec consider fairness between groups and, in the case of SINRB-Bisec, balances the SINR. Thus, the difference between the worst SINR and the highest SINR in a pair is not high and the performance improvement of ZP approach is not as much as for ZF, which does not include fairness.

In contrast to [14], we do not intend to maximize the sum rate of the multi-user two-way relaying but to provide fairness among the pairs. Therefore, algorithms maximizing the sum rate as provided in [14] will outperform the ZF, MAZF and SINRB-Bisec at the cost of fairness.

Point to point Multiple Input Multiple Output (MIMO) communication is expected to achieve higher sum rate than the multi-user two-way relaying. However, in multi-user two-way relaying it is assumed that there exists no direct link such that point to point communication among nodes in a pair can not be performed, e.g., due to high attenuation. To perform a fair comparison between point to point MIMO and multi-user two-way relaying, the attenuation factor must be taken into account.

6. CONCLUSION

In this work, we propose to apply multi-group multicast beamforming for the second phase transmission in a multi-user two-way relaying protocol. We extend three multi-group multicast beamforming algorithms, namely ZF, MAZF and SINRB-Bisec to the case of multi-user two-way relaying. The overall two-phase sum rate expressions are derived for both cases, with and without ZP, as a way to measure the performance of the multi-user two-way relaying. From the simulation analysis, the ZP approach outperforms the one without ZP. Comparing the three multicast beamforming strategies, the fairness oriented MAZF and SINRB-Bisec algorithms outperform the non multicast aware ZF, which requires the lowest computational complexity. Although MAZF and SINRB-Bisec provide similar performances, MAZF requires less computational complexity compared to SINRB-Bisec.

7. REFERENCES

[1] H. Bolcskei, R.U. Nabar, Ö. Oyman, and A. J. Paulraj, ”Capacity scaling laws in MIMO relay networks,” *IEEE Transactions on Wireless Communications*, vol.5, no. 6, pp. 1433-1444, June 2006.

[2] B. Rankov and A. Wittneben, ”Spectral efficient protocols for half-duplex relay channels,” *IEEE Journal on Selected Areas Communications*, vol. 25, no. 2, pp. 379-389, Feb. 2007.

[3] S. Katti, I. Maric, A. Goldsmith, D. Katabi, and M. Medrad, ”Joint relaying and network coding in wireless networks,” in *Proc. IEEE International Symposium on Information Theory*, pp. 1101-1105, Nice, June 2007.

[4] S. Katti, S. Gollakota, and D. Katabi, ”Embracing wireless interference: analog network coding,” in *Proc. ACM Special Interest Group on Data Communication*, pp. 397-408, Kyoto, Aug. 2007.

[5] P. Popovski and H. Yomo, ”Wireless network coding by amplify and aorward for bi-directional traffic flows,” *IEEE Communications Letters*, vol. 11, no. 1, pp. 16-18, Jan. 2007.

[6] P. Popovski and H. Yomo, ”The anti-packets can increase the achievable throughput of a wireless multi-hop network,” in *Proc. IEEE International Conference on Communications,* pp. 3885-3890, Istanbul, June 2006.

[7] S. Zhang, S. Liew and P.P. Lam, ”Physical-layer network coding,” in *Proc. ACM Mobile Computing and Networking*, pp. 358-365, Los Angeles, September, 2006.

[8] T. Unger and A. Klein, ”Applying relay stations with multiple antennas in the one- and two-way relay channel,” in *Proc. International Symposium on Personal, Indoor and Mobile Radio Communications*, Athens, Greece, Sep. 2007.

[9] T. Unger and A. Klein, ”Duplex schemes in multiple-antenna two-hop relaying,” *EURASIP Journal on Advances in Signal Processing, (special issue: ”Multihop-Based Cellular Networks”)*, Vol. 8 , Issue 2, pp. 1-14, Jan. 2008.

[10] I. Hammerström, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, ”MIMO two-way relaying with transmit CSI at the relay,” in *Proc. IEEE Signal Processing Advances in Wireless Communications*, Helsinki, Finland, June 2007.

[11] R. F. Wyrembelski, T. J. Oechtering, and H. Boche, ”Decode-and-forward strategies for bidirectional relaying,” in *Proc. 19th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Cannes, Sept. 2008.

[12] C. Esli and A. Wittneben, ”Multiuser MIMO Two-way Relaying for Cellular Communications,” in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Cannes, Sept. 2008.

[13] M. Chen and A. Yener, ”Multiuser two-way relaying for interference limited systems,” in *Proc. IEEE International Conference on Communications*, pp.3883-3887, Beijing, May 2008.

[14] C. Esli and A. Wittneben, ”One- and two-way decode-and-forward relaying for wireless multiuser MIMO networks,” in *Proc. IEEE Global Communications Conference*, New Orleans, Nov. 2008.

[15] Y. Silva, *Adaptive beamforming and power allocation in multi-carrier multicast wireless network*, PhD Dissertation, TU Darmstadt, 2008.

[16] Y. Gao and M. Schubert, ”Power allocation for multi-group multicasting with beamforming,” in *Proc. ITG/IEEE Workshop on Smart Antennas*, Ulm, Mar. 2006.

[17] E. Karipidis, N. D. Sidiropoulos, Z. Q. Luo, ”Far-field multicast beamforming for uniform linear antenna arrays,” *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 4916-4927, Oct. 2007.

[18] M. Joham, *Optimization of linear and nonlinear transmit signal processing*, PhD Dissertation, TU München, 2004.

[19] Q. H. Spencer and M. Haardt, ”Capacity and downlink transmission algorithms for a multi-user MIMO channel,” in *Proc. Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 1384-1388, Nov. 2002.

[20] C. Suh and C.-S. Hwang, ”Dynamic subchannel and bit allocation for multicast OFDM systems,” in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, pp. 2102-2106, Sep. 2004.

[21] R. F. Wyrembelski, T. J. Oechtering, I. Bjelakovic, C. Schnurr, and H. Boche, ”Capacity of Gaussian MIMO Bidirectional Broadcast Channels,” in *Proc. IEEE International Symposium on Information Theory*, pp. 584-588, Toronto, Canada, July 2008.

[22] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

see the program

Throughput maximization through Network-Based Stream-Number Decision for MIMO HSDPA

Govinda Lilley, Martin Wrulich and Markus Rupp
Institute of Communications and RF Engineering, Vienna University of Technology
{glilley,mwrulich,mrupp}@nt.tuwien.ac.at

***Abstract*—3GPP standardized double-stream transmit antenna array (D-TxAA) in Release 7 as the successor of classical single-input single-output (SISO) HSDPA, supporting a maximum of 42.2 Mbit/s [1]. To achieve these high throughputs, the number of parallel streams that are utilized during a transmission time interval (TTI) should be chosen to suit the current channel conditions. Thus, either the mobile or the network (NodeB) has to derive a decision upon the stream-utilization based on this information. Although the standard permits the evaluation of the number of streams on the mobile (and signaling it to the NodeB), the necessary SINR estimation for both the single- and double-stream case is difficult. Furthermore, the network operator may opt to have the number of streams being decided in the NodeB to keep control of the spatial interference structure in the cell. In this paper, we introduce two network-based stream decision algorithms and investigate them by means of system-level simulations to compare their performance to the mobile-based stream number decision.**

I. Introduction

During the last years, high-speed downlink packet access (HSDPA) has led to a significant increase in mobile data usage, and the throughput demand in modern communication systems is expected to grow even faster in the next years. In order to satisfy this demand, 3GPP standardized a successor of the WCDMA-based single-input single-output HSDPA that supports multiple antennas on both the transmitter and the receiver side.

As part of the 3GPP Release 7 [2], double-stream transmit antenna array (D-TxAA) has been defined as the transmission scheme for multiple-input multiple-output (MIMO) HSDPA communication. This scheme allows for the parallel transmission of up to two independent data streams, supporting a theoretical maximum data rate of 42.2 Mbit/s [1], [3]. The next generation communication system—Long Term Evolution (LTE)—will even outperform this limit [4], [5]. However, these networks need a completely new wireless infrastructure, and accordingly network operators will most probably introduce MIMO HSDPA before going towards LTE.

The D-TxAA scheme or Release 7 HSDPA relies on the parallel transmission of two data streams, where the number of utilized streams depends on the current channel conditions, similar to the concept of adaptive modulation and coding (AMC). The standard [2] does not fully specify the feedback scheme, but given the basic requirements it permits the evaluation and signaling of the number of streams from the mobile side. Typically such an evaluation would be based on a suitable SINR estimation [6].

Such an estimation depends on a number of factors and can be very challenging [7], [8]. In addition, the interference structure in the cell - a factor that depends on the number of active streams - is determined by by the equipment, which in general is undesired by the network operators [9]. Individual enhancements of the user channel quality by means of interference-canceling [10], [11] or similar techniques do not impose such problems. Still, for an efficient network operation, network entities should assign the resources according to overall goals, e.g. the average cell-throughput [12]–[14], and therefore actively manage the interference situation in the cells.

In this paper we introduce two different (*network-based*) stream number decision algorithms that are based upon the CQI and ACK/NACK reports of the user equipment. We compare the performance of our algorithms with the performance of the user-based stream number decision where full CSI is available. Furthermore, the necessary thresholds of the algorithms are evaluated by system-level simulations [15] to identify the optimum operation points.

The paper is organized as follows: Section II explains the system-level modeling and some details about our simulator. Section III and IV introduce the reference and network-based stream decision algorithms, with their performance being evaluated in Section V. Finally, we conclude the paper in Section VI.

II. System Model and Simulator Details

To assess the performance of the proposed algorithms, we utilized a MIMO HSDPA system level simulator [16], developed in a MATLAB-environment. The simulator can basically be split in three different parts, (1) channel quality evaluation, (2) performance prediction (i.e. BLER evaluation), and (3) network/user handling including all relevant algorithms, depicted in Figure 1. It has to be noted that the channel quality evaluation relies on the utilization of so-called *fading-parameters* that account for a computationally efficient evaluation [15].

The system-level simulator allows for the investigation of MIMO HSDPA networks in terms of block error rates, throughput, scheduler performance and other network-related performance measures. The investigated networks are of 3GPP type one layout [17] but can be configured in terms of simulated users, timing, interference structure, size and many more. Besides the physical layer abstraction, all layer-functionalities

see the program

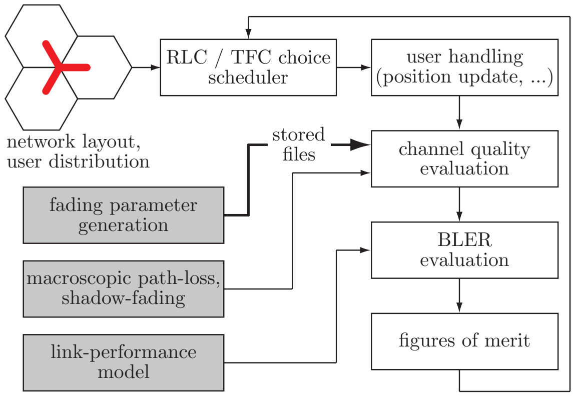


Fig. 1. Overview of D-TxAA MIMO HSDPA system level simulator [16]. The functionality can be split in three main parts, (1) channel quality evaluation, (2) performance prediction, and (3) network/user handling.

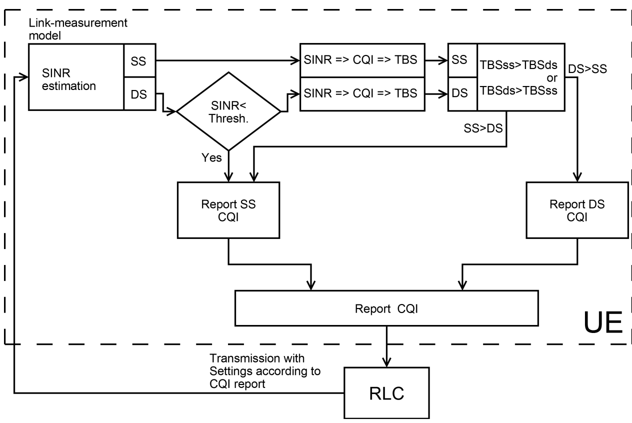


Fig. 2. UE Stream-Number Decision algorithm.

up to the MAC-d are implemented, with a special focus on MAC-hs algorithms, i.e. scheduling and physical layer resource management.

III. User Equipment Stream Decision Algorithm

The User Equipment (UE) Stream Decision algorithm is based on the fact that the UE has quasi full channel state information (CSI) and hence should be able to calculate the optimal transmission settings. Therefore this algorithm will be used as a reference to which our other algorithms shall be compared. The algorithm is shown in Figure 2.

The UE calculates the SINR for single stream (SS) and double stream (DS) transmission scenarios. As long as the SINR of stream 2 in DS transmission mode $\mathrm{SINR}_{\mathrm{DS},2}$ is above the threshold $\mathrm{SINR}_{thr.}$, the UE is allowed to decide how many streams to use. If the UE is permitted a decision and the double stream transport block size $\mathrm{TBS}_{\mathrm{DS}}$ is greater than the single stream transport block size $\mathrm{TBS}_{\mathrm{SS}}$, two streams are used. In all other cases the transmission is single-streamed. Note that $\mathrm{TBS} = \mathrm{TBS}(\mathrm{CQI}(\mathrm{SINR}))$ where the mapping CQI(SINR) is extracted from AWGN link level simulations. Essentially the stream number decision is given by

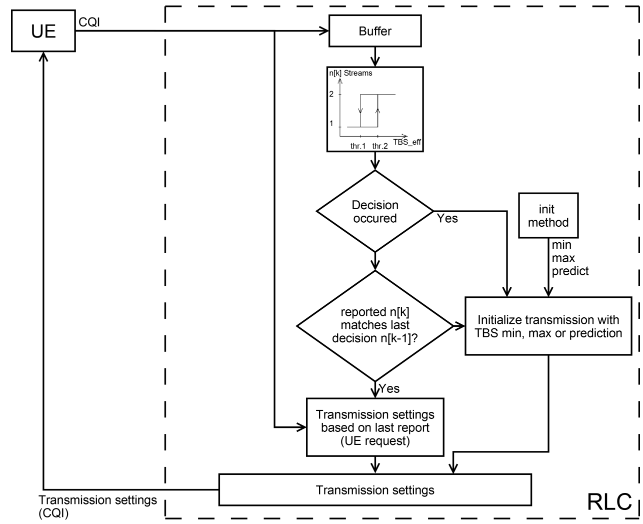


Fig. 3. RLC Stream-Number Decision algorithm based on CQI reports.

$$n_k = \begin{cases} 1, & \text{if } \mathrm{SINR}_{\mathrm{DS},2} < \mathrm{SINR}_{thr.} \vee \mathrm{TBS}_{\mathrm{DS}} < \mathrm{TBS}_{\mathrm{SS}} \\ 2, & \text{if } \mathrm{SINR}_{\mathrm{DS},2} \geq \mathrm{SINR}_{thr.} \wedge \mathrm{TBS}_{\mathrm{DS}} \geq \mathrm{TBS}_{\mathrm{SS}}. \end{cases} \tag{1}$$

After the stream number decision has been made, the CQI corresponding to the decision is reported to the radio link control (RLC), which then uses the transmission settings associated with it for the next transmission.

IV. Radio Link Control Stream Decision Algorithms

Signal processing in handheld devices is usually constrained in complexity due to limited battery capacity. Thus, if a user-based stream number decision were to be implemented, it would consume even more computational resources and battery. The network operator also has less control over the spatial interference situation, which is influenced by the number of used streams and their precoding. Furthermore, advanced schedulers utilizing the number of streams as an optimization parameter are rendered impossible. These issues do not exist when the stream number decision is moved to the RLC. We propose two algorithms, with the first one using only CQI reports from the UE to calculate the optimal transmission settings. The second algorithm uses a combination of CQI and ACK/NACK reports for transmission settings computation in order to optimize the BLER.

Figure 3 shows our first RLC Stream Decision algorithm which bases its decision on the CQI report history of the UE. The reports are saved in up- and downgrade FIFO buffers of lengths N_u and N_d, respectively. These reports are converted to corresponding TBS values, which are then compared to up- and downgrade thresholds in order to decide how many streams will be used. The stream number decision is given by

see the program

$$n_k = \begin{cases} 1, & \text{if } \mathrm{TBS}_i(\mathrm{CQI}_i) < \mathrm{TBS}(\mathrm{CQI}_d),\ i = 1 \ldots N_d \\ 2, & \text{if } \mathrm{TBS}_j(\mathrm{CQI}_j) > \mathrm{TBS}(\mathrm{CQI}_u),\ j = 1 \ldots N_u \\ n_{k-1}, & \text{otherwise,} \end{cases} \tag{2}$$

where CQI_i is taken from the downgrade FIFO buffer, CQI_j is taken from the upgrade buffer and k represents the current TTI index. $\mathrm{TBS}(\mathrm{CQI}_d)$ and $\mathrm{TBS}(\mathrm{CQI}_u)$ are the down- and upgrade thresholds which we generate from CQI thresholds. Since the CQI to TBS mapping is unique and CQIs are the same for all mapping tables, using CQIs rather than TBS values as thresholds is practical.

After up- or downgrading, the transmission settings that will be utilized have to be determined. We implemented three initialization methods in order to compare their impact on throughput performance and ACK/NACK ratio. For the first two methods the CQI is set to either the maximum or the minimum CQI value available within the selected CQI mapping table (e.g. 30 or 1 for single stream). The third method is based on extrapolation of TBS values via linear least square approximation. The TBS values are extracted from the CQI reports within the upgrade buffer, because this buffer contains the most recent reports and hence will deliver a more accurate prediction. In case of a downgrade, the predicted TBS value TBS_P is directly used to calculate the transmission settings. In case of an upgrade, TBS_P is distributed over both streams where $\mathrm{TBS}_{stream1} = \alpha \cdot \mathrm{TBS}_P$ and $\mathrm{TBS}_{stream2} = (1-\alpha) \cdot \mathrm{TBS}_P$ are the TBS values for stream 1 and 2, respectively, and α is the distribution factor. $\mathrm{TBS}_{stream1}$ and $\mathrm{TBS}_{stream2}$ are then used to gain the transmission settings for stream 1 and 2. In an effort to keep complexity low, α is kept constant and computed as follows

$$\alpha = \frac{\mathrm{TBS}(\mathrm{CQI}_1)}{\mathrm{TBS}(\mathrm{CQI}_1) + \mathrm{TBS}(\mathrm{CQI}_1 - \Delta\mathrm{CQI})} \tag{3}$$

$\mathrm{TBS}(\mathrm{CQI}_1)$ is the TBS of the first stream, and $\mathrm{TBS}(\mathrm{CQI}_1 - \Delta\mathrm{CQI})$ represents the TBS of the second stream, which in average is $\Delta\mathrm{CQI}$ lower than the first due to the precoding favouring stream 1, see [15]. Note that the precoding is optimised for stream 1 consequently resulting in a higher BLER for stream 2. To find an appropriate value for $\Delta\mathrm{CQI}$ we ran a simulation of our UE sided (reference) algorithm. The CQI trace of the simulation and average CQIs are shown in Figure 4 and a plot of α over CQI_1 is shown in Figure 5. The mapping $\mathrm{TBS}(\mathrm{CQI})$ is done according to [2].

It is also important to note, that immediately after an up- or downgrade decision, due to the feedback delay, no reports of the UE for the current stream number decision are available. Accordingly, the RLC transmits with the new transmission settings until valid reports have been recieved.

Our second algorithm is very similar to the first, save for the fact that the ACK/NACK ratio is also considered in order to improve the overall BLER, and that transmission initialization after an up- or downgrade is always done via linear least square extrapolation of the TBS weighted with the average BLER.

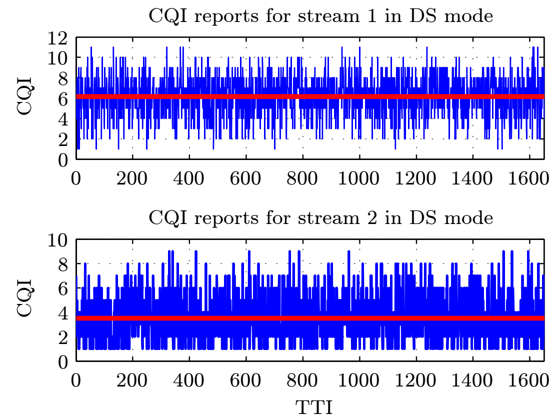


Fig. 4. CQI trace of a double stream transmission for the UE-based Stream-Number decision algorithm.

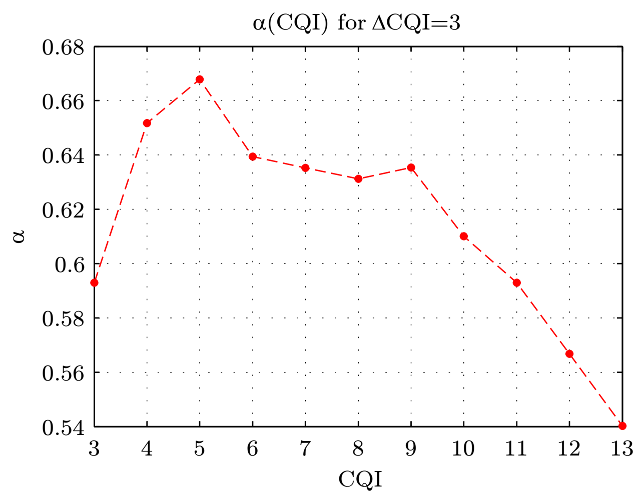


Fig. 5. α for CQI of the first stream for $\Delta\mathrm{CQI} = 3$.

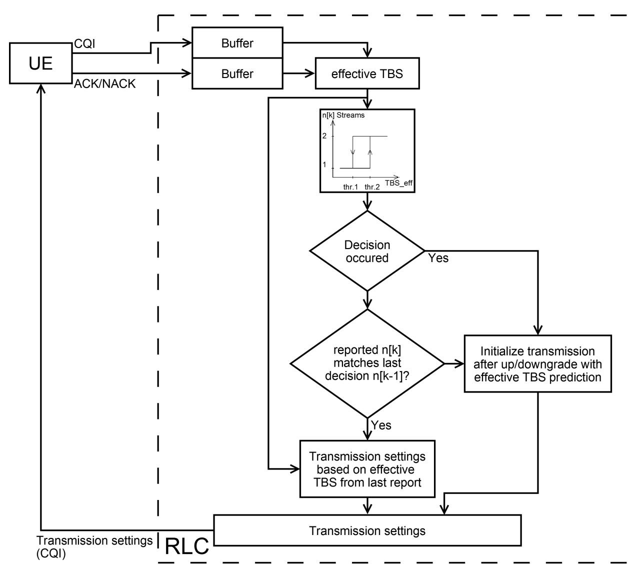


Fig. 6. RLC Stream-Number Decision algorithm based on effective TBS.

see the program

Figure 6 shows our second algorithm with the decision being expressed as

$$n_k = \begin{cases} 1, & \text{if}\, \mathrm{TBS}_{\mathrm{eff},i}(\mathrm{CQI}_i) < \mathrm{TBS}_{\mathrm{eff},d}(\mathrm{CQI}_d),\ i = 1 \ldots N_d \\ 2, & \text{if}\, \mathrm{TBS}_{\mathrm{eff},j}(\mathrm{CQI}_j) > \mathrm{TBS}_{\mathrm{eff},u}(\mathrm{CQI}_u),\ j = 1 \ldots N_u \\ n_{k-1}, & \text{otherwise}, \end{cases} \tag{4}$$

where the effective TBS of element i within the up/downgrade buffer $\mathrm{TBS}_{\mathrm{eff},i} = \mathrm{TBS}_i \cdot (1 - \mathrm{BLER})$ and BLER is the average of the most recent 40 ACK/NACK reports. The effective TBS threshold $\mathrm{TBS}_{\mathrm{eff},u/d}(\mathrm{CQI}_{u/d}) = \mathrm{TBS}_{u/d}(\mathrm{CQI}_{u/d}) \cdot (1 - \mathrm{BLER}_{\mathrm{target}})$ is the product of the TBS threshold and the desired overall transmission efficiency $(1 - \mathrm{BLER}_{\mathrm{target}})$ the system should have. In addition the transmitted TBS is reduced by the averaged ACK/NACK ratio to further decrease BLER. Note that as in the previous algorithm, after an up- or downgrade, the RLC transmits with the new transmission settings until valid reports are available.

V. Simulation Details

In this section we present simulation results for various thresholds of the algorithms discussed above and compare them to the performance of our reference algorithm. Noteworthy results can be expected for the up- and downgrade thresholds, which therefore will be covered here. In Figure 7, the average cell throughput is plotted for the TBS-based algorithm and an increasing downgrade threshold CQI_d. The low throughput at 0 and 1 results from an effective downgrade prevention caused by $\mathrm{TBS}(\mathrm{CQI}_{\mathrm{DS},0}) = \mathrm{TBS}(\mathrm{CQI}_{\mathrm{DS},1})$, where $\mathrm{TBS}(\mathrm{CQI}_j) > \mathrm{TBS}(\mathrm{CQI}_u)$ is always valid, thereby greatly increasing BLER, as can be seen in Figure 8. Splitting of the throughput and BLER curves for the three initialization methods, observed at downgrade thresholds greater than 2, is the result of a drastic reduction of DS transmissions due to the more aggressive downgrading. The low amount of DS transmissions coupled with most transmissions being downgraded shortly after being upgraded gives transmissions with optimized transmission initialization an obvious advantage, which can be seen in Figures 7 and 8. For the $\mathrm{TBS}_{\mathrm{eff}}$-based algorithm, the BLER is shown in Figure 9. The BLER is very close to the target value of 10% due to the feedback of transmission settings based on effective TBS. Since this algorithm keeps the BLER of stream 1 close to 10%, the benefit is a high throughput, even under suboptimal conditions. In a further simulation with a suboptimal mapping of $\mathrm{CQI}(\mathrm{SINR}_{\mathrm{opt.}} + 1\mathrm{dB})$, we observed a maximum throughput gain of 650 kBit/s, which is depicted in Figure 10. This case is of particular interest, when the estimators necessary for evaluating the feedback in the UE do not work properly or are poorly implemented.

For a variable upgrade threshold CQI_u the results are shown in Figures 11 and 12. The simulations show an upward trend of the throughput, which is a result of the decreasing BLER associated with increasingly conservative upgrading. The effect of higher upgrade thresholds is shown in Figures 13 and 14.

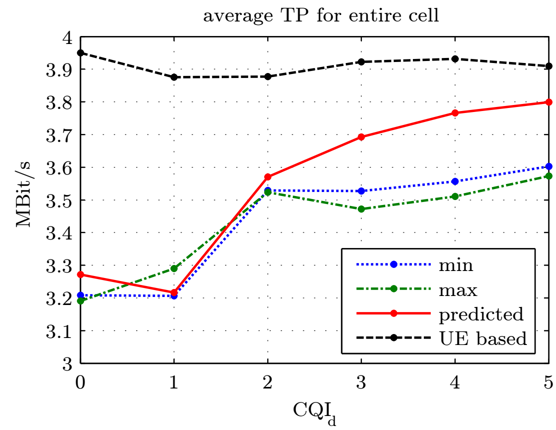


Fig. 7. Throughput simulation of the RLC-TBS-based algorithm for $\mathrm{CQI}_d = 0 \ldots 5$.

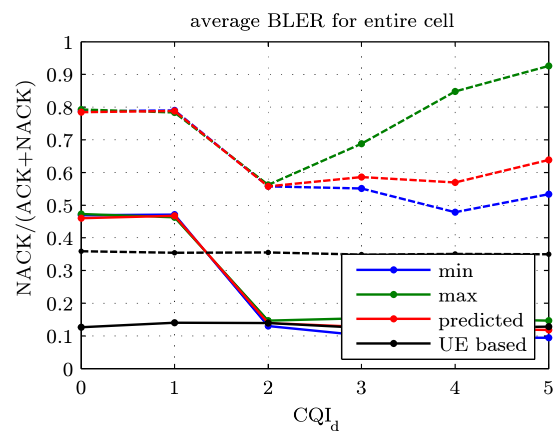


Fig. 8. Simulation results for the BLER of the RLC-TBS-based algorithm with $\mathrm{CQI}_d = 0 \ldots 5$. Full lines represent stream 1 and dashed lines represent stream 2. The blue lines (min) are for initialization of the transmission with the minimum value, green lines (max) are for initialization with the maximum value, red (predicted) is for initialization based on linear least square extrapolation and black (UE-based) is the average BLER for the UE-based reference algorithm.

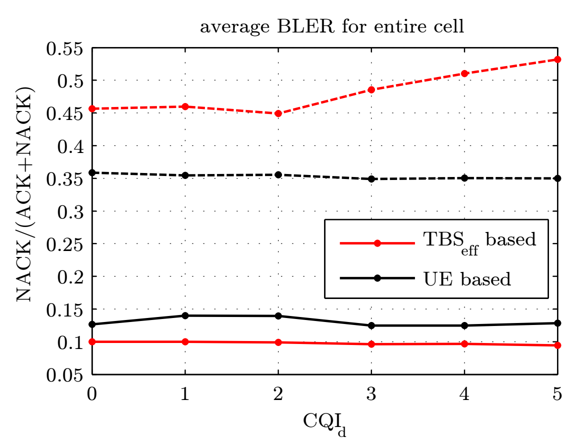


Fig. 9. Simulation results for the BLER of the RLC-$\mathrm{TBS}_{\mathrm{eff}}$-based algorithm with $\mathrm{CQI}_d = 0 \ldots 5$. Full lines represent stream 1 and dashed lines represent stream 2.

see the program

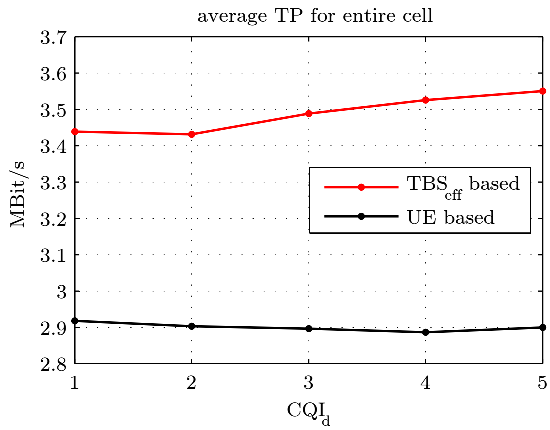


Fig. 10. Simulation results for the troughput of the RLC-TBS$_{eff}$-based algorithm with CQI$_d$ = 1...5 and suboptimal mapping of CQI(SINR$_{opt.}$ + 1dB).

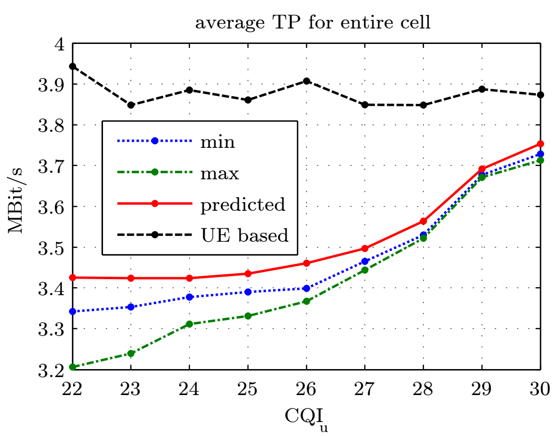


Fig. 11. Throughput results of the RLC-TBS-based algorithm with CQI$_u$ = 22...30.

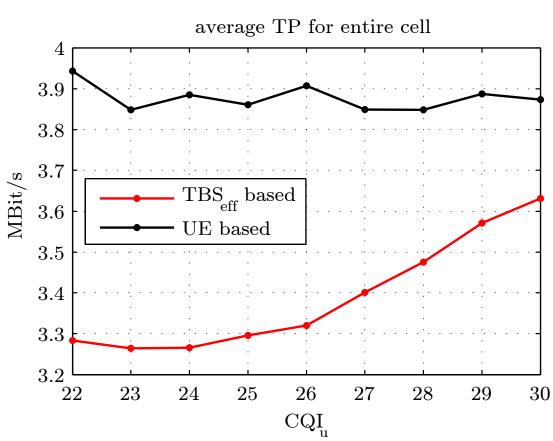


Fig. 12. Throughput results of the RLC-TBS$_{eff}$-based algorithm with CQI$_u$ = 22...30.

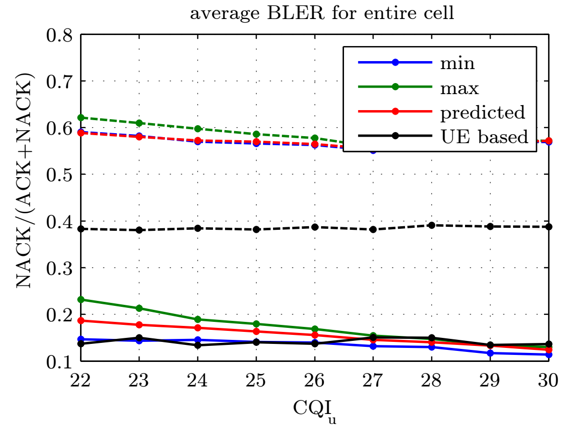


Fig. 13. Simulation results for the BLER of the RLC-TBS-based algorithm with CQI$_u$ = 22...30. Full lines represent stream 1 and dashed lines represent stream 2.

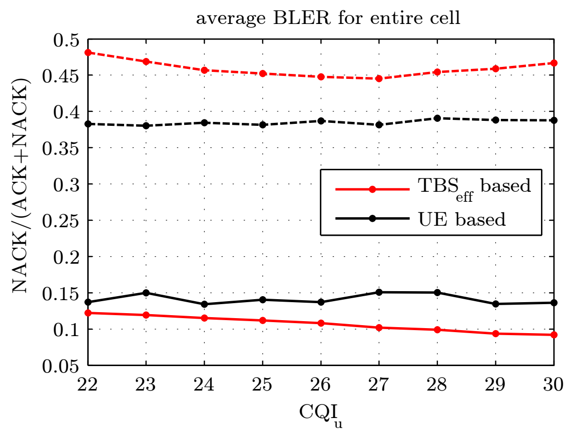


Fig. 14. Simulation results for the BLER of the RLC-TBS$_{eff}$-based algorithm with CQI$_u$ = 22...30. Full lines represent stream 1 and dashed lines represent stream 2.

VI. Conclusions

In this paper we discuss and investigate two RLC-based stream number decision algorithms and compare them to UE-based stream number decision performance. Simulations show good average throughput and BLER values for the TBS-based algorithm with linear least square predictive initialization at aggressive downgrade thresholds and low DS counts. Also downgrade blocking at downgrade thresholds CQI$_d$ of 0 and 1 was observed for the TBS-based algorithm, which lead to significant degradation of throughput performance. Finally we investigated the behaviour of our algorithms for suboptimal conditions. The TBS$_{eff}$-based algorithm proved to have the highest throughput, while retaining performance similar to the TBS-based algorithm under improved conditions. This case is particularly interesting when the estimators necessary for evaluating the feedback do not work properly or are poorly implemented.

Acknowledgment

This work has been funded by *mobilkom Austria AG* and the Institute of Communications and Radio-Frequency Engineering. The authors would like to thank C. Mehlführer for many fruitful discussions. The views expressed in this paper are those of the authors and do not necessarily reflect the views within *mobilkom Austria AG*.

References

[1] T. Nihtila and V. Haikola, "HSDPA MIMO system performance in macro cell network," in *Proc. IEEE Sarnoff Symposium*, 2008, pp. 1–4.

[2] Members of 3GPP, "Technical specification group radio access network; physical layer procedures (FDD)," 3GPP, Tech. Rep. 3GTS 25.214 Version 7.4.0, Mar. 2007.

[3] H. Holma and A. Toskala, *HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications*. John Wiley & Sons, Ltd., 2006, ISBN 0-470-01884-4.

[4] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G Evolution: HSDPA and LTE for Mobile Broadband*. Academic Press, Jul. 2007.

[5] C. Mehlführer, M. Wrulich, J. Colom Ikuno, D. Bosanska, and M. Rupp, "Simulating the long term evolution physical layer," submitted to EUSIPCO 2009.

[6] C. Mehlführer, S. Caban, M. Wrulich, and M. Rupp, "Joint throughput optimized CQI and precoding weight calculation for MIMO HSDPA," in *Proc. 42nd Asilomar Conference on Signals, Systems and Computers*, Oct. 2008.

[7] M. Wrulich and M. Rupp, "Efficient link measurement model for system level simulations of Alamouti encoded MIMO HSDPA transmissions," in *Proc. ITG International Workshop on Smart Antennas (WSA)*, Darmstadt, Germany, Feb. 2008.

[8] J. Colom Ikuno, M. Wrulich, and M. Rupp, "Performance and modeling of LTE H-ARQ," in *Proc. ITG International Workshop on Smart Antennas (WSA)*, Berlin, Germany, Feb. 2009.

[9] A. Mäder, D. Staehle, and M. Spahn, "Impact of HSDPA radio resource allocation schemes on the system performance of UMTS networks," in *Proc. IEEE 66th Vehicular Technology Conference Fall (VTC)*, 2007, pp. 315–319.

[10] M. Wrulich, C. Mehlführer, and M. Rupp, "Interference aware MMSE equalization for MIMO TxAA," in *Proc. IEEE 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, 2008, pp. 1585–1589.

[11] C. Mehlführer, M. Wrulich, and M. Rupp, "Intra-cell interference aware equalization for TxAA HSDPA," in *Proc. IEEE 3rd International Symposium on Wireless Pervasive Computing*, 2008, pp. 406–409.

[12] K. I. Pedersen, T. F. Lootsma, M. Stottrup, F. Frederiksen, T. Kolding, and P. E. Mogensen, "Network performance of mixed traffic on high speed downlink packet access and dedicated channels in WCDMA," in *Proc. IEEE 60th Vehicular Technology Conference (VTC)*, vol. 6, 2004, pp. 4496–4500.

[13] M. Wrulich, W. Weiler, and M.Rupp, "HSDPA performance in a mixed traffic network," in *Proc. IEEE Vehicular Technology Conference Spring (VTC)*, May 2008, pp. 2056–2060.

[14] H. Chao, Z. Liang, Y. Wang, and L. Gui, "A dynamic resource allocation method for HSDPA in WCDMA system," in *Proc. IEE International Conference on 3G Mobile Communication Technologies*, 2004, pp. 569–573.

[15] M. Wrulich, S. Eder, I. Viering, and M. Rupp, "Efficient link-to-system level model for MIMO HSDPA," in *Proc. IEEE 4th Broadband Wireless Access Workshop*, 2008.

[16] M. Wrulich and M. Rupp, "Computationally efficient MIMO HSDPA system-level evaluation," 2009, submitted to EURASIP Journal on Wireless Communications and Networking.

[17] Members of 3GPP, "Technical specification group radio access network; spatial channel model for multiple input multiple output (MIMO) simulations," 3GPP, Tech. Rep. 3GPP TS 25.996 Version 7.0.0, Jun. 2007.

see the program

SECRECY RATE REGION OF MISO INTERFERENCE CHANNEL: PARETO BOUNDARY AND NON-COOPERATIVE GAMES

Eduard A. Jorswieck and Rami Mochaourab

Dresden University of Technology
Communications Laboratory
Communications Theory
01062 Dresden
{jorswieck,mochaourab}@ifn.et.tu-dresden.de

ABSTRACT

Competition in interference channels is modeled using non-cooperative game theory. The well known Nash equilibrium often leads to inefficient solutions. In the current work, we study the achievable secrecy rate of a multiple-antenna (MISO) interference channel (IFC). First, we characterize the Pareto boundary of the secrecy rate region. Next, the parameterization is exploited to compute the Nash equilibrium. In contrast to the usual case in which the rate is considered as utility function, the Nash outcome is not maximum ratio transmission (MRT), it depends on the SNR and the channel realizations and it can be shown to be almost efficient at all SNR values. All results are illustrated by sample region plots, reaction curves, and average price of anarchy figures.

1. INTRODUCTION

Interference channels are one of the basic elements of complex networks. Future wireless communication systems will suffer from interference since the number of subscribers as well as the required data rate increases. Therefore, it is important to exploit carefully the spatial dimension by using multiple transmit or receive antennas. In the current work, we focus on the two user multiple-input single-output (MISO) interference channel (IFC) [1], i.e., two transmitters with multiple antennas (say each has n_T) want to send their independent and private information to their intended receivers equipped with a single antenna each. We presume that the receivers eavesdrop, i.e. try to decode the message of the other link.

For the MISO IFC without privacy, recent results characterize the Pareto boundary of the achievable rate region for this channel model by a single real valued parameter $\lambda_k, 0 \leq \lambda_k \leq 1$ per user k [2]. The game theoretic interpretation of this parameter λ_k is that it ranges from completely selfish behavior ($\lambda_k = 1$) which is maximum ratio transmission (MRT) to completely altruistic behavior ($\lambda_k = 0$). In the non-cooperative case, it is shown in [3], that the unique Nash equilibrium (NE) is that all users perform MRT. Unfortunately, this NE has poor performance for larger SNR values.

The idea of the current work is to take the privacy of the information into account and reconsider the non-cooperative solution of this two-player game since the utility function has changed. We consider the security on the physical layer: Already Shannon studied the notion of perfect secrecy in his seminal paper [4]. Later, the theoretical basis for an information-theoretic approach was laid first by Wyner [5] and then by Csiszár and Körner [6], who proved in two seminal papers that channel codes exist which guarantee both robustness to transmission errors and a prescribed degree of data confidentiality.

Recently, many research groups began to study resource allocation in wire-tapped fading channels. The interesting observation in [7] for flat-fading is that even if the average channel quality between transmitter and eavesdropper is better than the average channel between transmitter and intended receiver, the secrecy capacity can still be positive. A similar behavior is described in [8, 9] for non-ergodic fading. In multiuser systems, the secrecy capacity region for broadcast channels is studied in [10]. Achievable secrecy rate regions are obtained for the discrete memoryless IFC with private information in [11].

The contribution and organization of the paper is as follows: In Section 2.1, we describe the information theoretic model of the MISO IFC with private information. In Section 2.2, a recent result of an achievable secrecy rate region is specialized to the MISO IFC and the additional degrees of freedom are discussed compared to the single-input single-output (SISO) IFC. The problem statements are given in

The work is in part supported by the German Research Foundation (DFG).

see the program

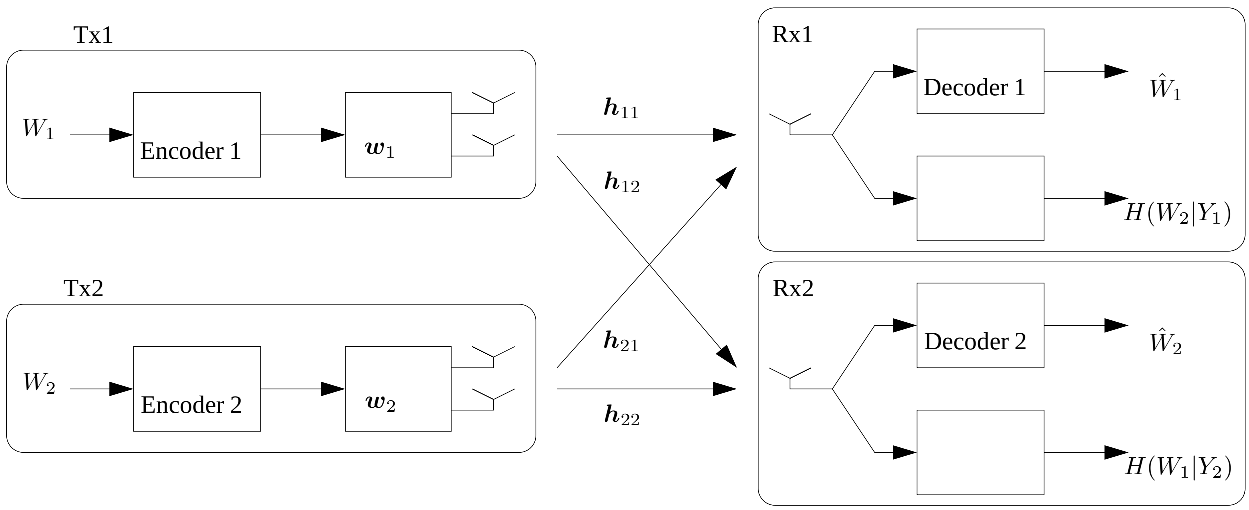


Fig. 1. The two-user MISO interference channel with private messages under study (illustrated for $n = 2$ transmit antennas).

Section 2.3. The Pareto boundary of the MISO IFC secrecy rate region is described in Section 3. The existence and uniqueness of the Nash equilibrium is shown in Section 4.1 where also an iterative algorithm to compute the NE solution is described. In Section 4.2, the efficiency of the NE is characterized using the average price of anarchy (POA). Finally, in Section 4.3, the Stackelberg formulation is briefly discussed. The paper is concluded in Section 5.

2. PRELIMINARIES, MODELS AND BASIC CHARACTERIZATIONS

2.1. Information theoretic model

The transmitter $Tx1$ and $Tx2$ intend to send independent messages $W_1 \in \mathcal{W}_1$ and $W_2 \in \mathcal{W}_2$ to the desired receivers $Rx1$ and $Rx2$, respectively. This is done in n channel uses while ensuring information theoretic security (to be defined rigorously below). The channel realizations $\boldsymbol{h}_{kl}$ for $1 \leq k, l \leq 2$ are quasi-static memoryless block flat-fading with additive white Gaussian noise at the receivers with variance σ_n^2. Both transmitters have a power constraint P. We define ρ as the average SNR, i.e. $\rho = \frac{P}{\sigma_n^2}$. Hence, we can characterize the MISO IFC by its input alphabets $\mathcal{X}_1, \mathcal{X}_2$, its output alphabets $\mathcal{Y}_1, \mathcal{Y}_2$ and its conditional probability distribution $p(\boldsymbol{y}_1, \boldsymbol{y}_2 | \boldsymbol{x}_1, \boldsymbol{x}_2)$ describing the channel. At the transmitters we have additional beamforming vectors $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ with $|\boldsymbol{w}_1|^2 = |\boldsymbol{w}_2|^2 = 1$. We assume that both transmitters have n_T transmit antennas.

In the definitions, we follow closely [11]. At the transmitters we have stochastic encoders mapping from the message sets to codewords, i.e. $f_t(\boldsymbol{x}_t|w_t), \boldsymbol{x}_t \in \mathcal{X}_t^n, w_t \in \mathcal{W}_t$, and $\sum_{\boldsymbol{x}_t \in \mathcal{X}_t^n} f_t(\boldsymbol{x}_t|w_t) = 1$. The decoding functions are mappings $\psi_t : \mathcal{Y}_t^n \mapsto \mathcal{W}_t$, $t \in \{0, 1\}$. The secrecy levels at the receivers are measured by their equivocation rates $\frac{1}{n}H(W_2|Y_1)$ and $\frac{1}{n}H(W_1|Y_2)$, respectively.

An $(M_1, M_2, n, P_e^{(n)})$-Code for the IFC consists of two encoding functions f_1, f_2, two decoding functions ψ_1, ψ_2, and the maximum error probability $P_e^{(n)} = \max\{P_{e,1}^{(n)}, P_{e,2}^{(n)}\}$ with

$$P_{e,t}^{(n)} = \sum_{w_1, w_2} \frac{1}{M_1 M_2} Pr\left[\psi_t(\boldsymbol{y}_t) \neq w_t | (w_1, w_2) \text{ sent}\right]. \quad (1)$$

A rate pair $({}_SR_1, {}_SR_2)$ is achievable if for any $\epsilon_0 > 0$ there is a $(M_1, M_2, n, P_e^{(n)})$-Code such that $M_t \geq 2^{n_S R_t}$ and $P_e^{(n)} \leq \epsilon_0$ and security constraints $n_S R_1 - H(W_1|\boldsymbol{y}_2) \leq n\epsilon_0$, $n_S R_2 - H(W_2|\boldsymbol{y}_1) \leq n\epsilon_0$ are satisfied.

2.2. Preliminaries and basic characterizations

We need a result from [11] which gives an achievable secrecy rate region for the IFC. We will specialize this result to our MISO IFC at hand.

Theorem 1 (Theorem 2 in [11]) *Any rate pair*

$$\begin{aligned} 0 &\leq R_1 \leq I(V_1; Y_1|U) - I(V_1; Y_2|V_2, U) \quad (2) \\ 0 &\leq R_2 \leq I(V_2; Y_2|U) - I(V_2; Y_1|V_1, U) \end{aligned}$$

over all $p(u, v_1, v_2, x_1, x_2, y_1, y_2) = p(u)p(v_1|u)p(v_2|u) \cdot p(x_1|v_1)p(x_2|v_2)p(y_1, y_2|x_1, x_2)$ *is achievable.*

The achievable rates in Theorem 1 correspond to the worst case scenario, i.e., the intended messages are decoded treating the interference as noise but the eavesdropped message is decoded without any interference (which is assumed to be subtracted before).

Lemma 1 *The Gaussian MISO IFC in Fig. 1 has the following achievable secrecy rate pair*

$$
\begin{aligned}
{}_SR_1 &= \underbrace{\log\left(1+\frac{\rho|\boldsymbol{w}_1^H\boldsymbol{h}_{11}|^2}{1+\rho|\boldsymbol{w}_2^H\boldsymbol{h}_{21}|^2}\right)}_{\textit{information term}} \\
&\quad \underbrace{-\log\left(1+\rho|\boldsymbol{w}_1^H\boldsymbol{h}_{12}|^2\right)}_{\textit{secrecy term}} \qquad (3)\\
{}_SR_2 &= \log\left(1+\frac{\rho|\boldsymbol{w}_2^H\boldsymbol{h}_{22}|^2}{1+\rho|\boldsymbol{w}_1^H\boldsymbol{h}_{12}|^2}\right) \\
&\quad -\log\left(1+\rho|\boldsymbol{w}_2^H\boldsymbol{h}_{21}|^2\right). \qquad (4)
\end{aligned}
$$

Proof: We apply Theorem 1 with the choice U is deterministic, $X_1 = V_1$, $X_2 = V_2$ and $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ are independently complex Gaussian distributed with zero-mean and covariance matrix $\boldsymbol{w}_1\boldsymbol{w}_1^H$, $\boldsymbol{w}_2\boldsymbol{w}_2^H$, respectively, i.e. $\boldsymbol{x}_1 \sim \mathcal{CN}(0, \boldsymbol{w}_1\boldsymbol{w}_1^H)$ and $\boldsymbol{x}_2 \sim \mathcal{CN}(0, \boldsymbol{w}_2\boldsymbol{w}_2^H)$. Next, we evaluate the mutual information expressions in (2).

$$
\begin{aligned}
I(X_1;Y_1) &= \log\left(1+\frac{\rho|\boldsymbol{w}_1^H\boldsymbol{h}_{11}|^2}{1+\rho|\boldsymbol{w}_2^H\boldsymbol{h}_{21}|^2}\right), \\
I(X_1;Y_2|X_2) &= \log\left(1+\rho|\boldsymbol{w}_1^H\boldsymbol{h}_{12}|^2\right). \qquad (5)
\end{aligned}
$$

Analog expressions are obtained for the second user which gives (4). ■

This set of achievable secrecy rates follows from the pessimistic assumption that the intended data is disturbed by interference from the other link while the eavesdropper has previously decoded and subtracted its own data (successive interference cancellation) before trying to eavesdrop the other user. The first term in (3) is the achievable rate of user one, and the second term in (3) is the information that user two obtains regarding user one's data. Therefore, these utility functions correspond to a type of worst case scenario as discussed below Theorem 1 already.

Definition 1 *A secrecy rate tuple $({}_SR_1, {}_SR_2)$ is Pareto optimal if there is no other tuple $(\tilde{{}_SR_1}, \tilde{{}_SR_2})$ with*

$$(\tilde{{}_SR_1}, \tilde{{}_SR_2}) \geq ({}_SR_1, {}_SR_2) \textit{ and } (\tilde{{}_SR_1}, \tilde{{}_SR_2}) \neq ({}_SR_1, {}_SR_2).$$

Note that the definition above corresponds to strong Pareto optimality. We will not characterize the weak Pareto optimal boundary of the secrecy rate region. These are the vertical or horizontal lines connecting the region with the axis.

2.3. Problem statements

At first, all users are interested in choosing a beamforming vector which corresponds to an efficient secrecy rate. Therefore, the first problem statement is: *Characterize the Pareto boundary of the achievable secrecy rate region of the two-user MISO IFC.*

Based on the characterization, the next question concerns the non-cooperative behavior of the two transmitters. We assume that they know their channels to receiver one and two perfectly. The second problem statement is: *Does there exist a unique Nash equilibrium? If yes, what are the corresponding beamforming vectors and how efficient is the solution?*

3. PARETO BOUNDARY OF THE ACHIEVABLE SECRECY RATE REGION

The Pareto optimal beamforming vectors are described in the following result. Denote the MRT beamforming vector of user k as $\boldsymbol{w}_k^{\text{MRT}}$ and the zero-forcing (ZF) beamforming vector as $\boldsymbol{w}_k^{\text{ZF}}$.

Theorem 2 *Any point on the Pareto boundary of the secrecy rate region is achievable with the beamforming strategies*

$$
\begin{aligned}
\boldsymbol{w}_1(\lambda_1) &= \frac{\lambda_1\boldsymbol{w}_1^{\text{MRT}}+(1-\lambda_1)\boldsymbol{w}_1^{\text{ZF}}}{||\lambda_1\boldsymbol{w}_1^{\text{MRT}}+(1-\lambda_1)\boldsymbol{w}_1^{\text{ZF}}||} \quad \textit{and}\\
\boldsymbol{w}_2(\lambda_2) &= \frac{\lambda_2\boldsymbol{w}_2^{\text{MRT}}+(1-\lambda_2)\boldsymbol{w}_2^{\text{ZF}}}{||\lambda_2\boldsymbol{w}_2^{\text{MRT}}+(1-\lambda_2)\boldsymbol{w}_2^{\text{ZF}}||} \qquad (6)
\end{aligned}
$$

for some $0 \leq \lambda_1, \lambda_2 \leq 1$.

Proof: The proof works along similar lines as the proof of Proposition 1 and Theorem 1 in [12]. It has two steps. In the first step, we show that all $\boldsymbol{w}_1$ that correspond to points on the Pareto boundary have the form

$$\boldsymbol{w}_1 = \alpha\frac{\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*||} + \beta\frac{\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*||} \qquad (7)$$

with real-valued non-negative $\alpha, \beta : \alpha^2 + \beta^2 = 1$. Any vector $\boldsymbol{w}_1 \in \mathbb{C}^{n_T}$ can be written as

$$\boldsymbol{w}_1 = e^{j\phi}\left(\alpha\frac{\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*||} + \beta e^{j\theta}\frac{\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*||}\right) + \sum_{i=1}^{n_T-2}\lambda_i\boldsymbol{u}_i$$

with $\{\boldsymbol{u}_i\}_{i=1}^{n_T-2}$ is a set of orthogonal vectors that span the orthogonal complement of $\left(\alpha\frac{\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*||}, \frac{\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*||}\right)$. Thus $\left(\alpha\frac{\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}\boldsymbol{h}_{11}^*||}, \frac{\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*}{||\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}^*||}, \boldsymbol{u}_1, ..., \boldsymbol{u}_{n_T-2}\right)$ form a complete orthogonal basis for $\mathbb{C}^{n_T}$.

see the program

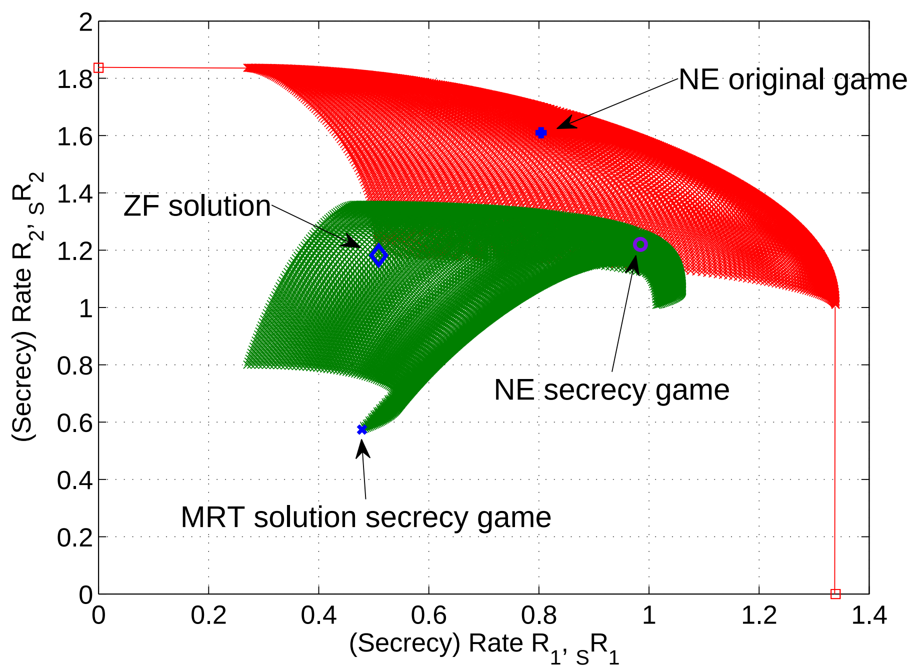


Fig. 2. Achievable rate region (red) and achievable secrecy rate region (green) and some operating points at 0 dB SNR.

We want to show that on the Pareto boundary we must have $\alpha^2+\beta^2=1$ and $\theta=0$. The proof is by contradiction. Consider the desired-signal part in $_SR_1$:

$$|\boldsymbol{w}_1^H\boldsymbol{h}_{11}|^2 = |\alpha\sqrt{\boldsymbol{h}_{11}^H\Pi_{h_{12}^*}\boldsymbol{h}_{11}} + \beta e^{j\theta}\sqrt{\boldsymbol{h}_{11}^H\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}}|^2 \quad (8)$$

and the eavesdrop signal part in $_SR_1$ which is identical to the interference term in the expression for $_SR_2$:

$$|\boldsymbol{w}_1^H\boldsymbol{h}_{12}|^2 = |\alpha|^2\frac{|\boldsymbol{h}_{11}^H\boldsymbol{h}_{12}|^2}{\boldsymbol{h}_{11}^H\Pi_{h_{12}^*}\boldsymbol{h}_{11}}. \quad (9)$$

Suppose we are on the Pareto boundary for some $\alpha,\beta,\lambda_i,\theta$ but $\alpha^2+\beta^2<1$. Then we can increase the magnitude of β and if necessary adjust θ (set $\theta=0$) in order to make (8) larger without increasing (9), i.e., we can increase $_SR_1$ (in two ways: increasing the power of the desired signal part and decreasing the power of the eavesdrop signal part) without decreasing $_SR_2$. This is a contradiction. Next, suppose we are on the Pareto boundary and $\alpha^2+\beta^2=1$ but $\theta\neq 0$. Then again we can increase (8) without decreasing (9). Therefore, we must have at the boundary $\alpha^2+\beta^2=1$, $\theta=0$, and $\lambda_i=0$ for all $i, 1\le i\le n_T-2$.

The second step of the proof maps the parameterization in (7) to the linear combination of the MRT beamforming vector $\boldsymbol{w}_1^{\text{MRT}}$ and the ZF beamforming vector $\boldsymbol{w}_1^{\text{ZF}}$. We define $\ell_1=||\Pi_{h_{12}^*}\boldsymbol{h}_{11}||^2$, $\ell_2=||\Pi_{h_{12}^*}^{\perp}\boldsymbol{h}_{11}||^2$. Note that $\ell_1+\ell_2=||\boldsymbol{h}_{11}||^2$. Then we can write expression (7) as

$$\boldsymbol{w}_1=\alpha\sqrt{\frac{\ell_1+\ell_2}{\ell_1}}\boldsymbol{w}_1^{\text{MRT}}+\left(\sqrt{1-\alpha^2}-\sqrt{\frac{\ell_2}{\ell_1}}\right)\boldsymbol{w}_1^{\text{ZF}} \quad (10)$$

where $\alpha\in[0,1]$. ■

Note that the characterization in Theorem 2 is very helpful for plotting the secrecy rate region as well as for the further analysis. In Fig. 2, we illustrate a secrecy rate region and also the rate region without secrecy constraints.

Fig. 2 compares the achievable rate region (from [2]) with the achievable secrecy rate region based on (3). It can be observed that in the range around the ZF strategies, the difference between the regions is small and at the ZF solution they are identical because the secrecy term in (3) vanishes. In the figure, the NE and ZF solution of the original rate game (utility function is the achievable rate) is shown. For small SNR, it is known that the NE is quite efficient and is close to the Pareto boundary.

4. NON-COOPERATIVE GAMES

The secrecy rate as the utility function in (3) and (4) can also be interpreted in another way. The secrecy term is the penalty term for the user if he acts selfishly. The rate is

see the program

reduced if and only if the users are not completely altruistic (ZF solution). Let us now consider the non-cooperative scenario and study the NE for these new utility functions.

The game is formally defined as

$$\mathcal{G} = \left(\{1,2\}, [0,1]^2, \{{}_SR_1, {}_SR_2\}\right)$$

with the set of players (two players), the strategy space is $[0,1]$ for each player and the utility functions are ${}_SR_1$ and ${}_SR_2$, respectively.

4.1. Existence and uniqueness of Nash equilibrium

Theorem 3 *There exists a unique NE for the MISO IFC secrecy rate game. The beamforming vectors that achieve the NE can be obtained from Algorithm 1.*

Result: Beamforming vectors corresponding to NE
Input: Channel realizations $\boldsymbol{h}_{11}, \boldsymbol{h}_{12}, \boldsymbol{h}_{21}, \boldsymbol{h}_{22}$ and noise variance σ_n^2
initialization: $\lambda_1^0 = 0$ and $\lambda_2^0 = 1$, ${}_SR_1^0 = {}_SR_2^0 = 0$, $\ell = 1$;
while $|{}_SR_1^\ell - {}_SR_1^{\ell-1}| + |{}_SR_2^\ell - {}_SR_2^{\ell-1}| > \epsilon$ **do**
  $\lambda_1^\ell = \arg \max_{0\leq\lambda_1\leq 1} {}_SR_1(\lambda_1, \lambda_2^{\ell-1})$;
  $\lambda_2^\ell = \arg \max_{0\leq\lambda_2\leq 1} {}_SR_2(\lambda_1^\ell, \lambda_2)$;
  $\ell = \ell + 1$;
end
Output: Optimal $\lambda_1^{\ell-1}, \lambda_2^{\ell-1}$

Algorithm 1: Cooperative and centralized beamforming optimization for the two user MISO IFC

Proof: The proof of this result is based on the characterization of the best response of user one to the choice of user two

$$\lambda_1^*(\lambda_2) = \arg \max_{0\leq\lambda_1\leq 1} {}_SR_1(\lambda_1, \lambda_2) \tag{11}$$

and vice versa $\lambda_2^*(\lambda_1)$. The first step is to show that the maximum in (11) is unique. We analyze the utility function ${}_SR_1$ for fixed λ_2 and show that it is unimodal, i.e. it has only one maximum at the mode λ_1^*. Furthermore, this mode can be computed in closed form.

Let us define the following quantities for further use

$$\begin{aligned} \gamma_{11}^2 &= ||\boldsymbol{h}_{11}||^2, & \gamma_{12}^2 &= ||\boldsymbol{h}_{11}^H \Pi^\perp_{\boldsymbol{h}_{12}}||^2 \\ \gamma_{22}^2 &= ||\boldsymbol{h}_{11}||^2, & \gamma_{21}^2 &= ||\boldsymbol{h}_{22}^H \Pi^\perp_{\boldsymbol{h}_{21}}||^2. \end{aligned} \tag{12}$$

Obviously, it holds

$$\gamma_{11} \geq \gamma_{12} \quad \text{and} \quad \gamma_{22} \geq \gamma_{21}. \tag{13}$$

We obtain

$$\begin{aligned} \alpha_1(\lambda_1) &= |\boldsymbol{w}_1^T(\lambda_1)\boldsymbol{h}_{11}|^2 \\ &= \frac{|\left(\lambda_1 \boldsymbol{w}_1^{\text{MRT}} + (1-\lambda_1)\boldsymbol{w}_1^{\text{ZF}}\right)^T \boldsymbol{h}_{11}|^2}{||\lambda_1 \boldsymbol{w}_1^{\text{MRT}} + (1-\lambda_1)\boldsymbol{w}_1^{\text{ZF}}||^2} \\ &= \frac{\left(\lambda_1 ||\boldsymbol{h}_{11}|| + \frac{(1-\lambda_1)}{||\Pi^\perp_{\boldsymbol{h}_{12}} \boldsymbol{h}_{11}||} \boldsymbol{h}_{11}^H \Pi^\perp_{\boldsymbol{h}_{12}} \boldsymbol{h}_{11}\right)^2}{\lambda_1^2 + (1-\lambda_1)^2 + 2\lambda_1(1-\lambda_1)\frac{||\boldsymbol{h}_{11}^H \Pi^\perp_{\boldsymbol{h}_{12}}||}{||\boldsymbol{h}_{11}||}} \\ &= \frac{\lambda_1^2\gamma_{11}^2 + (1-\lambda_1)^2\gamma_{12}^2 + 2\lambda_1(1-\lambda_1)\gamma_{11}\gamma_{12}}{\lambda_1^2 + (1-\lambda_1)^2 + 2\lambda_1(1-\lambda_1)\frac{\gamma_{12}}{\gamma_{11}}} \\ &= \frac{(\lambda_1\gamma_{11} + (1-\lambda_1)\gamma_{12})^2}{1 - 2\lambda_1(1-\lambda_1)(1-\frac{\gamma_{12}}{\gamma_{11}})}. \end{aligned} \tag{14}$$

Similarly, we obtain

$$\begin{aligned} \alpha_2(\lambda_2) &= |\boldsymbol{w}_2^T(\lambda)\boldsymbol{h}_{22}|^2 \\ &= \frac{(\lambda_2\gamma_{22} + (1-\lambda_2)\gamma_{21})^2}{1 - 2\lambda_2(1-\lambda_2)(1-\frac{\gamma_{21}}{\gamma_{22}})}, \end{aligned} \tag{15}$$

$$\begin{aligned} \beta_1(\lambda_1) &= |\boldsymbol{w}_1^T(\lambda)\boldsymbol{h}_{12}|^2 \\ &= \frac{\lambda_1^2\gamma_{11}^2}{1 - 2\lambda_1(1-\lambda_1)(1-\frac{\gamma_{12}}{\gamma_{11}})} \end{aligned} \tag{16}$$

$$\begin{aligned} \beta_2(\lambda_2) &= |\boldsymbol{w}_2^T(\lambda)\boldsymbol{h}_{21}|^2 \\ &= \frac{\lambda_2^2\gamma_{22}^2}{1 - 2\lambda_2(1-\lambda_2)(1-\frac{\gamma_{21}}{\gamma_{22}})}. \end{aligned} \tag{17}$$

Next, the first derivatives with respect to λ_1 or λ_2 are computed directly as

$$\begin{aligned} \frac{d\alpha_1(\lambda_1)}{d\lambda_1} &= \frac{2}{(\sim)^2}(\lambda_1(\gamma_{11}-\gamma_{12}) + \gamma_{12})\gamma_{11}(\gamma_{11}-\gamma_{12}) \\ &\quad \cdot(\gamma_{11}(1-\lambda_1) + \gamma_{12}(1-\lambda_1)) \geq 0 \end{aligned} \tag{18}$$

where the last inequality follows from (13). The monotonicity of $\alpha_2(\lambda_2)$ follows similarly. The first derivatives of $\beta_1(\lambda_1)$ with respect to λ_1 is given by

$$\frac{d\beta_1(\lambda_1)}{d\lambda_1} = \frac{2\lambda_1\gamma_{11}^3(\gamma_{11}(1-\lambda_1) + \lambda_1\gamma_{12})}{(\sim)^2} \tag{19}$$

where $(\sim)$ in (18) and (19) is given by $\gamma_{11} - 2\lambda_1\gamma_{11} + 2\lambda_1\gamma_{12} + 2\lambda_1^2\gamma_{11} - 2\lambda_1^2\gamma_{12}$. The secrecy rate of user one can be written using $\alpha_1, \alpha_2, \beta_1, \beta_2$ for fixed λ_2 as

$${}_SR_1(\lambda_1, \lambda_2) = \log\left(1 + \rho_1\alpha_1(\lambda_1)\right) - \log\left(1 + \rho\beta_1(\lambda_1)\right)$$

where $\rho_1 = \frac{\rho}{1+\rho\beta_2(\lambda_2)}$. The first derivative of ${}_SR_1$ with respect to λ_1 at the point $\lambda_1 = 0$ and $\lambda_1 = 1$ is given by

$$\begin{aligned} \left.\frac{\partial {}_SR_1(\lambda_1,\lambda_2)}{\partial\lambda_1}\right|_{\lambda_1=0} &= 2\frac{\rho_1\gamma_{12}(\gamma_{11}^2 - \gamma_{12}^2)}{\gamma_{11}(1+\rho_1\gamma_{12}^2)} \geq 0, \\ \left.\frac{\partial {}_SR_1(\lambda_1,\lambda_2)}{\partial\lambda_1}\right|_{\lambda_1=1} &= -2\frac{\rho\gamma_{12}^3}{\gamma_{11}(1+\rho\gamma_{12}^2)} \leq 0, \end{aligned} \tag{20}$$

see the program

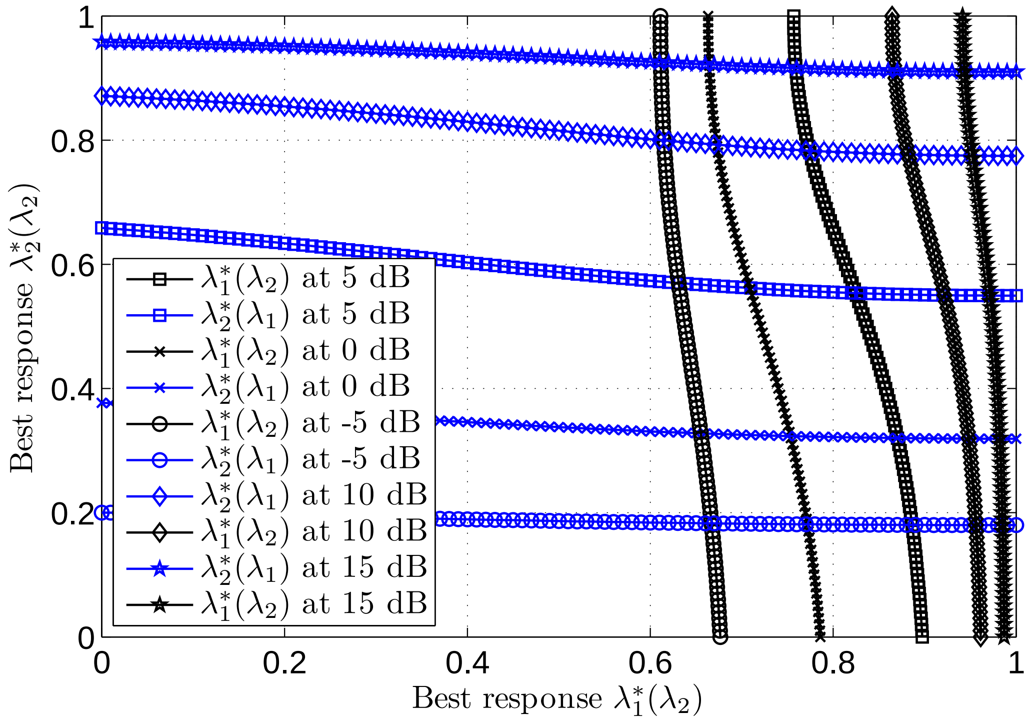


Fig. 3. Reaction curves for different SNR values and a fixed channel realization with $n_T = 2$.

because of (13). Since the first derivative of ${}_SR_1$ with respect to λ_1

$$\frac{\partial_S R_1(\lambda_1,\lambda_2)}{\partial \lambda_1} = \frac{\rho_1 \alpha_1'(\lambda_1)}{1+\rho_1\alpha_1(\lambda_1)} - \frac{\rho \beta_1'(\lambda_1)}{1+\rho\beta_1(\lambda_1)} = 0 \quad (21)$$

has at most two zeros (polynomial of degree two). This observation in connection with (20) proves the unimodality and the uniqueness of the best response. Furthermore, we can solve (21) in order compute the best response in closed form

$$\begin{aligned}\lambda_1^*(\lambda_2) \;=\; & \frac{1}{2}\Big(-2\gamma_{11}^2\gamma_{12}\rho_1 - \rho_1\gamma_{11}\gamma_{12}^2 \qquad (22)\\ & +\rho_1\gamma_{11}^3 - \gamma_{11}\rho_1\gamma_{12}^4\rho - \gamma_{11}\rho\gamma_{12}^2 + 2\rho_1\gamma_{12}^3 \\ & +\Big(-2\rho_1\gamma_{11}^4\rho\gamma_{12}^2 + \rho_1^2\gamma_{11}^6 - 2\rho_1\gamma_{11}^4\gamma_{12}^2 \\ & +\gamma_{11}^2\gamma_{12}^4\rho_1^2 + \gamma_{11}^2\rho^2\gamma_{12}^4 - 4\rho\gamma_{12}^6\rho_1 \\ & +\rho_1^2\rho^2\gamma_{11}^2\gamma_{12}^8 + 6\rho_1\gamma_{11}^2\gamma_{12}^4\rho + 2\rho_1^2\gamma_{11}^4\gamma_{12}^2\rho_2 \\ & +2\gamma_{11}^2\rho_1\gamma_{12}^6\rho^2 - 2\gamma_{11}^2\rho_1^2\gamma_{12}^6\rho\Big)^{1/2}\Big) / \Big(\rho_1\gamma_{11}^3 \\ & +\gamma_{11}^2\rho_1\gamma_{12}^3\rho_2 - \gamma_{11}^2\gamma_{12}\rho_1 - \gamma_{11}\rho\gamma_{12}^2 \\ & -\gamma_{11}\rho_1\gamma_{12}^4\rho - \rho_1\gamma_{11}\gamma_{12}^2 + \rho\gamma_{12}^3 + \rho_1\gamma_{12}^3\Big).\end{aligned}$$

A similar closed form expression can be derived for $\lambda_2^*(\lambda_1)$. Before we complete the proof of Theorem 3, we illustrate the results by the reaction curves in the strategy space of the players [13, Page 33] in Fig. 3 below.

Either by analyzing (22) or by looking at Fig. 3, it can be seen that the best response of user one is monotonically decreasing with λ_2, i.e. $\frac{\partial \lambda_1^*(\lambda_2)}{\partial \lambda_2} \le 0$ and the best response of user two is monotonically decreasing with λ_1, respectively. Note that $\lambda_1^*(\lambda_2), \lambda_2^*(\lambda_1) \in (0,1)$. Therefore, there exists always a unique intersection point $(\bar{\lambda}_1, \bar{\lambda}_2)$ where

$$\lambda_1^*(\bar{\lambda}_2) = \bar{\lambda}_1 \text{ and } \lambda_2^*(\bar{\lambda}_1) = \bar{\lambda}_2. \quad (23)$$

This shows the uniqueness of the NE. In order to show the convergence of the proposed Algorithm 1, we illustrate the convergence steps in Fig. 4 below.

The convergence is shown by observing either in (22) or in Fig. 4 that $\lambda_1^\ell = \lambda_1^*(\lambda_2^*(\lambda_1^{\ell-1}))$ is monotonic increasing whereas $\lambda_2^\ell = \lambda_1^*(\lambda_2^*(\lambda_1^{\ell-1}))$ is monotonic decreasing. Since the strategy set is closed and bounded, the series converges to the unique fixed point. ■

4.2. Efficiency of Nash equilibrium

One characterization of the efficiency of the Nash equilibrium is the so-called Price of Anarchy (PoA) [14, 15]. The PoA measures the cost that a system pays for operating

see the program

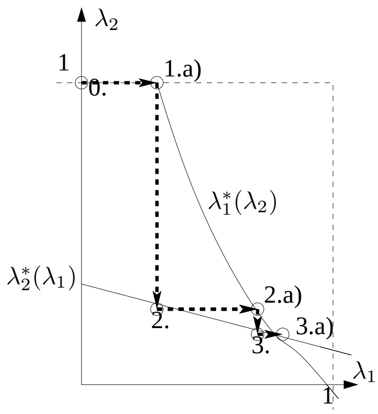


Fig. 4. Reaction curve example with convergence steps of Algorithm 1.

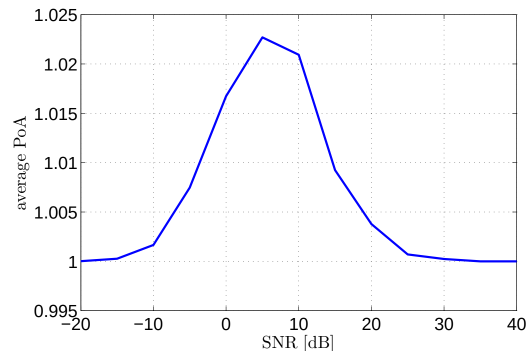


Fig. 5. Average price of anarachy for MISO IFC with two antennas at different SNR.

without cooperation. It is defined as the ratio of the profit obtained at the optimal operating point, over the profit when functioning at the worst-case Nash equilibrium.

Here, we use the utilitarian social welfare function to express the PoA

$$\text{PoA} = \frac{\max\limits_{\lambda_1,\lambda_2} {}_SR_1(\lambda_1,\lambda_2) + {}_SR_2(\lambda_1,\lambda_2)}{{}_SR_1(\lambda_1^*,\lambda_2^*) + {}_SR_2(\lambda_1^*,\lambda_2^*)}. \tag{24}$$

The PoA is always greater than or equal to one. If PoA $= 1$, the NE achieves the utilitarian optimal solution. The PoA can be interpreted as follows: If e.g., PoA $= 2$, the optimal solution is twice as good as the selfish NE solution.

In the case in which the utility functions are the usual rates of the users, the PoA was analyzed in [3] and it was shown that it grows unbounded for increasing SNR. The NE was thus a very inefficient working point for medium to high SNR.

In the case in which we use the achievable secrecy rate as utility function, the efficiency of the PoA is improved. This can be observed in Fig. 2. More quantitatively the behavior is shown in Figure 5. The PoA is shown over the SNR averaged over thousand independent channel realizations.

Interestingly, the efficiency of the NE is good at low and high SNR, i.e., the PoA is close to one. This is a significant difference to the PoA in the case in which the rate region is considered as described above. The interesting point to note here is that the efficiency of the NE depends on the utility function. By carefully selecting the utility function, the NE can be designed to become efficient.

4.3. Stackelberg equilibrium

In the Stackelberg formulation of the conflict, there is an order in which the players have to choose their strategies. The Stackelberg leader moves first. It is assumed that he or she is aware of the rules, and of all necessary information to compute how the follower will decide. Therefore, it is possible for the leader to compute the best response of the follower and hence choose his or her own strategy to maximize his or her own utility.

If player one goes first, the following optimization problem is solved

$$\max_{0\leq\lambda_1\leq 1} {}_SR_1(\lambda_1, \lambda_2^*(\lambda_1)) \tag{25}$$

whereas for the second player first, the optimization problem

$$\max_{0\leq\lambda_2\leq 1} {}_SR_2(\lambda_1^*(\lambda_2), \lambda_2) \tag{26}$$

has to be solved. The simulation results indicate that to be the leader always gives larger secrecy rates than to be the follower. We leave the analysis of the Stackelberg formulation for future research.

5. CONCLUSIONS

The achievable secrecy rate region of the MISO IFC is studied. First, the strong Pareto boundary is characterized and then the non-cooperative game outcome computed by an iterative algorithm. The resulting Nash equilibrium is more efficient than the NE in the case with rates as utilities. The system does not end in a Prisoners dilemma and in fact for all SNR values, the price of anarchy is close to one.

The approach is one way to modify a resource or beamforming game by carefully selecting a suitable utility function. The advantage of the proposed utility function is that it also has an operational meaning in terms of secrecy rate.

see the program

6. ACKNOWLEDGEMENTS

The first author acknowledges fruitful discussions with Erik G. Larsson about the MISO IFC. Both authors thank Siddharth Naik for interesting discussions about the application of non-cooperative game theory to the MISO IFC.

7. REFERENCES

[1] S. Vishwanath and S. A. Jafar, "On the capacity of vector Gaussian interference channels," *IEEE ITW*, 2004.

[2] E. Jorswieck and E. G. Larsson, "Complete characterization of the pareto boundary for the MISO interference channel," *IEEE Trans. on Signal Processing*, vol. 56, no. 10, Oct. 2008.

[3] E. Larsson and E. Jorswieck, "Competition versus collaboration on the MISO interference channel," *IEEE Journal on selected areas in Communications*, vol. 26, no. 7, pp. 1059 – 1069, September 2008.

[4] C. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. Journal*, vol. 28, pp. 656–715, 1949.

[5] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. Journal*, vol. 54, pp. 1355–1387, 1975.

[6] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. on Information Theory*, vol. 24, no. 3, pp. 339–348, Mar. 1978.

[7] M. Bloch, J. Barros, M. R. S. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. on Information Theory*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[8] P. K. Gopala, L. Cai, and H. El Gamal, "On the secrecy capacity of fading channels," *Proc. IEEE ISIT*, 2007.

[9] Y. Liang and H. V. Poor, "Secure communication over fading channels," *Proc. Allerton Conference on Communication, Control and Computing*, 2006.

[10] Y. Liang, H. V. Poor, and S. Shamai, "Secrecy capacity region of fading broadcast channels," *Proc. IEEE ISIT*, 2007.

[11] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. on Information Theory*, vol. 54, no. 6, pp. 2439–2507, June 2008.

[12] E. Jorswieck and E. G. Larsson, "The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves pareto optimality," *Proc. of ICASSP*, 2008, invited.

[13] M. J. Holler and G. Illing, *Einführung in die Spieltheorie*, Springer, 4th edition, 2000.

[14] C. Papadimitriou, "Algorithms, games, and the internet," *Proc. of the ACM symposium on Theory of computing*, pp. 749–753, 2001.

[15] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*, Cambridge University Press, 2007.

see the program

COORDINATION ON THE MISO INTERFERENCE CHANNEL USING THE VIRTUAL SINR FRAMEWORK

Randa Zakhour and David Gesbert

Mobile Communications Department
EURECOM
06560 Sophia Antipolis, France
{zakhour, gesbert}@eurecom.fr

ABSTRACT

This paper addresses the problem of coordination on the interference channel (IC), which has attracted a lot of attention in the research community recently. More precisely, coordinated beamforming is considered, in a multi-cell/link environment where base stations equipped with multiple antennas each attempt to serve a separate user despite the interference generated by the other bases. With single antenna users, this corresponds to the so-called MISO IC considered among others in [1, 2]. In this paper, we propose a distributed approach for designing the beamforming vectors to be used at the transmitters, which relies only on local channel state information (CSI) at each base. The technique exploits a metric which is reminiscent of the virtual uplink method proposed in [1]. We demonstrate analytically the optimality of the proposed approach in terms of achieving the outer bound of the rate region in certain cases. We conduct simulations showing the gains for general settings.

1. INTRODUCTION

The introduction of multiple antennas at transmitters and receivers in communication systems promises great improvements in terms of error resilience and rates achieved. Depending on how much channel state information (CSI) is available at the terminals involved, different degrees of such gains may be achieved in single link transmission, as well as in multiple access (MAC, corresponding to a cellular uplink) and broadcast (BC, corresponding to cellular downlink) channels. In scenarios involving multiple transmitters and receivers, such as a multi-cell scenario, the performance attained will depend on how much information may be shared at the nodes involved. Thus, if either all transmitters or all receivers share their entire data and as a result perform joint transmission or joint decoding respectively, the situation will be equivalent to a BC and a MAC, respectively, for which interference mitigation is well understood. However, if this is not the case (i.e. a distributed optimization scenario where the exchange of CSI among transmitters is limited), then an interference channel (IC) is obtained. This is the situation considered in this paper, as sharing data may put too much strain on the backhaul of the system. More precisely we deal with the downlink direction and propose a transmission strategy based on the so-called "virtual SINR framework" (explained below).

Assuming each transmitter has multiple antennas and each receiver a single antenna, the setting is the MISO interference channel, considered for example in [1, 2] (the more general MIMO IC, which corresponds to receivers also having multiple antennas, is considered in [3, 4], among others). In particular, [2] and subsequent publications [5, 6] of the same authors have focused on the case of two transmitters and full CSI at the transmitters (CSIT). Considering the scenario from the viewpoint of game theory, with transmitters as players, a parametrization of the Pareto boundary of the rate region was found, and different algorithms suggested for finding different points on the boundary. [7] provides a parametrization of the Pareto boundary in a more general case.

Here we argue that it may not be reasonable to assume that all the CSI is shared by all transmitters, and consider the case where each transmitter has local channel CSI knowledge: it only knows the channel between itself and all receivers that are within its range. In a TDD system, this information may be gained from those users' transmission in the uplink. If reciprocity may not be assumed, one could consider that each receiver feeds back his full CSI to his serving base which is partially shared with other base stations, thereby saving on signaling. This scenario has been tackled in [8] where an iterative method is proposed to achieve rates at all receivers involved that are higher than those achieved without cooperation. In contrast, what we develop in the present work is a one-shot algorithm. Given the local information at each transmitter, we propose a simple transmission scheme based on having each transmitter maximize what we refer to as a virtual SINR. For certain choices of parameters, the virtual SINR can be seen as the SINR achieved in the uplink if the same fil-

This work was funded in part by ETRI Korea, the European project COOPCOM and the French project ORMAC.

see the program

ters were used, in the TDD case, or in the virtual uplink (see [1]) in case there is no actual reciprocity.

Organization The rest of the paper is organized as follows. Section 2 defines the system model and performance measures considered. Section 3 introduces the virtual SINR framework. The approach of maximizing a virtual SINR at each transmitter is justified by relating it to uplink-downlink duality and more importantly to the full-CSIT results for the two-link case, in Section 4 below. Based on the given analysis, Section 5 states the proposed algorithm. Simulations in Section 6 show the value of the proposed algorithm in realistic scenarios when more links are considered.

Notation Throughout what follows we use the following common notation. $\mathbb{E}$ denotes statistical expectation. $\mathbb{C}$ denotes the complex number field. Boldface lowercase letters are used to denote vectors, and boldface uppercase denote matrices. $\mathcal{CN}(m, \sigma^2)$ is the probability distribution of a circularly symmetric complex Gaussian random variable of mean m and variance σ^2.

2. SYSTEM MODEL

We consider the MISO interference channel where K transmitters (e.g. base stations in a cellular system) with $N_t \geq 2$ antennas each, each communicate with a single receiver (mobile terminal) having a single antenna. This is illustrated in Figure 1 for $K = 2$, $N_t = 3$.

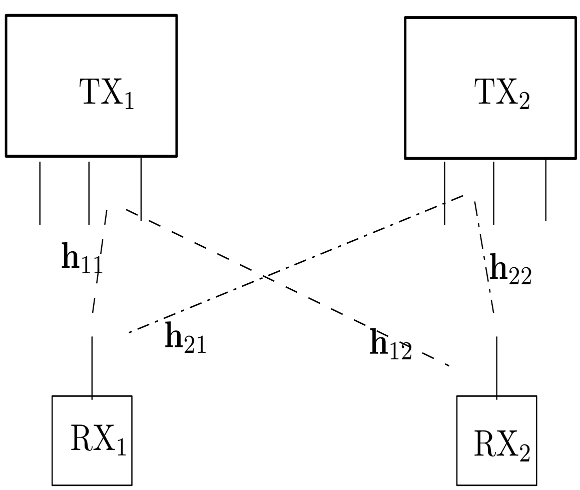


Fig. 1. Scenario considered for $K = 2$, $N_t = 3$. $\mathbf{h}_{11}, \mathbf{h}_{12}$ are known at TX$_1$, $\mathbf{h}_{21}, \mathbf{h}_{22}$ at TX$_2$.

We adopt a narrow-band channel model with frequency-flat block fading. Under linear precoding at each transmitter (no joint multibase precoding since BSs do not share the data symbols), the signal transmitted by base station k, $\mathbf{x}_k$ is given by:

$$\mathbf{x}_k = \sqrt{p_k}\mathbf{w}_k s_k \tag{1}$$

where $s_k \sim \mathcal{CN}(0, 1)$ is the symbol being transmitted intended for user k, $\mathbf{w}_k$ is the unit-norm beamforming vector used to carry this symbol and p_k is the transmit power used. A power constraint holds at each transmitter whereby $p_k \leq P$, P being the peak transmit power at each of the base stations. The signal received at user k is given by:

$$y_k = \sum_{j=1}^{K} \sqrt{p_j}\mathbf{h}_{jk}\mathbf{w}_j s_j + n_k \tag{2}$$

where $\mathbf{h}_{jk} \in \mathbb{C}^{N_t}$ is the channel between that user and base station j, $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the noise at the considered receiver. We assume that receivers have full CSI (CSIR) and do not attempt to decode the interfering signals (single-user decoding). Under these assumptions, the rate achieved at user k is given by:

$$R_k = \log_2(1 + \gamma_k) \tag{3}$$

where the SINR γ_k is equal to:

$$\gamma_k = \frac{p_k|\mathbf{h}_{kk}\mathbf{w}_k|^2}{\sigma^2 + \sum_{j\neq k} p_j|\mathbf{h}_{jk}\mathbf{w}_j|^2} \tag{4}$$

The rate region $\mathcal{R}$ is defined as the set of rates that may be achieved simultaneously at the different base stations, given the power constraints at each base station. I.e.:

$$\mathcal{R} = \{(R_1, \ldots, R_K) \in \mathbb{R}_+^K \\ |R_k \text{ as in (3)}, p_k \leq P \forall k \in \{1, \ldots, K\}\} \tag{5}$$

Beamforming under distributed CSIT

In this work, each transmitter's knowledge is limited to the channel between itself and all users [1]. We would like to achieve a set of rates which is as close as possible to the boundary of $\mathcal{R}$, while yielding the best sum rate possible. Moreover, we would like to do so in a distributed fashion, relying only on *locally available* CSI as just defined, which leads us to an optimization problem solved at each base station within the framework of virtual SINR, detailed in the next section.

3. VIRTUAL SINR

In its most general form, a virtual SINR at base station k is defined as the ratio between the useful signal power received

[1]Strictly speaking, each transmitter only needs to know the channels between itself and users that are close enough to suffer from interference.

see the program

at its served user and the sum of noise plus a weighted sum of the interference powers it causes at the remaining users. Thus:

$$\gamma_k^{virtual} = \frac{p_k|\mathbf{h}_{kk}\mathbf{w}_k|^2}{\sigma^2 + \sum_{j \neq k} \alpha_{kj} p_k |\mathbf{h}_{kj}\mathbf{w}_k|^2}, \qquad (6)$$

where $\alpha_{kj} \in \mathbb{R}_+, j, k = 1, \ldots, K$ are a given set of weights. This can be seen as the SINR achieved on the uplink of a system where at the kth base station, receive vector $\mathbf{w}_k$ is used to process the received signal, mobile station k transmits its signal with power p_k, and mobile $j, \forall j \neq k$ transmit with power $\alpha_{kj} p_k$: the 'virtual uplink' was first introduced in [1] in the context of downlink power control and beamforming in a multicell environment.

When transmitting at full power, Equation (6) becomes:

$$\gamma_k^{virtual} = \frac{|\mathbf{h}_{kk}\mathbf{w}_k|^2}{\frac{1}{\rho} + \sum_{j \neq k} \alpha_{kj} |\mathbf{h}_{kj}\mathbf{w}_k|^2}, \qquad (7)$$

where $\rho = \frac{P}{\sigma^2}$.

As the objective is to have a distributed algorithm which relies only on information local to each base station, we propose that each transmitter solve a virtual SINR maximization problem, which can be stated as follows:

$$\mathbf{w}_k = \arg \max_{\|\mathbf{w}\|^2 = 1} \frac{|\mathbf{h}_{kk}\mathbf{w}_k|^2}{\frac{1}{\rho} + \sum_{j \neq k} \alpha_{kj} |\mathbf{h}_{kj}\mathbf{w}_k|^2}. \qquad (8)$$

This is justified in the following section.

4. ANALYSIS

As first noted in [1], the same rate region may be achieved in the UL (for a reciprocal channel, in the virtual UL otherwise) and DL directions using the same set of vectors for receive and transmit beamforming respectively, but with different power levels in both directions that satisfy the same total power constraint. This is one form of what is referred to as uplink-downlink duality. In what follows, we do not pursue our formulation in the context of duality any further, but use it here to argue that considering virtual SINRs bears some relation to actual SINRs.

4.1. Two-link case

Transmission in the MISO IC may be viewed as a game, where each of the transmitters is a player trying to optimize his rate in some way. One can then define the Pareto boundary of the channel, as the set of Pareto-optimal rate-tuples: thus, a given tuple belongs to the Pareto boundary if it is not possible to increase any rate within that tuple without decreasing at least one of the others. As shown in [2, 7], rates on the Pareto boundary of the MISO interference channel are achieved by transmitting at full power (at least for the case where $N_t \geq K$). Thus, we restrict ourselves to the virtual SINR formulation of (7). We show that, for the two-link case, which is the most well understood, the following two theorems hold.

Theorem 1. *Any point on the Pareto boundary may be attained by solving the virtual SINR optimization problem, as given in* (8)*, for an appropriate choice of* $\alpha_{12}, \alpha_{21} \in \mathbb{R}_+$.

Proof. The proof follows by showing that the same parametrization of the Pareto boundary given in [6] can be retrieved for the beamforming vectors specified in this fashion. Details are given in Appendix A. □

Theorem 2. *The rate pair obtained by beamforming using the solutions to problem* (8) *with* $\alpha_{12} = \alpha_{21} = 1$ *lies on the Pareto boundary of the two-link rate region.*

Proof. See Appendix B. □

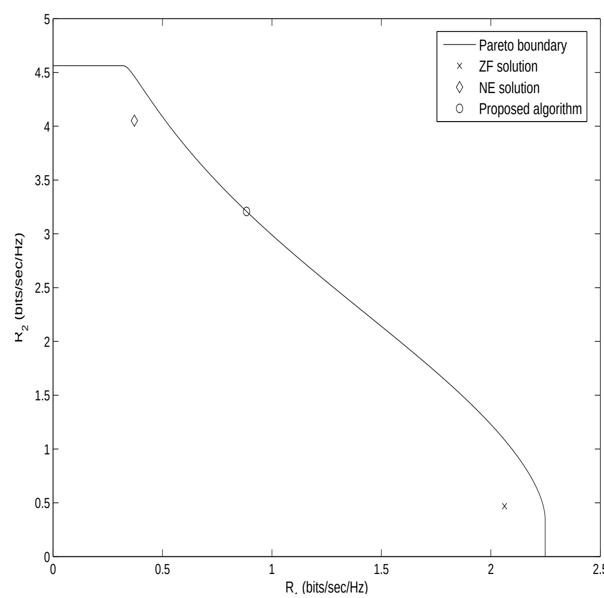


Fig. 2. Pareto rate boundary, MRT, ZF and $\alpha_{12} = \alpha_{21} = 1$ points for a channel instance sampled from a channel with independent identically distributed $\mathcal{CN}(0, 1)$ coefficients, $N_t = 3$, $K = 2$.

This is illustrated in Figure 2, which also shows the rate pairs corresponding to the Nash Equilibrium (NE) or Maximum Ratio Transmission (MRT) and Zero-Forcing (ZF) solutions, which correspond to the most selfish and the most altruistic strategies, respectively, and whose beamforming vectors are given in equation (15) of Appendix A.

see the program

5. PROPOSED ALGORITHM

The performance of the set of precoding vectors obtained in a distributed way by maximizing a virtual SINR at each of the transmitters will depend on the α_{ij}'s selected. Motivated by Theorem 2 above, we propose to set all of these to 1. Thus at base station k:

$$\mathbf{w}_k = \arg \max_{\|\mathbf{w}\|^2=1} \frac{|\mathbf{h}_{kk}\mathbf{w}|^2}{\frac{1}{\rho} + \sum_{j\neq k} |\mathbf{h}_{kj}\mathbf{w}|^2}. \qquad (9)$$

5.1. Two-link case: Further Analysis

Theorem 2 states that the rate pair achieved by setting $\alpha_{12} = \alpha_{21} = 1$ lies on the Pareto boundary of the rate region, but says nothing about the achieved sum rate. An idea of the performance is gained by analyzing the SINR's at low and high SNR, and comparing them with the optimal strategies at those extreme regimes, which are known to be the NE solution and the ZF solution, respectively.
The resulting SINR's from applying our algorithm in the two-link case are given by, where the parameters involved are defined in Appendix A:

$$\gamma_i = \rho \frac{(a_i + b_i(1+c_i))^2}{a_i + b_i(1+c_i)^2} \frac{a_{\bar{i}} + b_{\bar{i}}(1+c_{\bar{i}})^2}{(1+c_{\bar{i}})(a_{\bar{i}} + b_{\bar{i}}(1+c_{\bar{i}}))} \qquad (10)$$

At the ZF solution:

$$\gamma_i^{ZF} = \rho b_i \qquad (11)$$

At the NE solution:

$$\gamma_i^{NE} = \rho \frac{a_i + b_i}{1 + \frac{a_{\bar{i}}}{a_{\bar{i}}+b_{\bar{i}}} c_{\bar{i}}} \qquad (12)$$

At low SNR, $1 + c_i \approx 1$, and both γ_i and γ_i^{NE} may be approximated by $\rho(a_i + b_i)$. On the other hand, at high SNR, $\gamma_i \approx \rho b_i = \gamma_i^{ZF}$. Thus, at both extremes, this approach performs as good as the best out of these two schemes.

5.2. Comparison with Full CSIT Case

We would like to have an idea of the loss due to the distributed nature of our algorithm. The simplest way to define loss is in terms of total power consumption: in our algorithm, all transmitters always use full power. Alternatively, if full CSIT was available at all transmitters or these were allowed to share channel information, then it may be possible to achieve the same rates at all users with a lower total transmit power. The corresponding optimization problem may then be formulated as:

$$\text{minimize} \sum_{k=1}^{K} \mathbf{u}_k^H \mathbf{u}_k \qquad (13)$$

$$\text{subject to } \mathbf{u}_k^H \mathbf{u}_k \leq P, \quad k = 1, \ldots, K$$

$$\frac{|\mathbf{h}_{kk}\mathbf{u}_k|^2}{\sigma^2 + \sum_{j\neq k} |\mathbf{h}_{jk}\mathbf{u}_j|^2} \geq \gamma_k^{(alg)}, \quad k = 1, \ldots, K$$

Parameter	Value
Path loss model	Cost-231, small/medium city
K	3, 7
Cell radius	1000 m
Transmit antenna gain, G_{tx}	16 dB
Shadowing mean	0 dB
Shadowing variance	10 dB
Receive antenna gain, G_{rx}	6 dB
Edge SNR	0-15 dB

Table 1. Simulation setup parameters

where, in relation to our previous notation, $\mathbf{u}_k = \sqrt{p_k}\mathbf{w}_k$, and $\gamma_k^{(alg)}$ is the SINR achieved at user k when beamforming is done using our algorithm.
This can easily be shown to be a convex optimization problem (see [9] for example), and as such algorithms exist to solve it.

6. NUMERICAL RESULTS

Figure 3 illustrates the performance of our approach in a more realistic scenario, the parameters of which are specified in Table 1. User locations in a cell follow a uniform distribution.

We show the average sum rates achieved and compare them with the selfish scheme corresponding to MRT, and the altruistic scheme corresponding to minimizing the total interference caused to other users. Our scheme clearly surpasses both. As the figure illustrates the 7-cell case, and there are always fewer antennas than that, interference caused can never be eliminated completely and eventually the rates would saturate. However, for the cell edge SNR range considered the performance gains are still quite significant. Note that, when $N_t \geq K$, interference can be eliminated completely and at high SNR the rates achieved with our scheme and the interference minimizing solution would differ by at most a constant (in favor of our scheme).

Figure 4 illustrates the power loss due to the distributed nature of our scheme, again for the 7-cell case. More precisely, for different number of antennas, the minimum power needed, under the individual power constraints at each base station, to achieve the same rates as those achieved by our distributed algorithm is computed and the figure illustrates the difference between this power and the power consumed by our scheme as a percentage of the total power available across the system. As the number of antennas at each transmitter increases, the difference decreases: this is because with more degrees of freedom afforded by the higher number of antennas, even in our distributed algorithm, more efficient power use is done automatically thereby reducing the benefit of centralized knowledge, at least for achieving the same rates as our algorithm. From our other simulations, not shown here, for the 3-cell case, for $N_t \geq 3$, this difference is almost negligible, which leads us to conjecture that we are quite close,

see the program

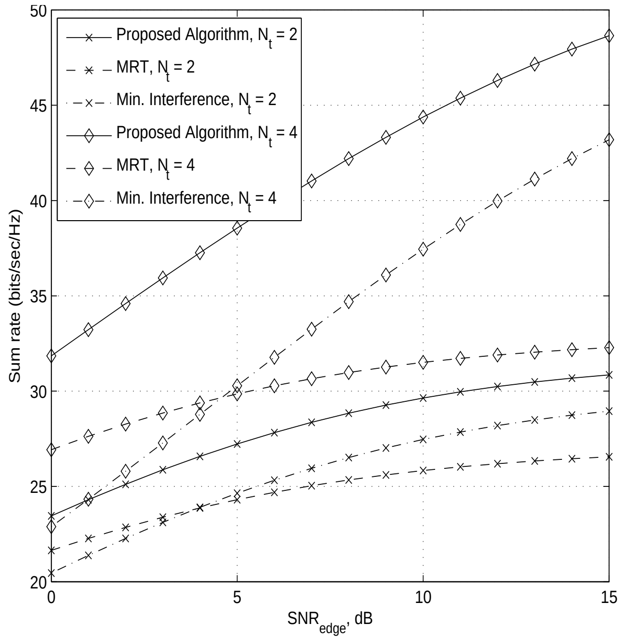


Fig. 3. Sum rates vs. cell-edge SNR for $N_t = 2, 4$, for the 7-cell case.

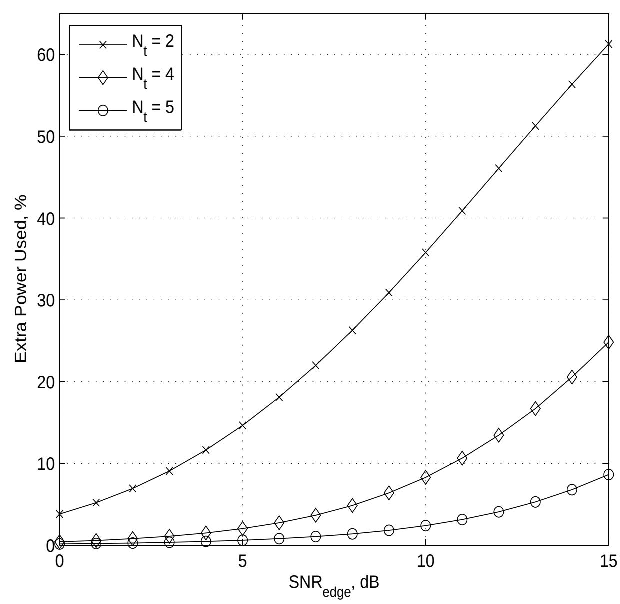


Fig. 4. Extra power consumed due to the distributed nature of our scheme vs. cell-edge SNR for $N_t = 2, 4, 5$, for the 7-cell case.

if not on, the Pareto boundary in this case, for most channel instances.

7. CONCLUSION

In this paper, a distributed coordinated beamforming approach, based on partial CSI at each BS, was proposed. Analytical justification for the algorithm was given for the case of two base stations, and numerical simulations illustrated its performance for more general cases.

A. PROOF OF THEOREM 1

To simplify expressions, in what follows $\bar{i}$ is used to denote the 'other' user/base station index (i.e., $\bar{i} = \mod(i, 2) + 1$, for $i \in \{1, 2\}$).

From Theorem 1 from [6], for the two-link case, any point on the Pareto boundary is achievable with the beamforming strategies:

$$\mathbf{w}_i(\lambda_i) = \frac{\lambda_i \mathbf{w}_i^{NE} + (1-\lambda_i)\mathbf{w}_i^{ZF}}{\|\lambda_i \mathbf{w}_i^{NE} + (1-\lambda_i)\mathbf{w}_i^{ZF}\|}, i = 1, 2 \tag{14}$$

for some $0 \le \lambda_i \le 1$, where

$$\mathbf{w}_i^{NE} = \frac{\mathbf{h}_{ii}^H}{\|\mathbf{h}_{ii}\|} \quad \text{and} \quad \mathbf{w}_i^{ZF} = \frac{\Pi_{i\bar{i}}^{\perp}\mathbf{h}_{ii}^H}{\|\Pi_{i\bar{i}}^{\perp}\mathbf{h}_{ii}^H\|} \tag{15}$$

are the NE or MRT and ZF solutions, respectively. $\mathbf{\Pi}_{i\bar{i}}^{\perp}$ is the projection matrix onto the null space of $\mathbf{h}_{i\bar{i}}$.

We show that the rate region achieved by this parametrization of the Pareto boundary can also be achieved by varying the α's in their feasible region ($\mathbb{R}_+$) and maximizing the corresponding virtual SINRs.
Maximizing the virtual SINR of (7):

$$\mathbf{w}_i = \arg \max_{\|\mathbf{w}\|^2=1} \frac{|\mathbf{h}_{ii}\mathbf{w}|^2}{1/\rho + \alpha_{i\bar{i}}|\mathbf{h}_{i\bar{i}}\mathbf{w}|^2}. \tag{16}$$

Proposition 1. *The solution of problem* (16) *can be written as:*

$$\mathbf{w}_i = \sqrt{\zeta_i}\frac{\Pi_{i\bar{i}}\mathbf{h}_{ii}^H}{\|\Pi_{i\bar{i}}\mathbf{h}_{ii}^H\|} + \sqrt{1-\zeta_i}\frac{\Pi_{i\bar{i}}^{\perp}\mathbf{h}_{ii}^H}{\|\Pi_{i\bar{i}}^{\perp}\mathbf{h}_{ii}^H\|} \tag{17}$$

where $0 \le \zeta_i \le 1, i = 1, 2$.

Proof. Similar to that of Proposition 1 in [6]. □

Define:

$$\begin{aligned} a_i &= \|\Pi_{i\bar{i}}\mathbf{h}_{ii}^H\|^2 \\ b_i &= \|\Pi_{i\bar{i}}^{\perp}\mathbf{h}_{ii}^H\|^2 \\ c_i &= \rho\|\mathbf{h}_{i\bar{i}}\|^2 \end{aligned} \tag{18}$$

Proposition 2. ζ_i *that solves* (16) *is given by:*

$$\zeta_i = \frac{a_i}{a_i + b_i(1+\alpha_{i\bar{i}}c_i)^2} \tag{19}$$

see the program

Proof. With $\mathbf{w}_i$ as in (17),

$$|\mathbf{h}_{ii}\mathbf{w}_i|^2 = \left| \sqrt{\zeta_i}\frac{\mathbf{h}_{ii}\Pi_{i\bar{i}}\mathbf{h}_{ii}^H}{\|\Pi_{i\bar{i}}\mathbf{h}_{ii}^H\|} + \sqrt{1-\zeta_i}\frac{\mathbf{h}_{ii}\Pi_{i\bar{i}}^{\perp}\mathbf{h}_{ii}^H}{\|\Pi_{i\bar{i}}^{\perp}\mathbf{h}_{ii}^H\|}\right|^2 = \left(\sqrt{a_i\zeta_i} + \sqrt{b_i(1-\zeta_i)}\right)^2 \quad (20)$$

Similarly,

$$|\mathbf{h}_{i\bar{i}}\mathbf{w}_i|^2 = \zeta_i\|\mathbf{h}_{i\bar{i}}\|^2 \quad (21)$$

Thus the virtual SINR is equal to:

$$\gamma_i^{virtual} = \frac{\rho\left(\sqrt{a_i\zeta_i} + \sqrt{b_i(1-\zeta_i)}\right)^2}{1+\alpha_{i\bar{i}}\zeta_i c_i} \quad (22)$$

One can easily verify that this ratio is maximized for the value specified in (19). □

Proposition 3. *In terms of the NE and ZF beamforming vectors,* (17) *can be rewritten as:*

$$\mathbf{w}_i = \frac{\lambda_i\mathbf{w}_i^{NE} + (1-\lambda_i)\mathbf{w}_i^{ZF}}{\|\lambda_i\mathbf{w}_i^{NE} + (1-\lambda_i)\mathbf{w}_i^{ZF}\|} \quad (23)$$

where

$$\lambda_i = \frac{1}{\sqrt{\frac{a_i}{a_i+b_i}}\left(\frac{1}{\zeta_i}-1\right) + \left(1-\sqrt{\frac{b_i}{a_i+b_i}}\right)} \quad (24)$$

Proof. $\mathbf{w}_i$, as expressed by (17), can be rewritten in terms of $\mathbf{w}_i^{NE}$ and $\mathbf{w}_i^{ZF}$ as:

$$\mathbf{w}_i = \sqrt{\zeta_i}\sqrt{\frac{a_i+b_i}{a_i}}\mathbf{w}_i^{NE} + \left[\sqrt{1-\zeta_i} - \sqrt{\zeta_i}\sqrt{\frac{b_i}{a_i}}\right]\mathbf{w}_i^{ZF} \quad (25)$$

We need to show that this is in fact of the form given in (23), i.e. that the following equalities hold for some λ_i:

$$\frac{\lambda_i^2}{\lambda_i^2\frac{a_i}{a_i+b_i} + \left(1-\lambda_i+\lambda_i\sqrt{\frac{b_i}{a_i+b_i}}\right)^2} = \zeta_i\frac{a_i+b_i}{a_i}, \text{ and}$$

$$\frac{1-\lambda_i}{\sqrt{\lambda_i^2\frac{a_i}{a_i+b_i} + \left(1-\lambda_i+\lambda_i\sqrt{\frac{b_i}{a_i+b_i}}\right)^2}} = \sqrt{1-\zeta_i} - \sqrt{\zeta_i}\sqrt{\frac{b_i}{a_i}}$$

where we replaced the denominator of (23) by its value in terms of the parameters defined in (18).

One can verify that λ_i as given by (24) above satisfies both these equations. □

Combining propositions 2 and 3, we complete the proof. Plugging (19) into (24), we get:

$$\lambda_i = \frac{1}{\alpha_{i\bar{i}}c_i\sqrt{\frac{b_i}{a_i+b_i}} + 1} \quad (26)$$

Clearly this is a decreasing function of $\alpha_{i\bar{i}}$. It is easy to check that for $\alpha_{i\bar{i}} = 0$, $\lambda_i = 1$ and that as $\alpha_{i\bar{i}} \to \infty$, $\lambda_i \to 0$.

B. PROOF OF THEOREM 2

For any $\mathbf{w}_1, \mathbf{w}_2$ of the form (17), one can show that SINRs are given by:

$$\gamma_i = \rho\frac{(\sqrt{a_i\zeta_i} + \sqrt{b_i(1-\zeta_i)})^2}{1+\zeta_{\bar{i}}c_{\bar{i}}} \quad (27)$$

A rate pair is Pareto optimal if one cannot increase one of the rates without necessarily decreasing the other. Note that any point on Pareto boundary has to have the corresponding (ζ_1, ζ_2) pair in the region defined by $\zeta_i \in \left[0, \frac{a_i}{a_i+b_i}\right], i = 1, 2$: this is so since for higher ζ_i it is always possible to achieve higher useful signal at user i while causing less interference at user $\bar{i}$ (cf. (27)).

Denote by $\gamma_i^{1,1}$ the SINR values achieved by setting $\alpha_{12} = \alpha_{21} = 1$. To show that the corresponding rates belong to the Pareto boundary, we solve the following optimization problem:

$$\begin{aligned} &\text{maximize } \gamma_1 && (28)\\ &\text{such that } 0 \le \zeta_i \le \frac{a_i}{a_i+b_i}, \quad i = 1, 2\\ &\gamma_2 \ge \gamma_2^{1,1} \end{aligned}$$

This can be formalized as the following convex optimization problem:

$$\begin{aligned} &\text{minimize } -t && (29)\\ &\text{such that } 0 \le \zeta_i \le \frac{a_i}{a_i+b_i}, \quad i = 1, 2\\ &t \ge 0\\ &\gamma_2^{1,1}(1+\zeta_1c_1) - \rho\left(\sqrt{a_2\zeta_2} + \sqrt{b_2(1-\zeta_2)}\right)^2 \le 0\\ &t(1+\zeta_2c_2) - \rho\left(\sqrt{a_1\zeta_1} + \sqrt{b_1(1-\zeta_1)}\right)^2 \le 0 \end{aligned}$$

This problem is strictly feasible and consequently Slater's condition for strong duality holds [10].

Let $\mu_i, i = 1, \ldots, 3$ be the Langrange multipliers associated with the positivity constraints, $\xi_i, i = 1, 2$ the Lagrange multipliers associated with the upper bounds on the ζ_i, and $\lambda_i, i = 1, 2$ the Lagrange multipliers associated with the SINR constraints, the corresponding Karush-Kuhn-Tucker

see the program

(KKT) conditions [10] are given by:

$$
\begin{aligned}
&-1-\mu_3+\lambda_2(1+\zeta_2 c_2)=0\\
&-\mu_1+\xi_1+\lambda_1\gamma_2^{1,1}c_1\\
&\quad=\lambda_2\rho\left(a_1-b_1+\sqrt{a_1b_1}\left(\sqrt{\frac{1-\zeta_1}{\zeta_1}}-\sqrt{\frac{\zeta_1}{1-\zeta_1}}\right)\right)\\
&-\mu_2+\xi_2+\lambda_2 t c_2\\
&\quad=\lambda_1\rho\left(a_2-b_2+\sqrt{a_2b_2}\left(\sqrt{\frac{1-\zeta_2}{\zeta_2}}-\sqrt{\frac{\zeta_2}{1-\zeta_2}}\right)\right)\\
&\quad\mu_1,\mu_2,\mu_3,\lambda_1,\lambda_2,\xi_1,\xi_2\geq 0\\
&\quad\mu_i\zeta_i=0,\quad \xi_i\left[\zeta_i-\frac{a_i}{a_i+b_i}\right]=0, i=1,2\\
&\quad\mu_3 t=0\\
&\quad\lambda_1\left[\gamma_2^{1,1}\left(1+\zeta_1c_1\right)-\rho\left(\sqrt{a_2\zeta_2}+\sqrt{b_2(1-\zeta_2)}\right)^2\right]=0\\
&\quad\lambda_2\left[t\left(1+\zeta_2c_2\right)-\rho\left(\sqrt{a_1\zeta_1}+\sqrt{b_1(1-\zeta_1)}\right)^2\right]=0
\end{aligned}
\tag{30}
$$

For $\zeta_i, i=1,2$ given by (19), with $\alpha_{i\bar{i}}=1$, one can verify that these values, together with the values of t and the Lagrange multipliers given in equation (31) below provide a consistent solution of the KKT conditions. This guarantees optimality. Noting that the optimal value of problem (28) is indeed that achieved by our algorithm completes the proof.

$$
\begin{aligned}
&\mu_1=\mu_2=\mu_3=0, \xi_1=\xi_2=0,\\
&\lambda_2=\frac{1}{1+\zeta_2c_2},\\
&\lambda_1=\frac{1}{1+\zeta_2c_2}\frac{(a_1+b_1(1+c_1))^2(a_2+b_2(1+c_2)^2)}{(a_2+b_2(1+c_2))^2(a_1+b_1(1+c_1)^2)},\\
&t=\gamma_1^{1,1}.
\end{aligned}
\tag{31}
$$

C. REFERENCES

[1] F. Rashid-Farrokhi, K.J.R. Liu, and L. Tassiulas. Transmit beamforming and power control for cellular wireless systems. *IEEE Journal on Selected Areas in Communications*, 16:1437–1450, October 1998.

[2] E.G. Larsson and E.A. Jorswieck. The MISO interference channel: Competition versus collaboration. In *Proc. Allerton Conference on Communication, Control and Computing*, September 2007.

[3] G. Scutari, D. P. Palomar, and S. Barbarossa. Asynchronous iterative water-filling for gaussian frequency-selective interference channels. *IEEE Transactions on Information Theory*, 54:2868–2878, July 2008.

[4] V. R. Cadambe and S. A. Jafar. Interference alignment and the degrees of freedom for the K user interference channel. *IEEE Transactions on Information Theory*, 54:3425–3441, August 2008.

[5] E.G. Larsson and E.A. Jorswieck. Competition versus cooperation on the MISO interference channel. *IEEE Journal on Selected Areas in Communications*, 26:1059–1069, September 2008.

[6] E.A. Jorswieck and E.G. Larsson. The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves Pareto optimality. In *Proc. Int'l Conf. Acoustics, Speech and Sig. Proc. (ICASSP)*, Las Vegas, March 31 - April 4 2008.

[7] E.A. Jorswieck and E.G. Larsson. Complete characterization of the Pareto boundary for the MISO interference channel. *IEEE Transactions on Signal Processing*, 56:5292–5296, October 2008.

[8] K. Ming Ho and D. Gesbert. Spectrum sharing in multiple antenna channels: A distributed cooperative game theoretic approach. In *Proc. IEEE International Symposium on Personal, Indoor, Mobile Radio Communications (PIMRC)*, Cannes, 15-18 September 2008.

[9] Wei Yu and T. Lan. Transmitter optimization for the multi-antenna downlink with per-antenna power constraints. *IEEE Transactions on Signal Processing*, 55:2646–2660, June 2007.

[10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, first edition, 2004.

see the program

LIMITED TRANSMITTER COOPERATION IN ADJACENT BROADCAST CHANNELS

David A. Schmidt, Wolfgang Utschick

Associate Institute for Signal Processing, Technische Universität München
80290 Munich, Germany, Telephone: +49 89 289-28508
{dschmidt|utschick}@tum.de

ABSTRACT

We discuss the scenario of two interfering broadcast channels that are not able to jointly precode the data symbols, but are allowed to cooperate in choosing their precoders. The multiplexing gain achievable by zero-forcing with this type of limited cooperation is derived. When the transmitters do not cooperate at all, but compete for the resources, the system is shown to be interference limited and the multiplexing gain is always zero. We furthermore present a distributed algorithm based on exchanging *interference prices*, that achieves the full multiplexing gain, but also performs well if the additive noise is not asymptotically low. It is numerically shown that limited cooperation yields great performance gains over competitive behavior if the interference is strong compared to the additive noise.

1. INTRODUCTION

In the downlink of a cellular system, inter-cell interference can be a severely limiting factor. Conventionally, the base stations do not cooperate in order to achieve a system-wide optimal transmission strategy; at most, they are able to react to the interference the mobiles experience from neighboring base stations and adjust their transmission strategies accordingly, with the goal of better serving their own associated mobile receivers. The interference caused to other cells is not considered. Thus, the base stations are acting on conflicting interests, when, from the point of view of the network operator, they should be cooperating towards a common good.

To overcome this, a *fully cooperative* approach has been proposed (e. g., [1] and the references therein): multiple base stations are connected by means of a highly performant backbone network and are thereby able to jointly precode the transmit symbols for all associated mobiles, thus essentially forming one transmitter with antennas located at many different sites.

While full cooperation yields great performance gains and may be an option in some cases, the existing backbone network will often not allow for every bit of information intended for every mobile to be available at each base station for joint precoding. We therefore investigate approaches that only require *limited cooperation*, where the base stations exchange channel state information and may cooperate in choosing their transmission strategy, but are not able to jointly process the actual transmit symbols. One such scheme is explored in [6], where the interference is *aligned* in certain subspaces at the receivers. This technique, however, requires multiple degrees of freedom at the receivers, whereas in our scenario we assume simple single-antenna receivers.

In this paper, we instead propose a limited cooperation scheme based on exchanging *interference prices* (cf. [2, 3, 4, 5]) and demonstrate how allowing for cooperation in the precoder design leads to a sum throughput which—in contrast to non-cooperative approaches—grows without bound for increasing *signal to noise ratio* (SNR).

2. SYSTEM MODEL

We consider a system consisting of two interfering broadcast channels, as depicted in Fig. 1. Transmitter i has N_i antennas and M_i associated single-antenna receivers. The channel coefficients between the N_i antennas of transmitter i and the m-th receiver associated to transmitter ℓ are collected in the row vector $\boldsymbol{h}_{\ell i}^{m,\mathrm{T}} \in \mathbb{C}^{1\times N_i}$. The additive white Gaussian noise at each receiver has the power σ^2.

The i-th transmitter precodes the unit-variance data symbols for its m-th associated receiver with the vector $\boldsymbol{t}_i^m \in \mathbb{C}^{N_i}$. All precoding vectors of one transmitter are subject to the sum power constraint

$$\sum_{m=1}^{M_i} \|\boldsymbol{t}_i^m\|^2 \leq P_i.$$

We assume linear precoding only (i. e., no *dirty paper coding*), therefore the *signal to interference plus noise ratio*

see the program

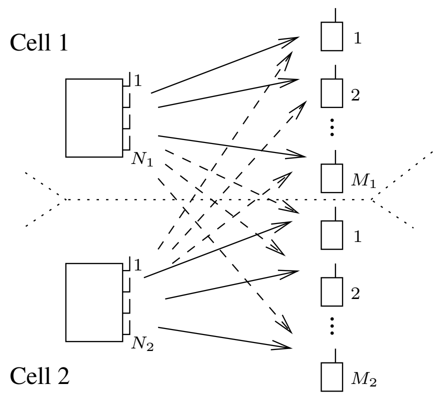


Fig. 1. Scenario with Two Interfering Broadcast Channels

(SINR) at receiver m associated to transmitter i is

$$\begin{aligned}\mathrm{SINR}_i^m &= \frac{|\boldsymbol{h}_{ii}^{m,\mathrm{T}}\boldsymbol{t}_i^m|^2}{\sum_{k\neq m}|\boldsymbol{h}_{ii}^{m,\mathrm{T}}\boldsymbol{t}_i^k|^2 + \sum_k|\boldsymbol{h}_{ij}^{m,\mathrm{T}}\boldsymbol{t}_j^k|^2 + \sigma^2} \\ &= \frac{S_i^m}{I_i^m},\end{aligned}$$

where we abbreviated the received signal power with S_i^m and the sum of all interference and noise powers with I_i^m. In the above equation and throughout the rest of the paper, we use j to denote the transmitter which is not i, i. e., $j = 2$ if $i = 1$ and vice versa. Our figure of merit is the achievable throughput

$$R_i^m = \log(1 + \mathrm{SINR}_i^m) = \log(S_i^m + I_i^m) - \log I_i^m, \quad (1)$$

where Gaussian codebooks are assumed and interference cancellation at the receivers is not possible.

3. ACHIEVABLE DEGREES OF FREEDOM

The *degrees of freedom* (DoF) of a system are defined as the slope of the sum capacity curve over the logarithm of the SNR for asymptotically high SNR, e. g., [7, 8]. The DoF can also be interpreted as the number of independent data streams that can be transmitted in parallel at high SNR.

One way to analyze the DoF of a system is to transform the fully coupled system into a system of parallel channels by means of linear *zero-forcing* (ZF) at either transmitters, receivers, or both. In our case, ZF is only possible at the transmitter side, as the receivers only have a single antenna. For the following derivation, we use the fact that a transmitter with N antennas is able to serve at most N single antenna receivers in parallel using ZF precoding. Note that ensuring that a receiver from the other cell receives no interference also counts as 'serving' this receiver.

We begin by assuming $N_1 \leq N_2$, w. l. o. g. The transmitter in the first cell transmits separate data streams to $0 < n_1 \leq \min(N_1, M_1)$ receivers, and thus is able to keep $N_1 - n_1$ receivers from the second cell free of interference. The second transmitter can now serve at most $n_{2,\max} = \max(\min(N_2 - n_1, N_1 - n_1, M_2), 0)$ of its receivers, where the term $N_2 - n_1$ ensures that the n_1 active receivers of the first cell can still be kept free of interference.

For the total number of receivers that can be served depending on the choice of n_1, we consequently obtain

$$\begin{aligned}n_1 + n_{2,\max} &= \max(\min(N_2, N_1, M_2 + n_1), n_1) \\ &= \min(N_1, M_2 + n_1)\end{aligned}$$

where the second equality is due to $n_1 \leq N_1 \leq N_2$. By choosing $n_1 = \min(N_1, M_1)$ we see that we can serve at most $\min(N_1, M_1 + M_2)$ receivers in parallel. By symmetry, for the case of $N_1 > N_2$, we can serve $\min(N_2, M_1 + M_2)$ receivers. Consequently, by employing ZF at the transmitters,

$$\eta_{\mathrm{AdjBC}} = \min(N_1, N_2, M_1 + M_2)$$

parallel, interference-free data streams can be transmitted. η_{AdjBC} thus is a lower bound on the DoF of our system.

Note that we assumed that the transmitter with the lower number of antennas transmits at least one data stream ($n_1 > 0$ in the above derivation). If $n_1 = 0$, however, $\min(N_2, M_2)$ independent data streams can be transmitted, which may be higher than η_{AdjBC}, as in the simple example $N_1 = M_1 = 1$ and $N_2 = M_2 = 2$. In this case it is advantageous in terms of total system performance (at high SNR) to completely shut off the first cell. Therefore, if we also allow the option of shutting off one transmitter, a multiplexing gain of

$$\max(\eta_{\mathrm{AdjBC}}, \min(N_1, M_1), \min(N_2, M_2))$$

is achievable.

It remains to be shown that the DoF of our system are information theoretically upper bounded by this expression. It should be noted, however, that for both the MIMO interference channel and the MIMO broadcast channel, the multiplexing gain achievable by ZF has been shown to coincide with the information theoretical DoF.

If the transmitters are able to fully cooperate on a symbol level and thus form a broadcast channel with $N_1 + N_2$ transmit antennas and M_1+M_2 single antenna receivers, the expression for the DoF is $\min(N_1 + N_2, M_1 + M_2)$, which is in general higher than above expression for our system with separate precoding of the data streams.

4. NON-COOPERATIVE APPROACH

In this section, we view the two transmitters as uncooperative entities with conflicting interests, where it is the goal of

see the program

transmitter i to maximize its own sum throughput:

$$\max_{\boldsymbol{t}_i^1,\ldots,\boldsymbol{t}_i^{M_i}} \sum_{m=1}^{M_i} R_i^m \qquad \text{s.t.:} \qquad \sum_{m=1}^{M_i} \|\boldsymbol{t}_i^m\|_2^2 \le P_i. \qquad (2)$$

Recall that R_i^m is function of SINR_i^m and thus also depends on the interference power received from transmitter j. We assume that the magnitude of the interference power experienced at the receivers is known to the transmitter and can be taken into account when solving the optimization (2).

When transmitter i adjusts its transmit strategy $\boldsymbol{t}_i^1,\ldots,\boldsymbol{t}_i^{M_i}$ according to (2), the interference power experienced at the receivers of cell j changes, necessitating transmitter j to react by readjusting $\boldsymbol{t}_j^1,\ldots,\boldsymbol{t}_j^{M_j}$. We thus consider an iterative process in which the transmitters alternatingly optimize their strategies taking into account the current interference conditions caused by the other transmitter.

The problem of maximizing the sum throughput of a broadcast channel with linear precoding (2), as must be solved in each step of our iterative procedure, is non-convex and cannot be solved analytically. A number of iterative and successive approaches exist that find good suboptimal precoding vectors (e. g., [9, 10, 11, 12]). In order to avoid nested iterative loops, we employ the successive LISA algorithm from [12], as well as the ZFS algorithm from [10] in this work. As will be shown in the numerical evaluation, the performance of these two approaches is nearly identical, and although we do not show it in this paper, we note that the other schemes differ only minimally in their performance.

The qualitative behavior of the uncooperative scheme is easily analyzed: depending on the power of the noise plus interference, transmitter i will serve between one and $\min(N_i, M_i)$ of its associated receivers. When the noise power σ^2 dominates the interference level, it is intuitively reasonable to consider the two cells in separate, and the uncooperative outcome should be close to the optimum, both in terms of individual cell rate and total system rate.

If the noise power is low compared to the interference power, however, the achievable throughput saturates, even for completely vanishing σ^2. This is a consequence of the fact that the interference *caused* to the other cell is not considered in the optimization (2); therefore, every receiver experiences a non-zero interference level from the neighboring cell and, no matter how low the noise power is, a certain SINR (and thus rate) per user cannot be surpassed.

Consequently, even though the system allows for the transmission of at least η_{AdjBC} parallel, interference-free data streams (as shown in the previous section), with the non-cooperative scheme all streams are in general interference limited and the multiplexing gain is zero.

5. APPROACHES WITH LIMITED COOPERATION

For the approaches with limited cooperation, we need to take into account the interference caused to the unintended receivers as well. To this end, it is necessary to assume that transmitter i has full knowledge of the channels to the receivers of cell j, which incurs a certain information exchange overhead between the cells. We do not, however, assume that the cells exchange payload data, i. e., joint precoding of the data symbols is not possible.

5.1. Zero-Forcing with Optimal Power Allocation

Depending on the antenna configuration, it is not possible to serve all $M_1 + M_2$ receivers with zero-forcing precoding (cf. Section 3). We assume that $0 < n_1 \le \min(N_1, M_1)$ receivers from the first cell and $0 < n_2 \le \min(N_2, M_2)$ receivers from the second cell have been selected for service, where $n_1 + n_2 \le \eta_{\mathrm{AdjBC}}$. This way, we have sufficient degrees of freedom to completely cancel out all interference on the transmitter side. Note that the method of user selection is left open at this point; current channel conditions, fairness, or higher layer demands could be taken into consideration.

For notational convenience, we relabel the receivers so that the indices of the served receivers in cell i run from 1 to n_i. The precoding vector $\boldsymbol{t}_i^m$ for receiver m of cell i is designed to maximize $|\boldsymbol{h}_{ii}^{m,\mathrm{T}}\boldsymbol{t}_i^m|$, subject to the zero-forcing constraints

$$\begin{aligned} &\boldsymbol{h}_{ii}^{k,\mathrm{T}}\boldsymbol{t}_i^m = 0 \quad \forall k \in \{1,\ldots,n_i\} \setminus \{m\} \qquad \text{and} \\ &\boldsymbol{h}_{ji}^{k,\mathrm{T}}\boldsymbol{t}_i^m = 0 \quad \forall k \in \{1,\ldots,n_j\} \end{aligned} \qquad (3)$$

thereby also maximizing SINR_i^m and R_i^m.

The direction of $\boldsymbol{t}_i^m$ can be found by projecting the matched filter vector $\boldsymbol{h}_{ii}^{m,*}$ into the subspace orthogonal to all channel vectors in (3). As this choice of precoding vectors eliminates interference completely and decouples the data streams, the available power P_i can be optimally distributed among the data streams in a second step. For our goal of maximizing the sum of rates over decoupled channels, the optimal power allocation is the waterfilling solution.

5.2. Interference Pricing with Fixed Power Allocation

Clearly, for a given selection of receivers and for vanishing σ^2, the zero-forcing scheme is optimal. For a technique that performs well regardless of the noise power, we employ the concept of *interference pricing*, as presented in [2, 3, 4, 5]. Again, we assume that an appropriate number n_i of receivers of cell i has been selected for service, where the method of user selection is left open.

see the program

The pricing technique is based on iteratively updating the individual precoding vectors $\boldsymbol{t}_i^m$ of both users, one after another, with the goal of maximizing the sum rate of the system

$$R_{\text{sum}} = \sum_{m=1}^{n_1} R_1^m + \sum_{m=1}^{n_2} R_2^m.$$

For the update of one precoding vector $\boldsymbol{t}_i^m$ all other precoding vectors are assumed to be fixed.

Maximizing R_{sum} over $\boldsymbol{t}_i^m$ is not an easy task; taking the derivative and writing out the optimality conditions reveals that there is no solution in closed form. Furthermore, each precoder update requires knowledge of *all* channel coefficients of the system, requiring the transmitters to exchange this information.

Therefore, instead of maximizing R_{sum} directly, we maximize a linear approximation of R_{sum}, which will turn out to have more desirable properties. When updating the precoder $\boldsymbol{t}_i^m$, we linearize the own rate contribution R_i^m in the received signal power $|\boldsymbol{h}_{ii}^{m,\mathrm{T}}\boldsymbol{t}_i^m|^2$, and all other rates R_j^k for $k \in \{1,\ldots,n_j\}$ and R_i^k for $k \in \{1,\ldots,n_i\} \setminus \{m\}$ in the received interference power $|\boldsymbol{h}_{ji}^{k,\mathrm{T}}\boldsymbol{t}_i^m|^2$ and $|\boldsymbol{h}_{ii}^{k,\mathrm{T}}\boldsymbol{t}_i^m|^2$, respectively, yielding the approximation

$$\begin{aligned} R_{\text{sum}} \approx{}& \rho_i^m|\boldsymbol{h}_{ii}^{m,\mathrm{T}}\boldsymbol{t}_i^m|^2 \\ & - \sum_{k\neq m}\pi_i^k|\boldsymbol{h}_{ii}^{k,\mathrm{T}}\boldsymbol{t}_i^m|^2 - \sum_k \pi_j^k|\boldsymbol{h}_{ji}^{k,\mathrm{T}}\boldsymbol{t}_i^m|^2 + c_i^m \end{aligned}$$

with the positive scalars

$$\rho_i^m = \frac{\partial R_i^m}{\partial(|\boldsymbol{h}_{ii}^{m,\mathrm{T}}\boldsymbol{t}_i^m|^2)} = \frac{\partial R_i^m}{\partial S_i^m} = \frac{1}{S_i^m + I_i^m}$$

and

$$\begin{aligned} \pi_\ell^k &= -\frac{\partial R_\ell^k}{\partial(|\boldsymbol{h}_{\ell i}^{k,\mathrm{T}}\boldsymbol{t}_i^q|^2)} \qquad (i,q) \neq (\ell,k) \\ &= -\frac{\partial R_\ell^k}{\partial I_\ell^k} = \frac{1}{I_\ell^k} - \frac{1}{S_\ell^k + I_\ell^k}, \end{aligned}$$

both evaluated at the current operating point. The constant c_i^m does not depend upon $\boldsymbol{t}_i^m$ and is therefore not relevant when maximizing the expression over $\boldsymbol{t}_i^m$. The *price* π_ℓ^k signifies how much a marginal increase in interference experienced by receiver k of cell ℓ will lead to a decrease in its rate R_ℓ^k. Note that the price does not depend on who is causing this interference, i. e., it is the same regardless of which of the precoding vectors we are updating.

The so approximated sum rate can be easily maximized; the necessary optimality condition

$$\boldsymbol{A}_i^m\boldsymbol{t}_i^m = \lambda\boldsymbol{t}_i^m \qquad \text{with} \qquad \lambda > 0$$

and

$$\boldsymbol{A}_i^m = \rho_i^m\boldsymbol{h}_{ii}^{m,*}\boldsymbol{h}_{ii}^{m,\mathrm{T}} - \sum_{k\neq m}\pi_i^k\boldsymbol{h}_{ii}^{k,*}\boldsymbol{h}_{ii}^{k,\mathrm{T}} - \sum_k \pi_j^k\boldsymbol{h}_{ji}^{k,*}\boldsymbol{h}_{ji}^{k,\mathrm{T}}$$

is fulfilled if $\boldsymbol{t}_i^m$ is an eigenvector of $\boldsymbol{A}_i^m$ associated with a positive eigenvalue. Assuming linearly independent channel vectors and $n_1 + n_2 \leq N_i$, the matrix $\boldsymbol{A}_i^m$ has rank $n_1 + n_2$ and exactly one positive eigenvalue. Therefore our precoder update is unique. Note that for calculating $\boldsymbol{A}_i^m$, cell i must have knowledge of the channel vectors $\boldsymbol{h}_{ji}^k$, but not of $\boldsymbol{h}_{ij}^k$ or $\boldsymbol{h}_{jj}^k$.

The question remains how to allocate the power among the precoding vectors. As the pricing algorithm does not decouple the data streams, waterfilling does not yield the optimal power allocation. In this paper, we use the simplest possible scheme and divide the power equally among the beams of one transmitter, so that $\|\boldsymbol{t}_i^m\|_2^2 = P_i/n_i$ for all $m \in \{1,\ldots,n_i\}$.

We sum up our iterative procedure as follows:

1. Select users to be served in each cell, so that multiplexing gain η_{AdjBC} can be achieved. Initialize precoders with zero-forcing solution, distribute the available power uniformly over precoders of each cell.
2. Calculate all π_i^m and ρ_i^m. Communicate prices π_i^m to cell j and vice versa.
3. Compute all matrices $\boldsymbol{A}_i^m$.
4. Update the precoders $\boldsymbol{t}_i^m$ by determining the eigenvector belonging to the only positive eigenvalue of each $\boldsymbol{A}_i^m$. Scale the updated precoders $\boldsymbol{t}_i^m$ to have the same Euclidean norm as before the update.
5. Check for convergence. If convergence is not achieved, repeat from step 2.

Alternatively, the precoders could be updated sequentially with the prices being recomputed after each precoder update. This leads to slightly faster convergence at the cost of more pricing information being exchanged between the cells.

It is straightforward to show that when this algorithm reaches a stationary point, the conditions necessary for optimality of the problem

$$\max_{\boldsymbol{t}_1^1,\ldots,\boldsymbol{t}_2^{n_2}} R_{\text{sum}} \quad \text{s. t.:} \quad \|\boldsymbol{t}_i^m\|_2^2 = P_i/n_i$$
$$\forall i \in \{1,2\} \quad \text{and} \quad \forall m \in \{1,\ldots,n_i\}$$

are fulfilled. While numerical experiments suggest that convergence is always achieved within a small number of iterations, an analytical convergence result is available for the special case of $n_1 = n_2 = 1$ for certain initializations [4].

6. NUMERICAL RESULTS

For the simulation results in this section, we express the sum rate in *bits per channel use*, i. e., we apply the logarithm

see the program

with base 2 in (1). We simulated two scenarios: in the first case, no user selection is necessary, i. e., all $M_1 + M_2$ receivers can be served at the same time with zero-forcing, while in the second a subset of receivers must be selected in order to achieve the full multiplexing gain.

For the results shown in Fig. 2, we averaged over 1000 realizations of a Gaussian i. i. d. channel model with $N_1 = N_2 = 4$ transmit antennas at each base station and $M_1 = M_2 = 2$ users per base station. The interference channels were assumed to be as strong as the desired channels. We define the *signal to noise ratio* (SNR) as the transmit power per base station divided by the noise power at each receiver.

The uncooperative scheme clearly is interference limited; the cooperative schemes, on the other hand, achieve full multiplexing gain, i. e. all four receivers can be served in parallel at high SNR, as can be seen from the slope of the curves at high SNR. The pricing algorithm converges towards the ZF solution for high SNR, for medium to low SNR it performs significantly better.

The scenario for Fig. 3 has $N_1 = N_2 = 4$ transmit antennas per base station and $M_1 = M_2 = 4$ receivers per base station. Thus, only four of the eight receivers can be served with zero-forcing. We employed the simplest scheme possible and randomly selected $n_1 = n_2 = 2$ receivers per base station for service. The non-cooperative schemes gain slightly from being able to serve more users; at low SNR, they now visibly outperform our simple cooperative schemes. Note that more advanced user selection techniques are expected further improve the performance of the zero-forcing and pricing schemes.

As can be seen in Fig. 4 where the interference channels are attenuated by 10 dB, reducing the magnitude of the cross channel coefficients causes the saturation level of the uncooperative schemes to rise; in the moderate SNR regime in particular, it becomes clear that the proposed simple cooperative techniques are not sophisticated enough to also perform close to optimally in an environment with weak interference.

7. CONCLUSION AND OUTLOOK

In a cellular system in which inter-cell interference is the limiting factor and the background noise level is relatively low, cooperation between the interfering base stations can yield great performance gains. As shown in this paper, even when full cooperation between the base stations is not possible, high gains can be achieved by exchanging some channel state information between the base stations and allowing them to cooperatively design their precoders. Depending on the antenna configuration of the system, the optimal strategy may involve serving fewer receivers and spending some degrees of freedom on avoiding interference for the users in the other cell. While for high SNR it is best to completely

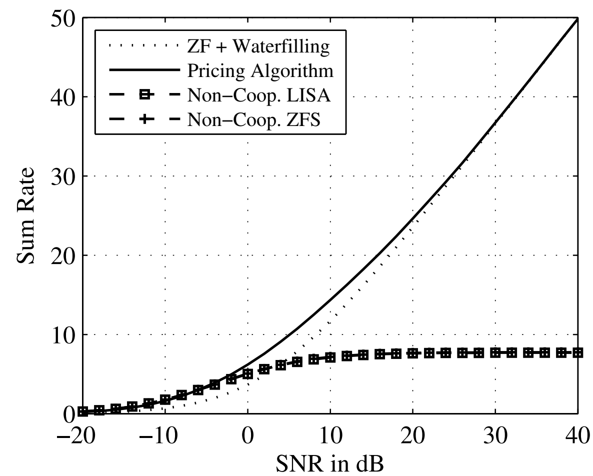


Fig. 2. $N_1 = N_2 = 4$ Transmit Antennas and $M_1 = M_2 = 2$ Receivers per Base Station.

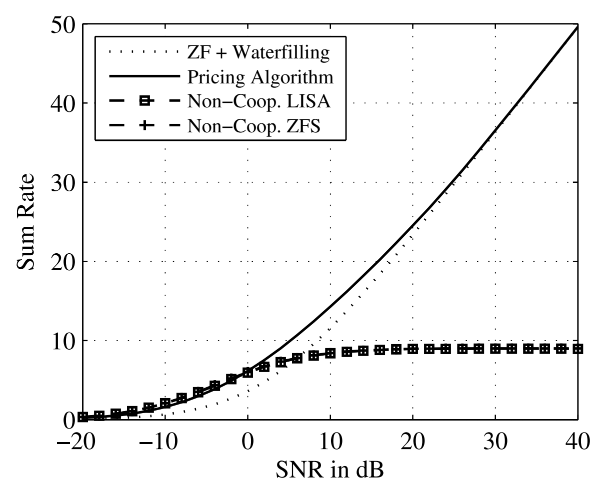


Fig. 3. $N_1 = N_2 = 4$ Transmit Antennas and $M_1 = M_2 = 4$ Receivers per Base Station.

see the program

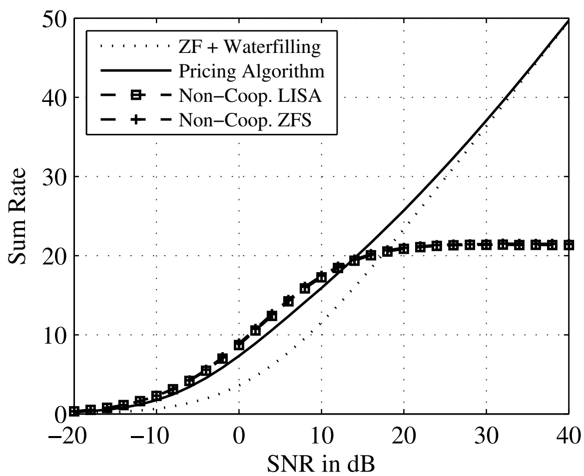


Fig. 4. $N_1 = N_2 = 4$ Transmit Antennas and $M_1 = M_2 = 4$ Receivers per Base Station, Desired Channels 10 dB Stronger than Interference Channel.

decouple the served users with zero-forcing, at moderate to low SNR, where allowing some interference is beneficial, the proposed pricing algorithm achieves better performance.

The pricing algorithm, as presented in this work, has potential for further improvement: adaptive beam powers and a variable number of served users [5] per base station are the subject of future work, as well as a user selection technique aimed at maximizing the sum rate. Hopefully, the performance of the non-cooperative approach can be reached even in weak-interference scenarios.

8. REFERENCES

[1] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming Intercell Interference in Indoor Wireless Systems," in *Proc. 41st Asilomar Conference on Signals, Systems and Computers*, Nov. 2007, pp. 83–87.

[2] J. Huang, R. A. Berry, and M. L. Honig, "Distributed Interference Compensation for Wireless Networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[3] D. A. Schmidt, A. Gründinger, W. Utschick, and M. L. Honig, "Distributed Precoder Optimization for Interfering MISO Channels," in *Proc. WSA 2008*, Feb. 2008, pp. 285–290.

[4] C. Shi, R. A. Berry, and M. L. Honig, "Distributed Interference Pricing with MISO Channels," in *Proc. 46th Annual Allerton Conference 2008*, Sep. 2008.

[5] C. Shi, D. A. Schmidt, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed Interference Pricing for the MIMO Interference Channel," *to appear in Proc. IEEE International Conference on Communications, 2009*.

[6] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the k-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[7] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR Power Offset in Multiantenna Communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

[8] S. A. Jafar and M. J. Fakhereddin, "Degrees of Freedom for the MIMO Interference Channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, Jul. 2007.

[9] M. Stojnic, H. Vikalo, and B. Hassibi, "Rate Maximization in Multi-Antenna Broadcast Channels with Linear Preprocessing," in *Proc. GLOBECOM 2004*, vol. 6, Nov. 2004, pp. 3957–3961.

[10] G. Dimić and N. D. Sidiropoulos, "On Downlink Beamforming With Greedy User Selection: Performance Analysis and a Simple New Algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.

[11] D. A. Schmidt, M. Joham, R. Hunger, and W. Utschick, "Near Maximum Sum-Rate Non-Zero-Forcing Linear Precoding with Successive User Selection," in *Proc. 40th Asilomar Conference on Signals, Systems and Computers*, Nov. 2006.

[12] C. Guthy, W. Utschick, G. Dietl, and P. Tejera, "Efficient Linear Successive Allocation for the MIMO Broadcast Channel," in *Proc. of 42nd Asilomar Conference on Signals, Systems and Computers*, 2008.

see the program

OPTIMIZED BEAMFORMING FOR THE TWO STREAM MIMO INTERFERENCE CHANNEL AT HIGH SNR

Christoph Hellings, David A. Schmidt, Wolfgang Utschick

Associate Institute for Signal Processing, Technische Universität München
80290 Munich, Germany, Telephone: +49 89 289-28508
christoph.hellings@mytum.de, {dschmidt|utschick}@tum.de

ABSTRACT

We investigate the high SNR behavior of the two-user multiple-input/multiple-output (MIMO) interference channel with two antennas at each terminal, where the receivers treat the interference as noise. It is known that the maximum multiplexing gain of such a system is two, and that it can be achieved with zero-forcing. We introduce the *high SNR rate offset* as a performance measure to further differentiate between solutions achieving the full multiplexing gain and, by maximizing the high SNR rate offset, formulate a system of multivariate polynomial equations necessary for the globally optimal high SNR transmit strategy in terms of sum rate. We furthermore define a two-player game in which the two users are constrained to transmit exactly one data stream, but compete for the beamformer design; the *Nash equilibria* of this game can be explicitly calculated. The so-found sum rate optimal and semi-competitive solutions are also shown in simulations to outperform the conventional competitive approach to the MIMO interference channel even at finite SNR.

1. INTRODUCTION

The information theoretic capacity region of the multiple-input/multiple-output interference channel (MIMO IFC) is a largely unsolved problem (e. g. [1]). A common suboptimal approach to finding a region of achievable rates is to assume Gaussian codebooks and no interference cancellation at the receivers, i. e. to restrict the receivers to only be able to treat the interference as noise. In this scenario, the optimization variables are the covariance matrices of the transmit signals, constrained to not exceed a given maximum power per user.

In [2], a gradient-based numerical approach to designing the covariance matrices of the transmitted signals was proposed, with the aim of finding operating points that are locally optimal in terms of sum rate or weighted sum rate. As the sum rate is in general non-concave in the transmit covariances, it is possible (and in fact very common) that many local optima exist, and therefore any gradient-based algorithm might converge to a solution that is far away from the global optimum. While in the case of only a single receive antenna the asymptotically optimal strategy for high signal to noise ratio (SNR) is known [3], the result cannot be easily generalized to the MIMO IFC. In Section 4 we will show how the global optimum can be found for high SNR and two antennas at each transmitter and receiver (and some other antenna configurations that allow for a multiplexing gain of two) in a two user MIMO IFC.

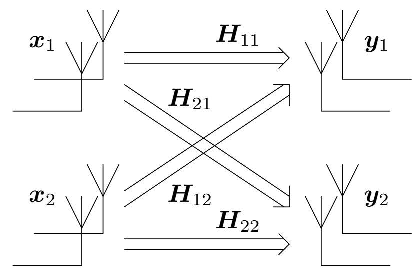


Fig. 1. A MIMO IFC with two Antennas at Each Terminal

The MIMO IFC has also been discussed from a game theoretic point of view (e. g. [4] and the references therein). In [2], the authors discuss Nash equillibria by modelling the choice of transmit covariance matrices as a non-cooperative game. In their conclusion they state that in order to achieve high data rates, cooperation between users is highly desirable when interference is the limiting factor. In Section 4.4 of this paper, we will discuss a middle course between competition and cooperation: while the users agree to only use one beam each, they compete for the beamformer design.

In this work vectors are typeset in boldface lowercase letters and matrices in boldface uppercase letters. We write $\mathbf{0}$ for the zero matrix or vector and $\mathbf{1}$ for the identity matrix. $[\boldsymbol{A}]_{k\ell}$ is used to denote the element in the k-th row and l-th column of the matrix $\boldsymbol{A}$. We use $\bullet^{\mathrm{T}}$ to denote the transpose of a vector or matrix, $\bullet^{*}$ for the conjugate complex and $\bullet^{\mathrm{H}}$ for the conjugate transpose.

For notational compactness, throughout the paper we will use j to denote the user which is not user i. As we consider a system with only two users, j is unambiguous, i. e. $j = 2$ whenever $i = 1$ and vice versa.

see the program

2. SYSTEM MODEL AND PROBLEM FORMULATION

The two user MIMO IFC (Fig. 1) consists of two transmitters and two receivers. Each receiver is connected to both transmitters by a MIMO channel. Assuming frequency flat channels, the received signals of both users can be written as

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{bmatrix}$$

where $\boldsymbol{x}_i$ is the signal transmitted by user i, $\boldsymbol{y}_i$ the signal received by user i, $\boldsymbol{\eta}_i$ the noise at the i-th receiver, and $\boldsymbol{H}_{i\ell}, \ell \in \{i, j\}$ denotes the channel between transmitter ℓ and receiver i. The channel matrices $\boldsymbol{H}_{ii}$ and $\boldsymbol{H}_{ij}$ are assumed to have full rank and to be perfectly known. Note that $\boldsymbol{H}_{i\ell} \in \mathbb{C}^{M_i \times N_\ell}$, $\boldsymbol{x}_i \in \mathbb{C}^{N_i}$, and $\boldsymbol{y}_i, \boldsymbol{\eta}_i \in \mathbb{C}^{M_i}$ where N_i denotes the number of antennas at the i-th transmitter while M_i denotes the number of antennas at the i-th receiver. We consider only antenna configurations that allow a maximum multiplexing gain [5] of two, i. e. systems with $N_i = 2$, $N_j \geq 2$, $M_i \geq 2$, and $M_j = 2$ (cf. Section 3.1).

The additive noise $\boldsymbol{\eta}_i$ is assumed to be circularly symmetric complex Gaussian with $\mathrm{E}[\boldsymbol{\eta}_i] = \boldsymbol{0}$ and $\mathrm{E}[\boldsymbol{\eta}_i \boldsymbol{\eta}_i^{\mathrm{H}}] = \sigma^2 \mathbf{1}$, where the noise power σ^2 is assumed to be equal at each receive antenna. Furthermore, $\boldsymbol{\eta}_1$, $\boldsymbol{\eta}_2$, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are assumed to be pairwise statistically independent.

We assume that receiver i is not able to decode and subtract the signal from transmitter j, so that the interference acts as additional noise. We decide that the transmitters are to use Gaussian Codebooks, so that their transmit signals are zero-mean and circularly symmetric complex Gaussian distributed, i. e. $\boldsymbol{x}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{Q}_i)$. If $\mathrm{Rank}[\boldsymbol{Q}_i] = r_i$, such a Gaussian vector $\boldsymbol{x}_i$ can be created by multiplying a beamforming matrix $\boldsymbol{T}_i \in \mathbb{C}^{M_i \times r_i}$ that fulfills $\boldsymbol{T}_i \boldsymbol{T}_i^{\mathrm{H}} = \boldsymbol{Q}_i$ with an r_i-dimensional vector $\boldsymbol{s}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{1})$.

We assume identical constraints on the average transmit power for both transmitters, which can be written as $\mathrm{tr}[\boldsymbol{Q}_i] \leq P_{\max}$. For simplicity we choose $P_{\max} = 1$ without loss of generality.

We will focus on the question what data rates can be achieved in the presence of interference. Given the transmitter covariance matrix $\boldsymbol{Q}_i$ and the interference plus noise covariance matrix $\boldsymbol{R}_i = \boldsymbol{H}_{ij} \boldsymbol{Q}_j \boldsymbol{H}_{ij}^{\mathrm{H}} + \sigma^2 \mathbf{1}$, the mutual information of the i-th transmitter receiver pair of a MIMO IFC (and hence the highest rate at which information can reliably be transmitted) is[1] [2]

$$R_i = \log\det\left(\mathbf{1} + \boldsymbol{H}_{ii} \boldsymbol{Q}_i \boldsymbol{H}_{ii}^{\mathrm{H}} \boldsymbol{R}_i^{-1}\right) \qquad (1)$$

The task is now to find pairs of $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ such that the sum rate $R_1 + R_2$ is maximized while $\mathrm{tr}[\boldsymbol{Q}_i] \leq P_{\max}$ is fulfilled for $i \in \{1, 2\}$. However, there seems to be no closed form solution and numerical methods might converge to suboptimal local maxima because the optimization problem is non-convex.

[1] As we will later take derivatives of expressions involving rates, we use the natural logarithm for the sake of notational brevity. Thus, rates are expressed in natural units per channel use. For numerical results this can be converted to bits per channel use by dividing by a factor of log 2.

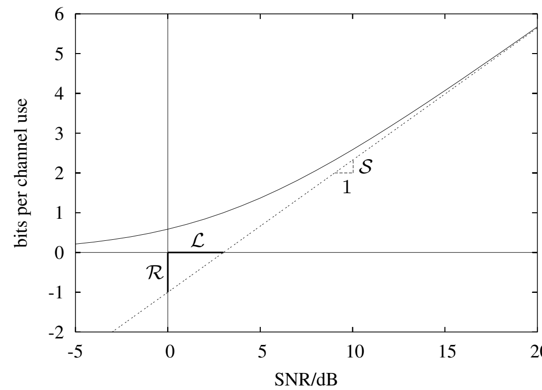


Fig. 2. Graphical Interpretation of the High SNR Slope $\mathcal{S}$, the High SNR Power Offset $\mathcal{L}$ and the High SNR Rate Offset $\mathcal{R}$

In order to develop alternative methods of finding solutions close to the global optimum, we introduce an additional assumption, namely that the signal to noise ratio (SNR) $\rho = P_{\max}/\sigma^2 = \sigma^{-2}$ is high. To be more precise, we will attempt to derive transmit strategies that are globally optimal if $\rho \to \infty$, as we assume that these strategies will also perform well for high, but finite values of the SNR.

3. HIGH SNR PERFORMANCE MEASURES FOR THE MIMO INTERFERENCE CHANNEL

It is well known that for asymptotically high SNR the achievable sum rate of a MIMO IFC grows without bound. Therefore, in the case of high SNR, performance measures other than the sum rate are needed. The authors of [6] propose describing the behavior of a single user MIMO channel in the high SNR regime by the high SNR slope $\mathcal{S}$ and the high SNR power offset $\mathcal{L}$. These two measures explained below can be easily extended to the MIMO IFC by defining a high SNR sum slope $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ and a high SNR sum power offset $\mathcal{L}$ with $\mathcal{S}\mathcal{L} = \mathcal{S}_1\mathcal{L}_1 + \mathcal{S}_2\mathcal{L}_2$. Additionally, we will present a related measure, the high SNR rate offset $\mathcal{R}$.

Fig. 2 shows the graphical interpretation of $\mathcal{S}$, $\mathcal{L}$ and $\mathcal{R}$. Any combination of two of them is sufficient to describe a unique line, which is the asymptote to the rate curve at high SNR. A detailed discussion of each of the performance measures will be presented in the following sections.

3.1. The High SNR Slope

The high SNR slope is the slope of the asymptote of the log(SNR)-versus-rate curve at high SNR and is also re-

see the program

ferred to as *degrees of freedom* (DOF) or *multiplexing gain*:

$$\mathcal{S} = \lim_{\rho\to\infty} \frac{R(\rho)}{\log \rho}, \tag{2}$$

where $R(\rho)$ is the rate which can be achieved at the SNR ρ [6]. The high SNR slope is often given in the pseudo unit $\mathrm{bits} / \mathrm{dB}$. Alternatively, it can be expressed without a unit by using the same logarithm for the rate and for the SNR. In the latter case, the obtained value indicates the ratio to the high SNR slope of a single user single-input/single-output link justifying the name multiplexing gain. Thus, the high SNR slope can also be interpreted as the number of independent data streams that can be transmitted.

In [5] it was shown that the maximum possible high SNR slope for the MIMO IFC with N_i antennas at the i-th transmitter and M_i antennas at the i-th receiver is $\mathcal{S}_{\max} =$

$$\min\{N_1{+}N_2, M_1{+}M_2, \max\{N_1,M_2\}, \max\{N_2,M_1\}\}.$$

As was also claimed in [5], this high SNR slope $\mathcal{S}_{\max}$ can be achieved by *zero-forcing* (ZF). Depending on the dimensions of the channel matrices, ZF can be performed by the transmitter, by the receiver, or by a combination of both.

3.2. The High SNR Power Offset

Altough the maximum possible high SNR slope can be easily achieved by ZF, we note that this leads to a whole class of solutions. As we expect the various possible solutions to differ in terms of sum rate, we will now introduce the high SNR power offset that is able to reflect such differences in terms of a high SNR performance measure.

The high SNR power offset is the distance between the origin of the $\log(\mathrm{SNR})$-versus-rate coordinate system and the point where the high SNR asymptote of the rate curve intersects the $\log(\mathrm{SNR})$-axis. We can see from Fig. 2 that [6]

$$\mathcal{L} = \lim_{\rho\to\infty}\left(\log\rho - \frac{R(\rho)}{\mathcal{S}}\right). \tag{3}$$

To calculate the high SNR power offset of the i-th link in a MIMO IFC, we rewrite (1) as

$$R_i = \log D_i \quad \text{with} \quad D_i = \det\left(\mathbf{1} + \boldsymbol{S}_i\left(\sigma^2\mathbf{1} + \boldsymbol{I}_i\right)^{-1}\right)$$

where $\boldsymbol{I}_i = \boldsymbol{H}_{ij}\boldsymbol{Q}_j\boldsymbol{H}_{ij}^{\mathrm{H}}$ is the covariance matrix of the received interference and $\boldsymbol{S}_i = \boldsymbol{H}_{ii}\boldsymbol{Q}_i\boldsymbol{H}_{ii}^{\mathrm{H}}$ is the covariance matrix of the received intended signal. From (3) we get

$$\mathcal{L}_i = -\mathcal{S}_i^{-1}\log\lim_{\sigma^2\to 0}\left(\left(\sigma^2\right)^{\mathcal{S}_i} D_i\right) \tag{4}$$

To see how D_i behaves in the limit of $\sigma^2 \to 0$, we apply the matrix inversion lemma and make use of the eigenvalue decomposition (EVD) $\boldsymbol{I}_i = \boldsymbol{U}_i\boldsymbol{\Lambda}_i\boldsymbol{U}_i^{\mathrm{H}}$:

$$\begin{aligned}\left(\sigma^2\mathbf{1} + \boldsymbol{I}_i\right)^{-1} &= \sigma^{-2}\left(\mathbf{1} - \boldsymbol{I}_i\left(\sigma^2\mathbf{1}{+}\boldsymbol{I}_i\right)^{-1}\right) &(5)\\ &= \sigma^{-2}\left(\mathbf{1} - \boldsymbol{U}_i\boldsymbol{\Lambda}_i\left(\sigma^2\mathbf{1}{+}\boldsymbol{\Lambda}_i\right)^{-1}\boldsymbol{U}_i^{\mathrm{H}}\right) &(6)\end{aligned}$$

Now we can calculate the behaviour for vanishing noise:

$$\lim_{\sigma^2\to 0}\boldsymbol{U}_i\boldsymbol{\Lambda}_i\left(\sigma^2\mathbf{1} + \boldsymbol{\Lambda}_i\right)^{-1}\boldsymbol{U}_i^{\mathrm{H}} = \boldsymbol{U}_i\boldsymbol{J}_i\boldsymbol{U}_i^{\mathrm{H}} \tag{7}$$

where $\boldsymbol{J}_i = \operatorname{diag}\{\mathrm{j}_{i,k}\}$ with $\mathrm{j}_{i,k} = 0$ if the k-th eigenvalue of $\boldsymbol{I}_i$ is zero, i. e. $\lambda_{i,k} = 0$, and $\mathrm{j}_{i,k} = 1$ otherwise.

By making use of (4) through (7) and inserting the definition of $\boldsymbol{S}_i$ we obtain

$$\mathcal{L}_i \;=\; -\mathcal{S}_i{}^{-1}\log\det\left(\boldsymbol{T}_i^{\mathrm{H}}\boldsymbol{H}_{ii}^{\mathrm{H}}(\mathbf{1} - \boldsymbol{U}_i\boldsymbol{J}_i\boldsymbol{U}_i^{\mathrm{H}})\boldsymbol{H}_{ii}\boldsymbol{T}_i\right)$$

with $\boldsymbol{J}_i$ defined as in (7). Note that we applied Sylvester's determinant theorem $\det(\mathbf{1} + \boldsymbol{AB}) = \det(\mathbf{1} + \boldsymbol{BA})$ in order to change the dimension of the matrix inside the det-operator to $\mathcal{S}_i \times \mathcal{S}_i$. Thus, when excluding a factor of σ^{-2} from the determinant, it gets an exponent $\mathcal{S}_i$ such that $(\sigma^2)^{\mathcal{S}_i}$ from (4) cancels out.

Finally, the high SNR sum power offset of the MIMO IFC can be obtained using $\mathcal{SL} = \mathcal{S}_1\mathcal{L}_1 + \mathcal{S}_2\mathcal{L}_2$ and the sum rate curve at high SNR can be approximated by the asymptote $R \approx \mathcal{S}(\log\rho - \mathcal{L})$.

3.3. The High SNR Rate Offset

We also define the more convenient high SNR rate offset $\mathcal{R}_i = -\mathcal{S}_i\mathcal{L}_i$ which describes the axis intercept of the high SNR asymptote on the rate axis:

$$\mathcal{R}_i = \log\det\left(\boldsymbol{T}_i^{\mathrm{H}}\boldsymbol{H}_{ii}^{\mathrm{H}}(\mathbf{1} - \boldsymbol{U}_i\boldsymbol{J}_i\boldsymbol{U}_i^{\mathrm{H}})\boldsymbol{H}_{ii}\boldsymbol{T}_i\right) \tag{8}$$

For the high SNR asymptote we now get $R \approx \mathcal{R} + \mathcal{S}\log\rho$ where $\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2$ is the high SNR sum rate offset. The high SNR slope $\mathcal{S}$ and the high SNR rate offset $\mathcal{R}$ are the performance measures used in the following sections.

3.4. Summary and Application of the High SNR Performance Measures

The high SNR slope is obviously the dominant high SNR performance measure. Only for systems with the same high SNR slope, the high SNR rate offset is a criterion worth considering because the former gives information about the number of independent data streams that can be transmitted, while the latter indicates the quality of the streams. For infinite SNR, even a very high offset can not compensate the difference in rate caused by a stream more or less.

Hence, in order to maximize the sum rate in a MIMO IFC at high SNR, we can instead maximize the high SNR sum power offset while ensuring that the maximum high SNR slope is achieved. This maximization is difficult to solve for several reasons: Firstly we impose a constraint on the sum slope which does not state how the slope should be divided between users. It seems that the problem of finding the globally optimal distribution of slope is NP-hard. Secondly the slope constraint makes the constraint set non-convex. Finally the fact that $\boldsymbol{U}_i\boldsymbol{J}_i\boldsymbol{U}_i^{\mathrm{H}}$ is a function of the

see the program

eigenvalue decomposition of the matrix $\boldsymbol{I}_i = \boldsymbol{H}_{ij}\boldsymbol{Q}_j\boldsymbol{H}_{ij}^{\mathrm{H}}$ complicates the derivative of $\mathcal{R}_i$ with respect to $\boldsymbol{Q}_j$.

However, (8) gets quite compact whenever the interference vanishes. As will be seen in the following section, we can make use of this property by introducing interference free effective channels after ZF has been applied.

4. THE TWO STREAM MIMO IFC AT HIGH SNR

As stated previously, we consider only antenna configurations where $\mathcal{S}_{\max} = 2$, as is the case if one user (assume that it is user i) has two transmit antennas while the other user has two receive antennas. Since it only makes sense to receive inside the span of $\boldsymbol{H}_{ii}$, the i-th receiver can be written as $\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{B}_i^{\mathrm{H}}$ where $\boldsymbol{g}_i \in \mathbb{C}^2$ and the columns of $\boldsymbol{B}_i$ are an orthonormal basis of the span of $\boldsymbol{H}_{ii}$. Accordingly, transmitter j has to send outside of the nullspace of $\boldsymbol{H}_{jj}$, so that its beamformer can be written as $\boldsymbol{B}_j\boldsymbol{t}_j$ where $\boldsymbol{t}_j \in \mathbb{C}^2$ and the columns of $\boldsymbol{B}_j$ are an orthonormal basis of the orthogonal complement of the nullspace of $\boldsymbol{H}_{jj}$. Hence, every system covered by our assumption can be transformed into a 2×2 MIMO IFC, i. e. a MIMO IFC with two antennas at each transmitter and receiver, by setting

$$\begin{aligned} \boldsymbol{H}'_{ii} &= \boldsymbol{B}_i^{\mathrm{H}}\boldsymbol{H}_{ii} & \qquad \boldsymbol{H}'_{ij} &= \boldsymbol{B}_i^{\mathrm{H}}\boldsymbol{H}_{ij}\boldsymbol{B}_j \\ \boldsymbol{H}'_{ji} &= \boldsymbol{H}_{ji} & \qquad \boldsymbol{H}'_{jj} &= \boldsymbol{H}_{jj}\boldsymbol{B}_j. \end{aligned}$$

Note that this transformation changes neither the noise power nor the power of the transmitted signals because the columns of $\boldsymbol{B}_i$ and $\boldsymbol{B}_j$ have unit norm. In the following we will only discuss the 2×2 MIMO IFC, keeping in mind that all results can be extended to the more general two stream MIMO IFC.

To achieve $\mathcal{S} = 2$, we either can shut one user off and treat the other link as a single user MIMO link (i. e. $\mathcal{S}_i = 0$ and $\mathcal{S}_j = 2$), or share the degrees of freedom (i. e. $\mathcal{S}_1 = \mathcal{S}_2 = 1$). The optimal strategy for the former case can be easily found with waterfilling [7]. The latter case is more challenging and is the subject of the following sections.

4.1. An equivalent 2x1 MISO IFC

The mutual information of a communication system cannot be increased by a receive filter. Consequently, most efforts to reach a good rate concentrate only on the choice of the beamforming vectors and calculate the rate from the mutual information between transmitter and receiver. An example can be found in [2]. Nevertheless, we propose a method of jointly choosing the beamforming vectors $\boldsymbol{t}_1$, $\boldsymbol{t}_2$ and the receive filters $\boldsymbol{g}_1^{\mathrm{H}}$, $\boldsymbol{g}_2^{\mathrm{H}}$. The latter are row vectors, as we assume one stream per user. The reason for this approach is that it enables us to introduce an equivalent system model in which the receive filters are viewed as part of the channel, permitting the use of well known results for the multiple-input/single-output (MISO) IFC.

We assume that receiver i is equipped with a receive filter $\boldsymbol{g}_i^{\mathrm{H}}$ so that $y'_i = \boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{y}_i$. In order to be able to use the results from (8) we have to ensure that the assumption of noise power σ^2 is still fulfilled for the filtered noise $\eta'_i = \boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{\eta}_i$. Therefore $\boldsymbol{g}_i^{\mathrm{H}}$ must have unit norm in order to not change the noise power.

Considering the receive filters as a part of the channel, we obtain an equivalent communication system, which is a MISO IFC with two antennas at each transmitter:

$$\boldsymbol{h}'^{\mathrm{H}}_{ii} = \boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ii} \qquad\qquad \boldsymbol{h}'^{\mathrm{H}}_{ij} = \boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ij},$$

As was shown in [3, 8], the only sensible strategy in a MISO IFC at high SNR is to transmit orthogonally to the interference channel, i. e. to avoid causing interference to the unintended receiver. The ZF conditions are

$$\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{12}\boldsymbol{t}_2 = 0 \qquad\qquad \boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{21}\boldsymbol{t}_1 = 0.$$

This way, we ensure that the maximum high SNR slope $\mathcal{S}_{\max}$ is achieved. After choosing the receive filters of both systems the beamforming vectors are fixed and can be explicitly expressed by means of the channel coefficients and the receive filters using the formula from [3, 8]:

$$\boldsymbol{t}_i = \frac{\boldsymbol{P}_{ji}^{\perp}\boldsymbol{h}'_{ii}}{\sqrt{\boldsymbol{h}'^{\mathrm{H}}_{ii}\boldsymbol{P}_{ji}^{\perp}\boldsymbol{h}'_{ii}}} \quad \text{with} \quad \boldsymbol{P}_{ji}^{\perp} = \mathbf{1} - \frac{\boldsymbol{h}'_{ji}\boldsymbol{h}'^{\mathrm{H}}_{ji}}{\boldsymbol{h}'^{\mathrm{H}}_{ji}\boldsymbol{h}'_{ji}}. \tag{9}$$

The beamforming vector $\boldsymbol{t}_i$ has unit norm to fulfill the power constraint. Note that we do not have to consider transmit powers smaller than $P_{\max} = 1$, as it has been shown in [9] that all Pareto optimal points in a MISO IFC use full power. Since no interference is received in the resulting MISO channel, (8) reduces to

$$\mathcal{R}_i = \log\left(\boldsymbol{t}_i^{\mathrm{H}}\boldsymbol{h}'_{ii}\boldsymbol{h}'^{\mathrm{H}}_{ii}\boldsymbol{t}_i\right) = \log\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ii}\boldsymbol{P}_{ji}^{\perp}\boldsymbol{H}_{ii}^{\mathrm{H}}\boldsymbol{g}_i\right) \tag{10}$$
$$\text{with} \quad \boldsymbol{P}_{ji}^{\perp} = \mathbf{1} - \frac{\boldsymbol{H}_{ji}^{\mathrm{H}}\boldsymbol{g}_j\boldsymbol{g}_j^{\mathrm{H}}\boldsymbol{H}_{ji}}{\boldsymbol{g}_j^{\mathrm{H}}\boldsymbol{H}_{ji}\boldsymbol{H}_{ji}^{\mathrm{H}}\boldsymbol{g}_j}$$

Note that in (10), the projector $\boldsymbol{P}_{ji}^{\perp}$ depends only on $\boldsymbol{H}_{ji}$ and on the j-th user's choice for the receive filter, but not on $\boldsymbol{g}_i$. In the following, we optimize $\mathcal{R}_1$ and $\mathcal{R}_2$ according to different criteria. Having found $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ with any of these methods, the beamforming vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ can be calculated using equation (9).

4.2. Optimal Signaling for Cases with a Prioritized User

Obviously, an upper bound for $\mathcal{R}_i$ is $\log(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ii}\boldsymbol{H}_{ii}^{\mathrm{H}}\boldsymbol{g}_i)$. Equality holds, when the projection does not have any effect, i. e. when $\boldsymbol{g}_j$ is chosen such that it lies in the nullspace of $\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ii}\boldsymbol{H}_{ji}^{\mathrm{H}}$. However, this bound is not constant, but it depends on the i-th user's choice for his receive filter. To maximize the upper bound, we have to choose $\boldsymbol{g}_i$ as the first left singular vector of $\boldsymbol{H}_{ii}$.

see the program

It is not possible to maximize both users' bounds and simultaneously make sure that the bound is achieved for both users because we only have two free variables, namely the directions of the two receive filters. If we assume that we would like to give one user priority in terms of rate offset, under the restriction that the degrees of freedom are shared, i. e. both users transmit one data stream, this leads to an explicit calculation rule for the optimal strategy: the prioritized user maximizes his upper bound while the other user ensures that the bound is reached. This can be interpreted as follows: the prioritized user may transmit over the principal mode of the channel to his intended receiver, while the other user is responsible for fulfilling both ZF conditions, i. e. he has to perform ZF at both transmitter and receiver.

4.3. Optimal High SNR Sum Rate Offset

If both users have equal priority, our optimization criterion is the high SNR sum rate offset $\mathcal{R}$, which complicates the optimization with respect to $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ significantly. Not only is it impossible to find a closed-form solution satisfying the Karush-Kuhn-Tucker (KKT) conditions necessary for local optimality, there also appear to be many such local optima in general. Therefore it is not our goal to develop an algorithm based on gradient projection, which might end up in a local, but not global optimum.

In order to transform the constrained optimization into an unconstrained one, we choose a parametrization for $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ that guarantees that the constraint is met:

$$\boldsymbol{g}_1 = \begin{bmatrix} \cos\alpha_1 \cdot \mathrm{e}^{\mathrm{j}\,\beta_1} \\ \sin\alpha_1 \end{bmatrix} \qquad \boldsymbol{g}_2 = \begin{bmatrix} \cos\alpha_2 \cdot \mathrm{e}^{\mathrm{j}\,\beta_2} \\ \sin\alpha_2 \end{bmatrix}$$

As $\boldsymbol{g}_i$ and $\boldsymbol{g}_j$ can be multiplied by a complex phase without changing the value of the cost function (10), we can assume a non-negative real valued second entry without loss of generality. Furthermore, we only have to consider $\alpha_i \in [0, \pi]$ and $\beta_i \in [0, \pi]$. [2]

As we only consider two dimensional vectors, an alternative formulation of (9) is

$$\boldsymbol{t}_i = \frac{\underline{\boldsymbol{P}}\boldsymbol{h}'^{*}_{ji}}{\sqrt{\boldsymbol{h}'^{\mathrm{H}}_{ji}\boldsymbol{h}'_{ji}}} \qquad \text{with} \qquad \underline{\boldsymbol{P}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{11}$$

This allows us to rewrite the high SNR sum rate offset:

$$\begin{aligned} \mathcal{R} = & \log\left(\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{11}\underline{\boldsymbol{P}}^{\mathrm{H}}\left(\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{g}_2\boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{21}\right)^{*}\underline{\boldsymbol{P}}\boldsymbol{H}_{11}^{\mathrm{H}}\boldsymbol{g}_1\right) \\ & + \log\left(\boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{22}\underline{\boldsymbol{P}}^{\mathrm{H}}\left(\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{12}\right)^{*}\underline{\boldsymbol{P}}\boldsymbol{H}_{22}^{\mathrm{H}}\boldsymbol{g}_2\right) \\ & - \log\left(\boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{21}\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{g}_2\right) - \log\left(\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{12}\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{g}_1\right) \end{aligned} \tag{12}$$

In order to find extremal points we have to take the derivatives with respect to the real valued parameters α_1, β_1, α_2, and β_2, and set them to zero. Applying the chain rule, the conditions for extremal points can be written as

$$0 = \frac{\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ii}\boldsymbol{g}_i\right)_{\alpha_i}}{\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ii}\boldsymbol{g}_i} + \frac{\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ij}\boldsymbol{g}_i\right)_{\alpha_i}}{\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ij}\boldsymbol{g}_i} - \frac{\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{K}_{i}\boldsymbol{g}_i\right)_{\alpha_i}}{\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{K}_{i}\boldsymbol{g}_i} \tag{13}$$

$$0 = \frac{\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ii}\boldsymbol{g}_i\right)_{\beta_i}}{\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ii}\boldsymbol{g}_i} + \frac{\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ij}\boldsymbol{g}_i\right)_{\beta_i}}{\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}_{ij}\boldsymbol{g}_i} - \frac{\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{K}_{i}\boldsymbol{g}_i\right)_{\beta_i}}{\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{K}_{i}\boldsymbol{g}_i} \tag{14}$$

where $(\ldots)_x$ is an abbreviation for $\frac{\partial \ldots}{\partial x}$ and the matrices $\boldsymbol{D}_{ii}$, $\boldsymbol{D}_{ij}$ and $\boldsymbol{K}_i$ are given by

$$\boldsymbol{D}_{ii} = \boldsymbol{M}_{ii}^{\mathrm{H}}\boldsymbol{g}_j^{*}\boldsymbol{g}_j^{\mathrm{T}}\boldsymbol{M}_{ii} \qquad \boldsymbol{M}_{ii} = \boldsymbol{H}_{ji}^{*}\underline{\boldsymbol{P}}\boldsymbol{H}_{ii}^{\mathrm{H}} \tag{15}$$

$$\boldsymbol{D}_{ij} = \boldsymbol{M}_{ij}^{\mathrm{H}}\boldsymbol{g}_j^{*}\boldsymbol{g}_j^{\mathrm{T}}\boldsymbol{M}_{ij} \qquad \boldsymbol{M}_{ij} = \boldsymbol{H}_{jj}^{*}\underline{\boldsymbol{P}}^{\mathrm{H}}\boldsymbol{H}_{ij}^{\mathrm{H}} \tag{16}$$

$$\boldsymbol{K}_i = \boldsymbol{H}_{ij}\boldsymbol{H}_{ij}^{\mathrm{H}}.$$

A Hermitian 2×2 matrix $\boldsymbol{D}$ with $\mathrm{Rank}[\boldsymbol{D}] = 1$ and a Hermitian 2×2 matrix $\boldsymbol{K}$ with $\mathrm{Rank}[\boldsymbol{K}] = 2$ can be parametrized by

$$\boldsymbol{D} = \mu \begin{bmatrix} 1 & d \\ d^{*} & |d|^2 \end{bmatrix} \quad \text{with} \quad \mu \in \mathbb{R},\ d \in \mathbb{C}, \tag{17}$$

$$\boldsymbol{K} = \begin{bmatrix} k_{11} & k_{12} \\ k_{12}^{*} & k_{22} \end{bmatrix} \quad \text{with} \quad k_{11}, k_{22} \in \mathbb{R},\ k_{12} \in \mathbb{C},$$

respectively, so that the following equations hold:

$$\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{D}\boldsymbol{g}_i = \left|\cos\alpha_i + d\sin\alpha_i\,\mathrm{e}^{-\mathrm{j}\,\beta_i}\right|^2\mu \tag{18}$$

$$\begin{aligned} \boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{K}\boldsymbol{g}_i = & k_{11}\cos^2\alpha_i \\ & + 2\cos\alpha_i\sin\alpha_i\Re\left\{k_{12}^{*}\,\mathrm{e}^{\mathrm{j}\,\beta_i}\right\} + k_{22}\sin^2\alpha_i \end{aligned} \tag{19}$$

From the definition of $\boldsymbol{D}_{ii}$ and $\boldsymbol{D}_{ij}$ in (15) and (16), respectively, it can be seen that the matrix parameters d_{ii} and d_{ij} are functions of $\boldsymbol{g}_j$ so that they can also be written as functions of α_j and β_j. After some calculations we get

$$d_{i\ell} = \frac{[\boldsymbol{D}_{i\ell}]_{12}}{[\boldsymbol{D}_{i\ell}]_{11}} = \frac{[\boldsymbol{M}_{i\ell}]_{12} + [\boldsymbol{M}_{i\ell}]_{22}\,t_j^{*}}{[\boldsymbol{M}_{i\ell}]_{11} + [\boldsymbol{M}_{i\ell}]_{21}\,t_j^{*}} \tag{20}$$

with $t_j := \tan\alpha_j\,\mathrm{e}^{\mathrm{j}\,\beta_j}$. We also define $t_i := \tan\alpha_i\,\mathrm{e}^{\mathrm{j}\,\beta_i}$.

In the following, we will ignore the measure zero event that the cosine of the optimal angle $\hat{\alpha}_i$ equals $\pi/2$, enabling us to reduce the fractions in (13) and (14) by powers of $\cos\alpha_i$. After multiplying those equations by $|t_i|/2$ and $1/2$, respectively, and inserting (18) through (20) and the respective derivatives, we get the equations (21) and (22), which are shown on the top of the next page.

For $\ell \in \{i, j\}$ we now define $t_{\ell,\mathrm{R}} = \Re\{t_\ell\}$, and $t_{\ell,\mathrm{I}} = \Im\{t_\ell\}$ and replace t_ℓ and $|t_\ell|^2$ by $t_{\ell,\mathrm{R}} + \mathrm{j}\,t_{\ell,\mathrm{I}}$ and $t_{\ell,\mathrm{R}}^2 + t_{\ell,\mathrm{I}}^2$, respectively. By expanding, carrying out the $\Re$-operations and multiplying by the three denominators, (21) and (22) can be converted to multivariate polynomial equations with four real valued variables $t_{i,\mathrm{R}}$, $t_{i,\mathrm{I}}$, $t_{j,\mathrm{R}}$ and $t_{j,\mathrm{I}}$. Since these polynomials have a total degree of 11 and consist of almost

[2] It is not necessary to consider any $\beta_i > \pi$ because instead of adding a value of π to β_i we can replace $\alpha_i \in [0, \pi]$ by $\alpha'_i = \pi - \alpha_i \in [0, \pi]$.

$$0 = \Re\left\{\frac{-|t_i|^2\left([\boldsymbol{M}_{ii}]_{11}^* + [\boldsymbol{M}_{ii}]_{21}^* t_j\right) + t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^* t_j\right)}{\left([\boldsymbol{M}_{ii}]_{11}^* + [\boldsymbol{M}_{ii}]_{21}^* t_j\right) + t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^* t_j\right)}\right\} + \Re\left\{\frac{-|t_i|^2\left([\boldsymbol{M}_{ij}]_{11}^* + [\boldsymbol{M}_{ij}]_{21}^* t_j\right) + t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^* t_j\right)}{\left([\boldsymbol{M}_{ij}]_{11}^* + [\boldsymbol{M}_{ij}]_{21}^* t_j\right) + t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^* t_j\right)}\right\} - \frac{|t_i|^2(k_{i,22} - k_{i,11}) + (1-|t_i|^2)\Re\left\{k_{i,12}^* t_i\right\}}{\left(k_{i,11} + k_{i,22}|t_i|^2 + 2\Re\left\{k_{i,12}^* t_i\right\}\right)} \tag{21}$$

$$0 = \Re\left\{\frac{\mathrm{j}\, t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^* t_j\right)}{\left([\boldsymbol{M}_{ii}]_{11}^* + [\boldsymbol{M}_{ii}]_{21}^* t_j\right) + t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^* t_j\right)}\right\} + \Re\left\{\frac{\mathrm{j}\, t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^* t_j\right)}{\left([\boldsymbol{M}_{ij}]_{11}^* + [\boldsymbol{M}_{ij}]_{21}^* t_j\right) + t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^* t_j\right)}\right\} - \frac{\Re\left\{\mathrm{j}\, k_{i,12}^* t_i\right\}}{k_{i,11} + k_{i,22}|t_i|^2 + 2\Re\left\{k_{i,12}^* t_i\right\}} \tag{22}$$

a thousand terms, they can neither be expanded by hand nor be written down here. To generate the expanded forms we used the Symbolic Math Toolbox of MATLAB. With two polynomials for each $i \in \{1,2\}$, we have in total four equations and four unknowns so that the solutions of the multivariate polynomial system are distinct intersection points.

Solving systems of polynomial equations is well investigated and can be handled by a solver like PHCpack [10], which computes approximations to all isolated solutions of the polynomial system. In doing so, we obtain all local extrema and only have to compare their rate offsets to find the global optimum. Simulating 1,000 i. i. d. complex Gaussian channels, we have obtained between 3 and 14 solutions of the polynomial system. On average we have found 5.36.

While calculation is fast for real valued channels where $\beta_i = 0$ and the optimality conditions reduce to a system of two bivariate polynomials, our method might be difficult to apply in practice for complex valued channels since computing the solutions to the full system of four polynomials takes much time.

4.4. A Game Theoretic View

We now use the same parametrization as above to interpret the process of choosing the receive filters as a noncooperative game with two players. In order to maximize his payoff, i. e. his own rate offset $\mathcal{R}_i$, each player i may pick a strategy $(\alpha_i, \beta_i) \in \mathbb{A}_i$ from the set of available strategies $\mathbb{A}_i = [0,\pi] \times [0,\pi]$. Thus, each possible combination of receive filters $(\boldsymbol{g}_1, \boldsymbol{g}_2)$ can be interpreted as a strategy $(\alpha_1, \beta_1, \alpha_2, \beta_2) =: a \in \mathbb{A} = [0,\pi] \times [0,\pi] \times [0,\pi] \times [0,\pi]$. A strategy $\tilde{a}$ is called a *Nash equilibrium* (NE), if no player can improve his payoff by changing his own strategy from $(\tilde{\alpha}_i, \tilde{\beta}_i)$ to a different (α_i, β_i), assuming that the other player sticks to his strategy $(\tilde{\alpha}_j, \tilde{\beta}_j)$ [11]. In other words, the combination of the angle $\tilde{\alpha}_1$ and the phase $\tilde{\beta}_1$ has to be a global maximizer of $\mathcal{R}_1(\alpha_1, \beta_1, \tilde{\alpha}_2, \tilde{\beta}_2)$ while $(\tilde{\alpha}_2, \tilde{\beta}_2)$ has to maximize $\mathcal{R}_2(\tilde{\alpha}_1, \tilde{\beta}_1, \alpha_2, \beta_2)$.

To find the NE, we now maximize $\mathcal{R}_i$ with respect to α_i and β_i. According to (12) we have

$$\mathcal{R}_i = \log\left(\boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}_{ii} \boldsymbol{P}^{\mathrm{H}} \left(\boldsymbol{H}_{ji}^{\mathrm{H}} \boldsymbol{g}_j \boldsymbol{g}_j^{\mathrm{H}} \boldsymbol{H}_{ji}\right)^* \boldsymbol{P} \boldsymbol{H}_{ii}^{\mathrm{H}} \boldsymbol{g}_i\right) - \log\left(\boldsymbol{g}_j^{\mathrm{H}} \boldsymbol{H}_{ji} \boldsymbol{H}_{ji}^{\mathrm{H}} \boldsymbol{g}_j\right)$$

with the permutation matrix $\boldsymbol{P}$ defined as in (11). The second summand depends neither on α_i nor on β_i and can therefore be dropped from the maximization. To optimize the first summand, we only have to maximize the argument of the logarithm, as the logarithm is a monotonic function:

$$\max_{\alpha_i, \beta_i} \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{D}_i \boldsymbol{g}_i$$
$$\text{with} \quad \boldsymbol{D}_i = \boldsymbol{M}_i^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{M}_i, \quad \boldsymbol{M}_i = \boldsymbol{H}_{ji}^* \boldsymbol{P} \boldsymbol{H}_{ii}^{\mathrm{H}}. \tag{23}$$

As $\boldsymbol{D}_i$ is Hermitian and has $\mathrm{Rank}[\boldsymbol{D}_i] = 1$, we can use (17) and (18). To maximize the absolute value of a sum, both summands must point in the same direction, i. e. they must have the same phase. Therefore, to maximize (18) it must hold that $\mathrm{e}^{\mathrm{j}(\arg(d_i) - \beta_i)} = \pm 1$.[3] Without loss of generality we choose $\beta_i = \arg(d_i)$.[4] We insert this into the derivative of (18) with respect to α_i and set the derivative to zero:

$$0 = \tan\alpha_i(|d_i|^2 - 1) + (1 - \tan^2\alpha_i)|d_i|$$

where we have divided by $\cos^2\alpha_i$ ignoring the measure zero event that the optimal angle $\tilde{\alpha}_i$ equals $\pi/2$. This condition is fulfilled if $\tan\alpha_i = |d_i|$ or if $\tan\alpha_i = -1/|d_i|$. By calculating the second derivative we can see that the former possibility is the maximizer while the latter minimizes $\mathcal{R}_i$. Hence, below we always assume that $|d_i| = \tan\alpha_i$. Thus, with $t_i := \tan\alpha_i \cdot \mathrm{e}^{\mathrm{j}\,\beta_i}$ and $\mu_i' := (\cos^2\alpha_i)/\mu_i$ the dyadic product $\boldsymbol{g}_i \boldsymbol{g}_i^{\mathrm{H}}$ can be written as

$$\boldsymbol{g}_i \boldsymbol{g}_i^{\mathrm{H}} = \cos^2\alpha_i \begin{bmatrix} 1 & t_i \\ t_i^* & |t_i|^2 \end{bmatrix} = \mu_i' \boldsymbol{D}_i.$$

This can be inserted twice into the definition of $\boldsymbol{D}_i$ in (23):

$$\boldsymbol{D}_1 = \boldsymbol{M}_1^{\mathrm{H}} \boldsymbol{g}_2^* \boldsymbol{g}_2^{\mathrm{T}} \boldsymbol{M}_1 = \mu_2' \boldsymbol{M}_1^{\mathrm{H}} \boldsymbol{D}_2^* \boldsymbol{M}_1$$
$$= \mu_1' \mu_2' \boldsymbol{M}_1^{\mathrm{H}} \boldsymbol{M}_2^{\mathrm{T}} \boldsymbol{D}_1 \boldsymbol{M}_2^* \boldsymbol{M}_1 =: \mu \boldsymbol{M}^{\mathrm{H}} \boldsymbol{D}_1 \boldsymbol{M} \tag{24}$$

[3]The arg-operation gives us the phase of a complex number.

[4]The resulting vector $\boldsymbol{g}_i$ can later be transformed according to Footnote 2 such that the assumption $\beta_i \in [0;\pi]$ is fulfilled.

see the program

In (24), $\boldsymbol{M} = \boldsymbol{M}_2^* \boldsymbol{M}_1$ and $\mu = \mu_1' \mu_2'$. Due to $d_1 = [\boldsymbol{D}_1]_{12} / [\boldsymbol{D}_1]_{11}$ we get after some calculations:

$$[\boldsymbol{M}]_{21} d_1^2 + ([\boldsymbol{M}]_{11} - [\boldsymbol{M}]_{22}) d_1 - [\boldsymbol{M}]_{12} = 0 \qquad (25)$$

From this quadratic equation we can explicitly calculate two solutions for d_1.[5] For d_2 we then get

$$d_2 = \frac{[\boldsymbol{D}_2]_{12}}{[\boldsymbol{D}_2]_{11}} = \frac{\left[\boldsymbol{M}_2^{\mathrm{H}} \boldsymbol{D}_1^* \boldsymbol{M}_2\right]_{12}}{\left[\boldsymbol{M}_2^{\mathrm{H}} \boldsymbol{D}_1^* \boldsymbol{M}_2\right]_{11}}. \qquad (26)$$

Two NE $(\tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\alpha}_2, \tilde{\beta}_2)$ can now be explicitly determined by calculating[6] $\tilde{\alpha}_i = \arctan(|d_i|)$ and $\tilde{\beta}_i = \arg(d_i)$ for both solutions d_1 of (25) with d_2 given by (26).

Although suboptimal in terms of sum rate offset, the NE are of practical interest, as they are a stable outcome from which no user will deviate. This solution differs from the NE discussed in [2, 4] as we changed the rules of the game such that each user is constrained to transmit only one data stream. It was stated in [2] that in order to achieve high data rates, cooperation is desirable when interference is the limiting factor. Thus, this new rule obviously makes sense. In fact, our middle course between competition and cooperation guarantees that the optimal high SNR slope is achieved, which is not the case when full competition is allowed.

5. THE HIGH SNR RATE OFFSET REGION

We now extend the well-established concept of the rate region to the high SNR rate offset by plotting the possible combinations of both users' rate offsets in the $\mathcal{R}_1$-$\mathcal{R}_2$-plane. The resulting area is called the high SNR rate offset region. As the high SNR slope is the dominant high SNR performance criterion (cf. Section 3.4), the high SNR rate offset region only makes sense if a certain combination $(\mathcal{S}_1, \mathcal{S}_2)$ of high SNR slopes is specified in advance, for which the region shall be drawn. In our case this is $(\mathcal{S}_1, \mathcal{S}_2) = (1, 1)$.

We can approximate the convex hull of the high SNR rate offset region by calculating weighted sum rates using a slightly modified version of (21) and (22). In Fig. 3, the points maximizing the weighted sum rate offset for $w_1 = \sin\theta$, $w_2 = \cos\theta$ with $\theta = n\Delta\theta \in [0; \pi/2]$, $n \in \mathbb{N}_0$ and $\Delta\theta = \pi/64$ can be seen for a certain channel realization. As approximation for the region we used the convex hull of these points, extended up to $-\infty$ in both directions.[7] Brute force simulations show that the high SNR rate offset region does not necessarily have to be convex. However, any point within the convex hull can be achieved by performing time sharing between two points inside the rate offset region.

[5] The equation can always be solved because we are operating $\mathbb{C}$.

[6] Strictly speaking, after the arg-operation, we have to add or subtract π such that $\beta_i \in [0; \pi]$ and α_i has to be updated according to Footnote 2.

[7] In contrast to the rate, the high SNR rate offset can reach negative values so that the high SNR rate offset region is not bounded at the bottom and at the left side.

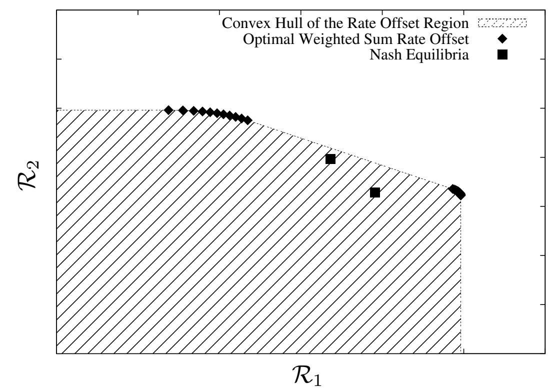


Fig. 3. Example for the Approximation of the Convex Hull of a High SNR Rate Offset Region.

Strategy	$\mathcal{R}$ (bits/channel use)
Globally Optimal $\mathcal{R}$	2.13
User 1 Prioritized	0.74
Nash Equilibrium (NE)	1.27

Table 1. Sum Rate Offset $\mathcal{R}$ Achieved on Average by Different Strategies

The rhombs corresponding to $\theta = 0$ and $\theta = \pi/2$ are equivalent to the optima for systems with a prioritized user. We also depict the NE in terms of Section 4.4, which are obviously suboptimal in terms of high SNR rate offset.

6. NUMERICAL RESULTS

To study the average behaviour of the proposed schemes we have computed the high SNR sum rate offset achieved by different strategies for 1,000 i. i. d. complex Gaussian channel realizations with zero mean und unit variance. The averaged values can be seen in Table 1. As expected, the optimal strategy for systems with a prioritized user performs suboptimal when sum rate offset is the performance criterion. The NE scheme performs quite well, when the best out of the two possible NE is chosen as has been done in the simulations. All proposed suboptimal schemes are clearly outperformed by the globally optimal solution.

However, since all of them can be explicitly calculated, while the global optimum exhibits a considerable computational complexity, the proposed suboptimal schemes may be a good choice for practical implementation.

Although the methods proposed in this paper are based on the assumption of asymptotically high SNR, simulations show that they also perform well for finite SNR values. For the results in Fig. 4, three of the proposed strategies were computed for 1,000 i. i. d. complex Gaussian channel realizations, and the actual sum rate resulting from the use of the obtained beamformers and receive filters at finite SNR

see the program

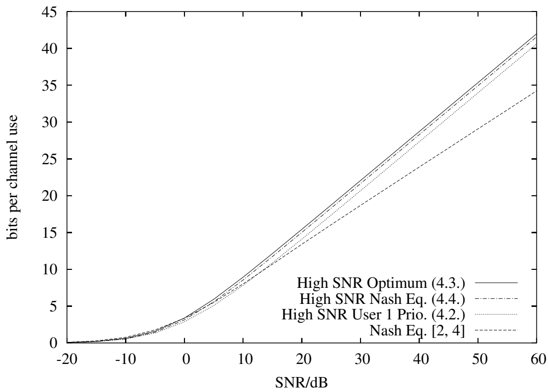


Fig. 4. Sum Rate Performance of the High SNR Solutions in a Two Stream MIMO IFC.

is plotted. For all three strategies we assumed the constraint that the slope has to be distributed equally among the two users because this was the assumption of the whole paper. However, for the rate offset optimal solution, a possible variation for finite SNR would be to compare the achieved sum rate with that of the two single user points $\mathcal{S}_1 = 2$ and $\mathcal{S}_2 = 2$ and use the strategy out of those three that yields the highest sum rate. For comparison, the game theoretic strategy from [2,4] is shown, which clearly does not achieve the optimal high SNR slope.

7. SUMMARY AND OUTLOOK

We have studied the class of MIMO interference channels that allow a maximum number of two data streams with the aim of maximizing the sum rate and finding a region of achievable rates. We have shown that all systems covered by our assumptions can be transformed to a 2×2 MIMO IFC and further, by considering the receive filters as a part of the channel, to a 2x1 MISO IFC. Assuming asymptotically high SNR, this has enabled us to find globally optimal filters for systems with a prioritized user and for systems where both users have equal priority. Furthermore, we derived a formula to explicitly calculate the Nash Equilibria when the process of choosing the beamforming vectors is interpreted as a non-cooperative game. Instead of drawing a rate region, we introduced the concept of the high SNR rate offset region. Numerical results show that the proposed algorithms perform well for high, but finite SNR values although they are only optimal for asymptotically high SNR.

A crucial step in the derivation of our method was the fact that the optimal beamformers for the MISO IFC at high SNR can be written in an explicit form. Therefore, another interesting question for further research would be if the recently proposed parametrization of Pareto optimal beamformers for the MISO IFC [12] can be used to develop a similar method without making use of the high SNR assumption. This would introduce one additional scalar optimization variable per user and the objective function would have to be changed from the high SNR sum rate offset to the sum rate at a given SNR.

8. REFERENCES

[1] E. Akuiyibo, O. Lévêque, and C. Vignat, "High SNR Analysis of the MIMO Interference Channel," in *Proc. ISIT 2008*, Jul. 2008, pp. 905–909.

[2] S. Ye and R. S. Blum, "Optimized Signaling for MIMO Interference Systems with Feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[3] E. G. Larsson, D. Danev, and E. A. Jorswieck, "Asymptotically Optimal Transmit Strategies for the Multiple Antenna Interference Channel," in *Proc. of Allerton Conference on Communication, Control, and Computing*, Sep. 2008.

[4] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive Design of Multiuser MIMO Systems based on Game Theory: A Unified View," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, Sep. 2008.

[5] S. A. Jafar and M. J. Fakhereddin, "Degrees of Freedom for the MIMO Interference Channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, Jul. 2007.

[6] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR Power Offset in Multiantenna Communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

[7] I. E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[8] D. A. Schmidt, A. Gründinger, W. Utschick, and M. L. Honig, "Distributed Precoder Optimization for Interfering MISO Channels," in *Proc. WSA '08*, Feb. 2008, pp. 285–290.

[9] E. G. Larsson and E. A. Jorswieck, "Competition versus cooperation on the MISO interference channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1059–1069, Sep. 2008.

[10] J. Verschelde, "PHCpack: A general-purpose solver for polynomial systems by homotopy continuation," *ACM Transactions on Mathematical Software*, vol. 25, no. 2, pp. 251–276, 1999.

[11] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, Eds., *Algorithmic Game Theory*, 1st ed. New York, NY, USA: Cambridge University Press, 2007.

[12] E. G. Larsson, E. A. Jorswieck, and D. Danev, "Complete Characterization of the Pareto Boundary for the MISO Interference Channel," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292–5296, Oct. 2008.

see the program

ANALYSIS OF FUSION AND COMBINING FOR WIRELESS SOURCE DETECTION

Rajesh K. Sharma and Jon W. Wallace

School of Engineering and Science, Jacobs University
Campus Ring 1, 28759, Bremen, Germany
ra.sharma@jacobs-university.de, wall@ieee.de

ABSTRACT

In wireless detection scenarios such as sensor networks or cognitive radio, "hidden nodes" can occur due to channel fading, which can be overcome by collaborative sensing. This work considers the performance of collaboration with combining and fusion methods for both static and fading environments, where each node employs energy detection for sensing an unknown source. The case of combined Rayleigh fading and shadowing is also considered, showing that AND fusion can have even lower performance than a single sensor. Our results show that choosing a fusion rule without considering the statistics of the fading environment can lead to very poor performance. It also proposes a new quantized equal gain combining (QEGC) method that strikes a good balance between performance and transmission overhead among the collaborating nodes.

1. INTRODUCTION

Source detection is an important component for sensor networks such as surveillance radar, sonar and many civilian applications related to public safety, security, health care, etc. Recently, a new cognitive radio technology has gained significant attention due to its potential to cope with spectrum scarcity [1]. An important component of cognitive radio is primary source detection (sensing) which is aimed at detecting the presence or absence of a primary user in the spectrum of interest. Collaborative sensing has been proposed as a technique to avoid the "hidden node" problem, which arises due to fading [2]. The information from different sensors is combined in order to make a global decision about the presence or absence of the wireless source.

Different combining techniques have been discussed in [3], which include equal gain combining (EGC), selection combining (SC), and switch and stay combining (SSC). Similarly in [4], maximal ratio combining (MRC) has been shown to be the optimal soft combining method whose performance has been compared with equal gain combining. However, for sensing unknown signals, the weighting factor needed for MRC may not be available, due to the lack of channel information. Recent work in [5] indicates that the optimal sensor fusion rule, in the sense of a locally most powerful (LMP) test, is to use EGC for the low SNR regime, which is a reasonable worst-case assumption when signal and channel characteristics are unknown.

In the presence of small-scale fading, the signal independence required by these soft combining schemes is obtained at sub-wavelength scale separation, allowing easy implementation with nodes having multiple antennas or wired connections. In the case of shadowing, however, fading may be correlated for many wavelengths, requiring nodes to have significant separation for independence, and complicating the transmission of node parameters to the decision center that makes the global decision. In such cases, decision fusion methods can be used which take only the *decisions* of the individual sensors and fuse them to obtain the global decision.

Fusion techniques normally employ a 'k out of n rule' (KNR) where n is the number of sensors. The special cases are OR ($k = 1$), AND ($k = n$) or majority gates $k = \lceil n/2 \rceil$, where $\lceil \cdot \rceil$ denotes rounding up to the next integer value. There exist conflicting claims on the superiority of OR fusion and AND fusion [6], [7]. Fefjar in [6] showed the superiority of OR fusion, which was refuted by Stearns in [7], claiming that OR and AND performances intersect, AND being superior for low false alarm probabilities. In [8], the performances of OR and AND fusion are compared, where the decision parameters are assumed to be exponentially distributed with different decay constants for sensors that are not identical, demonstrating that the superiority and intersection all depend upon the global false alarm probabilities as well as the distributions.

As an optimal data fusion technique based on a maximum likelihood ratio test, [9] has proposed a fusion rule in which the decisions 1 or -1 (for existence or absence of the source) are weighted according to sensor reliability, where reliability is a function of the probabilities of false alarm and missed detection. A decision fusion rule with the Neyman-Pearson (N-P) test is described in [10], showing that fusion can provide increased global probability of detection for a fixed or lower global probability of false alarm relative to a single sensor, when more than two sensors are available.

This work compares the performance of different cases of k out of n rules for the decision fusion in unknown source detection employing conventional energy detection at each node. The fusion techniques (hard combining) are also com-

see the program

pared with EGC (soft combining), which is the main feasible soft-combining method for an unknown source. Different possible distributions of the energy in individual sensors are considered in static and fading environments. Analytical results are also verified by Monte Carlo simulation for all the cases considered.

The rest of the paper is organized as follows: Section 2 introduces the problem of spectrum sensing, reviews energy detection, and discusses the effect of fading on signal sensing. Section 3 describes different collaborative approaches, such as equal gain combining and fusion methods. Section 4 compares the performance of different combining and fusion techniques in static and fading environments using closed form expressions and Monte Carlo simulations. Conclusion and some future extensions are discussed in Section 5.

2. SINGLE SENSOR DETECTION

2.1. Energy Detector-Based Sensing

First we review energy detection, being the most common method of spectrum sensing, due to its low computational complexity and ease of implementation [3], [11], [12].

Considering the bandpass noise within a fixed sensing bandwidth to have flat power spectral density (PSD), the noise is represented as

$$\eta(t) = \eta_c(t)\cos 2\pi f_c t - \eta_s(t)\sin 2\pi f_c t, \qquad (1)$$

where f_c is the reference frequency, and $\eta_c(t)$ and $\eta_s(t)$ are the in-phase and quadrature modulation components, respectively. If the bandpass noise in (1) has bandwidth W, $\eta_c(t)$ and $\eta_s(t)$ have bandwidth $W/2$. The variance of $\eta_c(t)$, $\eta_s(t)$ and $\eta(t)$ are all equal to the noise power.

The energy in a continuous sample of finite duration T is often approximated as [11]

$$U = \int_0^T \eta^2(t)\, dt \approx \frac{1}{2W}\sum_{i=1}^{TW}(a_{ci}^2 + a_{si}^2) = N_0 \sum_{i=1}^{TW}(b_{ci}^2 + b_{si}^2), \qquad (2)$$

where a_{ci} and a_{si} are the ith discrete samples (at a rate of $1/W$) of $\eta_c(t)$ and $\eta_s(t)$, respectively, $b_{ci} = a_{ci}/\sqrt{\sigma_i^2}$, $b_{si} = a_{si}/\sqrt{\sigma_i^2}$, $\sigma_i^2 = Var(a_{ci}) = Var(b_{si}) = 2N_0W$, and N_0 is a two-sided PSD. Scaling U in (2) by defining $U' = U/N_0$,

$$U' = \sum_{i=1}^{TW} b_{ci}^2 + \sum_{i=1}^{TW} b_{si}^2, \qquad (3)$$

which follows a central chi-square (χ^2) distribution with degree of freedom $2TW$.

When the signal is present, U' follows non-central chi-square distribution with $2TW$ degrees of freedom and a non-centrality parameter γ', which is the ratio of source signal energy to noise PSD (two-sided) [11]. In terms of SNR (γ), which is the ratio of source signal to noise power, γ' can be expressed as $2TW\gamma$. The decision statistic for this detector can be described compactly as

$$U' \sim \begin{cases} \chi^2_{2TW} & \mathsf{H}_0, \\ \chi^2_{2TW}(\gamma') & \mathsf{H}_1, \end{cases} \qquad (4)$$

where hypotheses H_0 and H_1 refer to the absence or presence of the source, respectively. The probability of false alarm P_f and the probability of correct detection P_d for this scheme can be calculated as [3]

$$P_f = Pr(U' > \lambda|\mathsf{H}_0) = \frac{\Gamma\left(u, \frac{\lambda}{2}\right)}{\Gamma(u)} \qquad (5)$$

and

$$P_d = Pr(U' > \lambda|\mathsf{H}_1) = Q_u\left(\sqrt{\gamma'}, \sqrt{\lambda}\right), \qquad (6)$$

where λ is the decision threshold, $\Gamma(.\,,\,.)$ is the incomplete gamma function, $u = TW$, and $Q_u(.\,,\,.)$ is the generalized Marcum Q-function. As expected, P_f is independent of γ' since under H_0 there is no transmission from the source. For longer sensing duration the distributions in both hypotheses can be approximated as Gaussian. The decision statistic can be described as

$$U' \sim \begin{cases} N(\mu_n, \sigma_n^2) & \mathsf{H}_0, \\ N(\mu_{sn}, \sigma_{sn}^2) & \mathsf{H}_1, \end{cases} \qquad (7)$$

where

$$\mu_n = 2TW,$$
$$\sigma_n^2 = 4TW,$$
$$\mu_{sn} = 2TW(1+\gamma),$$

and

$$\sigma_{sn}^2 = 4TW(1+2\gamma), \qquad (8)$$

respectively. For this Gaussian-distributed case, P_f and P_d can be calculated as

$$P_f = Pr(U' > \lambda|\mathsf{H}_0) = \frac{1}{2}\mathrm{erfc}\left[\frac{(\lambda - \mu_n)}{\sigma_n\sqrt{2}}\right] \qquad (9)$$

and

$$P_d = Pr(U' > \lambda|\mathsf{H}_1) = \frac{1}{2}\mathrm{erfc}\left[\frac{(\lambda - \mu_{sn})}{\sigma_{sn}\sqrt{2}}\right], \qquad (10)$$

where λ is the decision threshold and $\mathrm{erfc}(\cdot)$ is the complimentary error function.

2.2. Sensing in a Fading Channel

In practice, the sensing problem involves both small-scale (multipath) and large-scale (shadow) fading. These two effects are commonly treated as independent processes that combine to produce the overall fading effect [13], and this effect usually degrades the performance of spectrum sensing methods. The hidden node problem in cognitive radio, for example, poses a challenge to spectrum sensing, which happens when the primary source is not detected by a cognitive radio receiver due to fading, but the transmit power from this cognitive radio interferes with a primary receiver.

The distribution of the envelope of received signal in small-scale fading is commonly taken to be Rayleigh distributed, making the received power distribution exponential [14]. The Rayleigh probability density function (pdf) is

$$f_R\left(r\right) = \frac{r}{\sigma_r^2} e^{-\frac{r^2}{2\sigma_r^2}}, \quad r \geq 0, \tag{11}$$

where $2\sigma_r^2$ is the average power and r is the faded envelope of the signal.

The slow varying local mean power due to shadowing is commonly assumed to follow the lognormal distribution

$$f_{P_o}\left(p_o\right) = \frac{1}{\sqrt{2\pi}\sigma p_o} e^{-\frac{\left(\ln\, p_o - \mu\right)^2}{2\sigma^2}}, \quad p_o \geq 0, \tag{12}$$

where μ and σ^2 are the mean and variance of the logarithm of local mean power. If the power is expressed in dB, the distribution becomes Gaussian, with the mean and standard deviation reflecting the average power and shadowing variation, both expressed in dB.

The P_f and P_d we discussed in 2.1 do not consider the fading effects. If the channel gain is varying due to shadowing or multipath, (6) gives the probability of detection conditioned on the instantaneous SNR γ'. The average probability of detection is obtained by averaging (6) over the fading statistics [2], which is given as

$$P_d = \int_0^\infty Q_u\left(\sqrt{\gamma'}, \sqrt{\lambda}\right) f\left(\gamma'\right) d\gamma', \tag{13}$$

where $f\left(\gamma'\right)$ is the pdf of γ' under fading. For a Rayleigh fading channel $f\left(\gamma'\right)$ is given as

$$f\left(\gamma'\right) = \frac{1}{\overline{\gamma'}} e^{-\frac{\gamma'}{\overline{\gamma'}}}, \tag{14}$$

where $\overline{\gamma'}$ is the average value of γ'. Letting $x = \sqrt{\gamma'}$, we obtain

$$P_d = \frac{2}{\overline{\gamma'}} \int_0^\infty Q_u\left(x, \sqrt{\lambda}\right) x\, e^{-\frac{x^2}{\overline{\gamma'}}} dx. \tag{15}$$

This can be derived in closed form, using (12) from [15], as

$$P_d = e^{-\frac{\lambda}{2}} \sum_{m=0}^{u-2} \frac{1}{m!} \left(\frac{\lambda}{2}\right)^m + \left(\frac{2+\overline{\gamma'}}{\overline{\gamma'}}\right)^{u-1} \left[e^{-\frac{\lambda}{2+\overline{\gamma'}}} - e^{-\frac{\lambda}{2}} \sum_{m=0}^{u-2} \frac{1}{m!} \left(\frac{\lambda\overline{\gamma'}}{2\left(2+\overline{\gamma'}\right)}\right)^m \right]. \tag{16}$$

A similar result is also obtained in [2] and [3]. However, note that due to our definition of SNR (a two-sided PSD is used), the expression is slightly modified. Also note that [3] appears to have a missing exponent, which is corrected in [2] and in (16). The analytical expression for P_d in the combined fading environment, consisting of both Rayleigh fading and lognormal shadowing, is not yet available in the literature, and simulation studies on the effect of this combined fading is an important contribution of this work, providing additional insight on the sensing problem for realistic environments.

3. MULTIPLE SENSOR DETECTION

3.1. Fusion (Hard Combining Methods)

Fusion techniques normally employ k out of n rule (KNR), where n is the number of sensors. In this fusion rule the global decision is 1 if any k or more sensor has output 1, where the decisions 1 or -1 refer to presence or absence of the source. Therefore, KNR fusion is represented mathematically as

$$\mathcal{F}\left(u_1, \cdots, u_n\right) \sim \begin{cases} 1, & \sum_{i=1}^n u_i \geq 2k-n, \\ -1, & \text{otherwise.} \end{cases} \tag{17}$$

For OR fusion, at least one output must be 1 to have the global decision 1, so the decision rule is given as

$$\mathcal{F}\left(u_1, \cdots, u_n\right) \sim \begin{cases} 1, & \sum_{i=1}^n u_i \geq 2-n, \\ -1, & \text{otherwise.} \end{cases} \tag{18}$$

Similarly, in AND fusion the decision is 1 only if all sensors have output 1, or

$$\mathcal{F}\left(u_1, \cdots, u_n\right) \sim \begin{cases} 1, & \sum_{i=1}^n u_i = n, \\ -1, & \text{otherwise.} \end{cases} \tag{19}$$

It is obvious that OR and AND cases are the special cases of KNR with $k=1$ and $k=n$, respectively. In the majority gate rule, k is taken to be $\lceil n/2 \rceil$.

$$\mathcal{F}\left(u_1, \cdots, u_n\right) \sim \begin{cases} 1, & \sum_{i=1}^n u_i \geq 0, \\ -1, & \text{otherwise.} \end{cases} \tag{20}$$

The decision fusion rules discussed above do not consider the distribution of the decision statistics, the individual P_f and P_d, nor the global desired false alarm probability $P_{f,T}$ when making the global decision. The global receiver operating characteristics (ROC) obtained from theses rules may not be optimal for all values of $P_{f,T}$. The optimal data fusion rule given in [9] is

$$\mathcal{F}\left(u_1, \cdots, u_n\right) \sim \begin{cases} 1, & a_0 + \sum_{i=1}^n a_i u_i > 0, \\ -1, & \text{otherwise,} \end{cases} \tag{21}$$

see the program

where the optimum weights are given by

$$a_0 = \log \frac{P_1}{P_0} \tag{22a}$$

$$a_i = \begin{cases} \log \dfrac{P_{d,i}}{P_{f,i}}, & u_i = +1, \\ \log \dfrac{1-P_{f,i}}{1-P_{d,i}}, & u_i = -1, \end{cases} \tag{22b}$$

where P_0 and P_1 are the a-priori probabilities of the two hypotheses. If the a priori probabilities are unknown, then the optimal decision scheme should be based on an N-P test which has been discussed in [10], obtaining a similar result as that in (17), except for the a_0 term. This optimal data fusion rule is different from the majority gate rule in the sense that the value of k here changes according to the P_f and P_d pairs under consideration. The global false alarm and detection probabilities with KNR are given as

$$P_{f,T} = \sum_{i=k}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) P_f^i \left(1-P_f\right)^{n-i}, \tag{23}$$

and

$$P_{d,T} = \sum_{i=k}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) P_d^i \left(1-P_d\right)^{n-i}, \tag{24}$$

respectively, where $\left(\begin{array}{c} n \\ i \end{array} \right) = \dfrac{n!}{(n-i)!i!}$ is the number of combinations of the sensor outputs.

3.2. Decision Parameter Combining (Soft Combining) Methods

Several combining techniques for iid Rayleigh environments are developed in [3], where for the case of equal gain combining, the sum of Rayleigh-distributed random variables is modeled as a Nakagami distribution with suitable parameter values. The important results are also derived here, which is necessary due to our specific definition of SNR. Also, since the threshold after combining should clearly change, we account for this in our derivation.

When equal gain combining is used with n independent sensors, the energy is increased as $\gamma_T' = \sum_{i=1}^{n} \gamma_i'$. Here the concept of equal gain combining is different than that used in multiple antenna diversity schemes. The decision parameter (energy in our case) of the individual nodes is simply added here instead of co-phasing and adding the signals from different antennas in conventional diversity. Diversity-like maximum ratio or equal gain combining is usually not possible for unknown source detection since SNRs are typically very low and there is no cooperation from the source.

Considering energy detection, the total output energy of the combiner again has a non-central chi-square distribution with degree of freedom $2nTW$ and non-centrality parameter γ_T'. Similarly when no signal is present, the combiner output has a central chi-square distribution with degree of freedom $2nTW$. The $P_{f,T}$ and $P_{d,T}$ at the equal gain combiner output for the AWGN channel can be evaluated in a manner analogous to (5) and (6) as

$$P_{f,T} = Pr\left(U_T' > \lambda_T | \mathsf{H}_0\right) = \frac{\Gamma\left(nu, \frac{\lambda_T}{2}\right)}{\Gamma\left(nu\right)} \tag{25}$$

and

$$P_{d,T} = Pr\left(U_T' > \lambda_T | \mathsf{H}_1\right) = Q_{nu}\left(\sqrt{\gamma_T'}, \sqrt{\lambda_T}\right), \tag{26}$$

where λ_T is the decision threshold of the combiner. In the Rayleigh fading channel, the distribution of γ_T' is that of the sum of n independent exponentially distributed random variables, each having a mean of $\overline{\gamma'}$, which is represented as

$$f\left(\gamma_T'\right) = \gamma_T'^{n-1} \frac{e^{\frac{-\gamma_T'}{\overline{\gamma'}}}}{\Gamma\left(n\right) \overline{\gamma'}^{n}}, \tag{27}$$

where $\Gamma\left(n\right) = (n-1)!$ for integer n. The average $P_{d,T}$ for Rayleigh fading can then be obtained as

$$P_{d,T,Ray} = \int_0^{\infty} Q_{nu}\left(\sqrt{\gamma_T'}, \sqrt{\lambda_T}\right) f\left(\gamma_T'\right) d\gamma_T'. \tag{28}$$

Changing the variable $x = \sqrt{\gamma_T'}$, we obtain

$$\begin{aligned} P_{d,T,Ray} &= \frac{2}{\Gamma(n)\, \overline{\gamma'}^{n}} \int_0^{\infty} Q_{nu}\left(x, \sqrt{\lambda_T}\right) x^{2n-1} e^{-\frac{x^2}{\overline{\gamma'}}} dx \\ &= \alpha \left[\Psi_1 + \beta \sum_{i=1}^{nu-1} \frac{\left(\frac{\lambda_T}{2}\right)^i}{2 \times i!} \, {}_1F_1\left(n; i+1; \frac{\lambda_T}{2} \frac{\overline{\gamma'}}{2+\overline{\gamma'}}\right) \right], \end{aligned} \tag{29}$$

where $\alpha = 2/[\Gamma(n)\overline{\gamma'}^{n}]$, ${}_1F_1$ is the confluent hypergeometric function defined as

$${}_1F_1\left(a; b; z\right) = \sum_{i=1}^{\infty} \frac{(a)_i \, z^i}{(b)_i \, i!} \tag{30}$$

with $(a)_i$ given as $(a)_i = a\left(a+1\right)\left(a+2\right)...\left(a+i-1\right)$, $\beta = \Gamma\left(n\right) \left(\dfrac{2\overline{\gamma'}}{2+\overline{\gamma'}}\right)^n e^{-\frac{\lambda_T}{2}}$,

and

$\Psi_1 = \displaystyle\int_0^{\infty} Q_1\left(x, \sqrt{\lambda_T}\right) x^{2n-1} e^{-\frac{x^2}{\overline{\gamma'}}} dx$, where $Q_1\left(\cdot, \cdot\right)$ is the first order Marcum Q-function. Ψ_1 can be evaluated

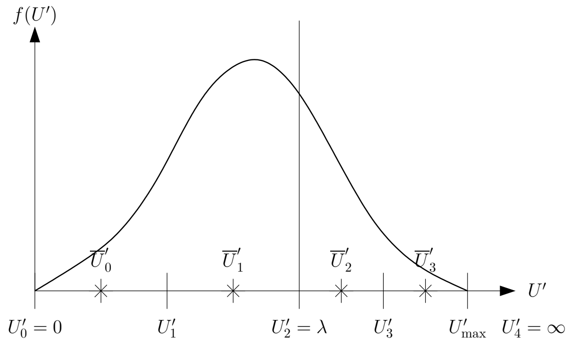


Figure 1: Computation of parameters for two-bit energy quantization using the energy distribution of noise

using [15] as

$$\Psi_1 = \frac{(n-1)!\overline{\gamma'}^{n+1}}{2\left(2+\overline{\gamma'}\right)} e^{-\frac{\lambda_T}{2+\overline{\gamma'}}} \times \left(\sum_{k=0}^{n-2} \left(\frac{2}{2+\overline{\gamma'}}\right)^k L_k\left(-\frac{\lambda_T}{2}\frac{\overline{\gamma'}}{2+\overline{\gamma'}}\right) + \frac{2^n}{\overline{\gamma'}}\left(2+\overline{\gamma'}\right)^{2-n} L_{n-1}\left(-\frac{\lambda_T}{2}\frac{\overline{\gamma'}}{2+\overline{\gamma'}}\right)\right), \tag{31}$$

where $L_k(x)$ is known as the Laguerre polynomial of degree k defined as

$$L_k(x) = \frac{e^x}{k!}\frac{d^k}{dx^k}\left(\frac{e^{-x}}{x^k}\right). \tag{32}$$

The values of $L_0(x)$ and $L_1(x)$ are 1 and $-x+1$, respectively. Using these values, the value of $L_k(x)$ can be computed using the recursive formula [16]

$$L_{k+1}(x) = \frac{1}{k+1}\left[(2k+1-x)L_k(x) - kL_{k-1}(x)\right]. \tag{33}$$

3.3. Decision Parameter Combining with Quantization (QEGC) Method

To improve the performance of fusion techniques without large investment in the cooperating bandwidth, an equal gain combining scheme with quantized decision parameters is proposed. Here the energy of each sensor is quantized using four levels as depicted in Figure 1, where two levels are used above and below the threshold. The identical quantization levels are known by all sensors as they are chosen based only on the noise distribution.

After choosing a suitable maximum energy level $U'_{\max}$, below which most of the probability falls (for the noise-only case), the midpoints below and above the threshold are computed as $U'_1 = \lambda/2$ and $U'_3 = (U'_{\max}+\lambda)/2$. When a node observes the energy on the interval $[U'_i, U'_{i+1}]$, the two-bit binary representation of i is transmitted, which is decoded as the quantized level $\overline{U}'_i = (U'_i + U'_{i+1})/2$ at the combiner. As with EGC, the combiner adds these quantized values from all nodes and compares to a threshold. Due to the combinations of the four levels with different sensors the output of the combiner has a distribution having a larger number of possible outcomes, which can improve performance significantly compared to one bit fusion.

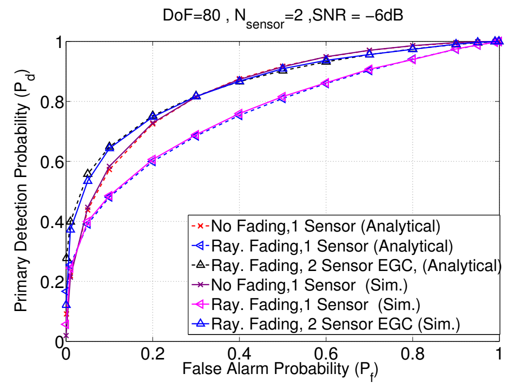


Figure 2: ROC curves comparing the analytical and simulated results in different scenarios

4. RESULTS

The analytical performance of energy detection is compared to Monte Carlo simulations for single sensor and 2-sensor parameter combining (EGC) for static and Rayleigh fading channels. The results are shown in Figure 2 in the form ROC curves. It is seen that the analytical model and simulations are in excellent agreement. It is interesting that when P_f tends to zero, the P_d for Rayleigh fading is better than in the case without fading. The reason is that for a given SNR, no signal is received above threshold in the static environment, but due to occasional constructive Rayleigh fading, there are still some events where the received signal is above the threshold.

In the following results, the performance of the various sensing methods is obtained both from closed-form expressions (when available) as well as Monte-Carlo simulations. Since the results are virtually identical in all cases, only the analytical results are plotted, except for combined fading, where no analytical expression is available. Since it is speculated that the relative performance of the methods will change as the distribution changes, various distributions of the signal energy are considered, ranging from 20 (low) to 160 (sufficiently high) degrees of freedom, having chi-square and approximately Gaussian distributions, respectively. Note that the degree of freedom depends directly on the sensing time available for a given application.

see the program

In all of the environments, the performance of EGC is found to be the highest. For a 2-sensor system in a static environment, OR and AND intersect each other with OR performing better for high $P_{f,T}$ values and AND performing better for lower $P_{f,T}$ (Figure 3). For a higher number of sensors, the performance of OR and AND show the same behavior (Figures 4 and 5), which becomes more pronounced for higher degrees of freedom (Gaussian approximation) as seen in Figure 6. Among the fusion techniques with more than 2 sensors, the performance of majority rule is the best for both the chi-square and Gaussian distributed cases (Figures 4, 5 and 6). For the similar sensors, the optimal (Chair) rule proposed in [9] ensures performance improvement compared to a single sensor giving the best performance in particular value of $P_{f,T}$, but it does not give the best performance for all global false alarm probability ($P_{f,T}$) values.

An interesting outcome of this study is that the detection performances are quite different for Rayleigh fading compared to the static case, which results from the difference in the distributions of the energy. In Rayleigh fading, OR fusion gives the best performance for almost all $P_{f,T}$ values, whereas AND fusion performs the worst (Figures 7, 8, and 9). Moreover, ROCs of AND and OR never intersect in Rayleigh fading, OR being far better than AND for all $P_{f,T}$ values. Also, in Rayleigh fading the performance of OR fusion is the best of all fusion rules and this approaches that of EGC when sensors are decreased to two. At the same time the AND fusion loses its performance and does not help at all when sensors are reduced to two.

The poor performance of AND fusion is higher when there is shadowing in addition to Rayleigh fading. This is apparent from Figure 10 where the collaboration with AND fusion degrades the performance rather than improving.

These results suggest that it is inappropriate to choose any fusion method without knowledge of the distribution of the decision parameter (energy in our case) in the collaborating sensors and the desired global false alarm probability $P_{f,T}$.

The performance improvement to fusion by implementing EGC with 2-bit quantization is shown in Figures 11 and 12, where $U'_{\max} = 45$ was chosen such that $Pr(U' > U'_{\max}) = 0.001$. This simple choice of the quantization regions gives performance much closer to optimal EGC than the best performing fusion rule. This result is encouraging given the simplicity of the method, and further gains are likely to be possible by optimizing the quantization regions. Note that for three sensors, the same false alarm probability was obtained for more than one threshold, and the threshold giving larger detection probability must be chosen.

5. CONCLUSIONS

This work presented different collaboration techniques for signal sensing in fading environments. Comparisons were

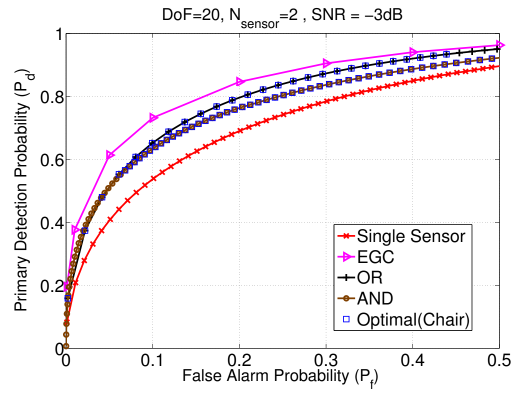


Figure 3: Performance of EGC and fusion techniques for low degree of freedom with 2-sensors

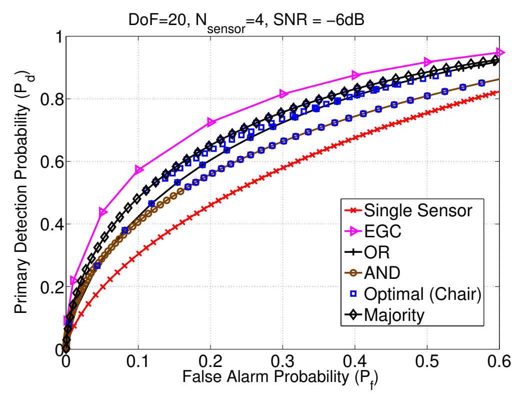


Figure 4: Performance of EGC and fusion techniques for low degree of freedom with 4-sensors

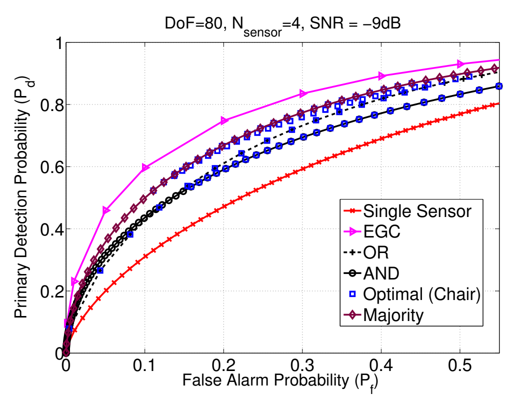


Figure 5: Performance of EGC and fusion techniques for higher degree of freedom

see the program

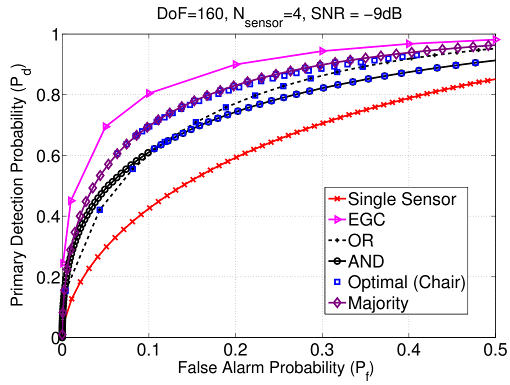


Figure 6: Performance of EGC and fusion techniques for sufficiently high degree of freedom (Gaussian distribution for energy)

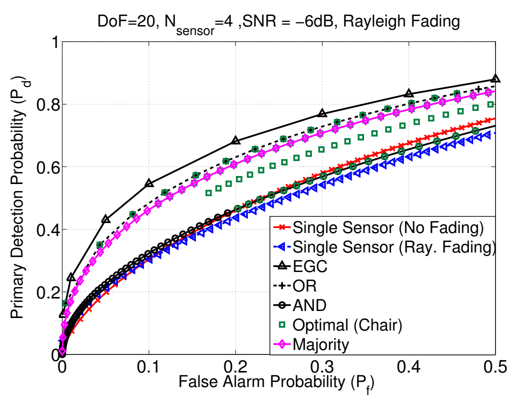


Figure 7: Performance of EGC and fusion techniques in a Rayleigh fading environment with 4-sensors and low degree of freedom

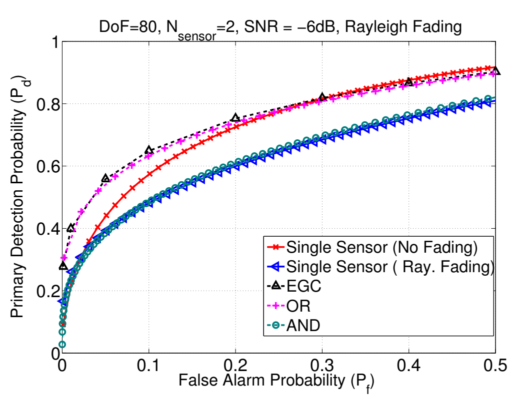


Figure 8: Performance of EGC and fusion techniques in a Rayleigh fading environment with 2-sensors and higher degree of freedom

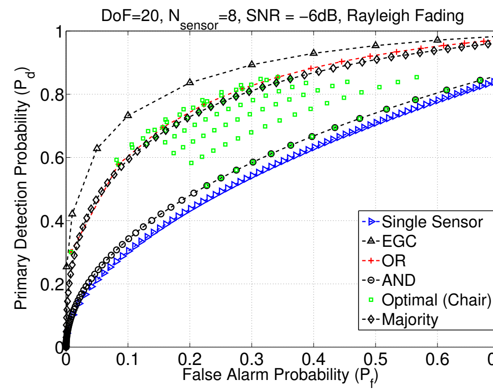


Figure 9: Performance of EGC and fusion techniques in Rayleigh fading environment with 8-sensors

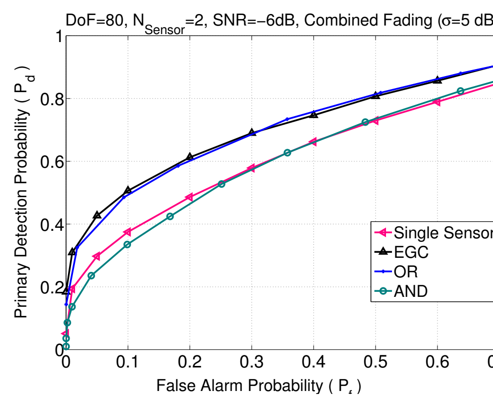


Figure 10: Performance of EGC and fusion techniques in a combined fading environment ($\sigma = 5$ dB) with 2-sensors

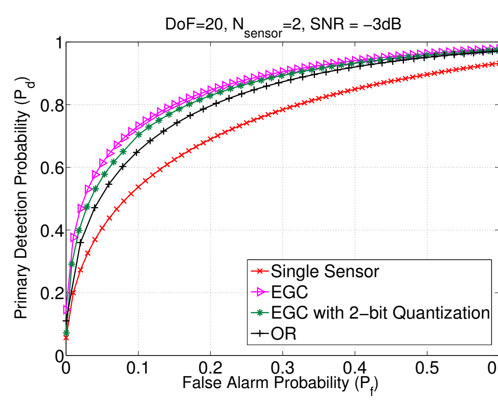


Figure 11: Comparison of the performance of 2-bit quantized EGC with other combining and fusion methods for 2-sensor collaboration

see the program

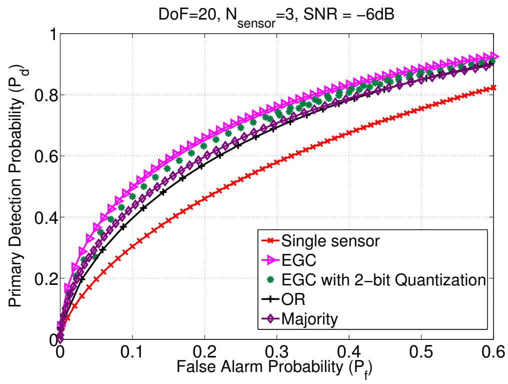


Figure 12: Comparison of the performance of 2-bit quantized EGC with other combining and fusion methods for 3-sensor collaboration

made of the various techniques for both analytical and simulation results. The results indicated that the superiority of a particular fusion method fully depends upon the distribution of the decision parameter at the sensors. Since this distribution in turn depends on the fading environment, a method that is superior for one environment may actually be inferior for a different environment. It appears that in general, AND fusion does not help in a Rayleigh fading environment. As the number of sensors is reduced to two, AND fusion does not give any performance gain, whereas OR fusion performs the best, approaching equal gain combining. Since OR fusion is simpler than EGC, it should be preferred for Rayleigh fading environments with two sensors. This result is even more pronounced for combined fading with increasing levels of lognormal shadowing. Finally, the fusion performance is significantly improved using a quantized EGC method.

In future work, we plan to investigate extending the fusion methods to find optimal performance with small increase in the data overhead. We also plan to study the performance of the these methods with more realistic path-based propagation channel models and actual measured channel data.

6. REFERENCES

[1] J. Mitola III and G. Q. Maguire Jr., “Cognitive radio: Making software radios more personal,” *IEEE Personal Communications Magazine*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] A. Ghasemi and E. S. Sousa, “Collaborative spectrum sensing for opportunistic access in fading environments,” in *Proc. of DySPAN’05*, Nov. 2005, pp. 131–136.

[3] F. F. Digham, M. Alouini, and M. K. Simon, “On the energy detection of unknown signals over fading channels,” in *Proc. 2003 IEEE Intl. Conf. Commun.*, May 2003, pp. 3575–3579.

[4] J. Ma and Y. G. Li, “Soft combination and detection for cooperative spectrum sensing in cognitive radio networks,” in *Proc. 2007 IEEE Global Telecomm. Conf.*, Washington, DC, Nov. 2007, pp. 3139–3143.

[5] S. Gezici, Z. Sahinoglu, and H. V. Poor, “On the optimality of equal gain combining for energy detection of unknown signals,” *IEEE Communications Letters*, vol. 10, no. 11, pp. 772–774, Nov. 2006.

[6] A. Fefjar, “Combining techniques to improve security in automated entry control,” in *Proc. Carnahan Conf. on Crime Countermeasures*, USA, May 1978, p. 89, abstract Only.

[7] S. D. Stearns, “Optimum detection using multiple sensors,” in *Proc. Carnahan Conf. on Security Technology*, Lexington, KY, USA, 1983.

[8] A. M. Aziz, M. Tummala, and R. Cristi, “Optimum detection using multiple sensors,” in *Proc. Thirty-First Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, CA, USA, Nov. 1997, pp. 941–945.

[9] Z. Chair and P. K. Varshney, “Optimal data fusion in multiple sensor detection systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 98–101, 1986.

[10] S. C. A. Thomopoulos, R. Vishwanathan, and D. C. Bougoulias, “Optimal decision fusion in multiple sensor systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, no. 5, pp. 644–653, Sept. 1987.

[11] H. Urkowitz, “Energy detection of unknown deterministic signals,” *Proc. IEEE*, vol. 55, no. 4, pp. 523–531, Apr. 1967.

[12] V. I. Kostylev, “Energy detection of a signal with random amplitude,” *IEEE Int. Conf. on Communications*, vol. 3, pp. 1606– 1610, Apr. 2002.

[13] R. Prasad and A. Kegel, “Effect of Rician faded and log-normal shadowed signals on spectrum efficiency in microcellular radio,” *IEEE Trans. Veh. Technol.*, vol. 42, no. 3, pp. 274–281, Aug. 1993.

[14] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[15] A. H. Nuttall, “Some integrals involving the Q_M function,” *IEEE Trans. Inf. Theory*, pp. 95–96, Jan. 1975.

[16] M. Petkovsek, H. Wilf, and D. Zeilberger, *A=B*. 888 Worcester St., Suite 230 Wellesley, MA 02482-3717: A K Peters. Ltd., April 1997, ch. 4, http://www.math.upenn.edu/%7Ewilf/AeqB.pdf.

see the program

DEVICE COLLABORATION IN AD-HOC MIMO NETWORKS

Ari Hottinen[1], *Tiina Heikkinen*[2], *Emanuele Viterbo*[3]

[1] Nokia Research Center, P.O.Box 407, FI-00045 Nokia Group, Finland. ari.hottinen@nokia.com
[2] Dept. Computer Science, University of Helsinki, Helsinki, Finland. tiina.heikkinen@cs.helsinki.fi
[3] DEIS – Universita della Calabria, via P. Bucci, Cubo 42C 87036 Rende (CS), Italy. viterbo@deis.unical.it

ABSTRACT

In this paper we consider methods for determining device coalitions for collaborative signal transmission, where different devices act as relay nodes to peers. The problem is to determine for R total number of users and R transmission slots the subsets of at most two devices that are allowed to transmit simultaneously. The subset selection problem is shown to be equivalent to an assignment problem. We consider both optimal assignment and greedy assignment and demonstrate the performance benefit due to device cooperation with simulations in a network model that models path loss between devices.

1. INTRODUCTION

In future networks different devices could potentially help each other in signal transmission, using each others hardware in an opportunistic way. Amplify-forward (AF) relaying is a potential candidate for such systems, since with AF, the relaying node need not know all transport parameters of the source node (as it does not decode the signal). On the other hand, AF relays are known to enhance also noise. Therefore, a randomly selected AF device can amplify noise to the extent that it has detrimental effect on network capacity.

In a practical network there are typically multiple AF-relaying devices and a limited number of orthogonal subchannels (time-frequency slots). The device population needs to be divided into subsets of active devices for each transmission subchannel. In addition, the roles (if a device acts as source or as a relay) for each device in each subset and channel use need to be determined.

Related subset selection and scheduling problems have appeared in uplink MU-MIMO [7], relay scheduling [2, 3], and in sensor networks [5]. Here, the subset selection problem considered from a MIMO relay network viewpoint, where a source and a co-channel relay jointly form a MIMO channel to a common destination node.

In the current application, we allow at most two devices to collaborate in a given channel use. We use sum-throughput (mutual information) of a MIMO relay channel as a performance measure when determining cooperative user coalitions. Unpaired devices are also allowed, if deemed beneficial. Unpaired devices transmit directly to the destination node (no relaying). A paired device transmits a part of its signal to a peer device during one channel use. In the next channel use, the paired devices transmit simultaneously to the destination node.

The work is supported by the STREP project No. IST-026905 (MASCOT) within the sixth framework programme of the European Commission.

2. SYSTEM MODEL

2.1. Relay model

We have a population of R devices each with one transmit antenna. Signal transmission is divided into two hops. In the first hop a source is allowed to communicate with the selected $K < R$ peers. In the second transmission hop the source and the selected peers transmit simultaneously to the destination node, which is assumed to have $N_{\mathrm{r}} \geq K$ receive antennas. The second hop channel is a Multiple Input Multiple Output (MIMO) channel. Formally, the signal model follows that of a MIMO relay network.

During the first hop, the source device transmits signal vector $\mathbf{x}$ with power P_1 through a $K \times K$ first hop channel $\mathbf{F}$, where K designates the number of active devices in the second hop channel. The off-diagonal terms of $\mathbf{F}$ (i.e. $|f_{k,n}|^2, n \neq k$) designate interference power due to source n at relay k input. Obviously, interference power vanishes for all relays if matrix $\mathbf{F}$ is diagonal. In this case, each device receives and retransmits a fraction $1/K$ of signal vector $\mathbf{x}$ during the second hop.

The $N_{\mathrm{r}} \times K$ second hop MIMO channel from the (selected) K devices to the destination is given by $\mathbf{H}$. During the second hop, each of the K devices multiply the signal with a relay-specific weighting coefficient w_k to satisfy a transmit power constraint at relay. We let

$$w_k = \sqrt{\frac{P_2/K}{\sum_{n=1}^{K} |f_{k,n}|^2 + \sigma_k^2}} \tag{1}$$

where σ_k^2 designates noise power at kth relay and P_2 is the desired sum transmit power over all K relay nodes. Note that

see the program

if interference terms and noise power vanish, the relay only modifies the transmit power of the original signal. In the second hop channel, for notational simplicity, the original source device is modelled a special AF relay with zero noise and interference power at relay input.

We collect the relay weights into a diagonal matrix

$$\Lambda = \mathrm{diag}(w_1, ..., w_K).$$

The destination receives

$$\mathbf{y} = \mathbf{H}\Lambda\mathbf{F}\mathbf{x} + \mathbf{H}\Lambda\mathbf{n}_r + \mathbf{n}_d$$

where the elements of complex Gaussian vector

$$\mathbf{n}_r = (n_1, ..., n_K)^T$$

designate noise with variance σ_k^2 at k'th relay node, and elements of

$$\mathbf{n}_d = (n_1, ..., n_{N_r})^T$$

designate complex Gaussian noise in each destination antenna. We assume that noise power is identical in each receiver antenna, i.e each has variance σ_d^2. The mutual information with i.i.d. Gaussian sources (in terms of bits-per-channel-use (bpcu)) for the considered signal model is [6]

$$\alpha = \frac{1}{2}\log_2\det(\mathbf{I} + \mathbf{H}\Lambda\mathbf{F}\mathbf{F}^\dagger\Lambda^\dagger\mathbf{H}^\dagger\mathbf{C}_{nn}^{-1}), \tag{2}$$

where the noise correlation matrix is

$$\mathbf{C}_{nn} = (\sigma_d^2\mathbf{I} + \mathbf{H}\Lambda\mathrm{diag}(\sigma_1^2, ..., \sigma_K^2)\Lambda^\dagger\mathbf{H}^\dagger).$$

Factor $1/2$ in model (2) is due to two-hop relaying.

2.2. Subset selection

We consider a special case of the subset selection problem to reduce computational burden of the optimization algorithm. Instead of allowing arbitrary-sized subsets, we determine identities of only $K \le 2$ second-hop devices for each channel use. We assume that each of the R devices is a source in exactly one of R channel uses. Assuming that $N_\mathrm{r} = 2$, each second hop MIMO channel supports $K \le 2$ simultaneously transmitting devices. Moreover, each of the devices acts as a relay exactly once in the R channel uses, to incorporate a notion of fairness to relay selection. That is, we determine for R sources and R transmission slots the distinct ordered subsets of at most two devices. We first describe the optimal (in system throughput sense) algorithm used for subset selection and then summarize the reference cases, greedy subset selection and random subset selection.

Since, $K \le 2$, we need to compute the mutual information α_{r_1,r_2} when device $r_1 \in \{1, ..., R\}$ is the source device and device $r_2 \in \{1, ..., R\}$ is the relay device. In general, $\alpha_{r_1,r_2} \neq \alpha_{r_2,r_1}$ since $\mathbf{F}$,$\mathbf{H}$ and Λ matrices also depend on these indices (omitted to simplify notation). When $r_1 = r_2$ ($K = 1$), the source transmits directly to destination with double power.

Optimal selection: Consider the selection of devices over R channel uses (via the following linear programming problem ([4]):

$$\arg\max_{(z_{r_1,r_2})} \sum_{r_2}^{R}\sum_{r_1}^{R} \alpha_{r_1,r_2} z_{r_1,r_2} \tag{3}$$

subject to

$$\sum_{r_1=1}^{R} z_{r_1,r_2} = 1, \forall r_2 \tag{4}$$

$$\sum_{r_2} z_{r_1,r_2} = 1, \forall r_1, \tag{5}$$

$$z_{r_1,r_2} \ge 0, \forall r_1, r_2, \tag{6}$$

The variables z_{r_1,r_2}, solved from above problem, dictate which devices become active source and relay nodes in each of the R slots. The model implicitly assumes all assignments involve either direct transmission or device pairing. When considering matrix (z_{r_1,r_2}), the solution to problem (3)-(6) dictates that there is exactly one non-zero element in each row and column, thus ensuring that all nodes act as sources equal number of times. When two nodes are active, either node may take the role of a source, while the other functions as a relay node. Whenever the $z_{r_1,r_2} = 1$, and $r_1 < r_2$, r_1 acts as source and r_2 relays. This convention results from the way the indices in eq. 3 are enumerated. When $z_{r_1,r_2} = 1$, with $r_1 = r_2$, only the direct link is activated, and relaying is disabled. Recall that problem (3)-(6) and the resulting permutation matrix can be solved efficiently (with polynomial complexity) applying transportation algorithm [4].

Naturally, considerably simpler subset selection algorithms exist:

Random selection: In random subset selection, the matrix (z_{r_1,r_2}) is defined as a random permutation matrix.

Greedy selection: In the first iteration of a Greedy subset selection the column and row indices of the largest element of (α_{r_1,r_2}) determine an element of the solution matrix. Then, the elements of these rows and columns are set to zero and maximum indices are sought in the following iteration from the modified matrix. This guarantees that the indices are unique for each iteration and that after R iterations a permutation matrix emerges.

3. NUMERICAL RESULTS

We study the arising collaboration patterns in a simple two-dimensional network. The R devices are placed randomly (uniformly) on a 20×20 rectangular area (meter units) with lower-left corner at coordinate $(0, 25)$. The destination position is $(30, 50)$. We assume $K = 2, N_\mathrm{r} = 2$, so that only

see the program

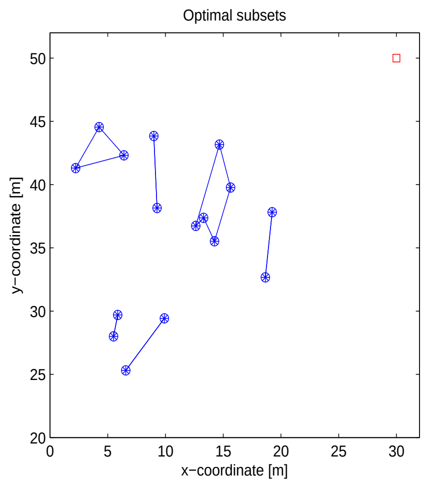


Fig. 1. Example of device collaboration patterns for *optimal subset selection* with $R = 16$. Destination is located on top-right corner, marked with character '□'. Collaboration patterns include 4 cycles of length 2, 1 cycle of length 3, and one cycle of length 5.

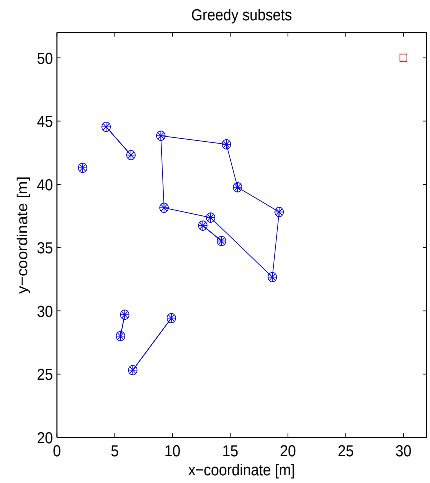


Fig. 2. Example of device collaboration patterns for *greedy subset selection* with $R = 16$. Destination is located on top-right corner, marked with character '□'. Collaboration patterns include 1 cycle of length 1 (no relaying), 4 cycles of length 2, and 1 cycle of length 7.

device pairing or direct transmission is allowed. The 2×2 network matrices $\mathbf{F}$ and $\mathbf{H}$ are computed using a simple path-loss model as follows: the distance between nodes r_1 and r_2 is d_{r_1,r_2} meters and the first-hop link matrix is set to

$$\mathbf{F} = \mathrm{diag}(1, \sqrt{P_1}/d_{r_1,r_2}^{2.3/2})$$

when devices r_1 and r_2, $r_1 \neq r_2$ are paired. The transmit power $P_1 = 27$ dB. For direct transmission ($r_1 = r_2$) the path-loss model is obviously neither applicable or relevant due to the weighting method given in eq. 1. Thus, to model direct transmission in the relay framework, we set $\mathbf{F} = \mathrm{diag}(1, 1)$.

Due to applied weighting, the total second-hop transmit power is identical for direct and paired transmission. The second-hop matrix is of form

$$\mathbf{H} = \mathrm{diag}(\sqrt{P_2}/d_{r_1,d}^{2.3/2}, \sqrt{P_2}/d_{r_2,d}^{2.3/2})\tilde{\mathbf{H}},$$

where $d_{r_1,d}$ and $d_{r_2,d}$ is the distance device r_1 and r_2 and the destination node, respectively, and P_2 is the transmit power on second hop. We set $P_2 = 31.7$ dB. Matrix $\tilde{\mathbf{H}}$ is an i.i.d. complex Gaussian-distributed MIMO matrix, where each element has unit power.

3.1. Collaboration patterns

The optimization schemes in previous section each determine a permutation of matrix of dimension R. The non-zero value on the rth row of the permutation matrix is mapped to element $\sigma(r)$, i.e. $z_{r,\sigma(r)} = 1$ in terms of notation in section 2.2. We say that devices $(r, \sigma(r))$ form a collaboration pair. The permutation matrices arising from optimal, greedy or random subset selection can each be mapped to a permutation σ of R elements of the form

$$\sigma : \begin{pmatrix} 1 & 2 & \cdots & R \\ \sigma(1) & \sigma(2) & \cdots & \sigma(R) \end{pmatrix}. \qquad (7)$$

If $r = \sigma(r)$, device r is unpaired. The unpaired devices correspond to the fixed elements of the permutation. In our relay model, this corresponds to the case, where a device transmits directly to the destination node.

If two devices, say r_1 and r_2, use each other as their respective relays, these devices form a pair (r_1, r_2). If in addition, r_2 uses r_1 as a relay, the corresponding permutation includes columns (7) $(r_1, r_2 = \sigma(r_1))^T$ and $(r_2, r_1 = \sigma(r_2))^T$.

In terms of [1, 7], unpaired devices correspond to cycles of length 1, while paired users that use each other as relays correspond to cycles of length 2. Naturally, an arbitrary permutation σ, e.g.

$$\sigma : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 1 & 4 \end{pmatrix}$$

can have longer cycles. Above we have a cycle $(1, 5, 4)$ of length 3. In the relay model, device 1 uses devices 5 as relay in the first channel use, device 2 is unpaired in the second channel use, and so on. The collaboration pattern is

see the program

thus $\{(1,5)(2,2)(3,3)(4,1)(5,4)\}$ in 5 channel uses, and it includes two unpaired users.

We first illustrate the emerging cooperation patterns using one realization of device locations and channels. In Fig. 1 the optimal device collaboration patterns for each transmission slot are computed by solving problem (3)-(6). In Fig. 2 the same is done for greedy heuristics. In both figures, the destination receiver is located on top-right corner with character '□'. In the two figures a line is drawn between two devices cooperative devices. For cycles of length 2, the two devices act as source and relay nodes for each other in alternate channel uses. For cycles with length 3 or higher, a device acts as source and relay for two different devices in separate channel uses. For example, in Fig. 1 a cycle of length 3 appears in top-left corner. It takes three channel uses to serve all three devices.

3.2. Performance

Fig. 3 depicts the ergodic performance (average mutual information) for four different subset selection schemes (optimal, greedy, random, direct/no pairing) with $R \in \{2, 4, 8, 12, 16\}$ single-antenna devices and one dual-antenna destination node. The results are averaged for each R over 1000 device locations each with independently generated MIMO channel. For optimal subset selection, the device collaboration patterns for each are computed from problem (3)-(6) and related mutual information is recorded. The mutual information arising from optimal subsets are shown in figures with legend 'Optimal'. For comparison, we also depict the performance with random subsets - these results are associated with legend 'Random'. The following observations are in order:

- Channel-aware subset selection provides a substantial capacity gain over both direct transmission and random device pairing, thanks to its ability to select network-optimal MIMO relays for the second-hop channel.
- The gain due to optimal subset increases with increasing number of devices. This is in part due to the fact that network is denser and cooperation occurs with devices that are closer.

4. CONCLUSIONS

We have considered device cooperation as means to form relay-based MIMO uplink. In the considered scheme optimal device collaboration patters (device subsets) are computed (up to pairs) using optimal and greedy matching algorithms. The subset selection algorithms determine which of the R devices should be paired and which should transmit directly to the destination in R channel uses. We demonstrated the performance gain (in terms sum mutual information) with simulations. It is observed that the device subsets have cyclic structure. If the cycle length is three, three devices need to form a

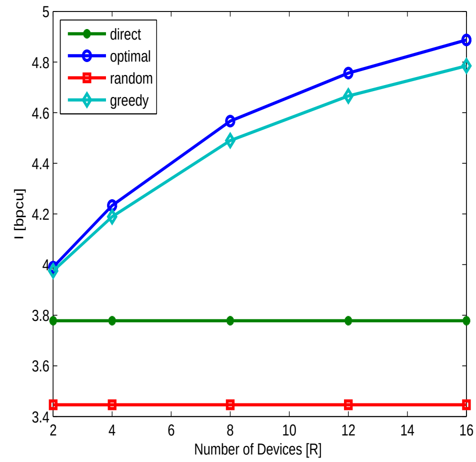


Fig. 3. Average mutual information (I) at destination node with different number of devices (R) and different pairing schemes. Direct transmission is depicted as reference.

coalition when forming source-relay pairs. A topic for future work is to consider subset selection from the point of view of cooperative game theory.

5. REFERENCES

[1] J.D. Dixon and B. Mortimer: *Permutation Groups*, Number 163 in Graduate Texts in Mathematics. Springer-Verlag, 1996.

[2] A. Hottinen, T. Heikkinen, “Delay-differentiated scheduling in a randomized relay network,” *Proc. EUSIPCO 2007*, Pozdam, Poland, Sept. 2007

[3] T. Heikkinen and A. Hottinen, “Delay-differentiated scheduling in a fading channel,” *IEEE Tr. Wireless Communications,* Vol. 7, No. 3, pp. 848–856, March 2008

[4] H.W. Kuhn, “The Hungarian method for the assignment problem,” *Naval Research Logistic Quarterly*, 2:83-97, 1955.

[5] A. Roumy and D. Gesbert, “Optimal matching in wireless sensor networks,” *IEEE Journal on Selected Topics in Signal Processing*, Vol. 1, Issue 4, pp. 725–735, Dec. 2007

[6] B. Rankov and A. Wittneben, “Impact of cooperative relays on the capacity of rank-deficient MIMO channels,” *Proc. 12th IST Summit on Mobile and Wireless Communications*, Aveiro, Portugal, pp. 421-425, June 2003.

[7] E. Viterbo and A. Hottinen, “Optimal user pairing for multiuser MIMO,” *Proc. IEEE ISSSTA 2008*, Bologna, Italy, August 2008

see the program

Advanced Physical Layer Techniques for Wireless Mesh Networks with Network Coding

Andreas Dotzler, Johannes Brehmer, and Wolfgang Utschick
Associate Institute for Signal Processing, Technische Universität München, 80290 München
Email: {dotzler,brehmer,utschick}@tum.de, Phone: +49-89-289-28510

***Abstract*—We consider wireless mesh networks where information is multicasted to multiple terminals in a multi-hop fashion. Due to their strong interdependence, we seek a joint optimization of network and physical layer that are coupled by the per link flow constraint. A common approach is to dualize this constraint and decompose the dual problem into a layered structure; routing at the network layer and rate assignment at the physical layer. For the network layer subproblem, linear network coding is an optimal routing strategy and the solution can be computed by solving a linear or convex program. The physical layer subproblem turns out to be more challenging, due to the nature of the wireless medium and the resulting diminishing effect of multiple access interference. Existing approaches try to avoid interference by full orthogonalization of the channels or building on the concept of conflict graphs. Contrary to these approaches, we are taking into account interference management, for example by exploiting the advanced abilities of multiple antenna systems. Our approach is the factorization of the achievable edge rate region into known rate regions of subgraphs, called Elementary Capacity Graphs (ECGs), which allows for taking into account the half duplex constraint implicitly. The parametrization of the achievable rate region of an ECG depends on the transmission technique used and is in general nonconvex. We demonstrate how the nonconvexity of the physical layer parametrization can be handled within a primal-dual framework without loss of optimality. As our solution is optimal for a given factorization we show by numerically simulations the advances compared to non-optimal schemes.**

I. Introduction and Problem Statement

Communication over a wireless mesh network needs transmission strategies to provide link rates at the physical layer, and a scheme for routing traffic at the network layer. While originally being developed for wired networks with fixed link capacities, network coding, as one possible routing scheme, has recently attracted a lot of attention for being used in wireless networks [1], [2]. At the physical layer, we employ advanced physical layer techniques and utilize the gained flexibility and increased link capacities. This potential gain can only be exploited if network and physical layer are optimized jointly, commonly done via a dual approach, see [1], [3]. Toumpis and Goldsmith [4] give a very general physical layer characterization by scheduling link configurations with fixed link capcities called basic rate matrices. A similar concept is used by Wu et al. [2], who coined the term Elementary Capacity Graphs (ECGs). We adopt this term for our work and extend it to ECGs with variable rates, where each ECG is fully described by its achievable rate region. Having fixed link rates renders the network optimization problem into a linear program, while taking into account variable rates is more challenging and, to the best of our knowledge, only suboptimal solutions are available. Xiao et al. [3] assume that the link rate is only a function of local resources, which implies that links have to be orthogonalized. Cruz and Santhanam [5] assume a linear dependence of the link rate on the SINR which is only true for small SINR values. Whereas Yuan et al. [1] use a convex approximation of how the link rate depends on the SINR. Multiple antenna systems are considered by Liu et al. [6] who construct the network by MIMO-BC systems from each node to its neighbors. The BC systems are orthogonalized by fixed frequency assignment, which is in general suboptimal. In our work we present a major algorithmic framework without loss of optimality, for a given factorization.

We consider a mesh network with graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ is the set of nodes and $\mathcal{L}$, $L = |\mathcal{L}|$, is the set of all wireless links in the network. A multicast session is described by its source $s \in \mathcal{N}$ and the set of terminals $\{t_1, \ldots, t_K\} \subset \mathcal{N}$. The decision of a routing scheme at the network layer determines the throughput $r \in \mathbb{R}_+$, and the actual traffic flows on the links $\boldsymbol{f} \in \mathbb{R}_+^L$ that are necessary for obtaining it. Choosing an operating point of the network layer is to select a valid pair of session throughput and traffic assignment $(r, \boldsymbol{f}) \in \mathcal{F}$, where all possible routing decisions are characterized by the routing region $\mathcal{F} \subset \mathbb{R}_+ \times \mathbb{R}_+^L$. For network coding the routing region has a explicit formulation in linear (in)equalities, and therefore $\mathcal{F}$ forms a polyhedron. An extention to multiple multicast sessions that are coded separately is straightforward, see [1].

At the physical layer link rates are assigned to the links in the network by resource allocation, where due to interference and jointly used resources link rates are traded off against each other, described by an achievable edge rate region $\mathcal{R} \subset \mathbb{R}_+^L$. Clearly, the traffic rates established by the network layer are limited to the link rates $\boldsymbol{c} \in \mathcal{R}$ that the physical layer provides, which results in the per link flow constraint $\boldsymbol{f} \leq \boldsymbol{c}$.

In this work we are aiming at the maximization of throughput and the optimization problem can be formulated as

$$\begin{aligned} \max_{r,\boldsymbol{f},\boldsymbol{c}} \quad & r \\ \text{subject to} \quad & (r, \boldsymbol{f}) \in \mathcal{F} \\ & \boldsymbol{f} \leq \boldsymbol{c} \\ & \boldsymbol{c} \in \mathcal{R}. \end{aligned} \tag{1}$$

A. The Network Layer Characterization

This section is concerned about the constraint $(r, \boldsymbol{f}) \in \mathcal{F}$ of the problem statement (1). A fundamental result of network information theory is that information flows from one source to different terminals do not compete for link capacities, and the maximal throughput is given by the max-flow min-cut theorem and can be achieved by network coding [7]. In other words, a throughput r is achievable if it is achievable for each of the terminals individually. Li et al. [8] prove that optimal throughput can be achieved by linear codes, subsequently Ho et al. [9] show that random linear codes are sufficient. This allows us to exclude code construction in this work, and the optimal routing can be found by a flow allocation problem. For modelling we use additional variables per terminal and link $\boldsymbol{e}^{t_1}, \ldots, \boldsymbol{e}^{t_K} \in \mathbb{R}_+^L$, the so called conceptional flows. The actual traffic flow caused on a link $\ell \in \mathcal{L}$ is the maximum of conceptional flows on the link:

$$\boldsymbol{e}^i \leq \boldsymbol{f}, \; \forall i \in \{t_1, \ldots, t_K\}. \tag{2}$$

A node cannot send more information than it received. Consequently, the "Kirchhoff law" for each node, where $\mathcal{I}(n)$ is the set of incoming links of node n and $\mathcal{O}(n)$ the set of outgoing links, reads

$$\sum_{\ell \in \mathcal{O}(n)} e_\ell^i = \sum_{\ell' \in \mathcal{I}(n)} e_{\ell'}^i, \quad \forall n \in \mathcal{N} \setminus \{s, i\}, \; i \in \{t_1, \ldots, t_K\}.$$

The throughput for a sink is obviously determined by the sum of incoming information flows. By introducing rate incidences $\boldsymbol{a}^i$ $i = t_1, \ldots, t_K$ that represent links from the terminals to the sink, we can conveniently express the flow constraints via the incidence matrix $\boldsymbol{A}$ of the network:

$$\begin{bmatrix} \boldsymbol{a}^{t_1} & \boldsymbol{A} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{a}^{t_2} & \boldsymbol{0} & \boldsymbol{A} & \ldots & \boldsymbol{0} \\ \vdots & & & \ddots & \\ \boldsymbol{a}^{t_K} & \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{A} \end{bmatrix} \begin{bmatrix} r \\ \boldsymbol{e}^{t_1} \\ \boldsymbol{e}^{t_2} \\ \vdots \\ \boldsymbol{e}^{t_K} \end{bmatrix} = \boldsymbol{0}. \tag{3}$$

As it is fully characterized by the linear (in)equalities (2) and (3), $\mathcal{F}$ forms a polyhedron.

B. Factorization of the Rate Region into ECGs

As there exists a huge manifold of transmission techniques with complex parametrization, we factorize the achievable edge rate region into known rate regions of subgraphs, called Elementary Capacity Graphs (ECGs). As we will see later this factorization allows to reduce the algorithmic complexity and when established via timesharing, we can take into account the half duplex constraint, which prohibits that a node receives and transmits simultaneously. Formally, an ECG is denoted as $\mathcal{B}_i \subseteq \mathcal{L}$, and the set of ECGs we decide for is given by $\mathcal{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_B\}$. A transmission schedule alternates between ECGs by assigning the fraction of time t_i, which the ECG $\mathcal{B}_i$ is active. The vector $\boldsymbol{t} = (t_1, \ldots, t_B)^\top \in \mathcal{T}$, where $\mathcal{T} = \{\boldsymbol{t} \geq \boldsymbol{0} : \|\boldsymbol{t}\|_1 = 1\}$, formally describes the transmission schedule. The individual rate regions of the ECGs are given by $\mathcal{R}_1, \ldots, \mathcal{R}_B$ and a parameter set $\boldsymbol{x}_i \in \mathcal{X}_i$ determines a rate point $\boldsymbol{R}_i(\boldsymbol{x}_i)$ from a rate region $\mathcal{R}_i$, i.e. $\mathcal{R}_i = \{\boldsymbol{R}_i(\boldsymbol{x}_i) : \boldsymbol{x}_i \in \mathcal{X}_i\}$. Thus the factorized overall edge rate region $\mathcal{R}$ is obviously defined by the convex hull of all involved rate regions,

$$\begin{aligned} \mathcal{R} &= \{(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_B)\, \boldsymbol{t} : \boldsymbol{R}_1 \in \mathcal{R}_1, \ldots, \boldsymbol{R}_B \in \mathcal{R}_B, \boldsymbol{t} \in \mathcal{T}\} \\ &= \operatorname{co}(\mathcal{R}_1, \ldots, \mathcal{R}_B). \end{aligned} \tag{4}$$

An operating point of the physical layer $\boldsymbol{c}$ is determined by the scheduling vector $\boldsymbol{t}$ and the parameter vector $\boldsymbol{x} = (\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_n^\top)^\top \in \mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_B$, i. e.

$$\boldsymbol{c} = (\boldsymbol{R}_1(\boldsymbol{x}_1), \ldots, \boldsymbol{R}_B(\boldsymbol{x}_B))\, \boldsymbol{t}.$$

Dealing with Interference: Physical layer configurations are constructed by timesharing between ECGs that may contain multiple links, which are exposed to destructive interference. Interference is treated as additional noise and no attempt to decode it is made. We consider three basic ways of dealing with interference, which describe a strategy on how the parameters $\boldsymbol{x} = (\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_n^\top)^\top \in \mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_B$ are selected.

a) Avoid Interference: The only way to definitely avoid interference within one ECG is to have only a single active link, which is operated at its best transmit strategy.

b) Selfish Transmission: Interference causes a decrease of rate at the nonintended receivers. This strategy simply operates each link assuming there are no other links and takes the decrease in rates into account.

For interference avoidance and selfish transmission we end up with ECGs that correspond to exactly one achievable rate point $\boldsymbol{R}_i$, that might not be optimal. Considering only ECGs that are fixed to a single rate point leads to an physical layer configuration where the edge rate region $\mathcal{R} = \{(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_B)\, \boldsymbol{t} : \boldsymbol{t} \in \mathcal{T}\}$ forms a polytope, and the optimization problem (1) is a linear program. Note that for the special case of one antenna systems where $\boldsymbol{x}$ are the transmit power levels, these rate points correspond to the basic rate matrices in [4].

c) Interference Management: By adjusting the parameter vector $\boldsymbol{x}$ we can trade off the link capacities against each other, which requires cooperation of the transmitters. By opting for interference management we have to include the parameter vector $\boldsymbol{x}$ into the joint optimization of physical and network layer, so this approach demands for major algorithmic solutions.

II. Algorithmic Solutions

Opting for the simple schemes avoid interference and selfish transmission results in an edge rate region given by a polytope where the extreme points are given by the fixed rate vectors of the ECGs, which renders the optimization problem into a linear program. Operating each ECG in an interference management mode requires to handle complex parametrizations that are in general nonlinear and nonconvex. However, the resulting individual rate regions of ECGs can always be made convex by timesharing. Having a convex edge rate region $\mathcal{R}$ and the polyhedral routing region $\mathcal{F}$ the problem (1) is a convex problem and the solution may be found via a dual problem, as for this kind of convex optimization problem the duality gap is zero. A common approach is to dualize the

see the program

constraint $\boldsymbol{f} \leq \boldsymbol{c}$ and solve the dual problem by a primal-dual algorithm. The algorithm iteratively evaluates the dual function which is a function of the Lagrangian multipliers $\boldsymbol{\lambda}$. For the i-th iteration and the corresponding $\boldsymbol{\lambda}^{(i)}$ we have to solve

$$\begin{aligned} \max_{r,\boldsymbol{f},\boldsymbol{c}} \quad & r - \boldsymbol{\lambda}^{(i),\top}(\boldsymbol{f} - \boldsymbol{c}) \\ \text{subject to} \quad & (r, \boldsymbol{f}) \in \mathcal{F} \\ & \boldsymbol{c} \in \mathcal{R}. \end{aligned} \tag{5}$$

Decomposed into two subproblems, for routing at the network layer we obtain

$$\begin{aligned} \max_{r,\boldsymbol{f}} \quad & r - \boldsymbol{\lambda}^{(i),\top} \boldsymbol{f} \\ \text{subject to} \quad & (r, \boldsymbol{f}) \in \mathcal{F}, \end{aligned} \tag{6}$$

which is a linear program. Rate assignment at the physical layer is

$$\begin{aligned} \max_{\boldsymbol{c}} \quad & \boldsymbol{\lambda}^{(i),\top} \boldsymbol{c} \\ \text{subject to} \quad & \boldsymbol{c} \in \mathcal{R}. \end{aligned} \tag{7}$$

We present a Theorem that allows an elegant reformulation of the physical layer subproblem.

Theorem 1: The optimum solution of the i-th physical layer subproblem (7) is always met by exclusively activating a single ECG.

Proof: Plugging (4) into the physical layer subproblem (7) leads to

$$\max_{\boldsymbol{c}} \quad \boldsymbol{\lambda}^{(i),\top} \boldsymbol{c} \quad \text{subject to } \boldsymbol{c} \in \operatorname{co}(\mathcal{R}_1, \ldots, \mathcal{R}_B).$$

It is well known that optimizing a linear function over the convex hull of a set can as well be solved over the set itself. Therefore, we can write

$$\begin{aligned} \max_{\boldsymbol{c}} \quad & \boldsymbol{\lambda}^{(i),\top} \boldsymbol{c} \\ \text{subject to} \quad & \boldsymbol{c} \in \bigcup_{i=1,\ldots,B} \mathcal{R}_i. \end{aligned}$$

We now can search for the optimal weighted sum rate point in each of the rate regions $\mathcal{R}_1, \ldots, \mathcal{R}_B$ and select the best point (or one of the best points) as solution to the physical layer subproblem, which corresponds to exclusively activating a single ECG. ■

With this Theorem we can reformulate the physical layer subproblem as

$$\max_{n=1,\ldots,B} \max_{\boldsymbol{c} \in \mathcal{R}_n} \quad {\boldsymbol{\lambda}^{(i)}}^{\top} \boldsymbol{c}. \tag{8}$$

The new physical layer subproblem has some profound advantages:

- The factorization into smaller problems constituted by ECGs, of which tractable parametrizations and algorithms are available.
- The reformulation provides a clear interface for any type of ECG that has a parametrization of its rate region which allows for the optimization of the weighted sum rate cost function. Weighted sum rate maximization is a well-researched problem and efficient solutions exist for a wide range of physical layer setups.
- The optimum scheduling $\boldsymbol{t}^*$ is found by primal recovery, which avoids an explicit parametrization of the convex hull of the edge rate region of the overall network. The time sharing within the individual ECGs is found via primal recovery as well.

A. Primal-Dual Algorithms and Primal Recovery

A primal-dual algorithm iteratively evaluates the dual function which means to solve an optimization in the primal variables. These optimal primal values are used to update the dual variables, for example by making an adequate step into the direction of a subgradient. For our problem (1) and the chosen dual function (5) the subgradient update rule is given by

$$\boldsymbol{\lambda}^{(i)} = \left[\boldsymbol{\lambda}^{(i)} + v^{(i)} \left(\boldsymbol{f}^{*(i)} - \boldsymbol{c}^{*(i)}\right)\right]^+,$$

where $[\bullet]^+$ denotes $\max(0, \bullet)$, $v^{(i)}$ is determined by a stepsize rule and $\boldsymbol{f}^{*(i)}$ and $\boldsymbol{c}^{*(i)}$ are the solutions to (6) and (7). However, subgradient methods tend to be slow in practice and other update rules for the dual variables should be considered. For our numerical simulations we used a variant of the well known cutting-plane algorithm [10].

Primal-dual algorithms guarantee to find the optimal dual variables $\boldsymbol{\lambda}$, but the primal variables found by evaluating the dual function are in general not feasible to the primal problem. To be explicit, the activation of a single ECG is in general not a feasible physical layer configuration. Feasible primal solutions can be constructed by a convex combination of the solutions found in each iteration. For details on recovering the primal solution of convex optimization problems we refer to [11]. The optimal timesharing of physical layer configurations then equals the convex combining parameters, which are conveniently calculated as a byproduct by the cutting-plane algorithm.

III. An ECG with Two Interfering Links

Having the algorithmic framework at hand, this Section gives an example for advanced physical layer techniques, featuring multiple antenna systems. We consider two types of ECGs, the single link or peer-to-peer connection and the two-user *Interference Channel* (IFC). Using ECGs with a single link effectively represents interference avoidance, whereas in the case of two links we can employ selfish transmission and interference management which requires cooperation of the senders. Figure 1 shows some exemplary ECGs. Other configurations are considered in [12], where ECGs are constituted by *Multicast Channels*, *Broadcast Channels*, and *Multiple Access Channels*.

In this work, without loss of our general conclusion we limit our investigation to multiple-input single-output (MISO) transmission instead of utilizing the enhanced capabilities of the full multiple-input multiple-output (MIMO) channel properties. The MISO two-user interference channel is described by the four channel vectors $\boldsymbol{h}_{ij}$ $i, j = 1, 2$. The transmit symbols $\boldsymbol{x}_i \in \mathbb{C}^N$ for the senders $i = 1, 2$ are constructed by the scalar data symbol $s_i \in \mathbb{C}$ and the beamforming vector $\boldsymbol{u}_i \in \mathbb{C}^N$ such that $\boldsymbol{x}_i = \boldsymbol{u}_i s_i$. The data symbols

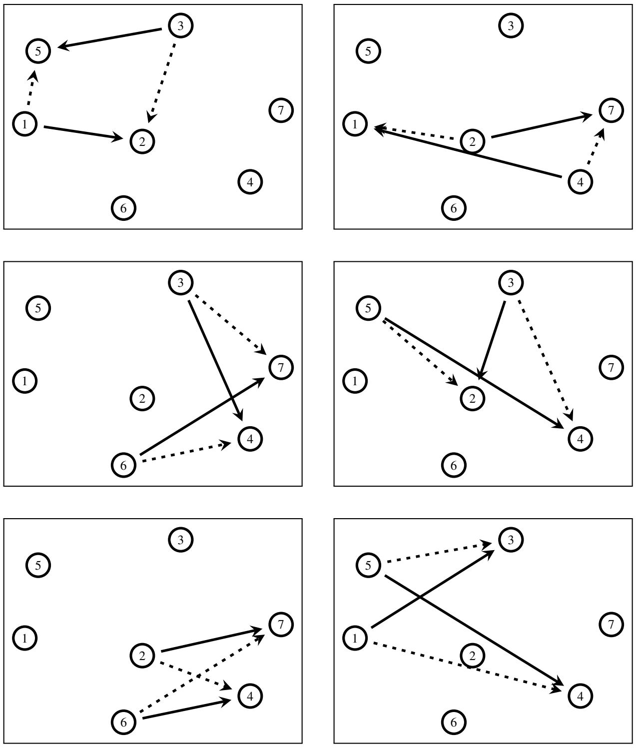


Fig. 1. Exemplary ECGs with the Two User Interference Channel

s_i are circularly symmetric Gaussian with unit variance. The transmitted symbols interfere additively and noise n_i is added at the receiver. The received symbols are

$$\begin{aligned} y_1 &= \boldsymbol{h}_{11}^\top \boldsymbol{x}_1 + \boldsymbol{h}_{12}^\top \boldsymbol{x}_2 + n_1 \\ y_2 &= \boldsymbol{h}_{22}^\top \boldsymbol{x}_2 + \boldsymbol{h}_{21}^\top \boldsymbol{x}_1 + n_2. \end{aligned}$$

Assuming that the receivers treat all interference as additional noise, the achievable rates are given by

$$R_1 = \log\left(1 + \frac{|\boldsymbol{h}_{11}^\top \boldsymbol{u}_1|^2}{\sigma^2 + |\boldsymbol{h}_{12}^\top \boldsymbol{u}_2|^2}\right) \quad (9)$$

$$R_2 = \log\left(1 + \frac{|\boldsymbol{h}_{22}^\top \boldsymbol{u}_2|^2}{\sigma^2 + |\boldsymbol{h}_{21}^\top \boldsymbol{u}_1|^2}\right). \quad (10)$$

The power of the noise $\sigma^2 = \mathrm{E}[|n_1|^2] = \mathrm{E}[|n_2|^2]$ is assumed to be the same at both receivers. The achievable rate region is the union of all beamforming vectors that fulfill a power constraint $||\boldsymbol{u}_1||_2^2, ||\boldsymbol{u}_2||_2^2 \le P_{\max}$ and can be written as

$$\mathcal{R} = \bigcup_{\substack{\boldsymbol{u}_1, \boldsymbol{u}_2 \\ ||\boldsymbol{u}_1||_2^2 \le P_{\max} \\ ||\boldsymbol{u}_2||_2^2 \le P_{\max}}} (R_1(\boldsymbol{u}_1, \boldsymbol{u}_2), R_2(\boldsymbol{u}_1, \boldsymbol{u}_2)).$$

For user i, using $\boldsymbol{u}_i^{\mathrm{MRT}} = \boldsymbol{h}_{ii}^*$ is maximum ratio transmission (MRT) beamforming. Altruistic or zero-forcing (ZF) beamforming, $\boldsymbol{u}_i^{\mathrm{ZF}} = \boldsymbol{h}_{ii}^* - \frac{\boldsymbol{h}_{ji}^\top \boldsymbol{h}_{ii}^*}{||\boldsymbol{h}_{ji}^*||_2^2} \boldsymbol{h}_{ji}^*$, causes no interference to the second user j while the own gain is reduced. Shi et al. [13] and Jorswieck et al. [14] show that the optimal beamforming vector can be written as a combination of $\boldsymbol{u}_i^{\mathrm{MRT}}$ and $\boldsymbol{u}_i^{\mathrm{ZF}}$:

$$\boldsymbol{u}_1(\gamma_1) = P^{\max} \cdot \frac{\gamma_1 \boldsymbol{u}_1^{\mathrm{MRT}} + (1-\gamma_1)\boldsymbol{u}_1^{\mathrm{ZF}}}{||\gamma_1 \boldsymbol{u}_1^{\mathrm{MRT}} + (1-\gamma_1)\boldsymbol{u}_1^{\mathrm{ZF}}||_2} \quad (11)$$

$$\boldsymbol{u}_2(\gamma_2) = P^{\max} \cdot \frac{\gamma_2 \boldsymbol{u}_2^{\mathrm{MRT}} + (1-\gamma_2)\boldsymbol{u}_2^{\mathrm{ZF}}}{||\gamma_2 \boldsymbol{u}_2^{\mathrm{MRT}} + (1-\gamma_2)\boldsymbol{u}_2^{\mathrm{ZF}}||_2}, \quad (12)$$

with $\gamma_1, \gamma_2 \in [0, 1]$. By plugging (11)–(12) into (9)–(10) the weighted sum rate problem for the weights λ_1, λ_2 can be formulated as

$$\begin{aligned} \max_{\gamma_1,\gamma_2} \quad & \lambda_1 \cdot \log\left(\sigma^2 + |\boldsymbol{h}_{12}^\top \boldsymbol{u}_2(\gamma_2)|^2 + |\boldsymbol{h}_{11}^\top \boldsymbol{u}_1(\gamma_1)|^2\right) + \\ & \lambda_2 \cdot \log\left(\sigma^2 + |\boldsymbol{h}_{21}^\top \boldsymbol{u}_1(\gamma_1)|^2 + |\boldsymbol{h}_{22}^\top \boldsymbol{u}_2(\gamma_2)|^2\right) - \\ & \lambda_1 \cdot \log\left(\sigma^2 + |\boldsymbol{h}_{12}^\top \boldsymbol{u}_2(\gamma_2)|^2\right) - \\ & \lambda_2 \cdot \log\left(\sigma^2 + |\boldsymbol{h}_{21}^\top \boldsymbol{u}_1(\gamma_1)|^2\right) \\ \text{subject to} \quad & \gamma_1, \gamma_2 \in [0, 1]. \end{aligned}$$

The objective can be split into two functions monotonic in γ_1, γ_2 and the problem fits into the framework of optimizing the difference of increasing functions. Having a similar structure, we adopted the approach suggested by Jorswieck and Larsson [15] based on the *Polyblock Algorithm*, which is a global optimization method proposed by Tuy [16]. Selfish transmission coresponds to choose $\gamma_1 = \gamma_2 = 1$ and for interference avoidance by single links the MRT beamformer is chosen.

IV. Results

By numerical simulations we compare the three strategies described in Section I-B: avoid interference, selfish transmission, and interference management. The simulations were made for a fully connected network of seven nodes, which exhibits 42 links. From those we can construct 420 ECGs with two interfering links each as described in Section III, an exemplary selection of these is illustrated in Figure 1. In general the number of ECGs of this type in a network with N nodes is given by:

$$\#\text{ECGs} = 12 \cdot \binom{N}{4}. \quad (13)$$

Each node is equipped with two antennas, and the channel coefficients are complex Gaussian distributed with unit variance. The results are averaged over 500 channel realizations per SNR value. We include simulation results for 4 and 6 terminals, see Figure 2 and Figure 3 respectively. Additionally the solutions of the network optimization problem for one channel realization at 10 dB is given, once for six terminals (flooding), Figure 4, and once for two terminals, Figure 5. For this example the established link capacities are equal to the traffic assigned to it, which is not necessarily always the case. The thickness of the arrows is proportional to the assigned rate.

Interference management by advanced physical layer techniques increases system complexity, so it is a fair question to ask if it is actually worth all the effort. The simulation results give a clear answer by showing a significant increase of the multicast throughput when utilizing the interference

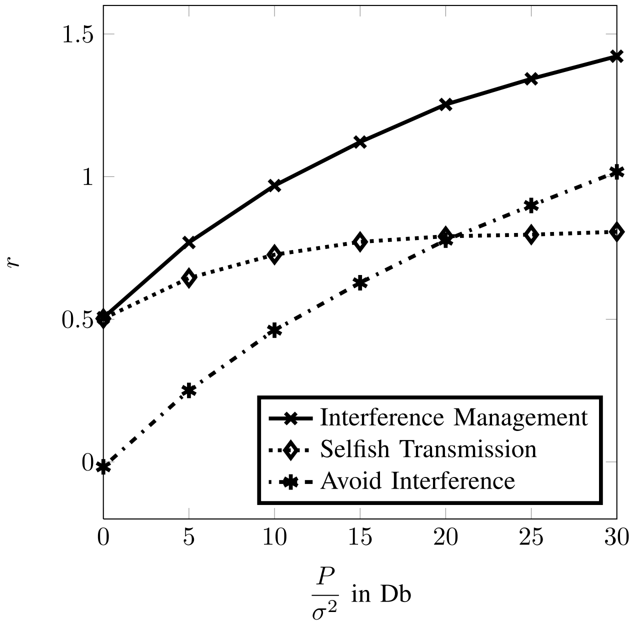


Fig. 2. Multicast Throughput vs. SNR for 4 Terminals

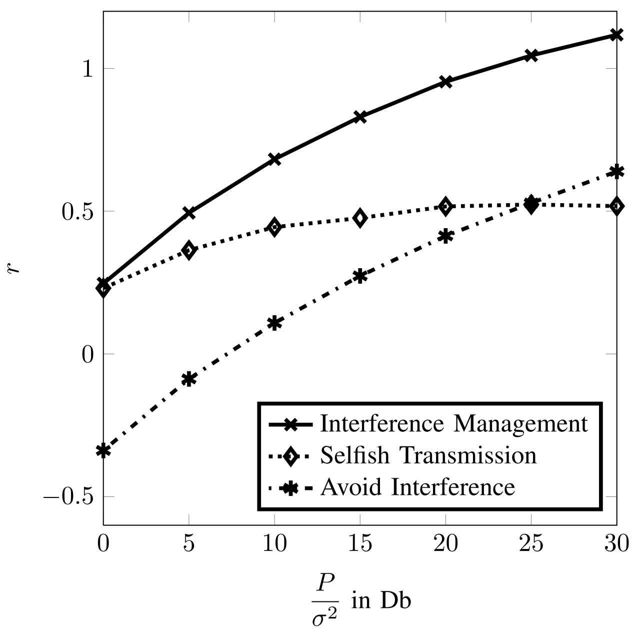


Fig. 3. Multicast Throughput vs. SNR for 6 Terminals

management capabilities in each ECG. For low SNR, selfish transmission is a good strategy, as noise dominates and the impact of interference is almost irrelevant. Obviously, for higher SNR the selfish transmission strategy suffers from being interference limited.

V. Conclusions and Future Work

In this paper, we investigated the potential benefit of advanced physical layer techniques to multicast throughput enhancement in wireless mesh networks. To this end, in a first step we introduced a factorization of the edge capacity region into multiple elementary capacity graphs, each operating in an interference management mode. The proposed optimization approach allows to exploit the dual decomposition framework, although the individual rate regions of the introduced ECGs do not fulfill the required convexity properties. For the physical layer, we used the two user interference channel to illustrate the enhancement of throughput by exploiting the advanced interference management abilities of multiple antenna systems. Although finding the optimal configuration of the two user interference channels requires to run the polyblock algorithm, a global optimization method, a solution to the network optimization problem is found in polynomial time. The ability to decompose the physical layer subproblem into a problem per ECGs keeps the number of variables of the polyblock algorithm constant, while the number of ECGs grows polynomial with the number of nodes, cf. (13). In contrast the naive approach to solve the physical layer subproblem jointly for all ECGs by the polyblock algorithm would result in non-polynomial complexity. By numerical simulations we show a significant gain in throughput compared to systems that do not manage interference. The framework presented in this work is very general with respect to the transmission techniques chosen for the ECGs, as long as the ECG has a parametrization of the rate region and an algorithm to solve the weighted sum rate problem. In ongoing work we will consider the a factorization of the edge rate region in ECGs of various other types. A further direction of future research is on the *Wireless Multicast Advantage* (WMA), which describes the fact that other nodes then the intended receiver might be able to decode the transmitted message, allowing nodes to simultaneously transmit identical data to many receivers. The benefits of considering the WMA haven been shown in [12], where each node has one antenna and ECGs are constructed by BC systems. Motivated by the degradedness of the SISO-BC channel and the superposition coding used, an adequate model for the WMA is derived. However, the WMA is difficult to model in general, and especially for MIMO systems where channels are in general not degraded.

VI. Acknowledgements

We would like to thank Maximilian Riemensberger for providing his expertise and help on implementing the software framework for the numerical simulations.

References

[1] J. Yuan, Z. Li, W. Yu, and B. Li, “A cross-layer optimization framework for multihop multicast in wireless mesh networks,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 11, pp. 2092–2103, 2006.
[2] Y. Wu, P. Chou, Q. Zhang, K. Jain, W. Zhu, and S.-Y. Kung, “Network planning in wireless ad hoc networks: a cross-layer approach,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 136–150, Jan. 2005.
[3] L. Xiao, M. Johansson, and S. Boyd, “Simultaneous routing and resource allocation via dual decomposition,” *IEEE Transactions on Communications*, vol. 52, no. 7, pp. 1136–1144, July 2004.
[4] S. Toumpis and A. Goldsmith, “Capacity regions for wireless ad hoc networks,” *Wireless Communications, IEEE Transactions on*, vol. 2, no. 4, pp. 736–748, July 2003.

see the program

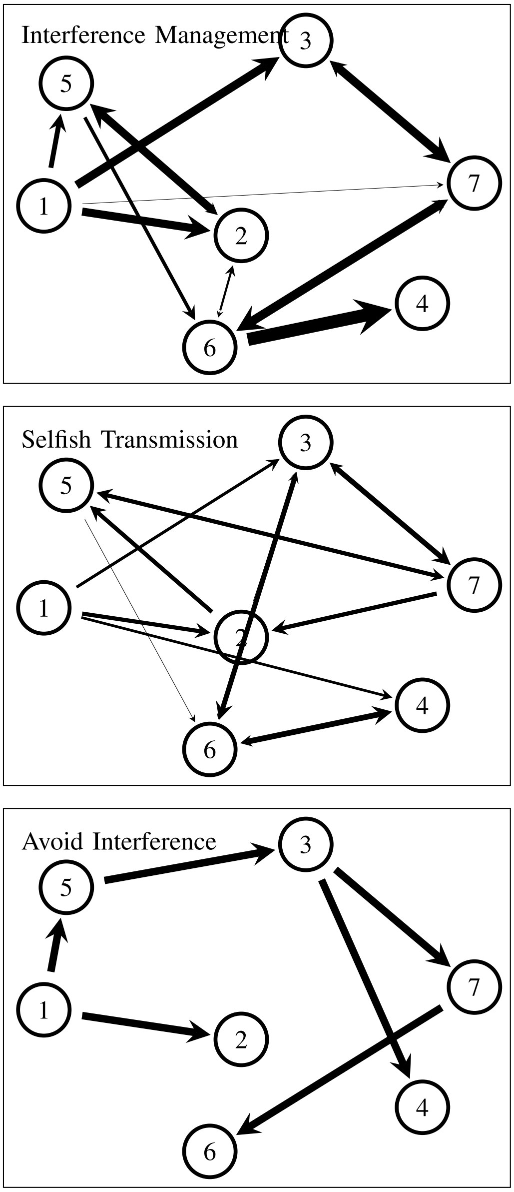


Fig. 4. Solution of the Network Optimization Problem, Source = 1,Terminals = {2,3,4,5,6,7}

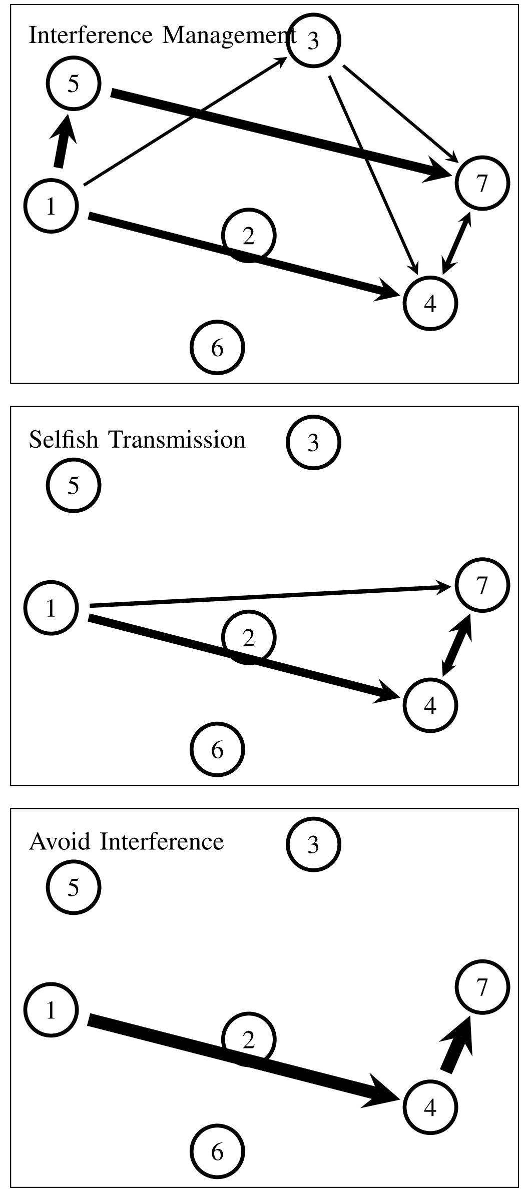


Fig. 5. Solution of the Network Optimization Problem (Source = 1,Terminals = {4,7})

[5] R. Cruz and A. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 1, pp. 702–711 vol.1, March-3 April 2003.
[6] J. Liu, Y. Hou, Y. Shi, and H. Sherali, "Cross-layer optimization for mimo-based wireless ad hoc networks: Routing, power allocation, and bandwidth allocation," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 913–926, August 2008.
[7] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
[8] S.-Y. Li, R. Yeung, and N. Cai, "Linear network coding," *Information Theory, IEEE Transactions on*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
[9] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.
[10] J. Kelley, J. E., "The Cutting-Plane Method For Solving Convex Programs," *J. Soc. Indust. Appl. Math.*, vol. 8, no. 4, pp. 703–712, Dec. 1960.
[11] T. Larsson, M. Patriksson, and A. Strömberg, "Ergodic, primal convergence in dual subgradient schemes for convex programming," *Mathematical Programming*, vol. 86, no. 2, pp. 283–312, 1999.
[12] M. Riemensberger, A. Dotzler, and W. Utschick, "A rigorous approach to factorization in wireless communication networks," 2009, *Submitted to the IEEE International Symposium on Information Theory 2009*.
[13] C. Shi, R. Berry, and M. L. Honig, "Distributed Interference Pricing with MISO Channels," in *Forty-Sixth Annual Allerton Conference, University of Illinois at Urbana-Champaign, IL, USA*, 2008.
[14] E. Jorswieck and E. Larsson, "The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves pareto optimality," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 5364–5367.
[15] ——, "Linear precoding in multiple antenna broadcast channels: Efficient computation of the achievable rate region," *International ITG Workshop on Smart Antennas, 2008*, pp. 21–28, 2008.
[16] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000. [Online]. Available: http://link.aip.org/link/?SJE/11/464/1

see the program

PERFORMANCE INVESTIGATION ON SCAN-ON-RECEIVE AND ADAPTIVE DIGITAL BEAM-FORMING FOR HIGH-RESOLUTION WIDE-SWATH SYNTHETIC APERTURE RADAR

F. Bordoni, M. Younis, E. Makhoul Varona, N. Gebert, G. Krieger

Microwaves and Radar Institute (IHR), German Aerospace Center (DLR), Oberpfaffenhofen, Germany;
E-mail: federica.bordoni@dlr.de; Tel: +49 (0)8153 28 3301; Fax: +49 (0)8153 28 1449

ABSTRACT

Intensive research is currently ongoing in the field of Smart Multi-Aperture Radar Technique (SMART) for Synthetic Aperture Radar (SAR). This work investigates the performance of the SMART SAR system for high-resolution wide-swath imaging based on Scan-on-Receive (SCORE) algorithm for receive beam steering. SCORE algorithm works under model mismatch conditions in presence of topographic height. A study on the potentiality of an adaptive approach for receive beam steering based on spatial spectral estimation is presented. The impact of topographic height on SCORE performance in different operational scenarios is examined, with reference to a realistic SAR system. The SCORE performance is compared to that of the adaptive approach by using the Cramèr Rao lower bound analysis.

1. INTRODUCTION

Spaceborne SAR for remote sensing applications is experiencing a golden age, as testified by the number of the recent and forthcoming missions, e.g. ALOS PALSAR, TerraSAR-X, COSMO-SkyMed, RADARSAT-2, TanDEM-X, Sentinel-1. Nevertheless, the current generation of spaceborne SAR sensors suffers a basic limitation: it does not allow for high resolution imaging and, simultaneously, wide coverage and high radiometric resolution [1]. For instance, a spatial resolution around 1 m could be achieved over a swath width of 10 km; whereas coverage of 200 km allows for SAR final products[1] with a resolution in the order of 100 m [2]. The importance for many remote sensing applications to overcome this limitation has motivated an intensive research within the frame of SMART (see [3, 4] and the references therein).

Main characteristics of SMART SAR systems are the employment of Smart antennas [5], i.e. the use of multiple transmit/receive channels and the introduction of digital signal processing techniques, such Digital Beam-Forming (DBF), in the conventional SAR processing [3, 4, 6]. In fact, Smart antennas allow a relaxation of SAR system design constraints by increasing the degrees of freedom, which results in lower ambiguity level, higher signal-to-noise ratio (SNR) and improved radiometric resolution, and a mitigation of the trade-off between swath width and spatial resolution [3, 4, 6]. It is worth noting that the intrinsic huge quantity of information associated with high-resolution and wide-swath imaging, together with redundancies involved by the multichannel acquisition, could place critical requirements on the downlink data rate.

Among SMART SAR, the system proposed by Suess *et al.* [7, 8], denoted as HRWS, merges the advantages of an extensive illumination capability with the high gain and directivity of a large antenna, and combines the flexibility offered by a multi-channel architecture with a limited download data volume. The HRWS SAR system is based on an algorithm for steering of the elevation beam pattern, called SCORE: a wide swath is illuminated by using a small transmit antenna; whereas in reception a large multi-channel antenna and DBF are employed in order to obtain a sharp and high gain pattern, which follows the pulse echo as it travels along the ground swath. The steering direction of the receive pattern corresponds to the expected direction of arrival (DOA) of the echo, which is assumed *a priori* known. In particular, according to [8], it is computed based on the vertical slant-range plane acquisition geometry, under the hypothesis of a stringent spherical Earth model, i.e. no topographic height is taken into account. Nevertheless, in real acquisition scenarios, characterized by mountains and relief, there will be a displacement between the actual DOA and the steering (i.e. maximum gain) direction; which results in a gain loss with respect to the ideal operational conditions (see Figure 1). Moreover, SCORE steering approach neglects not only the effect of the actual topographic profile along the slant-range elevation plane, but also the effect of surface variations along the azimuth direction.

These observations suggest the option of an Adaptive Digital Beam-Forming (ADBF), i.e. to compute adaptively the steering direction of the receive beam, by (digitally) processing the signals available from the vertical sub-apertures of the multi-channel receive antenna. In fact, the vertical sampling provides a “spatial history” of the signal, which could be used to evaluate the distribution of the received energy as a function of the DOA; then the receive beam steering direction could be selected as the one associated with the strongest

[1] The spatial resolution of a final product is further degraded by the multilook processing, mainly necessary to obtain satisfactory radiometric resolution.

see the program

signal, eventually within a roughly expected spatial sector.

According to the ADBF approach, the receive beam steering algorithm is cast into the frame of *spatial spectral estimation* and *DOA estimation*. This topic has been extensively studied in array signal processing theory [9, 10], and also with reference to the Interferometric SAR application [11]. Nevertheless, the HRWS SAR spaceborne application shows specific challenges. First, the processing of the signals available from the vertical sub-apertures should be performed onboard, in order to reduce the downlink data volume. This requires to dealing with wideband signals and imposes additional constraints on the complexity of the processing method [10, 12]. Moreover, in case of wide illuminated swaths, the useful signal could be superimposed to range-ambiguous echoes having a power comparable with that of the signal of interest. Finally, instrument parameters, such as dimension of the antenna, number of elements, noise level (NESZ), whose values strongly affect the ultimate estimation performance, do not allow for many degrees of freedom, due to imaging requirements and physical/economical constraints.

This paper shows the effect of topographic height on conventional, not adaptive, SCORE performance: the steering displacement and the corresponding gain loss introduced by topography are analyzed as a function of the acquisition geometry and of the receive antenna architecture and pattern shape, with reference to a realistic SAR system operational scenario. Moreover, in order to evaluate the potentiality of an ADBF for spaceborne high-resolution wide-swath SAR systems, a performance analysis based on Cramér Rao Lower Bound (CRLB) is developed [9, 10]. The performance of SCORE and ADBF are compared vs. the main SAR system parameters.

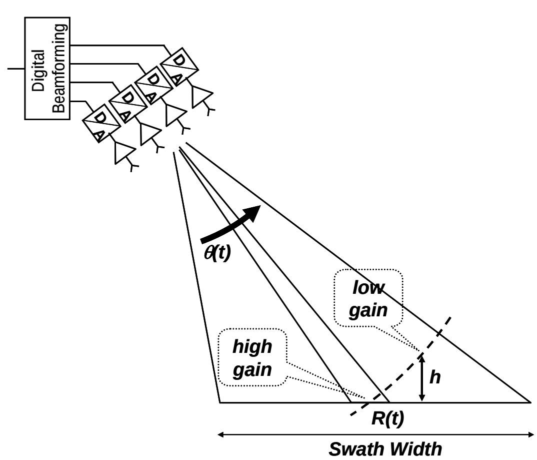


Fig. 1. SCORE in presence of topographic height.

2. REVIEW OF SCAN-ON-RECEIVE

The HRWS SAR system employs a small transmit antenna and a large receive antenna, which is split in multiple sub-apertures, arranged according to a uniform linear array (ULA) geometry in azimuth and elevation [7, 8]. The small transmit aperture is used to illuminate a large swath; the large receiving antenna allows for ambiguity suppression and compensation for the reduced transmit antenna gain. In particular, the conflict between swath width and azimuth resolution is overcome by using M sub-apertures located along the azimuth direction, according to the displaced phase center antenna (DPCA) technique; whereas the SAR radiometric resolution is improved by using multiple sub-apertures in elevation, and by processing the corresponding signals by SCORE algorithm. In detail, according to [8], the echo DOA is computed based on the vertical slant-range plane acquisition geometry, assuming a stringent spherical Earth model (no topographic elevation). Under this assumption, in fact, the DOA of the echo received from a point-like target is univocally associated to the two-way time delay, τ [8]:

$$\vartheta(\tau) = \operatorname{acos}\left\{ \frac{4\left(H_{orb}+R_E\right)^2 - 4R_E^2 + \left(c_0\tau\right)^2}{4\left(H_{orb}+R_E\right) c_0\, \tau} \right\} \qquad (1)$$

where, ϑ denotes the DOA measured w.r.t. the nadir angle, H_{orb} is the orbit height; R_E the Earth radius; c_0 the light speed. Then a time-varying DBF is used to combine the signals received by the elevation sub-apertures, in order to obtain, at each instant, a sharp and high gain beam, steered towards the expected DOA of the backscattering echo[2]. The DBF reduces the data rate by eliminating the redundancies; the high gain SCORE beam results in an increased SNR, compensating the low gain (wide beam) of the transmit antenna. Specifically at the swath edges (half-power beamwidth angles) the typical two-way loss of a conventional system is reduced.

3. BEHAVIOUR OF SCORE IN PRESENCE OF TOPOGRAPHIC HEIGHT

In presence of topographic height, h, SCORE works under model mismatch conditions: there is an *angular displacement*, $\Delta\vartheta$, between the actual DOA of the received echo, ϑ_{act}, and the SCORE steering direction, ϑ_s, computed according to eq. (1):

$$\Delta\vartheta = \vartheta_{act} - \vartheta_s(\tau_{act}), \qquad (2)$$

[2] In case of long chirp pulses an additional frequency dependent beam steering is included in SCORE [8, 9]. In the rest of the paper, we assume that SCORE frequency dependent beam steering allows recovering all the pulse energy with the maximum gain, when the DOA of the pulse centre is correctly recovered [8, 9].

where, τ_{act} denotes the actual two-way time delay of the received echo. This results in a loss of performance. In fact, the echo impinging on the receive antenna is not weighted by the maximum of the receive beam, as expected in ideal (no topographic height) conditions. The degradation of SCORE performance due to topographic height is here quantified by the parameter *Pattern Loss (PL)*, defined as the value of the normalized receive beam pattern, which weights the echo backscattered from a source located at a topographic height, h:

$$PL = \frac{C^R(\vartheta_s + \Delta\vartheta)}{C^R(\vartheta_s)} \qquad (3)$$

where, $C^R(\vartheta)$ indicates the value of the elevation receive beam pattern generated by DBF corresponding to the angle ϑ.

In order to evaluate the effect of surface elevation on a realistic HRWS SAR system, the reference HRWS SAR system described in Table 1 has been considered. This system allows for a spatial resolution and swath width in the order of 1 m and 70 km, respectively, a NESZ below – 22 dB, range ambiguity (RASR) below – 30 dB and azimuth ambiguity (AASR) below -26 dB.

Figures 2 and 4 show the angular displacement and the PL as a function of the ground-range position (of the projection on the Earth surface) of the backscattering source, parameterized vs. the source topographic height, h. Values of h between 0 m and 8000 m have been considered.

As shown in Figure 2, the angular displacement depends on the acquisition geometry. The main effect is related to the surface height: for the reference system, the angular displacement varies between 0.15° for h=1000 m and 1.45° for h=8000 m. For a fixed height, the displacement weakly decrease for increasing ground range positions. The extent of the PL depends on the receive beam shape.

Tab. 1. Reference HRWS SAR system parameters.

PARAMETER	UNIT	VALUE
Geometry		
Orbit Height	km	520
Antenna Tilt Angle, β	deg	32.25
Swath Limits (look ang., ground range)	deg km	[29.6, 34.9] [300, 370]
Radar Parameters		
PRF	Hz	1775
RF Center Frequency	GHz	9,65
Pulse Bandwidth	MHz	250
Av. Tx Power	W	1100
TX Antenna		
Height	m	0.50
Length	m	2.45
RX Antenna		
Height	m	1.5
Nr. of sub-apert. in el.		15 (x 0.10 m)
Length	m	9.8
Nr. of sub-apert. in az.		7

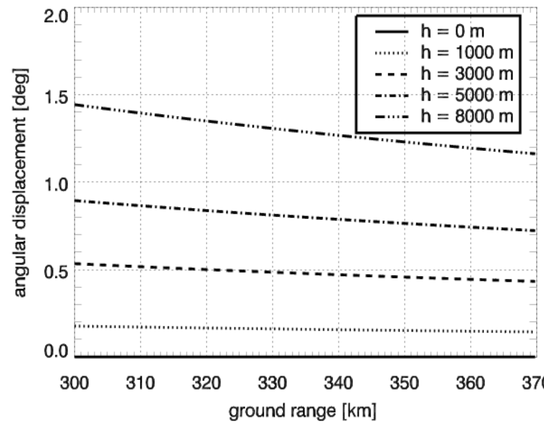


Fig. 2. SCORE angular displacement, in presence of topographic height, h.

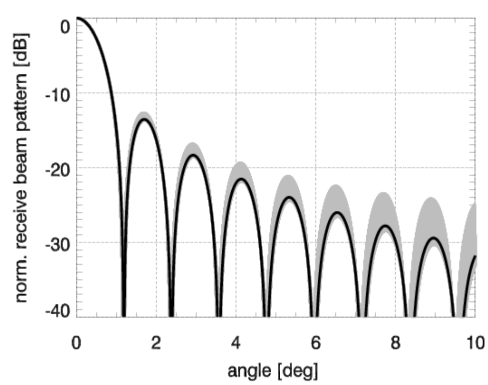


Fig. 3. Elevation receive beam patterns generated by DBF during the scansion of the swath vs. the elevation angle (instantaneous steering direction reported to the origin).

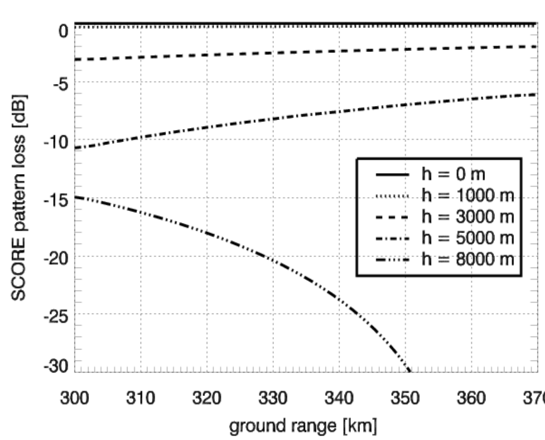


Fig. 4. SCORE pattern loss in presence of topographic height, h.

In particular, for the reference HRWS SAR, the half power beam width (HPBW) of the receive beam is in the order of 1° (see Figure 3). SCORE PL, obtained as the value of the normalized receive pattern corresponding to

see the program

the steering displacement, is plotted in Figure 4. The PL curves show that loss is moderate for topographic height below 1000 m; nevertheless a loss of several dB could be reached when no information about topographic height is used to steer the beam.

4. ADAPTIVE DIGITAL BEAM-FORMING

According to SCORE, the received pattern is steered towards the DOA of the received signal, assumed *a priori* known. Nevertheless, a critical loss of performance could occur when no information about the actual backscattering geometry is conveyed in the steering mechanism. An alternative approach is to steer the receive beam towards an estimate of the actual DOA, obtained by processing the signal echo impinging on the multiple sub-apertures, arranged according to a ULA geometry along the elevation direction.

As explained in the introduction, the wideband nature of the SAR pulses and the constraints on the algorithm complexity do not allow a direct application of simple spectral estimation methods, such as Beamforming or Capon, which have been successfully applied in radar array processing [9, 10] and also in multichannel Interferometric SAR (InSAR) [11]. On the other hand, the methods conceived for broadband signal are computationally demanding for onboard processing [10, 12].

In the following, a pre-processing procedure, which allows to trace the SAR signal back to a signal model typically used in narrowband spectral estimation [9, 10], is proposed. For simplicity, the explanation refers initially to a point-like target and is afterward generalized.

Assume that the transmitted SAR pulse is a chirp:

$$s(t) = rect\left(\frac{t}{T}\right)\cos\left[2\pi\left(f_c\,t + \frac{\kappa}{2}t^2\right)\right], \qquad (4)$$

where T is the chirp duration, κ the chirp rate, f_c the RF center frequency. The two-way time delay of the echo, received by the $k-th$ sub-aperture in elevation from a point-like target located in the vertical slant-range plane, could be written as:

$$\tau_k = \tau_0 + \frac{d\cdot k\cdot\sin(\beta-\vartheta)}{c_0}, \qquad k = 0,\cdots,K-1, \qquad (5)$$

where τ_0 denotes the two-way time delay of the pulse at the first ($k=0$) receive sup-aperture, d is the distance between the phase centers of two adjacent sub-apertures, K the number of elevation sub-apertures of the receive antenna, ϑ is the echo DOA and β the antenna tilt angle, both measured respect to the nadir. Then, the equivalent baseband signal of the echo at the $k-th$ sub-aperture could be written as:

$$r_k(t) = \alpha\cdot rect\left(\frac{t-\tau_k}{T}\right)\exp\left\{j\pi\kappa(t-\tau_k)^2\right\}\exp\left\{-j2\pi f_c\tau_k\right\},$$

$$k = 0,\cdots,K-1 \qquad (6)$$

where, α is the complex amplitude accounting for the propagation and backscattering mechanism. Note that from here on t indicates a discrete time variable.

After range compression with a matched filter to the transmit chirp and the application of a coregistration processing [13], the signal in (6) could be written as:

$$r_k(t) = \alpha\cdot \mathrm{sinc}\left\{\kappa\cdot T\cdot(t-\tau_0)\right\}\cdot\exp\left\{-j\cdot 2\pi f_c\tau_k\right\}, \qquad (7)$$

where, for simplicity, we use the same symbols as in (6) for the signal and the complex constant. Then considering the time-sample corresponding to the time delay, $t=\tau_0$:

$$r_k = \alpha\cdot\exp\left\{-j2\pi\frac{d}{\lambda}\sin(\beta-\vartheta)k\right\}, k = 0,\cdots,K-1 \qquad (8)$$

where λ denotes the radar wavelength, and without loss of generality, the factor $\exp\{-j2\pi f_c\tau_0\}$ has been included in the complex constant α.

Eq. (8) could be rewritten by using a vector notation and taking into account the additive thermal noise, which corrupts the useful signal:

$$\mathbf{y} = \alpha\cdot\mathbf{a}(\vartheta) + \mathbf{v}, \qquad (9)$$

where, $\mathbf{y}$, $\mathbf{a}(\vartheta)$, and $\mathbf{v}$ are K-dimensional complex vectors, and the element $k-th$ is associated to the $k-th$ sub-aperture. In particular, $\mathbf{v}$ collects the thermal noise contribute, $\mathbf{a}(\vartheta)$, whose $k-th$ element is given by

$$[\mathbf{a}(\vartheta)]_k = \exp\left\{-j2\pi\frac{d}{\lambda}\sin(\beta-\vartheta)k\right\}, \qquad (10)$$

denotes the steering vector of the pulse impinging on the array, which collects the information of the DOA.

It is worth comparing the echo received at the same sub-aperture from different, subsequent pulse transmissions: the echo DOA changes in a negligible way for $N<100$ multiple consecutive azimuth acquisitions[3]; moreover, the effect of range cell migration on the time delay, i.e. on the position of the recorded samples corresponding to the same backscattering source, could be easily computed and compensated onboard. These observations allow considering different equivalent *snapshots* or *looks*, n, of the signal in eq. (9):

$$\mathbf{y}(n) = \alpha\cdot\mathbf{a}(\vartheta) + \mathbf{v}(n),\ n = 1,\cdots,N\ . \qquad (11)$$

Eq. (11) could be generalized by introducing multiple, N_s, backscattering sources:

$$\mathbf{y}(n) = \sum_{i=1}^{N_s}\alpha_i\cdot\mathbf{a}(\vartheta_i) + \mathbf{v}(n),\ \ n = 1,\cdots,N\cdot \qquad (12)$$

[3] For the typical satellite parameters system, the maximum DOA variation due to the azimuth platform displacement is in the order of 10^{-5} deg.

see the program

In fact, in general, multiple echoes of the transmitted SAR pulses impinge on the receive antenna simultaneously, due to preceding or succeeding pulses with time-delays separated by a multiple of the PRF or backscattering sources in layover. When extended homogeneous backscattering sources are considered, the signal component in eq. (12) could be written as [11, 14]:

$$\alpha_i \cdot \mathbf{a}(\vartheta_i) = \sqrt{\tau_i}\mathbf{a}(\vartheta_i) \otimes \mathbf{x}_i(n), \; n = 1, \cdots, N, \qquad (13)$$

where, $\otimes$ denotes the Hadamard product; τ_i denotes the texture, i.e. the mean power level of each source, and $\mathbf{x}_i$ the speckle [11].

The problem of evaluating the distribution of the received energy as a function of the DOA is now traced back to the narrowband spectral estimation, and can be expressed as the estimation of the spatial power spectral density (PSD) of the data in eq. (12).

In order to statistically describe the signal in eq. (12-13), it is useful to recall to the model used in multichannel InSAR for the pixel complex amplitude, $\mathbf{y}(n)$, collected by the K sensors array [11, 14]:

$$\mathbf{y}(n) = \sum_{i=1}^{N_S} \sqrt{\tau_i}\mathbf{a}(\varphi_i) \otimes \mathbf{x}_i(n) + \mathbf{v}(n), \; n = 1, \cdots, N \qquad (14)$$

where, N the number of independent looks; $\mathbf{v}$ the thermal noise contribute, modeled as a complex Gaussian spatially white process, with zero mean and power σ_v^2; N_s denotes the number of extended homogeneous backscattering sources in layover; $\mathbf{a}(\varphi)$ the spatial steering vector; φ_i the source interferometric phase[4] (directly related to the DOA), modeled as an unknown constant; τ_i, is the radar reflectivity or texture, modeled as a real, positive, unknown deterministic parameter; $\mathbf{x}_i$ the speckle, modeled as complex correlated Gaussian random vector with a zero mean, unit variance and covariance matrix $\mathbf{C}_i$; with $\mathbf{x}_i(n_p)$ independent of $\mathbf{x}_i(n_q)$ when $n_p \neq n_q$, for $i = 1, \cdots, N_s$.

In the following analysis we will refer to the statistical model of eq. (14). This means that we consider an illuminated swath characterized, along the iso-range lines, by a homogeneous backscattering surface and constant topographic height. Though simple, this reference surface allows for a first comparison between SCORE and ADBF achievable performance.

5. BEHAVIOUR OF ADBF: CRAMÉR RAO LOWER BOUND

The achievable performance of the ADBF could be evaluated by CRLB analysis, based on the model in eq. (14). In fact, the value of the CRLB on the estimate of the DOA provides the minimum variance of any unbiased estimator of ϑ_i [9, 10]. Then, in analogy with the SCORE angular displacement in eq. (2), it could be defined an angular displacement between the actual DOA of the echo and that estimated by ADBF as:

$$\Delta\vartheta = \sqrt{CRLB\{\hat{\vartheta}_i\}}, \qquad (15)$$

where $\sqrt{CRLB\{\hat{\vartheta}_i\}}$ denotes the square root of the CRLB on the estimate of ϑ_i, i.e. the standard deviation of the estimated DOA, $\hat{\vartheta}_i$. The pattern loss can be computed according to eq. (3) for the angular displacement in eq. (15).

For the CRLB computation and performance analysis, it is assumed that the speckle covariance matrix is:

$$\left[\mathbf{C}_i\right]_{u,v} = E\left\{\left[\mathbf{x}_i(n)\right]_u \cdot \left[\mathbf{x}_i(n)\right]_v^H\right\} = \begin{cases} 1 - \dfrac{|u-v|}{K-1} \cdot H_i & \text{for } |u-v| \leq \dfrac{(K-1)}{H_i} \\ 0 & \text{otherwise} \end{cases} \qquad (16)$$

where $E\{\cdot\}$ denotes the mean statistical value, $(\cdot)^H$ the conjugate transpose operator, $H_i = H_{ant}/H_{ci}$ the normalized antenna's height, given by the ratio between the size of the receive antenna along the elevation direction, H_{ant}, and the critical antenna height, H_{ci}, for which the $i-th$ source decorrelates at the extremities of the array [11, 14, 15].

The CRLB on the DOAs, $\{\vartheta_i\}_{i=1}^{N_s}$, can be derived as in [11], particularizing the data model (14) for the DOA, ϑ_i, through the relationship with the interferometric phase, φ_i as:

$$\varphi_i = -4\pi \frac{d(K-1)}{\lambda} \sin(\beta - \vartheta_i), \qquad (17)$$

and considering the speckle covariance matrix, $\mathbf{C}_i$, as defined in (16).

6. PERFORMANCE COMPARISON

This section shows a comparison between the performance achievable with SCORE and ADBF. The performance is evaluated in terms of angular displacement and pattern loss. The reference HRWS SAR described in Table 1 is considered. Moreover, a reference acquisition scenario is assumed, which is characterized by the following parameters: two extended homogeneous backscattering sources, $N_s = 2$, located respectively at 304.42 km and 317.87 km ground range

[4] The interferometric phase is defined as the phase difference between the furthest phase centers of the array [11].

see the program

(30.0° and 31.05° off-nadir angle[5]), with a topographic height of 3000 m; both sources have the same mean power, such that the corresponding signal-to-noise ratio at the output of the array, $SNR_i = K\ \tau_i / \sigma_v^2$, is equal to 9 dB; the normalized antenna height, H_i, is in the order of 1.8e-4 for both sources[6]; the number of consecutive azimuth acquisitions, or *independent looks*, is $N = 5$ (this value allows to completely neglect the range cell migration effect). Starting from this reference acquisition scenario/system, the performance of SCORE and ADBF are evaluated as a function of single parameters of interest, by keeping constant the other parameters value. The results corresponding to the reference scenario are indicated by an asterisk in each figure.

Figures 5 show the performance as a function of the topographic height. For the ADBF the angular displacement is about 0.11°, with a trend *almost* non sensitive to the topographic variation; the corresponding pattern loss is negligible. SCORE angular displacement increases proportionally to the height value, till 1.4° for a topographic height of 8000 m. SCORE pattern loss increases till a value of -25 dB for a height of 7000 m. It must be noted that for 7000 m the PL is higher than for 8000 m, because the displacement for the case of 7000 m approaches the first null of the receive pattern, whereas in the case of 8000 m the signal is acquired with the first secondary lobe (see Figure 3). It is worth noting that the performance of ADBF is negligibly affected by the location of the sources along the swath. Figure 6 reports the angular displacement as a function of the SNR of the first source (# 1), $SNR_1 = K\ \tau_1 / \sigma_v^2$. For ADBF, the accuracy on the estimation of ϑ_1 improves as a function of the SNR_1; whereas for SCORE it remains unaltered. The SNR threshold, where the performance of the ADBF becomes better w.r.t. the conventional SCORE is around 0 dB. Figure 7 investigates the effect of source angular separation. As expected, SCORE performance is constant. The performance of ADBF is better than that of SCORE, for a source separation greater than about half of the HPBW. Figures 8 investigate the impact of the height of the receive antenna. The number of sub-apertures is constant, $K = 15$, i.e. the antenna size variation implies also a variation of the sub-aperture spacing (inter-element spacing). The angular displacement of ADBF degrades as the antenna becomes shorter. In fact, the reduced separation between the sub-apertures induces a low sensitivity (all the sub-apertures "observe" approximately the same signal). The antenna height $H_{ant} = 1m$ corresponds to the lower limit, where the performance of ADBF is better than that of SCORE. The angular displacement of SCORE is constant, nevertheless the corresponding PL degrades as the antenna height increases due to the reduced mainlobe width of the receive beam. Figure 9 shows the performance as a function of the number of sub-apertures in elevation, while the antenna size in elevation is constant, $H_{ant} = 1.5$ m. The availability of multiple sub-apertures does not affect the accuracy of ADBF estimation, at least when the sources are sufficiently distant (see Figure 7). For instance, when the DOA separation is about 0.5 HPBW, the use of $K = 4$ in spite than 15 sub-apertures degrades the angular displacement from 0.45° to 0.69°. It is worth noting that a low number of sub-apertures reduces the angular unambiguous range (with $H_{ant} = 1.5m$ it is about 17° for $K = 15$, and 5° for $K = 4$). As a consequence, possible echos of preceding or succeeding pulses could fold over the source of interest, reducing the estimation performance. Figure 10 investigates the effect of decorrelation, showing the performance as a function of the normalized antenna height. The performance of ADBF degrades for increasing normalized antenna height: for values greater than 0.8 the ADBF performance is worse than that of SCORE. The previous results are obtained by considering the reference value of number of looks, $N = 5$, which allows neglecting any effect of range cell migration and avoiding the corresponding onboard processing. It is worth noting that the square root of the CRLB, i.e. ADBF angular displacement, reduces for increasing number of looks as the square root of N [9, 10]. Figure 11 shows the angular displacement as a function of the number of looks: for the reference system parameters, the use of a higher number of looks it is not justified. Nevertheless, the use of a higher number of looks could be useful in critical acquisition conditions, such as low SNR and high decorrelation.

7. CONCLUSIONS

The impact of topographic height on SCORE performance has been examined with reference to a realistic SAR system. The numerical results show that the loss is moderate for topographic height below 1000 m; nevertheless a loss of several dB could occur when no information about the topographic height is used to steer the receive beam. An adaptive approach, ADBF, for receive beam steering based on spatial spectral estimation has been proposed. The achievable performance of ADBF have been investigated by the Cramèr Rao lower bound analysis and compared to those of SCORE versus the main system parameters. The numerical analysis shows that ADBF outperforms SCORE and reaches promising results in most of the analyzed scenarios.

[5] The angular separation between the DOAs of the two sources corresponds to the HPBW of the receive beam.

[6] The critical height, H_{ci}, is in the order of 8.3 km, assuming no local slope [15].

see the program

8. REFERENCES

[1] J. C. Curlander, R. N. McDonough, "Synthetic Aperture Radar: Systems and Signal Processing", NY, Wiley, 1991.

[2] S. Mezzasoma, A. Gallon, F. Impagnatiello, G. Angino, S. Fagioli, A. Capuzi, F. Caltagirone, R. Leonardi, U. Ziliotto, "COSMO-SkyMed system commissioning: End-to-end system performance verification", *Proc. IEEE Radar Conference, RADAR '08*, pp. 1092-1096, May 2008.

[3] G. Krieger, N. Gebert, A. Moreira, "Multidimensional Waveform Encoding: A New Digital Beamforming Technique for Synthetic Aperture Radar Remote Sensing", *IEEE Trans. on Geosci. and Remote Sens.*, vol. 46, no. 1, pp.31-46, Jan. 2008.

[4] M. Younis, F. Bordoni, N. Gebert, G. Krieger, "Smart Multi-Aperture Radar Techniques for Spaceborne Remote Sensing", *Proc. IEEE Geosci. and Remote Sens. Symp., IGARSS´08*, July 2008.

[5] M. Chryssomallis, "Smart Antennas", *IEEE Antennas and Propagation Magazine*, Vol. 42, No. 3, pp. 129-136, June 2000.

[6] N. Gebert, G. Krieger, and A. Moreira, "Digital Beamforming on Receive: Techniques and Optimization Strategies for High-Resolution Wide-Swath SAR Imaging", *IEEE Trans. on Aerosp. and Electron. Syst.*, 2008, accept. for publication.

[7] M. Suess, B. Grafmueller, R. Zahn, "A novel high resolution, wide swath SAR system," *Proc. IEEE Int. Geosci. and Remote Sens. Symp.*, IGARSS´01, vol. 3, pp. 1013-1015, July 2001.

[8] M. Suess and W. Wiesbeck, "Side-Looking Synthetic Aperture Radar System", European Patent Application, EP 1 241 487 A1, Sept. 18, 2002.

[9] P. Stoica, R. Moses, "Introduction to Spectral Analysis", Englewood Cliffs, NJ, Prentice-Hall, 1997.

[10] H. L. Van Trees, "Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing", John Wiley & Sons, 2002.

[11] F. Gini, F. Lombardini, M. Montanari, "Layover Solution in Multibaseline SAR Interferometry," *Trans. on Aerosp. and Electron. Syst.*, vol. 38, no. 4, pp. 1344-1356, Oct. 2002.

[12] G. Wang, X. -G. Xia, "Iterative Algorithm for Direction of Arrival Estimation with Wideband Chirp Signals", *IEE Proc. Radar, Sonar and Navig.*, vol. 147, no. 5, pp. 233-238, Oct. 2000.

[13] A. Moreira, J. Mittermayer, R. Scheiber, "Extended Chirp Scaling Algorithm for Air- and Spaceborne SAR Data Processing in Stripmap and Scansar Imaging Modes", *Geosci. and Remote Sens.*, vol. 34, no. 5, pp. 1123 – 1136, Sept. 1996.

[14] F. Bordoni, F. Gini, L. Verrazzani, "Capon-LS for Model Order Selection of Multicomponent Interfermetric SAR Signals", *IEE Proceeedings Part-F, Radar, Sonar, and Navigation*, vol. 151, no. 5, pp.299 – 305, Oct. 2004.

[15] F. Gatelli, A. M. Guarnieri, F. Parizzi, P. Pasquali, C. Prati, F. Rocca, "The Wavenumber Shift in SAR Interferometry", *IEEE Trans. on Geosc. and Remote Sens.*, vol.32, no. 4, pp. 855 – 865, July 1994.

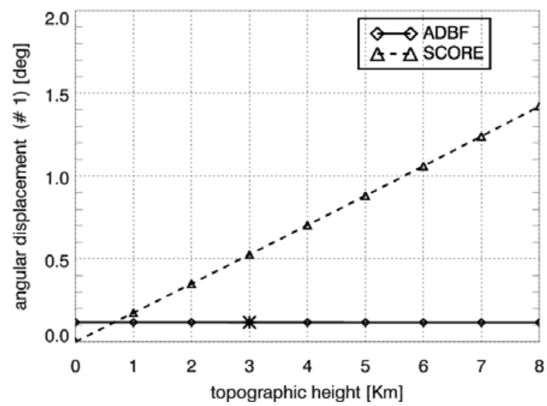


Fig. 5.a. Angular displacement vs. topographic height (const. projection on Earth surface of the source positions).

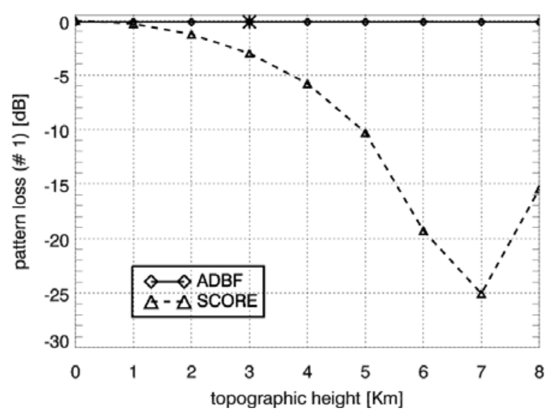


Fig. 5.b. Pattern loss vs. topographic height (const. projection on Earth surface of the source positions).

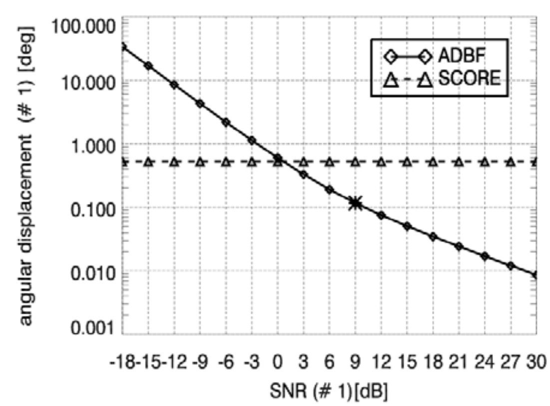


Fig. 6. Angular displacement vs. variation of the SNR of the source #1.

see the program

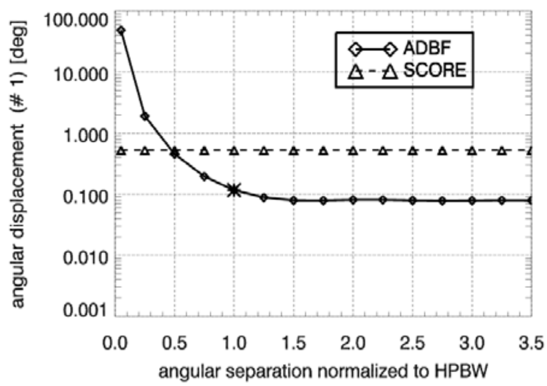


Fig. 7. Angular displacement vs. the angular separation between the two sources normalized to the HPBW (const. position of the source #1).

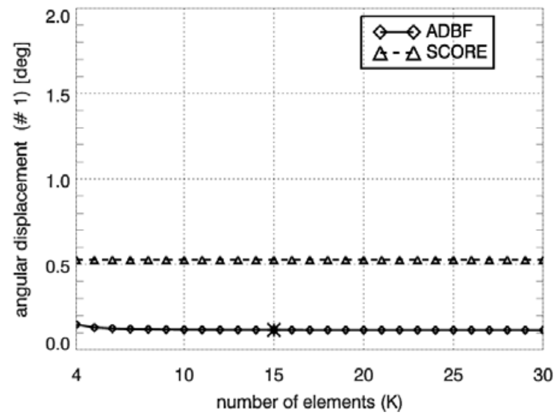


Fig. 9. Angular displacement vs. number of sub-apertures (const. elevation size of the receive antenna).

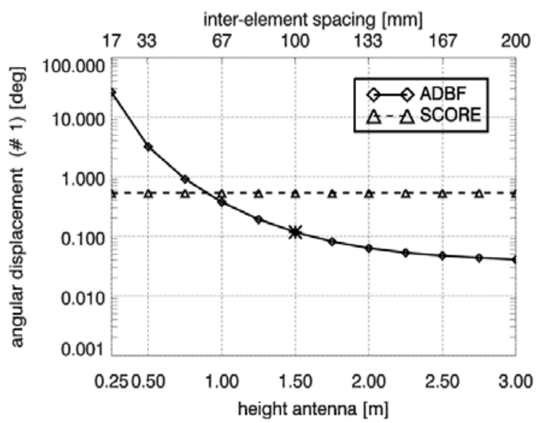


Fig. 8.a. Angular displacement vs. elevation size of the receive antenna (const. number of sub-apertures).

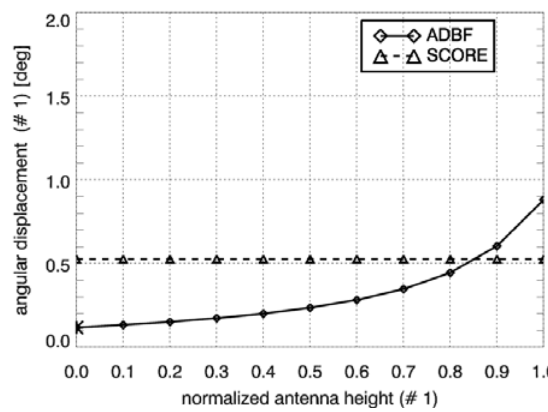


Fig. 10. Angular displacement vs. the normalized antenna height of the source #1.

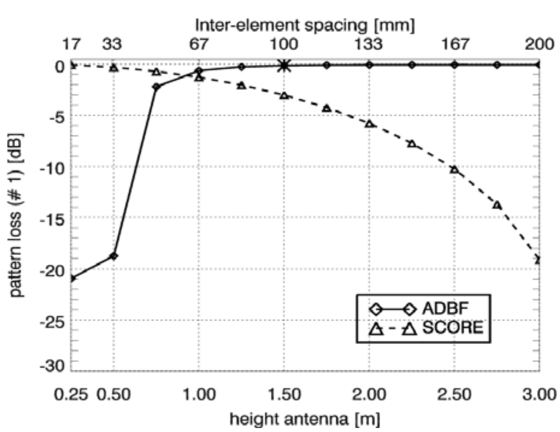


Fig. 8.b. Pattern loss vs. elevation size of the receive antenna (const. number of sub-apertures).

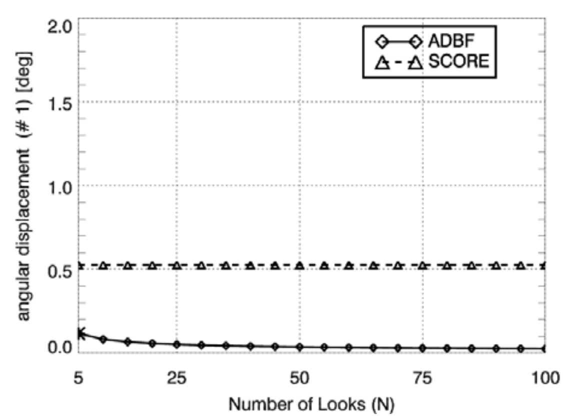


Fig. 11. Angular displacement vs. number of looks.

see the program

EXPERIMENTAL STUDY ON THE IMPACT OF THE BASE STATION HEIGHT ON THE CHANNEL PARAMETERS

Aihua Hong and Reiner S. Thomae

Technische Universitaet Ilmenau
PSF 100 565, D-98684 Ilmenau, Germany
Tel: 0049 3677 691157. Fax: 0049 3677 691113
aihua.hong, reiner.thomae@tu-ilmenau.de

ABSTRACT

In this paper, we investigated the experimental results about the statistical behavior of the channel parameters, based on the multiple-input multiple-output (MIMO) measurements performed in Ilmenau downtown, Germany. Our special interest focuses on the impact of the base station (BS) height on the statistical properties of the large scale channel parameters in two types of propagation condition: line of sight (LOS) and none line of sight (NLOS). Furthermore, the dependence of the inter-sector correlation of the large scale channel parameters on the BS height is investigated in the overlapped area of two sectors of the same BS.

1. INTRODUCTION

In the conventional cellular networks, base station (BS) is highly elevated by mounting on the top of high-buildings or high masks since large coverage by single BS has been expected. However, highly-elevated BS introduces simultaneously significant interference, which impairs the system performance. To find the trade off between coverage and interference, optimized deployment BS height is especially interesting. Different BS heights result in different propagation phenomena. As a consequent, the channel metrics, reflecting the propagation characteristics, show different statistical behavior.

In this paper, we consider the impact of the BS height on the statistical properties of the large scale parameters (LSPs), which have been intensively studied in the European WINNER project [6] [5] and are the key channel parameters for the system-level channel modeling. The impact of the statistical properties of the large scale fading (LSF) on the system performance has been investigated by Fraile in [7] and by Saunder in [1]. The results indicate that an accurate modeling of LSF is a key point for the system-level performance assessment. The statistical properties of the LSPs have been intensively studied by authors in [8] in the urban hot-spot scenario, in [9] [11] in the indoor scenario, and in [10] in the car2bridge scenario. In these comprehensive works, experimental results are presented based on the measurement data. However, the BS height is fixed in each of these measurement campaigns. The primary goal of this paper is to extend these studies to three dimensions. Whereby, the additional dimension is the height of the BS, besides the moving dimension of the mobile station (MS).

In this paper, we include propagation loss, LSF, delay spread (DS), and cross-polarization ratio (XPR) [2] [3] in the terminology LSPs. Using the multiple-input multiple-output (MIMO) measurement data gathered in Ilmenau downtown, Germany, we studied the impact of the BS height on the statistical properties and the inter-sector correlation properties of the channel parameters.

This paper is organized as follows. Section 2 provides a detailed description to the measurement campaign. Section 3 gives a short description to the definitions and extraction procedure of the LSPs, including propagation loss, LSF, DS, and XPR. Section 4 shows the experimental results, comprised of the propagation loss along the main street and two perpendicular streets, the statistical distributions of the LSPs, and their inter-sector correlation properties in the overlapped area. Finally, Section 5 concludes the paper.

2. MEASUREMENT CAMPAIGN

The measurement campaign has been conducted in Ilmenau downtown, Germany, using HyEff sounder [4]. The measurement area can be characterized as a canonical small urban scenario with 1-3 floor stone buildings with the average height around 12 [m] at both sides of the street. The measurements were performed for downlink with approximately 10 [w] transmitting power at the BS antenna at 5.2 GHz in a bandwidth of 120 MHz. The maximum multipath delay has been set to 3.2 μs. Dedicated antenna arrays are used both at the BS and at the MS. 8-element uniform linear patch array (ULA8) is used as the BS antenna with 0.49 λ element spacing, dual-polarization, approximate 120 degree directivity. At

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) within the TakeOFDM project

see the program

the MS, 24 x 4-element stacked dual-polarized uniform circular patch array (SPUCPA24x4) is used with radius about λ and omni-directivity. The MS antenna array is put on the top of a trolley which is about 1.8 [m] above the ground [1].
During the measurement campaign, the height of the BS has been changed from 10 [m] to 16 [m] with the help of a lifting ramp. Figure 1 shows the measurement routes and the orientation of the BS antenna sector in a 2-dimensional map. The names BS2 and BS5 stand for the 10 [m] BSs while BS3 and BS6 stand for the 16 [m] BSs. Two BS sectors, orthogonal with each other, have been set to the each BS height, as indicated in Fig. 1 as arrow. Since the surrounding buildings are around 12 [m] in height in the measurement area, the two BS heights correspond to two kinds of propagation. The one with 10 [m] BS height, below rooftop, represents the micro-cell propagation while the other with 16 [m] BS height, above rooftop, represents the marco-cell propagation. The MS moves both along the main street and along two perpendicular streets. The former is show as a solid curve while the latter is shown as dashed curves in Fig. 1. They correspond to the LOS and NLOS propagation conditions, respectively.

3. PARAMETER DEFINITIONS AND EXTRACTION PROCEDURE

At the beginning of this section, a short description to the definitions of the LSPs studied in this paper is given. After this, the methods used to extract the LSP values from the measurement data are described.

3.1. Definitions of propagation loss and LSF

The received signal power in mobile communication is often modeled as a produce of four factors (seen Eqn. 1): the transmitted power (in [dBm]), the distance dependent path loss (in [dB]) [2], log-normal distributed LSF (in [dB]) [14], and small scale fading (in [dB]) due to the superposition of multipath propagation. Small scale fading is also called as fast fading.

$$P_{Rx} = P_{Tx} - PL + LSF + FF. \tag{1}$$

The difference between the transmitted signal power P_{Tx} and the received signal power P_{Rx} is defined as propagation loss. Note that the antenna gain is included in propagation loss in this paper. The path loss characterizes the dependence of the signal attenuation on the distance between the transmitter and the receiver. As a result of multiplication of large number of random attenuating factors in the propagation, the LSF has a log-normal distribution with zero mean value and σ^2 variance [14].

[1] Even though both the BS and the MS antenna arrays are dual-polarized, only the horizontal (H) polarization is used in the measurement data post-processing due to its larger receive power compared with the vertical (V) polarization.

3.2. Definition of DS

Generally speaking, the root mean square (RMS) of a given probability density function (PDF) $f(x)$ is computed as the second central moment of $f(x)$. The RMS can be expressed as

$$\sigma = \sqrt{\frac{\int_{x_{min}}^{x_{max}} (x-\overline{x})^2 f(x) d_x}{\int_{x_{min}}^{x_{max}} f(x) d_x}}, \tag{2}$$

where $\overline{x}$ is the first order moment of $f(x)$ and can be calculated as

$$\overline{x} = \frac{\int_{x_{min}}^{x_{max}} x f(x) d_x}{\int_{x_{min}}^{x_{max}} f(x) d_x}. \tag{3}$$

The DS is computed over the delay power spectral density or over the power delay profile (PDP).

3.3. Definition of XPR

XPR is defined as the ratio of the power between the co-polar link and the cross-polar link. Since the H-polarization is considered in the measurement data post-processing, the co-polar link stands for the HH-polarized link while the cross-polar stands for the HV-polarized link.

3.4. Extraction procedure of the LSP values from measurement data

To get the values of propagation loss, LSF, DS, and XPR from the measurement data, the same method and extraction procedure as presented in [10] are used in this paper. At first, a space-time averaging is performed to the PDP, squaring of the channel impulse response, to remove the small scale effect. In this paper, the time duration that the MS moves a distance of 10λ, corresponds to the time averaging length while the averaging over 8 BS antennas together with 24x4 MS antennas is the space averaging. After that, the measurement thermal noise is removed from the averaged PDP. The value of the averaged PDP at delay bin will be set to zero when it is smaller than the noise level margin. 9 [dB] is used as the threshold of the noise level margin. The averaged PDP after removal of noise is the basis for the calculation of the Propagation loss, LSF, DS, and XPR values. The sum of this PDP in delay domain is the propagation loss. After removal of the distance dependent path loss values from the propagation loss as described in [11], LSF values can be returned. The difference between the propagation losses of the HH-polarized link and the HV-polarized link is the XPR value. Substituting the averaged PDP after removal of noise into Eqn. 2, the DS value can be obtained.

4. EXPERIMENTAL RESULTS

4.1. Propagation loss

The experimental results of propagation loss from 4 BSs are presented in Fig. 3. Figure 3(a) shows the results along the main street in the LOS propagation condition while Fig. 3(b) presents the results along two perpendicular streets in the NLOS propagation condition.
By comparing Fig. 3(a) with Fig. 3(b), it can be observed that the MS undergoes significant propagation loss when it disappears from the main street and enters into the perpendicular streets. It is due to the disappearance of the LOS propagation path. This loss is defined as corner loss. It is around 20 [dB] in the studied environment.
By comparing the propagation losses between two BS sectors, it is found that the MS experiences almost the same propagation loss from two BS sectors both along the main street and along two perpendicular streets. Even though the BS antenna has been set to two sectors, the main street is located in the main beam of the both BS antenna sectors. This is the reason why there is no difference in propagation loss between two sectors along the main street. As the MS enters the perpendicular streets, the signal from the BS is reflected or diffracted to the perpendicular streets by scatters around the cross-road on the main street. These scatters work as virtual sources. Therefore, the MS receives also the same signal strength from two BS sectors along two perpendicular streets due to the same virtual source position on the main street, even though the perpendicular streets are located in the side beam of one BS sector but in the main beam of another BS sector.
By comparing the propagation loss results of 10 [m] and 16 [m] BS heights, it can be observed that the propagation loss with 16 [m] BS is higher than the propagation loss with 10 [m] BS both in the LOS and in the NLOS propagation conditions, but not at a remarkable degree. In the LOS propagation condition, the difference is higher when the MS is located near to the BS. It is due to the difference in propagation distance. As the MS moves aways from the BS, the 6 [m] BS height difference contributes fewer to the total final propagation distance. As a consequence, the propagation losses of two BS heights show more similarity.

4.2. Statistical distributions of the LSPs

Figure 2 shows the cdf curves of the LSPs from the measurement data for the LOS and NLOS propagations. As stated in Subsection 2, the height of the BS varies from 10 [m] to 16 [m] during the measurement campaign. Figure 2 shows the dependence of the LSPs on the BS height. It is observed in Fig. 2 that, both in the LOS and in the NLOS propagations, the higher the BS antenna array, the higher the LSF and DS values are, whereas, the lower the XPR values are. The reason is that the propagation between the BS and the MS varies from a two-ring or ellipse model to a one-ring model [12] [13] when the BS antenna array goes from below rooftop to above rooftop. When the BS antenna array is below rooftop, both the BS and the MS are surrounded by the local objects such as buildings and traffic lights. The wave propagation between the MS and the BS interacts with these objects. As a consequence, a two-ring or ellipse model is formed with the BS and the MS being the two centers of the two rings or two foci of the ellipse. The geometric range of the two-ring or ellipse model is limited by the local objects. As the BS height increases, the number of the local objects around the BS decreases. In the extreme case, there is no object around the BS when the BS is highly elevated. However, the MS is surrounded by objects which interact with the propagation waves. Furthermore, more objects are involved into the propagation including some far objects. Therefore, a one-ring model is formed between the MS and the BS with the MS being the center of the ring. The geometric range of the ring goes from the local objects to the far objects. Therefore, the propagation between the BS and the MS includes multipath components with long propagation delay which can not happen in the two-ring or ellipse model. As a consequence, both the LSF and the DS values are increased with increasing BS height and their cdf curves shift to the right hand side. Simultaneously, the additional multipath components with long propagation delay lead to the reduction of the XPR values.

4.3. Inter-sector correlation of the LSPs

Table 1 shows the experimental results of the inter-sector correlation of the LSPs. The sector pair $\langle BS2, BS5\rangle$ has 10 [m] BS height while the sector pair $\langle BS3, BS6\rangle$ has 16 [m] BS height. It is found in Table 1 that all LSPs are highly correlated in the overlapped area with correlation coefficient around 0.9. Furthermore, in the overlapped area, the inter-sector correlation coefficients of the LSPs with 16 [m] BS height are a little bit higher than these with 10 [m] BS height. The reason is that more common scatters will be involved into the propagation when the BS antenna array becomes higher. As a consequence, the similarity in the propagation from two sectors in the overlapped area is increased. However, the inter-sector correlation could not be one, even though in the overlapped area where botho sectors have main beam. The fact is that the MS can still receive the signal coming from one sector's main beam area which is outside the overlapped area by reflection /diffraction /refraction. This signal reflected /diffracted /refracted by the objects outside the overlapped main beam leads to the reduction of the inter-sector correlation. The larger the ratio of the overlapped main beam area to the whole main beam width of two sectors, the larger the inter-sector correlation is. The inter-sector correlation could be one only if two sectors are fully overlapped to each other.

see the program

Table 1. Inter-sector correlation of the LSPs with different BS heights

BS pair	BS height	LSF	DS	XPR
$\langle BS2, BS5\rangle$	10 [m]	0.81	0.78	0.84
$\langle BS3, BS6\rangle$	16 [m]	0.85	0.84	0.88

5. CONCLUSIONS

In this paper, based on the field measurement data collected in an urban area of Ilmenau, Germany, we studied the impact of the BS height on the propagation loss, on the statistical distributions of the LSPs, and on the inter-sector correlation properties of the LSPs. The experimental results show that the BS height has impact on the propagation loss in the case when the MS is located in the near of the BS. Both in the LOS and in the NLOS propagations, higher BS antenna array results in higher LSF and DS values but lower XPR values. The inter-sector correlation of the LSPs is around 0.9 in the overlapped area. Furthermore, the inter-sector correlation coefficients of the LSPs become higher with increasing BS height.

Acknowledgment

The authors would like to thank the colleagues from Technische Universitaet Ilmenau for their efforts in conducting the measurement campaign.

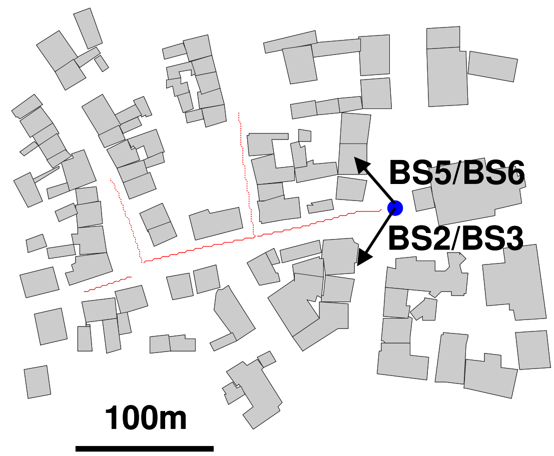


Fig. 1. Measurement routes in 2-dimensional Ilmenau city map

6. REFERENCES

[1] S. Saunders, *Antenna and propagation for communication systems concept and design*, Wiley, 1999.

[2] T. S. Rappaport, *Wireless Communications, Principles and Practice*, 2nd ed, Prentice Hall, New Jersey, 2002.

[3] M. Steinbauer, *The radio propagation channel - a non-directional, directional, and double-directional point-of-view*, Technische Universitt wien (TUW), Austria, PhD thesis, Sep. 2001.

[4] http://www.channelsounder.de/.

[5] www.ist-winner.org

[6] P. Kysti et al., *WINNER II Channel Models*, IST-4-027756 WINNER II D1.1.2 V1.2, Sep. 2007.

[7] R. Fraile , J. Monserrat, N. Cardona, and J. Nasreddine, *Impact of shadowing modelling on TD-CDMA system-level simulations*, In Proc.3rd International Symposium on Wireless Communication Systems (ISWCS), Valencia, Spain, Sep. 2006.

[8] A. Hong, G. Sommerkorn, R. S. Thomae, and W. Zirwas , *Considerations on the relationship between path loss and spatial characteristics based on MIMO measurements*, In Proc. ITG Workshop on Smart Antennas, Ulm, Germany, Mar. 2006.

[9] A. Hong, C. Schneider, G. Sommerkorn, M. Milojević, R. S. Thomae, and W. Zirwas , *Experimental Evaluation of Correlation Properties of Large Scale Parameters in Indoor Pico-cell Environments*, In Proc. 3rd International Symposium on Wireless Communication Systems (ISWCS), Valencia, Spain, Sep. 2006.

[10] A. Hong, M. Narandzic, C. Schneider, and R. S. Thomae, *Estimation of the correlation properties of large scale parameters from measurement data*, In Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Athens, Greece, Sep. 2007.

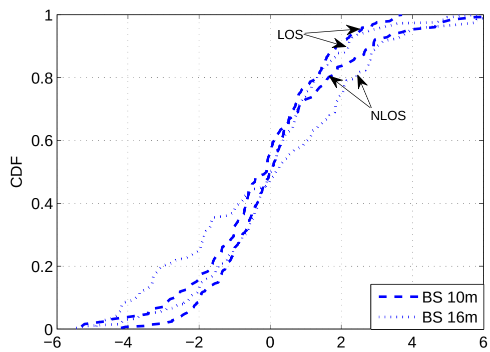
[11] N. Jalden, A. Hong, P. Zetterberg, P. Ottersten, and R. S. Thomae, *Correlation Properties of Large Scale Fading Based on Indoor Measurements*, In Proc. IEEE Wireless Communication and Networking Conference (WCNC), Hongkong, China, Mar. 2007.

[12] D-S. Shiu, G.J. Foschini, M.J. Gans, and J.M. Kahn, *Fading correlation and its effect on the capacity of multielement antenna systems*, IEEE Transactions on Communications, Vol. 48, Num. 3, pp. 502-513, Mar. 2000.

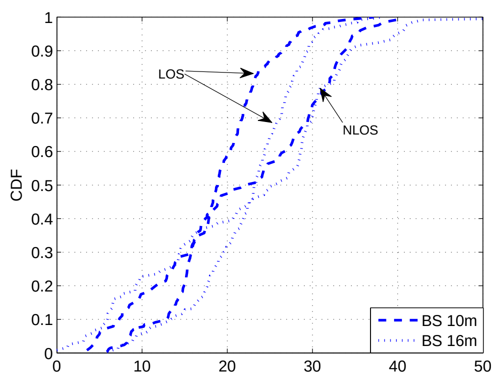
[13] Y. Kai, *Modeling of multiple-input multiple-output radio propagation channels*, Royal institute of Technology (KTH), Sweden, PhD thesis, Oct. 2002.

[14] J. Salo, L. Vuokko, and P. Vainikainen, *why is shadow fading lognormal?*, In Proc. the 8th International Symposium on Wireless Personal Multimedia Communications, Aalborg, Denmark, pp. 522-526, Sep. 2005.

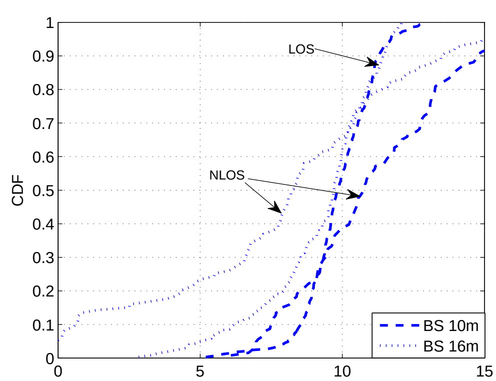
see the program



(a) LSF (dB)

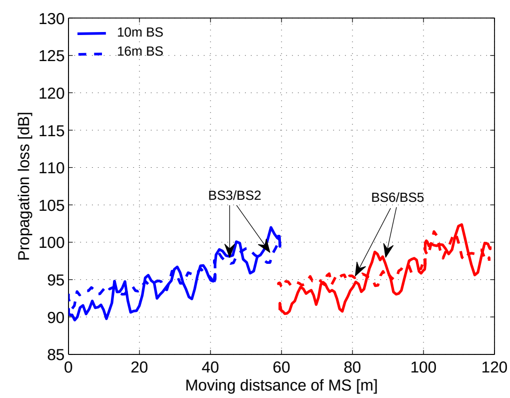


(b) DS (ns)

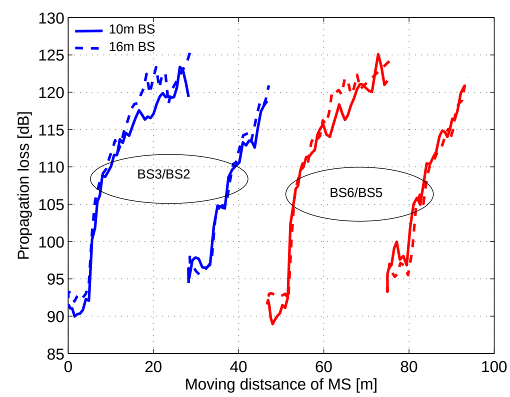


(c) XPR (dB)

Fig. 2. Cdf curves of the LSPs from the measurement data in the LOS and NLOS propagation conditions with two BS heights



(a) Propagation loss along the main street (LOS propagation condition)



(b) Propagation loss along two perpendicular streets (NLOS propagation condition)

Fig. 3. Propagation loss with two BS heights and two BS sectors in the LOS and NLOS propagation conditions

see the program

Performance of Printed Dipoles on Dual Band High-Impedance Surface

D. Kornek [1], S. Hampel , I. Kiral , I. Rolfes

Institut für Hochfrequenztechnik und Funksysteme, Leibniz Universität Hannover
Appelstr. 9A, 30167 Hannover, Germany
[1]kornek@hft.uni-hannover.de

***Abstract*— This paper investigates the characteristics of a printed dual band cross-dipole mounted on a High-Impedance Surface (HIS) designed for $f_{Low} = 2.4$ GHz and $f_{Hi} = 5.4$ GHz. The HIS operates as an artificial magnetic conductor for both frequencies by using a rectangular periodic structure at a fixed distance of 1 mm. Thus, the farfield for both frequencies of the cross-dipole radiates with an increased gain in the upper hemisphere compared to free space dipoles, whereas in the back of the antenna geometry electronical components can be placed.**

I. Introduction

During the last years the requirements on antenna designs have grown rapidly. On the one hand, especially in terms of wireless devices, a size reduction of the whole antenna is demanded to increase the number of antennas on the same area to cover various frequencies or to use diversity and MIMO concepts. On the other hand minimized coupling to the environment is desired for applications, in which the antenna is mounted on top of the system. In this case a significant front to back ratio of the antenna is required to shield the circuitry against unwanted radiation, which can be achieved by conventional microstrip antennas, e.g patch structures or antennas with complex substrates.

A novel possibility to fulfil the requirements is the use of High-Impedance Surfaces, also known as artificial magnetic conductors (AMC), in the reactive antenna near-field, which leads to both miniaturization and performance enhancement of the antenna in terms of gain and bandwidth, reported in [1] for a dipole and a patch antenna in combination with a High-Impedance Surface. Additionally, in [2] the shielding of the HIS is used to design a low profile antenna in a handset, whereby in [3] a bow-tie antenna in combination with an AMC is investigated for 4G communication services. As diversity schemes are mandatory in modern wireless systems, in [4] the diversity performance of a printed cross-dipole over a HIS is investigated. While these publications mainly discuss monoband applications the authors in [5] focus on design methodologies for multiband High-Impedance surfaces by modifying the shapes geometry.

This paper highlights the advantages due to the combination of a printed dual band cross-dipole with a dual band HIS. This design leads to an antenna supporting different wireless services including the advantages in terms of antenna size reduction, coupling and gain.

The article is organized as follows: Section II introduces the HIS design based upon the underlying theory and provides a dual band reflector for $f_{Low} = 2.4$ GHz and $f_{Hi} = 5.4$ GHz. Section III briefly describes the simulation setup of the complete antenna including HIS and presents results regarding antenna characterization. Conclusions are presented in Section IV.

II. Principle of Dual Band HIS

The HIS is a composite material consisting of a capacitive frequency selective surface (FSS) on a conductor backed dielectric substrate. In general, the textured surface be considered as a two-dimensional filter, which conducts DC currents but suppresses the propagation of AC currents within a forbidden frequency band. The periodic surface with a periodicity of a can be described by the transmission line theory assuming that the geometry parameters of a unit cell $a, g, t \ll \lambda_{eff}$ are much smaller than the effective wavelength of the design frequency, see Fig. 1 a).

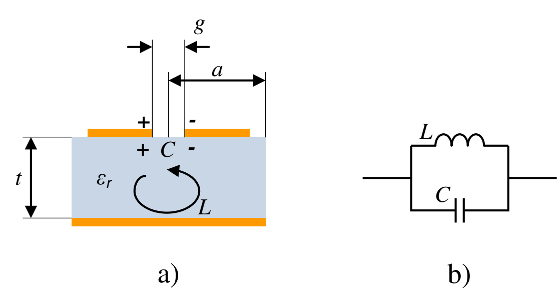


Fig. 1. Side view of two unit cells a), equivalent parallel resonant circuit b)

According to [6] the HIS can be simulated as a parallel resonant circuit with its resonant frequency $f_{res} = 1/\left(2\pi\sqrt{LC}\right)$ as depicted in Fig. 1 b), wherein the effective capacitance C and the effective inductance L depend on the gap width g, the substrate height t, the periodicity a and the effective dielectric constant ϵ_{eff}. The resonant frequency is determined by calculating the inductance and the capacitance based on the equations presented in [6] and [7]. At this frequency the extremely high surface impedance of the structure leads to a reflection coefficient of $\Gamma = 1$ and a reflection phase

see the program

of $\Phi_\Gamma = 0°$. Furthermore the reflection phase is used to define the operational bandwidth, while the corner frequencies correspond to values of $\Phi_\Gamma = \pm 90°$, see [6].
To obtain a planar dual band HIS for $f_{Low} = 2.4$ GHz and $f_{Hi} = 5.4$ GHz the independency of the geometric parameters in the orthogonal direction of the incident TE field reported in [7] is exploited, leading to a rectangular shape of the unit cell, see Fig. 2.

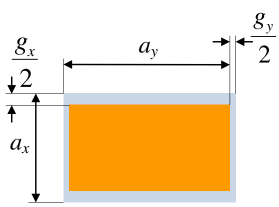


Fig. 2. Side view of rectangular shaped unit cells

The corresponding geometry parameters are evaluated for a substrate height of $t = 6$ mm and a permittivity $\epsilon_r = 3.55$ (Rogers4003) resulting in $a_x = 2.2$ mm, $g_x = 0.3$ mm, $a_y = 9.5$ mm and $g_y = 0.1$ mm. This unit cell is simulated with a commercial 3D fieldsover and the simulation results of the reflection phase are depicted in Fig. 2 for an incident TE wave parallel to the x- and y-axis, respectively.

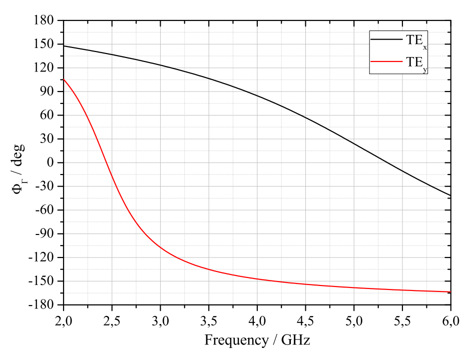


Fig. 3. Reflection phase of the rectangular shaped unit cell

At both frequency points $f_{Low} = 2.4$ GHz and $f_{Hi} = 5.4$ GHz the reflection phase of the structure equals $\Phi_\Gamma = 0°$, which enables a periodic surface consisting of these patches to be an ideal near field reflector for an antenna.

III. Antenna Setup and Simulation Results

In the following an antenna and the HIS are combined, in which the printed cross-dipole is mounted on a substrate with a height of $t_{Ant} = 1$ mm and a permittivity $\epsilon_r = 3.55$ for practical reasons, see Fig. 4. As this affects the effective medium of the HIS and modifies the ϵ_r used for the single HIS design, the geometry parameters have to be evaluated for the complete setup.

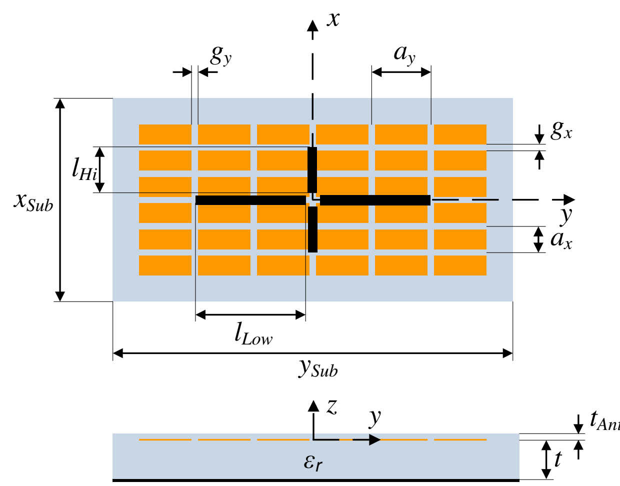


Fig. 4. Dipole on HIS

The resulting HIS consists of six patches in x- and y-direction and has maximal dimensions of $x_{sub} = 32.1$ mm and $y_{sub} = 68.6$ mm. Furthermore, the following geometric parameters $a_x = 3.3$ mm, $g_x = 0.3$ mm, $a_y = 9.7$ mm, $g_y = 0.1$ mm, $l_{Hi} = 7.35$ mm and $l_{low} = 19$ mm are used in the simulations. The corresponding input reflection coefficient S_{11} for both dipoles is shown in Fig. 5 and the resonance behavior of the complete setup approximately meets the design frequencies with $S_{11} \approx -11$ dB at $f_{Low} = 2.4$ GHz, see the red curve, and $S_{11} \approx -11$ dB at $f_{Hi} = 5.4$ GHz. The black curve yields a better S_{11} for higher frequencies with -22 dB at $f = 5.9$ GHz, what might be explained by the capacitive load of the dipoles. Additionally, the black curve shows a second resonance at $f = 4.25$ GHz with $S_{11} \approx -19$ dB.

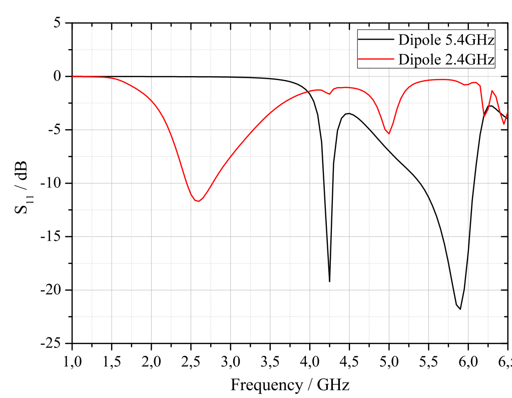


Fig. 5. Dipole on HIS

To prove the functionality of the HIS, in Fig. 6 and Fig. 7 the simulated realized gain for the design frequencies of the HIS and $f = 4.25$ GHz are depicted in the xz- and yz-plane.

see the program

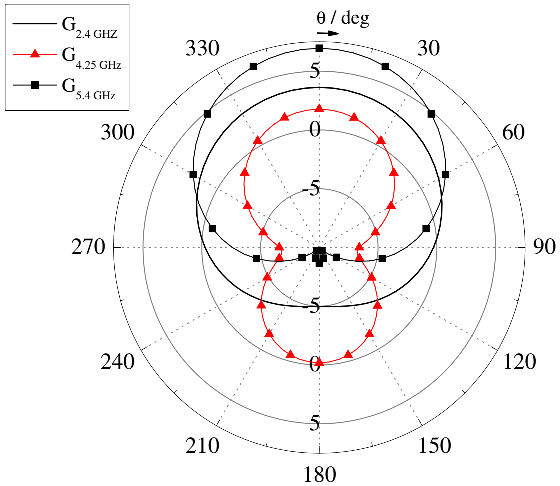


Fig. 6. Simulated realized gain in dBi in elevation, xz-plane

For the smaller dipole parallel to the x-axis, the gain reaches approximately 6.3 dBi at $f_{Hi} = 5.4$ GHz in z-direction, while the amount in the negative hemisphere is negligible, compare Fig. 6. For this frequency the shape of the gain pattern in xz-plane shows nearly the ideal constructive superposition caused by the reflecting HIS, while it is slightly disturbed in the yz-plane caused by the orthogonal dipole. At the lower resonance frequency $f_{Low} = 2.4$ GHz the gain in xz- and yz-plane has its maximum at $\theta = 0°$ with 3.8 dBi. Also, this curves indicate that the HIS operates well, as in negative z-direction the gain decreases heavily.

Finally, the gain pattern at $f = 4.25$ GHz will be discussed shortly, since the shape of the gain pattern in the yz-plane with nearly three equal maxima at $\theta = -90°, 0°, 90°$ leads to the assumption, that a patch mode radiates at this frequency.

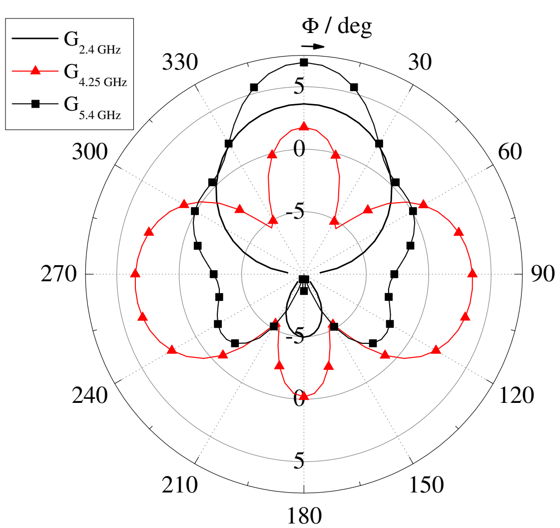


Fig. 7. Simulated realized gain in dBi in elevation, yz-plane

Comparing the radiation patterns of higher order modes of a rectangular patch with the curve in Fig. 7 results in a possible TM_{30} mode. Therefore, the wavelength in the effective medium for $f = 4.25$ GHz is estimated to approximately $\lambda = 38$ mm. The side length including all patches parallel to the y-axis is 58.2 mm, which equals nearly $\frac{3}{2}\lambda$ and for this reason confirms the patch theory. One possibility to suppress this parasitic mode can be to introduce vias from HIS to ground and will be investigated further.

IV. Conclusion

The paper presents a compact dual band antenna for WLAN applications consisting of a dual band HIS and a cross-dipole. The simulation results indicate, that the antenna offers a gain of 3.8 dBi at $f_{Low} = 2.4$ GHz and 6.3 dBi at $f_{Hi} = 5.4$ GHz in z-direction, which is achieved by adjusting the reflection phase of the HIS to this frequency points. Due to the size of the HIS a parasitic patch mode is excited at $f = 4.25$ GHz, which needs further investigation.

References

[1] H. Mosallaei and K. Sarabandi, "Antenna miniaturization and bandwidth enhancement using a reactive impedance substrate," *IEEE Trans. Antennas Propagat.*, vol. 52, pp. 2403–2414, Sept. 2004.
[2] R. F. J. Broas, D. F. Sievenpiper, and E. Yablonovitch, "A high-impedance ground plane applied to a cellphone handset geometry," *IEEE Trans. Antennas Propagat.*, vol. 49, pp. 1262–1265, July 2001.
[3] E. Carrubba, S. Genovesi, A. Monorchio, and G. Manara, "Amc-based low profile antennas for 4g communication services," in *Proc. IEEE Antennas and Propagation International Symposium APS'07)*, Honolulu, USA, June 2007, pp. 3364–3367.
[4] O. Schmitz, S. K. Hampel, I. Rolfes, and H. Eul, "Impact of sievenpiper high impedance surfaces on the performance of planar cross-dipole polarization diversity antennas," in *Proc. IEEE of the 10th European Conference on Wireless Technology (ECWT'07)*, Munich, Germany, Oct. 2007, pp. 118–121.
[5] D. Kern *et al.*, "The design synthesis of multiband artificial magnetic conductors using high impedance frequency selective surfaces," *IEEE Trans. Antennas Propagat.*, vol. 53, pp. 8–17, Jan. 2005.
[6] D. Sievenpieper *et al.*, "High-impedance electromagnetic surfaces with a forbidden frequency band," *IEEE Trans. Microwave Theory Tech.*, vol. 47, pp. 2059–2074, Nov. 1999.
[7] C. R. Simovski, P. de Maagt, and I. V. Melchakova, "High-impedance surfaces having stable resonance with respect to polarization and incidence angle," *IEEE Trans. Antennas Propagat.*, vol. 53, pp. 908–913, Mar. 2005.

see the program

PERFORMANCE AND MODELING OF LTE H-ARQ

Josep Colom Ikuno, Martin Wrulich, Markus Rupp

Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology, Austria
Gusshausstrasse 25/389, A-1040 Vienna, Austria
{jcolom, mwrulich, mrupp}@nt.tuwien.ac.at

ABSTRACT

The 3rd Generation Partnership Project (3GPP)'s Release'8 Long Term Evolution (LTE) defines the next step of 3G technology. LTE offers significant improvements over previous technologies such as UMTS/HSPA. Higher downlink and uplink speeds, lower latency and simpler network architecture are among the new features that are provided. One of the central features that provides transmission robustness is hybrid-ARQ, which in LTE provides physical layer retransmission using incremental redundancy and soft combining.

In this paper we propose a low-complexity model based on AWGN link level simulations that is capable of reliably predicting the BLER improvement due to the use of incremental redundancy H-ARQ in LTE.

1. INTRODUCTION

The increase of mobile data usage, fueled by bandwidth demanding applications such as mobile TV, Web 2.0 and online gaming has led to the development and standardization of UTRA Long Term Evolution (LTE) [1]. LTE is intended to be a mobile-communication system that can take the telecom industry into the 2020s. LTE leaves legacy technology behind and sports a radio interface being purely optimized for IP transmissions and offering operators a high spectrum flexibility, supporting from 1.4 MHz to 20 MHz bands [2].

Also the core network is focused on the packed-switched domain, migrating away from the circuit-switched network. The System Architecture Evolution (SAE) core network architecture significantly enhances the core network performance while ensuring interoperability with previous networks such as High Speed Packet Access (HSPA). Hence LTE will be very convincing for network operators that already have HSPA networks running.

Several techniques have been used in order to provide significant improvements over its predecessor technologies. Improvements on the air interface include OFDMA in the downlink, SC-OFDMA in the uplink, and Multiple-Input Multiple Output (MIMO) enhancements, while on the network side the SAE provides an all-IP network capable of supporting higher throughput and lower latency. To be specific, user plane latency has been reduced to below 5 ms while supporting speeds of up to 350 km/h (500 km/h in some cases). As a result of these improvements, LTE is capable of offering download speeds of up to 100 Mbps and upload speeds of up to 50 Mbps within a 20 MHz spectrum [2].

Transmission integrity is guaranteed by the application of an hybrid automatic repeat request (H-ARQ) retransmission scheme, very similar to the one utilized in High Speed Downlink Packet Access (HSDPA). The performance of the retransmission handling strongly influences the performance of LTE both on the physical layer [3] as well as from a network perspective.

For performance evaluations, in particular on system-level, low-complexity models for the physical layer procedures are needed to keep the simulation times manageable. H-ARQ — since positioned on layer 1— has been of interest for modeling already for quite some time. For other technologies, e.g. HSDPA such models already exist [4] and have been utilized in a variety of applications [5, 6].

In this paper we focus on deriving a low-complexity model capable of accurately predicting the H-ARQ gains on the physical layer. Our model can be easily utilized in system-level simulations [7], thus allowing for the investigation of retransmission strategies in LTE networks. The paper is organized as follows: in Section 2 the LTE downlink shared channel will be introduced, after which we explain the principles of H-ARQ in Section 3. With these findings, we are able to derive our model, describing the simulation results, and its application in Section 4. Finally, Section 5 concludes the paper.

2. LTE DOWNLINK SHARED CHANNEL STRUCTURE

The LTE Downlink Shared Channel (DL-SCH) uses OFDM to mitigate frequency selectivity and to multiplex the resources for multiple users. The spectrum is divided in 15 kHz subcarriers (a 7.5 kHz spacing is also possible) and arranged in groups of twelve contiguous subcarriers each forming a Resource Block (RB) which are the basic schedulable unit in

see the program

the frequency domain. For a 1.4 MHz downlink bandwidth space, six RBs are allocated, which we express as $N_{\mathrm{RB}}^{\mathrm{DL}} = 6$. Alternatively, for a bandwith of 20 MHz $N_{\mathrm{RB}}^{\mathrm{DL}} = 100$.

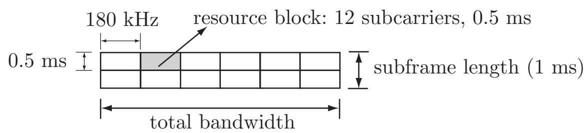


Fig. 1. Resource block grid structure (1 TTI) for a 1.4 MHz bandwidth.

In the time domain, the length of a radio frame is 10 ms, which is divided in 10 subframes, each 1 ms long. Each subframe consists of two 0.5 ms long slots. The Transmission Time Interval (TTI) of LTE is defined to be 1 ms long, so every 1 ms the scheduler must multiplex the users in grid of $N_{\mathrm{RB}}^{\mathrm{DL}} \times 2$ RBs [1]. This also sets 1 ms as the reaction time of the network. Fig. 1 depicts a resource block grid for a 1.4 MHz bandwidth.

The channel coding for the DL-SCH consists of Transport Block (TB) CRC attachment, Code Block (CB) segmentation, a rate 1/3 turbo encoder, rate matching and code block concatenation [8]. Fig. 2 depicts the processing structure for the DL-SCH, naming the bitstream after each process. Data a arrives to the coding unit in form of a maximum of one TB every TTI and a CRC is then attached to it, thus obtaining b. The resulting b bits are then segmented in r c_r CBs. The maximum CB size is 6144 bits, as that is the maximum size allowed by the turbo coder. In case segmentation is performed, an extra CRC is attached to c_r. The bits are then turbo coded and rate-matched, thus obtaining d_r and e_r respectively. Finally, the r CBs are concatenated to obtain the coded TB f.

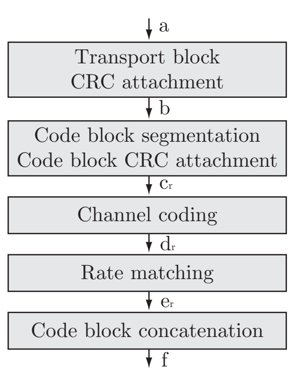


Fig. 2. The LTE DL-SCH channel coding processing structure. The data bits (a) are input and then coded (f).

Depending on channel conditions, different coding schemes and modulations can be used for each RB. Possible modulations for the DL-SCH are 4-QAM, 16-QAM and 64-QAM [9]. While the process is in essence identical to previous technologies such as HSDPA, due to the high data rates involved the structure of the turbo encoding and rate matching is slightly different such as to ease parallelization [10].

The LTE rate matcher will fit the rate 1/3 turbo-coded blocks into the available space assigned by the scheduler in the RB grid via a process of puncturing or repetition, thus obtaining the Effective Code Rate (ECR).

$$\mathrm{ECR} = \frac{c_r}{e_r} \cdot 1024. \tag{1}$$

LTE defines different Modulation and Coding Schemes (MCS) suitable for 15 different steps of the Channel Quality Indicator (CQI) which is feedback by the User Equipment (UE). ECRs between 78 and 948 are specified for LTE [11].

Fig. 3 depicts the rate-matching process for $e_r < d_r$ (ECR increase). The rate matcher plays a crucial role in H-ARQ, as it is performed by controlling the puncturing process depending on the retransmission number.

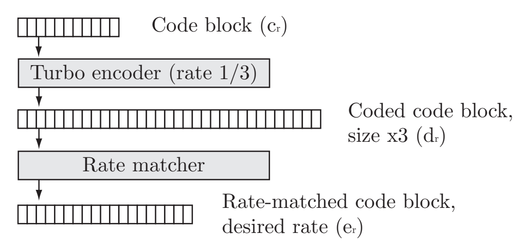


Fig. 3. Rate matching. Bits in the code block are punctured or repeated to achieve de desired ECR.

LTE provides both ARQ and H-ARQ functionalities. The ARQ functionality provides error correction by retransmissions in acknowledged mode at Layer 2, while the H-ARQ functionality ensures delivery between peer entities at Layer 1. In this paper we will focus on H-ARQ with Incremental Redundancy (IR) and how to model its performance in a simple way to use it in system level simulations.

Such models already exist for other technologies, such as the proposal in [4] for HSDPA and can improve the accuracy of current system level simulations such as [5], but so far none has been proposed for LTE.

3. LTE H-ARQ

A transmission scheme based on H-ARQ combines detection and Forward Error Correction (FEC) plus a retransmission of the erroneous packet. LTE additionally uses soft combining, in which a given received packet is combined with the previously received packets and the resulting more powerful FEC code is then decoded [12].

When soft combining the packets received in several H-ARQ retransmissions, basically two combining strategies can be used. These are Chase Combining (CC), where each retransmission is identical to the original transmission, and Incremental Redundancy (IR), where each retransmission con-

see the program

sists of new redundancy bits from the channel encoder. In IR, instead of sending simple repeats of the coded data packet, new information is sent in each subsequent transmission of the packet. The decoder then combines all the transmissions and decodes the packet at a lower code rate.

LTE utilizes IR H-ARQ with a 1/3 turbo encoder as FEC code and the TB CRC for error detection. Since in this case the receiver obtains just differently punctured versions of the same turbo-encoded data, each of these retransmission is self-decodable. Such an H-ARQ scheme falls into the category of a type-III Hybrid H-ARQ.

In order for IR to work, the H-ARQ functionality must be able to create appropriate redundancy versions from a given CB and prevent terminal buffer overflow. This is achieved through the rate matcher located after the fixed-rate channel encoder. LTE implements this functionality in a one-step rate matching as opposed to HSDPA, where it is accomplished in a two-step process [13].

Fig. 4 depicts CC in the case when ECR $> 1024/3$, where the received identical packets are recombined, while Fig. 5 depicts the IR case. The white part in the reconstructed code block represents bits that were punctured from the original CB.

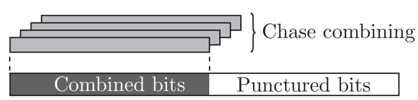


Fig. 4. Chase combining on four received transmissions $\left(\mathrm{ECR} > \frac{1024}{3}\right)$. Punctured bits are still missing after the combining.

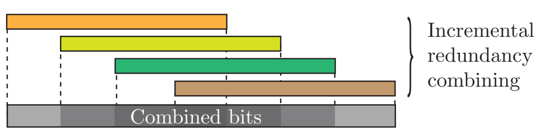


Fig. 5. Incremental redundancy combining on four received transmissions $\left(\mathrm{ECR} > \frac{1024}{3}\right)$. Every retransmission adds extra information bits.

While HSDPA provides both CC and IR H-ARQ, LTE only provides IR. IR can give large performance gains for high channel-coding rates and high modulation orders. For low MCSs, the link-level performance gains with IR are less significant, and there are only a few artificial scenarios in which IR performs worse than CC [14].

H-ARQ transmissions are indexed with the redundancy version $\mathrm{rv_{idx}}$ parameter, which tells the receiver whether the currently transmitted TB is new (0), or the n-th retransmission, up to a maximum of 3 (1,2 or 3). For a given target TB size G, the rate matcher can produce 4 different punctured versions of the original coded TB, depending on the value of $\mathrm{rv_{idx}}$ (see Fig. 6). Other parameters that fine tune the rate matching process are Q_m (modulation order), N_{IR} (buffer size at receiver end) and N_l (number of layers used in the channel coding) [8].

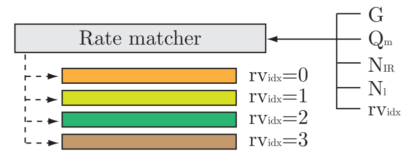


Fig. 6. Through the rate-matching process, up to four different rate-matched versions of a given code block can be produced by varying the $\mathrm{rv_{idx}}$ parameter.

Link level simulations with H-ARQ showed that the LTE turbo code rate matcher will even in the case of most extreme puncturing (CQI $= 15$) still introduce 13% of parity bits for $\mathrm{rv_{idx}} = 0$. For the second, third and fourth transmissions, these ratios change to 83%, 0% and 34% respectively (see Fig. 7).

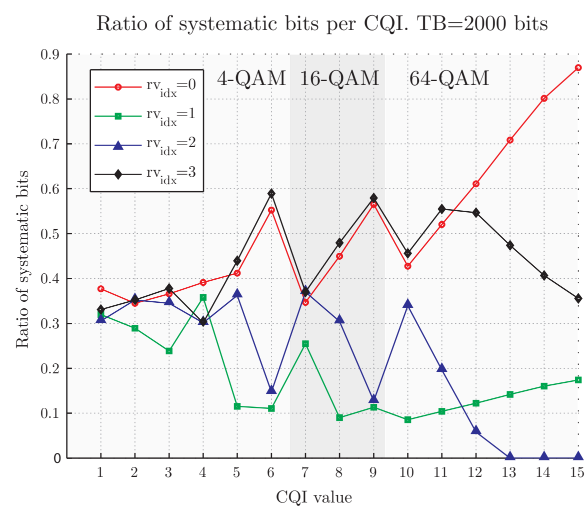


Fig. 7. Ratio of systematic bits in the rate-matched CB for each possible MCS, as defined by the CQI.

When using IR, some degradation can be expected when losing the systematic bits in the original transmission while receiving the parity bits [14]. Including systematic bits in two out of three possible retransmissions aims at reducing this problem. In order to improve the efficiency of H-ARQ, mapping techniques such as the one in [15] have also been used. The scheme used in LTE, depicted in Fig. 8, is based on transmitting systematic and parity bits in the same sent TB.

The rate matching process works by, depending on $\mathrm{rv_{idx}}$, choosing a starting point in the interleaved CB depicted in Fig. 8 and then choosing G consecutive bits. The $v^{(0)}$ bits are the interleaved systematic bits from the turbo encoder, while $v^{(1)}$ and $v^{(2)}$ the first and second parity bits respectively. This

see the program

Fig. 8. Coded bits before the puncturing/repetition in the rate matcher. A contiguous subset of these bits is the output of the rate matcher.

sub-block interleaving process is part of the rate-matching process, and separately spreads the systematic and parity bits over a wider area in the CB.

4. H-ARQ MODELING

For each H-ARQ retransmission that the LTE system can use, an improvement of the Block Error Rate (BLER) is expected. In order to model these improvements, link level simulations have been used [3].

The LTE link level simulations for the H-ARQ evaluation process were performed for a single-user scenario corresponding to the simulation parameters shown in Table 1. All of the 15 MCS defined by the CQI values in the LTE standard [11] have been used. As to have reliable data at BLERs of 10^{-3}, 10,000 subframe-long simulations were performed. MIMO benefits are open for investigation in future work, where we will put an emphasis on investigating different space-time coding schemes (similar to HSDPA [16]), as proposed by the standard.

Parameter	Value
System bandwidth	1.4 MHz
Subcarrier spacing	15 kHz
Cyclic prefix	normal
Channel profile	AWGN
Max. number of retransmissions	3
CQI values	1-15
Simulation length	10,000 subframes

Table 1. Link level simulation parameters.

Simulations were performed with the minimum specified bandwidth for LTE, which is 1.4 MHz. Since the LTE turbo encoder has a maximum CB size of 6144 bits, the channel coding process segments the received TBs when needed. Because of this segmentation, it is not expected that varying the available bandwidth (thus increasing $N_{\mathrm{RB}}^{\mathrm{DL}} = 6$) would change the results presented in this paper.

Fig. 9 shows typical BLER curves generated by the H-ARQ simulations. Each curve represents the BLER of the subframes with a specific $\mathrm{rv}_{\mathrm{idx}}$. The bold curve represents the overall BLER, not taking into account $\mathrm{rv}_{\mathrm{idx}}$. As each retransmission provides the receiver extra information, BLER improves for each additional retransmission. Our proposed H-ARQ model serves at estimating the BLER SNR shift that occurs when H-ARQ is used compared to a transmission that is not using H-ARQ.

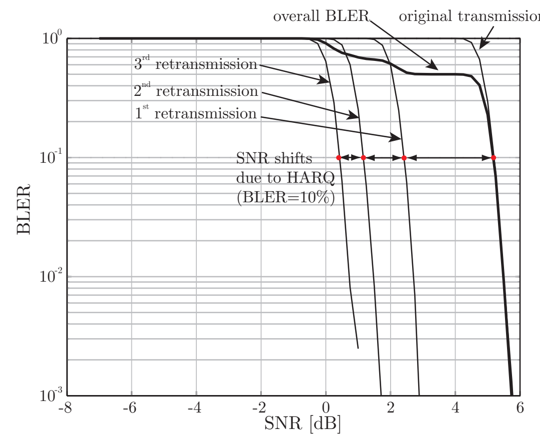


Fig. 9. H-ARQ BLER curves for CQI 7. 16-QAM modulation and 738 ECR.

Simulations were performed for the MCSs corresponding to each CQI value in LTE. Fig. 9 shows only the plot for a CQI of seven, but simulations were performed for all 15 CQI values. From each of these simulations the BLER for each retransmission was evaluated and the shifts between the curves (SNR gain due to H-ARQ) was noted. The shift was defined as the difference between the curve of the n-th retransmission and the original transmission at the $\mathrm{BLER} = 10\%$ point. The results of our investigation are illustrated in Fig. 10, which depicts the SNR gain depending on the number of retransmissions ($\mathrm{rv}_{\mathrm{idx}}$), the modulation (4-QAM, 16-QAM or 64-QAM) and the ECR. A linear regression has been performed to fit the simulation results, which is also overlayed onto the plot.

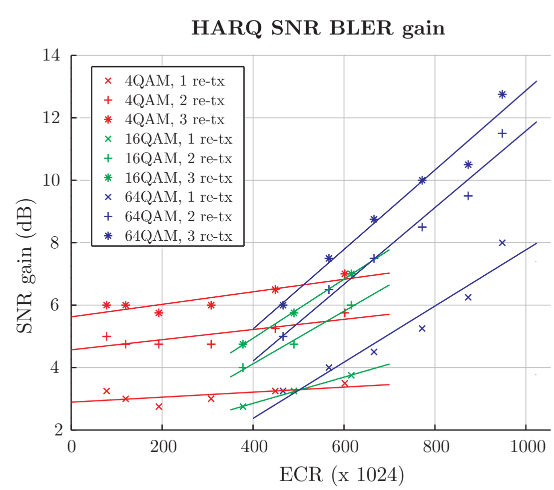


Fig. 10. SNR gain when using H-ARQ and the proposed analytical model.

The SNR gain can thus then be modeled this linear fit, shown in Eq. (2), where $\mathrm{SNR}_{\mathrm{gain}}(\mathrm{rv}_{\mathrm{idx}})$ is expressed in dB. The possible values for μ_{mod} and $\varepsilon_{\mathrm{mod}}$, depending the retransmission index $\mathrm{rv}_{\mathrm{idx}}$ and the modulations specified by the LTE standard can be found in Table 2.

$$\mathrm{SNR}_{\mathrm{gain}}(\mathrm{rv}_{\mathrm{idx}}) = \mu_{\mathrm{mod}}(\mathrm{rv}_{\mathrm{idx}}) \cdot \mathrm{ECR} + \varepsilon_{\mathrm{mod}}(\mathrm{rv}_{\mathrm{idx}}) \quad (2)$$

Modulation	$\mathrm{rv}_{\mathrm{idx}}$	$\mu_{\mathrm{mod}} \cdot 10^{-2}$	$\varepsilon_{\mathrm{mod}}$
4-QAM	1	0.0804	2.89
	2	0.1628	4.57
	3	0.2006	5.62
16-QAM	1	0.0420	1.17
	2	0.8435	0.74
	3	0.9464	1.15
64-QAM	1	0.8996	-1.23
	2	1.2288	-0.71
	3	1.2728	0.15

Table 2. H-ARQ model parameters for 4-QAM, 16-QAM and 64-QAM modulations.

When using our proposed model in a system level simulation, the link level model can then adjust the post-equalization symbol SINR to account for the effect of H-ARQ by means of a simple sum, as shown in Eq. (3), where i denotes the i-th retransmission.

$$\mathrm{SINR}(i) = \mathrm{SINR} + \mathrm{SNR}_{\mathrm{gain}}(i) \quad (3)$$

In system level simulations, the SINR of a transmission without taking H-ARQ into account is calculated by the link measurement model. Afterwards, in order to account for the effect of H-ARQ, a gain of $\mathrm{SNR}_{\mathrm{gain}}(i)$ dBs will be applied to the SINR according to the redundancy version index i and using the correct μ_{mod} and $\varepsilon_{\mathrm{mod}}$ parameters for the given MCS (see Eq. 2), thus obtaining an adjusted SINR value that includes the gain due to H-ARQ retransmissions.

Typical link to system level models utilized for the evaluation of multicarrier systems such as LTE are Exponential Effective SIR Mapping (EESM) and Packet Error Rate (PER) indicator [17]. Such models can be used to obtain the effective SINR of the received TB to which afterwards the H-ARQ gains can then be added.

Although there are already extensions for EESM to deal with H-ARQ [18] showing results for uncoded simulations, our proposed simulation-based model takes into account the whole LTE processing chain, providing a more tailored model for LTE. Details such as the number of new and systematic bits in each retransmission and the performance of the channel coding, which are specific to the LTE system are then taken into account.

Our model shows that, as expected, the highest SNR gains are observed for high modulation orders and ECRs, as IR performs at best in this region [14] (see Fig. 10). Up to a 12.75 dB improvement has been observed for CQI 15 (64-QAM, 948 ECR).

The relative mean square error (MSE) of the estimated SNR gains is shown in Table 3). As it can be seen, relative MSEs show errors below 1% except for the 64-QAM modulation. This high error is due to a mismatch of the model with the actual measured SNR gains for the last two CQI values (873 and 948 ECRs).

Modulation	$\mathrm{rv}_{\mathrm{idx}} = 1$	$\mathrm{rv}_{\mathrm{idx}} = 2$	$\mathrm{rv}_{\mathrm{idx}} = 3$
4-QAM	0.374	0.717	0.805
16-QAM	0.002	0.133	0.016
64-QAM	3.983	6.569	11.273

Table 3. Relative MSE $(\%)$ of the proposed LTE H-ARQ model for the different possible modulations and $\mathrm{rv}_{\mathrm{idx}}$.

5. CONCLUSIONS

In this paper, we propose a low-complexity H-ARQ SNR-gain model that is capable of predicting the SNR gain that using the H-ARQ scheme defined for LTE provides. For any of the defined modulation and coding schemes defined for LTE, our model is capable predicting the SNR shift of the block error rate curve that will be obtained for each H-ARQ retransmission.

The proposed model can be used in system level simulations to represent the impact of H-ARQ not only accurately but also with a very low computational complexity. Future work will deal with improving the model for high modulation and coding schemes, frequency selective channels, and MIMO benefits and channels such as the Spatial Channel Model (SCM) [19].

Acknowledgment

The authors would like to thank mobilkom Austria AG for supporting their research and C. Mehlführer for numerous discussions and helpful suggestions. The views expressed in this paper are those of the authors and do not necessarily reflect the views within mobilkom Austria AG.

6. REFERENCES

[1] *Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Physical Layer - General Description*, 3GPP Std. TS 36.201. Release 8, November 2007.

[2] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G Evolution*. Academic Press, July 2007.

see the program

[3] C. Mehlführer, M. Wrulich, J. C. Ikuno, D. Bosanska, and M. Rupp, "Simulating the long term evolution physical layer," in *Proc. of the 17th European Signal Processing Conference (EUSIPCO 2009)*, Glasgow, Scotland, Aug. 2009, submitted.

[4] F. Frederiksen and T. Kolding, "Performance and modeling of WCDMA/HSDPA transmission/H-ARQ schemes," *Vehicular Technology Conference, 2002. Proceedings. VTC 2002-Fall. 2002 IEEE 56th*, 2002.

[5] M. Wrulich, W. Weiler, and M.Rupp, "HSDPA performance in a mixed traffic network," in *Proc. IEEE Vehicular Technology Conference (VTC) Spring 2008*, May 2008, pp. 2056–2060.

[6] M. Wrulich, S. Eder, I. Viering, and M. Rupp, "Efficient link-to-system level model for MIMO HSDPA," in *Proc. of the 4th IEEE Broadband Wireless Access Workshop*, 2008.

[7] J. C. Ikuno, M. Wrulich, and M. Rupp, "Simulating long term evolution networks," 2009, to be submitted.

[8] *Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding*, 3GPP Std. TS 36.212. Release 8, March 2008.

[9] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation*, 3GPP Std. TS 36.211. Release 8, March 2008.

[10] A. Nimbalker, Y. Blankenship, B. Classon, and T. Blankenship, "PHY 40-1 - ARP and QPP interleavers for LTE turbo coding," *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*, 31 2008-April 3 2008.

[11] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures*, 3GPP Std. TS 36.213. Release 8, May 2008.

[12] E. Malkamaki, D. Mathew, and S. Hamalainen, "Performance of hybrid ARQ techniques for WCDMA high data rates," *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, 2001.

[13] M. Dottling, J. Michel, and B. Raaf, "Hybrid ARQ and adaptive modulation and coding schemes for high speed downlink packet access," *Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on*, Sept. 2002.

[14] P. Frenger, S. Parkvall, and E. Dahlman, "Performance comparison of HARQ with chase combining and incremental redundancy for HSDPA," *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, 2001.

[15] Panasonic, "Proposal of bit mapping for type-III HARQ," Panasonic, Tech. Rep. input paper TSC+R1#18(01)0031, January 2001.

[16] M. Wrulich and M. Rupp, "Efficient link measurement model for system level simulations of Alamouti encoded MIMO HSDPA transmissions," in *International ITG Workshop on Smart Antennas (WSA)*, Darmstadt, Germany, Feb. 2008.

[17] Y. Blankenship, P. Sartori, B. Classon, V. Desai, and K. Baum, "Link error prediction methods for multicarrier systems," *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, Sept. 2004.

[18] B. Classon, P. Sartori, Y. Blankenship, K. Baum, R. Love, and Y. Sun, "Efficient OFDM-HARQ system evaluation using a recursive EESM link error prediction," *Wireless Communications and Networking Conference, 2006. WCNC 2006. IEEE*, April 2006.

[19] D. Baum, J. Hansen, and J. Salo, "An interim channel model for beyond-G systems: extending the 3GPP spatial channel model (SCM)," *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, May-1 June 2005.

see the program

ON THE CHARACTERIZATION OF MU-MIMO CHANNELS

Florian Kaltenberger [1], *Laura Bernadó* [2], *Thomas Zemen* [2]

[1]Eurecom, 2229, Route des Cretes - B.P. 193, 06904 Sophia Antipolis, France
[2]ftw. Forschungszentrum Telekommunikation Wien, Donau-City-Strasse 1, 1220 Wien, Austria
Email: florian.kaltenberger@eurecom.fr, laura.bernado@ftw.at, thomas.zemen@ftw.at

ABSTRACT

In this work we study the divergence of different links in wide-band multi-user multiple-input multiple-output (MU-MIMO) channels. The divergence is measured on several levels: (i) spatial separation of the user's correlation matrices, (ii) co-linearity of the MIMO channel matrices, and (iii) correlation of large scale fading. The measurement data has been acquired using Eurecom's MIMO Openair Sounder (EMOS). The EMOS can perform real-time MIMO channel measurements synchronously over multiple users. For this work we have used an outdoor measurement with two transmit antennas and two users with two antennas each. Several measurements with different distances between users were acquired. We find that the structure of the MIMO channel matrices changes significantly with the inter-user distance. This is best captured by the co-linearity measure. The transmit and the full correlation matrix also show some dependence on the inter-user distance whereas the receive correlation matrices are independent of the inter-user distance. The shadowing correlation was found to be very low in all cases. These findings are important for MU-MIMO precoding and scheduling algorithms.

1. INTRODUCTION

In a cellular network, cooperation between users can be used to greatly increase power efficiency, reliability and throughput. Cooperation can be achieved by using the antennas of multiple users to form a virtual antenna array and by using MIMO transmission/reception techniques. The development and realistic performance assessment of such distributed MIMO systems requires measurement and characterization of the different channel links in these systems. To this end, only a limited amount of channel measurements and analysis of such distributed MIMO systems are available.

This work was supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM++ (contract n. 216715), the FP7 project SENDORA (contract n. 216076) and Eurecom, as well as the Vienna Science and Technology Fund (WWTF) in the ftw. project COCOMINT. The Telecommunications Research Center Vienna (ftw.) is supported by the Austrian Government and the City of Vienna within the competence center program COMET.

In [1] realistic MU-MIMO channel measurements have been obtained using Eurecom's MIMO Openair Sounder (EMOS). The EMOS can perform real-time channel measurements synchronously over multiple users moving at vehicular speed. The measured channels are used to calculate the capacity of the MU-MIMO broadcast channel. One of the findings of [1] was that the performance of MU-MIMO precoding drops drastically when the users are close together in an outdoor scenario. It was further noted that this decline in performance is due to the strong correlation at the transmitter.

In this paper we investigate this phenomenon further by studying the distance of the correlation matrices with respect to the inter-user distance. Different measures to characterize the divergence of MIMO channels are available in the literature. Some of them can be applied directly on the MIMO channel matrices, while others are only applicable to correlation matrices. In this work we use the co-linearity measure applied on the channel matrices and the correlation matrix distance [2] as well as the geodesic distance [3] applied on the transmit, the receive and the full correlation matrix.

Another important phenomenon studied in this paper is the large scale shadow fading correlation. Shadowing correlation has a big impact on the performance of such cooperative communication schemes. Therefore, it is necessary to identify scenarios where shadowing correlation occurs.

An alternative correlation measure is the spectral divergence (SD) [4]. It measures the distance between strictly positive, non-normalized spectral densities. The SD was used in [5] to characterize the similarity between scattering functions of different links in a MIMO channel and in [6] to characterize the similarity between local scattering functions of a time- and frequency-selective vehicular channel. As an addition to the matrix distances we also evaluate the applicability of the SD to MU-MIMO channels.

Related work. In [7], measurements were conducted using a MEDAV-LUND channel sounder with its corresponding receiver as well as the receiver of an Elektrobit channel sounder. The two receivers are perfectly synchronized. The authors present capacity with interference results, based on the dynamic multilink measurements, as well as path-loss and delay spreads for the measured scenarios. Distributed MIMO measurements have also been described in [8]. They were

see the program

conducted with a single RUSK channel sounder using long cables between the antennas and the channel sounder. In [9] these measurements were used to characterize the spatial separation of MU-MIMO channels. The shadowing correlation between users has been studied in several papers [10–15]. However, a clear dependence of the shadowing correlation on the user distance can not be deduced.

Contribution of the paper. We show how the structure of the MIMO channel matrices changes with the inter-user distance. Further we show that the transmit and the full correlation matrix also depend on the inter-user distance whereas the receive correlation matrices are independent of the inter-user distance.

2. DESCRIPTION OF THE MEASUREMENT PLATFORM

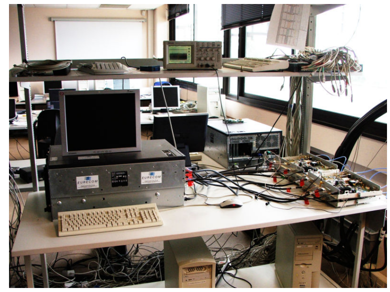
2.1. Hardware Description

The EMOS is based on the OpenAirInterface[1] hardware/software development platform at Eurecom. The platform consists of a BS that continuously sends a signaling frame, and one or more UEs that receive the frames to estimate the channel. The BS consists of a workstation with four PCI baseband data acquisition cards, which are connected to four PLATON RF boards (see Fig. 1(a)). The RF signals are amplified and transmitted by a Powerwave 3G broadband antenna (part no. 7760.00) composed of four elements which are arranged in two cross-polarized pairs (see Fig. 1(b)). The UEs consist of a laptop computer with Eurecom's dual-RF CardBus/PCMCIA data acquisition card (see Fig. 1(c)) and two clip-on 3G Panorama Antennas (part no. TCLIP-DE3G, see Fig. 1(d)). The platform is designed for a full software-radio implementation, in the sense that all protocol layers run on the host PCs under the control of a Linux real time operation system.

2.2. Sounding Signal

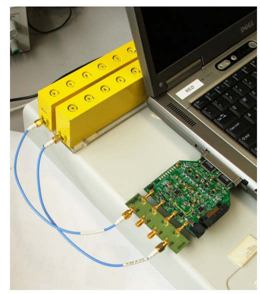
The EMOS is uses an OFDM modulated sounding sequence with 256 subcarriers (out of which 160 are non-zero) and a cyclic prefix length of 64. One transmit frame is 64 OFDM symbols (2.667 ms) long and consists of a synchronization symbol (SCH), a broadcast data channel (BCH) comprising 7 OFDM symbols, a guard interval, and 48 pilot symbols used for channel estimation (see Fig. 2). The pilot symbols are taken from a pseudo-random QPSK sequence defined in the frequency domain. The subcarriers of the pilot symbols are multiplexed over the M transmit antennas to ensure orthogonality in the spatial domain. We can therefore obtain one full MIMO channel estimate for one group of M subcarriers. The BCH contains the frame number of the transmitted frame that is used for synchronization.

[1] http://www.openairinterface.org

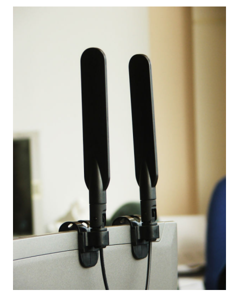


(a) Base Station with PLATON boards

(b) Power Amplifiers and Powerwave Antenna



(c) Dual-RF CardBus/PCMCIA Card



(d) Panorama Antennas

Fig. 1. EMOS base-station and user equipment [16]

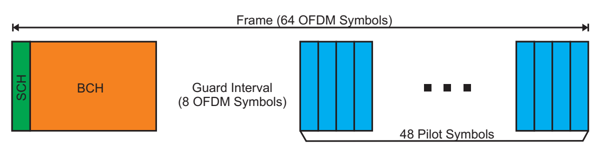


Fig. 2. Frame structure of the OFDM Sounding Sequence. The frame consists of a synchronization channel, (SCH), a broadcast channel (BCH), and several pilot symbols used for channel estimation.

2.3. Channel Estimation Procedure

Each UE first synchronizes to the BS using the SCH. It then tries to decode the data in the BCH. If the BCH can be decoded successfully, i.e., the cyclic redundancy check (CRC) is positive, then the channel estimation procedure is started. The channel estimation procedure consists of two steps. Firstly, the pilot symbols are derotated with respect to the first pilot symbol to reduce the phase-shift noise generated by the dual-RF CardBus/PCMCIA card. Secondly, the pilot symbols are averaged to increase the measurement SNR. The estimated MIMO channel is finally stored to disk. For a more detailed description of the synchronization and channel estimation procedure see [16, 17].

see the program

2.4. Multi-user Measurement Procedure

In order to conduct multi-user measurements, all the UEs need to be frame-synchronized to the BS. This is achieved by storing the frame number encoded in the BCH along with the measured channel at the UEs. This way, the measured channels can be aligned for later evaluations. The frame number is also used to synchronize the data acquisition between UEs. One measurement is 50 sec long.

3. POST PROCESSING AND PARAMETER EXTRACTION

The recorded measurement data of all users are normalized before further processing. In this section we describe the normalization, the calculation of the delay profile (PDP) and the correlation matrices. Further we define the different distance measures, i.e., the geodesic distance, the correlation matrix distance, the co-linearity and the spectral divergence.

3.1. Normalization

One measurement results in the set of MIMO matrices $\{\mathbf{H}_{k,m,q} \in \mathbb{C}^{N\times M}; k = 0,\dots,K-1; m = 0,\dots,N_\mathrm{F}-1; q = 0,\dots,Q-1\}$, where k denotes the user index, m the snapshot index, and q the frequency (or subcarrier) index. N, M, and K are the number of receive antennas, number of transmit antennas and number of users respectively. N_F is the total number of snapshots per measurement after removing erroneous frames (on average $N_\mathrm{F} \approx 18.000$). The total number of channel estimates in the frequency domain is given by $Q = 160/M$, since there are 160 subcarriers in total and the pilots are multiplexed over the M transmit antennas. The MIMO matrices are normalized by

$$\mathbf{H}'_{k,m,q} = \mathbf{H}_{k,m,q}\sqrt{\frac{N N_\mathrm{F} Q}{\sum_{m,q}\|\mathbf{H}_{k,m,q}\|_F^2}} \tag{1}$$

such that $\mathbb{E}\{\|\mathbf{H}'_k\|_F^2\} = N$.

3.2. PDP Estimation

The power delay profile (PDP) is estimated by averaging the channel impulse responses of every link of the MU-MIMO channel over $A = 200$ consecutive snapshots (this corresponds to a movement of the user of apx. 4λ at the maximum speed of 5km/h). We thus introduce a new time variable $n = \lfloor m/A \rfloor$ and write

$$P_{i,j,k}[n,\tau] = \frac{1}{A}\sum_{m=An}^{A(n+1)-1} |h_{i,j,k,m,\tau}|^2, \tag{2}$$

where $h_{i,j,k,m,\tau}$ is the (i,j)-th element of the time-delay domain MIMO matrix $\mathbf{H}_{k,m,\tau}$ of user k at time m.

3.3. Correlation Matrices

The correlation matrices are also estimated by averaging over $A = 200$ consecutive snapshots and all the frequency bins. The *per-user* transmit, receive, and full correlation matrices of the MU-MIMO channel are defined as

$$\mathbf{R}_\mathrm{Tx}^{(k)}[n] = \frac{1}{AQ}\sum_{m=An}^{A(n+1)-1}\sum_{q=0}^{Q-1}\{\mathbf{H}_{k,m,q}^H\mathbf{H}_{k,m,q}\}, \tag{3}$$

$$\mathbf{R}_\mathrm{Rx}^{(k)}[n] = \frac{1}{AQ}\sum_{m=An}^{A(n+1)-1}\sum_{q=0}^{Q-1}\{\mathbf{H}_{k,m,q}\mathbf{H}_{k,m,q}^H\} \tag{4}$$

$$\mathbf{R}^{(k)}[n] = \frac{1}{AQ}\sum_{m=An}^{A(n+1)-1}\sum_{q=0}^{Q-1}\{\mathrm{vec}(\mathbf{H}_{k,m,q})\,\mathrm{vec}(\mathbf{H}_{k,m,q})^H\}. \tag{5}$$

3.4. Correlation Matrix Distances

Correlation matrices are by definition Hermitian and positive definite. The space of Hermitian and positive definite matrices forms a convex cone [3]. A natural distance measure on this cone is given by the geodesic distance

$$d_\mathrm{Geod}(\mathbf{R}^{(k_1)},\mathbf{R}^{(k_2)}) = \left(\sum_i |\log\lambda_i|^2\right)^{1/2}, \tag{6}$$

where λ_i are the eigenvalues of $(\mathbf{R}^{(k_1)})^{-1}\mathbf{R}^{(k_2)}$ [3]. This distance measure has been successfully used in [18] to derive a differential limited feedback scheme for MIMO communications.

Another distance for correlation matrices was introduced by [2]. It is given by

$$d_\mathrm{Corr}(\mathbf{R}^{(k_1)},\mathbf{R}^{(k_2)}) = \frac{\mathrm{tr}(\mathbf{R}^{(k_1)}\mathbf{R}^{(k_2)})}{\|\mathbf{R}^{(k_1)}\|_F\|\mathbf{R}^{(k_2)}\|_F}. \tag{7}$$

It becomes one if the correlation matrices are equal up to a scaling factor and zero if they differ to a maximum extent.

3.5. Channel Matrix Distance Measures

Instead of looking at the distance of correlation matrices we can also find distance measures that apply to the channel matrices directly. A simple distance measure can be derived from the zero forcing (ZF) precoder used in MU-MIMO communications [1, 19]. To keep things simple, we will use only one antenna at the receivers ($N = 1$) and denote the corresponding channel of user k with the vector $\boldsymbol{h}_k$. If the channels of two users are orthogonal, i.e., if $\boldsymbol{h}_k\boldsymbol{h}_j^H = 0$, the ZF precoder achieves the optimal sum rate. The more aligned the vectors are, i.e, the larger the scalar product $\boldsymbol{h}_k\boldsymbol{h}_j^H$, the worse the performance of ZF will be. We thus define the co-linearity measure as

$$d_\mathrm{Colin}(\boldsymbol{h}_{k_1},\boldsymbol{h}_{k_2}) = \frac{\boldsymbol{h}_{k_1}\boldsymbol{h}_{k_2}^H}{\|\boldsymbol{h}_{k_1}\|\|\boldsymbol{h}_{k_2}\|}. \tag{8}$$

see the program

Parameter	Value
Center Frequency	1917.6 MHz
Usefull Bandwidth	4.0625 MHz
BS Transmit Power	33 dBm
Number of Antennas at BS	2 (co-polarized)
Number of UEs	2
Number of Antennas at UE	2

Table 1. EMOS Parameters

This co-linearity measure is a special case of the co-linearity measure introduced in [9], when the number of receive antennas is one.

3.6. Shadowing Correlation

Last but not least we evaluate the correlation of the shadow fading between users. Let $s_{k,n}$ denote the shadowing component of the received signal strength in dB. This is equivalent to the digital signal strength after power control, i.e., after removing the path loss component. The shadow fading correlation coefficient between users k_1 and k_2 is defined as

$$\rho[k_1, k_2] = \frac{\mathbb{E}\{s_{k_1,n} s_{k_2,n}\}}{\sigma_{k_1}\sigma_{k_2}}, \tag{9}$$

where σ_k is the standard deviation of $s_{k,n}$ [11].

3.7. Spectral Divergence

The SD measures the distance between strictly positive, non-normalized spectral densities [4]. From our measurements we can define per each user three power spectral densities: the power delay profile, the Doppler spectral density and the angular spectral density. We define the time-dependent SD between these three time-varying power spectral divergences, where the index n denotes the time dependency. The SD between power delay profiles $P_{i,j}[n,\tau]$, between Doppler spectral densities $S_{i,j}[n,\nu]$ and between angular spectral densities $A[n,\phi,\theta]$ read

$$\gamma_n^P[k_1,k_2] = log\left(\frac{1}{(TNM)^2}\sum_{\tau,i,j}\frac{P_{i,j}^{(k_1)}[n,\tau]}{P_{i,j}^{(k_2)}[n,\tau]}\sum_{\tau,i,j}\frac{P_{i,j}^{(k_2)}[n,\tau]}{P_{i,j}^{(k_1)}[n,\tau]}\right),$$

$$\gamma_n^S[k_1,k_2] = log\left(\frac{1}{(PNM)^2}\sum_{\nu,i,j}\frac{S_{i,j}^{(k_1)}[n,\nu]}{S_{i,j}^{(k_2)}[n,\nu]}\sum_{\nu,i,j}\frac{S_{i,j}^{(k_2)}[n,\nu]}{S_{i,j}^{(k_1)})[n,\nu]}\right),$$

$$\gamma_n^A[k_1,k_2] = log\left(\frac{1}{(NM)^2}\sum_{\phi,\theta}\frac{A^{(k_1)}[n,\phi,\theta]}{A^{(k_2)}[n,\phi,\theta]}\sum_{\phi,\theta}\frac{A^{(k_2)}[n,\phi,\theta]}{A^{(k_1)}[n,\phi,\theta]}\right),$$

respectively. The user indices are k_1 and k_2, T and P denote the number of samples in the delay and Doppler domain respectively. The number of antenna elements at the receiver and transmitter side is denoted by M and N. The angle of arrival and angle of departure are ϕ and θ and ν is de discrete Doppler shift.

Fig. 3. Map of the measurement scenario. The position of the BS antenna as well as the five measurement points are indicated.

4. MEASUREMENTS AND RESULTS

4.1. Measurement Description

The Eurecom MIMO OpenAir Sounder (EMOS) has been used to conduct measurements in the vicinity of Eurecom, Sophia-Antipolis, France. In all measurements there were 2 Tx antennas and 2 UEs with two antennas each. A map of the scenario is depicted in Fig. 3. The measurement parameters are summarized in Table 1.

Two different sets of measurements were taken. In the first set, the first user was stationary and position x_1 and the second one was being pushed on a trolley from position x_1 to x_5 with a constant speed. The terminals were also equipped with GPS receivers, so that their distance can be evaluated. In the second set the first user is always at position x_1 and the second user is at position $x_i, i = 1, \ldots, 5$. Positions x_1, x_2, and x_5 are LOS while positions x_3 and x_4 are behind an office building. During the measurements the users were moving only within a few wavelengths to get sufficient statistics for the evaluations.

4.2. Results

Correlation Matrix Distance. In Fig. 4 and 5 we show the geodesic distance (6) and the correlation matrix distance (7) for the transmit, the receive and the full correlation matrix over the distance between the users. The correlation matrices have been calculated as described in Equations (3)–(5) averaging over $A = 200$ frames and all frequency bins. For every such estimate, we also evaluate the distance between the users.

see the program

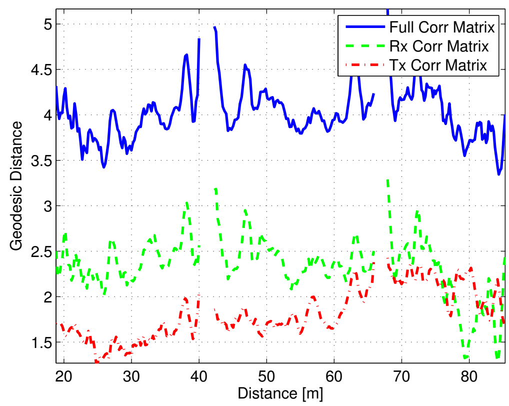


Fig. 4. Geodesic distance in dependence of the inter-user distance for the first measurement.

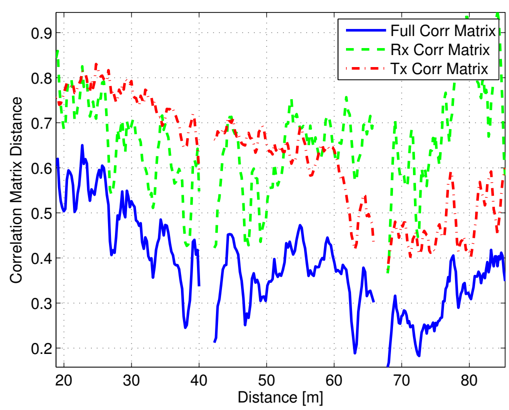


Fig. 5. Correlation matrix distance in dependence of the inter-user distance for the first measurement.

It can be seen that the geodesic distance does not show a clear dependence on the inter-user distance. The correlation matrix distance on the other hand changes significantly over distance. The full and the transmit correlation matrices are more similar when the users are close together and differ when the users are far apart. The distance between the receive correlation matrices on the other hand can not be related to the inter-user distance.

Channel Co-linearity. In Fig. 6 we plot the co-linearity between channel vectors (we use only one receive antenna). We first calculate the co-linearity for every frame m and every subcarrier q and then average over the subcarriers q. In the figure we also plot a linear function fitted to the data. A clear dependence of the co-linearity on the inter-user distance can be seen.

Shadow Fading Correlation. In Table 2 shows the shadowing correlation of second set of measurements. It can be seen that the shadowing correlation coefficient is rather low for all of the measurements.

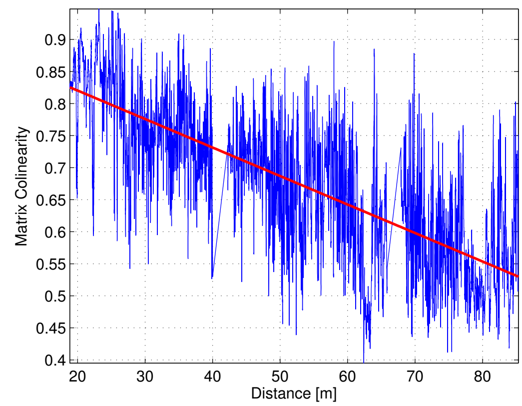


Fig. 6. Co-linearity in dependence of the inter-user distance for the first measurement. The red line is a linear function fitted to the data.

Distance [m]	1.9	15.7	32.3	53.4	76.3
Shadowing Corr.	-0.11	-0.13	-0.20	-0.08	0.01

Table 2. Shadow fading correlation in dependence of the inter-user distance for the second measurement.

Spectral Divergence. The presented results omit the SD analysis, which does not allow us to relate it to the distance between users. For the SD between power delay profiles, the value of the SD depends only on the distance from user to base station and not between users. The SD is lower when two users are at the same distance to the base station, independently of how far they are from each other. We analyse the Doppler spectral densities of two users, one of them is moving and the other remains static. We observe a constant nonzero Doppler shift for the moving user whereas the Doppler shift for the static user is zero. In this case, the SD analysis shows that the Doppler spectra between users is different, but does not present changes over distance. And finally, the analysis which could give us some meaningful results, the SD between angular spectral densities, does not have enough resultion, we only have two antenna elements at each terminal (user and base station). Because of these reasons we drop the SD analysis for this paper.

5. CONCLUSIONS

We have presented an analysis of measured MU-MIMO channels using several measures to characterize the (dis-)similarity of the channels of different users. The data was acquired us-

see the program

ing Eurecom's MU-MIMO channel sounder EMOS. The results show that the structure of the MIMO channel matrices changes significantly with the inter-user distance. This is best captured by the co-linearity measure. The transmit and the full correlation matrix also show some dependence on the inter-user distance whereas the receive correlation matrices are independent of the inter-user distance. The geodesic distance on the other hand does not show a clear dependence on the inter-user distance.

These findings are important for MU-MIMO precoding and scheduling algorithms. For example a MU-MIMO ZF precoder performs optimally if the channels of two users are orthogonal. The more aligned the channels are the worse the performance of ZF will be.

Last but not least we found that the shadowing correlation is quite low even when the nodes are quite close. This fact was also observed in other measurements [20]. However, the measurements are rather specific and thus more measurements are needed.

6. REFERENCES

[1] F. Kaltenberger, M. Kountouris, D. Gesbert, and R. Knopp, "Correlation and capacity of measured multi-user MIMO channels," in *Proc. IEEE Intl. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Cannes, France, Sep. 2008.

[2] M. Herdin, N. Czink, H. Ozcelik, and E. Bonek, "Correlation matrix distance, a meaningful measure for evaluation of non-stationary mimo channels," in *Proc. VTC 2005-Spring Vehicular Technology Conference 2005 IEEE 61st*, vol. 1, 2005, pp. 136–140.

[3] M. Talih, "Geodesic markov chains on covariance matrices," Statistical and Applied Mathematical Sciences Institute, Tech. Rep. 2007-4, mar 2007. [Online]. Available: http://www.samsi.info/TR/tr2007-04.pdf

[4] T. T. Georgiou, "Distances between power spectral densities," Tech. Rep. arXiv:math/0607026v2, Jul 2006. [Online]. Available: http://arxiv.org/abs/math.OC/0607026

[5] G. Matz, "Characterization and analysis of doubly dispersive MIMO channels," in *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Oct./Nov. 2006, pp. 946–950.

[6] L. Bernadó, T. Zemen, A. Paier, G. Matz, J. Karedal, N. Czink, C. Dumard, F. Tufvesson, M. Hagenauer, A. F. Molisch, and C. F. Mecklenbräuker, "Non-WSSUS vehicular channel characterization at 5.2 GHz - coherence parameters and channel correlation function," in *Proc. XXIX General Assembly of the International Union of Radio Science (URSI)*, Chicago, Illinois, USA, Aug. 2008.

[7] J. Koivunen, P. Almers, V.-M. Kolmonen, J. Salmi, A. Richter, F. Tufvesson, P. Suvikunnas, A. F. Molisch, and P. Vainikainen, "Dynamic multi-link indoor mimo measurements at 5.3 GHz," in *Proc. 2nd European Conference on Antennas and Propagation (EuCAP 2007)*, Edinburgh, UK, Nov. 2007.

[8] N. Czink, B. Bandemer, G. V. Vilar, L. Jalloul, and A. Paulraj, "Can multi-user MIMO measurements be done using a single channel sounder?" COST 2100, Lille, France, Tech. Rep. TD(08) 621, Nov. 2008.

[9] ——, "Spatial separation of multi-user mimo channels," COST 2100, Lille, France, Tech. Rep. TD(08) 622, Nov. 2008.

[10] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, no. 23, pp. 2145–2146, 7 Nov. 1991.

[11] F. Graziosi and F. Santucci, "A general correlation model for shadow fading in mobile radio systems," *IEEE Commun. Lett.*, vol. 6, no. 3, pp. 102–104, 2002.

[12] E. Perahia, D. C. Cox, and S. Ho, "Shadow fading cross correlation between basestations," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 1, 2001, pp. 313–317.

[13] R. Wang and D. Cox, "Channel modeling for ad hoc mobile wireless networks," in *Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th*, vol. 1, 2002, pp. 21–25.

[14] J. Weitzen and T. J. Lowe, "Measurement of angular and distance correlation properties of log-normal shadowing at 1900 mhz and its application to design of pcs systems," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 265–273, 2002.

[15] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, and A. Paulraj, "Experimental characterization of outdoor-to-indorr and indoor distributed channels," Universite de Louvain, Tech. Rep., 2009, in preparation.

[16] R. de Lacerda, L. S. Cardoso, R. Knopp, M. Debbah, and D. Gesbert, "EMOS platform: real-time capacity estimation of MIMO channels in the UMTS-TDD band," in *Proc. International Symposium on Wireless Communication Systems (IWCS)*, Trondheim, Norway, Oct. 2007.

[17] F. Kaltenberger, L. Bernadó, and T. Zemen, "Characterization of measured multi-user mimo channels using the spectral divergence measure," COST 2100, Lille, France, Tech. Rep. TD(08) 640, Nov. 2008.

[18] D. Sacristán-Murga, F. Kaltenberger, A. Pascual-Iserte, and A. I. Pérez-Neira, "Differential feedback in mimo communications: Performance with delay and real channel measurements," in *Workshop on Smart Antennas (WSA 2009)*, Berlin, Germany, Feb. 2009.

[19] F. Kaltenberger, M. Kountouris, L. S. Cardoso, R. Knopp, and D. Gesbert, "Capacity of linear multi-user MIMO precoding schemes with measured channel data," in *Proc. IEEE Intl. Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Recife, Brazil, Jul. 2008.

[20] P. Kyosti, J. Meinila, L. Hentila, X. Zhao, T. Jamsa, C. Schneider, M. Narandzic, M. Milojevic, A. Hong, J. Ylitalo, V.-M. Holappa, M. Alatossava, R. Bultitude, Y. de Jong, and T. Rautiainen, "WINNER II Channel Models," European Commission, Deliverable IST-WINNER D1.1.2 ver 1.1, Sep. 2007. [Online]. Available: https://www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf

MULTI-USER MIMO DOWNLINK PRECODING FOR TIME-VARIANT CORRELATED CHANNELS

Bin Song, Martin Haardt

Communications Research Laboratory
Ilmenau University of Technology
P. O. Box 100565, D-98694 Ilmenau, Germany
bin.song@tu-ilmenau.de
martin.haardt@tu-ilmenau.de

Tarcisio Ferreira Maciel, Anja Klein

Communications Engineering Lab
Darmstadt University of Technology
Merckstr. 25, D-64283 Darmstadt, Germany
t.maciel@nt.tu-darmstadt.de
a.klein@nt.tu-darmstadt.de

ABSTRACT

Multi-user multiple-input multiple-output (MU-MIMO) systems provide a significantly increased capacity and spectral efficiency by exploiting the benefits of space division multiple access (SDMA). The channel state information (CSI) at the base station (BS) is used to precode the transmit signals and to simplify the processing at the users' terminals. If perfect CSI is available at the transmitter, the multi-user interference (MUI) can be effectively eliminated at the BS. If the channel varies too fast to obtain short-term CSI, long-term CSI can be used alternatively to improve the system performance. In this paper we propose a new approach to multi-user precoding based on long-term CSI, which can be applied to previously defined precoding techniques originally requiring perfect CSI at the BS. It is shown that a significant performance improvement is achieved by the new approach as compared to a state of the art approach [1] to multi-user precoding with long-term CSI, especially for the case when a user has a line of sight (LOS) channel.

1. INTRODUCTION

In a multi-user multiple-input multiple-output (MIMO) communication system, multiple antennas at both ends of the link offer us the benefit of using space division multiple access (SDMA) to simultaneously transmit multiple data streams to a group of users, which results in a significant improvement of the system capacity. Obviously, this benefit comes from the awareness of channel state information (CSI) at the transmitter.

Linear precoding, as a sub-optimal SDMA strategy, has attracted much attention due to its lower complexity compared to dirty paper coding (DPC) [2–5]. In [2], a linear precoding technique called block diagonalization (BD) is proposed. With perfect CSI at the transmitter, multi-user interference (MUI) can be completely eliminated by choosing the precoding matrix of each user such that it lies in the null space of all other users' channel matrices. The main disadvantages of BD are the performance loss especially in the low SNR regime due to the strict zero MUI constraint and the dimensionality limitation that the aggregate number of receive antennas has to be less than or equal to the number of transmit antennas.

In [3], the authors introduce a regularized block diagonalization (RBD) linear precoding technique, which relaxes the limitation on the aggregate number of receive antennas. The precoding matrix of each user does not only lie in the null space of all other users' channel matrices, but also lies in the signal space of all other users' channel matrices with a power that is inversely proportional to the singular values of all other users' channels. As a result, some MUI is allowed. With perfect CSI, this technique can provide a higher data rate than BD.

By exploiting perfect CSI at the transmitter, the capacity of a multi-user MIMO system with linear precoding can be significantly improved. If it is impossible to acquire perfect instantaneous CSI at the transmitter, the spatial channel correlation can alternatively be used to reduce the MUI and improve the system performance. In this paper we consider the multi-user MIMO downlink and assume that the channel is correlated and varies too rapidly to obtain short-term CSI. We propose a new approach to exploit the knowledge of the spatial correlation at the base station (BS) that allows us to use existing precoding techniques (e.g., BD and RBD) designed for perfect CSI at the BS.

In this paper, upper case and lower case boldface letters are used to denote matrices $\boldsymbol{A}$ and column vectors $\boldsymbol{a}$, respectively. We use $\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{H}}$ to indicate the transpose and Hermitian transpose of the matrix $\boldsymbol{A}$. Moreover, $\boldsymbol{A}(i,j)$ is the matrix element in the ith row and the jth column. $\boldsymbol{A}(:,j)$ represents the jth column vector of the matrix $\boldsymbol{A}$.

This paper is organized as follows: The system model is described in Section 2. A new approach called rank-one ap-

The authors gratefully acknowlegde the partial support of the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under contract no. HA 2239/1-2.

see the program

proximated long-term CSI is introduced in Section 3, while the simulation results are presented in Section 4. A short conclusion follows in Section 5.

2. SYSTEM MODEL

We model the multi-user MIMO downlink channel as a perfectly tuned OFDM channel without any inter subcarrier interference. There are K users in the system. The BS is equipped with M_{T} transmit antennas and the ith user has M_{R_i} receive antennas. The total number of receive antennas of all users is denoted by M_{R} (i.e., $M_{\mathrm{R}} = \sum_{i=1}^{K} M_{\mathrm{R}_i}$). We use $\boldsymbol{H}_i(N_f, N_t) \in \mathbb{C}^{M_{\mathrm{R}_i} \times M_{\mathrm{T}}}$ to denote the propagation channel between the BS and the user i at subcarrier N_f and OFDM symbol N_t. Then the combined MIMO channel matrix of all users can be defined as

$$\boldsymbol{H}(N_f, N_t) = \left[\boldsymbol{H}_1^{\mathrm{T}}(N_f, N_t)\ \boldsymbol{H}_2^{\mathrm{T}}(N_f, N_t) \ldots \boldsymbol{H}_K^{\mathrm{T}}(N_f, N_t)\right]^{\mathrm{T}} . \tag{1}$$

We assume that it is not possible to track fast variations of users' channels but the information about spatial correlations of the channels can be obtained.

The downlink input output data model with linear precoding matrix $\boldsymbol{F}$ and decoding matrix $\boldsymbol{D}$ can be expressed as

$$\boldsymbol{y} = \boldsymbol{D}\big(\boldsymbol{H}(N_f, N_t)\boldsymbol{F}\boldsymbol{x} + \boldsymbol{n}\big) , \tag{2}$$

where the vectors $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{n}$ represent the vectors of transmitted symbols, received signals at all users, and additive noise at the receive antennas, respectively. $\boldsymbol{F} = [\boldsymbol{F}_1, \ldots, \boldsymbol{F}_K]$ denotes the joint precoding matrix which is used to mitigate MUI, and $\boldsymbol{D} \in \mathbb{C}^{r \times M_{\mathrm{R}}}$ is a block-diagonal decoding matrix containing each user's receive filter $\boldsymbol{D}_i \in \mathbb{C}^{r_i \times M_{\mathrm{R}_i}}$ which is designed to combine the signals of the user's antennas efficiently. The dimensions r and r_i denote the total number of data streams and the number of data streams at the ith user terminal, respectively.

We define a chunk as the basic resource element, which contains N_T consecutive OFDM symbols in the time direction and N_F subcarriers in the frequency direction. Therefore, the number of $N_{\mathrm{chunk}} = N_F \cdot N_T$ symbols are available within each chunk. Chunk-wise precoding and decoding are performed.

3. MULTI-USER LINEAR PRECODING

In a time division duplex (TDD) system, by taking into account the reciprocity principle it is possible to use the estimated uplink channel for downlink transmission. This information can be used as short-term CSI to perform precoding at the BS.

If we assume that the channel varies too rapidly to be trackable, only the information relative to the geometry of the propagation paths is captured by a spatial correlation matrix. In order to effectively perform precoding based on the available CSI at the BS, in this section we propose to exploit the knowledge of the spatial correlation with a new approach called rank-one approximated long-term CSI (ROLT-CSI).

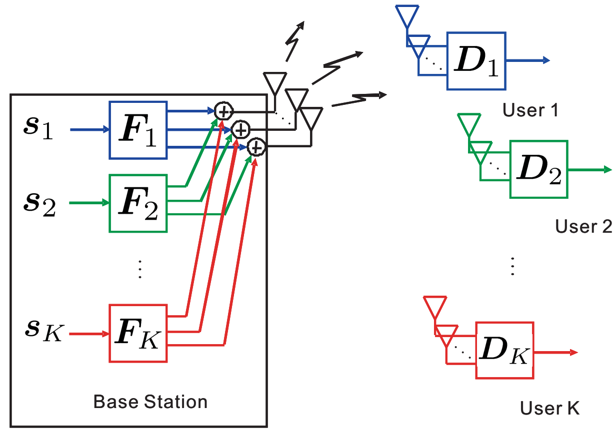


Fig. 1. Block diagram of a multi-user MIMO downlink system.

Based on ROLT-CSI, any linear precoding technique, which is designed for perfect CSI at the BS, can be modified for long-term CSI. In this paper, we present this modification for BD and RBD precoding as instructive examples.

3.1. Previous Long-term CSI Method

The authors in [1, 6] introduce a method to exploit the long-term CSI for multi-user precoding. They define the singular value decomposition (SVD) of the ith user's spatial correlation matrix estimate $\hat{\boldsymbol{R}}_i$ as

$$\hat{\boldsymbol{R}}_i = \boldsymbol{V}_i \boldsymbol{\Lambda}_i \boldsymbol{V}_i^{\mathrm{H}} \in \mathbb{C}^{M_{\mathrm{T}} \times M_{\mathrm{T}}} . \tag{3}$$

Then the equivalent channel of user i can be represented as

$$\hat{\boldsymbol{H}}_i = \boldsymbol{\Lambda}_i^{1/2} \boldsymbol{V}_i^{\mathrm{H}} . \tag{4}$$

The spatial correlation matrix estimate $\hat{\boldsymbol{R}}_{i,b}$ for user i and chunk b can be expressed as

$$\hat{\boldsymbol{R}}_{i,b} = \frac{1}{N_{\mathrm{chunk}}} \sum_{N_f=1}^{N_F} \sum_{N_t=1}^{N_T} \boldsymbol{H}_i^{\mathrm{H}}(N_f, N_t)\boldsymbol{H}_i(N_f, N_t) . \tag{5}$$

Its SVD is

$$\hat{\boldsymbol{R}}_{i,b} = \boldsymbol{V}_{i,b} \boldsymbol{\Lambda}_{i,b} \boldsymbol{V}_{i,b}^{\mathrm{H}} . \tag{6}$$

The multi-user MIMO precoding is now performed on the equivalent channel defined as follows

$$\hat{\boldsymbol{H}}_{i,b} = \boldsymbol{\Lambda}_{i,b}^{1/2} \boldsymbol{V}_{i,b}^{\mathrm{H}} . \tag{7}$$

see the program

3.2. ROLT-CSI

The ROLT-CSI approach is designed to effectively represent the channel by exploiting the knowledge of the estimated long-term channel spatial correlation.

For each receive antenna of user i, the spatial correlation matrix is represented as

$$\boldsymbol{R}_{i,l}(N_f,N_t) = \mathrm{E}\left\{\boldsymbol{h}_{i,l}(N_f,N_t)\boldsymbol{h}_{i,l}^{\mathrm{H}}(N_f,N_t)\right\} \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{T}}}\,. \tag{8}$$

Here $\boldsymbol{h}_{i,l}^{\mathrm{H}}(N_f,N_t)$ denotes the lth row of the channel matrix $\boldsymbol{H}_i(N_f,N_t) \in \mathbb{C}^{M_{\mathrm{R}_i}\times M_{\mathrm{T}}}$. The index l indicates the lth receive antenna of user i. In this paper we estimate the spatial correlation matrix of the lth receive antenna of user i by averaging over one chunk. Let $\hat{\boldsymbol{R}}_{i,b,l}$ denote the estimated spatial correlation matrix of user i, chunk b, and receive antenna l. Then we have

$$\hat{\boldsymbol{R}}_{i,b,l} = \frac{1}{N_{\mathrm{chunk}}}\sum_{N_f=1}^{N_F}\sum_{N_t=1}^{N_T}\boldsymbol{h}_{i,l}(N_f,N_t)\boldsymbol{h}_{i,l}^{\mathrm{H}}(N_f,N_t) \tag{9}$$

and its SVD as

$$\hat{\boldsymbol{R}}_{i,b,l} = \boldsymbol{V}_{i,b,l}\boldsymbol{\Lambda}_{i,b,l}\boldsymbol{V}_{i,b,l}^{\mathrm{H}}\,, l = 1,\ldots,M_{\mathrm{R}_i}\,. \tag{10}$$

According to [7], when only second-order channel statistics are available at the transmitter, the optimum strategy is to transmit along the dominant eigenmode of the matrix $\hat{\boldsymbol{R}}_{i,b,l}$. Therefore, we define the equivalent channel matrix of user i in chunk b as

$$\hat{\boldsymbol{H}}_{i,b} = \boldsymbol{A}_{i,b}\boldsymbol{B}_{i,b} \in \mathbb{C}^{M_{\mathrm{R}_i}\times M_{\mathrm{T}}}\,, \tag{11}$$

where

$$\boldsymbol{A}_{i,b} = \begin{bmatrix} \sqrt{\boldsymbol{\Lambda}_{i,b,1}(1,1)} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \sqrt{\boldsymbol{\Lambda}_{i,b,2}(1,1)} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \sqrt{\boldsymbol{\Lambda}_{i,b,M_{\mathrm{R}_i}}(1,1)} \end{bmatrix}$$

and

$$\boldsymbol{B}_{i,b} = \begin{bmatrix} \boldsymbol{V}_{i,b,1}^{\mathrm{H}}(:,1) \\ \boldsymbol{V}_{i,b,2}^{\mathrm{H}}(:,1) \\ \vdots \\ \boldsymbol{V}_{i,b,M_{\mathrm{R}_i}}^{\mathrm{H}}(:,1) \end{bmatrix}.$$

Here $\boldsymbol{\Lambda}_{i,b,l}(1,1)$ indicates the largest eigenvalue of $\hat{\boldsymbol{R}}_{i,b,l}$ and $\boldsymbol{V}_{i,b,l}^{\mathrm{H}}(:,1)$ denotes the corresponding eigenvector of $\hat{\boldsymbol{R}}_{i,b,l}$.

The multi-user MIMO precoding can now be performed on the equivalent channel as defined in equation (11). Clearly, the rank-one approximation in equation (11) effectively represents the channel only if its spatial correlation matrix in equation (9) also has a low rank.

3.3. Block Diagonalization Precoding

We define $\boldsymbol{F}_i \in \mathbb{C}^{M_{\mathrm{T}}\times r_i}$ as the ith user's precoding matrix. In [2], the optimal $\boldsymbol{F}_i$ of BD precoding lies in the null space of the other users' channel matrices. Thereby, a multi-user MIMO downlink channel is decomposed into multiple parallel independent single-user MIMO channels.

Let us define $\widetilde{\boldsymbol{H}}_i$ as [1]

$$\widetilde{\boldsymbol{H}}_i = \begin{bmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_{i-1} \\ \boldsymbol{H}_{i+1} \\ \vdots \\ \boldsymbol{H}_K \end{bmatrix} \in \mathbb{C}^{(M_{\mathrm{R}}-M_{\mathrm{R}_i})\times M_{\mathrm{T}}}\,. \tag{12}$$

The zero MUI constraint forces the matrix $\boldsymbol{F}_i$ to lie in the null space of $\widetilde{\boldsymbol{H}}_i$. By using the singular value decomposition (SVD), $\widetilde{\boldsymbol{H}}_i$ is written as

$$\widetilde{\boldsymbol{H}}_i = \widetilde{\boldsymbol{U}}_i\widetilde{\boldsymbol{\Sigma}}_i\left[\widetilde{\boldsymbol{V}}_i^{(1)}\ \widetilde{\boldsymbol{V}}_i^{(0)}\right]^{\mathrm{H}} \tag{13}$$

where $\widetilde{\boldsymbol{V}}_i^{(1)}$ holds the first $\widetilde{L}_i$ right singular vectors, and $\widetilde{\boldsymbol{V}}_i^{(0)}$ holds the last $(M_{\mathrm{T}}-\widetilde{L}_i)$ right singular vectors. Here $\widetilde{L}_i$ indicates the rank of $\widetilde{\boldsymbol{H}}_i$. Thus, $\widetilde{\boldsymbol{V}}_i^{(0)}$ forms an orthogonal basis for the null space of $\widetilde{\boldsymbol{H}}_i$. The equivalent channel of user i after eliminating the MUI is represented as $\boldsymbol{H}_i\widetilde{\boldsymbol{V}}_i^{(0)} \in \mathbb{C}^{M_{\mathrm{R}_i}\times(M_{\mathrm{T}}-\widetilde{L}_i)}$ which is equivalent to a system with $M_{\mathrm{T}}-\widetilde{L}_i$ transmit antennas and M_{R_i} receive antennas. Each of these equivalent single-user MIMO channels has the same properties as a conventional single-user MIMO channel.

We define the SVD of

$$\boldsymbol{H}_i\widetilde{\boldsymbol{V}}_i^{(0)} = \boldsymbol{U}_i\boldsymbol{\Sigma}_i\left[\boldsymbol{V}_i^{(1)}\ \boldsymbol{V}_i^{(0)}\right]^{\mathrm{H}} \tag{14}$$

and denote the rank of the ith user's equivalent channel matrix as L_i. Now the BD precoding matrix of user i can be defined as the product of the first L_i singular vectors $\boldsymbol{V}_i^{(1)}$ and $\widetilde{\boldsymbol{V}}_i^{(0)}$ with proper power loading.

If there is only long-term CSI available at the BS, we use the equivalent channel (11) from the ROLT-CSI approach instead of the exact channel $\boldsymbol{H}_i$ in equations (12) and (14).

3.4. Regularized Block Diagonalization Precoding

RBD precoding is designed to relax the limitation on the aggregate number of receive antennas and has a significantly improved data rate and diversity order compared to BD precoding [3]. The RBD precoding design is performed in two steps. In the first step, we balance the MUI suppression

[1] In subsections 3.3 and 3.4, we use $\boldsymbol{H}_i$ instead of $\boldsymbol{H}_i(N_f,N_t)$ in order to simplify the introduction to BD and RBD precoding.

see the program

which is achieved by reducing the overlap of the row spaces spanned by the effective channel matrices of different users and any MIMO processing gain which requires that the users use as much as possible the available subspace. In the second step, we optimize the system performance assuming parallel single-user MIMO channels.

Let us define the joint precoding matrix as

$$\boldsymbol{F} = [\boldsymbol{F}_1 \quad \boldsymbol{F}_2 \quad \dots \quad \boldsymbol{F}_K] \in \mathbb{C}^{M_{\mathrm{T}} \times r} , \qquad (15)$$

where $\boldsymbol{F}_i \in \mathbb{C}^{M_{\mathrm{T}} \times r_i}$ is the ith user's precoding matrix.

In [3], the matrix $\boldsymbol{F}$ of RBD precoding is proposed as

$$\boldsymbol{F} = \beta \boldsymbol{F}_a \cdot \boldsymbol{F}_b , \qquad (16)$$

where

$$\boldsymbol{F}_a = [\boldsymbol{F}_{a_1} \, \boldsymbol{F}_{a_2} \dots \boldsymbol{F}_{a_K}] \in \mathbb{C}^{M_{\mathrm{T}} \times r} ,$$

and

$$\boldsymbol{F}_b = \begin{bmatrix} \boldsymbol{F}_{b_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}_{b_2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{F}_{b_K} \end{bmatrix} \in \mathbb{C}^{r \times r} .$$

The matrix $\boldsymbol{F}_a$ is used to suppress MUI while balancing it with noise enhancement first, and then the matrix $\boldsymbol{F}_b$ is used to further optimize the system performance by optimal power loading. Finally, the parameter β is chosen to set the total transmit power to the power constraint.

The equivalent combined channel matrix of all users after precoding is equal to

$$\boldsymbol{H}\boldsymbol{F}_a = \begin{bmatrix} \boldsymbol{H}_1\boldsymbol{F}_{a_1} & \boldsymbol{H}_1\boldsymbol{F}_{a_2} & \cdots & \boldsymbol{H}_1\boldsymbol{F}_{a_K} \\ \boldsymbol{H}_2\boldsymbol{F}_{a_1} & \boldsymbol{H}_2\boldsymbol{F}_{a_2} & \cdots & \boldsymbol{H}_2\boldsymbol{F}_{a_K} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}_K\boldsymbol{F}_{a_1} & \boldsymbol{H}_K\boldsymbol{F}_{a_2} & \cdots & \boldsymbol{H}_K\boldsymbol{F}_{a_K} \end{bmatrix} , \qquad (17)$$

where $\boldsymbol{H} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}}}$ represents the combined channel matrix of all users. The ith user's effective channel is given by $\boldsymbol{H}_i\boldsymbol{F}_{a_i}$ and the interference generated to the other users is determined by $\widetilde{\boldsymbol{H}}_i\boldsymbol{F}_{a_i}$, where $\widetilde{\boldsymbol{H}}_i$ is defined in equation (12).

The matrix $\boldsymbol{F}_a$ is chosen such that the off-diagonal block matrices of equation (17) converge to zero as the SNR increases. Then we have

$$\boldsymbol{F}_{a_i} = \widetilde{\boldsymbol{V}}_i \left(\widetilde{\boldsymbol{\Sigma}}_i^{\mathrm{T}} \widetilde{\boldsymbol{\Sigma}}_i + \mu \boldsymbol{I}_{M_{\mathrm{T}}} \right)^{-1/2} , \qquad (18)$$

where the SVD of $\widetilde{\boldsymbol{H}}_i$ is given by

$$\widetilde{\boldsymbol{H}}_i = \widetilde{\boldsymbol{U}}_i \widetilde{\boldsymbol{\Sigma}}_i \widetilde{\boldsymbol{V}}_i^{\mathrm{H}} . \qquad (19)$$

After suppressing MUI by $\boldsymbol{F}_a$, we optimize the system performance by setting $\boldsymbol{F}_{b_i}$ as

$$\boldsymbol{F}_{b_i} = \boldsymbol{V}_i \boldsymbol{M}_{b_i} \qquad (20)$$

where $\boldsymbol{V}_i$ is obtained from the SVD of the ith user's equivalent channel

$$\boldsymbol{H}_i \boldsymbol{F}_{a_i} = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \boldsymbol{V}_i^{\mathrm{H}} . \qquad (21)$$

The choice of the power loading matrix $\boldsymbol{M}_{b_i}$ depends on the optimization criteria. In this work we assume that $\boldsymbol{M}_{b_i}$ is unitary.

If there is only long-term CSI available at the BS, we use the equivalent channel (11) from the ROLT-CSI approach instead of the exact channel $\boldsymbol{H}_i$ in equations (12), (17) and (21).

4. SIMULATION RESULTS

In this section we evaluate the throughput performance of the BD and RBD precoding techniques, when only long-term CSI is available. We consider a 3 users MIMO downlink system. The simulation scenario is illustrated in Figure 2. The channels between each user and the BS are generated by a geometry-based channel model called *IlmProp*, which has been developed at Ilmenau University of Technology [8] and is capable of dealing with time variant frequency selective scenarios.

There are 8 transmit antennas at the BS and each user is equipped with 2 receive antennas. We simultaneously transmit two data streams to each user. User 1 and user 2 always have non-line of sight (NLOS) channels and user 3 alway has a line of sight (LOS) channel. The velocities of the three users are 10 km/h. In Table 1, the important OFDM parameters are listed.

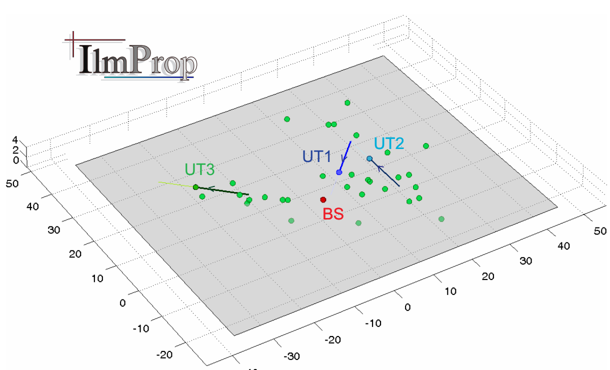


Fig. 2. The geometrical representation of the simulation scenario. Each green point represents a fixed scatter. The channel impulse responses (CIR) are generated as a sum of propagation rays. The channel is computed from the superposition of the LOS component and a number of rays which represent the multi-path components. User 1 and user 2 always have NLOS channels and user 3 alway has a LOS channel.

We use *uplink dedicated pilots* to estimate the channel between the user terminal and all BS antennas. For each chunk, there are several pilots available. We compute one channel estimate per pilot and then interpolate between these estimates

Table 1. OFDM Parameters

Parameters	Values
Carrier Frequency	5 GHz
Subcarrier Spacing	0.50196 MHz
Useful Symbol Duration	1.9922 μs
System Bandwidth	128.5 MHz
Used Subcarriers	$[-128 : +128]$, 0 not used
Chunk Size	8 subcarriers, 15 OFDM symbols
Duplexing Mode	TDD

for every symbol in the chunk. Then we calculate the equivalent channel of the chunk with equation (11) for the ROLT-CSI approach and with equation (7) for the long-term CSI method of [1], respectively. Then the BS can compute the precoding matrix $\boldsymbol{F}$ for each chunk. The linear precoding schemes used in the simulation are BD precoding and RBD precoding.

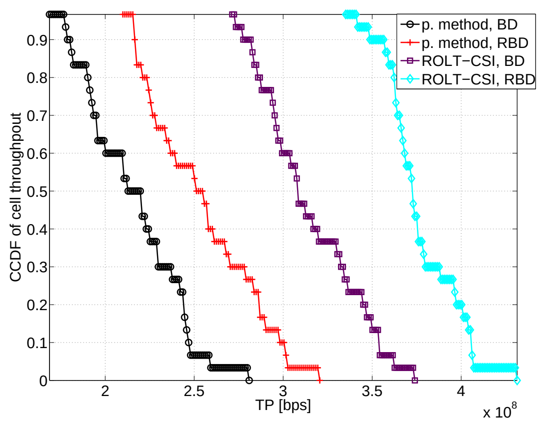


Fig. 3. CCDF of the sum rates with BD and RBD precoding based on long-term CSI at the transmitter, respectively. p. method indicates the previous long-term CSI method.

In Figures 3, 4, and 5 we assume that the channel estimate per pilot of each chunk is perfectly performed. In Figure 3, we compare the throughput of the system with precoding based on ROLT-CSI proposed in this paper to the throughput based on the state of the art long-term CSI method in [1]. We can see that RBD precoding can achieve a higher data rate than BD precoding. When linear precoding is performed based on long-term CSI, a significant performance gain can be achieved by our new approach relative to the previous long-term CSI method.

In Figures 4 and 5 the individual user throughputs based on ROLT-CSI and the previous long-term CSI approach are compared. It is shown that the ROLT-CSI approach is particularly efficient for the user who has the LOS channel. Even for the users who only have NLOS channels, which means

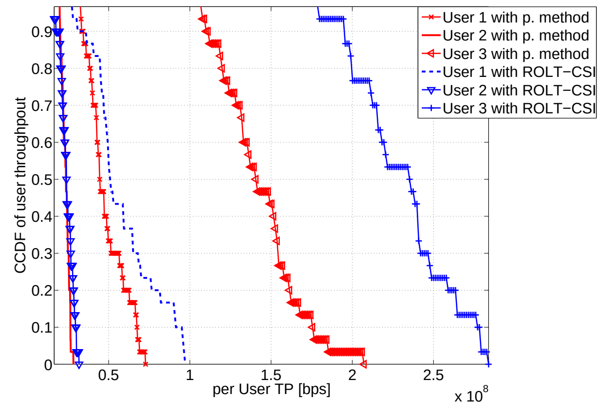


Fig. 4. CCDF of the individual user throughput with BD precoding based on long-term CSI at the transmitter, p. method indicates the previous long-term CSI method.

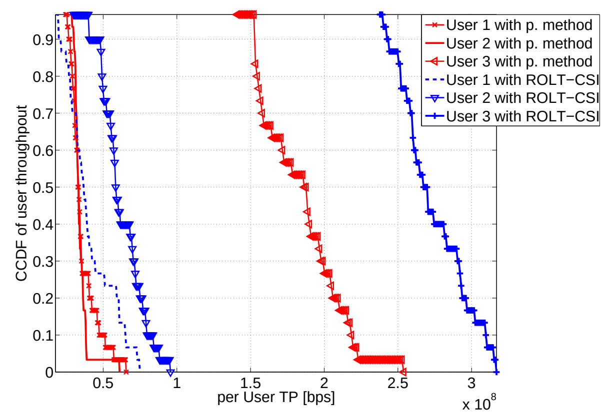


Fig. 5. CCDF of the individual user throughput with RBD precoding based on long-term CSI at the transmitter, p. method indicates the previous long-term CSI method.

that the spatial correlation matrix of these user channels have a high rank, relative to the previous long-term CSI method there are still some performance gains available for the presented ROLT-CSI approach.

Taking into account realistic channel propagation conditions, for Figure 6 and 7 we assume that the channel estimate per pilot of each chunk is imperfectly performed. We consider a channel estimation error, a channel interpolation error, and the delay resulting from the fact that the available CSI of chunk k will be used to optimize the transmission over the channel realization of chunk $(k + n)$. One chunk and three chunks delay are considered separately in the simulation. According to Table 1, the duration of one chunk is equal to the duration of 15 OFDM symbols.

see the program

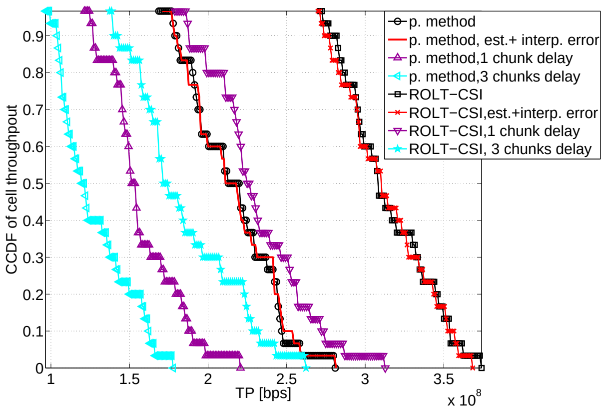


Fig. 6. CCDF of the sum rates with BD precoding based on long-term CSI at the transmitter, p. method indicates the previous long-term CSI method.

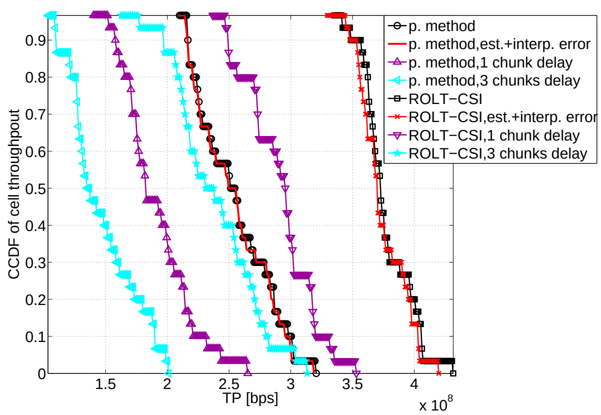


Fig. 7. CCDF of the sum rates with RBD precoding based on long-term CSI at the transmitter, p. method the indicates previous long-term CSI method.

For the CSI imperfection, the channel estimation error and interpolation error are modeled according to [9], but we increase the interpolation error variance to -20 dB. It is found that the delay is still the predominant cause of a performance degradation in a precoded multi-user MIMO system with long-term CSI.

5. CONCLUSIONS

In this paper we propose a new precoding approach that allows the use of previously defined linear precoding techniques originally requiring perfect CSI at the transmitter in cases when only long-term CSI is available. The new approach to exploit the long-term CSI is called rank-one approximated long-term CSI (ROLT-CSI). Using ROLT-CSI we evaluate the throughput performance of the two multi-user MIMO precoding techniques BD and RBD. RBD precoding permits some MUI and has no restrictions considering the number of antennas at the user. In contrast, BD has zero MUI and the aggregate number of receive antennas has to be less than or equal to the number of transmit antennas. Furthermore, we compare the throughput of the system with precoding based on ROLT-CSI to the system throughput based on the state of the art long-term CSI method in [1]. A significant performance gain can be achieved by our new approach. From the individual user throughput comparison we can see that our new approach is particularly efficient when the user's spatial correlation matrix has a low rank. If the user's spatial correlation matrix has a high rank, our new approach still works well.

To take into account realistic channel propagation conditions, we also consider in the simulations a channel estimation error, a channel interpolation error, and the delay resulting from the fact that the available CSI of chunk k will be used to optimize the transmission over the channel realization of chunk $(k+n)$. It is found that the delay is still the predominant cause of a performance degradation in a precoded multi-user MIMO system with long-term CSI.

REFERENCES

[1] V. Stankovic and M. Haardt, "Multi-user MIMO downlink beamforming over correlated MIMO channles," in *Proc. International ITG/IEEE Workshop on Smart Antennas (WSA'05)*, 2005.

[2] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multi-user MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 461–471, Feb. 2004.

[3] V. Stankovic and M. Haardt, "Generalized design of multi-user MIMO precoding matrices," *IEEE Trans. on Wireless Communications*, vol. 7, pp. 953–961, 2007.

[4] T. F. Maciel and A. Klein, "A low-complexity resource allocation strategy for SDMA/OFDMA systems," in *Proc. IST Mobile and Wireless Communications Summit*, 2007.

[5] T. F. Maciel and A. Klein, "A convex quadratic SDMA grouping algorithm based on spatial correlation," in *Proc. IEEE International Conference on Communications (ICC'07))*, 2007.

[6] F. Roemer, M. Fuchs, and M. Haardt, "Distributed MIMO systems with spatial reuse for high-speed-indoor mobile radio access," in *of the 20-th Meeting of the Wireless World Research Forum (WWRF)*, (Ottawa, ON, Canada), Apr. 2008.

[7] M. Bengtsson and B. Ottersten, "Optimum and suboptimum transmit beamforming," in *Handbook of antennas*

in wireless communications (L. C. Godara, eds.), CRC Press, 2002.

[8] G. Del Galdo, M. Haardt, and C. Schneider, "Geometry-based channel modelling of MIMO channels in comparison with channel sounder measurements," *Advances in Radio Science - Kleinheubacher Berichte*, pp. 117–126, October 2003, more information on the model, as well as the source code and some exemplary scenarios can be found at http://tu-ilmenau.de/ilmprop.

[9] WINNER II IST-4-027756, "D6. 13. 7, WINNER II test scenarios and calibration cases issue 2," Framework Programme 6, Tech. Rep. v1.0, 2007. [online]. Available: https://www.ist-winner.org/.

see the program

PROBABILITY OF ERROR FOR BPSK MODULATION IN DISTRIBUTED BEAMFORMING WITH PHASE ERRORS

Shuo Song, John S. Thompson, Pei-Jung Chung, Peter M. Grant

Institute for Digital Communications
Joint Research Institute for Signal & Image Processing
School of Engineering and electronics
University of Edinburgh
{s.song, john.thompson, p.chung, peter.grant}@ed.ac.uk

ABSTRACT

This paper presents an investigation into the error probability performance for binary phase-shift keying modulation in distributed beamforming with phase errors. The effects of the number of nodes on the beamforming performance are examined as well as the influences of the cumulative phase errors and the total transmit power. Simulation results show a good match with the mathematical analysis of error probability in both static and time-varying channels.

1. INTRODUCTION

Recently, there has been interest in applying beamforming techniques into wireless sensor networks. The motivation is to reduce the energy requirement for each sensor node in signal transmission, and extend the communication range to a far field receiver. The individual sensor nodes share the collected information and transmit it in such a way that the signals add coherently at the destination. Transmit beamforming requires accurate synchronization in frequency and phase among sensors, and accurate channel estimation between each sensor node and the receiver. Although certain techniques have been designed in [1], [2], [3] to minimize the phase errors among sensor nodes, phase errors cannot be eliminated due to hardware constraints. Minimizing total transmit power using quantized channel state information has been studied in [4]. The beam pattern performance of distributed beamforming has been studied in [5] and [6] with synchronous phase errors among sensor nodes. From a more practical view, in this paper, we investigate the probability of error for binary phase-shift keying (BPSK) modulation in distributed beamforming with synchronous phase errors and noise.

Shuo Song thanks China Scholarship Council/University of Edinburgh Joint Scholarship Program for supporting his PhD studies.

We acknowledge the support of the Scottish Funding Council for the Joint Research Institute with the Heriot-Watt University which is a part of the Edinburgh Research Partnership.

The rest of the paper is organized as follows. Section 2 introduces the system model. In Section 3 we give an equivalent channel concept to simplify the whole beamforming process. In Section 4 the mathematical analysis of the average bit error ratio (BER) for BPSK in both static and time-varying channels are presented. In Section 5 we analyze the beamforming gain with constant total transmit power. Section 6 then presents simulation results to compare with the theoretical analysis and Section 7 draws conclusions for the paper.

2. SYSTEM MODEL

We consider a system of N sensor nodes collaboratively beamforming a narrowband message signal $s(t) = A \cdot m(t)$ to a distant coherent receiver, where A is the amplitude of the message signal. This is performed in a distributed manner by each sensor node modulating $s(t)$ with a RF carrier signal, as illustrated in Fig. 1.

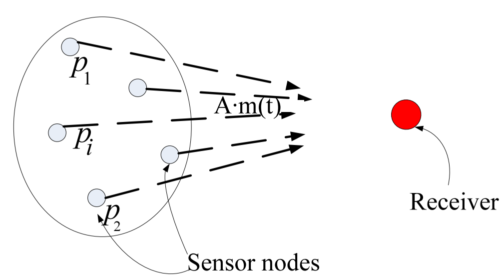


Fig. 1. System model for distributed beamforming

We assume that each sensor node and the receiver are equipped with one single ideal omnidirectional antenna, and there are no mutual coupling effects among the antennas. The receiver has the ability to retrieve the overall channel phase from the received signal. All sensor nodes are synchronized so that they can transmit at the same carrier frequency, and

see the program

signals transmitted from each sensor node will be added coherently at the receiver. The complex baseband model of the received signal is given by

$$r(t) = \sum_{i=1}^{N} |h_i(t)p_i(t)|e^{j\phi_i(t)}s(t) + n(t) \tag{1}$$

where $p_i(t)$ is the amplification factor and $h_i(t)$ is the channel gain for sensor node i, $\phi_i(t)$ is the cumulative phase error of the carrier signal from the synchronization process among sensor nodes and the estimation of the channel gain for sensor node i, $n(t) \sim CN(0, \sigma_n^2)$ is additive white Gaussian noise. We assume all phase errors $\phi_i(t)$ are independently and uniformly distributed within the range $(-\phi_0, \phi_0)$, which is the assumption adopted in previously reported investigations [1], [2].

A. Static Channel

In a static channel scenario, $h_i(t)$ is set equal to a constant. For simplicity, we set coefficients $h_i(t)$, $p_i(t)$ to be unity. Then the system model is expressed as:

$$r(t) = \sum_{i=1}^{N} e^{j\phi_i(t)}s(t) + n(t) \tag{2}$$

B. Time-Varying Channel

In our time-varying model, the channel coefficients are independent circularly symmetric complex Gaussian distributed, denoted as $h_i(t) \sim CN(0, 1)$, which corresponds to non-line of sight or Rayleigh fading channels. By applying maximal ratio combining, where the pre-amplification gain of each channel is made proportional to the received signal level, we set $|p_i(t)| = |h_i(t)|$ and the system model is then expressed as:

$$r(t) = \sum_{i=1}^{N} |h_i(t)|^2 e^{j\phi_i(t)}s(t) + n(t) \tag{3}$$

3. ANALYSIS OF THE EQUIVALENT CHANNEL

If we view the whole beamforming process as an equivalent channel, denoted as $H(t)$, the system model becomes:

$$r(t) = H(t)s(t) + n(t) \tag{4}$$

where $H(t) = \sum_{i=1}^{N} e^{j\phi_i(t)}$ for the static channel scenario, and $H(t) = \sum_{i=1}^{N} |h_i(t)|^2 e^{j\phi_i(t)}$ for the Rayleigh fading channel scenario. With a coherent receiver, the signal-to-noise ratio (SNR) gain, $\|H(t)\|^2$, is the key element deciding the error probability for distributed beamforming and the communication range for power limited sensor networks.

A. Static Channel

By the central limit theorem, with a large number of sensor nodes N, and the independent identically distributed (i.i.d.) random variables $\phi_i(t)$, we have:

$$\begin{aligned} \|H(t)\|^2 &= \left\| \sum_{i=1}^{N} e^{j\phi_i(t)} \right\|^2 \\ &= \left\| \sum_{i=1}^{N} \cos\phi_i(t) + j\sum_{i=1}^{N} \sin\phi_i(t) \right\|^2 \\ &= \|a_S + jb_S\|^2 \\ &= a_S^2 + b_S^2 \end{aligned} \tag{5}$$

where $a_S = \sum_{i=1}^{N} \cos\phi_i(t) \sim N(\mu_{a_S}, \sigma_{a_S}^2)$, $b_S = \sum_{i=1}^{N} \sin\phi_i(t) \sim N(\mu_{b_S}, \sigma_{b_S}^2)$, using the subscript S for the static channels.

Since the variables $\phi_i(t)$ are independently and uniformly distributed within the range $(-\phi_0, \phi_0)$, the means and variances of a_S and b_S can be obtained as:

$$\begin{aligned} \mu_{a_S} &= N \cdot E[\cos\phi_i(t)] \\ &= N\frac{\sin\phi_0}{\phi_0} \end{aligned} \tag{6}$$

$$\mu_{b_S} = 0 \tag{7}$$

$$\begin{aligned} \sigma_{a_S}^2 &= N\Big(\quad E[\cos^2\phi_i(t)] - (E[\cos\phi_i(t)])^2 \quad \Big) \\ &= N\left(\frac{1}{2} + \frac{\sin 2\phi_0}{4\phi_0} - (\frac{\sin\phi_0}{\phi_0})^2\right) \end{aligned} \tag{8}$$

$$\begin{aligned} \sigma_{b_S}^2 &= N\Big(\quad E[\sin^2\phi_i(t)] - (E[\sin\phi_i(t)])^2 \quad \Big) \\ &= N\left(\frac{1}{2} - \frac{\sin 2\phi_0}{4\phi_0}\right) \end{aligned} \tag{9}$$

From (8) and (9), we see, for the equivalent channel $H(t)$, the variance of the real part $\sigma_{a_S}^2$ and the variance of the imaginary part $\sigma_{b_S}^2$ are not equal, which means that the probability density function (PDF) of $\|H(t)\|^2$ is not easily obtained from the joint PDF of $H(t)$, $p(a_S, b_S)$.

B. Rayleigh Fading Channel

For the Rayleigh fading channels, similarly, with a large number of sensor nodes N, and the i.i.d. random variables $h_i(t)$ which are independent from the i.i.d. random variables $\phi_i(t)$, we have:

see the program

$$\begin{aligned}\|H(t)\|^2 &= \left\| \sum_{i=1}^{N} |h_i(t)|^2 e^{j\phi_i(t)} \right\|^2 \\ &= \left\| \sum_{i=1}^{N} |h_i(t)|^2 \cos\phi_i(t) + j \sum_{i=1}^{N} |h_i(t)|^2 \sin\phi_i(t) \right\|^2 \\ &= \|a_R + jb_R\|^2 \\ &= a_R^2 + b_R^2 \end{aligned} \tag{10}$$

where $a_R = \sum_{i=1}^{N} |h_i(t)|^2 \cos\phi_i(t) \sim N(\mu_{a_R}, \sigma_{a_R}^2)$, and $b_R = \sum_{i=1}^{N} |h_i(t)|^2 \sin\phi_i(t) \sim N(\mu_{b_R}, \sigma_{b_R}^2)$, using the subscript R for the Rayleigh fading channels.

Based on the previous assumptions that the channel coefficients $h_i(t)$ are independent circularly symmetric complex Gaussian distributed $h_i(t) \sim CN(0,1)$, and $\phi_i(t) \sim (-\phi_0, \phi_0)$, we derived the means and variances of a_R and b_R as follows:

$$\begin{aligned}\mu_{a_R} &= N \cdot E[|h_i(t)|^2 \cos\phi_i(t)] \\ &= N \cdot E[|h_i(t)|^2] \cdot E[\cos\phi_i(t)] \\ &= N \frac{\sin\phi_0}{\phi_0} \end{aligned} \tag{11}$$

$$\mu_{b_R} = 0 \tag{12}$$

$$\begin{aligned}\sigma_{a_R}^2 &= N\Big(\quad E[(|h_i(t)|^2 \cos\phi_i(t))^2] - (E[|h_i(t)|^2 \cos\phi_i(t)])^2 \quad \Big) \\ &= N\left(E[|h_i(t)|^4] \cdot E[\cos^2\phi_i(t)] - (E[|h_i(t)|^2 \cos\phi_i(t)])^2\right) \\ &= N\left(1 + \frac{\sin 2\phi_0}{2\phi_0} - (\frac{\sin\phi_0}{\phi_0})^2\right) \end{aligned} \tag{13}$$

$$\begin{aligned}\sigma_{b_R}^2 &= N\Big(\quad E[(|h_i(t)|^2 \sin\phi_i(t))^2] - (E[|h_i(t)|^2 \sin\phi_i(t)])^2 \quad \Big) \\ &= N\left(E[|h_i(t)|^4] \cdot E[\sin^2\phi_i(t)] - (E[|h_i(t)|^2 \sin\phi_i(t)])^2\right) \\ &= N\left(1 - \frac{\sin 2\phi_0}{2\phi_0}\right) \end{aligned} \tag{14}$$

Similarly, from (13) and (14) we see, for the Rayleigh fading channel scenario, $\sigma_{a_R}^2$ and $\sigma_{b_R}^2$ are not equal, thus the expression of the PDF of $\|H(t)\|^2$ is difficult to compute.

4. MATHEMATICAL ANALYSIS OF ERROR PROBABILITY

The BER of BPSK over a fixed channel in the presence of AWGN is given by [7] in Chapter 5:

$$P_e(\gamma) = \frac{1}{2}\mathrm{erfc}(\sqrt{\gamma}) \tag{15}$$

where γ is the received signal-to-noise ratio per bit, and erfc(.) is the complementary error function.

When the channel gain is random, the average BER for BPSK over all values of γ is given by [7] in Chapter 14:

$$P_e = \int_0^\infty P_e(\gamma) p(\gamma) d\gamma \tag{16}$$

where $\gamma = \|H(t)\|^2 \frac{A^2}{\sigma_n^2}$ in our system model described in Section 2.

In Section 3, we have analyzed the SNR gain $\|H(t)\|^2$ of the distributed beamforming system, and the expression of the PDF of $\|H(t)\|^2$ was not obtained due to the variances of the real part and the imaginary part of the equivalent channel being unequal. Consequently, $p(\gamma)$ is not available in either the static channel scenario or the Rayleigh fading channel scenario. Formula (16) cannot be solved directly to get a closed-form expression of the integration for our model, and can only be evaluated by numerical techniques. Instead, we provide another method to approximate the BER results as follows.

Method 1:

For both the static channel scenario and the Rayleigh fading channel scenario, we set the variances of the real part and the imaginary part of $H(t)$ to be equal and use the maximum value between them:

$$\sigma_S^2 = \max(\sigma_{a_S}^2, \sigma_{b_S}^2) \tag{17}$$

for the static channels, and

$$\sigma_R^2 = \max(\sigma_{a_R}^2, \sigma_{b_R}^2) \tag{18}$$

for the Rayleigh fading channels.

Because the real part and the imaginary part of the equivalent channel $H(t)$ now have different means but same variances, the magnitude gain of $H(t)$ is approximated as a Rician distribution.

The closed-form of BER for BPSK through Rician fading channel with a coherent receiver is given by [8]:

$$P_E = Q_1(u,w) - \frac{1}{2}\left(1 + \sqrt{\frac{d}{1+d}}\right)\exp\left(-\frac{u^2+w^2}{2}\right) I_0(uw) \tag{19}$$

where

$$d = 2\sigma^2 \frac{A^2}{\sigma_n^2} \tag{20}$$

$$u = \sqrt{\frac{\mu_a^2 + \mu_b^2}{2\sigma^2} \cdot \frac{1 + 2d - 2\sqrt{d(1+d)}}{2(1+d)}} \tag{21}$$

see the program

$$w = \sqrt{\frac{\mu_a^2 + \mu_b^2}{2\sigma^2} \cdot \frac{1 + 2d + 2\sqrt{d(1+d)}}{2(1+d)}} \qquad (22)$$

and $I_0(x)$ is the zeroth-order-modified Bessel function of the first kind, defined as:

$$I_0(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{2k}}{k!\Gamma(k+1)}, \qquad x \geq 0 \qquad (23)$$

$Q_1(x, y)$ is the Marcum Q-function, defined as:

$$Q_1(x, y) = \int_y^{\infty} z \cdot \exp\left(-\frac{z^2 + x^2}{2}\right) I_0(xz)dz \qquad (24)$$

Using (19) to (24), we can get the BER for our static channel scenario by substituting (6), (7), (17) for μ_a, μ_b, σ^2 in (20), (21), (22), and get the BER for our Rayleigh fading channel scenario by substituting (11), (12), (18) for μ_a, μ_b, σ^2 in (20), (21), (22).

An approximation of $I_0(x)$ is given by [9] in Chapter 6:

$$I_0(x) \approx \frac{1}{\sqrt{2\pi x}} \exp(x), \qquad x \gg 0 \qquad (25)$$

and after manipulation, (19) can be simplified as:

$$P_E = Q_1(u, w) - \frac{1}{2\sqrt{2\pi uw}}\left(1 + \sqrt{\frac{d}{1+d}}\right)\exp\left(-\frac{(u-w)^2}{2}\right) \qquad (26)$$

Method 2:

We are currently investigating the approximation of the BER performance by an additive white Gaussian noise formula. This is the subject of ongoing work.

5. DISTRIBUTED BEAMFORMING GAIN WITH CONSTANT TOTAL TRANSMIT POWER

As the received signal-to-noise ratio cannot show the advantages of beamforming gain, and is uncertain due to independent and random phase errors ϕ_i, our simulation results are plotted as BER vs total transmit power. Before we present our simulation results, we first analyze the beamforming gain with constant total transmit power. We use P to represent the total transmit power of all the sensor nodes. In the static channel scenario,

$$\begin{aligned} P &= \sum_{i=1}^{N} A^2 \\ &= A^2 \cdot N \end{aligned}$$

In the Rayleigh fading channel scenario,

$$\begin{aligned} P &= \sum_{i=1}^{N} (A|p_i(t)|)^2 \\ &= A^2 \sum_{i=1}^{N} |p_i(t)|^2 \end{aligned}$$

With large N, by the law of large numbers, it becomes:

$$P \approx A^2 \cdot N$$

Generally, with a constant P, we can represent A as:

$$A = \sqrt{\frac{P}{N}} \qquad (27)$$

Putting (27) and $s(t) = A \cdot m(t)$ into (4), we obtain:

$$r(t) = H(t)\sqrt{\frac{P}{N}} m(t) + n(t) \qquad (28)$$

and

$$\gamma = \frac{1}{N} \|H(t)\|^2 \frac{P}{\sigma_n^2} \qquad (29)$$

Since the mean of $\|H(t)\|^2$ grows linearly with N^2, the mean of the received signal-to-noise ratio per bit $\gamma_{mean} \propto P \cdot N$, and with a constant P, γ_{mean} is proportional to N.

6. SIMULATION RESULTS

In this section, we present some simulation results in accordance with our previous assumptions, and compare them with our mathematical analysis given in Section 4.

Fig. 2 shows the comparison of the simulation results with the mathematical analysis based on method 1 for BPSK modulation over *static channels* with phase errors. The simulation results are conducted over 10^5 symbols with different number of nodes $N = 10, 100, 1000$, and different phase error ranges $\phi_0 = 18°, 36°, 54°, 72°$. We set $n(t) \sim CN(0, 1)$. All curves in Fig. 2 are drawn by (19) except the curves for $\phi_0 = 18°$ with $N = 100$ in part (b) and $\phi_0 = 18°, 36°, 54°$ with $N = 1000$ in part (c). These four curves cannot be drawn out by (19) because of the overflow caused by the function $I_0(x)$ in (23) used in MATLAB. Instead of (19), We use (26) to draw these four curves in Fig. 2.

By comparing the simulation results plotted in parts (a), (b), (c) in Fig. 2 and noting the order of magnitude difference of total transmit power in (a), (b), (c), we find that, with similar BER performance in each part, when increasing the number of nodes N by a factor of 10, the total transmit power is reduced by a factor of 10, which means the energy transmitted by each node is reduced by a factor of 10^2. Thus, we have the conclusion that increasing the number of nodes N

see the program

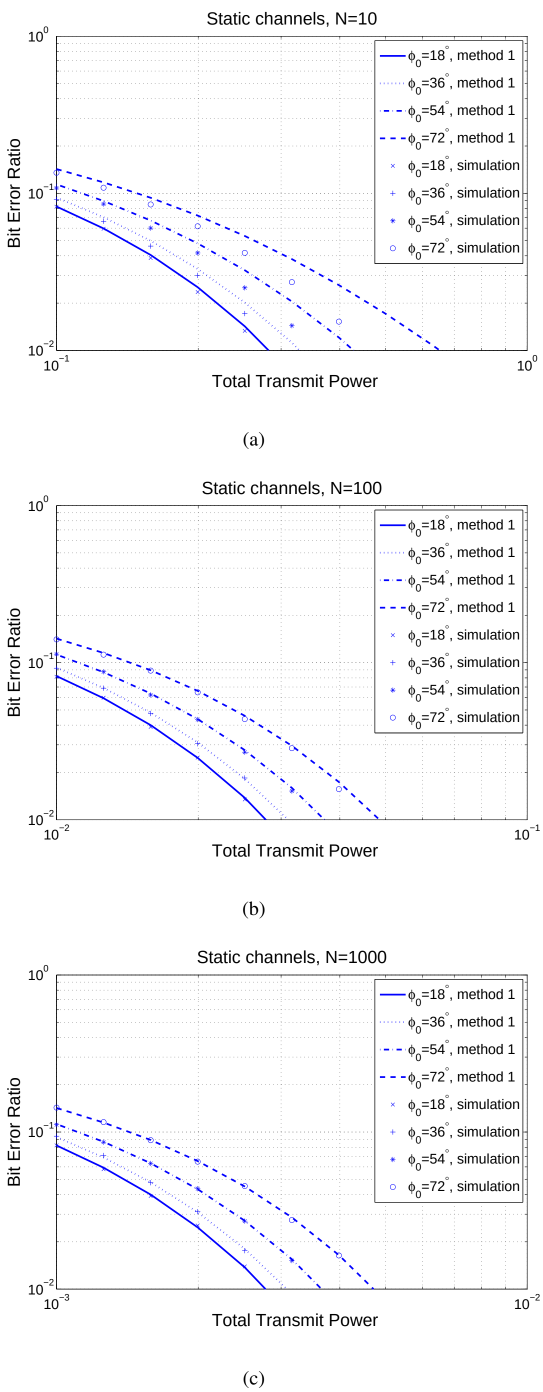


Fig. 2. Comparison of mathematical analysis based on method 1 with simulation results of BER versus total transmit power in the static channel scenario with different numbers of nodes N =(a)10, (b)100, and (c)1000, and different phase error ranges $\phi_0 = 18°, 36°, 54°, 72°$.

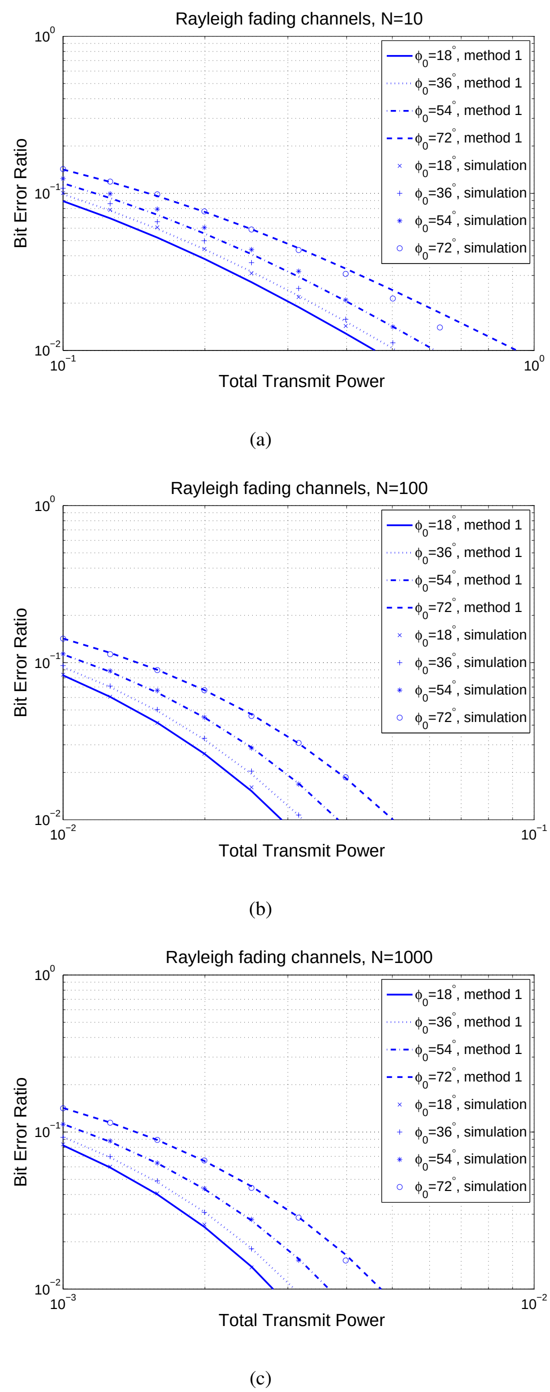


Fig. 3. Comparison of mathematical analysis based on method 1 with simulation results of BER versus total transmit power in the Rayleigh fading channel scenario with different numbers of nodes N =(a)10, (b)100, and (c)1000, and different phase error ranges $\phi_0 = 18°, 36°, 54°, 72°$.

see the program

can dramatically reduce the energy requirement for each sensor node subject to the same BER performance, and the number of nodes N has a much larger effect on BER performance than the phase error range ϕ_0.

From Fig. 2, we see, on the one hand, with a large number of nodes $N = 1000$, the BER analysis based on method 1 matches the simulation results accurately. On the other hand, with a small number of nodes $N = 10$, the BER analysis based on method 1 has a slight difference with the simulation results. This is due to the limitation that central limit theorem does not apply for a small number of nodes.

Fig. 3 shows the comparison of the simulation results with the mathematical analysis based on method 1 for BPSK modulation over *Rayleigh fading channels* with phase errors. The simulation results are also conducted over 10^5 symbols with different number of nodes $N = 10, 100, 1000$, and different phase error ranges $\phi_0 = 18°, 36°, 54°, 72°$. We also set $n(t) \sim CN(0, 1)$.

Similarly, from Fig. 3 we see, with large N, method 1 gives an accurate prediction of the BER, but with small N, method 1 gives a better prediction in the Rayleigh fading channel scenario than that in the static channel scenario.

By comparing the simulation results plotted in parts (a), (b), (c) in Fig. 3, we can also have the conclusion that increasing the number of nodes N can dramatically reduce the energy requirement for each sensor node subject to the same BER performance, and the number of nodes N has a much larger effect on BER performance than the phase error range ϕ_0.

By comparing the simulation results plotted in Fig. 2 with those in Fig. 3, we see when increasing the number of nodes N, the BER performance in the Rayleigh fading channel scenario comes close to that in the static channel scenario, which highlights the ability to mitigate fading through path diversity.

7. CONCLUSION

We have simulated the BER performance for BPSK modulation in distributed beamforming with phase errors in the static channel scenario and the Rayleigh fading channel scenario, where the results show a good match with our mathematical analysis. The whole beamforming process has been viewed as an equivalent channel and the system performance has been analyzed for different numbers of nodes and different phase error ranges. As the closed-form expression of BER is not easily obtained, we provide a method to approximate the BER results. Generally, method 1 gives a better prediction in the Rayleigh fading channel scenario than the static channel scenario. We are currently working on other approximations of the BER performance, such as method 2 outlined above. The effect of the energy limitation of each sensor node on the BER performance, and BER analysis for other modulation schemes in distributed beamforming with phase errors are also of particular interest for future work.

8. REFERENCES

[1] R. Mudumbai, G. Barriac, U. Madhow. *On the Feasibility of Distributed Beamforming in Wireless Networks*. IEEE Transactions on Wireless Communications, Vol. 6, No. 5, pp. 1754-1763, May 2007.

[2] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac. *Scalable Feedback Control for Distributed Beamforming in Sensor Networks*. Proceedings. International Symposium on Information Theory, Adelaide(SA), pp. 137-141, September 2005.

[3] P. Jeevan, S. Pollin, A. Bahai, and P.P. Varaiya. *Pairwise Algorithm for Distributed Transmit Beamforming*. IEEE International Conference on Communications, Beijing, pp. 4245-4249, May 2008.

[4] A.G. Marques, Xin Wang, and G.B. Giannakis. *Minimizing Transmit Power for Coherent Communications in Wireless Sensor Networks With Finite-Rate Feedback*. IEEE Transactions on Signal Processing, vol. 56, No. 9, pp. 4446-4457, September 2008.

[5] M. Tummala, C.C. Wai and P. Vincent. *Distributed Beamforming in Wireless Sensor Networks*. Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, California(USA), pp793-797, October 2005.

[6] H. Ochiai, P. Mitran, H.V. Poor, and V. Tarokh. *Collaborative Beamforming for Distributed Wireless Ad Hoc Sensor Networks*. IEEE Transactions on Signal Processing, Vol. 53, No. 11, pp. 4110-4124, November 2005.

[7] J.G. Proakis. Digital Communications 4^{th} ed., McGraw-Hill, Boston, 2001.

[8] W.C. Lindsey. *Error Probabilities for Rician Fading Multichannel Reception of Binary and N-ary Signals*. IEEE Transactions on Information Theory, Vol. 10, No. 4, pp. 339-350, October 1964.

[9] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. Numerical Recipes in C 2^{nd} ed., Cambridge University Press, 1992.

see the program

IMPLEMENTATION OF THE LEAST SQUARES CHANNEL ESTIMATION ALGORITHM FOR MIMO-OFDM SYSTEMS

Samuli Tiiro, Jari Ylioinas, Markus Myllylä, and Markku Juntti

Centre for Wireless Communications
University of Oulu
P.O. Box 4500, FI-90014 University of Oulu, Finland
{samuli.tiiro,jari.ylioinas,markus.myllyla,markku.juntti}@ee.oulu.fi

ABSTRACT

The least squares (LS) channel estimation algorithm for a multiple-input multiple-output (MIMO) system with orthogonal frequency division multiplexing (OFDM) is considered in this paper. Two architectures for the algorithm are presented and the architectures are implemented using fixed point arithmetic. The minimum word lengths for the implementations are determined through computer simulations for 2×2 and 4×4 MIMO systems with quadrature phase-shift keying (QPSK) modulation. Field-programmable gate array (FPGA) synthesis simulation is done for the architectures using the obtained word lengths which provides complexity and latency results. The algorithm implementations are shown to consume reasonably small amount of hardware resources.

1. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) [1] is an attractive air interface for high-rate communication systems with large bandwidths. OFDM divides the frequency selective fading channel into several parallel flat fading subchannels and, thus, efficiently mitigates the multipath delay spread. The use of multiple transmit and receive antennas, i.e., multiple-input multiple-output (MIMO) channel provides the potential for tremendous capacity increase [2, 3]. Consequently, MIMO techniques [4, 5] have gained remarkable attention during the recent years. Combining MIMO processing [6, 7] with OFDM is the key enabling technology for several current and future broadband wireless access systems and standards. Important examples include wireless local area network (WLAN) or Wireless Fidelity (WiFi) systems [8], wireless metropolitan area network (WMAN), like Worldwide Interoperability for Microwave Access (WiMAX) [9], and the Third Generation (3G) cellular system Long Term Evolution (LTE) [10].

This research was financially supported in part by Tekes, the Finnish Funding Agency for Technology and Innovation, Nokia, Texas Instruments, Nokia Siemens Networks and Elektrobit. Catapult C Synthesis license was provided by Mentor Graphics.

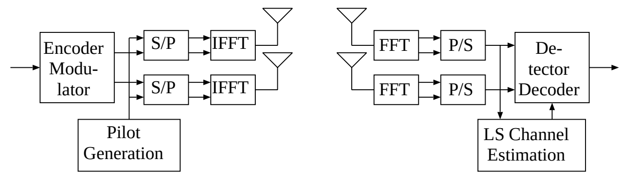


Fig. 1. 2×2 MIMO-OFDM system.

The channel state information (CSI) between each transmit and receive antenna pair is required at the receiver to coherently detect the information. Therefore, channel estimation is an important task that is required in wireless communication systems. Pilot-based channel estimation is a popular technique due to its simple implementation. In pilot-based channel estimation [11], known training symbols are inserted in the data stream, and the receiver can then use these symbols to perform channel estimation and synchronization. The least squares (LS) estimation is a common method for pilot-based channel estimation, as it offers good performance with reasonable complexity.

In this paper, a pilot-based channel estimator design for a MIMO-OFDM system is described and two channel estimator architectures are presented for the LS estimator. A word length study is done for both architectures in order to determine the minimum fixed point word lengths for the field-programmable gate array (FPGA) hardware implementation. The minimum word lengths are then used for obtaining hardware complexity results for the presented estimator architectures.

The rest of this paper is organized as follows. Section 2 describes the system model and the channel estimation algorithm is covered in Section 3. The proposed channel estimator architectures and the word length study can be found in Section 4. Synthesis results are presented in Section 5, and finally, paper is concluded in Section 6.

see the program

2. SYSTEM MODEL

In this paper, we consider a system model similar to the one presented in [12], and it is described as follows. Let us assume a MIMO-OFDM system with N_T transmit antennas and N_R receive antennas. As an example, a block diagram of a system with $N_\mathrm{T} = N_\mathrm{R} = 2$ is illustrated in Fig. 1. The discrete time channel impulse response (CIR) from the ith transmit antenna to the jth receive antenna at time n is denoted by $\boldsymbol{h}_{ij}[n] \in \mathbb{C}^L$ where L denotes the number of channel tap coefficients. The CIR for the entire $N_\mathrm{T} \times N_\mathrm{R}$ system at time n is defined in matrix form as

$$\boldsymbol{H}[n] = \begin{pmatrix} \boldsymbol{h}_{11}[n] & \dots & \boldsymbol{h}_{1N_\mathrm{R}}[n] \\ \vdots & \ddots & \vdots \\ \boldsymbol{h}_{N_\mathrm{T}1}[n] & \dots & \boldsymbol{h}_{N_\mathrm{T}N_\mathrm{R}}[n] \end{pmatrix} \in \mathbb{C}^{N_\mathrm{T}L \times N_\mathrm{R}}. \tag{1}$$

Correspondingly, we define the channel frequency response (CFR) matrix at time n

$$\tilde{\boldsymbol{H}}[n] = \tag{2}$$
$$\begin{pmatrix} \boldsymbol{F}_K^L\boldsymbol{h}_{11}[n] & \dots & \boldsymbol{F}_K^L\boldsymbol{h}_{1N_\mathrm{R}}[n] \\ \vdots & \ddots & \vdots \\ \boldsymbol{F}_K^L\boldsymbol{h}_{N_\mathrm{T}1}[n] & \dots & \boldsymbol{F}_K^L\boldsymbol{h}_{N_\mathrm{T}N_\mathrm{R}}[n] \end{pmatrix} \in \mathbb{C}^{N_\mathrm{T}K \times N_\mathrm{R}},$$

where $\boldsymbol{F}_K^L$ denotes the L first left columns of a K-point DFT matrix and K is the number of subcarriers used for transmitting data. The received signal block at the receiver at time n after the DFT can be presented as

$$(\boldsymbol{r}_1[n] \dots \boldsymbol{r}_{N_\mathrm{R}}[n]) = \boldsymbol{R}[n] = \boldsymbol{X}[n]\tilde{\boldsymbol{H}}[n] + \boldsymbol{\eta}[n], \tag{3}$$

where $\boldsymbol{X}[n] = \big(\mathrm{diag}(\boldsymbol{x}_1[n]) \dots \mathrm{diag}(\boldsymbol{x}_{N_\mathrm{T}}[n])\big) \in \mathbb{C}^{K \times N_\mathrm{T}K}$ and $\boldsymbol{x}_i[n] \in \Xi^K, i = 1, \dots, N_\mathrm{T}$ are the data vectors sent from the ith transmit antenna, Ξ denotes the modulation alphabet, $\boldsymbol{r}_j[n] \in \mathbb{C}^K$ is the data vector received at the jth receive antenna, and $\boldsymbol{\eta}[n] \in \mathbb{C}^{K \times N_\mathrm{R}}$ is the noise matrix at time n. White Gaussian noise is assumed with covariance matrix $\boldsymbol{\Sigma}_\eta = \mathrm{E}(\mathrm{vec}(\boldsymbol{\eta})\mathrm{vec}(\boldsymbol{\eta})^H) = \sigma_\eta^2 \boldsymbol{I}_{KN_\mathrm{R}}$. From (2) and (3), it can be seen that by combining the DFT matrices with the data matrix $\boldsymbol{X}[n]$, the received signal can also be expressed in terms of CIR as

$$\boldsymbol{R}[n] = \boldsymbol{\mathcal{C}}[n]\boldsymbol{H}[n] + \boldsymbol{\eta}[n], \tag{4}$$

where $\boldsymbol{\mathcal{C}}[n] = (\boldsymbol{\mathcal{C}}_1[n] \dots \boldsymbol{\mathcal{C}}_{N_\mathrm{T}}[n]) \in \mathbb{C}^{K \times N_\mathrm{T}L}$ and $\boldsymbol{\mathcal{C}}_i[n] = \mathrm{diag}(\boldsymbol{x}_i[n])\boldsymbol{F}_K^L$.

The channel model used in this work is a Ricean fading channel with temporal, spatial and spectral correlation proposed in [13].

3. PILOT-BASED LEAST SQUARES CHANNEL ESTIMATION

We consider an OFDM frame structure where each frame consists of N_SYMB consecutive OFDM symbols. The first and the third last are reserved for transmitting the pilot sequences. Out of the N_C available subcarriers, K are used for transmitting data and the pilot sequences are transmitted over all K tones. The pilot symbols are used for obtaining an estimate of the channel impulse response which is then used for the coherent detection of the transmitted symbols. From (4), the LS estimate of the CIR becomes

$$\hat{\boldsymbol{H}}_\mathrm{LS}[n] = (\boldsymbol{\mathcal{C}}^H\boldsymbol{\mathcal{C}})^{-1}\boldsymbol{\mathcal{C}}^H\boldsymbol{R}[n], \tag{5}$$

where n denotes the time indices of the pilot symbols. The time index of $\boldsymbol{\mathcal{C}}$ can be neglected as the pilot sequences do not depend on the block index n. It should be noted that if the noise $\boldsymbol{\eta}$ in (4) is white and Gaussian, then the LS estimator is also the maximum likelihood (ML) estimator and optimal in the sense that it attains the Cramer-Rao lower bound (CRLB).

The computation of the LS estimate requires an $N_\mathrm{T}L \times N_\mathrm{T}L$ matrix inversion which is computationally an intensive task. It was shown in [14] that with a proper pilot sequence design, the LS channel estimator for MIMO-OFDM can be simplified so that the matrix inversion is not needed while the estimator still achieves the minimum mean square error (MSE). In this work, local orthogonal pilot sequences were used and the sequence for ith transmit antenna is defined as

$$\boldsymbol{c}_i[k] = c_0[k]e^{j(2\pi\bar{L}/K)ik} \in \mathbb{C}^K, \tag{6}$$

where $\bar{L} \triangleq \lfloor \frac{K}{N_\mathrm{T}} \rfloor \geq L$, $i = 1, \dots, N_\mathrm{T} - 1$ and $k = 0, \dots, K-1$. The training sequence at the first antenna is set to a random sequence of symbols for which $|\boldsymbol{c}_0[k]| = 1\, \forall\, k$. The presented pilot sequence design reduces the matrix $(\boldsymbol{\mathcal{C}}^H\boldsymbol{\mathcal{C}})$ to a diagonal matrix $K\boldsymbol{I}_{N_\mathrm{T}L}$ and, thus, (5) can be rewritten as

$$\hat{\boldsymbol{H}}_\mathrm{LS}[n] = \frac{1}{K}\boldsymbol{\mathcal{C}}^H\boldsymbol{R}[n]. \tag{7}$$

4. ARCHITECTURE DESIGN AND WORD LENGTH STUDY

4.1. LS Estimator Architecture

As can be seen from (7), the channel estimator only needs to compute a complex matrix-matrix multiplication followed by scaling with a real value $\frac{1}{K}$. In OFDM systems, the number of subcarriers is usually chosen to be a power of two due to the FFT used in modulation and demodulation. In the case where all subcarriers are used for data transmission, i.e., $K = N_\mathrm{C}$ the multiplication can therefore be performed with a simple shift operation. An architecture for performing the computation in (7) is designed for 2×2 and 4×4 MIMO systems using quadrature phase-shift keying (QPSK). The top level block diagram for the proposed architecture for a 2×2 system is shown in Fig. 2. The channel estimator consists of $N_\mathrm{T}N_\mathrm{R}$ vector product units, and it

see the program

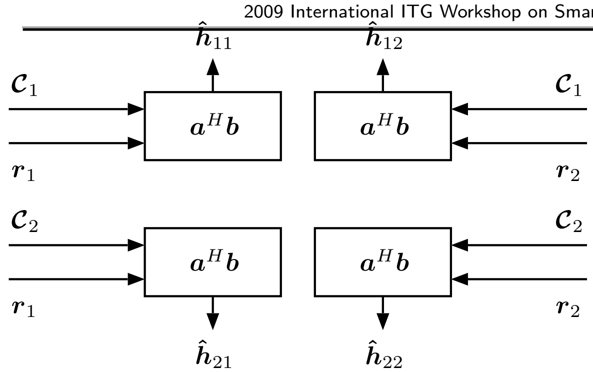


Fig. 2. Top level structure of the LS estimator for a 2×2 MIMO system.

is therefore easily scalable for different MIMO configurations. Each of the vector product units compute the CIR tap estimates for the corresponding MIMO link as

$$\hat{\boldsymbol{h}}_{ij}[n;l] = \frac{1}{K}\mathcal{C}_{il}^H \boldsymbol{r}_j[n] \tag{8}$$

where $i = 1,\ldots,N_\mathrm{T}$, $j = 1,\ldots,N_\mathrm{R}$, $l = 0,\ldots,L-1$ and $\mathcal{C}_{il}$ denotes the lth column of $\mathcal{C}_i$. In one computing cycle, each of the units can compute one channel tap estimate. Therefore, the computation of a single CIR vector $\hat{\boldsymbol{h}}_{ij}$ takes L computing cycles in total. As there are no data dependencies between the different CIR vectors, it is possible to compute all of them simultaneously. In the proposed architecture, all of the $N_\mathrm{T}N_\mathrm{R}$ vector product units are ran in parallel, and, therefore, the entire CIR matrix $\hat{\boldsymbol{H}}$ can be computed in L computing cycles.

A block diagram of the vector product unit is shown in Fig. 3. The unit has four scalar inputs and two scalar outputs. The input vectors' elements are read from inputs $1-4$, where 1 and 2 are for the real and imaginary parts of $\mathcal{C}_{il}$ and 3 and 4 are for the real and imaginary parts of $\boldsymbol{r}_j$, respectively. The vector product is computed in a straightforward manner by multiplying the elements of the two complex input vectors and then adding the product to the sum of previous elements' products. After all K elements of the input vectors are multiplied, the final sum is multiplied by the scaling factor $\frac{1}{K}$. The final result, i.e., the real and imaginary parts of $\hat{\boldsymbol{h}}_{ij}[n;l]$ can then be read from outputs 14 and 15, respectively. The complex multiplication of the vector elements is done with three real multipliers in order to simplify the implementation. For two complex numbers $z_1 = a + bi$ and $z_2 = c + di$, where $a,b,c,d \in \mathbb{R}$ and $i = \sqrt{-1}$, the product $z_1^* z_2$ can be expressed in terms of real and imaginary parts as $\Re\{z_1^* z_2\} = ac + bd$ and $\Im\{z_1^* z_2\} = ad - bc$ which requires four real multiplications. Same result can be achieved with only three real multiplications by rearranging the equations as

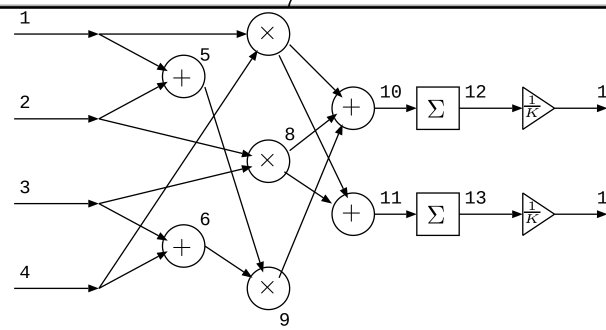


Fig. 3. Block diagram of the vector product unit.

$$\begin{aligned}\Re\{z_1^* z_2\} &= (a+b)(c+d) - ad - bc \\ \Im\{z_1^* z_2\} &= ad - bc.\end{aligned} \tag{9}$$

In (9), one multiplication is exchanged to three additions. This approach still results in a lower overall complexity as multiplication is computationally much more expensive operation than an addition. It should be noted that in Fig. 3 the signal sign flips required for the subtraction operations are left out.

Computing the LS channel estimate by using the precomputed matrix $\mathcal{C}$ requires memory to store the $N_\mathrm{T}LK$ complex matrix coefficients. As matrix $\mathcal{C}$ consists of diagonal matrices diag$(\boldsymbol{c}_i), i = 1,\ldots,N_\mathrm{T}$ multiplied by the partial DFT matrix $\boldsymbol{F}_K^L$, it is possible to reduce the memory needed for the precomputed data. By adding one complex multiplication into the structure shown in Fig. 3, the LS estimate can be computed as

$$\hat{\boldsymbol{h}}_{ij}[l] = \frac{1}{K}(\boldsymbol{c}_i \odot (\boldsymbol{F}_K^L)_l)^H \boldsymbol{r}_j \tag{10}$$

where $(\boldsymbol{F}_K^L)_l$ is the lth column of the partial DFT matrix, $i = 1,\ldots,N_\mathrm{T}$, $j = 1,\ldots,N_\mathrm{R}$, $l = 0,\ldots,L-1$ and $\odot$ denotes elementwise vector multiplication. With this approach, only the pilot sequences $\boldsymbol{c}_i$ and the partial DFT matrix $\boldsymbol{F}_K^L$, i.e., $K(L+N_\mathrm{T})$ complex coefficients in total need to be stored in memory. This is, however, done on the expense of the computational complexity as an additional complex multiplication is required in each vector product unit. A block diagram of the vector product unit using this alternative method is shown in Fig. 4.

4.2. Word Length Study

In practical applications, fixed point presentation of the numbers is usually preferred over floating point because of its lower complexity. The downside of using fixed point numbers is the limited range of signals and loss of precision. A

see the program

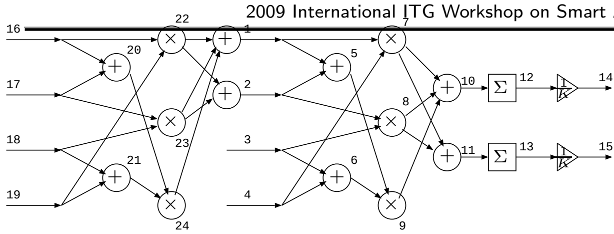


Fig. 4. Block diagram of the alternative vector product unit.

Table 1. System parameters for different channel profiles.

	EPA	ETU
CP length	4.69 μs	16.67 μs
Symbols per frame	7	6
OFDM symbol duration	71.35 μs	83.33 μs
Number of channel taps	9	7

word length study of the LS estimator was done in order to obtain the minimum word lengths for which the fixed point implementation does not exhibit significant performance degradation.

The word lengths were determined with computer simulations using 3G LTE system parameters with two different channel profiles. The channel profiles considered in this work were Extendend Typical Urban (ETU) and Extended Pedestrian-A (EPA) [15]. The simulation parameters for both channel profiles are presented in Table 1. The systems were simulated with mobile velocity of 70 km/h. 2×2 and 4×4 MIMO systems with QPSK were considered. Out of the $N_\mathrm{C} = 512$ available subcarriers $K = 300$ were used for transmitting data. System bandwidth was 5 MHz with the carrier frequency of 2.4 GHz. The data was coded using a $\frac{1}{2}$-rate turbo code with 8 decoder iterations and the coding was performed over a single OFDM frame. As the receiver front-end, a K-best list sphere detector (LSD) [16] was used with list sizes of 8 for 2×2 QPSK system and 32 for 4×4 QPSK system.

In a system where there are no null subcarriers present, it is sufficient to consider only the L first taps of the CIR as the rest of the taps only contain noise energy. By excluding these noise taps, the estimation accuracy is improved. This also results in simpler channel estimation as less taps are required to be estimated. Due to the null subcarriers used in the system considered here the data is demodulated with a 512-point FFT while the LS estimation is performed using a K-point DFT matrix. This causes some of the energy to leak over all K taps. By only estimating the L first taps of the CIR some performance loss will result due to the leakage energy. It was shown in [17], that if some additional taps beyond the L first are also included in the estimation, the MSE of the LS estimator can be improved. This will increase the size of the data matrix used in the LS estimation and, thus, results in higher estimator complexity. However, most of the leakage energy is contained only in a few taps outside the L first ones and therefore the increase in complexity is minor. An optimal number of additional taps exist for each SNR point that results in the smallest MSE [17] but in order to keep the estimator design simple, a fixed value for each antenna configuration was determined with computer simulations.

Table 2. Additional channel taps considered in the LS estimation.

MIMO	Channel profile	Channel length L	Taps used L'
2×2	EPA	2	$0, \ldots, 4, 297, \ldots, 299$
4×4	EPA	2	$0, \ldots, 4, 297, \ldots, 299$
2×2	ETU	23	$9, \ldots, 26, 296, \ldots, 299$
4×4	ETU	23	$0, \ldots, 25, 297, \ldots, 299$

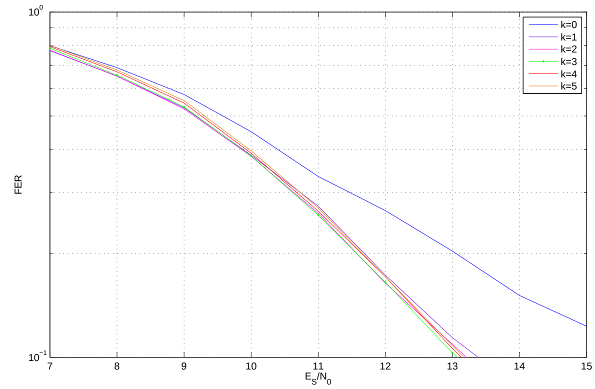


Fig. 5. The FER vs. SNR of the LS estimator for 2×2 QPSK in EPA channel with additional channel taps included in the estimation.

Estimators with different numbers of additional taps were simulated. The taps considered in each case are $\boldsymbol{h}[n] = (h[n;0] \ldots h[n;L+k-1] h[n;K-k] \ldots h[n;K-1])$, i.e., k additional taps on each side and, thus, $2k$ additional taps in total. The simulation length in each case was 1000 frames and the frame error rate (FER) was measured. As an example, the results for EPA channel with 2×2 antenna configuration can be found in Fig. 5, where it can be seen that $k = 3$ results in the best FER performance. It should be noted that the performance difference between $k = 1, \ldots, 5$ is minor and therefore even smaller values of k could be used to obtain a better trade-off between the performance and the complexity. The number of taps that resulted in the best FER performance in each case were used in this work and the taps considered can be found in Table 2.

To obtain the fixed point word lengths, a C language implementation of the channel estimation algorithm was written. The fixed point variables were implemented using Mentor Graphics algorithmic C data types [18]. The integer

see the program

widths, the fraction widths and the signedness were defined for all signals in the implementation. The range for a signed fixed point signal is from $(-0.5) \times 2^I$ to $(0.5 - 2^{-W}) \times 2^I$, where W is the total word length and I is the number of bits reserved to represent the integer part.

To determine the integer widths for the input and output signals of the LS estimator, a floating point simulation was performed. Using the signal samples collected during the simulation, a histogram was drawn for each signal from which the signal ranges could be determined. The internal word lengths as well as the fraction widths of the input and output signals were determined from the fixed point simulations. Signal word lengths were studied one by one by comparing the fixed point performance to the corresponding floating point implementation. The word length of each signal was reduced until the minimum that did not exhibit noticeable performance loss was found. The length of each simulation was 10000 frames and the frame error rate (FER) was used as the performance measure. The minimum word lengths were determined so that the fixed point performance was close to the floating point performance up to the FER of 10^{-1}. The obtained minimum word lengths can be found in Table 3. The word lengths therein are presented as (a, b), where a represents the number of bits reserved for the integer part and the sign and b denotes the number of bits reserved for the fraction part.

Table 3. Word lengths for the LS estimator.

	EPA		ETU	
Signal in Figure 3	2×2	4×4	2×2	4×4
1,2	(2,5)	(2,5)	(2,6)	(2,7)
3,4	(2,5)	(4,4)	(3,4)	(4,5)
5	(2,5)	(2,5)	(2,7)	(2,8)
6	(4,3)	(4,4)	(4,5)	(4,6)
7,8	(3,7)	(4,9)	(3,10)	(4,10)
9	(4,7)	(5,8)	(4,10)	(5,11)
10	(3,12)	(4,9)	(4,12)	(5,10)
11	(3,12)	(4,9)	(4,12)	(4,10)
12,12	(10,7)	(10,9)	(9,9)	(8,10)
14,15	(2,8)	(2,8)	(0,10)	(0,10)
Signal in Figure 4				
16,17	(2,5)	(2,5)	(2,5)	(2,7)
18,19	(2,5)	(2,6)	(2,5)	(2,7)
20	(2,5)	(2,5)	(3,5)	(3,6)
21	(2,5)	(2,6)	(3,5)	(3,6)
22	(1,7)	(1,7)	(1,5)	(1,8)
23	(2,5)	(2,6)	(2,5)	(2,7)
24	(2,5)	(2,6)	(3,4)	(2,7)

The performance of the fixed point implementations using EPA channel profile can be found in Fig. 6, where fixed point 1 denotes the architecture using the vector product units shown in Fig. 3 and fixed point 2 denotes the second presented architecture using the vector product units described in Fig. 4. Correspondingly, the results for the estimator using ETU channel profile can be seen in Fig. 7. All of the presented FER results are average values over the 10000 simulated frames and they were simulated using the minimum word lengths found in Table 3. It can be seen that the fixed point implementations provide near floating point performance above the FER of 10^{-1} which was chosen as the threshold value as it corresponds to the operating point of the system.

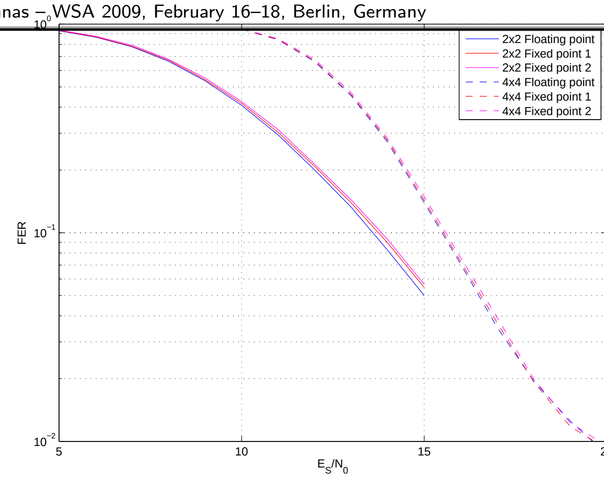


Fig. 6. FER vs. SNR for the LS estimator with EPA channel profile.

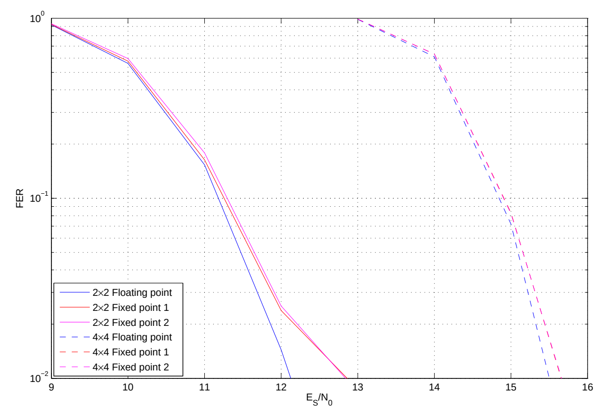


Fig. 7. FER vs. SNR for the LS estimator with ETU channel profile.

5. SYNTHESIS RESULTS

The channel estimation algorithm was implemented using the presented architectures and synthesis simulation was done.

Table 4. Memory requirements for the LS estimator implementations.

Estimator	$\mathcal{C}$	$\boldsymbol{c}$	$\boldsymbol{F}_K^{L'}$	$\boldsymbol{R}$	Total [kbit]
LS 1 (2×2 EPA)	2×33600	-	-	2×4200	75.6
LS 2 (2×2 EPA)	-	2×4200	33600	2×4200	50.4
LS 1 (4×4 EPA)	4×33600	-	-	4×4800	153.6
LS 2 (4×4 EPA)	-	4×4200	38400	4×4800	74.4
LS 1 (2×2 ETU)	2×148800	-	-	2×4800	307.2
LS 2 (2×2 ETU)	-	2×4200	130200	2×4800	148.2
LS 1 (4×4 ETU)	4×156600	-	-	4×5400	648.0
LS 2 (4×4 ETU)	-	4×5400	156600	4×5400	199.8

The system parameters used with the implementations were those presented in Section 4.2. Implementation of the estimator architectures was studied using Mentor Graphics Catapult C tool. The algorithm was implemented in C code and fixed point representation with the obtained minimum word lengths were used for defining the signals. From the C code implementation, the tool created a VHDL code for which a synthesis simulation was performed using Mentor Graphics Precision RTL Synthesis software. The synthesis simulation for the architectures was done by using Xilinx Virtex-4 FPGA with 75 MHz operation frequency as the target platform. The device utilization for each architecture are shown in Table 5. The target latency for the implementations was chosen to be less than the OFDM burst duration $T_{\mathrm{B}} = T_{\mathrm{CP}} + T_{\mathrm{OFDM}}$ so that the channel estimate for one symbol could be computed during the reception of the next symbol. The OFDM burst durations are $T_{\mathrm{B}_{\mathrm{EPA}}} = 71.354\mu$s and $T_{\mathrm{B}_{\mathrm{ETU}}} = 83.333\mu$s for the EPA and ETU channel models, respectively. The vector product units for the EPA channel are pipelined in order to result in a faster implementation and to meet the latency requirements. Due to the longer length of the ETU channel, more channel taps need to be computed than in the case of the EPA channel. Therefore, the vector product units for the ETU implementation also employ partial unrolling of the computation loops in addition to pipelining. In partial unrolling, some iterations of the computation loop are executed in parallel and it will therefore significantly speed up the implementation. However, this is done on the expense of increased complexity. For the ETU channel, an unrolling factor of 2 was found to be sufficient to meet the latency requirements. Due to the unrolling, a larger memory bandwidth is required so that the vector product units can read two coefficients from the same memory during one clock cycle.

In the synthesis simulation, the memory required for the precomputed matrices was assumed to be available outside the LS estimator and therefore was not included in the device utilization figures. Using the minimum word lengths and the data matrix dimensions, the amount of required memory can be computed. The memory requirements for each design are shown in Table 4. It can be seen that the estimator 1 has lower complexity with every configuration as there is less computation required. The reduced memory requirements of estimator 2, however, are more significant, especially with 4×4 MIMO. As the memory requirements for each design do not exceed the available memory on the platforms [19] used in the synthesis simulations, the choice of architecture can be done based on whether a faster execution or less computation is desired.

Table 5. Synthesis results for the LS estimator implementations.

Estimator	Device	CLB Slices	DSP48s	Latency
LS 1 (2×2 EPA)	XC4VFX12	277	14	32.76 μs
LS 2 (2×2 EPA)	XC4VFX12	317	20	32.87 μs
LS 1 (4×4 EPA)	XC4VLX40	1919	32	34.15 μs
LS 2 (4×4 EPA)	XC4VLX40	2424	32	34.25 μs
LS 1 (2×2 ETU)	XC4VLX25	519	26	64.91 μs
LS 2 (2×2 ETU)	XC4VLX40	749	32	65.32 μs
LS 1 (4×4 ETU)	XC4VLX100	5647	32	65.75 μs
LS 2 (4×4 ETU)	XC4VLX100	7372	32	66.13 μs

The simulated architectures were designed for a system with $K = 300$ data subcarriers. The architectures are quite easily scalable to support different kind of configurations. However, in systems that use more subcarriers for data transmission, maintaining the same latency might become an issue. More intensive loop unrolling can be done for the vector product units to maintain or even to reduce the latency but this will dramatically increase the hardware complexity. The combination of pipelining and unrolling of the computation loops will also increase the number of memory reads per clock cycle and, therefore, larger unrolling factors will also require larger memory bandwidths. Systems with more data subcarriers also require larger DFT matrix and longer pilot sequences to be used and, thus, more memory needs to be available for the precomputed data.

6. CONCLUSIONS

The LS channel estimation algorithm for MIMO-OFDM was considered. A channel estimator architecture based on precomputed data was designed. A modified architecture that

see the program

requires less precomputed data to perform the estimation was also presented. The architectures were implemented using fixed point arithmetic for 2×2 and 4×4 MIMO-OFDM systems with 300 data subcarriers and QPSK modulation. Minimum word lengths for each design were studied and they were shown to provide near floating point performance up to 10^{-1} FER. A synthesis simulation was done for the architectures from which latency and complexity results were obtained. The designs were shown to be implementable on the available FPGA platforms as well as to meet the latency requirements based on the system parameters.

7. REFERENCES

[1] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, Artech House, London, 2000.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications, Kluwer Academic Publishers*, vol. 6, pp. 311–335, 1998.

[3] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[5] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct 1998.

[6] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 281–302, 2003.

[7] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications — A key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[8] ANSI/IEEE Standard 802.11,1999 Edition (R2003), "Information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," 2003.

[9] IEEE Standard 802.16, "IEEE standard for local and metropolitan area networks part 16: Air interface for fixed broadband wireless access systems," 2004.

[10] 3rd Generation Partnership Project (3GPP), "Physical layer aspect for evolved universal terrestrial radio access (UTRA)," Tech. Rep., 3rd Generation Partnership Project (3GPP), 2006.

[11] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[12] H. Miao and M. Juntti, "Space-time channel estimation and performance analysis for wireless MIMO-OFDM systems with spatial correlation," *IEEE Trans. on Vehicular Technology*, vol. 54, no. 6, pp. 2003 – 2016, Nov. 2005.

[13] E. Kunnari and J. Iinatti, "Stochastic modelling of rice fading channels with temporal, spatial and spectral correlation," *IET Communications*, vol. 1, no. 2, pp. 215–224, Apr. 2007.

[14] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 67–75, Jan. 2002.

[15] 3rd Generation Partnership Project (3GPP); Technical Specification Group Radio Access Network, "Evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception (release 8)," Tech. Rep., 3rd Generation Partnership Project (3GPP), 2007.

[16] K. Wong, C. Tsui, R. K. Cheng, and W. Mow, "A VLSI architecture of a K-best lattice decoding algorithm for MIMO channels," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Scottsdale, AZ, May 26-29 2002, vol. 3, pp. 273–276.

[17] J.-J. van de Beek, O. Edfors, M. Sandell, S.K. Wilson, and P.O. Borjesson, "On channel estimation in OFDM systems," in *Proceedings of the IEEE Vehicular Technology Conference*, Chicago, IL, USA, July 25–28 1995, vol. 2, pp. 815–819.

[18] "Mentor Graphics (accessed on 17 September 2008) Algorithmic C Data Types. URL: http://www.mentor.com/products/esl/high_level_synthesis/ac_datatypes.cfm," .

[19] "Xilinx (accessed on 28 September 2008) Virtex-4 Product Table. URL: http://www.xilinx.com/publications/matrix/matrix_v4.pdf," .

see the program

A LOW COMPLEXITY SPACE-FREQUENCY MULTIUSER SCHEDULING ALGORITHM

Ana I. Pérez-Neira[1,2], Pol Henarejos[1], Velio Tralli[3], Miguel A. Lagunas[1,2]

email: anuska@gps.tsc.upc.edu, pol@redyc.com, vtralli@ing.unife.it, m.a.lagunas@cttc.es
(1) Dept. of Signal Theory and Communications - Universitat Politècnica de Catalunya (UPC) – Spain
(2) Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) - Spain
(3) ENDIF - Engineering Department University of Ferrara - CNIT - Italy

ABSTRACT

This work presents a resource allocation algorithm in K-user, M-subcarrier and NT-antenna systems for on-line scheduling. To exploit temporal diversity and to reduce complexity, the ergodic sum rate is maximized instead of the instantaneous one. Dual optimization is applied to further diminish complexity together with a stochastic approximation, which is more suitable for online algorithms. Weighted sum rate is considered so that users can be either prioritized by higher layers or differentiated by proportional rate constraints. The performance and complexity of this algorithm is compared with well-known benchmarks and also evaluated under real system conditions for the MIMO Broadcast channel.

1. INTRODUCTION

In a Multi-User MIMO (MU-MIMO) spatial multiplexing scheme, multiple users are scheduled in the same resource block. MU-MIMO is a promising way to increase system throughput and there is a growing interest on the topic as [1,2, 3, 4] shows. Recently attention has been paid to the combination of spatial diversity multiple access systems and frequency domain packet scheduling [5,6,7,8,9]. Specifically, in [6] the authors present a low complexity sum-power constraint iterative waterfilling that is capacity achieving. It improves the convergence of [3] and is probably convergent. In [8] the authors address the problem of feedback reduction. The present paper aims at both, low complexity and reduced feedback. In contrast to [7], in order to further reduce complexity for on line implementation we follow a dual decomposition strategy and a stochastic approximation. In order to reduce feedback load the paper resorts to opportunistic strategies that solve the spatial scheduling. More specifically, an efficient algorithm for optimal beam subset and user selection is performed to find the best trade-off between the multiplexing gain and the multiuser interference in the opportunistic scheme when the number of users is not high. In summary, this paper proposes a joint spatial and frequency scheduler that allows on-line implementation and only requires partial or low feedback and a low-complexity implementation.

This paper is organized as follows. The space-frequency scheduler is formulated in Section 2 and the distributed scheme that is proposed based on dual optimization is presented in Section 3. Section 4 explains the low complexity ergodic algorithm, together with an evaluation of its complexity. The numerical results are presented in Section 5, and Section 6 concludes the paper.

2. PROBLEM FORMULATION

We consider an OFDMA scenario with M subcarriers and K users. Each user k is single antenna and receives simultaneously N_T signals, which can come from different spatial locations, antennas or beams. Only one of the N_T signals is intended for user k. The received signal by user k on the subcarrier m is given by

$$y_{k,m} = a_{k,m}\sqrt{p_{k,m}}\alpha_{k,m,k_q}s_{k,m} + \sum_{s\neq k} a_{s,m}\sqrt{p_{s,m}}\alpha_{k,m,s_q}s_{s,m} + w_{k,m} \quad (1)$$

where, $\mathbf{a}=\{a_{k,m}\}$ is the set of binary allocation variables, i.e. $a_{k,m}=1$ if user k is scheduled on frequency m, $a_{k,m}=0$ otherwise, and $\mathbf{p}=\{p_{k,m}\}$ is the set of allocated powers, $s_{k,m}$ is the information signal of user k through frequency m, $E\{|s_{k,m}|^2\}=1$. Finally, α_{k,m,k_q} denotes the equivalent channel seen by the kth user at frequency m with respect to the qth beam, antenna or transmitter associated with user k. For instance, in the case of a MISO (Multiple input single output) broadcast channel, $|\alpha_{k,m,k_q}|^2 = c_{k,m,k_q} = |\mathbf{h}_{k,m}^T\mathbf{b}_{m,k_q}|^2$, where $\mathbf{b}_{m,k_q}$ is the beamforming vector that is associated with user k and that is obtained from the set of beams $k_q \in \{1,\dots,N_T\}$. In this case the number of interference terms in (1) is equal to N_T-1. From a viewpoint of information theory, the model in (1) could correspond either to a broadcast channel or to an interference channel with N_T transmitters and K receivers. In spite of the big gains in spectral efficiency that can be obtained by incorporating multiantenna transmission to a multicarrier system, an evident drawback of this scenario is the increased design complexity. In other words, multiantenna, multiuser and multicarrier channels significantly increase the set of design parameters and degrees of freedom at the PHY layer. In this work, the focus is on the optimization of the PHY layer parameters with low complexity burden. Concerning the optimality criteria, we consider the problem of rate maximization in (2) with power constraints and also proportional rate constraints.

see the program

$$\begin{aligned}
&\max_{\mathbf{a},\mathbf{p}} R(\mathbf{a},\mathbf{p}) = \mathrm{E}_{\gamma}\left\{\sum_{k=1}^{K}\sum_{m=1}^{M}\log_2\left(1+\gamma_{k,m,k_q}(\mathbf{a},\mathbf{p})\right)\right\}\\
&s.t.\quad \mathrm{E}\left\{\sum_{m=1}^{M}\log_2\left(1+\gamma_{k,m,k_q}(\mathbf{a},\mathbf{p})\right)\right\} \geq \phi_k R, \quad k=1,\ldots,K\\
&\qquad \sum_{k=1}^{K}\phi_k = 1\\
&\qquad \mathrm{E}_{\gamma}\left\{\sum_{m=1}^{M}\sum_{k=1}^{K} a_{k,m}p_{k,m}\right\} \leq \overline{P}\\
&\qquad a_{k,m} \in \{0,1\} \quad m=1,\ldots,M\\
&\qquad p_{k,m} \geq 0 \qquad m=1,\ldots,M \quad k=1,\ldots,K
\end{aligned} \tag{2}$$

with

$$\gamma_{k,m,k_q}(\mathbf{a},\mathbf{p}) = \frac{a_{k,m}p_{k,m}c_{k,m,k_q}}{\sigma^2 + \sum_{s\neq q}^{N_T} a_{s,m}p_{s,m}c_{k,m,k_s}} \tag{3}$$

where **a** and **p** are vectors whose components are $a_{k,m}$ and $p_{k,m}$, respectively. $\gamma_{k,m,k_q}(\mathbf{a},\mathbf{p})$ is the SINR (Signal to Interference and Noise Ratio) of user k at frequency m and associated with beam q and c_{k,m,k_q} denotes the equivalent channel power gain seen by the kth user at frequency m with respect to the qth beam, antenna or transmitter. The formulation of the sum rate in (2) indicates that at each frequency up to N_T transmissions can be spatially multiplexed. We assume that $\gamma_{k,m,k_q}(\mathbf{a},\mathbf{p})$ are known by the N_T transmitters by means of partial channel feedback. For instance, this would be the case of a broadcast channel where the Base Station (BS) has perfect SINR feedback. Other possible scenario is that of N_T BS's in a cellular system; in this case the assumption would be that all BS know the equivalent channel magnitude. ϕ_k are the weights that allow prioritizing the users. The problem to solve deals with scheduling of users and powers, the spatial precoder is fixed and part of the initial conditions of the problem.

Rate optimization is a reasonable choice for utility, reflecting the various coding rates implemented in the system. We assume an idealized link adaptation protocol. Proportional rate constraints allow a more definitive priorization among the users, which is quite useful for service class differentiation. Theoretically, this formulation also traces out the boundary of the capacity region similar to the weighted sum-rate maximization. The main difference is that it actually identifies the points on the capacity region boundary that satisfy the rate proportionally constraints. Furthermore, the max-min rate formulation is a special case of this formulation, i.e., when $\phi_1=\ldots=\phi_K$. Finally, by enforcing the average power constraint we allow instantaneous power levels to exceed the average power when necessary. Sum power constraint is needed in scenarios such as BC channel, but it is not usually imposed in multi-cell scenarios.

Note that ergodic optimization is considered because of twofold: i) it reduces the complexity of the resulting algorithm and ii) it incorporates the time dimension in the resulting resource allocation. In other words, in the case of instantaneous rate allocation only, the OFDMA algorithms are re-run every symbol (or several symbols). In this paper, we can capture the idea of "time slot allocation" by using the ergodicity assumption, and determine power allocation functions that are parameterized by the channel knowledge.

Note also that if there is no frequency structured components or noise, the maximal sum rate signaling does not require introducing correlation between subcarriers (cooperative subcarrier transmission or joint frequency-space processing). Therefore, the problem is separable across the subcarriers, and is tied together only by the power constraint. In these problems, it is useful lo approach the problem using duality principles. In addition, the utility function is non convex and by solving the dual problem and formulating it as a canonical distributed algorithm [10], the algorithm is simplified and also convergence to the globally optimal rate allocation can be achieved.

Finally and as notational convention vectors are set in boldface.

3. DUAL OPTIMIZATION

The proposed algorithm is based on a dual optimization framework. In other words, it is based on a Lagrangian relaxation of the power constraints and (possibly) rate constraints. This relaxation retains the subcarrier assignment exclusivity constraints, but "dualizes" the power/rate constraints and incorporates them into the objective function, thereby allowing us to solve the dual problem instead. This dual optimization is much less complex as we explain next.

To derive the dual problem we first write the Lagrangian. In order to simplify notation we define

$$\begin{aligned}
r_k &\triangleq \sum_{m=1}^{M}\mathrm{E}_{\gamma}\left\{\log_2\left(1+\gamma_{k,m,k_q}\right)\right\}\\
\hat{p}_k &\triangleq \sum_{m=1}^{M}\mathrm{E}_{\gamma}\left\{a_{k,m}p_{k,m}\right\}
\end{aligned} \tag{4}$$

where we do not explicitly write the dependence of r_k and $\hat{p}_k$ on the optimization variables **a**, **p**. Based on this definitions, the Lagrangian is

$$L = R\left(1-\boldsymbol{\mu}^T\phi\right) + \lambda\overline{P} - \lambda\mathrm{E}_{\gamma}\left\{\sum_{k=1}^{K}\hat{p}_k\right\} + \sum_{k=1}^{K}\mu_k r_k \tag{5}$$

λ, **μ** are the dual variables (also called prices) that relax the cost function, $R=\sum_{k=1}^{K} r_k$ and $\overline{P}$ is the power constraint.

Focusing on the first term in the maximization, we observe that if $\left(1-\boldsymbol{\mu}^T\phi\right)>0$ then the optimal solution would be $R^*=\infty$, since R is a free variable. This is clearly an infeasible solution for ergodic sum rate. Furhtermore, if $\left(1-\boldsymbol{\mu}^T\phi\right)<0$ then the optimal solution would be $R^{opt}=0$. Thus, we would like to constrain the multiplier to satisfy $\boldsymbol{\mu}^T\phi=1$. Thus, (5) can be simplified to

$$L = \lambda\overline{P} - \lambda\mathrm{E}_{\gamma}\left\{\sum_{k=1}^{K}\hat{p}_k\right\} + \sum_{k=1}^{K}\mu_k r_k \tag{6}$$

Note that the weights μ_k are the dual multipliers that enforce the proposed rate constraints. Additivity of the utility and linearity of the constraints lead to the following Lagrangian dual decomposition into individual user terms

$$L = \sum_k L_k + \lambda \bar{P} \tag{7}$$

where, for each user k,

$$L_k = \sum_k \mu_k r_k - \lambda_k \sum_k \hat{p}_k \qquad \lambda_k = \lambda \tag{8}$$

only depends on local rate r_k and the prices λ, **μ**.
The dual function g(λ, **μ**) is defined as:

$$\begin{aligned} g(\lambda,\boldsymbol{\mu}) &= \max_{\mathbf{a},\mathbf{p}} \left(\sum_k L_k(\mathbf{a},\mathbf{p},\lambda_k,\boldsymbol{\mu}) + \lambda \bar{P} \right) = \\ &= \sum_k L_k^*(\mathbf{a}^*,\mathbf{p}^*,\lambda_k,\boldsymbol{\mu}) + \lambda \bar{P} \end{aligned} \tag{9}$$

Evidently, this dual problem involves only K+1 variables and it is, therefore much easier to solve than the primal problem. Moreover, the maximization in (9) can be conducted parallel by each user, as long as the aggregate link price λ is feedback to source user *k*. Note that if there were no global constraint, as it is the case in multicell systems where $p_{k,m}$ stands for the power that Base Sation *k* has to allocate in frequency *m* and there is only per BS power constraints, the problem is further simplified.
The dual problem is defined as:

$$\begin{aligned} &\min g(\lambda,\boldsymbol{\mu}) \\ &s.t. \quad \lambda \geq 0, \boldsymbol{\mu} \in D \quad D = \left\{ \boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\mu}^T \phi = 1 \right\}. \end{aligned} \tag{10}$$

Since g(λ, **μ**) is the pointwise supremum of a family of affine functions in λ, **μ**, it is convex and (10) is a convex minimization problem (even if the primal is not a concave maximization problem).
Since g(λ) may be non differentiable, an iterative subgradient method can be used to update the dual variable λ to solve the dual problem. The computation of the subgradient requires knowing the individual weighted ergodic rates per user. Note that the “weights” in this case are no longer predetermined constants, but are effectively the multipliers that enforce the proportional rate constraints. From an initial guess λ^o and μ^o, the subgradient method generates a sequence of dual feasible parts according to the iteration

$$\lambda^{i+1} = \left[\lambda^i - s^i g_\lambda^i \right]^+ \qquad \boldsymbol{\mu}^{i+1} = \Pi_D \left[\boldsymbol{\mu}^i - s^i \mathbf{g}_{\boldsymbol{\mu}}^i \right] \tag{11}$$

Where g_λ^i denotes the subgradient of $g\left(\lambda^*\left(\boldsymbol{\mu}^i\right),\boldsymbol{\mu}^i\right)$ with respect to λ

$$g_\lambda^i = \bar{P} - E_\gamma \left\{ \sum_k \hat{p}_k^* \right\} \tag{12}$$

and s^i is a positive scalar step-size. $\mathbf{g}_{\boldsymbol{\mu}}^i$ denotes the subgradient of $g\left(\lambda^*\left(\boldsymbol{\mu}^i\right),\boldsymbol{\mu}^i\right)$ with respect to **μ**

$$\begin{aligned} &g_\mu^i = \bar{\mathbf{R}}^i - \phi \bar{R}^i \\ &with \quad \left[\bar{\mathbf{R}}^i\right]_k = \sum_m E_\gamma \left\{ \log_2\left(1+\gamma_{k,m,k_q}\left(\mathbf{a}^*,\mathbf{p}^*,\lambda^i,\mu_k^i\right)\right) \right\} \\ &\qquad \bar{R}^i = \sum_k \left[\bar{\mathbf{R}}^i\right]_k \end{aligned} \tag{13}$$

Finally, $\Pi_D[\cdot]$ denotes projection onto the set D.

Concerning convergence, for a primal problem that is a convex optimization, the convergence is towards a global optimum. Otherwise, global maximum of non concave functions is an intrinsically difficult problem on non convex optimization. In [10] the authors show that the sequence of the maximization of (8) and the computation of (11) forms a canonical distributed algorithm that solves (2) and the dual problem (10). Even for non concave utilities the canonical distributed algorithm may still converge to a globally optimal solution if L_k^* is continuous at optimum λ^*. Based on this property, an analytical proof of convergence for the algorithm that is proposed next is left for further work. Simulation results have proved good convergence for it.

4. ALGORITHM

In the rest of the paper we deal with the specific case of Broadcast (BC) channel, where the global power constraint is needed. The maximization of (9) could have been formulated only with respect to the powers, **p**. In this way, whenever any of the optimal components $p_{k,m}^* = 0$, this would mean that user *k* should not be scheduled in frequency *k*. In addition to the complexity of this multiuser frequency power allocation problem, note that $p_{k,m}$ depends on the spatial channel at frequency *m*, which, in the case of the BC channel, depends on the spatial precoder that is associated with user *k*. The purpose of this work is to design a low complexity scheduler; this fact motivates the simplification of the complex space-frequency multiuser scheduler that has been described by using the multibeam opportunistic scheme [11]. In this case, the Base Station uses a set of orthonormal beams that are associated with the users depending on their reported SINR. This scheduler is designed such that it works without interacting with power allocation and dual optimization. The problem formulation of (9) accounts for this explicit user scheduling by incorporating the discrete variables ***a***. As shown in Fig. 1, the first step in the proposed algorithm is the spatial scheduler, which obtains ***a*** in a low complexity way, as it is described in Section A.

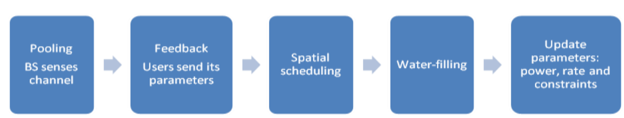


Fig. 1. Block diagram of the proposed algorithm.

Once user *k* has been selected by being associated with beam *q* at each frequency *m*, the next step in the algorithm (see Fig. 1) is the power allocation, which should be derived from the equation

$$\frac{\partial L_k}{\partial p_{k,m}} = 0 \qquad \forall k \tag{14}$$

Assuming the interference with constant power $\bar{P}$, a suboptimal solution is provided by the waterfilling in (15)

$$p_{k,m}^* = \left[\frac{\mu_k}{\lambda \ln(2)} - \frac{1}{\gamma_{k,m,k_q}^1} \right]^+ \tag{15}$$

where

see the program

$$\gamma^1_{k,m,k_q} = \frac{c_{k,m,k_q}}{\sigma^2 + \sum_{s\neq q}^{N_T} a_{s,m} P c_{k,m,k_s}} \tag{16}$$

The dual variables are obtained by solving the dual problem (10) with low complexity in a stochastic way. The detailed waterfilling procedure is described in Section C.

A. Spatial scheduler

The spatial scheduler is obtained for each frequency m by introducing a simplified SINR that considers uniform power allocation (i.e. no interaction neither with the primal variables $\mathbf{p}$ nor with the dual variables λ, $\boldsymbol{\mu}$). The proposed spatial scheduler is based on the multibeam opportunistic strategy, which considers an orthonormal random beamforming set as precoder and assigns users to beams based only on SINR feedback. In order to counteract the losses that opportunistic schemes present when the number of users is moderate or low we extend the beam and user set optimization proposed in [12] to OFDMA. Let $Q_m \in \{1,\ldots,N_T\}$ denotes the number of users served or active beams at mth frequency bin, $U_m^{(Q_m)} \subset \{1,\ldots,K\}$ and $S_m^{(Q_m)} \subseteq \{1,\ldots,N_T\}$ are the user and beam set, respectively, with Q_m elements without repetition. For the BC the SINR in (3) is equivalent to

$$\gamma^{(Q_m)}_{k,m,q} = \frac{p_{k,m}\left|\mathbf{h}^T_{k,m}\mathbf{b}_{m,q}\right|^2}{\sigma_n^2 + \sum_{\substack{q'\neq q\\ q'\in S_m^{(Q_m)}}} p_{q',m}\left|\mathbf{h}^T_{k,m}\mathbf{b}_{m,q'}\right|^2} \quad m\in S_m^{(Q_m)} \quad k\in U_m^{(Q_m)} \tag{17}$$

where $\mathbf{b}_{m,q}$ is qth beam at mth frequency and $\sigma_n^2=\sigma_w^2/P$ is the noise variance. In (17) there are $\binom{K}{Q^m}Q^m!$ permutations of $\gamma^{(Q_m)}_{k,m}$ for $S^{(Q_m)}$. The optimal beam subset S_m^* and user subset U_m^* at mth frequency bin are obtained by (9) using exhaustive search. However we apply a suboptimal approach to reduce the complexity. Next, we assume that equal power P is allocated among beams and subcarriers, $j=1,\ldots,\binom{N_T}{Q_m}$ is the index of all beam combinations.

Fixing jth combination, the optimal user satisfies

$$k^*_{m,j,q} = \arg\max_k \gamma^{(Q_m)}_{k,m,j,q} \tag{18}$$

where subindex k_q in (3) has become indexes k,j,q in (18) due to the beam and user search. The optimal value of (18) is added to $U^{(Q_m)}_{j,m}$. For each frequency m, optimal jth index is found by expression

$$j^* = \arg\max_{1\leq j\leq \binom{N_T}{Q_m}} \sum_{q=1}^{Q_m} \log_2\left(1+\gamma^{(Q_m)}_{k^*_{m,j,q},m,j,q}\right). \tag{19}$$

Hence, the optimal user and beam set that serves Q_m users simultaneously are $U_m^{*(Q_m)} = U^{(Q_m)}_{j^*_m,m}$ and $S_m^{*(Q_m)} = S^{(Q_m)}_{j^*_m,m}$. Finally, optimal Q_m is given by

$$Q_m^* = \arg\max_{1\leq Q_m\leq NT} \sum_{k\in U_m^{*(Q_m)}} \sum_{q\in S_m^{*(Q_m)}} \log_2\left(1+\gamma^{(Q_m)}_{k,m,j^*,q}\right). \tag{20}$$

Once this spatial scheduling is finished for each frequency m, we simplify indexes $k^{(Q_m^*)}_{m,j,q},m,j^*,q \triangleq k,m,k_q$ as they are used in (3), ready to be applied in the frequency waterfilling and dual optimization step. The corresponding user selection variable $a_{k,m}$ is set to 1 if the user has been chosen.

Regarding fairness, this solution has the drawback of scheduling users on the available beams and frequencies by only looking at channel gains. In this way, users with good channel conditions, i.e. users located near the base station with a small path loss, tend to monopolize channel resources. The lack of resources for weak users may constrain the behavior of dual optimization. An improved spatial scheduler can be designed by considering the maximization of

$$\max_{k_q} \sum_k \sum_q \mu_k \log_2\left(1+\gamma_{k,m,k_q}\right) \quad \forall m \tag{21}$$

by using the equal power approximation of $p_{k,m}$, i.e. $p_{k,m}=P$. This can be simply implemented by inserting μ_k in (19) and (20). This scheduler interacts with the dual optimization algorithm. It releases beams and frequencies to users according to rate constraints, but preserves both the light requirements on feedback parameters and the distributed implementation.

Finally, instead of MOB other spatial precoders (such as Zero Forzing) can be used, however, at the expense of complexity increase in the feedback.

B. Feedback complexity

The number of feedback parameters to be transmitted to the BS by each kth user are N_TxM and they correspond to all possible values of $\left|\mathbf{h}^T_{k,m}\mathbf{b}_{m,q}\right|^2$ for $q=1,\ldots,N_T$ and $m=1,\ldots,M$. This amount can be reduced by fixing Q_m previously. Users must know the value Q_m. In this case, feedback is reduced to 3xM and they correspond to q^*, j^*, best beam and best permutation, and γ_{k,m,j^*,q^*}. That is because fixing Q_m causes the number of permutations is known a priori by all users and it is fixed. Thus, each user can compute γ_{k,m,j^*,q^*} and send it to BS, jointly with indexes. In addition, fixing $Q_m = Q = N_T$ the amount of feedback is also reduced to 2xM, q^* and $\gamma_{k,m,1,q^*}$ since $j=1$.

Finally, depending on the delay spread of the channel, the number of parameters to feedback can be further reduced by frequency grouping or chunk processing [5].

C. Frequency power allocation

Once the user selection problem has been solved, $\mathbf{a}$ is known in (2), and the next step is to obtain the solution for the power allocation $\mathbf{p}$ in (12).

To compute (12), ergodic maximization through stochastic approximation is introduced [14]. In other words, maximization occurs through time or iterations. At nth iteration, power assigned to kth user at mth carrier becomes

$$p^*_{k,m}[n] = \left[\frac{\mu_k[n]}{\lambda[n]\ln(2)} - \frac{1}{\gamma^1_{k,m,k_q}[n]}\right]^+ \tag{22}$$

see the program

where $[x]^+ = \max(x,0)$.

Per-user rate and total power are given by

$$\begin{aligned} R_k[n] &= \sum_{m\in M} R_{k,m}\left(\gamma_{k,m,k_q}(p^*_{k,m}[n])\right) \\ P[n] &= \sum_{m\in M}\sum_{k\in K} p^*_{k,m}[n] \end{aligned} \qquad (23)$$

Finally, λ, **μ** are updated using subgradient method as in (11)-(13) but with a stochastic approximation. In fact, these parameters are given by expressions

$$\begin{aligned} \lambda[n+1] &= \left[\lambda[n] - \delta\left(\bar{P} - P[n]\right)\right]^+ \\ \boldsymbol{\mu}[n+1] &= \Pi_{\mathcal{D}}\left(\boldsymbol{\mu}[n] - \delta\left(\mathbf{R}[n] - \phi R[n]\right)\right) \end{aligned} \qquad (24)$$

where $\Pi_{\mathcal{D}}$ is the projection onto set $\mathcal{D} = \left\{\boldsymbol{\mu} \geq \mathbf{0} \,|\, \boldsymbol{\mu}^T\phi = 1\right\}$, $\mathbf{R}[n] = \left[R_1[n] \ \cdots \ R_K[n]\right]^T$ and $R[n] = \sum_{k\in K} R_k[n]$. To obtain a good performance, a suitable values could be $\lambda[0] = 1$, $\boldsymbol{\mu}[0] = \frac{\phi}{\phi^T\phi}$ and $\delta = 0.01$.

Subgradient search methods have been used to obtain the solution for the dual variables. As it is important to be able to perform resource allocation in real-time, we obtain an on-line adaptive algorithm by performing the iterations of the subgradient across time.

D. Complexity

Using ergodic sumrate relaxes complexity since constraints are not instantaneous but ergodic. Note that though objective is ergodic, feedback parameters contain instantaneous information. This algorithm has several stages. First of all there is the pooling stage. During this step, BS sends a pilot signal through all beams and only one beam is active. The complexity in the beamforming is $O(MN_T)$.

Next step is computing all γ parameters. Its complexity depends on fixing Q_m. Leaving it as a free parameter, complexity is $O(MK2^{NT}N_T^2)$. Otherwise, fixing it, complexity becomes $O(MK2^{NT}N_T)$. Moreover, adjusting $Q_m = Q = N_T$ it is reduced to $O(MKN_T)$.

Finally, there is the power allocation stage, water-filling has complexity O(MK) and it is followed by O(K) updates for the rates, power and multipliers.

In general, complexity could be very low, as $O(MKN_T)$, or higher, as $O(MK2^{NT}N_T^2)$, depending on how optimum is desirable.

Next table shows algorithm step-by-step and its complexity.

Note that in step 2 there are three possibilities: 2.a has N_TxM parameters of feedback; 2.b, 3xM and 2.c, 2xM.

Other solutions such as [13] find the optimal bound of capacity rate at cost of complexity, $O(K^2M\log N)$. Others have less complexity, as O(M) in [8], but they lose in performance. See fig. 4 for more details in the comparison.

5. RESULTS

We organize numerical results in two parts: the first part illustrated in section 1 refers to a simple cellular scenario and has the aim of showing the main behavior of the algorithm; the second part refers to a more realistic scenario.

1. Pooling: BS transmits pilot signal to sense each equivalent channel.	$O(MN_T)$
2.a Non-fixed Q_m Feedback: each user sends $\left\|\mathbf{h}^T_{k,m}\mathbf{u}_{m,q}\right\|^2$ BS schedules users spatially	 $O(MN_T)$ $O(MK2^{NT}N_T^2)$
2.b Fixed Q_m Feedback: each user sends $q^*, j^*, \gamma_{k,m,j^*,q^*}$ BS schedules users spatially	 O(M) $O(MK2^{NT}N_T)$
2.c Q_m=N_T Feedback: each user sends q^*, $\gamma_{k,m,1,q^*}$ BS schedules users spatially	 O(M) $O(MKN_T)$
3. Water-filling	O(MK)
4. Updating parameters λ, **μ**	O(K)

Table 1. Algorithm step-by-step and its complexity.

A. Results for simple scenario

All simulations in this scenario consider M=64 subcarriers, power constraint $\bar{P} = 10dB$ and power parameter P=1 in equation (16). The channel model includes normalized Rayleigh fading and does not take care of path-loss or shadowing components. All carriers have frequency spacing of 1Hz. All users are located at same distance from BS. A linear array of NT antenna isconsidered at the base station and distance between sensors is 0.5λ.

Fig. 2 shows how every user converges to its weight with few iterations. Note that user rates are normalized with sumrate. Fig. 3 shows power convergence to average power constraint. Hence the good convergence properties of dual optimization algorithm are confirmed by the results.

In order to compare the spatial scheduling algorithm with other algorithms of the literature, we show in Fig. 4 the rate region for two users, compared with those obtained with the algorithms in [13] (DPC) and [8] (KOU), and to uniform power allocation strategy (UPA). Note that [13] has higher complexity and fixes a theoretical maximum sum-rate that can be achieved, whereas [8] has lower complexity, but worse performance.

Fig. 5 shows the computation complexity in terms of simulation time for different system configurations and strategies of choosing Q_m. Note how complexity increases with the usage of dynamic Q_m. In order to compare the complexity with that of [13], based on DPC, we provide the following table that shows the computation complexity for K=8 and K=32 users, and N_T=2. Time is expressed in seconds. This gives a clear idea of the complexity of [13], $O(K^2M\log N)$, and how it increases with number of users.

	K=8	K=32
[13]	283	3349
Our algorithm (Dyn Q_m / Q_m=N_T)	28 / 31	50 / 62

Table 2. Complexity comparison with [13].

see the program

When compared to solution [8], for N_T=2 and K=2, our algorithm requires roughly the same computation time, but achieves best performance, as shown in Fig. 4.

Finally, Fig. 6 shows the impact on sum-rate of using a dynamic Q_m or a fixed Q_m. Dynamic choice of Q_m is useful for a large number of antennas, whereas the other choice is good for few users.

B. Results in a realistic scenario

In this section we present and discuss simulation results obtained for a scenario which incorporate some characteristic aspects of practical application in next generation wireless systems, In fact, 3GPP-LTE and WiMAX employ OFDMA as their main multiple access mechanism (although other options are also defined in the standards).

We are considering here a single cell of the downlink of an OFDM wireless system with M=128 subcarriers working on a bandwidth of 1.25 Mhz. Base station is equipped with multiple antennas. The system is TDD and it is assumed that 2/5 of frame interval is used for downlink transmission. The CSI coming from users is updated every 10ms. Two options for user distribution are considered: in the first option the users have a position which is uniformly distributed in circular area of radius 500m; in the second option the users are placed at the same distance of 250 m from the base station.

Channel model includes path loss, correlated shadowing (not present in the second option for user distributions) and time and frequency correlated fast fading. Path loss is modeled as a function of distance as L(db)= k0 + k1 log(d) (k1=40, k1=15.2 for results). Shadowing is superimposed to path-loss, with classical lognormal model (sigma= 6 dB) and exponential correlation in space (correlation distance equal to 20m). Fast fading on each link of the MIMO broadcast channel is complex Gaussian, independent across antennas and is modeled according to a 3GPP Pedestrian model [11]. This model has a finite number of complex multipath components with fixed delay (delay spread around 2-3 microseconds) and power (average normalized to 1). Time correlation is obtained according to a Jakes' model [12] with given Doppler bandwidth (6 Hz in the results). At the base station orthogonal beamforming is adopted, where beam vectors change randomly at each frame. In the simulated system the total average power constraint is fixed to 1W.

Note that, in realistic conditions, channel variations in time due to Doppler effects have a non negligible impact on the feedback quality. In fact, at the scheduling time n the algorithm uses feedback parameters measured at time n-1, which can be changed in the meanwhile. Therefore, due to outdated feedback the transmission at the scheduled rate may fail sometimes. This aspect is left for future investigation

In the first two figures, fig. 7 and fig. 8, the dynamic behavior of sum-rate and total allocated power is illustrated for a system with 10 users in fixed position at distance 250 m. In the scheduling algorithm equal weights ϕ_k = 1/K are used. We note that although total power and sum rate change frame by frame due to fast channel variations, the algorithm for dual variable optimization correctly tracks the constraint on the average power. We also observed from a wide set of results that the range of variations enlarges when the users have

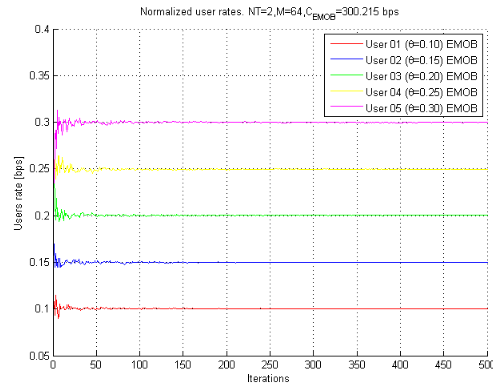


Fig. 2. Different rates for 5 users using this algorithm

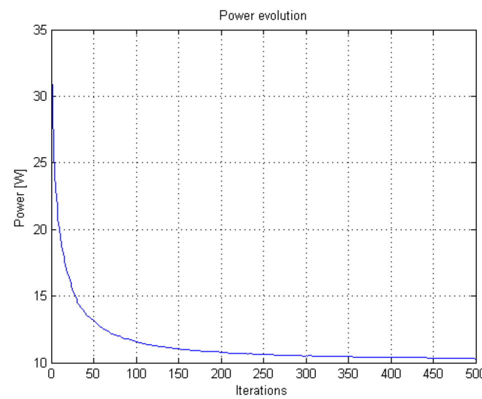


Fig. 3. Power evolution vs. iterations.

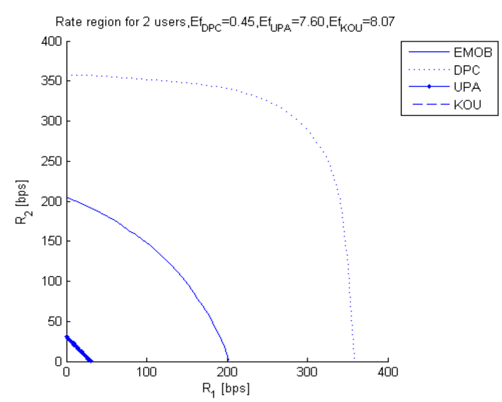


Fig. 4. Rate region for two users. DPC from [8], and KOU from [10] are plotted jointly with this algorithm EMOB. UPA is also plotted. Note that UPA and KOU are practically coincident.

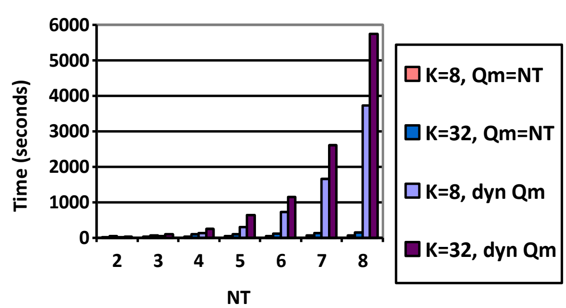


Fig. 5. Simulation time in seconds for K=8 and K=32, dynamic Q_m or fixed Q_m.

see the program

different path-loss, but again the algorithm tracks correctly the average.

In spite of rate and power variations, we also checked the robustness of algorithm to ensure fair average rate allocation among users. This is shown in the following two histograms in fig. 9 and fig. 10, which illustrates the distribution of average user rates in a system with 10 users, 3 antennas and 3 different user classes (class 1: weight 0.5/K - class 2: weight 1/K - class 3: weight 1.5/K). Fig. 9 refers to equal distance users at 250 m, whereas fig. 10 refers to uniformly distributed users in a circular area of radius 500 m. We can note that the algorithms are quite fair to assign rates to different users, even when they have with different weights and different path-loss conditions.

Finally, fig. 11 shows sum rate and user rate vs. number of user, in a system with 3 antennas, users at equal distance 250 m from BS, and 3 different user classes (class 1: weight 0.5/K - class 2: weight 1/K - class 3: weight 1.5/K). We note that sum rate increases with the number of users, meaning that the scheduling algorithm capture the available multiuser diversity while preserving average rate fairness. Per-user rate decreases since the sum-rate needs to be shared among an increasing number of users.

6. CONCLUSIONS

This paper has presented a low complexity space-frequency scheduler that allocates power among users. Ergodic objective and ergodic constraints are purposed to relax complexity. Moreover many strategies had been presented and low complexity has been explained. In addition, weights are purposed in order to set rate priorities or several QoS. Finally, some benchmarks are presented to compare the performance. Aspects such as robustness to imperfect CSIT, discrete rate allocation, modification of the algorithm to incorporate jointly encoded sub-channels (e.g. space-time codes) and cross-layer design for user scheduling improvement are possible topics for further research. The space-frequency multiuser scheduler has been presented in a general formulation such that the proposed distributed strategy (as a result of the dual optimization and opportunistic user selection) together with the low complexity of the proposed ergodic scheduler can be applied to different space-frequency scheduling scenarios.

7. REFERENCES

[1] J.Brehmer, W. Utschick, “Nonconcave Utility Maximisation in the MIMO Broadcast Channel,” Eurasip JASP, vol. 2009.

[2] M. Haardt, V. Stankovic, G. Del Galdo, “Efficient Multi-user MIMO Downlink Precoding and Scheduling,” CAMSAP 05, Dec. 2005.

[3] Sriram Vishwanath, Nihar Jindal, and Andrea Goldsmith, “On the Capacity of Multiple Input Multiple Output Broadcast Channels,” IEEE Trans. Inf. Theory, vol. 51, pp. 1570–1580, Apr. 2005.

[4] W. Yu, “Uplink–Downlink Duality Via Minimax Duality,” IEEE Trans. on IT, February 2006.

[5] E. Jorswieck, A. Sezgin, B. Ottersten, A. Paulraj, “Feedback reduction in uplink MIMO OFDM Systems by Chunk Optimization,” Eurasip JASP, Vol. 2008.

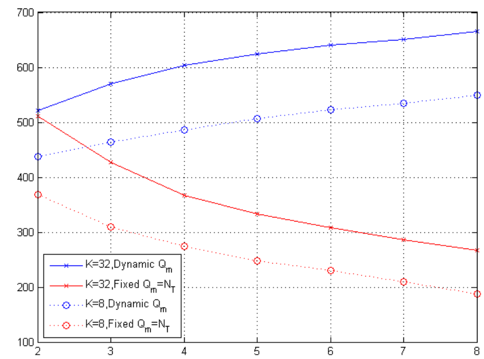


Fig. 6. Contribution to sum-rate of dynamic Q_m or fixed Q_m. plotted for K=8 and K=32 users with different number of antennas.

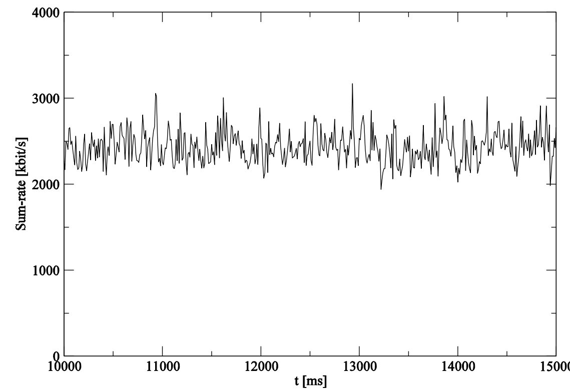


Fig. 7. Dynamic behavior of sum-rate in a cell with 10 users in fixed positions at distance 250m from BS.

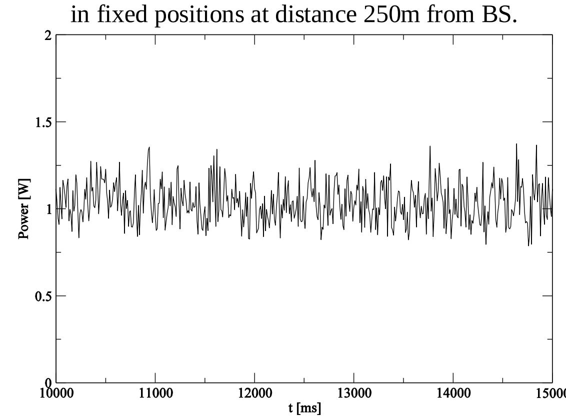


Fig. 8. Dynamic behavior of total allocated power in a cell with 10 users in fixed positions at distance 250m from BS.

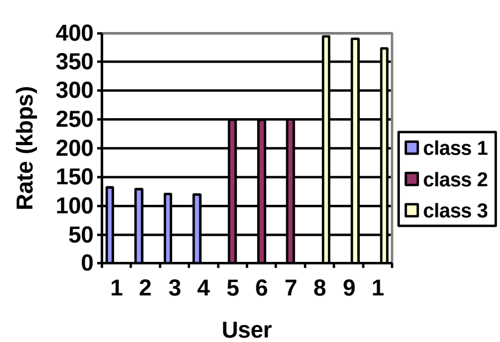


Fig. 9. Per-user average rate distribution in a system with 10 users at equal distance 250 m from BS, belonging to 3 different user classes.

see the program

[6] N. Wei, A. Pokhariyal, T. B. Sorensen, T. E. Kolding, P.E. Mogensen, IEEE JSAC, vol.26, no. 6, Aug. 2008.
[7] M. Codrenau, M. Juntti, M. Latva-Aho, “Low-complexity iterative algorithm for finding the MIMO-OFDM broadcast channel sum capacity”, IEEE Trans. on Comm., vol. 55, no. 1, jan 2007.
[8] Issam Toufik, Marios Kountouris, “Power allocation and feedback reduction for MIMO-OFDMA opportunistic Beamforming,” VTC Spring 2006: 2568-2572
[9] G. Liu; J. Zhang; F. Jiang; W. Wang, “Joint Spatial and Frequency Proportional Fairness Scheduling for MIMO OFDMA Downlink,” Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on Volume , Issue , 21-25 Sept. 2007 Page(s):491 - 494
[10] M. Chiang, S. Zhang, P. Hande, “Distributed rate allocation for inelastic flows: optimization frameworks, optimality conditions, and optimal algorithms,” IEEE JSAC, vol. 23, no.1, Jan. 2005.
[11] M.Sharif, B.Hassibi, “On the capacity of Broadcast Channels with Partial Side Information,” IEEE Trans. on IT, Feb. 2005.
[12] T. Kang and H. Kim, “Optimal Beam Subset and User Selection for Orthogonal Random Beamforming”, IEEE Communications letters, vol. 12, no. 9, Sep. 2008.
[13] M. Kobayashi, G. Caire, “An Iterative Water-Filling Algorithm for Maximum Weighted Sum-Rate of Gaussian MIMO-BC,” IEEE JSAC, vol. 24, Aug. 2006.
[14] I. Wong, B. Evans, “Optimal OFDMA resource allocation with linear complexity to maximize ergodic weighted sum capacity,” ICASSP 07, Hawaii.

8. ACKNOWLEDGMENTS

This work was supported by the European Commission under project NEWCOM++ (216715), Optimix (Grant Agreement 214625) and by Spanish Government TEC2008-06327-C03-01. The work has been done during the 6 months stay of A. Perez-Neira at ACCESS/Signal Processing Lab, KTH (Stockholm).

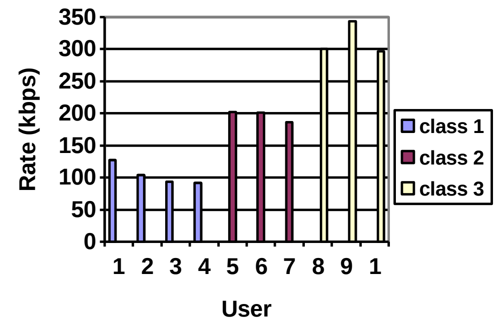


Fig. 10. Per-user average rate distribution in a system with 10 users uniformly distributed in a cell of radius 500 m, belonging to 3 different user classes.

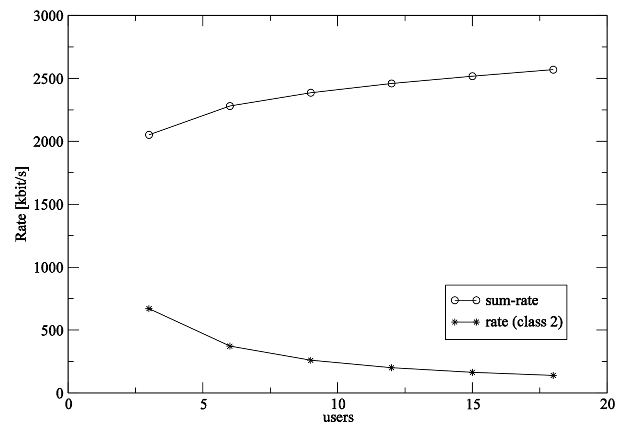


Fig. 11. Sum rate and per-user rate (user in class 2) vs. number of users, for a system with users at equal distance from BS.

see the program

RESOURCE ALLOCATION IN MULTIANTENNA SYSTEMS UNDER GENERAL POWER CONSTRAINTS

Angela Feistel[*], *Sławomir Stańczak*[*†] *and Michał Kaliszan*[†]

[*] Heinrich-Hertz Chair, EECS,
University of Technology Berlin,
Einsteinufer 25, 10587 Berlin, Germany
angela.feistel@mk.tu-berlin.de

[†] Fraunhofer German-Sino Lab
for Mobile Communications,
Einsteinufer 37, 10587 Berlin, Germany
{stanczak, kaliszan}@hhi.fhg.de

ABSTRACT

We address the problem of jointly optimizing powers and receive beamformers in a wireless network subject to general power constraints. More precisely, we solve the max-min SIR-balancing problem by maximizing a weighted sum of utilities of the SIRs over the joint space of powers and receive beamformers. Thus, the paper extends the work of [1] to noisy channels under general power constraints and to a larger class of utility functions. The proposed algorithm is shown to converge to a max-min SIR-balancing solution. It is not amenable to distributed implementation but the results presented in this paper provide a basis for the development of decentralized algorithmic solutions.

1. INTRODUCTION

In wireless networks, power control and beamforming are two central mechanisms for resource allocation and interference management. So far, most work in this area has focused on the max-min SIR-balancing problem (see, for instance, [2], [3], [4, Sections 3.1, 5.6], [1], [5, Section 5.5]) and references therein). Recent results [6] have obtained a simple characterization of the max-min SIR-balanced power vector under general power constraints. The authors formulate an eigenvalue problem of the same dimension as the original problem and give a procedure to compute the max-min SIR power vector.

Another known approach aims at maximizing a weighted sum of utilities of the SIRs (or balanced SIRs) [7],[5]. Here one can trade efficiency for fairness by increasing the so-called relative concavity of the utility function [5]. This tradeoff also depends on the choice of the weight vector. A utility-based joint power control and receive beamforming algorithm can be implemented distributedly, yet converges only to a local optimum [8].

In this paper, we consider the possibility of solving the max-min SIR-balancing problem by maximizing a certain aggregate utility function over the joint space of power vectors and receive beamformers. We apply the connection between the max-min SIR-balancing and the utility-based power control problem established in [6]. This connection is briefly explained in section 3, where we show how to choose the weight vector so that the sum utility optimization achieves the max-min SIR-balancing solution. In addition we discuss the properties of this connection in order to derive an algorithmic solution. Then, in section 4, we propose an iterative algorithm that solves the optimization problem. We prove important properties as monotonicity and convergence to the max-min SIR-balancing solution. Thus, this paper generalizes the results of [1] to noisy channels under general power constraints and to a larger class of utility functions. Finally we discuss implementation aspects, propose a heuristic algorithm that additionally updates transmit beamformers and give some simulation results to illustrate the convergence behaviour and the gain achievable by resource allocation.

2. SYSTEM MODEL

We consider an arbitrary wireless network with K active links that are referred to as users. Let $\mathcal{K} = \{1, \dots, K\}$ denote the set of all users. Each user is equipped with $M \geq 2$ antenna elements at the transmitter side and the receiver side. Let $\mathbf{u}_l^{(k)} = (u_{1,l}^{(k)}, \dots, u_{M,l}^{(k)}) \in \mathbb{C}^M$ with $\mathbf{u}_k = \mathbf{u}_k^{(k)}$ be the effective transmit beamformer of user l that is associated with the receiver of user k. The effective transmitter is determined by the transceiver structure, the channel state, etc. In case of transmit beamforming the effective transmitter $\mathbf{u}_l^{(k)}$ is the product of the channel matrix between the l-th transmitter and the k-th receiver and the transmit beamformer of the l-th user. In this paper we assume that the effective transmitters are *arbitrary but fixed*. In contrast, the receive beamformers, which act as linear receivers, should be jointly optimized with transmit powers of the users. The transmit powers p_k and the receive beamformers $\mathbf{v}_k \in \mathbb{C}^M$ of all users are collected in the power vector $\mathbf{p} = (p_1, \dots, p_K) \in \mathbb{R}_+^K$ and in the receive beamforming matrix $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_K) \in \mathbb{C}^{M\times K}$, re-

see the program

spectively.[1] The transmit powers are subject to power constraints

$$\mathrm{P} = \{\mathbf{p} \in \mathbb{R}_+^K : \mathbf{C}\mathbf{p} \le \mathbf{p}', \mathbf{C} \in \{0,1\}^{N\times K}\} \subset \mathbb{R}^K \quad (1)$$

for some given $\mathbf{p}' = (P_1, \dots, P_N) > 0$ and $\mathbf{C}$ with at least one 1 in each column so that P is a compact set. Throughout the paper we use $\mathcal{N} = \{1, \dots, N\}$ where N is the number of power constraints. Thus, in the special cases of individual power constraints and sum power constraints we have $N = K$, $\mathbf{C} = \mathbf{I}$ and $N = 1$, $\mathbf{C} = \mathbf{1}^T$, respectively. Furthermore, since the SIR is independent of the norm of the receive beamformers, we can assume that $\|\mathbf{v}_k\|_2 = 1$ for each $1 \le k \le K$ and use $\mathrm{V} = \{(\mathbf{x}_1, \dots, \mathbf{x}_K) : \|\mathbf{x}_k\|_2 = 1\}$ to denote the set of all receive beamforming matrices.

The main figure of merit is the SIR given by

$$\begin{aligned}\mathrm{SIR}_k(\mathbf{p}, \mathbf{v}_k) :&= \frac{p_k}{I_k(\mathbf{p}, \mathbf{v}_k)} \\ &= \frac{p_k}{\sum_{l=1}^K g_{k,l}(\mathbf{v}_k)p_l + z_k(\mathbf{v}_k)}.\end{aligned} \quad (2)$$

Here and hereafter, for any $\mathbf{V} \in \mathrm{V}$, the nonnegative matrix $\mathbf{G} := \mathbf{G}(\mathbf{V}) = (g_{k,l}) \in \mathbb{R}_+^{K\times K}$ is called the *gain matrix* where $g_{k,l} := g_{k,l}(\mathbf{v}_k) = G_{k,l}/G_{k,k} \ge 0$ if $l \ne k$ and 0 if $l = k$. Given $\mathbf{V}$, $G_{k,k} := G_{k,k}(\mathbf{v}_k) > 0$ is the signal power gain of user k and $G_{k,l} := G_{k,l}(\mathbf{v}_k) \ge 0$, $l \ne k$ is the interference power gain between the transmitter of user l and the receiver of user k. The kth entry of $\mathbf{z} := \mathbf{z}(\mathbf{V}) = (z_1, \dots, z_K)$ is $z_k := z_k(\mathbf{v}_k) = \sigma_k^2/G_{k,k} > 0$, where $\sigma_k^2 > 0$ is the noise variance at the receiver output. Throughout the paper we assume that

(A.1) each entry of $\mathbf{G}(\mathbf{V})$ is a *continuous* function of $\mathbf{V} \in \mathrm{V}$. Moreover, for any fixed $\mathbf{V} \in \mathrm{V}$, $\mathbf{G} = \mathbf{G}(\mathbf{V}) \ge 0$ is an *irreducible* matrix.

Remark 1 *In practice, the irreducibility property can be guaranteed by approximating the gain matrices* $\mathbf{G}(\mathbf{V}), \mathbf{V} \in \mathrm{V}$*, with matrices of the form* $\mathbf{G}(\mathbf{V}) + \epsilon\mathbf{A}$ *for some sufficiently small* $\epsilon > 0$ *where* $\mathbf{A} = (\mathbf{1}\mathbf{1}^T - \mathbf{I}) \in \{0,1\}^{K\times K}$.

Let $\gamma_1, \dots, \gamma_K > 0$ be some given positive constants, and let $\mathbf{\Gamma} := \mathrm{diag}(\gamma_1, \dots, \gamma_K)$.

Definition 1 *Given any* $\mathbf{\Gamma}$*, we say that* $(\mathbf{p}^*, \mathbf{V}^*)$ *is a max-min SIR-balanced solution if*

$$(\mathbf{p}^*, \mathbf{V}^*) := \arg\max_{(\mathbf{p},\mathbf{V})\in \mathrm{P}\times\mathrm{V}} \min_{k\in\mathcal{K}} (\mathrm{SIR}_k(\mathbf{p}, \mathbf{v}_k)/\gamma_k). \quad (3)$$

The maximum exists as the SIRs are continuous on the compact set $\mathrm{P} \times \mathrm{V}$. Moreover, it may be easily verified that $\mathbf{p}^* > 0$, allowing us to focus on $\mathrm{P}_+ = \mathrm{P} \cap \mathbb{R}_{++}^K$.

[1] $\mathbb{R}_+$ and $\mathbb{R}_{++}$ are nonnegative reals and positive reals, respectively.

Notice that γ_k is not necessarily met under a max-min SIR-balanced solution (3). However, it is well known [5] that $\mathrm{SIR}_k(\mathbf{p}, \mathbf{v}_k) \ge \gamma_k$ for each k and some $(\mathbf{p}, \mathbf{V}) \in \mathrm{P}\times\mathrm{V}$ if and only if $\mathrm{SIR}_k(\mathbf{p}^*, \mathbf{v}_k^*) \ge \gamma_k$ for each k. In other words, $\mathbf{\Gamma}$ is feasible if and only if each γ_k is met under a max-min SIR-balanced solution. Thus, γ_k can be interpreted as the SIR target of user k.

3. MAX-MIN SIR BALANCING PROBLEM FOR FIXED BEAMFORMERS

Throughout this section we assume that in addition to the transceivers, the receive beamformers $\mathbf{V} \in \mathrm{V}$ are arbitrary but fixed. Note that by (A.1), $\mathbf{G} := \mathbf{G}(\mathbf{V})$ is a nonnegative *irreducible* matrix. In what follows, we briefly review the connection between the max-min SIR-balancing power control problem (3) and a utility-based power control problem [5]. More precisely, we show how to choose the weight vector in the utility maximization problem so that a solution to this problem is a max-min SIR-balancing solution. For the proofs, the reader is referred to [6, 5].

Let us first define the max-min SIR-balancing power control problem. For any given $\gamma_k > 0, k \in \mathcal{K}$, a max-min SIR-balanced power vector is defined to be

$$\begin{aligned}\bar{\mathbf{p}} := \bar{\mathbf{p}}(\mathbf{V}) = \arg\max_{\mathbf{p}\ge 0} &\min_{k\in\mathcal{K}}\big(\mathrm{SIR}_k(\mathbf{p}, \mathbf{v}_k)/\gamma_k\big) \\ \text{s.t.} \quad &\max_{n\in\mathcal{N}} g_n(\mathbf{p}) \le 1\,,\end{aligned} \quad (4)$$

where

$$g_n(\mathbf{p}) := (1/P_n)\mathbf{c}_n^T\mathbf{p}, \quad n \in \mathcal{N} \quad (5)$$

and $\mathbf{c}_n \in \{0,1\}^K$ is a (column) vector equal to the nth row of the matrix $\mathbf{C}$. Note that $\mathbf{p} \ge 0$ and $\max_{n\in\mathcal{N}} g_n(\mathbf{p}) \le 1$ holds if and only if the power constraints are satisfied: $\mathbf{p} \in \mathrm{P}$. Moreover, for any power vector $\bar{\mathbf{p}}$ that solves (4), we have $\max_{n\in\mathcal{N}} g_n(\bar{\mathbf{p}}) = 1$, which means that at least one power constraint is active.

The irreducibility assumption of $\mathbf{G}(\mathbf{V})$ implies that the max-min SIR-balanced power vector $\bar{\mathbf{p}}$ is unique and that

$$\forall_{k\in\mathcal{K}}\ \mathrm{SIR}_k(\bar{\mathbf{p}})/\gamma_k = \beta \quad (6)$$

for some $\beta > 0$. From these observations, it follows that [5]

$$\frac{1}{\beta}\bar{\mathbf{p}} = \mathbf{\Gamma}\mathbf{G}\bar{\mathbf{p}} + \mathbf{\Gamma}\mathbf{z} \qquad \max_{n\in\mathcal{N}} g_n(\bar{\mathbf{p}}) = 1\,. \quad (7)$$

Now let

$$\mathcal{N}(\mathbf{p}) := \big\{m \in \mathcal{N} : m = \arg\max_{n\in\mathcal{N}} g_n(\mathbf{p}) = 1\big\}$$

be the set of power constraints that are active under the power vector $\mathbf{p}$. Note that due to the fact that at least one

see the program

power constraint is active in the optimum, $\mathcal{N}(\bar{\mathbf{p}})$ is not empty. An important observation is that (see [6])

$$\bar{\mathcal{N}} := \mathcal{N}(\bar{\mathbf{p}}) = \{m \in \mathcal{N} : m = \arg\max_{n\in\mathcal{N}} \rho(\mathbf{B}^{(n)})\} \tag{8}$$

where $\rho(\mathbf{B}^{(n)})$ is used to denote the spectral radius of $\mathbf{B}^{(n)}$.[2] Using this definition, we can rewrite (7) as

$$\frac{1}{\beta}\bar{\mathbf{p}} = \mathbf{\Gamma}\mathbf{G}\bar{\mathbf{p}} + \mathbf{\Gamma}\mathbf{z}\cdot g_n(\bar{\mathbf{p}}),\ n \in \bar{\mathcal{N}} \tag{9}$$

which is simply because $g_n(\bar{\mathbf{p}}) = 1$ for each $n \in \bar{\mathcal{N}}$. Thus, with (5), we have (for each $n \in \bar{\mathcal{N}}$)

$$\frac{1}{\beta}\bar{\mathbf{p}} = \mathbf{B}^{(n)}\bar{\mathbf{p}}, \quad \beta > 0, \bar{\mathbf{p}} \in \mathbb{R}^K_{++}, g_n(\bar{\mathbf{p}}) = 1 \tag{10}$$

where $\mathbf{B}^{(n)} \in \mathbb{R}^{K\times K}_{+}$ is defined to be (for each $n \in \mathcal{N}$)

$$\mathbf{B}^{(n)} := \mathbf{\Gamma}\mathbf{G} + \frac{1}{P_n}\mathbf{\Gamma}\mathbf{z}\mathbf{c}_n^T = \mathbf{\Gamma}\big(\mathbf{G} + 1/P_n\mathbf{z}\mathbf{c}_n^T\big)\,. \tag{11}$$

So, given $\bar{\mathbf{p}}$ defined by (4), the condition (10) must hold for each $n \in \bar{\mathcal{N}}$. In other words, if $\mathbf{G}$ is irreducible, a max-min SIR-balanced power vector (4) must satisfy (10) for each node $n \in \mathcal{N}$ whose power constraint is active at the maximum. Since $\mathbf{B}^{(n)}, n \in \mathcal{N}$, is irreducible, the converse can be concluded from the Perron-Frobenius theorem [9, 10].

Proposition 1 ([6, 5]) *Let (A.1) be satisfied, and let $\beta > 0$ be given by (6). Then, $\mathbf{p} = \bar{\mathbf{p}} > 0$ if and only if $\mathbf{p}$ is the right principle eigenvector of $\mathbf{B}^{(n)}, n \in \bar{\mathcal{N}}$, normalized such that $g_n(\mathbf{p}) = 1$. Moreover, $1/\beta = \rho(\mathbf{B}^{(n)}), n \in \bar{\mathcal{N}}$.*

Now we are in a position to establish the aforementioned connection between the max-min SIR-balancing problem and a utility maximization problem for any fixed beamformers. To this end, given any $\mathbf{V} \in \mathrm{V}$, consider the following aggregate utility function

$$U(\mathbf{p}, \mathbf{w}) := \sum_{k\in\mathcal{K}} w_k \phi\Big(\frac{\mathrm{SIR}_k(\mathbf{p}, \mathbf{v}_k)}{\gamma_k}\Big) \tag{12}$$

where ϕ fulfills (A.2) and (A.3):

(A.2) $\phi : \mathbb{R}_{++} \to \mathrm{Q} \subseteq \mathbb{R}$ is continuously differentiable and a strictly increasing function.

(A.3) $g(x) := \phi^{-1}(x)$ is log-convex [5, 11].

Moreover, without loss of generality, it is assumed that

(A.4) $\mathbf{w} \in \Pi^+_K := \{\mathbf{u} \in \mathbb{R}^K_{++} : \|\mathbf{u}\|_1 = 1\}$.

[2] Since $\mathbf{B}^{(n)}, n \in \mathcal{N}$, is irreducible, it follows from Perron-Frobenius theorem [9] that $\rho(\mathbf{B}^{(n)})$ is a simple eigenvalue of $\mathbf{B}^{(n)}$. Moreover, left and right eigenvectors of $\mathbf{B}^{(n)}$ associated with $\rho(\mathbf{B}^{(n)})$ are positive and unique up to positive multiples. These eigenvectors are called (left and right) principle eigenvectors of $\mathbf{B}^{(n)}$.

The utility maximization problem of interest is to maximize the aggregate utility function (12) over P_+:

$$\mathbf{p}(\mathbf{w}) := \arg\max\nolimits_{\mathbf{p}\in\mathrm{P}_+} U(\mathbf{p}, \mathbf{w}) \tag{13}$$

where the maximum can be shown to exist [5]. Furthermore, it is obvious that in the maximum at least one power constraint is active:

$$\forall_{\mathbf{w}\in\Pi^+_K} \exists_{n\in\mathcal{N}}\ g_n(\mathbf{p}(\mathbf{w})) = 1\,. \tag{14}$$

Now, the following connection between (4) and (13) can be observed.

Proposition 2 ([6, 5]) *Suppose that $\mathbf{y}^{(n)}$ and $\mathbf{x}^{(n)}$ are left and right principle eigenvectors of $\mathbf{B}^{(n)}, n \in \mathcal{N}$. Define*

$$\mathbf{w}^{(n)} := \mathbf{y}^{(n)} \circ \mathbf{x}^{(n)},\ (\mathbf{y}^{(n)})^T\mathbf{x}^{(n)} = 1, n \in \mathcal{N}\,. \tag{15}$$

Then, $\bar{\mathbf{p}} = \mathbf{p}(\mathbf{w})$ whenever $\mathbf{w} = \mathbf{w}^{(m)}$ for some $m \in \bar{\mathcal{N}}$.

Interestingly, this connection is independent of the choice of the utility function ϕ, provided that ϕ fulfills (A.2) and (A.3). The basic idea behind Proposition 2 is illustrated in Figure 1. For some ϕ satisfying (A.2)–(A.3), the figure depicts the set $\mathrm{F} \subset \mathrm{Q}^K$ (for the two-user case) defined to be

$$\mathrm{F} = \{\mathbf{q} \in \mathrm{Q}^K : q_k = \phi\big(\mathrm{SIR}_k(\mathbf{p})/\gamma_k\big), k \in \mathcal{K}, \mathbf{p} \in \mathrm{P}_+\}\,. \tag{16}$$

Note that F can be interpreted as a feasible QoS region where the QoS value for link k is $\phi(\mathrm{SIR}_k(\mathbf{p})/\gamma_k)$. We define the boundary of F (denoted by $\partial\mathrm{F}$) to be the set of all points of F such that, if $\mathbf{p}$ is the corresponding power vector in (16), then $\mathbf{C}\mathbf{p} \le \mathbf{p}'$ holds with at least one equality. Let

$$\bar{q}_k = \phi(\mathrm{SIR}_k(\bar{\mathbf{p}})/\gamma_k), k \in \mathcal{K}$$

and

$$q_k(\mathbf{w}) = \phi\big(\mathrm{SIR}_k(\mathbf{p}(\mathbf{w}))/\gamma_k\big), k \in \mathcal{K}$$

where $\bar{\mathbf{p}}$ and $\mathbf{p}(\mathbf{w})$ are defined by (4) and (13), respectively. Both

$$\bar{\mathbf{q}} = (\bar{q}_1, \dots, \bar{q}_K) \in \partial\mathrm{F}$$

and

$$\mathbf{q}(\mathbf{w}) = (q_1(\mathbf{w}), \dots, q_K(\mathbf{w})) \in \partial\mathrm{F}$$

are boundary points as at least one power constraint is active in the optimum. The figure shows an example of the feasible QoS region under individual power constraints ($\mathbf{C} = \mathbf{I}$). The boundary is given by the following condition: $\mathbf{q} \in \partial\mathrm{F}$ if and only if $\max_{n\in\mathcal{N}} \rho(\mathbf{Q}(\mathbf{q})\mathbf{B}^{(n)}) = 1$, where $\mathbf{Q}(\mathbf{q}) := diag(g(q_1), \dots, g(q_K))$ with $g(x)$ defined by (A.3). The point $\bar{\mathbf{q}} = \mathbf{q}(\mathbf{w}^{(m)})$ for some $m \in \bar{\mathcal{N}}$, corresponds to the unique max-min SIR-balanced power allocation. The weight vector $\mathbf{w}^{(m)}, m \in \bar{\mathcal{N}}$, is normal to a hyperplane which supports the feasible QoS region at $\bar{\mathbf{q}} \in \partial\mathrm{F}$. Note that in Figure 1, $\bar{\mathcal{N}} = \{1\}$.

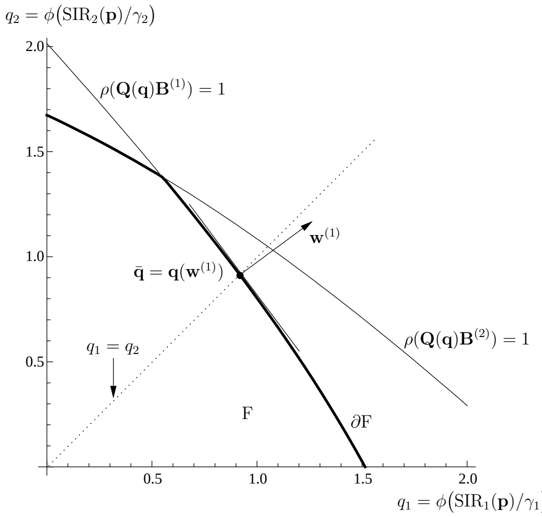


Fig. 1. An illustration of Proposition 2 [6]. We have $\mathbf{w} = \mathbf{w}^{(1)}$ and $\bar{\mathcal{N}} = \{1\}$: The power constraint of the first link is active under the max-min SIR-balanced power vector $\bar{\mathbf{p}} = \mathbf{p}(\mathbf{w}^{(1)})$.

Proposition 2 shows how to choose the weight vector $\mathbf{w} \in \Pi_K^+$ in (12) so that the max-min SIR-balanced power vector $\bar{\mathbf{p}}$ given by (4) is equal to $\mathbf{p}(\mathbf{w})$ given by (13). We have the following interesting result that leads to Proposition 2 and is a key ingredient in the proof of a saddle point of the Perron roots of $\mathbf{B}^{(n)}, n \in \mathcal{N}$ (which are irreducible matrices by assumption): If $\psi(x) = -\phi(1/x), x > 0$ (with (A.2) and (A.3)), then [5] (with $n \in \mathcal{N}$)

$$\psi(\rho(\mathbf{B}^{(n)})) \le G_n(\mathbf{p},\mathbf{w}) := \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{(\mathbf{B}^{(n)}\mathbf{p})_k}{p_k}\Big) \quad (17)$$

holds for *all* $\mathbf{p} > 0$ if and only if $\mathbf{w} = \mathbf{w}^{(n)} \in \Pi_K^+$ given by (15). Moreover, if $\mathbf{w} = \mathbf{w}^{(n)} \in \Pi_K^+$, then (17) holds with equality if and only if $\mathbf{p} = \mathbf{x}^{(n)} > 0$, that is, if and only if $\mathbf{p}$ is the right principal eigenvector of $\mathbf{B}^{(n)}$.

On the other hand, since $\max_{n\in\mathcal{N}} g_n(\mathbf{p}) \le 1$ for all $\mathbf{p} \in \mathrm{P}_+$, $\psi(x) = -\phi(1/x), x > 0$, is strictly increasing if (A.2) and (A.3) are satisfied, the diagonal matrix $\mathbf{\Gamma}$ is positive definite and $\mathbf{z} > 0$, we have

$$\begin{aligned} G_n(\mathbf{p},\mathbf{w}) &= \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{[\mathbf{\Gamma}(\mathbf{G} + 1/P_n \mathbf{z}\mathbf{c}_n^T)\mathbf{p}]_k}{p_k}\Big) \\ &= \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{(\mathbf{\Gamma}\mathbf{G}\mathbf{p} + \mathbf{\Gamma}\mathbf{z} g_n(\mathbf{p}))_k}{p_k}\Big) \\ &\le \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{(\mathbf{\Gamma}\mathbf{G}\mathbf{p} + \mathbf{\Gamma}\mathbf{z})_k}{p_k}\Big) \\ &= F(\mathbf{p},\mathbf{w}),\ n \in \mathcal{N}, \mathbf{p} \in \mathrm{P}_+, \mathbf{w} \in \Pi_K^+ \end{aligned} \quad (18)$$

with equality if and only if $g_n(\mathbf{p}) = 1$ where

$$F(\mathbf{p},\mathbf{w}) := \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{\gamma_k}{\mathrm{SIR}_k(\mathbf{p})}\Big). \quad (19)$$

So, considering (17), (18) and (15) shows that (note that $G_n(\mathbf{p},\mathbf{w})$ has the ray property with respect to $\mathbf{p}$)

$$\begin{aligned} \psi(\rho(\mathbf{B}^{(n)})) &= \min_{\mathbf{p}>0} G_n(\mathbf{p},\mathbf{w}^{(n)}) \\ &= \max_{\mathbf{w}\in\Pi_K^+} \min_{\mathbf{p}>0} G_n(\mathbf{p},\mathbf{w}) \\ &= \max_{\mathbf{w}\in\Pi_K^+} \min_{\mathbf{p}\in\mathrm{P}_+} G_n(\mathbf{p},\mathbf{w}) \\ &\le \max_{\mathbf{w}\in\Pi_K^+} \min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p},\mathbf{w}) = \min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p},\mathbf{w}^*), n \in \mathcal{N} \end{aligned} \quad (20)$$

with equality if and only if $g_n(\mathbf{p}(\mathbf{w}^*)) = 1$, where $\mathbf{w}^* \in \Pi_K^+$ is a maximizer of $\min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p},\mathbf{w})$ and $\mathbf{p}(\mathbf{w}^*)$ is a minimizer of $\mathbf{p} \mapsto F(\mathbf{p},\mathbf{w}^*)$ over P_+ (see also (13)[3]). By (14), we know that $\mathcal{N}' = \{m \in \mathcal{N} : g_m(\mathbf{p}(\mathbf{w}^*)) = 1\}$ is a nonempty set. Thus, by (18) and (20),

$$\psi(\rho(\mathbf{B}^{(n)})) < \psi(\rho(\mathbf{B}^{(m)})) = \min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p},\mathbf{w}^*) \quad (21)$$

for each $m \in \mathcal{N}'$ and $n \notin \mathcal{N}'$. So, by (8), we obtain $\mathcal{N}' = \bar{\mathcal{N}}$, which together with (17) and (21) implies that

$$\mathbf{w}^{(m)} \in \mathrm{W} := \big\{\mathbf{w}^* = \arg\max_{\mathbf{w}\in\Pi_K^+} \min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p},\mathbf{w})\big\}.$$

for each $m \in \bar{\mathcal{N}}$.

Consequently, the max-min SIR-balancing power vector can also be obtained by minimizing $F(\mathbf{p},\mathbf{w}^{(m)})$ for any $m \in \bar{\mathcal{N}}$ over P_+. From (21), we see that $\rho(\mathbf{B}^{(m)}) > \rho(\mathbf{B}^{(n)})$ with $m \in \bar{\mathcal{N}}$, $n \notin \bar{\mathcal{N}}$ and $\beta^{(m)} < \beta^{(n)}$. Figure 1 illustrates these facts. The line $q_1 = q_2$ intersects both individual boundaries but only the intersection point on $\rho(\mathbf{B}^{(1)}) = 1$ satisfies both power constraints so that we have $\bar{\mathcal{N}} = \{1\}$.

Finally, note that by the above discussion, we can relax the power constraint $\mathbf{p} \in \mathrm{P}_+$ to $g_m(\mathbf{p}) \le 1, \mathbf{p} > 0$, for any $m \in \bar{\mathcal{N}}$. Thus, instead of minimizing $\mathbf{p} \mapsto F(\mathbf{p},\mathbf{w}^{(m)}), m \in \bar{\mathcal{N}}$ over P_+, we can consider

$$\mathbf{p}^{(m)} = \arg\min_{\mathbf{p}>0, g_m(\mathbf{p})\le 1} F(\mathbf{p},\mathbf{w}^{(m)}),\ m \in \bar{\mathcal{N}}. \quad (22)$$

Alternatively, one can minimize $\mathbf{p} \mapsto G_m(\mathbf{p},\mathbf{w}^{(m)})$ over $\mathbb{R}_{++}^K$, which can be also deduced from (20).

[3]Note that any power vector minimizing $\mathbf{p} \mapsto F(\mathbf{p},\mathbf{w})$ over P_+ for some given weight vector $\mathbf{w} > 0$ is optimal in the sense of the utility maximization problem (13) where the utility is defined by (12).

see the program

4. JOINT POWER CONTROL AND BEAMFORMING

4.1. Joint Power Control and Receive Beamforming

Based on the insights provided by the previous section we propose an algorithmic solution to find the max-min SIR-balanced resource allocation defined in (3) and prove its convergence to the optimum. Instead of directly solving problem (3) we find the jointly optimal transmit powers and receive beamformers by solving the following utility-based optimization problem

$$\min_{(\mathbf{p},\mathbf{V})\in \mathrm{P}_+\times \mathrm{V}} F(\mathbf{p},\mathbf{V},\mathbf{w})\,, \tag{23}$$

where $F(\mathbf{p},\mathbf{V},\mathbf{w})$ is defined in (19) with $\mathrm{SIR}_k(\mathbf{p},\mathbf{v}_k)$. The connection between the max-min SIR-balancing power control problem (4) and the utility-based power control problem (13) assuming a fixed gain matrix $\mathbf{G}(\mathbf{V})$ (i.e. fixed receive beamforming matrix $\mathbf{V}$) was already established in [6] and has been summarized in the previous section. Now, in order to solve (23), the basic idea is to perform an iterative algorithm that updates the power vector and the receive beamforming matrix in an alternating way. More precisely, the algorithm keeps one of the variables $\mathbf{V}$ and $\mathbf{p}$ fixed while optimizing with respect to the other. The alternating process is repeated until the termination condition is satisfied. The termination condition may be for instance a predefined number of iterations or a sufficiently small relative improvement $\delta > 0$. The proposed iterative algorithm is summarized in Alg. (1). In the following we describe its single steps in detail.

Algorithm 1 Joint power control and receive beamforming achieving the max-min SIR-balancing solution

Input: $\delta > 0, t = 0, \mathbf{V}(0) \in \mathrm{V}$

repeat
- $t = t + 1$
- determine some $m \in \bar{\mathrm{N}}$
- $\mathbf{w}^{(m)}(t) = c \cdot \mathbf{y}(\mathbf{B}^{(m)}(t-1)) \circ \mathbf{x}(\mathbf{B}^{(m)}(t-1))$
- $\mathbf{p}(t) = \arg\min_{\mathbf{p}>0, g_m(\mathbf{p})\le 1} F(\mathbf{p}, \mathbf{w}^{(m)}(t))$
- $\mathbf{v}_k(t) = \arg\max_{||\mathbf{v}_k||_2=1} \mathrm{SIR}_k(\mathbf{p}(t), \mathbf{v}_k) \quad \forall_k$

until termination condition is satisfied

Consider iteration t. For some given beamforming matrix $\mathbf{V}(t-1)$ we first update the power vector $\mathbf{p}(t)$. Remember, if we fix the receive beamformers problem (23) reduces to a pure power control problem. In order to obtain the max-min SIR-balancing power vector solving the pure power control problem is equivalent to solving problem (22) as shown in the previous section. Here, the active power constraint, say m, is determined by (8). The weight vector $\mathbf{w}^{(m)}(t)$ is chosen applying proposition 2. The power control problem can be solved by means of a gradient projection algorithm [5].

Then we fix the power vector $\mathbf{p}(t)$ and recalculate the receive beamforming matrix $\mathbf{V}(t)$. Since the k-th receiver influences only the k-th SIR and has no impact on other SIRs we can focus on the single link k. An optimal linear receive beamformer of user k is exactly that beamformer for which the k-th SIR attains its maximum (for a fixed power vector)

$$\mathbf{v}_k(t) = \arg\max_{||\mathbf{v}_k||_2=1} \mathrm{SIR}_k(\mathbf{p}(t),\mathbf{v}_k) \quad \forall_k. \tag{24}$$

Then we obtain

$$\min_k \frac{\mathrm{SIR}_k(\mathbf{p}(t),\mathbf{v}_k(t-1))}{\gamma_k} \le \min_k \frac{\mathrm{SIR}_k(\mathbf{p}(t),\mathbf{v}_k(t))}{\gamma_k}.$$

The proposed algorithm converges to the optimum max-min SIR-balanced resource allocation $(\mathbf{p}^*, \mathbf{V}^*)$ defined in (1). This will be proven in the remainder of this subsection. We start by showing monotonicity.

Lemma 1 *Assume that $\mathbf{G}(\mathbf{V}(t))$ is irreducible. Then, the sequence*

$$\beta(t) := 1/\rho(\mathbf{B}^{(m)}(t)) \quad m \in \bar{\mathrm{N}} \tag{25}$$

is monotonically increasing in t.

We have the following

$$\begin{aligned}
\psi(1/\beta(t)) :&= \psi(\rho(\mathbf{B}^{(m)}(t))) \qquad m \in \bar{\mathrm{N}}\\
&= F(\mathbf{p}(t), \mathbf{w}^{(m)}(t), \mathbf{V}(t))\\
&\ge F(\mathbf{p}(t), \mathbf{w}^{(m)}(t), \mathbf{V}(t+1)))\\
&\ge \min_{\mathbf{p}>0, g_m(\mathbf{p})\le 1} F(\mathbf{p}, \mathbf{w}^{(m)}(t+1), \mathbf{V}(t+1))\\
&= F(\mathbf{p}(t+1), \mathbf{w}^{(m)}(t+1), \mathbf{V}(t+1))\\
&= \psi(\rho(\mathbf{B}^{(m)}(t+1))) = \psi(1/\beta(t+1))
\end{aligned}$$

The first inequality follows from the fact that the SIR of user k depends only on the k-th receive beamformer. Thus maximizing the SIR (see (24)) decreases the utility measure for a fixed power and weight vector. The second inequality follows from the previous section where it was shown that the max-min SIR-balancing power vector can be obtained by solving (22). Now, due to the fact that $\psi(x)$ is strictly increasing the sequence $\beta(t)$ is monotonically increasing.

Next it will be shown that the monotone sequence $\beta(t)$ converges to the optimum.

Theorem 1 *The proposed algorithm converges to the optimum max-min SIR-balanced value (27), i.e.*

$$\lim_{t\to\infty} \beta(t) = \beta^* \tag{26}$$

see the program

where

$$\beta^* = \min_{k \in \mathcal{K}} \left(\mathrm{SIR}_k(\mathbf{p}^*, \mathbf{v}_k^*)/\gamma_k \right) \tag{27}$$

is the minimum SIR value under a max-min SIR-balanced solution (3).

The theorem can be proven by proceeding essentially as in [1], with the matrix $\mathbf{\Gamma G}$ substituted by $\mathbf{B}^{(m)}$ for some $m \in \bar{\mathcal{N}}$ and the power vectors $\mathbf{p}(t), t \in \mathbb{N}$, normalized such that $g_m(\mathbf{p}(t)) = 1, m \in \bar{\mathcal{N}}$.

4.2. Implementation Issues

In this section we discuss practical implementation issues concerning the receive beamformer and power updates. First we review aspects of the receive beamformer optimization. It is widely known that an explicit computation of the receive beamformers is intricate due to the lack of perfect synchronization and multipath propagation in many wireless networks. Adaptive algorithms that converge to the optimal receiver seem to be the only suitable option, especially in distributed wireless networks. Distributed algorithms for computing optimal receive beamformers defined by (24) are widely established. These algorithms are based either on blind or pilot-based estimation methods.

Next we consider the power updates. Under the assumption that each user knows its weight, the power control problem can be solved using a gradient projection algorithm that can be implemented distributedly applying the concept of the adjoint network [5]. Unfortunately the calculation of the weight vector is not amenable to distributed implementation. However, the results presented in this paper and the saddle point characterization of the Perron root of $\mathbf{B}^{(n)}, n \in \bar{\mathcal{N}}$ in [6] provide a basis for the development of novel decentralized algorithmic solutions.

Finally we emphasize, that in practice it is sufficient to perform some $L_p \geq 1$ steps of the gradient projection algorithm and some $L_v \geq 1$ beamformer updates during the t-th iteration of Alg. 1 to limit the algorithm complexity.

4.3. Transceiver Optimization

The effective transmit vector is affected by the transceiver structure, the channel state etc. These vectors determine how the transmit signals direct into and through the space. Assuming the transceiver structure to be a transmit beamformer any user can influence its transmit vectors $\mathbf{u}_k^{(l)}$ by changing its transmit beamformer. The optimization of transmit beamforming is crucial, because in contrast to receive beamforming it impacts the interference power to all other receivers and thus the performance of the other users. However, depending on the interference scenario a transmit beamformer optimization may achieve high performance gains.

Related work to both transmitter and joint transmitter and receiver optimization can be found in [12, 13, 4]. However, it has been shown that the virtual uplink concept to calculate transmit beamformers and powers cannot be applied in networks with general power constraints [12]. Next we present a heuristic scheme that recalculates the transmit beamformers as long as the balanced SIR values improve.

As in [13] the basic idea is to optimize the transmit and receive beamformers in an alternating way. For a better understanding we use the notion of the primal and the reversed network. In contrast to the primal network the reversed network is the network that is obtained by reversing the roles of transmitters and receivers. More precisely, transmitters and transmit beamformers are assumed to be receivers and receive beamformers, respectively. Vice versa receivers and receive beamformers are assumed to be transmitters and transmit beamformers. Each iteration of the joint optimization of powers, receive and transmit beamformers consists of 4 steps, where 2 steps are executed in the primal and in the reversed network, respectively. These 2 steps include the receive beamformer optimization and the update of the power vector. The proposed algorithm is summarized in Alg. 2. Next we describe the single steps in detail.

The algorithm performs the following steps. First each user updates its receive beamformers $\mathbf{v}_k(t)$ in the primal network such that the SIR is maximized. Therefore each user fixes its transmit power $\mathbf{p}(t-1)$ and transmit beamforming vector $\mathbf{s}_k(t-1), k \in \mathcal{K}$. The transmit beamforming vectors are collected in the transmit beamforming matrix $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_K) \in \mathbb{C}^{M \times K}, \mathbf{S} \in \mathrm{S}$ where $\mathrm{S} = \{(\mathbf{x}_1, \dots, \mathbf{x}_K) : \|\mathbf{x}_k\|_2 = 1\}$ denotes the set of all transmit beamforming matrices. Then we switch to the reversed network. Remember, in the reversed network the transmit beamformers $\mathbf{s}_k, k \in \mathcal{K}$ are assumed to be receive beamformers and the receive beamformers $\mathbf{v}_k, k \in \mathcal{K}$ are assumed to be transmit beamformers. The transmit powers of the reversed network are collected in the power vector $\mathbf{q} = (q_1, \dots, q_K) \in \mathbb{R}_+^K$ that is subject to some power constraints $\mathbf{q} \in \mathrm{Q}$. We will discuss this issue later. Now, in the reversed network first we perform the power update for fixed transmit and receive beamformers. The power allocation $\mathbf{q}(t)$ for the reversed network is computed such that the minimum balanced SIR-value is maximized. The power computation was explained in section 4.1 and can be found in Alg. 1. Then the receive beamformers of the reversed network $\mathbf{s}_k(t), k \in \mathcal{K}$ are updated such that the SIR of each user is maximized. Finally we switch back to the primal network and calculate the power allocation $\mathbf{p}(t)$. The algorithm terminates if the balanced SIR values decrease, a sufficiently small relative improvement is achieved. In addition the algorithm may also be stopped after a predefined number of iterations.

In a network with general power constraints a transceiver

see the program

optimization has some inherent problems. As mentioned before the uplink-downlink duality does not hold. More considerable is the fact that the power constraints in a general network may not be equivalent in the reversed network. For example, if we consider a downlink that has a sum power constraint, the corresponding reversed network is usually represented by an uplink with individual power constraints for each user. The power constraints are defined by $\mathbb{Q} = \{\mathbf{q} \in \mathbb{R}_+^K : \mathbf{C}'\mathbf{q} \leq \mathbf{q}', \mathbf{C}' \in \{0,1\}^{N' \times K}\} \subset \mathbb{R}^K$ for some given $\mathbf{q}' = (Q_1, \ldots, Q_{N'}) > 0$ and $\mathbf{C}'$ with at least one 1 in each column so that $\mathbb{Q}$ is a compact set. N' is the number of power constraints. To be more precise, compared to the primal network, different subsets of links may have (different) sum power constraints. So we cannot ensure an improvement with each iteration step. This is incorporated in the termination condition.

Algorithm 2 Heuristic Transceiver Optimization

Input: $\delta > 0, t = 0, \mathbf{p}(0) \in \mathbb{P}_+, \mathbf{V}(0) \in \mathbb{V}, \mathbf{S}(0) \in \mathbb{S}$

repeat
  $t = t + 1$
  primal network: given $\mathbf{p}(t-1)$ and $\mathbf{S}(t-1)$ compute receive beamformers $\mathbf{v}_k(t), k \in \mathcal{K}$
  reversed network: compute power allocation $\mathbf{q}(t)$
  reversed network: given $\mathbf{q}(t)$ and $\mathbf{V}(t)$ compute receive beamformers $\mathbf{S}(t)$
  primal network: compute power allocation $\mathbf{p}(t)$
until some termination condition is satisfied

5. NUMERICAL RESULTS

We consider a network with K users, $M = 4$ transmit and receive antennas and a channel matrix which entries are iid complex Gaussian distributed. Each user is subject to the same individual power constraint and operates at SNR$= 30$dB. The SIR targets are chosen to be $\mathbf{\Gamma} = \mathbf{I}$.

Figure 2 shows the average number of iterations t that is required to fulfill the termination condition $\delta = 10^{-2}$ (relative improvement) provided that the power and beamformer updates during each iteration are executed with the same accuracy. The number of iterations is averaged over 1000 channel realizations and is depicted over the number of users in the network. Note, with an increasing number of users the interference in the network increases. As can be seen the algorithm seems to converge quite fast in approximately 3 iterations to the optimum. Note, if we relax the accuracy of the power and beamformer updates or limit the number of updates L_p, L_v during each iteration the total number of iterations t increases. This, however, strongly depends on the choice of the step size.

Figures 3 and 4 illustrate for an exemplarily network the absolute and relative gain (relative with respect to the simplest strategy) that can be achieved by different resource allocation schemes. More precisely, the max-min SIR-balanced value averaged over 1000 channel realizations is depicted over the number of users in the network for the following 4 resource allocation strategies. If not otherwise stated the beamformers of the schemes adapt to the channel. The first strategy, that is also used as reference strategy in figure 4, allocates the maximum power to each user (max. power) whereas the second scheme allocates the max-min SIR power vector to the users (max-min power). The third strategy solves the joint receive beamforming and power allocation problem applying algorithm 1 (Alg. 1) and the fourth scheme additionally optimizes the transceivers applying algorithm 2 (Alg. 2). As can be seen a sophisticated resource allocation scheme may significantly improve the network performance.

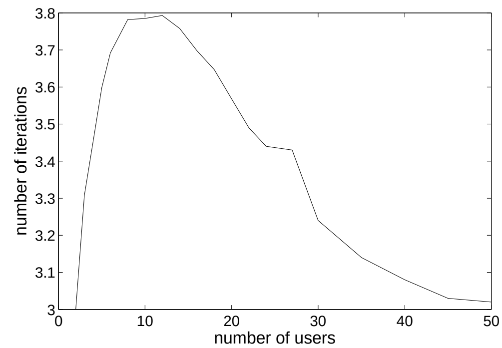


Fig. 2. Average number of iterations

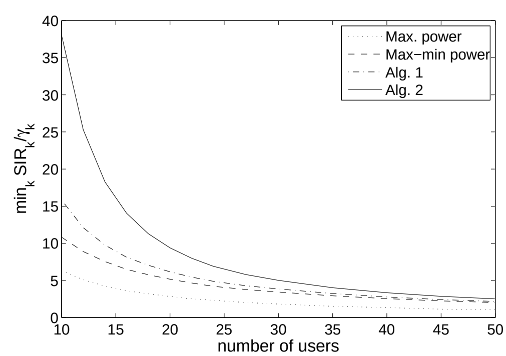


Fig. 3. Balanced SIR over the number or users for different resource allocation schemes

6. CONCLUSIONS

The paper has addressed the problem of jointly optimizing powers and receive beamformers in a wireless network

see the program

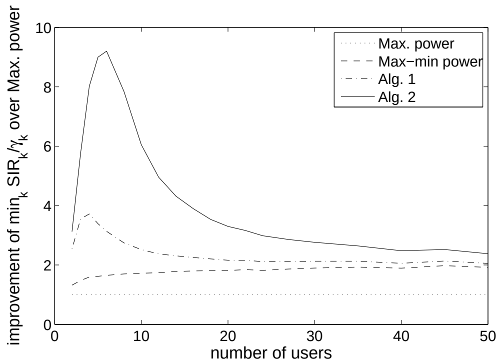


Fig. 4. Relative improvement with respect to the maximum power scheme (Max. power) for different resource allocation schemes

subject to general power constraints. More precisely, we solved the max-min SIR-balancing problem by maximizing a weighted sum of utilities of the SIRs over the joint space of powers and receive beamformers. The authors of [1] described a similar alternating scheme to solve the joint max-min fair power control and pre-equalizer problem in the noiseless case for a multiuser downlink channel. They achieved the max-min SIR-balancing solution by minimizing the sum of weighted inverse SIRs. Thus, due to their problem structure the single steps of the algorithm are different. In fact, the generalization provided by our paper is in two directions. We extended the results to noisy channels under general power constraints and considered a larger class of utility functions. However, we only provided optimal powers and receive beamformers. For transmit beamformer optimization we described a heuristic approach. This is among other things due to the fact that the theory of uplink-downlink duality is not applicable in case of general power constraints.

Our proposed algorithm converges to a max-min SIR-balancing solution. It is not amenable to distributed implementation but provides a basis for the development of novel decentralized algorithms.

7. REFERENCES

[1] H. Boche and M. Schubert, “Resource allocation in multi-antenna systems - Achieving max-min fairness by optimizing a sum of inverse SIRs,” *IEEE Trans. Signal Processing*, vol. 54, no. 6, June 2006.

[2] J. Zander and S.-L. Kim, *Radio Resource Management for Wireless Networks*, Artech House, Boston, London, 2001.

[3] W. Yang and G. Xu, “Optimal downlink power assignment for smart antenna systems,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Seattle, WA, USA, May 1998.

[4] M. Schubert and H. Boche, “Qos-based resource allocation and transceiver optimization,” *Foundation and Trends in Communications and Information Theory*, vol. 2, no. 6, 2006.

[5] S. Stanczak, M. Wiczanowski, and H. Boche, *Fundamentals of Resource Allocation in Wireless Networks*, vol. 3 of *Foundations in Signal Processing, Communications and Networking*, Springer, Berlin, 2009, (to appear).

[6] S. Stanczak, M. Kaliszan, N. Bambos, and M. Wiczanowski, “A characterization of max-min SIR-balanced power allocation with applications,” Available at http://arxiv.org/abs/0901.0824.

[7] J. Huang, R. Berry, and M. L. Honig, “Distributed interference compensation for wireless networks,” *IEEE J. Select. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[8] S. Stanczak, A. Feistel, and D. Tomecki, “On utility-based power control and receive beamforming,” in *Proc. 41st Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, March 14-16 2007.

[9] E. Seneta, *Non-Negative Matrices and Markov Chains*, Springer, Berlin, 1981.

[10] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[12] F. Rashid-Farrokhi, K. J. Liu, and L. Tassiulas, “Transmit beamforming and power control for cellular wireless systems,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 16–27, Jan. 2002.

[13] J. Chang, L. Tassiulas, and F. Rashid-Farrokhi, “Joint transmitter receiver diversity for efficient space division multiaccess,” *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1437–1449, Oct. 1998.

see the program

On Proportional Fairness in Nonconvex Wireless Systems

Johannes Brehmer, Wolfgang Utschick
Associate Institute for Signal Processing
Technische Universität München
{brehmer, utschick}@tum.de

***Abstract*—We investigate the concept of proportional fairness in the context of nonconvex rate regions. For convex rate regions, Kelly's definition of proportional fairness is equivalent to maximizing a sum of logarithms. For nonconvex rate regions, this equivalence does not always hold. In particular, a proportionally fair point in the sense of Kelly's definition may not exist. We propose the concept of local proportional fairness (LPF) and characterize its properties. Under mild conditions, an LPF point always exists, and all LPF points are Pareto optimal. Moreover, we show that in a multiantenna downlink with linear precoding, computing an LPF point represents a practically solvable problem.**

I. Introduction

We consider a wireless system in which $K > 1$ users compete for limited resources. A fundamental problem in such a system is to find a criterion on which resource allocation is to be based. In his seminal paper [1], Kelly proposed the concept of *proportional fairness* (PF) as a criterion for resource allocation. Let $\mathcal{R} \subset \mathbb{R}_+^K$ denote the set of feasible rates. Throughout this paper, it is assumed that $\mathcal{R}$ is compact and nontrivial. We say that $\mathcal{R}$ is nontrivial if there exists $\boldsymbol{r} \in \mathcal{R}$ such that $r_k > 0, \forall k$. Kelly defines PF as follows [1]: A vector of rates $\boldsymbol{r}^* \in \mathcal{R} \cap \mathbb{R}_{++}^K$ is proportionally fair if for any other feasible vector $\boldsymbol{r}$, the aggregate of proportional changes is zero or negative, i.e.,

$$\sum_{k=1}^{K} \frac{r_k - r_k^*}{r_k^*} \leq 0, \forall \boldsymbol{r} \in \mathcal{R}. \tag{1}$$

Kelly also showed the relationship between proportional fairness and maximizing a sum of logarithms,

$$\max_{\boldsymbol{r} \in \mathcal{R}} \sum_{k=1}^{K} \ln r_k. \tag{2}$$

Let $\mathcal{R}^*$ denote the set of global maximizers of (2). If $\mathcal{R}$ is convex, compact, and nontrivial, then $\mathcal{R}^*$ has a single element, and this rate vector is proportionally fair [1]. From this result, it follows that if $\mathcal{R}$ is convex, compact, and nontrivial, a proportionally fair vector always exists, and the proportionally fair rate allocation is unique.

The performance of wireless systems can be improved significantly by allowing for interference between users. This particularly holds true in multiantenna systems, where optimum transmission strategies often do not result in a complete interference suppression. On the other hand, interference can result in nonconvex rate regions, see, e.g., [2]. While any rate region can be convexified by time-sharing, a time-sharing mode can significantly complicate the problem of finding an optimum resource allocation [3]. This paper investigates the concept of proportional fairness in the context of compact, but not necessarily convex rate regions.

If $\mathcal{R}$ is not convex, the existence of a feasible rate vector that fulfills the definition of proportional fairness (1) cannot be guaranteed – in other words, there exist nontrivial compact sets $\mathcal{R}$ such that no vector $\boldsymbol{r} \in \mathcal{R}$ is proportionally fair. We provide an example of such a set in Section II. Based on this observation, it is desirable to extend the concept of proportional fairness in a way such that a proportionally fair rate vector (in the extended sense) always exists.

For compact and nontrivial $\mathcal{R}$, the set $\mathcal{R}^*$ is always well-defined and nonempty . Accordingly, one approach to extend the concept of proportional fairness to nonconvex $\mathcal{R}$ is to declare all points in $\mathcal{R}^*$ as proportionally fair. This approach was taken in [4], [5], [6], [7], [8]. As pointed out before, however, none of the maximizers in $\mathcal{R}^*$ may fulfill (1). Moreover, Kelly's definition is based on the notion of aggregates of proportional changes – that the proportionally fair point maximizes a sum of logarithms (if $\mathcal{R}$ is convex) is a property of PF, but not its definition. Based on this observation, we propose the concept of *local* proportional fairness: Let $\hat{\mathcal{R}}(\boldsymbol{r}^*)$ denote a local approximation of $\mathcal{R}$ at $\boldsymbol{r}^*$ based on the tangent cone of $\mathcal{R}$ at $\boldsymbol{r}^*$. We say that a vector of rates $\boldsymbol{r}^*$ is *locally proportional fair* (LPF) if $\boldsymbol{r}^* \in \mathcal{R} \cap \mathbb{R}_{++}^K$ and if for any other vector $\boldsymbol{r} \in \hat{\mathcal{R}}(\boldsymbol{r}^*)$, the aggregate of proportional changes is zero or negative.

We show the following results:

1) If $\mathcal{R}$ is compact and nontrivial, there always exists an LPF point.
2) If $\mathcal{R}$ is compact, nontrivial, and normal, all LPF points are Pareto optimal.
3) If $\mathcal{R}$ is convex, LPF and PF are equivalent.

In addition, we demonstrate that in a multiantenna downlink with linear precoding, an LPF point can be found by standard local methods for unconstrained optimization. This result shows that in terms of problem complexity, there is a fundamental difference between LPF and maximizing a sum of logarithms. For nonconvex $\mathcal{R}$, maximizing a sum of logarithms over $\mathcal{R}$ is a nonconvex problem. Thus, in general, computing a

see the program

point in $\mathcal{R}^*$ is practically feasible only if the number of users is small [9].

A. Notation and Preliminaries

Let $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the set of nonnegative and positive real numbers, respectively. Given a vector $\boldsymbol{r} \in \mathbb{R}^K$ and a set $\mathcal{R} \subseteq \mathbb{R}^K$, the set $\boldsymbol{r} + \mathcal{R}$ is defined as follows:

$$\boldsymbol{r} + \mathcal{R} = \{\boldsymbol{r} + \boldsymbol{r}', \boldsymbol{r}' \in \mathcal{R}\}.$$

Order relations $\geq, >$, and $\gg$ between two vectors $\boldsymbol{r}, \boldsymbol{r}' \in \mathbb{R}^K$ are defined as follows:

$$\begin{aligned} \boldsymbol{r}' \geq \boldsymbol{r} &\Leftrightarrow \boldsymbol{r}' \in \boldsymbol{r} + \mathbb{R}_+^K, \\ \boldsymbol{r}' > \boldsymbol{r} &\Leftrightarrow \boldsymbol{r}' \in \boldsymbol{r} + (\mathbb{R}_+^K \setminus \{\boldsymbol{0}\}), \\ \boldsymbol{r}' \gg \boldsymbol{r} &\Leftrightarrow \boldsymbol{r}' \in \boldsymbol{r} + \mathbb{R}_{++}^K. \end{aligned}$$

Given a set $\mathcal{R} \subset \mathbb{R}^K$, the Pareto boundary of $\mathcal{R}$ is given by

$$\mathcal{P} = \left\{\boldsymbol{r} \in \mathcal{R} : \mathcal{R} \cap \boldsymbol{r} + \mathbb{R}_+^K = \{\boldsymbol{r}\}\right\}.$$

A function $f : \mathcal{D} \subseteq \mathbb{R}^K \to \mathbb{R}$ is strictly increasing if

$$\boldsymbol{r}, \boldsymbol{r}' \in \mathcal{D}, \boldsymbol{r}' > \boldsymbol{r} \Rightarrow f(\boldsymbol{r}') > f(\boldsymbol{r}).$$

A set $\mathcal{R} \subseteq \mathbb{R}_+^K$ is normal if

$$\boldsymbol{r}' \in \mathcal{R}, \boldsymbol{r} \in \mathbb{R}_+^K, \boldsymbol{r}' \geq \boldsymbol{r} \Rightarrow \boldsymbol{r} \in \mathcal{R}.$$

Note that this property is also known as comprehensive [8] or coordinate-convex [7].

The closure of a set $\mathcal{S}$ is denoted by $\mathrm{cl}(\mathcal{S})$, and ∇f denotes the gradient of a function f. Finally, $\boldsymbol{I}$ is the identity matrix, and $\boldsymbol{1}$ the vector of all ones.

II. An Example

In this section, we show with an example that there exist compact and nontrivial rate regions for which no proportionally fair rate vector exists.

Define a function $f : \mathbb{R}_{++}^K \to \mathbb{R}$ as follows:

$$f(\boldsymbol{r}) = \sum_{k=1}^{K} \ln r_k. \tag{3}$$

Note that f is strictly increasing, and

$$\nabla f(\boldsymbol{r}) \in \mathbb{R}_{++}^K, \forall \boldsymbol{r} \in \mathbb{R}_{++}^K. \tag{4}$$

Moreover,

$$\nabla f(\boldsymbol{r}^*)^{\mathrm{T}}(\boldsymbol{r} - \boldsymbol{r}^*) = \sum_{k=1}^{K} \frac{r_k - r_k^*}{r_k^*}. \tag{5}$$

Define the following level sets:

$$\underline{\mathcal{S}}_f(\gamma) = \left\{\boldsymbol{r} \in \mathbb{R}_{++}^K : f(\boldsymbol{r}) \leq \gamma\right\}, \tag{6}$$

$$\overline{\mathcal{S}}_f(\gamma) = \left\{\boldsymbol{r} \in \mathbb{R}_{++}^K : f(\boldsymbol{r}) \geq \gamma\right\}. \tag{7}$$

Due to the fact that f is strictly concave and smooth on $\mathbb{R}_{++}^K$,

$$\nabla f(\boldsymbol{r}^*)^{\mathrm{T}}(\boldsymbol{r} - \boldsymbol{r}^*) > 0, \forall \boldsymbol{r} \in \overline{\mathcal{S}}_f(f(\boldsymbol{r}^*)) \setminus \{\boldsymbol{r}^*\}. \tag{8}$$

Now let $K = 2$ and define a rate region $\mathcal{R}$ as follows:

$$\mathcal{R} = \left\{\boldsymbol{r} \in \mathbb{R}_+^2 : r_k \leq \exp(2)\right\} \cap \mathrm{cl}(\underline{\mathcal{S}}_f(3)). \tag{9}$$

The set $\mathcal{R}$ is compact and nontrivial. Only points on the Pareto boundary of $\mathcal{R}$ are candidates for PF points, as for all other points in $\mathcal{R}$ the rate of at least one user can be increased without decreasing the rate of any other user, leading to a positive aggregate proportional change. The Pareto boundary of $\mathcal{R}$ is given by

$$\mathcal{P} = \mathcal{R} \cap \overline{\mathcal{S}}_f(3) = \{(\exp(t+1), \exp(2-t)) : t \in [0,1]\}. \tag{10}$$

Accordingly, any point $\boldsymbol{r}^* \in \mathcal{R}$ with $f(\boldsymbol{r}^*) = 3$ is a candidate for PF. But according to (5) and (8), for any such point there exists another point $\boldsymbol{r} \in \mathcal{P}$ such that

$$\nabla f(\boldsymbol{r}^*)^{\mathrm{T}}(\boldsymbol{r} - \boldsymbol{r}^*) = \sum_{k=1}^{2} \frac{r_k - r_k^*}{r_k^*} > 0. \tag{11}$$

Consequently, no point in $\mathcal{R}$ fulfills (1) – in other words, there exists no proportionally fair rate vector. In contrast, note that any point in $\mathcal{P}$ is a global maximizer of (2), i.e., $\mathcal{R}^* = \mathcal{P}$.

Notably, in [8] it was shown that the classical Nash Bargaining Solution (NBS) [10], which was originally defined for convex utility sets only, can be extended to so-called strictly log-convex utility sets. By slightly modifying the above example, however, it can be shown that there exist strictly log-convex sets that do not contain a proportionally fair point. In particular, for such sets the NBS point that results if the disagreement point is the zero vector is not proportionally fair. This result constitutes another example for the difference between maximizing a sum of logarithms (or, equivalently, the Nash product with disagreement point $\boldsymbol{0}$) and proportional fairness on the family of nonconvex compact sets.

III. Local Proportional Fairness

As shown in the previous section, for general $\mathcal{R}$, there may not exist a rate vector that is proportionally fair. The idea underlying local proportional fairness is to approximate $\mathcal{R}$ at a candidate point $\boldsymbol{r}^*$ by a set $\hat{\mathcal{R}}(\boldsymbol{r}^*)$. The approximation of $\mathcal{R}$ at $\boldsymbol{r}^*$ is defined as follows:

$$\hat{\mathcal{R}}(\boldsymbol{r}^*) = \{\boldsymbol{r}^* + \boldsymbol{v}, \boldsymbol{v} \in \mathcal{T}(\boldsymbol{r}^*)\} \cap \mathbb{R}_+^K, \tag{12}$$

where $\mathcal{T}(\boldsymbol{r}^*)$ denotes the tangent cone of $\mathcal{R}$ at $\boldsymbol{r}^*$ [11]. The tangent cone $\mathcal{T}(\boldsymbol{r}^*)$ provides information about the local structure of $\mathcal{R}$ near $\boldsymbol{r}^*$.

Our definition of *local proportional fairness* is as follows: A point $\boldsymbol{r}^* \in \mathcal{R} \cap \mathbb{R}_{++}^K$ is locally proportional fair in $\mathcal{R}$ if for any other vector $\boldsymbol{r} \in \hat{\mathcal{R}}(\boldsymbol{r}^*)$, the aggregate of proportional changes is nonpositive:

$$\sum_{k=1}^{K} \frac{r_k - r_k^*}{r_k^*} \leq 0, \forall \boldsymbol{r} \in \hat{\mathcal{R}}(\boldsymbol{r}^*). \tag{13}$$

Obviously, $\boldsymbol{r}^*$ is LPF in $\mathcal{R}$ if it is proportionally fair in $\hat{\mathcal{R}}(\boldsymbol{r}^*)$.

see the program

Again, LPF is closely related to the maximization of a sum of logarithms: If f is chosen as in (3), then (13) is equivalent to

$$\nabla f(\boldsymbol{r}^*)^{\mathrm{T}}\boldsymbol{v} \leq 0, \forall \boldsymbol{v} \in \mathcal{T}(\boldsymbol{r}^*). \tag{14}$$

Eq. (14) represents a necessary condition for $\boldsymbol{r}^*$ being a local maximizer of f over $\mathcal{R} \cap \mathbb{R}^K_{++}$ [11]. It is important to note, however, that Eq. (14) is only a necessary, not a sufficient condition – as a consequence, any local maximizer of f over $\mathcal{R} \cap \mathbb{R}^K_{++}$ is LPF, but an LPF point is not necessarily a local maximizer. Let $\mathcal{R}^{\mathrm{LPF}}$ denote the set of LPF points in $\mathcal{R}$. From the perspective of maximizing a sum of logarithms, the set $\mathcal{R}^{\mathrm{LPF}}$ can be partitioned into three sets: the set of points that satisfy (14), but are not local maximizers, the set of local maximizers that are not global maximizers, and the set of global maximizers. If a sum of logarithms is to be maximized, there is an obvious hierarchy among these sets. In contrast, from the perspective of aggregate proportional changes over an approximated rate region, the three sets are equivalent.

According to (14), every local maximizer of f over $\mathcal{R}\cap\mathbb{R}^K_{++}$ is LPF. As $\mathcal{R}^*$ contains the global maximizers,

$$\mathcal{R}^* \subseteq \mathcal{R}^{\mathrm{LPF}}. \tag{15}$$

Moreover, $\mathcal{R}^*$ is nonempty for compact and nontrivial $\mathcal{R}$, thus $\mathcal{R}^{\mathrm{LPF}}$ is nonempty. This proves the existence of an LPF point: *If $\mathcal{R}$ is compact and nontrivial, there always exists a feasible rate vector that is LPF.*

Let $\mathcal{P}$ denote the Pareto boundary of $\mathcal{R}$. From the strict monotonicity of f defined in (3), it follows that $\mathcal{R}^* \subseteq \mathcal{P}$ [12]. Under the assumption that $\mathcal{R}$ is normal, the same holds for the set of LPF points:

$$\mathcal{R}^{\mathrm{LPF}} \subseteq \mathcal{P}.$$

In other words, *if $\mathcal{R}$ is compact, nontrivial, and normal, all LPF points are Pareto optimal*. The proof is by contradiction: Let $\boldsymbol{r}^*$ in $\mathcal{R}^{\mathrm{LPF}}$. Moreover, assume that $\boldsymbol{r}^*$ is not Pareto optimal. Accordingly, there exists $\boldsymbol{r}' \in \mathcal{R}$ such that $\boldsymbol{r}' > \boldsymbol{r}^*$. Thus, $\boldsymbol{v} = \boldsymbol{r}' - \boldsymbol{r}^* > \boldsymbol{0}$. Due to the fact that $\mathcal{R}$ is normal,

$$\boldsymbol{r}^* + \alpha\boldsymbol{v} \in \mathcal{R}, \forall \alpha \in [0,1].$$

Consequently, $\boldsymbol{v} \in \mathcal{T}(\boldsymbol{r}^*)$. But $\nabla f(\boldsymbol{r}^*) \in \mathbb{R}^K_{++}$ and $\boldsymbol{v} \in \mathbb{R}^K_{+}$, thus

$$\nabla f(\boldsymbol{r}^*)^{\mathrm{T}}\boldsymbol{v} > 0,$$

which contradicts the assumption that $\boldsymbol{r}^*$ in $\mathcal{R}^{\mathrm{LPF}}$.

If $\mathcal{R}$ is convex, (14) is equivalent to [11]

$$\nabla f(\boldsymbol{r}^*)^{\mathrm{T}}(\boldsymbol{r} - \boldsymbol{r}^*) \leq 0, \forall \boldsymbol{r} \in \mathcal{R}. \tag{16}$$

By choosing f as in (3), (16) is identical to (1). As a result, *for convex $\mathcal{R}$, LPF and PF are equivalent.*

IV. LPF in a MU-MISO Downlink

In this section, an achievable rate region in a K-user multiple-input, single-output (MISO) downlink with linear precoding is considered. The transmitter has N antennas, while each of the K receivers has a single antenna. Interference is treated as noise. The transmitter is subject to a power constraint. Using the duality relation between uplink and downlink [13], a parameterization of the rate region $\mathcal{R}$ can be given as follows:

$$\mathcal{R} = \left\{\boldsymbol{r}(\boldsymbol{p}) : \boldsymbol{p} \in \mathbb{R}^K_{+}, \|\boldsymbol{p}\|_1 \leq P\right\}, \tag{17}$$

with the uplink rates

$$r_k(\boldsymbol{p}) = \log_2 \det\left(\boldsymbol{I} + \Big(\sigma^2\boldsymbol{I} + \sum_{q\neq k}\boldsymbol{h}_q\boldsymbol{h}_q^{\mathrm{H}}p_q\Big)^{-1}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{H}}p_k\right). \tag{18}$$

We assume that $\mathcal{R}$ is nontrivial.

Due to the nonconvexity of the function $f(\boldsymbol{r}(\boldsymbol{p}))$, determining an element of $\mathcal{R}^*$ represents a hard problem. As illustrated in [2], the rate region $\mathcal{R}$ may be nonconvex. Accordingly, the elements of $\mathcal{R}^*$ correspond to global optima of a nonconvex problem. As shown in [3], an element of $\mathcal{R}^*$ can be found by using methods from global optimization, but such a solution is practically feasible for small K only [9]. In contrast, computing an LPF point is simple, as shown in the following.

The rate region $\mathcal{R}$ is normal. Thus, all LPF points lie on the Pareto boundary of $\mathcal{R}$. The Pareto boundary corresponds to the rate vectors where full transmission power is used [13]. Define a parameterization $\boldsymbol{\phi}$ as follows:

$$\boldsymbol{\phi}(\boldsymbol{\mu}) = \boldsymbol{r}(\boldsymbol{p}^0 + \boldsymbol{Q}\boldsymbol{\mu}),$$

where $\boldsymbol{p}^0 = \frac{P}{K}\boldsymbol{1}$ and the columns of $\boldsymbol{Q}$ constitute a basis of the nullspace of $\boldsymbol{1}^{\mathrm{T}}$, i.e., $\boldsymbol{1}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{0}$. A parameterization of the Pareto boundary is then given by

$$\mathcal{P} = \left\{\boldsymbol{\phi}(\boldsymbol{\mu}) : \boldsymbol{\mu} \in \mathbb{R}^{K-1}, \boldsymbol{p}^0 + \boldsymbol{Q}\boldsymbol{\mu} \geq \boldsymbol{0}\right\}. \tag{19}$$

Let $\boldsymbol{J}_\phi(\boldsymbol{\mu})$ denote the Jacobian of $\boldsymbol{\phi}$ at $\boldsymbol{\mu}$. The parameterization $\boldsymbol{\phi}$ is differentiable, thus the tangent space at $\boldsymbol{r}^* \in \mathcal{P}\cap\mathbb{R}^K_{++}$ is given by $\mathrm{span}\boldsymbol{J}_\phi(\boldsymbol{\phi}^{-1}(\boldsymbol{r}^*))$, and the tangent cone at $\boldsymbol{r}^*$ corresponds to the half-space

$$\mathcal{T}(\boldsymbol{r}^*) = \left\{\boldsymbol{v} \in \mathbb{R}^K : \boldsymbol{n}^{\mathrm{T}}\boldsymbol{v} \leq 0\right\}, \tag{20}$$

where $\boldsymbol{n} \in \mathbb{R}^K_{+}$ is a nonnegative solution of

$$\boldsymbol{J}_\phi(\boldsymbol{\phi}^{-1}(\boldsymbol{r}^*))^{\mathrm{T}}\boldsymbol{n} = \boldsymbol{0}.$$

Due to the fact that $\mathcal{R}$ is normal, such a solution exists.

From (14), (20) and $\nabla f(\boldsymbol{r}) \in \mathbb{R}^K_{++}$, it follows that $\boldsymbol{r}^* \in \mathcal{P} \cap \mathbb{R}^K_{++}$ is LPF if $\nabla f(\boldsymbol{r}^*) = \alpha\boldsymbol{n}$ for some $\alpha \in \mathbb{R}_{++}$, or, equivalently,

$$\boldsymbol{J}_\phi(\boldsymbol{\phi}^{-1}(\boldsymbol{r}^*))^{\mathrm{T}}\nabla f(\boldsymbol{r}^*) = \boldsymbol{0}. \tag{21}$$

Define a function $g : \mathbb{R}^{K-1} \to \mathbb{R}$,

$$g(\boldsymbol{\mu}) = f(\boldsymbol{\phi}(\boldsymbol{\mu})), \tag{22}$$

see the program

Note that

$$\nabla g(\boldsymbol{\mu}) = \boldsymbol{J}_{\phi}(\boldsymbol{\mu})^{\mathrm{T}} \nabla f(\boldsymbol{\phi}(\boldsymbol{\mu})).$$

With Eq. (21), it follows that any $\boldsymbol{\mu}^*$ satisfying

$$\nabla g(\boldsymbol{\mu}^*) = \mathbf{0} \tag{23}$$

yields an LPF point

$$\boldsymbol{r}^* = \boldsymbol{\phi}(\boldsymbol{\mu}^*).$$

Accordingly, any of the standard local methods for unconstrained optimization (gradient ascent, (quasi-)Newton) can be used to find an LPF point.

V. Conclusions

For convex rate regions, the concepts of proportional fairness and maximizing a sum of logarithms are equivalent. This equivalence does not hold for nonconvex rate regions. In particular, for nonconvex rate regions, there may not exist a proportionally fair rate vector. We proposed the concept of local proportional fairness (LPF), which, similar to the original definition of proportional fairness, is based on the aggregate of proportional changes. If the rate region is compact, nontrivial, and normal, an LPF solution always exists, and all LPF points are Pareto optimal. A significant advantage of LPF is that it represents a local property. In the MISO downlink with linear precoding, an LPF point can be found by standard local methods – as a result, finding an LPF point is a practically solvable problem in this setup. In contrast, finding a maximizer of a sum of logarithms over a nonconvex rate region is a global optimization problem that is practically solvable only if the number of users is small.

References

[1] F. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, January 1997.

[2] H. Boche and M. Schubert, "Analysis of different precoding/decoding strategies for multiuser beamforming," *Proc. IEEE Vehicular Technology Conference*, vol. 1, pp. 39–43, April 2003.

[3] J. Brehmer and W. Utschick, "Utility maximization in the multi-user MISO downlink with linear precoding," in *Proc. IEEE International Conference on Communications (ICC)*, June 2009.

[4] C. Touati, H. Kameda, and A. Inoie, "Fairness in non-convex systems," University of Tsukuba, Ibaraki, Japan, Tech. Rep. CS-TR-05-4, September 2005.

[5] T. Bonald and A. Proutiere, "Flow-level stability of utility-based allocations for non-convex rate regions," in *Proc. Conference on Information Sciences and Systems (CISS)*, March 2006, pp. 327–332.

[6] H. Boche and M. Schubert, "On the existence of a proportionally fair operating point for wireless communication systems," in *Proc. IEEE 8th Workshop on Signal Processing Advances in Wireless Communications SPAWC 2007*, 17–20 June 2007, pp. 1–5.

[7] J. Liu, A. Proutière, Y. Yi, M. Chiang, and H. V. Poor, "Flow-level stability of data networks with non-convex and time-varying rate regions," in *Proc. ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*. New York, NY, USA: ACM, 2007, pp. 239–250.

[8] M. Schubert and H. Boche, "Nash bargaining and proportional fairness for log-convex utility sets," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2008, pp. 3157–3160.

[9] S. A. Vavasis, "Complexity issues in global optimization: a survey," in *Handbook of Global Optimization*. Kluwer, 1995, pp. 27–41.

[10] J. Nash, "The bargaining problem," *Econometrica*, vol. 18, pp. 155–162, 1950.

[11] D. Bertsekas, A. Nedic, and A. Ozdaglar, *Convex analysis and optimization*. Athena Scientific, 2003.

[12] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.

[13] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, January 2004.

DFT-based vs. Cooperative MET-based MU-MIMO in the Downlink of Cellular OFDM Systems

L. Thiele, M. Schellmann, T. Wirth and V. Jungnickel
Fraunhofer Institute for Telecommunications
Heinrich-Hertz-Institut
Einsteinufer 37, 10587 Berlin, Germany
{thiele, schellmann, thomas.wirth, jungnickel}@hhi.fraunhofer.de

F. Boccardi and H. Huang
Bell Labs,
Alcatel-lucent
{fb,hchuang}@alcatel-lucent.com

***Abstract*—In this work, we compare different pre-coding techniques and their performance in a cellular MIMO OFDM downlink. The first target is to determine the achievable additional beamforming gain by using a near-optium pre-coder instead of DFT-based pre-coding. In the second step, we include multi-user selection diversity, which turned out to be advantageous especially for downlink cooperation. An extended score-based scheduler, which is known to asymptotically reach proportional fairness, is used for the user grouping and resource allocation in the cellular downlink. Further, we use minimum mean square error equalization at the terminal side to combat residual cochannel interference.**

I. Introduction

To enable ubiquitous broadband wireless access in future cellular systems supporting multiple-input multiple-output (MIMO) technology, transmission must be made robust against multi-cell interference. Recently, it was shown that the capacity scaling law, known from an isolated cell, also holds for the interference limited case of a multi-cellular radio system [1] with $N_T = N_R$ transmit and receive antennas. This work mainly focused on the optimization at the receiver side. However, cochannel interference (CCI) is still the dominant source of performance degradation in the cellular network, especially if $N_T > N_R$. Removing CCI may lead to a additional performance gain.

In this work we compare different pre-coding techniques and their performance in a cellular MIMO orthogonal frequency division multiplexing (OFDM) downlink. As a baseline concept, we consider independent DFT-based pre-coding and multi-user MIMO (MU-MIMO) service in each sector of the system. As a first target, we determine the available additional beamforming gain by using a near-optium pre-coder instead. Therefore, we employ multi-user eigenmode transmission (MET), known to realize 90% of the dirty paper coding (DPC) capacity in an isolated cell context [2], based on the dominant eigenmodes of the mobile terminals (MTs) in the serving area.
A second target of this work is to realize limited localized cooperative transmission in a multi-cellular network. Recent results obtained for a cellular MIMO OFDM downlink with a sparse user distribution per sector, show potential performance gains for cooperation [3]. In this context it turned out, that downlink cooperation based on DFT beams is unsuitable and MET should get the precedence. Additionally to preceding work, we include multi-user selection diversity in the user grouping process. Multi-user diversity from a large number of users per cell may help to increase the system performance further. From [2] it is well known that this would require a brute force or at least a greedy search over the user space to choose appropriate users for the active set $\mathcal{M}$. To combat residual CCI from surrounding cells we employ MMSE equalization, also known as optimum combining (OC) [4], at the terminal side.

II. Downlink system model

The downlink MIMO-OFDM transmission system with N_T transmit and N_R receive antennas per MT is described on each subcarrier by

$$\mathbf{y} = \mathbf{HCx} + \mathbf{n}\,, \tag{1}$$

where $\mathbf{H}$ is the $N_R \times N_T$ channel matrix and $\mathbf{C}$ the unitary $N_T \times N_T$ pre-coding matrix; $\mathbf{x}$ denotes the $N_T \times 1$ vector of transmit symbols; $\mathbf{y}$ and $\mathbf{n}$ denote the $N_R \times 1$ vectors of the received signals and of the additive white Gaussian noise (AWGN) samples, respectively, with covariance $E\{\mathbf{nn}^H\} = \sigma^2\mathbf{I}$.

Assume that a group of α cooperating base station (BS) sectors provides a beam set $\mathbf{C}_i$. The beam set contains αN_T pre-coding beams $\mathbf{b}_{i,u}$ with $u \in \{1, ..., \alpha N_T\}$. In the following we denote $\mathbf{b}_{i,u}$ as the u-th pre-coding vector provided by the i-th cell cluster. The received downlink signal $\mathbf{y}^m$ at the MT m in the cellular environment is given by

$$\begin{aligned}\mathbf{y}^m = \; & \underbrace{\mathbf{H}_i^m \mathbf{b}_{i,u}}_{\bar{\mathbf{h}}_{i,u}} x_{i,u} + \underbrace{\sum_{\substack{j=1\\j\neq u}}^{\alpha N_T} \mathbf{H}_i^m \mathbf{b}_{i,j} x_{i,j}}_{\zeta_{i,u}} \\ & + \underbrace{\sum_{\substack{\forall l\\l\neq i}} \sum_{j=1}^{N_T} \mathbf{H}_l^m \mathbf{b}_{l,j} x_{l,j} + \mathbf{n}}_{\mathbf{z}_{i,u}}\,,\end{aligned} \tag{2}$$

The desired data stream $x_{i,u}$ transmitted on the u-th beam from the i-th cluster is distorted by the intra-cluster and inter-cluster interference aggregated in $\zeta_{i,u}$ and $\mathbf{z}_{i,u}$, respectively. $\mathbf{H}_i^m$ spans the $N_R \times \alpha N_T$ channel matrix for user m formed by the cluster i. Thus, $\zeta_{i,u}$ denotes the interference generated in

see the program

the cooperation area. In the scope of this paper, it is assumed that all αN_T beams in the beam set $\mathbf{C}_i$ are simultaneously active, whereby the total available power p_i is assumed to be uniformly distributed over the αN_T beams. Thus, $E\left\{|x_{i,j}|^2\right\} = p_i/(\alpha N_T)$ holds, and $p_i = \sum_{j=1}^{\alpha N_T} \mathrm{E}\left\{|x_{i,j}|^2\right\} = \alpha p_s$ with p_s being the transmit power per sector.

A. Determine serving BS or cooperative BS cluster

The general assumption for single-cell operation is that each MT is assigned to the BS sector yielding the highest receive power over the entire frequency band, which is denoted as top-1 signal. Thus, the BS assignment is based on broadband power conditions, and a fast cell handover is assumed. For downlink cooperation, we extend this scheme by evaluating the top-α strongest signals and grouping the users selecting the same set of α BSs for joint signal transmission.

By subdividing the signal bandwidth into single sub-bands confined to a fixed number of consecutive subcarriers, we define sets of contiguous transmission resources, which are denoted as resource blocks (RBs) in the following. Each RB is processed independently, and thus the sectors to form a group for cooperation can be selected individually per RB.

III. Downlink pre-coding

A. Baseline: DFT-based fixed pre-coding

As a baseline, we consider a system concept from [1], where all sectors operate independently, while the inter-cell interference is accounted for at the multi-antenna MTs only. The following evaluation is carried out for each RB independently (refer to Fig. 1): Each BS provides a fixed matrix $\mathbf{C}$ consisting of unitary DFT beams. Assuming that the inter-cell interference is completely known at the MTs, the MTs evaluate the achievable rate per beam and convey this information to their BS. At the BS, the feedback from the different MTs is collected, and the DFT beams from matrix $\mathbf{C}$ are assigned individually to the MTs. This simple approach has the convenient property that with the fixed beam set $\mathbf{C}$ used for all BSs, the CCI, i.e. $\zeta_{i,u} + \mathbf{z}_{i,u}$, becomes fully predictable, enabling interference-aware scheduling in a cellular system.

In combination with fair, interference-aware scheduling policies, it has been shown that users profit from almost doubled spectral efficiencies in the MIMO 2×2 system, as compared to the single-input single-output (SISO) setup [1].

B. Downlink cooperation: MET-based pre-coding

There are several concepts for cooperative downlink transmission, all imposing different demands on the system architecture. As a basic requirement, coherent downlink transmission is mandatory. Thus, downlink transmission from all BSs has to be synchronized with respect to the carrier frequency and the frame start. A basic concept to achieve this kind of synchronization has been presented in [5].

As a reference design for downlink cooperation in cellular systems, we use an approach based on MET, which is known to achieve a near-optimum pre-coding performance for an isolated cell cluster [2]. For MET, the MTs report the $\lambda \leq N_R$

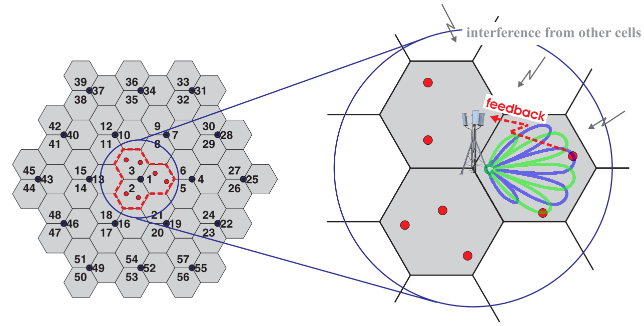


Fig. 1. System concept assuming multiple antennas at the base station for the purpose unitary fixed DFT-based pre-coded beamforming. SINR feedback is provided by the terminal for possible transmission modes using a narrow band feedback channel.

dominant eigenmodes of their channel to the cell cluster together with the corresponding eigenvalues. MET supports to simultaneously transmit up to αN_T data streams via unitary precoding beams, while up to N_R beams may be assigned to a single MT. However, it has been indicated in [2] that MU-MIMO service for distinct terminals using MET is more efficient than time-multiplexing multi-stream transmission to a single user. Thus, for our investigations, we let all MTs report their dominant eigenmode only ($\lambda = 1$), which keeps the required amount of feedback per user limited.

Determination of the MET-based pre-coding beams in matrix $\mathbf{C}$ is briefly sketched as follows: Consider a fixed set $\mathcal{M}$ of users, which should be served in a RB. Each user m decomposes its $N_R \times \alpha N_T$ channel matrix $\mathbf{H}_i^m$ according to the singular value decomposition (SVD), yielding $\mathbf{H}_i^m = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{V}_i^H$. The dominant eigenmode is the first column vector of matrix $\mathbf{V}_i$, denoted as $\mathbf{v}_{i,1}$. Together with the dominant eigenvalue in $\mathbf{\Sigma}_i$, denoted as $\Sigma_{i,1}$, we obtain the user's effective eigenmode channel $\mathbf{\Gamma}_m = \Sigma_{i,1} \mathbf{v}_{i,1}^H$. This measure needs to be fed back from each MT to the cell cluster.

User orthogonalization at the BS: To obtain the pre-coding vector for the m-th user, the BS cluster aggregates the interfering eigenmodes $\mathbf{\Gamma}_n$ with $n \in \{1, ..., (m-1), (m+1), ..., \alpha N_T\}$ from the other terminals in the active set $\mathcal{M}$, yielding a matrix of dimension $(\alpha N_T - 1) \times \alpha N_T$

$$\tilde{\mathbf{\Gamma}}_m = [\mathbf{\Gamma}_1^H \ldots \mathbf{\Gamma}_{m-1}^H \, \mathbf{\Gamma}_{m+1}^H \ldots \mathbf{\Gamma}_{\alpha N_T}^H]^H \tag{3}$$

Performing the SVD of $\tilde{\mathbf{\Gamma}}_m$ yields

$$\tilde{\mathbf{\Gamma}}_m = \tilde{\mathbf{U}}_m \left[\begin{array}{cc} \tilde{\mathbf{\Sigma}}_m & \mathbf{0} \end{array} \right] \left[\begin{array}{cc} \tilde{\mathbf{V}}_m^1 & \tilde{\mathbf{v}}_m^0 \end{array} \right]^H, \tag{4}$$

where $\tilde{\mathbf{v}}_m^0$ corresponds to the eigenvector associated with the null space of $\tilde{\mathbf{\Gamma}}_m$. Note, in principle the null space is represented by a matrix of dimensions $\lambda \times \alpha N_T$. Since we limit each user to be served on its dominant eigenmode only, i.e. $\lambda = 1$, $\tilde{\mathbf{v}}_m^0$ is of dimension $1 \times \alpha N_T$ and thus a vector.

This vector is used for pre-coded transmission to user m, which ensures that all other users in $\mathcal{M}$ do not experience any interference from this beam under ideal conditions. Note that the block-diagonalization constraint $\alpha N_T \geq \sum_{\forall m \in \mathcal{M}} N_R(m)$

is relaxed by the use of dominant eigenmodes, resulting in $\alpha N_T \geq |\mathcal{M}|$ instead [2].

The selected pre-coding matrix on a RB and time slot is given by

$$\mathbf{C} = [\tilde{\mathbf{v}}_1^0 \ldots \tilde{\mathbf{v}}_m^0 \ldots \tilde{\mathbf{v}}_{\alpha N_T}^0] \,, \tag{5}$$

where $\mathrm{tr}\left[\mathbf{C}\mathbf{C}^H\right] = p_i$. As the beamforming vectors are unitary, $\mathbf{C}$ implicitly includes the constraint of equal transmit power per beam and sum power per cell cluster.[1]

IV. Linear MMSE receiver

Assuming a linear equalizer $\mathbf{w}_u$ at the MTs, which is required to extract the useful signal $x_{i,u}$ from $\mathbf{y}^m$ according to (2), yields a post-equalization signal to interference and noise ratio (SINR) at the MT for stream $x_{i,u}$ given by

$$\mathrm{SINR}_u = p_i \frac{\mathbf{w}_u^H \overline{\mathbf{h}}_{i,u} \overline{\mathbf{h}}_{i,u}^H \mathbf{w}_u}{\mathbf{w}_u^H \mathbf{Z}_u \mathbf{w}_u} \,, \tag{6}$$

where $\mathbf{Z}_u$ is the covariance matrix of the interfering signals aggregated in $\zeta_{i,u}$ and $\mathbf{z}_{i,u}$, i.e. $\mathbf{Z}_u = \mathrm{E}\left[(\zeta_{i,u} + \mathbf{z}_{i,u})(\zeta_{i,u} + \mathbf{z}_{i,u})^H\right]$, with $E[.]$ being the expectation operator.

In this work, we consider minimum mean square error (MMSE) equalization for the purpose of inter-cell interference suppression at the receiver side according to

$$\mathbf{w}_m^{\mathrm{MMSE}} = \frac{p_i \mathbf{R}_{yy}^{-1} \mathbf{H}_i^m \mathbf{b}_{i,u}}{\alpha N_T} \,, \tag{7}$$

where $\mathbf{R}_{yy}$ denotes the covariance matrix of $\mathbf{y}^m$ from (2), i.e. $\mathbf{R}_{yy} = \mathrm{E}\left[\mathbf{y}^m (\mathbf{y}^m)^H\right]$. This receiver yields a post-equalization SINR on a given RB for user m

$$\mathrm{SINR}_m = \frac{p_i}{\alpha N_T} \left[\mathbf{H}_i^m \mathbf{b}_{i,u}\right]^H \mathbf{Z}_u^{-1} \mathbf{H}_i^m \mathbf{b}_{i,u} \tag{8}$$

In a practical context, the covariance matrix

$$\mathbf{R}_{yy} = \mathbf{Z}_u + \mathbf{H}_i^m \mathbf{b}_{i,u} \left(\mathbf{H}_i^m \mathbf{b}_{i,u}\right)^H \tag{9}$$

may be obtained by using multi-cell channel estimates based on common and dedicated reference signals [6]. Thus, each terminal is able to combat residual CCI from surrounding cells aggregated in $\mathbf{z}_{i,u}$ from (2).

V. Resource allocation and fair user selection

Resource allocation is conducted based on a score-based scheduling process at a central processing unit (CPU) controlling the cell cluster of size α. For the baseline system, each MT firstly evaluates the achievable SINR conditions per fixed DFT-beam according to (6) based on the current channel conditions in each RB. After conveying this information to the BS[2] , the collected per-beam rates from any user are ranked over all available RBs according to their quality, and corresponding scores are assigned (score-based scheduling policy as proposed in [7]). Resource assignment is then done for each RB individually by assigning each beam to the user providing minimum score for that beam.

For the MET-based pre-coding, each MTs determines the effective eigenmode channel $\mathbf{\Gamma}_m$ per RB and conveys this information to the CPU. At the CPU, an equivalent score-based scheduling process is carried out for each of the αN_T spatial layers that are available for transmission in a RB. For the first layer, only the effective eigenmode channel $\mathbf{\Gamma}_m$ per user is considered, whereas for all successive layers the correlation of the effective channels from all possible user pairings is taken into account. The applied algorithm is closely related to the greedy user selection approach from [2], but has been modified to be applicable with the score-based scheduling technique. Thus, a direct comparison with the baseline system presented in [1] is enabled. After the proper user selection together with the corresponding precoding beams in (5) have been found per RB, the SINR according to (6) can be determined, yielding the achievable performance in terms of the spectral efficiency of the system.

Note that for both precoding concepts, each user is served on a single spatial stream in a RB only, i.e. only pure MU-MIMO access is considered without allowing spatial multiplexing transmission to any of the multi-antenna MTs. This constraint is motivated by findings from [2], [8], [9], where the authors emphasize that MU-MIMO transmission mode is preferential as long as the number of users K is sufficiently large.

A. Note on the applied scheduler

The objective of the score-based resource allocation process is to assign each user his best RBs and therein his favoured precoding beams, thus guaranteeing a certain amount of fairness for users during scheduling. Clearly, the process is of heuristic nature, and hence the global scheduling target of assigning each user an equal amount of resources is achieved on average only or if the number of available resources tends to infinity. However, its convenient property for practical applications is its flexible utilization, as the set of resources can be defined over arbitrary dimensions (time/frequency/space). Thus, fairness can be established on a small time scale, e.g. even for the scheduling of resources contained within a single OFDM symbol.

VI. Simulation environment

The performance is investigated in a triple-sectored hexagonal cellular network with 19 BSs in total. The extended spatial channel model (SCME) with urban macro scenario parameters is used [10] yielding an user's geometry for the center cell, refer to Fig. 2 (right), which is equivalent to [11]. The basic system settings for our simulations are summarized in Table I. For the evaluation of cooperative transmission strategies, we have to take users from surrounding cells into account. For a realistic performance evaluation of the cellular network, we employ a wrap-around, which ensures that the interference scenario follows independent identically distributed (i.i.d.) statistics for all users.

[1]Note that we are not considering any optimal power allocation scheme here.

[2]As for the baseline system $\alpha = 1$ holds, the CPU may be located directly at the BS.

see the program

TABLE I
SIMULATION ASSUMPTIONS.

parameter	value
channel model	3GPP SCME
type	Monte Carlo
scenario	urban-macro
additional modifications	LOS-NLOS propagation mix
traffic model	full buffer
f_c	2 GHz
frequency reuse	1
signal bandwidth	18 MHz, 100 RBs
intersite distance	500m
number of BSs	19 having 3 sectors each
N_T ; spacing	1,2 ; 4λ
transmit power	46 dBm
sectorization	triple, with FWHM of $68°$
BS height	32m
N_T ; spacing	1,2 ; $\lambda/2$
MT height	2m

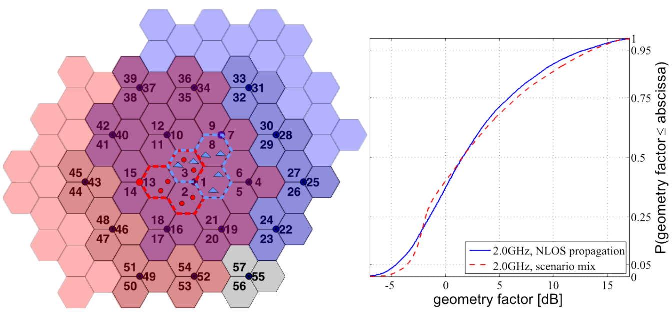


Fig. 2. Left: Triple-sectored cellular setup. The region bounded by the red and blue dashed line indicates a cooperative cell cluster jointly serving their terminals on different RBs. Right: User geometries obtained from the center cell with id={1,2,3} and parameters according to Table I.

As an initial step, we place up to 20 terminals per sector inside the inner cell with cell ID={1,2,3}. Since evaluation in this work is limited to linear pre-coded beams, up to αN_T user may be served in the same RB. N_T corresponds to the available spatial dimensions per sector antenna array. If the number of terminals per sector exceeds N_T, interference-aware score-based scheduling is applied, which increases the system throughput further by exploiting multi-user diversity [1].

The users in the cooperating cell of a cluster are generated by dropping the users in the center cell constituted from the sectors with ID={1,2,3} and then shifting the origin of the cell topology into the desired direction. In particular, if cells 1,3 and 8 (indicated by the blue framed region in Fig. 2 (left)) form a cooperation cluster, the users distributions in cell 1 and 3 are generated without any shift of the origin, while for the user in cell 8, the center cell is shifted to the north-east direction, as illustrated by the light blue region in Fig. 2 (left). The figure depicts another clustering given as red framed region. The active user set $\mathcal{M}$ in each cluster is indicated as blue triangles and red dots, respectively.

Performance is evaluated for the sum throughput in a specific cell cluster of size α. This value is divided by the signal bandwidth and α, yielding an effective spectral efficiency per sector. The achievable rates are determined from Shannon's formula, which represent theoretical limits in a practical system. The receiver is assumed to to have perfect channel state information (CSI), and feedback given by the MTs is considered to be error free and not affected by any delays.

VII. RESULTS

Beamforming and cooperation gains: Fig. 3(a) depicts the spectral efficiency of a system with a sparse user distribution, in particular $K = 2$ users per sector. For reference purpose, we include results obtained for a non-cooperationg SISO system, i.e. without any optimizations at receiver and transmitter side. The baseline is configured as a MIMO 2×2 system. The spectral efficiency for this setup is given as a blue line. By comparing this performance with a non-cooperative MET-based pre-coded system ($\alpha = 1$, green line), we observe a similar throughput. In the practical context, DFT-based pre-coding would be favorable due to complexity issues. However, this DFT-based pre-coding approach lacks the ability of downlink cooperation, which is intended to provide additional performance gains due to intra-cluster interference mitigation of $\zeta_{i,u}$. These gains are shown by the results from MET-based downlink BS cooperation using $\alpha \in \{2, 3\}$ BS sectors. We observe that MET benefits from additional degrees of freedom due to an increased number of transmit antennas spread over different BS locations, and thus enables additional beamforming gains. The relative gain with respect to the DFT system amounts to 59% for MET and $\alpha = 3$.

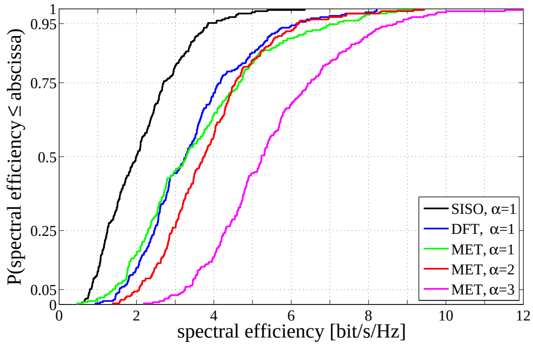
Multi-user selection diversity: In the following, the performance of the same system setup, but with $K = 10$ users per sector is given in Fig. 3(b). For all system settings considered in this work, we clearly observe a significant performance gain due to multi-user selection diversity.

In case of DFT-based pre-coding, the system beneficially uses the selection diversity for the purpose of interference-aware scheduling. This scheduling approach can be seen as a simple way of interference avoidance applicable in a cellular system. Therefore, the scheduler is taking channel quality identifier (CQI) information into account, which is extracted from the feedback provided by the MTs. This results in an intra-cluster interference $\zeta_{i,u} \neq 0$. The median spectral efficiency amounts to approx. 5.1 bit/s/Hz in this case.

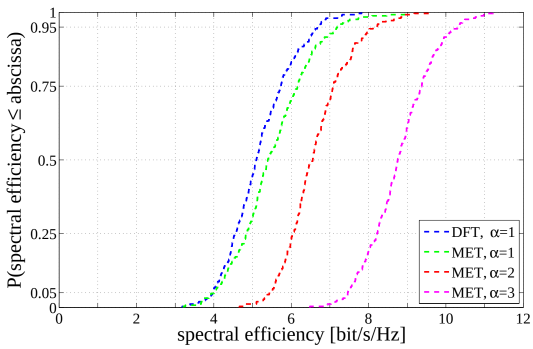
Turning the focus to MET-based pre-coding, we observe that this pre-coder realizes higher benefits from using multi-user selection diversity, refer to blue and green dashed lines in Fig. 3(b). The near-optimum pre-coder mainly benefits from the diversity within the user grouping process, which results in a orthogonal service to different users. Thus, the MET system mitigates the intra-cluster interference[3], while ensuring user orthogonalization with low costs in receive power reduction. In contrast to the DFT-based pre-coded system, the scheduler in

[3]Since the MMSE equalizer considers full CCI, $\zeta_{i,u} + \mathbf{z}_{i,u}$, optimum receiver weights are determined, so that the intra-cluster interference $\zeta_{i,u}$ cannot completely be forced to zero.

see the program



(a) $K = 2$ users per sector



(b) $K = 10$ users per sector

Fig. 3. Achievable spectral efficiency (Shannon) for different types of pre-coding; with and without multi-user selection diversity.

the BS takes the correlation of different users' effective eigenmode channels $\mathbf{\Gamma}_m$ into account. Further, Fig. 3(b) depicts the achievable spectral efficiency for the cooperative MIMO 2×2 for $\alpha \in \{2, 3\}$. The relative gains compared to the DFT system increase to 72% for MET and $\alpha = 3$, substantiating that the cooperation gains increase in the context of muti-user diversity. In this case the median achievable spectral efficiency amounts to approx. 8.8 bit/s/Hz.

VIII. Conclusion

In this work we compared different pre-coding techniques and their achievable performance in a cellular MU-MIMO OFDM downlink. As a baseline, we considered DFT-based beamforming as suggested in [1]. To evaluate the additional beamforming gains enabled by a near-optimum pre-coder, we use MET-based beamforming under the assumption of ideal knowledge of the users' effective eigenmode channels. In non-cooperative system with a sparse user distribution, we observe equivalent spectral efficiencies for both, DFT- and MET-based pre-coding. However, the downlink cooperation among $\alpha = 3$ sectors using MET is able to provide up to 59% performance gain compared to the non-cooperative DFT pre-coded system. Multi-user selection diversity was shown to enhance the system performance significantly, while MET turned out to benefit more from this diversity than DFT-based pre-coding. The gains for DFT beams are obtained by interference-aware scheduling, which is a simple way of interference avoidance in cellular systems. In contrats to that, the multi-user diversity in the MET system enables to determine most suited user groups with lowest correlation between different users' eigenmodes. Thus, the MET system mitigates the intra-cluster interference, while ensuring user orthogonalization with low costs in receive power reduction per data stream thanks to multi-user selection diversity.

Acknowledgements

The authors are grateful for financial support from the German Ministry of Education and Research (BMBF) in the national collaborative project EASY-C under contract No. 01BU0631.

References

[1] L. Thiele, M. Schellmann, W. Zirwas, and V. Jungnickel, "Capacity scaling of multi-user MIMO with limited feedback in a multi-cell environment," in *41st Asilomar Conference on Signals, Systems and Computers*. Monterey, USA: IEEE, Nov. 2007, invited.

[2] F. Boccardi and H. Huang, "A near-optimum technique using linear precoding for the MIMO broadcast channel," *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, pp. III–17–III–20, April 2007.

[3] L. Thiele, M. Schellmann, T. Wirth, and V. Jungnickel, "Cooperative multi-user MIMO based on reduced feedback in downlink OFDM systems," in *42nd Asilomar Conference on Signals, Systems and Computers*. Monterey, USA: IEEE, Nov. 2008.

[4] J. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 4, pp. 528–539, 1984.

[5] V. Jungnickel, T. Wirth, M. Schellmann, T. Haustein, and W. Zirwas, "Synchronization of cooperative base stations," in *IEEE International Symposium on Wireless Communication Systems 2008 (ISWCS08)*, Oct. 2008.

[6] L. Thiele, M. Schellmann, S. Schiffermüller, and V. Jungnickel, "Multi-cell channel estimation using virtual pilots," in *IEEE 67th Vehicular Technology Conference VTC2008-Spring*, Singapore, May 2008.

[7] T. Bonald, "A score-based opportunistic scheduler for fading radio channels," in *5th European Wireless Conference*, Feb. 2004.

[8] L. Thiele, M. Schellmann, T. Wirth, and V. Jungnickel, "Interference-Aware Scheduling in the Synchronous Cellular Multi-Antenna Downlink," Apr. 2009, invited.

[9] M. Schellmann, L. Thiele, T. Wirth, T. Haustein, and V. Jungnickel, "Resource Management in MIMO-OFDM systems," in *OFDMA: Fundamentals and Applications*, T. Jiang, L. Song, and Y. Zhang, Eds. CRC Press, Taylor&Francis Group, 2009.

[10] 3GPP TR 25.996 V7.0.0, "Spatial channel model for multiple input multiple output (MIMO) simulations (release 7)," July 2007. [Online]. Available: http://www.tkk.fi/Units/Radio/scm/

[11] H. Huang, S. Venkatesan, A. Kogiantis, and N. Sharma, "Increasing the peak data rate of 3G downlink packet data systems using multiple antennas," vol. 1, april 2003, pp. 311–315 vol.1.

CHANNEL ADAPTIVE OFDM SYSTEMS WITH PACKET ERROR RATIO ADAPTATION

Dagmar Bosanska, Christian Mehlführer, and Markus Rupp

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology

Gusshausstrasse 25/389, A-1040 Vienna, Austria

Email: {dbosansk, chmehl, mrupp}@nt.tuwien.ac.at

ABSTRACT

MIMO-OFDM systems with pre-filtering as well as bit and power loading have gained a lot of interest recently. These systems are based on theoretically optimal solutions maximizing the uncoded throughput and constraining the maximum uncoded BER. In this paper, we extend the existing frameworks to coded systems with bit and packet error ratio adaptation. We evaluate the performance of this approach in a standard compliant WiMAX link level simulator.

1. INTRODUCTION

The transmitter design adapted to the propagation channel is capable of improving both performance and throughput of a network. Therefore, an increased interest has been given to channel adaptive transmissions in MIMO-OFDM, e.g. in [1] there is a channel adaptive transmitter with Alamouti coding, in [2] they applied a beamformer with an appropriate power allocation policy but without bit loading. The application goes even further to the cross-layer optimized algorithms for OFDM systems [3, 4] where the user power and bit allocation policy works similarly as in [1]. To the best of our knowledge, all investigations have not yet assumed channel coding. Our algorithm is based on the problem formulation and solution from [1], but we extend this to coded systems and packet error ratio adaptation.

The rest of the paper is organized as follows. In Section 2, we present the MIMO-OFDM system model, followed by the formulation of the constrained optimization problem for the coded user throughput maximization in Section 3. In Section 4, the power cost and threshold metric required by the adaptive algorithm are derived for the $[133\ 171]_8$ half-rate convolutional code. This derivation is based on the bit error probability approximation. The optimal Greedy Power Allocation Algorithm is described in Section 5. The simulation results are obtained using a standard compliant WiMAX link level simulator and presented in Section 6. Finally, Section 7 concludes this paper.

2. MIMO-OFDM SYSTEM MODEL

We consider MIMO-OFDM systems equipped with $K = 192$ data subcarriers (FFT size is 256), N_T transmit, and N_R receive antennas. Figure 1 shows the discrete time baseband model that is obtained under the following considerations. Let $\mathbf{h}_{n_\mathrm{r}n_\mathrm{t}}$ be the baseband equivalent FIR channel between the n_t-th transmit and the n_r-th receive antenna during the given block with channel order N. Let $k \in \{1, \ldots, K\}$ denote the subcarrier index. The frequency flat channel $H_{n_\mathrm{r}n_\mathrm{t}}[k]$ between the n_t-th transmit and the n_r-th receive antenna on the kth subcarrier is given by the Discrete Fourier Transform (DFT). The full MIMO frequency flat channel on the kth subcarrier is then modeled by the $N_\mathrm{R} \times N_\mathrm{T}$ channel matrix $\mathbf{H}[k]$ with rank R

$$\mathbf{H}[k] = \begin{bmatrix} H_{11}[k] & \cdots & H_{1N_\mathrm{T}}[k] \\ \vdots & \ddots & \vdots \\ H_{N_\mathrm{R}1}[k] & \cdots & H_{N_\mathrm{R}N_\mathrm{T}}[k] \end{bmatrix}. \tag{1}$$

We assume that the channel $\mathbf{H}[k]$ changes slowly and is fully known at the transmitter and at the receiver. If the latter condition cannot be satisfied, almost ideal performance can be achieved by applying advanced LMMSE channel estimators with acceptable complexity [5]. By stacking the received symbols of all N_R receive antennas

$$\mathbf{y}[k] = [y_1[k], y_2[k], \ldots, y_{N_\mathrm{R}}[k]]^\mathrm{T} \tag{2}$$

we can calculate the received symbol vector as

$$\mathbf{y}[k] = \mathbf{H}[k]\mathbf{V}[k]\tilde{\mathbf{a}}[k] + \mathbf{n}[k], \tag{3}$$

where $\mathbf{n}[k]$ is zero mean circularly symmetric complex Additive White Gaussian Noise (AWGN) with covariance $\mathrm{E}\{\mathbf{n}[k]\mathbf{n}[k]^\mathrm{H}\} = N_0\mathbf{I}_{N_\mathrm{R}}$, $\mathbf{V}[k]$ is a pre-filtering matrix, and $\tilde{\mathbf{a}}[k]$ is a power-loaded transmit symbol vector obtained as

$$\tilde{\mathbf{a}}[k] = \mathbf{p}[k] \odot \mathbf{a}[k], \tag{4}$$

where $\odot$ denotes an element-wise multiplication between the power loading vector $\mathbf{p}[k] = \left[\sqrt{P_1[k]}, \ldots, \sqrt{P_R[k]}\right]^\mathrm{T}$ and the vector $\mathbf{a}[k]$. Each information symbol $a_r[k], r = 1, \ldots, R$ is an element of a given symbol alphabet $\mathcal{A}_r[k]$, consisting of $M_r[k] = 2^{2m}, m = 1, 2, \ldots$ symbols for square QAMs.

For exploiting the channel knowledge even further, linear pre-filtering is used in (3). The pre-filtering matrix $\mathbf{V}[k]$ contains analyzing basis vector directions obtained by the Singular Value Decomposition (SVD) of the channel matrix $\mathbf{H}[k] = \mathbf{U}[k]\mathbf{\Sigma}[k]\mathbf{V}^\mathrm{H}[k]$. When (3) is left-multiplied by $\mathbf{U}^\mathrm{H}[k]$ and the SVD of the channel matrix per subcarrier k is applied, (3) reduces to

$$\tilde{\mathbf{y}}[k] = \mathbf{\Sigma}[k]\tilde{\mathbf{a}}[k] + \tilde{\mathbf{n}}[k], \tag{5}$$

where $\tilde{\mathbf{n}}[k] = \mathbf{U}^\mathrm{H}[k]\mathbf{n}[k]$ and $\tilde{\mathbf{y}}[k] = \mathbf{U}^\mathrm{H}[k]\mathbf{y}[k]$.

see the program

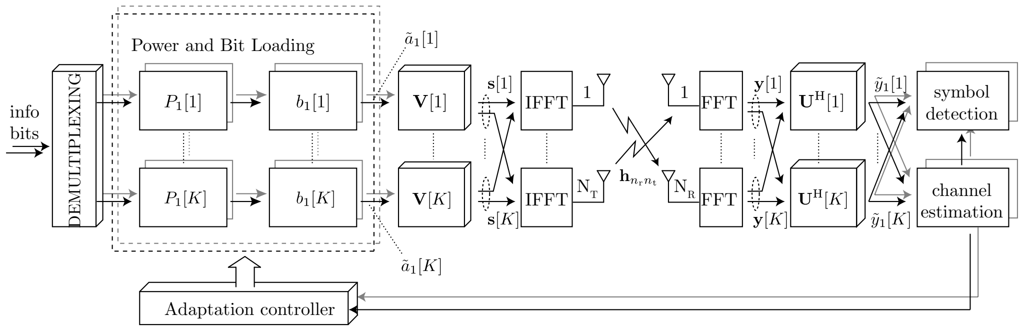


Fig. 1. Discrete-time equivalent baseband MIMO OFDM model.

3. PROBLEM FORMULATION

The goal is to optimize the coded throughput of the MIMO-OFDM system depicted in Figure 1. The power levels $P_r[k]$ to be assigned to the subcarriers with index k and to the spatial sub-channels with index r are grouped into the transmit power matrix

$$\mathbf{P} = \begin{bmatrix} P_1[1] & \cdots & P_1[K] \\ \vdots & \ddots & \vdots \\ P_R[1] & \cdots & P_R[K] \end{bmatrix}, \tag{6}$$

so are the number of bits $b_r[k]$ grouped into the matrix

$$\mathbf{B} = \begin{bmatrix} b_1[1] & \cdots & b_1[K] \\ \vdots & \ddots & \vdots \\ b_R[1] & \cdots & b_R[K] \end{bmatrix}. \tag{7}$$

The objective is to maximize the coded user throughput T subject to a total transmit power constraint $\bar{P}$ and discrete modulation levels while maintaining a target bit error ratio $\mathrm{BER}_r^{(0)}[k]$ on each subcarrier and in each sub-channel. This objective can thus be formulated as the following constrained optimization problem:

$$\begin{aligned} \text{maximize } T(\mathbf{P},\mathbf{B}) &= \sum_{r=1}^{R}\sum_{k=1}^{K} b_r[k] && (8)\\ \text{subject to coded } \overline{\mathrm{BER}}_r[k] &\leq \mathrm{BER}_r^{(0)}[k] && (9)\\ \sum_{r=1}^{R}\sum_{k=1}^{K} P_r[k] &\leq \bar{P} && (10)\\ P_r[k] &\geq 0 && (11)\\ b_r[k] &\in \{0,1,2,3,4,\ldots\} && (12) \end{aligned}$$

The target BER is assumed to be identical across subcarriers and sub-channels and equal to $\mathrm{BER}^{(0)}$. The relationship between the initial $\mathrm{BER}^{(0)}$ and the resulting coded $\mathrm{PER}^{(1)}$ in the first iteration of the PER adaptation to the target packet error ratio $\mathrm{PER}^{(0)}$ is given by

$$\mathrm{PER}^{(1)} = 1-(1-\mathrm{BER}^{(0)})^{\left(\sum_{r=1}^{R}\sum_{k=1}^{K} b_r^{(1)}[k]\cdot R_c\cdot N_s\right)}, \tag{13}$$

where N_s is the number of OFDM data symbols in a WiMAX frame representing one codeword, and R_c is the coding rate.

4. THRESHOLD METRIC AND POWER COST FOR A CONVOLUTIONAL CODE AND M-QAM

Note that in (12) both rectangular ($b[k] \in \{1,3,5,\ldots\}$) and square ($b[k] \in \{2,4,6,\ldots\}$) QAMs are allowed ($b[k] = 0$ refers to the case of an unused subcarrier). However, we want to reduce the possible constellations just to square ones, what significantly simplifies an adaptive demapper. This excludes also the BPSK modulation, which is used in the first AMC mode of the WiMAX standard. By doing this, one bit at most is lost from the throughput per one OFDM symbol. The proof is given in [1]. Let $d_{\mathrm{min}}[k]$ denote the minimum square Euclidean distance for the given constellation on the kth subcarrier. For QAM constellations, it holds that [6]

$$d_{\mathrm{min}}^2[k] = 4g(b[k])E_s[k] = 4g(b[k])P[k]T_s, \tag{14}$$

where $E_s[k] = P[k]T_s$ is the average energy of the constellation chosen for the kth subcarrier, T_s is the useful symbol duration (OFDM symbol duration without cyclic prefix) and the constant $g(b[k])$ depends on whether the given constellation is square or rectangular QAM:

$$g(b[k]) := \begin{cases} \frac{6}{5\cdot 2^{b[k]}-4}, & b[k] = 1,3,5,\ldots; \\ \frac{6}{4\cdot 2^{b[k]}-4}, & b[k] = 2,4,6,\ldots. \end{cases} \tag{15}$$

From the definition of the constant $g(b[k])$ in (15) it is clear that the square QAMs are more power efficient than the rectangular QAMs. Thus, with K subcarriers, it is always possible to avoid the usage of less efficient rectangular QAMs and save the remaining power for other subcarriers to use higher order square QAM [9]. Our exclusion of the rectangular constellations is thus justified.

Equation (14) simply gives us the decreasing minimum Euclidean distance in the constellation with the increasing modulation order and can be easily verified. According to (14) replacing $d_{\mathrm{min}}^2[k]$ by term $d_{\mathrm{avg}}^2[k] = c(b[k])\cdot g(b[k])P[k]$, where T_s is equal to one second, the required power $P[k]$ to transmit $b[k]$ bits/s/Hz is equal to

$$P[k] = \frac{d_{\mathrm{avg}}^2[k]}{c(b[k])\cdot g(b[k])} = \frac{d_0^2[k]}{g(b[k])}, \tag{16}$$

because $d_{\mathrm{avg}}^2[k]/d_0^2[k] = c(b[k])$ and $d_0^2[k] = d_{\mathrm{min}}^2[k]/4$, where $d_0[k]$ is our threshold metric or the unit grid in the

see the program

M-QAM constellation. The constant $c(b[k])$ that depends on the modulation order is set to be higher than four and it expresses the deviation from the worst-case upper bound on a bit error probability of the convolutional code. This worst-case upper bound refers to the case when $d^2_{\text{avg}}[k] = d^2_{\min}[k]$. In this way, the bit error probability approximation is obtained for the $[133\ 171]_8$ half-rate convolutional code with M-QAM modulation for an AWGN channel

$$P_{\text{b}}[k] \approx \frac{1}{n} \sum_{d=d_{\text{free}}}^{d_{\text{free}}+4} \beta(d)[k] \frac{1}{5} \exp\left(\frac{-c(b[k]) \cdot d_0^2[k] \cdot \hat{d}(b[k])}{4N_0} \right), \tag{17}$$

where n is the $n = 1$ input bits of the convolutional encoder and $\hat{d}(b[k]) = d_{\text{free}} + 4$, thus making the exponential term of this approximation independent from d. More details about this probability approximation and its verification can be found in [7, 8]. We decided to use for the simulations $c(b[k]) = 8, 10, 20$ for 4-,16-, and 64-QAM, respectively, in order to achieve the bit error ratio in the range of 10^{-6} to 10^{-5} [8]. The threshold metric or the unit grid $d_0^2[k]$ can be expressed from (17) as follows

$$d_0^2[k] = \left[\ln\left(\sum_{d=d_{\text{free}}}^{d_{\text{free}}+4} \beta(d)[k] \right) - C \right] \frac{4N_0}{c(b[k]) \cdot \hat{d}(b[k]) \cdot \sigma^2[k]}, \tag{18}$$

where $C = \ln(5P_{\text{b}})$ and $\sigma[k]$ is a nonzero eigenvalue of the channel on a subcarrier k. This constant C is replaced by $C = \ln 5\text{BER}^{(0)}$ for practical purposes, as the bit error probability can be only achieved in the simulations when the number of transmitted bits goes to infinity. The power cost incurred when loading the lth and $(l-2)$th bits (two bits are always loaded in one step) to the kth subcarrier and the rth sub-channel is from (18) and (16) for coded systems

$$p_r[k,l] = \frac{d^2_{0,r}[k,l]}{g_r[l]} - \frac{d^2_{0,r}[k,l-2]}{g_r[l-2]},\ l = 2,4,6;\ \forall k,r. \tag{19}$$

So this cost is quantified by the additional power needed to maintain the target bit error ratio performance. For $l = 2$, $g_r[l-2]$ is set to ∞ and $c_r[0] = \hat{d}_r[0] = 1$.

5. GREEDY POWER ALLOCATION ALGORITHM

In this section, the optimal Greedy Power Allocation Algorithm with the packet error ratio adaptation is described. This greedy algorithm converts the frequency-selective channel into a frequency-flat one, what is the main idea behind the water-filling algorithm. Therefore, the performance of the variable-rate M-QAM modulation with convolutional encoding in an AWGN channel may be considered per subcarrier, as it was already done in Section 4. Moreover, the target packet error ratio $\text{PER}^{(0)}$ per each codeword can be maintained. Remember that the same target bit error ratio $\text{BER}^{(0)}$ and target packet error ratio $\text{PER}^{(0)}$ across all subcarriers are assumed.

Two iteration loops are given, in the inner nth iteration, it is decided whether to load two more bits or not, in the outer mth iteration it is checked if the target packet error ratio is achieved. In the first iteration step (m = 1), we start with the target bit error ratio $\text{BER}^{(0)}$ set to 10^{-5}, if the target packet error ratio equal to 10^{-2} is of interest. It is important to have the starting value of BER to obtain the first estimation of the codeword length. In addition, the expression for the coded power cost in (19) depends on the target bit error ratio. It follows from the simulation results, plotted in [8], that for the $\text{PER}^{(0)} = 10^{-2}$, the bit error ratio has to be in the range of 10^{-6} to 10^{-5} (also the bit error ratio approximation in (17) was previously optimized for this BER region).

In this greedy algorithm, P_{rem} denotes the remaining power after each iteration, $b_r^{(n)}[k]$ the number of bits loaded at the kth subcarrier and the rth sub-channel in the nth iteration, and $P_r^{(n)}[k]$ to denote the power level in iteration step n on the kth subcarrier and in the rth sub-channel. The total number of loaded data bits N_{total} in the nth iteration can then be calculated by summing the loaded bits over R sub-channels, K subcarries and N_{s} OFDM symbols and taking also the coding rate R_{c} into account

$$N_{\text{total}} = \sum_{r=1}^{R} \sum_{k=1}^{K} b_r^{(n)}[k] \cdot R_{\text{c}} \cdot N_{\text{s}}. \tag{20}$$

The optimal Greedy Power Allocation Algorithm can be performed according to the following steps:

1. In the first (outer) iteration step $m = 1$, $\text{BER}^{(1)}$ is equal to the initial target bit error ratio $\text{BER}^{(0)}$, otherwise

$$\text{BER}^{(m)} = 1 - (1 - \text{PER}^{(0)})^{1/N_{\text{total}}}. \tag{21}$$

2. Initialization step $n = 1$: Set the remaining power equal to the power constraint $P_{\text{rem}} = \bar{P}$. For each subcarrier and sub-channel, set $b_r^{(n)}[k] = P_r^{(n)}[k] = 0$.

3. Compute $p_r(k, b_r^{(n)}[k] + 2])$ for all subcarriers and sub-channels, where $b_r^{(n)}[k] \neq 6$. If $b_r^{(n)}[k] = 6$, then set $p_r(k, b_r^{(n)}[k]) = \infty$. Choose the subcarrier and sub-channel that needs the least power to load two additional bits, i.e. select

$$\{r_0, k_0\} = \arg\min_{r,k} p_r(k, b_r^{(n)}[k] + 2). \tag{22}$$

4. If there is not enough power remaining, i.e. if $P_{\text{rem}} < p_{r_0}(k_0, b_{r_0}^{(n)}[k_0] + 2)$ (it always happens if all subcarriers and sub-channels are loaded with 6 bits ($b_r^{(n)}[k] = 6\ \forall r, k$)), then jump to Step 5. Otherwise, load two bits to the k_0th subcarrier and the r_0th sub-channel, and update iteration variables:

$$P_{\text{rem}} = P_{\text{rem}} - p_{r_0}(k_0, b_{r_0}^{(n)}[k_0] + 2) \tag{23}$$
$$P_{r_0}^{(n)}[k_0] = P_{r_0}^{(n)}[k_0] + p_{r_0}(k_0, b_{r_0}^{(n)}[k_0] + 2) \tag{24}$$
$$b_{r_0}^{(n)}[k_0] = b_{r_0}^{(n)}[k_0] + 2. \tag{25}$$

Loop back to step 3 with $n = n + 1$.

see the program

5. Calculate

$$\mathrm{PER}^{(\mathrm{m})} = 1 - (1 - \mathrm{BER}^{(m)})^{1/N_{\mathrm{total}}} \tag{26}$$

If $\mathrm{PER}^{(m)} \leq \mathrm{PER}^{(0)}$, then exit with

$$\left\{ P_r[k] = P_r^{(n)}[k], b_r[k] = b_r^{(n)}[k] \right\}_{r=1,k=1}^{R,K}, \tag{27}$$

otherwise loop back to step 1 with $m = m + 1$.

Applying this iterative algorithm, the maximum number of iterations for the packet error ratio adaptation is usually equal to two (m = 2). It may happen that the pre-set target bit error ratio is already sufficient to obtain the target packet error ratio after the first outer iteration. Otherwise, the pre-set target bit error ratio is corrected accordingly to meet the PER requirement after the second outer iteration.

The optimal power levels obtained from this algorithm can be only then assigned when the individual SISO channels of the MIMO system are accessible. This is guaranteed by the optimal linear pre-filtering as can be seen in (5)

6. SIMULATION RESULTS

The simulation results presented in this section are obtained using a standard compliant IEEE 802.16-2004 WiMAX link level simulator. First, we derive the expression for the ergodic capacity that will be used as a performance bound for the simulated data throughput. For this purpose, the capacities $\bar{C}'[k]$ and $\bar{C}''[k]$ per subcarrier k of a MIMO channel are defined as the ensemble average of the information rate over the N_{sim} simulated random realizations of the channel matrix $\mathbf{H}[k]$. It is assumed that the channel matrix $\mathbf{H}[k]$ is full-rank, that is, $R = N_{\mathrm{T}} = N_{\mathrm{R}} = N$ and the channel eigenvalues are $\lambda_r = \sigma_r^2 > 0$, for $r = 1, \ldots, N$, what is always fulfilled in the simulations with an uncorrelated Pedestrian B channel.

The ergodic capacity $\bar{C}'$ of the OFDM system can be calculated by summing the individual capacities $\bar{C}'[k]$ of all data subcarriers (192 data subcarriers are used in the WiMAX standard)

$$\bar{C}' = \frac{F}{N_{\mathrm{sim}}} \sum_{s=1}^{N_{\mathrm{sim}}} \sum_{k=1}^{192} \sum_{r=1}^{N} \log_2 \left(1 + \frac{P_r^{(s)}[k]\lambda_r^{(s)}[k]}{N_0} \right), \tag{28}$$

in case of the channel knowledge at the transmitter, and

$$\bar{C}'' = \frac{F}{N_{\mathrm{sim}}} \sum_{s=1}^{N_{\mathrm{sim}}} \sum_{k=1}^{192} \sum_{r=1}^{N} \log_2 \left(1 + \frac{\lambda_r^{(s)}[k]}{N_0} \right), \tag{29}$$

when no channel knowledge is assumed at the transmitter. The ergodic capacity in (28) and the mutual information in (29) was derived from the capacity of the MIMO channel expressed as the sum of the capacities of N SISO channels and can be found in [9]. Since the transmission of an OFDM signal requires also the transmission of a cyclic prefix to avoid inter-symbol interference, and a preamble for the synchronization and channel estimation, the ergodic capacity and mutual information given by Equations (28) and (29) include the

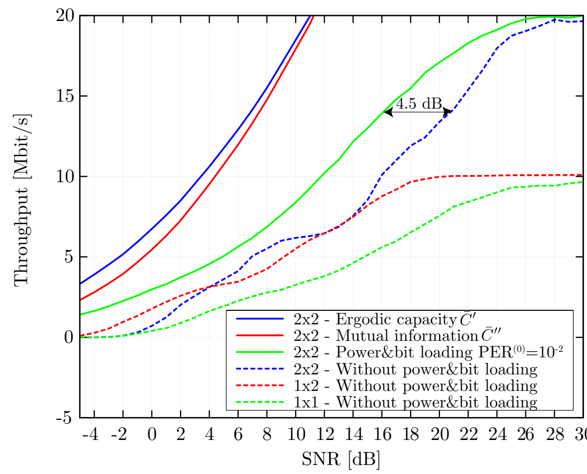


Fig. 2. Performance of channel adaptive algorithms in an uncorrelated Pedestrian B channel with $\mathrm{PER}^{(0)} = 10^{-2}$.

correction factor F. This correction term accounts for these inherent system losses and is defined according to [10] as follows

$$F = \frac{1}{1+G} \cdot \frac{1/T_{\mathrm{sample}}}{256} \cdot \frac{N_{\mathrm{s}}}{N_{\mathrm{total}}}, \tag{30}$$

where G (1/4 in the simulations) corresponds to the ratio of the cyclic prefix time and useful OFDM symbol time T_{s}, N_{s} (44 in the simulations) is the number of OFDM data symbols, N_{total} (47 in the simulations) is the total number of OFDM symbols in one transmission frame (also with the preamble), and T_{sample} is the sampling rate of the transmit signal. Therefore, the factor $(1/T_{\mathrm{sample}})/256$ is equivalent to the available bandwidth per subcarrier (the FFT size is equal to 256) [10].

In Figure 2, the channel adaptive algorithm for a 2×2 MIMO system with power and bit loading is compared to the ergodic capacity $\bar{C}'$ and mutual information $\bar{C}''$, as well as a 2×2, a 1×2, and a 1×1 system without power and bit loading. The half rate convolutional code of the WiMAX standard is used in combination with 4-,16- and 64-QAM. The 2×2 system using this adaptive algorithm outperforms the standard non adaptive 2×2 system by maximum 4.5 dB. The difference in the low SNR region from the ergodic capacity $\bar{C}'$ is as low as 5 dB. This loss is due to the non-optimal channel coding. With increasing SNR, the loss is growing significantly because of the low rate of the convolutional code and could be reduced by larger symbol alphabets and/or higher coding rates. As expected, the ergodic capacity $\bar{C}'$ with the channel known at the transmitter is higher than the mutual information $\bar{C}''$ in the case of unknown channel. This advantage reduces at higher SNR, as can be observed in Figure 2 and is proved in [9]. Interestingly, although the capacity advantage is only small, the gain in throughput is quite significant.

In Figure 3 the resulting frame error ratio is plotted for the target packet error ratio equal to 10^{-2}. The adaptation performance of the optimal Greedy Power Allocation Algorithm is satisfactory over a large SNR range of 34 dB.

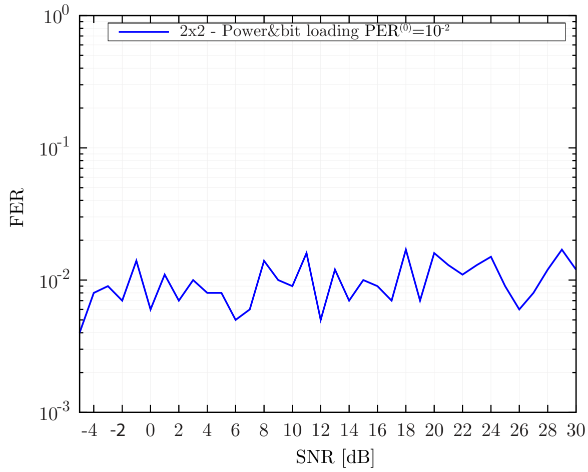


Fig. 3. Frame error ratio adaptation for the target packet error ratio $\mathrm{PER}^{(0)} = 10^{-2}$.

7. CONCLUSION

In this paper, we create a framework for the bit error ratio and packet error ratio adaptation. In slow fading channels, we find this to be a very promising approach, because the achievable capacity can be reached by applying an appropriate modulation and coding scheme. Moreover, the PER of the system can be controlled. To the best of our knowledge, this is the first attempt to evaluate the performance of the water-pouring based solution in the more realistic scenario of a coded WiMAX system. It remains to further explore the behavior of the proposed algorithm and the possible gains when more efficient channel coding is deployed. We also plan to implement this framework into our Long Term Evolution (LTE) link level simulator [11] that is also based on MIMO-OFDM and is currently under development.

Finally, the complexity of the optimal Greedy Power Allocation Algorithm grows linearly with the number of bits and the number of subcarriers. Obviously, the complexity is considerably large when the number of bits and subcarriers is large. The maximum number of iterations for one particular bit and power level assignment is for example for the $2{\times}2$ MIMO OFDM system with 192 data subcarriers equal to $3 \cdot 2 \cdot 192 = 1152$, where 3 is the maximum number of bit loading steps for the kth subcarrier and the rth subchannel (4-QAM → 16-QAM → 64-QAM) . For every assignment of two bits, the power cost has to be calculated across all subcarriers and sub-channels. For the adaptive transmitter design, it is recommended in [1] to use in practice the fast Lagrange bi-sectional search proposed in [12] that also provides an optimal solution with lower complexity.

ACKNOWLEDGEMENT

This work has been funded by the Christian Doppler Laboratory for Design Methodology of Signal Processing Algorithms and the Austrian Research Centers GmbH.

References

[1] P. Xia, S. Zhou, and G.B. Giannakis, "Adaptive MIMO-OFDM based on partial channel state information," *IEEE transactions on signal processing*, vol. 52, no. 1, pp. 202–213, 2004.

[2] S. Pascual-Iserte, AI Perez-Neira, and MA Lagunas, "On power allocation strategies for maximum signal to noise and interference ratio in an MIMO-OFDM system," *Wireless Communications, IEEE Transactions on*, vol. 3, no. 3, pp. 808–820, 2004.

[3] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks-part I: theoretical framework," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 2, pp. 614–624, 2005.

[4] G. Song and YG Li, "Cross-layer optimization for OFDM wireless networks-part II: algorithm design," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 2, pp. 625–634, 2005.

[5] Christian Mehlführer and Markus Rupp, "Approximative LMMSE channel estimation for OFDM WiMAX," in *Proc. Third International Symposium on Communications, Control, and Signal Processing (ISCCSP 2008)*, St. Julians, Malta, Mar. 2008.

[6] S. Zhou and GB Giannakis, "Adaptive Modulation for multiantenna transmissions with channel mean feedback," *Wireless Communications, IEEE Transactions on*, vol. 3, no. 5, pp. 1626–1636, 2004.

[7] R.C. Manso, "Performance analysis of M-QAM with Viterbi soft-decision decoding," M.S. thesis, Naval Postgraduate School, California, March 2003.

[8] D. Bosanska, "Channel adaptive OFDM systems with quantized feedback," M.S. thesis, Vienna Universtity of Technology, Austria, November 2008, `http://publik.tuwien.ac.at/files/PubDat_168072.pdf`.

[9] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 2003.

[10] Christian Mehlführer, Sebastian Caban, and Markus Rupp, "Experimental evaluation of adaptive modulation and coding in MIMO WiMAX with limited feedback," *EURASIP Journal on Advances in Signal Processing, Special Issue on MIMO Systems with Limited Feedback*, vol. 2008, Article ID 837102, 2008.

[11] Christian Mehlführer, Martin Wrulich, Josep Colom Ikuno, Dagmar Bosanska, and Markus Rupp, "Simulating the long term evolution physical layer," in *Proc. of the 17th European Signal Processing Conference (EUSIPCO 2009)*, Glasgow, Scotland, Aug. 2009, submitted.

[12] BS Krongold, K. Ramchandran, and DL Jones, "Computationally efficient optimal power allocation algorithm for multicarrier communication systems," in *Communications, 1998. ICC 98. Conference Record. 1998 IEEE International Conference on*, 1998, vol. 2.

TIME-DELAY ESTIMATION APPLYING THE EXTENDED INVARIANCE PRINCIPLE WITH A POLYNOMIAL ROOTING APPROACH

Felix Antreich, Josef A. Nossek⋆, Gonzalo Seco†, and A. Lee Swindlehurst‡*

German Aerospace Center (DLR), Inst. for Communications and Navigation, 82234 Wessling, Germany, phone: +49 8153 28 2119 *

Inst. for Circuit Theory & Signal Processing, Munich University of Technology (TUM), Germany⋆

Dept. of Telecommunications & Systems Engineering, Universitat Autonoma de Barcelona (UAB), Spain†

Dept. of Electrical Engineering & Computer Science, The University of California, Irvine CA, U.S.A.‡

e-mail: felix.antreich@dlr.de*, nossek@nws.ei.tum.de⋆, gonzalo.seco@uab.es†, and swindle@uci.edu‡

ABSTRACT

This work treats the problem of joint estimation of time-delay and spatial (direction-of-arrival, DOA) parameters of several replicas of a known signal in an unknown spatially correlated field. Unstructured and structured data models have been proposed for maximum likelihood (ML) estimators, whereas the former suffers from a severe performance degradation in some scenarios, and the latter involves huge complexity. In this work it is shown how the extended invariance principle (EXIP) can be applied to obtain estimates with the quality of those of the structured model, but with much lower complexity than directly utilizing the structured model. We present how to improve the quality of the time-delay estimates obtained with an unstructured spatial model by introducing DOA estimates. DOA estimates are derived either directly applying EXIP with a polynomial rooting approach or with decoupled estimators for temporal and spatial parameters applying Unitary ESPRIT. Both methods are compared with respect to estimation accuracy and complexity. Exemplarily, simulation results for time-delay estimation for GPS (Global Positioning System) are shown and confirm that our proposals both approach the Cramer-Rao lower bound (CRLB) of the structured model.

1. INTRODUCTION

Channel estimation is important in many applications as MIMO channel characterization, radar, synchronization, and Global Navigation Satellite Systems (GNSS) like GPS (Global Positioning System). A simple unstructured data model has been used for such problems in order to have low complexity [1] and a solution for an unknown spatial field was given in [2]. On the other hand the structured data model provides better results [3, 4] but has a high complexity and only seems to be easily manageable in the single path case [5]. The latter work already shows how the extended invariance principle (EXIP) can be applied to refine maximum likelihood (ML) estimates for an unstructured model to achieve the performance available using a structured model.

This work extends the approach given in [5] to a multipath case. We present a two-step approach to improve the quality of time-delay estimates obtained with an unstructured spatial model. As a first step we, apply the Space Alternating Generalized Expectation Maximization (SAGE) algorithm [4] to obtain estimates for an unstructured model. As a second step, we adopt EXIP to refine the time-delay estimates. The important new result with respect to [6] is the derivation of a low complexity polynomial rooting approach, an iterative quadratic maximum likelihood (IQML) [7] algorithm in order to solve the resulting direction-of-arrival (DOA) estimation problem. This new approach provides automatic pairing between the DOA estimates and parameter estimates for the unstructured model. In [6] we applied Unitary ESPRIT for the estimation of the DOAs based on the estimates of the spatial covariance matrix and the spatial signatures for the unstructured model which showed that the proposed two-step approach attains the Cramer-Rao lower bound (CRLB) of the structured model even when suboptimal DOA estimates are introduced. However, for this method the pairing between the DOA estimates and the parameter estimates for the unstructured model also has to be estimated. Exemplarily, simulation results for time-delay estimation of the line-of-sight signal (LOSS) for GPS (Global Positioning System) are included and confirm that our proposal approaches the CRLB of the structured model.

2. DATA MODEL

We assume that L narrowband planar wavefronts, $1 \leq \ell \leq L$ are impinging on an antenna array of M isotropic sensor elements. The noise-plus-interference corrupted baseband signal at the antenna output $\mathbf{y}(t) \in \mathbb{C}^{M\times 1}$ can be modelled as a superposition of L wavefronts and additional temporally white Gaussian noise $\mathbf{n}(t) \in \mathbb{C}^{M\times 1}$,with zero-mean and unknown spatial covariance matrix $\mathbf{Q} \in \mathbb{C}^{M\times M}$.

2.1. Structured Model

For the structured data model we define the parameter vector $\boldsymbol{\theta} = \left[\mathrm{Re}\{\boldsymbol{\gamma}\}^{\mathrm{T}}, \mathrm{Im}\{\boldsymbol{\gamma}\}^{\mathrm{T}}, \boldsymbol{\phi}^{\mathrm{T}}, \boldsymbol{\tau}^{\mathrm{T}}\right]^{\mathrm{T}}$ on the domain $D_{\boldsymbol{\theta}}$ with the vector of complex amplitudes $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_\ell, \ldots, \gamma_L]^{\mathrm{T}}$, the vector of azimuth angles $\boldsymbol{\phi} = [\phi_1, \ldots, \phi_\ell, \ldots, \phi_L]^{\mathrm{T}}$, and the vector of time-delays $\boldsymbol{\tau} = [\tau_1, \ldots, \tau_\ell, \ldots, \tau_L]^{\mathrm{T}}$. Thus we

see the program

can write

$$\mathbf{y}(t) = \sum_{\ell=1}^{L} \mathbf{s}_\ell(t) + \mathbf{n}(t), \tag{1}$$

where $\mathbf{s}_\ell(t)$ is given by

$$\mathbf{s}_\ell(t) = \mathbf{a}_\ell\left(\phi_\ell\right)\ \gamma_\ell\ c(t-\tau_\ell). \tag{2}$$

Here, $\mathbf{a}_\ell\left(\phi_\ell\right)$ denotes the steering vector of an antenna array and $c(t-\tau_\ell)$ denotes the pseudo-noise (PN) binary sequence with delay τ_ℓ.

2.2. Unstructured Model

For the simpler unstructured model we define the parameter vector $\boldsymbol{\xi} = \left[\mathrm{Re}\{\mathrm{vec}\{\mathbf{H}\}\}^{\mathrm{T}}, \mathrm{Im}\{\mathrm{vec}\{\mathbf{H}\}\}^{\mathrm{T}}, \boldsymbol{\tau}^{\mathrm{T}}\right]^{\mathrm{T}}$ on the domain $D_{\boldsymbol{\xi}}$ with the matrix of spatial signatures $\mathbf{H} = [\mathbf{h}_1 \ \cdots \ \mathbf{h}_\ell \ \cdots \ \mathbf{h}_L] \in \mathbb{C}^{M\times L}$. Here, $\mathrm{vec}\{\cdot\}$ denotes the vec operator, which vectorizes a matrix by stacking its columns. Thus we can alternatively write

$$\mathbf{s}_\ell(t) = \mathbf{h}_\ell\ c(t-\tau_\ell). \tag{3}$$

3. MAXIMUM LIKELIHOOD (ML) ESTIMATION FOR THE UNSTRUCTURED MODEL

The spatial observations are collected at N time instances, as $\mathbf{y}[n] = \mathbf{y}(n\cdot T_s)$ with $n = 1, 2, \ldots, N$. The channel parameters are assumed constant during the observation interval. Collecting the samples of the observation interval leads to

$$\mathbf{Y} = [\mathbf{y}[1], \mathbf{y}[2], \ldots, \mathbf{y}[N]] \in \mathbb{C}^{M\times N}, \tag{4}$$

$$\mathbf{N} = [\mathbf{n}[1], \mathbf{n}[2], \ldots, \mathbf{n}[N]] \in \mathbb{C}^{M\times N}, \tag{5}$$

$$\mathbf{S}(\boldsymbol{\xi}) = [\mathbf{s}[1], \mathbf{s}[2], \ldots, \mathbf{s}[N]] \in \mathbb{C}^{M\times N}. \tag{6}$$

Thus, the unstructured signal model can be written in matrix notation

$$\mathbf{Y} = \mathbf{S}(\boldsymbol{\xi}) + \mathbf{N} = \sum_{\ell=1}^{L} \mathbf{S}_\ell(\boldsymbol{\xi}_\ell) + \mathbf{N} = \mathbf{H}\,\mathbf{C} + \mathbf{N}. \tag{7}$$

Here, $\boldsymbol{\xi}_\ell = \left[\mathrm{Re}\{\mathbf{h}_\ell\}^{\mathrm{T}}, \mathrm{Im}\{\mathbf{h}_\ell\}^{\mathrm{T}}, \tau_\ell\right]^{\mathrm{T}}$ contains the parameters of one wave, and $\mathbf{C} = [\mathbf{c}(\tau_1)\cdots\mathbf{c}(\tau_\ell)\cdots\mathbf{c}(\tau_L)]^{\mathrm{T}} \in \mathbb{R}^{L\times N}$ contains the sampled PN binary sequence for each impinging wavefront $\mathbf{c}(\tau_\ell)$. The negative log-likelihood function for the unstructured data model, neglecting additive constants is [5]

$$\begin{aligned}\Lambda(\boldsymbol{\xi}, \mathbf{Q}) &= N\cdot\ \log\ (\det\,(\mathbf{Q})) \\ &\quad + \mathrm{tr}\left(\mathbf{Q}^{-1}\,(\mathbf{Y}-\mathbf{S}(\boldsymbol{\xi}))\,(\mathbf{Y}-\mathbf{S}(\boldsymbol{\xi}))^{\mathrm{H}}\right),\end{aligned} \tag{8}$$

where $(\cdot)^{\mathrm{H}}$ denotes complex conjugate transposition, $\det(\cdot)$ the determinant operation, and $\mathrm{tr}(\cdot)$ the trace operator. The ML estimate for $\mathbf{Q}$ is given by

$$\hat{\mathbf{Q}} = \frac{1}{N}\,(\mathbf{Y}-\mathbf{S}(\boldsymbol{\xi}))\,(\mathbf{Y}-\mathbf{S}(\boldsymbol{\xi}))^{\mathrm{H}}. \tag{9}$$

4. SPACE-ALTERNATING GENERALIZED EXPECTATION MAXIMIZATION (SAGE) ALGORITHM

In order to obtain the ML estimates $\hat{\boldsymbol{\xi}}$ for the unstructured model we apply an iterative method, the SAGE algorithm [4]. We assume that L is given. The expectation step (E-Step) can be expressed as

$$\hat{\mathbf{X}}_\ell = \mathbf{Y} - \sum_{\substack{\ell'=1 \\ \ell'\neq\ell}}^{L} \mathbf{S}_{\ell'}(\hat{\boldsymbol{\xi}}_{\ell'}), \tag{10}$$

$$\hat{\mathbf{Q}} = \frac{1}{N}\,(\mathbf{Y}-\mathbf{S}(\hat{\boldsymbol{\xi}}))\,(\mathbf{Y}-\mathbf{S}(\hat{\boldsymbol{\xi}}))^{\mathrm{H}}, \tag{11}$$

and for the maximization step (M-Step) we get

$$\hat{\tau}_\ell = \arg\max_{\tau_\ell}\left\{||\,\hat{\mathbf{Q}}^{-\frac{1}{2}}\,\hat{\mathbf{X}}_\ell\,\mathbf{c}^*(\tau_\ell)||_2^2\right\}, \tag{12}$$

$$\hat{\mathbf{h}}_\ell = \frac{\hat{\mathbf{X}}_\ell\,\mathbf{c}^*(\hat{\tau}_\ell)}{\mathbf{c}^{\mathrm{H}}(\hat{\tau}_\ell)\mathbf{c}(\hat{\tau}_\ell)}. \tag{13}$$

The parameters of each wavefront are estimated sequentially. The E-step and the M-step are performed iteratively for each wavefront until the algorithm converges. Instead of solving a L-dimensional non-linear optimization problem only 1-dimensional optimization procedures need to be solved. Initialization of the SAGE algorithm is carried out by successive interference cancellation starting with $\hat{\boldsymbol{\xi}} = [0, \ldots, 0]^{\mathrm{T}}$ as described in [3]. The initial estimate for $\mathbf{Q}$ is

$$\hat{\mathbf{Q}} = \frac{1}{N}\,\mathbf{Y}\,\mathbf{Y}^{\mathrm{H}}. \tag{14}$$

5. EXTENDED INVARIANCE PRINCIPLE (EXIP)

As a second step we invoke the EXIP [5, 8] in order to refine the ML estimates, $\hat{\boldsymbol{\xi}}$, to achieve the performance using an ML estimate for the structured model, $\hat{\boldsymbol{\theta}}$, for which we would need to solve a $2\cdot L$-dimensional non-linear problem, if solved directly. We briefly recall the EXIP following [5].

Assuming that there exists a function f which is one to one, satisfying

$$\boldsymbol{\xi} = f(\boldsymbol{\theta})\ \in D_{\boldsymbol{\xi}}\quad, \forall_{\boldsymbol{\theta}}\ \in D_{\boldsymbol{\theta}}, \tag{15}$$

and

$$\lim_{N\to\infty}\hat{\boldsymbol{\xi}} = \lim_{N\to\infty} f(\hat{\boldsymbol{\theta}}). \tag{16}$$

Then

$$\hat{\hat{\boldsymbol{\theta}}} = \arg\min_{\boldsymbol{\theta}}\left[\hat{\boldsymbol{\xi}} - f(\boldsymbol{\theta})\right]^{\mathrm{T}}\ \mathbf{W}\ \left[\hat{\boldsymbol{\xi}} - f(\boldsymbol{\theta})\right], \tag{17}$$

is asymptotically (for large N) equivalent to the structured ML estimate $\hat{\boldsymbol{\theta}}$, where

$$\mathbf{W} = \mathrm{E}\left\{\frac{\partial\Lambda(\boldsymbol{\xi}, \mathbf{Q})}{\partial\boldsymbol{\xi}\ \partial\boldsymbol{\xi}^{\mathrm{T}}}\right\}\Bigg|_{\substack{\boldsymbol{\xi}=\hat{\boldsymbol{\xi}} \\ \mathbf{Q}=\hat{\mathbf{Q}}}}, \tag{18}$$

and $\mathrm{E}\{\cdot\}$ denotes expectation.

see the program

For the problem at hand we can write

$$[\hat{\boldsymbol{\xi}}-f(\boldsymbol{\theta})]=\left[\left[\begin{array}{c}\mathrm{Re}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\\ \mathrm{Im}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\\ \hat{\boldsymbol{\tau}}\end{array}\right]-\left[\begin{array}{c}\boldsymbol{\Psi}(\boldsymbol{\phi})\left[\begin{array}{c}\mathrm{Re}\{\boldsymbol{\gamma}\}\\ \mathrm{Im}\{\boldsymbol{\gamma}\}\end{array}\right]\\ \boldsymbol{\tau}\end{array}\right]\right], \quad (19)$$

where

$$\boldsymbol{\Psi}(\boldsymbol{\phi})=\left[\begin{array}{cc}\mathrm{Re}\{\mathbf{I}_L\,\square\,\mathbf{A}\} & -\mathrm{Im}\{\mathbf{I}_L\,\square\,\mathbf{A}\}\\ \mathrm{Im}\{\mathbf{I}_L\,\square\,\mathbf{A}\} & \mathrm{Re}\{\mathbf{I}_L\,\square\,\mathbf{A}\}\end{array}\right], \quad (20)$$

and, $\square$ denotes the Khatri-Rao product, $\mathbf{I}_L$ is a $L\times L$ identity matrix, and $\mathbf{A}=[\mathbf{a}_1(\phi_1)\ \cdots\ \mathbf{a}_\ell(\phi_\ell)\ \cdots\ \mathbf{a}_L(\phi_L)]\in\mathbb{C}^{M\times L}$. The weighting matrix $\mathbf{W}$ can be expressed as

$$\mathbf{W}=\left[\begin{array}{ccc}\mathrm{Re}\{\mathbf{W}_1\} & -\mathrm{Im}\{\mathbf{W}_1\} & \mathrm{Re}\{\mathbf{W}_3\}\\ \mathrm{Im}\{\mathbf{W}_1\} & \mathrm{Re}\{\mathbf{W}_1\} & \mathrm{Im}\{\mathbf{W}_3\}\\ \mathrm{Re}\{\mathbf{W}_3\}^{\mathrm{T}} & \mathrm{Im}\{\mathbf{W}_3\}^{\mathrm{T}} & \mathrm{Re}\{\mathbf{W}_2\}\end{array}\right], \quad (21)$$

with

$$\mathbf{W}_1 = 2\cdot\left(\mathbf{C}\,\mathbf{C}^{\mathrm{H}}\right)\otimes\hat{\mathbf{Q}}^{-1}, \quad (22)$$

$$\mathbf{W}_2 = 2\cdot(\boldsymbol{\Upsilon}^{\mathrm{H}}\,\boldsymbol{\Upsilon})\odot(\hat{\mathbf{H}}^{\mathrm{H}}\,\hat{\mathbf{Q}}^{-1}\,\hat{\mathbf{H}}), \quad (23)$$

$$\mathbf{W}_3=\left[\begin{array}{c}2\cdot(\hat{\mathbf{Q}}^{-1}\,\hat{\mathbf{H}})\,\mathrm{diag}\{(\mathbf{c}^{\mathrm{H}}(\hat{\tau}_1)\,\boldsymbol{\Upsilon}\}\\ \vdots\\ 2\cdot(\hat{\mathbf{Q}}^{-1}\,\hat{\mathbf{H}})\,\mathrm{diag}\{(\mathbf{c}^{\mathrm{H}}(\hat{\tau}_\ell)\,\boldsymbol{\Upsilon}\}\\ \vdots\\ 2\cdot(\hat{\mathbf{Q}}^{-1}\,\hat{\mathbf{H}})\,\mathrm{diag}\{(\mathbf{c}^{\mathrm{H}}(\hat{\tau}_L)\,\boldsymbol{\Upsilon}\}\end{array}\right]. \quad (24)$$

Here, $\otimes$ denotes the Kronecker product, $\mathrm{diag}\{\cdot\}$ defines a diagonal matrix, and

$$\boldsymbol{\Upsilon}=\left[\frac{\partial\mathbf{c}(\hat{\tau}_1)}{\partial\hat{\tau}_1}\cdots\frac{\partial\mathbf{c}(\hat{\tau}_\ell)}{\partial\hat{\tau}_\ell}\cdots\frac{\partial\mathbf{c}(\hat{\tau}_L)}{\partial\hat{\tau}_L}\right]. \quad (25)$$

Thus, introducing (19) in (17) and then minimizing (17) with respect to $\boldsymbol{\tau}$, for fixed $\boldsymbol{\gamma}$ and $\boldsymbol{\phi}$ leads to the refined time-delay estimates

$$\hat{\hat{\boldsymbol{\tau}}}=\hat{\boldsymbol{\tau}}+\mathrm{Re}\{\mathbf{W}_2\}^{-1}\left[\mathrm{Re}\{\mathbf{W}_3\}^{\mathrm{T}}\,\mathrm{Im}\{\mathbf{W}_3\}^{\mathrm{T}}\right]\cdot\left[\left[\begin{array}{c}\mathrm{Re}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\\ \mathrm{Im}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\end{array}\right]-\boldsymbol{\Psi}(\boldsymbol{\phi})\left[\begin{array}{c}\mathrm{Re}\{\boldsymbol{\gamma}\}\\ \mathrm{Im}\{\boldsymbol{\gamma}\}\end{array}\right]\right] \quad (26)$$

Note that the second term in (26) acts like a correction of the unstructured ML estimates $\hat{\boldsymbol{\tau}}$.

Replacing (26) in (19) and then minimizing (17) with respect to $\boldsymbol{\gamma}$ leads to

$$\left[\begin{array}{c}\mathrm{Re}\{\hat{\boldsymbol{\gamma}}\}\\ \mathrm{Im}\{\hat{\boldsymbol{\gamma}}\}\end{array}\right]=\left(\boldsymbol{\Psi}^{\mathrm{T}}(\boldsymbol{\phi})\,\boldsymbol{\Omega}\,\boldsymbol{\Psi}(\boldsymbol{\phi})\right)^{-1}\cdot\boldsymbol{\Psi}^{\mathrm{T}}(\boldsymbol{\phi})\,\boldsymbol{\Omega}\left[\begin{array}{c}\mathrm{Re}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\\ \mathrm{Im}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\end{array}\right], \quad (27)$$

where $\boldsymbol{\Omega}$ denotes the Schur complement of block $\mathrm{Re}\{\mathbf{W}_2\}$ in matrix $\mathbf{W}$, namely

$$\boldsymbol{\Omega}=\left[\left[\begin{array}{cc}\mathrm{Re}\{\mathbf{W}_1\} & -\mathrm{Im}\{\mathbf{W}_1\}\\ \mathrm{Im}\{\mathbf{W}_1\} & \mathrm{Re}\{\mathbf{W}_1\}\end{array}\right]-\left[\begin{array}{c}\mathrm{Re}\{\mathbf{W}_3\}\\ \mathrm{Im}\{\mathbf{W}_3\}\end{array}\right]\mathrm{Re}\{\mathbf{W}_2\}^{-1}\left[\mathrm{Re}\{\mathbf{W}_3\}^{\mathrm{T}}\,\mathrm{Im}\{\mathbf{W}_3\}^{\mathrm{T}}\right]\right]. \quad (28)$$

Finally, $\hat{\hat{\boldsymbol{\phi}}}$ can be obtained by substituting $\boldsymbol{\tau}$ and $\boldsymbol{\gamma}$ in (17) for (26) and (27)

$$\hat{\hat{\boldsymbol{\phi}}}=\arg\min_{\boldsymbol{\phi}}\left\{\left[\begin{array}{c}\mathrm{Re}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\\ \mathrm{Im}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\end{array}\right]^{\mathrm{T}}\left(\boldsymbol{\Omega}-\boldsymbol{\Omega}\,\boldsymbol{\Psi}(\boldsymbol{\phi})\cdot\left(\boldsymbol{\Psi}^{\mathrm{T}}(\boldsymbol{\phi})\,\boldsymbol{\Omega}\,\boldsymbol{\Psi}(\boldsymbol{\phi})\right)^{-1}\boldsymbol{\Psi}^{\mathrm{T}}(\boldsymbol{\phi})\,\boldsymbol{\Omega}\right)\left[\begin{array}{c}\mathrm{Re}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\\ \mathrm{Im}\{\mathrm{vec}\{\hat{\mathbf{H}}\}\}\end{array}\right]\right\}. \quad (29)$$

Hence, (29) provides estimates of the DOAs that are asymptotically equivalent to the ML estimates of the structured model. The estimate $\hat{\hat{\boldsymbol{\phi}}}$ is introduced in (27) and (26) improving the delay estimates, thus deriving $\hat{\hat{\boldsymbol{\tau}}}$.

In the following we first derive a complex formulation of (29) and then we develop an IQML approach [7] in order to obtain low complexity ML DOA estimates. The problem given in (29) can be transformed to a complex formulation using the unitary matrices

$$\mathbf{T}_1=\frac{1}{\sqrt{2}}\left[\begin{array}{cc}\mathbf{I}_{LM} & \mathrm{j}\,\mathbf{I}_{LM}\\ \mathbf{I}_{LM} & -\mathrm{j}\,\mathbf{I}_{LM}\end{array}\right], \quad (30)$$

and

$$\mathbf{T}_2=\frac{1}{\sqrt{2}}\left[\begin{array}{cc}\mathbf{I}_{L} & \mathrm{j}\,\mathbf{I}_{L}\\ \mathbf{I}_{L} & -\mathrm{j}\,\mathbf{I}_{L}\end{array}\right]. \quad (31)$$

We define the complex matrices $\tilde{\boldsymbol{\Omega}}=\mathbf{T}_1\,\boldsymbol{\Omega}\,\mathbf{T}_1^{\mathrm{H}}$ with $\tilde{\boldsymbol{\Omega}}\in\mathbb{C}^{2LM\times 2LM}$ and $\tilde{\boldsymbol{\Psi}}(\phi)=\mathbf{T}_1\,\boldsymbol{\Psi}(\phi)\,\mathbf{T}_2^{\mathrm{H}}$ with $\tilde{\boldsymbol{\Psi}}(\phi)\in\mathbb{C}^{2LM\times 2L}$. Thus, (29) can be reformulated as

$$\hat{\hat{\boldsymbol{\phi}}}=\arg\min_{\boldsymbol{\phi}}\left\{\mathrm{tr}\left(\mathbf{P}^{\perp}_{\tilde{\boldsymbol{\Omega}}^{\frac{1}{2}}\tilde{\boldsymbol{\Psi}}(\boldsymbol{\phi})}\,\tilde{\boldsymbol{\Omega}}^{\frac{1}{2}}\left[\begin{array}{c}\mathrm{vec}\{\hat{\mathbf{H}}\}\\ \mathrm{vec}\{\hat{\mathbf{H}}^*\}\end{array}\right]\left[\begin{array}{c}\mathrm{vec}\{\hat{\mathbf{H}}\}\\ \mathrm{vec}\{\hat{\mathbf{H}}^*\}\end{array}\right]^{\mathrm{H}}\tilde{\boldsymbol{\Omega}}^{\frac{1}{2}}\right)\right\}, \quad (32)$$

whereas

$$\mathbf{P}^{\perp}_{\tilde{\boldsymbol{\Omega}}^{\frac{1}{2}}\tilde{\boldsymbol{\Psi}}(\boldsymbol{\phi})}=\mathbf{I}_{2LM}-\tilde{\boldsymbol{\Omega}}^{\frac{1}{2}}\,\tilde{\boldsymbol{\Psi}}(\boldsymbol{\phi})\left(\tilde{\boldsymbol{\Psi}}^{\mathrm{H}}(\boldsymbol{\phi})\,\tilde{\boldsymbol{\Omega}}\,\tilde{\boldsymbol{\Psi}}(\boldsymbol{\phi})\right)^{-1}\tilde{\boldsymbol{\Psi}}^{\mathrm{H}}(\boldsymbol{\phi})\tilde{\boldsymbol{\Omega}}^{\frac{1}{2}}. \quad (33)$$

In case of an uniform linear array (ULA), where the array steering matrix $\mathbf{A}$ has a Vandermonde structure we can solve (32) applying a polynomial reparametrization approach [7]. Therefore, we define

$$\boldsymbol{\Xi}=\left[\begin{array}{cccccc}\mathbf{B}_1^{\mathrm{H}} & & & & & \\ & \ddots & & & & \\ & & \mathbf{B}_L^{\mathrm{H}} & & & \\ & & & \mathbf{B}_1^{\mathrm{T}} & & \\ & & & & \ddots & \\ & & & & & \mathbf{B}_L^{\mathrm{T}}\end{array}\right]\in\mathbb{C}^{2L(M-1)\times 2LM}, \quad (34)$$

see the program

such that

$$\mathbf{\Xi}^{\mathrm{H}}\, \tilde{\mathbf{\Psi}}(\boldsymbol{\phi}) = \mathbf{0}, \tag{35}$$

whereas

$$\mathbf{B}_\ell^{\mathrm{H}} = \begin{bmatrix} b_{\ell,1} & b_{\ell,0} & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & b_{\ell,1} & b_{\ell,0} \end{bmatrix} \in \mathbb{C}^{(M-1)\times M}, \tag{36}$$

and

$$\mathbf{B}_\ell^{\mathrm{H}} \mathbf{a}_\ell(\phi_\ell) = \mathbf{0}. \tag{37}$$

Here, $\mathbf{b}_\ell = [b_{\ell,0}, b_{\ell,1}]^{\mathrm{T}} \in \mathbb{C}^{2\times 1}$ with $\ell = 1,\dots,L$ comprises the coefficients of a first order polynomial. Due to the very special structure of $\tilde{\mathbf{\Psi}}(\boldsymbol{\phi})$ given in (20), we use a first order polynomial for each of the L wavefronts for reparametrization. Since $\mathbf{\Xi}^{\mathrm{H}}\, \tilde{\mathbf{\Omega}}^{-\frac{\mathrm{H}}{2}} \in \mathbb{C}^{2LM\times 2L(M-1)}$ has rank $2L(M-1)$ and $\tilde{\mathbf{\Omega}}^{\frac{1}{2}}\, \tilde{\mathbf{\Psi}}(\boldsymbol{\phi}) \in \mathbb{C}^{2LM\times 2L}$ has rank $2L$ it follows from (35) that $\mathbf{\Xi}^{\mathrm{H}}\, \tilde{\mathbf{\Omega}}^{-\frac{\mathrm{H}}{2}}$ spans the null space of $\tilde{\mathbf{\Omega}}^{\frac{1}{2}}\, \tilde{\mathbf{\Psi}}(\boldsymbol{\phi})$. Thus,

$$\mathbf{P}^{\perp}_{\tilde{\mathbf{\Omega}}^{\frac{1}{2}}\tilde{\mathbf{\Psi}}(\boldsymbol{\phi})} = \mathbf{P}_{\tilde{\mathbf{\Omega}}^{-\frac{1}{2}}\mathbf{\Xi}} = \tilde{\mathbf{\Omega}}^{-\frac{1}{2}}\, \mathbf{\Xi}\, (\mathbf{\Xi}^{\mathrm{H}}\, \tilde{\mathbf{\Omega}}^{-1}\, \mathbf{\Xi})^{-1} \mathbf{\Xi}^{\mathrm{H}}\, \tilde{\mathbf{\Omega}}^{-\frac{\mathrm{H}}{2}}. \tag{38}$$

Hence, (32) can be rewritten in terms of the coefficients of the L first order polynomials as

$$\hat{\mathbf{b}} = \arg\min_{\mathbf{b}} \operatorname{tr}\left\{ \left(\mathbf{\Xi}^{\mathrm{H}}\, \tilde{\mathbf{\Omega}}^{-1}\, \mathbf{\Xi}\right)^{-1} \mathbf{\Xi}^{\mathrm{H}} \begin{bmatrix} \operatorname{vec}\{\hat{\mathbf{H}}\} \\ \operatorname{vec}\{\hat{\mathbf{H}}^*\} \end{bmatrix} \begin{bmatrix} \operatorname{vec}\{\hat{\mathbf{H}}\} \\ \operatorname{vec}\{\hat{\mathbf{H}}^*\} \end{bmatrix}^{\mathrm{H}} \mathbf{\Xi} \right\} \tag{39}$$

subject to

$$\forall_{\ell=1}^{L} b_{\ell,0} = 1, \tag{40}$$

with

$$\mathbf{b} = [\mathbf{b}_1, \dots, \mathbf{b}_L, \mathbf{b}_1^*, \dots, \mathbf{b}_L^*]^{\mathrm{T}}. \tag{41}$$

This problem finally can be solved with very low complexity by applying an IQML algorithm [7] which results to iteratively solving a quadratic problem with linear constraints. The estimates of the DOAs can derived by $\hat{\hat{\phi}}_\ell = -\arcsin\left(\frac{\arg\{-\hat{b}_{\ell,1}\}}{\pi}\right)$. The IQML algorithm is initialized with $\mathbf{b} = [1, 0, 1, 0, \dots, 1, 0]^{\mathrm{T}}$. Thus, (39) provides estimates of the DOAs that are asymptotically equivalent to the ML estimates of the structured model, and hence are asymptotically efficient. The estimate $\hat{\hat{\boldsymbol{\phi}}}$ is introduced to (27) and (26) improving the delay estimates $\hat{\boldsymbol{\tau}}$.

6. SIMULATION RESULTS

We assume a centro-symmetric [9] uniform linear array (ULA) with $M = 8$ isotropic sensor elements with half-wavelength spacing. The one-sided bandwidth of the signal is $B = 1.023$ MHz. For the PN sequence we apply Gold codes [10] as used for the GPS C/A code with code period $T = 1$ ms, 1023 chips per code period each with a time duration $T_c = 977.52$ ns. Signal-to-noise ratio (SNR) denotes the LOSS-to-noise ratio. The effective SNR in dB can be obtained by

$$\mathrm{SNR} = \mathrm{C/N_0} - 10 \cdot \log_{10}(2\cdot B) + 10 \cdot \log_{10}(N_c), \tag{42}$$

whereas $\mathrm{C/N_0}$ in dB-Hz denotes the carrier-to-noise density ratio and $N_c \in \mathbb{N}$ is the number of code periods within the observation period. We assume $\mathrm{C/N_0}$ = 40.3 dB-Hz and $N_c = 6$ which leads to SNR = -15.03 dB. In this analysis we consider a single reflective multipath as a function of its relative delay to the LOSS ($L = 2$). In the following parameters with the subscript 1 refer to the LOSS and parameters with the subscript 2 refer to the reflection. The reflected multipath and the LOSS are considered to be in-phase, which means $\arg(\gamma_1) = \arg(\gamma_2)$, and the signal-to-multipath ratio (SMR) is 5 dB. Further, we assume a temporally white Gaussian interference with interference-to-signal ratio (ISR) of 40 dB which is responsible for the spatial covariance of the noise plus interference field. The interference is uncorrelated with the signals and the noise. The DOAs for the LOSS and the multipath are $\phi_1 = -30°$ and $\phi_2 = 62°$, and for the interference is $\phi_I = 10°$. Further, we define the relative time-delay between the LOSS and the multipath as $\Delta\tau = |\tau_1 - \tau_2|/T_c$.

In this work a two-step approach is assessed. As a first step the SAGE algorithm is used in order to obtain estimates for the unstructured data model as described in section 4. In the second step EXIP is applied in order to refine the time-delay estimates for the unstructured data model as given in (26) by introducing DOA estimates. The DOA estimates are derived either directly applying EXIP with an IQML algorithm as given in (39) or with decoupled estimators for temporal and spatial parameters applying Unitary ESPRIT as discussed in [6]. However, for the latter approach the pairing between the estimates derived for the unstructured model and the DOA estimates obtained with Unitary ESPRIT additionally has to be estimated [6].

For the SAGE algorithm the quantization precision is chosen to 0.1955 ns ($0.0002 \cdot T_c$) for the time-delay estimates $\hat{\boldsymbol{\tau}}$ and a maximum of 10 iteration cycles is set. Determination of the parameter vector $\hat{\boldsymbol{\xi}}_\ell$ which refers to the LOSS is achieved by evaluating the detection criterion

$$\max_{\ell=1,\dots,L} \frac{||\hat{\mathbf{Q}}^{-\frac{1}{2}}\, \hat{\mathbf{h}}_\ell||_2^2}{M}. \tag{43}$$

In order to derive DOA estimates solving (29) we use the proposed IQML algorithm as given in (39). The convergence condition is satisfied either if $\forall_{\ell=1}^{L} |\hat{\hat{\phi}}_\ell^{(k-1)} - \hat{\hat{\phi}}_\ell^{(k)}| < 0.01°$ or if $k > 50$, whereas $k \in \mathbb{N}$ denotes the k-th iteration of the IQML algorithm. Detection of the parameter vector $\hat{\boldsymbol{\theta}}_\ell$ referring to the LOSS is achieved through (43), as for the estimates for the unstructured model.

Instead of solving (29) we apply the approach as described in [6] using Unitary ESPRIT with structured least squares (SLS) with one iteration [9] and spatial smoothing. The prewhitened channel estimate

$$\hat{\mathbf{R}} = \hat{\mathbf{Q}}^{-\frac{1}{2}}\, \hat{\mathbf{H}}\, (\hat{\mathbf{Q}}^{-\frac{1}{2}}\, \hat{\mathbf{H}})^{\mathrm{H}}, \tag{44}$$

is used to follow the covariance approach as given in [9]. For the spatial smoothing with $L = 2$ we define the following selection matrices for two subarrays with maximum overlap

see the program

[9]:

$$\mathbf{J}_1 = [\mathbf{0}\ \mathbf{I}_{M-1}] \in \mathbb{R}^{M-1\times M},\ \mathbf{J}_2 = [\mathbf{I}_{M-1}\ \ \mathbf{0}] \in \mathbb{R}^{M-1\times M}. \tag{45}$$

Thus, the spatially smoothed covariance matrix is given as

$$\hat{\mathbf{R}}_{ss} = \frac{1}{2}\left(\mathbf{J}_1\,\hat{\mathbf{R}}\,\mathbf{J}_1^{\mathrm{T}} + \mathbf{J}_2\,\hat{\mathbf{R}}\,\mathbf{J}_2^{\mathrm{T}}\right), \tag{46}$$

which is to be used by Unitary ESPRIT with SLS for the DOA estimation. We detect the parameter vector $\hat{\boldsymbol{\theta}}_\ell$ which refers to the LOSS using the detection criterion (43) as for the estimates for the unstructured model and the estimates derived by the IQML algorithm.

In Fig. 1 the probability of miss-detection of the parameter vector $\boldsymbol{\theta}_\ell$ referring to the LOSS is depicted. The decision on the LOSS is taken by evaluating (43). Here, IQML denotes the approach solving (39) and ESPRIT denotes the approach using Unitary ESPRIT for DOA estimation with decoupled estimators for temporal and spatial parameters. Miss-detection of the LOSS for the unstructured model using the SAGE algorithm did not occur.

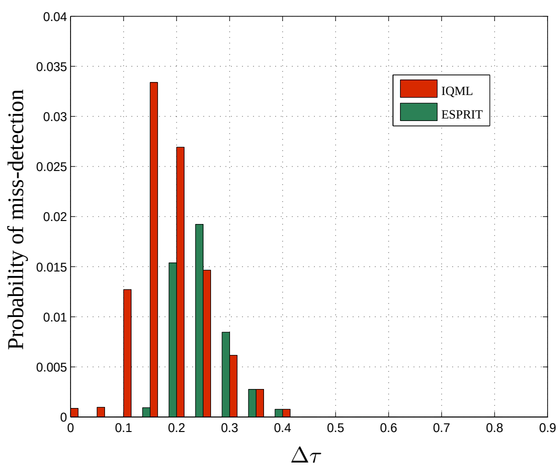


Fig. 1. Probability of miss-detection of LOSS versus $\Delta\tau$

Miss-detection of the LOSS occurs due to several reasons. This includes miss-match of the pairing of the estimates for the unstructured model and the DOA estimates derived by Unitary ESPRIT [6], non-convergence of the IQML algorithm when $k > 50$, and in some rare cases failure of resolving the LOSS. All these cases of miss-detection of the LOSS are excluded from the error statistics which are presented in the following. In a practical implementation miss-detection can be avoided by consistency check of the DOA estimates of the LOSS and for GNSS applications the navigation almanac data can be used to avoid miss-detection. However, simulation results confirmed that miss-detection of the LOSS did not have a significant impact on the time-delay estimates of the LOSS.

In Fig. 2 the root mean square error (RMSE) of $\hat{\tau}_1$ and $\hat{\hat{\tau}}_1$ are depicted. RMSE $\hat{\hat{\tau}}_1$, ESPRIT denotes the approach where the DOA estimates are obtained by Unitary ESPRIT and RMSE $\hat{\hat{\tau}}_1$, IQML denotes the case where the DOA estimates are obtained by solving (39).

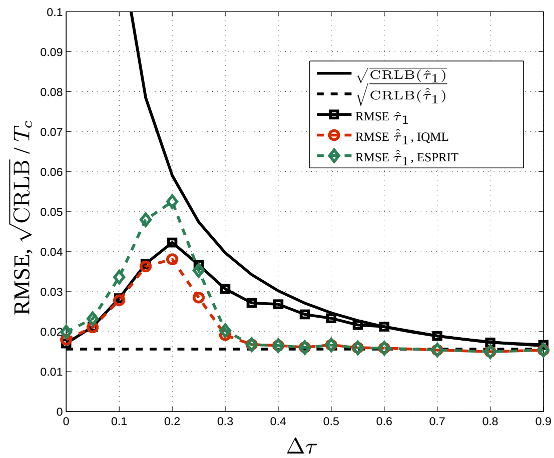


Fig. 2. RMSE of $\hat{\tau}_1$ and $\hat{\hat{\tau}}_1$ versus $\Delta\tau$

In Fig. 2 the SAGE algorithm for small $\Delta\tau$ becomes biased and the RMSE $\hat{\tau}_1$ is below $\sqrt{\mathrm{CRLB}(\hat{\tau}_1)}$, since the interference cancellation in the E-step (10) is not able to separate the two waves if $\Delta\tau$ is very small [4]. In these cases the estimate of one hidden data space, $\hat{\mathbf{X}}_1$ is an estimate for a superposition of the two wavefronts and the other, $\hat{\mathbf{X}}_2$ includes only noise.

In Fig. 3 and Fig. 4 the RMSE $\hat{\hat{\phi}}_1$ and the RMSE $\hat{\hat{\phi}}_2$ for the two methods are depicted.

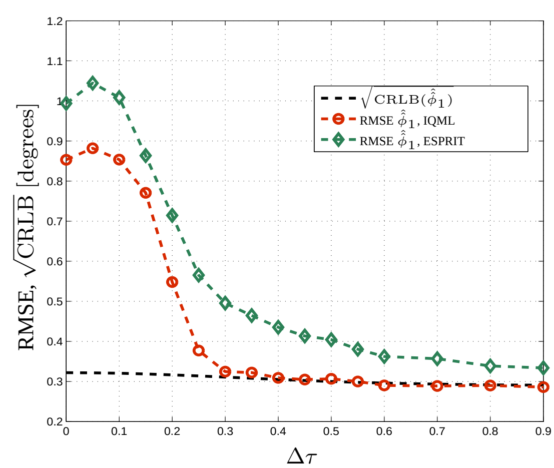


Fig. 3. RMSE of $\hat{\hat{\phi}}_1$ versus $\Delta\tau$

In case we solve (39) with the proposed IQML algorithm, the DOA of the multipath can not be resolved for $\Delta\tau < 0.3$ as depicted in Fig. 4 . This is consistent with the behavior of the SAGE algorithm, which towards small $\Delta\tau$ more and more

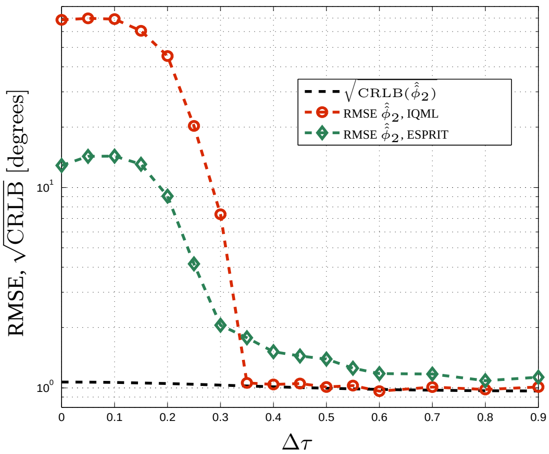


Fig. 4. RMSE of $\hat{\hat{\phi}}_2$ versus $\Delta\tau$

loses its capability of resolving the two paths as described above (cf. Fig. 2). Unitary ESPRIT shows better capability to resolve the DOA of the multipath due to the use of spatial smoothing. However, this does not improve the corresponding RMSE $\hat{\tau}$.

7. CONCLUSIONS

In this work we proposed a two-step approach to achieve estimates for a structured model for the multipath case in an unknown spatially colored field. As a first step the SAGE algorithm is used to obtain estimates for the unstructured data model. In the second step EXIP is applied in order to refine the time-delay estimates for the unstructured data model as given in (26) by introducing DOA estimates. The DOA estimates are derived either directly applying EXIP with an IQML algorithm as given in (39) or with decoupled estimators for temporal and spatial parameters applying Unitary ESPRIT as discussed in [6]. Exemplarily, simulation results for time-delay estimation in GPS are shown which confirm that the proposed two-step approach attains the CRLB of the structured model. The developed low complexity IQML algorithm to solve the DOA estimation problem resulting from EXIP shows better estimation accuracy for the time-delay of the LOSS than the approach using Unitary ESPRIT, however its probability of miss-detection of the LOSS being slightly higher compared to the other approach.

References

[1] A. L. Swindlehurst, "Time Delay and Spatial Signature Estimation Using Known Asynchronous Signals," *IEEE Transactions on Signal Processing*, vol. 46, no. 2, February 1998.

[2] G. Seco, A. L. Swindlehurst, and D. Astely, "Exploiting Antenna Arrays for Synchronization," in *Signal Processing Advances in Wireless and Mobile Communications, Volume 2: Trends in Single- and Multi-User Systems*, G. B. Giannakis, P. Stoica, Y. Hua, and L. Tong, Eds., pp. 403 – 430. Prentice-Hall, 2000.

[3] F. Antreich, J. A. Nossek, and W. Utschick, "Maximum Likelihood Delay Estimation in a Navigation Receiver for Aeronautical Applications," *Aerospace Science and Technology, doi:10.1016/j.ast.2007.06.005*, 2008.

[4] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel Parameter Estimation in Mobile Radio Environments Using the SAGE Algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, March 1999.

[5] A. L. Swindlehurst and P. Stoica, "Maximum Likelihood Methods in Radar Array Signal Processing," *Proceedings of the IEEE*, vol. 86, no. 2, February 1998.

[6] F. Antreich, J. A. Nossek, G. Seco, and L. A. Swindlehurst, "Time delay estimation in a spatially structured model using decoupled estimators for temporal and spatial parameters," in *Proceedings of the International ITG Workshop on Smart Antennas WSA*, Darmstadt, Germany, February 2008.

[7] P. Stoica and K. Sharman, "Maximum Likelihood Methods for Direction-of-Arrival Estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 7, July 1990.

[8] P. Stoica and T. Söderström, "On Reparametrization of Loss Functions Used in Estimation and the Invariance Principle," *Signal Processing*, vol. 17, pp. 383-387, August 1989.

[9] M. Haardt, *Efficient One-, Two-, and Multidimensional High-Resolution Array Signal Processing*, Ph.D. thesis, Department Electrical Engineering of Munich University of Technology, 1996.

[10] B. W. Parkinson and J. J. Spilker, *Global Positioning System: Theory and Applications*, vol. 1, Progress in Astronautics and Aeronautics, 1996.

see the program

STRUCTURED LEAST SQUARES (SLS) BASED ENHANCEMENTS OF TENSOR-BASED CHANNEL ESTIMATION (TENCE) FOR TWO-WAY RELAYING WITH MULTIPLE ANTENNAS

Florian Roemer and Martin Haardt

Ilmenau University of Technology, Communications Research Laboratory
P.O. Box 100565, D-98684 Ilmenau, Germany, http://tu-ilmenau.de/crl
florian.roemer@ieee.org, haardt@ieee.org

***Abstract* — In this paper, we develop a novel tensor-based channel estimation algorithm for two-way relaying with amplify and forward (AF) relays. In two-way relaying two terminals transmit simultaneously to one relay station which then amplifies the received signal and transmits it back to the terminals. With sufficient channel knowledge the terminals can subtract the interference their own transmissions have caused to their received signal and subsequently decode the data from the other terminal. While this relaying scheme uses the radio resources in a particularly efficient way, reliable channel knowledge is crucial to obtain an acceptable link quality.**

It has been shown that we can estimate all relevant channel parameters via the purely algebraic TENCE algorithm. In this paper, we introduce a novel tensor-based approach to enhance these estimates even further via an iterative procedure based on Structured Least Squares (SLS).

The improvement in channel estimation accuracy is demonstrated via computer simulations at the end of the paper. We also show that the number of required iterations is only between one and four, even in critical scenarios.

1. INTRODUCTION

A crucial aspect in designing future mobile communication systems is to enhance transmission rates, reliability, and coverage of the mobile radio access. A promising system concept to meet these demands is the deployment of intermediate relay stations that support the transmissions between communication partners in the radio network.

Relay stations can be divided into two categories: Regenerative or decode-and-forward (DF) relays which decode the received transmissions and reencode them for the second hop and non-regenerative or amplify-and-forward (AF) relays which retransmit an amplified version of their received signal without decoding the individual transmissions. We focus on AF relays since their hardware complexity is significantly lower than for DF relays which facilitates the mass deployment of simple and cheap devices.

A large variety of relaying schemes has been studied in the literature. Two-way relaying is one particular scheme that uses the radio resources very efficiently [7]. The basic two-way relaying scenario which we consider in this paper consists of two user terminals equipped with M_1 and M_2 antennas that want to exchange data and one AF relay equipped with M_{R} antennas that assists this communication link as depicted in Figure 1. The transmission is performed in two subsequent time slots in a TDD fashion: In the first time slot both terminals transmit their data to the relay which receives the superimposed transmissions. In the second time slot, the relay sends an amplified version of the received signal back to both terminals. If the terminals possess channel knowledge, they can subtract the interference that their own transmission has caused from their received signal and then decode the data from the other communication partner.

Two-way relaying has been studied in many recent publications. However, the channel estimation is often ignored by considering perfect channel state information (CSI) [5], [11]. Moreover, the authors usually assume that the relay station actively suppresses interference. This can be achieved by using decode and forward (DF) relays [6] or by exploiting channel state information at the AF relays [10]. In contrast to this we focus on a two-way relaying scenario where the relay station does not need any channel state information.

In [9] we have introduced the TENCE algorithm which is a tensor-based algebraic channel estimation technique that provides both terminals with knowledge of all relevant channel parameters and is applicable to arbitrary antenna configurations. In this contribution we demonstrate that the initial estimate obtained by TENCE in a closed-form can be further improved via a tensor-based iterative algorithm.

The derivation of the algorithm is inspired by Structured Least Squares (SLS) [3], which is an efficient technique to solve the overdetermined shift invariance equations that appear in ESPRIT-based high-resolution parameter estimation algorithms. Since SLS exploits the structure inherent in the invariance equations, it outperforms alternative schemes such as Total Least Squares (TLS) [2, 12]. The SLS method has been extended for the solution of the tensor shift invariance equations which appear in Tensor-ESPRIT type algorithms [4]. This extended method is known as Tensor-Structure SLS (TS-SLS) [8].

As we show in this paper, we can enhance TENCE by exploiting an additional symmetry in the data model in an iterative fashion inspired by the ideas used in SLS and TS-SLS. Certainly, this iterative procedure results in an increase in the computational complexity as compared to the closed-form TENCE. However, we demonstrate by simulations that between one and four iterations are sufficient since no additional improvement is achieved for more iterations, even in critical scenarios.

2. NOTATION

To facilitate the distinction between scalars, vectors, matrices, and tensors, the following notation is used throughout the paper: Scalars

see the program

are denoted as italic letters ($a, b, \ldots, A, B, \ldots$), vectors as lower-case bold-faced letters ($\boldsymbol{a}, \boldsymbol{b}, \ldots$), matrices are written as upper-case bold-faced letters ($\boldsymbol{A}, \boldsymbol{B}$), and tensors as bold-faced calligraphic letters ($\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}$).

The superscripts $^{\mathrm{T}}, ^{\mathrm{H}}, ^{+}, ^{*}$ represent matrix transposition, Hermitian transposition, the Moore-Penrose pseudo inverse, and complex conjugation, respectively. The Kronecker product between two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is denoted by $\boldsymbol{A} \otimes \boldsymbol{B}$. Likewise, $\boldsymbol{A} \diamond \boldsymbol{B}$ symbolizes the Khatri-Rao (columnwise Kronecker) product and $\boldsymbol{A} \odot \boldsymbol{B}$ the Schur (elementwise) product.

A three-dimensional tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$ is a three-way array with size M_r along mode (or dimension) r for $r = 1, 2, 3$. The n-mode vectors of $\boldsymbol{\mathcal{A}}$ are obtained by stacking the elements of $\boldsymbol{\mathcal{A}}$ into a vector where the n-th index varies in its range and all the other indices are kept fixed. Collecting all the n-mode vectors into a matrix we obtain the n-mode unfolding of $\boldsymbol{\mathcal{A}}$, which is represented by $[\boldsymbol{\mathcal{A}}]_{(n)} \in \mathbb{C}^{M_n \times M_1 \cdot M_2 \cdot M_3 / M_n}$. The order of the columns in $[\boldsymbol{\mathcal{A}}]_{(n)}$ is chosen in accordance with [1].

The n-mode product between a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$ and a matrix $\boldsymbol{U} \in \mathbb{C}^{P_n \times M_n}$ is denoted as $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_n \boldsymbol{U}_n$ and computed by multiplying all n-mode vectors of $\boldsymbol{\mathcal{A}}$ from the left-hand side by the matrix $\boldsymbol{U}_n$, i.e., $[\boldsymbol{\mathcal{B}}]_{(n)} = \boldsymbol{U}_n \cdot [\boldsymbol{\mathcal{A}}]_{(n)}$.

To concatenate two tensors along the n-th mode we use the symbol $[\boldsymbol{\mathcal{A}} \sqcup_n \boldsymbol{\mathcal{B}}]$.

The vectorization operators $\mathrm{vec}\{.\}$ aligns all the elements of a matrix or a tensor into a vector, such that the row index is varied first, then the column index, and then the third index. For arbitrary three-dimensional tensors $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I \times J \times K}$, permutation matrices $\boldsymbol{P}^{(n)}_{I,J,K}$ of size $I \cdot J \cdot K \times I \cdot J \cdot K$ are uniquely defined via the following property [8]

$$\boldsymbol{P}^{(n)}_{I,J,K} \cdot \mathrm{vec}\left\{[\boldsymbol{\mathcal{X}}]_{(n)}\right\} = \mathrm{vec}\left\{\boldsymbol{\mathcal{X}}\right\}, \; n = 1, 2, 3, \forall\, \boldsymbol{\mathcal{X}}. \tag{1}$$

The higher-order norm of a tensor $\boldsymbol{\mathcal{A}}$ is written as $\|\boldsymbol{\mathcal{A}}\|_{\mathrm{H}}$. It represents a generalization of the Frobenius norm of matrices expressed as $\|\boldsymbol{A}\|_{\mathrm{F}}$ in the sense that both are computed via the square root of the sum of the squared magnitudes of all elements. It is easy to see that for a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I \times J \times K}$ and for a matrix $\boldsymbol{A} \in \mathbb{C}^{I \times J}$, the norms satisfy the following identities

$$\|\boldsymbol{\mathcal{A}}\|_{\mathrm{H}}^2 = \|\mathrm{vec}\{\boldsymbol{\mathcal{A}}\}\|_2^2 = \left\|[\boldsymbol{\mathcal{A}}]_{(n)}\right\|_{\mathrm{F}}^2, \; n = 1, 2, 3 \tag{2}$$

$$\|\boldsymbol{A}\|_{\mathrm{F}}^2 = \|\mathrm{vec}\{\boldsymbol{A}\}\|_2^2, \tag{3}$$

where $\|\boldsymbol{x}\|_2$ is the Euclidean norm (2-norm) of the vector $\boldsymbol{x}$.

Finally, $\boldsymbol{0}_{p \times q}$ indicates a zero matrix of size $p \times q$, $\boldsymbol{I}_p$ is the identity matrix of size $p \times p$, and $\boldsymbol{\mathcal{I}}_{3,r}$ is the $r \times r \times r$ identity tensor, which is equal to one if all three indices are equal and zero otherwise.

3. SYSTEM DESCRIPTION

The system under investigation is depicted in Figure 1. Two terminals equipped with M_1 and M_2 antennas, respectively, exchange data. This transmission is supported by an intermediate relay station having M_{R} antennas. In the first transmission phase the terminals transmit simultaneously to the relay. The relay receives the superimposed transmissions, amplifies the received signal vector by a complex matrix $\boldsymbol{G} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{R}}}$, and transmits the amplified signal back to the terminals in the second time slot in a TDD fashion.

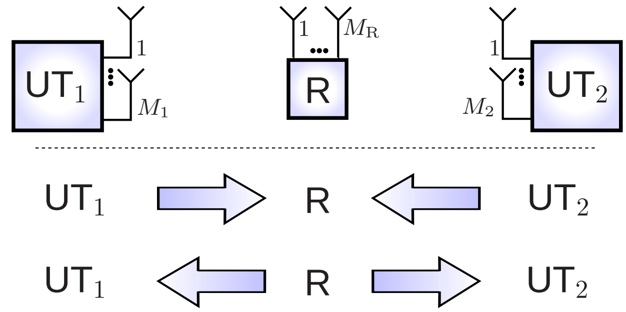


Fig. 1. Two-way relaying system model.

Assuming frequency-flat fading, the signal received at the relay can be expressed as

$$\boldsymbol{r} = \boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \boldsymbol{n}_{\mathrm{R}}, \tag{4}$$

where $\boldsymbol{x}_1 \in \mathbb{C}^{M_1}$ and $\boldsymbol{x}_2 \in \mathbb{C}^{M_2}$ are the transmitted vectors from terminal 1 and terminal 2, the matrices $\boldsymbol{H}_1 \in \mathbb{C}^{M_{\mathrm{R}} \times M_1}$ and $\boldsymbol{H}_2 \in \mathbb{C}^{M_{\mathrm{R}} \times M_2}$ represent the MIMO channels between the relay and the two terminals, and the vector $\boldsymbol{n}_{\mathrm{R}}$ is the additive noise vector at the relay station.

In the second time slot, the relay amplifies the vector $\boldsymbol{r}$ by a complex matrix $\boldsymbol{G} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{R}}}$ and transmits the amplified signal back to the two terminals. The received signals therefore obey the following model

$$\boldsymbol{y}_1 = \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot (\boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \boldsymbol{n}_{\mathrm{R}}) + \boldsymbol{n}_1 \tag{5}$$

$$\boldsymbol{y}_2 = \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot (\boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \boldsymbol{n}_{\mathrm{R}}) + \boldsymbol{n}_2, \tag{6}$$

where we have assumed that reciprocity holds and the channels have not changed between the two transmission phases. The vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ represent the additive noise at the first and second terminal, respectively. Note that (5) and (6) can also be expressed in the following fashion

$$\boldsymbol{y}_1 = \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \tilde{\boldsymbol{n}}_1 \tag{7}$$

$$\boldsymbol{y}_2 = \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \tilde{\boldsymbol{n}}_2. \tag{8}$$

Here $\tilde{\boldsymbol{n}}_1$ and $\tilde{\boldsymbol{n}}_2$ represent the effective noise terms which are given by

$$\tilde{\boldsymbol{n}}_1 = \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{n}_{\mathrm{R}} + \boldsymbol{n}_1 \tag{9}$$

$$\tilde{\boldsymbol{n}}_2 = \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{n}_{\mathrm{R}} + \boldsymbol{n}_2. \tag{10}$$

We conclude from (7) and (8) that each terminal receives the data from the other terminal via the effective channels $\boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2$ and $\boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1$, respectively. Moreover, these transmissions are superimposed by interference caused via the terminals' own transmitted data. However, since this data is known, the terminals can subtract this self-interference term if they have knowledge of the channel matrices $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$. Therefore we now focus on the acquisition of reliable channel state information at both terminals.

In order to facilitate the channel estimation we require a training phase which consists of M_{R} frames. In each frame the terminals transmit known pilot sequences of length N_{P} described by $\boldsymbol{X}_1 = [\boldsymbol{x}_{1,1}, \ldots, \boldsymbol{x}_{1,N_{\mathrm{P}}}] \in \mathbb{C}^{M_1 \times N_{\mathrm{P}}}$ and $\boldsymbol{X}_2 = [\boldsymbol{x}_{2,1}, \ldots, \boldsymbol{x}_{2,N_{\mathrm{P}}}] \in \mathbb{C}^{M_2 \times N_{\mathrm{P}}}$. Moreover, for the i-th frame, the relay amplification matrix $\boldsymbol{G}^{(i)} \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{R}}}$ is used. Since the relay is assumed to have no

see the program

channel state information, all matrices $\boldsymbol{G}^{(i)}$ can be designed beforehand and are therefore known to both terminals.

Consequently, the signal received by the user terminals in the i-th frame from the transmission of the j-th pilot can be expressed as

$$\begin{aligned}\boldsymbol{y}_{1,i,j} &= \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G}^{(i)} \cdot \boldsymbol{H}_1 \cdot \boldsymbol{x}_{1,j} + \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G}^{(i)} \cdot \boldsymbol{H}_2 \cdot \boldsymbol{x}_{2,j} + \tilde{\boldsymbol{n}}_{1,i,j} \\ \boldsymbol{y}_{2,i,j} &= \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G}^{(i)} \cdot \boldsymbol{H}_1 \cdot \boldsymbol{x}_{1,j} + \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G}^{(i)} \cdot \boldsymbol{H}_2 \cdot \boldsymbol{x}_{2,j} + \tilde{\boldsymbol{n}}_{2,i,j}.\end{aligned} \tag{11}$$

Let us introduce the following definitions

$$\boldsymbol{H} = [\boldsymbol{H}_1,\ \boldsymbol{H}_2] \in \mathbb{C}^{M_{\mathrm{R}} \times (M_1+M_2)} \tag{12}$$

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} \in \mathbb{C}^{(M_1+M_2)\times N_{\mathrm{P}}} \tag{13}$$

$$\boldsymbol{\mathcal{G}} = \left[\boldsymbol{G}^{(1)} \sqcup_3 \dots \sqcup_3 \boldsymbol{G}^{(M_{\mathrm{R}})}\right] \in \mathbb{C}^{M_{\mathrm{R}}\times M_{\mathrm{R}}\times M_{\mathrm{R}}}, \tag{14}$$

where $\sqcup_n$ represents the concatenation operator along the n-th dimension. Using these definitions we can express (11) in tensor form

$$\boldsymbol{\mathcal{Y}}_1 = \boldsymbol{\mathcal{G}} \times_1 \boldsymbol{H}_1^{\mathrm{T}} \times_2 (\boldsymbol{H}\cdot\boldsymbol{X})^{\mathrm{T}} + \boldsymbol{\mathcal{N}}_1 \in \mathbb{C}^{M_1\times N_{\mathrm{P}}\times M_{\mathrm{R}}} \tag{15}$$

$$\boldsymbol{\mathcal{Y}}_2 = \boldsymbol{\mathcal{G}} \times_1 \boldsymbol{H}_2^{\mathrm{T}} \times_2 (\boldsymbol{H}\cdot\boldsymbol{X})^{\mathrm{T}} + \boldsymbol{\mathcal{N}}_2 \in \mathbb{C}^{M_2\times N_{\mathrm{P}}\times M_{\mathrm{R}}}, \tag{16}$$

where $\times_n$ represents the n-mode product between a tensor and a matrix [1].

4. SLS-BASED CHANNEL ESTIMATION

In [9] we have shown that each terminal can estimate both channel matrices $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ in closed-form from its own received training data with the help of the TENCE algorithm. Via TENCE we are able to resolve all scaling ambiguities except for one sign ambiguity, i.e., instead of $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ we might estimate $-\boldsymbol{H}_1$ and $-\boldsymbol{H}_2$. However, since the transmissions take place via the effective channels $\boldsymbol{H}_i^{\mathrm{T}}\cdot\boldsymbol{G}\cdot\boldsymbol{H}_j,\ i,j=1,2$, this sign ambiguity is canceled and therefore does not have any impact on the transmission.

In the presented SLS-based improvement of TENCE we can additionally exploit the fact that in (15) and (16) the channel matrix that appears in the first factor is also present in the second factor and therefore improve the TENCE solution. Here, we show the solution for terminal 1. Due to the symmetry of the data model in (15) and (16) the solution for terminal 2 is similar.

Let $\hat{\boldsymbol{H}}_1$ and $\hat{\boldsymbol{H}}_2$ be the initial channel estimates for $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$, respectively, and $\hat{\boldsymbol{H}} = [\hat{\boldsymbol{H}}_1,\ \hat{\boldsymbol{H}}_2]$. Moreover, define $\tilde{\boldsymbol{\mathcal{Y}}}_1 = \boldsymbol{\mathcal{Y}}_1 \times_2 \left(\boldsymbol{X}^{\mathrm{T}}\right)^{+}$, where the superscript $^{+}$ represents the Moore-Penrose pseudo inverse. In the absence of noise, $\tilde{\boldsymbol{\mathcal{Y}}}_1 = \boldsymbol{\mathcal{G}}\times_1\hat{\boldsymbol{H}}_1^{\mathrm{T}}\times_2\hat{\boldsymbol{H}}^{\mathrm{T}}$. Therefore, an indicator for the channel estimation accuracy is the norm of the tensor $\tilde{\boldsymbol{\mathcal{Y}}}_1-(\boldsymbol{\mathcal{G}}\times_1\hat{\boldsymbol{H}}_1^{\mathrm{T}}\times_2\hat{\boldsymbol{H}}^{\mathrm{T}})$.

Finding the channel matrices $\hat{\boldsymbol{H}}_1$ and $\hat{\boldsymbol{H}}_2$ that minimize this norm requires solving a non-linear optimization problem which is not guaranteed to have a unique solution. However, we can take advantage of the fact that we already have an initial estimate for the channel matrices which is obtained via TENCE. We therefore introduce update terms $\Delta\boldsymbol{H}_1$ and $\Delta\boldsymbol{H}_2$ for the estimates $\hat{\boldsymbol{H}}_1$ and $\hat{\boldsymbol{H}}_2$, similarly to the SLS algorithm [3]. By including penalty terms for the norm of the updates we confine the final solution to be close to the original estimate obtained via TENCE. This can be seen as a form of regularization which enhances the numerical stability.

The cost function that we minimize in an iterative fashion can be expressed in the following way

$$\begin{aligned}J(\Delta\boldsymbol{H}_k) = &\left\|\tilde{\boldsymbol{\mathcal{Y}}}_1 - \boldsymbol{\mathcal{G}}\times_1\left(\hat{\boldsymbol{H}}_1+\Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}}\times_2\left(\hat{\boldsymbol{H}}+\Delta\boldsymbol{H}_k\right)^{\mathrm{T}}\right\|_{\mathrm{H}}^2 \\ &+\kappa_1^2\left\|\Delta\boldsymbol{H}_{1,k}\right\|_{\mathrm{F}}^2+\kappa_2^2\left\|\Delta\boldsymbol{H}_{2,k}\right\|_{\mathrm{F}}^2,\end{aligned} \tag{17}$$

where $\Delta\boldsymbol{H}_{1,k}$ and $\Delta\boldsymbol{H}_{2,k}$ represent the update terms for $\hat{\boldsymbol{H}}_1$ and $\hat{\boldsymbol{H}}_2$ after the k-th iteration and $\Delta\boldsymbol{H}_k = [\Delta\boldsymbol{H}_{1,k},\ \Delta\boldsymbol{H}_{2,k}]$. Moreover, $\kappa_1^2 = M_1/\alpha$ and $\kappa_2^2 = M_2/\alpha$. Here $\alpha\in\mathbb{R},\alpha>0$ is a design parameter that allows to controls the amount of regularization used: The larger α is chosen, the less weight is given to the penalty terms and therefore less regularization is included.

The solution of the non-linear least squares problem (17) is established in an iterative fashion by local linearization inspired by the derivations of SLS in [3] and TS-SLS in [8]. To this end, introduce the residual tensor $\boldsymbol{\mathcal{R}}_k$ after the k-th iteration as

$$\boldsymbol{\mathcal{R}}_k = \tilde{\boldsymbol{\mathcal{Y}}}_1 - \boldsymbol{\mathcal{G}}\times_1\left(\hat{\boldsymbol{H}}_1+\Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}}\times_2\left(\hat{\boldsymbol{H}}+\Delta\boldsymbol{H}_k\right)^{\mathrm{T}}. \tag{18}$$

Using this definition and the identities (2) and (3), we can express (17) in the following form

$$J(\Delta\boldsymbol{H}_k) = \left\|\begin{bmatrix}\operatorname{vec}\{\boldsymbol{\mathcal{R}}_k\} \\ \kappa_1\cdot\operatorname{vec}\{\Delta\boldsymbol{H}_{1,k}\} \\ \kappa_2\cdot\operatorname{vec}\{\Delta\boldsymbol{H}_{2,k}\}\end{bmatrix}\right\|_2^2. \tag{19}$$

The matrices $\Delta\boldsymbol{H}_{1,k}$ and $\Delta\boldsymbol{H}_{2,k}$ are updated in each iteration according to

$$\Delta\boldsymbol{H}_{1,k+1} = \Delta\boldsymbol{H}_{1,k} + \Delta\Delta\boldsymbol{H}_{1,k} \tag{20}$$

$$\Delta\boldsymbol{H}_{2,k+1} = \Delta\boldsymbol{H}_{2,k} + \Delta\Delta\boldsymbol{H}_{2,k}, \tag{21}$$

starting from the initial values

$$\Delta\boldsymbol{H}_{1,k} = \boldsymbol{0}_{M_{\mathrm{R}}\times M_1} \tag{22}$$

$$\Delta\boldsymbol{H}_{2,k} = \boldsymbol{0}_{M_{\mathrm{R}}\times M_2}. \tag{23}$$

Therefore, our goal is to find $\Delta\Delta\boldsymbol{H}_{1,k}$ and $\Delta\Delta\boldsymbol{H}_{2,k}$ that minimize (19) in the k-th iteration, which requires a series of algebraic manipulations. As a first step, we use (20) and (21) in (18) in order to express the residual tensor in the $(k+1)$-th iteration in the following form

$$\begin{aligned}\boldsymbol{\mathcal{R}}_{k+1} = \boldsymbol{\mathcal{R}}_k &- \boldsymbol{\mathcal{G}}\times_1\Delta\Delta\boldsymbol{H}_{1,k}^{\mathrm{T}}\times_2\left(\hat{\boldsymbol{H}}+\Delta\boldsymbol{H}_k\right)^{\mathrm{T}} \\ &- \boldsymbol{\mathcal{G}}\times_1\left(\hat{\boldsymbol{H}}_1+\Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}}\times_2\Delta\Delta\boldsymbol{H}_k^{\mathrm{T}} \\ &- \boldsymbol{\mathcal{G}}\times_1\Delta\Delta\boldsymbol{H}_{1,k}^{\mathrm{T}}\times_2\Delta\Delta\boldsymbol{H}_k^{\mathrm{T}}.\end{aligned} \tag{24}$$

In order to linearize (24), the last term is neglected since it contains the product of two update terms which we assumed to have a small Frobenius norm. The next step is to vectorize (24) and to rearrange the terms applying the permutation matrices defined in (1). We obtain

$$\begin{aligned}&\operatorname{vec}\{\boldsymbol{\mathcal{R}}_{k+1}\} \approx \operatorname{vec}\{\boldsymbol{\mathcal{R}}_k\} \\ &- \boldsymbol{P}^{(3)}_{M_1,M_1+M_2,M_{\mathrm{R}}}\cdot\left(\boldsymbol{I}_{M_1}\otimes\left[\boldsymbol{\mathcal{G}}\times_2\left(\hat{\boldsymbol{H}}+\Delta\boldsymbol{H}_k\right)^{\mathrm{T}}\right]_{(1)}^{\mathrm{T}}\right) \\ &\qquad\cdot\operatorname{vec}\{\Delta\Delta\boldsymbol{H}_{1,k}\} \\ &- \boldsymbol{P}^{(1)}_{M_1,M_1+M_2,M_{\mathrm{R}}}\cdot\left(\boldsymbol{I}_{M_1+M_2}\otimes\left[\boldsymbol{\mathcal{G}}\times_1\left(\hat{\boldsymbol{H}}_1+\Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}}\right]_{(2)}^{\mathrm{T}}\right) \\ &\qquad\cdot\operatorname{vec}\{\Delta\Delta\boldsymbol{H}_k\}.\end{aligned} \tag{25}$$

To separate the terms depending on $\Delta\Delta\boldsymbol{H}_{1,k}$ and $\Delta\Delta\boldsymbol{H}_{2,k}$ we observe that

$$\operatorname{vec}\{\Delta\Delta\boldsymbol{H}_k\} = \operatorname{vec}\left\{\left[\Delta\Delta\boldsymbol{H}_{1,k},\ \Delta\Delta\boldsymbol{H}_{2,k}\right]\right\} \tag{26}$$

$$= \begin{bmatrix}\operatorname{vec}\{\Delta\Delta\boldsymbol{H}_{1,k}\} \\ \operatorname{vec}\{\Delta\Delta\boldsymbol{H}_{2,k}\}\end{bmatrix}, \tag{27}$$

see the program

which follows from the definition of the vec-operator. Inserting (27) into (25) we get

$$\operatorname{vec}\{\mathcal{R}_{k+1}\} \approx \operatorname{vec}\{\mathcal{R}_k\} - \tilde{\boldsymbol{F}}_k^{(1)} \cdot \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{1,k}\} - \tilde{\boldsymbol{F}}_k^{(2)} \cdot \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{2,k}\}, \tag{28}$$

where the matrices $\tilde{\boldsymbol{F}}_k^{(1)}$ and $\tilde{\boldsymbol{F}}_k^{(2)}$ are given by

$$\begin{aligned}
\tilde{\boldsymbol{F}}_k^{(1)} =& \boldsymbol{P}^{(3)}_{M_1,M_1+M_2,M_\mathrm{R}} \cdot \left(\boldsymbol{I}_{M_1} \otimes \left[\boldsymbol{\mathcal{G}} \times_2 \left(\hat{\boldsymbol{H}} + \Delta \boldsymbol{H}_k\right)^\mathrm{T}\right]^\mathrm{T}_{(1)}\right) \\
&+ \boldsymbol{P}^{(1)}_{M_1,M_1+M_2,M_\mathrm{R}} \cdot \left(\boldsymbol{I}_{M_1+M_2} \otimes \left[\boldsymbol{\mathcal{G}} \times_1 \left(\hat{\boldsymbol{H}}_1 + \Delta \boldsymbol{H}_{1,k}\right)^\mathrm{T}\right]^\mathrm{T}_{(2)}\right) \cdot \boldsymbol{J}_1 \\
\tilde{\boldsymbol{F}}_k^{(2)} =& \boldsymbol{P}^{(1)}_{M_1,M_1+M_2,M_\mathrm{R}} \cdot \left(\boldsymbol{I}_{M_1+M_2} \otimes \left[\boldsymbol{\mathcal{G}} \times_1 \left(\hat{\boldsymbol{H}}_1 + \Delta \boldsymbol{H}_{1,k}\right)^\mathrm{T}\right]^\mathrm{T}_{(2)}\right) \cdot \boldsymbol{J}_2 \\
\boldsymbol{J}_1 =& \begin{bmatrix} \boldsymbol{I}_{M_1 \cdot M_\mathrm{R}} \\ \boldsymbol{0}_{M_2 \cdot M_\mathrm{R} \times M_1 \cdot M_\mathrm{R}} \end{bmatrix} \quad \boldsymbol{J}_2 = \begin{bmatrix} \boldsymbol{0}_{M_1 \cdot M_\mathrm{R} \times M_2 \cdot M_\mathrm{R}} \\ \boldsymbol{I}_{M_2 \cdot M_\mathrm{R}} \end{bmatrix}.
\end{aligned}$$

We can rewrite the cost function in (19) by inserting (20), (21), and (28). As a result we have

$$J(\Delta \boldsymbol{H}_{k+1}) = \left\| \begin{bmatrix} \operatorname{vec}\{\mathcal{R}_k\} \\ \kappa_1 \cdot \operatorname{vec}\{\Delta \boldsymbol{H}_{1,k}\} \\ \kappa_2 \cdot \operatorname{vec}\{\Delta \boldsymbol{H}_{2,k}\} \end{bmatrix} + \boldsymbol{F}_k \cdot \begin{bmatrix} \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{1,k}\} \\ \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{2,k}\} \end{bmatrix} \right\|_2^2,$$

where the matrix $\boldsymbol{F}_k$ is given by

$$\boldsymbol{F}_k = \begin{bmatrix} -\tilde{\boldsymbol{F}}_k^{(1)} & -\tilde{\boldsymbol{F}}_k^{(2)} \\ \kappa_1 \cdot \boldsymbol{I}_{M_1 \cdot M_\mathrm{R}} & \boldsymbol{0}_{M_1 \cdot M_\mathrm{R} \times M_2 \cdot M_\mathrm{R}} \\ \boldsymbol{0}_{M_2 \cdot M_\mathrm{R} \times M_1 \cdot M_\mathrm{R}} & \kappa_2 \cdot \boldsymbol{I}_{M_2 \cdot M_\mathrm{R}} \end{bmatrix}. \tag{29}$$

Note that our goal is to find the update terms $\Delta\Delta \boldsymbol{H}_{1,k}$ and $\Delta\Delta \boldsymbol{H}_{2,k}$ that minimize the cost function in the k-th iteration. We have reduced this task to a linear least squares problem, for which the solution is well-known and given by

$$\begin{bmatrix} \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{1,k}\} \\ \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{2,k}\} \end{bmatrix} = -\boldsymbol{F}_k^+ \cdot \begin{bmatrix} \operatorname{vec}\{\mathcal{R}_k\} \\ \kappa_1 \cdot \operatorname{vec}\{\Delta \boldsymbol{H}_{1,k}\} \\ \kappa_2 \cdot \operatorname{vec}\{\Delta \boldsymbol{H}_{2,k}\} \end{bmatrix}. \tag{30}$$

The iteration (30) is repeated until convergence is achieved. We propose the following method to detect convergence: After the k-th iteration is complete, let us define $\rho_k = \|\mathcal{R}_{k-1}\|_\mathrm{H} - \|\mathcal{R}_k\|_\mathrm{H}$ as a measure for how much the fit of the estimates has improved. Then we terminate the iteration if $\rho_k < \delta$ for a predefined threshold δ. Moreover, if ρ_k should become negative, the last iteration is ignored and the estimates from the $(k-1)$-th iteration are used as final solutions instead.

Our simulations have shown that the algorithm is not very sensitive to the choice of the regularization parameter α as long as it is not chosen too small. For $\alpha < 1$ the amount of regularization is too large and the algorithm fails to converge to the desired minimum. Instead, the algorithm is terminated too early very close to the initial solution. For larger values of α the impact of α on the performance is negligible.

In critical scenarios, we found that a value of α in the range of 100 is beneficial to enhance the numerical stability. On the other hand, for less critical scenarios, regularization is not needed and α can be chosen bigger to reduce the number of required iterations. If the SNR is very high we can even set $\alpha = \infty$, i.e., skip the regularization completely. In this case, (30) is replaced by

$$\begin{bmatrix} \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{1,k}\} \\ \operatorname{vec}\{\Delta\Delta \boldsymbol{H}_{2,k}\} \end{bmatrix} = \left[\tilde{\boldsymbol{F}}_k^{(1)},\ \tilde{\boldsymbol{F}}_k^{(2)}\right]^+ \cdot \operatorname{vec}\{\mathcal{R}_k\}. \tag{31}$$

The threshold parameter δ represents a trade off between computational complexity and estimation accuracy. For very small values of δ, the number of iterations is increased. The experience from our simulations has shown that $\delta = 10^{-3}$ is a reasonable value. Smaller values lead to more iterations, however these do not result in a significant improvement in accuracy. Larger values of δ terminate the algorithm too early.

5. SIMULATION RESULTS

In this section we evaluate the achievable channel estimation accuracy of the proposed algorithm. The channel matrices are generated according to the following model

$$\boldsymbol{H}_i = \sqrt{\frac{K_i}{K_i+1}} \cdot \boldsymbol{H}_{\mathrm{LOS},i} + \sqrt{\frac{1}{K_i+1}} \cdot \boldsymbol{H}_{\mathrm{NLOS},i},\ i = 1,2, \tag{32}$$

where $\boldsymbol{H}_{\mathrm{LOS},i} \in \mathbb{C}^{M_\mathrm{R} \times M_i}$ is a rank-1 matrix capturing the line of sight (LOS) component, $\boldsymbol{H}_{\mathrm{NLOS},i} \in \mathbb{C}^{M_\mathrm{R} \times M_i}$ contains the non line of sight (NLOS) component, and K_i is the Rician K-factor. The matrix $\boldsymbol{H}_{\mathrm{NLOS},i}$ consists of zero mean circularly symmetric complex Gaussian random variables. Moreover, a Kronecker spatial correlation model is assumed with factors given by

$$\mathrm{E}\{\boldsymbol{H}_{\mathrm{NLOS},i} \cdot \boldsymbol{H}_{\mathrm{NLOS},i}^\mathrm{H}\} = \boldsymbol{R}_\mathrm{R} \in \mathbb{C}^{M_\mathrm{R} \times M_\mathrm{R}},\ i = 1,2 \tag{33}$$

$$\mathrm{E}\{\boldsymbol{H}_{\mathrm{NLOS},i}^\mathrm{H} \cdot \boldsymbol{H}_{\mathrm{NLOS},i}\} = \boldsymbol{R}_i \in \mathbb{C}^{M_i \times M_i},\ i = 1,2, \tag{34}$$

where $\boldsymbol{R}_i$ is the spatial correlation matrix at the user terminal i and $\boldsymbol{R}_\mathrm{R}$ represents the spatial correlation matrix at the relay. To model correlated channels we construct $\boldsymbol{R}_i$ and $\boldsymbol{R}_\mathrm{R}$ such that their main diagonal elements are equal to one and all off-diagonal elements have magnitude ρ_i and ρ_R, respectively. Consequently, $K_1 = K_2 = 0$ and $\rho_1 = \rho_2 = \rho_\mathrm{R} = 0$ corresponds to uncorrelated Rayleigh fading.

To estimate the channel we use training data which is constructed following the rules derived in [9]: The pilot matrix $\boldsymbol{X} \in \mathbb{C}^{(M_1+M_2) \times N_\mathrm{P}}$ in (13) is obtained from the first $M_1 + M_2$ rows of a DFT matrix of size $N_\mathrm{P} \times N_\mathrm{P}$, scaled in such a way that the transmit power constraint is satisfied. If not stated otherwise, we use the smallest possible value for N_P which is $N_\mathrm{P} = M_1 + M_2$. Moreover, the tensor $\boldsymbol{\mathcal{G}}$ in (14) is constructed via

$$\boldsymbol{\mathcal{G}} = \gamma \cdot \boldsymbol{\mathcal{I}}_{3,M_\mathrm{R}} \times_2 \boldsymbol{D}_{M_\mathrm{R}} \times_3 \left(\boldsymbol{D}_{M_\mathrm{R}} \odot \boldsymbol{S}_{M_\mathrm{R},\min\{M_1,M_2\}}\right), \tag{35}$$

where $\gamma \in \mathbb{R}$ ensures the relay transmit power constraint and $\boldsymbol{D}_{M_\mathrm{R}}$ is a DFT matrix of size $M_\mathrm{R} \times M_\mathrm{R}$. The (i,j) element of the matrix $\boldsymbol{S}_{p,q}$ is equal to one if $\mod(p-i+j,p) < q$ and zero otherwise for $i,j = 1,2,\ldots,p$. Consequently, for $\min\{M_1,M_2\} \geq M_\mathrm{R}$ the matrix $\boldsymbol{S}$ is equal a matrix of ones and the Schur product $\odot$ in (35) vanishes.

To evaluate the channel estimation accuracy we compute the relative mean square error, which is defined as

$$\mathrm{rMSE} = \mathrm{E}\left\{\min_{s=1,-1} \frac{\left\|\hat{\boldsymbol{H}} + s \cdot \boldsymbol{H}\right\|_\mathrm{F}^2}{\|\boldsymbol{H}\|_\mathrm{F}^2}\right\}, \tag{36}$$

where $\hat{\boldsymbol{H}}$ is an estimate of $\boldsymbol{H}$ and s accounts for the sign ambiguity in the estimation. The estimation error curves are labeled as `He11, He12, He21, He22`, where the first number indicates the terminal which estimates the channel referenced by the second number. For example, `He12` represents the estimate of $\boldsymbol{H}_2$ at user terminal 1. If not stated otherwise, we set $\alpha = 100$ and $\delta = 10^{-3}$.

see the program

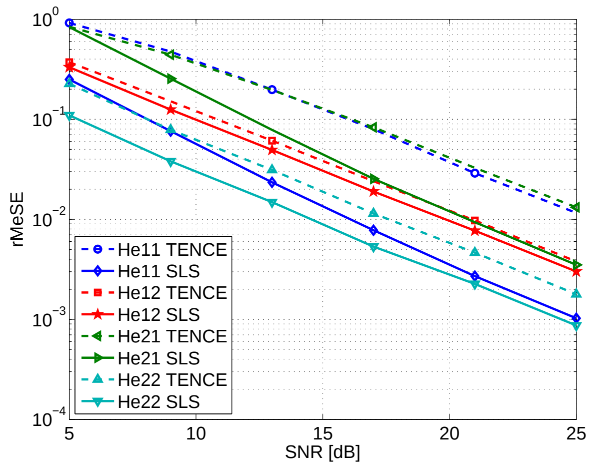


Fig. 2. Median of the channel estimation error vs. SNR in a mixed LOS/NLOS scenario where $M_1 = M_2 = M_{\mathrm{R}} = 2$, $\rho_1 = \rho_2 = \rho_{\mathrm{R}} = 0$, $K_1 = 20$, and $K_2 = 0$.

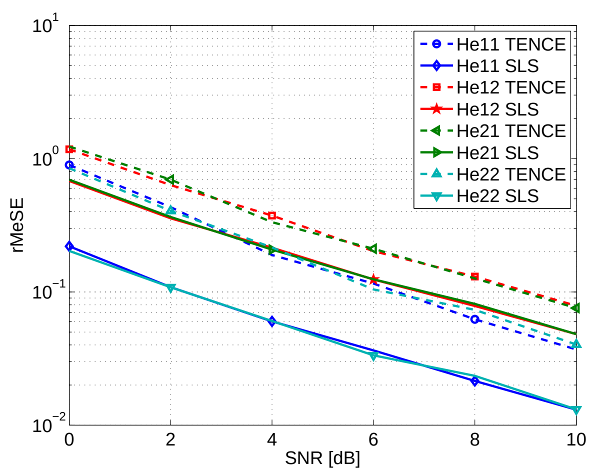


Fig. 3. Median of the channel estimation error vs. SNR in a pure NLOS scenario where $M_1 = M_2 = M_{\mathrm{R}} = 4$, $\rho_1 = \rho_2 = \rho_{\mathrm{R}} = 0$, $K_1 = 0$, and $K_2 = 0$.

The first simulation result in Figure 2 shows a scenario where $M_1 = M_2 = M_{\mathrm{R}} = 2$, $\rho_1 = \rho_2 = \rho_{\mathrm{R}} = 0$, $K_1 = 20$, and $K_2 = 0$. Consequently, the link from terminal one to the relay is a line of sight connection, whereas the link between terminal two and the relay is non line of sight. We observe that the improvements from our SLS-based technique are very large for the estimate of $\boldsymbol{H}_1$ at terminal 1. This is due to the fact that the high K-factor leads to a strong correlation which deteriorates the performance of TENCE. Via the iterative algorithm we can significantly improve this initial estimate by exploiting the structure of the training tensor.

In the second simulation shown in Figure 3 we study a pure NLOS scenario, i.e., $K_1 = K_2 = 0$. We consider $M_1 = M_2 = M_{\mathrm{R}} = 4$ antennas at the terminals and the relay and no spatial correlation, i.e., $\rho_1 = \rho_2 = \rho_{\mathrm{R}} = 0$. We observe that even in low SNR regimes, which are typical for mobile communication systems, the improvements from the SLS-based algorithm are significant.

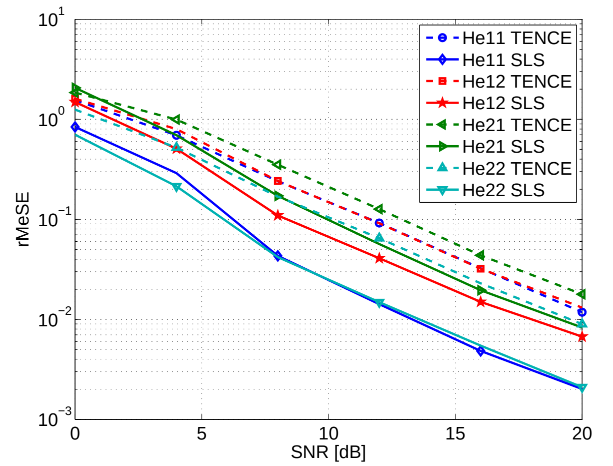


Fig. 4. Median of the channel estimation error vs. SNR in a pure NLOS scenario with spatial correlation where $M_1 = 4$, $M_2 = 5$, $M_{\mathrm{R}} = 3$, $\rho_1 = \rho_2 = 0$, $\rho_{\mathrm{R}} = 0.9$, $K_1 = 0$, and $K_2 = 0$.

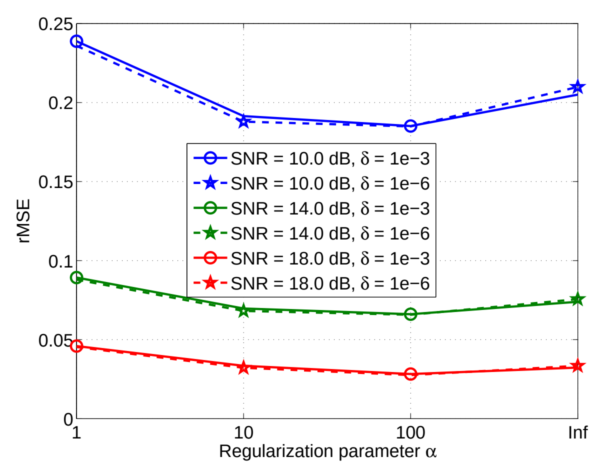


Fig. 5. Mean channel estimation error of the SLS-based channel estimation algorithm for $\boldsymbol{H}_1$ at terminal 1 vs. the regularization parameter α for different values of the SNR and two choices for the threshold parameter δ. Again, an NLOS scenario with $K_1 = K_2 = 0$ and with no spatial correlation ($\rho_1 = \rho_2 = \rho_{\mathrm{R}} = 0$) was considered. The number of antennas was chosen according to $M_1 = M_2 = 2$ and $M_{\mathrm{R}} = 4$.

Figure 4 depicts a scenario where $M_1 = 4$, $M_2 = 5$, $M_{\mathrm{R}} = 3$ and spatial correlation is introduced by setting $\rho_1 = \rho_2 = 0$, $\rho_{\mathrm{R}} = 0.9$. As before we consider the NLOS case by choosing $K_1 = K_2 = 0$. The strong correlation in both channel matrices deteriorate the performance of the TENCE algorithm. Therefore there is a lot of room for improvement via the SLS-based technique which leads to a significant gain especially in the estimation of $\boldsymbol{H}_1$ at terminal 1 and $\boldsymbol{H}_2$ at terminal 2. This is a reasonable observation that confirms the original goal to exploit the fact that the "own" channel matrix is present in the training tensor in the first as well as the second mode.

To show the effect of the regularization parameter α and the threshold parameter δ we study a scenario where each terminal is equipped with 4 antennas, there is no line of sight and also no spatial

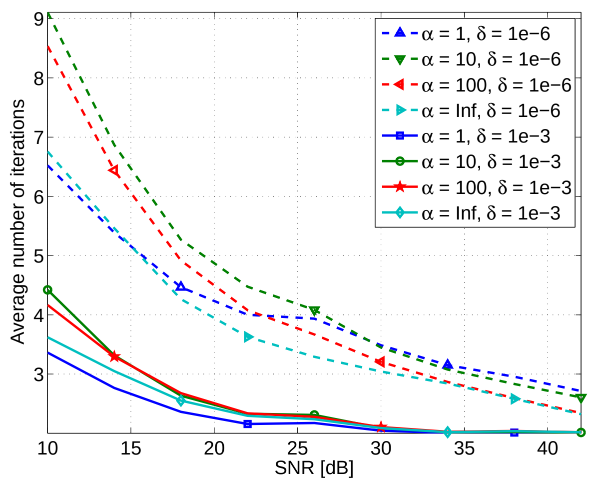


Fig. 6. Average number of iterations for the SLS-based channel estimation algorithm vs. the SNR for the same scenario as in Fig. 5.

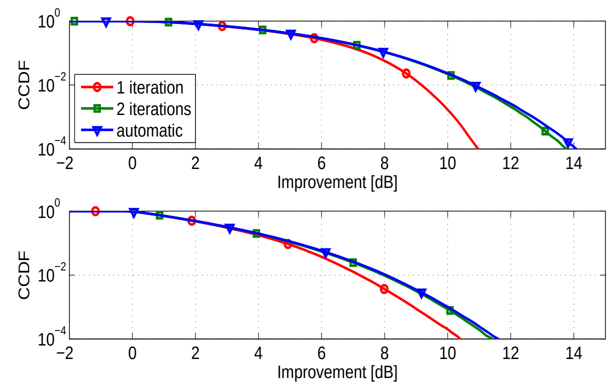


Fig. 8. CCDF of the improvement through the SLS-based refinement in a scenario where $M_1 = M_2 = 6$, $M_{\rm R} = 3$, SNR = 30 dB, $\rho_1 = \rho_2 = \rho_{\rm R} = 0.9$, and $K_1 = K_2 = 0$. The top graph shows $\boldsymbol{H}_1$, the bottom graph $\boldsymbol{H}_2$, both estimated at terminal 1.

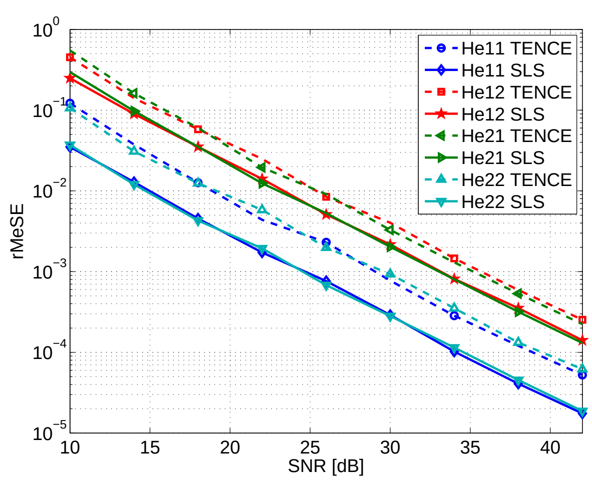


Fig. 7. Median of the channel estimation error vs. SNR for the same scenario as in Fig. 5. Here α was chosen to 100 and δ to 10^{-3}. The curves for $\delta = 10^{-6}$ are indistinguishable from these results (no improvement in accuracy). Also the curves for different values of α are exactly on the ones displayed here (except for $\alpha = 1$ which would be slightly worse).

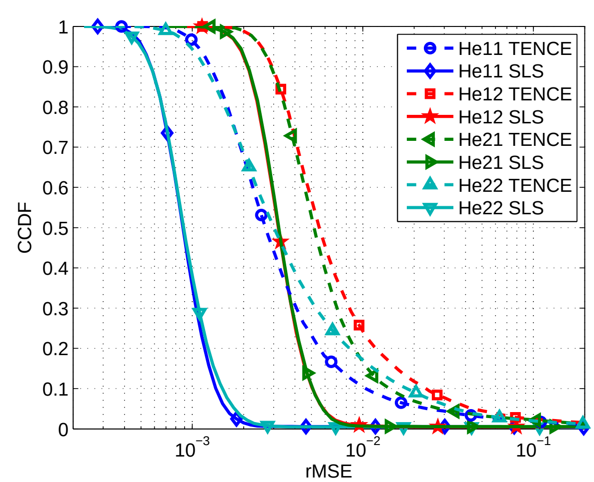


Fig. 9. CCDF of the estimation error for TENCE and the SLS-based technique in a scenario where $M_1 = M_2 = M_{\rm R} = 5$, SNR = 20 dB, $\rho_1 = \rho_2 = \rho_{\rm R} = 0.0$, and $K_1 = K_2 = 0$.

see the program

correlation. In Figure 5 we display the channel estimation error for $\boldsymbol{H}_1$ vs. the regularization parameter α. The last value for α corresponds to the case where no regularization is used. We observe that for a low SNR, α has an impact on the performance. If α is chosen too small, the correct minimum in the cost function is not achieved and if α is chosen too large the missing regularization has a negative effect on the performance. For an increasing SNR the impact of α diminishes. Consequently, for a very high SNR, no regularization is needed. We also display two values of the threshold parameter δ. As the difference between $\delta = 10^{-3}$ and $\delta = 10^{-6}$ is negligible we can conclude that $\delta = 10^{-3}$ is sufficient, since it requires less iterations. To confirm this in Figure 6 we demonstrate the average number of iterations for the same scenario versus the SNR. We observe that for $\delta = 10^{-3}$, two iterations are sufficient for high SNRs[1]. For low SNRs the number of iteration increases up to four. If δ is set to 10^{-6} the number of iterations increases significantly, however, as confirmed in Figure 5 this does not lead to a visible improvement in channel estimation accuracy. For comparison to the previous results, in Figure 7 we depict the estimation error vs. the SNR for the same scenario as in Figure 5 and Figure 6. Again, the accuracy of all channel estimates is enhanced and the improvement for the "own" channels is more pronounced.

In Figure 8 we show a simulation result in a scenario where both terminals are equipped with six antennas and the relay station with three antennas. We consider a spatial correlation of $\rho_1 = \rho_2 = \rho_\mathrm{R} = 0.9$ and no line of sight. Moreover, we set the SNR to 30 dB. In the previous figures we have only seen statistical moments of the estimation error but not its distribution. Therefore, in this simulation we estimate the complementary cumulative density function (CCDF) of the ratio of the channel estimation error before and after the SLS-based refinement, i.e., the probability that this ratio exceeds the abscissa. The top graph shows the estimation of the channel $\boldsymbol{H}_1$ and the bottom graph displays $\boldsymbol{H}_2$, both estimated at terminal 1. We observe that already after a single iteration (depicted in red), significant gains are achievable. The second iteration (green) is already almost indistinguishable from the blue line where up to 10 iterations are allowed and the SLS algorithm is terminated after convergence is detected. Consequently, two iterations are sufficient and the improvement ranges up to 14 dB.

Finally, Figure 9 displays the CCDF of the relative mean square estimation error for TENCE and for the SLS-based technique. The SNR is fixed to 20 dB, the number of antennas is set to 5 for both terminals and the relay. Also, NLOS with no spatial correlation ($\rho_1 = \rho_2 = \rho_\mathrm{R} = 0.0$) is considered. We observe that the improvement through the SLS-based refinement is particularly pronounced for the points where the estimation error obtained via TENCE is large. Moreover, the slope of the CCDF for the SLS-based techniques is significantly larger, which demonstrates that the estimates are more robust.

6. CONCLUSIONS

In this contribution an iterative channel estimation algorithm for two-way relaying based on Structured Least Squares is developed. We demonstrate that based on an initial channel estimate that can be obtained by the closed-form TENCE algorithm, the accuracy is improved further by exploiting the structure inherent in the received tensor during the training phase. Furthermore we show that only between one and four iterations are required, even for critical scenarios. Numerical computer simulations depict the achievable gain in terms of the channel estimation accuracy.

[1] Note that the algorithm is terminated if the change from the k-th to the previous iteration drops below the threshold. Consequently, in the case where two iterations are needed, only one iteration changed the result significantly and therefore one could limit the number of iterations to one without losing any performance in the high SNR regime.

REFERENCES

[1] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition", *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, 2000.

[2] G. Golub and C. F. Van Loan, "An analysis of the total least squares problem", *SIAM Journal Numer. Analysis*, vol. 17, pp. 883 – 893, Dec. 1980.

[3] M. Haardt, "Structured least squares to improve the performance of ESPRIT-type algorithms", *IEEE Transactions on Signal Processing*, vol. 45(3), pp. 792–799, Mar. 1997.

[4] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems", *IEEE Transactions on Signal Processing*, vol. 56, pp. 3198–3213, July 2008.

[5] I. Hammerstrom, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, "MIMO two-way relaying with transmit CSI at the relay", in *Proc. IEEE 8th Workshop on Sig. Proc. Adv. in Wireless Comm. (SPAWC 2007)*, Helsinki, Finland, June 2007.

[6] T. J. Oechtering and H. Boche, "Bidirectional relaying using interference cancellation", in *Proc. ITG/IEEE Workshop on Smart Antennas (WSA '07)*, Vienna, Austria, Feb. 2007.

[7] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels", in *Proc. 39th Asilomar Conference on Signals, Systems and Computers*, pp. 1066–1071, Pacific Grove, CA, USA, Oct. 2005.

[8] F. Roemer and M. Haardt, "Tensor-structure structured least squares (TS-SLS) to improve the performance of multi-dimensional ESPRIT-type algorithms", in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007)*, vol. II, pp. 893–896, Honolulu, HI, Apr. 2007.

[9] F. Roemer and M. Haardt, "Tensor-based channel estimation (TENCE) for two-way relaying with multiple antennas and spatial reuse", in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc. (ICASSP 2009)*, Apr. 2009.

[10] T. Unger and A. Klein, "Duplex schemes in multiple antenna two-hop relaying", *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2008, doi: 10.1155/2008/128592.

[11] T. Unger and A. Klein, "Maximum sum rate of non-regenerative two-way relaying in systems with different complexities", in *Proc. IEEE 9th Intl. Symp. on Personal, Indoor and Mobile Radio Comm. (PIMRC 2008)*, Cannes, France, Sept. 2008.

[12] S. van Huffel and J. Vanderwalle, *The total least squares problem: Computational Aspects and analysis*, Philadelphia: Society for Industrial and Applied Mathematics, 1991.

SEQUENTIAL GSVD BASED PREWHITENING FOR MULTIDIMENSIONAL HOSVD BASED SUBSPACE ESTIMATION

João Paulo C. L. da Costa, Florian Roemer, and Martin Haardt

Ilmenau University of Technology, Communications Research Laboratory
P.O. Box 100565, D-98684 Ilmenau, Germany.
{joaopaulo.dacosta,florian.roemer,martin.haardt}@tu-ilmenau.de

***Abstract* — Recently, R-dimensional subspace-based parameter estimation techniques have been improved by exploiting the tensor structure already in the subspace estimation step via a Higher Order Singular Value Decompostion (HOSVD) based low-rank approximation. Often this parameter estimation is performed in the presence of colored noise or interference, which can severely degrade the estimation accuracy. To avoid this degradation, prewhitening techniques are applied.**

In this contribution, we propose a Sequential Generalized Singular Value Decomposition (S-GSVD) based prewhitening scheme for multidimensional HOSVD based subspace estimation. By exploiting the Kronecker structure of the noise correlation matrix for the estimation of the correlation factors, we achieve an improved accuracy compared to matrix based prewhitening schemes. In addition, our S-GSVD approach is computationally more efficient than the classical matrix approach, since it has a lower complexity due to the n-mode GSVD operations.

I. INTRODUCTION

High-resolution parameter estimation from R-dimensional signals is a task required for a variety of applications, such as estimating the multidimensional parameters of the dominant multipath components from MIMO channel measurements, radar, sonar, seismology, and medical imaging.

One important application of the S-GSVD (Sequential Generalized Singular Value Decomposition) based prewhitening scheme is the parameter estimation in environments with multidimensional colored noise or interference with Kronecker correlation structure, i.e., the overall noise covariance matrix can be expressed as a Kronecker product of the noise covariance matrices in the separate dimensions. Since the S-GSVD based prewhitening approach requires knowledge of the noise correlation, we propose an efficient scheme to estimate the correlation factors which exploits this Kronecker structure.

The Kronecker noise or interference structure is present in many different applications. For example, in EEG/MEG applications [6], the noise is correlated in both space and time dimensions, and it has been shown that by combining these two correlations matrices using the Kronecker product a better model for the noise is obtained. Moreover, for MIMO systems the noise covariance matrix is often assumed to be the Kronecker product of the temporal and spatial correlation matrices [7].

Comparing the generalized HOSVD presented in [9] to our proposed S-GSVD based prewhitening approach, we observe that in the former the GSVD was only applied to one unfolding, while in our approach sequential GSVDs are applied to all unfoldings.

The remainder of this paper is organized as follows. In Section II, we describe the used notation. In Section III the data model and the colored noise structure are presented. A short description of the matrix based approach is shown in Section IV. In Section V, it is shown how the generalized HOSVD could be applied to the prewhitening step. The proposed S-GSVD is presented in Section VI. Simulations are shown in Section VII and conclusions are drawn in Section VIII.

II. TENSOR AND MATRIX NOTATION

In order to facilitate the distinction between scalars, matrices, and tensors, the following notation is used: Scalars are denoted as italic letters ($a, b, \ldots, A, B, \ldots, \alpha, \beta, \ldots$), column vectors as lower-case bold-face letters ($\boldsymbol{a}, \boldsymbol{b}, \ldots$), matrices as bold-face capitals ($\boldsymbol{A}, \boldsymbol{B}, \ldots$), and tensors are written as bold-face calligraphic letters ($\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \ldots$). Lower-order parts are consistently named: the (i, j)-element of the matrix $\boldsymbol{A}$, is denoted as $a_{i,j}$ and the (i, j, k)-element of a third order tensor $\boldsymbol{\mathcal{X}}$ as $x_{i,j,k}$. The n-mode vectors of a tensor are obtained by varying the n-th index within its range $(1, 2, \ldots, I_n)$ and keeping all the other indices fixed. The n-mode unfolding of $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ is represented by $[\boldsymbol{\mathcal{A}}]_{(n)} \in \mathbb{C}^{I_n \times I_1 \cdot I_2 \ldots \cdot I_N / I_n}$.

We use the superscripts $^{\mathrm{T}}$, $^{\mathrm{H}}$, $^{-1}$, $^{+}$, and * for transposition, Hermitian transposition, matrix inversion, the Moore-Penrose pseudo inverse of matrices, and complex conjugation, respectively. Moreover, the Khatri-Rao product (columnwise Kronecker product) and the trace operator are denoted by $\boldsymbol{A} \diamond \boldsymbol{B}$ and $tr(\cdot)$, respectively.

Here we use the n-mode product consistent with [1], i.e., the n-mode product between a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ and a matrix $\boldsymbol{U} \in \mathbb{C}^{J_n \times I_n}$ along the n-th mode is denoted as $\boldsymbol{\mathcal{A}} \times_n \boldsymbol{U} \in \mathbb{C}^{I_1 \times I_2 \ldots \times J_n \ldots \times I_N}$. It is obtained by multiplying all n-mode vectors of $\boldsymbol{\mathcal{A}}$ from the left-hand side by the matrix $\boldsymbol{U}$, i.e., $\boldsymbol{U} \cdot [\boldsymbol{\mathcal{A}}]_{(n)}$.

III. DATA MODEL

To demonstrate the effect of the enhanced multidimensional prewhitening schemes discussed in this paper, we consider its application to multidimensional harmonic retrieval based on ESPRIT [2], [3]. Therefore, we model our observations as a superposition of d undamped exponentials sampled on an R-dimensional grid of size $M_1 \times M_2 \times \ldots \times M_R$ at N subsequent time instants. The measurement samples are given by

$$\begin{aligned} x_{m_1, m_2, \ldots, m_R}(n) &= \sum_{i=1}^{d} s_i(n) \cdot \prod_{r=1}^{R} e^{j \cdot (m_r - 1) \cdot \mu_i^{(r)}} \\ &\quad + n^{c}_{m_1, m_2, \ldots, m_R}(n), \end{aligned} \tag{1}$$

where $m_r = 1, 2, \ldots, M_r$ for $r = 1, 2, \ldots, R$, $n = 1, 2, \ldots, N$, $s_i(n)$ denotes the complex amplitude of the i-th exponential at

João Paulo C. L. da Costa is a scholarship holder of the National Counsel of Technological and Scientific Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq) of the Brazilian Government and also a First Lieutenant of the Brazilian Army (Exército Brasileiro).

see the program

time instant n, $\mu_i^{(r)}$ symbolizes the spatial frequency of the i-th exponential in the r-th mode, and $n^c_{m_1,m_2,\ldots,m_R}(n)$ models the additive colored noise (and interference) component inherent in the measurement process.

In the context of array signal processing, each of the exponentials represents one planar wavefront and the complex amplitudes $s_i(n)$ are the symbols.

In the classical matrix approach, (1) is transformed into a matrix-vector equation by defining an array steering matrix [2]

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{A}^{(1)} \diamond \boldsymbol{A}^{(2)} \ldots \diamond \boldsymbol{A}^{(R)} \in \mathbb{C}^{M\times d} \\ \boldsymbol{A}^{(r)} &= \left[\boldsymbol{a}^{(r)}\!\left(\mu_1^{(r)}\right), \boldsymbol{a}^{(r)}\!\left(\mu_2^{(r)}\right), \ldots, \boldsymbol{a}^{(r)}\!\left(\mu_d^{(r)}\right)\right], \end{aligned} \tag{2}$$

where $M = \prod_{r=1}^{R} M_r$ and the vector $\boldsymbol{a}^{(r)}\!\left(\mu_i^{(r)}\right) \in \mathbb{C}^{M_r \times 1}$ denotes the array response in the r-th dimension for the i-th source. Here, all the spatial dimensions are stacked into column vectors. This stacking operation allows us to write the measurement equation in matrix form

$$\boldsymbol{X} = \boldsymbol{A} \cdot \boldsymbol{S} + \boldsymbol{N}^c, \tag{3}$$

where $\boldsymbol{X} \in \mathbb{C}^{M\times N}$ now contains the measurements stacked in a similar fashion as in $\boldsymbol{A}$, the matrix $\boldsymbol{S} \in \mathbb{C}^{d\times N}$ contains the symbols $s_i(n)$, and the colored noise samples are collected in the matrix $\boldsymbol{N}^c \in \mathbb{C}^{M\times N}$. It is obvious that the measurement matrix $\boldsymbol{X}$ does not capture the structure inherent in the R-D array that is used to sample the data.

We therefore replace the measurement matrix $\boldsymbol{X}$ by a measurement tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_1\times M_2\times\ldots\times M_R\times N}$ according to [3]. Similarly to (3), $\boldsymbol{\mathcal{X}}$ can be modeled as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{A}} \times_{R+1} \boldsymbol{S}^{\mathrm{T}} + \boldsymbol{\mathcal{N}}^c. \tag{4}$$

Here the matrix $\boldsymbol{S}$ is the same as in (3), the tensor $\boldsymbol{\mathcal{N}}^c$ contains the colored noise samples, and the tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1\times M_2\times\ldots\times M_R\times d}$ is termed the array steering tensor.

The multidimensional colored noise is assumed to have a Kronecker correlation structure, which can be written as

$$[\boldsymbol{\mathcal{N}}^c]_{(R+1)} = [\boldsymbol{\mathcal{N}}]_{(R+1)} \cdot (\boldsymbol{L}_1 \otimes \boldsymbol{L}_2 \otimes \ldots \otimes \boldsymbol{L}_R)^{\mathrm{T}}, \tag{5}$$

where $\otimes$ represents the Kronecker product. We can also rewrite (5) by using the n-mode products in the following fashion

$$\boldsymbol{\mathcal{N}}^c = \boldsymbol{\mathcal{N}} \times_1 \boldsymbol{L}_1 \times_2 \boldsymbol{L}_2 \ldots \times_R \boldsymbol{L}_R, \tag{6}$$

where $\boldsymbol{\mathcal{N}} \in \mathbb{C}^{M_1\times M_2\ldots\times N}$ is a tensor with uncorrelated ZMCSCG (zero-mean circularly symmetric complex Gaussian) elements with variance σ_n^2, and $\boldsymbol{L}_i \in \mathbb{C}^{M_i\times M_i}$ is the correlation factor of the i-th dimension of the colored noise tensor. The noise covariance matrix in the i-th mode is defined as

$$\mathrm{E}\left\{[\boldsymbol{\mathcal{N}}^c]_{(i)} \cdot [\boldsymbol{\mathcal{N}}^c]_{(i)}^{\mathrm{H}}\right\} = \alpha \cdot \boldsymbol{W}_i = \alpha \cdot \boldsymbol{L}_i \cdot \boldsymbol{L}_i^{\mathrm{H}}, \tag{7}$$

where α is a normalization constant, such that $\mathrm{tr}(\boldsymbol{L}_i \cdot \boldsymbol{L}_i^{\mathrm{H}}) = M_i$.

This equivalence between (5), (6), and (7) is shown in the Appendix, and it is the basis for our multidimensional approach. For notational convenience, we assume that the noise is uncorrelated in the time dimension, i.e., $\boldsymbol{L}_{R+1} = \boldsymbol{I}_N$.

IV. MATRIX BASED PREWHITENING SCHEME

In this section, we show briefly how the matrix based prewhitening approach can be used to prewhiten multidimensional data for the tensor-based signal subspace estimation schemes developed in [3].

IV-A. Prewhitening matrix estimation

In this subsection, we consider the case that no signal components are present in the data. Therefore, we can estimate the prewhitening matrix according to [4], [5], [8] from the noise covariance matrix in the following way

$$\begin{aligned} &\hat{\boldsymbol{W}}_{\mathrm{mtx}} = \hat{\boldsymbol{L}}_{\mathrm{mtx}} \cdot \hat{\boldsymbol{L}}_{\mathrm{mtx}}^{\mathrm{H}} = \\ &\frac{1}{\alpha \cdot N_l} \cdot [\boldsymbol{\mathcal{N}}^c]_{(R+1)}^{\mathrm{T}} \cdot [\boldsymbol{\mathcal{N}}^c]_{(R+1)}^{*} \in \mathbb{C}^{M\times M}, \end{aligned} \tag{8}$$

where α is chosen such that $tr(\hat{\boldsymbol{W}}_{\mathrm{mtx}}) = M$ and N_l is the number of snapshots without signal components. Note that $\hat{\boldsymbol{L}}_{\mathrm{mtx}}$ can, for example, be obtained by a Cholesky decomposition of $\hat{\boldsymbol{W}}_{\mathrm{mtx}}$ or by an eigenvalue decomposition. Since we assume an ergodic Kronecker noise model, if N_l approaches infinity, then $\hat{\boldsymbol{L}}_{\mathrm{mtx}}$ satisfies the following expression

$$\boldsymbol{L}_{\mathrm{mtx}} = \boldsymbol{L}_1 \otimes \boldsymbol{L}_2 \ldots \otimes \boldsymbol{L}_R. \tag{9}$$

where $\boldsymbol{L}_i$ is defined in Section III. However, (8) is applicable to an arbitrary noise correlation model.

IV-B. Matrix prewhitening

In the classical prewhitening approach to prewhiten the subspace tensor defined in [3], the prewhitening matrix $\hat{\boldsymbol{L}}_{\mathrm{mtx}}$ from (8) is applied and the prewhitening step is performed by

$$\left[\boldsymbol{\mathcal{X}}''\right]_{(R+1)}^{\mathrm{T}} = \hat{\boldsymbol{L}}_{\mathrm{mtx}}^{-1} \cdot [\boldsymbol{\mathcal{X}}]_{(R+1)}^{\mathrm{T}}, \tag{10}$$

where $\boldsymbol{\mathcal{X}}''$ represents the prewhitened data tensor using the matrix based approach. Computing the HOSVD-based low-rank approximation of $\boldsymbol{\mathcal{X}}''$ according to [3], we obtain the prewhitened subspace tensor $\boldsymbol{\mathcal{U}}^{[s]}$. Finally, for the estimation of the spatial frequencies, the noise correlation should be taken into account (dewhitening), which is performed by

$$\left[\boldsymbol{\mathcal{U}}^{[s]\prime\prime}\right]_{(R+1)}^{\mathrm{T}} = \hat{\boldsymbol{L}}_{\mathrm{mtx}} \cdot \left[\boldsymbol{\mathcal{U}}^{[s]}\right]_{(R+1)}^{\mathrm{T}}, \tag{11}$$

where $\boldsymbol{\mathcal{U}}^{[s]\prime\prime}$ can be used directly in the R-D Standard Tensor-ESPRIT algorithm [3] to obtain the harmonic retrieval parameters. Alternatively, a similar procedure can be used for the R-D Unitary Tensor ESPRIT [3].

Note that instead of the matrix inversion in (10), the GSVD or GEVD (Generalized Eigenvalue Decomposition) [4], [8] could have been applied. Except for critical cases that are not in the scope of this paper, the performance of using inversion, the GSVD, and the GEVD are the same in terms of accuracy. However, in terms of computational complexity the GSVD and GEVD are always better than using the matrix inversion.

V. GENERALIZED HOSVD OF [9]

The generalized HOSVD proposed in [9] could also be applied to the prewhitening task. The HOSVD is defined via the SVDs of all unfoldings of a tensor, while in the generalized HOSVD one of the SVDs is replaced by the GSVD of the t-th unfoldings of two tensors. In our case, we consider the tensors $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{N}}^c$.

Since the generalized HOSVD considers only one dimension for the GSVD [9], we can apply it only in the case that the t-th dimension of the noise is correlated. Therefore, applying the generalized HOSVD of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{N}}^c$ in the t-th dimension, we obtain

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{S}} \times_1 \boldsymbol{U}_1 \times_2 \boldsymbol{U}_2 \ldots \times_t \boldsymbol{G} \ldots \times_R \boldsymbol{U}_R \times_{R+1} \boldsymbol{U}_{R+1} \tag{12}$$

$$\boldsymbol{\mathcal{N}}^c = \boldsymbol{\mathcal{T}} \times_1 \boldsymbol{V}_1 \times_2 \boldsymbol{V}_2 \ldots \times_t \boldsymbol{G} \ldots \times_R \boldsymbol{V}_R \times_{R+1} \boldsymbol{V}_{R+1}, \tag{13}$$

Note that since $\boldsymbol{\mathcal{N}}^c$ and $\boldsymbol{\mathcal{X}}$ cannot be observed simultaneously, they represent different realizations of the noise process and the signal plus noise process, respectively.

Similarly to the HOSVD-based low-rank approximation of $\mathcal{X}$ [3], we might apply the generalized HOSVD-based low-rank approximation to (12) in order to obtain

$$\mathcal{U}^{[s]} = \mathcal{S}^{[s]} \times_1 \boldsymbol{U}_1^{[s]} \times_2 \boldsymbol{U}_2^{[s]} \ldots \times_t \boldsymbol{G}^{[s]} \ldots \times_R \boldsymbol{U}_R^{[s]}, \tag{14}$$

where $\mathcal{U}^{[s]} \in \mathbb{C}^{M_1 \times M_2 \times \ldots \times M_R \times d}$, $\mathcal{S}^{[s]} \in \mathbb{C}^{p_1 \times p_2 \times \ldots \times p_R \times d}$, $\boldsymbol{U}_i^{[s]} \in \mathbb{C}^{M_i \times p_i}$, and $\boldsymbol{G}^{[s]} \in \mathbb{C}^{M_t \times p_t}$, such that $p_i = \min(M_i, d)$. The subspace tensor $\mathcal{U}^{[s]}$ can be used in the R-D Standard Tensor-ESPRIT algorithm.

VI. MULTIDIMENSIONAL PREWHITENING SCHEMES

In this section, we present the proposed S-GSVD based prewhitening scheme.

VI-A. Prewhitening correlation factors estimation

In this subsection, we consider the case that no signal components are present in the data. Nevertheless, if the noise covariance has a Kronecker structure, we can obtain a better estimate of the noise model by considering each $[\mathcal{N}^c]_{(i)}$ for $i = 1, 2, \ldots, R$. Based on (7) it is possible to estimate the correlation factor $\boldsymbol{L}_i$ by applying an eigenvalue decomposition (EVD) or a Cholesky decomposition. For example, in case of using the EVD we can estimate the factors by performing

$$\hat{\boldsymbol{W}}_i = \frac{M_i}{\alpha \cdot N_l \cdot M} \cdot [\mathcal{N}^c]_{(i)} \cdot [\mathcal{N}^c]_{(i)}^{\mathrm{H}}, \tag{15}$$

where α is chosen such that $\operatorname{tr}(\hat{\boldsymbol{W}}_i) = M_i$. Since the EVD of $\hat{\boldsymbol{W}}_i$ is given by

$$\hat{\boldsymbol{W}}_i = \boldsymbol{Q}_i \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{Q}_i^{\mathrm{H}}, \tag{16}$$

we can choose

$$\hat{\boldsymbol{L}}_i = \boldsymbol{Q}_i \cdot \boldsymbol{\Lambda}^{\frac{1}{2}}. \tag{17}$$

VI-B. Tensor prewhitening scheme: n-mode products using matrix inversions

In this subsection, we show briefly how to apply the prewhitening using the correlation factors. In the simulations, the prewhitening techniques shown here are used in conjunction with R-D Standard Tensor-ESPRIT [3]. In the same way, it is also possible to use the multidimensional prewhitening schemes in conjunction with R-D Unitary Tensor-ESPRIT [3].

First let us propose the prewhitening scheme based on n-mode products using matrix inversions with the correlation factor estimation proposed in (17). The expression below shows the prewhitened tensor

$$\mathcal{X}' = \mathcal{X} \times_1 \hat{\boldsymbol{L}}_1^{-1} \times_2 \hat{\boldsymbol{L}}_2^{-1} \ldots \times_R \hat{\boldsymbol{L}}_R^{-1}. \tag{18}$$

Replacing (4) in (18), we obtain that:

$$\mathcal{X}' = \mathcal{A} \times_1 \hat{\boldsymbol{L}}_1^{-1} \times_2 \hat{\boldsymbol{L}}_2^{-1} \ldots \times_R \hat{\boldsymbol{L}}_R^{-1} \times_{R+1} \boldsymbol{S}^{\mathrm{T}} + \mathcal{N}, \tag{19}$$

where after prewhitening the multidimensional noise becomes white. Moreover, $\mathcal{X}'$ can be represented by its HOSVD as follows

$$\mathcal{X}' = \mathcal{S} \times_1 \boldsymbol{U}_1 \times_2 \boldsymbol{U}_2 \ldots \times_R \boldsymbol{U}_R \times_{R+1} \boldsymbol{U}_{R+1}, \tag{20}$$

where $\mathcal{S} \in \mathbb{C}^{M_1 \times M_2 \times M_R \times N}$ and $\boldsymbol{U}_i \in \mathbb{C}^{M_i \times M_i}$, such that $i = 1, 2, \ldots R$. The HOSVD-based low-rank approximation can be applied to (20). Therefore, the corresponding subspace tensor is equal to

$$\mathcal{U}^{[s]\prime} = \mathcal{S}^{[s]} \times_1 \boldsymbol{U}_1^{[s]} \times_2 \boldsymbol{U}_2^{[s]} \ldots \times_R \boldsymbol{U}_R^{[s]}, \tag{21}$$

where $\mathcal{S}^{[s]} \in \mathbb{C}^{p_1 \times p_2 \ldots \times p_R \times d}$, $\boldsymbol{U}_i^{[s]} \in \mathbb{C}^{M_i \times p_i}$, and $\mathcal{U}^{[s]\prime} \in \mathbb{C}^{M_1 \times M_2 \ldots \times M_R \times d}$, such that $p_i = \min(M_i, d)$ for $i = 1, 2, \ldots, R$. We assume that $d \leq N$.

Similarly to the derivation of the R-D Standard Tensor-ESPRIT in [3], there is an invertible matrix $\boldsymbol{T}$, such that

$$\mathcal{U}^{[s]\prime} = \mathcal{A} \times_1 \hat{\boldsymbol{L}}_1^{-1} \times_2 \hat{\boldsymbol{L}}_2^{-1} \ldots \times_R \hat{\boldsymbol{L}}_R^{-1} \times_{R+1} \boldsymbol{T}. \tag{22}$$

Therefore, to obtain the correct subspace is necessary to dewhiten the estimated subspace by performing

$$\mathcal{U}^{[s]} = \mathcal{U}^{[s]\prime} \times_1 \hat{\boldsymbol{L}}_1 \times_2 \hat{\boldsymbol{L}}_2 \ldots \times_R \hat{\boldsymbol{L}}_R. \tag{23}$$

VI-C. Tensor prewhitening scheme: S-GSVD

In the S-GSVD based prewhitening approach, the explicit prewhitening by matrix inversion can be avoided, and additionally the it allows the parallel implementation of (18). Let $\widetilde{\mathcal{X}}^{(i)}$ be the tensor which has been prewhitened up to the i-th dimension, where $i = 1, 2, \ldots, R$, and $\widetilde{\mathcal{X}}^{(0)} = \mathcal{X}$. To perform the prewhitening in the i-th dimension, we calculate the GSVD of $\left[\widetilde{\mathcal{X}}^{(i-1)}\right]_{(i)}^{\mathrm{H}}$ and $\hat{\boldsymbol{L}}_i^{\mathrm{H}}$, which gives us

$$\begin{aligned} \hat{\boldsymbol{L}}_i &= \boldsymbol{V}_i \cdot \boldsymbol{\Xi}_i^{(L)\mathrm{T}} \cdot \boldsymbol{U}_i^{(L)\mathrm{H}} \\ \left[\widetilde{\mathcal{X}}^{(i-1)}\right]_{(i)} &= \boldsymbol{V}_i \cdot \boldsymbol{\Xi}_i^{(X)\mathrm{T}} \cdot \boldsymbol{U}_i^{(X)\mathrm{H}}, \end{aligned} \tag{24}$$

where $\boldsymbol{U}_i^{(X)} \in \mathbb{C}^{\frac{M \cdot N}{M_i} \times \frac{M \cdot N}{M_i}}$ and $\boldsymbol{U}_i^{(L)} \in \mathbb{C}^{M_i \times M_i}$ are unitary matrices, and the matrices $\boldsymbol{\Xi}_i^{(X)} \in \mathbb{R}^{\frac{M \cdot N}{M_i} \times M_i}$ and $\boldsymbol{\Xi}_i^{(L)} \in \mathbb{R}^{M_i \times M_i}$ have all elements equal to zero, except for the main diagonals. Note that the S-GSVD is related to the technique of Subsection VI-B, since each n-mode GSVD would be equivalent to one n-mode multiplication by $\hat{\boldsymbol{L}}_i^{-1}$ of the technique of Subsection VI-B. Indeed, we have that

$$\begin{aligned} &\left[\widetilde{\mathcal{X}}^{(i)}\right]_{(i)} = \hat{\boldsymbol{L}}_i^{-1} \cdot \left[\widetilde{\mathcal{X}}^{(i-1)}\right]_{(i)} = \\ &\boldsymbol{U}_i^{(L)} \cdot \boldsymbol{\Xi}_i^{(L)-1} \cdot \boldsymbol{\Xi}_i^{(X)\mathrm{T}} \cdot \boldsymbol{U}_i^{(X)\mathrm{H}}, \end{aligned} \tag{25}$$

which is equivalent to

$$\widetilde{\mathcal{X}}^{(i)} = \widetilde{\mathcal{X}}^{(i-1)} \times_i \hat{\boldsymbol{L}}_i^{-1}, \tag{26}$$

After applying R GSVDs in (24) and rebuilding the prewhitened tensor using (25), we obtain $\widetilde{\mathcal{X}}^{(R)}$ which is equal to $\mathcal{X}'$ in (18).

Since $\widetilde{\mathcal{X}}^{(0)}$ is prewhitened sequentially in the S-GSVD approach, $\boldsymbol{U}_i^{(L)}$ in (25) is an approximation of $\boldsymbol{U}_i$ in (20). With every step in the sequential prewhitening, the approximation becomes better, until that for the last GSVD, e.g., when $i = R$, $\boldsymbol{U}_R^{(L)} = \boldsymbol{U}_R$. At the end of this section, it is explained how to improve the accuracy of the S-GSVD in order to obtain the same performance as in Subsection VI-B.

Since $\boldsymbol{U}_i^{(L)}$ is an unitary matrix, we can rewrite $\widetilde{\mathcal{X}}^{(R)}$ in the following fashion

$$\widetilde{\mathcal{X}}^{(R)} = \mathcal{S}' \times_1 \boldsymbol{U}_1^{(L)} \times_2 \boldsymbol{U}_2^{(L)} \ldots \times_R \boldsymbol{U}_R^{(L)} \times_{R+1} \boldsymbol{U}_{R+1}. \tag{27}$$

where $\mathcal{S}' \in \mathbb{C}^{M_1 \times M_2 \ldots \times M_R \times N}$. $\boldsymbol{U}_{R+1}$ is calculated by applying the SVD in the $R+1$-mode unfolding.

If we multiply (24) by $\boldsymbol{U}_i^{(L)}$ from the right side, we can conclude that

$$\hat{\boldsymbol{L}}_i \cdot \boldsymbol{U}_i^{(L)} = \boldsymbol{V}_i \cdot \boldsymbol{\Xi}_i^{(L)\mathrm{T}} = \boldsymbol{V}_i \cdot \boldsymbol{\Xi}_i^{(L)}. \tag{28}$$

In a similar fashion, the low-rank approximation can be applied to (27) and using (30) an approximated dewhitened subspace tensor can be directly obtained

$$\begin{aligned} \widetilde{\mathcal{U}}^{[s]} = \mathcal{S}^{[s]\prime} \times_1 \left(\boldsymbol{V}_1^{[s]} \cdot \boldsymbol{\Xi}_1^{(L)[s]}\right) \times_2 \left(\boldsymbol{V}_2^{[s]} \cdot \boldsymbol{\Xi}_2^{(L)[s]}\right) \ldots \\ \times_R \left(\boldsymbol{V}_R^{[s]} \cdot \boldsymbol{\Xi}_R^{(L)[s]}\right), \end{aligned} \tag{29}$$

see the program

where $\boldsymbol{\mathcal{S}}^{[s]'} \in \mathbb{C}^{p_1 \times p_2 \ldots \times p_R \times d}$, $\boldsymbol{V}_i^{[s]} \in \mathbb{C}^{M_i \times p_i}$, and $\boldsymbol{\Xi}_i^{(L)[s]} \in \mathbb{C}^{p_i \times p_i}$. This $\widetilde{\boldsymbol{\mathcal{U}}}^{[s]}$ is the subspace tensor which can be used in the R-D Standard Tensor-ESPRIT algorithm. Note that the order of the dimensions in the sequential prewhitening scheme can be chosen arbitrarily. The best accuracy is achieved if we begin in the dimension with the highest noise correlation and proceed to the less correlated dimensions sequentially.

As mentioned previously, since the S-GSVD is applied sequentially, the terms $\boldsymbol{U}_i^{(L)}$ are an approximation of $\boldsymbol{U}_i$. To obtain an equality, $\boldsymbol{U}_i$ can be obtained via an SVD as the left singular vectors of $\left[\widetilde{\boldsymbol{\mathcal{X}}}^{(R)}\right]_{(i)}$. Therefore, (29) has to be modified by

$$\boldsymbol{\mathcal{U}}^{[s]} = \boldsymbol{\mathcal{S}}^{[s]} \times_1 \left(\hat{\boldsymbol{L}}_1 \cdot \boldsymbol{U}_1^{[s]}\right) \times_2 \left(\hat{\boldsymbol{L}}_2 \cdot \boldsymbol{U}_2^{[s]}\right) \ldots$$
$$\times_{R-1} \left(\hat{\boldsymbol{L}}_{R-1} \cdot \boldsymbol{U}_{R-1}^{[s]}\right) \times_R \left(\boldsymbol{V}_R^{[s]} \cdot \boldsymbol{\Xi}_R^{(L)[s]}\right), \quad (30)$$

where $\boldsymbol{\mathcal{S}}^{[s]} \in \mathbb{C}^{p_1 \times p_2 \ldots \times p_R \times d}$ and $\boldsymbol{U}_i^{[s]} \in \mathbb{C}^{M_i \times p_i}$. By performing the modification, the accuracy is exactly the same as in Subsection VI-B, although additionally $R-1$ SVDs are computed, which leads to a mild increase in complexity as compared to the low complexity version of the S-GSVD.

VI-D. Comparing the computational complexity

In the proposed multidimensional approach in (18), we need to invert R matrices of a much smaller size than in (10) to prewhiten the data. Consequently, (18) is also computationally much more efficient than (10), since the complexity of an inversion of an $M \times M$ matrix is approximately $O(M^3)$. Moreover, in the S-GSVD based prewhitening approach in (24) no explicit matrix inversion is required.

Note that the S-GSVD in its low complexity form in (29) only requires the computation of GSVDs for the unfoldings with noise correlation and SVDs for the unfoldings without noise correlation. However, in its high accuracy form in (30) additional SVDs should be computed.

VII. SIMULATION RESULTS

In this section we present simulation results demonstrating the performance of the proposed methods. The spatial frequencies $\mu_i^{(r)}$ are drawn from a uniform distribution in $[-\pi, \pi]$. The source symbols are zero mean i.i.d. circularly symmetric complex Gaussian distributed with power equal to σ_s^2 for all the sources. We define at the receiver

$$\text{SNR} = 10 \cdot \log_{10}\left(\frac{\sigma_s^2}{\sigma_n^2}\right), \quad (31)$$

where σ_n^2 is the variance of the elements of the white noise tensor $\boldsymbol{\mathcal{N}}$. Since $\operatorname{tr}(\boldsymbol{W}_i) = M_i$ for the case that the correlation levels are not zero, this definition is still valid. In addition, $\boldsymbol{W}_i$ varies as a function of the correlation coefficient ρ_i.

In the simulations, we assume the noise correlation structure of equation (7), where $\boldsymbol{W}_i$ of the i-th factor for $M_i = 3$ is given by

$$\boldsymbol{W}_i = \begin{bmatrix} 1 & \rho_i^* & (\rho_i^*)^2 \\ \rho_i & 1 & \rho_i^* \\ \rho_i^2 & \rho_i & 1 \end{bmatrix}, \quad (32)$$

where ρ_i is the correlation coefficient. Note that also other types of correlation models different from (32) can be used.

In order to evaluate the prewhitening schemes integrated into the R-D Standard Tensor-ESPRIT [3], we compute the total RMSE of the estimated spatial frequencies $\hat{\mu}_i$ as follows

$$\text{RMSE} = \sqrt{\text{E}\left\{\sum_{r=1}^{R}\sum_{i=1}^{d}\left(\hat{\mu}_i^{(r)} - \mu_i^{(r)}\right)^2\right\}}. \quad (33)$$

Abbreviation	Algorithm
R-D SE Color	The Standard ESPRIT algorithm is applied without prewhitening.
R-D SE MP E	The Standard ESPRIT algorithm is applied together with the matrix based prewhitening scheme based on (10) and (11) with the matrix-based estimate of the correlation matrix from (8).
R-D SE TP E	The Standard ESPRIT algorithm is applied together with the matrix based prewhitening scheme based on (10) and (11), where the correlation factors matrix is computed via (9) based on the factor estimation from (17).
R-D STE Color	The R-D Standard Tensor-ESPRIT algorithm is applied without prewhitening.
R-D STE MP E	The matrix prewhitening is calculated according to (10) and (11) with the matrix-based estimate of the correlation matrix from (8).
R-D STE S-GSVD I	The proposed multidimensional schemes with correlation factors estimation (17), and we use the S-GSVD prewhitening based on (24), (25) and (29).
R-D STE S-GSVD II	The proposed multidimensional schemes with correlation factors estimation (17), and we use the S-GSVD prewhitening based on (24) and (25). However, we include $R-1$ SVDs at the end of the S-GSVD in order to improve the estimation, and (30) should be computed.
R-D STE TP CI	The correlation factors are known, and we apply (18) and (22) to perform the multidimensional prewhitening.
G-HOSVD	The generalized HOSVD-based low-rank approximation is applied in (14).

Table I. Notation of the legends used in all figures, where $R = 3$ for Fig. 3 and $R = 2$ for the other figures.

Initially let us describe the notation of the legends used in all figures. The prefix R-D STE stands for R-D Standard Tensor-ESPRIT [3], and SE stands for Standard ESPRIT. In Table I, the acronym of each applied technique is shown. Moreover, for $N_l \geq 10^2$, the correlation factors estimation based on (17) gives an estimate, which is very close to the true correlation factors as shown in all figures.

In Figure 1, 2-D SE Color and 2-D STE Color present almost the same performance, since for the colored noise the major part of the noise is concentrated in the signal eigenvectors, and consequently the HOSVD-based low-rank approximation does not lead to a significant denoising. For this reason, when a prewhitening scheme is applied, it is expected that the noise gets almost equally distributed over all eigenvectors. Therefore, since the prewhitening takes the multidimensional structure into account, the estimation is improved much more. The 2-D SE MP E and 2-D SE TP E give a similar performance, since the 2-D SE does not take into account the multidimensional structure. The 2-D STE MP E has a limited accuracy of approximately 10^{-2}, which is caused by the estimation of the prewhitening matrix which does not take into account the Kronecker structure. 2-D STE S-GSVD I has a performance close to the optimal, however it is slightly worse, since the prewhitening is applied sequentially. 2-D STE S-GSVD II achieves the optimal performance, since it is on top of 2-D STE TP CI, which considers the correlation factors as being known.

In Figure 2, the number of sources is increased from $d = 3$ to $d = 4$, and in general all the curves get closer to each other, which means that the gain obtained by using the prewhitening is reduced. Moreover, for the matrix based approach 2-D STE MP E, the limited accuracy is increased to $2 \cdot 10^{-2}$. Therefore, increasing

see the program

the number of sources, the performance of the matrix based approach gets drastically worse.

In Figure 3, the number of dimensions has been increased, and the array size is changed to $M_1 = 5$, $M_2 = 5$, and $M_3 = 5$. This requires an additional correlation factor ρ_3, which is set to 0.95. In general, we can observe that the gain caused by the multidimensional prewhitening is significantly increased, if we compare it to Fig. 2, for example. In addition, the matrix based approach 3-D STE MP E has an even worse performance, since its limited accuracy is approximately $9 \cdot 10^{-2}$. Moreover, for such scenarios, the multidimensional prewhitening enhances the accuracy significantly even for high SNR levels, while for scenarios in Fig. 2 for high SNR levels the improvement caused by the multidimensional prewhitening vs. no prewhitening is very small.

In Figure 4, the SNR is fixed to 30 dB. For the case that no prewhitening is applied, i.e., 2-D STE Color, the RMSE should be constant. Since we consider a high SNR regime, the matrix based prewhitening 2-D STE MP E has a limited accuracy for each N_l. Moreover, even for N_l very large, the accuracy is much worse than applying no prewhitening. Note also that the matrix based prewhitening for N_l lower than $M_1 \cdot M_2 = 25$ leads to a rank-deficient estimate of $\boldsymbol{L}_{\mathrm{mtx}}$. Therefore, for the simulations $\boldsymbol{L}_{\mathrm{mtx}}^{-1}$ in (10) is replaced by $\boldsymbol{L}_{\mathrm{mtx}}^{+}$. On the other hand, since the SNR is very high, the multidimensional prewhitening schemes show no accuracy limitation even for a very small N_l, since they remain almost constant for N_l varying from 10^2 to $2 \cdot 10^4$.

In Figure 5, the same scenario as in Fig. 4 is considered, but the SNR is fixed to 0 dB. Since we consider a low SNR regime, the matrix based prewhitening 2-D STE MP E has an improvement in its estimate until $N_l = 10^4$, after this value no improvement is observed. Similarly to Fig. 4, the multidimensional prewhitening schemes show no accuracy limitation even for a very small N_l, since they remain almost constant for N_l varying from 10^2 to $2 \cdot 10^4$.

In Figure 6, a scenario with an intermediate SNR = 10 dB is selected, and the correlation levels ρ_1 and ρ_2 are equal to ρ, which varies from 0 to 0.999. Since the matrix based prewhitening technique 2-D STE MP E has an accuracy limitation, then for high correlation levels, the performance is worse than the one using no prewhitening. The approaches using the correlation factors give a better estimate compared to the matrix based prewhitening or no prewhitening cases, when the correlation levels are higher. Therefore, for intermediate SNRs the gains are substantial only for high correlation levels, i.e., $\rho > 0.6$.

In Figure 7, we have the same scenario as in Fig. 6, except that instead of an intermediate SNR level, we have a low SNR level of −5 dB. For such scenarios, even for low correlations levels $\rho > 0.3$, there is a significant improvement.

In Figure 8, the array size varies, and as expected the RMSE becomes smaller. Note that there is always a significant improvement of the techniques based on correlation factors compared to the matrix based prewhitening and no prewhitening.

In Figure 9, the number of snapshots N increases, and as expected the RMSE becomes smaller for all approaches. Note that increasing the samples the gain obtained by the proposed multidimensional schemes is greater than the one obtained in 2-D STE Color and 2-D STE MP E.

In Figure 10, the G-HOSVD is included, and due to its definition we consider only one correlation factor in the first dimension. The correlation level is 0.9. The proposed multidimensional techniques outperform the G-HOSVD as shown in Fig. 10.

In Figure 11, we increase the number of sources from 2 to 4, and the G-HOSVD becomes even closer to the 2-D STE Color.

In Figure 12, the number of sources is increased from 4 to 5 and the G-HOSVD superposes the 2-D STE Color, as expected. Note that for $d \geq M_i$, there is no accuracy limitation for the 2-D STE MP E. The performance of the techniques with factors

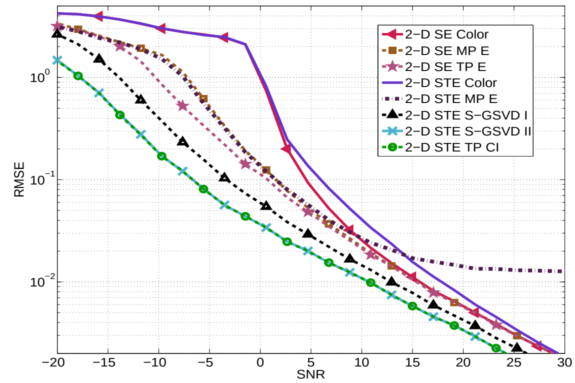


Fig. 1. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard ESPRIT and 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider an array of size $M_1 = 5$ and $M_2 = 5$. The number of snapshots N is set to 20 and the number of sources $d = 3$. The correlation levels are $\rho_1 = 0.9$ for $\boldsymbol{L}_1$ and $\rho_2 = 0.95$ for $\boldsymbol{L}_2$. The number of snapshots N_l is set to 10^4.

estimation, e.g., 2-D STE TP S-GSVD II, is the same as the one in which the factors are assumed to be known, i.e., 2-D STE TP CI. Note that the proposed multidimensional techniques outperforms all the other techniques.

In Figures 10, 11, and 12, N_l for the generalized HOSVD is set to 10^2, since for larger N_l, the computational complexity of the generalized HOSVD can be very high. Note that the generalized HOSVD is not included in Figs. 1- 9, since it is not possible to prewhiten more than one dimension using the generalized HOSVD. Therefore, the inclusion of the generalized HOSVD for such scenarios would not be a fair comparison.

VIII. CONCLUSIONS

In this paper, we propose the S-GSVD based prewhitening, where we take into account the Kronecker tensor structure of the colored noise. The first benefit of the S-GSVD is the superior estimation accuracy, since the S-GSVD is based on the estimation of the prewhitening correlation factors proposed in Subsection VI-A. The proposed prewhitening correlation factors can be integrated separately to other prewhitening approaches, e.g., in Subsection IV-B via (9), in Subsection VI-B, and in Subsection VI-C, and for all of them the same accuracy on the estimation is obtained. In addition, besides the improved estimation accuracy of the S-GSVD based prewhitening approach, our proposed scheme also has a lower computational complexity version, which requires only the computation of n-mode GSVDs.

APPENDIX

In this appendix, we show the derivation for the equivalence between (5), (6), and (7), which is also the basis for (17). It is possible to derive (7) by first calculating the covariance matrix of each unfolding of the noise tensor as follows

$$\begin{aligned} &\mathrm{E}\left\{[\boldsymbol{\mathcal{N}}^c]_{(i)} \cdot [\boldsymbol{\mathcal{N}}^c]_{(i)}^{\mathrm{H}}\right\} = \qquad (34)\\ &\boldsymbol{L}_i \cdot \mathrm{E}\{[\boldsymbol{\mathcal{N}}]_{(i)} \cdot (\boldsymbol{L}_{i+1} \otimes \ldots \otimes \boldsymbol{L}_R \otimes \boldsymbol{I}_N \otimes \boldsymbol{L}_1 \ldots \otimes \boldsymbol{L}_{i-1}) \cdot \\ &(\boldsymbol{L}_{i+1} \otimes \ldots \otimes \boldsymbol{L}_R \otimes \boldsymbol{I}_N \otimes \boldsymbol{L}_1 \ldots \otimes \boldsymbol{L}_{i-1})^{\mathrm{H}} \cdot [\boldsymbol{\mathcal{N}}]_{(i)}^{\mathrm{H}}\} \cdot \boldsymbol{L}_i^{\mathrm{H}}. \end{aligned}$$

Since the elements of $[\boldsymbol{\mathcal{N}}]_{(i)}$ are independent and identically

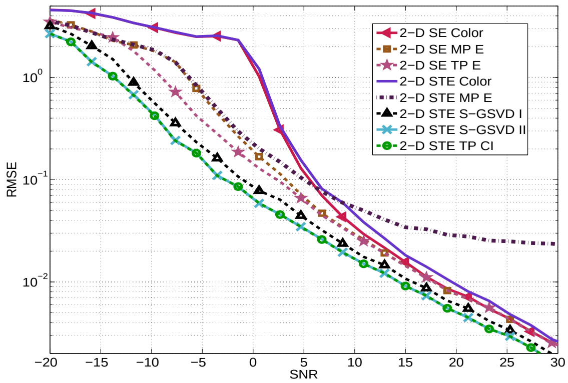


Fig. 2. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard ESPRIT and 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, number of snapshots N, correlation levels ρ_1 and ρ_2, and snapshots N_l as in Fig. 1. However, the number of sources is set to $d = 4$.

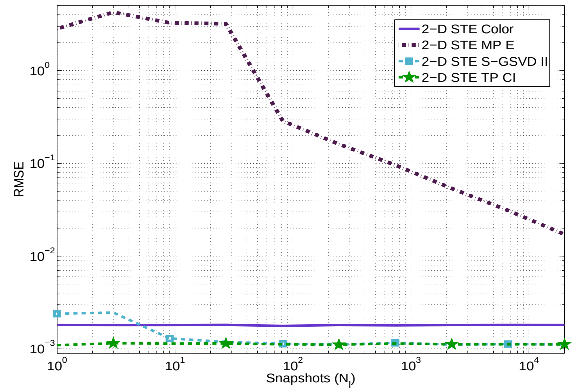


Fig. 4. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, number of snapshots N, and number of sources d as in Fig. 2. However, the SNR is fixed to 30 dB, and the snapshots N_l vary from 1 to $2 \cdot 10^4$ samples.

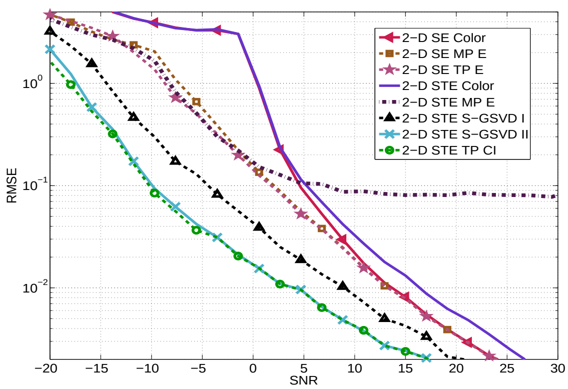


Fig. 3. Comparing the different types of prewhitening schemes in conjunction with 3-D Standard ESPRIT and 3-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the number of sources d, number of snapshots N, and snapshots N_l as in Fig. 2. However, the array size is set $M_1 = 5$, $M_2 = 5$ and $M_3 = 5$, and the correlation levels are $\rho_1 = 0.9$ for $\boldsymbol{L}_1$, $\rho_2 = 0.95$ for $\boldsymbol{L}_2$, and $\rho_3 = 0.95$ for $\boldsymbol{L}_3$.

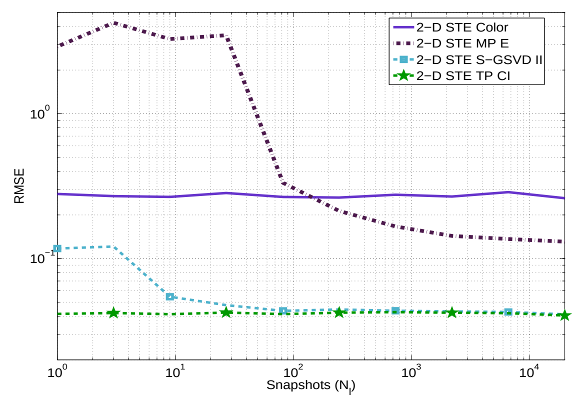


Fig. 5. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, number of snapshots N, and number of sources d as in Fig. 2. However, the SNR is fixed to 0 dB, and the snapshots N_l vary from 1 to $2 \cdot 10^4$ samples.

distributed (i.i.d.), we can show that

$$\mathrm{E}\{[\boldsymbol{\mathcal{N}}]_{(i)} \cdot (\boldsymbol{L}_{i+1} \otimes \ldots \otimes \boldsymbol{L}_R \otimes \boldsymbol{I}_N \otimes \boldsymbol{L}_1 \ldots \otimes \boldsymbol{L}_{i-1}) \cdot (\boldsymbol{L}_{i+1} \otimes \ldots \otimes \boldsymbol{L}_R \otimes \boldsymbol{I}_N \otimes \boldsymbol{L}_1 \ldots \otimes \boldsymbol{L}_{i-1})^{\mathrm{H}} \cdot [\boldsymbol{\mathcal{N}}]^{\mathrm{H}}_{(i)}\} = \alpha \cdot \boldsymbol{I}_{M_i}, \tag{35}$$

Therefore, replacing (35) in (34), we realize that $\mathrm{E}\left\{[\boldsymbol{\mathcal{N}}^c]_{(i)} \cdot [\boldsymbol{\mathcal{N}}^c]^{\mathrm{H}}_{(i)}\right\} = \alpha \cdot \boldsymbol{L}_i \cdot \boldsymbol{L}_i^H$. The proof for (35) is given in Lemma 1.

Lemma 1. *Let us consider a complex Gaussian random matrix $\boldsymbol{N} \in \mathbb{C}^{M \times N}$ and also a constant matrix $\boldsymbol{F} \in \mathbb{C}^{N \times T}$, where $n_{i,j}$ indicates the element of $\boldsymbol{N}$ in row i and column j, and $f_{i,j}$ indicates the element of $\boldsymbol{F}$ in row i and column j, respectively. The following equality holds*

$$\mathrm{E}\{\boldsymbol{N} \cdot \boldsymbol{F} \cdot \boldsymbol{F}^{\mathrm{H}} \cdot \boldsymbol{N}^{\mathrm{H}}\} = \alpha \cdot \boldsymbol{I}_M. \tag{36}$$

where $\alpha = \mathrm{tr}(\boldsymbol{F} \cdot \boldsymbol{F}^{\mathrm{H}}) \cdot \sigma_n^2$

Proof: First let us represent the product of two matrices as a sum of rank-one matrices as follows:

$$\boldsymbol{N} \cdot \boldsymbol{F} = \sum_{k=1}^{N} \begin{bmatrix} n_{1,k} \\ n_{2,k} \\ \vdots \\ n_{M,k} \end{bmatrix} \cdot [f_{k,1} \quad f_{k,2} \quad \ldots \quad f_{k,T}]. \tag{37}$$

Also the term $(\boldsymbol{N} \cdot \boldsymbol{F})^{\mathrm{H}}$ can be written as

$$(\boldsymbol{N} \cdot \boldsymbol{F})^{\mathrm{H}} = \boldsymbol{F}^{\mathrm{H}} \cdot \boldsymbol{N}^{\mathrm{H}} = \sum_{p=1}^{N} \begin{bmatrix} f^*_{p,1} \\ f^*_{p,2} \\ \vdots \\ f^*_{p,T} \end{bmatrix} \cdot [n^*_{1,p} \quad n^*_{2,p} \quad \ldots \quad n^*_{M,p}]. \tag{38}$$

Therefore, the product of (37) and (38) can be represented as the

see the program

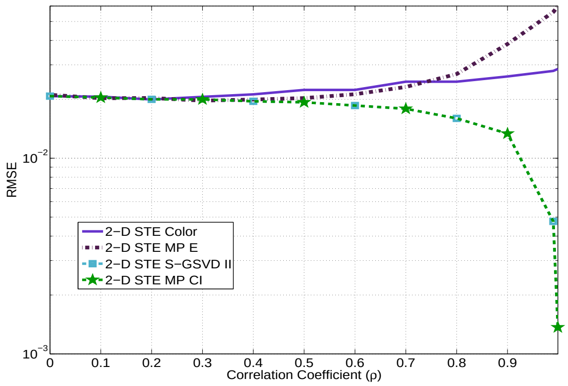


Fig. 6. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, number of snapshots N, snapshots N_l, and number of sources as in Fig. 2. However, the SNR is fixed to 10 dB, and the correlation levels ρ_1 and ρ_2 are equal to ρ, where $0 \leq \rho < 1$.

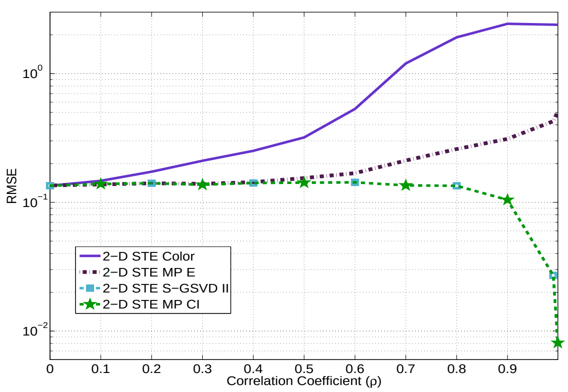


Fig. 7. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, number of snapshots N, snapshots N_l, and number of sources d as in Fig. 2. However, the SNR is fixed to -5 dB, and the correlation levels ρ_1 and ρ_2 are equal to ρ, where $0 \leq \rho < 1$.

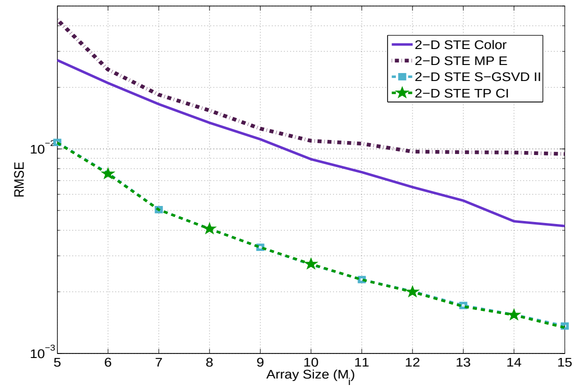


Fig. 8. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same correlation levels ρ_1 and ρ_2, number of snapshots N, snapshots N_l, and number of sources d as in Fig. 2. However, the SNR is fixed to 10 dB, and the array size M_1 and M_2 equal to M_i, which varies from 5 to 15.

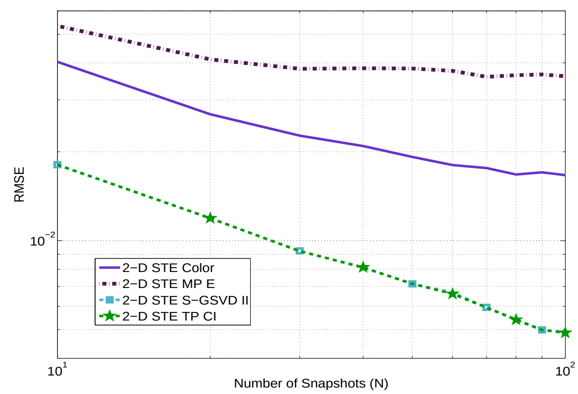


Fig. 9. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, snapshots N_l, correlation levels ρ_1 and ρ_2, and number of sources d as in Fig. 2. However, the SNR is fixed to 10 dB, and the number of snapshots N varies from 10 to 100.

product of two sums of matrices:

$$\boldsymbol{N} \cdot \boldsymbol{F} \cdot \boldsymbol{F}^{\mathrm{H}} \cdot \boldsymbol{N}^{\mathrm{H}} =$$
$$\sum_{k=1}^{N} \begin{bmatrix} n_{1,k} \\ n_{2,k} \\ \vdots \\ n_{M,k} \end{bmatrix} \cdot \begin{bmatrix} f_{k,1} \\ f_{k,2} \\ \vdots \\ f_{k,T} \end{bmatrix}^{\mathrm{T}} \cdot \sum_{p=1}^{N} \begin{bmatrix} f_{p,1}^{*} \\ f_{p,2}^{*} \\ \vdots \\ f_{p,T}^{*} \end{bmatrix} \cdot \begin{bmatrix} n_{1,p}^{*} \\ n_{2,p}^{*} \\ \vdots \\ n_{M,p}^{*} \end{bmatrix}^{\mathrm{T}} . \tag{39}$$

In order to facilitate the derivation, we separate (39) into two terms. The first term is represented by $\boldsymbol{B}_1$ for the cases when $p = k$, and the second term is represented by $\boldsymbol{B}_2$ for the cases when $p \neq k$.

$$\boldsymbol{N} \cdot \boldsymbol{F} \cdot \boldsymbol{F}^{\mathrm{H}} \cdot \boldsymbol{N}^{\mathrm{H}} = \boldsymbol{B}_1 + \boldsymbol{B}_2, \tag{40}$$

where the first term is given by

$$\boldsymbol{B}_1 = \sum_{k=1}^{N} \begin{bmatrix} n_{1,k} \\ n_{2,k} \\ \vdots \\ n_{M,k} \end{bmatrix} \cdot \begin{bmatrix} f_{k,1} \\ f_{k,2} \\ \vdots \\ f_{k,T} \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} f_{k,1}^{*} \\ f_{k,2}^{*} \\ \vdots \\ f_{k,T}^{*} \end{bmatrix} \cdot \begin{bmatrix} n_{1,k}^{*} \\ n_{2,k}^{*} \\ \vdots \\ n_{M,k}^{*} \end{bmatrix}^{\mathrm{T}} . \tag{41}$$

Since the product between the two vectors in the middle gives a scalar, we can rewrite (41) in the following way

$$\boldsymbol{B}_1 = \sum_{k=1}^{N} \sum_{t=1}^{T} |f_{k,t}|^2 \cdot \begin{bmatrix} n_{1,k} \\ n_{2,k} \\ \vdots \\ n_{M,k} \end{bmatrix} \cdot \begin{bmatrix} n_{1,k}^{*} \\ n_{2,k}^{*} \\ \vdots \\ n_{M,k}^{*} \end{bmatrix}^{\mathrm{T}} . \tag{42}$$

Applying the expected value operator, we obtain

$$\mathrm{E}\{\boldsymbol{B}_1\} = \sum_{k=1}^{N} \sum_{t=1}^{T} |f_{k,t}|^2 \cdot \sigma_n^2 \cdot \boldsymbol{I}_M . \tag{43}$$

While applying the expected value operator to $\boldsymbol{B}_2$

$$\mathrm{E}\{\boldsymbol{B}_2\} = \boldsymbol{0}_{M \times M}, \tag{44}$$

since the noise samples $n_{i,j}$ are uncorrelated. In addition, the expected value of the sum is equal to the sum of the expected

see the program

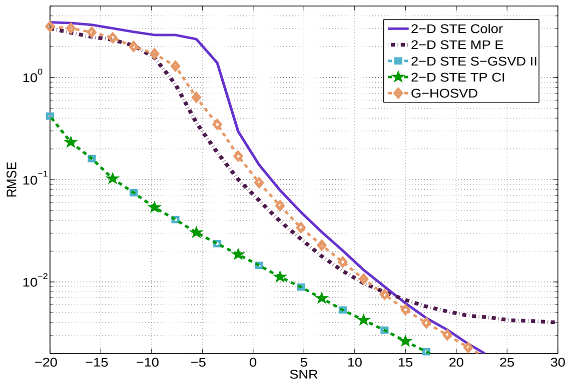


Fig. 10. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, snapshots N_l, and number of snapshots N as in Fig. 2. However, the number of sources d is equal to 2, and we set $\rho_1 = 0.9$ and $\rho_2 = 0$.

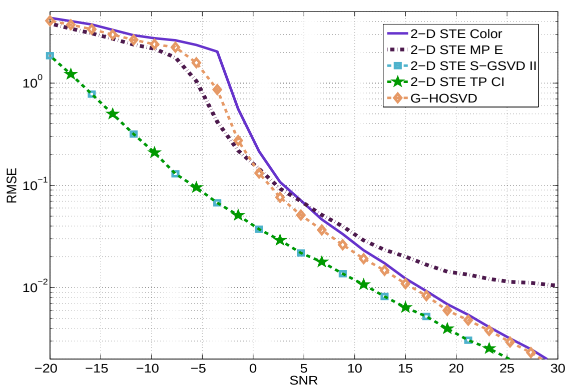


Fig. 11. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, snapshots N_l, number of snapshots N, and correlation levels ρ_1 and ρ_2 as in Fig. 10. However, the number of sources d is set to 4.

values. Therefore, we can apply the expected value to each matrix $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ in (40) separately. Moreover, using that $\sum_{k=1}^{N}\sum_{t=1}^{T}|f_{k,t}|^2 = \mathrm{tr}(\boldsymbol{F}\cdot\boldsymbol{F}^{\mathrm{H}})$, we find that

$$\mathrm{E}\{\boldsymbol{N}\cdot\boldsymbol{F}\cdot\boldsymbol{F}^{\mathrm{H}}\cdot\boldsymbol{N}^{\mathrm{H}}\} = \mathrm{tr}(\boldsymbol{F}\cdot\boldsymbol{F}^{\mathrm{H}})\cdot\sigma_n^2\cdot\boldsymbol{I}_M. \tag{45}$$

Since $\mathrm{tr}(\boldsymbol{F}\cdot\boldsymbol{F}^{\mathrm{H}})$ and σ_n^2 are constants, just one constant α can represent them. ■

The expression in (35) is shown by inserting $\boldsymbol{F} = (\boldsymbol{L}_{i+1}\otimes\ldots\otimes\boldsymbol{L}_R\otimes\boldsymbol{I}_N\otimes\boldsymbol{L}_1\ldots\otimes\boldsymbol{L}_{i-1})\cdot(\boldsymbol{L}_{i+1}\otimes\ldots\otimes\boldsymbol{L}_R\otimes\boldsymbol{I}_N\otimes\boldsymbol{L}_1\ldots\otimes\boldsymbol{L}_{i-1})^{\mathrm{H}}$ and $\boldsymbol{N} = [\boldsymbol{\mathcal{N}}]_{(i)}$ in Lemma 1.

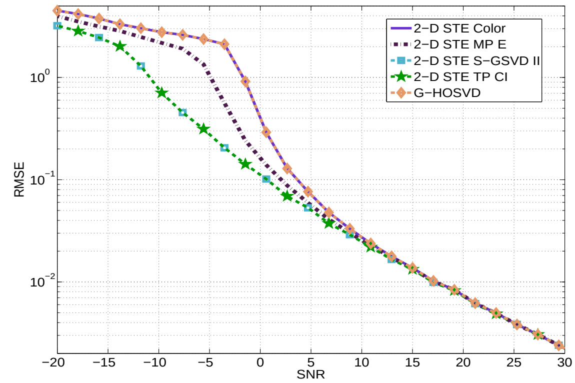


Fig. 12. Comparing the different types of prewhitening schemes in conjunction with 2-D Standard Tensor-ESPRIT for the estimation of the spatial frequencies. Here we consider the same array size M_1 and M_2, snapshots N_l, number of snapshots N, and correlation levels ρ_1 and ρ_2 as in Fig. 10. However, the number of sources d is set to 5.

REFERENCES

[1] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition", *SIAM J. Matrix Anal. Appl.*, vol. 21(4), 2000.

[2] M. Haardt and J. A. Nossek, "Simultaneous Schur decomposition of several non-symmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems", *IEEE Trans. Signal Processing*, vol. 46, pp. 161–169, Jan. 1998.

[3] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems", *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3198 – 3213, July 2008.

[4] M. Haardt, R. S. Thomä, and A. Richter, "Multidimensional high-resolution parameter estimation with applications to channel sounding", in *High-Resolution and Robust Signal Processing*, pp. 255–338. Marcel Dekker, New York, NY, 2004, Chapter 5.

[5] P. C. Hansen and S. H. Jensen, "Prewhitening for rank-deficient noise in subspace methods for noise reduction", *IEEE Trans. Signal Processing*, vol. 53, pp. 3718–3726, Oct. 2005.

[6] H. M. Huizenga, J. C. de Munck, L. J. Waldorp, and R. P. P. P. Grasman, "Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model", *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 6, pp. 533 – 539, June 2002.

[7] B. Park and T. F. Wong, "Training sequence optimization in MIMO systems with colored noise", in *Military Communications Conference (MILCOM 2003)*, Gainesville, USA, Oct. 2003.

[8] R. Roy and T. Kailath, "ESPRIT - Estimation of signal parameters via rotational invariance techniques", in *Signal Processing Part II: Control Theory and Applications*, L. Auslander, F. A. Grünbaum, J. W. Helton, T. Kailath, P. Khargonekar, and S. Mitter, Eds., pp. 369–411. Springer-Verlag, 1990.

[9] J. Vandewalle, L. De Lathauwer, and P. Comon, "The generalized higher order singular value decomposition and the oriented signal-to-signal ratios of pairs of signal tensors and their use in signal processing", in *European Conference on Circuit Theory and Design*, Cracow, Poland, Sept. 2003.

see the program

PERFORMANCE ASSESSMENT OF THE SAGE ALGORITHM IN UNKNOWN NOISE FIELDS

Michail Matthaiou[1], *Felix Antreich*[2], *and Josef A. Nossek*[1]

[1] Institute for Circuit Theory and Signal Processing, Technische Universität München (TUM),
Arcistrasse 21, 80 333, Munich, Germany
email:{matthaiou,nossek}@nws.ei.tum.de
[2] Institute for Communications and Navigation, German Aerospace Center (DLR),
82 234, Wessling, Germany
email:felix.antreich@dlr.de

ABSTRACT

In this paper, we explore the performance of the Space Alternating Generalized Expectation-Maximization (SAGE) algorithm in estimating the parameters of multipath components in wireless propagation channels. The main goal of this contribution is to extend the theoretical framework of the algorithm to account for the general case of an arbitrary unknown noise field with non-uniform structure. This scenario is of high practical importance since it encounters in several modern applications and, more importantly, inherently encompasses the common case of white uniform additive noise. In the sequel, it is thoroughly demonstrated how the accuracy and convergence rate of the estimators are affected by the spatio-temporal parameters of the incoming wavefronts.

1. INTRODUCTION

It is an indisputable fact the construction of efficient physical channel models as well as the optimum design of multi-antenna systems, requires a detailed knowledge of the statistical distributions of the multipath parameters. Hence, it is of vital importance to identify the dominant multipath components (MPCs) impinging on the receive unit and thereafter extract their spatio-temporal characteristics, such as Time of Arrival (ToA), Angle of Arrival (AoA), Doppler frequency and complex amplitude. At a next stage, we can obtain a geometrical projection of the signal's interactions with the surrounding environment and thereafter work out the dominant propagation mechanisms. Generally speaking, this demanding task can normally be accomplished with the aid of high-resolution multi-dimensional array processing algorithms that are directly applied on the stochastic channel response.

Among the numerous parametric methods for signal parameter estimation, the Space Alternating Generalized Expectation Maximization (SAGE) algorithm has been extensively employed thanks to its accuracy and rapid convergence [1]. In contrast to the conventional estimation techniques such as MUSIC (MUltiple Signal Classification) and ESPRIT (Estimation of Signal Parameters via Rotational Invariance Technique) which perform badly in the presence of correlated signals, the SAGE algorithm is robust and yields precise estimates even at low Signal-to-Noise ratio (SNR) [2]. What is more, no spatial smoothing is required to mitigate the effects of correlated signal sources. Equally importantly, the SAGE algorithm is applicable to any arbitrary antenna geometry provided that the array manifold is fully available.

The SAGE algorithm is in principle a twofold extension of the well-known Expectation Maximization (EM) algorithm which computes the maximum-likelihood (ML) estimators of the unknown parameters in a sequential way. In other words, it replaces the computationally prohibitive high-dimensional non-linear optimization process by several low-dimensional maximization procedures. In each SAGE iteration, only a subset of the parameters is updated, while keeping the estimators of the remaining parameters fixed. The derivation of the algorithm still relies on the key notions of *complete* (unobservable) and *incomplete* (observable) data.

The scheme was initially applied to the problem of multipath parameter estimation in [3, 4] and since then has undergone a plethora of extensions (see for instance [5, 6, 7, 8, 9] and references therein) to account for additional selectivity domains (polarization and elevation). Most studies dealing with the SAGE algorithm, consider the tractable case of uniformly distributed white noise which significantly simplifies the derivations included throughout the execution of the algorithm. In practical applications though (like passive sonar, aeronautical and mobile communications), where no signal-free samples can be used for estimating the noise-covariance, this condition is quite often violated [10].

Under these circumstances, the additive noise term may be not only spatially correlated (as for instance between adjacent antenna elements or in the presence of wideband interference), but also correlated in the time domain. The latter phenomenon may be caused by either weak residual compo-

see the program

nents with similar temporal characteristics as the dominant MPCs or by random perturbations of the propagated wavefronts as they travel through inhomogeneous media [11]. On this basis, the main scope of this paper is to extend the mathematical background of the SAGE algorithm to the generalized case of unknown noise fields with arbitrary covariance matrix and, further, assess its performance against various spatio-temporal parameters of interest.

The rest of the paper is organized as follows: In Section 2, the channel model used throughout the paper is introduced with a view to the propagation characteristics. Section 3 discusses the novel mathematical background of the extended SAGE algorithm. In Section 4, a set of numerical results is presented in order to get a deeper insight into the algorithm's performance. Finally, Section 5 concludes the paper and proposes some directions for additional research in the future.

A note on notation: We use upper and lower case boldfaces to denote matrices and vectors, respectively while $\Re\{\cdot\}$ will return the real part of a complex number. The symbol $\sim \mathcal{CN}(\mathbf{A},\mathbf{B})$ stands for a complex normally distributed matrix with mean $\mathbf{A}$ and covariance $\mathbf{B}$. The symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ and $(\cdot)^{-1}$ correspond to transposition, complex conjugate, Hermitian transposition and matrix inversion, respectively, whereas $\odot$ is the Hadamard-Schur product and $\otimes$ the Kronecker product. Finally, $\det(\cdot)$ and $\mathrm{tr}(\cdot)$ respectively express the determinant and trace of a matrix, whereas the $\mathrm{vec}(\cdot)$ operator stacks the columns of a matrix into a vector.

2. SYSTEM MODEL

Assuming that all signal sources and scatterers are located in the corresponding far field regions, the radio channel between the transmitter (Tx) and receiver (Rx) can be effectively modeled as the superposition of L propagation paths. For a channel with M receive antennas, the received signal matrix $\mathbf{Y} \in \mathbb{C}^{M\times N}$ across N time samples reads as

$$\mathbf{Y} = \sum_{\ell=1}^{L} \gamma_\ell \boldsymbol{\alpha}(\phi_\ell) \Big(\underbrace{\mathbf{d}(\nu_\ell) \odot \mathbf{c}(\tau_\ell)}_{=\mathbf{z}(\nu_\ell,\tau_\ell)} \Big)^T + \mathbf{N} \qquad (1)$$

where γ_ℓ, τ_ℓ, ϕ_ℓ, and ν_ℓ are the complex amplitude, ToA, AoA and Doppler frequency of the ℓ-th wave respectively. In addition, $\boldsymbol{\alpha}(\phi_\ell)$ is the azimuthal spatial steering vector, $\mathbf{d}(\nu_\ell)$ contains the sampled Doppler frequency and $\mathbf{c}(\tau_\ell)$ the sampled pseudo-random binary sequence of the ℓ-th incoming wavefront. It is worth mentioning that the delay and Doppler effects are jointly modeled through their so-called *Doppler-code dot product* $\mathbf{z}(\nu_\ell,\tau_\ell)$. With regard to the noise-plus-interference term $\mathbf{N}$, we hereafter assume it follows the following complex zero-mean Gaussian distribution

$$\mathbf{N} \sim \mathcal{CN}(\mathbf{0},\mathbf{Q}) \qquad (2)$$

where $\mathbf{Q} \in \mathbb{C}^{M\times M}$ denotes the spatial correlation matrix which is, in general, non-diagonal thereby indicating the non-uniformity of noise. Referring back to (1), we can rewrite it in a more compact form, using matrix notation, as follows

$$\mathbf{Y} = \sum_{\ell=1}^{L} \mathbf{S}_\ell(\boldsymbol{\theta}_\ell) + \mathbf{N} = \mathbf{S}(\boldsymbol{\theta}) + \mathbf{N} \qquad (3)$$

where

$$\mathbf{S}_\ell(\boldsymbol{\theta}_\ell) = \gamma_\ell \boldsymbol{\alpha}(\phi_\ell)\mathbf{z}(\nu_\ell,\tau_\ell)^T \qquad (4)$$

and the matrix $\mathbf{S}(\boldsymbol{\theta}) \in \mathbb{C}^{M\times N}$ contains the superimposed impinging wavefronts $\sum_{\ell=1}^{L}\mathbf{S}_\ell(\boldsymbol{\theta}_\ell)$. The spatio-temporal parameters of each distinct path are concatenated into the vector $\boldsymbol{\theta}_\ell \triangleq [\gamma_\ell,\tau_\ell,\phi_\ell,\nu_\ell]$ with $\boldsymbol{\theta} = [\boldsymbol{\theta}_1,\boldsymbol{\theta}_2,\ldots,\boldsymbol{\theta}_L]^T$. From inspection of (1) and (3), we can easily infer that $\mathbf{Y}$ is identified as the incomplete data and is related to the complete data, $\mathbf{X}_\ell$, via the following relationship

$$\mathbf{Y} = \sum_{\ell=1}^{L} \mathbf{X}_\ell \qquad (5)$$

where

$$\mathbf{X}_\ell = \mathbf{S}_\ell(\boldsymbol{\theta}_\ell) + \beta_\ell \mathbf{N}_\ell. \qquad (6)$$

In general, the real and positive coefficients β_ℓ represent the contribution of each noise component.

3. THE SAGE ALGORITHM

Using (6) as a starting point, we recall that in the SAGE context it is optimal to set $\beta_\ell = 1$ since empirical evidence has shown that this choice leads to a fast convergence of the algorithm even in the early iteration steps [1, 4]. At a next stage, the likelihood function for the signal model can be modeled via the conditional probability density function (PDF) of the complex multivariate Gaussian according to [12]

$$p(\mathbf{Y};\boldsymbol{\theta}) = \frac{\exp\left\{-\mathrm{vec}\left(\mathbf{Y}-\mathbf{S}(\boldsymbol{\theta})\right)^H \mathbf{Q}^{-1}\mathrm{vec}\left(\mathbf{Y}-\mathbf{S}(\boldsymbol{\theta})\right)\right\}}{\pi^{MN}\det(\mathbf{Q})}. \qquad (7)$$

The negative log-likelihood function, $\mathcal{L}(\mathbf{Y};\boldsymbol{\theta})$, then reads

$$\mathcal{L}(\mathbf{Y};\boldsymbol{\theta}) = N\log(\det(\mathbf{Q})) + \mathrm{tr}\left(\mathbf{Q}^{-1}\left(\mathbf{Y}-\mathbf{S}(\boldsymbol{\theta})\right)\left(\mathbf{Y}-\mathbf{S}(\boldsymbol{\theta})\right)^H\right) \qquad (8)$$

where we have dropped out all constant terms throughout the derivations and also made use of the following key matrix identities [13]

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{vec}\left(\mathbf{A}^H\right)^H \mathrm{vec}\left(\mathbf{B}\right) \qquad (9)$$

$$\mathrm{vec}(\mathbf{ABC}) = \left(\mathbf{C}^H \otimes \mathbf{A}\right)\mathrm{vec}(\mathbf{B}). \qquad (10)$$

Differentiating (8) with respect to $\mathbf{Q}$ and taking into account the following properties,

$$\frac{\partial \log|\det(\mathbf{A})|}{\partial \mathbf{A}} = \left(\mathbf{A}^{-H}\right) \qquad (11)$$

$$\frac{\partial}{\partial \mathbf{A}}\mathrm{tr}\left(\mathbf{AC}^{-1}\mathbf{B}\right) = -(\mathbf{C}^{-1})^H\mathbf{A}^H\mathbf{B}^H(\mathbf{C}^{-1})^H \qquad (12)$$

see the program

we end up with a ML estimator for the spatially colored covariance matrix of the noise term

$$\widehat{\mathbf{Q}} = \frac{1}{N}\left(\mathbf{Y} - \mathbf{S}(\hat{\boldsymbol{\theta}})\right)\left(\mathbf{Y} - \mathbf{S}(\hat{\boldsymbol{\theta}})\right)^H. \qquad (13)$$

As was previously underlined, the main reason for the development of the SAGE algorithm was to reduce the extensive computational burden imposed by the EM scheme which updates the parameters of all waves $\widehat{\boldsymbol{\theta}}$ simultaneously. This goal can be accomplished by the more efficient process of estimating the parameters of each wave $\widehat{\boldsymbol{\theta}}_\ell$ sequentially. By doing so, an enhanced convergence rate along with a reduced complexity are achieved [1]. For the specific problem under consideration, we settle to decompose the logarithm of the PDF of the received signal vector in (7), as shown in (14) at the top of the next page.

In order to estimate the parameters of one wave, as contained in the vector $\widehat{\boldsymbol{\theta}}_\ell$, it is a common approach to match its PDF in (14) with that of an exponential family via the Cameron-Martin formula. We emphasize that this approach was initially reported in [14] and since then has been extensively used in the field of signal detection with unknown parameters [3, 4, 5, 15]. After some basic algebraic manipulations, the ML solution can be iteratively derived by solving

$$\widehat{\boldsymbol{\theta}}_\ell = \arg\max_{\boldsymbol{\theta}_\ell}\Big(2\Re\left\{\mathrm{vec}\left(\mathbf{S}_\ell(\boldsymbol{\theta}_\ell)\right)^H \mathrm{vec}\left(\widehat{\mathbf{Q}}^{-1}\widehat{\mathbf{X}}_\ell\right)\right\} - \mathrm{vec}\left(\mathbf{S}_\ell(\boldsymbol{\theta}_\ell)\right)^H \widehat{\mathbf{Q}}^{-1}\mathrm{vec}\left(\mathbf{S}_\ell(\boldsymbol{\theta}_\ell)\right)\Big) \qquad (15)$$

which after taking into account (4) may be simplified to

$$\widehat{\boldsymbol{\theta}}_\ell = \arg\max_{\boldsymbol{\theta}_\ell}\Big(2\Re\left\{\gamma_\ell^*\boldsymbol{\alpha}(\phi_\ell)^H\widehat{\mathbf{Q}}^{-1}\widehat{\mathbf{X}}_\ell\mathbf{z}(\nu_\ell,\tau_\ell)^*\right\} - |\gamma_\ell|^2\mathbf{z}(\nu_\ell,\tau_\ell)^T\mathbf{z}(\nu_\ell,\tau_\ell)^*\boldsymbol{\alpha}(\phi_\ell)^H\widehat{\mathbf{Q}}^{-1}\boldsymbol{\alpha}(\phi_\ell)\Big). \qquad (16)$$

The kernel of the algorithm consists of the so-called Expectation step (E-step) and Maximization step (M-step). During the E-step the complete data of the ℓ-th path, $\widehat{\mathbf{X}}_\ell$, is obtained by subtracting the estimated contribution of all L paths, except that of the ℓ-th path, from the incomplete data space, $\mathbf{Y}$. The estimated noise covariance matrix is also updated during this step in order to increase the overall accuracy of the algorithm. Mathematically speaking, this can be expressed as

$$\widehat{\mathbf{X}}_\ell = \mathbf{Y} - \sum_{\substack{\ell'=1\\ \ell'\neq\ell}}^{L}\mathbf{S}_{\ell'}(\hat{\boldsymbol{\theta}}'_{\ell'}) \qquad (17)$$

$$\widehat{\mathbf{Q}} = \frac{1}{N}\left(\mathbf{Y} - \mathbf{S}(\hat{\boldsymbol{\theta}}'_\ell)\right)\left(\mathbf{Y} - \mathbf{S}(\hat{\boldsymbol{\theta}}'_\ell)\right)^H. \qquad (18)$$

We point out that this approach is usually referred to in the corresponding literature as Parallel Interference Cancellation (PIC). During the M-step, the coordinate-wise updating procedure for obtaining the parameters $\hat{\boldsymbol{\theta}}''_\ell$ of the ℓ-th wave is based on all previous estimates $\hat{\boldsymbol{\theta}}'_\ell$; hence,

$$\hat{\tau}''_\ell = \arg\max_{\tau}\left(\frac{\left|\boldsymbol{\alpha}(\phi'_\ell)\widehat{\mathbf{Q}}^{-1}\widehat{\mathbf{X}}_\ell\mathbf{z}(\nu'_\ell,\tau_\ell)^*\right|^2}{D\left(\tau_\ell,\phi'_\ell,\nu'_\ell\right)}\right) \qquad (19)$$

where

$$D\left(\tau_\ell,\phi_\ell,\nu_\ell\right) = \mathbf{z}(\nu_\ell,\tau_\ell)^T\mathbf{z}(\nu_\ell,\tau_\ell)^*\boldsymbol{\alpha}(\phi_\ell)^H\widehat{\mathbf{Q}}^{-1}\boldsymbol{\alpha}(\phi_\ell) \qquad (20)$$

and likewise,

$$\hat{\phi}''_\ell = \arg\max_{\phi}\left(\frac{\left|\boldsymbol{\alpha}(\phi_\ell)\widehat{\mathbf{Q}}^{-1}\widehat{\mathbf{X}}_\ell\mathbf{z}(\nu'_\ell,\tau''_\ell)^*\right|^2}{D\left(\tau''_\ell,\phi_\ell,\nu'_\ell\right)}\right) \qquad (21)$$

$$\hat{\nu}''_\ell = \arg\max_{\nu}\left(\frac{\left|\boldsymbol{\alpha}(\phi''_\ell)\widehat{\mathbf{Q}}^{-1}\widehat{\mathbf{X}}_\ell\mathbf{z}(\nu_\ell,\tau''_\ell)^*\right|^2}{D\left(\tau''_\ell,\phi''_\ell,\nu_\ell\right)}\right). \qquad (22)$$

The execution of this update process once defines one iteration cycle of the SAGE algorithm while at the μ-th iteration step the parameters of the path $\ell = \mu \mathrm{mod}(L) + 1$ are re-estimated. The parameter estimates are sequentially and cyclically updated until convergence is obtained. The complex amplitude is then computed as the output signal normalized by the total energy, as follows

$$\hat{\gamma}''_\ell = \frac{\boldsymbol{\alpha}(\phi''_\ell)^H\widehat{\mathbf{Q}}^{-1}\widehat{\mathbf{X}}_\ell\widehat{\mathbf{K}}^{-1}\mathbf{z}(\nu''_\ell,\tau''_\ell)^*}{D\left(\tau''_\ell,\phi''_\ell,\nu''_\ell\right)}. \qquad (23)$$

Please note that the same estimator of the complex amplitude can be alternatively obtained by differentiating (16) with respect to γ_ℓ^*, setting the derivative equal to zero and solving for $\hat{\gamma}_\ell$.

3.1. Initialization of the SAGE algorithm

It is well established that the convergence of any iteration technique to a global maximum is strongly dependent upon the initial conditions. In general, convergence to a global maximum has been observed in the simulation results when the initial estimate is within a significant range of the global maximum. In the present case, the initialization procedure is based on a successive cancellation scheme (SIC) with the iteration step μ ranging from $\{-(L-1),\ldots,0\}$ and the initial setting being $\hat{\boldsymbol{\theta}} = [\mathbf{0},\ldots,\mathbf{0}]^T$. At the initialization of the ℓ-th wave the estimates related to $\ell' \geq \ell$ remain equal to 0, i.e. $\hat{\boldsymbol{\theta}}' = [\hat{\boldsymbol{\theta}}'_1,\ldots,\hat{\boldsymbol{\theta}}'_{\ell-1},\mathbf{0},\ldots,\mathbf{0}]$ [3, 6, 16]. Finally, for the covariance matrix $\mathbf{Q}$, the initial estimate reads as

$$\widehat{\mathbf{Q}} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H. \qquad (24)$$

see the program

$$\begin{aligned}\log(p(\mathbf{Y};\boldsymbol{\theta})) &= \sum_{\ell=1}^{L}\log(p(\mathbf{X}_\ell;\boldsymbol{\theta}_\ell))\\ &= \sum_{\ell=1}^{L}\log\left(\frac{\exp\left\{-\mathrm{vec}\left(\mathbf{X}_\ell-\mathbf{S}(\boldsymbol{\theta}_\ell)\right)^H\mathbf{Q}^{-1}\mathrm{vec}\left(\mathbf{X}_\ell-\mathbf{S}(\boldsymbol{\theta}_\ell)\right)\right\}}{\pi^{MN}\det(\mathbf{Q})}\right)\end{aligned} \tag{14}$$

4. NUMERICAL RESULTS

In this section, the performance of the SAGE algorithm is assessed in depth as a function of different spatial and temporal parameters. As was previously underlined, we are particularly interested in the case of spatial and temporal correlation. The former effect is inherently taken into account within the SAGE update process through the additive noise term; on the other hand, the latter is assumed to encounter with no *a priori* knowledge. This means that the algorithm does not into account temporal correlation which it is very likely to occur between closely separated MPCs and, consequently, the optimization process is carried out without considering this very important implication.

In the following, our main focus will be on Global Navigation Satellite Systems (GNSS) applications where, in general, the case of multipath propagation represents a critical hindrance since it makes the synchronization at the ground Rx laborious and thus may introduce a bias up to a hundred meters in terms of range estimation [7]. Therefore, it is of vital importance to identify the impinging multipaths in order to get a precise positioning of the satellite Tx. Apart from multipath propagation though, satellite communications are prone to strong interferers which completely mask the LoS wavefront and deteriorate the overall acquisition procedure since they raise the noise floor of the system. For this reason, this effect is also modeled in the following evaluation.

The measurement setup consists of a receive uniform linear array (ULA) with $M = 8$ isotropic radiators and interelement distance of half wavelength. The channel parameters are assumed to remain constant during the observation interval T_N while the one-sided bandwidth of the signal is taken as $B = 1.023$ MHz. For the pseudo-binary sequence, we assume Gold Codes as used in typical Global Positioning Systems (GPS) with the code period being $T = 1$ ms [17]. In total, 1023 chips per code are taken with a time duration of $T_c = 977.52$ ns. The effective SNR in dB can be directly expressed as

$$\mathrm{SNR} = C/N_0 - 10\log_{10}(2B) + 10\log_{10}(N_c) \tag{25}$$

where C/N_0 in dB-Hz denotes the carrier-to-noise density ratio and N_c the number of code periods within $T_N = N_cT$. By setting $C/N_0 = 40.3$ dB-Hz and $N_c = 6$, the effective SNR becomes equal to -15.02 dB. The propagation scenario under investigation includes a direct LoS component and a reflected wavefront, i.e. $L = 2$. Hereafter, the parameters of the former will be denoted by the subscript 1 and those of the former by the subscript 2. The LoS signal parameters are then taken as $(\tau_1, \phi_1, \nu_1) = (0\text{ ns}, -10°, 2.25\text{ Hz})$. Likewise, the parameters of the multipath component are accordingly $(\tau_2, \phi_2, \nu_2) = (0.1T_c, 43°, 2.25\text{ Hz})$. The reflected multipath and the direct component are considered to be in-phase while the signal-to-multipath ratio (SMR) is set to 5 dB. It can be easily seen that the relative Doppler difference is set equal to zero which represents the worst-case scenario. The wideband temporally white interference term has an associated AoA of $\phi_I = 10°$ and yields an interference-to-signal ratio (ISR) of 40 dB; it is further responsible for the spatial covariance of the noise plus interference term. Following a widely used approach, we herein assume that the signal and noise sources are uncorrelated with the interferer [7, 17].

At a next stage, a weaker multipath is also assumed to be present in the communication link which may not be resolved by the Rx but induces temporal correlation. The corresponding ToA and Doppler shift are respectively $(\tau_3, \nu_3) = (0.2T_c, 2.25\text{ Hz})$ whereas the SMR $= 8$ dB in this case. From a physical viewpoint, this scenario can be caused by a rough diffuse scatterer at the vicinity of the Rx which creates a cluster of rays toward all non-specular directions with identical delay characteristics. In the following, we examine the performance of the algorithm against a varying relative azimuth separation, $\Delta\phi'$, between the two MPCs; that is $\Delta\phi' = |\phi_3 - \phi_2|$ where ϕ_3 is the actual AoA of the weak MPC (with an initial value of $\phi_3 = 73°$).

In Fig. 1, the root mean square error (RMSE) of the delay estimators is illustrated along with the associated Cramér-Rao lower bounds (CRLBs). It is easily seen that the delay estimator of the LoS signal, $\hat{\tau}_1$, remains remarkably robust across the whole range of azimuth separations and is very close to the statistical CRLB. Moreover, its accuracy remains unaffected by $\Delta\phi'$, which implies that the SAGE algorithm is able to precisely extract the temporal parameter of the dominant component regardless of the model order estimation. On the other hand, the delay estimator of the MPC, $\hat{\tau}_2$, is severely degraded when the angular separation distance gets larger. This phenomenon is attributed to the underestimation of the model order L which makes the SAGE estimators biased with respect to the considered noise model (2), i.e. the weak MPC

see the program

is treated as noise contribution. In particular, the scheme to track only the nominal ToA of this cluster of rays and inevitably leads to irreducible errors. It further seems to benefit from closely separated paths which are coherently superimposed, thereby yielding identical spatio-temporal characteristics, and manages to estimate the latter with sufficiently good accuracy.

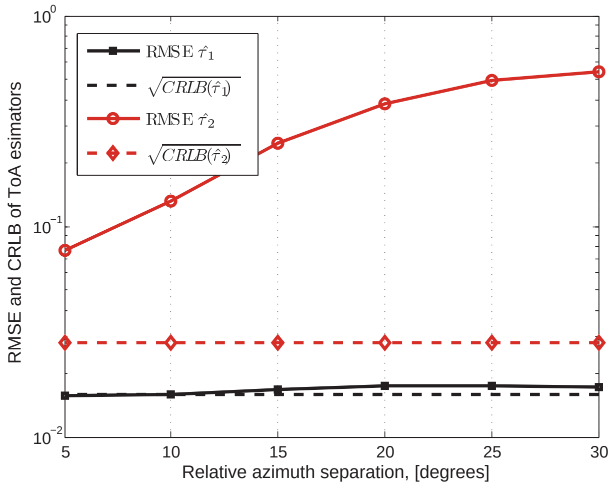


Fig. 1. RMSE of $\hat{\tau}_1$ and $\hat{\tau}_2$ against the relative azimuth separation, $\Delta\phi'$, in steps of $5°$.

In a similar manner, we now explore the AoA estimators against the relative azimuth separation (c.f. Fig. 2). It is noteworthy that the accuracy of both AoA estimators is consistently worse than the ToA estimators.

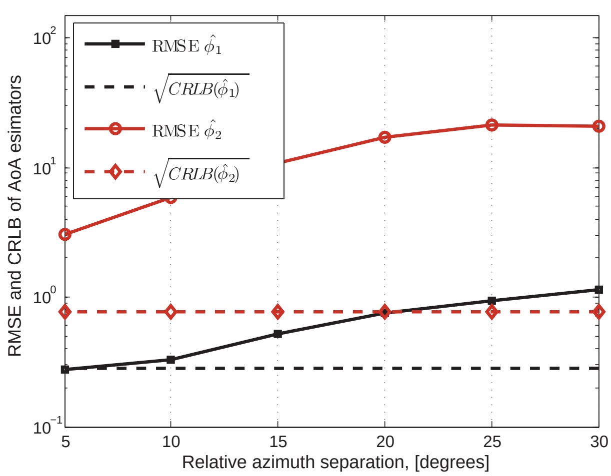


Fig. 2. RMSE of $\hat{\phi}_1$ and $\hat{\phi}_2$ against the relative azimuth separation, $\Delta\phi'$, in steps of $5°$.

This trend is anticipated as the spatial signal dimensions are inherently limited by the number of antennas whereas the temporal signal dimension is determined by the operating bandwidth. This fact is also responsible for the deteriorating performance of $\hat{\phi}_1$ with an increasing $\Delta\phi'$. In any case though, for $\Delta\phi' \leq 10°$ the RMSE$(\hat{\phi}_1)$ is able to satisfactorily approximate the CRLB.

Apart from estimation accuracy, a fundamental feature of array processing algorithms is their convergence rate. In light of this fact, the mean number of SAGE iterations has been computed as well and the obtained results are depicted in Fig. 3. It is readily observed that the algorithm benefits from low azimuth separation whereas for higher values of $\Delta\phi'$, the scheme needs more cycles to converge to the final solution due to the model order underestimation. However, it should be underlined that the number of iterations is systematically lower than six which implies that the SAGE algorithm can deliver a rapid convergence rate even under severe model mismatches.

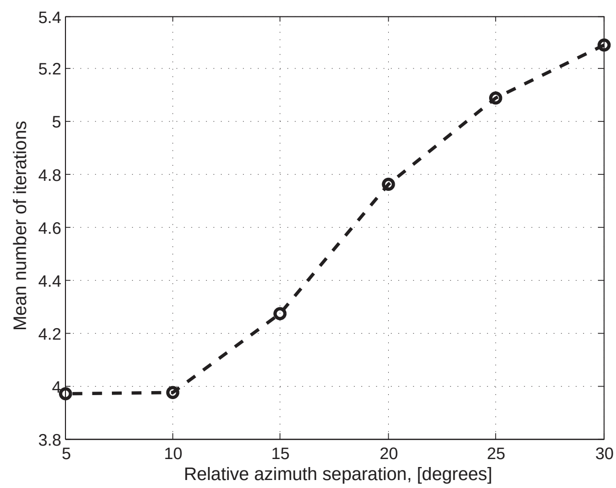


Fig. 3. Mean number of SAGE iterations against the relative azimuth separation, $\Delta\phi'$, in steps of $5°$.

Summarizing, we point out that in the presence of temporally correlated components, the estimation of the number of dominant wavefronts becomes of vital importance. In fact, in the worst-case of undetectable paths the algorithm can still yield a robust performance especially when the reflected wavefronts become coherent since in this extreme scenario the two MPCs are resolved as a single path.

5. CONCLUSION

In this paper, the mathematical background of the SAGE algorithm was extended to account for the case of arbitrary noise fields with unknown, non-diagonal underlying structure. Then, the generalized scheme was applied to the problem of multipath parameter estimation that arises in the vast majority of wireless applications and therefore is of high practical importance. Throughout the evaluation section, our main

see the program

focus was on GNSS applications where multipath propagation has been shown to cause a severe bias in range estimation. Apart from spatial correlation which was a priori included in the noise covariance matrix, we assumed that the estimation process is affected by temporal correlation as well via a residual weak MPC which is not detected by the Rx. Our analysis reveals that the LoS parameters, and especially its ToA, can be extracted with a great accuracy whereas the estimated parameters of the reflected wavefront are heavily dependent on the spatial separation $\Delta\phi'$. When the latter is getting smaller, the SAGE scheme is able to detect the spatio-temporal parameters of the impinging as the two paths are coherently superimposed. On the other hand, for an increasing $\Delta\phi'$ the scheme falls short of precisely estimating the MPC parameters due to the underestimation of the model order, L. Under these circumstances, only the nominal parameters of this cluster of rays can be resolved thereby resulting in a relatively high estimation error. Please note that similar conclusions were also drawn for the convergence rate of the algorithm.

As part of future research, a rather interesting topic to be addressed is the joint modeling of spatio-temporal correlation into the noise term. Evidently, this will represent the most general approach in the associated area and can serve as a very useful reference point to investigate the impact of time-varying, non-uniform colored noise fields on the performance of array processing algorithms.

6. REFERENCES

[1] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2664–2677, October 1994.

[2] H. Krima and M. Viberg, "Two decades of array signal processing: The parametric approach," *IEEE Signal Processing Magazine*, vol. 13, pp. 67–94, 1996.

[3] B. H. Fleury, D. Dahlhaus, R. Heddergott, and M. Tschudin, "Wideband angle of arrival estimation using the SAGE algorithm," in *Proc. International Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, Mainz, Germany, September 1996, pp. 79–85.

[4] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 438–450, March 1999.

[5] X. Yin, B. H. Fleury, P. Jourdan, and A. Stucki, "Polarisation estimation of individual propagation paths using the SAGE algorithm," in *Proc. International Symposium on Personal, Indoor and Mobile Radio Communication (PIMRC)*, Beijing, China, September 2003, pp. 1795–1799.

[6] M. Matthaiou, N. Razavi-Ghods, D. I. Laurenson, and S. Salous, "Characterization of an indoor MIMO channel in frequency domain using the 3D-SAGE algorithm," in *Proc. International Conference on Communications (ICC)*, Glasgow, U.K., June 2007, pp. 5868–5872.

[7] F. Antreich, J. A. Nossek, and W. Utschick, "Maximum likelihood delay estimation in a navigation receiver for aeronautical applications," *Elsevier Aerospace Science and Technology*, vol. 12, pp. 256–267, 2008.

[8] N. Czink, G. Del Galdo, X. Yin, and C. F. Mecklenbräuker, "A novel environment characterisation metric for clustered MIMO channels used to validate a SAGE parameter estimator," *IST Mobile and Wireless Summit*, Myconos, Greece, July 2006.

[9] M. Matthaiou, D. I. Laurenson, and J. S. Thompson, "Detailed characterisation of an indoor MIMO channel in the double directional spatial domain," *in press IET Microwaves, Antennas & Propagation*, vol. 3, no. 2, March 2009.

[10] H. Ye and R. D. Degroat, "Maximum likelihood DOA estimation and asymptotic Cramér-Rao bounds for additive unknown colored noise," *IEEE Transactions on Signal Processing*, vol. 43, no. 4, pp. 938–949, April 1995.

[11] J. Ringelstein, A. B. Gershman, and J. F. Böhme, "Direction finding in random inhomogeneous media in the presence of multiplicative noise," *IEEE Signal Processing Letters*, vol. 7, no. 10, pp. 269–272, October 2000.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1, Prentice-Hall PTR, 1993.

[13] A. Graham, *Kronecker products and matrix calculus with applications*, Ellis Horwood, London, 1981.

[14] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 2nd edition, 1994.

[15] F. A. Dietrich, "A tutorial on channel estimation with SAGE," *Technical Report TUM-LNS-TR-06-03*, 2006.

[16] C. C. Chong *et al.* "Joint detection-estimation of directional channel parameters using the 2-D frequency domain SAGE algorithm with serial interference cancellation," in *Proc. International Conference on Communications (ICC)*, vol. 2, New York, USA, April 2002, pp. 906–910.

[17] B. W. Parkinson and J. J. Spilker, *Global Positioning System: Theory and Applications*, vol. 1, Progress in Astronautics and Aeronautics, 1996.

see the program

BOUNDS ON THE CAPACITY OF MISO CHANNELS WITH DIFFERENT TYPES OF IMPERFECT CSI

Mario Castañeda, Amine Mezghani and Josef A. Nossek

Lehrstuhl für Netzwerktheorie und Signalverarbeitung
Technische Universität München, Arcisstraße 21, 80333 Munich
{castaneda,mezghani,nossek}@nws.ei.tum.de

ABSTRACT

Consider the downlink of an isolated cell with a *base station* (BS) equipped with M multiple antennas and a single user with a single antenna. For this *multiple-input single-output* (MISO) channel, the capacity with perfect *channel state information* (CSI) the transmitter and receiver is known. In practice, however, it is actually more realistic to assume *imperfect* CSI at the transmitter. For instance, in a *frequency divsion duplex* (FDD) system with limited feedback, the available transmit CSI is *estimated, quantized, outdated* and affected by *erroneous* feedback. The quality of the transmit CSI then depends on the downlink training length, the number of feedback bits, the feedback error probability and on the rate of variation of the downlink channel. However, the capacity of the MISO channel with imperfect CSI in general is unknown. In this work, nonetheless, we compute bounds of the MISO channel capacity employing *coherent beamforming* at the BS with different types of imperfect CSI at the transmitter.

1. INTRODUCTION

Multiple transmit and receive antennas increase substantially the capacity of wireless communication links but under the assumption of perfect channel knowledge at the transmitter and receiver. In practice, however, the channel knowledge available at the transmitter and receiver is not perfect. For instance, the receiver can obtain a *minimum mean square error* (MMSE) estimate of the channel through training [1], and hence, we assume the available receive CSI to be *estimated*. In case of a *time division duplex* (TDD) system, the BS can employ the receive CSI from the uplink by making use of the channel reciprocity, such that the available transmit CSI is *estimated* and possibly outdated. Then, the quality of the transmit CSI depends on the training length employed in the uplink to estimate the channel and on the rate of variation of the channel.

In an FDD system, the availabe receive CSI in the uplink cannot be employed as transmit CSI in the downlink, since we assume the downlink and uplink channels to be uncorrelated. In this case, we assume the transmitter obtains partial knowledge of the downlink channel through *limited* feedback of B bits in the uplink [2]. To this end, the receiver in the downlink, i.e., the user, first estimates the downlink channel and then *quantizes* the downlink channel estimate with B bits, which are relayed back to the BS. For this, we employ the *random vector quantization* (RVQ) scheme [2], where we have at the transmitter (BS) and receiver (user) a codebook with 2^B random beamforming vectors i.i.d over the M-dimensional unit sphere. Assuming a faded- and noise-prone feedback link, i.e., the uplink, the B relayed bits could be received *erroneously* at the BS, which reduces the quality of the available CSI at the transmitter. Besides the previous issues, the transmit CSI is further degraded due to the delay incurred in the feedback process, i.e., the partial knowledge the BS has obtained at a given time slot can only be employed at a later time slot, when the channel could have changed. Hence, the transmit CSI is also *outdated* under the assumption of a time-varying channel. In this way, the transmit CSI available at the BS for the downlink transmission is *estimated, quantized, outdated* and affected by *erroneous* feedback. Such a characterization of the imperfect transmit CSI in FDD systems has also been assumed in [3], but has not been considered in general in the literature.

The capacity with imperfect CSI, such as the different types described above, is unknown in general and in such cases, one can instead recur to the computation of bounds on the capacity. In this work, we present bounds on the capacity of a single-user MISO channel with *coherent beamforming* at the BS in an FDD system with imperfect transmit CSI, which is estimated, quantized, outdated and affected by erroneous feedback, as a function of the downlink training length, the number of feedback bits B, the feedback error probability in the uplink and the rate of variation of the downlink channel. Additionally, we compute bounds on the capacity in a TDD system with estimated transmit CSI as a function of the uplink training length and of the rate of variation of the channel. As a reference we also include the capacities with perfect CSI and without CSI at the transmitter. To this end, this paper is organized as follows. In Section 2, the system model and the types of imperfect CSI are discussed. The capacity bounds are derived in Section 3. Section 4 presents numerical results and finally, Section 5 concludes the paper.

see the program

2. SYSTEM MODEL WITH IMPERFECT CSI

We assume that the downlink channel is constant for T symbols, which is the duration of a time slot. The first T_{DL} symbols are used to obtain an MMSE estimate of the downlink channel and the remaining $D = T - T_{\text{DL}}$ symbols are used to transmit information from the BS to the user. With coherent beamforming at the BS, the equivalent *single-input single-output* (SISO) downlink system of the D data symbols at time slot n is given by

$$\mathbf{y}[n] = \sqrt{P}\,\mathbf{w}_{\text{BS}}^{\text{H}}[n]\mathbf{h}[n]\mathbf{r}[n] + \mathbf{v}[n], \tag{1}$$

where $\mathbf{y}[n] \in \mathbb{C}^D$ are the received signals, $\mathbf{r}[n] \in \mathbb{C}^D$ are the transmit symbols with unit variance, $\mathbf{w}_{\text{BS}}[n] \in \mathbb{C}^M$ is the beamforming vector with unit norm, $\mathbf{h}[n] \in \mathbb{C}^M$ is the downlink MISO channel and $\mathbf{v}[n] \in \mathbb{C}^D$ is the *additive white Gaussian noise* (AWGN) with zero mean and variance σ_v^2 at time slot n. Additionally, P is the transmit power available at the BS. We assume Rayleigh fading, i.e. the elements of the channel vector $\mathbf{h}[n]$ are i.i.d. complex Gaussian random variables with zero mean and unit variance. Contrary to the common block fading assumption found in the literature, in this paper we consider *temporally correlated block fading*, i.e. $\mathbf{h}[n]$ is assumed to be constant for the coherence time of T symbols and is correlated with the channel in the previous time slot $\mathbf{h}[n-1]$ according to a first order Markov model:

$$\mathbf{h}[n] = \sqrt{\alpha}\,\mathbf{h}[n-1] + \sqrt{1-\alpha}\,\mathbf{g}[n-1], \tag{2}$$

where the elements of $\mathbf{g}[n-1] \in \mathbb{C}^M$ are i.i.d. zero-mean unit-variance complex Gaussian random variables and are uncorrelated with $\mathbf{h}[n-1]$ and $\sqrt{\alpha}$ is the correlation coefficent, with $\alpha \in [0,1]$. For instance, with a Jake's spectrum of the Rayleigh fading we would have that $\sqrt{\alpha} = \mathrm{J}_0(2\pi T f_{\text{c}} s/c)$, where J_0 is the zeroth-order Bessel function of the first kind, f_{c} is the carrier frequency, c is the speed of light and s is the radial component of the velocity of the user, along a line from the BS to the user. Hence, α is a function of the velocity of the user. However, note that the autocorrelation function of the fading process with a Jake's spectrum cannot be modeled by a first order Markov model and it has solely been shown here to present a possible relation between α and the user's velocity. Also, we assume that α is unknown at the transmitter. In the following, we discuss different types of imperfect CSI in the context of the downlink of an FDD system and explain how the beamforming vector $\mathbf{w}_{\text{BS}}[n]$ is determined.

2.1. Estimated CSI

At each time slot, T_{DL} pilot symbols are employed to obtain an MMSE estimate $\hat{\mathbf{h}}$ of the downlink channel. Due to practical limitations, we assume an even power distribution, i.e. P, over the training and data payload phase. Based on the downlink channel estimate $\hat{\mathbf{h}}[n]$, the downlink channel $\mathbf{h}[n]$ can be written as

$$\mathbf{h}[n] = \hat{\mathbf{h}}[n] + \mathbf{e}[n], \tag{3}$$

where $\mathbf{e}[n]$ is the error vector whose elements are i.i.d. zero-mean complex Gaussian random variables with variance σ_e^2 given by [1]

$$\sigma_e^2 = \begin{cases} \dfrac{1+\rho(M-T_{\text{DL}})}{1+\rho M} & \text{for} \quad T_{\text{DL}} < M \\ \dfrac{1}{1+\rho T_{\text{DL}}} & \text{for} \quad T_{\text{DL}} \geq M \end{cases}, \tag{4}$$

where $\rho = \frac{P}{M\sigma_v^2}$. Additionally, the elements of the MMSE estimate $\hat{\mathbf{h}}$ are i.i.d. zero-mean Gaussian random variables with variance $(1-\sigma_e^2)$. Note that the quality of the estimated CSI depends on the training length T_{DL}. As the training length increases, i.e. $T_{\text{DL}} \to T$ the estimated CSI becomes closer to perfect CSI, i.e., $\sigma_e^2 \to 0$. Nonetheless, $D \to 0$ and there will be less symbols available to transmit data, thus reducing the downlink capacity. We then have a tradeoff between the training length and the data payload length as discussed in [1].

Remark 2.1: Estimation in TDD Systems. Consider a TDD system where the uplink utilizes the channel at time slot $n-1$ and the downlink makes use of the channel at time slot n. We assume that the BS does not know α and hence the BS employs the estimate of the uplink channel as transmit CSI for the downlink transmission. Denoting the uplink channel estimate at time slot $n-1$ as $\hat{\mathbf{h}}_{\text{UL}}[n-1]$, we have that the employed beamforming vector in the downlink is $\mathbf{w}_{\text{BS}}[n] = \frac{\hat{\mathbf{h}}_{\text{UL}}[n-1]}{\|\hat{\mathbf{h}}_{\text{UL}}[n-1]\|_2}$. However note that estimating the uplink *single-input multiple-output* (SIMO) channel is not exactly the same as estimating the MISO downlink channel. The uplink channels from the user to each antenna can be estimated independently at each antenna of the BS, since the M channels are orthogonal in space. This is contrast to the downlink, where at the single receiving antenna of the user we need to estimate M channels. Hence, the variance of the estimation error of $\hat{\mathbf{h}}_{\text{UL}}[n]$ and in turn of $\hat{\mathbf{h}}[n]$ in a TDD system is given by $\sigma_{e_{\text{TDD}}}^2 = \frac{1}{1+\frac{P_{\text{UL}}}{\sigma_u^2}T_{\text{UL}}}$, where $P_{\text{UL}}, \sigma_u^2$ and T_{UL} are the transmit power of the user in the uplink, the noise variance at the BS in the uplink and the uplink training length, respectively. Note, however, that only uplink resources are consumed for the estimation of the channel and in turn to obtain $\mathbf{w}_{\text{BS}}[n]$. Nevertheless when the $\mathbf{w}_{\text{BS}}[n]$ is employed in the downlink the channel could have already changed and hence the estimated CSI is also outdated.

2.2. Quantized CSI

The estimated channel is first available only at the receiver and in order to provide the transmitter with knowledge about the current channel in an FDD system, the estimated channel must be quantized and then relayed back to the BS with the limited feedback of B bits. For this, we employ the RVQ scheme, where we have at the transmitter (BS) and receiver (user) a codebook with 2^B random beamforming vectors $\mathbf{t}_j$, $j = 1, \ldots, 2^B$, i.i.d isotropically distributed over the M-dimensional unit sphere. For instance, the channel estimate

see the program

$\hat{\mathbf{h}}[n-1]$ is quantized by selecting the beamforming vector $\mathbf{w}[n]$ (to be used at time slot n) that best matches $\hat{\mathbf{h}}[n-1]$ [4]

$$\mathbf{w}[n] = \underset{\mathbf{t}_j}{\operatorname{argmax}} \ |\mathbf{t}_j^{\mathrm{H}}\hat{\mathbf{h}}[n-1]|^2. \tag{5}$$

Note that we have written the beamforming vector $\mathbf{w}$ without the subscript BS as $\mathbf{w}_{\mathrm{BS}}$ shown in (1). The reason behind this will explained in the next subsection. The user feeds back the B bits representing the index of the beamforming vector $\mathbf{w}[n]$ in the codebook. The loss due to quantization depends on the number of feedback bits B. As B increases, the quantization error becomes smaller, but this comes at the expense of the consumption of uplink resources accompanied with a higher feedback error probability as it will be seen next.

2.3. Erroneous CSI

The feedback channel in an FDD system is the uplink, which is not error-free and so, the relayed CSI can be received erroneously. If at least one feedback symbol is in error, we assume there is a *total* loss of the feedback as in [3], such that the index received by the BS corresponds to a different beamforming vector than the one intended by the user. We further assume no error detection scheme for the feedback and hence, after erroneous feedback the beamforming vector applied by the BS, i.e., $\mathbf{w}_{\mathrm{BS}}[n]$, would be different than the beamforming vector $\mathbf{w}[n]$, which was fedback by the user (c.f. (5)). In such a case, the beamforming vector applied by the BS would be completely *uncorrelated* with the actual downlink channel. Let us denote the probability of erroneous feedback, i.e., that $\mathbf{w}_{\mathrm{BS}}[n] \neq \mathbf{w}[n]$, as p_ϵ. We do not discuss how p_ϵ is computed and simply assume it is given. For instance, assuming that the feedback bits are sent uncoded with QPSK symbols similar to [3] with a symbol error probability p_{S} in the uplink, we have that the feedback error probability is given as $p_\epsilon = 1-(1-p_{\mathrm{S}})^{\frac{B}{2}}$. In this case, p_ϵ depends on the symbol error probability and on the number of feedback bits, i.e. on the quantization of the downlink channel estimate. A larger B would lead to a smaller quantization error, but also to a higher p_ϵ and hence, we have a tradeoff between the quantization loss and the feedback error probability. Furthermore, the probability p_{S} and of course p_ϵ depend on the uplink *signal to noise ratio* (SNR) per receive antenna and on the uplink training length [6].

2.4. Outdated CSI

Due to delay involved in the feedback process, the available CSI at the BS is outdated in a time-varying downlink channel. Between downlink channel estimates of successive time slots, a first order Markov model also holds similar to (2):

$$\hat{\mathbf{h}}[n] = \sqrt{\alpha'}\,\hat{\mathbf{h}}[n-1] + \sqrt{1-\alpha'}\,\hat{\mathbf{g}}[n-1], \tag{6}$$

where $\sqrt{\alpha'}$ is the correlation coefficient between channel estimates and the elements of $\hat{\mathbf{g}} \in \mathbb{C}^M$ are i.i.d. zero-mean complex Gaussian random variables with variance $1-\sigma_e^2$ and are uncorrelated with $\hat{\mathbf{h}}[n-1]$. To compute α', we need

$$\begin{aligned}
\mathrm{E}\Big[\hat{\mathbf{h}}[n]\hat{\mathbf{h}}^{\mathrm{H}}[n-1]\Big] &= \mathrm{E}\Big[\mathrm{E}\Big[\hat{\mathbf{h}}[n]\hat{\mathbf{h}}^{\mathrm{H}}[n-1] \mid \mathbf{h}[n],\mathbf{h}[n-1]\Big]\Big] \\
&= \mathrm{E}\left[\mathrm{E}\left[\hat{\mathbf{h}}[n]\,|\mathbf{h}[n]\right]\mathrm{E}\left[\hat{\mathbf{h}}^{\mathrm{H}}[n-1]\,|\mathbf{h}[n-1]\right]\right] \\
&= \mathrm{E}\left[(1-\sigma_e^2)\mathbf{h}[n](1-\sigma_e^2)\mathbf{h}^{\mathrm{H}}[n-1]\right] \\
&= \sqrt{\alpha}(1-\sigma_e^2)^2\mathbf{1}_M,
\end{aligned} \tag{7}$$

where $\mathbf{1}_M$ is an $M \times M$ identity matrix and the first step follows from the fact the estimates at different time slots conditioned on the actual channels are independent and the last step from (2). By definition, the correlation coefficient $\sqrt{\alpha'}$ is computed by dividing the covariance $\sqrt{\alpha}(1-\sigma_e^2)^2$ with the product of the standard deviation of the elements of $\hat{\mathbf{h}}[n]$, i.e. $\sqrt{1-\sigma_e^2}$, and the standard deviation of the elements of $\hat{\mathbf{h}}[n]$, i.e. $\sqrt{1-\sigma_e^2}$. Hence, we have that $\sqrt{\alpha'} = \sqrt{\alpha}(1-\sigma_e^2)$. Note that the correlation coefficient between the channel estimates depends not only on the correlation coefficient $\sqrt{\alpha}$ between the true channels but also on the estimation error. The reason for this can be explained with the following example. Even if the channels remain constant, i.e., $\alpha = 1$, such that $\mathbf{h}[n] = \mathbf{h}[n-1]$, this does not mean that $\hat{\mathbf{h}}[n] = \hat{\mathbf{h}}[n-1]$, since at each time slot there are different noise realizations!

3. CAPACITY BOUNDS WITH IMPERFECT CSI

The capacity with imperfect CSI is unknown and hence, we compute bounds on the MISO channel capacity with the discussed imperfect CSI. Let us first rewrite (1) using (3)

$$\begin{aligned}
\mathbf{y}[n] &= \sqrt{P}\,\mathbf{w}_{\mathrm{BS}}^{\mathrm{H}}[n]\left(\hat{\mathbf{h}}[n]+\mathbf{e}[n]\right)\mathbf{r}[n] + \mathbf{v}[n,k] \\
&= \sqrt{P}\,\mathbf{w}_{\mathrm{BS}}^{\mathrm{H}}[n]\hat{\mathbf{h}}[n]\mathbf{r}[n] + \mathbf{z}[n].
\end{aligned} \tag{8}$$

where $\mathbf{z}[n] = \sqrt{P}\,\mathbf{w}^{\mathrm{H}}[n]\mathbf{e}[n]\mathbf{r}[n] + \mathbf{v}[n]$ is the effective noise. The elements of $\mathbf{z}[n]$ are not necessarily Gaussian distributed and can be correlated with the signal $\mathbf{r}[n]$[1]. The variance of the elements of $\mathbf{z}[n]$ is

$$\sigma_z^2 = \sigma_e^2 P + \sigma_v^2. \tag{9}$$

Recall that the first T_{DL} symbols in each time slot are used to compute a downlink channel estimate at the user, i.e., $\hat{\mathbf{h}}$. Therefore, the user knows the actual channel estimate $\hat{\mathbf{h}}[n]$ for the data detection in the last $T - T_{\mathrm{DL}}$ symbols of the time slot. Additionally, at time slot n the user *expects* that the BS will be performing beamforming with $\mathbf{w}[n]$ (c.f. (5)), whose index was fedback from the user to the BS at time slot $n-1$. This will be the case if there was no feedback error and then $\mathbf{w}_{\mathrm{BS}}[n] = \mathbf{w}[n]$, such that the user's expectation about the beamforming vector is correct. However, if the feedback was received with errors then $\mathbf{w}_{\mathrm{BS}}[n] \neq \mathbf{w}[n]$ and the user's assumption about the beamforming vector is wrong! The user,

[1]If the $\mathbf{r}[n]$ are QPSK symbols, then $\mathbf{z}[n]$ is uncorrelated with $\mathbf{r}[n]$.

see the program

nevertheless, is not aware if a feedback error has occurred or not. But since the feedback error probability p_ϵ, for instance, depends only on the uplink SNR, which is a long term parameter, and on the uplink training length and B, which are fixed, we have that p_ϵ remains constant over many transmission blocks. Hence, even if the user is not aware of the feedback errors, it can be assumed that p_ϵ is known at the user.

3.1. Capacity Lower Bound with Imperfect CSI

In order to compute a lower bound of the capacity C with imperfect CSI we assume that transmitted symbols $\mathbf{r}[n]$ are Gaussian distributed, which is not necessary capacity achieving in this case. Let us denote the feedback error event by the binary random variable ϵ, where $\epsilon[n-1]=1$ in case of a feedback error at time slot $n-1$ and $\epsilon[n-1]=0$ otherwise. It is also known that replacing $\mathbf{z}[n]$ with a zero-mean Gaussian random variable with variance σ_z^2 which is independent of the signal $\mathbf{r}[n]$ minimizes the mutual information between the transmitted symbols $\mathbf{r}[n]$ and the received signal $\mathbf{y}[n]$ given in (8) [1]. A lower bound on the downlink capacity C_{lb} with imperfect CSI assuming the receiver knows $\hat{\mathbf{h}}[n]$ and $\mathbf{w}[n]$, (and not $\mathbf{w}_{\text{BS}}[n]$!) can be computed as follows

$$
\begin{aligned}
T\cdot C &\overset{\text{(a)}}{\geq} I(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n]) \\
&\overset{\text{(b)}}{=} I(\mathbf{r}[n];\mathbf{y}[n]|\hat{\mathbf{h}}[n],\mathbf{w}[n],\epsilon[n-1]) \\
&\quad +I(\epsilon[n-1];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n]) \\
&\quad -I(\epsilon[n-1];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\mathbf{r}[n]) \\
&\overset{\text{(c)}}{\geq} I(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\epsilon[n-1]) \\
&\quad -I(\epsilon[n-1];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\mathbf{r}[n]) \\
&\overset{\text{(d)}}{=} I(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\epsilon[n-1]) \\
&\quad -H(\epsilon[n-1]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\mathbf{r}[n]) \\
&\quad +H(\epsilon[n-1]\mid\mathbf{y}[n],\hat{\mathbf{h}}[n],\mathbf{w}[n],\mathbf{r}[n]) \\
&\overset{\text{(e)}}{\geq} I(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\epsilon[n-1]) \\
&\quad -H(\epsilon[n-1]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\mathbf{r}[n]) \\
&\overset{\text{(f)}}{\geq} I(\mathbf{r}[n];\mathbf{y}[n]\,|\,\hat{\mathbf{h}}[n],\mathbf{w}[n],\epsilon[n-1])-H(\epsilon[n-1])
\end{aligned}
$$

$$
C_{\text{lb}}=\frac{1}{T}I(\mathbf{r}[n];\mathbf{y}[n]\,|\,\hat{\mathbf{h}}[n],\mathbf{w}[n],\epsilon[n-1])-\frac{1}{T}h_{\text{b}}(p_\epsilon) \tag{10}
$$

where $I(p;q\mid m)$ denotes the mutual information between p and q given m and $H(p\mid q)$ denotes the conditional entropy of p given q. Inequality (a) follows by assuming that $\mathbf{r}[n]$ and $\mathbf{z}[n]$ are Gaussian distributed as explained above, step (b) from the chain rule of the mutual information, inequality (c) from the non-negativity of the mutual information, step (d) derives from the definition of the mutual information, inequality (e) from the non-negativity of the entropy and step (f) from the fact that conditioning reduces the entropy: $H(\epsilon[n-1]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\mathbf{r}[n])\leq H(\epsilon[n-1])$. Note that $H(\epsilon[n-1])=h_{\text{b}}(p_\epsilon)$, which is the binary entropy function with probability p_ϵ for the error event $\epsilon[n-1]=1$, i.e.,

$$
h_{\text{b}}(p_\epsilon)=-p_\epsilon\log_2 p_\epsilon-(1-p_\epsilon)\log_2(1-p_\epsilon). \tag{11}
$$

The first term in (10) represents the capacity when the user is aware of the feedback error events but does not know the beamforming vector $\mathbf{w}_{\text{BS}}[n]$ actually applied by the BS after a feedback error which can be written as

$$
\begin{aligned}
&I(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}[n],\epsilon[n-1])= \\
&\qquad(1-p_\epsilon)\;I\left(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}_{\text{BS}}[n]{=}\mathbf{w}[n]\right) \\
&\qquad p_\epsilon\;I\left(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}_{\text{BS}}[n]{\neq}\mathbf{w}[n]\right).
\end{aligned} \tag{12}
$$

$I\left(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}_{\text{BS}}[n]{=}\mathbf{w}[n]\right)$ corresponds to the mutual information assuming perfect feedback and that the effective estimated channel after beamforming, i.e., $\mathbf{w}_{\text{BS}}^{\text{H}}[n]\hat{\mathbf{h}}[n]$, is also perfectly known. The second term corresponds to the mutual information of a noncoherent SISO channel with coherence length D and *unknown* effective channel $h[n]=\mathbf{w}^{\text{H}}[n]\mathbf{h}[n]$. As explained in the previous section, after a feedback error, the applied beamforming vector $\mathbf{w}_{\text{BS}}[n]$ is totally uncorrelated with the channel and hence, it can be considered as if the beamforming vector were random. In this case the *unknown* channel $h[n]$ is a zero-mean unit-variance complex Gaussian random variable with $\mathrm{E}\left[|h[n]|^2\right]=1$. We can lower bound the second term using the following

$$
\begin{aligned}
I\left(\mathbf{r}[n];\mathbf{y}[n]\mid\hat{\mathbf{h}}[n],\mathbf{w}_{\text{BS}}[n]{\neq}\mathbf{w}[n]\right)&=I(\mathbf{r}[n];\mathbf{y}[n]) \\
&=I(\mathbf{r}[n];\mathbf{y}[n]\mid h[n])+I(h[n];\mathbf{y}[n]) \\
&\quad-I(h[n];\mathbf{y}[n]\mid\mathbf{r}[n]) \\
&\geq I(\mathbf{r}[n];\mathbf{y}[n]\mid h[n])-I(h[n];\mathbf{y}[n]\mid\mathbf{r}[n]),
\end{aligned} \tag{13}
$$

where the second step follows from the chain rule for the mutual information and the last step follows from the non-negativity of the mutual information. The first term in the expression above can be computed in closed form [7]

$$
\begin{aligned}
I(\mathbf{r}[n];\mathbf{y}[n]\mid h[n]) &= D\,\mathrm{E}\left[\log_2\left(1+\frac{P}{\sigma_v^2}|h[n]|^2\right)\right] \\
&= D\log_2(\mathrm{e})\,\mathrm{e}^{\sigma_v^2/P}\mathrm{E}_1(\sigma_v^2/P),
\end{aligned} \tag{14}
$$

whre $\mathrm{E}_k(z)$ is the generalized exponential integral defined as $\mathrm{E}_k(z)=\int_1^\infty\frac{\mathrm{e}^{-zt}}{t^k}\mathrm{d}t$. The second term in (13) is the mutual information between $h[n]$ and $\mathbf{y}[n]=h[n]\mathbf{x}[n]+\mathbf{v}[n]$ given $\mathbf{r}[n]$, which can be lower bounded as

$$
\begin{aligned}
I(h[n];\mathbf{y}[n]\mid\mathbf{r}[n]) &= \mathrm{E}\left[\log_2\left(\left|\mathbf{1}_D+\frac{P}{\sigma_v^2}\mathbf{r}[n]\mathbf{r}^{\text{H}}[n]\right|\right)\right] \\
&= \mathrm{E}\left[\log_2\left(1+\frac{P}{\sigma_v^2}\|\mathbf{r}[n]\|_2^2\right)\right] \\
&= \log_2(\mathrm{e})\,\mathrm{e}^{\sigma_v^2/P}\sum_{k=1}^{D}\mathrm{E}_k(\sigma_v^2/P),
\end{aligned} \tag{15}
$$

where the expectation is taken over $\mathbf{r}[n]$ and the last step follows similarly to the last step in (14).

3.2. Capacity Lower Bound with Perfect Feedback

Now we focus on the second term of (12), i.e. the mutual information under perfect feedback with known $\mathbf{w}_{\text{BS}}^{\text{H}}[n]\hat{\mathbf{h}}[n]$

$$I\left(\mathbf{r}[n];\mathbf{y}[n]|\hat{\mathbf{h}}[n],\mathbf{w}_{\text{BS}}[n]{=}\mathbf{w}[n]\right){=}I\left(\mathbf{r}[n];\mathbf{y}[n]|\mathbf{w}_{\text{BS}}[n]\hat{\mathbf{h}}[n]\right)$$
$$= \mathrm{E}[\log_2(1+\frac{P}{\sigma_v^2}|\mathbf{w}_{\text{BS}}[n]\hat{\mathbf{h}}[n]|^2)], \quad (16)$$

with a lower bound given by the following theorem.

Theorem 3.1. *A capacity lower bound with perfect feedback and estimated, quantized and outdated transmit CSI reads as*

$$\frac{D}{T}\left(1-\frac{1}{\sqrt{2\pi}}\frac{\sigma_\eta}{\mu_\eta}\right)\log_2\left(1+\frac{(\alpha'(M\mathrm{E}\left[\nu\right]-1)+1)(1-\sigma_e^2)}{\sigma_e^2+\frac{\sigma_v^2}{P}}\right),$$

Proof. Let us first define

$$\nu = \frac{|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n-1]|^2}{\|\hat{\mathbf{h}}[n-1]\|_2^2} \quad (17)$$
$$\eta = |\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n]|^2, \quad (18)$$

and present the following useful result from [8].

Theorem 3.2. *For a concave function $\Omega(\eta)$, we have that*

$$E\left[\Omega(\eta)\right] \geq \left(1-\frac{d_\eta}{2\mu_\eta}\right)\Omega(\mu_\eta),$$

where $\eta \in [0,\infty[$ and μ_η and $d_\eta = E\left[|\eta-\mu_\eta|\right]$ are the expected value and the absolute deviation of η, respectively.

From (16) note that $\Omega(\eta) = \log_2(1+\eta)$ and to employ Theorem 3.2, we first need the expected value of η which is

$$\begin{aligned}
\mu_\eta &= \mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n]|^2\right]\\
&= \alpha'\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n-1]|^2\right]+(1-\alpha')\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{g}}[n-1]|^2\right]\\
&= \alpha'\mathrm{E}\left[\|\hat{\mathbf{h}}[n-1]\|_2^2\frac{|\mathbf{w}^{\text{H}}\hat{\mathbf{h}}[n-1]|^2}{\|\hat{\mathbf{h}}[n-1]\|_2^2}\right]+(1-\alpha')(1-\sigma_e^2)\\
&= \alpha'\mathrm{E}\left[\|\hat{\mathbf{h}}[n-1]\|_2^2\right]\mathrm{E}\left[\frac{|\mathbf{w}^{\text{H}}\hat{\mathbf{h}}[n-1]|^2}{\|\hat{\mathbf{h}}[n-1]\|_2^2}\right]+(1-\alpha')(1-\sigma_e^2)\\
&= \left(\alpha' M\ \mathrm{E}\left[\nu\right]+1-\alpha'\right)(1-\sigma_e^2),
\end{aligned} \quad (19)$$

where the second step makes use of (6), the third steps follows since $\mathbf{w}$ has unit norm and the variance of the elements of $\hat{\mathbf{g}}[n-1]$ is $(1-\sigma_e^2)$, the fourth step derives from the fact that $\|\hat{\mathbf{h}}[n-1]\|_2^2$ and $\nu = |\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n-1]|^2/\|\hat{\mathbf{h}}[n-1]\|_2^2$ are independent [5]. Now we need $\mathrm{E}\left[\nu\right]$, which can be shown is given by [2]

$$\mathrm{E}\left[\nu\right] = 1-2^B\mathrm{Beta}\left(2^B, M/(M-1)\right), \quad (20)$$

where Beta$(\bullet,\bullet)$ is the Beta function and recall that B are the quantization bits. Let now characterize the variance of η

$$\sigma_\eta^2 = \mathrm{E}\left[(\eta-\mu_\eta)^2\right] = \mathrm{E}\left[\eta^2\right]-\mu_\eta^2, \quad (21)$$

for which we need the second moment of η, i.e.

$$\begin{aligned}
\mathrm{E}\left[\eta^2\right] &= \alpha'^2\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n-1]|^4\right]+(1-\alpha')^2\,\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{g}}[n-1]|^4\right]\\
&\quad+4\alpha'(1-\alpha')\,\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n-1]|^2\right]\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{g}}[n-1]|^2\right]\\
&= \alpha'^2\mathrm{E}\left[\|\hat{\mathbf{h}}[n-1]\|_2^4\right]\mathrm{E}\left[\nu^2\right]+(1-\alpha')^2\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n-1]|^4\right]\\
&\quad+4\alpha'(1-\alpha')\,\mathrm{E}\left[\|\hat{\mathbf{h}}[n-1]\|_2^2\right]\mathrm{E}\left[\nu\right]\mathrm{E}\left[|\mathbf{w}^{\text{H}}[n]\hat{\mathbf{h}}[n-1]|^2\right]\\
&= \alpha'^2(1-\sigma_e^2)^2M(M+1)\mathrm{E}\left[\nu^2\right]+2(1-\alpha')^2(1-\sigma_e^2)^2\\
&\quad+4\alpha'(1-\alpha')(1-\sigma_e^2)^2M\mathrm{E}\left[\nu\right],
\end{aligned} \quad (22)$$

where the second step follows by replacing $|\mathbf{w}^{\text{H}}[n]\ \hat{\mathbf{h}}[n-1]|^2$ for $\|\hat{\mathbf{h}}[n-1]\|_2^2\ \ \nu$ and then using the fact that $\|\hat{\mathbf{h}}[n-1]\|_2^2$ and ν are independent. The fourth step is derived from noting that $\mathrm{E}\left[\|\hat{\mathbf{h}}\|_2^4\right] = M(M+1)(1-\sigma_e^2)^2$, i.e. the second moment of the chi-square distributed random variable $\|\hat{\mathbf{h}}\|_2^2$ with $2M$ degrees of freedom, each degree of freedom with variance $(1-\sigma_e^2)/2$. To evaluate $\mathrm{E}\left[\eta^2\right]$ in the above expression, we still need $\mathrm{E}\left[\nu^2\right]$, which can be shown is given by

$$\mathrm{E}\left[\nu^2\right]{=}1-2^{B+1}\mathrm{Beta}\left(2^B,\frac{M}{M{-}1}\right){+}2^B\mathrm{Beta}\left(2^B,\frac{M{+}1}{M{-}1}\right). \quad (23)$$

Using (19)-(23) we can calculate the ratio $\frac{\sigma_\eta}{\mu_\eta} =$

$$\frac{\sqrt{\alpha'^2(M(M{+}1)\mathrm{E}[\nu^2]{-}M^2\mathrm{E}[\nu]^2)+(1{-}\alpha')^2{+}2\alpha'(1{-}\alpha')M\mathrm{E}[\nu]}}{\alpha'(M\mathrm{E}\left[\nu\right]{-}1){+}1}. \quad (24)$$

To compute a lower bound of (16) with Theorem 3.2, we need d_η, but exact evaluation of d_η appears intractable in general. However, we can find an upper bound for d_η based on σ_η. To this end let us compute μ_η, d_η and σ_η by setting $B\to\infty$ (no quantization, i.e. $\nu\to 1$) and $\alpha' = 1$ (no outdating)

$$\mu_\eta|_{\substack{B\to\infty\\ \alpha'=1}} = \sqrt{2}\sigma_e\,M$$
$$d_\eta\,|_{\substack{B\to\infty\\ \alpha'=1}} = \frac{2\sqrt{2}\sigma_e\,M^M}{(M-1)!\ \mathrm{e}^M}$$
$$\sigma_\eta^2\,|_{\substack{B\to\infty\\ \alpha'=1}} = 2\sigma_e^2\,M,$$

which we can use to obtain

$$\left.\frac{d_\eta}{\mu_\eta}\right|_{\substack{B\to\infty\\ \alpha'=1}} = \frac{2M^M}{M!\ \mathrm{e}^M} \leq \frac{2}{\sqrt{2\pi M}} \quad (25)$$
$$\left.\frac{2}{\sqrt{2\pi}}\frac{\sigma_\eta}{\mu_\eta}\right|_{\substack{B\to\infty\\ \alpha=1}} = \frac{2}{\sqrt{2\pi M}}, \quad (26)$$

where the inequality follows from Stirling's approximation. From (25) and (26) we can deduce that for the general case

$$\frac{d_\eta}{\mu_\eta} \leq \frac{2}{\sqrt{2\pi}}\frac{\sigma_\eta}{\mu_\eta}. \quad (27)$$

Substituting (24) in (27) and then plugging this with (19) in the result of Theorem 3.2 multiplied with the ratio of data symbols to total number of symbols concludes the proof. □

see the program

We are ready now to give C_{lb} from (10).

Theorem 3.3. *A lower bound on the capacity with imperfect transmit CSI which is estimated, quantized, outdated and affected by feedback errors is given by*

$$\begin{aligned}C_{lb} &= \frac{D}{T}\, p_\epsilon \log_2(\mathrm{e})\, \mathrm{e}^{\frac{\sigma_v^2}{P}} \mathrm{E}_1\left(\frac{\sigma_v^2}{P}\right)\\ &+ \frac{D}{T}(1-p_\epsilon)\left(1-\frac{1}{\sqrt{2\pi}}\frac{\sigma_\eta}{\mu_\eta}\right)\times\\ &\quad \log_2\left(1+\frac{(\alpha'(M\mathrm{E}[\nu]-1)+1)(1-\sigma_e^2)}{\sigma_e^2+\frac{\sigma_v^2}{P}}\right)\\ &- \frac{1}{T}\left(p_\epsilon \log_2(\mathrm{e})\, \mathrm{e}^{\frac{\sigma_v^2}{P}}\sum_{k=1}^{D}\mathrm{E}_k\left(\frac{\sigma_v^2}{P}\right)+H(\epsilon[n-1])\right).\end{aligned}$$

Proof. The result follows by plugging (11), (14), (15) and the result from Theorem 3.1 in (10) with $\frac{\sigma_\eta}{\mu_\eta}$ given by (24). □

3.3. Capacity Upper Bound with Imperfect CSI

We now draw our attention to computing an upper bound on the mutual information between $\mathbf{r}[n]$ and $\mathbf{y}[n]$ in given in (8).

Theorem 3.4. *An upper bound on the capacity under beamforming of the MISO channel with imperfect CSI in an FDD system is given by*

$$\begin{aligned}C_{ub} &= \frac{D}{T}(1-p_\epsilon)\log_2\left(1+\frac{P}{\sigma_v^2}(\alpha'(M\mathrm{E}[\nu]-1)(1-\sigma_e^2)+1)\right)\\ &+\frac{D}{T}\, p_\epsilon \log_2(\mathrm{e})\mathrm{e}^{\frac{\sigma_v^2}{P}}\mathrm{E}_1\left(\frac{\sigma_v^2}{P}\right).\end{aligned}$$

Proof. This upper bound can be derived by assuming the existence of a genie which provides extra information to the user namely, the exact channel vector $\mathbf{h}[n]$ and the actual beamforming vector applied by the BS $\mathbf{w}_{\text{BS}}[n]$, such that

$$\begin{aligned}C &\overset{(a)}{\leq} \frac{D}{T}\mathrm{E}\left[\log_2\left(1+\frac{P}{\sigma_v^2}\|\mathbf{w}_{\text{BS}}^{\mathrm{H}}[n]\mathbf{h}[n]\|_2^2\right)\right]\\ &\overset{(b)}{=} \frac{D}{T}(1-p_\epsilon)\,\mathrm{E}\left[\log_2\left(1+\frac{P}{\sigma_v^2}\|\mathbf{w}^{\mathrm{H}}[n]\mathbf{h}[n]\|_2^2\right)\right]\\ &\quad+\frac{D}{T}\, p_\epsilon\, \mathrm{E}\left[\log_2\left(1+\frac{P}{\sigma_v^2}|h[n]|^2\right)\right]\\ &\overset{(c)}{\leq} \frac{D}{T}(1-p_\epsilon)\log_2\left(1+\frac{P}{\sigma_v^2}\mathrm{E}\left[\|\mathbf{w}^{\mathrm{H}}[n]\hat{\mathbf{h}}[n]+\mathbf{w}^{\mathrm{H}}[n]\mathbf{e}[n]\|_2^2\right]\right)\\ &\quad+\frac{D}{T}\, p_\epsilon\, \log_2(\mathrm{e})\, \mathrm{e}^{\frac{\sigma_v^2}{P}}\,\mathrm{E}_1\left(\frac{\sigma_v^2}{P}\right)\\ &\overset{(d)}{=} \frac{D}{T}(1-p_\epsilon)\log_2\left(1+\frac{P}{\sigma_v^2}(\mu_\eta+\sigma_e^2)\right)\\ &\quad+\frac{D}{T}\, p_\epsilon\, \log_2(\mathrm{e})\, \mathrm{e}^{\frac{\sigma_v^2}{P}}\,\mathrm{E}_1(\sigma_v^2/P).\end{aligned}$$

where (a) follows from knowing $\mathbf{w}_{\text{BS}}^{\mathrm{H}}$ and $\mathbf{h}[n]$ at the user and (b) from the fact that $\mathbf{w}_{\text{BS}}[n] = \mathbf{w}[n]$ in case of no feedback errors and in case of a feedback error we have $h[n] = \mathbf{w}_{\text{BS}}^{\mathrm{H}}\mathbf{h}[n]$ which is Gaussian distributed. In Step (c) the first term results from Jensen's inequality and (3) and the second term from [7]. Step (d) arises from (18) and then the result of the theorem follows by plugging (19). □

3.4. Capacity Bounds in a TDD System

As discussed in Remark 2.1, in a TDD system the transmit CSI is estimated and outdated. The capacity bounds for this case can be derived based on the results of Theorems 3.3 and 3.4 by neglecting the effects of the quantization and of the feedback errors. First we present a lower bound on the downlink capacity C_{TDD} in a TDD system.

Theorem 3.5. *A lower bound on the downlink capacity of a TDD system with estimated and outdated CSI is*

$$C_{TDD,lb}=\left(1-\frac{1}{\sqrt{2\pi}}\frac{\sigma_{\eta,TDD}}{\mu_{\eta,TDD}}\right)\log_2\left(1+\frac{(\alpha'(M-1)+1)(1-\sigma_{e_{TDD}}^2)}{\sigma_{e_{TDD}}^2+\frac{\sigma_v^2}{P}}\right).$$

Proof. The proof follows by taking the result from Theorem 3.3 without quantization, with perfect feedback and without downlink resources employed for training, i.e. $T_{\text{DL}} = 0$. Additionally, as explained in Remark 2.1 the variance of the estimation error is different in an TDD system, i.e. $\sigma_{e_{\text{TDD}}}^2 = \frac{1}{1+\frac{P_{\text{UL}}}{\sigma_u^2}T_{\text{UL}}}$. Plugging $\mathrm{E}[\nu] = 1$ (no quantization), $p_\epsilon = 0$ (no feedback errors) and $D = T$ (no downlink training) in (24) and Theorem 3.3 gives us the result of the theorem. With $\mathrm{E}[\nu] = 1$ and $p_\epsilon = 0$ in (24) we have that

$$\frac{\sigma_{\eta,\text{TDD}}}{\mu_{\eta,\text{TDD}}} = \frac{\sqrt{\alpha'(2-\alpha')(M-1)+1}}{\alpha'(M-1)+1}. \tag{28}$$

□

An upper bound on the downlink capacity under beamforming in a TDD system can be derived by making use of Theorem 3.4.

Theorem 3.6. *An upper bound on the downlink capacity of a TDD system with estimated CSI is*

$$C_{TDD,ub} = \log_2\left(1+\frac{P}{\sigma_v^2}((M-1)(1-\sigma_{e_{TDD}}^2)+1)\right).$$

Proof. The proof makes use of Theorem 3.4, which corresponds to the capacity upper bound in an FDD system and hence the result needs to be modified for the TDD case. For the TDD case, we have no quantization and hence $\mathrm{E}[\nu] = 1$ (c.f. (17)). Additionally, we have no outdating and no feedback and therefore, $\alpha' = 1$ and $p_\epsilon = 0$. The final result follows by replacing $\mathrm{E}[\nu] = 1$, $\alpha' = 1$ and $p_\epsilon = 0$ in Theorem 3.4 and by setting $D = T$. □

see the program

3.5. Capacity with Perfect CSI

For comparison we present the capacity of the MISO downlink channel with perfect CSI at the transmitter and receiver. The beamforming vector applied by the BS in the downlink is $\mathbf{w}_{\text{BS}}[n] = \frac{\mathbf{h}[n]}{\|\mathbf{h}[n]\|_2}$ such that (1) can be rewritten as

$$\mathbf{y}[n] = \sqrt{P}\,\|\mathbf{h}[n]\|_2\mathbf{r}[n] + \mathbf{v}[n]. \qquad (29)$$

The capacity $C_{\text{Perf. CSI}}$ with perfect CSI can be computed in closed form and is given by [7]

$$C_{\text{Perf. CSI}} = \log_2(\mathrm{e})\ \mathrm{e}^{\frac{\sigma_v^2}{P}} \sum_{k=1}^{M} \mathrm{E}_k\left(\frac{\sigma_v^2}{P}\right). \qquad (30)$$

This capacity is similar to the result from Theorem 3.6 except that in (30) the factor $\frac{T-T_{\text{UL}}}{T} = 1$ and $\sigma^2_{e_{\text{TDD}}} = 0$ since there is no estimation loss as the user knows the channel perfectly.

3.6. Capacity without Transmit CSI but with Receive CSI

Also for comparison we present the capacity when the BS has no CSI but the receiver can estimate the downlink channel through training. Note that this is not the same capacity with non-coherent detection like in (13). Without CSI at the transmitter the use of multiple antennas are transparent to the user in a single-user MISO channel. If the BS either applies a random beam $\mathbf{w}_{\text{BS}}[n]$ or equivalently transmits with only one antenna to the user, the equivalent SISO channel $h[n]$ is a scalar which is Gaussian distributed. Assuming the receiver can estimate the scalar channel $h[n]$ with T_{DL} symbols, such that the channel estimate is $\hat{h}[n]$ we have that

$$\begin{aligned}\mathbf{y}[n] &= \sqrt{P}\,\hat{h}[n]\mathbf{r}[n] + \sqrt{P}\,e'[n]\mathbf{r}[n] + \mathbf{v}[n] \\ &= \sqrt{P}\,\hat{h}[n]\mathbf{r}[n] + \mathbf{z}'[n]. \end{aligned} \qquad (31)$$

where $e'[n]$ is the estimation error with variance given exactly like the variance of the estimation error in a TDD system (Remark 2.1), i.e. $\sigma^2_{e'} = \frac{1}{1+\frac{P}{\sigma_v^2}T_{\text{DL}}}$, and where the effective noise $\mathbf{z}'[n] = \sqrt{P}\ e'[n]\mathbf{r}[n] + \mathbf{v}[n]$ with variance $\sigma_v^2 + \sigma^2_{e'}$. The capacity without CSI C_{NoTxCSI} at the transmitter is unknown and so we compute lower and upper bounds on the capacity of (31). We can compute a lower bound by first assuming that $\mathbf{r}[n]$ and $\mathbf{z}'[n]$ are Gaussian distributed as has been explained before, such that

$$\begin{aligned} C_{\text{NoTxCSI}} \geq\ & \frac{D}{T}\ \mathrm{E}\left[\log_2\left(1 + \frac{P}{\sigma_v^2 + \sqrt{P}\sigma^2_{e'}}|h[n]|^2\right)\right] \\ \geq & \frac{D}{T}\left(1-\frac{1}{\mathrm{e}}\right)\log_2\left(1+\frac{1-\sigma^2_{e'}}{\frac{\sigma_v^2}{P}+\sigma^2_{e'}}\right) = C_{\text{NoTxCSI,lb}}, \end{aligned} \qquad (32)$$

where the second step follows by making use of Theorem 3.2 where $\eta = |h[n]|^2$ is an exponential distributed random variable with expected value given by $1-\sigma^2_{e'}$ and absolute deviation given by $\frac{2}{\mathrm{e}}\left(1-\sigma^2_{e'}\right)$. An upper bound can be computed by making use of Theorem 3.6 and setting $M = 1$

$$C_{\text{NoTxCSI,ub}} = \frac{D}{T}\log_2\left(1+\frac{P}{\sigma_v^2}\right).$$

4. NUMERICAL RESULTS AND DISCUSSION

In this section we present some results of the bounds computed in the previous section. Fig. 1 depicts the bounds as a function of $\frac{P}{\sigma_v^2}$ in dB for $M = 10$, $T = 500$, $T_{\text{DL}} = 100$, $B = 20$, $\alpha = 0.9$, $p_\epsilon = 0$, $\sigma_u^2 = 1$, $P_{\text{UL}} = P/M$ and $T_{\text{UL}} \in \{10, 100\}$ (the last 3 parameters have an impact on the TDD case). One can see the impact of the different types and degrees of the CSI on the capacity. The best case with imperfect CSI consists of having just estimated and outdated CSI at the transmitter, i.e. the TDD case. Nevertheless, in the TDD system, the performance of the downlink is directly related to the quality of the channel estimate in the uplink, i.e., T_{UL} and P_{UL}/σ_u, while the FDD downlink performance is influenced by the number of feedback bits and their reliability. For this reason, the FDD system could achieve higher downlink rates than the TDD system in some cases.

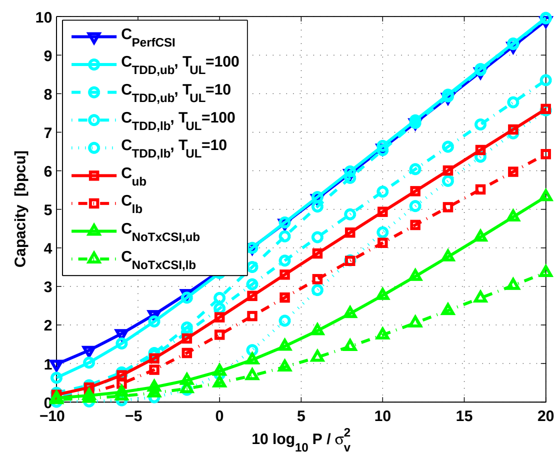


Fig. 1. Capacity Bounds vs $\frac{P}{\sigma_v^2}$ for the Different Scenarios.

The tradeoff between training and data payload can be observed in Fig. 2, where the bounds are shown as a function of T_{DL}. Only the bounds for the FDD case, i.e. C_{ub} and C_{lb}, and the lower bound for the case without CSI at the transmitter are functions of T_{DL}. The parameters are the same as before except that $\frac{P}{\sigma_v^2} = 6$ dB and $T_{\text{UL}} = 100$. Obviously, if $M \ll T$, then it is optimum to spent more than $T_{\text{DL}} = M$ symbols in order to reduce the estimation error.

Fig. 3 presents the FDD bounds as function of the feedback bits B. Now, we assume a feedback error probability given by $p_\epsilon = 1 - (1 - p_{\text{s}})^{\frac{B}{2}}$ with $p_{\text{s}} = 0.005$. This corresponds, for instance, to an uncoded QPSK (unfaded) transmission link with symbol error probability p_{s}. Again, the other parameters are the same as given in the beginning of the section except that $M = 4$, $\frac{P}{\sigma_v^2} = 6$ dB and $T_{\text{UL}} = 100$. One can observe an optimum number of feedback bits B. The reason for this is that, larger B decreases the quantization loss while leading to an increased feedback error probability p_ϵ.

see the program

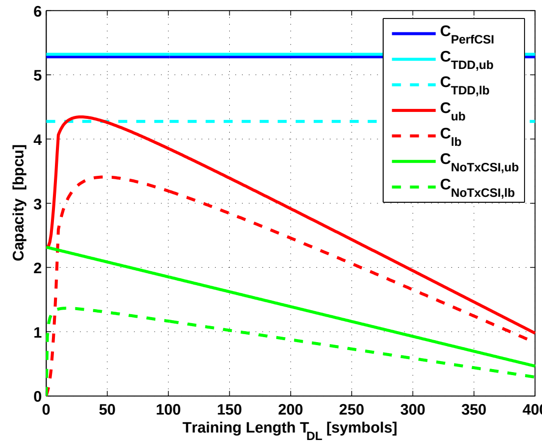


Fig. 2. Capacity Bounds vrs Training Length T_{DL}

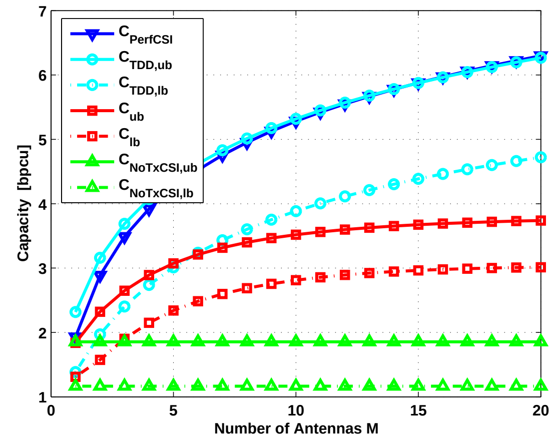


Fig. 4. Capacity Bounds vrs Number of Antennas M

The results are given for different α. Interestingly, the optimal values of B obtained, each for the lower and upper bounds of the capacity, are nearly the same, and thus, they probably maximize its exact value. Fig. 4 shows the bounds as a function of M for fixed B. The parameters are the same as in the previous figure, except that $B = 20$. As it can be observed from the upper as well as the lower bounds, as the number of antennas M approaches the number of feedback bits B or the training length T_{DL}, the increase in FDD capacity becomes marginal. Thus, in terms of FDD capacity, there is barely any benefit from having M beyond T_{DL} or B.

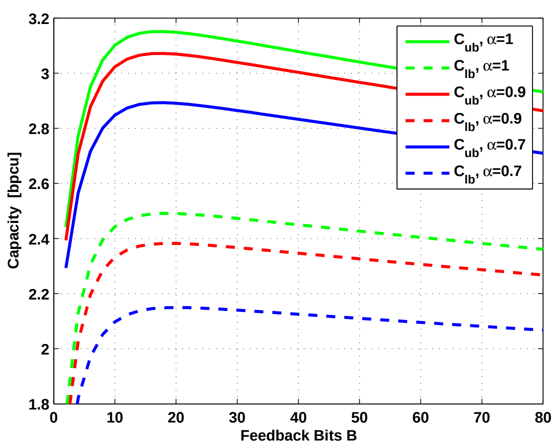


Fig. 3. FDD System Capacity Bounds vrs B

5. CONCLUSIONS

We developed a general MISO downlink framework that considers the inherent tradeoff between training, feedback and data in downlink communications. Based on it, we derived lower and upper bounds on the downlink sum capacity for various communication scenarios and settings, taking into account estimated, quantized and delayed CSI and erroneous feedback. Numerical results shows that, even under this conditions, the feedback of imperfect CSI is still beneficial. The capacity bounds may have interesting applications in the analysis and design of wireless systems, such as the optimization of resource allocation between data, training and feedback.

6. REFERENCES

[1] B. Hassibi and B. M. Hochwald, ”How much training is needed in a multiple-antenna wireless link?”, *IEEE Trans. Inform. Theory*, vol. 49, pp. 951-964, Apr. 2003.

[2] C. K. Au-Yeung and D. J. Love, ”On the performance of random vector quantization limited feedback beamforming in a MISO System”, *IEEE Trans. Wireless Commun.*, vol. 6, Feb. 2007.

[3] G. Caire, N. Jindal, M. Kobayashi and N. Ravindran, ”Multiuser MIMO Downlink Made Practical: Achievable Rates with Simple Channel State Estimation and Feedback Schemes,” *Arxiv preprint cs.IT/0710.2642*.

[4] W. Santipach and M. L. Honig, ”Capacity of Beamforming with Limited Training and Feedback”, *Proceedings IEEE ISIT*, pp. 376-380, Jul. 2006.

[5] J. C. Roh and B. D. Rao, ”Transmit beamforming in multiple-antenna systems with finite rate feedback: A VQ-based approach”, *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1101-1112, Mar. 2006.

[6] M. Castaneda, A. Mezghani, and J. A. Nossek, ”Optimal Resource Allocation in Downlink/Uplink of Single-User MISO/SIMO FDD Systems with Limited Feedback”, *submitted to the IEEE SPAWC* 2009.

[7] Mohamed-Slim Alouini, ”Adaptive and Diversity Techniques for Wireless Digital Communications over Fading Channels,” *Phd Thesis*, p. 128.

[8] A. Ben-Tal and E. Hochman, ”More bounds on the expectation of a convex function of a random variable”, *Journal of Applied Probability*, vol. 9, pp. 803 812, 1972.

see the program

IMPROVING THE PERFORMANCE OF BICM-OFDM SYSTEMS IN PRESENCE OF HPA NONLINEARITIES BY EFFICIENT BIT AND POWER LOADING

Mark Petermann, Carsten Bockelmann, Dirk Wübben and Karl-Dirk Kammeyer

Department of Communications Engineering
University of Bremen
Otto-Hahn-Allee, 28359 Bremen, Germany
{petermann, bockelmann, wuebben, kammeyer}@ant.uni-bremen.de

ABSTRACT

The usage of Orthogonal Frequency Division Multiplexing (OFDM) allows to adapt the available time-frequency grid to the current channel conditions. To exploit the offered degrees of freedom many bit and power loading algorithms were introduced. However, they do not include channel coding in the optimization. Similar to conventional multicarrier schemes in such link adaptation scenarios the peak-to-average power-ratio (PAPR) problem occurs. The combination of active constellation extension (ACE) with additional tone reservation (TR) is a promising approach to reduce this ratio without sacrificing the bit error rate (BER). Hence, in this contribution we propose an extended bisection method for coded bit and power loading in combination with ACE/TR techniques to jointly improve the BER and PAPR performance in high-rate OFDM systems. Results indicate that in presence of a memoryless nonlinear high power amplifier (HPA) device the BER can be lowered compared to existing loading algorithms.

1. INTRODUCTION

OFDM has become a popular modulation technique for broadband wireless communication systems [1]. The transmitter side adaptation of these systems to the current channel state is a crucial step towards higher spectral efficiencies and more robust systems. Although offering the use of adaptive algorithms OFDM also suffers from disadvantages like the high PAPR. Obviously, the whole system has to be considered in the optimization, where this paper deals with a combined approach of bit and power loading and PAPR reduction.

The bit and power allocation problem includes common parameters like modulation and power, but also the applied channel code needs to be considered to satisfy specific bit error rate (BER) or rate requirements. This, however, has usually been neglected in previous works, e.g., [2] and only recently some attention has been spent to the consideration of channel coding in bit and power loading algorithms with respect to the capacity of Bit Interleaved Coded Modulation (BICM) [3, 4] systems [5–7]. Especially BICM is suited for our purposes and allows for flexible allocation of code rate and modulation. Still, the information theoretical measure capacity does not describe the performance of a coded system completely and only holds for perfect capacity achieving codes. Therefore, we propose an efficient extension of the original loading approach of Krongold et al. [2], which uses the bisection method to solve the resulting convex optimization problem. Instead of analytical error rate expressions, in this paper the simulated AWGN performance of a set of code and modulation combinations, also known as transmission modes, will be used to form a look-up table describing the required signal-to-noise-ratio (SNR) for a specific subcarrier rate. The codes may be chosen from a common code family, e.g., convolutional codes with a fixed constraint length, realizing different code rates. Adapting the code rate additionally allows for higher flexibility in comparison to fixed code rate scenarios. A similar approach has been proposed by Stiglmayr et al. [8]. The authors solve the rate optimization problem formulated in terms of the BICM capacity by linearization, however they neglect a finer grained power control. Further investigations dealing with SNR mapping strategies based on clustered time-frequency grids (chunks) also neglect the power allocation [9, 10].

In OFDM systems potentially high peak values of the time-domain signal pose a big problem. As the transmit signal may be distorted by a high-power amplifier (HPA) nonlinear distortions and a spectral regrowth occur. This especially holds for adaptive scenarios as previous results showed a performance decrease of link adaptation algorithms if different modulation schemes are used for a certain output backoff (OBO) of the HPA [11]. Hence, several techniques were proposed to reduce the PAPR, a good overview can be found in [12]. The ideas of tone reservation (TR) to exploit unused subcarriers and tone injection (TI) to extend constellation points on used carriers motivate the combination of both for further PAPR reduction [13]. An efficient subclass of TI is the active constellation extension (ACE) method, which ex-

This work was supported in part by the German Research Foundation (DFG) under grants Ka841-18/1 and Ka841-21/1.

see the program

clusively alters outer constellation points of the used modulation alphabets on the subcarriers [14].

In this contribution we propose a combination of an extended bisection method for coded bit and power loading [15] with ACE/TR techniques [16] to jointly improve the BER and PAPR performance in high-rate OFDM systems.

The remainder of the paper is organized as follows. In Section 2 the OFDM system is described and the PAPR problem is stated. Section 3 deals with the performance of coded OFDM systems and the solution of the convex optimization problem using the proposed coded bisection method. Afterwards, Section 4 introduces a the new PAPR reduction approach without sacrificing the BER. Simulation results are shown in Section 5 and, finally, conclusions are given in Section 6.

2. SYSTEM MODEL AND PAPR PROBLEM

The considered OFDM system is assumed to be perfectly synchronized and free of intercarrier interference (ICI), i.e., the duration of the guard interval (GI) is sufficiently long. Furthermore, the channel state shall be perfectly known at both the transmitter and receiver side. The general system model including channel coding and interleaving is shown in Fig. 1.

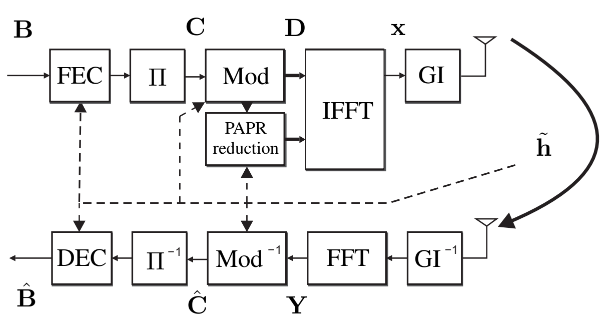


Fig. 1. Considered coded OFDM system model with additional PAPR reduction

The equivalent baseband system can be described in frequency domain as

$$Y_n = H_n \cdot \sqrt{P_n} \cdot D_n + N_n \,, \tag{1}$$

where H_n denotes the channel coefficient in frequency domain on subcarrier $n = 0, \cdots, N_C - 1$. P_n, D_n, N_n and Y_n denote the transmit power, transmit symbol, Gaussian noise and receive symbol, respectively. The total transmit power is given by $P_{\text{Total}} = \sum_{n=0}^{N_C-1} P_n$ and the power of the noise $N_n \sim \mathcal{N}_C(0, \sigma_N^2)$ is fixed to $\sigma_N^2 = 1$. The N_C frequency domain channel coefficients are determined by

$$H_n = \sum_{\ell=0}^{L_F-1} \tilde{h}(\ell) e^{-j2\pi \frac{n\ell}{N_C}} \,, \tag{2}$$

where the L_F taps of the time domain channel are defined as $\tilde{h}(\ell) \sim \mathcal{N}_C(0, 1/L_F)$.

Throughout this paper transmit symbols stemming from M-QAM modulation alphabets with Gray mapping are considered. To each subcarrier n an individual alphabet of cardinality $|M_n|$ will be assigned. Soft-Demapping via a-posteriori-probability (APP) detection is used to supply soft information to the decoder. In terms of forward error correction (FEC) non-systematic non-recursive convolutional codes of rates $R_C \in \{1/4, 1/3, 1/2, 2/3, 3/4\}$ and constraint length $L_C = 3$ are applied, leading to a variety of possible transmission modes, i.e. combinations of modulation and code. In all cases, the code word length is fixed to the number of bits in one OFDM symbol, leading to longer code words for higher data rates. Thus, no time diversity is exploited. A BCJR algorithm is used for soft-decoding and random interleaving is applied.

The time-domain transmit baseband signal for one OFDM symbol following (1) can be obtained using the IFFT

$$x_k = \frac{1}{\sqrt{N_C}} \sum_{n=0}^{N_C-1} X_n \exp\left(j2\pi \frac{kn}{N_C}\right) \,, \tag{3}$$

for time index $k = 0, \ldots, N_C - 1$, where $X_n = \sqrt{P_n} \cdot D_n$. Then, $\mathbf{x} = [x_0, \ldots, x_{N_c-1}]^T$ is the time-domain symbol vector at the IFFT output, whose elements, due to the central limit theorem, can be modeled as truncated zero-mean Gaussian random variables. This leads to a large probability of high peak values especially for an increasing number of subcarriers. Usually, an IFFT of a zero-padded input data vector of length wN_c is applied, where w is the oversampling factor. All peaks of the time-domain signal can be captured if an oversampling factor of $w \geq 4$ is used. Then, its corresponding discrete PAPR closely approximates that of the continuous-time signal. In this work all algorithms perform at Nyquist sampling rate, whereas the PAPR of the oversampled signal is measured in the investigations. Consequently, the PAPR for such an OFDM system is defined as [12]

$$\text{PAPR}\,(\mathbf{x}) = \frac{\|\mathbf{x}\|_\infty^2}{\mathrm{E}\left\{\|\mathbf{x}\|_2^2\right\}/wN_C} \,, \tag{4}$$

where $\mathrm{E}\{\cdot\}$ denotes the expectation and $\|\cdot\|_\alpha$ is the α-norm. In (4) $\|\mathbf{x}\|_\infty^2$ defines the peak value $\max_{\mathbf{x}} |x_k|^2$. This PAPR can be significantly reduced, if the transmit signal $\mathbf{x}$ is modified by an additive time-domain signal $\mathbf{c} \in \mathbb{C}^{N_c \times 1}$, which is optimized with respect to ACE and TR constraints such that the corresponding PAPR

$$\text{PAPR}\,(\mathbf{x}, \mathbf{c}) = \frac{\|\mathbf{x} + \mathbf{c}\|_\infty^2}{\mathrm{E}\left\{\|\mathbf{x} + \mathbf{c}\|_2^2\right\}/wN_c} \tag{5}$$

is minimized. The principles of ACE and TR are explained in Section 4.1. Following [14] the optimization problem using

see the program

the ACE and/or TR technique can be written as

$$\underset{\mathbf{c}}{\text{minimize}} \|\mathbf{x}+\mathbf{c}\|_\infty = \underset{\mathbf{C}}{\text{minimize}} \|\mathbf{x}+\mathbf{FC}\|_\infty \ , \qquad (6)$$

where $\mathbf{C}$ is the corresponding frequency-domain vector of the additive signal $\mathbf{c}$ and $\mathbf{F}$ denotes the IDFT matrix of size $N_c \times N_c$ with elements $f_{n,k} = \left(1/\sqrt{N_c}\right)\exp\left(j2\pi kn/N_c\right)$. This convex problem can be interpreted as minimizing the maximum squared magnitude of the resulting signal with respect to $\mathbf{c}$. Unfortunately, this is a minimization task in the complex plane and results in a quadratically constraint quadratic program (QCQP) [17]. Thus, a less complex solution is required.

As any square M-QAM can be represented by two $\sqrt{M}$-ASK without loss, we will restrict to ASK constellations in the following. This means the optimization of the bit and power loading and the PAPR optimization are carried out with an equivalent real-valued representation. For bit and power loading an expansion to QAM constellations is obtained by simply halving the power constraint and the maximum rate, while doubling the resulting powers and rates. Taking into account the ASK symbol representation, assume the real-valued system variables

$$\mathbf{X}_\text{R} = \left[\Re\{\mathbf{X}\}\,\Im\{\mathbf{X}\}\right]^T \in \mathbb{R}^{2N_c\times 1} \qquad (7a)$$

$$\mathbf{x}_\text{R} = \left[\Re\{\mathbf{x}\}\,\Im\{\mathbf{x}\}\right]^T \in \mathbb{R}^{2N_c\times 1} \qquad (7b)$$

$$\mathbf{C}_\text{R} = \left[\Re\{\mathbf{C}\}\,\Im\{\mathbf{C}\}\right]^T \in \mathbb{R}^{2N_c\times 1} \qquad (7c)$$

$$\mathbf{F}_\text{R} = \begin{bmatrix}\Re\{\mathbf{F}\} & -\Im\{\mathbf{F}\}\\ \Im\{\mathbf{F}\} & \Re\{\mathbf{F}\}\end{bmatrix} \in \mathbb{R}^{2N_c\times 2N_c}\ , \qquad (7d)$$

where $\mathbf{X}_\text{R}$ is a real-valued ASK symbol vector, $\mathbf{x}_\text{R}$ the equivalent time-domain represenation, $\mathbf{C}_\text{R}$ the frequency-domain correction term and $\mathbf{F}_\text{R}$ the real-valued representation of the IDFT matrix $\mathbf{F}$. Then, the PAPR optimization problem is given by

$$\underset{\mathbf{C}_\text{R}}{\text{minimize}} \|\mathbf{x}_\text{R}+\mathbf{F}_\text{R}\mathbf{C}_\text{R}\|_\infty \ . \qquad (8)$$

In the next section the coded bit and power loading is explained, whereas the PAPR reduction is addressed in Section 4.

3. CODED BIT AND POWER LOADING

3.1. Coded System Performance

The bit error rate performance of an uncoded $\sqrt{M}$-ASK transmission is well-known. It can be described as a function of the SNR γ by

$$\mathrm{P}_{b,\sqrt{M}-\text{ASK}} = \frac{2}{\log_2(M)}\left(1-\frac{1}{\sqrt{M}}\right)\text{erfc}\left(\sqrt{\frac{3}{M-1}\gamma}\right). \qquad (9)$$

Accordingly, the frame error rate (FER) given a certain frame length L_N - the number of channel uses - is defined as

$$\mathrm{P}_{f,\sqrt{M}-\text{ASK}} = 1-\left(1-\mathrm{P}_{b,\sqrt{M}-\text{ASK}}\right)^{L_N\log_2(\sqrt{M})} \qquad (10)$$

and can be used to derive the SNR γ required to achieve a given frame error rate for a certain modulation. This is the basis for many known bit loading algorithms, e.g., [2]. However, these results are limited to uncoded systems. To capture the behavior of the whole system - including an appropriate channel code - (9) and (10) are not sufficient. One way to obtain a quality indicator is to simulate the performance of the coded system.

In order to characterize the SNR, which is necessary to achieve a given bit or frame error rate performance on a single subcarrier, a system with equivalent block length of $L_N = N_C$ and AWGN noise is simulated. The motivation for this approach is, that given the SNR of a single subcarrier γ_n, we assume all symbols of the code word to have the SNR γ_n. The error rate performance of a specific code and modulation disturbed by AWGN at γ_n indicates, which transmission mode may be chosen to guarantee an error rate constraint. Using this heuristic, the overall error rate of an OFDM symbol with different SNRs and properly chosen subcarrier modes can be assumed to fulfill the error rate constraint.

Still, the AWGN assumption is optimistic in the sense that individual channel states have been compensated for properly. Nevertheless, it offers a good indication of the SNR which is necessary to support a target error rate on a single subcarrier at a specific data rate. Even though the applied channel code could cope with SNR variations over subcarriers, an ergodic Rayleigh fading channel would lead to far too pessimistic performance measures because of existing subcarriers with very low SNR. Such subcarriers will be compensated for in a perfect adaptive system by the assignment of more power, different modulation and stronger coding.

Fig. 2 shows the BER results of Monte-Carlo simulations for $\sqrt{M}$-ASK constellations up to $\sqrt{M} = 2^4$ and a variety of code rates versus E_b/N_0. It is quite clear, that only a subset of combinations will actually be used due to the fact that at a given rate one of the code-modulation combinations will lead to the best performance, e.g., 4-ASK with a half rate code compared to 16-ASK with a quarter rate code achieves the same spectral efficiency of 1 bit/s/Hz at a much lower E_b/N_0. Based on these simulation results, the system performance can be characterized as will be discussed in the next section.

3.2. Coded Bisection Approach

Consider the well-known optimization problem to enhance the error rate performance

$$\text{minimize} \qquad P_\text{Total} = \sum_{n=0}^{N_C-1} P_n \qquad (11a)$$

$$\text{subject to} \qquad \sum_{n=0}^{N_C-1} r_n = R_\text{Total} \qquad (11b)$$

$$\mathrm{P}_b < \mathrm{P}_\text{Target}\ . \qquad (11c)$$

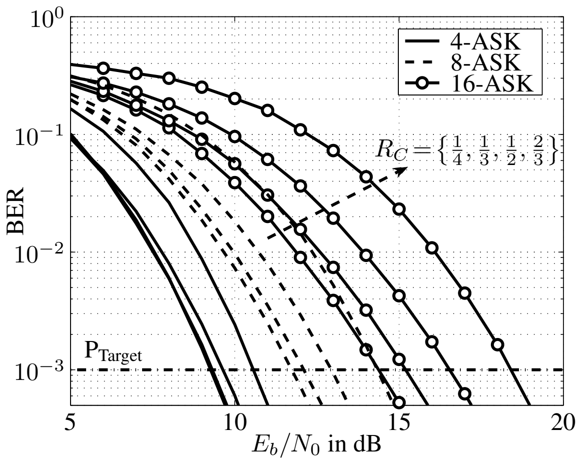


Fig. 2. BER vs. E_b/N_0 comparison of all combinations of $\sqrt{M}$-ASK with $\log_2(\sqrt{M}) = 2, 3, 4$ and convolutional codes $R_C \in \{1/4, 1/3, 1/2, 2/3\}$ (left to right) with $L_C = 3$ and frame length $L_N = 256$, $\mathrm{P}_{\mathrm{Target}} = 10^{-3}$

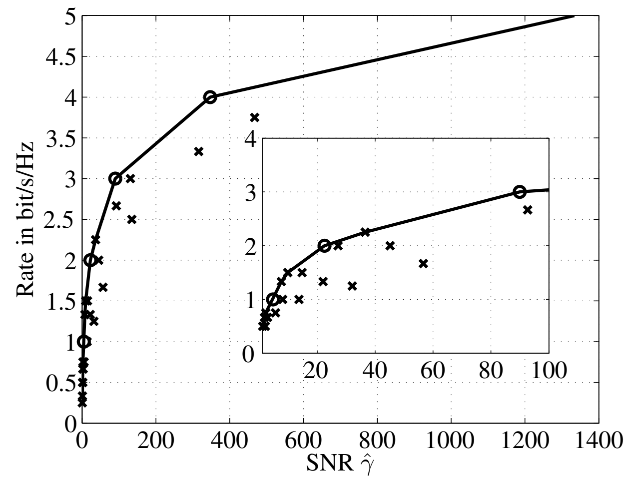


Fig. 3. Set of all rate-power points (x - coded modes; o - uncoded modes) and its convex hull (solid line) at $\mathrm{P}_{\mathrm{Target}} = 10^{-3}$, $L_N = 256$ and $L_C = 3$.

With this approach the transmit power P_{Total} is minimized given an overall target rate R_{Total} and error rate $\mathrm{P}_{\mathrm{Target}}$. A local code rate $R_{C,n}$ in conjunction with the applied modulation $\sqrt{M_n}$ define the bit rate $r_n = \log_2(\sqrt{M_n})R_{C,n}$ on subcarrier n. Note that the optimum solution at one target BER/FER scaled to the available transmit power can be used to show performance gains in terms of the error rate.

The error rate constraint $\mathrm{P}_{\mathrm{Target}}$ determines the power that is necessary to achieve a certain rate requirement. The results in Fig. 2 at the target BER lead to rate-power pairs defining potential subcarriers modes. Accordingly, Fig. 3 shows all rate-power points up to a maximum rate $R_{\mathrm{max}} = 5$ bit/s/Hz at a target BER of $\mathrm{P}_{\mathrm{Target}} = 10^{-3}$ found by Monte-Carlo simulations, stating the AWGN performance of an equivalent system with frame length N_C and the analytical bit error rate expression for uncoded ASK constellations in (9). In a perfect adaptive system subcarrier channel variations would be exploited or compensated by assignment of power, modulation alphabet and code rate, making the AWGN performance a good quality indicator to identify the SNR requirement of each mode.

The set of all rate-power points is therefore defined as $\mathcal{S} = \{(R_{C,i}, \hat{\gamma}_i) | f_{M_i, R_{C,i}}(p_i) = \mathrm{P}_{\mathrm{Target}}\}$, where $f_{M_i,R_{C,i}}$ denotes the utilized error rate function, e.g., the simulated BER in Fig. 2, parametrized by the transmission parameters and the power. The SNR $\hat{\gamma}_i$ for the real valued system is defined as

$$\hat{\gamma}_i = R_{C,i} \log_2(\sqrt{M_i}) \frac{E_{b,i}}{N_0/2} . \tag{12}$$

where i is the index over the elements of set $\mathcal{S}$.

To solve (11a) efficiently, convexity has to be ensured. To this end, a convex set of rate-power points $\mathcal{C} \subset \mathcal{S}$ has to be found, which can be constructed by the convex hull of the whole set as shown by the solid line in Fig. 3, where [18] has been used. A look-up table has to be generated once for a set of system parameters, i.e., N_C, code ensemble and maximum allowed rate. Based on such an easily storable look-up table the rate of each OFDM symbol can be optimized. More specifically, the look-up table has to return the mode with the greatest rate still being feasible. Feasibility in the bisection approach is connected to the slope of the rate-power curve, given by $\delta\mathbf{R}/\delta\hat{\mathbf{\Gamma}}$ with $\mathbf{R}$ being the vector of all rates on the convex hull and $\hat{\mathbf{\Gamma}}$ the vector of the respective SNRs $\hat{\gamma}$.

One very efficient way to find the optimal solution given the Lagrangian formulation of the convex optimization problem (11a)

$$J(\lambda) = \sum_{n=1}^{N_C} P_n + \lambda \left(\sum_{n=1}^{N_C} r_n - R_{\mathrm{Total}} \right) \tag{13}$$

is the bisection approach, which has been applied to the uncoded bit and power loading problem by Krongold et. al [2]. Considering the unconstrained problem, i.e., neglecting R_{Total}, each λ corresponds to an optimal power and rate distribution, which minimizes the cost function $J(\lambda)$. Accordingly, if $\delta J(\lambda)/\delta r_n = 0, \ \forall n$, the condition

$$\frac{\delta P_n}{\delta r_n} = -\lambda \quad \forall n , \tag{14}$$

has to be fulfilled, meaning that the optimal rates and powers have to be chosen through that point on the rate-power curve with slope λ.

The (scaled) look-up table, which provides the rate and power at a specific slope $\eta = \lambda/|H_n|^2$, has to be constructed

see the program

from $\mathcal{C}$ to calculate the optimal rates and powers on all subcarriers for a given λ. The optimal λ^*, which minimizes the power taking the target rate R_{Total} into account, can then be found iteratively. A pseudo-code description of the algorithm and more details are given in [15].

This bisection approach efficiently solves the convex optimization problem of the coded system by the previously explained steps, which is carried out once (1-step). However, the optimization results in local subcarrier code rates, whose applicability, e.g., by variable puncturing techniques is beyond the scope of this paper. Instead, a mean code rate $\bar{R}_C$ is calculated and chosen such that the code rate $\bar{R}_C$ fulfills

$$\bar{R}_C < \frac{1}{N_C} \sum_{n \in \mathcal{T}}^{N_C} R_{C,n} \,. \tag{15}$$

This code is used as the outer code for one OFDM symbol, where $\mathcal{T}$ denotes the set of all nonzero local code rates $R_{C,n}$. Due to this solution, though, the target rate cannot be guaranteed. As the subcarrier rates are changed by the application of a fixed global code rate, an overall rate loss is introduced, which means that stronger error protection than required is applied. This in turn violates the target rate constraint. A solution to this problem is a 2-step process, fixing the code rate in the first step followed by the optimization over a convex set of all modes applying this code rate in the second step. This procedure is used in our simulations as well.

The proposed bisection method used here is a very efficient and fast algorithm as simple calculations are necessary to obtain a feasible solution.

4. COMBINING THE CODED BISECTION METHOD WITH PAPR REDUCTION

4.1. Active Constellation Extension and Tone Reservation

The bit and power loading approach from Section 3.2 may decrease the error rate, whereas the possibly large PAPR value after the IFFT operation remains. Therefore, the principles of active constellation extension and tone reservation are combined in one common optimization approach to reduce the PAPR. The idea of ACE is to extend the outer constellation points of the utilized modulation alphabet. For higher order modulation the inner constellation points must remain at their positions, otherwise the Euclidean distance between these symbols is decreased. Fig. 4 depicts an ACE example of 16-QAM, where the points at the edges only have one degree of freedom (real *or* imaginary part) and the corner points have two (both real and imaginary part). The inner points are identical to the original constellation. Basically, such an extended constellation does not increase the bit error rate due to the constant minimum Euclidean distance. Instead, an increased Euclidean distance at the outer constellation points may decrease the error rate.

Furthermore, if bit and power loading procedures are applied free subcarriers may occur, i.e., no modulation is applied on these carriers. Now, tone reservation (TR) can be used to exploit these unused subcarriers. They can also be found in the origin of the left I/Q diagram in Fig. 4. Using TR methods means to find a complex additive signal on these unused subcarriers, where compared to ACE subcarriers no restrictions to the optimum additive components are made. According to (8), the QCQP problem in (6) is now stated as a real-valued problem with $2N_C$ optimization variables by separating real and imaginary part. However, as real and imaginary part are dependent optimization variables, this separation leads to a suboptimal solution with a remaining small error. The right part of Fig. 4 shows how the approach approximates the optimal solution circle of the power restricted complex envelope, which depicts the maximum achievable peak power reduction resulting from the optimum complex optimization problem. The real-valued approach packs all time-domain samples in an outer square, where the maximum possible error is indicated by the two dots. The variable t will be explained later. Furthermore, the circle can be closer approximated by using

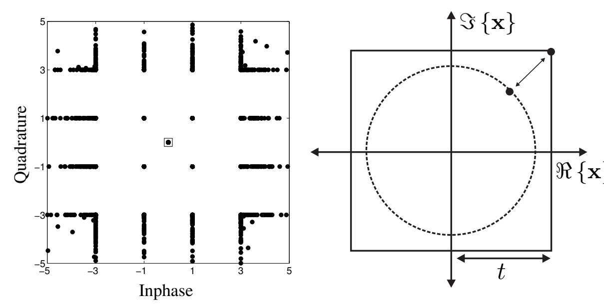


Fig. 4. (left) Exemplary active constellation extension for 16-QAM with zero symbols exploitable by TR techniques; (right) Approximation of the power restricted complex envelope

additional phase-shifted versions of the time-domain signal block. This in turn escalates the amount of constraint equations especially in a link adaption scenario and slows down convergence speed of the resulting algorithm [17]. That is quite an important aspect as the link adaption algorithms already introduce processing delays at the transmitter.

4.2. Joint ACE/TR Optimization for Bit Loading

Motivated by the fact that the QAM symbols are separated into two ASK symbols in our system model, the magnitudes of the real and imaginary parts are now optimized independently. Therefore, (8) can be written as a linear programming (LP) problem [19]. The combination of ACE and TR introduces the following constraints to this optimization problem. For all non-data bearing subcarriers that can be exploited by tone reservation, there are no restrictions on the feasible region of the symbols after optimization. For all other active carriers only those symbols corresponding to a corner or edge point of the selected complex symbol alphabet are considered. In summary, this implies that only outer constellation symbols of the ASK signal $\mathbf{X}_\mathrm{R}$ are incorporated in the following optimization process. The determined set of ASK symbol indices prior to the optimization is indicated by $\mathcal{I}_c = \left[i_1, \ldots, i_{|\mathcal{I}_c|}\right]$. Only those elements contained in this set are used for optimization purposes. It is worth mentioning that the zero amplitude symbols of the TR subcarriers are also included in $\mathcal{I}_c$. Hence, by optimizing with respect to the constraints on $\mathcal{I}_c$ the additive signal $\mathbf{C}$ including ACE and TR subcarriers is optimized. If the columns of the disregarded ASK symbols are excluded from the IDFT matrix such that $\tilde{\mathbf{F}}_\mathrm{R} = \left[\mathbf{f}_{i_1}, \mathbf{f}_{i_2}, \ldots, \mathbf{f}_{i_{|\mathcal{I}_c|}}\right] \in \mathbb{R}^{2N_c \times |\mathcal{I}_c|}$ and similarly arrange all frequency-domain correction symbols of the considered symbols in vector $\tilde{\mathbf{C}}_\mathrm{R} = \left[C_{i_1}, C_{i_2}, \ldots, C_{i_{|\mathcal{I}_c|}}\right]^T \in \mathbb{R}^{|\mathcal{I}_c| \times 1}$, it is easy to see that the relation $\mathbf{F}_\mathrm{R}\mathbf{C}_\mathrm{R} = \tilde{\mathbf{F}}_\mathrm{R}\tilde{\mathbf{C}}_\mathrm{R}$ holds. Then, if the objective function in (8) is upperbounded by a certain value t, the optimization problem can be reformulated by

$$\text{minimize} \quad t \tag{16a}$$

$$\text{subject to} \quad \left| x_{R,k} + \mathbf{F}_R^{(k)} \tilde{\mathbf{C}}_\mathrm{R} \right| \le t \quad \forall\, k \in \mathcal{I}_c\ , \tag{16b}$$

where $\mathbf{F}_R^{(k)}$ denotes the k-th row of matrix $\mathbf{F}_\mathrm{R}$. The element-wise inequality constraints for the absolute values in (16b) can be written in matrix form

$$\begin{bmatrix} \tilde{\mathbf{F}}_\mathrm{R} & \mathbf{I}_{2N_c \times 1} \\ -\tilde{\mathbf{F}}_\mathrm{R} & \mathbf{I}_{2N_c \times 1} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{C}}_\mathrm{R} \\ t \end{bmatrix} \ge \begin{bmatrix} -\mathbf{x}_\mathrm{R} \\ \mathbf{x}_\mathrm{R} \end{bmatrix}\ , \tag{17}$$

whereas the set limitation in (16b) can be stated with respect to the feasible regions of the transmit symbols

$$S_k C_{R,k} \ge 0 \quad \forall \quad k \in \mathcal{I}_c \tag{18a}$$

$$C_{R,k} \stackrel{!}{=} 0 \quad \forall \quad k \notin \mathcal{I}_c\ , \tag{18b}$$

where $S_k = \operatorname{sgn}\{X_{R,k}\} \in \{-1, 0, +1\}$ is the sign of the kth ASK symbol. The constraints of inner constellation points and the unconsidered real and imaginary parts of edge points of the QAM symbols are given in (18b). These constraints can be easily fulfilled if they are excluded from the optimization problem. The inequality constraints in (18a) can again be written in matrix form with element-wise inequality

$$\mathbf{S}\tilde{\mathbf{C}}_\mathrm{R} \ge \mathbf{0}_{|\mathcal{I}_c| \times 1}\ . \tag{19}$$

In (19) the sign variables are arranged in a diagonal matrix $\mathbf{S} = \operatorname{diag}\left\{S_1, \ldots, S_{|\mathcal{I}_c|}\right\}$. At last, by defining the vectors $\mathbf{d} = \left[\mathbf{0}_{1 \times |\mathcal{I}_c|}\ 1\right]^T \in \mathbb{R}^{|\mathcal{I}_c|+1 \times 1}$ and $\mathbf{y}_\mathrm{R} = \left[\tilde{\mathbf{C}}_\mathrm{R}\ t\right]^T \in \mathbb{R}^{|\mathcal{I}_c|+1 \times 1}$ the optimization problem in (16) becomes

$$\text{minimize} \quad \mathbf{d}^T \mathbf{y}_\mathrm{R} \tag{20a}$$

$$\text{subject to} \quad \mathbf{A}_\mathrm{R}\mathbf{y}_\mathrm{R} \ge \mathbf{b}_\mathrm{R}\ , \tag{20b}$$

with matrix $\mathbf{A}_\mathrm{R}$ and vector $\mathbf{b}_\mathrm{R}$ given by

$$\mathbf{A}_\mathrm{R} = \begin{bmatrix} \tilde{\mathbf{F}}_\mathrm{R} & \mathbf{I}_{2N_c \times 1} \\ -\tilde{\mathbf{F}}_\mathrm{R} & \mathbf{I}_{2N_c \times 1} \\ \mathbf{S} & \mathbf{0}_{|\mathcal{I}_c| \times 1} \end{bmatrix} \quad \text{and} \quad \mathbf{b}_\mathrm{R} = \begin{bmatrix} -\mathbf{x}_\mathrm{R} \\ \mathbf{x}_\mathrm{R} \\ \mathbf{0}_{|\mathcal{I}_c| \times 1} \end{bmatrix}\ . \tag{21}$$

This is a LP problem [19] with $2N_c + 1$ variables and $4N_c + |I_c|$ constraints that can be solved with standard tools, e.g., an efficient iterative Newton method [20], which works considerably well with a large number of constraints. The minimum solution value of t describes the size of the approximate square as shown in Fig. 4 [16].

One possible alternative method for approximating the previously described LP with less complexity is to use a gradient-project approach like in [14] for the ACE component and extend it with an update rule for the additive correction term related to the TR subcarriers [21]. This combines the properties of both iterative schemes in a single procedure, which is called enhanced gradient-project algorithm (EGPA) in the following. The corresponding update rule for the time-domain transmit vector at iteration i is [16]

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mu\mathbf{c}_\mathrm{ACE} - \nu\mathbf{c}_\mathrm{TR}\ , \tag{22}$$

where the scalars μ and ν are smart-gradient step sizes chosen as suggested for the smart gradient method in [14] and the algorithm in [21]. For more details refer to [16]. It is worth mentioning that this algorithm introduces additional IFFT and FFT operations, which in turn slightly increases the complexity compared to the original OFDM system. For high spectral efficiencies all subcarriers are data bearing subcarriers especially if a maximum modulation alphabet size is set. Then this algorithm inherently reduces to the one in [14].

5. SIMULATION RESULTS

For the investigations the number of subcarriers was set to $N_C = 256$, the channel has $L_F = 6$ taps, the spectral efficiency is 3 bit/s/Hz with a non-adaptive average code rate of 0.5, whereas the maximum modulation size is $\sqrt{M_{n,\max}} =$ 32 ASK (1024-QAM)[1]. The error rate constraint was set to a BER of $\mathrm{P}_\mathrm{Target} = 10^{-3}$. For the consideration of the memoryless nonlinear HPA, the well-known Rapp model with smoothing parameter $p = 2$ and an input power backoff (IBO) of 3 dB is used [22]. The EGPA method uses a clipping ratio of 3 dB and the maximum iteration number was set to 3. In Fig. 5 it is shown that for a high spectral efficiency the proposed coded bit and power loading outperforms the original biscetion method of Krongold, where a gain of approximately 1.5 dB is obtained at a BER of $2 \cdot 10^{-3}$. By applying PAPR reduction with ACE/TR techniques using a direct solution of the LP, the BER performance gain can be further improved up

[1]Higher order modulation is generally more often used in DSL as in wireless channels. Nevertheless, although occurring rarely (cf. Fig. 6), they allow for higher flexibility of the algorithm.

see the program

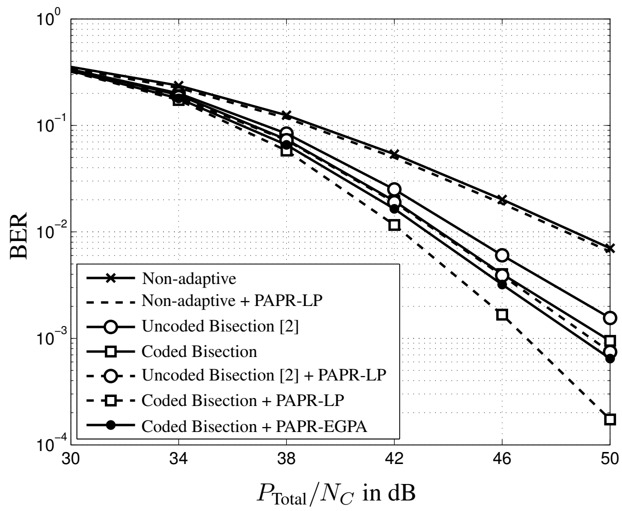


Fig. 5. BER vs. subcarrier SNR for 3 bit/s/Hz for $N_c = 256$ using the Rapp HPA model with $p = 2$ and an IBO of 3 dB

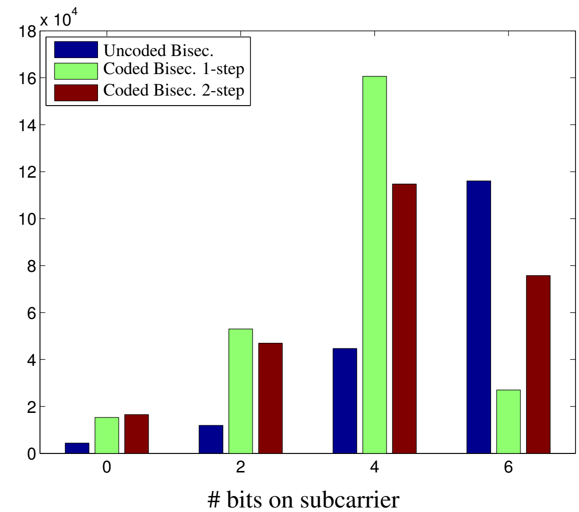


Fig. 6. Histogram of different bit allocation algorithms for 1000 channel realizations with $N_C = 256$ subcarriers, a maximum modulation of $M_{n,\max} = 64$ (QAM), $L_F = 6$ and $\mathrm{P}_{\text{Target}} = 10^{-3}$

to approximately 4 dB at the same BER. This improvement follows from the increased Euclidean distance of outer constellation points and is larger for the coded bisection method. This is due to the fact that the coded bisection method deactivates subcarriers more often by trend. In addition, lower order modulation is chosen more frequently compared to the uncoded bisection method. Instead, larger code rates are used to compensate for the rate loss. This behavior is depicted in a histogram in Fig. 6 recorded for 1000 channel realizations with a maximum modulation size of $M_{n,\max}=64$ QAM. The difference of the coded bisection methods (1-step or 2-step) was explained in Section 3.2, whereas the 2-step algorithm is applied in the BER curves in Fig. 5. Hence, as lower order modulation is used more often the ACE reduction algorithm has more degrees of freedom to decrease the PAPR and inherently increases the Euclidean distance of more constellation points. Furthermore, the proposed coded method needs less power to achieve the same target BER. Thus, SNR gains can be obtained.

The performance of the EGPA is degraded due to the imperfect constellation extension but still achieves better performance compared to the uncoded bisection method with PAPR reduction. For a spectral efficiency of 3 bit/s/Hz the OFDM scheme without loading suffers from the commonly chosen 64-QAM alphabet, which strongly degrades the BER performance under HPA influences. The ACE/TR scheme cannot compensate this behavior as for 64-QAM only a small number of symbols at the constellation edges can be used for improvements.

Fig. 7 shows the complementary cumulative distribution function (CCDF) of the uncoded and coded bisection methods with different PAPR reduction techniques. There, the curves indicate the probability that the PAPR of one OFDM symbol exceeds a certain threshold, i.e.,

$$\mathrm{CCDF}\left(\mathrm{PAPR}\left(\mathbf{x},\mathbf{c}\right)\right) = P\left(\mathrm{PAPR}\left(\mathbf{x},\mathbf{c}\right) > \beta\right) \ , \qquad (23)$$

where $P(\cdot)$ denotes probability and β is the threshold in dB. For the chosen IBO of the HPA the PAPR does neither increase nor decrease for different loading algorithms compared to the non-adaptive OFDM scheme. Hence, loading has no impact on the PAPR performance. Furthermore, applying additional ACE/TR techniques at this rate does not lead to any gains compared to the non-adaptive case as in the original bisection method too many subcarriers obtain a high order modulation. In contrast, due to the above-mentioned reasons the coded bisection method is able to reduce the PAPR. Here, a gain of 0.5 dB in terms of PAPR at 10^{-4} is visible for the proposed combination of coded bit and power loading and ACE/TR techniques. The EGPA algorithm instead is able to achieve a good compromise between PAPR/BER performance and complexity.

The common principle of clipping and filtering [23] usually achieves better PAPR performance with less complexity compared to our approach. Nevertheless, the increased complexity comes with a large BER gain in terms of efficient loading strategies and no increase in the out-of-band radiation. However, it is possible to apply another clipping device after our approach to further decrease the PAPR as the BER gain can be exploited for further PAPR improvements.

6. CONCLUSIONS

In this contribution a combination of a modified bisection approach and PAPR reduction using ACE and TR techniques is

see the program

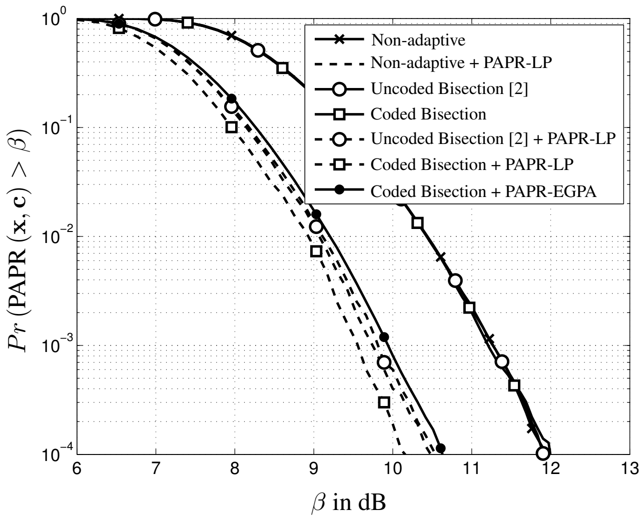


Fig. 7. PAPR CCDFs of different schemes with $N_c = 256$

proposed. Firstly, the channel code is included in the optimization of bit and power loading for multicarrier communication systems. In addition, the ACE technique can exploit the smaller modulation alphabets on the subcarriers allocated by the coded bisection method. Then, the rate loss is compensated for by larger code rate assignments. The proposed combination efficiently enhances the bit error rate performance for high-rate systems especially in presence of memoryless nonlinear distortions. Furthermore, the PAPR reduction is significantly improved with less complexity by an enhanced gradient-project algorithm, which offers a better trade-off between BER and PAPR performance.

7. REFERENCES

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.

[2] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," *IEEE Trans. Commun.*, vol. 48, pp. 23–27, Jan. 2000.

[3] G. Caire, G. Taricco, and E. Biglieri, "Capacity of bit-interleaved channels," *IEE Electronics Letters*, vol. 32, no. 12, pp. 1060–1061, June 1996.

[4] G. Caire, G. Taricco, and E. Biglieri, "Bit-Interleaved Coded Modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[5] C. Bockelmann and K.-D. Kammeyer, "Adaptive Interleaving and Modulation for BICM-OFDM," in *International OFDM-Workshop (InOWo)*, Hamburg, Germany, August 2008.

[6] Y. Li and W. Ryan, "Mutual-information-based adaptive bit-loading algorithms for LDPC-coded OFDM," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1670–1680, May 2007.

[7] C. Stierstorfer and R. F. H. Fischer, "Rate Loading in OFDM Based on Bit Level Capacities," in *International Symposium on Information Theory (ISIT)*, Toronto, Canada, July 2008.

[8] S. Stiglmayr, M. Bossert, and E. Costa, "Adaptive Coding and Modulation in OFDM Systems using BICM and Rate-Compatible Punctured Codes," in *European Wireless*, Paris, France, April 2007.

[9] K. Brueninghaus, D. Astely, T. Sölzer, S. Visuri, A. Alexiou, S. Karger, and G. Seraji, "Link performance models for system level simulations of broadband radio access systems," in *IEEE Intern. Symposium on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Berlin, Germany, Sept. 2005.

[10] S. Stiglmayr, M. Bossert, and E. Costa, "Mutual-information based adaptive coding and modulation in bit-interleaved OFDM systems using punctured LDPC codes," *European Trans. on Telecommun.*, vol. 19, no. 7, pp. 801–811, Sept. 2008.

[11] S. S. Das, F. Tariq, M. I. Rahman, F. B. Frederiksen, E. De Carvalho, and R. Prasad, "Impact of nonlinear power amplifier on link adaptation algorithm of OFDM systems," in *IEEE Vehicular Technology Conf. (VTC) Fall*, Baltimore, MD, USA, Oct. 2007.

[12] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Commun. Mag.*, vol. 12, no. 2, pp. 56–65, Apr. 2005.

[13] Y. J. Kou, W.-S. Lu, and A. Antoniou, "New peak-to-average power-ratio reduction algorithms for multicarrier communications," *IEEE Trans. Circuits Syst. I*, vol. 51, no. 9, pp. 1790–1800, Sept. 2004.

[14] B. S. Krongold and D. L. Jones, "PAR reduction in OFDM via active constellation extension," *IEEE Trans. Broadcast.*, vol. 3, pp. 258–268, Sept. 2003.

[15] C. Bockelmann, D. Wübben, and K.-D. Kammeyer, "Efficient coded bit and power loading for BICM-OFDM," in *IEEE Vehicular Technology Conf. (VTC) Spring*, Barcelona, Spain, Apr. 2009.

[16] M. Petermann, D. Wübben, and K.-D. Kammeyer, "Joint constellation extension and tone reservation for PAPR reduction in adaptive OFDM systems," *submitted to IEEE Intern. Workshop on Signal Proc. Advances for Wireless Comm. (SPAWC)*, 2009.

[17] B. S. Krongold and D. L. Jones, "An active-set approach for OFDM PAR reduction via tone reservation," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 495–509, Feb. 2004.

[18] C. Bradford Barber, D. P. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Trans. on Mathematical Software*, vol. 22, pp. 469–483, 1996.

[19] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[20] O. L. Mangasarian, "A Newton method for linear programming," *Journal of Optimization Theory and Applications*, vol. 121, no. 1, pp. 1–18, Apr. 2004.

[21] H. Schmidt and K.-D. Kammeyer, "Reducing the peak to average power ratio of multicarrier signals by adaptive subcarrier selection," in *IEEE Int. Conf. Universal Personal Commun. (ICUPC)*, Florence, Italy, Oct. 1998.

[22] C. Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal," in *Second European Conference on Satellite Communications (ECSC-2)*, Luettich, Belgium, Oct. 1991.

[23] J. Armstrong, "Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering," *IEE Electronics Letters*, vol. 38, no. 5, pp. 46–247, Feb. 2002.

On the Performance of Trellis Coded Spatial Modulation

Raed Mesleh*, Irina Stefan*, Harald Haas*‡ and Peter M. Grant‡

*Jacobs University Bremen, School of Engineering and Science, Campus Ring 1, Research 1, 28759 Bremen, Germany,
Email: r.mesleh & i.stefan@jacobs-university.de

‡Institute for Digital Communications, Joint Research Institute for Signal and Image Processing, The University of Edinburgh,
Edinburgh EH9 3JL, UK, Email: h.haas & peter.grant@ed.ac.uk

***Abstract*—Trellis coded modulation (TCM) is a well known scheme that applies mapping by set partitioning. The key idea is to group the constellation symbols into sets with each set having the maximum free distance between its symbols. In this paper, a similar approach is applied to antenna constellation points of spatial modulation (SM) in order to enhance its performance in correlated channel conditions. In SM, multiple antennas exist at the transmitter side, but only one of them is active at any particular time instant. The incoming data bits determine the active transmit antenna and the signal constellation point transmitted from it. At the receiver side, the active antenna index and the transmitted symbol are estimated and used together to decode the transmitted information bits. The locations of the transmit antennas are considered as spatial constellation points, and TCM is applied to enhance the bit error ratio (BER) performance of bits encoded into the physical location of an antenna within an antenna array. TCM partitions the entire set of transmit antennas into sub-sets such that the spacing between antennas within a particular sub-set is maximized. The performance of TCSM is analyzed in this paper and compared to the performance of SM and coded V-BLAST (vertical Bell Labs layered space-time) applying optimum sphere decoder algorithm. It is shown that under the same spectral efficiency, TCSM performs nearly the same as SM and V-BLAST in idealistic channel conditions. However, a significant enhancement is reported in the presence of realistic channels performance such as Rician fading and spatial correlation (SC).**

***Index Terms*—MIMO, Spatial modulation, Trellis coded modulation, V-BLAST, Sphere decoder, Correlated channel conditions, Rician fading, Spatial correlation.**

I. Introduction

Wireless radio frequency channel generally poses several challenges on the system design. The physical layer of such systems has to deal with multipath propagation, interference and limited available spectrum. MIMO (multiple-input-multiple-output) transmission technology constructively exploits multipath propagation to provide higher data throughput for the same given bandwidth [1].

One of the most promising MIMO techniques to achieve the expected high data rate is the proposed V-BLAST (vertical Bell Labs layered space-time) architecture [2]. In V-BLAST, the information bit stream is separated in substreams. All the symbols of a certain stream are transmitted through the same antenna (one stream per antenna). The substreams are co-channel signals, that is, they have the same frequency band. Therefore, as compared to a SISO system, a linear increase of the data rate with the number of transmit antennas is achieved. The major task at the receiver is to resolve the inter-channel-interference (ICI) between the transmitted symbols. There are several detection algorithms available in the literature for V-BLAST. The optimum solution is to use maximum likelihood (ML) decoder. The problem of ML algorithm is the high complexity required to search over all possible combinations. Therefore, other algorithms are proposed trying to achieve similar performance as ML detection but with a significant reduction in receiver complexity. A highly potential technique is the proposed sphere decoder (SD) algorithm in [3]. The main idea behind SD is to limit the number of possible codewords by considering only those codewords that are within a sphere centered at the received signal vector. The complexity of separating these signals should be small enough such that the overall complexity of the sphere decoding is lower than that of the full search. In this paper, coded V-BLAST with the SD algorithm is considered.

Traditional MIMO systems use all transmit antennas to simultaneously transmit data to the receiver side. The aim is to improve power efficiency by maximizing spatial diversity [4], or to boost the data rate by transmitting independent streams from each transmit antenna (as in V-BLAST) [1, 2, 5], or to achieve both of them at the same time at the expense of increasing complexity [6].

An alternative multiple antenna transmission technique, called spatial modulation (SM), utilizes the multiple transmit antennas in a different way [7, 8]. Multiple antennas are considered as additional constellation points that are used to carry information bits as seen in Fig. 1. At one time instant, only one transmit antenna is active. Part of the incoming bit sequence determines the active antenna. The active antenna transmits the data symbol and both the transmitted symbol and the active antenna index are estimated at the receiver and used to decode the original information bits.

TCM is a modulation scheme which allows highly efficient and reliable digital transmission without bandwidth expansion or data rate reduction [9]. TCM combines the function of convolutional encoder of rate $R = k/(k+1)$ and M-ary signal mapper that maps $M = 2^{(k+1)}$ constellation points. Unlike conventional coding techniques only certain sequences of successive constellation points are allowed (mapping by set partitioning). The key idea is to group symbols into sets

see the program

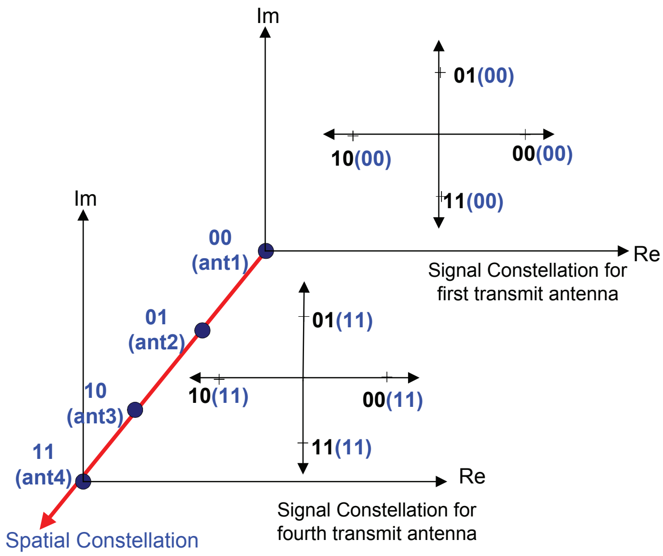


Fig. 1. Spatial modulation maps a sequence of bits into a signal constellation and into a spatial constellation point. The spatial constellation point data is shown inside brackets in the figure. In this example 4-PSK and four transmit antennas are considered. Each spatial constellation point defines an independent complex plane of signal constellation points. For illustration purpose, two such planes, for the sequences (00) and (11), are shown in the figure.

of equal sizes where each set maximizes the free distance between its symbols.

In this paper, the key idea of TCM is applied to the antenna constellation points of SM. This novel scheme is called trellis coded spatial modulation (TCSM). In TCSM, the transmit antennas are partitioned into sub-sets in such a way that the spatial spacing between antennas in the same sub-set is maximized. Therefore, the effect of correlated channels on the performance of SM is reduced. This fact is significant when considering portable devices with multiple antennas installed in compact space and enough separation between them cannot be guaranteed. The performance of the proposed idea is analyzed in the presence of Rician fading and spatial correlation (SC) channels and major enhancements in BER is reported as compared to SM and V-BLAST with the same spectral efficiency.

The rest of the paper is organized as follows: In Section II, the system model of TCSM is presented. V-BLAST system model is discussed in Section III. Section IV present the channel models. Simulation results are discussed in Section V. Finally, Section VI concludes the paper.

II. Trellis Coded Spatial Modulation (TCSM) System Model

The TCSM system model is depicted in Fig. 2. A MIMO system consisting of four transmit antennas ($N_t = 4$) and four receive antennas ($N_r = 4$) is considered as an example. The transmitted bits at each time instant are grouped as the row vectors of the matrix $\mathbf{x}(t)$. For illustration purposes, the incoming bit sequences are considered $\mathbf{x}(t) = \begin{bmatrix} 001 & 110 & 111 \end{bmatrix}^T$, where $(\cdot)^T$ denotes the transpose of a vector or a matrix. The first step is to split this matrix into two matrices. The first matrix $\mathbf{x_1}(t)$ contains the bits that are mapped to spatial constellation points. While the second matrix contains the bits that are mapped to signal constellation points[1]. In the considered example, $\mathbf{x_1}(t) = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^T$ and $\mathbf{x_2}(t) = \begin{bmatrix} 01 & 10 & 11 \end{bmatrix}^T$. Assuming 4-PSK (phase shift keying) constellation, as seen in Fig. 1, the second matrix is mapped to $\begin{bmatrix} i & -1 & -i \end{bmatrix}^T$, where each element in this matrix corresponds to the symbol that is transmitted from one antenna among the set of existing transmit antennas at one time instant. The first matrix, $\mathbf{x_1}(t)$, is then used to select the active transmit antenna. However, before mapping the bits in the first matrix to the spatial constellation points (the transmit antenna indexes), the bits are processed by a rate 1/2 TCM encoder. The TCM encoder block consists of a convolutional encoder followed by a random block interleaver. The TCM encoder, state transition, and spatial mapping are depicted in Fig. 3. TCM groups the antenna indexes in a tree like fashion, then separates them into two limbs of equal size. At each limb of the tree, the indexes are further apart. In other words, TCM partitions the transmit antennas into sub-sets with the constraint of maximizing the spacing of antennas belonging to the same sub-set. In the given example and assuming all antennas are equally spaced on a vertical line, antennas one and three form a set and antennas two and four form the other set. The output of the TCM encoder is then used to select the active antenna. In the above example, $\mathbf{x_1}(t)$ is transformed into another matrix $l(t) = \begin{bmatrix} 00 & 01 & 11 \end{bmatrix}^T$ by the encoder of Fig. 3(c) assuming the initial state of the encoder is 00. The SM mapper operates on both $l(t)$ and $\mathbf{x_2}(t)$ matrices creating the output matrix

$$s(t) = \begin{bmatrix} i & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -i \end{bmatrix}.$$

Each column from the output matrix is transmitted at a single time instant from the existing transmit antennas over the MIMO channel $\mathbf{H}(t)$. For instance, at the first time instant in the considered example, the elements of the first column are transmitted from the four transmit antennas. Since, however, only *one* element is different from zero, only *one* antenna emits a signal. This means, that only the first antenna is active at this particular time instant and is transmitting symbol i while all other antennas are switched off. The signal experiences an N_r- dim additive white Gaussian noise (AWGN). The channel and the noise are assumed to have independent and identically distributed (iid) entries according to $\mathcal{CN}(0, 1)$. In addition, Rician fading and Kronecker SC channel models are considered. The complete models for the channel with Rician fading and spatial correlation are discussed in Section IV.

At the receiver, the optimum SM decoder proposed in [10]

[1] This is the basic working mechanism of SM, the reader is kindly requested to refer to [7] for detailed information.

see the program

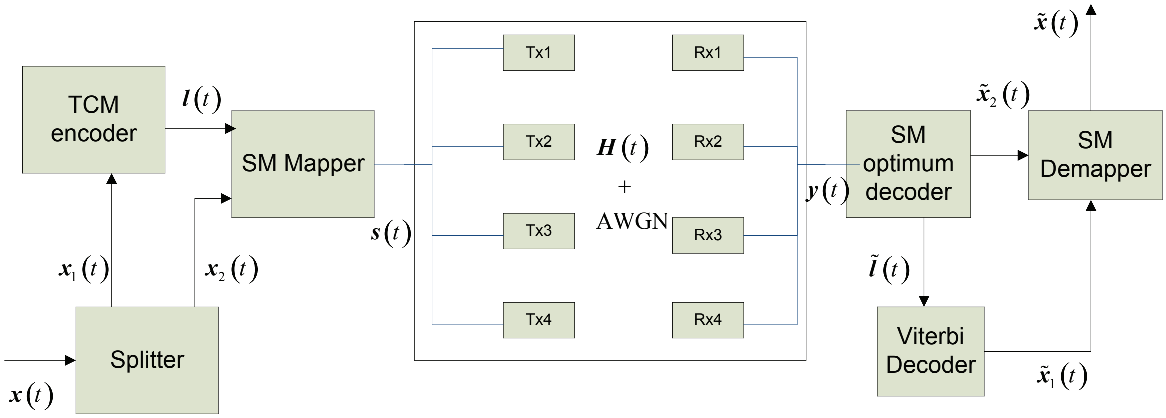


Fig. 2. Trellis coded spatial modulation (TCSM) system model

is considered to estimate the transmitted symbol $\tilde{\mathbf{x}}_2(t)$ and the transmit antenna index $\tilde{l}(t)$ as follows:

$$\begin{aligned}[\tilde{x}_2, \tilde{l}] &= \arg\max_{\tilde{x}_2, \tilde{l}} p_{\mathbf{y}}\left(\mathbf{y}|\mathbf{s}_{l,m}, \mathbf{H}\right) \\ &= \arg\max_{\tilde{x}_2, \tilde{l}} \sqrt{\rho}\|\mathbf{g}_{l,m}\|_{\mathrm{F}}^2 - 2\,\mathrm{Re}\left\{\mathbf{y}^{\mathbf{H}}\mathbf{g}_{l,m}\right\}, \quad (1)\end{aligned}$$

where $\mathbf{g}_{l,m} = \mathbf{h}_l s_{l,m}$ is the received vector when transmitting the symbol $s_{l,m}$ from antenna index l where $1 \leq l \leq N_t$ and $1 \leq m \leq M$ and $\mathbf{h}_l$ is the channel vector containing the channel path gains from transmit antenna l to all receive antennas; M is the size of the signal constellation diagram and Re is the real part of a complex number. In addition, ρ is the average signal to noise ratio (SNR) at each receive antenna, and

$$p_{\mathbf{y}}\left(\mathbf{y}|\mathbf{s}_{l,m}, \mathbf{H}\right) = \pi^{-N_t} \exp\left(-\|\mathbf{y} - \sqrt{\rho}\mathbf{H}\mathbf{s}_{l,m}\|_{\mathrm{F}}^2\right) \quad (2)$$

is the probability density function (pdf) of $\mathbf{y}$ conditioned on the transmitted symbol $\mathbf{s}_{l,m}$ from antenna index l and the channel $\mathbf{H}$. The notation $\|\cdot\|_{\mathrm{F}}$ stands for the Frobenius norm of a vector or a matrix.

The estimated antenna number is de-mapped to the corresponding bits and the incoming data sequence of one complete frame is applied to a random block deinterleaver and then decoded using a hard decision Viterbi decoder. The output from the Viterbi decoder together with the estimated symbols are used to retrieve the original information bits.

In this paper, the performance of TCSM scheme is compared to SM and V-BLAST. SM applies no channel coding and uses a smaller number of transmit antennas or lower modulation order to achieve the same spectral efficiency as TCSM. V-BLAST system model, on the other hand, is discussed in the following section.

III. V-BLAST System Model

The considered V-BLAST system in this paper is depicted in Fig. 4. The rate 1/2 convolutional encoder shown in Fig. 3 is considered. Coded V-BLAST system is generally termed

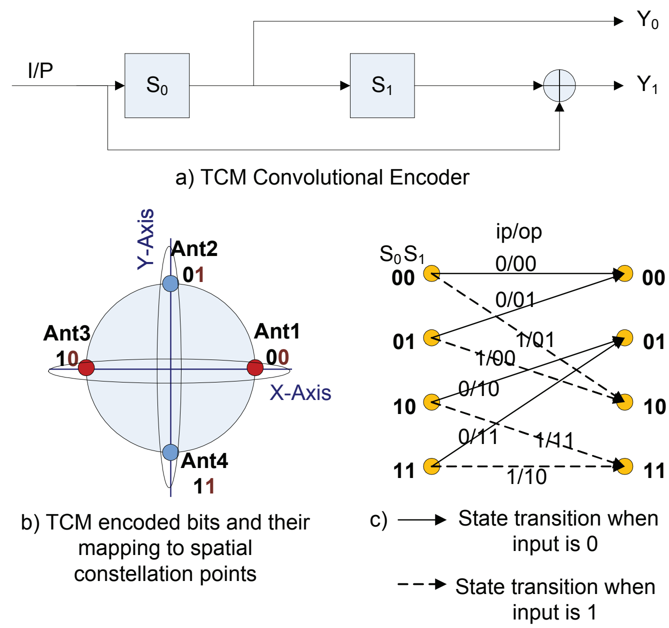


Fig. 3. Trellis coded encoder, spatial constellation mappings, and state transitions of the TCM encoder

horizontal BLAST (H-BLAST) [11]. In H-BLAST, the incoming bit stream is demultiplexed into N_t parallel substreams. Channel coding followed by interleaving is applied to each substream. The coded bits are modulated and then transmitted from the corresponding transmit antenna. If the interleaving depth is selected to be larger than the coherence time of the channel, additional diversity gain can be achieved [12].

Another way of applying channel coding to V-BLAST is to use a single channel code for all layers as shown in Fig. 4. This scheme is called single coded BLAST (SCBLAST) [12]. SCBLAST is simpler than H-BLAST in the sense that only a single channel encoder is needed for all layers. In addition, in correlated slow or block fading channel, SCBLAST outperforms H-BLAST since the demultiplexer (at the transmitter)

see the program

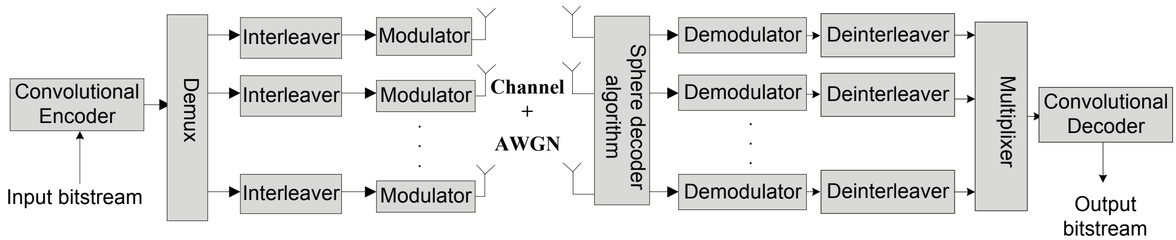


Fig. 4. V-BLAST system model

and multiplexer (at the receiver) act as spatial interleavers where they together help in breaking some of the correlation in the received signal [12]. In this paper, correlated slow fading channels are assumed and therefore, SCBLAST is considered for the comparisons used in this paper.

At the receiver, SD algorithm is employed to detect the transmitted symbols from all layers. In simulations, SD algorithm based on integer lattice theory is implemented. A complex MIMO system is decoupled into its real and imaginary parts so as to form an equivalent real-valued system. This approach is most appropriate for lattice-based modulation schemes such as quadrature amplitude modulation (QAM) or pulse amplitude modulation (PAM) [3, 13, 14]. For other complex constellations such as phase-shift keying (PSK), the SD based on integer lattice theory are inefficient due to the existence of invalid candidates. A solution is to avoid decoupling of the complex system by applying complex SD algorithms [15].

The SD algorithm avoids an exhaustive search by examining only those points that lie inside a sphere with radius C. The performance of the SD algorithm is closely tied to the choice of the initial radius. The radius should be chosen large enough so that the sphere contains the solution. However, the larger the radius is, the longer the search takes which increases the complexity. On the other hand, a small radius may cause the algorithm to fail finding any point inside the sphere. In this paper, the initial radius of the SD algorithm is adjusted according to the noise level assuming the knowledge of the SNR at the receiver side. If no point is found inside the sphere, the search is repeated with a larger radius ($C = C + 1$) [16]. This approach is shown to perform near optimum maximum likelihood detection [16].

SD receivers have been implemented in custom application-specific integrated circuits (ASICs) [17] and as simplified fixed complexity designs [18] conveniently realized in field-programmable gate arrays (FPGAs) [19].

The output symbols from SD are demodulated and the bits are deinterleaved. The bits from all layers are multiplexed and hard decision Viterbi decoder is then applied.

IV. Channel models

In this paper, $\mathbf{H}$ is an $N_r \times N_t$ flat fading channel matrix representing the path gains h_{ij} between transmit antenna j and receive antenna i.

$$\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & h_{22} & \cdots & h_{2N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r1} & h_{N_r2} & \cdots & h_{N_rN_t} \end{bmatrix} \tag{3}$$

In case of NLOS (non-line-of-sight), the sum of all scattered components of the received signal is modeled as a zero mean complex Gaussian random process given by $\alpha(t) = \alpha_1(t) + \sqrt{-1}\,\alpha_2(t)$, where $\alpha_1(t)$ and $\alpha_2(t)$ are assumed to be real valued statistically independent Gaussian random processes. As a result, the phase of the random process $\alpha(t)$ takes a uniform distribution and the amplitude takes a Rayleigh distribution. Therefore, a static fading Rayleigh channel matrix that is flat for all frequency components is modeled.

A. Rician fading channel

If a LOS path exists between the transmit and receive antennas, the channel amplitude gain is characterized by a Rician distribution and the channel is said to exhibit Rician fading. The Rician fading MIMO channel matrix can be modeled as the sum of the fixed LOS matrix and a Rayleigh fading channel matrix as follows [20]:

$$\mathbf{H}_{\text{Ricean}}(t) = \sqrt{\frac{K}{1+K}}\bar{\mathbf{H}}(t) + \sqrt{\frac{1}{1+K}}\mathbf{H}(t), \tag{4}$$

where $\sqrt{\frac{K}{1+K}}\bar{\mathbf{H}}$ is the LOS component, $\sqrt{\frac{1}{1+K}}\mathbf{H}$ is the fading component, and K is the Rician K-factor. The Rician K-factor is defined as the ratio of the LOS and the scatter power components and $\bar{\mathbf{H}}$ is a matrix with all elements being one.

B. Spatial correlation (Kronecker model)

The channel correlation depends on both the environment and the spacing of the antenna elements. It is assumed that correlations at the transmitter and receiver array are independent of each other because the distance between the transmit and receive array is large compared to the antenna element spacing.

To incorporate the SC into the channel model, the correlation among channels at multiple elements needs to be calculated. The correlated channel matrix is then modeled using the Kronecker model [21].

$$\mathbf{H}^{\text{corr}}(t) = \mathbf{R}_{\text{rx}}^{1/2}\mathbf{H}(t)\mathbf{R}_{\text{tx}}^{1/2} \tag{5}$$

see the program

The correlation matrices are computed analytically based on the power azimuth spectrum (PAS) distribution and array geometry [21]. A clustered channel model, in which groups of scatterers are modeled as clusters located around the transmit and receive antennas, is assumed. The clustered channel model is validated through measurements [22] and adopted by various wireless system standard bodies such as the IEEE 802.11n Technical Group (TG) [23] and the 3GPP/3GPP2 Technical Specification Group (TSG) [24].

V. Simulation Results

In order to validate the proposed TCSM idea, Monte Carlo simulation results for at least 10^6 channel realizations are obtained and the average BER is plotted versus the average SNR at each receiver input. In all simulations where Rician fading is considered, channel correlation due to antenna spacing is zero, but the Rician K factor is set to $K = 3$. This value is within the range of the measured values in indoor wireless communication [25]. For SC channel model, similar parameters as discussed in [7] are adopted here as well, except that the element spacing at the transmitter and the receiver are set to 0.1λ and 0.5λ, respectively. The 0.5λ separation between the antennas can achieve relatively low correlation assuming the receiver is surrounded by a large number of local scatterers [26]. The 0.1λ element spacing at the transmitter results in high correlation which models a small mobile device with multiple antennas where large separation between the antennas cannot be achieved.

A. TCSM and SM performance comparison

In the first results, depicted in Figs. 5, 6 and 7, TCSM performance under ideal, Rician fading, and SC channel conditions are plotted, respectively, and compared to SM performance under similar channel conditions. TCSM transmits 4QAM symbol from a 4x4 MIMO system and applies the rate 1/2 TCM encoder. Therefore, TCSM achieves 4 b/s/Hz but only 3 bits are data bits and the fourth one is a coding bit. SM transmits a BPSK symbol from a 4x4 MIMO system achieving 3 b/s/Hz spectral efficiency. Thus, the two systems have the same spectral efficiency.

In ideal channel conditions (the channel paths are uncorrelated), the BER of the two systems are compared in Fig. 5. SM performs slightly better than TCSM. The reason for this behavior is that the TCM coding gain and the set partitioning of the transmit antennas has no advantage since all channel paths are uncorrelated. However, the situation is different if correlated channel paths are considered, i.e. when Rician fading and SC channels are considered. The advantage of TCSM over SM is obvious from Figs. 6 and 7. A SNR gain of about 6 dB in Rician fading channel at a BER of 10^{-4} can be noticed in Fig. 6. In addition, similar gain in SNR at a the same BER is noticed in Fig. 7 for SC channel. The significant gains in the presence of channel correlations due to Rician fading or SC can be attributed to TCM encoding and the underlying partitioning of the transmit antennas. The fact that the transmit antennas with larger separation distance

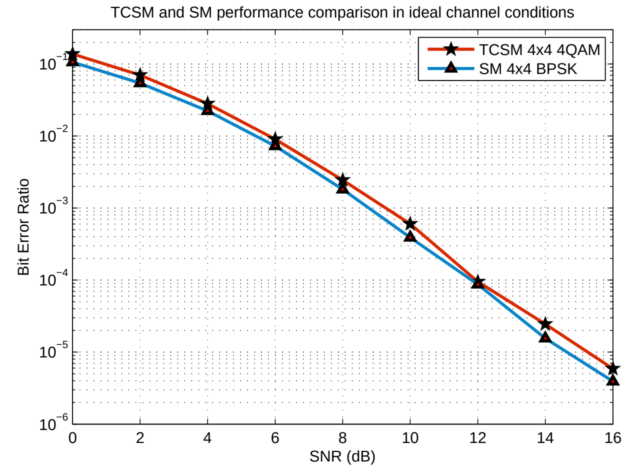


Fig. 5. Performance comparison of SM 4x4 BPSK and TCSM 4x4 4QAM for ideal channel

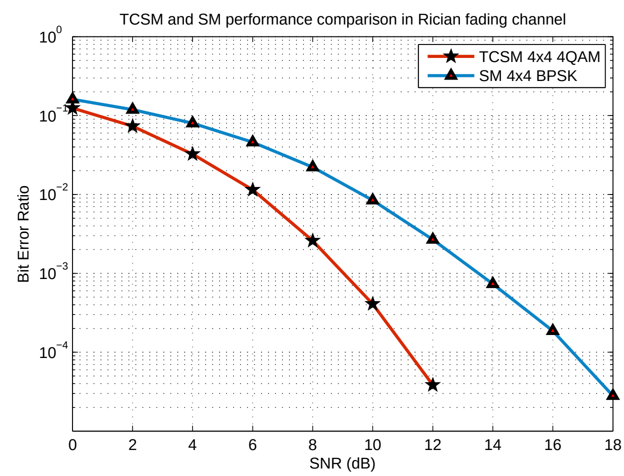


Fig. 6. Performance comparison of SM 4x4 BPSK and TCSM 4x4 4QAM in Rician fading channel with Rician K-factor of 3

are grouped in one set, reduces the effect of correlation and results in a better performance.

B. TCSM and V-BLAST performance comparison

In the second set of results, the BER of TCSM and V-BLAST[2] are compared in ideal, Rician fading, and SC channel conditions as depicted in Figs. 8, 9, and 10, respectively. Two spectral efficiencies are studied for each system. TCSM transmits 4QAM and 32QAM symbols from 4x4 MIMO system achieving a spectral efficiencies of 3 b/s/Hz and 6 b/s/Hz, respectively. V-BLAST system transmits 4QAM and 16QAM symbols from a 3x4 MIMO system and applying the rate 1/2 channel encoder, therefore, achieving 3 b/s/Hz and 6 b/s/Hz spectral efficiencies, respectively.

[2] V-BLAST here refers to SCBLAST discussed in Section III. The name of V-BLAST is retained here for consistency.

see the program

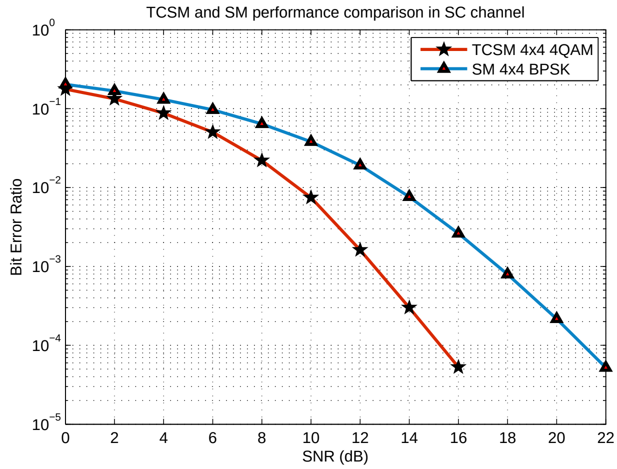


Fig. 7. Performance comparison of SM 4x4 BPSK and TCSM 4x4 4QAM in SC channel with transmit antenna element spacing of 0.1λ at the transmitter and 0.5λ at the receiver

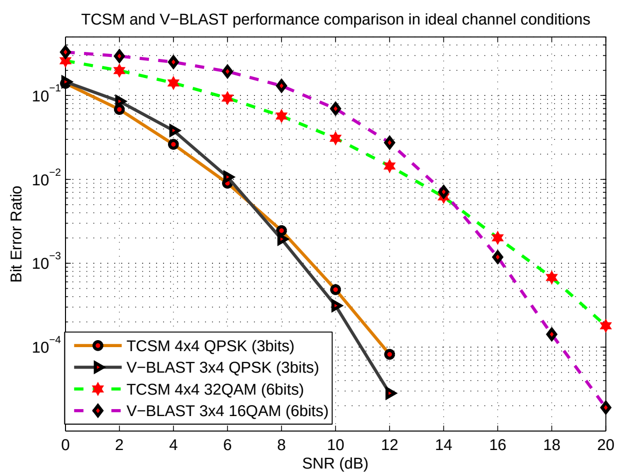


Fig. 8. Performance comparison of TCSM and V-BLAST in ideal channel conditions

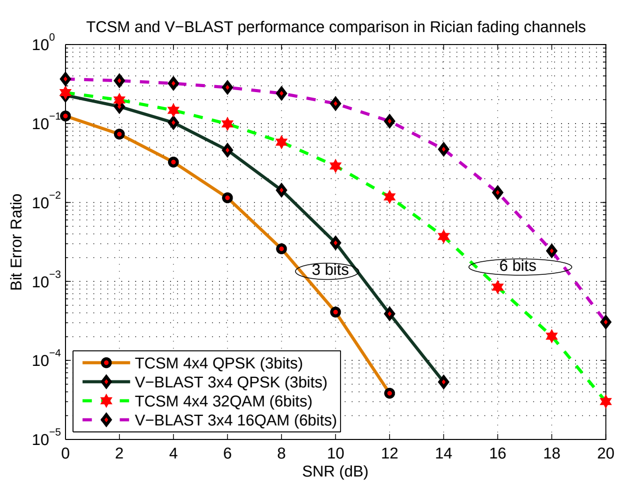


Fig. 9. Performance comparison of TCSM and V-BLAST in Rician fading channel with Rician K-factor of 3

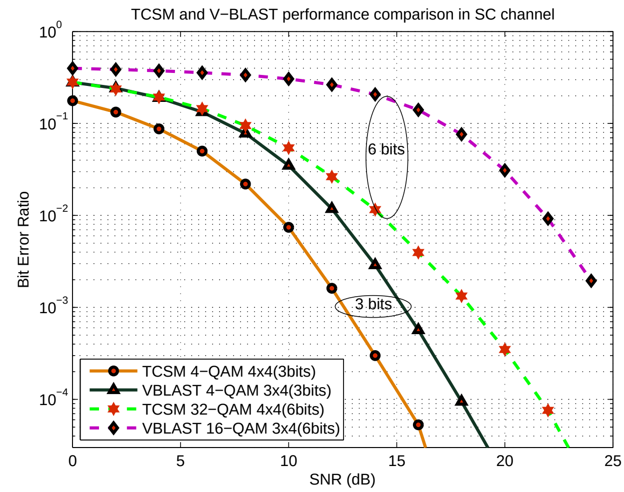


Fig. 10. Performance comparison of TCSM and V-BLAST in SC channel with transmit antenna element spacing of 0.1λ and receive antenna element spacing of 0.5λ

In ideal channel condition, TCSM and V-BLAST outperform each other in a range of SNRs. The BER curves intersect at 7 dB for 3 b/s/Hz and at 14 dB for 6 b/s/Hz as shown in Fig. 8. The channel coding gain in V-BLAST causes BER enhancements at high SNR. In addition, and as discussed previously, the effect of TCM coding and set partitioning are insignificant as the channel paths are uncorrelated in this scenario. The BER gain of V-BLAST over TCSM at high SNR is larger for the case of 6 b/s/Hz. This is mainly because the higher coding gain of V-BLAST at high SNR and the fact that TCSM uses higher order modulation in order to achieve similar spectral efficiency as V-BLAST. The improvements of TCSM over V-BLAST at low SNR is not related to the TCM coding and set partitioning. It is mainly due to the underlying working mechanism of SM and the fact that it completely avoids inter-channel-interference (ICI) at the receiver side [7].

In Fig. 9, the performance of V-BLAST and TCSM is compared in the presence of line of sight (LOS) path between transmitter and receiver (Rician fading channel). Rician fading enhances the SNR at the receiving antennas, but increases the correlation between the antenna elements [27]. Therefore, Rician fading significantly degrades the performance of V-BLAST and SM [7]. This degradation can be observed for V-BLAST in the results depicted in Fig. 9. As compared to the obtained results in ideal channel conditions, V-BLAST requires 2 dB and 3 dB increase in SNR to achieve a BER of 10^{-3} for 3 b/s/Hz and 6 b/s/Hz, respectively. However, TCSM seems to be less affected by the presence of Rician fading. In fact, it demonstrates even better performance as compared to ideal channel conditions results. For instance, the 32QAM 4x4 TCSM achieves a BER of 10^{-3} at a SNR of about 19 dB in ideal channel condition. Though, it achieves the same BER at a SNR of 16 dB in the presence of Rician fading channel. Indeed, TCSM demonstrates better performance in the presence

see the program

of LOS path between the transmitter and the receiver. This can be explained by the fact that Rician fading increases the SNR at the receiver side and the underlying set partitioning together with TCM coding eliminates the correlation between transmit antennas. Nevertheless, it should be mentioned that the performance of TCSM in Rician fading channels depends on the number of transmit antennas, the considered modulation order, and the Rician K-factor. However, for the systems considered in this paper and in the presence of Rician fading, TCSM outperforms V-BLAST by 2 dB and 3 dB in SNR at a BER of 10^{-3} for 3 b/s/Hz and 6 b/s/Hz, respectively.

Finally, the effect of SC on the performance of TCSM and V-BLAST is studied and the results are shown in Fig. 10. The presence of correlation degrades the performance of the two systems. Again, TCSM is significantly less affected by the presence of SC as compared to V-BLAST. As compared to the results obtained in ideal channel conditions, SC degrades the performance of TCSM by 3 dB and 1 dB in SNR at a BER of 10^{-3} for 3 b/s/Hz and 6 b/s/Hz, respectively. While V-BLAST system performance degrades by about 5 dB and 9 dB in SNR at a BER of 10^{-3} for 3 b/s/Hz and 6 b/s/Hz, respectively. TCSM outperforms V-BLAST by about 3 dB and 6 dB in SNR at a BER of 10^{-3} for 3 b/s/Hz and 6 b/s/Hz, respectively.

In summary, the basic idea of the proposed scheme is to divide the existing antennas into sets using TCM such that each set maximizes the spatial distance between its antennas, and therefore minimizes the effect of correlation fading.

VI. Summary and Conclusion

SM is a radically different and relatively new MIMO approach. It has the important feature that it fully *avoids* inter-channel interference while it still enables the system to benefit from spatial multiplexing gains. The key to achieving this goal is the exploitation of the relative dislocation of the antennas within an antenna array. This dislocation is used to *implicitly* convey extra data bits. At the receiver a new block, namely an antenna detector, is required. This paper, for the first time, proposes the use of trellis coded modulation for the correction of data errors that occur within the antenna detector block, *i.e.*, the correction of erroneous data bits that are implicitly encoded into the location of the antenna.

TCM concept is adopted in a novel way in this paper to combat performance degradation of SM in correlated channel conditions. TCM is applied to the spatial constellation points of SM. In the proposed TCSM, only certain sequences of successive spatial constellation points are allowed which reduces the correlation between neighboring antennas. TCSM performance is analysed in this paper and compared to the performance of SM and V-BLAST. Major enhancements in SNR are reported in Rician fading and spatially correlated channel conditions. The proposed TCSM allows the integration of multiple antenna system in small devices with antenna separations as low as 0.1λ. It is also suitable for indoor applications with direct LOS between transmitter and receivers.

Future work will consider different channel conditions and TCM encoders with different states.

Acknowledgment

Harald Haas acknowledges the Scottish Funding Council support of his position within the Edinburgh Research Partnership in Engineering and Mathematics between the University of Edinburgh and Heriot Watt University.

References

[1] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Transaction on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1996.

[2] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an Architecture for Realizing very High Data Rates over the Rich-Scattering Wireless Channel," in *Unino Radio-Scientifique Internationale (URSI) Intern. Symp. on Signals, Systems, and Electronics (ISSSE)*, Sep. 29–Oct. 2, 1998, pp. 295–300.

[3] E. Viterbo and J. Boutros, "A Universal Lattice Code Decoder for Fading Channels," *IEEE Transaction on Information Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

[4] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[5] G. J. Foschini, "Layered Space-Ttime Architecture for Wireless Communication in a Fading Environment when Using Multi-Element Antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Sep. 1996.

[6] M. Sellathurai and S. Haykin, "Turbo-BLAST for Wireless Communications: Theory and Experiments," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2538–2546, Oct. 2002.

[7] R. Mesleh, H. Haas, S. Sinanović, C. W. Ahn, and S. Yun, "Spatial Modulation," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2228 – 2241, Jul. 2008.

[8] S. Ganesan, R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "On the Performance of Spatial Modulation OFDM," in *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, 29 Oct. - 1 Nov. 2006, pp. 1825–1829.

[9] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Journal on Information Technology*, vol. 28, no. 1, pp. 55–67, 1982.

[10] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial Modulation: Optimal Detection and Performance Analysis," *IEEE Communication Letters*, vol. 12, no. 8, pp. 545–547, 2008.

[11] G. Foschini, D. Chizhik, M. Gans, C. Papadias, and R. Valenzuela, "Analysis and Performance of Some Basic Space-Time Architectures," *IEEE Journal on Selected Areas in Communications [Invited Paper]*, vol. 21, no. 3, pp. 303–320, Apr. 2003.

[12] T. M. Duman and A. Ghrayeb, *Coding for MIMO Communication Systems*. John Wiley and Sons, 2007.

[13] B. Hochwald and S. ten Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[14] O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice Coode Decoder for Space-Time Codes," *IEEE Communications Letters*, vol. 4, no. 5, pp. 161–163, May 2000.

[15] A. Chan and I. Lee, "A New Reduced-Complexity Sphere Decoder for Multiple Antenna Systems," in *Proc. of IEEE International Conference on Communications (ICC'02)*, vol. 1, NY, USA, Apr. 28–May 2, 2002, pp. 460–464.

[16] M. Damen, K. Abed-Meraim, and M. Lemdani, "Further Results on the Sphere Decoder," in *Proc. of the IEEE International Symposium on Information Theory*, Washington, DC, Jun. 24–29 2001, pp. 333–.

[17] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI Implementation of MIMO Detection Using the Sphere Decoding Algorithm," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.

[18] L. Barbero and J. Thompson, "Extending a Fixed-Complexity Sphere Decoder to Obtain Likelihood Information for Turbo-MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 5, pp. 2804–2814, Sep. 2008.

see the program

[19] ——, “FPGA Design Considerations in the Implementation of a Fixed-Throughput Sphere Decoder for MIMO Systems,” in *Proc. International Conference on Field Programmable Logic and Applications FPL ’06*, Madrid, Spain, Aug. 28–30 2006, pp. 1–6.

[20] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. U.K.: Cambridge University Press, 2003.

[21] A. Forenza, D. Love, and R. Heath Jr., “A Low Complexity Algorithm to Simulate the Spatial Covariance Matrix for Clustered MIMO Channel Models,” in *Proc. of the IEEE Vehicular Technology Conference (VTC 2004-Fall)*, vol. 2, Los Angeles, CA, USA, 17-19 May 2004, pp. 889–893.

[22] Q. Spencer, B. Jeffs, M. Jensen, and A. Swindlehurst, “Modeling the Statistical Time and Angle of Arrival Characteristics of an Indoor Multipath Channel,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 347–360, Mar. 2000.

[23] V. Erceg et al., *TGn Channel Models*, IEEE P802.11 Wireless LANs, IEEE Std. IEEE 802.11-03/940r4, May 10, 2004. Retrieved Jan. 12, 2007 from http://www.nari.ee.ethz.ch/ dsbaum/11-03-0940-04-000n-tgn-channel-models.pdf.

[24] G. T. S. Group, *Spatial Channel Model*, Spatial Channel Model AHG (Combined ad-hoc from 3GPP and 3GPP2), 3GPP Std. SCM-134 text V6.0, Apr. 22, 2003.

[25] M. Carroll and T. Wysocki, “Fading Characteristics for Indoor Wireless Channels at 5GHr Unlicensed Bands,” in *IEEE Joint First Workshop on Mobile Future and Symposium on Trends in Communications (SympoTIC’03)*, Bratislava, Slovakia, Oct. 26–28 2003, pp. 102–105.

[26] A. Hottinen, O. Tirkkonen, and R. Wichman, *Multi-Antenna Tansceiver Techniques for 3G and Beyond*. John Wiley & Sons Ltd., 2003.

[27] M. Godavarti, I. Hero, A.O., and T. Marzetta, “Min-Capacity of a Multiple-Antenna Wireless Channel in a Static Rician Fading Environment,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1715–1723, 2005.

see the program

SELECTED SORTING FOR PAR REDUCTION IN OFDM MULTI-USER BROADCAST SCENARIOS

Christian Siegl and Robert F.H. Fischer

Lehrstuhl für Informationsübertragung, Friedrich–Alexander–Universität Erlangen–Nürnberg
Cauerstraße 7/LIT, 91058 Erlangen, Germany, Email: {siegl,fischer}@LNT.de

ABSTRACT

OFDM suffers from a high peak-to-average power ratio (PAR) of the transmit signal. This issue becomes even more important when considering multi-antenna systems. In this paper PAR reduction schemes for the multi-antenna broadcast scenarios are assessed. Hereby, the scheme Selected Sorting (SLS) is introduced and analyzed in terms of PAR reduction performance, error performance, and computational complexity. The huge benefit of this scheme is that no side information needs to be signaled to the receiver. Numerical results shown in this paper, demonstrate that SLS offers significant gains in PAR reduction. Moreover, Selected Sorting is compared with simplified Selected Mapping (sSLM), whereby SLS outperforms sSLM with respect to all three parameters, PAR reduction, error performance, and computational complexity.

1. INTRODUCTION

Orthogonal frequency-division multiplexing (OFDM) is a popular scheme for equalizing the temporal interferences caused by frequency-selective channels. One essential drawback of OFDM systems is the high dynamic of the transmit signal. The occurrence of large signal peaks leads to signal clipping at the non-linear power amplifier, which in turn leads to very undesirable out-of-band radiation. In order to avoid violating spectral masks a transmitter sided algorithmic control of the peak power is essential. Moreover, the transmitters of modern communication systems will be equipped with multiple antennas (multiple-input/multiple-output (MIMO) systems). In this case the issue of out-of-band radiation gets even more serious and the reduction of the signal's peak power is more relevant. Recently, peak power reduction schemes, developed for single antenna systems, are extended to the MIMO case. For instance, this has been done for the popular scheme *Selected Mapping (SLM)* [1, 2, 3, 4]. However, in most cases these extensions have only been discussed for multi-antenna point-to-point scenarios where the equalization of the spatial interferences can be accomplished at the receiver side.

This paper deals with multi-antenna point-to-multipoint transmission (broadcast scenario). Hereby, it is essential to apply a transmitter sided precoding [5, 6] of the channel's spatial (or multi-user) interferences. The combination of transmitter sided precoding with peak-power reduction algorithms is not always possible and may lead to degradation of the error performance or to a significant increase of computational complexity.

In this paper a peak-power reduction technique named *Selected Sorting (SLS)*, originally introduced in [7], is examined in details and further improved. Moreover, it is compared to *simplified Selected Mapping (sSLM)* [2], the extension of SLM to broadcast scenarios, in terms of peak-power reduction, bit error performance, and computational complexity.

This paper is organized as follows: in Section 2 the system model is defined; Section 3 gives a short definition of *sorted Tomlinson-Harashima precoding (sTHP)* [5, 6], the precoding technique which is considered in this paper. Moreover, the impact of different sorting orders is assessed. In Section 4 peak-power reduction for broadcast scenarios are assessed analytically and by numerical simulations. Section 5 draws some conclusions.

2. SYSTEM MODEL

In this paper, we consider transmission over a frequency-selective multi-user broadcast channel. The transmitter (central unit) is equipped with N_{C} antennas where joint signal processing is possible; the receivers are K distributed users each equipped with a single antenna. For convenience we restrict to the case $N_{\mathrm{C}} \overset{!}{=} K$. The impulse response of the respective channel in the z-domain is given (in the equivalent complex baseband) by a matrix polynomial $\boldsymbol{H}(z) = \sum_{k=0}^{l_{\mathrm{H}}-1} \boldsymbol{h}_k \cdot z^{-k}$. The fading coefficient at delay step k is given by the complex

This work was supported by Deutsche Forschungsgemeinschaft (DFG) within the framework TakeOFDM under grants FI 982/1-1 and 1-2.

see the program

matrix $\boldsymbol{h}_k$ which describes the multi-user interferences; l_{H} is the length of the channel impulse response. In order to equalize the temporal interferences OFDM is applied, whereby D subcarriers are assumed. The remaining multi-user interferences at each subcarrier are described by the flat fading channel matrix $\boldsymbol{H}_d = \boldsymbol{H}(\mathrm{e}^{\mathrm{j}2\pi d/D})$, $d = 0, \dots, D-1$ and have to be equalized by transmitter-sided precoding. In this paper we consider *sorted Tomlinson-Harashima Precoding (sTHP)* [6].

The complex-valued modulation symbols (drawn from an M-ary QAM constellation) are collected in the $K \times D$ matrix $\boldsymbol{A} = [A_{k,d}]$, the frequency-domain MIMO OFDM frame. In order to equalize the multi-user interferences sTHP has to be applied on each column (vector $\boldsymbol{A}_d = [A_{(k=1,\dots,K),d}]$, $d = 1, \dots, D$) of $\boldsymbol{A}$ (see Section 3).

The resulting precoded frequency-domain MIMO OFDM frame is denoted by the matrix $\boldsymbol{X}$. The time-domain MIMO OFDM frame (matrix $\boldsymbol{x}$) is obtained via an *inverse discrete Fourier transform (IDFT)* [8] along each row of the matrix $\boldsymbol{X}$.

Assuming that the frequency-domain modulation symbols $A_{k,d}$ and hence $X_{k,d}$ are statistically independent, the time-domain symbols $x_{k,d}$ are (approximately) Gaussian distributed due to the central limit theorem. Hence, the transmit signal at the k^{th} antenna exhibits a large *peak-to-average power ratio (PAR)*. As usual in literature we consider the worst-case PAR[1], i.e., the maximum PAR over all antennas which is defined as

$$\mathrm{PAR} \stackrel{\text{def}}{=} \frac{\max_{\forall d, \forall k} |x_{k,d}|^2}{\mathrm{E}\{|x_{k,d}|^2\}}\,. \tag{1}$$

As performance measure of the PAR reduction schemes discussed in this paper, we consider the *complementary cumulative distribution function (ccdf)* of the PAR, i.e., the probability that the PAR of a given OFDM frame exceeds a certain threshold $\mathrm{PAR}_{\mathrm{th}}$:

$$\mathrm{ccdf}(\mathrm{PAR}_{\mathrm{th}}) \stackrel{\text{def}}{=} \Pr\{\mathrm{PAR} > \mathrm{PAR}_{\mathrm{th}}\}. \tag{2}$$

Assuming all samples of the time-domain signal $x_{k,d}$ to be i.i.d. Gaussian distributed the ccdf of the original signal is given by [3]

$$\mathrm{ccdf}_{\mathrm{MIMO}}(\mathrm{PAR}_{\mathrm{th}}) \stackrel{\text{Gauss}}{=} 1 - (1 - \mathrm{e}^{-\mathrm{PAR}_{\mathrm{th}}})^{DK}. \tag{3}$$

[1] In this paper we do not consider pulse shaping, modulation to radio frequency and the influence of the cyclic prefix. Moreover, we restrict ourselves to the PAR of the non-oversampled transmit signal. Considering oversampling would have no impact on the relation between the results.

3. SORTED TOMLINSON HARASHIMA PRECODING

3.1. Sorted THP for MIMO Flat Fading Channels

Subsequently, we restrict our considerations to MIMO flat fading channels, i.e., we only regard a certain subcarrier of the OFDM system. To be consistent with the definition of the OFDM system all symbols exhibit the index d to represent the d^{th} subcarrier.

As precoding strategy, we consider sorted *Tomlinson-Harashima Precoding (THP)* in each subcarrier. A block diagram of this scheme is given in Fig. 1. First the precoding order of the K users is affected by the permutation matrix $\boldsymbol{P}_d$. This precoding order can be optimized according to some optimization criterion, which will be further specified in Section 3.2. Then the signals of the users are successively precoded in the feedback-loop with the feedback matrix $\boldsymbol{B}_d$, a lower triangular matrix with unit main diagonal, and modulo reduced into the support of the signal constellation. Finally, the signals are processed via the feedforward matrix $\boldsymbol{F}_d$, a unitary matrix which ensures that the average power is equal at each transmit antenna. The modulo conversion in the feedback-loop leads to a slight increase of the transmit power. This effect is known as the precoding loss [9]. The remaining individual scaling factors, given by the diagonal matrix $\boldsymbol{\Gamma}_d = \mathrm{diag}(g_{k,d})$, $k = 1, \dots, K$, can be equalized within the receiver's *automatic gain control (agc)*.

Given a suited permutation matrix $\boldsymbol{P}_d$ the feedforward matrix $\boldsymbol{F}_d$ and the feedback matrix $\boldsymbol{B}_d$ can be calculated via QR decomposition [5, 10]

$$\boldsymbol{H}_d^{\mathsf{H}} \boldsymbol{P}_d^{\mathsf{T}} = \boldsymbol{Q}_d \cdot \boldsymbol{R}_d = \boldsymbol{F}_d \cdot \boldsymbol{B}_d^{\mathsf{H}} \boldsymbol{\Sigma}_d. \tag{4}$$

Hereby, $\boldsymbol{Q}_d$ describes a unitary and $\boldsymbol{R}_d$ an upper triangular matrix. The diagonal scaling matrix $\boldsymbol{\Sigma}_d$ ensures that the lower triangular feedback matrix $\boldsymbol{B}_d$ has unit main diagonal. The individual scaling factors at the receiver read

$$\boldsymbol{\Gamma}_d = \boldsymbol{P}_d^{\mathsf{T}} \boldsymbol{\Sigma}_d^{-1} \boldsymbol{P}_d. \tag{5}$$

3.2. Precoding Order

So far, the precoding order given by the permutation matrix $\boldsymbol{P}_d$ has not been specified. According to [6] a reasonable criterion is to maximize the performance of the worst user, i.e., the one exhibiting the minimum signal-to-noise ratio (SNR) as this one dominates the mean bit error ratio (BER) over all users. An almost optimum solution to this criterion can be found using the V-BLAST algorithm [11, 6], which finds the optimum detection order for decision-feedback equalization.

see the program

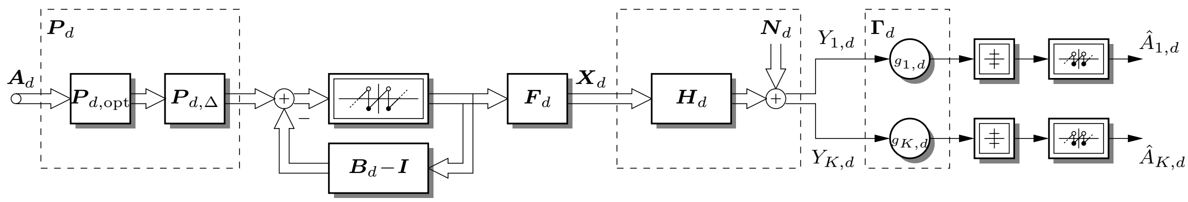


Fig. 1. Block diagram of sorted Tomlinson-Harashima precoding.

Considering the uplink-downlink duality [12], the reverse permutation order found by the V-BLAST algorithm can be applied for sorted THP [6, 13]. The resulting optimum permutation matrix is denoted as $\boldsymbol{P}_{d,\mathrm{opt}}$.

Subsequently, we always consider the application of the optimum encoding order given by $\boldsymbol{P}_{d,\mathrm{opt}}$. Starting from this solution it is possible to change the sorting order by an additional sorting through the matrix $\boldsymbol{P}_{d,\Delta}$, which can be regarded as an offset on the optimum one. The total permutation matrix $\boldsymbol{P}_d$ is then given by (see also Fig. 1)

$$\boldsymbol{P}_d = \boldsymbol{P}_{d,\Delta}\boldsymbol{P}_{d,\mathrm{opt}}. \tag{6}$$

Choosing an additional permutation matrix $\boldsymbol{P}_{d,\Delta}$ other than the identity matrix $\boldsymbol{I}$ will evidently decrease the performance in terms of the mean bit error ratio of the system. In total there exist $K!$ different additional sorting orders and hence permutation matrices $\boldsymbol{P}_{d,\Delta}$.

According to [6] the overall performance (in terms of the bit error ratio) is governed by the user encoded last. Hence, a rearrangement of the sorting with the matrix $\boldsymbol{P}_{d,\Delta}$ without changing the position of the last encoded user will hardly influence the overall performance. On the contrary, the degradation of the bit error ratio induced by $\boldsymbol{P}_{d,\Delta}$ is mainly determined by the new position of the originally (determined by $\boldsymbol{P}_{d,\mathrm{opt}}$) last encoded user. The sorting of all other users will have almost no impact on the overall performance.

Hence, the $K!$ different additional sortings given by the permutation matrices $\boldsymbol{P}_{d,\Delta}$ can be classified into K different classes of additional sortings which lead to different performance results in terms of the bit error ratio. These K classes are defined by the position within the precoding order of the originally (determined by $\boldsymbol{P}_{d,\mathrm{opt}}$) last encoded user. Within each class there exist $(K-1)!$ different sorting orders, which exhibit almost the same performance.

Fig. 2 shows all resulting mean bit error ratios obtained by all $K!$ different permutations for the special case of $K = 4$. The different classes of additional sorting orders are depicted by different colors. The green curve shows the results if the fourth encoded user (after $\boldsymbol{P}_{d,\mathrm{opt}}$) is not influenced by $\boldsymbol{P}_{d,\Delta}$. The blue, cyan, and red curves show the results if the fourth user is rearranged to the third, second, or first position, respectively. As can be seen from this plot, $K = 4$ different classes of bit error ratio results are present, whereby the differences within these classes are neglectable.

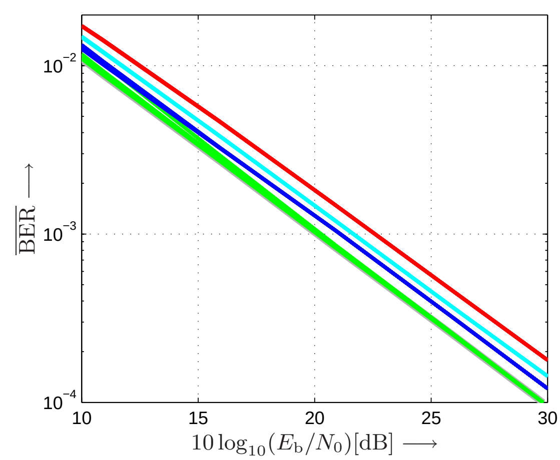


Fig. 2. Mean bit error ratios of sorted THP employing all $K!$ possible additional permutations $\boldsymbol{P}_{d,\Delta}$ considering transmission over a flat (Rayleigh) fading MIMO channel. The resulting curves are classified into four classes whereby the permutation matrix $\boldsymbol{P}_{d,\Delta}$ rearranges the fourth encoded user to the fourth (green), third (dark blue), second (light blue), or first (red) position, respectively. The best result (given with $\boldsymbol{P}_{d,\Delta} = \boldsymbol{I}$) is depicted in gray. $M = 4$-QAM, $K = 4$.

4. SELECTED SORTING FOR PAR REDUCTION

4.1. Review of (Simplified) Selected Mapping

Selected Mapping [1] is one of the most popular techniques for PAR reduction in OFDM systems. The idea behind this scheme is to generate out of the original OFDM frame several, say U_SLM, different signal representations via U_SLM different bijective mappings $\mathcal{M}^{(u)}$, $u = 1, \dots, U_\mathrm{SLM}$. Out of these signal candidates, the best one, i.e., the one exhibiting the lowest PAR is chosen for transmission. At the receiver af-

see the program

ter equalization, the original data has to be reconstructed by inverting the applied mapping. Hence, side information, in terms of an index of the applied mapping, has to be transmitted. The required redundancy has to be encoded with at least $\lceil \log_2(U_{\mathrm{SLM}}) \rceil$ bits. However, this index is extraordinary sensitive to transmission errors as the application of the wrong inverse mapping leads to the loss of the whole OFDM frame.

Possible schemes to transmit the side information have been proposed in [14, 15]. For the analysis of the bit error performance of SLM (Section 4.4) we will consider the scheme from [14]. The so-called scrambler variant of SLM distributes the side information inherently over the whole OFDM frame and does not require its explicit transmission.

For multi-antenna scenarios the SLM technique has been extended to the basic schemes *ordinary SLM (oSLM)* [2], *simplified SLM (sSLM)* [2], and *directed SLM (dSLM)* [3, 4]. Following the discussion in [7] it is not feasible to apply ordinary or directed SLM in a broadcast scenario. Due to the transmitter-sided precoding, which influences the data streams of all users, only the simplified approach can be applied here.

Hereby, sSLM is the simplest extension of SLM to MIMO systems. With sSLM the original frequency-domain MIMO OFDM frame $\boldsymbol{A}$ has to be mapped jointly onto U_{sSLM} different signal representations, whereby each row of $\boldsymbol{A}$ has to be mapped the same. Afterwards, each of the resulting signal candidates has to be precoded and transformed into time domain. Out of these the best one, i.e., the one exhibiting the lowest PAR, is then chosen for transmission.

As the individual signal candidates are assumed to be statistically independent, the ccdf of sSLM can be given with respect to the ccdf of the original signal (3) [2, 3, 4]

$$\begin{aligned} \mathrm{ccdf}_{\mathsf{sSLM}}(\mathrm{PAR}_{\mathrm{th}}) &= \left(\mathrm{ccdf}_{\mathsf{MIMO}}(\mathrm{PAR}_{\mathrm{th}})\right)^{U_{\mathrm{sSLM}}} \qquad (7)\\ &\overset{\text{Gauss}}{=} \left(1-(1-\mathrm{e}^{-\mathrm{PAR}_{\mathrm{th}}})^{DK}\right)^{U_{\mathrm{sSLM}}}. \end{aligned}$$

4.2. Selected Sorting

Another approach to generate different signal representations could be to combine the mapping with the precoding by applying different instances of sTHP in each subcarrier. Hereby, these different instances are generated by considering different permutations $\boldsymbol{P}_d$ of the users. A practical advantage of this approach is that no side information needs to be signaled to the receiver.

Subsequently, we pick a set of V different permutation matrices $\boldsymbol{P}_{d,\Delta}^{(v)}$, $v = 1, \dots, V$, out of the set of $K!$ possible ones. Starting with the optimum sorting order, the reverse of that obtained according to the V-BLAST criterion [6], we consider the next better (suboptimal, acc. Section 3.2) ones. Now, the information carrying signal $\boldsymbol{A}$ is precoded via all V different precoder instances, the resulting precoded signals are denoted as $\tilde{\boldsymbol{X}}^{(v)}$, $v = 1, \dots, V$. In oder to generate U_{SLS} different signal candidates $\boldsymbol{X}^{(u)}$, $u = 1, \dots, U_{\mathrm{SLS}}$, the respective columns (corresponding to the carriers) of $\tilde{\boldsymbol{X}}^{(v)}$ are combined in U_{SLS} different ways. Hence, every column of each of the U_{SLS} signal candidates $\boldsymbol{X}^{(u)}$ is drawn as the column from one of the V possible precoded signals. This is possible as the actual choice of the sorting order of THP at the d^{th} subcarrier influences the precoded signal only at this position.

Noteworthy, with this approach we are able to generate (much) more signal candidates than precoded candidates are present ($U_{\mathrm{SLS}} \geq V$). A principal example how the U_{SLS} signal candidates are generated is depicted in Fig. 3.

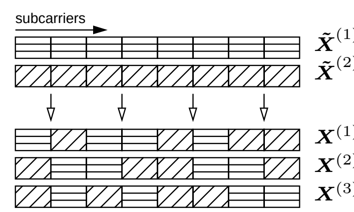


Fig. 3. Generation of $U_{\mathrm{SLS}} = 3$ candidates out of a set of $V = 2$ alternative precoded sequences.

Compared to sSLM, assuming perfect transmission of the side information, this scheme will lead to a loss in bit error performance as suboptimal sorting orders will be used within the generation the signal candidates. However, with SLS much less computational complexity is needed as the precoding has to be performed only V times to generate the U_{SLS} signal candidates.

To further reduce the computational complexity the SLS technique could only be applied on a subset of $D \cdot \rho$, with $0 < \rho \leq 1$, (randomly chosen) influenced subcarriers. All other subcarriers remain unaffected and the optimum sorting order is applied.

The PAR reduction performance of SLS will be the same as that of sSLM (7) if all U_{SLS} resulting signal candidates are statistically independent. However, this is only guaranteed if the number V of alternative precoders is sufficiently large and if the number of influenced subcarriers is adequate, i.e., the factor ρ is near one. Hence, the ccdf of SLS reads

$$\mathrm{ccdf}_{\mathsf{SLS}}(\mathrm{PAR}_{\mathrm{th}}) \xrightarrow{\text{large } V,\, \rho} \left(\mathrm{ccdf}_{\mathsf{MIMO}}(\mathrm{PAR}_{\mathrm{th}})\right)^{U_{\mathrm{SLS}}}. \qquad (8)$$

4.3. Analysis of Computational Complexity

For a reasonable comparison of SLS with sSLM the computational complexity of both schemes has to be taken into

see the program

account. Subsequently, we refer to complex operations as complexity measure and regard multiplications and divisions equally.

The decomposition of the channel matrix in each subcarrier into feedforward, feedback, and the optimum sorting order has to be accomplished with the V-BLAST algorithm [11]. In [16, 17] low complex implementations have been proposed which reduce the complexity from $\mathcal{O}(K^4)$ to $\mathcal{O}(K^3)$. However, as this decomposition appears in the overall complexities of SLS and sSLM equally, it has not to be specified in more details for a comparison.

The calculation of the feedforward and feedback matrices for the $V-1$ alternative sorting orders is usually implemented as a QR decomposition [10]. Using the result[2] from [18]

$$c_{\mathrm{QR}} = D \cdot \left(2K^3 - \frac{K^2}{2} - \frac{K}{2}\right) \tag{9}$$

complex operations are required here.

In order to calculate the alternative precoding matrices it is not necessary to accomplish $V-1$ times the QR decomposition per influenced subcarrier. Furthermore, it is possible to perform this calculation within the respective implementation of the V-BLAST algorithm by considering in each iteration not only the user exhibiting the lowest SNR but also the one exhibiting the second lowest SNR. Subsequently, we consider the exhaustive complexity of $V-1$ QR decompositions per influenced subcarrier which leads to an upper-bound on the complexity of SLS.

Each precoding of the transmit signal requires

$$c_{\mathrm{prec}} = D \cdot \left(\frac{3}{2}K^2 + \frac{K}{2}\right) \tag{10}$$

complex operations; the transformation into time domain (implemented as fast Fourier transform [8]) and the calculation of the decision metric (PAR) require

$$c_{\mathrm{FFT}} = K \cdot \frac{D}{2}\log_2(D); \qquad c_{\mathrm{met}} = K \cdot D \tag{11}$$

operations.

In the following, we assume that the channel remains constant for the duration of N_{B} OFDM symbols. Hence, for this block of OFDM symbols the calculation of the precoding matrices has to be performed only once, whereas the computation of the precoded signal, the FFT, and the decision metric have to be accomplished for each of the N_{B} OFDM symbols. Table 1 shows the factors with which these individual complexities appear in the overall complexities c_{sSLM} and c_{SLS} of sSLM and SLS.

[2] The result from [18] only considers the pure QR decomposition. In addition to that we need to accomplish the normalization of the feedback matrix to unit main diagonal.

Table 1. Contribution of the individual complexities to the entire complexities of sSLM and SLS.

complexity	sSLM	SLS
$c_{\text{V-BLAST}}$	1	1
c_{QR}	—	$(V-1)\rho$
c_{prec}	$N_{\mathrm{B}} \cdot U_{\mathrm{sSLM}}$	$N_{\mathrm{B}} \cdot [1+(V-1)\rho]$
c_{FFT}	$N_{\mathrm{B}} \cdot U_{\mathrm{sSLM}}$	$N_{\mathrm{B}} \cdot U_{\mathrm{SLS}}$
c_{met}	$N_{\mathrm{B}} \cdot U_{\mathrm{sSLM}}$	$N_{\mathrm{B}} \cdot U_{\mathrm{SLS}}$

Assuming both schemes evaluate the same number $U = U_{\mathrm{SLS}} = U_{\mathrm{sSLM}}$ of alternative signal representations the difference of both complexities is given by

$$\begin{aligned} \Delta c &= c_{\mathrm{sSLM}} - c_{\mathrm{SLS}} \\ &= (1-V)\rho \cdot c_{\mathrm{QR}} + [U - 1 + (1-V)\rho] N_{\mathrm{B}} \cdot c_{\mathrm{prec}}. \end{aligned} \tag{12}$$

Hence, as $c_{\mathrm{QR}} > c_{\mathrm{prec}}$ SLS gains especially for reasonable large block lengths N_{B}.

For a fair comparison of sSLM with SLS both schemes should have the same complexity. Given the parameters V and U_{SLS} for SLS then sSLM assessing

$$U_{\mathrm{sSLM}} = \left\lceil \frac{(V-1)\rho/N_{\mathrm{B}} \cdot c_{\mathrm{QR}} + V c_{\mathrm{prec}} + U_{\mathrm{SLS}}(c_{\mathrm{FFT}} + c_{\mathrm{met}})}{c_{\mathrm{prec}} + c_{\mathrm{FFT}} + c_{\mathrm{met}}} \right\rceil \tag{13}$$

signal candidates will exhibit approximately the same computational complexity. Hereby, the number U_{sSLM} of assessed candidates for sSLM is rounded to the next greater integer, whereby sSLM will exhibit a slightly larger complexity.

4.4. Numerical Results

The subsequent numerical simulations consider a $l_{\mathrm{H}} = 5$ tap (equal gain) MIMO channel with $N_{\mathrm{C}} = 4$ transmit and $K = 4$ receive antennas. The number of subcarriers in the OFDM system is chosen to $D = 512$ whereby all subcarriers are active. As modulation scheme $M = 4$-QAM is considered.

In the top plot of Fig. 4 numerical results of the ccdf of SLS for a various number of assessed signal candidates ($U_{\mathrm{SLS}} = 4$, 8, 16) and number of available alternative sorting orders ($V =$ 2, 4, 8) are shown. Hereby all subcarriers are influenced by SLS ($\rho = 1$). Compared to the PAR distribution of the original signal (red curve) SLS has the ability to reduce the peak-power significantly. Moreover, the numerical results are compared with the analytical ones (8) (gray curves). For $V = 2$ explicit differences are visible, which shows that in this case the

see the program

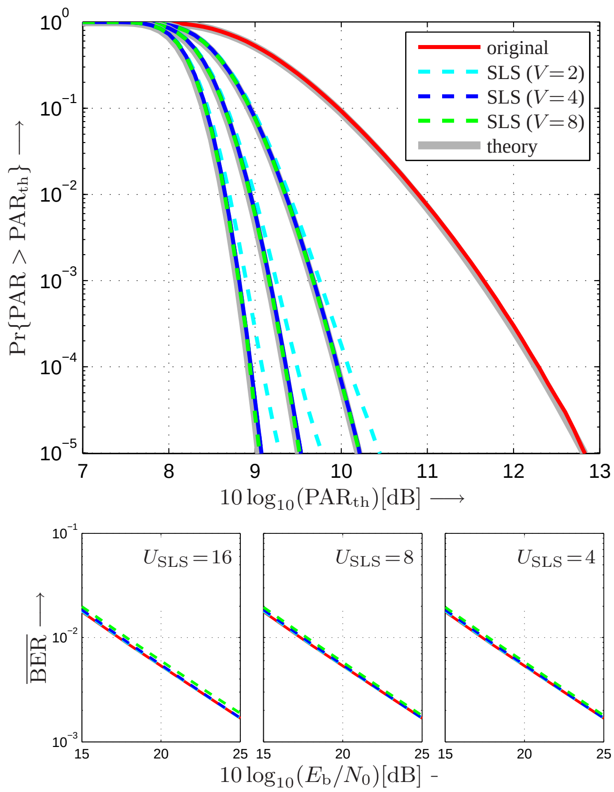


Fig. 4. Top: ccdf of PAR of the original signal (red) and SLS with $V = 2$ (blue), $V = 4$ (green), $V = 8$ (magenta) alternative precoding orders, each assessing (from left to right) $U_{\mathrm{SLS}} = 16, 8, 4$ signal candidates. Bottom: corresponding mean bit error ratios of SLS. As reference serves the result of the original signal (red curve) which represents pure sTHP. $M = 4$-QAM, $K = 4$, $l_{\mathrm{H}} = 5$, $D = 512$, $\rho = 1$.

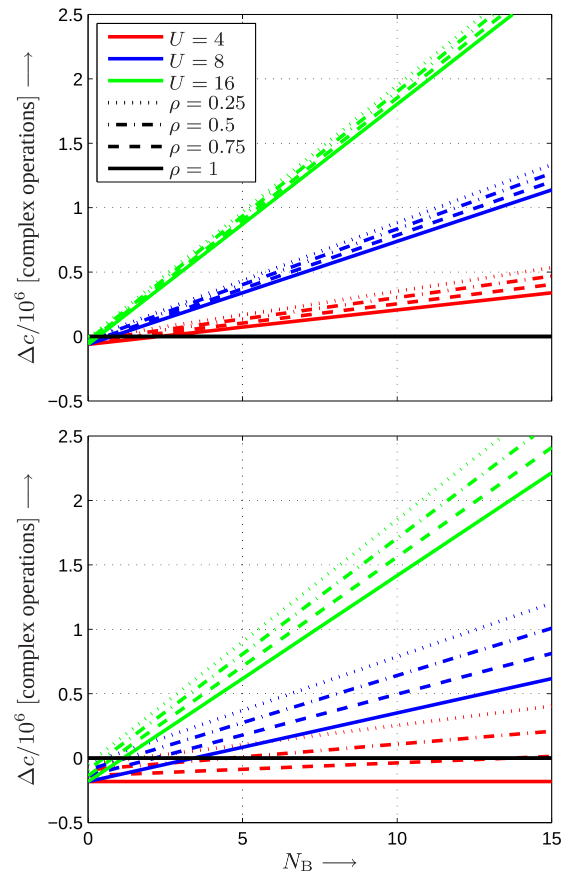


Fig. 6. Difference of complexities of sSLM and SLS (with $V = 2$ (top) and $V = 4$ (bottom)) depending on the block length N_{B}. Both schemes assess an equal number of signal candidates: $U = U_{\mathrm{sSLM}} = U_{\mathrm{SLS}}$. $K = 4$, $D = 512$.

number of degrees of freedom is not large enough to generate statistically independent signal candidates. If V is chosen to $V \geq 4$ the simulation results of SLS fit to the theory very well. The bottom plots of Fig. 4 show the bit error performance for SLS for $U_{\mathrm{SLS}} = 16, 8, 4$ (from left to right). The performance loss in terms of bit error ratios is rather neglectable. Even for $V = 8$ where alternative permutation orders have to be used which lead to significant degradation of the BER (see Section 3.2), the overall loss in BER is very small.

In order to reduce the computational complexity it is possible to apply SLS only on a subset of $D \cdot \rho$ subcarriers. As discussed in Section 4.2, choosing $\rho < 1$ will restrict the possibilities to generate statistical independent signal candidates, whereby the analytical result from (8) is not strict any more. Fig. 5 shows the influence of different values of $\rho = 1, 0.75, 0.5$, and 0.25 on the PAR reduction performance. Hereby it is obvious that only for large values of ρ (here $\rho \geq 0.75$) the performance is comparable with the maximum number of $\rho = 1$.

For the comparison of SLS with sSLM the difference Δc (as given in (12)) of computational complexity over the block length N_{B} is depicted in Fig. 6. Already for small N_{B} (in most scenarios values of $N_{\mathrm{B}} \approx 5$ are sufficient) a gain of SLS compared to sSLM in terms of computational effort can be recognized as the impact of the higher complexity, which occurs through the computation of the precoding matrices with SLS, becomes less important on the overall difference.

In addition to that, Fig. 6 shows the difference Δc for var-

see the program

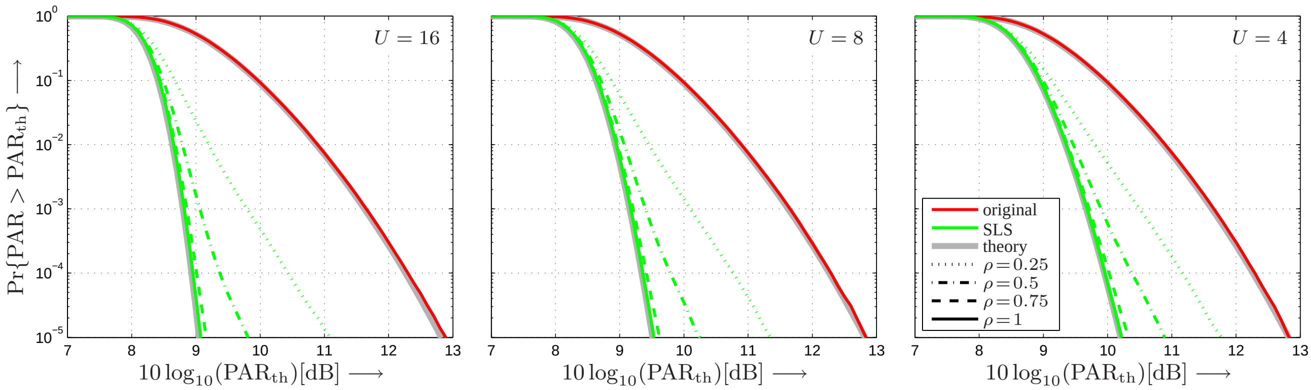


Fig. 5. Ccdf of PAR of the original signal (red) and SLS (green), with $V = 4$ and $U_{\mathrm{SLS}} = 16$ (left), 8 (middle), and 4 (right) for various factors $\rho = 0.25, 0.5, 0.75$, and 1. $M = 4$-QAM, $K = 4$, $D = 512$.

ious ratios ρ of influenced subcarriers. These results show that these gains in computational complexity are rather neglectable. Against the background of the PAR reduction results from Fig. 5 we recommend to choose $\rho = 1$ which is assumed for the analysis of Fig. 4 and 7.

According to (12) the difference $U - V$ determines the slope of Δc. Hence, SLS provides especially gains in terms of complexity if the difference $U - V$ is large. If U and V are chosen equal then SLS will not provide any gains compared to sSLM as in both cases the same number of precoding procedures are processed and SLS suffers from the more complex computation of the precoding matrices. This effect can also be observed form Fig. 6.

The top row of Fig. 7 compares the PAR reduction performance of SLS with sSLM on the bases that both schemes exhibit the same computational complexity. Hereby, the number of assessed candidates for sSLM is chosen according to (13). Providing a good trade-off between loss in bit error performance and computational complexity, the number of alternative precoders is chosen to $V = 2$ (left column) and $V = 4$ (right column). For the most relevant range of values of V and U_{SLS}, SLS offers significant gains in terms of PAR reduction performance.

The bottom row of Fig. 7 shows the corresponding bit error ratios of SLS and sSLM. To achieve a fair comparison, we consider the side information, required with sSLM, to be transmitted through the scrambler variant of SLM [14]. However, the descrambler at the receiver introduces some error propagation in this case. Due to this error propagation SLS outperforms sSLM in terms of error performance as the degradation of the error rate according to the suboptimal precoders is neglectable.

5. CONCLUSIONS

In this paper we have analyzed Selected Sorting for PAR reduction in broadcast scenarios. This scheme combines the channel precoding, which is essential in broadcast scenarios, with peak-power reduction. Hereby, SLS offers significant gains in PAR reduction, whereas the degradation of the error performance is neglectable.

The peak-power reduction performance of SLS is comparable with the one of simplified SLM. Comparing these two schemes shows that SLS exhibits much less complexity than sSLM. On the contrary, choosing the parameters of both PAR reduction schemes such that they exhibit the same complexity SLS will outperform sSLM. These significant gains of SLS are possible as no side information is necessary to be signaled to the receiver.

6. REFERENCES

[1] R. Bäuml, R.F.H. Fischer, and J.B. Huber, "Reducing the Peak-to-Average Power Ratio of Multicarrier Modulation by Selected Mapping," *IEE Electronics Letters*, pp. 2056–2057, Nov. 1996.

[2] M.-S. Baek, M.-J. Kim, Y.-H. You, and H.-K. Song, "Semi-Blind Channel Estimation and PAR Reduction for MIMO-OFDM System with Multiple Antennas," *IEEE Transactions on Broadcasting*, pp. 414–424, Dec. 2004.

[3] R.F.H. Fischer and M. Hoch, "Directed Selected Mapping for Peak-to-Average Power Ratio Reduction in MIMO OFDM," *IEE Electronics Letters*, pp. 1289–1290, Oct. 2006.

[4] R.F.H. Fischer and M. Hoch, "Peak-to-Average Power Ratio Reduction in MIMO OFDM," in *Proceedings of IEEE International Conference on Communications (ICC)*, Glasgow, Scotland, June 2007.

[5] R.F.H. Fischer, *Precoding and Signal Shaping for Digital Transmission*, John Wiley & Sons, New York, 2002.

see the program

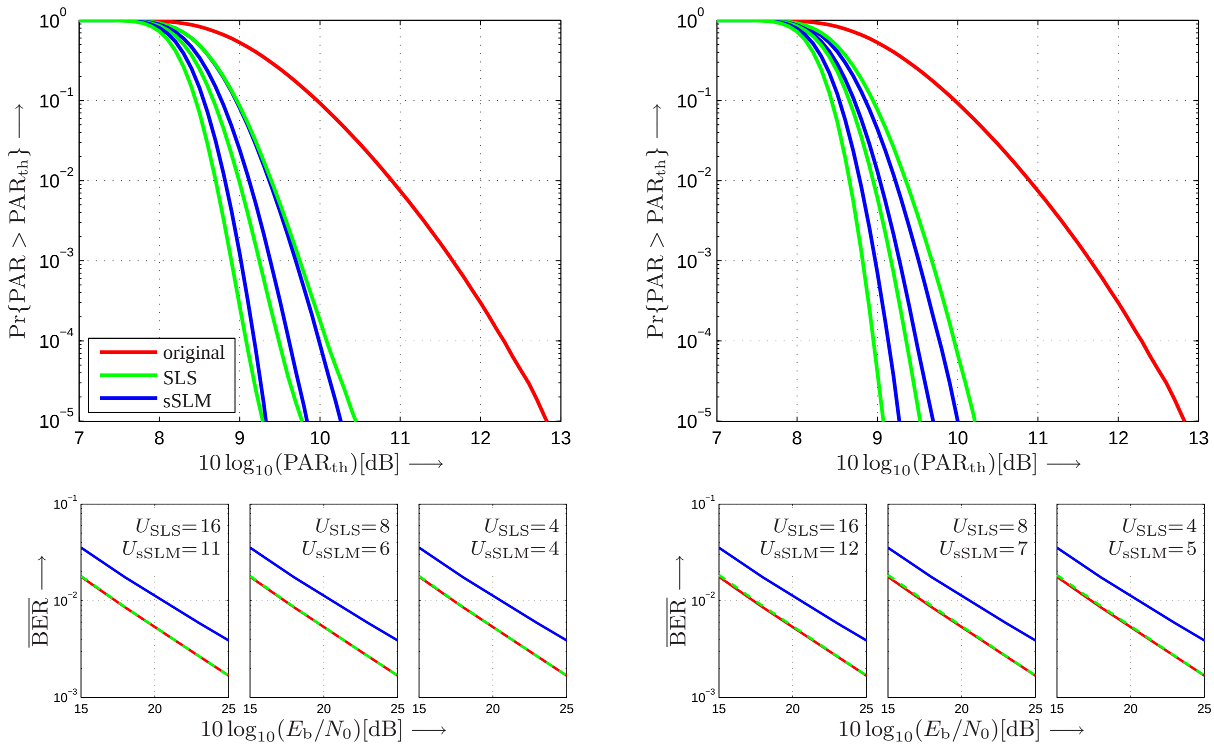


Fig. 7. Comparison ccdf of PAR (top row) and bit error ratio (bottom row) of SLS and sSLM; the number of alternative precoders for SLS is $V = 2$ (left column) or $V = 4$ (right column). The number of assessed signal candidates for SLS is (from right to left) $U_{\mathrm{SLS}} = 4$, 8, and 16; to exhibit almost same computational complexity the number of assessed candidates of sSLM is chosen to $U_{\mathrm{sSLM}} = 4$, 6, and 11 (left column) or $U_{\mathrm{sSLM}} = 5$, 7, and 12 (right column). For the bit error ratio analysis the side information, necessary with sSLM, is transmitted according the scheme from [14], whereby a scrambler with $N_{\mathrm{states}} = 16$ states has been used. $M = 4$-QAM, $K = 4$, $D = 512$, $\rho = 1$.

[6] C. Windpassinger, *Detection and Precoding for Multiple Input Multiple Output Channels*, Ph.D. thesis, Universität Erlangen-Nürnberg, 2004.

[7] C. Siegl and R.F.H. Fischer, "Peak-to-Average Power Ratio Reduction in Multi-User OFDM," in *Proceedings IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[8] A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Upper Saddle River, 1999.

[9] C. Windpassinger, R.F.H. Fischer, T. Vencel, and J.B. Huber, "Precoding in Multi-Antenna and Multi-User Communications," *IEEE Transactions on Wireless Communications*, pp. 1305–1316, July 2004.

[10] G.H. Golub and C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[11] G.J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multiple Antennas," *Bell Laboratories Technical Journal*, pp. 41–59, 1996.

[12] P. Viswanath and D.N.C. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, Aug. 2003.

[13] C. Siegl and R.F.H. Fischer, "Uplink-Downlink Duality with Regard to Constraints Imposed in Practice," *International Journal of Electronics and Communications (AEÜ)*, Sept. 2009, DOI 10.1016/j.aeue.2008.04.011.

[14] M. Breiling, S. Müller-Weinfurtner, and J.B. Huber, "SLM Peak-Power Reduction without Explicit Side Information," *IEEE Communications Letters*, pp. 239–241, June 2001.

[15] B.K. Khoo, S.Y. Le Goff, C.C. Tsimenidis, and B.S. Sharif, "OFDM PAPR Reduction Using Selected Mapping Without Side Information," in *Proceedings of IEEE International Conference on Communications (ICC)*, Glasgow, Scotland, June 2007.

[16] D. Wübben, J. Rinas, R. Böhnke, V. Kühn, and K.D. Kammeyer, "Efficient algorithm for detecting layered space-time codes," in *Proceedings of 4th International ITG Conference on Source and Channel Coding (SCC)*, Berlin, Germany, Jan. 2002.

[17] J. Benesty, Y. Huang, and J. Chen, "A Fast Recursive Algorithm for Optimum Sequential Signal Detection in a BLAST System," *IEEE Transactions on Signalprocessing*, pp. 1722–1730, July 2003.

[18] D. Cescato, M. Borgmann, H. Bölcskei, J. Hansen, and A. Burg, "Interpolation-Based QR Decomposition in MIMO-OFDM Systems," in *Proceedings of IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, June 2005.

see the program

SUM RATE IMPROVEMENT BY SPATIAL PRECODING IN A MULTI-USER MIMO SC-FDMA UPLINK

Hanguang Wu[1], Thomas Haustein[1], Eduard Axel Jorswieck[2], and Peter Adam Hoeher[3]

[1]Nokia Siemens Networks, St.-Martin-Str. 76, 81617 Munich, Germany
[2]Communications Laboratory, Dresden University of Technology, D-01062 Dresden, Germany
[3]Faculty of Engineering, University of Kiel, Kaiserstr. 2, 24143 Kiel, Germany

ABSTRACT

Conventional virtual MIMO SC-FDMA systems as proposed for the LTE uplink using single-antenna user equipments may suffer from channel rank deficiency and/or channel deep fade at some sub-carriers, which can lead to significant system sum rate degradation. In this work, we propose to use multiple antennas and spatial precoding to mitigate the problem. An efficient algorithm is developed to optimize the spatial precoder which achieves significant performance gain in terms of spectral efficiency over conventional virtual MIMO SC-FDMA systems. Furthermore, a simple spatial scheduler is proposed to effectively control the individual rates of the users.

1. INTRODUCTION

Single carrier FDMA (SC-FDMA) is the standard uplink transmission scheme in 3GPP Long Term Evolution (LTE). It can be viewed as a DFT precoded OFDMA system where the user's transmitted signal will be first transformed to frequency domain via an N point discrete Fourier transform (DFT) and the DFT output will be assigned to a number of selected sub-carriers for transmission. The main advantage of SC-FDMA is that its transmit waveform has lower peak to average power ratio (PAPR) than that of OFDMA for low order modulation schemes like BPSK and QPSK, which benefits the mobile terminal in terms of power efficiency [1]. However, this requires the selected sub-carriers to be consecutive or equidistantly distributed in the entire bandwidth [2]. These sub-carrier selection methods are usually referred to as localized FDMA (LFDMA) [3] and interleaved FDMA (IFDMA) [4], respectively. It has been shown that IFDMA has the lowest PAPR [3] which also provides large frequency diversity, while LFDMA can benefit more from frequency dependent scheduling to improve the throughput [5]. Another variant of DFT precoded OFDMA using regularly interleaved blocks of sub-carriers is denoted block-IFDMA (B-IFDMA), which provides robustness to frequency offsets at the expense of increased PAPR compared to IFDMA [6].

Similar to OFDMA systems, SC-FDMA can easily be combined with SDMA (spatial division multiple access), also called the virtual MIMO technique, by applying spatial processing on a sub-carrier basis, in the frequency domain after DFT precoding to improve the system data rate. This technique enables independent data streams from different user equipments (UEs) to use the same sub-carriers to communicate with the base station (BS) simultaneously. At the receiver, simple MIMO equalization, e.g., zero-forcing (ZF) equalization, can be carried out in the frequency domain to separate the data streams from different UEs, which are then transformed back to the time domain via IDFT for data decoding and detection. However, in a system with UEs each having a single antenna, the compound channel between the transmit antennas of the UEs and the receive antennas at the BS will be close to rank deficiency at some sub-carriers if the UEs have similar spatial signatures at those sub-carriers. Moreover, even the UEs can be well separated in spatial domain but they may experience deep fades at some sub-carrier due to the frequency selectivity of the mobile channel. Unlike OFDMA systems, those sub-carriers can not be excluded from transmission due to the sub-carrier mapping constraints. Consequently, the compulsory usage of those sub-carriers will lead to significant system rate loss [7], especially when the UEs occupy large bandwidth, which is a typical case in future broadband wireless systems.

In this work, we propose to employ multiple antennas and spatial precoding for each UE to mitigate the rank deficiency and channel deep fade problems. We develop an algorithm to optimize the spatial precoder design to maximize the system sum rate. Furthermore, a simple spatial scheduler is proposed to handle the user rate fairness using the spatial domain. The rest of the paper is organized as follows. Section 2 describes the system model under consideration. Section 3 discusses the proposed spatial precoder optimization algorithm and the spatial scheduling algorithm. Simulation results are presented in Section 4. Conclusions are drawn in Section 5.

2. SYSTEM MODEL

We consider an SC-FDMA uplink with two UEs, each having 2 antennas and the BS also equipped with 2 antennas.

see the program

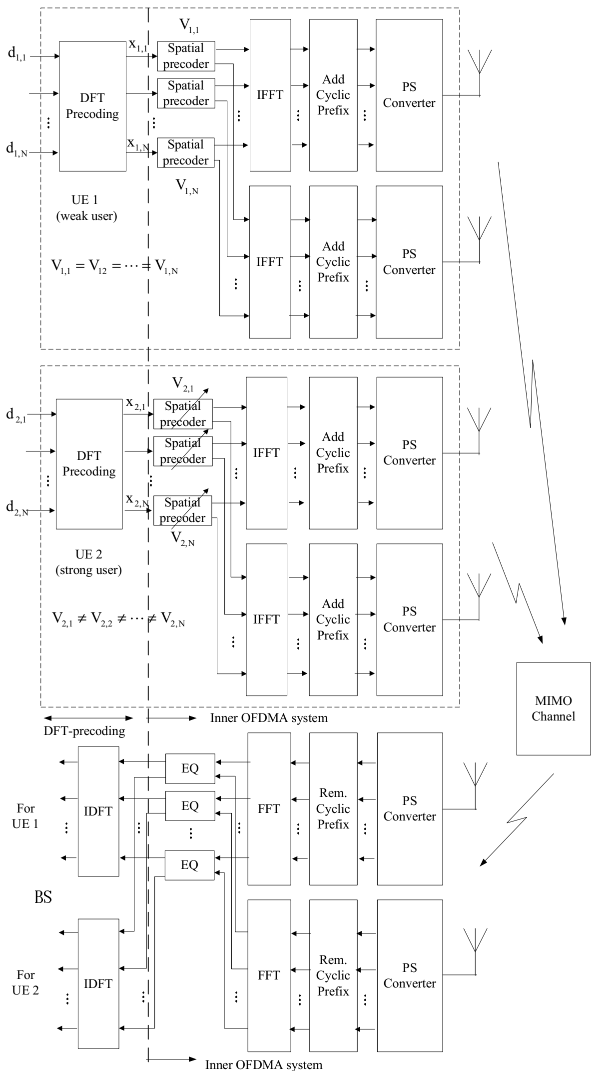


Fig. 1. Block diagram of the SC-FDMA MIMO system with two UEs under consideration

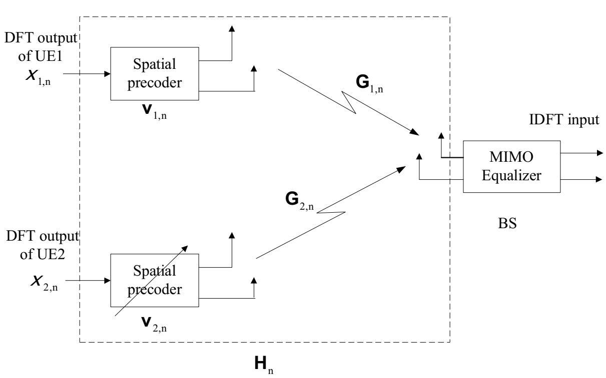


Fig. 2. Spatial precoder for MIMO uplink with multiple transmit antennas at each UE and multiple antennas at the BS

The block diagram of the system setup is shown in Fig. 1. The transmitted data streams $d_{u,1}, \cdots, d_{u,N}$ of UE u is transformed to the frequency domain via an N point DFT and the DFT output $x_{u,1}, \cdots, x_{u,N}$ is linear precoded by $\mathbf{v}_{u,1}, \cdots, \mathbf{v}_{u,N}$, respectively. We assume equal power of the transmitted signal, i.e., $E\{d_{u,n}d^*_{u,n}\} = P_u/N$, where P_u denotes the power constraint of UE u. Note that DFT precoding is unitary precoding which preserves the signal energy and it will not change the power distribution of the incoming signal, hence

$$E\{x_{u,n}x^*_{u,n}\} = P_u/N \tag{1}$$

The Nx2 outputs of the linear precoder represents 2 spatial data streams, each of which is processed at one antenna by an IFFT and cyclic prefix is inserted (CP-OFDM). Then the resulting signal is subsequently parallel to serial converted for transmission. The transmitted signals of both UEs undergo multiple path propagation and are received by the receiver at the BS. The receiver converts the incoming data streams from serial to parallel, removes the cyclic prefix, and processes them using an IFFT. Next, ZF-MIMO equalization is performed and the equalized signal is converted back to the time domain via an N point IDFT for detection. In Fig. 1 the block diagram without DFT precoding at the transmitter and IDFT at the receiver is referred to as the inner MIMO OFDMA system.

Our system model only considers single stream transmission on each sub-carrier for each UE. In principle, it is possible for a UE to transmit multiple data streams by applying spatial multiplexing (SM) as discussed in [8] either with or without spatial precoding. However, on the one hand, SM for a UE with spatial precoding will generally increases PAPR with respect to the single antenna transmission [8]. On the other hand, the performance of SM for a UE without spatial precoding will be degraded by spatial correlation between the antennas, which is mainly due to the limited antenna separation in the UE. Therefore, it is better to multiplex different data streams from different UEs than from different antennas of the same UE, since the compound virtual MIMO channel benefits from appropriate user grouping and may achieve good rank, even if the 2 antennas of each UE have high correlation. Moreover, the SC-FDMA transmission structure for each antenna can be preserved in our model hence maintains low PAPR if a common, say frequency independent spatial precoder is used for all the sub-carriers at a UE. This is especially critical for the cell edge users to be able to bridge the long distance. Furthermore, single stream transmission is preferred in terms of implementation complexity since coding and transmission can be simply done as in the single antenna system [9] and hence multiple antennas can be easily integrated into the conventional single antenna system without having 2 baseband chains.

On each sub-carrier, the relationship between the DFT output at the transmitter and the IDFT input at the receiver can be illustrated as in Fig. 2. Let $\mathbf{G}_{u,n}$ denote the chan-

see the program

nel matrix between the transmit antennas of the UE u on sub-carrier n and the receive antennas at the BS. The compound channel of UE u seen by the BS can be obtained by $\mathbf{h}_{u,n} = \mathbf{G}_{u,n}\mathbf{v}_{u,n}$. The multi-user MIMO channel matrix on sub-carrier n is written as

$$\mathbf{H}_n = [\mathbf{h}_{1,n} \quad \mathbf{h}_{2,n}] = [\mathbf{G}_{u,n}\mathbf{v}_{u,n} \quad \mathbf{G}_{u,n}\mathbf{v}_{u,n}]. \qquad (2)$$

Signals transmitted on sub-carrier n from all the UEs are collected in a vector and denoted by $\mathbf{x}_n = [x_{1,n} \quad x_{2,n}]^T$. The received signal on sub-carrier n at the BS is then given by $\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{n}_n$, where $\mathbf{n}_n$ is the white Gaussian noise with variance $E\{\mathbf{n}_n\mathbf{n}_n{}^{\mathbf{H}}\} = \sigma^2\mathbf{I}$. Denote the post-detection SNR of UE u on sub-carrier n in the inner OFDMA system by $\gamma_{u,n}$, which is

$$\gamma_{u,n} = \frac{E\{|x_{u,n}|^2\}}{\sigma^2\left[(\mathbf{H}_\mathrm{n}^{\mathbf{H}}\mathbf{H}_\mathrm{n})^{-1}\right]_{u,u}} \qquad (3)$$

the post-detection SNR for the nth the component at the IDFT outputs for UE u is related with $\gamma_{u,n}$ by [7]:

$$\hat{\gamma}_{u,n} = \frac{N}{\sum_{n=1}^{N}(1/\gamma_{u,n})} \qquad (4)$$

which is the harmonic mean of $\gamma_{u,n}$ and it is the same for all the components. Using Shannon's formula the achievable spectral efficiency of the sub-channel between each input and output in SC-FDMA system for UE u is then given by $\log_2(1+\hat{\gamma}_{u,n})$ and the system sum rate of the MIMO SC-FDMA system is the rate sum of all the sub-carriers of all the UEs [7], i.e.,

$$\begin{aligned} R &= \sum_{u=1}^{2}\sum_{n=1}^{N}\log_2(1+\hat{\gamma}_{u,n}) \\ &= \sum_{u=1}^{2}\sum_{n=1}^{N}\log_2\left(1+\frac{N}{\sum_{n=1}^{N}(1/\gamma_{u,n})}\right) \qquad (5) \end{aligned}$$

According to (1), (2), (3) and (5), our objective to maximize the system sum rate R can be formulated as follows:

$$\max_{\mathbf{v_{1,1}},\cdots,\mathbf{v_{2,N}}} \sum_{u=1}^{2} N\log_2\left(1+\frac{P_u}{\sigma^2\sum_{n=1}^{N}A_{u,u}}\right), \text{with}$$
$$A = \begin{bmatrix} \mathbf{v_{1,n}^H G_{1,n}^H G_{1,n} v_{1,n}} & \mathbf{v_{1,n}^H G_{1,n}^H G_{2,n} v_{2,n}} \\ \mathbf{v_{2,n}^H G_{2,n}^H G_{1,n} v_{1,n}} & \mathbf{v_{2,n}^H G_{2,n}^H G_{2,n} v_{2,n}} \end{bmatrix}^{-1}$$
$$\text{subject to} \quad ||\mathbf{v}_{u,n}||_2^2 = 1, \quad u = 1,2; n = 1,\cdots,N, \qquad (6)$$

where optimization is performed over all possible precoding vectors subject to the constraints that the precoder is normalized as (6) according to the transmit power constraint. Note that our system model also includes the special case that only one antenna is available at each UE (conventional virtual MIMO), by setting $\mathbf{v_{u,n}}$ to $[1 \quad 0]^T$ or $[0 \quad 1]^T$ for $u = 1,2; \quad n = 1,\cdots,N$ depending on which antenna is used by the UEs.

3. SPATIAL PRECODER OPTIMIZATION

A direct optimization of the objective function in (6) seems very difficult and therefore we look for an approximative solution. According to (5) a higher $\gamma_{u,n}$ for both UEs on sub-carrier n in the inner OFDMA system leads to a higher R, therefore, to maximize R, it is beneficial to maximize $\gamma_{u,n}$, or equivalently the data rate for both UEs in the inner OFDMA system and at the same time take the objective function (harmonic mean of $\gamma_{u,n}$'s) into account.

3.1. Beamforming Vector

If only a single UE, e.g., UE 1 is present in the system according to Fig. 2, it is known that the dominant eigenbeamforming transmission (DET) strategy maximizes the single user rate by applying the left dominant eigenvector (EV) at the transmitter and the right dominant eigenvector at the receiver. In the case that both UEs are present, if both UEs use the dominant eigenbeamfoming strategy for transmission, maximum power of both UEs are coupled into the channel but the UEs' signal will generally interfere with each user unless their compound channels happen to be orthogonal to each other, i.e., $<\mathbf{h}_{1,\mathrm{n}}, \mathbf{h}_{2,\mathrm{n}}>= 0$ and for this special case ZF-MIMO equalizer reduces to a match filter which maximizes the output SNR of both data streams [10] and thus also maximizes the achievable system sum rate. On the other hand, the transmitted signal from both UEs can always be made interference free to each other if one UE, i.e., UE 2, applies a precoding vector in a way such that its compound channel $\mathbf{h}_{2,n}$ is orthogonal to that of UE 1. For convenience, this precoder is referred to as the orthogonal precoder in the sequel. However, due to the limited degrees of freedom of the linear precoder, after precoding the compound channel orthogonal to UE 1 may exhibit bad channel conditions and cause UE 2 to suffer from low data rates. In this work, our proposal is to find an appropriate trade off between completely eliminating the interference (irrespective how much energy is lost for UE 2) and preserve as much energy as possible for both UEs (at the expense of possibly suffering from high interference between the inter data streams interference).

The principle of our proposal is as follows: First, we choose a precoder for a single UE, i.e., UE 1, to optimize its achievable rate in the inner OFDMA system on one or all the sub-carriers depending on the requirements. Then candidate precoders are generated on each sub-carrier for UE 2 with which its resulting transmission beam on that sub-carrier varies from its eigenbeam to a beam orthogonal to that of the first UE. Finally, a selection procedure is carried out to select the best precoder on each sub-carrier for UE 2 from the candidate precoders to optimize the objective function. In this work, we use the dominant eigenvector (EV) of the average correlation matrix $\frac{1}{N}\sum_{n=1}^{N}(\mathbf{G_{1,n}^H G_{1,n}})$ as the precoding vector for all sub-carriers or RBs for UE 1, in order not to destroy the single carrier structure hence maintain the low PAPR property. The candidate precoders for UE 2 are generated by a linear combination of the right dominant eigenvector of its channel matrix on sub-carrier n, termed $\mathbf{v}_{2,n}^{EV}$, and the orthogonal precoder $\mathbf{v}_{2,n}^{\perp}$, i.e.,

$$\mathbf{v}_{2,n}^{EV,\perp} = \frac{\alpha \mathbf{v}_{2,n}^{EV} + (1-\alpha)\mathbf{v}_{2,n}^{\perp}}{||\alpha \mathbf{v}_{2,n}^{EV} + (1-\alpha)\mathbf{v}_{2,n}^{\perp}||_2} \quad (7)$$

where $0 \leq \alpha \leq 1$ is the weighting factor to be optimized to maximize (5). We call the resulting optimum $\mathbf{v}_{2,n}^{EV,\perp}$ as the optimum complementary precoder (OCP) in the sequel. In (7), the denominator is for normalization purpose and for the special case of $\alpha = 0$ and $\alpha = 1$, the candidate vector for UE 2 corresponds to the orthogonal precoder and its dominant EV, respectively. Note that seeking an optimal α for (5) is still prohibitively difficult because α's have to be considered jointly for all the N used sub-carriers and hence the computational complexity increases exponentially with N. In order to further reduce the complexity of the selection procedure, optimization can be performed on an arbitrary sub-carrier first to obtain the best precoder for that sub-carrier and then it is considered fixed for the optimization in the next sub-carrier. As a result, the computational complexity is linear in the number of sub-carriers. A description of the algorithm with two UEs can be found in Algorithm 1. The principle of Algorithm 1 can be straightforward extended to the scenario with $U > 2$ UEs by constructing candidate precoders $\mathbf{v}_{u,n}^{EV,\perp}$ for UE u ($2 \leq u \leq U$) using the linear combination of its right dominant eigenvector $\mathbf{v}_{u,n}^{EV}$ of the channel matrix on sub-carrier n and its orthogonal precoder $\mathbf{v}_{u,n}^{\perp}$, i.e.,

$$\mathbf{v}_{u,n}^{EV,\perp} = \frac{\alpha_u \mathbf{v}_{u,n}^{EV} + (1-\alpha_u)\mathbf{v}_{u,n}^{\perp}}{||\alpha_u \mathbf{v}_{u,n}^{EV} + (1-\alpha_u)\mathbf{v}_{u,n}^{\perp}||_2} \quad (8)$$

where $\mathbf{v}_{u,n}^{EV}$ can be calculated by performing singular value decomposition (SVD) for the associated channel matrix and $\mathbf{v}_{u,n}^{\perp}$ can be obtained by using orthogonalization algorithms, e.g., the Gram-Schmidt process [11]. Clearly in (8) each UE is associated with an individual weighting factor $0 \leq \alpha_u \leq 1$ ($2 \leq u \leq U$) which should be jointly optimized.

In the above algorithm, UE 1 always utilizes its dominant eigenbeam direction and then UE 2 has to transmit in a direction such that the system sum rate is maximized. If the transmission directions (spatial signatures) are viewed as spatial resources, then UE 1 actually always occupies the best spatial resources but UE 2 does not. This fixed optimization order may cause the individual rate of the UEs to differ a lot from each other. Another factor which may affect the individual rate of the UEs is that the UE which uses OCP may produce strong interference to the UE which uses the dominant EV precoder. This will not cause any problem if both UEs have similar channel conditions. However, if their channel qualities are largely unbalanced, even the weaker UE always transmits in its dominant eigenbeam direction, a small amount of interference from the much stronger UE will have a strong impact on the rate of the weaker UE. In this situation, it is desirable to let the stronger UE transmit in the direction orthogonal to that of the weaker UE. This leads to our following simple scheduling algorithm to mitigate the aforementioned problems and to balance the individual rate of the UEs.

Input : Channel realization $\mathbf{G}_{1,1}, \cdots, \mathbf{G}_{2,N}$ and power constraint P_u for $1 \leq u \leq 2$

Initialization: For all $1 \leq u \leq 2$ and $1 \leq n \leq N$, perform SVD for $\mathbf{G}_{u,n}$, obtain $\mathbf{v}_{u,n}^{EV}$, set $\mathbf{h}_{1,n} = \mathbf{G}_{1,n}\mathbf{v}_{1,n}^{EV}$, calculate $\mathbf{v}_{2,n}^{\perp}$, construct $\mathbf{v}_{2,n}^{EV,\perp}$ according to (7), set $\mathbf{h}_{2,n} = \mathbf{G}_{2,n}\mathbf{v}_{1,n}^{EV,\perp}$ and $\mathbf{O}_{u,l} = 0 \cdot \mathbf{I}$

for $n = 1 : N$ **do**

calculate $\mathbf{H}_n$ according to (2);

$\alpha_{n,opt} = \max_\alpha \sum_{u=1}^{2} N \log(1 + \frac{P_u/\sigma^2}{\mathbf{O}_{u,l} + [(\mathbf{H_n^H H_n})^{-1}]_{u,u}})$;

$\mathbf{H}_{n,opt} = \left[\mathbf{h}_{1,n} \quad \mathbf{G_{2,n}}\left[\alpha_{\mathbf{n,opt}}\mathbf{v}_{2,\mathrm{n}}^{\mathrm{EV}} + (\mathbf{1} - \alpha_{\mathbf{n,opt}})\mathbf{v}_{2,\mathrm{n}}^{\perp}\right]\right]$;

$\mathbf{O}_{u,l} = \mathbf{O}_{u,l} + \left[(\mathbf{H_{n,opt}^H H_{n,opt}})^{-1}\right]_{u,u}, u = 1, 2;$

end

Output : $\mathbf{v}_{1,n}^{EV}$, $\alpha_{n,opt}$ and $\mathbf{v}_{2,n}^{EV,\perp}$ according to (7), for $1 \leq n \leq N$

Algorithm 1: Spatial Precoder Optimization Algorithm

3.2. Scheduling

The scheduler works as follows. It keeps track of the average rate $R_{avg,u}$ of each UE. In time slot t, the scheduling algorithm assigns the dominant EV precoder to UE u^* with smaller $R_{avg,u}$ in the system, which aims to give higher priority to the weaker UE to balance the individual user rate. In addition, in order to avoid interfering the rate of weaker UE if the UEs experience largely unbalanced channel conditions in the system, a weighting factor β is introduced to weight the compound channel orthogonal to that of the weaker UE by setting

$$h_{\tilde{u},n} = \begin{cases} \beta \mathbf{G}_{\tilde{u},n}\left(\frac{\mathbf{v}_{\tilde{u},n}^{\perp}}{||\mathbf{v}_{\tilde{u},n}^{\perp}||_2}\right), & \alpha = 0; \\ \mathbf{G}_{\tilde{u},n}\mathbf{v}_{\tilde{u},n}^{EV,\perp}, & \alpha \neq 0 \end{cases} \quad (9)$$

in the precoder optimization algorithm, where $\tilde{u}$ denotes the UE with higher average rate $R_{avg,u}$. In (9) choosing a bigger β means to virtually boost the quality of the channel orthogonal to the transmission direction of the weaker UE, so that it is treated as a good channel and the selection procedure preferably picks it up for the stronger UE. In other words, a bigger β indicates higher importance that the UE with higher average rate in the past should transmits in a direction which does not cause any interference to the UE with lower average rate in the past and vice versa.

4. SIMULATION RESULTS

To evaluate the performance of the proposed spatial precoders in a 2x2 uplink MIMO system with two UEs as shown in Fig. 2, simulations are conducted in the 3GPP LTE uplink with the parameter assumptions given in Table 1.

Table 1. Parameter assumptions for simulation

Parameters	**Assumption**
Carrier frequency	2.0GHz
Transmission bandwidth	3MHz, 20MHz
Transmit time interval	1ms
Sub-carrier spacing	15KHz
Number of sub-carriers	180, 1200
Number of sub-carriers per RB	12
Channel model	3GPP SCME urban macro [12]
Number of UEs	2
Number of BSs	1
Antennas per UE	2
Antennas per BS	2
UE antenna spacing	0.5 wavelength
BS antenna spacing	10 wavelengths
UE velocity	10 m/s

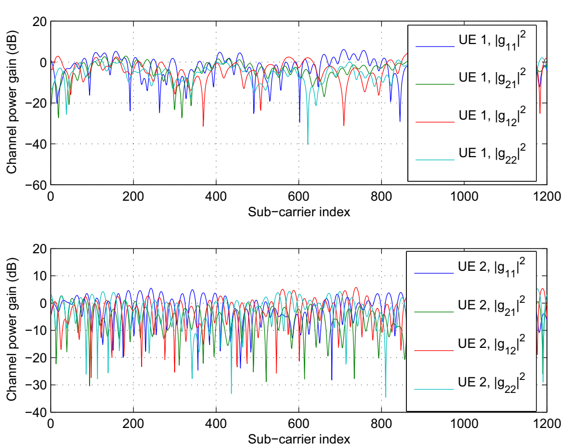


Fig. 3. Channel power gain snapshot between the UEs and the BS according to Fig. 2; g_{ij} denotes the sub-carrier frequency response between the jth transmitter of the UE and the ith receiver at the BS.

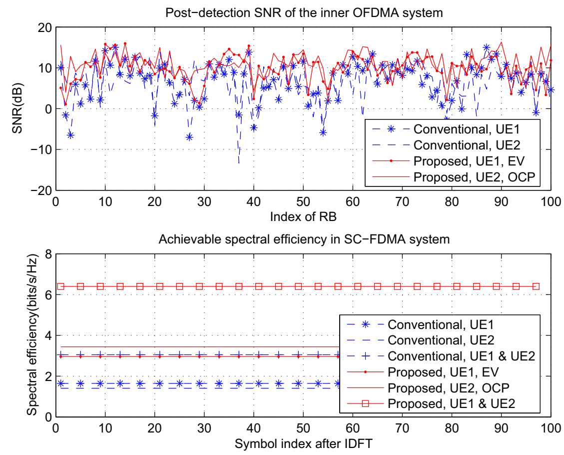


Fig. 4. Performance comparison between the conventional virtual MIMO SC-FDMA and the proposed MIMO SC-FDMA with precoding using the channel snapshot in Fig. 3. (Upper: post-detection SNR of the inner OFDMA system; Lower: achievable spectral efficiency in the SC-FDMA system)

A snapshot of the sub-carrier channel power gain between the UEs and the BS is illustrated in Fig. 3. For simplicity, it is further assumed that each RB experiences the same channel condition and its channel frequency response is represented by the middle, i.e., the $6th$, sub-carrier of the RB. Under this condition, performance evaluation can be conducted per RB basis and the concept meant for a sub-carrier in our previous discussion can be directly applied to a RB to reduce the computational complexity. In the following, first the performance is evaluated using a channel snapshot of 20MHz (totally 100 RBs) for illustrative purpose. Then we present results in terms of average spectral efficiency for different bandwidth and SNR conditions.

The upper part of Fig. 4 shows the post-detection SNR $\gamma_{u,n}$ of the inner OFDMA system using the channel snapshot in Fig. 3. The dashed lines represent the results for the conventional UEs with each having only a single antenna, i.e., setting $\mathbf{v}_{u,n} = [1 \quad 0]$ for all u's and n's in our model. The solid lines stand for the results obtained by the proposed precoding scheme where UE 1 employs the dominant EV of the average channel correlation matrix as precoders for all the sub-carriers and UE 2 uses the OCP for each sub-carrier. The spectral efficiency of each sub-channel between each input and output component in the SC-FDMA system is plotted in the lower part of Fig. 4. It can be observed that the proposed algorithm significantly improves the spectral efficiency, or in other words the rate of the UEs in the system.

Next, the statistics of the post-detection SNR is studied

see the program

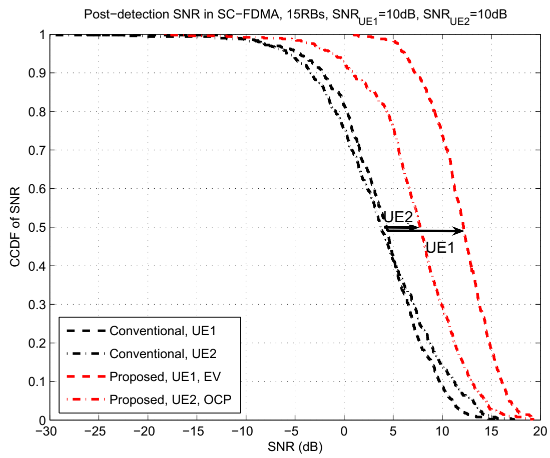


Fig. 5. Complementary cumulative distribution function (CCDF) of the post-detection SNR between the conventional MIMO SC-FDMA and the proposed MIMO SC-FDMA with spatial precoding. 15RBs are available in the system and both UEs have the same average received SNR of 10dB

over 5000 channel samples for the conventional virtual MIMO SC-FDMA and the proposed MIMO SC-FDMA with spatial precoding, where a setting of two UEs with the same average SNR=10dB over 15RBs (≈3MHz) is assumed. The complementary cumulative distribution function (CCDF) of the post-detection SNR for both schemes are compared in Fig. 5, which gives the probability that the post-detection SNR is larger than a certain value. It can seen that with the proposed spatial precoding scheme the post-detection SNRs of both UEs are significantly increased, where the improvement for the UE using dominant EV precoding is much larger than that for the UE using OCP.

For the same setting, Fig. 6 shows the cumulative distribution function (CDF) of the system spectral efficiency and the individual user spectral efficiency for the conventional virtual MIMO SC-FDMA and the proposed MIMO SC-FDMA with spatial precoding. For better illustration, the results are repeated in both the left and the right parts of Fig. 6, where the left part emphasizes the system spectral efficiency and the right part the individual user spectral efficiency. It can be observed that in comparison with the conventional virtual MIMO SC-FDMA the 50 percentile achievable system spectral efficiency almost doubles after applying the proposed scheme for different fixed optimization orders. It can also be clearly seen that with the fixed optimization orders, the UE using the dominant EV precoder always has higher average spectral efficiency than the one using OCP. If the proposed spatial scheduler is used, individual spectral efficiency of the UEs are balanced and the CDF of the achievable system spectral efficiency (green) lies between those obtained by fixed optimization orders (red and blue). For reference purpose, the spectral efficiency obtained by the virtual MIMO OFDMA system, where each UE has only one antenna, is also plotted which is about 1 bit/s/Hz less in the 50 percentile achievable system spectral efficiency with respect to the proposed schemes. Note that the impact that the OFDMA waveform requires higher input power backoff than the SC-FDMA waveform is not taken into account for performance evaluation here. Therefore, even higher gains can be expected from the proposed precoding schemes in reality.

Fig. 7 depicts the performance comparison assuming 100 RBs are available in the system and other parameters remain the same as in the previous simulation. It can be seen that both the conventional virtual MIMO SC-FDMA and the proposed precoding schemes have inferior performance to the case with 15RBs (cf. Fig. 6). The reason is as follows: with increasing bandwidth in the system, the probability that a deep fade happens in some sub-carriers and/or the probability that both UEs have similar spatial signatures in some sub-carriers also increase. Such deep fades and/or similar spatial signature cause a low post-detection SNR in some sub-carriers of the inner OFDMA system, which degrades the post-detection SNR of all the received components in the SC-FDMA system due to the harmonic mean property (cf. (4)). Nevertheless, the proposed scheme still achieves about twice as high system spectral efficiency as the conventional scheme. A comparison between Fig. 7 and Fig. 6 also indicates that for the same average received SNR the achievable system spectral efficiency of a virtual MIMO-OFDMA system is not sensitive to the occupied bandwidth.

Fig. 8 compares the performance with unequal average received SNR for different UEs, i.e., $\mathrm{SNR}_{\mathrm{UE1}} = 0\mathrm{dB}$ and $\mathrm{SNR}_{\mathrm{UE2}} = 10\mathrm{dB}$ in the 3MHz bandwidth. More than twice the system spectral efficiency with respect to the conventional scheme can be expected from the proposed spatial precoding schemes for different fixed optimization orders. It can be seen that our proposed scheduling algorithm (green) can effectively balance the individual rate of the UEs although they experience largely unbalanced channel conditions, while the scheme always assigning dominant eigenbeam direction to the weaker UE and OCP for the stronger UE results in significant unbalance of the individual rates. This is due to the strong amount of interference caused by the stronger UE.

5. CONCLUSIONS

We propose to use multiple antennas and spatial precoding to improve the system sum rate in MIMO SC-FDMA uplink. An efficient algorithm is developed to optimize the precoder design which is shown to significantly increases the system sum rate in 3GPP LTE uplink scenario with two UEs for different SNR conditions and bandwidths. Furthermore, a simple spatial scheduler is introduced to handle the user rate fairness in the framework of the proposed algorithm.

see the program

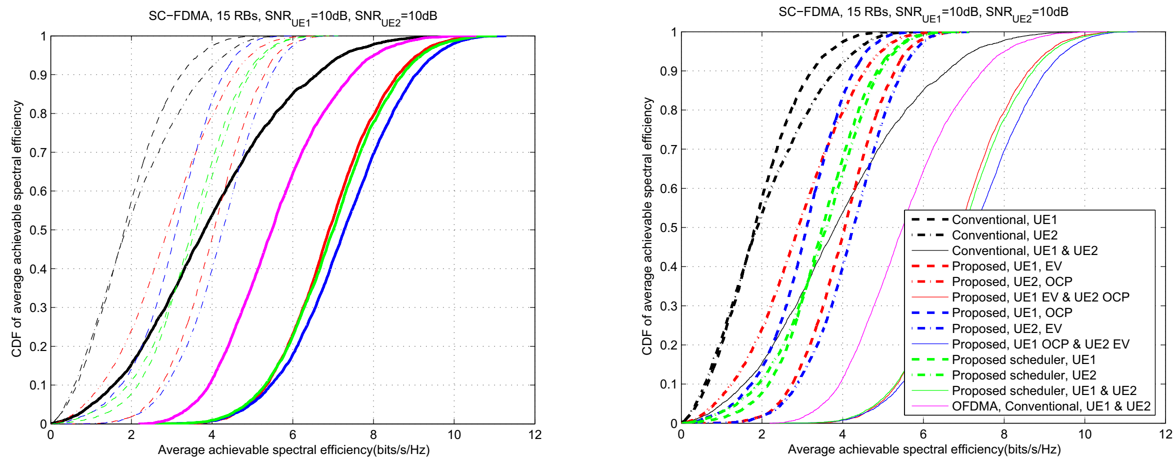


Fig. 6. Cumulative distribution function (CDF) of the achievable spectral efficiency by using conventional virtual MIMO with a single antenna per UE (black) and by using a spatial precoder according to Fig. 2 with fixed optimization order (red and blue) and a spatial scheduler (green) for an SC-FDMA system ($\beta = 1$). The inner OFDMA spectral efficiency with a single antenna per UE is plotted in magenta. 15RBs are available in the system and both UEs have the same average received SNR of 10dB.

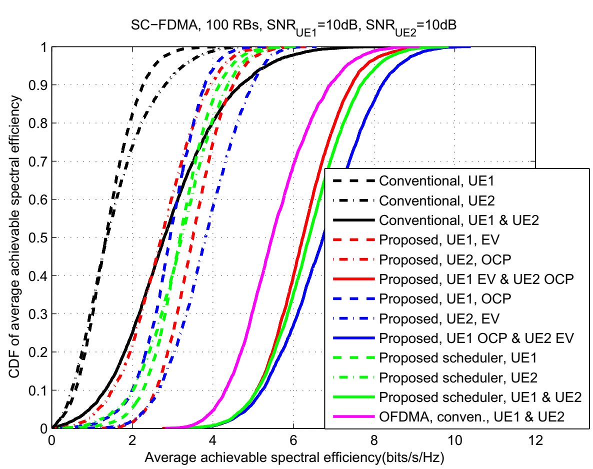


Fig. 7. Parameter setting is the same with Fig. 6 but 100RBs are available in the system

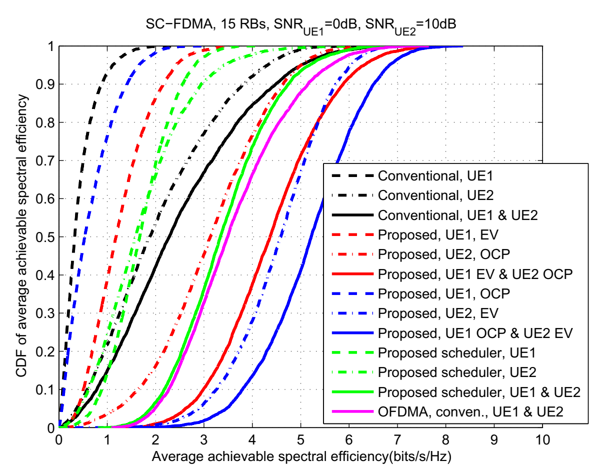


Fig. 8. Parameter setting is the same with Fig. 6 but with unequal average received SNR for the UEs: $\mathrm{SNR}_{\mathrm{UE1}} = 0\mathrm{dB}$ and $\mathrm{SNR}_{\mathrm{UE2}} = 10\mathrm{dB}$; $\beta = 10$

see the program

6. REFERENCES

[1] H.G. Myung, "Introduction to single carrier FDMA," in *15th European Signal Processing Conference (EUSIPCO)*, Poznan, Poland, Sept. 2007.

[2] 3GPP TSG RAN, *Some aspects of single-carrier transmission for E-UTRA*, R1-050765, Aug. 29-Sept. 2 2005.

[3] H.G. Myung, J. Lim, and D.J. Goodman, "Peak-to-average power ratio of single carrier FDMA signals with pulse shaping," in *PIMRC06*, Helsinki, Finland, Sept. 2006.

[4] U. Sorger, I. Broeck, and M. Schnell, "Interleaved FDMA-a new spread-spectrum multiple-access scheme," in *IEEE International Conference on Communications, ICC 98*, Atlanta, Georgia, USA, June 1998, vol. 2, pp. 1013–1017.

[5] J. Lim, H.G. Myung, K. Oh, and D.J. Goodman, "Channel-dependent scheduling of uplink single carrier FDMA systems," *IEEE VTC06 fall*, Sept. 2006.

[6] T. Svensson, T. Frank, D. Falconer, M. Sternad, E. Costa, and A. Klein, "B-IFDMA-a power efficient multiple access scheme for non-frequency-adaptive transmission," in *IST mobile and wireless communications summit*, Budapest, Hungary, July 2007.

[7] H. Wu and T. Haustein, "Sum rate analysis of SDMA transmission in single carrier FDMA system," in *Proc. 11th International Conference on Communication Systems*, Guangzhou, China, Nov. 2008.

[8] H.G. Myung, K.J.L. Pan, R. Olesen, and D. Grieco, "Peak power characteristics of single carrier FDMA MIMO precoding system," in *IEEE 66th Vehicular Technology Conference, VTC-2007*, Fall, 2007, pp. 477–481.

[9] A. Narula, M.D. Trott, and G.W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2418–2433, 1999.

[10] D. Tse and P. Viswanath, *Fundamentals of wireless communication*, Cambridge University Press, Cambridge, UK, 2005.

[11] G.H. Golub and C.F. Van Loan, *Matrix computations*, 1996.

[12] D. S. Baum, J. Hansen, G. Del Galdo, M. Milojevic, J. Salo, and P. Kyösti, "An interim channel model for beyond-3G systems: extending the 3GPP spatial channel model (SCM)," in *Proc. of the 61st IEEE VTC'05*.

see the program

ON THE ASYMPTOTIC OPTIMALITY OF BLOCK-DIAGONALIZATION FOR THE MIMO BC UNDER LINEAR FILTERING

Raphael Hunger and Michael Joham

Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany
Telephone: +49 89 289-28508, Fax: +49 89 289-28504, Email: hunger@tum.de

ABSTRACT

In this paper, we address the high SNR regime of the MIMO broadcast channel under linear filtering. For systems where the base station is equipped with more antennas than the user terminals have in sum, we prove that block-diagonalization is the asymptotically optimum transmission strategy for maximizing the sum rate. For this type of transmission strategy, the asymptotically optimum transmit covariance matrix in the broadcast channel is derived in closed form. In addition, we present an expression for the asymptotic sum capacity for an instantaneous channel realization which only depends on this particular channel realization. No precoders or singular-value-decompositions arise as they used to do in hitherto existing sum rate expressions in the multi-antenna terminal case. All results are deduced from the dual multiple access channel in which the optimum transmit covariance matrices can easily be computed. Our recent rate duality for multi-antenna systems where the individual streams of a user are not treated as self-interference allows us then to convert the solution of the dual uplink back to the downlink and to find the optimum transmit and receive filters in the broadcast channel.

1. INTRODUCTION

Block-diagonalization (BD) is a multi-user MIMO channel decomposition technique that orthogonalizes the overall MIMO broadcast channel (BC) of the different users into parallel single-user MIMO channels without inter-user interference. It can thus be thought of as the generalization of zero-forcing channel inversion algorithms for the case when the receivers have multiple antennas. Since capacity achieving transmission strategies do not seem to exist for the multi-user MIMO case under linear filtering, block-diagonalization was introduced as a constraint which leads to closed-form expressions for the resulting sum rate [1]. Depending on the number of available antennas at the base station and the number of antennas at the terminals, standard BD (e.g. [1–3]), BD with user selection (e.g. [4]), or generalized BD where the receive filters are taken into account (e.g. [5]), can be applied. While the complete suppression of inter-user interference, that was heuristically introduced to end up with closed form expressions [1], fails to reach the sum capacity in general, we will show that in the high SNR regime, block-diagonalization is indeed the asymptotically optimum transmission strategy in the multi-user MIMO BC with linear filtering. In order to prove above statement, we utilize the recently introduced rate duality [6] for multi-antenna users under linear filtering and investigate the sum rate maximization problem in the dual uplink multiple access channel (MAC) instead. Despite the simpler structure of the dual MAC (aligned channel and precoder indices), the simplest multi-user setup with single antenna terminals already allows for the presumption that closed form expressions for the sum capacity will remain infeasible even in the multiple access channel irrespective of whether linear or nonlinear filtering is considered. Fortunately, the high signal-to-noise ratio regime is an exception to this deflating circumstance, since there, asymptotic results on the sum capacity have been discovered for dirty paper coding and partly for linear filtering. In [7, 8] for example, the single user point-to-point MIMO case was decomposed into a supremum capacity term, an instantaneous SNR effect term, and an instantaneous capacity degradation term due to the eigenvalue spread. Outage capacity and throughput of a fading point-to-point MIMO system are analyzed in [9], whereas a lower and an upper bound on the sum rate of block-diagonalization was derived in [10]. Nonetheless, precoder-free expressions for the asymptotic sum capacity of a point-to-multipoint broadcast channel for an *instantaneous* channel realization did not exist so far for linear filtering in the multi-antenna terminals case, only sum rate expressions still containing the precoders have been derived yet when the individual users are equipped with several antennas, see [11, Eq. (21)] and [12, Eq. (10)]. Instead, *ergodic* statements can be found in the literature, see for example [11–13]. Therein, the affine approximation of the sum capacity introduced in [14] and elaborately discussed in [15] was applied.

Having derived the asymptotic sum capacity of the dual MAC, we can immediately conclude by means of the duality in [6] that the broadcast channel features the same high SNR sum capacity. Moreover, when we convert the simple solution for the precoders in the dual MAC that asymptotically achieve the sum capacity back to the broadcast channel, it turns out that the resulting BC system features a block-diagonal structure. Hence, we have a formal proof that block-

see the program

diagonalization is the asymptotically optimum transmission strategy when linear filtering is considered what so far has not been shown, but often been used and investigated, see e.g. [11, 12]. Of course, the MAC to BC conversion also delivers the asymptotically optimum transmit and receive filters.

1.1. Contributions

The main contributions of this paper are summarized in the following list:

1. We prove that block diagonalization is asymptotically optimum for sum rate maximization in the broadcast channel.
2. Optimum precoding and transmit covariance matrices in the broadcast channel are derived by means of our rate duality in [6]. In contrast to existing block-diagonalization algorithms, singular-value-decompositions do not arise in our notation and therefore do not have to be computed.
3. We derive an analytic expression for the sum capacity of the multi-user MIMO BC achievable with linear filtering in the high SNR regime which is the first precoder-free closed form solution depending only on the channel matrices and the antenna configuration.
4. A closed form solution of the covariance matrices in the dual uplink asymptotically achieving this sum capacity.

1.2. Organization

In Section 2, the system model underlying the multi-user scenario is described. The optimum signaling strategy for the sum-rate maximization with linear filtering is derived in Section 3 in the dual multiple access channel and afterwards converted to the broadcast channel in Section 4. Section 5 concludes this paper.

2. SYSTEM MODEL

We consider the communication between an N antenna base station and K multi-antenna terminals, where user k multiplexes B_k data streams over his r_k antennas. For a short notation, we define r as the sum of all antennas at the terminals, i.e., $r = \sum_{k=1}^{K} r_k$, and b as the total number of transmitted streams, i.e., $b = \sum_{k=1}^{K} B_k$. Recent results [6] on the rate duality of the BC and the MAC under linear filtering and a user-wise joint stream decoding allow us to investigate the dual MAC (with its simpler structure) instead of the BC. Afterwards, the obtained solutions can conveniently be transformed back to the BC. In this MAC, user k applies a precoding matrix $\boldsymbol{T}_k \in \mathbb{C}^{r_k \times B_k}$ generating his $r_k \times r_k$ transmit covariance matrix $\boldsymbol{Q}_k = \boldsymbol{T}_k \boldsymbol{T}_k^{\mathrm{H}}$. The precoded symbol vector propagates over the channel described by the matrix $\boldsymbol{H}_k \in \mathbb{C}^{N \times r_k}$. At the receiver side, zero-mean noise $\boldsymbol{\eta} \in \mathbb{C}^N$ with identity covariance matrix is added and the receive filter for user k is denoted by $\boldsymbol{G}_k \in \mathbb{C}^{B_k \times N}$. Due to the reversed signal flow in the BC, we characterize the transmission from the base station to terminal k by the Hermitian channel $\boldsymbol{H}_k^{\mathrm{H}}$ in the BC, and the precoder dedicated to the B_k streams of user k is denoted by $\boldsymbol{P}_k \in \mathbb{C}^{N \times B_k}$. Throughout this paper, we assume that the base station has at least as many antennas as the terminals have in sum, i.e., $N \geq r$.

3. OPTIMUM SIGNALING IN THE DUAL MAC

Introducing the composite channel matrix $\boldsymbol{H}$ and the composite block-diagonal precoder matrix $\boldsymbol{T}$ of all K users via

$$\begin{aligned} \boldsymbol{H} &= [\boldsymbol{H}_1, \ldots, \boldsymbol{H}_K] \in \mathbb{C}^{N \times r}, \\ \boldsymbol{T} &= \operatorname{blockdiag}\{\boldsymbol{T}_k\}_{k=1}^{K} \in \mathbb{C}^{r \times b}, \end{aligned} \tag{1}$$

the rate of user k seeing interference from all other users can be expressed as (see [6])

$$\begin{aligned} R_k &= \log_2 \left| \mathbf{I}_N + \Big(\mathbf{I}_N + \sum_{\ell \neq k} \boldsymbol{H}_\ell \boldsymbol{Q}_\ell \boldsymbol{H}_\ell^{\mathrm{H}}\Big)^{-1} \boldsymbol{H}_k \boldsymbol{Q}_k \boldsymbol{H}_k^{\mathrm{H}} \right| \\ &= -\log_2 \left| \mathbf{I}_{B_k} - \boldsymbol{T}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{X}^{-1} \boldsymbol{H}_k \boldsymbol{T}_k \right|, \end{aligned} \tag{2}$$

where the substitution $\boldsymbol{X}$ reads as

$$\boldsymbol{X} = \mathbf{I}_N + \sum_{\ell=1}^{K} \boldsymbol{H}_\ell \boldsymbol{Q}_\ell \boldsymbol{H}_\ell^{\mathrm{H}} = \mathbf{I}_N + \boldsymbol{H}\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}.$$

Note that (2) in general requires all streams of a single user to be decoded jointly. Reformulating the rate expression (2), we get

$$\begin{aligned} R_k &= -\log_2 \left| \boldsymbol{E}_k^{\mathrm{T}} \big(\mathbf{I}_b - \boldsymbol{T}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{X}^{-1} \boldsymbol{H} \boldsymbol{T}\big) \boldsymbol{E}_k \right| \\ &= -\log_2 \left| \boldsymbol{E}_k^{\mathrm{T}} \big(\mathbf{I}_b + \boldsymbol{T}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \boldsymbol{T}\big)^{-1} \boldsymbol{E}_k \right|, \end{aligned} \tag{3}$$

where the transposed kth block unit matrix is defined via

$$\boldsymbol{E}_k^{\mathrm{T}} = [\mathbf{0}, \ldots, \mathbf{0}, \mathbf{I}_{B_k}, \mathbf{0}, \ldots, \mathbf{0}] \in \{0,1\}^{B_k \times b}$$

with the identity matrix at the kth block and the ith block with $i \neq k$ corresponds to the zero matrix of dimension $B_k \times B_i$. Due to the assumption that the base station has more antennas than the terminals have in sum, all r streams can be activated leading to square precoders $\boldsymbol{T}_k$ with $B_k = r_k \ \forall k$. Raising P_{Tx}, all r streams become active, $\boldsymbol{T}$ becomes full rank, and all eigenvalues of $\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{T}$ become much larger than one. In the asymptotic limit, we obtain

$$\begin{aligned} R_k &\cong -\log_2 \left| \boldsymbol{T}_k^{-1} \boldsymbol{E}_k^{\mathrm{T}} \big(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\big)^{-1} \boldsymbol{E}_k \boldsymbol{T}_k^{-\mathrm{H}} \right| \\ &= \ \log_2 \left| \boldsymbol{Q}_k \right| - \log_2 \left| \boldsymbol{E}_k^{\mathrm{T}} \big(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\big)^{-1} \boldsymbol{E}_k \right|, \end{aligned} \tag{4}$$

since $\boldsymbol{E}_k^{\mathrm{T}} \boldsymbol{T}^{-1} = \boldsymbol{T}_k^{-1} \boldsymbol{E}_k^{\mathrm{T}}$. The notation $x \cong y$ means that the difference $x - y$ vanishes when the sum power P_{Tx} goes

see the program

to infinity. Interestingly, the rate of user k depends only on the determinant of his own transmit covariance matrix $\boldsymbol{Q}_k$, and not on the covariance matrices of the other users! Consequently, the eigenbases of all transmit covariance matrices do not influence the rates of the users, only the powers of the eigenmodes are relevant. Let the eigenvalue decomposition of $\boldsymbol{Q}_k$ read as $\boldsymbol{Q}_k = \boldsymbol{V}_k \boldsymbol{\Lambda}_k \boldsymbol{V}_k^{\mathrm{H}}$ with unitary $\boldsymbol{V}_k$ and the diagonal nonnegative power allocation $\boldsymbol{\Lambda}_k$. Due to the determinant operator, $\boldsymbol{V}_k$ can be chosen arbitrarily and therefore, we set $\boldsymbol{V}_k = \mathbf{I}_{r_k}\ \forall k$ without loss of generality. Let the power allocation matrix be composed by the entries $\boldsymbol{\Lambda}_k = \mathrm{diag}\{\lambda_k^{(i)}\}_{i=1}^{r_k}$. As only the traces of the covariance matrices are involved in the sum power constraint $\sum_{k=1}^{K} \mathrm{tr}(\boldsymbol{Q}_k) \le P_{\mathrm{Tx}}$, the determinant $|\boldsymbol{Q}_k| = |\boldsymbol{\Lambda}_k|$ is maximized by setting

$$\lambda_k^{(1)} = \ldots = \lambda_k^{(r_k)} := \lambda_k, \tag{5}$$

i.e., by evenly distributing the power allocated to that user onto his individual modes, so $\boldsymbol{Q}_k = \lambda_k \mathbf{I}_{r_k}\ \forall k$. Inserting (5) into the asymptotic rate equation of user k in (4) leads to the asymptotic sum rate expression

$$\sum_{k=1}^{K} R_k \cong \sum_{k=1}^{K} \left(r_k \log_2 \lambda_k - \log_2 \left| \boldsymbol{E}_k^{\mathrm{T}} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \right| \right). \tag{6}$$

Subject to the sum power constraint $\sum_{k=1}^{K} r_k \lambda_k \le P_{\mathrm{Tx}}$, the sum rate in (6) is maximized for

$$\lambda_k = \frac{P_{\mathrm{Tx}}}{r}, \tag{7}$$

so power is evenly allocated to the users (similar to the single-antenna case proven in [11]), and every user evenly distributes his fraction of power onto his modes. Finally, the asymptotic sum rate at high SNR reads as

$$\sum_{k=1}^{K} R_k \cong r \log_2 P_{\mathrm{Tx}} - r \log_2 r - \sum_{k=1}^{K} \log_2 \left| \boldsymbol{E}_k^{\mathrm{T}} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \right|, \tag{8}$$

and is interestingly achieved with the transmit covariance matrices $\boldsymbol{Q}_k = P_{\mathrm{Tx}}/r \cdot \mathbf{I}_{r_k}\ \forall k$. Using (8), we are able to quantify the asymptotic sum rate that can be achieved by means of linear filtering for every single channel realization and antenna/user profile in terms of the transmit power P_{Tx} and the channel itself as long as $N \ge r$ holds. Note that no precoders arise in (8) in contrast to [12, Eq. (10)] and [11, Eq. (21)]. In principle, the ergodic rate can be obtained by averaging corresponding to *any* distribution of the channel. In [12], results on the *ergodic* rate were presented for the specific case of Rayleigh fading only, where the channel entries of $\boldsymbol{H}_1, \ldots, \boldsymbol{H}_K$ all have the same distribution. More complicated fading models cannot be captured due to this restricting assumption. Moreover, the instantaneous rate expression is given by means of bases representing null spaces of shortened channel matrices taken from [1] and not as a function of the channel purely as we do in (8). Concerning the asymptotic rate expression, we have created a smooth transition from the r single-antenna-users system configuration in [13] where no cooperation exists between the antenna elements at the terminals, to the single-user point-to-point MIMO link where all r antennas fully cooperate, see [16] for example. In between, we can now specify any antenna/user profile we want and compute the feasible rate in the asymptotic limit under linear filtering.

4. OPTIMUM SIGNALING IN THE BC

Using our recent rate duality in [6], we can convert the simple solution for the covariance matrices $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_K$ in the dual MAC to covariance matrices $\boldsymbol{S}_1, \ldots, \boldsymbol{S}_K$ in the BC, where the Hermitian channels are applied. Since this duality explicitly uses the receive filters in the MAC as scaled transmit matrices in the BC, we first compute the MMSE receivers in the dual MAC, as they are optimum and generate sufficient statistics. This MMSE receive filter $\boldsymbol{G}_k$ for user k in the dual MAC reads as

$$\boldsymbol{G}_k = \boldsymbol{E}_k^{\mathrm{T}} \boldsymbol{T}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \left(\mathbf{I}_N + \boldsymbol{H} \boldsymbol{T} \boldsymbol{T}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \right)^{-1}.$$

Using asymptotically optimum precoders $\boldsymbol{T}_k = \sqrt{P_{\mathrm{Tx}}/r}\,\mathbf{I}_{r_k}$, above expression asymptotically converges to

$$\boldsymbol{G}_k \cong \sqrt{r/P_{\mathrm{Tx}}} \cdot \boldsymbol{E}_k^{\mathrm{T}} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{H}^{\mathrm{H}}. \tag{9}$$

Let $\boldsymbol{P}_k$ denote the precoder of user k in the BC, then the ith column $\boldsymbol{p}_{k,i}$ of $\boldsymbol{P}_k$ follows from the conjugate ith row $\boldsymbol{g}_{k,i}'^{\mathrm{T}}$ of the matrix $\boldsymbol{G}_k' = \boldsymbol{W}_k^{\mathrm{H}} \boldsymbol{G}_k$ via (see [6])

$$\boldsymbol{p}_{k,i} = \alpha_{k,i} \boldsymbol{g}_{k,i}'^{*} = \frac{\alpha_{k,i}}{\sqrt{P_{\mathrm{Tx}}/r}} \cdot \boldsymbol{H} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \boldsymbol{W}_k \boldsymbol{e}_i, \tag{10}$$

where the scaling factor $\alpha_{k,i}$ is obtained by the duality transformation and $\boldsymbol{W}_k$ is a unitary decorrelation matrix. This matrix $\boldsymbol{W}_k$ finally ensures that the individual streams of every user can be decoded separately instead of jointly without having to face the rate loss that usually has to be taken into account when streams are decoded separately rather than jointly, see [6]. Since we convert only the asymptotically optimum transmit precoders and receive filters, the duality transformation from the MAC to the BC in [6] drastically simplifies and can even be computed in closed form. In particular, the matrices $\boldsymbol{M}_{a,b}$ in [6, Eq. (23)] vanish for $a \ne b$ yielding a *diagonal* matrix $\boldsymbol{M}$ from which the scaling factors are derived. Those scalars now compute to

$$\alpha_{k,i} = \frac{\sqrt{P_{\mathrm{Tx}}/r}}{\|\boldsymbol{g}_{k,i}'\|_2}. \tag{11}$$

In combination with (10), the ith column of the precoder associated to user k reads as

$$\boldsymbol{p}_{k,i} = \sqrt{P_{\mathrm{Tx}}/r} \cdot \frac{\boldsymbol{H} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \boldsymbol{W}_k \boldsymbol{e}_i}{\left\| \boldsymbol{H} \left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \boldsymbol{E}_k \boldsymbol{W}_k \boldsymbol{e}_i \right\|_2},$$

see the program

generating the precoder matrix

$$\boldsymbol{P}_k = \sqrt{P_{\mathrm{Tx}}/r} \cdot \boldsymbol{H}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{E}_k\boldsymbol{W}_k\boldsymbol{D}_k^{-1}, \tag{12}$$

where the ith diagonal element of the diagonal matrix $\boldsymbol{D}_k$ is

$$[\boldsymbol{D}_k]_{i,i} = \sqrt{\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{W}_k^{\mathrm{H}}\boldsymbol{E}_k^{\mathrm{T}}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{E}_k\boldsymbol{W}_k\boldsymbol{e}_i}. \tag{13}$$

We can immediately see, that the precoding filters in (12) lead to a block diagonalization of the transmission, since $\boldsymbol{H}_\ell^{\mathrm{H}}\boldsymbol{P}_k = \boldsymbol{0}$ holds for $k \neq \ell$. Next, the decorrelation matrix $\boldsymbol{W}_k$ which enables the duality is usually chosen as the eigenbasis of $\boldsymbol{G}_k\boldsymbol{H}_k\boldsymbol{T}_k \cong \mathbf{I}_{r_k}$, which asymptotically coincides with the identity matrix due to (9). Since all eigenvalues are identical to one, the decorrelation matrices $\boldsymbol{W}_k$ are not given a priori, but can easily be computed such that the BC features the same sum rate as the dual MAC. By means of (12) and the block diagonalization property of the precoders, we obtain for user k's receive signal

$$\boldsymbol{y}_k = \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}_k\boldsymbol{s}_k + \boldsymbol{\eta}_k = \sqrt{P_{\mathrm{Tx}}/r} \cdot \boldsymbol{W}_k\boldsymbol{D}_k^{-1}\boldsymbol{s}_k + \boldsymbol{\eta}_k, \tag{14}$$

where $\boldsymbol{\eta}_k \in \mathbb{C}^{r_k}$ is the noise and $\boldsymbol{s}_k$ the symbol vector of user k both having an identity covariance matrix. From (14), the rate of user k achieved in the BC reads as

$$R_k = \log_2\left|\mathbf{I}_{r_k} + P_{\mathrm{Tx}}/r \cdot \boldsymbol{W}_k\boldsymbol{D}_k^{-2}\boldsymbol{W}_k^{\mathrm{H}}\right|, \tag{15}$$

which asymptotically converges to

$$R_k \cong r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2 |\boldsymbol{D}_k^2|. \tag{16}$$

For the asymptotic result in (16), the identity matrix $\mathbf{I}_{r_k}$ in (15) was omitted, so the true rate R_k will always converge to the right hand side from (16) *from above*. It remains to minimize the determinant of $\boldsymbol{D}_k^2$ by the choice of the unitary decorrelation filter $\boldsymbol{W}_k$, such that the asymptotic limit of R_k is maximized, see (16). From the definition of $\boldsymbol{D}_k$ in (13), *Hadamard's* inequality [17] tells us that $\boldsymbol{W}_k$ has to be chosen as the unitary eigenbasis of $\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k$. Choosing $\boldsymbol{W}_k$ this way, $\boldsymbol{D}_k^2$ contains the eigenvalues of $\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k$, and therefore, the elements of $\boldsymbol{D}_k^2$ are as different as possible since the eigenvalues of any positive definite matrix majorize its diagonal elements according to *Schur's* theorem [18]:

$$\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k = \boldsymbol{W}_k\boldsymbol{D}_k^2\boldsymbol{W}_k^{\mathrm{H}}. \tag{17}$$

The optimum transmit covariance matrix $\boldsymbol{S}_k = \boldsymbol{P}_k\boldsymbol{P}_k^{\mathrm{H}}$ of user k in the BC which asymptotically achieves the same sum rate as the dual MAC counterpart reads by means of (12) and (17) as

$$\boldsymbol{S}_k = \frac{P_{\mathrm{Tx}}}{r} \cdot \boldsymbol{H}^{+\,\mathrm{H}}\boldsymbol{E}_k\left(\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k\right)^{-1}\boldsymbol{E}_k^{\mathrm{T}}\boldsymbol{H}^{+}. \tag{18}$$

In (18), we make use of the channel pseudo-inverse $\boldsymbol{H}^{+}$ which is defined via $\boldsymbol{H}^{+} = (\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{H}}$. Note that r_k eigenvalues of $\boldsymbol{S}_k$ are P_{Tx}/r whereas the remaining $N - r_k$ ones are zero. Thus, $\boldsymbol{S}_k$ is a weighted orthogonal projector. Furthermore, $\mathrm{tr}(\boldsymbol{S}_k) = P_{\mathrm{Tx}} \cdot r_k/r\ \forall k$, so the power is uniformly allocated to the individual users in the broadcast channel as well. Comparing (18) with the simple solution of the transmit covariance matrix $\boldsymbol{Q}_k = P_{\mathrm{Tx}}/r \cdot \mathbf{I}_{r_k}$ in the dual MAC, it becomes obvious that the optimum covariance matrices are much more difficult to find directly in the BC without using the rate duality, than in the dual MAC. Plugging the optimum $\boldsymbol{D}_k^2$ containing the eigenvalues of $\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k$ into (16) finally yields

$$R_k \cong r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2\left|\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k\right|.$$

Hence, the maximum sum rate (8) in the dual MAC is also achieved in the BC. While the rate expression in (15) at first glance seems to require the joint detection of all the streams belonging to user k, the application of the simple matched filter receiver

$$\boldsymbol{B}_k = \boldsymbol{D}_k^{-1}\boldsymbol{W}_k^{\mathrm{H}}$$

to the receive signal $\boldsymbol{y}_k$ of every user k completely decorrelates the individual streams of user k. As a consequence, the individual streams of every user can be detected separately without any loss in rate, and the SINR of the ith stream belonging to user k reads as

$$\mathrm{SINR}_{k,i} = \frac{P_{\mathrm{Tx}}}{r} \cdot \frac{1}{[\boldsymbol{D}_k^2]_{i,i}}$$

entailing the rate $R_{k,i} = \log_2(1+\mathrm{SINR}_{k,i})$. Summing up the rates of all r_k streams of user k, we obtain

$$R_k = \sum_{i=1}^{r_k} R_{k,i} = \log_2 \prod_{i=1}^{r_k}\left(1 + \frac{P_{\mathrm{Tx}}}{r} \cdot \frac{1}{[\boldsymbol{D}_k^2]_{i,i}}\right)$$

in accordance to (15), and the rate R_k of user k asymptotically converges to

$$\begin{aligned} R_k &\cong r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2 |\boldsymbol{D}_k^2| \\ &= r_k \log_2 P_{\mathrm{Tx}} - r_k \log_2 r - \log_2\left|\boldsymbol{E}_k^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}\boldsymbol{E}_k\right|. \end{aligned}$$

Therefore, stream-wise detection achieves the same user rates for our special choice of the decorrelation matrices $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K$. While the transmission chain consisting of the precoder and the channel leads to a block-diagonalization, the inclusion of the matched filter receiver even leads to a total diagonalization of the transmission chain. Therefore, a joint diagonalization of the transmission chain by both the transmit filter and the matched filter receiver with appropriately chosen decorrelation matrices is asymptotically optimum in the same way as the block-diagonalization of the channel by the sender alone is.

see the program

5. CONCLUSION

We have shown that block-diagonalization is the asymptotically optimum transmission strategy in the broadcast channel and derived the asymptotic sum capacity when linear filtering is applied instead of dirty paper coding. Starting in the dual multiple access channel, we have found a very simple closed form solution for the optimum transmit covariance matrix which has then been converted back to the downlink by our recent rate duality. To the best of our knowledge, we have presented the first precoder-free asymptotic sum capacity expression of linear filtering which is only a function of the individual users' channel matrices, the available transmit power, and the antenna profile of the users.

6. REFERENCES

[1] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, February 2004.

[2] L.-U. Choi and R. D. Murch, "A Transmit Preprocessing Technique for Multiuser MIMO Systems Using a Decomposition Approach," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 20–24, January 2004.

[3] P. Komulainen, M. Latva-aho, and M. Juntti, "Block Diagonalization for Mulituser MIMO TDD Downlink and Uplink in Time-Varying Channel," in *Proc. ITG/IEEE WSA 2008*, February 2008, pp. 74–81.

[4] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low Complexity User Selection Algorithms for Multiuser MIMO Systems With Block Diagonalization," *IEEE Trans. on Signal Processing*, vol. 54, no. 9, pp. 3658–3663, September 2006.

[5] Z. Pan, K.-K. Wong, and T.-S. Ng, "Generalized Multiuser Orthogonal Space Division Multiplexing," *IEEE Trans. on Wireless Communications*, vol. 3, no. 6, pp. 1969–1973, November 2004.

[6] R. Hunger and M. Joham, "A General Rate Duality of the MIMO Multiple Access Channel and the MIMO Broadcast Channel," in *Global Telecommunications Conference (Globecom '08)*, November 2008.

[7] J. Salo, P. Suvikunnas, H. M. El-Sallabi, and P. Vainikainen, "Some results on MIMO mutual information: the high SNR case," in *Global Telecommunications Conference (Globecom '04)*, vol. 2, December 2004, pp. 943–947.

[8] J. Salo and P. Suvikunnas and H. M. El-Sallabi and P. Vainikainen, "Some Insights into MIMO Mutual Information: The High SNR Case," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 2997–3001, November 2006.

[9] N. Prasad and M. K. Varanasi, "Throughput analysis for MIMO systems in the high SNR regime," in *International Symposium on Information Theory (ISIT)*, July 2006, pp. 1954–1958.

[10] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Sum Capacity of Multiuser MIMO Broadcast Channels with Block Diagonalization," in *International Symposium on Information Theory (ISIT)*, July 2006, pp. 886–890.

[11] J. Lee and N. Jindal, "Dirty Paper Coding vs. Linear Precoding for MIMO Broadcast Channels," in *40th Asilomar Conference on Signals, Systems, and Computers (Asilomar 2006)*, October 2006, pp. 779–783.

[12] ——, "High SNR Analysis for MIMO Broadcast Channels: Dirty Paper Coding Versus Linear Precoding," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4787–4792, December 2007.

[13] N. Jindal, "High SNR Analysis of MIMO Broadcast Channels," in *International Symposium on Information Theory (ISIT 2005)*, September 2005, pp. 2310–2314.

[14] S. Shamai and S. Verdú, "The Impact of Frequency-Flat Fading on the Spectral Efficiency of CDMA," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[15] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR Power Offset in Multiantenna Communication," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4134–4151, December 2005.

[16] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, November/December 1999.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[18] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, ser. Mathematics in Science and Engineering, R. Bellman, Ed. Academic Press, 1979, vol. 143.

see the program

TOMLINSON HARASHIMA PRECODING FOR MIMO SYSTEMS WITH LOW RESOLUTION D/A-CONVERTERS

Amine Mezghani, Rafik Ghiat and Josef A. Nossek

Institute for Circuit Theory and Signal Processing
Technische Universität München, 80290 Munich, Germany
{mezghani,ghiat,nossek}@nws.ei.tum.de

ABSTRACT

We study the joint transmitter optimization for the flat multi-input multi-output (MIMO) channel under nonlinear distortion from the digital-to-analog converters (DACs). Our design is based on a *minimum mean square error* (MMSE) approach, taking into account the effects of the transmitter nonlinearities. The derivation does not make use of the assumption of uncorrelated white distortion (quantization) errors and considers the correlations of the quantization error with the other signals of the system. Through simulation, we compare the new optimized Tomlinson Harashima precoder to previously proposed transmitter designs when operating under DACs in terms of uncoded BER.

1. INTRODUCTION

The use of multiple antennas at both sides of the transmission link (MIMO systems) can improve the communication performance dramatically. If the channel state information is available at the transmitter and the receiver(s) have to be of low complexity, then transmit processing becomes advantageous. An other application of transmit processing is the broadcast scenario where no cooperation is possible between the receivers. The transmit filter will be then designed based on the knowledge of the a priori defined receiver (simple scalar receiver) and the Channel State Information (CSI). It is well known that nonlinear processing strategies can clearly improve the performance. In [1, 2] transmit filters combined with Tomlinson Harashima Precoding (THP) [3] were derived using a *minimum mean square error* (MMSE or Wiener) approach, where the receiver is restricted to be a simple common weight. However, most of these contributions on transmitter design for MIMO systems assume that the transmitted signals do not experience any nonlinear distortion. In practice, however, the signal undergoes many nonlinearities (e.g. Digital-to-Analog Conversion (DAC), Power Amplifier,...). For instance, the D/A-Conversion can be also regarded as a kind of quantization since it requires rounding the data to a suitable level of precision to be converted into the analog domain. The effect of the digital-to-analog conversion on the transmitter designs has been neglected by the research community up to now. This is due to the fact that the DAC is a non-linear operation which complicates theoretical analysis. In high speed applications, the DACs have to be of low resolution in order to save power and area. In this case the proposed design do not perform well when operating under coarse quantization. In fact, in order to reduce circuit complexity and save power and area, low resolution DACs have to be employed [4]. Therefore, the proposed designs do not necessarily have good performance when operating under DACs in a real system. Motivated by the same approach as in our recent works [5, 6], which concerns the linear and nonlinear MMSE receivers operating on quantized data under A/D conversion at the receiver side, we modify the Wiener (or MMSE) THP transmitter from [2] for the transmit-quantized flat MIMO channel (later denoted by WFQ-THP), taking into account the presence of the DACs. Under the choice of an optimally designed DAC we evaluate the resulting MSE between the estimated and the transmitted symbols and we minimize it subject to a Wiener THP transmitter and a common scalar receiver. In our model we assume perfect channel state information (CSI) at the transmitter.

2. SYSTEM MODEL

We consider a MIMO Gaussian channel where the transmitter employs N antennas and the receiver has B antennas (or users). Fig. 1 shows the general form of a THP quantized MIMO system, where $\boldsymbol{H} \in \mathbb{C}^{B\times N}$ is the channel matrix. The vector $\boldsymbol{s} \in \mathbb{C}^{B}$ comprises the M transmitted symbols, which are uncorrelated and have zero-mean and covariance matrix $\boldsymbol{R}_{ss} = \mathrm{E}[\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}] = \sigma_s^2\mathbf{1}$. The vector $\boldsymbol{\eta}$ refers to zero-mean complex circular Gaussian noise with covariance $\boldsymbol{R}_{\eta\eta} = \sigma_\eta^2\mathbf{1}$. Furthermore, $\boldsymbol{y} \in \mathbb{C}^{N}$ is the unquantized transmit signal.

In addition to the feedforward filter denoted by $\boldsymbol{P} \in \mathbb{C}^{M\times B}$ the transmitter is extended with a $modulo$ device $\mathrm{M}(\cdot)$ and a spatial feedback filter $\boldsymbol{F}$, where

$$\boldsymbol{P} = [\boldsymbol{p}_1 \cdots \boldsymbol{p}_B] \in \mathbb{C}^{N\times B}, \quad \text{and} \tag{1}$$

$$\boldsymbol{F} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ f_{2,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ f_{B,1} & \dots & f_{B,B-1} & 0 \end{bmatrix} \in \mathbb{C}^{B\times B}. \tag{2}$$

see the program

In other words, $\boldsymbol{p}_k$ is the k-th column of $\boldsymbol{P}$ and $f_{k,j}$ is the entry at the k-th row and j-th column ($k > j$) of the feedback matrix $\boldsymbol{F}$.
The data streams (or users) are successively encoded with order π, i.e. stream π_k (with order k) sees the interference caused by the data streams (or users) $\pi_{k+1} \dots \pi_B$. Then we construct the unquantized transmit signal

$$\boldsymbol{y} = \boldsymbol{P}\boldsymbol{u}, \tag{3}$$

where the outputs of the *modulo* devices are computed successively as

$$u_k = \big(s_{\pi_k} - \sum_{j<k} f_{k,j} u_j\big) \operatorname{mod}(\tau). \tag{4}$$

Through the DACs, the real parts $y_{i,R}$ and the imaginary parts $y_{i,I}$ of the transmit signals y_i, $1 \le i \le N$, are each quantized by a b-bit resolution scalar quantizer. Thus, the resulting analog signals read as

$$r_{i,l} = Q(y_{i,l}) = y_{i,l} + q_{i,l},\ l \in \{R, I\},\ 1 \le i \le N, \tag{5}$$

where $Q(\cdot)$ denotes the quantization operation and $q_{i,l}$ is the resulting quantization error. In the transmit processing architecture of Fig. 1, the receiver is restricted to a simple signal scaling with a common weight g. The matrices $\boldsymbol{P}$, $\boldsymbol{F}$ and the weight g, together, delivers each estimate $\hat{s}_{\pi k}$ as

$$\hat{s}_{\pi_k} = \big(g \boldsymbol{h}_{\pi_k}^{\mathrm{T}} \boldsymbol{r} + \eta_{\pi_k}\big) \operatorname{mod}(\tau), \tag{6}$$

$$\text{with, } \boldsymbol{r} = Q(\boldsymbol{y}) = \boldsymbol{y} + \boldsymbol{q}, \tag{7}$$

where $\boldsymbol{h}_{\pi_k}$ is the π_k-th row vector of the channel matrix $\boldsymbol{H}$. Our aim is to choose the DAC stepsizes, the matrix $\boldsymbol{P}$, and the scalar g, minimizing the MSE $=\mathrm{E}[\|\hat{\boldsymbol{s}} - \boldsymbol{s}\|_2^2]$, taking into account the quantization effect. Throughout this paper, $r_{\alpha\beta}$ denotes $\mathrm{E}[\alpha\beta^*]$. The operators $(\bullet)^{\mathrm{T}}$, $(\bullet)^{\mathrm{H}}$, $(\bullet)^*$, $\mathrm{Re}(\bullet)$, $\mathrm{Im}(\bullet)$ stand for transpose, Hermitian transpose, complex conjugate, real and imaginary parts of a complex number, respectively; $\mathbf{e}_k$ is the k-th column of the $B \times B$ identity matrix $\mathbf{1}_B$.

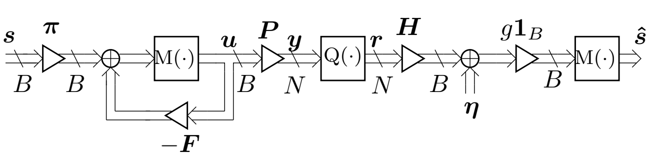


Fig. 1. MIMO-Quantizer THP Transmit System

3. REVIEW OF THE SPATIAL MMSE-THP

In this section, we review the MMSE-THP (or Wiener Filter-THP) algorithm [2], while ignoring the quantization (i.e. $\boldsymbol{r} \equiv \boldsymbol{y}$). Neglecting the *modulo-loss*, the error signal at the k-th encoding step can be approximated by

$$e_{\pi_k} = -u_k + g\boldsymbol{h}_{\pi_k}^{\mathrm{T}} \boldsymbol{y} - \sum_{j<k} f_{k,j} u_j + g\eta_{\pi_k}. \tag{8}$$

Then, under the assumption that the outputs of the *modulo* devices u_k are i.i.d. with variance σ_u^2, it is easy to show by the KKT conditions that, for given ordering π, the optimal precoding and feedback matrices minimizing the sum-MSE$=\sum_k \mathrm{E}[|e_k|^2]$ read as

$$\begin{aligned}
f_{k,j} &= g\boldsymbol{h}_{\pi_k}^{\mathrm{T}} \boldsymbol{p}_j, \\
\boldsymbol{p}_k &= \frac{1}{g}\Big(\sum_{j\le k} \boldsymbol{h}_{\pi_j}^* \boldsymbol{h}_{\pi_j}^{\mathrm{T}} + \xi \mathbf{1}_N\Big)^{-1} \boldsymbol{h}_{\pi_k}^*, \\
g &= \sqrt{\frac{\sigma_u^2}{E_{\mathrm{tr}}} \sum_k \boldsymbol{h}_{\pi_k}^{\mathrm{T}} \Big(\sum_{j\le k} \boldsymbol{h}_{\pi_j}^* \boldsymbol{h}_{\pi_j}^{\mathrm{T}} + \xi \mathbf{1}_N\Big)^{-2} \boldsymbol{h}_{\pi_k}^*},
\end{aligned} \tag{9}$$

where $\xi = \frac{\mathrm{tr}(\boldsymbol{R}_{\eta\eta})}{E_{tr}}$. With these optimum matrices, the sum *mean square error* reads as

$$\mathrm{MSE} = \sigma_u^2 \Big[B - \sum_k \boldsymbol{h}_{\pi_k}^{\mathrm{T}} \Big(\sum_{j\le k} \boldsymbol{h}_{\pi j}^* \boldsymbol{h}_{\pi j}^{\mathrm{T}} + \xi \mathbf{1}_N\Big)^{-1} \boldsymbol{h}_{\pi_k}^*\Big]. \tag{10}$$

Now, the problem of finding the optimal ordering π minimizing the sum-MSE is NP hard, since we must check all $B!$ possible permutations. To reduce the complexity, we minimize each summand in (10) successively, i. e., π_k is chosen under the assumption that $\pi_{k+1}, \dots, \pi_B$ are fixed

$$\pi_k = \underset{i \notin \{\pi_{k+1},\dots,\pi_B\}}{\operatorname{argmax}} \boldsymbol{h}_i^{\mathrm{T}} \Big(\sum_{j \notin \{\pi_{k+1},\dots,\pi_B\}} \boldsymbol{h}_j^* \boldsymbol{h}_j^{\mathrm{T}} + \xi \mathbf{1}_N\Big)^{-1} \boldsymbol{h}_i^*. \tag{11}$$

4. DAC OPTIMIZATION AND CHARACTERIZATION

D/A-conversion can be viewed as a quantization process since the resolution of the values of $\boldsymbol{y}$ has to be reduced to a smaller one, let say b bits, to be transformed into the analog domain, which is inevitable for high speed transmission. This operation can be characterized by a distortion factor $\rho_q^{(i,l)}$ to indicate the relative amount of quantization error generated, which is defined as follows

$$\rho_q^{(i,l)} = \frac{\mathrm{E}[q_{i,l}^2]}{r_{y_{i,l} y_{i,l}}}, \tag{12}$$

where $r_{y_{i,l} y_{i,l}} = \mathrm{E}[y_{i,l}^2]$ is the variance of $y_{i,l}$ and the distortion factor $\rho_q^{(i,l)}$ depends on the number of quantization bits b, the quantizer type (uniform or non-uniform) and the probability density function of $y_{i,l}$. Note that the signal-to-quantization noise ratio (SQNR) has an inverse relationship with regard to the distortion factor

$$\mathrm{SQNR}^{(i,l)} = \frac{1}{\rho_q^{(i,l)}}. \tag{13}$$

see the program

Similar to our work [5], the quantizer design is based on minimizing the *mean square error* (distortion) between the input $y_{i,l}$ and the output $r_{i,l}$ of each quantizer. In other words, the SQNR values are maximized. Under this optimal design of the scalar finite resolution quantizer, whether uniform or not, the following equations hold for all $0 \leq i \leq N$, $l \in \{R, I\}$ [7, 8]

$$\mathrm{E}[q_{i,l}] = 0 \tag{14}$$
$$\mathrm{E}[r_{i,l}q_{i,l}] = 0 \tag{15}$$
$$\mathrm{E}[y_{i,l}q_{i,l}] = -\rho_q^{(i,l)} r_{y_{i,l}y_{i,l}}. \tag{16}$$

Obviously, (16) follows from (12) and (15). For the uniform quantizer case, (14) holds only if the probability density function of $y_{i,l}$ is even.
Assuming a dense matrix $\boldsymbol{P}$ and for large number of antennas, the DAC input signals $y_{i,l}$ are weighted sums of weakly dependent random symbols and thus, due to central limit theorem, they are approximately Gaussian distributed. Besides, they undergo nearly the same distortion factor ρ_q, i.e., $\rho_q^{(i,l)} = \rho_q \; \forall i \forall l$. Furthermore, the optimal parameters of the uniform as well as the non-uniform quantizer and the resulting distortion factor ρ_q for Gaussian distributed signal are tabulated in [7] for different bit resolutions b. Recent research work on optimally quantizing the Gaussian source can be found in [9, 10].

Now, let $q_i = q_{i,R} + \mathrm{j}q_{i,I}$ be the complex quantization error. Under the assumption of uncorrelated real and imaginary part of y_i, we easily obtain

$$r_{q_iq_i} = \mathrm{E}[q_iq_i^*] = \rho_q r_{y_iy_i}, \text{ and } r_{y_iq_i} = \mathrm{E}[y_iq_i^*] = -\rho_q r_{y_iy_i}. \tag{17}$$

For the uniform quantizer case, it was shown in [10], that the optimal quantization step Δ for a Gaussian source decreases as $\sqrt{b}2^{-b}$ and that ρ_q is asymptotically well approximated by $\frac{\Delta^2}{12}$ and decreases as $b2^{-2b}$. On the other hand, the optimal non-uniform quantizer achieves, under high-resolution assumption, approximately the following distortion [11]

$$\rho_q \approx \frac{\pi\sqrt{3}}{2} 2^{-2b}. \tag{18}$$

This particular choice of the (non-)uniform scalar quantizer minimizing the distortion between $\boldsymbol{r}$ and $\boldsymbol{y}$ for the DACs, combined with the transmitter of the next section, is also optimal with respect to the total MSE between the transmitted symbol vector $\boldsymbol{s}$ and the estimated symbol vector $\hat{\boldsymbol{s}}$, as we will see later.

5. NEARLY OPTIMAL QUANTIZED THP-TRANSMITTER

In this section, we optimize the transmit matrix $\boldsymbol{P}$ and the scalar receiver g based on the MMSE criterion, taking into account the quantization process. To this end, we minimizes the sum-MSE under the power constraint $\mathrm{E}\left[\|\boldsymbol{r}\|_2^2\right] \leqslant E_{\mathrm{tr}}$

$$\min_{\{\boldsymbol{P},\boldsymbol{F},g\}} \mathrm{MSE}(\boldsymbol{P},\boldsymbol{F},g) \quad \text{s.t.:} \quad \mathrm{E}\left[\|\boldsymbol{r}\|_2^2\right] \leqslant E_{\mathrm{tr}}, \tag{19}$$

with

$$\mathrm{E}\left[\|\boldsymbol{r}\|_2^2\right] = \mathrm{tr}(\boldsymbol{R}_{rr}).$$

For this optimization we establish the Lagrangian function to find the global minimum

$$\mathrm{L}(\boldsymbol{P},g,\lambda) = \mathrm{MSE}(\boldsymbol{P},g) + \lambda\Big(\mathrm{tr}(\boldsymbol{R}_{rr}) - E_{\mathrm{tr}}\Big) \quad \text{with} \quad \lambda \in \mathbb{R}^+. \tag{20}$$

As the u_is are results of *modulo*-operations, they are nearly independent and have approximately a uniform distribution in the complex square with side τ. Thus we assume that the power of u_i is $\sigma_u^2 = \frac{\tau^2}{6}$. Let us first consider the received symbol $\hat{s}_{\pi_k}$ of the π_k-th user

$$\hat{s}_{\pi_k} = \left(g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{r} + g\eta_{\pi_k}\right) \mathrm{mod}(\tau).$$

After adding and subtracting the term $\sum_{j<k} f_{k,j}u_j$ we obtain[1]

$$\begin{aligned}\hat{s}_{\pi_k} &= \Big(\sum_{j>k} f_{k,j}u_j + g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{r} - \sum_{j<k} f_{k,j}u_j + g\eta_{\pi_k}\Big) \mathrm{mod}(\tau) \\ &= \Big(s_{\pi_k} - u_k + g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{r} - \sum_{j<k} f_{k,j}u_j + g\eta_{\pi_k}\Big) \mathrm{mod}(\tau).\end{aligned}$$

When we neglect the modulo-loss, the error $e_{\pi_k} = \hat{s}_{\pi_k} - s_{\pi_k}$ is nearly given by

$$e_{\pi_k} = -u_k + g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{r} - \sum_{j<k} f_{k,j}u_j + g\eta_{\pi_k}. \tag{21}$$

In order to compute each individual *mean square error* $\varepsilon_{\pi_k} = \mathrm{E}[e_{\pi_k}e_{\pi_k}^*]$, we assume that the u_k have the variance σ_u^2 and are mutually uncorrelated. Furthermore, we consider the correlation of the quantization error with the other signal of the system by means of (21)

$$\begin{aligned}\varepsilon_{\pi_k} = &|g|^2\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{R}_{rr}\boldsymbol{h}_{\pi_k}^* - 2\mathrm{Re}\{g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{R}_{ru}\mathbf{e}_k\} \\ &- 2\mathrm{Re}\{g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{R}_{ru}\sum_{j<k}\mathbf{e}_j f_{k,j}^*\} + \sigma_u^2(1 + \sum_{j<k}|f_{k,j}|^2) + |g|^2\sigma_\eta^2.\end{aligned} \tag{22}$$

with the correlation matrix

$$\boldsymbol{R}_{ru} = \mathrm{E}[\boldsymbol{r}\boldsymbol{u}^{\mathrm{H}}] = \mathrm{E}[(\boldsymbol{y}+\boldsymbol{q})\boldsymbol{u}^{\mathrm{H}}] = \boldsymbol{R}_{yu} + \boldsymbol{R}_{qu}, \tag{23}$$

and $\boldsymbol{R}_{rr}$ the covariance matrix of the quantized signal given by

$$\boldsymbol{R}_{rr} = \mathrm{E}[(\boldsymbol{y}+\boldsymbol{q})(\boldsymbol{y}+\boldsymbol{q})^{\mathrm{H}}] = \boldsymbol{R}_{yy} + \boldsymbol{R}_{yq} + \boldsymbol{R}_{yq}^{\mathrm{H}} + \boldsymbol{R}_{qq}. \tag{24}$$

5.1. Derivation of the Covariance Matrices involving the Quantization Error

Before investigating the MMSE optimization, we first derive all needed covariance matrices by using the fact that the quantization error q_i, conditioned on y_i, is statistically independent from

[1] Note that $(s_{\pi_k} - u_k)\,\mathrm{mod}(\tau) = (\sum_{j<k} f_{k,j}u_j)\,\mathrm{mod}(\tau)$ and $(a+b)\bmod \tau = (a \bmod \tau + b \bmod \tau) \bmod \tau$

all other random variables of the system.
First we calculate $r_{y_i q_j} = \mathrm{E}[y_i q_j^*]$ for $i \neq j$

$$\begin{aligned}\mathrm{E}[y_i q_j^*] &= \mathrm{E}_{y_j}[\mathrm{E}[y_i q_j^*|y_j]] \\ &= \mathrm{E}_{y_j}[\mathrm{E}[y_i|y_j]\mathrm{E}[q_j^*|y_j]] \\ &\approx \mathrm{E}_{y_j}\Big[r_{y_i y_j} r_{y_j y_j}^{-1} y_j \mathrm{E}[q_j^*|y_j]\Big] \qquad (25) \\ &= r_{y_i y_j} r_{y_j y_j}^{-1} \mathrm{E}[y_j q_j^*] \\ &= -\rho_q r_{y_i y_j}. \qquad (26)\end{aligned}$$

In (25), we approximate the Bayesian estimator $\mathrm{E}[y_i|y_j]$ with the linear estimator $r_{y_i y_j} r_{y_j y_j}^{-1} y_j$, which holds with equality if the vector $\boldsymbol{y}$ is jointly Gaussian distributed. Eq. (26) follows from (17). Summarizing the results of (17) and (26), we obtain

$$\boldsymbol{R}_{yq} \approx -\rho_q \boldsymbol{R}_{yy}. \qquad (27)$$

Similarly, we evaluate $r_{q_i q_j}$ for $i \neq j$ to end up in

$$\mathrm{E}[q_i q_j^*] = \mathrm{E}_{y_j}[\mathrm{E}[q_i q_j^*|y_j]] \approx \rho_q^2 r_{y_j y_i}^* = \rho_q^2 r_{y_i y_j}, \qquad (28)$$

where we used (27) and (17). From (28) and (17) we deduce the covariance matrix of the quantization error

$$\begin{aligned}\boldsymbol{R}_{qq} &\approx \rho_q \mathrm{diag}(\boldsymbol{R}_{yy}) + \rho_q^2 \mathrm{nondiag}(\boldsymbol{R}_{yy}) \\ &= \rho_q \boldsymbol{R}_{yy} - (1-\rho_q)\rho_q \mathrm{nondiag}(\boldsymbol{R}_{yy}),\end{aligned} \qquad (29)$$

with $\mathrm{diag}(\boldsymbol{A})$ denotes a diagonal matrix containing only the diagonal elements of $\boldsymbol{A}$ and $\mathrm{nondiag}(\boldsymbol{A}) = \boldsymbol{A} - \mathrm{diag}(\boldsymbol{A})$. Inserting the expressions (27) and (29) into (24), we obtain

$$\begin{aligned}\boldsymbol{R}_{rr} &\approx (1-\rho_q)(\boldsymbol{R}_{yy} - \rho_q \mathrm{nondiag}(\boldsymbol{R}_{yy})). \\ &= (1-\rho_q)^2 \boldsymbol{R}_{yy} + \rho_q(1-\rho_q)\mathrm{diag}(\boldsymbol{R}_{yy})\end{aligned} \qquad (30)$$

In a very similar way, we get the covariance matrix $\boldsymbol{R}_{qu} = \mathrm{E}[\boldsymbol{q}\boldsymbol{u}^{\mathrm{H}}]$ as

$$\boldsymbol{R}_{qu} = \mathrm{E}[\boldsymbol{q}\boldsymbol{u}^{\mathrm{H}}] \approx -\rho_q \boldsymbol{R}_{yu}. \qquad (31)$$

Thus, Equation (23) becomes

$$\boldsymbol{R}_{ru} \approx (1-\rho_q)\boldsymbol{R}_{yu}. \qquad (32)$$

Finally, $\boldsymbol{R}_{yy}$ and $\boldsymbol{R}_{uy}$ can be easily obtained from our system model

$$\boldsymbol{R}_{yy} = \sigma_u^2 \boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}, \qquad (33)$$
$$\boldsymbol{R}_{yu} = \sigma_u^2 \boldsymbol{P}. \qquad (34)$$

5.2. Derivation of the TxWFQ transmitter

Now, we return to our MMSE problem and insert these results into (22) to obtain

$$\begin{aligned}\varepsilon_{\pi_k} =& \sigma_u^2 \Big[\alpha|g|^2 \boldsymbol{h}_{\pi_k}^{\mathrm{T}}(\alpha \boldsymbol{P}\boldsymbol{P}^{\mathrm{H}} + \rho_q \mathrm{diag}(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}))\boldsymbol{h}_{\pi_k}^* - 2\alpha\mathrm{Re}\{g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{p}_k\} \\ &-2\alpha\mathrm{Re}\{g\boldsymbol{h}_{\pi_k}^{\mathrm{T}} \sum_{j<k} \boldsymbol{p}_j f_{k,j}^*\} + 1 + \sum_{j<k}|f_{k,j}|^2\Big] + |g|^2\sigma_\eta^2,\end{aligned} \qquad (35)$$

with $\alpha = 1 - \rho_q$.
Afterwards, we establish the Lagrangian function for the optimization (19) with the Lagrangian multiplier $\lambda \in \mathbb{R}_+$

$$\mathrm{L}(\boldsymbol{p}_{\pi_k}, g, f_{k,j}) = \sum_{j=1}^{B} \varepsilon_{\pi_k} + \lambda\Big(\sigma_u^2 \,\mathrm{tr}(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}) - \frac{E_{\mathrm{tr}}}{(1-\rho_q)}\Big). \qquad (36)$$

When differentiating this lagrangian from with respect to $f_{k,j}$, we obtain

$$\frac{\partial}{\partial f_{k,j}} \mathrm{L}(\boldsymbol{p}_{\pi_k}, g, f_{k,j}) = \alpha\sigma_u^2 g^* \boldsymbol{h}_{\pi_k}^{\mathrm{H}} \boldsymbol{p}_j^* + \sigma_u^2 f_{k,j}^*, \qquad (37)$$

which must be equal to zero in order to optimize the MSE. Thus we obtain the optimal feedback matrix

$$f_{k,j} = (1-\rho_q) g \boldsymbol{h}_{\pi_k}^{\mathrm{T}} \boldsymbol{p}_j. \qquad (38)$$

This delivers the following MSE

$$\begin{aligned}\varepsilon_{\pi_k} =& \sigma_u^2 \Big[\alpha|g|^2 \boldsymbol{h}_{\pi_k}^{\mathrm{T}} \Big(\alpha \sum_{j\geq k} \boldsymbol{p}_j \boldsymbol{p}_j^{\mathrm{H}} + \rho_q \mathrm{diag}(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}})\Big)\boldsymbol{h}_{\pi_k}^* \\ &-2\alpha\mathrm{Re}\{g\boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{p}_k\} + 1\Big] + |g|^2\sigma_\eta^2.\end{aligned} \qquad (39)$$

Then, we formulate the KKT equations with respect to g and $\boldsymbol{P}$

$$\begin{aligned}\frac{\partial \mathrm{L}(\boldsymbol{P}, g, \lambda)}{\partial g} =& \sigma_u^2 \sum_k \alpha \Big[g^* \boldsymbol{h}_{\pi_k}^{\mathrm{T}} \Big(\alpha \sum_{j\geq k} \boldsymbol{p}_j \boldsymbol{p}_j^{\mathrm{H}} \\ &+ \rho_q \mathrm{diag}(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}})\Big)\boldsymbol{h}_{\pi_k}^* - \boldsymbol{h}_{\pi_k}^{\mathrm{T}}\boldsymbol{p}_k\Big] + g^*\sigma_\eta^2 B \overset{!}{=} 0,\end{aligned} \qquad (40)$$

$$\begin{aligned}\frac{\partial \mathrm{L}(\boldsymbol{P}, g, \lambda)}{\partial \boldsymbol{p}_k} =& \sigma_u^2 \alpha \Big[|g|^2 (\alpha \sum_{j\leq k} \boldsymbol{h}_{\pi_j} \boldsymbol{h}_{\pi_j}^{\mathrm{H}} + \rho_q \mathrm{diag}(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}^*))\boldsymbol{p}_k^* \\ &- g\boldsymbol{h}_{\pi_k} + \lambda\sigma_u^2 \boldsymbol{p}_k^*\Big] \overset{!}{=} 0,\end{aligned} \qquad (41)$$

where we used the following identity for two matrices $\boldsymbol{A}, \boldsymbol{B}$

$$\mathrm{tr}(\boldsymbol{A}\,\mathrm{diag}(\boldsymbol{B})) = \mathrm{tr}(\boldsymbol{B}\,\mathrm{diag}(\boldsymbol{A})). \qquad (42)$$

Summing the complex conjugate of the derivative with respect to $\boldsymbol{p}_k$ (41), then multiplying it by $\boldsymbol{P}^{\mathrm{H}}$ from the right followed by the trace operation, and finally comparing it with (40) multiplied by g, we can show the following relation

$$\lambda = |g|^2 \frac{B\sigma_\eta^2}{\alpha\sigma_u^2 \,\mathrm{tr}(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}})} = |g|^2 \frac{\mathrm{tr}(\boldsymbol{R}_{\eta\eta})}{E_{\mathrm{tr}}} = |g|^2 \xi.$$

Considering the derivative with respect to $\boldsymbol{P}$ (41) and above expression for the Lagrangian multiplier λ, the filter can be derived as

$$\boldsymbol{p}_k = \frac{1}{g}\Big((1-\rho_q)\sum_{j\leq k} \boldsymbol{h}_{\pi_j}^* \boldsymbol{h}_{\pi_j}^{\mathrm{T}} + \rho_q \,\mathrm{diag}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}) + \xi \boldsymbol{1}_N\Big)^{-1} \boldsymbol{h}_{\pi_k}^* \qquad (43)$$

see the program

Note that the inverse always exists independently of $\boldsymbol{H}$. Due to the transmit energy constraint we finally get

$$g=\sqrt{\frac{\sigma_u^2(1-\rho_q)}{E_{\text{tr}}}}\cdot \sqrt{\sum_k \boldsymbol{h}_{\pi_k}^{\mathrm{T}}\Big((1-\rho_q)\sum_{j\le k}\boldsymbol{h}_{\pi_j}^{*}\boldsymbol{h}_{\pi_j}^{\mathrm{T}}+\rho_q\,\mathrm{diag}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})+\xi\mathbf{1}_N\Big)^{-2}\boldsymbol{h}_{\pi_k}^{*}}, \tag{44}$$

Notice that, when we set $\rho_q = 0$ (infinite precision DACs) in the expressions of $\boldsymbol{P}$ and g, we obtain exactly the same expressions as derived in (9) for the unquantized system.
Using (39), (43) and (44), the sum-MSE resulting from the optimal design of the receiver filter and feedback matrix becomes after some computations

$$\mathrm{MSE} = \sum_k \varepsilon_k = \sigma_u^2\Big[B-(1-\rho_q)\sum_k \boldsymbol{h}_{\pi_k}^{\mathrm{T}}\Big((1-\rho_q)\sum_{j\le k}\boldsymbol{h}_{\pi j}^{*}\boldsymbol{h}_{\pi j}^{\mathrm{T}} + \rho_q\mathrm{diag}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})+\xi\mathbf{1}_N\Big)^{-1}\boldsymbol{h}_{\pi_k}^{*}\Big]. \tag{45}$$

Now, the problem of finding the optimal ordering π minimizing the sum-MSE ($\sum_k \varepsilon_k$) in (45) is NP hard, since we must check all $B!$ possible permutations. To reduce the complexity, we minimize each summand of the sum-MSE successively, i. e., π_k is chosen under the assumption that $\pi_B,\ldots,\pi_{k+1}$ are fixed

$$\pi_k = \underset{i\neq\pi_{k+1}\ldots\pi_B}{\operatorname{argmax}}\ \boldsymbol{h}_i^{\mathrm{T}}\Big((1-\rho_q)\sum_{j\neq\pi_{k+1}\ldots\pi_B}\boldsymbol{h}_j^{*}\boldsymbol{h}_j^{\mathrm{T}}+\rho_q\mathrm{diag}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})+\xi\mathbf{1}_N\Big)^{-1}\boldsymbol{h}_i^{*}. \tag{46}$$

5.3. Effects of Quantization on the MSE

In order to verify whether the chosen quantizer minimizes the MSE of our system, we examine the first derivative of the MSE in (45) with respect to ρ_q

$$\frac{\partial\mathrm{MSE}}{\partial\rho_q} = g^2\sigma_u^2\sum_k \boldsymbol{p}_k^{\mathrm{H}}(\mathrm{diag}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})+\frac{\mathrm{tr}(\boldsymbol{R}_{\eta\eta})}{E_{\text{tr}}}\mathbf{1}_N)\boldsymbol{p}_k,$$

where we used (45). Obviously, the derivative of the MSE with respect to the distortion is positive. Therefore, the $\mathrm{MSE}^{\mathrm{WFQ-THP}}$ achieved by THP transmit processing is monotonically increasing with respect to ρ_q. Since we chose the quantizer to minimize the distortion factor ρ_q, our receiver quantizer designs are jointly optimum with respect to the total MSE.

Now, we expand the MSE expression (45) into a Taylor series around $\rho_q = 0$ up to the order one, to get an approximation of the MSE

$$\mathrm{MSE}^{\text{WFQ-THP}} \approx \mathrm{MSE}^{\text{WF-THP}} + \rho_q g^2\sigma_u^2\,\mathrm{tr}(\boldsymbol{P}^{\mathrm{H}}(\mathrm{diag}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})+\frac{\mathrm{tr}(\boldsymbol{R}_{\eta\eta})}{E_{\text{tr}}}\mathbf{1}_N)\boldsymbol{P}), \tag{47}$$

where $\mathrm{MSE}^{\text{WF-THP}} = \mathrm{MSE}^{\text{WFQ-THP}}|_{\rho_q=0}$ is the achievable MSE without quantization given in (10). The second term gives the increase in the MSE due to the quantization as a function of ρ_q and the channel parameters. It reveals also the residual error at infinite SNR.

6. SIMULATION RESULTS

The performance of the modified THP transmitter for a 4-bit quantized output MIMO system (WFQ-THP), in terms of BER averaged over 10^6 channel realizations, is shown in Fig. 2 for a 10×10 MIMO system (QPSK), compared with the conventional Tomlinson Harashima Precoder (WF-THP) from [2]. The symbols and the noise are assumed to be uncorrelated, that is $\boldsymbol{R}_{ss} = \sigma_s^2\mathbf{1}$ and $\boldsymbol{R}_{\eta\eta} = \sigma_\eta^2\mathbf{1}$. Besides, the BER curve of the modified linear MMSE transmitter from [12] is shown. Thereby, the entries of $\boldsymbol{H}$ are complex-valued realization of independent zero-mean Gaussian random variables with unit variance. Clearly, the modified transmitter outperforms the conventional ones at high SNR. This is because the effect of quantization error is more pronounced at higher SNR values when compared to the additive Gaussian noise variance. Since the conventional THP-precoder looses its regularized structure at high SNR values, its performance degrades asymptotically, when operating under DACs. For comparison, we also plotted the BER curve for the THP-precoder transmitter, if no quantization is applied (infinite precision DAC).

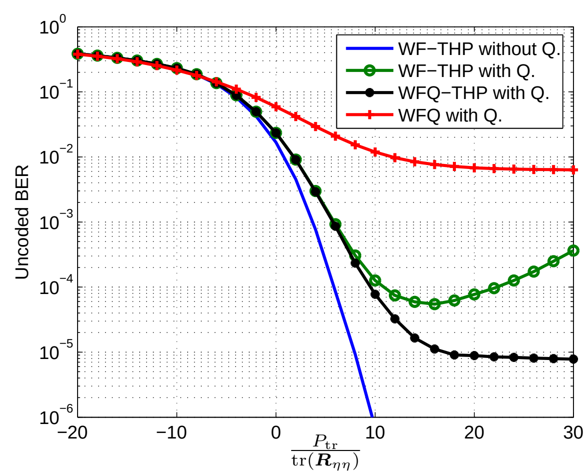


Fig. 2. WFQ-THP vs. WFQ, WF-THP with quantization and WF-THP without quantization; 4QAM modulation with $B = 10$, $N = 10$, 4-bit ($\rho_q = 0.01154$) uniform quantizer.

In Fig. 3, we consider a 4-bit quantized 4×4 MIMO configuration. In addition to the BER curve of the proposed design,

see the program

it shows the simulated BER with a THP transmitter considering the quantization error $\boldsymbol{q}$ as an additive white noise, which is uncorrelated with the other signals of the system. For the additive quantization noise model, we take as effective noise covariance matrix (see first equality of (17))

$$\bar{\boldsymbol{R}}_{\eta\eta} = \boldsymbol{R}_{\eta\eta} + \boldsymbol{H}\bar{\boldsymbol{R}}_{qq}\boldsymbol{H}^{\mathrm{H}} = \boldsymbol{R}_{\eta\eta} + \rho_q\sigma_u^2\boldsymbol{H}\mathrm{diag}(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}})\boldsymbol{H}^{\mathrm{H}}. \tag{48}$$

Thus the well-known solution (9) can be adjusted to the following effective noise power

$$\mathrm{tr}(\bar{\boldsymbol{R}}_{\eta\eta}) \approx B\sigma_\eta^2 + \frac{\rho_q}{1-\rho_q}\frac{E_{\mathrm{tr}}}{N}\,\mathrm{tr}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}). \tag{49}$$

Obviously, this model, which has been commonly used in the literature is also outperformed by the presented design at any SNR level and independently of the resolution.

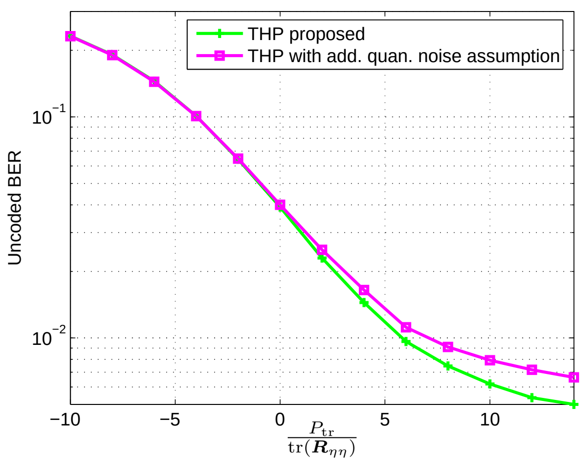


Fig. 3. THP proposed vs. the conventional THP [2] adjusted to an additive quantization noise term; 4QAM modulation with $B = 10$, $N = 10$, 3-bit (ρ_q = 0.03744) uniform quantizer.

7. CONCLUSION

We addressed the problem of designing a THP transmitter for MIMO channels under DAC nonlinearity. We provided an approximation for the mean squared error for each data stream, where the quantizer is optimized for a Gaussian input. Then, we proposed an optimized THP transmitter in terms of the MSE criteria that takes into account the quantization effects. The proposed precoder operating for quantized transmit signals shows better performance in terms of BER than simply applying the WF-THP (as well-known in literature [1, 2]) designed for an unquantized MIMO system. It also outperforms the most obvious approach of simply adjusting this WF-THP to an effective noise power of $B\sigma_\eta^2 + \frac{\rho_q}{1-\rho_q}\frac{E_{\mathrm{tr}}}{N}\,\mathrm{tr}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}})$. An essential aspect of our derivation is that we do not make the assumption of uncorrelated white quantization error. Moreover, our transmitter does not present any extra complexity from the implementation point of view.

8. REFERENCES

[1] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, “Space-time transmission using Tomlinson-Harashima precoding,” in *4th ITG Conference on Source and Channel Coding*, Jan. 2002.

[2] M. Joham, D. Schmidt, J. Brehmer, and W. Utschick, “Finite-length MMSE Tomlinson-Harashima precoding for frequency selective vector channels,” *IEEE Trans. on Sig. Proc.*, vol. 55, no. 6, pp. 3073–3088, June 2007.

[3] M. Tomlinson, “New automatic equaliser employing modulo arithmetic,” *Electronic Letters*, vol. 7, no. 5/6, pp. 138–139, March 1971.

[4] R. Schreier and G. C. Temes, “Understanding Delta-Sigma Data Converters,” *IEEE Computer Society Press*, 2004.

[5] A. Mezghani, M. S. Khoufi, and J. A. Nossek, “A Modified MMSE Receiver for Quantized MIMO Systems,” *In Proc. ITG/IEEE WSA, Vienna, Austria*, February 2007.

[6] A. Mezghani, M. S. Khoufi, and J. A. Nossek, “Spatial MIMO Decision Feedback equalizer Operating on Quantized Data,” *Submitted to ICASSP*, 2008.

[7] J. Max, “Quantizing for Minimum Distortion,” *IEEE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, March 1960.

[8] J. G. Proakis, *Digital Communications*, McGraw Hill, New York, third edition, 1995.

[9] N. Al-Dhahir and J. M. Cioffi, “On the Uniform ADC Bit Precision and Clip Level Computation for a Gaussian Signal,” *IEEE Trans. on Sig. Proc.*, vol. 44, no. 2, February 1996.

[10] D. Hui and D. L. Neuhoff, “Asymptotic Analysis of Optimal Fixed-Rate Uniform Scalar Quantization,” *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 957–977, March 2001.

[11] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Dordrecht, Niederlande, first edition, 1992.

[12] R. Ghiat, A. Mezghani, and J. A. Nossek, “Transmit Processing with Low Resolution D/A-Converters,” *in preparation*, 2009.

see the program

LINK ADAPTATION FOR MIMO-OFDM SYSTEMS USING MAXIMUM LIKELIHOOD DETECTOR

Tetsushi Abe

DOCOMO Euro-Labs
Munich, Germany
abe@docomolab-euro.com

Gerhard Bauch

DOCOMO Euro-Labs and Universität der Bundeswehr München
Munich, Germany
Gerhard.bauch@unibw.de

ABSTRACT

This paper investigates adaptive modulation and coding (AMC) in 2-Tx multiple-input and multiple output (MIMO)-orthogonal frequency division multiplexing (OFDM) spatial multiplexing systems using non-iterative maximum likelihood detector (MLD). We derive an approximate post-processing signal to interference-plus-noise ratio (SINR) for MLD, which is computed as post minimum mean-squared error (MMSE) SINR plus gain value, where the gain value is calculated by symbol error rate of the interfering stream. The post-MLD SINR is then applied to perform link adaptation for AMC. Simulation results show that the proposed link adaptation strategy achieves throughput performance close to that of the optimum approach.

1. INTRODUCTION

Multiple-input and multiple-output (MIMO)-orthogonal frequency division multiplexing (OFDM) spatial multiplexing is now an essential technology for future high data rate wireless communication systems. In the third generation partnership project (3GPP) long term evolution (LTE) system [1], 2-Tx MIMO is considered as a baseline configuration. Maximum likelihood detector (MLD) [2] is a promising receiver structure for MIMO-OFDM spatial multiplexing. Especially, MLD with no detection-and-decoding iteration would be a good option in initial deployments to achieve low decoding latency.

One practical issue to use MLD in MIMO systems is complexity reduction. There are significant amount of literatures that investigate this problem and provide various solutions that uses techniques, such as the sphere decoding and M-algorithm, e.g., [3–5].

Another practical issue regarding MLD, which has not yet been investigated as extensively as the complexity issue, is link adaptation for adaptive modulation and coding (AMC). In link adaptation, the receiver chooses the most preferable modulation and coding scheme (MCS) which is fed back to the transmitter. When the linear minimum mean-squared error (LMMSE) receiver is used, link adaptation is simple because the post-MMSE signal to interference-plus-noise ratio (SINR) can be explicitly computed for each spatial stream using channel estimates. However, when MLD is used, link adaptation is not straightforward because the exact post-processing SINR cannot be explicitly computed. One solution would be to use the pair-wise error probability (PEP) bound to estimate frame error rate (FER) [6]. However, in MIMO spatial multiplexing, link adaptation using PEP requires searching for multi-dimensional constellation points for multiple MCS combinations, which would become computationally expensive. Since AMC is an essential technology for modern wireless systems, such as LTE, alternative solutions could be investigated.

In this paper, we develop a suboptimum link adaptation scheme for 2-Tx MIMO-OFDM spatial multiplexing systems using non-iterative MLD. We derive an approximate post-MLD SINR for each spatial stream, which is then utilized to perform link adaptation for adaptive modulation and coding. The effectiveness of the proposed approach will be shown via computer simulations.

2. SYSTEM MODEL

Figure 1 shows the system model. This paper focuses on two transmit and N receive antenna ($N \geq 2$) MIMO-OFDM systems with K subcarriers where two independent spatial streams are transmitted simultaneously at each subcarrier. At the transmitter, independent coding and modulation is assumed for each spatial stream, and a common modulation and coding is applied for all subcarriers within a stream [1]. The receiver uses MLD for spatial signal separation followed by stream-wise channel decoders. We assume no iterative detection and decoding between MLD and channel decoders for low decoding latency. The received signal vector $\boldsymbol{r}_k \in \mathbb{C}^N$ at the k-th subcarrier ($k = 1 \dots K$) is defined as

$$\begin{aligned} \boldsymbol{r}_k &= \boldsymbol{h}_{k1} s_{k1} + \boldsymbol{h}_{k2} s_{k2} + \boldsymbol{n}_k && (1) \\ &= \boldsymbol{H}_k \boldsymbol{s}_k + \boldsymbol{n}_k && (2) \end{aligned}$$

where $\boldsymbol{h}_{kl} \in \mathbb{C}^N$ ($l = 1, 2$) is the spatial channel vector of the l-th stream. $\boldsymbol{n}_k$ is the AWGN vector of which covariance

see the program

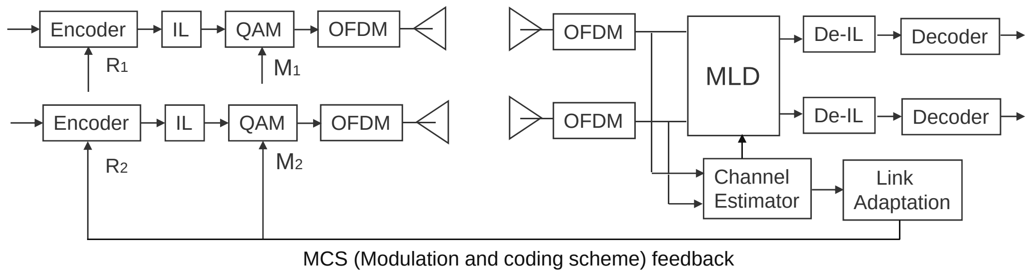


Fig. 1. 2Tx-antenna MIMO-OFDM spatial multiplexing systems using MLD

matrix is defined as $E[\boldsymbol{n}_k\boldsymbol{n}_k^H] = \sigma^2\boldsymbol{I}_N$. $\boldsymbol{I}_N$ is an $N \times N$ indentity matrix. $\boldsymbol{s}_k$ is the transmitted symbol vector which satisfies $E[\boldsymbol{s}_k\boldsymbol{s}_k^H] = \boldsymbol{I}_2$.

Adaptive modulation and coding is applied, where the receiver performs link adaptation based on the instantaneous channel $\boldsymbol{h}_{kl}$ and σ^2 to select one modulation and coding scheme (MCS) providing the highest throughput for each stream. Selected modulation order M_l and coding rate R_l for the l-th stream are fed back to the transmitter.

3. LINK ADAPTATION FOR MIMO-OFDM SPATIAL MULTIPLEXING SYSTEMS

3.1. LMMSE receiver

When the LMMSE receiver is used, the post-MMSE SINR for the l-th stream at the k-th subcarrier can be computed as,

$$\gamma_{kl}^{\text{mmse}} = \frac{\boldsymbol{h}_{kl}^H\boldsymbol{R}_k^{-1}\boldsymbol{h}_{kl}}{1-\boldsymbol{h}_{kl}^H\boldsymbol{R}_k^{-1}\boldsymbol{h}_{kl}}, \quad (3)$$

where $\boldsymbol{R}_k = \boldsymbol{H}_k\boldsymbol{H}_k^H + \sigma^2\boldsymbol{I}_N$. Once the stream-wise post-MMSE SINR value for each subcarrier is computed, the receiver chooses a single MCS matched to all subcarriers. Several techniques have been studied to select a single MCS from multiple SINR values. This paper employs the average mutual information (MI) scheme [7] that computes an average MI from multiple SINR values, then use it to estimate FER for each MCS with look-up-table, and select one MCS providing the highest throughput.

3.2. MLD

3.2.1. Post Processing SINR computation

We first assume without loss of generality that the first stream is the desired stream and second stream is the interfering stream. Since MLD performs joint detection by evaluating MIMO channel metric $-\frac{1}{\sigma^2}\|\boldsymbol{r}_k - \boldsymbol{H}_k\boldsymbol{s}_k\|^2$ for multiple hypothesis on $\boldsymbol{s}_k$, the exact post-processing SINR cannot be derived. Thus, we derive an approximate post-processing SINR. For this aim, the post-MLD SINR of the first stream is first lower-and-upper bounded as

$$\gamma_{k1}^{\text{mmse}} \le \gamma_{k1}^{\text{mld}} \le \gamma_{k1}^{\text{if}}, \quad (4)$$

where $\gamma_{k1}^{\text{if}} \equiv \frac{\|\boldsymbol{h}_{k1}\|^2}{\sigma^2}$ is the interference-free SNR of the first stream. An important observation here is as follows. When the second (interfering) stream becomes more detectable by MLD, the post processing SINR of the first stream becomes larger, i.e., $\gamma_{k1}^{\text{mld}} \to \gamma_{k1}^{\text{if}}$. This happens, for example, when the modulation order of the second stream is decreased while the channel condition is fixed. On the other hand, when the second stream becomes less detectable, the post processing SINR of the first stream becomes smaller, i.e., $\gamma_{k1}^{\text{mld}} \to \gamma_{k1}^{\text{mmse}}$. This occurs, for example, when the modulation order of the second stream is increased. Based on (4), we compute the post-MLD SINR of the first stream as

$$\gamma_{k1}^{\text{mld}} = \gamma_{k1}^{\text{mmse}} + \alpha_{k1}(\gamma_{k1}^{\text{if}} - \gamma_{k1}^{\text{mmse}}), \quad (5)$$

where the second term in the right hand side of (5) is the gain factor obtained by MLD compared with the LMMSE detector. According to the observation above, the coefficient α_{k1} $(0 \le \alpha_{k1} \le 1.0)$ is adaptively chosen according to the detectability of the second (interfering) stream. To measure the detectability of the second stream, symbol-error-rate (SER) can be employed. The SER upper-bound of the second stream can be computed by using its post-MMSE SINR value as [8],

$$P_{k2}^{\text{ser}} = 1 - \left\{1 - 2\left(1 - \frac{1}{\sqrt{2^{M_2}}}\right) Q\left(\sqrt{\frac{3\gamma_{k2}^{\text{mmse}}}{2^{M_2}-1}}\right)\right\}^2, \quad (6)$$

where M_2 is the modulation order of the second stream, and $Q(x) \equiv \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$. Note that in actual implementations, $Q(x)$ is realized by using a look-up-table.

The observation above also suggests that α_{1k} be a decreasing function of P_{k2}^{ser}. In this paper, we define α_{1k} as,

$$\alpha_{1k} = \begin{cases} 0 & P_{k2}^{\text{ser}} > 0.3 \\ 0.3 & 0.1 \le P_{k2}^{\text{ser}} \le 0.3 \\ 0.7 & 0.05 < P_{k2}^{\text{ser}} \le 0.1 \\ 1.0 & P_{k2}^{\text{ser}} \le 0.05 \end{cases}, \quad (7)$$

see the program

Table 1. Link Adaptation Algorithm for MLD

Step 1.	Compute $\gamma_{kl}^{\text{mmse}}$ and γ_{kl}^{if} ($k = 1 \ldots K$ and $l = 1, 2$)
Step 2.	Select a set of Q candidate modulation pairs:
	$(M_{11},M_{21}) \ldots (M_{1Q},M_{2Q})$
Step 3.	Find optimum MCS
	$Th_{\max} = 0.0$
	for q=1:Q
	(a) Compute the post-MLD SINR ($k = 1 \ldots K$):
	$\gamma_{kl}^{\text{mld}} = \gamma_{kl}^{\text{mmse}} + \alpha_{kl}(\gamma_{kl}^{\text{if}} - \gamma_{kl}^{\text{mmse}})$ (l =1,2)
	$\alpha_{k1} = f(\gamma_{k2}^{\text{mmse}}, M_{2q})$, $\alpha_{k2} = f(\gamma_{k1}^{\text{mmse}}, M_{1q})$
	(b) Coding rate selection and FER estimation using average MI scheme [7]:
	(R_{1q},R_{2q}) and $(P_{1q}^{\text{fer}},P_{2q}^{\text{fer}})$
	(c) Throughput estimation and comparison
	$Th_q = \sum_{l=1}^{2} M_{lq}R_{lq}(1 - P_{lq}^{\text{fer}})$
	if $Th_q > Th_{\max}$
	$(M_l, R_l) = (M_{lq}, R_{lq})$ (l =1,2)
	$Th_{\max} = Th_q$
	end
	end

which was optimized by off-line simulations. The optimum α_{k1} function could be different for different channel models, but in this paper we apply the common α_{k1} function to all channel models for simple implementation. Finally, (6) and (7) are integrated into a single function,

$$\alpha_{k1} = f(\gamma_{k2}^{\text{mmse}}, M_2). \tag{8}$$

In the same way as above, the post-MLD SINR value of the second stream is computed by regarding the first stream as interference,

$$\gamma_{k2}^{\text{mld}} = \gamma_{k2}^{\text{mmse}} + \alpha_{k2}(\gamma_{k2}^{\text{if}} - \gamma_{k2}^{\text{mmse}}) \tag{9}$$

with

$$\alpha_{k2} = f(\gamma_{k1}^{\text{mmse}}, M_1). \tag{10}$$

3.2.2. *Link adaptation algorithm*

As shown in (8) and (10), computing the post-MLD SINR value of each stream requires the modulation order of the interfering stream. Thus, in applying the post-MLD SINR to link adaptation, we first determine a set of candidate modulation pairs, and then compute the post-MLD SINR for each pair and select the one providing highest throughput. Table I summarizes the link adaptation algorithm. Our proposed post-MLD-SINR calculation is used in Step 3 (a). Steps 3 (b) and 3 (c) are standard procedures. Different techniques could be applied in Step 3(b). Regarding the choice of candidate modulation pairs in Step 2, we evaluate the two alternatives as follows:

Option 1: All possible modulation pairs.
Option 2: A modulation pair optimum for the LMMSE receiver plus two surrounding pairs.

In option 1, there will be nine pairs if QPSK, 16QAM and 64QAM are used. In option 2, we first select one modulation pair matched to the post-LMMSE SINR, and then we form additional candidate pairs by incrementing the modulation order of the first (second) stream by one level and decrementing the modulation order of the second (first) stream by one level. For example, if the modulation pair selected from the post-LMMSE SINR is (16QAM, 16QAM), the surrounding pairs are (QPSK, 64QAM) and (64QAM, QPSK).

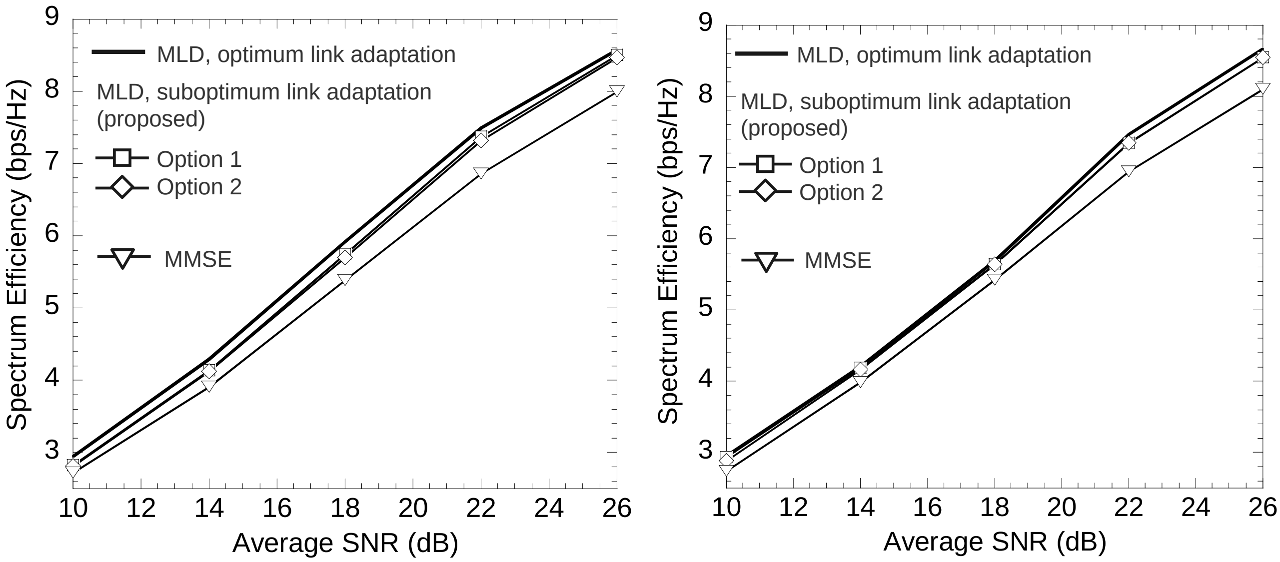
4. SIMULATION RESULTS

Simulations were conducted assuming 2-by-2 MIMO-OFDM spatial multiplexing systems, 1.0-msec transmission frame length with 14 OFDM symbols, 1.8 MHz bandwidth with 120 subcarriers and 15 kHz subcarrier spacing. For adaptive modulation and coding, the following 15 level MCS were used: QPSK with coding rate R=1/16, 1/8, 1/6, 1/3, 4/9, 3/5, 16QAM with R=2/5, 1/2, 3/5, 64QAM with R=4/9, 5/9, 2/3, 3/4, 6/7, 12/13. The turbo code with memory three and mother code rate 1/3 was used. Maximum Doppler frequency was 5.55 Hz. Ideal channel estimation was assumed.

As a performance reference, we evaluated the throughput performance of MLD with optimum link adaptation, where the receiver simulates FER (without using a look-up-table) for all possible MCS combinations of the two streams for each channel realization to find the MCS combination which would provide the highest throughput. This exhaustive approach would yields optimum performance when MCS feedback delay is negligible. In addition, the throughput performance of LMMSE was evaluated.

We first evaluate throughput performances with a small MCS feedback delay of 1.0 msec. Figures 2 (a) and 2 (b) show the throughput performance in uncorrelated MIMO channels with 1-path, and the 6-path typical urban (TU) [9] multi-path delay profiles, respectively. Figures 3 (a) and 3 (b) show the throughput performance in correlated MIMO channels with TU and SCM-C [10] channel models, respectively, where correlation factor 0.5 was assumed between adjacent transmit antennas for the TU channel model. The figures show that MLD with the proposed link adaptation can achieve about 2.0 dB gain over the LMMSE receiver in mid-to-high SNR regions. Option 1 and 2 achieve almost same throughput. The small performance loss of the proposed scheme compared with the optimum scheme is due to use of the SER lower-bound by

see the program



(a) Flat (1-path) channel

(b) TU (6-path) channel

Fig. 2. Throughput performance: 2-by-2 MIMO, uncorrelated channel, 1.0 msec feedback delay

using the post-MMSE SINR value (6), which might underestimate MCS.

We next evaluate the performance with a practical MCS feedback delay of 4.0 msec. Figures 4 shows the throughput performance in the correlated MIMO channels with TU delay profile. The figure show that in the presence of 4.0 msec channel ambiguity, the throughput performance of the proposed approach becomes closer to that of the optimum approach. This is because the proposed approach is more robust to the channel variation due to the MCS underestimation effect.

The results above also show that the throughput gain of MLD compared with LMMSE becomes smaller in lower SNR regions. To examine this, Figure 5 shows FER performance after decoding for various MCS without link adaptation, where the same MCS was assumed for both streams. Inter-stream symbol interleaving was applied to eliminate the performance loss due to use of common MCS format for both streams. The figure shows that when lower coding rates are used, the FER performance gain by MLD over LMMSE becomes smaller. This is because in lower coding rates, decoding is the dominant part that determines the overall FER performance especially when strong codes such as the turbo code is used. Therefore, in lower-SNR regions, the throughput performance gain of MLD over LMMSE detector becomes smaller because MCS with relatively lower coding rates are selected by link adaptation.

5. CONCLUSION

This paper developed a link adaptation technique for 2-Tx MIMO-OFDM systems using non-iterative MLD. We derived an approximate post-MLD SINR, which is computed by using post-MMSE SINR value plus gain factor, where the gain factor is compute by SER of the interfering stream. The throughput performance of MLD with the proposed link adaptation is close to that with the optimum approach, and about 2.0 dB better than the throughput performance of the LMMSE receiver. Extension of the proposed post-MLD SINR derivation to the MIMO channel with more than two spatial streams would be a future study.

6. REFERENCES

[1] 3GPP, TS36.211(V8.4.0), *Evolved Universal Terrestrial Radio Access (E-UTRA): Physical channels and modulation*, Sept. 2008.

[2] R. van Nee, A. van Zelst and G. Awater, *Maximum likelihood decoding in a space division multiplexing system*, Proc. IEEE Vehicular Technology Conference, vol. 1, pp. 6 - 10, May 2000.

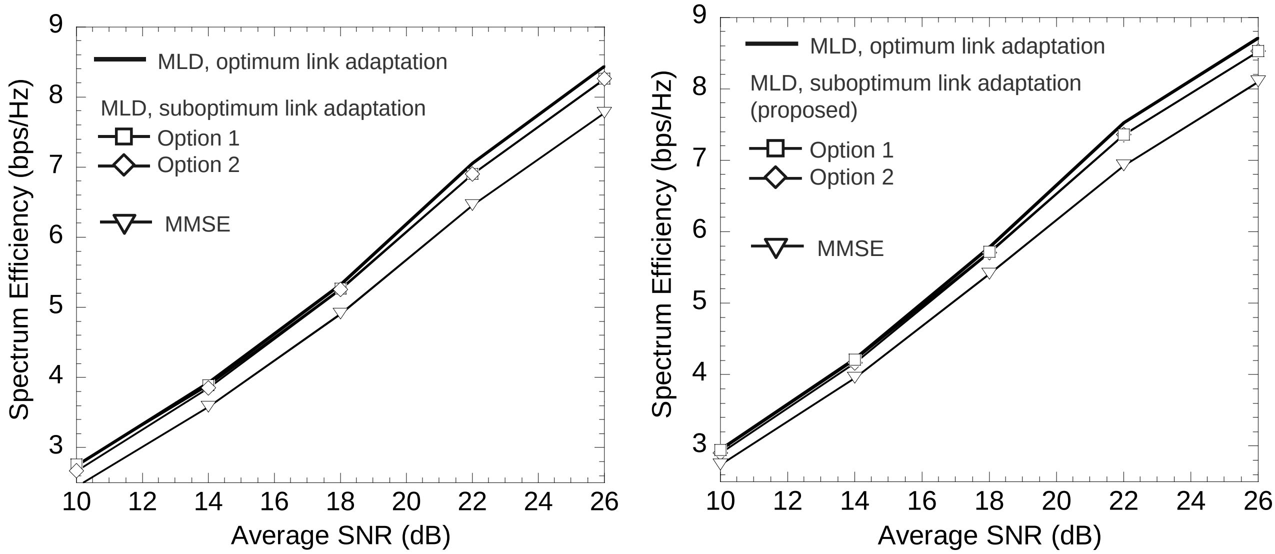
[3] Y. L. C. de Jong and T. J. Willink, *Iterative tree search detection for MIMO wireless systems*, Proc. IEEE Vehicular Technology Conference (VTC), vol. 2, pp. 1041 - 1045, Sept. 2002.

[4] B. M. Hochwald and S. ten Brink, *Achieving near-capacity on a multiple-antenna channel*, IEEE Trans. Communications, vol. 51, no. 3, pp. 389 - 399, March 2003.

[5] K. J. Kim and Y. Jiang, *Joint channel estimation and data detection algorithms for MIMO-OFDM systems*, Proc. Thirty-Sixth Asilomar Conference, vol. 2, pp. 1857 - 1861, Nov. 2002.

[6] F. Kharrat-Kammoun, S. Fontenelle and J. Boutros *Accurate Approximation of QAM Error Probability on Quasi-Static*

see the program



(a) TU channel with Tx correlation factor 0.5

(b) SCMC channel

Fig. 3. Throughput performance: 2-by-2 MIMO, correlated channel, 1.0 msec feedback delay

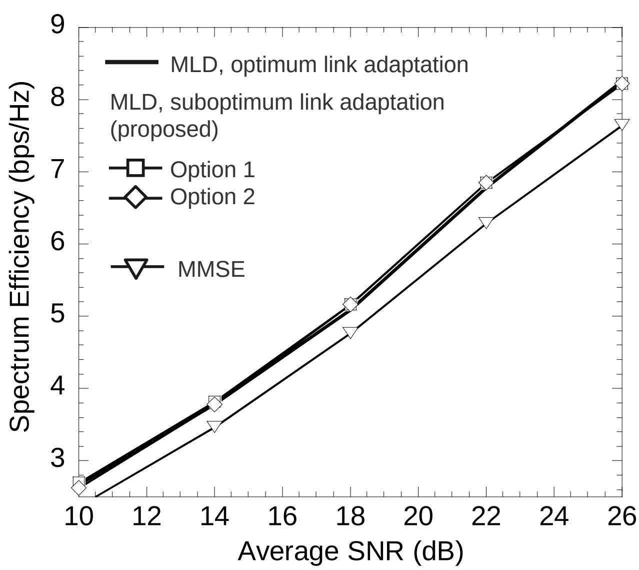


Fig. 4. Throughput performance: 2-by-2 MIMO, TU channel with Tx correlation factor 0.5, 4.0 msec feedback delay

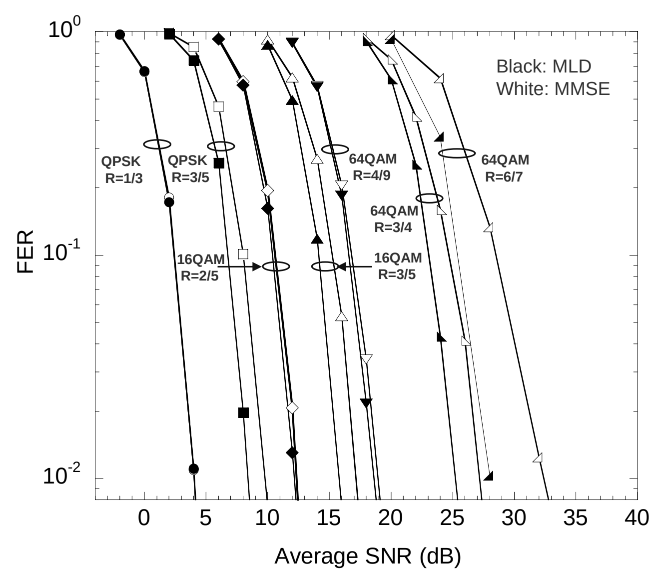


Fig. 5. FER performance: 2-by-2 MIMO, uncorrelated, 6-path TU channel

MIMO Channels and Its Application to Adaptive Modulation, IEEE Trans. Information Theory, vol. 53, no. 3, pp. 1151 - 1160, March 2007.

[7] L. Wan, S. Tsai and M. Almgren, *A fading-insensitive performance metric for a unified link quality model*, Proc. IEEE WCNC, vol. 4, pp. 2110 – 2114 , April 2006.

[8] J. Proakis, *Digital Communications, Third Edition*, McGraw-Hill, 1995.

[9] 3GPP, TS45.005 V7.3.0, Technical Specification Group GSM/EDGE Radio Access Network, *Radio Transmission and Reception*, Nov. 2005.

[10] 3GPP, R1-061001, Ericsson, *LTE Channel Models and link simulations*, March 2006.

see the program

IMPROVING CHANNEL ESTIMATION WITH TURBO FEEDBACK AND TIME DOMAIN FILTERING IN MIMO-OFDM

Jari Ylioinas, Samuli Tiiro, and Markku Juntti

Centre for Wireless Communications
University of Oulu
P.O. Box 4500, FI-90014 University of Oulu, Finland
{jari.ylioinas,samuli.tiiro,markku.juntti}@ee.oulu.fi

ABSTRACT

An improved channel estimation using a turbo decoder feedback and a time domain minimum mean square error (MMSE) filtering is considered in this paper. By using orthogonal preambles, the matrix inversion in the least-squares (LS) multiple-input multiple-output (MIMO) channel estimation can be avoided. The LS estimates can be further improved by the time domain MMSE filtering that exploits the time correlation of the channel. The MMSE filter is approximated by using pre-calculated filters stored in the memory. It is shown by analytical means and by simulations that only few pre-calculated filters are required resulting in tremendous decrease in complexity. The turbo decoder feedback can be used to obtain the channel estimates not only on pilot symbol positions but also on data symbol positions. The space-alternating generalized expectation-maximization (SAGE) algorithm is used to calculate the LS estimates iteratively to reduce the size or even to avoid the matrix inversion required in the LS estimation for data symbol positions. The performance evaluations using the LS estimation and the SAGE estimation together with the MMSE filtering are provided.

1. INTRODUCTION

Multiple-input multiple-output (MIMO) processing (see [1], [2], and the references therein), combined with orthogonal frequency-division multiplexing (OFDM) [3], is the key enabling technology for several current and future broadband wireless access systems and standards. Accurate channel state information is required in the detection when using MIMO channels. The least-squares (LS) method is the best linear unbiased channel estimator in Gaussian noise [3]. The preamble-based LS channel estimation can be further improved by post-processing the LS estimates with the minimum mean square error (MMSE) filter [4]. The MMSE filter exploits the time correlation of the channel taps to calculate the channel estimates for the data symbol positions as well.

A preamble-based channel estimation fails to exploit the channel information contained in the detected and decoded data symbols. To achieve better spectral efficiency under fast-fading channel conditions, the decision directed (DD) channel estimation algorithms (see, e.g., [5, Chapt. 15]) which take advantage of the detected and decoded data symbols in channel estimation, are a promising approach. With the DD LS channel estimation, calculation of a matrix inverse is required. The size of the matrix to be inverted depends on the number of transmit antennas and the length of the channel impulse response. The space-alternating expectation-maximization (SAGE) [6] channel estimator reduces the size of the matrix to be inverted or even completely avoids the matrix inversion due to dividing the MIMO channel estimation problem into multiple SISO channel estimation problems. Error propagation is the frailty of the DD channel estimation methods.

In this paper, we compare the performance of the preamble-based LS channel estimation to that of the DD SAGE channel estimator using pilot density comparable to the one used in mobile commucation system standards nowadays. Both estimators are studied together with the MMSE post-processing [4] as well. We show how the complexity of the MMSE filtering can be reduced by using only few pre-calculated filters with insignificant loss in the performance.

2. SYSTEM MODEL

2.1. OFDM Signal Model

Let us consider a MIMO-OFDM system with P subcarriers, M_{T} transmit, and M_{R} receive antennas. The received signal is the superposition of M_{T} distorted transmitted signals. Consequently, the received signal in the receive antenna m_{R} over sub-carriers at time n after performing a discrete Fourier transform (DFT) can be expressed as

$$\underline{\boldsymbol{y}}_{m_{\mathrm{R}}}(n) = \underline{\boldsymbol{X}}(n)\underline{\boldsymbol{F}\boldsymbol{h}}_{m_{\mathrm{R}}}(n) + \boldsymbol{\eta}_{m_{\mathrm{R}}}(n), \tag{1}$$

where $\underline{\boldsymbol{y}}_{m_{\mathrm{R}}} = [y_{m_{\mathrm{R}},1}, ..., y_{m_{\mathrm{R}},p}, ..., y_{m_{\mathrm{R}},P}]^{\mathrm{T}} \in \mathbb{C}^P$, $\underline{\boldsymbol{X}} = [\boldsymbol{X}_1, ..., \boldsymbol{X}_{M_{\mathrm{T}}}] \in \mathbb{C}^{P\times M_t P}$ consists of transmitted symbols,

This research was financially supported in part by Tekes, the Finnish Funding Agency for technology and Innovation, Nokia, Texas Instruments, Nokia Siemens Networks and Elektrobit.

see the program

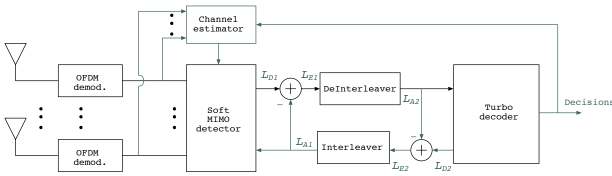


Fig. 1. Receiver structure.

$\boldsymbol{X}_{m_\mathrm{T}} \in \mathbb{C}^{P\times P}$ is a diagonal matrix with $\boldsymbol{X}_{m_\mathrm{T}}[p,p] = x_{m_\mathrm{T},p}$, $\underline{\boldsymbol{F}} = \boldsymbol{I}_{M_\mathrm{T}} \otimes \boldsymbol{F} \in \mathbb{C}^{M_\mathrm{T}P\times M_\mathrm{T}L}$ is DFT matrix, with $[\boldsymbol{F}]_{u,s} = \frac{1}{\sqrt{P}}e^{-j2\pi us/P}$, and $u = 0,1,...,P-1$; $s = 0,1,...,L-1$, $\underline{\boldsymbol{h}}_{m_\mathrm{R}} = [\underline{\boldsymbol{h}}^\mathrm{T}_{1,m_\mathrm{R}},...,\underline{\boldsymbol{h}}^\mathrm{T}_{M_\mathrm{T},m_\mathrm{R}}]^\mathrm{T} \in \mathbb{C}^{M_t L}$ is time domain channel vector, with $\underline{\boldsymbol{h}}_{m_\mathrm{T},m_\mathrm{R}} = [\underline{h}_{m_\mathrm{T},m_\mathrm{R},0},...,\underline{h}_{m_\mathrm{T},m_\mathrm{R},l},...,\underline{h}_{m_\mathrm{T},m_\mathrm{R},L-1}]^\mathrm{T} \in \mathbb{C}^L$, and $\underline{h}_{m_\mathrm{T},m_\mathrm{R},l}$ is the lth multipath component between the m_Tth transmit and the m_Rth receive antenna, L is the length of the channel impulse response for all channels, and $\boldsymbol{\eta}_{m_\mathrm{R}}$ is an $P\times 1$ noise vector containing samples from a complex zero-mean white Gaussian noise process with covariance $\sigma^2_{\boldsymbol{\eta}}\boldsymbol{I}_P$.

2.2. Channel Model

A wideband stochastic MIMO channel model [7,8] is adopted. It is assumed that the amplitude of $\underline{h}_{m_\mathrm{T},m_\mathrm{R},l}$ is Rayleigh distributed, and the fading gains are uncorrelated over the delay domain or $\langle \underline{h}_{m_\mathrm{T},m_\mathrm{R},l_1}, \underline{h}_{m_\mathrm{T},m_\mathrm{R},l_2}\rangle = 0$ for $l_1 \neq l_2$. The spatial correlation coefficient between the transmit antennas m_T and m'_T is given by $\rho^\mathrm{t}_{m_\mathrm{T},m'_\mathrm{T}} = \langle \underline{h}_{m_\mathrm{T},m_\mathrm{R},l}, \underline{h}_{m'_\mathrm{T},m_\mathrm{R},l}\rangle$, where $\langle a,b\rangle = E[ab^*]/\sqrt{E[|a|^2]E[|b|^2]}$ denotes the normalized correlation coefficient between the random variables a and b. The spatial correlation function at the transmitter is assumed to be independent of the receive antenna index m_R. The correlation coefficient between the receive antennas is correspondingly denoted by $\rho^\mathrm{r}_{m_\mathrm{T},m'_\mathrm{T}}$. The overall spatial correlation model is assumed to obey the Kronecker product model $\boldsymbol{C}_\mathrm{tr} = \boldsymbol{C}_\mathrm{t} \otimes \boldsymbol{C}_\mathrm{r} \in \mathbb{R}^{M_\mathrm{T}M_\mathrm{R}\times M_\mathrm{T}M_\mathrm{R}}$, where $\boldsymbol{C}_\mathrm{t}$ and $\boldsymbol{C}_\mathrm{r}$ are the transmit and receive correlation matrices with elements $\rho^\mathrm{t}_{m_\mathrm{T},m'_\mathrm{T}}$ and $\rho^\mathrm{r}_{m_\mathrm{R},m'_\mathrm{R}}$, respectively [8].

3. RECEIVER ALGORITHMS

3.1. Receiver Structure

The receiver structure is shown in Fig. 1. The feedback from the decoder to the channel estimator is required only when the SAGE estimator is applied. In the core of the receiver, iterative detection and decoding is performed by the soft MIMO detector and the turbo decoder. The iterations between the detector and the decoder are referred to as global iterations. Let us assume that one frame includes multiple OFDM symbols and the channel coding is over one OFDM symbol. In each frame, the preamble symbol index set is $\mathcal{S}_P = \{1, N_\mathrm{symb}/2+1\}$, where N_symb is the number of OFDM symbols in the frame. The cardinality of $\mathcal{S}_P$ is denoted by N_P. In the preamble based LS estimation, the estimate given by the first preamble is used in the detection of the symbols with indices $2...N_\mathrm{symb}/2-1$ and the estimate given by the second preamble is used in the detection of the symbols with indices $N_\mathrm{symb}/2+2...N_\mathrm{symb}$. In the preamble based LS estimation with the MMSE post-processing, the two LS estimates given by the preambles are filtered and as an output a channel estimate for each OFDM symbol is given. In the DD SAGE channel estimation, the LS estimate is used as an initial guess in the beginning of the frame and after the second OFDM symbol the previous estimate is used to initialize the SAGE algorithm. It is also possible to initialize the SAGE algorithm by predicting the channel state by filtering the previous channel estimates with the MMSE filter. The SAGE algorithm is performed after global iterations.

Let us consider coding over the whole frame. The LS estimation with the MMSE post-processing is performed as explained above in the case of coding over one OFDM symbol. However, the DD SAGE estimation is now applied within the global iterations. The initial channel estimate for the whole frame is given by the LS estimation with the MMSE post processing and after a certain number of global iterations the channel is re-estimated by the SAGE algorithm which is followed by a certain number of global iterations.

3.2. LS Algorithm

The LS method is the best linear unbiased channel estimator in Gaussian noise [3]. The LS channel estimate for the system model in [4] is expressed as

$$\hat{\underline{\boldsymbol{h}}}_{m_\mathrm{R}}(n) = \Bigg(\underbrace{\underline{\boldsymbol{F}}^\mathrm{H}\underline{\boldsymbol{X}}^\mathrm{H}(n)\underline{\boldsymbol{X}}(n)\underline{\boldsymbol{F}}}_{\boldsymbol{C}^\mathrm{H}\boldsymbol{C}}\Bigg)^{-1} \underline{\boldsymbol{F}}^\mathrm{H}\underline{\boldsymbol{X}}^\mathrm{H}(n)\underline{\boldsymbol{y}}_{m_\mathrm{R}}(n). \quad (2)$$

In the preamble based channel estimation, the preambles can be designed to diagonalize the matrix to be inverted. However, if the LS method is used with turbo feedback (in the decision directed (DD) mode), the inversion of a $M_\mathrm{T}L\times M_\mathrm{T}L$ matrix is inevitable.

3.3. SAGE Algorithm

The SAGE channel estimator [6] calculates iteratively the LS solution as in (2) resulting in lower complexity with imperceptible loss in the performance in the DD channel estimation. With the iterative processing, the size of the required matrix inversion is M_T times smaller per dimension in the case of non-constant envelope modulation and with a constant envelope modulation the matrix inversion can be avoided. In the SAGE channel estimator, the received signal $\underline{\boldsymbol{y}}_{m_\mathrm{R}}$ is viewed as the "incomplete" data and the "complete" data is defined as $\boldsymbol{z}_{m_\mathrm{T},m_\mathrm{R}}$, which is the component of the received signal at the m_Rth receive antenna transmitted by the m_Tth transmit antenna. The SAGE channel estimator is

see the program

initialized by

$$\hat{\underline{\boldsymbol{z}}}^{(0)}_{m_\mathrm{T},m_\mathrm{R}}(n) = \boldsymbol{X}_{m_\mathrm{T}}(n)\boldsymbol{F}\hat{\underline{\boldsymbol{h}}}^{(0)}_{m_\mathrm{T},m_\mathrm{R}}(n), \tag{3}$$

where $m_\mathrm{T} = 1, 2, ..., M_\mathrm{T}$ and $m_\mathrm{R} = 1, 2, ..., M_\mathrm{R}$. The SAGE algorithms, as expectation-maximization (EM) based methods in general, require and are sensitive to an initial guess of the parameters to be estimated. This is well addressed in the literature [9]. The channel estimate of the previous OFDM symbol can be used as the initial time domain channel estimate $\hat{\underline{\boldsymbol{h}}}^{(0)}_{m_\mathrm{T},m_\mathrm{R}}(n)$. After the initialization, the estimates of the channel impulse responses between the transmit antennas and the receive antenna m_R are refined by the following iterations (for notational simplicity the time index n is omitted)

$$\hat{\underline{\boldsymbol{z}}}^{(i)}_{m_\mathrm{T},m_\mathrm{R}} = \hat{\underline{\boldsymbol{z}}}^{(i)}_{m_\mathrm{T},m_\mathrm{R}} + \left[\underline{\boldsymbol{y}}_{m_\mathrm{R}} - \sum_{m'_\mathrm{T}=1}^{M_\mathrm{T}} \hat{\underline{\boldsymbol{z}}}^{(i)}_{m'_\mathrm{T},m_\mathrm{R}}\right] \tag{4}$$

$$\hat{\underline{\boldsymbol{h}}}^{(i+1)}_{m_\mathrm{T},m_\mathrm{R}} = (\boldsymbol{F}^\mathrm{H}\boldsymbol{X}^\mathrm{H}_{m_\mathrm{T}}\boldsymbol{X}_{m_\mathrm{T}}\boldsymbol{F})^{-1}\boldsymbol{F}^\mathrm{H}\boldsymbol{X}^\mathrm{H}_{m_\mathrm{T}}\hat{\underline{\boldsymbol{z}}}^{(i)}_{m_\mathrm{T},m_\mathrm{R}} \tag{5}$$

$$\hat{\underline{\boldsymbol{z}}}^{(i+1)}_{m_\mathrm{T},m_\mathrm{R}} = \boldsymbol{X}_{m_\mathrm{T}}\boldsymbol{F}\hat{\underline{\boldsymbol{h}}}^{(i+1)}_{m_\mathrm{T},m_\mathrm{R}} \tag{6}$$

$$\hat{\underline{\boldsymbol{z}}}^{(i+1)}_{m''_\mathrm{T},m_\mathrm{R}} = \hat{\underline{\boldsymbol{z}}}^{(i)}_{m''_\mathrm{T},m_\mathrm{R}}, \tag{7}$$

where $m_\mathrm{T} = 1 + (i \bmod M_\mathrm{T})$, $(i = 0, 1, 2, ...)$, and in (7) $1 \leq m''_\mathrm{T} \leq M_\mathrm{T}$ but $m''_\mathrm{T} \neq m_\mathrm{T}$. With one iteration in the case of the SAGE channel estimator, we refer to the cycles during which every link is updated once. The above algorithm transforms the LS channel estimation problem of a system with multiple transmit antennas as in (2) back to the multiple separate single transmit antenna LS channel estimation problems performed serially. Consequently, the matrix inversion needed is M_T times smaller per dimension than the one required with the LS estimation. Furthermore, with constant envelope constellation (5) can be simplified to $\hat{\underline{\boldsymbol{h}}}^{(i+1)}_{m_\mathrm{T},m_\mathrm{R}}(n) = \boldsymbol{F}^\mathrm{H}\boldsymbol{X}^{-1}_{m_\mathrm{T}}(n)\hat{\underline{\boldsymbol{z}}}^{(i)}_{m_\mathrm{T},m_\mathrm{R}}(n)$ leading to the trivial inversion of the diagonal matrix $\boldsymbol{X}_{m_\mathrm{T}}(n)$.

3.4. MMSE Filtering

The LS and SAGE algorithms do not exploit the time domain correlation of the channel in the estimation. Thus, the estimation performance can be further improved by post-processing the LS and SAGE channel estimates with a filter based on the MMSE criterion [4], which takes advantage of the temporal correlation of channel impulse response (CIR) taps between consecutive OFDM symbols. The autocorrelation of the CIR taps can be expressed as

$$\Sigma_{h_{m_\mathrm{T},m_\mathrm{R},l}[n],h_{m_\mathrm{T},m_\mathrm{R},l}[n']} = \rho_{m_\mathrm{T},m_\mathrm{R}}[n-n']\sigma^2_{h_{m_\mathrm{T},m_\mathrm{R},l}} \tag{8}$$

where $\rho_{m_\mathrm{T},m_\mathrm{R},l}[n-n']$ denotes the temporal correlation between the channel taps at times n and n' and $\sigma^2_{h_{m_\mathrm{T},m_\mathrm{R},l}} = h^*_{m_\mathrm{T},m_\mathrm{R},l}h_{m_\mathrm{T},m_\mathrm{R},l}$. The temporal correlation depends on the Doppler spectrum and it is distributed according to the Jakes' model. Thus, the temporal correlation can be written as

$$\rho_{m_\mathrm{T},m_\mathrm{R},l}[n-n'] = \mathrm{J}_0(2\pi f_\mathrm{d}(n'-n)T_\mathrm{B}) \tag{9}$$

where $\mathrm{J}_0(\cdot)$ denotes the zeroth-order Bessel function of the first kind, f_d is the Doppler frequency, and T_B is the duration of OFDM symbol. In order to obtain the MMSE estimates, we define an LS channel estimate vector as

$$\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l} = (\hat{h}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}[n_1] \ldots \hat{h}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}[n_{N_\mathrm{P}}])^\mathrm{T} \in \mathbb{C}^{N_\mathrm{P}}, \quad n \in \mathcal{S}_P. \tag{10}$$

The MMSE CIR estimate for the link between the m_Tth transmit antenna and the m_Rth receive antenna at time n can now be expressed as

$$\hat{h}^{<\mathrm{MMSE}>}_{m_\mathrm{T},m_\mathrm{R},l}[n] = \boldsymbol{w}_{m_\mathrm{T},m_\mathrm{R},l}[n]\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l} \in \mathbb{C}, \tag{11}$$

where the Wiener filtering vector $\boldsymbol{w}_{m_\mathrm{T},m_\mathrm{R},l}[n]$ is defined as

$$\boldsymbol{w}_{m_\mathrm{T},m_\mathrm{R},l}[n] = \boldsymbol{\Sigma}_{h_{m_\mathrm{T},m_\mathrm{R},l}[n],\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}}\boldsymbol{\Sigma}^{-1}_{\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}}. \tag{12}$$

The cross-covariance vector $\boldsymbol{\Sigma}_{h_{m_\mathrm{T},m_\mathrm{R},l}[n],\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}} \in \mathbb{C}^{N_\mathrm{P}}$ and the autocovariance matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}} \in \mathbb{C}^{N_\mathrm{P}\times N_\mathrm{P}}$ in (12) are defined as

$$\boldsymbol{\Sigma}_{h_{m_\mathrm{T},m_\mathrm{R},l}[n],\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}} = (\rho[n-n_1]\ldots\rho[n-n_{N_\mathrm{P}}])\sigma^2_{h_{m_\mathrm{T},m_\mathrm{R},l}}$$

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{h}}^{<\mathrm{LS}>}_{m_\mathrm{T},m_\mathrm{R},l}} = \begin{pmatrix} \rho[n_1-n_1] & \ldots & \rho[n_1-n_{N_\mathrm{P}}] \\ \vdots & \ddots & \vdots \\ \rho[n_{N_\mathrm{P}}-n_1] & \ldots & \rho[n_{N_\mathrm{P}}-n_{N_\mathrm{P}}] \end{pmatrix}\sigma^2_{h_{m_\mathrm{T},m_\mathrm{R},l}} + \boldsymbol{\Sigma}_{\boldsymbol{\eta}},$$

where the noise covariance matrix is $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = \sigma^2_\mathrm{N}[\boldsymbol{\mathcal{C}}^H\boldsymbol{\mathcal{C}}]_{u,s}\boldsymbol{I}_{N_\mathrm{P}}$, $u = (m_\mathrm{T}-1)M_\mathrm{T} + l$, and $s = (m_\mathrm{T}-1)M_\mathrm{T} + l$. To keep the complexity increment due to the filtering to a minimum, only a small set of filters are pre-calculated and stored in the memory.

In Fig. 2, the analytical mean square error (AMSE) [4] performance of the LS estimation with the time domain MMSE filtering is shown in the typical urban channel and with 2×2 MIMO system. The dashed lines are with fixed pre-calculated filters. Two pre-calculated filters are stored and the one closer to prevailing conditions is used. One filter assumes mobile speed of 55 km/h and signal to noise ratio (SNR) value of 8.5 dB and the other 85 km/h and SNR value of 8.5 dB. The SNR is defined as E_S/N_0, where E_S is symbol energy received by a receive antenna and $N_0\boldsymbol{I}$ is the covariance matrix of the noise. It can be noted that based on the analytical results, the mobile speed range from 40 to 100 km/h can be handled with two pre-calculated filters with only minor loss in the performance. In Fig. 3, the corresponding AMSE performance is shown in 4×4 MIMO system. Two pre-calculated filters are used in this case as well. The pre-calculated filters assume the same mobile speeds as in the previous case and the SNR value of 12.5 dB. Two pre-calculated filters seem to be enough also in this case.

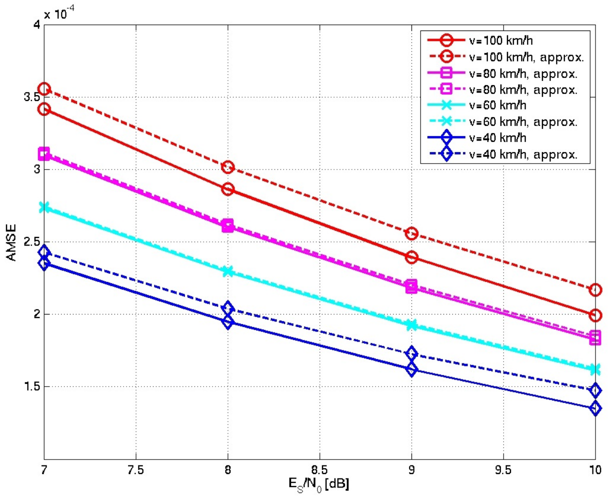


Fig. 2. Analytical MSE after time domain MMSE filtering in 2×2 system.

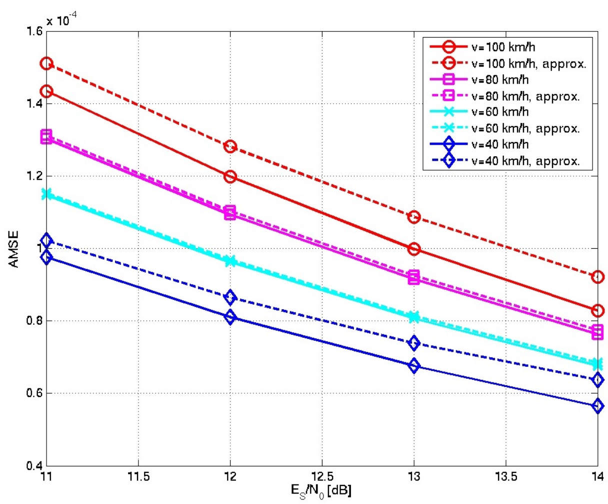


Fig. 3. Analytical MSE after time domain MMSE filtering in 4×4 system.

4. NUMERICAL EXAMPLES

Monte Carlo computer simulations were conducted to study the performance of the channel estimation algorithms. The entire channel bandwidth of 7.68 MHz was divided into 512 subcarriers. The symbol duration of 66.7 μs and the cyclic prefix length of 16.68 μs were assumed. The carrier frequency was 2.4 GHz. The simulations were performed with the typical urban (TU) channel model [10]. The key parameters of the stochastic MIMO radio channel model are summarized in [11, Table I]. The signal-to-noise ratio (SNR) is defined as E_S/N_0, where E_S is symbol energy received by a receive antenna and $N_0\boldsymbol{I}$ is the covariance matrix of the noise. Turbo coding of rate 1/2 was assumed. The turbo coding scheme is a parallel concatenated convolutional code (PCCC) with two 8-state encoders and an internal interleaver followed by puncturing, multiplexing, and bit-level

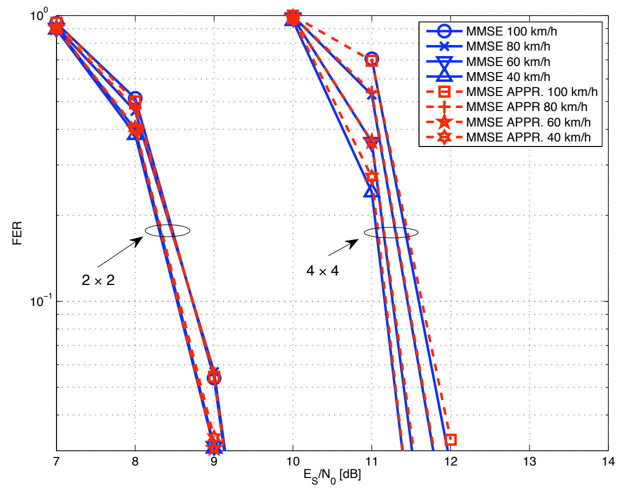


Fig. 4. Performance of pre-calculated MMSE filters in 2×2 and 4×4 antenna configurations assuming QPSK.

interleaving. The constituent codes (CCs) are rate 1/2 binary recursive systematic convolutional codes [12]. The turbo decoder consists of two component maximum *a posteriori* (MAP) decoders [13] and 8 decoder iterations were performed. The increasing radius (IR) list sphere detector (LSD) [14] was operating as a soft MIMO detector.

Figs. 4 and 5 show the frame error rate (FER) performances after the MMSE filtering in 2×2 and 4×4 antenna configurations and with quaternary phase shift keying (QPSK) and 16 quadrature amplitude modulation (16QAM), respectively. The frame length is 6 OFDM symbols and the channel coding is over one frame. The results with two precaluculated filters (MMSE approximations) are compared to those of given by the filters calculated using the ideal channel parameters. Mobile speed of 85 km/h was used in the pre-calculated filter for mobile speeds of 100 km/h and 80 km/h while mobile speed of 55 km/h was used in the pre-calculated filter for mobile speeds of 40 km/h and 60 km/h. The fixed SNR values were 8.6 dB and 11.4 dB in the case of QPSK in 2×2 and 4×4 antenna configurations, respectively, while in the case of 16QAM they were 17.4 dB and 21.7 dB. The results clearly show that in the presented cases the MMSE filtering can be approximated with two pre-calculated filters with insignificant loss and in 2×2 antenna configuration even one pre-calculated filter might be sufficient.

In Fig. 6, FER performances of the channel estimators are presented with preamble overheads of 33.3 % and 16.7 %. 2×2 antenna configuration is used. The data is coded over one OFDM symbol and 16QAM is used. The number of global iterations is 4. The channel estimate of the previous OFDM symbol is used as an initial guess for the SAGE estimator. Due to error propagation, the SAGE estimator loses roughly 1 dB to the preamble based LS estimator with the time domain MMSE post-processing when preamble overhead is 33.3 %. With preamble overhead of

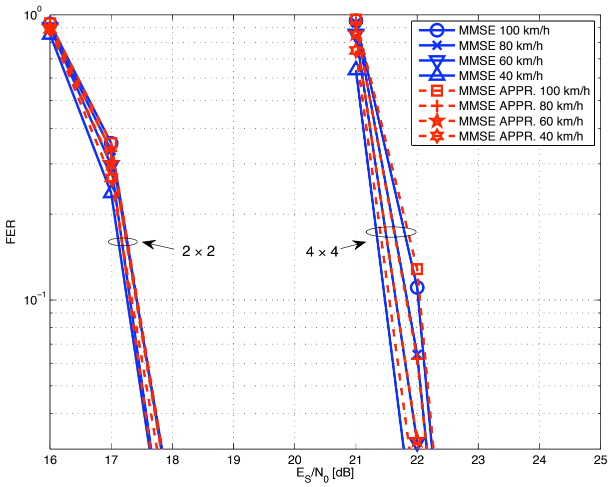


Fig. 5. Performance of pre-calculated MMSE filters in 2×2 and 4×4 antenna configurations assuming 16QAM.

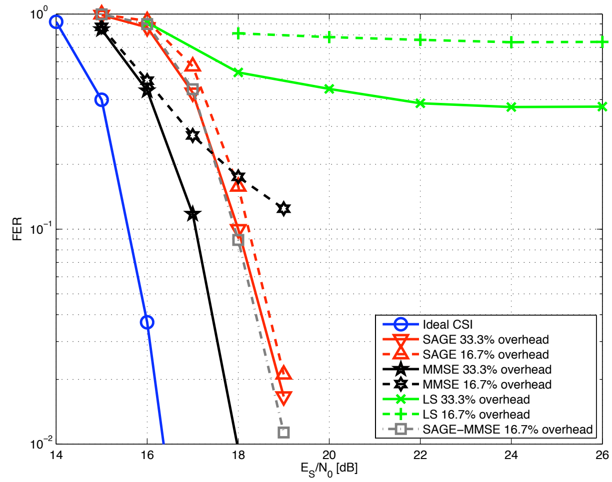


Fig. 6. FER performance with coding over one OFDM symbol.

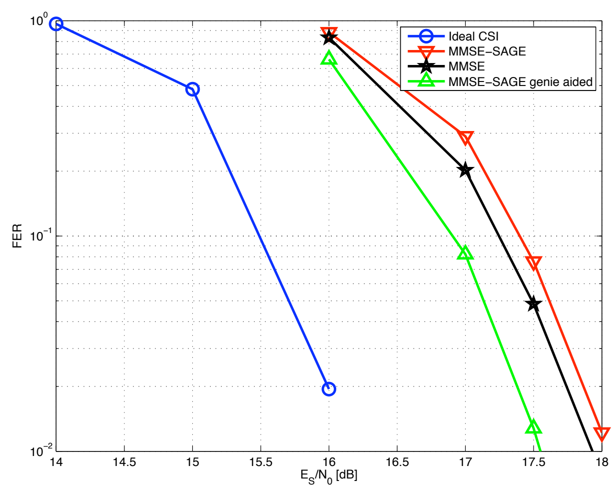


Fig. 7. FER performance with coding over 12 OFDM symbols.

16.7 %, one can note that the LS estimation with time domain post-processing starts to suffer from lower pilot density showing a slope degradation while the SAGE estimation gives almost the same performance as with preamble overhead of 33.3 %. The LS estimator performance is insufficient in both cases. The performance of the SAGE channel estimation can be slightly improved by using the MMSE filter to predict the initial guess. Five previous OFDM symbols were used in the prediction.

In Fig. 7, the data is coded over 12 OFDM symbols and the preamble overhead of 16.7 % is assumed. The studied antenna configuration is 2×2. The number of global iterations is 4. The performance of the LS estimation with the time domain MMSE post-processing is presented. In addition, the result in which the SAGE estimator is used to re-estimate the channel estimate after 2 global iterations is shown. One can note that using the SAGE estimator within the global iterations degrades the performance slightly. When the channel coding is over multiple OFDM symbols, the re-estimated channel estimate does not improve the symbol decisions on a symbol-by-symbol basis. Thus, the erroneous symbol decisions deteriorate the SAGE channel estimates. To verify this, the results in which known symbols are used in the SAGE estimation are presented. From those, one can see that ideally the SAGE estimation could provide a slightly improved performance.

5. CONCLUSIONS

An improved channel estimation was considered for a MIMO-OFDM system. It was shown that the time correlation of the channel can be exploited in the channel estimation to improve the performance with small increase in complexity by using only few pre-calculated filters. The performance of the preamble-based LS estimation with the time domain MMSE post-processing was compared to that of the DD SAGE estimator. According to the results, the use of the DD SAGE channel estimation is beneficial if the channel coding is over one OFDM symbol and with preamble overheads less than 16.7 %. The SAGE estimator performance can be slightly improved by using the MMSE filter and the previous channel estimates to predict the initial guess.

6. REFERENCES

[1] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, “From theory to practice: An overview of MIMO space-time coded wireless systems,” *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, 2003.

[2] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, “An overview of MIMO communications — A key to gigabit wireless,” *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[3] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[4] H. Miao and M. Juntti, “Space-time channel estimation and performance analysis for wireless MIMO-OFDM systems with spatial correlation,” *IEEE Trans. on Vehicular Technology*, vol. 54, no. 6, pp. 2003 – 2016, Nov. 2005.

see the program

[5] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadcasting Multi-User Communications, WLANs and Broadcasting*. New York: John Wiley & Sons, 2003.

[6] Y. Xie and C. N. Georghiades, "Two EM–type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 106–115, Jan. 2003.

[7] K. I. Pedersen, J. B. Andersen, J. P. Kermoal, and P. Mogensen, "A stohastic multiple-input-mobile-output radio channel model for evaluation of space-time coding algorithms," in *Proc. IEEE Veh. Technol. Conf.*, Boston, USA, Sept. 24–28 2000, pp. 893–897.

[8] J. P. Kermoal, L. Schumacher, K. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[9] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York, USA: Wiley, 1997.

[10] 3rd Generation Partnership Project (3GPP); Technical Specification Group Radio Access Network, "Radio transmission and reception (3G TS 45.005 version 5.4.0 (release 5))," 3rd Generation Partnership Project (3GPP), Tech. Rep., 2002.

[11] J. Ylioinas and M. Juntti, "Iterative joint detection, decoding, and channel estimation in turbo coded MIMO-OFDM," *IEEE Trans. Veh. Technol.to appear*.

[12] 3rd Generation Partnership Project (3GPP); Technical Specification Group Radio Access Network, "Multiplexing and channel coding (FDD) (3G TS 25.212 version 7.2.0)," 3rd Generation Partnership Project (3GPP), Tech. Rep., 2006.

[13] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[14] M. Myllylä, M. Juntti, and J. R. Cavallaro, "A list sphere detector based on Dijkstra's algorithm for MIMO-OFDM system," in *Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Athens, Greece, Sept. 3-7 2007, pp. 1–5.

see the program

RATE ALLOCATION FOR K-USER MAC WITH TURBO EQUALIZATION

Marcus Grossmann

Institute for Information Technology
Ilmenau University of Technology, Germany
marcus.grossmann@tu-ilmenau.de

Tad Matsumoto

Japan Advanced Institute of Science and Technology, Japan
Center for Wireless Communication at University of Oulu, Finland
matumoto@jaist.ac.jp

ABSTRACT

A primary objective of this paper is to examine the achievable rate region of frequency selective K-user multiple access channels (MACs) with soft-cancelling frequency domain minimum-mean-squared-error (SC FD-MMSE) turbo equalization. With the aid of the knowledge about the extrinsic information transfer (EXIT) characteristic of the equalizer, we study the problem of determining the optimal rate allocation (RAAL) to the K users, for which the sum rate is maximized, and introduce a heuristic approach based on the dynamic programming principle for solving this optimization problem. We also propose a practical code selection algorithm for the RAAL to the K users given the multidimensional EXIT functions of the equalizer. It is shown through simulations that with SC FD-MMSE turbo equalization, substantial throughput gain can be achieved with the proposed RAAL technique over automatic repeat request with fixed coding rate.

1. INTRODUCTION

Recently, iterative (turbo) techniques [1] have been recognized as practical solutions to the signal detection problem in *frequency selective* channels. In [2], the optimal detector based on the maximum a posteriori probability (MAP) criterion is replaced by simple signal processing that performs soft canceling and minimum mean squared error (SC MMSE) filtering. The turbo equalization technique for single carrier signaling over *frequency-selective* MIMO channels presented in [3] and [4] (referred to as SC FD-MMSE in the following) performs the equivalent processing in the frequency domain, by which the computational complexity is reduced to a logarithmic order of the frame length.

The Extrinsic Information Transfer (EXIT) chart [5] is a powerful tool to analyze the convergence property in terms of mutual information (MI) exchange between the soft input soft output (SfiSfo) components of turbo systems. In [6], the authors showed that the area under the decoder EXIT function is $1 - R$ with R being the rate of the code. This area property leads to the important conclusion that *link* design problems for adaptive transmission with single user turbo equalization, in general, reduce to a *curve-fitting* problem of the two-dimensional (2D) EXIT curves of the equalizer and decoder. It also invokes the idea to change the *link* parameters such as the code and modulation format to increase the information rates of the turbo system, depending on the equalizer EXIT function for a given channel realization. In [7], the authors determine the optimal degree distributions of low density parity check (LDPC) codes for spatially multiplexed MIMO systems to achieve decoding thresholds very close to information-theoretic performance limits. Based on a similar idea, the authors of [8] investigate the optimal code design for bit interleaved coded modulation with iterative detection.

The convergence property analysis for the *frequency selective* K-user MAC with multi-user turbo equalization is possible, in principle, with K-dimensional (KD) EXIT charts, similar to the single user link design described above. There are two detrimental events that dominate performance of such systems, in general, which are: (1) the convergence is stuck before relatively high MI between the transmitted signal and decoder output LLR can be achieved, and (2) an unreasonably low rate channel code is used, resulting in a rate loss. Furthermore, unlike the optimal *single user link* design, the equalizer EXIT function of each user depends on the feedback MI from the multiple decoders. This indicates that the code parameters, including code rate and generator polynomials, used by the multiple users have to be carefully chosen so that the two detrimental outcomes do not happen to the users.

To fully exploit the advantage of the *frequency selective* K-user MAC with SC FD-MMSE turbo equalization, we consider in this paper the problem of determining the rate allocation (RAAL) for the K users that can achieve the maximum sum rate, while properly avoiding the convergence stuck.

This paper is organized as follows: Section 2 introduces the system model. Section 3 examines the achievable rate region of the K-user MAC with turbo equalization. In Section 3.1.1, we formulate the optimal RAAL problem that maxi-

This work was supported in part by the German Research Foundation (DFG) under grant SPP1163 and in part by the Japanese government funding program, Grant-in-Aid for Scientific Research (B), No. 20360168.

see the program

mizes the total sum rate, and propose a heuristic approach solving this problem using a dynamic programming technique [10]. In Section 3.2 we propose a practical code *selection* algorithm for the RAAL to the K users given the EXIT characteristic of the equalizer. Section 4 presents the results of simulations conducted to evaluate the throughput enhancement with the proposed RAAL technique.

2. SYSTEM MODEL

We consider a symbol-synchronous single carrier CP assisted K-user uplink wireless communication system, where a base station having M receive antennas receives signals from K active users, each equipped with a single transmit antenna. The transmission scheme of each user is bit interleaved coded modulation, where the information bit sequence is independently encoded by a rate-$r_{c,k}$ (with $k = 1, ..., K$ being the user index) binary encoder, randomly bit-interleaved, binary phase-shift keying (BPSK) modulated, and demultiplexed into N transmit blocks, each having a length KQ in BPSK symbols. In this paper, the binary encoder can be a single convolutional code (SCC) or a serially concatenated convolutional code (SCCC) [13]. Note that throughout the paper, we use the index k to denote the k-th user.

The MIMO channel is assumed to be frequency selective, where each of the KM links is comprised of L statistically independent path components $h_{k,m}(l)$, $l = 1, .., L$, $m = 1, .., M$, which are modeled as complex Gaussian random variables with zero mean and unit variance. Further, it is assumed that the channel gains are constant during the transmission of one frame (comprised of N blocks), but varying independently frame-by-frame. We further assume that all channel gains are perfectly known at the receiver.

Employing a cyclic prefix of sufficient length to each transmit block, the received signals can be expressed as[1]

$$\mathbf{r}_n = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_{n,k} + \mathbf{v}_n, n = 1, .., N, \tag{1}$$

where $\mathbf{b}_{n,k} = [b_{n,k}(1), ..., b_{n,k}(Q)]^T$ denotes the transmitted vector of user k. $\mathbf{H}_k$ is the block-circulant channel matrix associated to the k-th user, and $\mathbf{v}_n$ is the AWGN with covariance $\sigma^2\mathbf{I}$. Note that $\mathbf{H}_k$ can be de-composed into a block-diagonal matrix $\mathbf{F}_K^H \mathbf{\Xi}_k \mathbf{F}$ with $\mathbf{F}_K = \mathbf{I}_K \otimes \mathbf{F}$, where $\mathbf{F}$ denotes the Fourier transform matrix of size $Q \times Q$.

At the receiver side, iterative processing for joint equalization and decoding employing an SfiSfo algorithm is performed. It is assumed that the receiver consists of an SC FD-MMSE equalizer and K decoders that perform APP decoding. However, it should be noted that the major outcomes of this paper are also applicable to other types of equalization techniques.

Within the iterative processing, *extrinsic* log-likelihood ratios (LLRs) of the coded bits are exchanged between the SC FD-MMSE equalizer and the K decoders, following the turbo principle [2].

Inputs to the SC FD-MMSE equalizer are the received signals $\mathbf{r}_n$, and the *a priori* LLRs

$$\zeta(n,k)(q) = \ln \frac{P(b_{n,k}(q) = 0)}{P(b_{n,k}(q) = 0)}, \forall n, \forall k, q = 1, .., Q.$$

The equalizer computes the extrinsic LLR for each transmitted bit $b_{n,k}(q)$ as

$$\lambda_{n,k}(q) = \ln \frac{P(z_{n,k}(q)|b_{n,k}(q) = 0)}{P(z_{n,k}(q)|b_{n,k}(q) = 1)}, \tag{2}$$

where $z_{n,k}(q)$ is the FD-MMSE filter output as defined in [4]. Note that during the first iteration of turbo equalization, $\zeta(n,k)(q)$ is zero for all n, k, q, and later on $\zeta(n,k)(q)$ is provided via the k-th interleaver by the extrinsic LLRs of the k-th decoder.

The receiver also selects for each user the code to be used from an available code set according to the criterion derived in the Section 3.2, where the K users are notified of the codes selected via separated feedback links. It is assumed that the feedback links have no delay and are error-free.

3. RATE ALLOCATION WITH TURBO EQUALIZATION

Let the MI between the transmitted bits $b_{n,k}(q)$ and the corresponding LLRs $\lambda_{n,k}(q)$ provided by the k-th equalizer output be denoted as [4]

$$I_{e,k} = \lim_{N\to\infty} \frac{1}{NQ} I(b_{n,k}(1), ..., b_{n,k}(Q); \lambda_{n,k}(1), ..., \lambda_{n,k}(Q)),$$

and let the MI between $b_{n,k}(q)$ and the LLRs $\zeta_{n,k}(q)$ at the k-th equalizer input be denoted as

$$I_{d,k} = \lim_{N\to\infty} \frac{1}{NQ} I(b_{n,k}(1), ..., b_{n,k}(Q); \zeta_{n,k}(1), ..., \zeta_{n,k}(Q)).$$

In the K-user case the convergence characteristic of the equalizer is defined by K EXIT functions, where each function is conditioned on the channel response and the receiver noise variance [4]. These K EXIT functions can be represented by a vector-function, $\mathbf{f}_e : \mathbf{I}_d \to \mathbf{f}_e = (f_{e,1}(\mathbf{I}_d), ..., f_{e,K}(\mathbf{I}_d)) \in \mathbb{E}^K$, which depends on the feedback MI $\mathbf{I}_d = (I_{d,1}, ..., I_{d,K}) \in \mathbb{E}^K$ from the K decoders. Furthermore, the convergence characteristics of the K decoders are defined by the K EXIT functions $f_{d,k} : I_{d,k} \to$

[1]The transpose operator is denoted by $(\cdot)^T$. The symbol $\otimes$ indicates the Kronecker product. We use the notation $\mathbb{E}$ and $\mathbb{E}^K$ to represent the closed sets $\mathbb{E} = \{x \in [0,1]\}$ and $\mathbb{E}^K = \{\mathbf{x} \in [0,1]^K\}$, respectively. We denote the two vertices given by the K-dimensional row vectors $\mathbf{0}_K$ and $\mathbf{1}_K$ of the region $\mathbb{E}^K$ by $a_0 := \mathbf{0}_K$, and $a_1 := \mathbf{1}_K$, respectively. $\mathcal{D}^1$ denotes the space of continuous and piecewise differentiable functions on the set $\mathbb{E}$. Finally, we use the notation $:=$ to indicate 'by definition'.

see the program

$f_{d,k}(I_{d,k}) \in \mathbb{E}$. In the following, we assume that $\mathbf{f}_e$ and $f_{d,k}$, $\forall k$ are monotonically increasing, continuous and piecewise differentiable.

An example of the two equalizer EXIT functions in the $K = 2$-user case, obtained through simulations using LLR histogram measurements [5] and the decoder EXIT function of user 1, is shown in Fig. 1. Also shown is a possible decoding trajectory of the MI exchange, which is plotted as a projection onto the plane region $\mathbb{E}^2$. Note that the decoder EXIT function of user 2 (not shown) is drawn in the $I_{d,2}$-coordinate. For the computation of the decoding trajectory, the codes of the both users were assumed to be identical, and hence the shapes of their EXIT functions are exactly the same.

As observed in Fig. 1, the MI $\mathbf{I}_d$ exchange between the equalizer and the two decoders is limited to a plane region $\mathcal{S}$. This region (referred to as the convergence region in the following) defines the set of all possible $\mathbf{I}_d$-vectors, for which the equalizer EXIT function of each user is above the corresponding decoder EXIT function, i.e.,

$$\mathcal{S} := \left\{\mathbf{I}_d \in \mathbb{E}^2 | f_{e,k}(\mathbf{I}_d) > f_{d,k}(I_{d,k}), \forall k\right\}. \quad (3)$$

Due to the monotonicity of the EXIT functions, the sequence of MI $\mathbf{I}_d$ resulting from the MI exchange between the SfiSfo components over the iterations, converges monotonically to an achievable maximum point $\mathbf{I}_d^* = (I_{d,1}^*, I_{d,2}^*, ..., I_{d,K}^*)$, which is independent of the activation ordering of the equalizer and decoder iterations [11]. The turbo equalizer can possibly converge to a nearly zero bit error rate (BER) if asymptotically $\mathbf{I}_d^* = a_1$ can be attained. Thus, to achieve near zero BER, the following constraints must hold[2]:

$$\mathcal{S} \text{ is pathwise connected and } a_1 \in \mathcal{S}. \quad (4)$$

Let $A_{\mathcal{S}} := \iint_{\mathcal{S}} \mathrm{d}\mathbf{I}_d$ be the K-dimensional volume of $\mathcal{S}$. Assume that each decoder EXIT function $f_{d,k}$ is matched to the corresponding equalizer EXIT function $f_{e,k}$ so that only an infinitesimally low open-tunnel area between all EXIT functions is maintained. Note that such decoder EXIT functions imply (1) an ideally designed channel code for each user of infinite length to achieve infinitesimally small BER and (2) an infinite number of performed turbo iterations. Further note, it does not lead to the conclusion that the whole transmission chain can achieve the capacity region, since the use of sub-optimal (SC FD-MMSE) equalization algorithms already incurs a loss in rate [6]. Under this assumption, the area $A_{\mathcal{S}}$ asymptotically approaches the value zero and $\mathcal{S}$ reduces to a simple curve. This curve is referred in the following to as convergence curve, which is parameterized by a class $\mathcal{D}^1$-vector-function, $\mathbf{u} := \big(u_1(t), u_2(t), ..., u_K(t)\big) : \mathcal{I} \to \mathcal{S}, \mathcal{I} \subset \mathbb{R}$, where each function $u_k(t)$ is monotonically increasing in the parameter t. The monotonicity of $\mathbf{u}$ is a direct result of the monotonically

[2] A set $\mathcal{A}$ is said to be pathwise-connected if every pair of points in $\mathcal{A}$ can be joined by a path in $\mathcal{A}$.

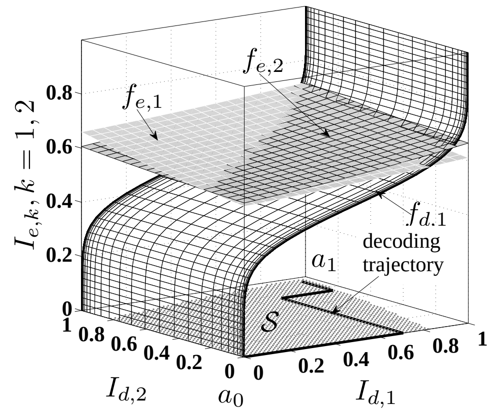


Fig. 1. Equalizer EXIT functions of an 2-user turbo system a single random channel realization and decoder EXIT function of user 1 for a constraint length 5 rate-1/2 SCC. $L = 10$, $Q = 128$, and $E_s/N_0 = 0$ dB.

increasing characteristic of $\mathbf{f}_e$ and $f_{d,k}$, $\forall k$, which ensures that the sequence of MI $\mathbf{I}_d$ does not decrease by the iterations. Let A_k be the area under $f_{d,k}$, $A_k := \int_0^1 f_{d,k}(I_{dk,k})\mathrm{d}I_{d,k}$. In [6] it has been shown that an APP-based decoder for a rate-R_k code satisfies the property $A_k = R_k$[3]. Thus, with the above assumptions on $f_{d,k}$, the rate R_k can be upper bounded by the area under the $(K+1)$-D equalizer EXIT space curve $\mathbf{w}_k(\mathbf{u}(t)) := (u_1(t), u_2(t), ..., u_K(t), f_{e,k}(\mathbf{u}(t)))$,

$$R_k < \int_{\mathcal{S}} f_{e,k}(\mathbf{I}_d)\mathrm{d}I_{d,k} = \int_{\mathcal{I}} f_{e,k}\left(\mathbf{u}(t)\right) u_k'(t)\mathrm{d}t, \quad (5)$$

where $u_k'(t)$ denotes the first derivative of $u_k(t)$. Equation (5) can be equivalently expressed as $R_k < \int_0^1 f_k^{(\mathbf{u})}(I_{d,k})\mathrm{d}I_{d,k}$, where $f_k^{(\mathbf{u})}(I_{d,k})$ denotes the projection of the space curve $\mathbf{w}_k(\mathbf{u}(t))$ onto the $I_{e,k}, I_{d,k}$-plane, where $I_{d,k} = u_k(t)$. An example of the convergence curve for an $K = 2$-user system with the corresponding equalizer EXIT space curves for the two users and the related areas defining the achievable rates R_1 and R_2 for user 1 and user 2, respectively, is shown in Fig. 2.

Let

$$\mathcal{P} := \Big\{\mathbf{p} : \mathcal{I} \to \mathbb{E}^K | \mathbf{p} \text{ of class } \mathcal{D}^1, p_k'(t) \geq 0, \forall t \in \mathcal{I}, \forall k, \\ \mathbf{p}(t_a) = a_0, \mathbf{p}(t_b) = a_1\Big\} \quad (6)$$

be the set of all admissible parametric curves in the region $\mathbb{E}^K$. Each curve $\mathbf{p} \in \mathcal{P}$ is a monotonically increasing func-

[3] The area property has been proved so far only for the case when the a-priori LLRs are assumed to be from a binary erasure channel. Nevertheless, the area property appears also to work well for Gaussian likelihoods, although it does not have a theoretical justification.

see the program

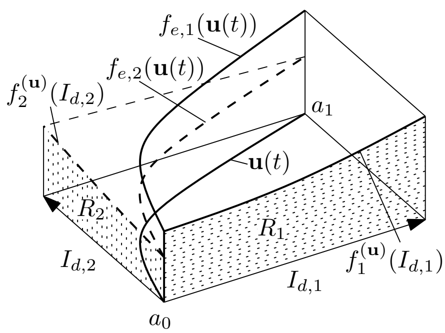


Fig. 2. Example of the parametric convergence curve $\mathbf{u}(t)$ with the two corresponding EXIT space curves and the two areas defining the achievable rates for an $K = 2$-user system.

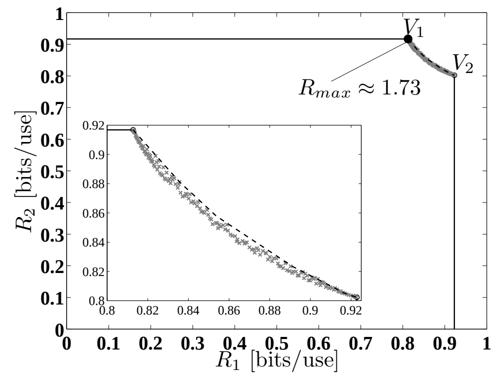


Fig. 3. Approximated rate region of 2-user MAC with SC FD-MMSE turbo equalization for a single random channel realization (a gray dot corresponds to one curve). $L = 10$, $Q = 128$, and $E_s/N_0 = 3$ dB.

tion of $t \in \mathcal{I} = [t_a, t_b]$ that has the prescribed boundary values $\mathbf{p}(t_a) = a_0$ and $\mathbf{p}(t_b) = a_1$. With (5) and (6), the expression of the achievable rate region for the frequency selective K-user MAC with turbo equalization can be expressed as

$$\bigcup_{\mathbf{p}\in\mathcal{P}} \Bigg\{ (R_1, .., R_K) : \sum_{l\in\mathcal{Y}} R_l < \sum_{l\in\mathcal{Y}} \int_{\mathbf{p}} f_{e,l}\left(\mathbf{I}_d\right) \mathrm{d}I_{d,l}, \forall\mathcal{Y} \subseteq \{\forall k\} \Bigg\}. \quad (7)$$

Fig. 3 illustrates an example of the achievable rate region of an 2-user system with SC FD-MMSE turbo equalization for a single random channel realization. In the following, we study the problem of determining the rate tuple and the corresponding convergence curve, for which the sum rate is maximized.

3.1. Sum Rate Maximization

3.1.1. Problem formulation

To identify the rate tuple and the corresponding convergence curve that maximizes the sum rate, we are interested to solve the following optimization problem:

$$R_{max} := \max_{\mathbf{p}\in\mathcal{P}} \sum_{k=1}^{K} \int_{\mathbf{p}} f_{e,k}\left(\mathbf{I}_d\right) \mathrm{d}I_{d,k}. \quad (8)$$

Using (5), the objective can also be written as

$$\max_{\mathbf{p}\in\mathcal{P}} R[\mathbf{p}(\cdot)] := \int_{\mathcal{I}} I(\mathbf{p}(t), \mathbf{p}'(t))\ \mathrm{d}t, \quad (9)$$

where $I(\mathbf{p}(t), \mathbf{p}'(t)) = \sum_{k=1}^{K} f_{e,k}(\mathbf{p}(t))p'_k(t)$. The integral optimization problem (9) falls into the category of the calculus of variations, where the first-order necessary condition of the optimality is given by the Euler-Lagrange differential equations [9]:

$$\frac{\partial I(\mathbf{p}(t), \mathbf{p}'(t))}{\partial p_k} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial I(\mathbf{p}(t), \mathbf{p}'(t))}{\partial p'_k} \overset{!}{=} 0. \quad (10)$$

However, one can easily check that, in general, the candidate curves satisfying the optimality requirement (10) do not satisfy the boundary and monotonicity condition (6). Thus, we cannot use the direct methods of the calculus of variations [9] to solve (9).

3.1.2. An approximate solution using dynamic programming

An approximate solution to (8) can be obtained by transforming the continuous problem into an equivalent discrete dynamic programming problem [10]. In particular, we can discretize the region $\mathbb{E}^K$ into a finite K-dimensional grid of points, which allows us to formulate the variational problem (8) by a path-search problem over a finite weighted graph. Each edge of the graph has its own specific cost and corresponds to a linear curve segment in the region $\mathbb{E}^K$. The admissible curves in this model are assumed to be piecewise linear functions, constrained by the grid points that have to be passed through.

For the ease of notation, we consider in the following only the $K = 2$-user case. But, note that the proposed algorithm can easily be extended to $K > 2$.

Let us denote by $\mathcal{T}$ an uniform 2D grid of $\mathbb{E}^2$,

$$\mathcal{T} := \{ih, jh, 0 \le i \le B, 0 \le j \le B\} \subset \mathbb{E}^2, \quad (11)$$

where the constant $h \in \mathbb{R}^+$ determines the grid spacing in the respective coordinate, and i, j take integer values. We next define a *directed* graph $G = (\mathcal{V}, \mathcal{N})$ for the discrete problem with $\mathcal{V}$ denoting a set of $(B+1)^2$ vertices, $\mathcal{N}$ denoting a finite

see the program

set of directed edges connecting neighboring pairs of vertices, and $c : \mathcal{N} \to \mathbb{E}$ being a cost function over the edges. Each vertex $(i, j) \in \mathcal{V}$ of G is an ordered 2-tuple and corresponds to a point (ih, jh) in the cartesian 2D grid $\mathcal{T}$ as defined above. Moreover, we define, for any vertex $(i, j) \in \mathcal{V}$ of G, the set $\text{Neighb}[(i, j)]$ of its neighbors by:

$$\text{Neighb}[(i,j)] = \Big\{(i',j') \in \mathcal{V} | (i',j') = (i-1,j), (i,j-1), (i-1,j-1)\Big\}. \quad (12)$$

An edge of this graph is therefore of the form $\big((i',j'),(i,j)\big) \in \mathcal{N}$ with $(i,j) \in \mathcal{V}$, and $(i',j') \in \text{Neighb}[(i,j)]$.

Having defined the vertices and edges of G, we can now specify the cost c for each edge. According to (9), a reasonable cost-definition is the sum rate increment that may be approximated for a small step-width τ between neighboring vertices (i', j') and (i, j) as

$$\int_{t'}^{t'+\tau} I(\mathbf{p}(t), \mathbf{p}'(t))\mathrm{d}t \approx \tau f_{e,1}\left(p_1^{(i',i)}, p_2^{(j',j)}\right)\Delta p_1^{(i',i)} + \tau f_{e,2}\left(p_1^{(i',i)}, p_2^{(j',j)}\right)\Delta p_2^{(j',j)} =: c[((i',j'),(i,j))], \quad (13)$$

where the terms $p_k^{(l',l)}$ and $\Delta p_k^{(l',l)}$ are defined as

$$p_k^{(l',l)} := \frac{(l'+l)h}{2}, \ \Delta p_k^{(l',l)} := \frac{(l-l')h}{\tau}. \quad (14)$$

With the above definition of $\text{Neighb}[(i,j)]$, any discrete convergence curve through the 2D grid $\mathcal{T}$ from the bottom left corner (vertex $(0,0)$) to the top right corner (vertex (B,B)) is restricted to the direct neighboring grid points, as illustrated in Fig. 4. The neighbors of a point $r \in \mathcal{T}$ are the three nearest grid points from which the point r is accessible in one step. Thus, given two neighboring grid points, their relative path angle $\varphi = \arctan(\Delta p_2^{(j',j)}/\Delta p_1^{(i',i)})$ is restricted to a multiple of $\pi/4$. Furthermore, any admissible path on G satisfies the monotonicity condition (6), since

$$\Delta p_1^{(i',i)} \geq 0 \text{ and } \Delta p_2^{(j',j)} \geq 0, \forall((i',j'),(i,j)) \in \mathcal{N}. \quad (15)$$

The specified neighborhood structure may be further extended beyond the nearest grid points by extending the definition of the set $\text{Neighb}[(i,j)]$. Naturally, this extended definition will increase the path angular resolution, but also reduce the accuracy of approximation (13). However, the optimal choice of the destination grid is out of the scope of this paper.

Based on the above description of the graph G, we next propose a simple recursive dynamic programming algorithm that provides an efficient solution to the problem of finding the path with maximum cost from vertex $(0,0)$ to (B,B) on G. A pseudo-code implementation of the algorithm is shown in Algorithm 1.

For each vertex $(i,j) \in \mathcal{V}$, the algorithm stores the total cost $d[(i,j)]$ of the maximum-cost path found so far between vertex (i,j) and $(0,0)$. Initially, $d[(i,j)] = 0$ for all $(i,j) \in \mathcal{V}$. The computation of $d[(i,j)]$ is performed by evaluating the edge costs of all neighboring vertices $(i',j') \in \text{Neighb}[(i,j)]$ to vertex (i,j), and selecting the one yielding the highest edge cost, i.e., $d[(i,j)] = d[(i^*,j^*)] + c[(i^*,j^*),(i,j)]$. Additionally, a predecessor label is stored for each vertex (i,j), which represents the previous vertex (i^*,j^*) in the maximum-cost path to the current vertex (i,j). When the algorithm terminates, the value $d[(B,B)]$ at vertex (B,B) represents the maximum cost corresponding to the optimal path on G. Based on the stored predecessor information, back-tracing is then performed to construct the maximum-cost path to vertex $(0,0)$.

Algorithm 1 Pseudo-code for graph-based path-search

```
Initialize d[(i,j)] = 0, 0 ≤ i ≤ B, 0 ≤ j ≤ B.
Initialize pre[(i,j)] = none, 0 ≤ i ≤ B, 0 ≤ j ≤ B.
for i = 0 to B do
  for j = 0 to B do
    if Neighb[(i,j)] ≠ ∅ then
      Find the predecessor to vertex (i,j) with the highest cost value:
        (i*,j*) = arg max_{(i',j')∈Neighb[(i,j)]} { d[(i',j')] + c[((i',j'),(i,j))] }
      Save the predecessor to vertex (i,j):
        pre[(i,j)] = (i*,j*)
      Update the cost d[(i,j)] of vertex (i,j):
        d[(i,j)] = d[(i*,j*)] + c[((i*,j*),(i,j))]
    end if
  end for
end for
Initialize the sequence p̄ = {(Bh,Bh)}.
(i,j) ← (B,B)
while (i,j) ≠ (0,0) do
  (i,j) = pre[(i,j)]
  Insert the grid point (ih,jh) at the beginning of p̄.
end while
Use linear interpolation between the grid points in p̄ and
output the corresponding convergence curve as result.
```

To estimate the computational complexity of Algorithm 1, let us denote by s the maximal number of neighbors of each vertex $v \in \mathcal{V}$. For the specified neighborhood structure (12),

see the program

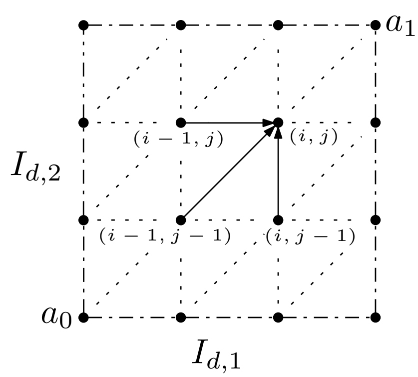


Fig. 4. Discretization of the plane region $\mathbb{E}^2$. Each vertex $(i,j) \in \mathcal{V}$ (corresponding to grid point $(ih, jh) \in \mathcal{T}$) has the three (feasible) neighbors $(i-1,j)$, $(i,j-1)$ and $(i-1,j-1)$.

$s \leq 3$. At each vertex v, the edge cost-values of the s possible neighbors to v have to be computed and compared, which results in $O(s)$ computations. With a total number of $(B+1)^2$ vertices, we can conclude that Algorithm 1 needs $O(sB^2)$ computations to solve the maximum-cost path search problem on G.

3.2. Rate Allocation using EXIT charts

With Algorithm 1, we can calculate for any particular realization of $\mathbf{f}_e$ the convergence curve $\bar{\mathbf{p}} \in \mathcal{P}$ that approximately solves (8). The maximal achievable rate of the k-th user, corresponding to $\bar{\mathbf{p}}$, is then given by the area under the projected equalizer EXIT function $f_k^{(\bar{\mathbf{p}})}$. To closely approach this rate, the decoder EXIT function $f_{d,k}$ of the k-th user should be as close as possible to $f_k^{(\bar{\mathbf{p}})}$. Furthermore, a narrow tunnel between the both functions for all $k = 1, .., K$ should be open so that the conditions for convergence, as stated in (4), are satisfied. In practice, however, optimizing $f_{d,k}$ by adjusting the available code parameters such that the code optimality in a strict sense is always guaranteed is not possible, since the code parameters are presumably limited. Thus, a practical approach for rate allocation is to *select*, for each k, the code $\mathcal{C}_n \in \mathcal{C}$ with the highest possible rate from a finite code set $\mathcal{C} = \{\mathcal{C}_1, ..., \mathcal{C}_m\}$, with m being the number of the codes in the set, with which convergence is achieved, while its decoder EXIT function $f_{d,k}^{(\mathcal{C}_n)}$ best fits to $f_k^{(\bar{\mathbf{p}})}$.

Now, we propose a simple scheme for rate allocation to the K users, which is summarized as follows.

1. Determine $\mathbf{f}_e$ for given ($\mathbf{H}_1$,..,$\mathbf{H}_K$, σ^2).

2. Calculate the convergence curve $\bar{\mathbf{p}} \in \mathcal{P}$ that solves (8) using Algorithm 1.

3. Project the corresponding EXIT space curve $\mathbf{w}_k(\bar{\mathbf{p}}(t))$, for each k, onto the $I_{e,k}$-$I_{d,k}$-plane: $f_k^{(\bar{\mathbf{p}})}(I_{d,k}) = f_{e,k}(\bar{\mathbf{p}}), I_{d,k} = \bar{p}_k(t), \forall t \in \mathcal{I}$.

4. To obtain a high code rate, select the channel code that satisfies:

$$r_{c,k} = \max_{\mathcal{C}_n \in \mathcal{C}} \Big\{ r(\mathcal{C}_n) \Big| f_k^{(\bar{\mathbf{p}})}(I_{d,k}) > f_{d,k}^{(\mathcal{C}_n)}(I_{d,k}) \ , \quad \forall I_{d,k} \in [0,1) \Big\}, \quad (16)$$

   where $r(\mathcal{C}_n)$ denotes the code rate of $\mathcal{C}_n$.

5. If $r_{c,k}$ is null, select the code with the lowest possible rate in $\mathcal{C}$, $r_{c,k} = \min_{\mathcal{C}_n \in \mathcal{C}} r(\mathcal{C}_n)$.

6. Output the selected code rates $\{r_{c,1}, ..., r_{c,K}\}$.

The above scheme requires the computation of the equalizer EXIT vector-function $\mathbf{f}_e$ (step 1) *for each channel realization*, which is not feasible, in practice, because of the high computational effort [12]. Therefore, we use in this paper the semi-analytical technique proposed in [4] to compute $\mathbf{f}_e$. For more details, please refer to [4].

4. NUMERICAL RESULTS

In this section, results of the simulations conducted to evaluate the throughput enhancement of the proposed RAAL technique are presented. We consider an 2-user single-carrier block-cyclic transmission system with each transmit block having $Q = 128$ coded BPSK symbols over Rayleigh fading channels with $L = 5$ and $L = 32$ path components and equal average power delay profile.

4.1. Comparison with Capacity

Fig. 5 shows the average (maximal) sum rate of the 2-user turbo system when ideally designed channel codes are allocated to the both users with respect to Algorithm 1. The rate of each user has been calculated by averaging $R_k = \int_0^1 f_k^{(\bar{\mathbf{p}})}(I_{d,k})\mathrm{d}I_{d,k}$ over a large number of channel realizations. For comparison, the ergodic capacity with i.i.d Gaussian inputs and the ergodic constellation constrained capacity with binary-input distribution are also shown. The constrained capacity result has been obtained by Monte Carlo simulations, as described in [14]. Comparing the numerical results in Fig 5, we find that the maximal sum rate achieved with the proposed RAAL technique is only slightly worse than the constrained capacity at moderate to high E_s/N_0-values. In addition, we observe that there is almost no difference between the unconstrained capacity and the constrained capacity as well as the maximal sum rate of the 2-user system in the low E_s/N_0 region.

see the program

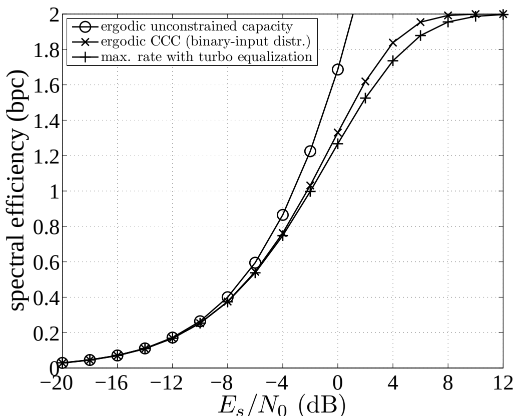


Fig. 5. Comparison of ergodic unconstrained capacity and constrained capacity with binary-input distribution to the maximal achievable rate of an 2-user system with turbo SC FD-MMSE equalization over frequency selective ($L = 5$) Rayleigh channels.

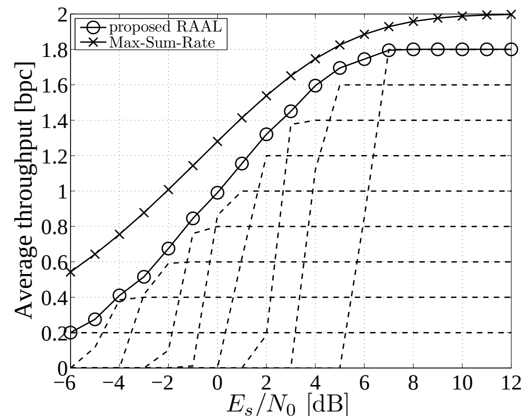


Fig. 6. Average total throughput of the both users versus E_s/N_0 for the proposed RAAL scheme and for ARQ with fixed coding rates $r_{c,1/2} = 0.1(n)$ with $n = 1, .., 9$ (dashed curves, from bottom to top). $L = 32$, $Q = 128$.

4.2. RAAL with practical codes

The rate-compatible punctured SCCCs proposed in [13], consisting of a rate-r_c outer encoder and a recursive rate-1 inner encoder with polynomials $(g_r, g_0) = (3, 2)$ (g_r denotes the feedback polynomial) in octal notation, were assumed for the RAAL to the both users. The outer encoder of the SCCC was selected from a set of $m = 17$ subcodes with rates of $r_c = 0.05(1 + n)$, $n = 1, .., m$. The 17 subcodes are constructed from a systematic, rate-1/2 mother code with polynomials $(g_r, g_1) = (23, 35)$, where higher rate codes are obtained through puncturing, and lower rate codes are obtained by adding more generators and using puncturing, as specified in [13]. We generated the EXIT functions of the SCCCs using Monte Carlo simulations, where 30 iterations between the inner and outer decoder were assumed. The block length of the input sequence to the outer encoder was 16384. Thus, the channel gains of the four links were assumed to be static over a frame of $N = 128$ transmitted blocks. The turbo equalizer performed 20 iterations between the equalizer and the both SCCC decoders. The grid spacing parameter in Algorithm 1 is set at $h = 0.02$ ($B = 50$), which is sufficiently small, yielding a fine grid of $(B + 1)^2 = 2601$ points. Note that in general h is a design parameter to be chosen appropriately depending on the particular realization of the equalizer EXIT vector function $\mathbf{f}_e$. Optimizing the grid spacing parameter h, given the particular realization of $\mathbf{f}_e$, is an interesting research topic, but beyond the primary scope of this paper.

When evaluating the throughput efficiency, a selective-repeat ARQ system with infinite buffer size [12] was assumed. Fig. 6 shows the average total throughput (in bpc) of the 2-user system versus the average E_s/N_0 per receive antenna. For comparison, the maximal rate with respect to Algorithm 1 (ideally designed channel codes are allocated to the both users) is shown. Also shown is the average throughput performance for an ARQ scheme with fixed coding rates $r_{c,1/2} = 0.1(n)$, $n = 1, .., 9$ of the both users. As observed in Fig. 6, substantial throughput gain is obtained with the proposed RAAL technique over the fixed rate ARQ. Further, we find that the throughput performance is only 1.4 dB away (in the high E_s/N_0 region) from the maximal information-rate achieved with ideally designed codes for each channel realization. Notably, however, there is still a gap of 2.3 dB in the low E_s/N_0 region between the throughput achieved by the proposed RAAL scheme and the maximal rate. By using more flexible coding techniques, such as irregular SCCCs that allow a code design which is optimal for the corresponding channel realization, the loss in E_s/N_0, or equivalently in throughput, can be further reduced.

5. CONCLUSION

We have examined the achievable rate region of frequency selective K-user MACs with SC FD-MMSE turbo equalization. It has been shown that with the aid of the equalizer EXIT functions the optimal RAAL to the K users can be formulated as an integral optimization problem. Further, we have proposed a heuristic approach based on the dynamic programming principle for solving this problem. In addition, we have also proposed a practical RAAL scheme based on a code *selection* algorithm. Numerical results for the 2-user MAC with SC FD-MMSE turbo equalization show that substantial throughput gain can be observed with the proposed RAAL technique over the fixed rate ARQ. Further, we have found that the spectral efficiency of the proposed practical RAAL technique using rate-compatible punctured SCCCs for each user, is only 1.4 dB away from the maximal rate achieved with ideally designed codes for each channel realization in the high E_s/N_0 region.

6. REFERENCES

[1] C. Douillard et al., ”Iterative correction of intersymbol interference: Turbo equalization,” *European Trans. Telecomm.*, vol. 6, pp. 507-511, Sept. 1995.

[2] X. Wang, and H. V. Poor, ”Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” *IEEE Trans. Commun.*, vol. 47, pp. 1046-10611, July 1999.

[3] M. Tüchler, and J. Hagenauer, ”Linear time and frequency domain Turbo equalization,” *Proc. IEEE Veh. Technol. Conf. 2001*, vol. 2, pp. 1449-1453, May 2001.

[4] K. Kansanen, and T. Matsumoto, ”An analytical method for MMSE MIMO turbo equalizer EXIT chart computation,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 59-63, Jan. 2007.

[5] S. ten Brink, ”Convergence behavior of iteratively decoded parallel concatenated codes,” *IEEE Trans. Comm.*, vol. 49, pp. 1727-1737, Oct. 2001.

[6] A. Ashikhmin, G. Kramer, and S. ten Brink, ”Extrinsic information transfer functions: model and erasure channel properties,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 2657-2673, Nov. 2004.

[7] S. ten Brink, G. Kramer, and A. Ashikhmin, ”Design of low-density parity-check codes for modulation and detection,” *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670-678, April 2004.

[8] S. ten Brink, and G. Kramer, ”Design of repeat-accumulate codes for iterative detection and decoding,” *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2764-2772, Nov. 2003.

[9] J. Jost, X. Li-Jost, ”Calculus of variations,” *Cambridge Studies in Advanced Mathematics*, vol. 64, Cambridge University Press, Cambridge, 1998.

[10] D.P. Bertsekas, ”Dynamic programming and optimal control,” *3th ed. Athena Scientific*, Belmont, MA, 2005.

[11] F. Brännström, L.K. Rasmussen, A.J. Grant, ”Convergence analysis and optimal scheduling for multiple concatenated codes,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 3354-3364, Sept. 2005.

[12] S. IBI et al., ”EXIT chart-aided adaptive coding for MMSE turbo equalization with multilevel BICM in frequency selective MIMO channels,” *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3757-3769, Nov 2007.

[13] M. Tüchler, ”Design of serially concatenated systems depending on the block length,” *IEEE Trans. Comm.*, vol. 52, pp. 209-218, Feb. 2004.

[14] Z. Zhang et al., ”Achievable information rates and coding for MIMO systems over ISI channels and frequency-selective fading channels,” *IEEE Trans. Comm.*, vol. 52, pp. 1698-1710, Oct. 2004.

see the program

Cooperative Regions For Coded Cooperation Over Time-Varying Fading Channels

P. Castiglione[1], M. Nicoli[2], S. Savazzi[2] and Thomas Zemen[1]
[1] Forschungszentrum Telekommunikation Wien, Donau-City-Strasse 1, A-1220 Vienna, Austria
[2] Dip. di Elettronica e Informazione, Politecnico di Milano, Piazza L. da Vinci 32, I-20133 Milano, Italy
e-mail: {castiglione, thomas.zemen}@ftw.at, {nicoli,savazzi}@elet.polimi.it

***Abstract*—The performance analysis of coded cooperation has been mainly focused on two extreme cases of channel variability, i.e. the block-fading (BF) and the fast-fading (FF) model. In more practical propagation environments the fading correlation across time depends on the level of user mobility. This paper analyzes the effects of time-selective fading on the performance of coded cooperation by providing an analytical framework for the error rate evaluation as a function of the mobility degree of the mobile station (MS) and of the quality of the inter-MS channel. The purpose is to evaluate the conditions on the propagation settings where the additional exploitation of spatial diversity (when time-diversity is available) provided by cooperative transmission is able to enhance substantially the performance of the non-cooperative transmission. We show that coded cooperation can outperform the non-cooperative (coded and bit-interleaved) transmission only up to a certain degree of mobility. The *cooperative region* is defined as the collection of mobility settings for which coded cooperation can be regarded as a competitive strategy compared to non-cooperative transmission. Contrary to what has been previously shown for BF channels, we demonstrate that the inter-MS channel quality plays a key role in the definition of the cooperative region.**

I. Introduction

Cooperative communication [1] was originally motivated as a method to bring *spatial diversity* gain in networks of single antenna terminals through redundant transmissions from multiple mobile stations (MSs). The integration of the user cooperation idea with channel coding was then proposed in [2] to further improve the cooperative link performance. Forward error correcting (FEC) codes are used by two or more mobile stations that cooperate by transmitting to the base station (BS) incremental redundancy for the partners. This approach was shown to provide significant performance enhancements, compared to conventional non-cooperative transmission, primarily in networks with fixed or nomadic terminals: in such a situation *time diversity* cannot be exploited, while spatial diversity is beneficial as it reduces the fading impairments.

In the literature the performance analysis of coded cooperation has been mainly focused on two extreme cases of channel variability, i.e. the block-fading (BF) and the fast-fading (FF) model. However, more practical propagation environments, e.g. in vehicular networks, experience fading variations across time with a degree of correlation depending on the level of MS mobility. In these cases, temporal diversity can be exploited to improve the link reliability by the joint use of channel coding and bit-interleaved modulation [3]. Hence, it is crucial to investigate to what extent the additional exploitation of

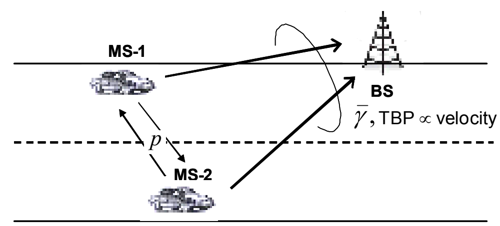


Fig. 1. System setting: the average signal-to-noise ratio (SNR) $\bar{\gamma}$ and the time-bandwidth product (TBP) of the uplink channels; the block error probability p of the inter-MS channels.

spatial diversity through collaborative transmission is able to provide substantial performance enhancements compared to non-cooperative transmission.

The contribution of this paper is twofold. Firstly, the effects of MS mobility on the performance of coded cooperation are analyzed by providing an analytical framework for the error rate evaluation as a function of the degree of fading variations. The variation degree is here measured in terms of the time-bandwidth product (TPB), that is the product of the codeword duration (for bit interleaved coded modulation) and the Doppler bandwidth. Our performance analysis is carried out in closed form and extends the results obtained by simulations in [4]. The analysis takes into account the reliability of the communication link between the cooperating MSs, by considering the block error probability p over the inter-MS link as a penalty factor that limits the cooperation performance (an overview of the system setting is shown in Fig. 1).

The above analysis is then used to evaluate the performance gain provided by coded cooperation with respect to a non-cooperative system for varying degree of MS mobility and for different channel state conditions. We show that coded cooperation can outperform the non-cooperative (coded and bit-interleaved) transmission only up to a certain degree of mobility. Beyond this limiting threshold, the temporal diversity gain, which is made available by channel coding and bit-interleaving, dominates the performances, while the spatial diversity gain offered by collaborative transmissions provides only marginal improvements. Analytical and numerical results show that the mobility degree threshold, beyond which coded cooperation is no more advantageous, decreases with the

see the program

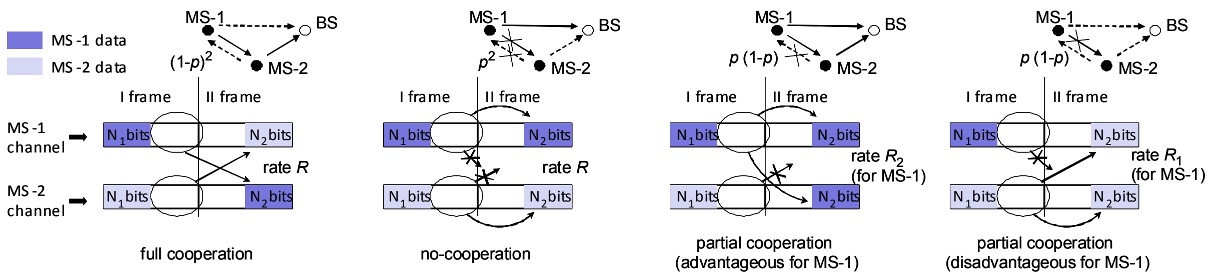


Fig. 2. Possible configurations of the coded cooperation scheme.

quality of the inter-MS channel and strongly depends on the quality of the relayed links (the uplink signal-to-noise ratio $\bar{\gamma}$ - SNR) from the cooperating MSs towards the common BS. To better highlight these results, we evaluate the *cooperative region* [5] as the collection of mobility and channel settings for which coded cooperation can be regarded as a competitive option compared to non-cooperative transmission.

The paper is organized as follows. In Section II we give a brief description of the system and the channel model. Section III presents the derivation of the analytical upperbound on the bit error rate (BER). We derive the statistical distribution of the SNR at the decision variable (here referred to as *effective SNR*) [6], based on the knowledge of the fading channel autocorrelation over the transmitted data block (a similar approach has been used for performance evaluation in frequency-selective OFDM systems [7], [8]). Section IV contains both numerical and analytical results, for the validation of the analytical BER and the investigation of the MSs' mobility effects on coded cooperation. Conclusions are drawn in Section V.

II. System model

We consider the transmission of rate-R coded data from two mobile stations, MS-1 and MS-2, towards a common BS through two orthogonal channels by frequency division multiple access (FDMA). The two channels are assumed to be subject to *independent* time-selective (due to MS mobility) frequency-flat fading. Coded cooperation is carried out according to the scheme introduced in [2], by sending portions of each MS data over the two independent channels so that a diversity gain is provided, as briefly summarized below. Each MS encodes its data block of K information bits by means of a rate-compatible punctured convolutional (RCPC) code [9] that yields an overall codeword of $N = K/R$ bits. This codeword is divided through puncturing into two sub-codewords of length N_1 and N_2, with $N = N_1 + N_2$: the first subset is the punctured codeword of rate $R_1 = K/N_1$, the second one is the set of removed parity bits. The sub-codewords are then transmitted into two subsequent time frames. In the first frame each MS broadcasts the first sub-codeword, that is received by the cooperating partner and the BS. If the partner successfully decodes the first sub-codeword (this is determined by a cyclic redundancy check - CRC - code or any other error detection code), then it will compute and transmit the N_2 additional parity bits in the second frame. At the BS this incremental redundancy is used for de-puncturing the rate-R_1 codeword received in the first frame, thus obtaining the initial rate-R codeword. Hence, the level of cooperation is quantified by $\alpha = \frac{N_2}{N}$. If the partner cannot successfully decode the MS' first-frame data, it will transmit its own N_2 code bits during the second frame. The latter rule avoids error propagation: each MS is forced to stop cooperating if the inter-MS transmission fails due to various possible reasons related to the inter-MS channel conditions (e.g. a deep fade, channel estimation and/or synchronization errors, etc.). Notice that even if the MSs are close and no obstacle stands on their line of sight (e.g. two vehicles running adjacently on the motorway), the media access control (MAC) protocol at one MS could anyhow decide to stop the ongoing cooperation with the partner.

Characterizing the different causes of no cooperation goes beyond the scope of the present work. Here we assume that no cooperation occurs only due to a block error event in the inter-MS transmission, with block error probability p being the same for both MSs. Furthermore, we assume, as worst case, that the inter-MS channels are independent (which is true in FDMA). Under these assumptions, four different configurations of coded cooperation can occur [2] with probability depending on the inter-MS block error rate p, as specified below (see also Fig. 2)[1]:

- *Full cooperation* ($\Theta = 1$). Both MSs successfully decode each other during the first frame and can transmit the partner's code bits during the second frame. Probability: $\Pr(\Theta = 1) = (1-p)^2$.
- *No-cooperation* ($\Theta = 2$). Both MSs fail to decode the partner's first frame and transmit their own code bits during the second frame. Probability: $\Pr(\Theta = 2) = p^2$.
- *Partial cooperation - advantageous for MS-1* ($\Theta = 3$). MS-1 cannot decode MS-2's data during the first frame, while MS-2 successfully decodes MS-1's data. During the second frame both MSs transmit the N_2 parity bits for MS-1. Probability: $\Pr(\Theta = 3) = p(1-p)$.
- *Partial cooperation - disadvantageous for MS-1* ($\Theta = 4$). The same as the previous case, with switched MSs.

[1] To simplify the reasoning, the cases of partial cooperation are described from the point of view of MS-1, while MS-2 is the partner.

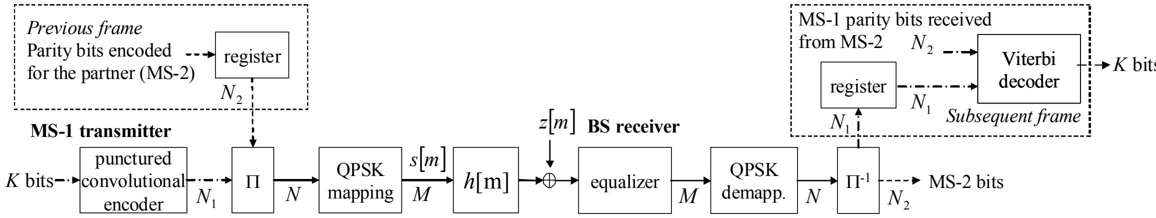


Fig. 3. Uplink system model for MS-1 successfully cooperating with MS-2.

Probability: $\Pr(\Theta = 4) = \Pr(\Theta = 3) = p(1 - p)$.

Notice that the transmission of one codeword is temporally concatenated with the transmission of the previous one (see Fig. 2). Therefore, splitting the codeword transmission into two frames does not generally modify the overall code rate R, apart from the case of partial cooperation in which the code-rate raises to R_1 for the disadvantaged MS and lowers to $R_2 = K/(N + N_2)$ for the advantaged MS.

The complete system model is depicted in Fig. 3 for each MS-BS link (only the case $\Theta = 1$ is represented for simplicity). Bit-interleaved quadrature phase shift keying (QPSK) modulation is assumed [11] with symbol rate $1/T_\mathrm{S}$. The baseband-equivalent discrete-time signal transmitted by MS-i, with $i \in \{1, 2\}$, is $s_i[m] = \sqrt{E_\mathrm{S}}q_i[m]$, $m = 1, \dots, M$, where $M = N/2$ is the number of symbols per frame, E_S is the transmitted energy per symbol, and $q_i[m] = (\pm 1 \pm j)/\sqrt{2}$ is the QPSK symbol at time mT_S. The corresponding signal received at the BS is then

$$y_i[m] = h_i[m]s_i[m] + z[m], \tag{1}$$

where $z[m] \sim \mathcal{CN}(0, \sigma_n^2)$ denotes the complex additive white Gaussian noise (AWGN) at all receivers, with variance σ_n^2. The time-variant Rayleigh-fading channel for the link between MS-i and the BS is $h_i[m] \sim \mathcal{CN}(0, \Omega_i)$ with variance Ω_i. The fading process is assumed to be wide-sense stationary (up to the second-order statistics) with Clarke's auto-correlation function given by [10]

$$R_i[k] = \mathrm{E}\{h_i[m]h_i^*[m + k]\} = \Omega_i J_0(2\pi k \nu_{\mathrm{D}i}), \tag{2}$$

where J_0 is the zeroth-order Bessel function of the first kind, $\nu_{\mathrm{D}i}$ is the one-sided normalized Doppler bandwidth $\nu_{\mathrm{D}i} = \frac{v_i f_\mathrm{C}}{c_0} T_\mathrm{S}$, v_i is the MS-i velocity, f_C the carrier frequency and c_0 the speed of light. The parameter $\nu_{\mathrm{D}i}$ is a measure of the temporal variability of the channel. A more meaningful parameter for coded transmissions is the time-bandwidth product, here defined as

$$\mathrm{TBP}_i = 2M\nu_{\mathrm{D}i}, \tag{3}$$

where $2M$ is the temporal duration of the codeword expressed in symbol times (two frames), i.e. the time interval in which the interleaved code can exploit the temporal diversity. By definition (3), TBP represents the velocity of MS-i in terms of number of wavelengths travelled during the transmission of two frames.

According to the Rayleigh fading assumption, the instantaneous SNR, defined as

$$\gamma_i[m] = |h_i[m]|^2 \frac{E_\mathrm{S}}{\sigma_n^2}, \tag{4}$$

exhibits an exponential distribution with mean $\overline{\gamma_i} = \Omega_i \frac{E_\mathrm{S}}{\sigma_n^2}$ [6].

At the receiver side, coherent equalization is carried out using perfect knowledge for the channel $h_i[m]$, followed by de-mapping, de-interleaving and decoding, as illustrated in Fig. 3.

III. Performance analysis

In the following, the derivation of the analytical upperbound on the average BER is presented. We first assume that the MSs are always cooperating, as for an error-free inter-MS channel with $p = 0$ (Sect. III-A). Then in Sect. III-B the analysis is extended to include the other cases of coded cooperation.

A. Analysis for ideal (error-free) inter-MS channel

According to the union bound approach [11], the average bit error probability, for the full cooperation case ($\Theta = 1$), P_b at the Viterbi decoder output is:

$$P_\mathrm{b} \leq \frac{1}{k} \sum_{d \geq d_\mathrm{free}} \sum_{\mathbf{c} \in \mathcal{E}(d)} \beta(\mathbf{c}) P(\mathbf{c}), \tag{5}$$

where k is the number of input bits for each branch of the convolutional code trellis, d_free is the free distance, $\mathcal{E}(d)$ is the set of error events $\mathbf{c}$ at a certain Hamming distance d, $\beta(\mathbf{c})$ is the Hamming weight of the input sequence corresponding to $\mathbf{c}$ and $P(\mathbf{c})$ is the average pairwise error probability (PEP). The average PEP $P(\mathbf{c})$ is the probability of detecting the codeword $\mathbf{c}$ instead of the transmitted all-zero codeword.

Let $\mathcal{T}_\mathbf{c} = \{\tau_{\mathbf{c},1}, \dots, \tau_{\mathbf{c},d}\}$ be the set of time instants associated with the d error bits in $\mathbf{c}$, and $\tilde{\mathbf{h}} = \sqrt{E_\mathrm{s}}\, [h(\tau_{\mathbf{c},1}) \cdots h(\tau_{\mathbf{c},d})]^\mathrm{T} / \sigma_\mathrm{n}$ be the vector that gathers the corresponding channel gains scaled by $\sqrt{E_\mathrm{s}}/\sigma_\mathrm{n}$. The average PEP $P(\mathbf{c})$ can be calculated as [3]:

$$P(\mathbf{c}) = \int_0^\infty Q\left(\sqrt{2\gamma_\mathrm{eff}}\right) p\left(\gamma_\mathrm{eff}\right) d\gamma_\mathrm{eff}, \tag{6}$$

where γ_eff (effective SNR) is the sum of the SNR variates that are experienced over the time instants $\mathcal{T}_\mathbf{c}$ [7], or, equivalently, the sum of the squared magnitudes of the vector $\tilde{\mathbf{h}}$'s entries:

$$\gamma_\mathrm{eff} = \sum_{k \in \mathcal{T}_\mathbf{c}} \gamma(k) = \left\|\tilde{\mathbf{h}}\right\|^2. \tag{7}$$

see the program

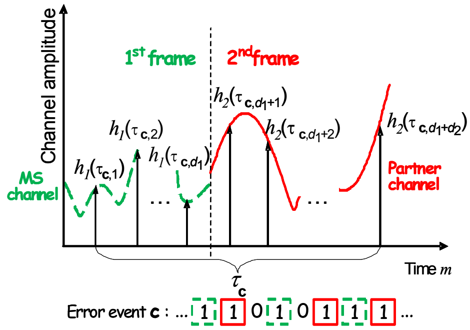


Fig. 4. Example of the correlated fading observed along the error event in the case of full cooperation ($\Theta = 1$).

Its probability density function (pdf), $p(\gamma_{\text{eff}})$, clearly depends on the correlation of the channel gains contained in $\tilde{\mathbf{h}}$. We observe that $\tilde{\mathbf{h}}$ is a zero-mean complex Gaussian random vector, $\tilde{\mathbf{h}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R_c})$, with covariance $\mathbf{R_c} = \mathrm{E}[\tilde{\mathbf{h}} \cdot \tilde{\mathbf{h}}^{\mathrm{H}}]$ whose entries are samples of the auto-correlation function (2). The distribution of the effective SNR is here derived based on the knowledge of the correlation matrix $\mathbf{R_c}$, by extending the approach in [8] to the cooperative scenario with time-selective fading channels.

Recalling that the codeword is partitioned into two frames due to the coded cooperation scheme, it should be observed that, for the case of full cooperation ($\Theta = 1$), the d error bits in $\mathbf{c}$ are split into two groups of bits coming from the MS's and the partner's uplink channels (see Fig. 4). Let us consider for instance the codeword of MS-1, the time instants associated with the first and the second groups are here indicated as $\mathcal{T}_{\mathbf{c};1}$ (d_1 elements) and $\mathcal{T}_{\mathbf{c};2}$ (d_2 elements), respectively, with $\mathcal{T}_{\mathbf{c}} = \mathcal{T}_{\mathbf{c};1} \bigcup \mathcal{T}_{\mathbf{c};2}$. Accordingly, the channel vector is $\tilde{\mathbf{h}} = [\tilde{\mathbf{h}}_{1,d_1}^{\mathrm{T}}, \tilde{\mathbf{h}}_{2,d_2}^{\mathrm{T}}]^{\mathrm{T}}$, where $\tilde{\mathbf{h}}_{1,d_1} = \sqrt{E_{\mathrm{s}}}\,[h_1(\tau_{\mathbf{c},1}) \cdots h_1(\tau_{\mathbf{c},d_1})]^{\mathrm{T}}/\sigma_{\mathrm{n}}$ and $\tilde{\mathbf{h}}_{2,d_2} = \sqrt{E_{\mathrm{s}}}\,[h_2(\tau_{\mathbf{c},d_1+1}) \cdots h_2(\tau_{\mathbf{c},d_1+d_2})]^{\mathrm{T}}/\sigma_{\mathrm{n}}$ gather the channel coefficients for, respectively, the MS-1 and MS-2 uplink channels, at time instants $\mathcal{T}_{\mathbf{c};1}$ and $\mathcal{T}_{\mathbf{c};2}$. In order to derive the effective SNR's statistical distribution, we introduce the eigenvalue decomposition (EVD) of the covariance matrices of the two channel vectors:

$$
\begin{aligned}
[\mathbf{R}_{\mathbf{c},1}]_{d_1 \times d_1} &= \mathrm{E}[\tilde{\mathbf{h}}_{1,d_1} \cdot \tilde{\mathbf{h}}_{1,d_1}^{\mathrm{H}}] = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^{\mathrm{H}}, \qquad (8)\\
[\mathbf{R}_{\mathbf{c},2}]_{d_2 \times d_2} &= \mathrm{E}[\tilde{\mathbf{h}}_{2,d_2} \cdot \tilde{\mathbf{h}}_{2,d_2}^{\mathrm{H}}] = \mathbf{U}_2 \mathbf{\Lambda}_2 \mathbf{U}_2^{\mathrm{H}}. \qquad (9)
\end{aligned}
$$

The matrices $\mathbf{\Lambda}_1 = \mathrm{diag}\,[\lambda_{1,1}, \ldots, \lambda_{1,r_1}]$ and $\mathbf{\Lambda}_2 = \mathrm{diag}\,[\lambda_{1,1}, \ldots, \lambda_{1,r_2}]$ contain the non-zero eigenvalues, with $r_1 = \mathrm{rank}\,[\mathbf{R}_{\mathbf{c},1}]_{d_1 \times d_1} \leq d_1$ and $r_2 = \mathrm{rank}\,[\mathbf{R}_{\mathbf{c},2}]_{d_2 \times d_2} \leq d_2$. $\mathbf{U}_1$ and $\mathbf{U}_2$ gather the corresponding eigenvectors. We recall that the two MSs' channels are assumed to be independent, hence it is $\mathrm{E}[\tilde{\mathbf{h}}_{1,d_1} \cdot \tilde{\mathbf{h}}_{2,d_2}^{\mathrm{H}}] = \mathbf{0}$ and the correlation matrix can be written as

$$
\begin{aligned}
\mathbf{R_c} &= \begin{bmatrix} [\mathbf{R}_{\mathbf{c},1}]_{d_1 \times d_1} & \mathbf{0} \\ \mathbf{0} & [\mathbf{R}_{\mathbf{c},2}]_{d_2 \times d_2} \end{bmatrix} \qquad (10)\\
&= \underbrace{\begin{bmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix}}_{\mathbf{\Lambda}} \underbrace{\begin{bmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 \end{bmatrix}^{\mathrm{H}}}_{\mathbf{U}^{\mathrm{H}}},
\end{aligned}
$$

where $\mathbf{\Lambda} = \mathrm{diag}\,[\lambda_{1,1}, \ldots, \lambda_{1,r_1}, \lambda_{2,1}, \ldots, \lambda_{2,r_2}]$ collects the eigenvalues of $\mathbf{R_c}$ and $\mathbf{U}$ the corresponding eigenvectors. Notice that it is $r = \mathrm{rank}[\mathbf{R_c}] = r_1 + r_2$.

Using the EVD (10), the effective SNR can now be rewritten as $\gamma_{\text{eff}} = \|\mathbf{b}\|^2 = \sum_{i=1}^{r} b_i^2$, in terms of the projection of the channel onto the r-dimensional column-space of $\mathbf{R_c}$: $\mathbf{b} = \mathbf{U}^{\mathrm{H}}\tilde{\mathbf{h}} = [b_1 \cdots b_r]^{\mathrm{T}}$. Notice that $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda})$, thus the effective SNR is the sum of r independent exponentially distributed variates having as mean values the eigenvalues of $\mathbf{R_c}$. It follows that the pdf of γ_{eff} exhibits the moment-generating function (MGF) [6], [12] :

$$
\mathbf{M}_{\gamma_{\text{eff}}}(s) = \prod_{i=1}^{r_1} \frac{1}{1 - \lambda_{1,i} s} \prod_{j=1}^{r_2} \frac{1}{1 - \lambda_{2,j} s}. \qquad (11)
$$

The integral over γ_{eff} in (6) can now be derived using the alternate integral form of the Q-function [13] and the well known MGF method [12]. We get the average PEP

$$
\begin{aligned}
P(\mathbf{c}) &= \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{i=1}^{r_1} \left(1 + \frac{\lambda_{1,i}}{\sin^2 \vartheta}\right)^{-1} \prod_{j=1}^{r_2} \left(1 + \frac{\lambda_{2,j}}{\sin^2 \vartheta}\right)^{-1} d\vartheta \quad (12)\\
&\leq \frac{1}{2} \prod_{i=1}^{r_1} \frac{1}{1 + \lambda_{1,i}} \prod_{j=1}^{r_2} \frac{1}{1 + \lambda_{2,j}}, \qquad (13)
\end{aligned}
$$

upperbounded in (13) using $\sin^2 \vartheta \leq 1$.

We observe that each MS interleaves its own bits and the parity bits computed for the other MS before mapping them into symbols. It follows that the $d = d_1 + d_2$ non-zero bits of the error event $\mathbf{c}$ can appear within the two time frames in several possible configurations, each corresponding to a different shift of $\mathbf{c}$ at the input of the Viterbi decoder. To get an upperbound, we will select for each error event $\mathbf{c}$ the most probable configuration among all these possible shifts by finding the one that maximizes (12).

B. Extension to imperfect inter-MS channel

In Sect. III-A, the average BER performance (for MS-1) has been analyzed only for the full cooperation case ($\Theta = 1$). The average BER over all the possible cooperation configurations $\Theta = \{1, 2, 3, 4\}$ can be evaluated as:

$$
P_{\mathrm{b}} \leq \frac{1}{k} \sum_{d \geq d_{\text{free}}} \sum_{\mathbf{c} \in \mathcal{E}(d)} \beta(\mathbf{c}) \left[\sum_{\Theta=1}^{4} \Pr(\Theta) P(\mathbf{c}\,|\Theta)\right]. \qquad (14)
$$

The conditioned PEP $P(\mathbf{c}\,|\Theta = 1)$ for the full cooperation case is evaluated as in (12) - (13) based on the eigenvalue decomposition of the matrix (10). For the other cases $\Theta =$

see the program

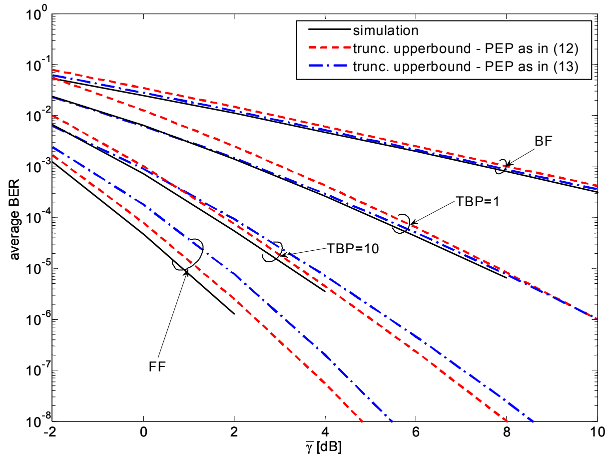


Fig. 5. Performance of full cooperation ($\Theta = 1$). Code-rate $R = 1/4$.

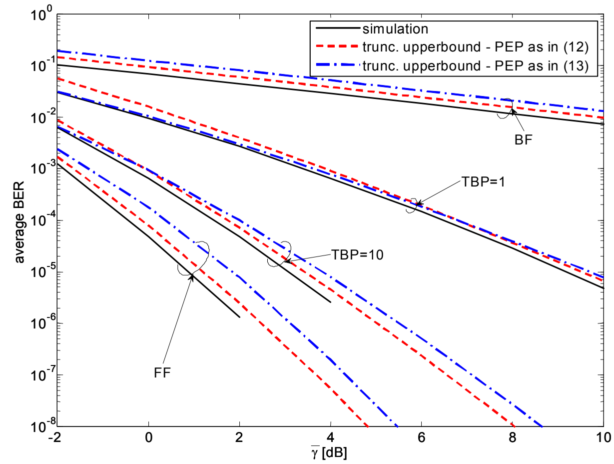


Fig. 6. Performance of no-cooperation ($\Theta = 2$). Code-rate $R = 1/4$.

$\{2,3,4\}$, the conditioned PEP can be obtained similarly using the eigenvalues of the following matrices:

$$\mathbf{R_c}(\Theta=2) = \begin{bmatrix} [\mathbf{R_{c,1}}]_{d_1\times d_1} & [\mathbf{R_{c,1}}]_{d_1\times d_2} \\ \left[[\mathbf{R_{c,1}}]_{d_1\times d_2}\right]^{\mathrm{H}} & [\mathbf{R_{c,1}}]_{d_2\times d_2} \end{bmatrix}, \tag{15}$$

$$\mathbf{R_c}(\Theta=3) = \begin{bmatrix} \mathbf{R_c}(\Theta=2) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & [\mathbf{R_{c,2}}]_{d_2\times d_2} \end{bmatrix}, \tag{16}$$

$$\mathbf{R_c}(\Theta=4) = [\mathbf{R_{c,1}}]_{d_1\times d_1}, \tag{17}$$

where $[\mathbf{R_{c,1}}]_{d_2\times d_2} = E[\tilde{\mathbf{h}}_{1,d_2}\tilde{\mathbf{h}}_{1,d_2}^{\mathrm{H}}]$ is the auto-correlation of the MS-1 uplink channel gains associated to the d_2 second-frame error bits: $\tilde{\mathbf{h}}_{1,d_2}=\sqrt{E_{\mathrm{s}}}\,[h_1(\tau_{\mathbf{c},d_1+1})\cdots h_1(\tau_{\mathbf{c},d_1+d_2})]^{\mathrm{T}}/\sigma_{\mathrm{n}}$. The matrix $[\mathbf{R_{c,1}}]_{d_1\times d_2} = E[\tilde{\mathbf{h}}_{1,d_1}\tilde{\mathbf{h}}_{1,d_2}^{\mathrm{H}}]$ denotes the cross-correlation between $\tilde{\mathbf{h}}_{1,d_1}$ and $\tilde{\mathbf{h}}_{1,d_2}$. Notice that the sum in (16) is due to the assumption that the two sets of N_2 bits received from MS-1 and MS-2 during the second frame are optimally combined at the receiver using a maximum ratio combiner (MRC).

IV. Simulation Results

In this Section we provide both numerical and analytical results on the performance of coded cooperation. The numerical results are obtained by Monte-Carlo simulations, BER results are obtained by averaging over a large number of frames. In Sect. IV-A, we validate the exact derivation (12) and the upperbound (13) on the average PEP by evaluating the average BER performance of coded cooperation for ideal inter-MS channel ($\Theta = 1$). Next, by removing the assumption of error-free inter-MS channel, more insight is given on the conditions for which coded cooperation can be regarded as competitive in terms of perfomance compared to non-cooperative (direct) transmission (Sect. IV-B): the cooperative regions for coded cooperation over time-varying channels are defined through examples.

A. Performance limits of coded cooperation

The results presented here refer to the case of ideal inter-MS channel, the best case for coded cooperation. The aim is to show the mobility conditions for which coded cooperation gains significantly. We employ a rate $R = 1/4$ RCPC mother-code [9], with octal generators (23, 35, 27, 33) and free distance $d_{\text{free}} = 15$. The mother-code is punctured, obtaining a rate $R_1 = 1/2$ sub-codeword for the first frame transmission ($\alpha = 50\%$). A soft-input hard-output Viterbi decoder is implemented at the receiver-side [11] and the error events $\mathbf{c}$ are found via computer-enumeration. The coded-block length is $N = 512$ bits, resulting in $M = 256$ QPSK symbols. Before symbol mapping, the coded bits are interleaved by a block bit-interleaver, which writes the input codeword row by row in a (128×4) matrix, and then reads it column by column. The two MSs transmit on independent time-varying flat-fading channels with average SNR $\bar{\gamma}_i$ and time-bandwidth product TBP_i, with $i \in \{1,2\}$. In Fig. 5, 6 and 7 the fading statistics are the same for both uplink channels, i.e. $\bar{\gamma} = \bar{\gamma}_i$ and $\mathrm{TBP} = \mathrm{TBP}_i$, which is almost true if the two MSs are moving at the same speed and are close to each other with respect to the location of the BS. The Clarke's model is implemented as in [14], [15, App. A].

In Fig. 5 and 6, the average BER performance of cooperative and non-cooperative systems is plotted versus the average SNR $\bar{\gamma}$ for different values of TBP. A vehicle-to-vehicle communication system is considered with carrier frequency $f_{\mathrm{C}} = 5.2$GHz. The symbol duration is set to $T_{\mathrm{S}} = 10\mu$s, this ensure that the channel' spectrum is flat[2]. The time-bandwidth product TBP is chosen as perfomance metric to assess the degree of temporal variability of the channel (compared to the length of the codeword), as an example, when the MSs exhibit velocities up to $v = 160$km/h, the time-bandwidth product TBP goes proportionally up to 4. The analytical BER

[2]Recent channel measurements presented in [16] show that the delay spread at $f_{\mathrm{C}} = 5.2$GHz is around 1μs.

see the program

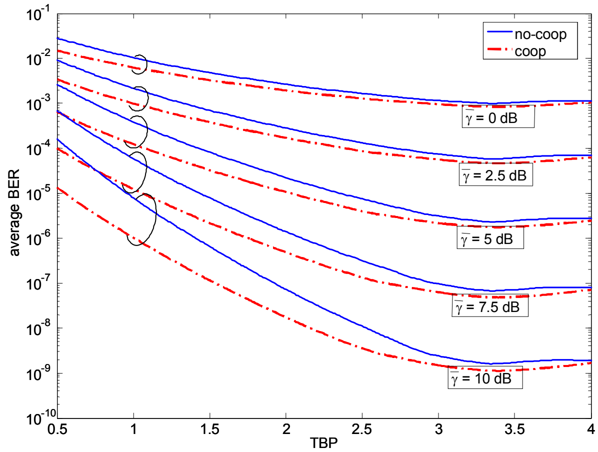


Fig. 7. Full cooperation ($\Theta = 1$) and no-cooperation ($\Theta = 2$) performance comparison. Code-rate $R = 1/4$.

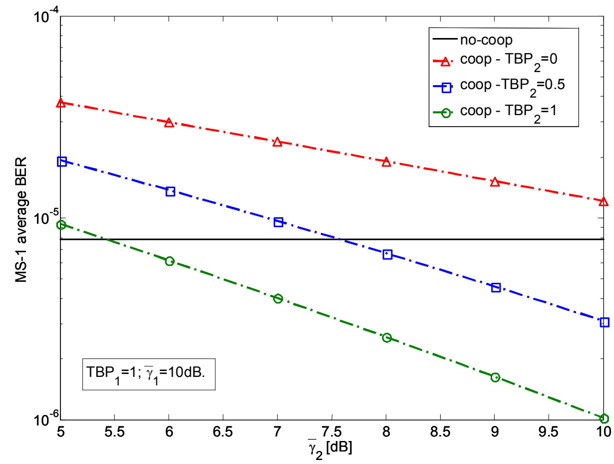


Fig. 8. Performance of MS-1 for asymmetric uplink channel conditions: MS-1 moves with $\mathrm{TBP}_1 = 1$ and $\bar{\gamma}_1 = 10$dB; MS-2 moves at different velocities with varying $\bar{\gamma}_2$. Full coded cooperation and no-cooperation are compared. Code-rate $R = 1/4$.

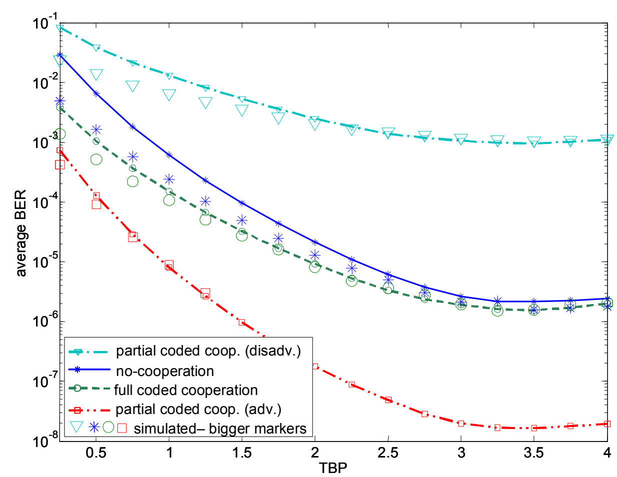


Fig. 9. Performance of the four cases of coded cooperation at $\bar{\gamma} = 7.5$dB: truncated bounds and simulation results. Partial cooperation is adavantageous or disadvantageous from the point of view of one MS. Code-rate $R = 1/3$.

bounds are computed by truncating the first summation in (5) at $d = 23$, or at values that are smaller but sufficient to upper bound the simulation results. The average PEP is computed both according to (12) and (13), (for the latter we truncate (5) at $d_{\mathrm{free}} = 15$). We observe that the performance for increasing TBP moves to the one obtained for FF.

In Fig. 7 the bounds on the average BER are plotted versus the TBP. Coded cooperation and no-cooperation are compared at different average SNR values $\bar{\gamma}$. Up to $\mathrm{TBP} \approx 3$ ($v = 120\mathrm{km/h}$), the performance gain of coded cooperation increases with increasing average SNR. At higher velocities the gain is almost negligible, which means that benefits of coded cooperation vanish for $\mathrm{TBP} \gtrsim 3$.

We analyze now the case where the uplink channels are unbalanced (thus showing different values of average SNR and/or velocity). In Fig. 8 MS-1 moves at $\mathrm{TBP}_1 = 1$ and transmits on a channel with average SNR $\bar{\gamma}_1 = 10$dB. On the other hand, average SNR for MS-2 varies from 5dB to 10dB, and time-bandwidth product $\mathrm{TBP}_2 = \{0, 0.5, 1\}$. The performance results suggest that coded cooperation outperforms remarkably no-cooperation only if the MSs are moving approximately at the same speed. The larger is the difference between MSs' velocities, the less advantageous it is for the fastest MS to cooperate. This result can be useful in case partner selection can be allowed [4].

We argue that the conclusions drawn in this Section are valid for every good RCPC code, because the diversity gain is carried by the cooperation scheme (space-diversity) and the bit-interleaving (time-diversity), independently from the specific code.

B. Cooperative regions

We now consider the more realistic case of imperfect inter-MS channel. We also adopt a RCPC code with higher rate than the one in Sect. IV-A to allow for a reasonable simulation complexity. The adopted RCPC mother-code, with octal generators (15, 17, 13), has rate $R = 1/3$ and free distance $d_{\mathrm{free}} = 10$. The coded-block length is $N = 384$ bits, resulting in $M = 192$ QPSK symbols. The punctured code has rate $R_1 = 1/2$, with code-length $N_1 = \frac{2N}{3}$ ($\alpha \simeq 33\%$). The coded bits are interleaved by a block bit-interleaver (96×4). Average BER performance for the four cases of coded cooperation (see Sect. II) are depicted in Fig. 9 for an average SNR of both uplink channels $\bar{\gamma} = 7.5$dB and varying degree of temporal variability TBP. In accordance with the results in Sect. IV-A, the coded cooperation (for the case $\Theta = 1$) does not provide any significant gain with respect to non-cooperative case for $\mathrm{TBP} \geq 3$. The upperbound (5), with PEP computed as in (12) and error event autocorrelation matrices according to (10) and (15) - (17), are truncated at $d = 14$,

see the program

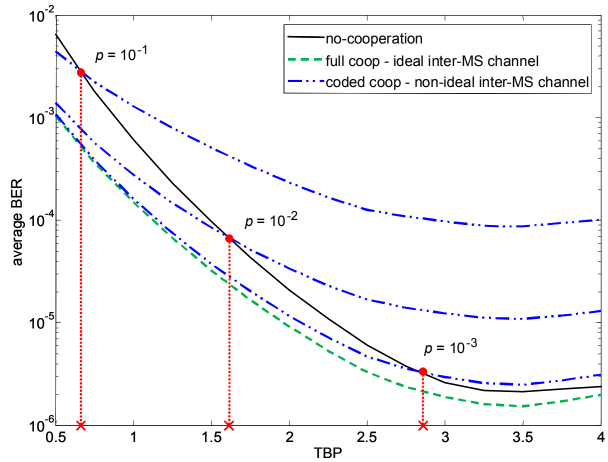


Fig. 10. Performance of coded cooperation at different inter-MS block error probability p compared to no-cooperation at $\bar{\gamma} = 7.5$dB. Code-rate $R = 1/3$.

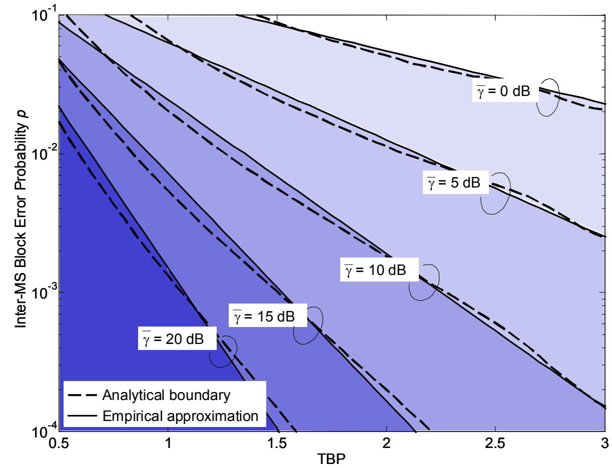


Fig. 11. *Cooperative regions* for coded cooperation at different $\bar{\gamma}$. Coded cooperation outperforms no-cooperation in the region below the analytical boundary. Code-rate $R = 1/3$.

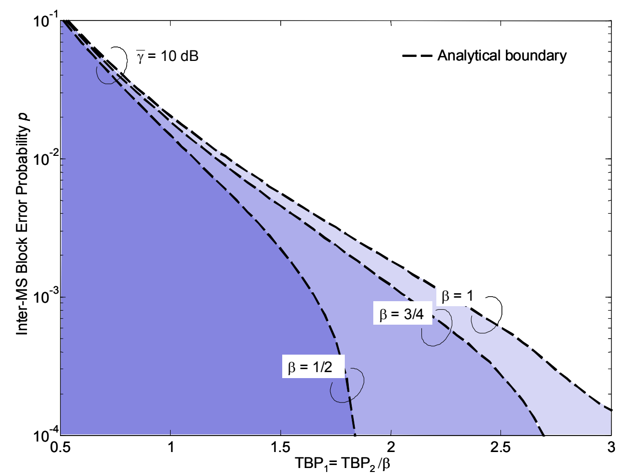


Fig. 12. *Cooperative regions* for MS-1. Uplink channels are both at $\bar{\gamma} = 10$dB, but MS-2 velocity is β times smaller than MS-1 velocity. Code-rate $R = 1/3$.

or at smaller values, in order to get tighter bounds for the subsequent analysis. These bounds are weighted as in (14) for different values of the inter-MS block error probability p and compared to no-cooperation in Fig. 10. As far as the mobility degree becomes large enough, the average BER performance of coded cooperation with non-ideal inter-MS channel is increasingly dominated by the worst case partial cooperation ($\Theta = 4$), as expected. The comparison between cooperative and non-cooperative transmission in terms of BER performances shows that the mobility degree threshold (cross markers in Fig. 10), beyond which coded cooperation is no more advantageous, decreases with decreasing quality of the inter-MS channel. For instance, coded cooperation with inter-MS block error probability $p = 10^{-3}$ is advantageous only up to $\text{TBP} \simeq 2.5$ for the considered code and channel settings. Furthermore, analytical and numerical results reveal that this threshold strongly depends on the quality of the relayed links ($\bar{\gamma}$).

The *cooperative region* is the collection of mobility (TBPs) and channel ($\bar{\gamma}$, p) settings for which coded cooperation is beneficial in providing enhanced average BER performance with respect to the non-cooperative case. The cooperative regions are illustrated as shaded areas delimited by solid lines in Fig. 11. For different values of the average SNR $\bar{\gamma}$, shaded areas contain the collection of values (TBP, p) for which coded cooperation provides superior performances compared to non-cooperative transmission. Interestingly, we observe that, for $\bar{\gamma} < 10$dB, the cooperative region spans the entire TBP range considered: the most promising opportunities to exploit the benefits of coded cooperation (in time-varying fading) arise for those applications where energy efficiency (for low SNR) is a key issue. Reasonably, the cooperative region size increases (decreases) with decreasing (increasing) average SNR $\bar{\gamma}$. In Fig. 12, the cooperative region for MS-1 is depicted by assuming the uplink channels with average SNR $\bar{\gamma} = 10$dB, while the degree of mobility of MS-2 is lower compared to MS-1 as $\text{TBP}_2 = \beta \times \text{TBP}_1$ with $\beta < 1$. As expected from the analysis in Sect. IV-A, the cooperative region for MS-1 becomes smaller if the partner moves at lower velocities ($\text{TBP}_2 < \text{TBP}_1$).

V. Conclusions

We have provided an analytical method to evaluate the average BER performance of coded cooperation over time-varying flat fading channels. The key idea lies in recognizing the algebraic structure of the fading channel autocorrelation matrix associated to the decision variable. The theoretical results have been corroborated and validated by simulation results. The present work has focused on a generic single-carrier transmission system with narrowband channels, but

see the program

the methodology can be transposed to broadband frequency-selective OFDM systems taking into account the correlation of the fading channel over the subcarriers. The analysis has at first encompassed the widest range of temporal variability of the fading process, from the BF to the FF model. However, the temporal variability strictly depends on the velocity of the mobile stations, which is necessarily limited. This physical limitation has been taken into account by circumscribing the performance evaluations to a more realistic range of temporal variability of the channel.

Analytical results, validated by simulations, have shown that coded cooperation, in the best case of error-free inter-MS channel, outperforms significantly a comparable non-cooperative transmission only up to a certain degree of mobility, approximately when the time-bandwidth product $\mathrm{TBP} \lesssim 3$ (corresponding to the speed $v = 120$km/h in a system with carrier frequency $f_{\mathrm{C}} = 5.2$GHz, symbol duration $T_{\mathrm{S}} = 10\mu$s and code block length $M = 256$ symbols). Beyond this limit, coded cooperation and non-cooperative transmissions perform similarly, since the gain offered by the time diversity is now dominant.

We have also investigated how the MSs' speed difference and the degradation of the inter-MS channel quality affect the BER performance. As expected, the larger is the speed difference the less advantageous it is to cooperate for the fastest MS. However, it has been shown that, for the MS moving at twice the velocity of the partner, coded cooperation improves significantly the performance with respect to no-cooperation up to $\mathrm{TBP} \simeq 1.5$ ($v = 60$km/h). Furthermore, the increase of the inter-MS block error probability has been proven to be a penalty factor for coded cooperation performance. For instance, coded cooperation in symmetric uplink channels' conditions performs better than no-cooperation at $\mathrm{TBP} = 2.5$, average SNR $\bar{\gamma} = 10$dB and inter-MS block error probability $p = 10^{-4}$, while performing worse as the inter-MS block error probability raises to $p = 10^{-3}$. The *cooperative region* have been derived in order to provide more insight to the analysis by defining the mobility and channel settings for which coded cooperation provides better performance than no-cooperation. The proposed approach can be used in network design in order to define algorithms for the selection of the cooperating MSs and the optimization of the cooperation level [4].

We believe that the present work contributes to build the base for future evaluations of coded cooperation in real mobile communication systems, with the support of detailed channel and mobility models.

Acknowledgements

This work was supported by the European Project #IST-216715 Network of Excellence in Wireless Communications (NEWCOM++), the Vienna Science and Technology Fund in the ftw. project COCOMINT, and the Austria Science Fund (FWF) through grant NFN SISE (S106). The Telecommunications Research Center Vienna (ftw.) is supported by the Austrian Government and the City of Vienna within the competence center program COMET.

References

[1] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," IEEE Trans. Inform. Theory, vol.50, pp.3062-3080, Dec. 2004.
[2] T. E. Hunter and A. Nosratinia, "Cooperation diversity through coding," in *Proc. IEEE ISIT*, July 2002, p. 220.
[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
[4] S. Valentin and H. Karl, "Effect of user mobility in coded cooperative systems with joint partner and cooperation level selection," in *Proc. IEEE WCNC*, March 2007, pp. 896-901.
[5] Z. Lin, E. Erkip, A. Stefanov, "Cooperative regions and partner choice in coded cooperative systems," *IEEE Trans. Commun.*, vol. 54, pp. 1323-1334, Jul. 2006.
[6] J. Proakis, *Digital Communications*, 4th Ed., McGraw Hill, 2001.
[7] K.Witrisal, Y.Kim, R.Prasad, "A novel approach for performance evaluation of OFDM with error correction coding and interleaving," in *Proc. IEEE VTC Fall*, Sep. 1999, Vol. 1, pp. 294-299.
[8] D. Molteni, M. Nicoli, R. Bosisio, L. Sampietro, "Performance analysis of multiantenna WiMax systems over frequency-selective fading channels," in *Proc. IEEE PIMRC*, Athens, Sep. 2007.
[9] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 38, pp. 389-400, Nov. 1988.
[10] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst.Tech. J.*, p. 957, Jul.–Aug. 1968.
[11] S. Lin and J. D. Costello, *Error Control Coding - Fundamentals and Application*, Prentice-Hall, 2003.
[12] M.K. Simon, M.S. Alouini, *Digital communications over fading channels: a unified approach to performance analysis*, Wiley, 2000.
[13] J. W. Craig, "A new, simple, and exact result for calculating the probability of error for two-dimensional signal costellations," in *Proc. IEEE MILCOM*, Nov. 1991, vol. 2, pp. 571-575.
[14] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 920-928, June 2003.
[15] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Processing*, vol. 53, no. 9, pp. 3597-3607, Sep. 2005.
[16] A. Paier et al, "Non-WSSUS vehicular channel characterization in highway and urban scenarios at 5.2 GHz using the local scattering function," in *Proc. IEEE WSA*, Darmstadt, Germany, Feb. 2008.

see the program

Performance of *Sp(n)* Codes in Two-Way Wireless Relay Networks

Gilbert Yammine, Zoran Utkovski and Juergen Lindner
Institute of Information Technology, University of Ulm, Germany
Email: {gilbert.yammine, zoran.utkovski, juergen.lindner}@uni-ulm.de

***Abstract*—Here we study the performance of *Sp(n)* codes as distributed space-time codes for two-way relaying systems. The *Sp(n)* codes are space-time codes based on the *Sp(n)* group. The performance is evaluated in the Amplify-and-Forward and Decode-and-Forward scenario, with protocols requiring 4, 3 and 2 time slots for information exchange between the terminals. The analysis shows that these codes are appropriate for the two-way relaying scenario, since they achieve the performance (throughput) gain promised by two-way relaying.**

I. Introduction

In this work we focus on a two-way relaying system where two wireless terminals want to exchange their information with the assistance of relays. Each terminal acts as a transmitter and a receiver, and works in a half-duplex mode. The terminals may, in general, be equipped with more antennas. The two-way channel was first studied by Shannon [1], where he found inner and outer bounds on the capacity.

With the increased interest in ad-hoc networks, several cooperation based methods which exploit spatial diversity provided by antennas of different users have been developed, e.g. [2]–[5]. The improvement is called cooperative diversity and is achieved by having different users in the network cooperate in some way. Among the first cooperative strategies are amplify-and-forward [4], [5] and decode-and-forward [2]–[4]. An enhanced cooperative strategy, based on distributed space-time coding was introduced in [6]. These schemes assume half-duplex communication, i.e. the nodes in the network do not transmit and receive simultaneously. Additionally, the communication takes place in one direction (one-way relaying), from the transmitter to the receiver, in two stages.

The one-way relaying protocols suffer from a loss in spectral efficiency due to the half-duplex constraint of the terminals. In order to increase the spectral efficiency of such relay network, a bi-directional (two-way) communication between two terminals where the relay assists in the two-way communication was introduced in [7]. The connection between two-way relaying and network coding was established in [8], where protocols for two-way (bi-directional) relaying based on network coding on symbol level were introduced. The main idea is that the relays combine the information from the terminals and broadcast it in the next stage. Each terminal then subtracts its own contribution and decides about the signal transmitted from the other terminal. The two-way relaying schemes promise throughput gain compared to one-way relaying, by saving time slots in the broadcast stage.

Here, we will analyze the performance of $Sp(n)$ codes in the context of two-way relaying. The $Sp(n)$ codes are based on the symplectic Lie group $Sp(n)$ and were first introduced in [9] as space-time codes for MIMO systems and later in the distributed space-time coding scenario for one-way relaying [6]. We will compare distributed space-time codes performance in the context of Amplify-and-Forward (AF) as well as Decode-and-Forward (DF) scenarios. We will consider the cases where the communication takes place in 4, 3, and 2 time slots respectively, aiming at showing the throughput/performance gain of relaying when less than 4 time slots are used.

The rest of the paper is organized as follows. In Section II, we describe the two-way relaying system model along with the different exchange protocols. In Section III, we show how can codes be built from the $Sp(n)$ group, and present the $Sp(2)$ space-time code. In Section IV we describe the deployment of distributed space-time codes in Amplify-and-Forward (AF) and Decode-and-Forward (DF) scenarios and show their system equations. In Section V, we present worked-out examples with $Sp(2)$ codes implemented in wireless relay networks. Simulation results are given in Section VI. Section VII concludes the paper.

The following notation is used. Vectors are denoted by bold face lower case letters. Matrices are denoted by normal upper case letters. Sets are denoted by caligraphic letters. $(\cdot)^{\mathrm{H}}$ is the hermitian operator, $(\cdot)^*$ is complex conjugation and $(\cdot)^{\mathrm{T}}$ is matrix transpose. $\exp(\cdot)$ is the matrix exponential and $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm of a matrix.

II. Two-Way Relaying

A. System Model

We consider a wireless network with $R+2$ nodes, where R nodes act as relays (in parallel). The communication scheme is half-duplex, meaning that the nodes do not transmit and receive simultaneously. We assume block Rayleigh fading model where the channel is constant in a certain time block. The terminals are equipped with M antennas. Each terminal sends a time sequence (vector) of duration of T symbols (channel uses). With this, the signal transmitted from terminal 1 (T_1) is a $T \times M$ matrix S. We normalize the transmitted signal such that $E\{\mathrm{tr}(S^{\mathrm{H}}S)\} = M$. Similarly, terminal 2 (T_2) sends a $T \times M$ transmit matrix D, normalized such that $E\{\mathrm{tr}(D^{\mathrm{H}}D)\} = M$. We denote by P_1 and P_2 the average

see the program

transmit power for one transmission of T_1 and T_2 respectively. The T_1's transmit signal is $\sqrt{P_1 T/M}S$ and the T_2's transmit signal is $\sqrt{P_2 T/M}D$. Further, we denote the average power for one transmission for each relay as P_3 which gives a total power of $P_3 R$ for the R relays.

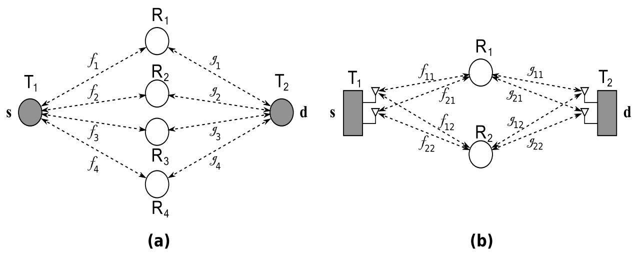


Fig. 1. Two-Way Relaying Networks

We concentrate on a distributed space-time coding scenario, where the space-time codes are distributively built at the relays. The usual requirement for space-time block coding is that the channel is constant for at least T time instants (channel uses), which in the literature is also denoted as "time slot", for example when 2, 3 and 4 time slots protocols are compared. Here we use both "time slot" and "stage" to denote the period of duration of T time instants.

In the following, we consider three exchange protocols, which use 4, 3 and 2 time slots for the data exchange, respectively. In the first protocol (Fig. 2(a)), T_1 sends in time slot 1 to the relays, which forward their signals to T_2 during time slot 2. The other way communication takes part in time slots 3 and 4. In the 3 time slots protocol (Fig. 2(b)), Terminal 1 transmits its signal during time slot 1 to the relays. In time slot 2, Terminal 2 transmits its signal to the relays. In time slot 3 the relays broadcast their transmit signals to both terminals which, knowing what they have sent, can decide on the received data from the other terminal. In the 2 time slots protocol (Fig. 2(c)), both terminals transmit simultaneously in the first time slot. In the second time slot, the relays broadcast the information to both terminals where the decision is made similarly to the case of 3 time slots.

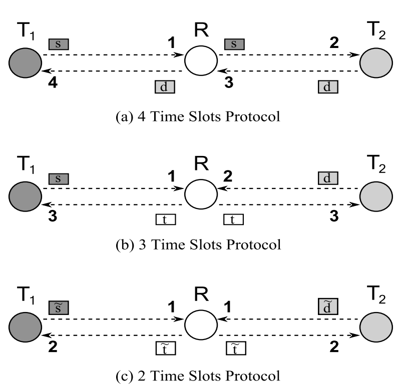


Fig. 2. Two-Way Relaying Protocols

The channel coefficient between the m-th antenna of Terminal 1 (T_1) and the i-th relay during time slot τ is denoted as $f_{mi}^{(\tau)}$ and the channel coefficient between the m-th antenna of Terminal 2 (T_2) and the i-th relay is denoted as $g_{mi}^{(\tau)}$.

The multicast and the broadcast stage save time slots compared to the case where the information exchange between the terminals requires four stages, two for each direction. This leads to increased throughput in the wireless network. If instead of throughput we like to gain reliability, we can fix the throughput to be the same in both protocols. In the two and three time slots protocols, this will allow for using alphabets of lower cardinality at the terminals compared to the four-stage protocol.

III. Codes for Two-Way Relaying based on the $Sp(n)$ Lie Group

Basically, distributed space-time codes designed for one-way relaying systems are also appropriate for tow-way relaying. An extensive analysis of distributed space-time codes in two-way wireless relay networks was performed in [10]. Here we concentrate on the codes obtained from the $Sp(n)$ Lie group as promising candidates for two-way relaying, due to their diversity advantages and the possibility for simplified decoding. The application in the two-way relaying scenario is not straightforward and therefore the usual Amplify-and-Forward and Decode-and-Forward schemes should be modified in order to accomodate these codes. We will start with the codes from the $Sp(2)$ group introduced first for point-to-point MIMO block Rayleigh fading channels. Later in this section we discuss on the possible extension to codes from the more general $Sp(n)$ $(n > 2)$ group.

A. $Sp(2)$ codes

The $Sp(2)$ codes were first introduced by in [6], [11] as space-time codes for coherent block Rayleigh channels. Due to their properties they can also be used as distributed space-time codes in wireless networks with relays. Here we extend the construction to distributed codes for two-way relaying and evaluate their performance in AF and DF scenarios, with protocols using 4, 3 and 2 time slots for information exchange between the terminals. $Sp(2)$ codes are based on the symplectic Lie group $Sp(2)$. They can be used in a system with transceivers having one transmit antenna and four relays, or a system with transceivers with two antennas and two relays assisting the communication. The $Sp(2)$ code is given as

$$\mathcal{U} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} UV & UV^* \\ -U^*V & U^*V^* \end{bmatrix} \right\} \quad (1)$$

where

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j\frac{2\pi k}{P}} & e^{j\frac{2\pi l}{P}} \\ -e^{-j\frac{2\pi l}{P}} & e^{-j\frac{2\pi k}{P}} \end{bmatrix}$$
$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j(\frac{2\pi m}{Q}+\theta)} & e^{j(\frac{2\pi n}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n}{Q}+\theta)} & e^{-j(\frac{2\pi m}{Q}+\theta)} \end{bmatrix} \quad (2)$$

and $0 \le k, m < P$, $0 \le l, n < Q$. When P and Q do not have common divisors, it can be shown that the codes achieve

see the program

full diversity for all values of θ. If we define the symbols in U and V as

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} a_1 & a_2 \\ -a_2^* & a_1^* \end{bmatrix}, \qquad V = \frac{1}{\sqrt{2}} \begin{bmatrix} b_1 & b_2 \\ -b_2^* & b_1^* \end{bmatrix}, \quad (3)$$

and if we set

$$u_1 = \frac{1}{2\sqrt{2}}(a_1 b_1 - a_2 b_2^*) \qquad u_2 = \frac{1}{2\sqrt{2}}(a_2^* b_1 + a_1^* b_2^*)$$
$$u_3 = -\frac{1}{2\sqrt{2}}(a_1^* b_1 - a_2^* b_2^*) \quad u_4 = \frac{1}{2\sqrt{2}}(a_2 b_1 + a_1 b_2^*), \quad (4)$$

the code in (1) becomes a unitary STBC defined as

$$\mathcal{U} = \begin{bmatrix} u_1 & -u_2^* & -u_3^* & u_4 \\ u_2 & u_1^* & -u_4^* & -u_3 \\ u_3 & -u_4^* & u_1^* & -u_2 \\ u_4 & u_3^* & u_2^* & u_1 \end{bmatrix}. \quad (5)$$

In order to distributively build the $Sp(2)$ code, the relays transmit signals are designed to be linear functions of their received signals and their conjugate in AF mode, and a linear function of the decoded symbols and their conjugate in DF mode. For that, we define specially designed unitary matrices at each relay that provide different combinations of the symbols, and lead to a distributed space-time code. For the DF mode, we use a model in which the transmit signal is defined to resemble the received signal but without noise. This can be achieved by weighting the decoded symbols at the relays by the channel coefficients that affected their decoding.

B. Extension to codes from the $Sp(n)$ Lie group

If we want to extend the construction to codes which are appropriate to networks with different number of relays, we have to look at the symplectic Lie group $Sp(n)$. The $Sp(n)$ group is the same as the quaternionic unitary group $U(n, \mathbb{H})$. This group is a Lie group and can be equivalently be represented as the group of $2n \times 2n$ matrices S which satisfy the unitary condition $SS^{\mathrm{H}} = S^{\mathrm{H}}S = I_{2n}$ and the symplectic condition $S^{\mathrm{T}} J_{2n} S = J_{2n}$, where $J_{2n} = \begin{bmatrix} \mathbf{0} & I_n \\ -I_n & \mathbf{0} \end{bmatrix}$.

Every Lie group is also a differentiable manifold, meaning that a tangent space can be constructed at every element of the group (manifold). The tangent space at the identity element of a Lie group has the structure of a Lie algebra [1]. The Lie algebra of the $Sp(n)$ group, $\mathfrak{sp}(n)$, contains $n \times n$ quaternionic matrices of the type $X = -X^{\mathrm{H}}$. An $n \times n$ matrix with entries from the quaternions can be equivalently represented by an $2n \times 2n$ matrix with complex entries by using the matrix representation of quaternions

$$a + bi + cj + dk \equiv \begin{bmatrix} a + bi & c + di \\ -c + di & a - bi \end{bmatrix}. \quad (6)$$

This means that the $Sp(n)$ group can be parameterized with the tangent space at the identity element consisting of the quaternionic matrices fulfilling $X = -X^{\mathrm{H}}$. The connection between the Lie group $Sp(n)$ and the Lie algebra $\mathfrak{sp}(n)$ is given by the exponential map, which for matrix Lie groups (such as $Sp(n)$) is the matrix exponential $\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}$.

The tangent space at the identity element of $Sp(n)$ is a vector space of real dimension $n(2n+1)$ which, of course, coincides with the dimension of the group $Sp(n)$. Therefore, it can be embedded in the Euclidean space $\mathbb{R}^{n(2n+1)}$. Therefore, elements of the tangent space can be obtained by sampling a lattice in $\mathbb{R}^{n(2n+1)}$, as in [12] for example. In order to maximize the diversity of the code obtained after the mapping to the group $Sp(n)$, lattice rotation can be used [13]. Finally, the constellation in the tangent space is scaled by a factor which controls the distance profile of the code after the mapping to the Lie group (manifold).

[1] Lie algebra is a nonassociative algebra, i.e. a vector space $\mathfrak{g}$ over a field $\mathbb{F}$ with the Lie bracket as binary operation

IV. Relay Cooperative Strategies

A. Amplify-and-Forward (AF)

In the AF scenario, the relays act as simple signal repeaters that amplify and apply simple operations on the received signals, such as addition, multiplication, and conjugation. Using the AF cooperative strategy, we present next the three data exchange protocols, using 4, 3, and 2 time slots, respectively.

1) 4 Time Slots Protocols: The conventional way of data exchange between T_1 and T_2 is to use 4 time slots, as shown in Fig. 2(a). At time slot 1, T_1 transmits $\sqrt{P_1 T/M} S$ to the relays, each receiving the following vector

$$\mathbf{r}_i^{(1)} = \sqrt{P_1 T/M} \sum_{m=1}^{M} f_{mi}^{(1)} \mathbf{s}_m + \mathbf{v}_i^{(1)}, \qquad i = 1 \cdots R \quad (7)$$

where $\mathbf{r}_i^{(\tau)}$ denotes the received vector at the i-th relay, at time slot τ, and $\mathbf{v}_i^{(\tau)}$ denotes the additive noise at the i-th relay, at time slot τ. Using the idea of distributed space-time coding [14], each relay transmit signal is designed as a linear function of its received signal or its conjugate as

$$\mathbf{t}_i = \sqrt{\frac{P_3}{P_1 + 1}} A_i \hat{\mathbf{r}}_i, \quad (8)$$

where A_i is a precoding $T \times T$ unitary matrix, and $\hat{\mathbf{r}}_i$ is even $\mathbf{r}_i$ or $\mathbf{r}_i^*$, depending on the relay. The normalization factor is such that the average transmit power is $P_3 T$ at each relay.

At time slot 2, the relays forward their signals to T_2. The received vector at the m-th antenna of T_2 is given by

$$\mathbf{y}_m = \sum_{i=1}^{R} g_{im}^{(2)} \mathbf{t}_i + \mathbf{w}_m, \quad (9)$$

where $\mathbf{w}_m$ is the additive noise at the m-th receive antenna of T_2. If we denote $\mathbf{f}_i = [\hat{f}_{1i}^{(1)}, \cdots, \hat{f}_{Mi}^{(1)}]^T$, where $\hat{f}_{mi}^{(1)}$ is equal to $f_{mi}^{(1)}$ when the i-th relay doesn't conjugate the received signal, and equal to $f_{mi}^{(1)*}$ when the i-th relay conjugates its received signal, equation (9) can then be rewritten as

$$\mathbf{y}_m = \rho_1 [A_1 \hat{S} \ \cdots \ A_R \hat{S}] \begin{bmatrix} \mathbf{f}_1 g_{m1}^{(2)} \\ \cdots \\ \mathbf{f}_R g_{mR}^{(2)} \end{bmatrix} + \rho_2 \sum_{i=1}^{R} g_{mi}^{(2)} A_i \hat{\mathbf{v}}_i^{(1)} + \mathbf{w}_m, \quad (10)$$

where the *hat* means a conjugate when a relay conjugates its received signal, $\rho_1 = \sqrt{\frac{P_1P_3T}{(P_1+1)M}}$, and $\rho_2 = \sqrt{\frac{P_3}{P_1+1}}$. By defining $Y_{T_2} = [\mathbf{y}_1, \cdots, \mathbf{y}_M]$, $\Phi = [A_1\hat{S} \;\cdots\; A_R\hat{S}]$, $\mathbf{g}_i = [g_{1i}^{(2)}, \cdots, g_{Mi}^{(2)}]$, $H = [\mathbf{f}_1\mathbf{g}_1, \cdots, \mathbf{f}_R\mathbf{g}_R]^T$, and $W = \left[\rho_2 \sum_{i=1}^R g_{1i}^{(2)} A_i \hat{\mathbf{v}}_i^{(1)} + \mathbf{w}_1, \cdots, \rho_2 \sum_{i=1}^R g_{Mi}^{(2)} A_i \hat{\mathbf{v}}_i^{(1)} + \mathbf{w}_M\right]$, we get the general system model at T_2 as

$$Y_{T_2} = \sqrt{\frac{P_1P_3T}{(P_1+1)M}}\Phi H + W, \tag{11}$$

where Φ is a $T \times MR$ distributively built space-time code, and Y_{T_2} is the $T \times M$ received matrix at T_2.

The ML decoding at T_2 is given by

$$\tilde{\mathbf{s}}_\mathbf{i} = \underset{\mathbf{s}_i}{\operatorname{argmin}} \left\| Y_{T_2} - \sqrt{\tfrac{P_1P_3T}{(P_1+1)M}}\Phi_i H \right\|_F^2. \tag{12}$$

During time slots 3 and 4, the data transfer from T_2 to T_1 is done similarly. Due to the symmetry of the network, we do not show the equations of the other way system, knowing that they can be easily derived.

2) 3 Time Slots Protocol: In this section, we consider the three-time slots protocol shown in Fig. 2(b). Here, the relays combine the data they receive from the terminals, and broadcast the designed transmit signals to both terminals. At time slot 1, T_1 transmits $\sqrt{P_1T/M}S$ to the relays, and in time slot 2, T_2 transmits $\sqrt{P_2T/N}D$. Consequently, the relay received signals at the first and second time slots are given respectively as

$$\mathbf{r}_i^{(1)} = \sqrt{P_1T/M}\sum_{m=1}^M f_{mi}^{(1)}\mathbf{s}_m + \mathbf{v}_i^{(1)}, \qquad i = 1\cdots R \tag{13}$$

$$\mathbf{r}_i^{(2)} = \sqrt{P_2T/M}\sum_{m=1}^M g_{mi}^{(2)}\mathbf{d}_m + \mathbf{v}_i^{(2)}. \qquad i = 1\cdots R \tag{14}$$

The relay transmit vectors are designed as a linear function of the sum or the conjugate of the sum of both received signals, and obtained as

$$\mathbf{t}_i = \sqrt{\frac{P_3}{P_1+P_2+2}}A_i(\hat{\mathbf{r}}_i^{(1)} + \hat{\mathbf{r}}_i^{(2)}), \tag{15}$$

where the power normalization factor ensures an average transmit power of P_3T at each relay.

At time slot 3, the relays broadcast their signals to both terminals, and looking at T_1, the received signal at the m-th antenna is given as

$$\begin{aligned}\mathbf{y}_m = {} & \rho_1[A_1\hat{S}\cdots A_R\hat{S}]\begin{bmatrix}\mathbf{f}_1 f_{m1}^{(3)} \\ \cdots \\ \mathbf{f}_R f_{mR}^{(3)}\end{bmatrix} + \rho_2\sum_{i=1}^R f_{mi}^{(3)}A_i\hat{\mathbf{v}}_i^{(1)} + \\ & \rho_3[A_1\hat{D}\cdots A_R\hat{D}]\begin{bmatrix}\mathbf{g}_1 f_{m1}^{(3)} \\ \cdots \\ \mathbf{g}_R f_{mR}^{(3)}\end{bmatrix} + \rho_4\sum_{i=1}^R f_{mi}^{(3)}A_i\hat{\mathbf{v}}_i^{(2)} + \mathbf{w}_m,\end{aligned} \tag{16}$$

where $\mathbf{f}_i = [\hat{f}_{1i}^{(1)}, \cdots, \hat{f}_{Mi}^{(1)}]^T$, $\mathbf{g}_i = [\hat{g}_{1i}^{(2)}, \cdots, \hat{g}_{Mi}^{(2)}]^T$, $\rho_1 = \sqrt{\frac{P_1P_3T}{(P_1+P_2+2)M}}$, $\rho_2 = \sqrt{\frac{P_3}{P_1+P_2+2}}$, $\rho_3 = \sqrt{\frac{P_2P_3T}{(P_1+P_2+2)M}}$, and $\rho_4 = \sqrt{\frac{P_3}{P_1+P_2+2}}$.

Consequently, the overall system equation at T_1 can be written as

$$Y_{T_1} = \rho_1\Phi H + \rho_3\Psi G + W, \tag{17}$$

where $Y_{T_1} = [\mathbf{y}_1, \cdots, \mathbf{y}_M]$, $\Phi = [A_1\hat{S} \;\cdots\; A_R\hat{S}]$, $\Psi = [A_1\hat{D} \;\cdots\; A_R\hat{D}]$, $\mathbf{f}_i' = [f_{1i}^{(3)}, \cdots, f_{Mi}^{(3)}]$, $H = [\mathbf{f}_1\mathbf{f}_1', \cdots, \mathbf{f}_R\mathbf{f}_R']^T$, $G = [\mathbf{g}_1\mathbf{f}_1', \cdots, \mathbf{g}_R\mathbf{f}_R']^T$ and $W = [\rho_2\sum_{i=1}^R f_{1i}^{(3)}A_i\hat{\mathbf{v}}_i^{(1)} + \rho_4\sum_{i=1}^R f_{1i}^{(3)}A_i\hat{\mathbf{v}}_i^{(3)} + \mathbf{w}_1, \cdots, \rho_2\sum_{i=1}^R f_{Mi}^{(3)}A_i\hat{\mathbf{v}}_i^{(1)} + \rho_4\sum_{i=1}^R f_{Mi}^{(3)}A_i\hat{\mathbf{v}}_i^{(2)} + \mathbf{w}_M]$.

Φ and Ψ are two distributively built space-time codes with symbol data from T_1 and T_2, respectively. From the overall system equation, and having in mind that T_1 knows its sent data and the channel coeficients, the ML decoding at T_1 can be written as follows

$$\tilde{\mathbf{d}} = \underset{\mathbf{d}_i}{\operatorname{argmin}} \|Y_{T_1} - \rho_1\Phi H - \rho_3\Psi_i G\|_F^2. \tag{18}$$

One advantage of 3 time slots protocols is that each terminal can exploit the weak direct path between the other terminal and itself to improve the decoding. For example, in time slot 1, T_1 transmits to the relays, and T_2 also receives a weak signal, which can help in the decoding later. More on this issue can be found in [10].

3) 2 Time Slots Protocol: In the 2 time slots protocol, both terminals transmit simultaneously during time slot 1. Hence, the relays receive an additive interference signal given by

$$\mathbf{r}_i^{(1)} = \sqrt{P_1T/M}\sum_{m=1}^M f_{mi}^{(1)}\mathbf{s}_m + \sqrt{P_2T/M}\sum_{m=1}^M g_{mi}^{(1)}\mathbf{d}_m + \mathbf{v}_i^{(1)}. \tag{19}$$

The relays then apply the same operations as in the previous cases. Therefore, the relay transmit vectors are designed as

$$\mathbf{t}_i = \sqrt{\frac{P_3}{P_1+P_2+1}}A_i\hat{\mathbf{r}}_i^{(1)}, \tag{20}$$

where again, the power normalization factor ensures an average transmit power of P_3T at each relay.

At time slot 2, the relays broadcast their signals to both terminals. The received signal at the m-th antenna of T_1 is given by

$$\begin{aligned}\mathbf{y}_m = {} & \rho_1[A_1\hat{S}\cdots A_R\hat{S}]\begin{bmatrix}\mathbf{f}_1 f_{m1}^{(2)} \\ \cdots \\ \mathbf{f}_R f_{mR}^{(2)}\end{bmatrix} + \rho_2\sum_{i=1}^R f_{mi}^{(2)}A_i\hat{\mathbf{v}}_i^{(1)} + \\ & \rho_3[A_1\hat{D}\cdots A_R\hat{D}]\begin{bmatrix}\mathbf{g}_1 f_{m1}^{(2)} \\ \cdots \\ \mathbf{g}_R f_{mR}^{(2)}\end{bmatrix} + \mathbf{w}_m,\end{aligned} \tag{21}$$

where $\mathbf{f}_i = [\hat{f}_{1i}^{(1)}, \cdots, \hat{f}_{Mi}^{(1)}]^T$, $\mathbf{g}_i = [\hat{g}_{1i}^{(1)}, \cdots, \hat{g}_{Mi}^{(1)}]^T$, $\rho_1 = \sqrt{\frac{P_1P_3T}{(P_1+P_2+1)M}}$, $\rho_2 = \sqrt{\frac{P_3}{P_1+P_2+1}}$, and $\rho_3 = \sqrt{\frac{P_2P_3T}{(P_1+P_2+1)M}}$.

By defining $Y_{T_1} = [\mathbf{y}_1, \cdots, \mathbf{y}_M]$, $\Phi = [A_1\hat{S} \;\cdots\; A_R\hat{S}]$, $\Psi = [A_1\hat{D} \;\cdots\; A_R\hat{D}]$, $\mathbf{f}_i' = [f_{1i}^{(2)}, \cdots, f_{Mi}^{(2)}]$, $H =$

see the program

$[\mathbf{f}_1\mathbf{f}_1', \cdots, \mathbf{f}_R\mathbf{f}_R']^T$, $G = [\mathbf{g}_1\mathbf{f}_1', \cdots, \mathbf{g}_R\mathbf{f}_R']^T$ and $W = [\rho_2 \sum_{i=1}^R f_{1i}^{(2)} A_i \hat{\mathbf{v}}_i^{(1)} + \mathbf{w}_1, \cdots, \rho_2 \sum_{i=1}^R f_{Mi}^{(2)} A_i \hat{\mathbf{v}}_i^{(1)} + \mathbf{w}_M]$, we can write the general system equation as

$$Y_{T_1} = \rho_1 \Phi H + \rho_3 \Psi G + W, \quad (22)$$

where Φ and Ψ are two distributively built space-time codes. The ML decoding at T_1 can be obtained as

$$\tilde{\mathbf{d}} = \underset{\mathbf{d}_i}{\operatorname{argmin}} \|Y_{T_1} - \rho_1 \Phi H - \rho_3 \Psi_i G\|_F^2 . \quad (23)$$

By looking at the accumulated noise at the terminal, we can see that the power of this noise when using the 2 time slots protocol is less than that of the 3 time slots protocol, which shows up as a performance advantage in the simulations.

B. Decode-and-Forward (DF)

In the Decode-and-Forward strategy, the relays decode the signals received from the terminals, then encode them again, and forward the generated signals to the terminals. Furthermore, in this section, we use the *Partial-Decode-and-Forward* (PDF) strategy introduced in [10], where the relays weight the decision on the decoded symbols with the channel coefficients of the terminal-relay link. By this, more value is given to the decisions when the link is more reliable (the channel is good). As for the AF case, we also present here the four exchange protocols, using 4, 3 and 2 time slots.

1) 4 Time Slots Protocol: During time slot 1, T_1 sends it data to the relays, which receive the signal given in equation (7). The vector $\mathbf{s}$ containing the data from T_1 is then decoded using the ML decoder

$$\tilde{\mathbf{s}}_i = \underset{\mathbf{s}_l}{\operatorname{argmin}} \left\| \mathbf{r}_i - \sqrt{\tfrac{P_1 T}{M}} \textstyle\sum_{m=1}^M f_{mi}^{(1)} \mathbf{s}_m^l \right\|_F^2 , \quad (24)$$

where the search is done over all possible sent vectors $\mathbf{s}_l$.

After decoding, the relays set the following transmit vectors

$$\mathbf{t}_i = \sqrt{\frac{P_3 T}{M}} A_i \sum_{m=1}^M \hat{f}_{mi}^{(1)} \hat{\tilde{\mathbf{s}}}_m, \quad (25)$$

where the *hat* denotes a conjugation when the relay conjugates, and the *tilde* denotes a decoded symbol (estimated). As it can be seen, the built transmit vector resembles the received one but without noise, which allows to assume the used scheme as a denoising process.

At time slot 2, the relays forward their signals to T_2. Assuming correct decoding at the relays (so we can remove the estimation), the received vector at the m-th antenna of T_2 is given by

$$\mathbf{y}_m = \rho_1 [A_1 \hat{S} \ \cdots \ A_R \hat{S}] \begin{bmatrix} \mathbf{f}_1 g_{m1}^{(2)} \\ \cdots \\ \mathbf{f}_R g_{mR}^{(2)} \end{bmatrix} + \mathbf{w}_m, \quad (26)$$

where $\rho_1 = \sqrt{\frac{P_3 T}{M}}$ and $\mathbf{f}_i = [\hat{f}_{1i}^{(1)}, \cdots, \hat{f}_{Mi}^{(1)}]^T$.

Consequently, the overall system equation at T_2 can be written as

$$Y_{T_2} = \sqrt{\frac{P_3 T}{M}} \Phi H + W, \quad (27)$$

where $\Phi = [A_1 \hat{S} \ \cdots \ A_R \hat{S}]$, $\mathbf{g}_i = [g_{1i}^{(2)}, \cdots, g_{Mi}^{(2)}]$, $H = [\mathbf{f}_1\mathbf{g}_1, \cdots, \mathbf{f}_R\mathbf{g}_R]^T$, and $W = [\mathbf{w}_1, \cdots, \mathbf{w}_M]$.

T_2 can then use the following ML decoder to decode the symbols sent from T_1

$$\tilde{\mathbf{s}} = \underset{\mathbf{s}_i}{\operatorname{argmin}} \left\| Y_{T_2} - \sqrt{\tfrac{P_3 T}{M}} \Phi_i H \right\|_F^2 . \quad (28)$$

During time slot 3 and 4, the data transmission from T_2 to T_1 is performed similarly.

2) 3 Time Slots Protocol: During time slots 1 and 2, the relays receive both terminals' data separated, as in equations (13) and (14). The relays then decode the symbols using the decoder in (24) and build the following transmit vectors

$$\mathbf{t}_i = \sqrt{\frac{P_3 T}{2M}} A_i \left(\sum_{m=1}^M \hat{f}_{mi}^{(1)} \hat{\tilde{\mathbf{s}}}_m + \sum_{m=1}^M \hat{g}_{mi}^{(2)} \hat{\tilde{\mathbf{d}}}_m \right), \quad (29)$$

where again, the decoded symbols are weighted by the channel coefficients. Assuming correct decoding at the relays, the broadcasted signal at time slot 3 obtained at the m-th antenna of T_1 is given by

$$\begin{aligned} \mathbf{y}_m = & \rho_1 [A_1 \hat{S} \ \cdots \ A_R \hat{S}] \begin{bmatrix} \mathbf{f}_1 f_{m1}^{(3)} \\ \cdots \\ \mathbf{f}_R f_{mR}^{(3)} \end{bmatrix} + \\ & \rho_1 [A_1 \hat{D} \ \cdots \ A_R \hat{D}] \begin{bmatrix} \mathbf{g}_1 f_{m1}^{(3)} \\ \cdots \\ \mathbf{g}_R f_{mR}^{(3)} \end{bmatrix} + \mathbf{w}_m, \end{aligned} \quad (30)$$

where $\rho_1 = \sqrt{\frac{P_3 T}{2M}}$, $\mathbf{f}_i = [\hat{f}_{1i}^{(1)}, \cdots, \hat{f}_{Mi}^{(1)}]^T$ and $\mathbf{g}_i = [\hat{g}_{1i}^{(2)}, \cdots, \hat{g}_{Mi}^{(2)}]^T$.

The overall system equation at T_1 can be then written as

$$Y_{T_1} = \sqrt{\tfrac{P_3 T}{2M}} \Phi H + \sqrt{\tfrac{P_3 T}{2M}} \Psi G + W, \quad (31)$$

where $\Phi = [A_1 \hat{S} \ \cdots \ A_R \hat{S}]$, $\Psi = [A_1 \hat{D} \ \cdots \ A_R \hat{D}]$, $\mathbf{f}_i' = [f_{1i}^{(3)}, \cdots, f_{Mi}^{(3)}]$, $H = [\mathbf{f}_1\mathbf{f}_1', \cdots, \mathbf{f}_R\mathbf{f}_R']^T$, $G = [\mathbf{g}_1\mathbf{f}_1', \cdots, \mathbf{g}_R\mathbf{f}_R']^T$, and $W = [\mathbf{w}_1, \cdots, \mathbf{w}_M]$.

The decoder is very similar to that of the AF case given in (23), with the change in the definition of the matrices.

3) 2 Time Slots Protocol: Having received an interference of both terminals at time slot 1, the relays try to decode a pair of symbol vectors $\{\tilde{\mathbf{s}}_i, \tilde{\mathbf{d}}_i\}$ received from both terminals, via the following decoder

$$\{\tilde{\mathbf{s}}_i, \tilde{\mathbf{d}}_i\} = \underset{\mathbf{s}_k, \mathbf{d}_l}{\operatorname{argmin}} \left\| \mathbf{r}_i - \rho_1 \sum_{m=1}^M f_{mi}^{(1)} \mathbf{s}_m^k - \rho_2 \sum_{m=1}^M g_{mi}^{(1)} \mathbf{d}_m^l \right\|_F^2 , \quad (32)$$

where $\rho_1 = \sqrt{\frac{P_1 T}{M}}$ and $\rho_2 = \sqrt{\frac{P_2 T}{N}}$.

Applying the PDF strategy in [10], the relay transmit vectors are designed as follows

$$\mathbf{t}_i = \sqrt{\frac{P_3 T}{2M}} A_i \left(\sum_{m=1}^M \hat{f}_{mi}^{(1)} \hat{\tilde{\mathbf{s}}}_m + \sum_{n=1}^N \hat{g}_{ni}^{(1)} \hat{\tilde{\mathbf{d}}}_n \right). \quad (33)$$

The received signals at the terminals are very similar to those of the 3 time slots case, and can be easily derived.

V. Examples Using $Sp(2)$

In this section, we show how can the $Sp(2)$ space-time code be applied to a two-way relaying network. First, we consider a network with four relays, and assume that the terminals are equipped with just one antenna each (Fig. 1(a)). Then, in the following section, we apply the $Sp(2)$ space-time code to a two-way relaying network containing two relays and having terminals with two antennas each (Fig. 1(b)).

A. Network (a): $R = 4$, $M = N = 1$

As depicted in Fig. 1(a), T_1 and T_2 have each $M = N = 1$ antenna, and the data exchange between them is assisted using $R = 4$ relays, with one antenna each. T_1 sends the vector $\mathbf{s} = [u_1^s, u_2^s, u_3^s, u_4^s]^T$, while T_2 transmits the vector $\mathbf{d} = [u_1^d, u_2^d, u_3^d, u_4^d]^T$, where the definition of u_l^s and u_k^d is given in (4).

We would like to get the $Sp(2)$ code at the terminals. For that, we choose the four relay matrices as follows

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
$$A_3 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \tag{34}$$

and ensure that relay 2 and 3 conjugate their transmit signals. Consequently, the space-time code Φ would have the following structure

$$\begin{aligned} \Phi &= [A_1\mathbf{s} \quad A_2\mathbf{s}^* \quad A_3\mathbf{s}^* \quad A_4\mathbf{s}] \\ &= \begin{bmatrix} u_1^s & -u_2^{s*} & -u_3^{s*} & u_4^s \\ u_2^s & u_1^{s*} & -u_4^{s*} & -u_3^s \\ u_3^s & -u_4^{s*} & u_1^{s*} & -u_2^s \\ u_4^s & u_3^{s*} & u_2^{s*} & u_1^s \end{bmatrix}, \end{aligned} \tag{35}$$

which is the $Sp(2)$ space-time code defined previously (5).

B. Network (b): $R = 2$, $M = N = 2$

In this section, we assume that each terminal is equipped with $M = N = 2$ antennas, and consider $R = 2$ relays that help in the data exchange. T_1 and T_2 send the following symbol matrices, respectively

$$S = \begin{bmatrix} u_1^s & -u_2^{s*} \\ u_2^s & u_1^{s*} \\ u_3^s & -u_4^{s*} \\ u_4^s & u_3^{s*} \end{bmatrix} \qquad D = \begin{bmatrix} u_1^d & -u_2^{d*} \\ u_2^d & u_1^{d*} \\ u_3^d & -u_4^{d*} \\ u_4^d & u_3^{d*} \end{bmatrix}, \tag{36}$$

where each column of the matrix is sent by one antenna.

By setting the relay matrices as follows

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \tag{37}$$

and by ensuring that relay 2 conjugates its transmit vector, we would distributively build the $Sp(2)$ space-time code at both terminals.

VI. Simulation Results

In this section, we present the simulation results of the two examples shown in the previous section. On our generated plots, we show the Block Error Rate (BLER) curves versus the power used in one time slot. To fairly compare the different schemes, we consider an equal data rate over the 4, 3, and 2 time slots protocols. Consequently, the symbols used in the 4 time slots protocol are chosen from 5-PSK and 7-PSK constellations, which gives a bit rate of 2.56 bit/TS. For the 3 time slots case, the symbols are chosen from 3-PSK and 5-PSK constellations, which results in a 2.6 bit/TS bit rate. Similarly, the symbols for the 2 time slots protocol are chosen from 2-PSK and 3-PSK constellations, which gives a rate of 2.58 bit/TS. We note here that for the $Sp(2)$ code to achieve a full diversity, the symbol constellation sizes should be relatively prime [6], [11]. The channel coefficients are chosen as $f_{mi}^{(\tau)} \sim \mathcal{CN}(0,1)$ and $g_{mi}^{(\tau)} \sim \mathcal{CN}(0,1)$, $\forall i, m$.

In Fig. 3, we compare the performance of the $Sp(2)$ STC with that of the Real Orthogonal STC. The symbols for the real orthogonal STC were chosen from a PAM constellation, with the same sizes as in the $Sp(2)$ case. Moreover, we simulate the performance in AF, and using the 2 time slots protocol. The results show that the performance of the $Sp(2)$ is 1-dB better than that of the real orthogonal. Also, one can see that the PDF strategy that we applied outperforms the AF scheme by approximately 2-dB, but at the cost of more complexity at the relays, which need to decode and estimate the channels. At higher rates, the $Sp(2)$ code shows an even higher diversity gain than that of the real orthogonal. Fig. 4 compares the

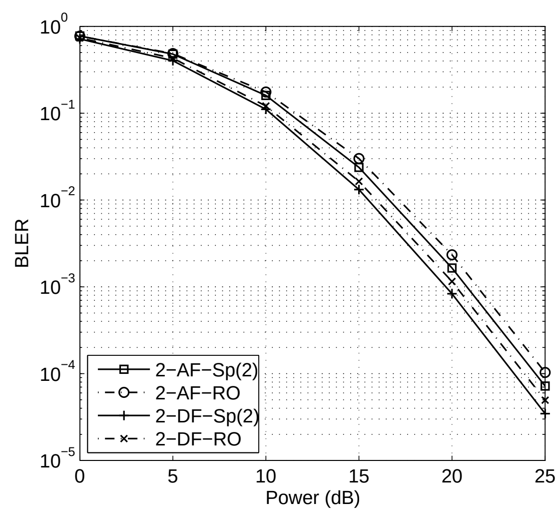


Fig. 3. Performance comparison of $Sp(2)$ and Real Orthogonal STCs using the 2 time slots protocol in network (a).

performance of the $Sp(2)$ STC in network (a) and (b) when using the AF cooperative strategy and the 4, 3, and 2 time slots protocols. The 2 time slots protocol is shown to be the best among the other protocols in both networks, and outperforms the conventional 4 time slots protocol by approximately 3-dB. As discussed before, the 2 time slots protocol shows a better performance than the 3 time slots protocol due to the less accumulated noise power at the terminals. The figure also depicts the gain when the terminals are equipped with more than one antenna. As it can be seen, this shifts the BLER curves by approximately 4-dB to the left.

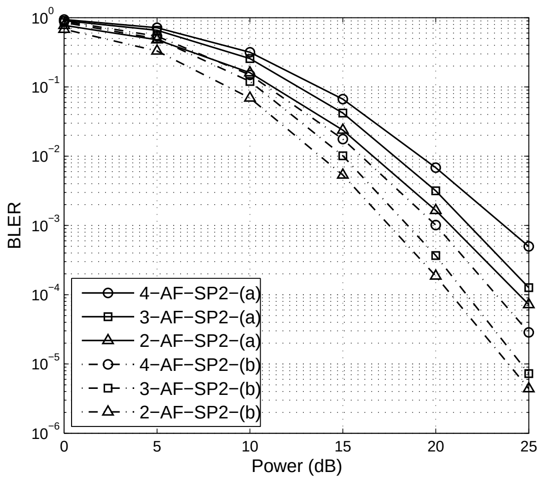


Fig. 4. AF performance comparison of $Sp(2)$ in 4, 3, and 2 time slots protocols in networks (a) and (b).

In Fig. 5, we compare the performance of the $Sp(2)$ code using the PDF strategy and the 4, 3, and 2 time slots protocols in the two networks (a) and (b). The simulations show that the 3 time slots protocol in this case performs better than the 2 time slots protocol. In fact, in the 3 time slots case, the relays receive the symbols of the terminals separated, and decode each one alone, while in the 2 time slots case, the relays receive both terminals' data combined and try to decode a pair of symbols. This can badly affect the decoding and that's why the 3 time slots protocol performs better. Again, one can see the improvement in the BLER when using network (b).

VII. Conclusion

We have evaluated the performance of distributed space-time codes for two-way relaying based on the symplectic Lie Group $Sp(n)$. Worked-out examples for both single antenna systems with 4 relays and two antenna systems with two relays based on $Sp(2)$ codes were presented. We have included Amplify-and-Forward and Decode-and-Forward scenarios, with protocols requiring 4, 3 and 2 time slots for information exchange between the terminals. The results show that these codes are appropriate for two-way relaying since they achieve the performance (throughput) gain. Further, we elaborated on the possible extension to systems with more

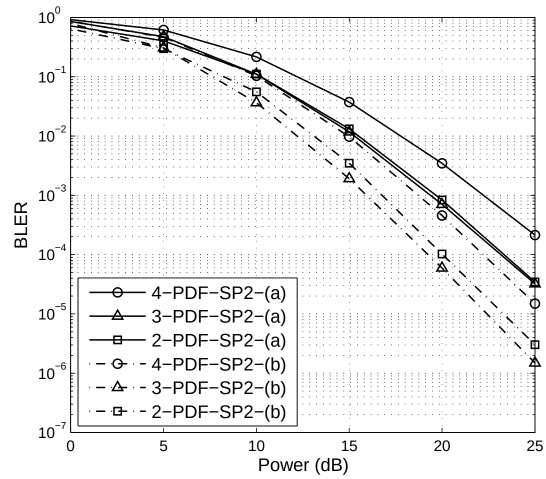


Fig. 5. PDF performance comparison of $Sp(2)$ in 4, 3, and 2 time slots protocols in networks (a) and (b).

relays/transmit antennas based on the geometry of the Lie group $Sp(n)$.

References

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. and Prob.*, 1961, pp. 611–644.

[2] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - part i: system description," *IEEE Trans. on Communications*, vol. 51, no. 1927:1938, Nov. 2003.

[3] ——, "User cooperation diversity - part ii: implementation aspects and performance analysis," *IEEE Trans. on Communications*, vol. 51, pp. 1939–1948, Nov. 2003.

[4] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. on Information Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[5] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: performance limits and space-time signal design," *IEEE Jour. on Selected Areas in Communications*, vol. 22, pp. 1099–1109, Aug. 2004.

[6] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wirel. Commun.*, 2006.

[7] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE Int. Symposium on Inf. Theory*, 2006, (Seattle, (WA)).

[8] P. Popovski and H. Yomo, "Wireless network coding by amplify-and-forward for bi-directional traffic flows," *IEEE Communications Letters*, 2007.

[9] Y. Jing and B. Hassibi, "Design of fully diverse multiple-antenna codes based on sp(2)," *IEEE Trans. Inform. Theory*, 2004.

[10] T. Cui, T. Ho, and A. Nallanathan, "Distributed space-time coding for two-way wireless relay networks," in *Proc. of IEEE International Conference on Communications (ICC)*, 2008.

[11] Y. Jing, "Space-time code design and its applications in wireless networks," Ph.D. dissertation, California Institute of Technology, 2004.

[12] Z. Utkovski, P. Chen, and J. Lindner, "Some geometric methods for construction of space-time codes in grassmann manifolds," in *Proc. ALLERTON Conference on Communication, Control and Computing*, 2008, (Urbana-Champaign).

[13] O. Henkel, "Space time codes from permutation codes," in *Proc. IEEE GlobeCom*, 2006, (San Francisco).

[14] Y. Jing and B. Hassibi, "Cooperative diversity in wireless relay networks with multiple-antenna nodes," *IEEE Transactions on Signal Processing*.

see the program

Arithmetic Extended-Mapping for BICM-ID with Repetition Codes

Takashi Yano

Central Research Laboratory, Hitachi, Ltd.
1-280 Higashi-koigakubo, Kokubunji,
Tokyo 185-8601, Japan
E-mail: takashi.yano.rh@hitachi.com

Tad Matsumoto

Information Theory and Signal Processing Laboratory,
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
E-mail: matumoto@jaist.ac.jp
and
Center for Wireless Communications, University of Oulu, Finland
E-mail: tadashi.matsumoto@ee.oulu.fi

Abstract — **A novel arithmetic extended mapping technique is proposed in this paper for bit-interleaved coded modulation with iterative decoding (BICM-ID) with repetition codes as well as for their modifications. With the proposed mapping scheme, the computational complexity of BICM-ID is significantly reduced compared with conventional quadrature amplitude modulation (QAM) mapping schemes Performances of BICM-ID using the proposed extended mapping rules are also verified both by the extrinsic information transfer (EXIT) analysis and bit error rate (BER) simulations. The results show that with the proposed BICM-ID scheme, the code rate, at which, given the channel signal-to-noise power ratio (SNR) bit error rate can made arbitrarily small, is almost 90% of channel capacity, and such excellent property holds at relatively wide range of SNR values.**

Keywords- extended mapping; repetition code; EXIT; BICM-ID; turbo coding; channel capacity;

I. INTRODUCTION

It has long been believed that when designing wireless communication system using high order modulation such as 16 quadrature amplitude modulation (QAM), bit-wise separability from a symbol corresponding to the bit patterns allocated to the symbol is most important. Gray mapping is the most widely used modulation rule because of the reason described above.

Recently, a technique called bit interleaved coded modulation with iterative decoding (BICM-ID) [1] has been proposed, of which transmitter is a concatenation of an encoder and bit-to-symbol mapper, separated by an interleaver. In the receiver side, the iterative processing for demapping and decoding technique takes place. This is quite reasonable, because BICM-ID systems can be viewed as a serially concatenated turbo system, where the inner component is mapper and the outer component is the channel code used. Therefore, by iterations, bit separability can be enhanced, and bit error rate performance, as a whole can be significantly improved. In [2], it is shown that a BICM-ID scheme comprised of a combination of a constraint length 2 (K=2) convolutional code and 8 amplitude shift keying (ASK) with (non-Gray) mapping rule, optimized assuming the availability of full *a priori* information, can achieve near Shannon capacity. Since the K=2 (memory 1) convolutional code is a very "weak" code, it was rarely used in practical communication systems that do not use the iterative detection. In [2], it is shown through extrinsic information transfer (EXIT) analysis [3] that the reason why such a simple structure can achieve excellent performance is because the EXIT functions of those two components are well-matched while the tunnel still opens until *a posteriori* mutual information reaches very close to one.

Another simpler BICM-ID scheme has been proposed in [4], where encoder is a concatenation of a single parity check code and a repetition code. The coded sequence is interleaved, and mapped on to a signal point in the utilized modulation format, where so called extended mapping [8] is assumed; with extended mapping more than M bits are allocated to one constellation point, when 2^M-QAM is used. It is shown that the BICM-ID proposed in [4] can also achieve near capacity performance even the structure is extremely simple. However, since determining the optimal bit patterns allocated to the each constellation point for extended mapping that best matches the code structure belongs to the combinatorial problem, it is not practical to directly apply the Ref. [4]'s technique to higher order modulation schemes.

This paper proposes a new BICM-ID scheme, arithmetic mapping BICM-ID, for the systems using higher order modulation. Our proposed technique can also achieve very close-capacity performance. Furthermore, our proposed scheme replace the symbol-level extended mapping by bit-reduction encoding, which can be seen as alternative way of mapping multiple patterns to the same constellation point. This technique eliminates the likelihood calculations for more than 2^M patterns, resulting in significant reduction in computational complexity. This paper also proposes the use of per-dimension non-Gray ASK to eliminate the difficulty in obtaining the optimal allocation pattern, which, as noted above, falls into a combinatory optimization problem.

Results of the EXIT analysis and BER simulation results, conducted to verify the effectiveness of the proposed technique are presented in this paper. It is shown in this paper that the results are exactly consistent with each other. It is noted that this contribution forms a sister paper of another contribution [11] made by one of the authors of this paper.

II. BICM-ID WITH REPETITION CODE

A. System Model

In this paper, BICM-ID scheme proposed in [4] is assumed. The schematic diagram of Ref [4]'s proposed technique is

see the program

shown in Fig. 1, but because of the space limitation, this paper only provide important properties/observations of the Ref. [4]'s technique. For more details, please see [4].

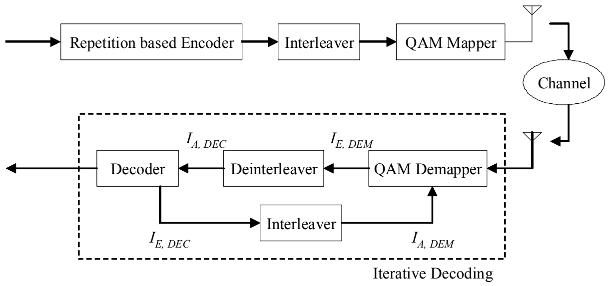


Fig. 1. System Model.

B. Designing BICM-ID with Repetition Code

As mentioned above, even a "weak" code can be utilized in BICM-ID if the EXIT curves are well matched to that of the demapper. The EXIT function of the repetition code is expressed as

$$I_{E,VND}(I_A) = J\left(\sqrt{d_v - 1} \cdot J^{-1}(I_A)\right) \tag{1}$$

where d_v denotes the repetition times, $J(\cdot)$ is the function introduced in [3] that converts the standard deviation of the log likelihood ratio (LLR) to mutual information in AWGN channels, $J^{-1}(\cdot)$ is its inverse function, and I_A is *a priori* mutual information between the coded bit and the decoder input LLR [5], [6]. Fig. 2 shows EXIT curves of repetition codes with, as a parameter, several different repetition times, referred to as variable node degree d_v. It is found that the EXIT curves exhibit convex shape with values between the mutual information value being 0 and 1, of which tendency is common to the variable node degree d_v larger than 2, and the higher the repetition time is, the sharper the convexity of the shape.

Optimization of QAM mapping rule for BICM-ID schemes are discussed in [2], [7] and [8]. The modulation schemes used in these references are non-Gray mappings, of which EXIT curves exhibits relatively sharp rising decay so that they are matched to the EXIT curve of short memory code decoders. This paper however, their decay is not sharp enough in convexity to better match the repetition codes.

Fig. 3 shows EXIT curves of an non-Gray mapped 256-QAM demepper at SNR=0, 5, 10, 15 and 20dB. It is found that non-Gray mapping, in general, with per-symbol bits allocation larger than two times the channel capacity has its EXIT curve exhibiting rising convex shape. In the case of Fig. 3, the mapping rule is optimized such that with full *a priori* information, the extrinsic mutual information has the largest value. Hence, extrinsic mutual information at the EXIT curve's right most point is increased, while that the left most pint is decreased. However, the area below demapper's EXIT curve is equal to averaged constellation constrained capacity (CCC) per bit, according to the area property [9]. Thus the number of bits assigned to one constellation point is larger than two times the channel capacity, EXIT curve becomes convex shape. With 256-QAM, for example, 8 bits are assigned to one modulation symbol while the channel capacities are 1.00, 2.06, 3.46, 5.23 and 6.66 bits/symbol at SNR=0, 5, 10, 15 and 20 dB, respectively. At SNR=0, 5 and 10 dB, number of bits assigned to one modulation symbol is larger than two times the capacity. As a result, the EXIT curves of 256-QAM optimized with *a priori* information exhibit convex shapes at SNR=10, 5 and 0 dB, as shown in Fig. 3.

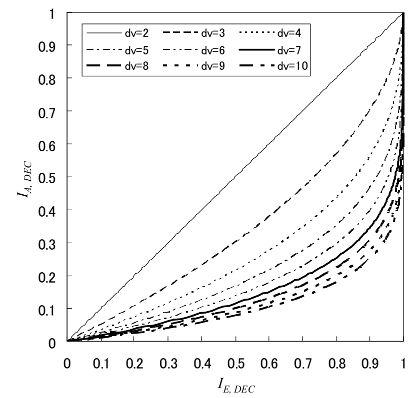


Fig. 2. EXIT curves of repetition codes

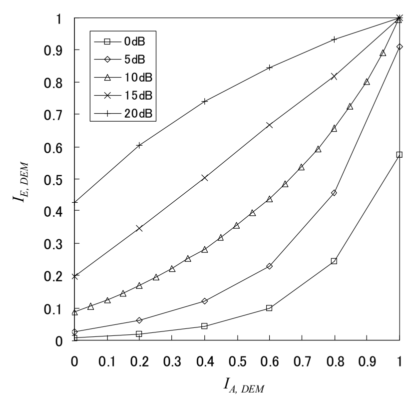


Fig. 3. EXIT curves of a 256-QAM demepper at SNR=0, 5, 10, 15 and 20dB.

Ref. [4] introduces two additional techniques to achieve flexibility in changing the shape of the EXIT curve; one is irregular degree allocation to the variable nodes (=repetition times take different values in one frame to be transmitted), and the other serial concatenation of the repetition code (inner code) and single parity check code (outer code).

The EXIT function of variable node with irregular degree allocation is expressed as

$$I_{E,VND}(I_A) = \frac{\sum_i a_i d_{v,i} \cdot J\left(\sqrt{d_{v,i} - 1} \cdot J^{-1}(I_A)\right)}{\sum_i a_i d_{v,i}}, \tag{2}$$

where a_i denotes the number of variable node of which degree is $d_{v,i}$.

see the program

When the serially concatenating the single parity check code as the outer code, parity bits are appended, according to the information bits. Afterwards, information bits as well as the parity bits are encoded by the repetition code encoder. Fig. 4 shows effect of single parity check codes, where it is found that the right part of the EXIT curve is further shifted to the right. On the contrary, it is found from Eq. (2) that irregular degree allocation affects the whole portion of the curve.

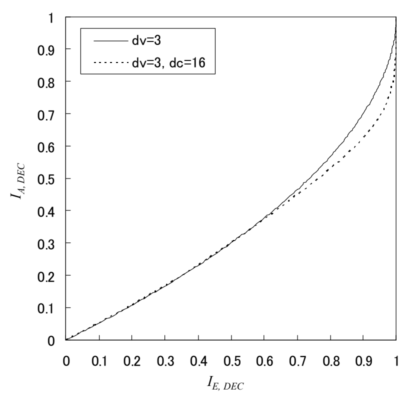


Fig. 4. Effect of single parity check code for repetition code.

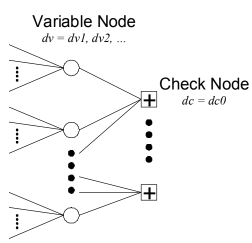


Fig. 5. Tanner graph of repetition-based code in [4].

The whole decoder structure presented in [4] is show in Fig. 5. The EXIT function of the decoder is expressed as

$$I_{E,DEC}(I_{A,DEC}) = \frac{\sum_i a_i \cdot d_{v,i} \cdot J\left(\sqrt{(d_{v,i}-1)\cdot J^{-1}(I_{A,DEC})^2 + J^{-1}(I_{E,CND,i})^2}\right)}{\sum_i a_i \cdot d_{v,i}}, \quad (3)$$

where $I_{E,CND,i}$ is the extrinsic mutual information output from the check node connected to variable node with degree $d_{v,i}$, and $I_{E,CND,i}$ is the output of the check node decoder. It is well known that $I_{E,CND,i}$ can be approximated by

$$I_{E,CND,i} \approx 1 - J\left(\sqrt{d_c - 1}\cdot J^{-1}(1 - I_{A,CND,i})\right), \quad (4)$$

where $I_{A,CND,i}$ is *a priori* information input to the check node and d_c is degree of the check node, as described in [5], [6]. If each check node is connected uniformly to variable nodes the mutual information of the variable nodes decoder output can be expressed as

$$I_{A,CND,i} = \frac{\sum_j a_j \cdot J\left(\sqrt{d_{v,j}}\cdot J^{-1}(I_{A,DEC})\right)}{\sum_j a_j}. \quad (5)$$

If each check node is connected to only the variable nodes having the same degree $d_{v,i}$, $I_{A,CND,i}$ is expressed as

$$I_{A,CND,i} = J\left(\sqrt{d_{v,i}}\cdot J^{-1}(I_{A,DEC})\right). \quad (6)$$

As described above, Ref. [4]'s proposed coding scheme enables flexible design of its EXIT curve shape so as to match the decoder EXIT curve shape of demapper. Furthermore, local iteration in the decoder is not required, because each variable node is connected to only one check node. If not, iteration between variable nodes and check nodes is necessary, like in the LDPC codes. This significantly reduces the computational complexity of the decoding process.

C. *Coding rate of the Repetition Code with Single Parity Check Code.*

First, one single parity check bit is appended to each set of $d_c - 1$ information bits, and thus the coding rate of the single parity check code is $(d_c - 1)/d_c$. Afterwards, a_i bits are repeated by repetition encoder with degree $d_{v,i}$, thus the total number of coded bits are $\sum_i a_i d_{v,i}$ while the number of the bits before repetition is $\sum_i a_i$. Thus the total coding rate of the repetition code with single parity check code, described above, is shown as

$$\frac{\sum_i a_i}{\sum_i a_i d_{v,i}} \cdot \frac{d_c - 1}{d_c}. \quad (7)$$

III. Extended Mapping Design

Even though the coding scheme proposed in [4] provides us with a lot of flexibility in designing EXIT curve, we need further degrees of freedom in designing the EXIT curve for the QAM mapping rule so that demapper and decoder EXIT curves have exact matching. To achieve this goal, QAM with per-symbol bit number larger than twice the channel capacity, optimized with full *a priori* information, provides us with a good solution. This section proposes a new, but very simple coding scheme that can provides further design flexibility.

A. *Demapping Operations*

If an M bit pattern $(b_0, b_1, \ldots, b_{M-1})$ is mapped onto one of the 2^M constellation points $(c_0, c_1, \ldots, c_{2^M-1})$, the extrinsic LLR of the coded bit, output from the demapper, can be calculate as

$$L_e(b_i) = \log \frac{\sum_{c_k | c_k(b_i)=0} p(x|c_k)\cdot p(c_k | b_i = 0)}{\sum_{c_k | c_k(b_i)=1} p(x|c_k)\cdot p(c_k | b_i = 1)} \quad (8)$$

for the bit b_i, where x is the received sampled value of the symbol. The probability density function (pdf) of x is given by

$$p(x|c_k) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{|x - c_k|^2}{2\sigma^2}\right) \quad (9)$$

see the program

in additive white Gaussian noise (AWGN) channel with a noise variance σ^2 and

$$p(c_k \mid b_i = b) = \prod_{j \neq i} \frac{\exp(-c_k(b_j)L_a(b_j))}{1+\exp(-L_a(b_j))}, \tag{10}$$

where $c_k(b_j)$ is the j-th bit value (0 or 1) on the constellation point c_k and $L_a(b_j)$ is the *a priori* LLR of the bit b_j. Eq. (8) can be simplified to

$$L_e(b_i) = \log \frac{\sum_{c_k | c_k(b_i)=0} \exp\left(-\frac{|x-c_k|^2}{2\sigma^2} - \sum_j c_k(b_j)L_a(b_j)\right)}{\sum_{c_k | c_k(b_i)=1} \exp\left(-\frac{|x-c_k|^2}{2\sigma^2} - \sum_j c_k(b_j)L_a(b_j)\right)} - L_a(b_i). \tag{11}$$

In practice, the demapper has to calculate

$$\frac{|x-c_k|^2}{2\sigma^2} + \sum_j c_k(b_j)L_a(b_j) \tag{12}$$

for every c_k first, then $L_e(b_i)$ using (11). This means that the complexity of demapper is in exponential order of number of bits assigned to one modulation symbol. Though this is common to any BICM-ID schemes, the BICM-ID with repetition code is even more complex because it requires larger number of bits assigned to modulation symbol. While it does not affect much at low SNR, it is computationally expensive at relatively high SNR. If SNR is 20dB, for example, 12~16 bits have to be allocated to one symbol because the capacity is more than 6 bits/symbol. Thus, demapper have to calculate the values of (12) 4096~65536 (2^{12}~2^{16}) times per each symbol, which may impose unacceptable computational burden for practical applications. The complexity problem due to exp() and log() in (11) may well be avoided by the Max-Log-Map algorithm, but that for (12) still can not be avoided.

B. *Extended Mapping*

In [8], a mapping technique, "Extended Mapping," is introduced, where the number of bits mapped to each constellation point is larger than M, withthe modulation having 2^M signal points. The bit patterns allocated to one symbol is determined such that the demapper EXIT function has a preferable shape. Ref. [4] presents a BICM-ID technique with extended mapping only with 4-QAM modulation. In this paper, it is shown that using extended mapping, BICM-ID can perform well even if the number of constellation size is larger. Fig. 6 compares EXIT curves of 256-QAM with extended 16-QAM having 8 bits per one symbol. Both are optimized such that with full *a priori* information, the extrinsic mutual information takes the largest value. It is found that both curves are almost the same at SNR=7dB. This is because the number of constellation points 16 (=2^4) is enough to deliver ~2.6 bits/symbol, which is the channel capacity at SNR=7dB. Therefore, extended mapping is a good choice because with if extended mapping, the total number of the first terms of (12) is reduced. However, the second terms are still unchanged.

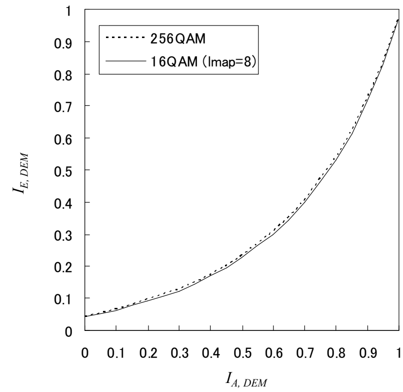


Fig. 6. Comparison of EXIT curves of a 256-QAM and of an extended 16-QAM (l_{map}=8) demapper (SNR=7dB)

C. *Arithmetic Extended Mapping*

To reduce the complexity due to the second term of (12), we proposed a novel extended mapping technique, named "Arithmetic Extended Mapping." Our proposed scheme further extends the mapping rule according to an "arithmetical" rule. First, the number of bits to be mapped on to a symbol is reduced by a set of check node encoders, referred to as "bit-reduction encoder" in this paper. The purpose of the "bit-reduction encoding" is to reduce the number of the bits to be mapped onto one symbol to the number which standard (i.e. not extended) mapping rule allocates to each symbol. An example of bit-reduction encoder is shown in Fig. 7, where the symbol boxed-plus denotes check sum (exclusive-OR) operation. Eight bits (b_0, b_1,..., b_7) are input to the bit-reduction encoder and four bits (m_0, m_1,..., m_3) are output. Afterwards, the bits (m_0, m_1,..., m_3) are mapped onto one 16-QAM symbol by the normal (non-extended) mapping rule. With this operation, still 8 bits (b_0, b_1,..., b_7) can be transmitted using only 16 combinations, with help of *a priori* information at the receiver.

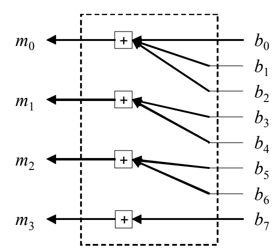


Fig. 7. Exmample of bit-reduction encoder.

In the receiver, first the demapper calculates from the received symbol with no *a priori* information the extrinsic information for the bits (m_0, m_1,..., m_3). The extrinsic information of (b_0, b_1,..., b_7) is then calculated using extrinsic information for the bits (m_0, m_1,..., m_3) provided by the bit-reduction decoder, as

$$L(u_1 \oplus u_2 \oplus \cdots \oplus u_n) = 2\tanh^{-1}\left(\prod_{i=1}^{n} \tanh\frac{L(u_i)}{2}\right), \tag{13}$$

which is exactly the same operation as a check node decoder of the LDPC codes. For example, the extrinsic output for b_0 in Fig.

see the program

7 can be calculated by (13), using extrinsic LLR for m_0 from the demapper and *a priori* information for b_1 and b_2. In decoding, the *a priori* information from the decoder of the outer code (=Ref. [4]'s proposed structure), which is provided via the interleaver, is used by the check node decoder to update the LLR, according to (13). The output of the check node decoder is input to demapper as *a priori* information. The demapper calculates the updated extrinsic information using the *a priori* information from the check node decoders using (11) and (12).

It should be noted here that for any codeword, bit b_i should not affect more than one modulation bit m_j in the bit-reduction encoder. This is because (13) is correct only if the all inputs to the check node decoder are considered as independent. Because the outputs of the QAM demapper corresponding to the identical received signal are not independent, another interleaver should be placed between QAM demappers and check node decoders, similarly to the MIMO case, as shown in [5]. This requires local iterations to be performed, which increases computational complexity.

In the proposed scheme, the computational complexity of the demapper is reduced significantly because the number of the second term in (12) is not affected by the length extension for mapping. The number of constellation points should be only necessary enough to achieve near Shannon capacity performance. The convexity of the EXIT curve can be realized by bit-reduction encoder, as well as the QAM mapping rule. In other words, the shape of EXIT curve can be designed by both mapping rule of the modulator and check node degree allocations in the bit-reduction encoder.

Though the flexibility of design of the proposed scheme is less than the extended mapping introduced in [8], still it preserves large degrees-of-freedom in flexibly changing the shape of the EXIT curve so that it is better-matched to the EXIT curve of the repetition code, as described later.

The computational complexity of the demapping can further be greatly reduced if the mapping is performed independently in-phase and quadrature dimensions. In Fig. 8, two 8-bit codewords (each denoted by b_0, b_1, ..., b_7 and b_8, b_9, ..., b_{15}) are reduce to two 4-bits (m_0, m_1,..., m_3) codewords, respectively by the bit-reduction encoder described above. The each 4-bits codeword is mapped on to a non-Gray 16-ASK modulation, independently between in-phase and quadrature dimensions (I-ch and Q-ch, respectively), which are then combined into a complex I+*j*Q symbol. This technique is very suitable especially when high order modulation is required at a relatively large SNR range, say, SNR=20dB. This is because the complexity for demaping of 256-QAM may still be too large for practical use. Therefore, splitting the complex modulation into two one-dimensional modulation is very effective in reducing the complexity. In this case, the calculation of (12) for only 32 points (16 points for each 16-ASK demapper) is needed, and the processing for check node decoder is easy even if 16 bits are mapped onto one modulation symbol. The complexity is extremely low and it is no longer "exponential order" of the modulation multiplicity.

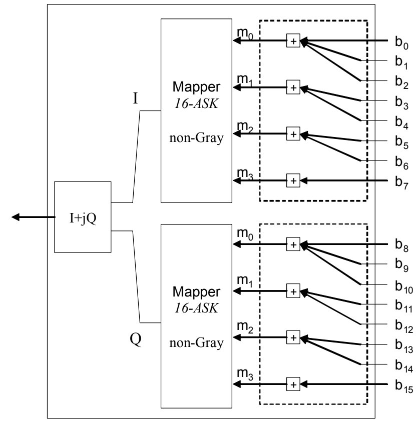


Fig. 8. Arithmetic extended mapping 256-QAM (l=16)

D. Designing Arithmetic Extended Mapping

Even with the simplified scheme shown in Fig. 8, our proposed arithmetic extended mapper still has enough flexibility to the EXIT curve shape so that it is well-matched to the repetition code. We also preserve the flexibility in determining the degree allocation of check nodes in bit-reduction encoder as well as mapping rule of ASK (or QAM). The extrinsic information from the decoder for arithmetic extended mapper, without *a priori* information, is expressed as

$$I_{E,DEM}(0) = \frac{\sum_{i|d_{c,i}=1} I_{E,m_i}(0)}{\sum_{i|d_{c,i}=1} 1} \qquad (14)$$

where $I_{E,m_i}(0)$ and $d_{c,i}$ denotes the extrinsic output of the non-Gray demapper without *a priori* information, and the degree of a check node connected to the modulation bit m_i. The extrinsic information $I_{E,m_i}(0)$ of the bit m_i, to which the degree of the check node connected is more than 1, does not affect the extrinsic information output from arithmetic extended demapper without *a priori* information; this is because the check node decoder output is zero if one of its inputs is zero, as shown in (13). This means at least one check node has to have degrees d_c=1, as shown in Fig. 7. Otherwise, the demapper cannot extract any information from received signal at the first iteration. The less number of check node decoders with d_c=1, the less extrinsic information at the left most point. Moreover, the smaller the $I_{E,m_i}(0)$ values for check node decoders with d_c=1, the less extrinsic information at the left most point.

However, on the contrary, the extrinsic information with full *a priori* information from the arithmetic extended demapper is described as

$$I_{E,DEM}(1) = \frac{\sum_i d_{c,i} I_{E,m_i}(1)}{\sum_i d_{c,i}}, \qquad (15)$$

where $I_{E,m_i}(1)$ is the extrinsic output of the non-Gray demapper with full *a priori* information. The outputs of non-Gray demapper are weighted by $d_{c,i}$ in (15). This means that the

see the program

extrinsic information at the right most point gets larger if the check node decoder having higher degrees, is connected to the modulation bit m_i of which extrinsic output $I_{E,m_i}(1)$ from non-Gray demapper is larger. Thus, to maximize the extrinsic output with full *a priori* information, maximum allowed number of the check node degree has to be assigned to, at least, one modulation bit, and check node degrees=1 is assigned to the others. Then, optimization of mapping rule, which maximizes (15), is completed. Any results of the optimization directly conducted to the extended mapping in [8], are supposed to be equivalent to the result of the optimization above, if the optimization is perfectly performed for maximizing extrinsic output. This optimization can be performed by using binary switching algorithm (BSA) introduced in [10]. BSA often stacks on local maximums especially in the case of the total number of bits assigned is large. However, optimization of mapping rule of non-Gray mapping for arithmetic extended mapping rarely stacks on local maxima because the number of bits of the non-Gray mapping is smaller than that of extended mapping. Moreover, optimization of the arithmetic extended mapping using BSA converges in much less steps than direct optimization for extended mapping.

Fig. 9 shows an example of the result of the optimization. The optimization with weights of $d_{c,i}$={5, 1, 1, 1} and $d_{c,i}$={3, 2, 2, 1} using BSA result in the same mapping rule shown in Fig. 9. EXIT curves of the arithmetic extended mapping, which is a combination of non-Gray 16-ASK shown in Fig. 9 and the bit-reduction encoders with check node degree of $d_{c,i}$={5, 1, 1, 1} and $d_{c,i}$={3, 2, 2, 1}, are shown in Fig. 10 and Fig. 11, at SNR=15dB and 5dB, respectively. It is found that the extrinsic outputs without *a priori* information (at left most point) are reduced, while they are increased with full *a priori* information (at right most point), compared with that of the demapper itself for 16-ASK shown in Fig. 9. It is also found that the bit-reduction encoder with check node degrees allocation of {3, 2, 2, 1} is better matched with the EXIT curves of the repetition code at SNR=15dB, while {5, 1, 1, 1} is slightly better at SNR=5dB. As mentioned above, the result with $d_{c,i}$={5, 1, 1, 1} is supposed to maximize the extrinsic output with full *a priori* information. However, it is not always the best solution for the BICM-ID with the repetition code. In the case the output at right most point is high enough, it is better to reduce the left most hand, as shown in Fig. 10. Thus, we choose the degree allocation of $d_{c,i}$={3, 2, 2, 1} with non-Gray 16-ASK presented in Fig. 9 as an example in evaluations hereafter. The EXIT curves of the demapper for Fig. 8 with non-Gray 16-ASK mapping in Fig. 9 at several SNR points are shown in Fig. 12. It is shown that the EXIT curves exhibit rising convex shapes in a wide range of SNR. The demapper EXIT curve can be designed more flexibly using modulation doping technique proposed in our sister paper [11].

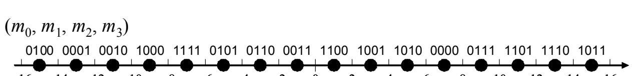


Fig. 9. Base Mapping (Optimized non-Gray 16-ASK)

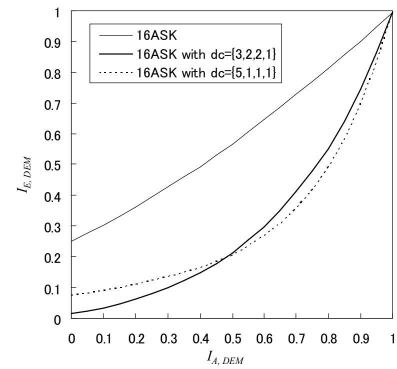


Fig. 10. EXIT curve of arithmetic extended mapping at SNR=15dB

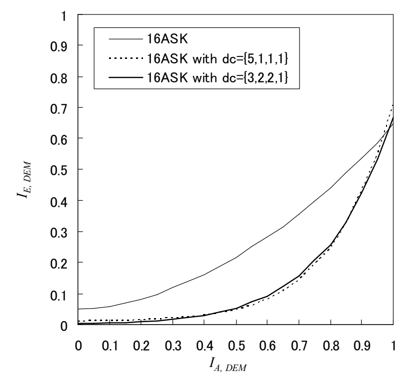


Fig. 11. EXIT curve of arithmetic extended mapping at SNR=5dB

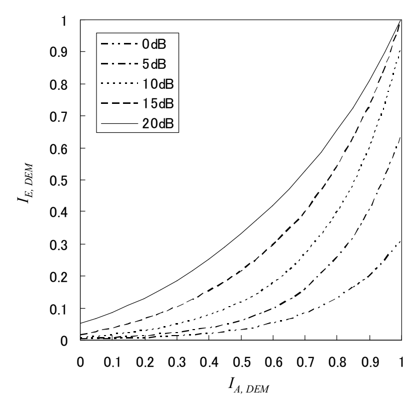


Fig. 12. EXIT curves of a demapper of proposed mapping at SNR=0, 5, 10, 15 and 20dB.

E. Design Example of the Repetition Codes for proposed mappings

A block diagram of the whole structure of the proposed BICM-ID scheme is shown in Fig. 13. The multiplexers/ demultiplexers for I-ch and Q-ch in Fig. 8 are not shown explicitly. The code design example for each SNR in Fig. 12 are listed in TABLE I. The EXIT curves of the codes, listed in TABLE I, are calculated using (3). Their shapes are well-matched to that of the demapper in Fig. 12 at SNR=20, 15, 10, 5 and 0dB as shown in Fig. 14~Fig. 18 respectively. The

see the program

combination of the arithmetic extended mapping, proposed in this paper, and repetition code with single parity check code, proposed in [4], is expected to perform near Shannon capacity performance, because the gap between EXIT curves of the decoder and demapper, which corresponds to the rate loss, is extremely small, as shown in Fig. 14~Fig. 18. The rate loss can not be reduced anymore, especially at SNR=15dB and 10dB. This means the bit rates achieved by the proposed BICM-ID almost reach the CCC of the QAM constellation.

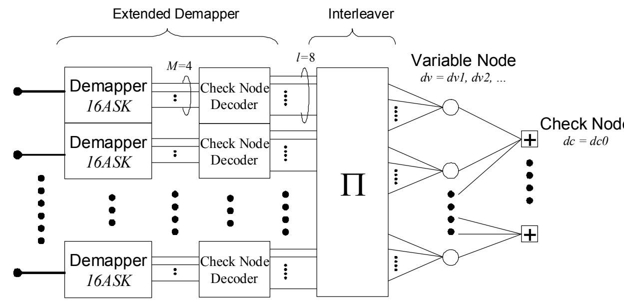


Fig. 13. Proposed BICM-ID scheme using arithmetic extended mapping.

TABLE I. CODE DESIGN EXAMPLES

Code Type	Code design		
	Variable node degrees	***Check node degree***	***Coding rate***
I	d_{v1}=11 (93 %), d_{v2}=100 (7 %)	d_c=8	0.0508
II	d_{v1}=5 (93 %), d_{v2}=50 (7 %), d_{v3}=100 (0.1 %)	d_c=32	0.1175
III	d_{v1}=3 (87 %), d_{v2}=15 (13 %), d_{v3}=80 (0.6 %)	d_c=512	0.1992
IV	d_{v1}=2 (44 %), d_{v2}=3 (35 %), d_{v3}=5 (14 %), d_{v4}=12 (7 %)	d_c=512	0.2876
V	d_{v1}=2 (55 %), d_{v2}=3 (33 %), d_{v3}=5 (12 %)	d_c=1024	0.3714

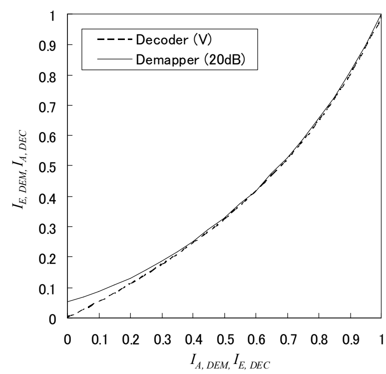


Fig. 14. EXIT chart of proposed mapping with the Code V (SNR=20dB)

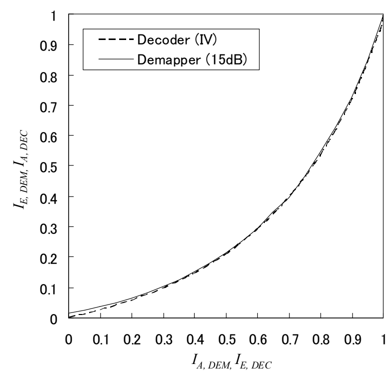


Fig. 15. EXIT chart of proposed mapping with the Code IV (SNR=15dB)

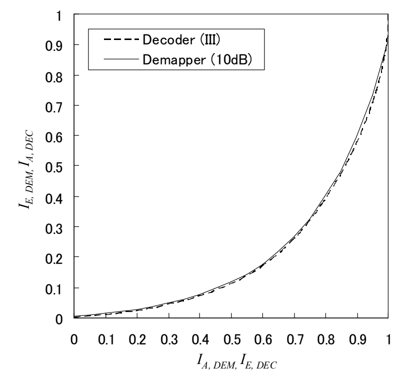


Fig. 16. EXIT chart of proposed mapping with the Code III (SNR=10dB)

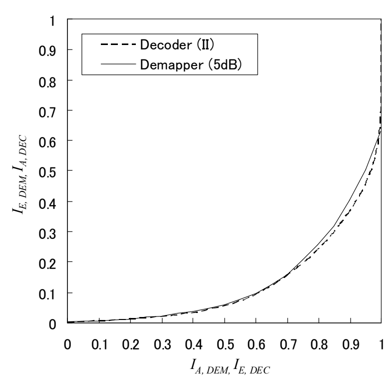


Fig. 17. EXIT chart of proposed mapping with the Code II (SNR=5dB)

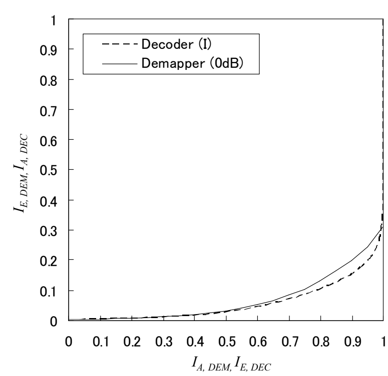


Fig. 18. EXIT chart of proposed mapping with the Code I (SNR=0dB)

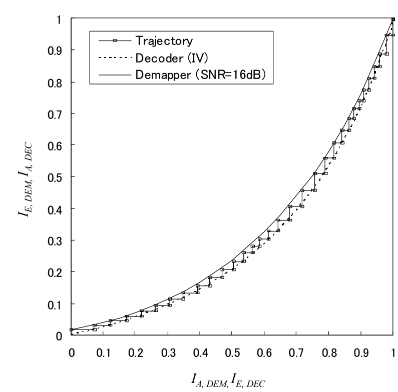


Fig. 19. Exemplifying trajectory of the proposed BICM-ID

IV. CHAIN SIMULATION RESULTS

A series of chain simulations assuming the conditions summarized in TABLE II were conducted. The block size of about 5000 information bits is chosen in the simulations for practicality. The codes listed in TABLE I are used, which are designed to match the arithmetic extended mapping shown in Fig. 8 with non-Gray 16-ASK mapping presented in Fig. 9 at

see the program

SNR=0, 5, 10, 15 and 20dB. An example of the snapshot trajectory on EXIT chart is shown in Fig. 19, together with its corresponding EXIT chart. Since the code IV is designed at SNR=15dB, narrow tunnel is constantly open between the mutual information 0 to almost 1 at SNR=16dB. The trajectory shown in Fig. 19 is very well matched with the EXIT curves. Fig. 20 shows the BER simulation results. Sharp turbo cliffs appear at the targeted SNRs. If the code size is larger, the turbo cliffs will be even shaper. It is also shown that BER floor at below 10^{-3} appears with code I and II, which are caused only by the intersection points of EXIT curves, because there are no cycles in the decoder. The BER floor can be reduced by redesigning the mapping rule since the mapper used in this simulation is designed to match at relatively high SNRs, thus the gaps between EXIT curves at SNR=0dB around right most point is lager than with higher SNRs, as shown in Fig. 18. If an arithmetic extended mapper with less number of bits assigned is used, the curves should be better-matched with repetition codes at low SNRs, say, SNR=0 and 5dB. Although proposed BICM-ID scheme has BER floor more or less due to the intersection point in EXIT chart, it can be designed so as to reduce the floor to a smaller level than the values allowed practically, using both flexibilities of the repetition code and the mapping scheme. This is one of the significant advantageous points of the proposed BICM-ID scheme.

The bit rates achieved by the proposed BICM-ID are shown in Fig. 21, as well as the bit rates achieved by the BICM-IDs presented in [2], [7] and [8] with “plus”, “cross” and “triangle” marks, respectively. Fig. 21 shows that our proposed BICM-ID outperforms the other BICM-ID schemes, even with its extremely simple structure.

TABLE II. SIMULATION CONDITIONS

Item	Values
Number of Information bits	about 5000 bits (Minimum multiple of d_c−1, which is greater than or equal to 5000)
Modulation	Arithmetic Extended Mapping of Fig. 8 with 16-ASK in Fig. 9.
Coding	Code I, II, III, IV and V in TABLE I.
Interleaver	Random interleaver
Data	Random binaries
Channel	AWGN
Number of Iterations	100

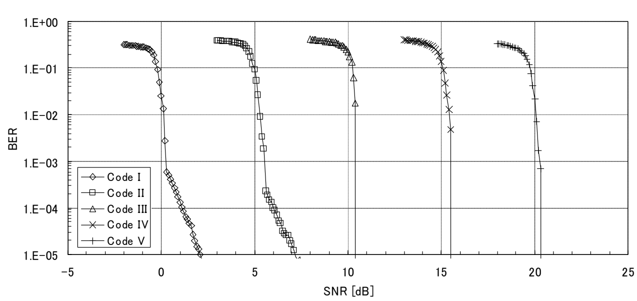


Fig. 20. Bit error rates simulation resuts of the proposed BICM-ID scheme.

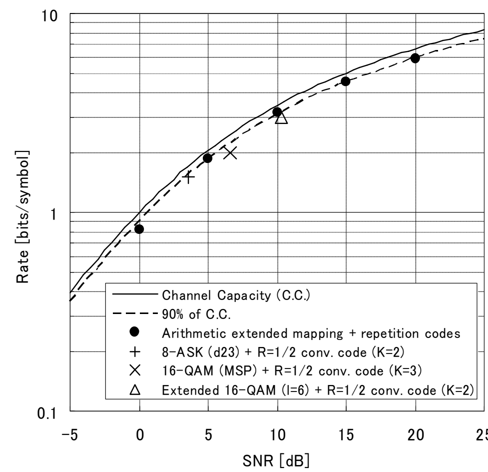


Fig. 21. Bit rates achieved by the proposed BICM-ID

V. CONCLUSION

A novel mapping technique, named arithmetic extended mapping, has been proposed for BICM-ID with repetition codes and their modifications. The computational complexity in the demapper can be reduced significantly using proposed extended mapper. Performance of proposed BICM-ID scheme has been verified both by EXIT analysis and BER simulations. The results show that the proposed BICM-ID can achieve almost 90% of channel capacity even at relatively high SNRs.

REFERENCES

[1] X. Li and J. A. Ritcey, “Bit-interleaved coded modulation with iterative decoding,” *IEEE Communications Letters*, vol.1, pp. 169–171, 1997.

[2] S. ten Brink, “Designing iterative decoding schemes with the extrinsic infomation transfer chart,” *AEÜ International Jurnal of Electronics and Communications*, vol. 54, no. 6, pp. 389-398, Novenber 2000.

[3] S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Transactions on Communications*, Vol. 49, No. 10, pp.1727-1737, October 2001.

[4] D. Zhao, A. Dauch and T. Matsumoto, “BICM-ID Using Extended Mapping and Repetition Code with Irregular Node Degree Allocation,” *2009-Spring IEEE Vehicular Technology Conference*, April 2009.

[5] S. ten Brink, “Design of Repeat-Accumulate Codes for Iterative Detection and Decoding,” *IEEE Transactions on Signal Processing*, Vol. 51, No. 11, pp.2764-2772, November 2003.

[6] S. ten Brink, G. Kramer and A. Ashikhmin, “Design of Low-Density Parity-Check Codes for Modulation and Detection,” *IEEE Transactions on Communications*, Vol. 52, No. 4, pp.670-678, April 2004.

[7] F. Schreckenbach, N. Gortz, J. Hagenauer and G. Bauch, “Optimized Symbol Mappings for Bit-Interleaved Coded Modulation with Iterative Decoding,” *IEEE GLOBECOM 2003*, pp.3316-3320, 2003.

[8] P. Henkel, “Extended Mappings for Bit-Interleaved Coded Modulation,” *IEEE PIMRC 2006*.

[9] A. Ashikhmin, G. Kramer and S. ten Brink, “Code Rate and the Area under Extrinsic Information Transfer Curves,” *IEEE International Symposium on Information Theory*, June 2002.

[10] K. Zeger and A. Gersho, “Pseudo-Gray coding,” *IEEE Transactions on Communications*, vol. 38, pp. 2147-2158, December 1990.

[11] D. Zhao, A. Dauch and T. Matsumoto, "Modulation Doping for Repetition Coded BICM-ID with Irregular Degree Allocation," *WSA 2009*, Febrrary 2009.

see the program

Modulation Doping for Repetition Coded BICM-ID with Irregular Degree Allocation

Dan Zhao
Japan Advanced Institute of Science and Technology, Japan
Email: dan.zhao@jaist.ac.jp

Axel Dauch
Information Technology
University of Ulm, Germany
Email: Axel.Dauch@gmx.de

Tad Matsumoto
Japan Advanced Institute of Science and Technology, Japan
Email: matumoto@jaist.ac.jp
and
Center for Wireless Communication at University of Oulu, Finland
Email: tadashi.matsumoto@ee.oulu.fi

Abstract—A new BICM-ID technique that makes efficient use of a mixture of standard Gray and anti-Gray extended mapping, combined with irregular degree-allocated repetition code with check nodes. It is shown that with the proposed technique, very low rate codes that can achieve turbo cliff at a very low E_s/N_0 value range can well be designed, and that the EXIT chart analysis and chain simulation results are exactly consistent with each other.

Index terms - extended mapping; repetition code; irregular node allocation; EXIT analysis; turbo cliff; capacity;

I. Introduction

In Bit-Interleaved Coded Modulation with Iterative Detection/Decoding (BICM-ID) [1], transmitter is comprised of a concatenation of encoder and bit-to-symbol mapper separated by a bit interleaver, and at the receiver, iterative processing is invoked, according to the standard turbo principle.

Performances of BICM-ID have to be evaluated by the convergence and asymptotic properties [2], which are represented by the threshold signal-to-noise power ratio (SNR) and bit error rate (BER) floor, respectively. In principle, since BICM-ID is a serially concatenated system, analyzing its performances can rely on the area property [3] of the extrinsic information transfer (EXIT) chart. Therefore, the transmission link design based on BICM-ID falls into the issue of matching between the demapper and decoder EXIT curves. Various efforts have been made in the seek of better matching between the two curves for minimizing the gap while still keeping the tunnel open, aiming, without requiring heavy detection/decoding complexity, at achieving lower threshold SNR and BER floor. Reference [4] introduces a technique that makes good matching between the detector and decoder EXIT curves using low density parity check (LDPC) code in multiple input multiple output (MIMO) spatial multiplexing systems.

It has been believed that a combination of Gray mapping and Turbo or LDPC codes achieves the best performance comparing with other combinations for BICM-ID. However, Ref. [5] proposes another approach towards achieving good matching between the two curves by introducing different mapping rules, such as anti-Gray mapping, which allows the use of even simpler codes, such as convolutional codes, and with which still it is possible to achieve BER pinch-off (corresponding to the threshold SNR). Another good idea that can provide us with design flexibility is extended mapping (EM) presented by [6], [7], [8] where with 2^m-QAM, more than m bits, l_{map}, are allocated to one signal point in the constellation. With this mapping rule, the left-most (LM) point of the demapper EXIT function has a lower value than with the Gray mapping, but the right-most (RM) point becomes higher. With this setting, the demapper EXIT function achieves better matching with even weaker codes such as small memory length convolution codes, and of the most practical importance is the fact that hardware complexity of the modulator and demodulator remains the same as that with the original 2^m- QAM.

In Ref [9], we proposed techniques that combine EM with an extremely simple code, repetition code, concatenated with single parity check code with irregular degree allocation. Even with this very simple combination, the EXIT curves of the demapper and decoder match well, and we can achieve BER pinch-off exactly at a designed SNR. The complexity for the Ref [9]'s proposed technique is extremely low, because EM does not require higher order modulation format and no iterations are needed in the decoder itself. Even with such simple structure, near-capacity performance can be achieved [9].

However, this technique is not suitable in designing BICM-ID with low rate codes that achieve BER pinch-off at very low SNR. This is because the LM point of the demapper EXIT function, corresponding to decoder feedback mutual information MI=0, is already very low, and there is not enough room to further lower the EXIT LM point while avoiding the intersection between the demapper and decoder curves. To solve this problem, this paper introduces modulation doping technique which mixes the labeling rules (=bit patterns allocated to each constellation point) for the extended anti-Gray mapping and the standard Gray mapping, at a certain ratio. Since the demapper EXIT function with Gray mapping is relatively flat, its LM point has the higher MI value than EM mapping rules. Therefore, with modulation doping the demapper EXIT LM point is lifted up from that without doping, and the amount depends on the doping ratio.

With the modulation doping technique combined with repetition coded EM BICM-ID with irregular degree allocation, we can achieve more degrees-of-freedom in choosing the parameter values so that we can achieve better matching between the demapper and decoder EXIT curves.

This work was supported by the Japanese government funding program, Grant-in-Aid for Scientific Research (B), No. 20360168..

see the program

The paper is organized as follows. The system model is described in Section II. The basic structures of the technique presented by [9] are summarized in Section II; The structures introduced in [9] has three forms; (A) is comprised of only EM and repetition code, (B) comprised of variable and check nodes with regular degree allocation, and (C) comprised of EM BICM-ID with irregular degree allocation. The notations, (A), (B), and (C), are commonly used to describe Sub-section Indexes in this paper to identify the structures. Their convergence performances are evaluated in Section III using EXIT chart, and chain simulation results are shown in Section IV. Section V further extends the structures described in Section III to the case where modulation doing is combined. Results of the EXIT analysis and BER simulations are presented to verify the advantageous points of the combined use of EM and modulation doping. It is shown that the EXIT analysis and BER results are exactly consistent each other even with such low rate code cases where BER threshold happens at very low SNR range. Finally Section VI concludes this paper.

II. System Model without Doping

A schematic diagram of the BICM-ID system is shown in Fig. 1.

Transmitter

The binary bit information sequence $\underline{u}$ to be transmitted is encoded by (A) a simple repetition code, (B) a single parity check code where a single parity bit is added to every d_c-1 information bits, followed by a repetition code. d_c is referred to as check node degree. The structure (C) is the same as (B), but the repetition times d_v, referred to as variable node degree, take different values in one whole block (transmission frame); if d_v takes several different values in a block, such code is referred to as having irregular degree allocation. It is assumed that throughout the paper d_c takes only one identical value as in [4].

The coded bit sequence, encoded by the encoder (A), (B), or (C), is first bit-interleaved, segmented into l_{map}–bit segments, and then the each segment is mapped on to one of the 2^m constellation points for modulation. Since $l_{map}>m$ with the EM technique, more than one labels having different bit patterns in the segment are mapped on to each constellation point. However, there are many possible combinations of bit patterns to be allocated to the constellation points. This paper uses the labels assigned to the each constellation point, obtained by [7], so that, with full *a* priori information, the mutual information (MI) between the coded bit and the demapper output extrinsic LLR at the right most point of the demapper EXIT curve is maximized. Fig. 2 shows the optimal labeling for 4-QAM with l_{map}=3 and 5. The complex-valued signal modulated according to the mapping rule, referred to as Ψ, is finally transmitted to the wireless channel.

Channel

This paper assumes frequency flat block fading additive white Gaussian noise (AWGN) channel. If the channel exhibits frequency selectivity due to the multipath propagation, the receiver needs an equalizer to eliminate the

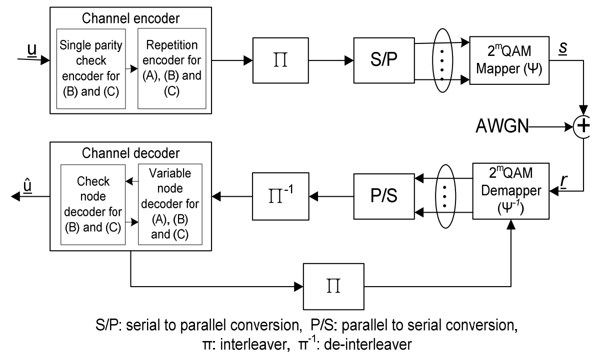


Fig. 2. System model

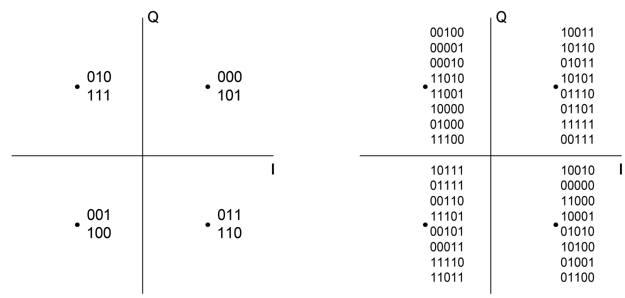


Fig. 1. Extended mapping 4-QAM l_{map}=3 and l_{map}=5, optimized with *a* priori information

inter-symbol interference. However, combining the technique presented in this paper with the turbo equalization framework [10]-[13] is rather straightforward.

It is assumed that transmission chain is properly normalized so that the received symbol energy-to-noise power spectral density ratio $E_s/N_0= 1/\sigma_n^2$; with this normalization, we can properly delete the channel complex gain term from the mathematical expression of the channel. The discrete time description of the received signal $y(k)$ is then expressed by

$$y(k)=x(k)+n(k), \tag{1}$$

where, with k being the symbol timing index, $x(k)$ is the transmitted modulated signal with unit power, and $n(k)$ the zero mean complex AWGN component with variance σ_n^2 (i.e., $<|x(k)|^2>=1$, $<n(k)>=0$, and $<|n(k)|^2>=\sigma_n^2$).

Receiver

At the receiver side, the iterative processing is invoked, where extrinsic information is exchanged between the demapper and decoder. The extrinsic LLRs calculated by soft input soft output (SISO) decoding/demapping are fed back and used for demapping/decoding as *a* priori LLR; the iteration continues until no more relevant gains in extrinsic MI can be achieved [14]; when such convergence point is reached, binary decisions are made on the information bits based on the *a* posteriori LLR at the variable node. Therefore,

see the program

the larger the MI at the convergence point, the lower the BER floor, which is depending on the matching between the encoder and the mapping rule.

The demapper calculates from the received signal point $y(k)$, corrupted by AWGN, the extrinsic LLR of the μ^{th} bit in the symbol transmitted at the k^{th} symbol timing, by

$$L_e[b_\mu(k)] = \ln \frac{\sum_{s\in S_0} \exp\left\{-\frac{|y-s|^2}{\sigma_N^2}\right\} \prod_{v=1,v\neq\mu}^{l_{map}} \exp(-b_v(s)L_a(b_v(s)))}{\sum_{s\in S_1} \exp\left\{-\frac{|y-s|^2}{\sigma_N^2}\right\} \prod_{v=1,v\neq\mu}^{l_{map}} \exp(-b_v(s)L_a(b_v(s)))}. \tag{2}$$

where s denotes a signal point in the constellation, $S_0(S_1)$ indicates the set of the labels having the μ^{th} bit being $0(1)$, and $L_a(b_v(s))$ the *a* priori LLR fed back from the decoder corresponding to the v^{th} position in the label allocated to the signal point s.

A. Repetition Coded EM BICM-ID

Fig. 3 (A) shows a block diagram of the decoder with the structure (A). The d_v bits constituting one segment, output from the de-interleaver, are connected to one variable node, where the a priori LLR to be updated is calculated by summing up the $(d_v\text{-}1)$ LLRs, as

$$L_{e,j} = \sum_{i=1,i\neq j}^{d_v} L_{a,i} \quad , \tag{3}$$

to produce the extrinsic LLR for the j^{th} bit in the segment. This process is performed for all the other bits in the same segment as well as for all the other segments independently in the frame. Finally, the updated extrinsic LLRs are interleaved, and fed back to the demapper. With the *a* priori LLRs provided in the form of the decoder's output extrinsic LLR, demapper again performs the processing of Eq. (2) to update the demapper output extrinsic LLRs. This process is repeated. Obviously, the rate of this code is $1/dv$, and the spectrum efficiency is l_{map}/d_v bits per symbol.

B. Repetition Coded EM BICM-ID with Check Nodes

Fig. 3 (B) shows a block diagram of the decoder with the structure (B). The d_v bits constituting one segment, output from the de-interleaver, are connected to a variable node, and d_c variable nodes are further segmented and connected to a check node decoder; those demapper output bits in one segment, connected to the same variable node decoder, are not overlapping with other segments, and so is the case of the variable node segmentation. Therefore, no iterations in the decoder are required. The extrinsic LLR update for a bit at the check node is exactly the same as the check node operation in the LDPC codes, as

$$L_{e,cnd,k} = \sum_{i=1,i\neq k}^{d_c} \boxplus L_{a,cnd,i} \tag{4}$$

where $\sum\boxplus$ indicates the box-sum operator [15].

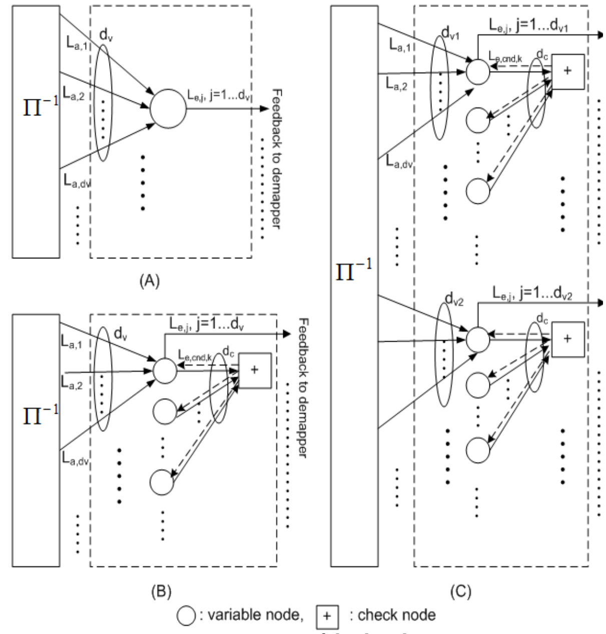


Fig. 3. Structures of the decoders

The extrinsic LLR, calculated by the check node decoder, is fed back to its connected variable node decoder, where it is further combined with $(d_v\text{-}1)$ apriori LLRs forwarded from the demapper via the de-interleaver, as

$$L_{e,j} = L_{e,cnd,k} + \sum_{i=1,i\neq j}^{d_v} L_{a,i} \quad . \tag{5}$$

This process is performed for the other bits in the same segment, and also for all the other segments independently in the same transmission block. Finally, the updated extrinsic LLRs obtained at the each variable node are interleaved, and fed back to the demapper. With the *a* priori LLRs provided in the form of decoder's output extrinsic LLR, demapper again performs the processing of (2) to update the demapper output extrinsic LLRs. This process is repeated. The rate of this code is $(d_c-1)/(d_c\cdot d_v)$, and the spectrum efficiency is $l_{map}\cdot(d_c-1)/(d_c\cdot d_v)$ bits per symbol.

C. Irregular Degree Allocations

Fig. 3 (C) shows a block diagram of the decoder with the structure (C). The segment-wise structure of (C) is exactly the same as that of (B), but the variable node degrees $d_{v,i}$ may have different values segment-by-segment. The equations for the LLR update are also the same as that for (B), but the calculations have to reflect different values $d_{v,i}$ of the variable node degrees. The rate of the code is $(d_c\text{-}1)/(d_c\cdot\sum a_i\cdot d_{v,i})$, and the spectrum efficiency is $l_{map}\cdot(d_c\text{-}1)/(d_c\cdot\sum a_i\cdot d_{v,i})$ bits per symbol, where a_i represents the ratio of variable nodes having degree $d_{v,i}$ in a block.

III. Design Based on Exit Chart

Fig. 4 shows EXIT curves for 4-QAM demappers with l_{map} as a parameter, where the labels were determined so that the EXIT curve has the largest decay. Also, the EXIT curve

see the program

with Gray mapping is presented in the figure. It is found that the Gray mapping has a completely flat EXIT curve, while anti-Gray mapping has a decay. With $l_{map}=m$, since mapping rule does not change the constellation constraint capacity (CCC), the area under the EXIT curves has to stay almost the same so far as the same modulation is used (m=*constant*). With EM, the left most point is further decreased, and the right most point increased. However, the area under the curve is smaller than without EM, because even with $l_{map}>m$, still the spectrum efficiency of the modulation itself stays the same (=m).

(A) Repetition Code

With the Gaussian assumption for the LLR distribution, the EXIT function of the repetition code is given by

$$I_{e,v}=J\left(\sqrt{(d_v-1)\cdot J^{-1}(I_{a,v})^2}\right). \tag{6}$$

where $I_{a,v}$ is the variable node input *a* priori MI and $I_{e,v}$ is its output extrinsic MI. $J()$ and $J^{-1}()$ are the functions that convert the square-root variance σ of LLR to its corresponding MI, and its inverse, respectively [2]. Obviously, (6) is corresponding to (3) for LLR update, with which $I_{a,v}= I_{e,dem}$ with $I_{e,\,dem}$ being the demapper output extrinsic MI.

(B) Repetition Code with Check Node

The check node EXIT function can be approximated by [16]

$$I_{e,cnd}=1-J\left(\sqrt{d_c-1}\cdot J^{-1}(1-I_{a,cnd})\right), \tag{7}$$

where

$$I_{a,cnd}=J\left(\sqrt{d_v\cdot J^{-1}(I_{a,dec})^2}\right). \tag{8}$$

The EXIT function of the whole decoder comprised of the variable and check node decoders can be calculated by combining Eq. (6) and (7) [4], as

$$I_{e,dec}=J\left(\sqrt{(d_v-1)\cdot J^{-1}(I_{a,dec})^2+J^{-1}(I_{e,cnd})^2}\right), \tag{9}$$

with $I_{a,dec}= I_{e,dem}$.

(C) Irregular Repetition Code with Check Node

The EXIT function of the whole decoder with the structure (C) can be obtained by weighting the segment-wise EXIT functions given by (9), by a_i corresponding to their distributions, as

$$I_{e,dec}=\frac{\sum_i a_i\cdot d_{v,i}\cdot J\left(\sqrt{(d_{v,i}-1)\cdot J^{-1}(I_{a,dec})^2+J^{-1}(I_{e,cnd})^2}\right)}{\sum_i a_i\cdot d_{v,i}}. \tag{10}$$

Hence, the shape of the combined code EXIT function can be flexibly controlled so that better matching between demapper and decoder curves can be achieved.

Fig. 5 shows EXIT functions of the decoders (A), (B), and (C) for the degree allocations presented in the box in the

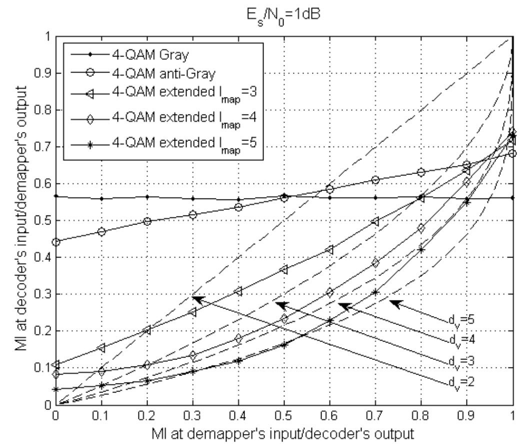


Fig. 4. EXIT chart: 4-QAM extended mapping $l_{map}=2,\ldots,5$ combined with repetition code $d_v=2,\ldots,5$

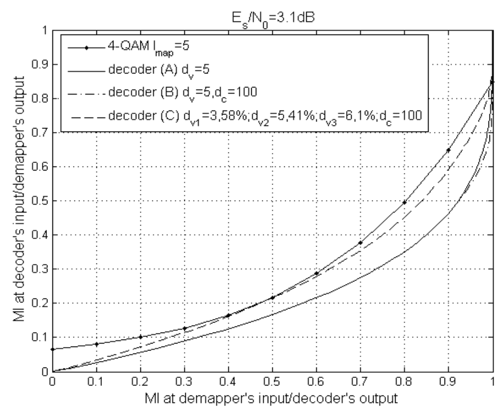


Fig. 5. EXIT chart: 4-QAM EM lmap=5 combined with 3 kinds of repetition code due to structures (A), (B) and (C)

figure, where the code rates are *1*, *0.99*, and *1.29*, respectively. It is found by comparing Figs. 4 and 5 that the EXIT functions of EM and the repetition code with check nodes can be well matched by changing the degree allocations a_i, by which we can expect sharp turbo cliff to happen at their corresponding threshold E_s/N_0, even though the structures of (A), (B), and (C) are very simple and easy to implement.

IV. Numerical Results

A series of simulations was conducted with enough number of bits transmitted to verify the advantageous characteristics of the techniques proposed. As described above, the EM technique does not change the CCC, and therefore the capability of achieving near capacity performance even with the very simple structure, which is the most significant advantage of the proposed technique, is mostly in a low E_s/N_0 value range (or equivalently low rate code case). If the CCC is much lower than the Gaussian capacity at the system's operation E_s/N_0 point, no significant merit can be expected, and obviously larger modulation format should be used in such case.

see the program

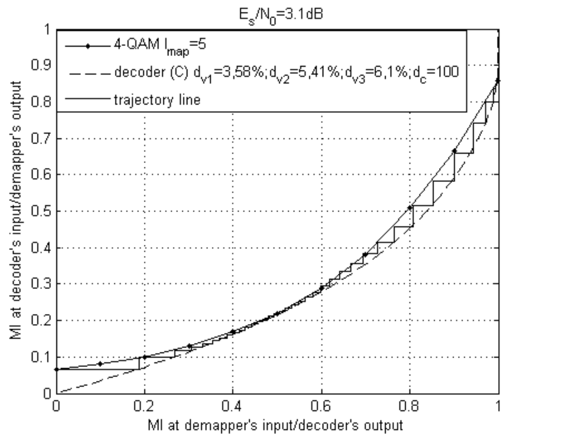


Fig. 6. EXIT chart: 4-QAM EM $l_{map}=5$ combined with repetition code $d_v=5$

With irregular degree allocation of (C), as shown by Fig. 6, the trajectory sneaks through the narrowly tiny open tunnel, and finally reaches the intersection point. With the structure (A), the adjustable parameters are only l_{map} and d_v. We therefore have introduced the check node in (B), and furthermore, irregular degree allocation in (C), both to (A). Two BER curves with $l_{map}=5$ EM 4-QAM with structure (C) are shown in Fig. 7, together with $l_{map}=5$ EM 4-QAM and $d_v=5$ to demonstrate the design flexibility. The degree allocations are shown under the figure caption. In fact, the allocation parameters for the rate *1.29* code, of which BER is indicated by "o" in the figure, are exactly the same as that shown in Fig. 5 for (C) with $E_s/N_0=3.1dB$. It is found that the EXIT and BER curves, presented in Figs. 5 and 8, are exactly consistent with each other (at $E_s/N_0=3.1dB$, EXIT tunnel opens and BER turbo cliff happens). These observations indicate that with the design flexibility made available by introducing the irregular structure, we can flexibly control the threshold E_s/N_0 and the error floor.

V. Modulation Doping

The technique described above is flexible in making good matching between the demapper and decoder EXIT curves, but is not suitable in designing BICM-ID with very low rate codes that achieve the BER pinch-off at very low SNR. This is because the LM point of the demapper EXIT function, corresponding to decoder feedback mutual information MI=0, is already very low, and there is not enough room to further lower the EXIT LM point while avoiding the intersection between the demapper and decoder curves. To solve this problem, this paper applies the idea of mixing modulation symbols having different labeling rules for the extended anti-Gray mapping and the standard Gray mapping, at a certain ratio. Since the demapper EXIT function with Gray mapping is relatively flat, its LM point has the higher MI value than EM mapping rules. Therefore, with modulation doping the demapper EXIT LM point is lifted up from that without doping, and the amount depends on the doping ratio.

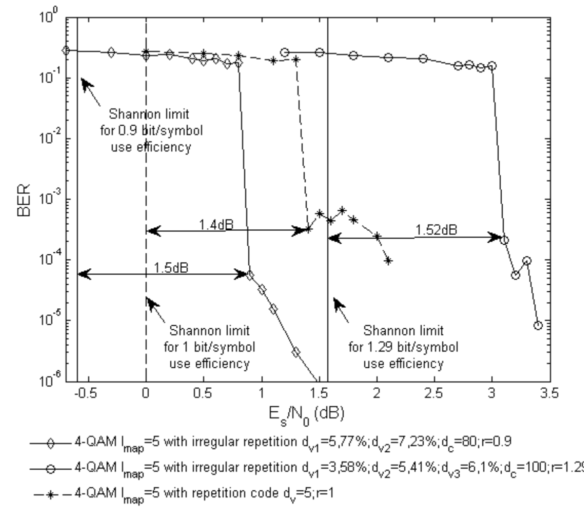


Fig. 7. BER curves optimized by irregular repetition code

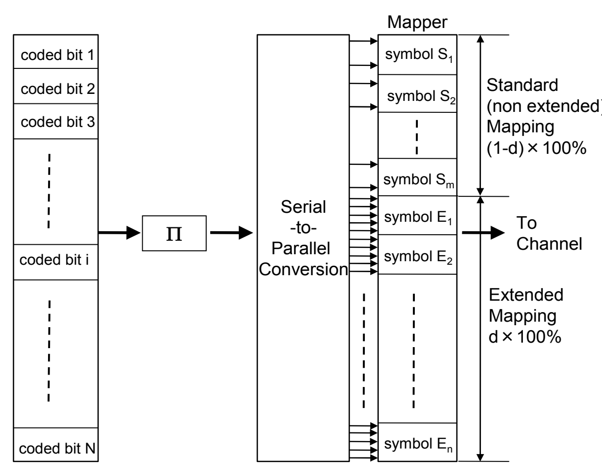


Fig. 8. Block Diagram of Modulation Doping

With EM, more than one labels having different bit patterns in the segment are mapped on to each constellation point ($l_{map}>m$), while without mapping extension, such as standard Gray mapping, only one bit pattern in the segment is mapped on to each constellation point ($l_{map}=m$). Fig. 8 shows a block diagram of the modulation doping system. The spectrum efficiency of the system then becomes $\eta = \{d \bullet m + (1-d) \bullet l_{map}\} R_c$, where d and (1-d) are the ratios of the symbols with doped (standard) mapping and extended mappings, respectively, in a transmission frame. R_c is the code rate, given by $R_c=(d_c-1)/(d_c \cdot \sum a_i \cdot d_{v,i})$, where a_i is the ratio of the variable node having degree $d_{v,i}$. It has been assumed that the interleaver is long and random enough so that the repetition encoder output bits are uniformly spread over the whole transmission frame.

Fig. 9 shows EXIT curves of some representative design examples using EM 4QAM with l_{map}=5: Their corresponding

see the program

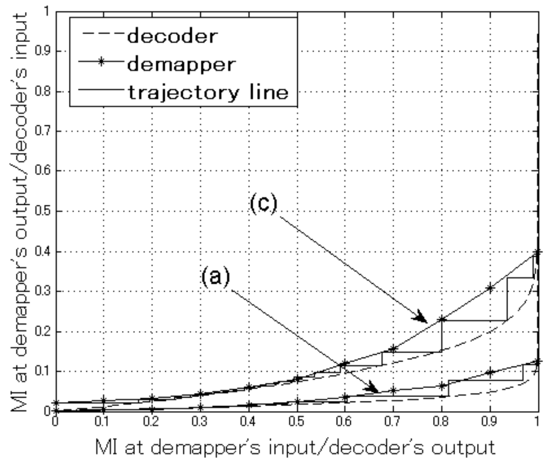


Fig. 9. EXIT chart: Low rate design case

Table 1. Degree allocations of the codes shown in Fig. 9

	E_s/N_0 (dB)	d_{v1}, a_1	d_{v2}, a_2	d_{v3}, a_3	d_c	d	bits/symbol
(a)	-9.2	30, 0.5	40, 0.5		3	0.03	0.0935
(b)	-6.3	18, 0.5	20, 0.5		5	0.04	0.2055
(c)	-3.4	9, 0.5	10, 0.05	12, 0.45	8	0.05	0.4081

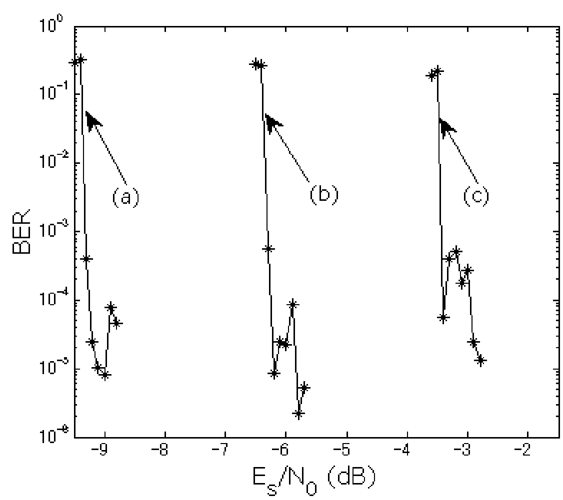


Fig. 10. BER Performance for the code design of Table 1

code design parameters and their threshold E_s/N_0 values are summarized in Table 1. A series of chain simulations for the code designs shown in Table 1 has been conducted to verify the advantageous characteristics of the techniques proposed in this paper. The results of the chain simulations conducted to obtain the trajectories for the code parameters shown in Table 1are also shown in Fig. 9. It should be emphasized that the trajectory and EXIT curve are exactly consistent each other. Fig. 10 shows the BER performance for the code design of Table 1. It is found that the BER pinch-off happens exactly at the values indicated by the EXIT chart analysis.

VI. CONCLUSIONS

In this paper, a new BICM-ID technique that makes efficient use of a mixture of standard Gray and anti-Gray extended mapping, combined with irregular degree-allocated repetition code with check nodes has been proposed. Mathematical details as well as EXIT analysis and BER simulations results have been provided,. It has been shown that with the proposed technique, very low rate code that can achieve a turbo cliff at a very low E_s/N_0 value range can well be designed, and that the EXIT chart analysis and chain simulation results are exactly consistent with each other.

REFERENCES

[1] G.Caire, G.Taricco and E.Biglieri, "Bit-Interleaved Coded Modulation", IEEE Trans. on Inform. Theory, vol.44, no.3, pp. 927-946, May 1998.

[2] J. Hagenauer, "The EXIT Chart - introduction to extrinsic information transfer in iterative processing", 12th European Signal Processing Conference (EUSIPCO), pp. 1541–1548, 2004.

[3] S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes", IEEE Trans. on Comm., vol.49, pp. 1727–1737, Oct. 2001.

[4] S.ten Brink, G.Kramer, A.Ashikmin, "Design of low-density parity-check codes for modulation and detection", IEEE Trans. on Comm., vol.52, pp. 670-678, June 2004.

[5] F. Schreckenbach, N. Görtz, J. Hagenauer, G. Bauch. "Optimized Symbol Mappings for Bit-Interleaved Coded Modulation with iterative decoding", IEEE GLOBECOM'03, vol.6, pp. 3316-3320, Dec. 2003

[6] T. Clevorn and P. Vary, "Iterative Decoding of BICM with Non-Regular Signal Constellation Sets", International ITG Conference on Source and Channel Coding, 2004.

[7] P. Henkel, "Extended Mappings for Bit-Interleaved Coded Modulation", IEEE PIMRC, pp. 1-4, 2006.

[8] P. Henkel, "Doping of Extended Mapping for Signal Shaping" In Vehicular Technology Conference, 2007. VTC2007-Spring, IEEE 65th.

[9] D. Zhao, A. Dauch, and T. Matsumoto, "BICM-ID Using Extended Mapping and Repetition Code with Irregular Node Degree Allocation", Accepted for presentation in IEEE VTC 2009 Spring, Barcelona

[10] M. Tuchler, R. Koetter, and A. Singer, "Turbo equalization: principles and new results", IEEE Trans. Comm., vol. 50, pp. 754-767, 2002a.

[11] X. Wang and H.V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA", IEEE trans. on Comm., vol. 47, pp. 1046-1061, 1999.

[12] T. Abe and T. Matsumoto, "Space-time turbo equalization in frequency-selective MIMO channels", IEEE Trans. Veh. Tech., vol.52, pp. 469-475, 2003.

[13] L. Hanzo, T.H. Liew, and B.L. Yeap, "Turbo Coding, Turbo Equalization and Space-Time Coding for Transmission over Fading Channels", Wiley-IEEE Press, 2002.

[14] S. ten Brink, "Convergence of iterative decoding", Electronics Letters, vol.35, pp.806–808, May 1999.

see the program

[15] J. Hagenauer, E. Offer, L. Papke, "Iterative decoding of binary block and convolutional codes", IEEE Trans. on Inform. Theory, vol. 42, pp. 429-445, Mar. 1996

[16] Shu L., Daniel J., Costello Jr., "Error Control Coding (2nd Edition)", Prentice Hall, June 2004.

[17] F. Schreckenbach and G. Bauch, "Bit-interleaved coded irregular modulation", European Transactions on Telecommunications, vol.17, pp. 269–282, March/April 2006.

see the program

PERFORMANCE EVALUATION OF JOINT POWER AND TIME ALLOCATIONS FOR ADAPTIVE DISTRIBUTED MIMO MULTI-HOP NETWORKS

Yidong Lang and Dirk Wübben

Department of Communications Engineering, University of Bremen, Germany
Phone: +49 421 218 7434, {lang, wuebben}@ant.uni-bremen.de

ABSTRACT

Distributed multiple input multiple output (MIMO) multi-hop networks are proven to achieve superior performance in terms of data throughput and communication reliability. In this paper a low-complexity adaptive relaying scheme is considered in order to achieve robust end-to-end (e2e) communications. For this system optimal as well as near-optimal efficient resource allocation strategies that reduce the total transmit power while satisfying a given e2e outage probability are proposed. It will be shown, that notable power savings can be achieved by the adaptive scheme compared to non-adaptive distributed MIMO multi-hop networks. Moreover, the proposed joint power and time allocation in closed form allows simple implementation and achieves near optimal power consumption.

1. INTRODUCTION

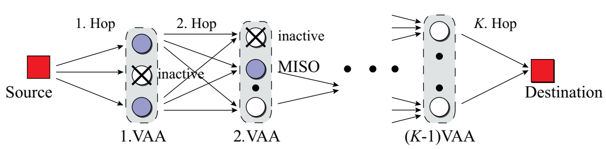


Figure 1: Topology of adaptive distributed MIMO multi-hop relaying systems.

Recently, the remarkable capacity improvement of multi-hop systems by the concept of virtual antenna arrays (VAA) was revealed [1]. The VAA allows the application of traditional MIMO techniques performed on spatially distributed relaying nodes with only one antenna, e.g., distributed space-time codes. A generic realization of a distributed MIMO multi-hop system is depicted in Fig. 1. Here, one source communicates with one destination through several VAAs in multiple hops. Spatially adjacent nodes in a VAA receive data from the previous VAA and relay data to the consecutive VAA until the destination is reached. At each relay, the decode-and-forward protocol [2] and space-time coding will be applied.

In [1], Dohler derived resource allocation strategies to maximize the e2e throughput with respect to the ergodic capacity. In contrast, we consider e2e outage probability due to its higher practical relevance to real wireless communications [3–5]. However, these investigations are based on the assumption that the e2e communication is determined by the weakest link in the network in terms of either ergodic capacity or outage probability. This strong assumption degrades the e2e performance drastically. In order to fully exploit the potential of VAA assisted multi-hop networks, a simple adaptive scheme for distributed MIMO multi-hop networks is introduced here, where one relay stops sending the data when it is in outage and other nodes from the same virtual antenna array (VAA) adapt to a new space-time code. Furthermore, the performance of optimal as well as near-optimal joint power and time allocation strategies will be investigated.

The remainder of the paper is organized as follows. In Section 2 the system model of the adaptive transmission scheme is introduced. The mathematical description of the outage probability will be given in Section 3 and the joint power and time allocation is formulated as a convex optimization problem in Section 4. In order to achieve an efficient solution, a simplified problem is formulated and the corresponding closed-form solution for this approximated optimization problem is derived in Section 5. Finally, the performance is investigated in Section 6 and conclusions will be given in Section 7.

2. SYSTEM DESCRIPTION

We consider a K-hop network with t_k transmit nodes and r_k receive nodes at hop k. Several relays are grouped to a VAA at each hop to apply a distributed space-time code. The data is then transmitted from the source to the destination through $K-1$ VAAs. It is assumed that no interference between the hop occur. Thus, the bandwidth or time has to be divided into non-overlapping parts for each hop such that at any time they are occupied by only one hop, i.e., FDMA or TDMA respectively. Without loss of generality, the TDMA based adaptive scheme will be considered here.

At the first time fraction α_1 the source transmits data to the relays of the 1. VAA. The nodes of the 1. VAA de-

see the program

code the received signals separately in order to avoid enormous information exchanges, i.e., *separately decoding* is performed, which decomposes this hop to several SISO links (or MISO links at the next hops). The t'_k relays successfully decoding the message (or being not in outage) are denoted as *active nodes* and the others failing to decode the message (or being in outage) are denoted as *inactive nodes*, respectively. The inactive nodes will stop transmission at the next time fraction. The t'_k active nodes will *adapt* to transmit the decoded message *cooperatively* according to a space-time code with respect to t'_k, i.e., $0 \le t'_k \le t_k$. To this end, each node transmits a spatial fraction of a space-time code. If all relays within one VAA fail to decode the message, the e2e connection is considered to be in outage, denoted by the probability P_{e2e}. Otherwise, the t'_k active nodes send the data to the next VAA at the next time fraction α_{k+1}. This adaptive transmission continues at each VAA until the message reaches the destination. Note that a given fixed network topology is assumed and the task of grouping the VAAs is beyond the scope of this paper.

It is assumed that each relay transmits signals with the same data rate R but with individual time fraction α_k, of which $\sum_{k=1}^{K} \alpha_k = 1$ holds. All the hops use the total bandwidth W that is available to the network. We define $\mathbf{S}_k \in \mathbb{C}^{t'_k \times L_k}$ as the space-time encoded signal with length L_k from the t'_k active nodes at the kth hop, i.e., $0 \le t'_k \le t_k$. The received signal $\mathbf{y}_{k,j} \in \mathbb{C}^{1 \times L_k}$ at the jth node at the kth VAA is given by

$$\mathbf{y}_{k,j} = \sqrt{\theta_k \mathcal{P}_k / t_k}\,\mathbf{h}_{k,j} \mathbf{S}_k + \mathbf{n}_{k,j} \ , \tag{1}$$

where $\mathbf{n}_{k,j} \sim \mathcal{N}_C(0,\, N_0) \in \mathbb{C}^{1 \times L_k}$ is the Gaussian noise with power spectral density N_0. Each active node from one VAA transmits data with power $\mathcal{P}_k / t_k$ equally. This permits simple power control and hardware implementation at each relay which is especially important for relaying nodes with minimal processing functionality. The channel from the t'_k active nodes to the jth receive node within the kth hop is denoted as $\mathbf{h}_{k,j} \in \mathbb{C}^{1 \times t'_k}$, whose elements obey the same uncorrelated Rayleigh fading statistics with unit variance. Note that the relaying nodes belonging to the same VAA are assumed be spatially sufficiently close as to justify a common path loss θ_k between two VAAs, which is known as *symmetric network*. The path loss is simply described as $\theta_k = d_k^{-\epsilon}$, where d_k is the distance between two nodes and ϵ is the path loss exponent within range of 2 to 5 for most wireless channels.

In order to meet a given Quality-of-Services (QoS) requirement, the transmit power $\mathcal{P}_k$ and the time fraction α_k per hop need to be optimized. In the next section the e2e outage probability is introduced as the QoS parameter and optimum as well as near-optimum solutions to the occurring optimization problem are proposed subsequently.

3. END-TO-END OUTAGE PROBABILITY

In order to describe the outage behavior of the adaptive multi-hop network the varying number of active nodes t'_k per hop k has to be considered. To this end, the probability $p_{\text{out},k,j}(t'_k)$ that a *given* MISO link with t'_k active nodes leads to an outage in the jth node in the kth hop is calculated. In combination with the probability $\Pr(t'_k)$ of t'_k active nodes in hop k the effective outage probability $P_{\text{out},k,j}$ of node j in hop k is derived.

The instantaneous achievable rate from t'_k active nodes to the receive node j at hop k is given by

$$C_{k,j}(t'_k) = \alpha_k W \log\left(1 + \frac{\mathcal{P}_k}{\alpha_k W t_k d_k^{\epsilon} N_0} \|\mathbf{h}_{k,j}\|^2\right) \ , \tag{2}$$

with $\|\mathbf{h}_{k,j}\|^2 = \sum_{i=1}^{t'_k} |h_{k,j,i}|^2$. The outage probability $p_{\text{out},k,j}(t'_k)$ is then given by the probability that the channel can not support an error-free transmission at rate R

$$\begin{aligned} p_{\text{out},k,j}(t'_k) &= \Pr(R > C_{k,j}(t'_k)) \\ &= \Pr\left(\|\mathbf{h}_{k,j}\|^2 < \frac{\left(2^{\frac{R}{\alpha_k W}} - 1\right) \alpha_k W N_0 d_k^{\epsilon} t_k}{\mathcal{P}_k}\right) . \end{aligned} \tag{3}$$

For simplicity, the approximation $\log(1+x) \approx \sqrt{x}$ is applied to the achievable rate in (2) as assessed in [1]. Thus, (2) can be simplified by

$$C_{k,j}(t'_k) \approx \sqrt{\frac{\alpha_k W \mathcal{P}_k}{d_k^{\epsilon} N_0 t_k} \|\mathbf{h}_{k,j}\|^2} \tag{4}$$

and the outage probability (3) becomes

$$\begin{aligned} p_{\text{out},k,j}(t'_k) &\approx \Pr\left(\|\mathbf{h}_{k,j}\|^2 < \frac{R^2 N_0 d_k^{\epsilon} t_k}{\alpha_k W \mathcal{P}_k}\right) \\ &= \Pr\left(\|\mathbf{h}_{k,j}\|^2 < x_k\right) \ . \end{aligned} \tag{5}$$

For ease notation the variables $x_k = Q_k / (\alpha_k \mathcal{P}_k)$ and $Q_k = R^2 N_0 d_k^{\epsilon} t_k / W$ were introduced, where the parameter x_k is proportional to the inverse signal-to-noise ratio, i.e., $x_k \sim 1/\text{SNR}_k$. In (5), $\|\mathbf{h}_{k,j}\|^2$ obeys a Gamma distribution [6], whose CDF can be described by the lower incomplete Gamma function $\gamma(t'_k, x_k) = \int_0^{x_k} e^{-u} u^{t'_k - 1}\, \mathrm{d}u$ normalized by Gamma function $\Gamma(t'_k)$, i.e.,

$$p_{\text{out},k,j}(t'_k) \approx \frac{\gamma(t'_k, x_k)}{\Gamma(t'_k)} \ . \tag{6}$$

In order to determine the outage probability $P_{\text{out},k,j}$ the probability of active nodes has to be considered. This depends itself on the success of decoding in the previous hop, which is given by the outage probability $P_{\text{out},k-1,j}$. Furthermore, the outage probabilities of all nodes within one VAA are

see the program

equal under the assumption of symmetric networks, i.e., $P_{\text{out},k,1} = \cdots = P_{\text{out},k,r_k} = P_{\text{out},k,j'}$ where j' indexes an arbitrary $j \in [1, \cdots, r_k]$. Thus, the number of active relaying nodes t'_k follows the binomial distribution $\mathcal{B}$ with parameters t_k and $P_{\text{out},k-1,j'}$ [6], i.e.,

$$t'_k \sim \mathcal{B}(t_k, 1 - P_{\text{out},k-1,j'}) \ . \tag{7}$$

More general, the probability of i nodes being active at hop k is given by the probability mass function as

$$\Pr(t'_k = i) = \binom{t_k}{i}(1 - P_{\text{out},k-1,j'})^i P^{t_k - i}_{\text{out},k-1,j'} \ , \ \forall\, i \tag{8}$$

where $\binom{t_k}{i} = \frac{t_k!}{i!(t_k-i)!}$ is the number of combinations of i active nodes over t_k. Hence, $\Pr(t'_k = i) \cdot p_{\text{out},k,j}(i)$ describes the probability that i nodes are active and lead to an outage event. The outage probability $P_{\text{out},k,j'}$ of node j' in the hop k is given by the sum of the outage probabilities over all possible i, namely,

$$\begin{aligned} P_{\text{out},k,j'} &= \sum_{i=1}^{t_k} \Pr(t'_k = i) \cdot p_{\text{out},k,j}(i) \qquad (9) \\ &= \sum_{i=1}^{t_k} \binom{t_k}{i}(1 - P_{\text{out},k-1,j'})^i P^{t_k-i}_{\text{out},k-1,j'} \frac{\gamma(i, x_k)}{\Gamma(i)} \ . \end{aligned}$$

Clearly, an outage event occurs in one hop if all nodes of one VAA cannot decode the message successfully. Thus, the outage probability of hop k is given by the product of $P_{\text{out},k,j}$, $1 \le j \le r_k$,

$$P_{\text{out},k} = \prod_{j=1}^{r_k} P_{\text{out},k,j} = P^{r_k}_{\text{out},k,j'} \ . \tag{10}$$

Consequently the e2e connection is in outage if any hop is broken and the e2e outage probability corresponds to

$$P_{\text{e2e}} = 1 - \prod_{k=1}^{K}(1 - P_{\text{out},k}) = 1 - \prod_{k=1}^{K}\left(1 - P^{r_k}_{\text{out},k,j'}\right) \ . \tag{11}$$

4. OPTIMUM JOINT POWER AND TIME ALLOCATION (JPTA)

In order to minimize the total power consumption $\mathcal{P}_{\text{total}}$ while satisfying a given e2e outage probability constraint e, the optimization problem for joint power and time allocation is formulated as

$$\text{minimize } \mathcal{P}_{\text{total}} = \sum_{k=1}^{K} \alpha_k \mathcal{P}_k (1 - P_{\text{out},k-1,j'}) \tag{12a}$$

$$\text{subject to } P_{\text{e2e}} \le e, \tag{12b}$$

$$\sum_{k=1}^{K} \alpha_k = 1 \ . \tag{12c}$$

To calculate $\mathcal{P}_{\text{total}}$, the inactive nodes stopping the transmission to save power with the probability $P_{\text{out},k-1,j'}$ is taken into account. Moreover, the duration of each time fraction α_k is also considered. The optimization problem (12) can be shown to be convex for low outage probability requirements by proving the Hessian matrix of $P_{\text{e2e}}(\mathcal{P}_k, \alpha_k, \forall\, k)$ to be positive semi-definite. By using standard optimization methods [7], the optimal solution $\mathcal{P}^\star_k$, $\alpha^\star_k$ for (12) can be achieved leading to considerable complexity. An approximated problem for joint power and time allocation is investigated in the following text.

5. APPROXIMATE JOINT POWER AND TIME ALLOCATION

5.1. Problem Simplification

Following the approximation method given in [3,8], the outage probability $p_{\text{out},k,j}(t'_k)$ in (6) is approximated for high SNR as

$$p_{\text{out},k,j}(t'_k) = \frac{\gamma(t'_k, x_k)}{\Gamma(t'_k)} \lessapprox \frac{t'^{-1}_k x^{t'_k}_k}{\Gamma(t'_k)} = \frac{x^{t'_k}_k}{\Gamma(t'_k + 1)} \ . \tag{13}$$

Hence the outage probability per node (9) and the outage probability per hop (10) are approximated by $\tilde{P}_{\text{out},k,j'}$ and $\tilde{P}_{\text{out},k,j'}$, respectively,

$$P_{\text{out},k,j'} \lessapprox \sum_{i=1}^{t_k} \Pr(t'_k = i) \frac{x^i_k}{\Gamma(i+1)} \triangleq \tilde{P}_{\text{out},k,j'} \tag{14}$$

$$\tilde{P}_{\text{out},k} = \prod_{j=1}^{r_k} \tilde{P}_{\text{out},k,j} = \tilde{P}^{r_k}_{\text{out},k,j'} \ . \tag{15}$$

Furthermore, the e2e outage probability (11) is union bounded by [3]

$$P_{\text{e2e}} \le \sum_{k=1}^{K} P_{\text{out},k} = \sum_{k=1}^{K} P^{r_k}_{\text{out},k,j'} \le \sum_{k=1}^{K} \tilde{P}^{r_k}_{\text{out},k,j'} \triangleq \tilde{P}_{\text{e2e}} \ . \tag{16}$$

Finally, for small $P_{\text{out},k-1,j'}$ the objective function of (12) can be relaxed to $\mathcal{P}_{\text{total}} \approx \sum_{k=1}^{K} \alpha_k \mathcal{P}_k$. Thus, the approximated optimization problem is obtained

$$\text{minimize } \mathcal{P}_{\text{total}} \approx \sum_{k=1}^{K} \alpha_k \mathcal{P}_k \tag{17a}$$

$$\text{subject to } \tilde{P}_{\text{e2e}} = \sum_{k=1}^{K} \tilde{P}^{r_k}_{\text{out},k,j'} \le e, \tag{17b}$$

$$\sum_{k=1}^{K} \alpha_k = 1 \ . \tag{17c}$$

By neglecting the time fraction constraint $\sum_{k=1}^{K} \alpha_k = 1$ in (17), the optimization problem depends only on the product

see the program

$\alpha_k \mathcal{P}_k$, which is therefore approximately symmetric with respect to α_k and $\mathcal{P}_k$. In other word, one of the optimal power allocations $\mathcal{P}_k^\star$ is proportional to the optimal time fraction $\alpha_k^\star$, i.e., $\mathcal{P}_k^\star \sim \alpha_k^\star$. With consideration of the constraint $\sum_{k=1}^K \alpha_k = 1$, the relation between the optimal power and time fraction can be achieved

$$\alpha_k^\star = \frac{\mathcal{P}_k^\star}{\sum_{k=1}^K \mathcal{P}_k^\star} . \tag{18}$$

By introducing the auxiliary variable $\beta_k = \alpha_k \mathcal{P}_k$ the optimization problem (17) is relaxed to

$$\text{minimize } \mathcal{P}_{\text{total}} \approx \sum_{k=1}^{K} \beta_k \tag{19a}$$

$$\text{subject to } \tilde{P}_{\text{e2e}} \leq e . \tag{19b}$$

5.2. Closed-Form Solution (JPTA-CF)

Similar to the solution introduced in [9], (19) can be used to derive approximative solution by the means of Lagrangian

$$L(\beta_k, \lambda) = \sum_{k=1}^{K} \beta_k + \lambda(\tilde{P}_{\text{e2e}} - e) . \tag{20}$$

To obtain the near optimal solution, the derivatives of $L(\beta_k, \lambda)$ with respect to β_k has to be zero for all $1 \leq k \leq K$, i.e.,

$$\frac{\partial L(\beta_k, \lambda)}{\partial \beta_k} = 0, \quad \forall\, k . \tag{21}$$

Furthermore, the equality of the constraint function in (19) must be fulfilled,

$$\tilde{P}_{\text{e2e}} = \sum_{k=1}^{K} \tilde{P}_{\text{out},k} = e . \tag{22}$$

By several further approximations as outlined in the Appendix, a closed-form solution for β_k can be achieved. Furthermore, according to (18), β_k is given by

$$\beta_k = \alpha_k^\star \mathcal{P}_k^\star = \frac{\mathcal{P}_k^{\star 2}}{\sum_{k=1}^K \mathcal{P}_k^\star} . \tag{23}$$

Rewriting this form we achieve

$$\mathcal{P}_k^\star = \sqrt{\sum_{k=1}^{K} \mathcal{P}_k^\star \sqrt{\beta_k}} , \tag{24}$$

$$\sum_{k=1}^{K} \mathcal{P}_k^\star = \sqrt{\sum_{k=1}^{K} \mathcal{P}_k^\star \sum_{k=1}^{K} \sqrt{\beta_k}} , \tag{25}$$

$$\sum_{k=1}^{K} \mathcal{P}_k^\star = \left(\sum_{k=1}^{K} \sqrt{\beta_k} \right)^2 . \tag{26}$$

Inserting (26) to (23), the following theorem is obtained.

Theorem 1. [Joint Power and Time Allocation in Closed Form (JPTA-CF)] *The joint power and time (or bandwidth) allocation for outage restricted adaptive distributed MIMO multi-hop networks in closed form is given by*

$$\mathcal{P}_k^\star = \sqrt{\beta_k} \sum_{k=1}^{K} \sqrt{\beta_k} \quad \textit{and} \quad \alpha_k^\star = \frac{\sqrt{\beta_k}}{\sum_{k=1}^K \sqrt{\beta_k}} , \tag{27}$$

with outage probability per hop $\tilde{P}_{out,k} = \tilde{P}_{out,k,j'}^{r_k} \approx \frac{\delta_k \cdot e}{\sum_{k=1}^K \delta_k}$, *where the parameters* δ_k *and* β_k *are given by*

$$\delta_k = \frac{\left(2 t_k^{\frac{2}{t_k+1}} \left(\prod_{i=1}^{t_k} \frac{\binom{t_k}{i} (1 - e^{\frac{1}{r_{k-1}}})^i e^{\frac{t_k - i}{r_{k-1}}} Q_k^i}{\Gamma(i+1)} \right)^{\frac{2}{t_k+1}} \right)^{\frac{2+(t_k+1) r_k}{(t_k+1) r_k}}}{(r_k(t_k+1))^{\frac{2+(t_k+1) r_k}{(t_k+1) r_k}}} ,$$

$$\beta_k = \frac{t_k^{\frac{2}{t_k+1}}}{\tilde{P}_{out,k}^{\frac{2}{r_k(t_k+1)}}} \left(\prod_{i=1}^{t_k} \frac{\binom{t_k}{i} (1 - \tilde{P}_{out,k-1}^{\frac{1}{r_{k-1}}})^i \tilde{P}_{out,k-1}^{\frac{t_k - i}{r_{k-1}}} Q_k^i}{\Gamma(i+1)} \right)^{\frac{2}{t_k(t_k+1)}} .$$

6. PERFORMANCE EVALUATION

The performance of joint power and time allocation for adaptive distributed MIMO multi-hop schemes is evaluated here for various network configurations. It is assumed that the e2e communication over $W = 5$ MHz should meet an e2e outage probability constraint of $e = 1\%$ with the path loss exponent $\epsilon = 3$ and $N_0 = -174$ dBm/Hz.

6.1. Total power: Adaptive v.s. non-adaptive.

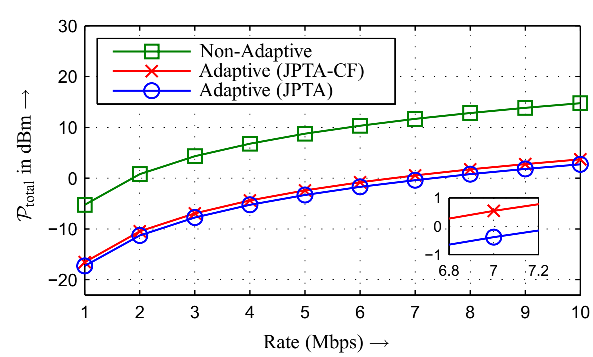


Figure 2: $\mathcal{P}_{\text{total}}$ in dBm for non-adaptive transmission, closed-form and optimal resource allocation solution.

Fig. 2 shows the total power versus the data rate R for non-adaptive and adaptive transmissions both with optimized resource allocations for a 3-hop system with $t_k = 3$ nodes per VAA. The distance between the VAAs is $d_k =$

see the program

$[1, 1, 1]$km. Note that the optimal solution *JPTA* is solved by means of standard optimization methods [7] and the closed-form solution *JPTA-CF* is given by Theorem 1. It can be observed that the proposed closed-form solution yields near-optimum total power consumption and almost 15 dBm gain in comparison with the non-adaptive scheme, where the e2e connection is in outage if any node in the network is in outage [3, 5].

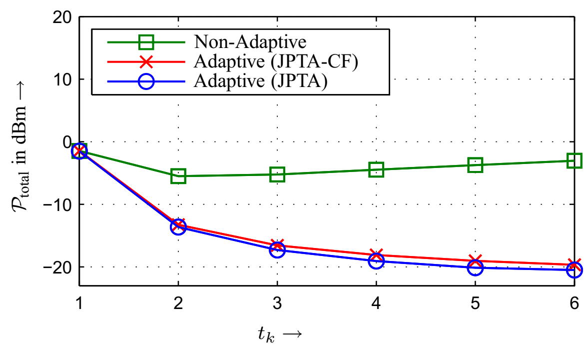


Figure 3: $\mathcal{P}_{\text{total}}$ in dBm versus different number of nodes per VAA from 1 to 6 for non-adaptive transmission, closed-form and optimal resource allocation solution, $R = 1$Mbps.

Fig. 3 shows the total power consumption versus number of nodes per VAA for non-adaptive and adaptive schemes at data rate $R = 1$Mbps. The number of nodes per VAA t_k varies from 1 to 6. It is observed that the performance gap between non-adaptive and adaptive scheme grows for increasing number of nodes t_k. The optimal number of t_k for non-adaptive scheme for this case is 2. Moreover, along with the increasing t_k, the power consumption of the proposed closed-form solution is slightly increased. However, it still achieves near-optimal performance with significantly reduced complexity.

6.2. Total power: Joint power and time allocation (*JPTA*) v.s. only power allocation ($\alpha_k = 1/K$) (*PA*).

In order to reveal the benefits of joint power and time allocation, the optimal power allocation with equal time for each hop is considered for comparison [9], i.e., $\alpha_k = 1/K, \forall k$ which is denoted as *PA* for brevity. Fig. 4 depicts the total transmission power of *PA* and *JPTA* schemes versus data rate. The network configuration is the same as in Section 6.1. *JPTA* improves the communication over 5 dBm in comparison with *PA* for this case.

We consider the impact of different distance d_k on the total power of *JPTA* and *PA* at data rate $R = 1$Mbps. A 3-hop system with 2 VAAs each containing 2 nodes is assumed. At first, the distances $d_2 = d_3 = 1$km are constant, where the distance d_1 from the source to the first VAA varies

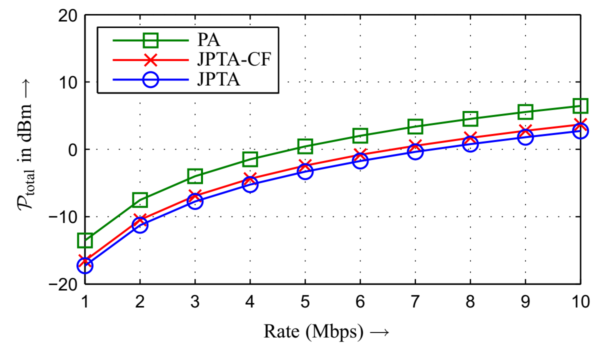


Figure 4: $\mathcal{P}_{\text{total}}$ for only power optimized (equal time $\alpha_k = 1/K, \forall k$) and joint power and time allocation for adaptive scheme.

from 0.5km to 5km. The result is shown in Fig. 5(a). In contrast, for the second case, d_3 varies from 0.5km to 5km and d_1, d_2 are constantly 1km, shown in Fig. 5(b). Clearly, if the distance is increased, the total power of each approach is also enlarged due to the path loss. For the first case, i.e., increasing the distance between the source and the first VAA, the gains due to joint power and time allocation is vanished slowly. This can be explained that in comparison with only power optimization the additional time allocation can not compensate the increased path loss at the first hop, which only has diversity degree 1. In Fig. 5(b) we observe an diametrically opposed phenomenon. When the distance d_3 is increased, the performance gap due to *JPTA* is also increased. Here *JPTA* dominates *PA* by flexible time allocation.

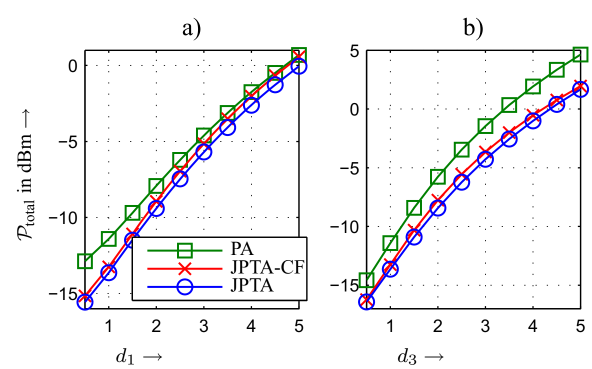


Figure 5: Impact of distance on $\mathcal{P}_{\text{total}}$ in dBm for only power optimized ($\alpha_k = 1/K, \forall k$) and joint power and time allocation for adaptive scheme. a) d_1 varies from 0.5km to 5km; b) d_3 varies from 0.5km to 5km.

see the program

6.3. Power of each hop: Adaptive v.s. non-adaptive.

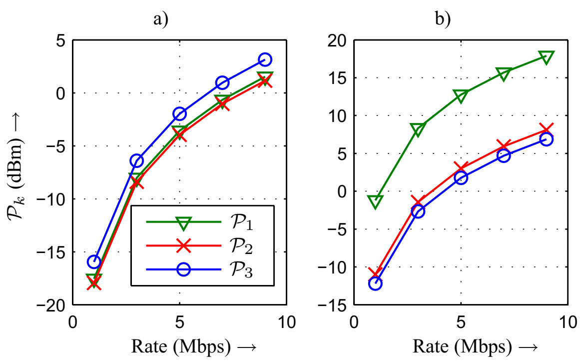


Figure 6: Power $\mathcal{P}_k$ per hop for a) *JPTA* and b) non-adaptive scheme.

Fig. 6 depicts the power $\mathcal{P}_k$ of each hop by *JPTA* and non-adaptive scheme. The network configuration is the same as in Fig. 2, i.e., $t_k = [1, 3, 3, 1]$ and $d_k = [1, 1, 1]$km. It is observed that the power consumed at the third hop is the largest for *JPTA*. The reason for this is that there is only one node at the destination which has to decode the data correctly otherwise an outage event occurs. Similar at the source, there are no nodes to transmit the data cooperatively with high diversity degrees. The source consumes the second most power. Due to the adaptive scheme and space-time coding at the second hop, the second hop uses the least power. In contrast, for non-adaptive scheme the first hop uses the most power, which is due to the lack of diversity degrees at the first hop.

6.4. Time fraction of each hop: *JPTA* v.s. *JPTA-CF*.

As mentioned before, the optimization problem (12) depends only on the product of the variables $\mathcal{P}_k$ and α_k. It can be proven that this problem is not strictly convex with respect to $\mathcal{P}_k$ and α_k [7], i.e., the optimal solution is not unique. There are many optimal combinations of $\mathcal{P}_k$ and α_k to satisfying the optimization problem. This is verified by our simulation results shown in Fig. 7(b)(c). We choose two different starting points for the exact optimization problem (12) which are denoted as *JPTA*$_1$ and *JPTA*$_2$. Different time fractions results can be obtained which do not lead to any increased total power consumption, as the same as that shown in Fig. 2. In other words, these results are also optimal. Fig. 7(a) shows the time fraction derived by *JPTA-CF*, which is near to the optimal solution *JPTA*$_1$.

Fig. 8 shows the relative power allocation per hop $\mathcal{P}'_k = \frac{\mathcal{P}_k}{\sum_{k=1}^{K} \mathcal{P}_k}$ versus data rate for *JPTA* and *JPTA-CF*. Comparing Fig. 8(a)(b) with Fig. 7(a)(b), they are exact the same which verifies the the relation (18), i.e., $\alpha_k^\star = \frac{\mathcal{P}_k^\star}{\sum_{k=1}^{K} \mathcal{P}_k^\star}$.

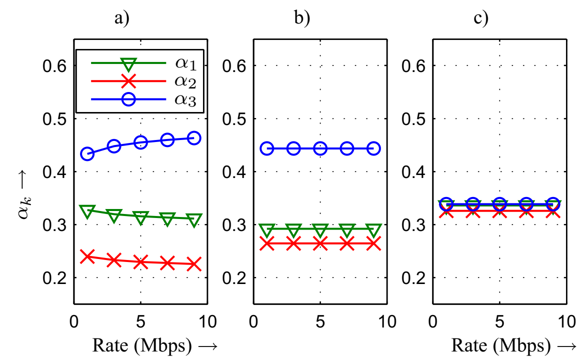


Figure 7: The time fraction α_k per hop for a) *JPTA-CF*, b) *JPTA*$_1$ and c) *JPTA*$_2$.

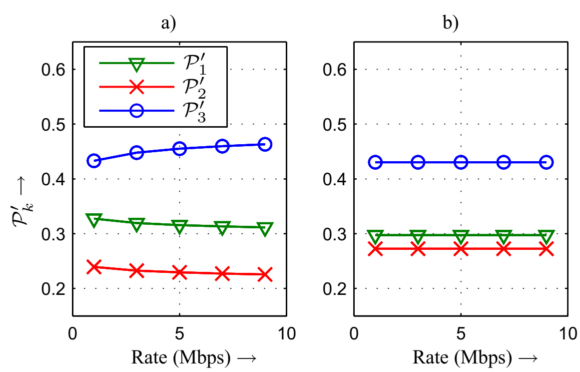


Figure 8: The time fraction α_k per hop for a) *JPTA-CF*, b) *JPTA*.

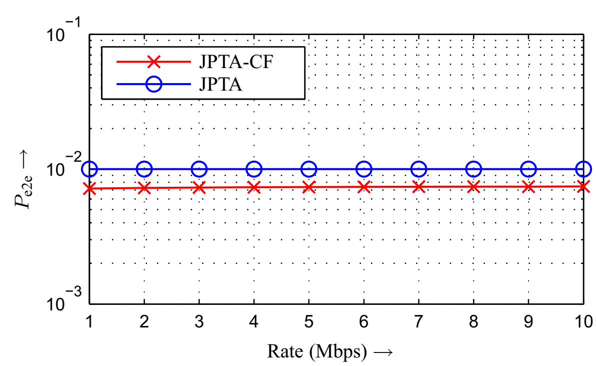


Figure 9: The e2e outage probability with *JPTA-CF* and *JPTA*.

6.5. Outage probability: *JPTA* v.s. *JPTA-CF*.

Fig. 9 shows the e2e outage probability P_{e2e} versus data rate by the proposed *JPTA* and *JPTA-CF* solutions. It is observed that both e2e outage probabilities are independent

see the program

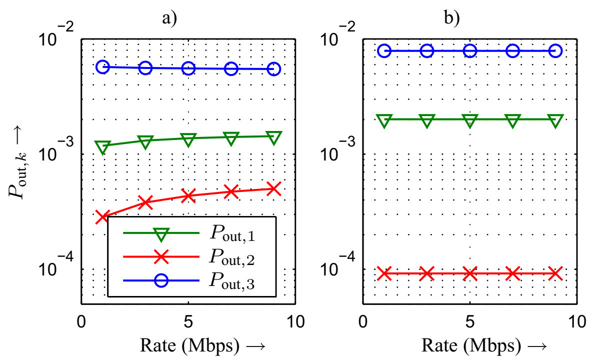


Figure 10: The outage probability per hop with a) *JPTA-CF*, b) *JPTA*.

from the data rate. Moreover, the *JPTA-CF* solution leads to a lower outage level as required which results in slightly higher power consumption as dedicated in Fig. 2. Fig. 10 depicts the outage probability of each hop $P_{\text{out},k}$ versus data rate. The most outage events happen at the third hop where the most power is used as shown in Fig. 8.

7. CONCLUSION

In this paper, we proposed optimal as well as near-optimal joint power and time (or bandwidth) allocation for adaptive distributed MIMO multi-hop networks, which is required to support an given outage probability level with minimized total transmission power. As shown in simulation results, the adaptive scheme outperforms the non-adaptive scheme significantly. Furthermore, the closed form of joint power and time allocation *JPTA-CF* achieves near-optimal performance with lower complexity comparing to the optimal solution *JPTA*.

8. APPENDIX

[Proof of Theorem 1] The first derivative of $L(\beta_k, \lambda)$ in (21) with respect to β_k relates to $\tilde{P}_{\text{out},k}$ as well as $\tilde{P}_{\text{out},k+1}$, expressed as follows

$$\frac{\partial L(\beta_k,\lambda)}{\partial \beta_k} = 1 + \lambda \left(\frac{\partial \tilde{P}_{\text{out},k}}{\partial \beta_k} + \frac{\partial \tilde{P}_{\text{out},k+1}}{\partial \beta_k} \right) = 0 \,, \quad (28)$$

which is due to the dependence between $\tilde{P}_{\text{out},k}$ and $\tilde{P}_{\text{out},k+1}$ indicted in (9). This makes the further analysis involved. To simplify the analysis, we replace $\tilde{P}_{\text{out},k-1,j'}$ by $e^{\frac{1}{r_{k-1}}}$ which is motivated by the fact that $\tilde{P}_{\text{out},k} < e, \forall k$. Thus, (14) becomes

$$\tilde{P}_{\text{out},k,j'} \approx \sum_{i=1}^{t_k} \binom{t_k}{i} \left(1 - e^{\frac{1}{r_{k-1}}}\right)^i e^{\frac{t_k-i}{r_{k-1}}} \frac{x_k^i}{\Gamma(i+1)} \,. \quad (29)$$

Furthermore, (29) can be approximated by its geometric mean

$$\tilde{P}_{\text{out},k,j'} \approx t_k \left(\prod_{i=1}^{t_k} \frac{\binom{t_k}{i}(1-e^{\frac{1}{r_{k-1}}})^i e^{\frac{t_k-i}{r_{k-1}}} x_k^i}{\Gamma(i+1)} \right)^{\frac{1}{t_k}} \quad (30)$$

$$= x_k^{\frac{t_k+1}{2}} t_k \left(\prod_{i=1}^{t_k} \frac{\binom{t_k}{i}(1-e^{\frac{1}{r_{k-1}}})^i e^{\frac{t_k-i}{r_{k-1}}}}{\Gamma(i+1)} \right)^{\frac{1}{t_k}}$$

$$= \left(\frac{Q_k}{\beta_k}\right)^{\frac{t_k+1}{2}} t_k \left(\prod_{i=1}^{t_k} \frac{\binom{t_k}{i}(1-e^{\frac{1}{r_{k-1}}})^i e^{\frac{t_k-i}{r_{k-1}}}}{\Gamma(i+1)} \right)^{\frac{1}{t_k}}$$

Hence, β_k can be expressed by $\tilde{P}_{\text{out},k,j'}$ as

$$\beta_k = \frac{t_k^{\frac{2}{t_k+1}}}{\tilde{P}_{\text{out},k,j'}^{\frac{2}{t_k+1}}} \left(\prod_{i=1}^{t_k} \frac{\binom{t_k}{i}(1-e^{\frac{1}{r_{k-1}}})^i e^{\frac{t_k-i}{r_{k-1}}} Q_k^i}{\Gamma(i+1)} \right)^{\frac{2}{t_k(t_k+1)}} . \quad (31)$$

As the dependence between $\tilde{P}_{\text{out},k}$ and $\tilde{P}_{\text{out},k+1}$ has been removed, equation (28) simplifies to

$$\frac{\partial L(\beta_k,\lambda)}{\partial \beta_k} = 1 + \lambda \frac{\partial \tilde{P}_{\text{out},k}}{\partial \beta_k} = 0 \,. \quad (32)$$

Differentiating (30) along β_k yields

$$0 = 1 + \lambda r_k \tilde{P}_{\text{out},k}^{\frac{r_k-1}{r_k}} \frac{\partial \tilde{P}_{\text{out},k,j'}}{\partial \beta_k} \quad (33a)$$

$$= 1 - \frac{\lambda r_k (t_k+1) \tilde{P}_{\text{out},k}^{\frac{r_k-1}{r_k}}}{2\beta_k} \tilde{P}_{\text{out},k,j'} \quad (33b)$$

$$= 1 - \frac{\lambda r_k (t_k+1) \tilde{P}_{\text{out},k}^{\frac{r_k-1}{r_k}}}{2\beta_k} \tilde{P}_{\text{out},k}^{\frac{1}{r_k}} \quad (33c)$$

$$= 1 - \frac{\lambda r_k (t_k+1) \tilde{P}_{\text{out},k}}{2\beta_k} \,. \quad (33d)$$

Inserting (31) in (33d), $\tilde{P}_{\text{out},k}$ is given by

$$\tilde{P}_{\text{out},k} = \tilde{P}_{\text{out},k,j'}^{r_k} = \lambda^{-\frac{2+(t_k+1)r_k}{(t_k+1)r_k}} \cdot \delta_k \,, \quad (34)$$

where δ_k is introduced to simply the notation

$$\delta_k = \frac{\left(2 t_k^{\frac{2}{t_k+1}} \left(\prod_{i=1}^{t_k} \frac{\binom{t_k}{i}(1-e^{\frac{1}{r_{k-1}}})^i e^{\frac{t_k-i}{r_{k-1}}} Q_k^i}{\Gamma(i+1)} \right)^{\frac{2}{t_k+1}} \right)^{\frac{2+(t_k+1)r_k}{(t_k+1)r_k}}}{(r_k(t_k+1))^{\frac{2+(t_k+1)r_k}{(t_k+1)r_k}}} \,. \quad (35)$$

Since $\lambda^{-\frac{2+(t_k+1)r_k}{(t_k+1)r_k}}$ can be approximated by λ^{-1} for large t_k, inserting (34) in (22) yields

$$\lambda^{-1} \approx \frac{e}{\sum_{k=1}^{K} \delta_k} \,. \quad (36)$$

see the program

Hence the approximated outage probability is written as

$$\tilde{P}_{\text{out},k} = \tilde{P}^{r_k}_{\text{out},k,j'} \approx \frac{\delta_k}{\sum_{k=1}^{K} \delta_k} \cdot e \ , \tag{37}$$

insert it into (31), replacing e by $\tilde{P}_{\text{out},k-1}$ we finally achieve β_k given in Theorem 1. This concludes the proof.

9. REFERENCES

[1] M. Dohler, *Virtual Antenna Arrays*, Ph.D. thesis, King's College London, U.K., November 2003.

[2] J. N. Laneman, D. Tse, and G.W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. on Information Theory*, vol. 50, no. 12, pp. 3062–3080, February 2004.

[3] Y. Lang, D. Wübben, C. Bockelmann, and K.-D. Kammeyer, "A Closed Power Allocation Solution for Outage Restricted Distributed MIMO Multi-hop Networks," in *Workshop on Resource Allocation in Wireless Networks (RAWNET)*, Berlin, Germany, March 2008.

[4] Y. Lang, D. Wübben, and K.-D. Kammeyer, "Efficient power allocation for outage restricted asymmetric distributed mimo multi-hop networks," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Cannes, France, September 2008.

[5] D. Wübben and Y. Lang, "Near-optimum Power Allocation for Outage Restricted Distributed MIMO Multi-hop Networks," in *IEEE Proc. Global Communications Conference (GLOBECOM)*, New Orleans, USA, November 2008.

[6] M. Kendall and A. Stuart, *The Advanced Theory of Statistics*, vol. 1, Griffen, London, U.K., 4. edition, 1979.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambrigde University Press, 2004.

[8] M. Dohler and M. Arndt, "Inverse Incomplete Gamma Function and Its Application," *IEE Electronics Letters*, vol. 42, no. 1, pp. 35–36, January 2006.

[9] Y. Lang, D. Wübben, and K.-D. Kammeyer, "Power allocations for adaptive distributed mimo multi-hop networks," in *IEEE International Conference on Communications (ICC)*, Dresden, Germany, June 2009.

see the program

ADAPTIVE MODULATION FOR A REDUCED FEEDBACK STRATEGY IN THE MULTIBEAM OPPORTUNISTIC SCHEMES

Nizar Zorba

Centre Tecnològic de Telecomunicacions de Catalunya
Av. Canal Olímpic, 08860-Castelldefels, Barcelona, Spain
nizar.zorba@cttc.es

Ana I. Pérez-Neira

Universitat Politècnica de Catalunya (UPC)
C/ Jordi Girona, Mòdul D5, 08034-Barcelona, Spain
anuska@gps.tsc.upc.edu

ABSTRACT

Scheduling in a multiantenna downlink channel based on partial Channel State Information at the Transmitter (CSIT) is carried out through a Multibeam Opportunistic Beamforming (MOB) technique. To enable MOB implementation in realistic systems, a threshold is applied at the receivers side to decrease the feedback load, so that only users with a channel indicator above a predefined threshold are allowed to ask for service. A closed form expression for the throughput is obtained for such scenario, where an adaptive modulation strategy is considered to align with the current broadband wireless standards. A comparison with other more sophisticated but also complex strategies is considered.

1. INTRODUCTION

One of the major transmission techniques within partial CSIT scenarios is the MOB scheme, where random beams are generated at the Base Station (BS) to serve several users at the same time, obtaining an additional multiplexing gain [1] [2]. Its partial channel state information is in the form of Signal-to-Noise-Interference-Ratio (SNIR) that decreases the feedback load. Even this strategy is a partial CSIT one, but the SNIR value for all active users drives a considerable feedback load in the system, which decreased the system efficiency and can stand as a handicap for its consideration in standards.

A possible approach to decrease such load is by using a threshold philosophy [3] on the feedback value, where such a scheme considerably mitigates the feedback load. The performance decrease is negligible in comparison to the feedback load enhancement [3], which motivates this strategy to be an attractive option for its inclusion in commercial systems, where the opportunistic systems are already included in the UMTS-HSDPA systems, and standing as a potential scheme for LTE. Another strategy that is already implemented in realistic systems is the Adaptive Coding and Modulation (ACM) [4] where the transmitter is continuously changing the employed modulation and channel code to match them to the instantaneous channel conditions; where ACM is also shown to increase the system performance.

Along this paper we focus on the MOB strategy and we study the feedback quantization [5] as well as the application of the feedback threshold. Closed form expressions that help in the system design are obtained thanks to the Cumulative Density Function (CDF) characterization that we provide for MOB. With the consideration of ACM, the paper also presents the closed form expression for the average system throughput to enable the system evaluation and its performance optimization; where up to the authors' knowledge, no such formulation is previously obtained in literature.

To properly evaluate the MOB performance under the proposed scenario, another transmitting technique is also tackled: the Zero Forcing Beamforming (ZFB) [6], that shows good rate performance and low complexity implementation, but it requires for full CSIT, which constitutes a huge amount of feedback information in multiuser scenarios. For a fair comparison, we will employ the same amount of feedback load for both techniques.

The remainder of this paper is organized as follows: while section II deals with the system model, in section III a review of the MOB procedure is presented, where the feedback quantization and threshold application are introduced. Section IV exposes the ZFB scheme, while the ACM strategy with its closed form expression are tackled in section V. The numerical results and simulations are in section VI, followed by the conclusions in section VII.

2. SYSTEM MODEL

We focus on the single cell Downlink channel where N receivers, each one of them equipped with a single receiving antenna, are being served by a transmitter at the BS provided with n_t transmitting antennas, and supposing that N is greater than n_t. A wireless multiantenna channel $\mathbf{h}_{[1\times n_t]}$ is considered between each of the users and the BS, where a quasi-static block fading model is assumed, which keeps constant

This work was partially supported by the Catalan Government under grant SGR2005-00996, by the Spanish Government under project TEC2008-06327-C03-01 (FBMC-MULAN), and by the European project ICT-2007.1.1 NEWCOM++ 216715.

see the program

through the coherence time, and independently changes between consecutive time intervals with independent and identically distributed (i.i.d.) complex Gaussian entries $\sim \mathcal{CN}(0,1)$. Therefore this model captures the instantaneous channel fluctuations over each coherence time, where all users are assumed to keep fixed during each fading block, and allowed to move from block to block. Let $\mathbf{x}(t)$ be the $n_t \times 1$ transmitted vector, while denote $y_i(t)$ as the i^{th} user received signal, given by

$$y_i(t) = \mathbf{h}_i(t)\mathbf{x}(t) + z_i(t) \tag{1}$$

where $z_i(t)$ is an additive Gaussian complex noise component with zero mean and $E\{|z_i|^2\} = \sigma^2$. The transmitted signal $\mathbf{x}(t)$ encloses the independent data symbols $s_i(t)$ to all the selected users with $E\{|s_i|^2\} = 1$. A total transmitted power constraint P_t is considered, and for ease of notation, time index is dropped whenever possible.

3. MULTIBEAM OPPORTUNISTIC BEAMFORMING (MOB)

One of the main transmission techniques in multiuser MIMO scenarios is the MOB technique [1], where random beams are generated at the BS to simultaneously serve more than one user. The beam generation follows an orthogonal manner to decrease the interference among the served users, where n_t beams are generated. Within the acquisition step, each one of the users sequentially calculates the SNIR related to each beam, and feeds back only the best value to the BS together with an integer indicating the index of the selected beam. The BS scheduler chooses the user with the largest SNIR value for each one of the beams, so that it extracts the multiuser gain from the scenario. After that, the BS enters the transmission stage and simultaneously forwards every one of the n_t selected users with its intended data, where no user can obtain more than one beam at a time.

Through this low-complexity transmitter processing and at the same time, an opportunistic user selection based on the instantaneous SNIR values, the MOB strategy achieves high system sum rate by spatially multiplexing several users at the same time, making the transmitted signal to enclose the data symbols for the n_t selected users as

$$\mathbf{x} = \sqrt{\frac{1}{n_t}} \sum_{m=1}^{n_t} \mathbf{b}_m \, s_m \tag{2}$$

with $\mathbf{b}_m$ as the unit-power beam assigned to the m^{th} user, where the square root term is due to a total power constraint of $P_t = 1$.

This scheme is characterized by its SNIR term due to the interference that each beam generates to its non-intended users, representing a major drawback of this system; thus stating the SNIR formulation for the i^{th} user through the m^{th} beam as

$$SNIR_{i,m} = \frac{\frac{1}{n_t}\left|\mathbf{h}_i \mathbf{b}_m\right|^2}{\sigma^2 + \sum\limits_{u \neq m}^{n_t} \frac{1}{n_t}\left|\mathbf{h}_i \mathbf{b}_u\right|^2} \tag{3}$$

where a uniform power allocation is considered.

Based on the MOB philosophy to deliver service to the best users, the *serving SNIR* value is the maximum SNIR over the active users in the system corresponding to each generated beam. Using the SNIR equation in (3) with n_t transmitted beams, note that the numerator follows a Chi-square $\chi^2(2)$ distribution while the interference terms in the denominator are modeled as $\chi^2(2(n_t-1))$, which allows to obtain the SNIR Probability Distribution Function (PDF) as [1] [2]

$$f(x) = \frac{e^{-(x \cdot n_t \sigma^2)}}{(1+x)^{n_t}} \Big(n_t \sigma^2 (1+x) + n_t - 1 \Big) \tag{4}$$

and the CDF is then formulated as

$$F(x) = 1 - \frac{e^{-(x \cdot n_t \sigma^2)}}{(1+x)^{n_t - 1}} \tag{5}$$

and since the serving SNIR is the maximum over all the users' SNIR values, then the serving SNIR CDF is stated as

$$FF(x) = (F(x))^N = \left[1 - \frac{e^{-(x \cdot n_t \sigma^2)}}{(1+x)^{n_t - 1}}\right]^N \tag{6}$$

3.1. Quantization

To allow the feedback process from each one of the users to the BS, some quantization on the SNIR values is required in practical systems to limit the number of feedback bits [7]. A scalar uniform quantization scheme is the easiest and most practical case, where the system designer must define two parameters: the maximum quantization error ϵ (e.g. $\epsilon = 0.01$), and the quantization range η (i.e. 95% whole range coverage). For the latter, it is required to define the largest SNIR value M that is covered with η, where the relation that links between M and η is through the CDF formulation as

$$\eta = \left[1 - \frac{e^{-(M \cdot n_t \sigma^2)}}{(1+M)^{n_t - 1}}\right]^N \tag{7}$$

With the two quantization parameters defined, the quantization process is straightforward, as each user compares the maximum SNIR value to the quantization intervals, thus it can find the interval where its SNIR value fits, and then it feeds back the interval middle value.

3.2. Feedback Threshold

MOB is a partial CSIT scheme, as only the maximum SNIR is fed back to the BS but nevertheless, such a feedback load is large for commercial systems, motivating the seek for alternative approaches to decrease the feedback, to make MOB

see the program

being attractive for its introduction in communication standards. A smart technique to decrease the feedback load is to allow only the users with good enough channels to enter the feedback process [3], so that only the users with their SNIR value above a predefined threshold are asked to feedback their SNIR value. Based on the SNIR distribution, the feedback load decrease can be very large, while the performance decrease is insignificant [3]. Moreover, this selective feedback strategy can be also employed to meet each user's Quality of Service (QoS) demands.

Therefore, the system administrator can impose an SNIR threshold γ to the feedback process, but a main problem for the application of the feedback threshold relates to the outage possibility P_{out} of no-feedback in the system, as the γ value can be too high, so that no-user participates in the feedback process. The probability for such phenomena is also obtained from the CDF as follows

$$P_{out} = 1 - \left[1 - \frac{e^{-(\gamma \cdot n_t \sigma^2)}}{(1+\gamma)^{n_t - 1}}\right]^N \tag{8}$$

where if the threshold is too low, then a high number of user participate in the feedback process, making the P_{out} to be negligible, but its price is a large feedback load. The most suitable value is left for the system administrator to select the one that satisfies the system demands and requirements.

3.3. Feedback Load

To calculate the feedback load is of critical importance to evaluate any scheme probabilities to be included in any commercial system, and thanks to the MOB CDF formulation, the feedback load is now presented.

Through the quantization step, we just defined the system maximum range of SNIR values that will be quantized. Also, the maximum quantization error was denoted to allow the system to adapt its processing to the received SNIR value at each user. Therefore, the number of quantization steps Q_s can be calculated as

$$Q_s = \frac{\eta}{2\epsilon} \tag{9}$$

and as each user not only has to feedback its quantized SNIR value, but also the beam indicator for the corresponding maximum SNIR value, then the feedback from all users (i.e. the system feedback load F_L) is as

$$F_L = N \cdot \Big(\lceil \log_2(Q_s) \rceil + \lceil \log_2 n_t \rceil \Big) \tag{10}$$

where $\lceil x \rceil$ refers to the ceil of x.

4. ZERO FORCING BEAMFORMING

A multiantenna technique that its complexity is affordable for its implementation in practical systems is the *Zero-Forcing Beamforming (ZFB)*, where the cancellation of inter-user interference is accomplished through a channel inversion [6] [8]. Two main drawbacks affect this scheme: the requirement for full CSIT which constitutes a very large feedback load; and its noise enhancement and low performance when the users' channels show some correlation, thus falling in the rank deficiency problem. A special feature of the ZFB scheme is that it does not consider the multiuser availability in their performance, so that some sort of users selection techniques, independently performed from the spatial processing, are required as an initial step before the spatial processing. Such a user selection process helps to obtain uncorrelated users' channels, and ZFB is shown to be optimal if an infinite number of users is available [8].

A main difference between MOB and ZFB relates to their training step [5] where in the MOB, as the training is done with the same transmitting beams, then each user has the knowledge of what will be its serving rate (in case it is selected for transmission), as the effect of the other serviced users is known, all of that through the received SNIR to the user. On the other hand in ZFB, the training is performed so that each user knows its full channel $\mathbf{h}$, but no information at all is provided in relation to the exact rate that will be received (actually no information about the effect of the other serviced users is provided). This has a major effect in feedback quantization and the application of feedback threshold. In MOB, it is obvious that the feedback threshold will be applied to the SNIR values, but in ZFB, the threshold can be only applied to the user channel norm $|\mathbf{h}|^2$, which, as already commented, does not provide information related to the effect of the other serviced users.

5. ADAPTIVE CODING AND MODULATION IN MOB

The wireless channel is continuously fluctuating, where an adaptation to overcome the changes in the channel characteristics is required in order to achieve good performance in the wireless systems. The Adaptive Coding and Modulation (ACM) strategy [4] is accomplished by an instantaneous change in the employed modulation and coding kinds, to match the BS parameters to the channel conditions subject to a predefined Packet Error Rate (PER) system performance. The introduction of ACM in commercial standards is a fact with its presence in WLAN, WiMAX and LTE among others.

Notice that the selection for the users with the best SNIR values for each one of the transmitting beams is actually another way to adapt the transmitter processing to the channel properties. In this aspect, it follows the same strategy as ACM schemes [4], with both strategies looking forward improving the wireless channel performance.

This paper concentrates on the MOB strategy as we think it is more appropriate for its implementation, and we con-

see the program

sider it together with ACM for the transmitter adaptation to the channel characteristics. Then, a closed form expression for the resultant throughput is obtained, where no previous contribution in literature has presented the throughput mathematical expression. Moreover, the proposed feedback threshold can be also considered in our results.

Considering a set of modulation levels with convolutional codes [4] as those presented in Table I, the transmitter selects a modulation kind when the SNIR value γ is in the interval $\gamma_{i-1} \leq \gamma < \gamma_i$, where the resulting rate t_i is also shown for the corresponding SNIR bounding values. Thanks to the user selection process, the BPSK is not employed as almost all the times, the scheduler can find users that have good enough channels to support higher modulation than BPSK. Therefore, the threshold philosophy is applied on the γ value, as the scheduler can decide to ask users with an SNIR value enabling QPSK or higher modulations to participate in the feedback process, so that all users with $\gamma < -4$ are blocked, with the consequent saving in the feedback load. If the threshold is too large, there exists the possibility that no user participating in the feedback process, where this situation is previously discussed and attached with a closed form expression to optimize it according to the system demands and requirements.

Modulation, Code rate	Rate t_i (Mbps)	SNIR γ (dB)
BPSK, 1/2	5	$\gamma <$ -4
QPSK, 1/2	10	-4 $\leq \gamma < 0$
8QAM, 3/4	22.5	$0 \leq \gamma < 3$
16QAM, 3/4	30	$3 \leq \gamma < 7$
32QAM, 5/6	41.6	$7 \leq \gamma < 10$
64QAM, 5/6	50	$10 \leq \gamma$

Table I: SNIR Thresholds

It is worth noting that the SNIR value guarantees that the user's decoding process is successful. In that case a unit step function is used for the detection procedure, making the Packet Success Rate (PSR) for the i^{th} modulation to relate to SNIR as

$$PSR = \begin{cases} 1 & \text{if } SNIR \geq \gamma_i \\ 0 & \text{if } SNIR < \gamma_i \end{cases} \tag{11}$$

Having a look on Table I, if we can calculate the probability of the serving SNIR to be within each modulation interval, then we can formulate the closed form expression for the system aggregate throughput, mainly from the obtained CDF for the SNIR service distribution.

So that with a consideration of ACM that is defined in S steps, with each step related to SNIR thresholds (e.g. those presented in Table I), the throughput is directly related to the transmission rate that is employed in each step. Considering the CDF in (6), the probability $U_{i(m)}$ to have an SNIR value below a threshold γ_i and above a threshold γ_{i-1} for the m^{th} served user, is

$$U_{i(m)} = FF(\gamma_i) - FF(\gamma_{i-1}). \tag{12}$$

As the system provides service to n_t simultaneous and mutually independent users, the total probability for the k^{th} group of users to be served through a set of rates is

$$P(set_k) = \prod_{m=1}^{n_t} U_{i(m)} \tag{13}$$

The resultant throughput T_k for the k^{th} set option is obtained as the sum of the rates corresponding to the m users within that set as

$$T_k = \sum_{m=1}^{n_t} t_m \tag{14}$$

but the total throughput T is formulated as the sum over all possible combinations of sets, each set multiplied by its corresponding probability $P(set_k)$, as follows

$$T = \sum_{k=1}^{K} T_k \, P(set_k) \tag{15}$$

with K as the total number of possible sets. Even this formulation seems to be straightforward, but it has not been presented in literature due to the absence of the probability formulation in (12) that was obtained thanks to the CDF MOB formulation.

6. SIMULATIONS

The performance of the discussed scheme is presented by Monte Carlo simulations, where the serving policy tackles the system sum rate maximization. We consider a Downlink single cell wireless scenario with $n_t = 2$ transmitting antennas and a variable number of active users, each one equipped with a single receiving antenna. The channel of each user is assumed to be a block fading complex Gaussian channel $\sim \mathcal{CN}(0, 1)$. The transmitter either runs the MOB technique where two orthogonal beams are set up, or ZFB with beamforming to two users at the same time. A system bandwidth of 10MHz is considered together with a total transmitted power $P_t = 1$ and a noise variance of $\sigma^2 = 1$.

We start with a study of the system aggregate throughput performance under the ACM application, and whether the closed form expression that we provide is valid or not. Fig. 1 shows the throughput performance for a variable number of users in the system. The performance of MOB, ZFB and the closed form theoretical expression for MOB are shown all together; where we notice the similar performance of MOB and the obtained closed form expression. Obviously for a fair comparison, the feedback load on both strategies is the same. First of all, we fix the quantization to cover 95% of the whole range, and a maximum quantization error of 0.01; with these values and following the expression in (7), we obtain that 200 quantization levels are required for MOB to fulfil the above demands. Therefore, the feedback load is obtained as 9 bits per user. Considering that the number of users is 10 for example, then 90 bits are required for the feedback load. We

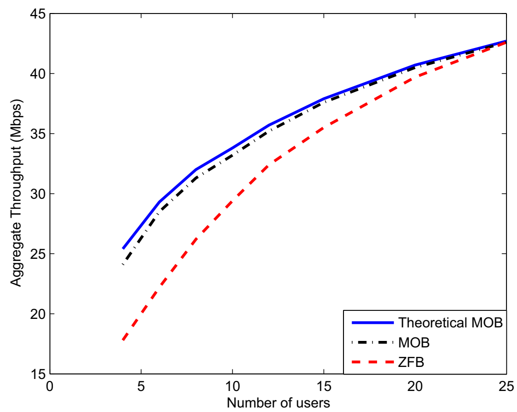


Fig. 1. *The ACM aggregate throughput performance.*

consider this total amount of bit for MOB and we notice that to feedback the full CSIT required for ZFB, only 5 users can feedback their information. As a maximum of two users can be serviced (as $n_t = 2$), then a users selection process is required for ZFB. The great performance and low complexity strategy proposed in [8] for ZFB is tackled. From the results, we see that the ZFB comes closer to MOB as we increase the number of users, and we notice that for all realistic number of users (with a maximum of 20 active users per cell), MOB always outperforms ZFB.

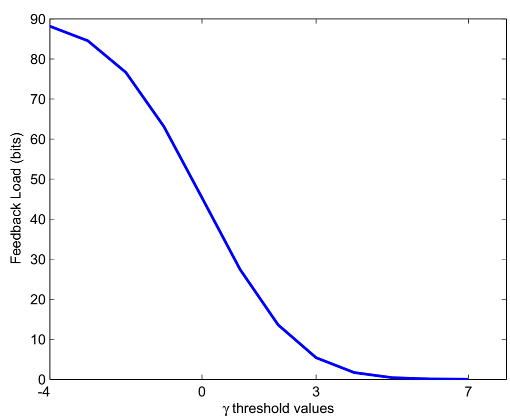


Fig. 2. *The feedback load for a variable threshold.*

The effect of the feedback threshold plays a major role in the system, as it directly affects both the feedback load and the probability of no-user, that we defined it as the probability of outage P_{out}. Fig. 2 plots the decrease in the feedback load by increasing the threshold, where this is an obvious result, so that the system administrator can fix some modulation threshold value as the feedback threshold on the basis of the allowed maximum feedback load.

As previously commented, the increase in the feedback load can drive an undesired effect: the increase of the probability of no-user participating in the feedback process, thus wasting all the multiuser diversity in the system. Fig. 3 shows how the probability of outage is higher by an increase in the threshold value, where the operating point depends on the system demands and requirements.

7. CONCLUSION

This paper has presented the MOB scheme under practical considerations for its implementation, where the feedback quantization is accomplished. To further decrease the feedback load in this scheme, a feedback threshold is proposed, where the exact equations that control its performance are presented. An Adaptive Coding and Modulation strategy is tackled to align with current broadband wireless systems, and the paper introduced a closed form expression for the average system throughput, under the above considerations.

To evaluate the MOB performance, it is compared to the ZFB scheme, where both of them are run under same conditions and within the same feedback load. The results had shown the higher performance of MOB in practical scenarios. The simulations also considered the effect of the feedback threshold on both the feedback load and the probability of outage, where a tradeoff between the two metrics appeared.

8. REFERENCES

[1] M. Sharif and B. Hassibi, “On the Capacity of MIMO Broadcast Channel with Partial Side Information”, *IEEE Trans. Inform. Theory*, Feb. 2005.

[2] N. Zorba and A.I. Pérez-Neira, “Robust Power Allocation Schemes for Multibeam Opportunistic Transmission Strategies Under Quality of Service Constraints”, *IEEE JSAC*, Aug. 2008.

[3] V. Hassel, D. Gesbert, M.S. Alouini, and G. Oien, “A Threshold Based Channel State Feedback Algorithm for Modern Cellular Systems”, *IEEE Trans. Wireless Comm.* July 2007.

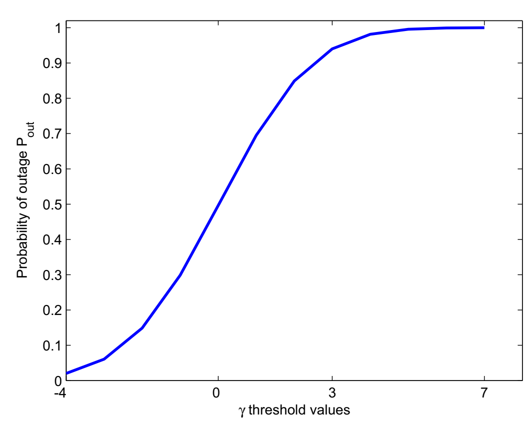


Fig. 3. *The probability of outage for a variable feedback threshold.*

see the program

[4] X. Wang, Q. Liu, and G.B. Giannakis, “Analyzing and Optimizing Adaptive Modulation Coding Jointly With ARQ for QoS-Guaranteed Traffic”, *IEEE Trans. Vehicular Tech.* Mar. 2007

[5] G. Caire, N. Jindal, M. Kobayashi and N. Ravindran, “Achievable Throughput of MIMO Downlink Beamforming with Limited Channel Information”, *IEEE-PIMRC*, Athens-Greece, 2007.

[6] D. Bartolomé and Ana I. Pérez-Neira, “Spatial Scheduling in Multiuser Wireless Systems: from Power Allocation to Admission Control,” *IEEE Transactions on Wireless Communications*, vol. 5, August 2006.

[7] J.L. Vicario, R. Bosisio, U. Spagnolini, and C. Anton-Haro, “A Throughput Analysis for Opportunistic Beamforming with Quantized Feedback,” *IEEE-PIMRC*, Helsinki-Finland, 2006.

[8] T. Yoo and A. Goldsmith, “On the Optimality of Multi-Antenna Broadcast Scheduling Using Zero-Forcing Beamforming,” *IEEE JSAC*, vol. 24, March 2006.

see the program

COMPARISON OF LOCALIZATION ALGORITHMS USING TIME DIFFERENCE OF ARRIVAL

Dirk Czepluch, Franz Demmel

Rohde & Schwarz
dirk.czepluch@rohde-schwarz.com
franz.demmel@rohde-schwarz.com

Stefan Schmidt

Rohde & Schwarz
stefan-a.schmidt@rohde-schwarz.com

ABSTRACT

This paper performs a study on different TDOA based localization algorithms, in dispersive radio channels. The cross-correlation between the signals is performed under COST259 [1] channel model conditions. The scenario is an non-cooperative system and the spectrum of the signal is unknown. The different methods are localization based on intersecting hyperbolas, location on the conic axis (LOCA) and a divide and conquer (DAC) approach.

1. INTRODUCTION

This paper is a summary of different localization methods based on the time difference of arrival. We consider a uncooperative system. This means there is no communication between the transmitter and the receivers or trainings sequence of a given communication standard. This is the standard case in localization systems using by federal authorities.

The paper is organized in six sections. In the second Section we will show, how to get the time difference of arrival. The next three Sections discusses different methods to estimate the position of the unknown transmitter. This means in detail the localization method with intersecting hyperbolas, Section three, followed by the localization based on conic axes in Section four and finished with a divide and conquer approach in Section five. Section 6 presented the simulation results in detail. The last Section yields the conclusion of the paper.

2. MEASUREMENT OF TIME DIFFERENCE OF ARRIVAL

If a electromagnetic wave occurs at a transmitter, this will arrive at a set of different receivers at a different times. To make this more explicit let us define a set of receivers and a signal source (transmitter) given by its coordinates $R_i \triangleq [x_i, y_i]^T$ for $i = 1, \ldots, N_r$, where N_r denotes the number of receivers, and $T \triangleq [x_0, y_0]^T$ for the transmitter respectively. The range distances r_i for $i = 1, \ldots, N_r$ from each receiver to the transmitter T are given as

$$r_i \triangleq \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2} \text{for} i = 1, \ldots, N_r \quad (1)$$

Now the time difference of arrival between receiver i and receiver j can be written as

$$t_{ij} \triangleq \frac{r_j - r_i}{c} \text{for} i = 1, \ldots, N_r \text{and} j > i \quad (2)$$

where c denotes the speed of light. Now let $x_i(t)$ and $x_j(t)$ for $t = 1 \ldots K$ be time series measured at receivers R_i and R_j respectively. One common method to determine t_{ij} is given by means of the standard cross-correlation function

$$R_{x_i x_j}(\tau) = E[x_i(t) x_j^*(t - \tau)] \quad (3)$$

The argument τ that maximizes equation (3) provides an estimate of t_{ij}. In future the t_{ij} will be described as Δt. Note that in literature there is a concept called generalized cross correlation which aims at improving the accuracy of the time delay estimate through appropriate pre-filtering of the time series $x_i(t)$ and $x_j(t)$ before cross-correlation. For example, one common approach is to choose the pre-filters in a way that the signals passed to the correlator are accentuated at those frequencies at which the signal-to-noise ratio (SNR) is highest. Unfortunately the generalized cross-correlation methods require knowledge of the signal spectrum, which is not given in case of the non-cooperative location problem considered in this work. Therefore the signal spectrum would have to be estimated in order to apply these methods, which in turn would lead to a large amount of additional signal processing. This is the reason why most practical non-cooperative location systems rely on the standard cross-correlation procedure as stated in (3) and we do not consider the generalized cross-correlation further in this work. Nevertheless we refer to [2] for a very good summary of the most important generalized cross-correlation techniques.

3. INTERSECTING HYPERBOLAS

The first localization method based on time difference of arrival we presented in this paper based on the intersecting of

see the program

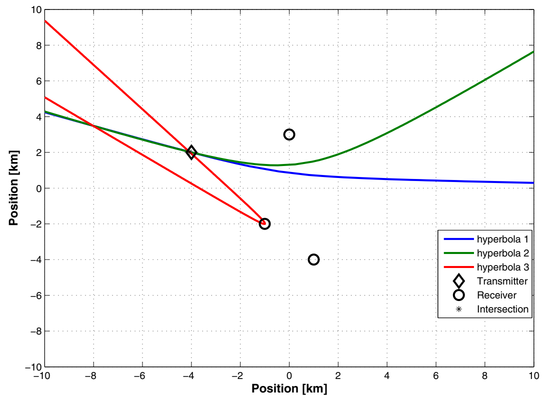


Fig. 1. Intersection of three hyperbola axis

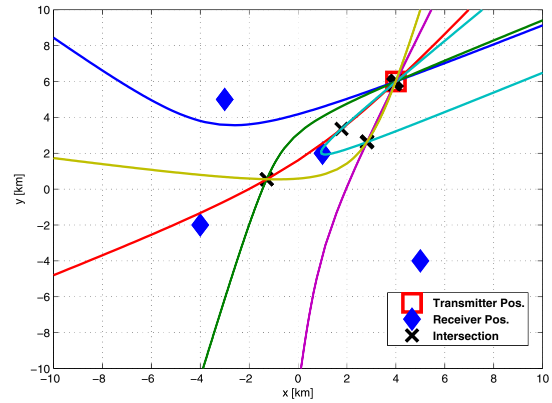


Fig. 2. Intersection of four hyperbola axis

hyperbolas. It can be easily shown that the calculated time difference between two receivers the transmitter possible source location is existing of a hyperbola. The sign of the time difference leads to the hyperbola axis.

Without loss of generality we assume the receivers to be located at x-axis with distance of a and the transmitter is located at $[x_0, y_0]$. Then the time difference between the two receivers is determined through

$$\Delta t = \frac{\sqrt{(-\frac{a}{2} - x_0)^2 + y_0^2} - \sqrt{(\frac{a}{2} - x_0)^2 + y_0^2}}{c} \tag{4}$$

where c denotes the speed of light. After transforming this equation this finally ends up with

$$1 = \frac{x_0^2}{c^2\delta t^2/4} - \frac{y_0^2}{a^2/4 - c^2\delta t^2/4} \tag{5}$$

which is recognized as the equation of a hyperbola. In figure below the hyperbola is plotted for the two cases. The first scenario is a three receiver scenario, which shows the problem that three receivers are not enough for unambiguous estimation. The figure () shows the solution of the problem, while using four receivers.

4. LOCATION ON THE CONIC AXIS - LOCA

In this section we will present another localization method based of time differences. The method is called the "Location on the Conic Axis (LOCA)" algorithm given in [2]. Schmidt shows that the range differences from the transmitter to three known receivers, which can be obtained by multiplying the measured TDOAs with the speed of light c, provide a straight line of position, which is the major axis of a general conic (conic axis). The conic axis connects the foci of a conic. Moreover it is shown that the three receivers lie on

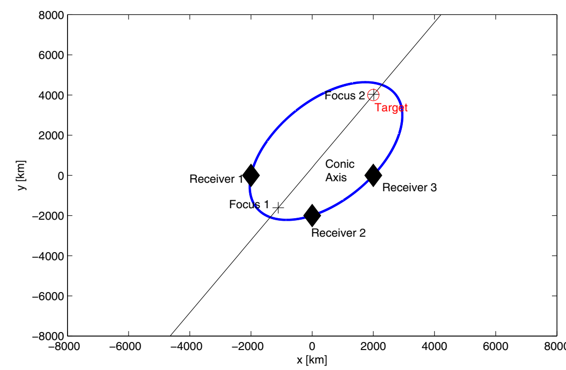


Fig. 3. The conic is an ellipse

the conic and one of the foci is the location in question, see figures (4)and (**??**). Therefore his algorithm is referred to as LOCA (Location on the Conic Axis).

No we describe the approach in more detail. Let $R_i \triangleq [x_i, y_i]^T$ for $i = 1, 2, 3$ and $T \triangleq [x_0, y_0]^T$ define the coordinates of the receivers and the transmitter respectively. Moreover let r_i denote the range between receiver R_i and transmitter T. Then with

$$\begin{aligned} r_1^2 &\triangleq (x_0 - x_1)^2 + (y_0 - y_1)^2 \\ &= x_0^2 + y_0^2 - 2x_0x_1 - 2y_0y_1 + a_1^2 \end{aligned} \tag{6a}$$

$$\begin{aligned} r_2^2 &\triangleq (x_0 - x_2)^2 + (y_0 - y_2)^2 \\ &= x_0^2 + y_0^2 - 2x_0x_2 - 2y_0y_2 + a_2^2 \end{aligned} \tag{6b}$$

$$\begin{aligned} r_3^3 &\triangleq (x_0 - x_3)^2 + (y_0 - y_3)^2 \\ &= x_0^2 + y_0^2 - 2x_0x_3 - 2y_0y_3 + a_3^2 \end{aligned} \tag{6c}$$

$\Delta_{ij} \triangleq r_j - r_i$ and $a_i^2 \triangleq x_i^2 + y_i^2$ for $i = 1, 2, 3$ one has

see the program

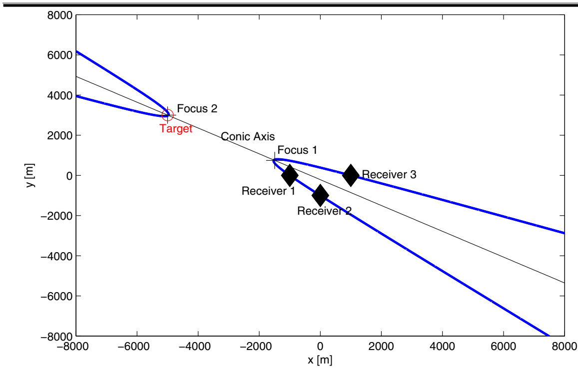


Fig. 4. The conic is an hyperbel

$$\begin{aligned} r_2^2 - r_1^2 &= \Delta_{12}(r_1 + r_2) \\ &= 2x_0(x_1 - x_2) + 2y_0(y_1 - y_2) + (a_2^2 - a_1^2) \end{aligned} \tag{7a}$$

$$\begin{aligned} r_3^2 - r_2^2 &= \Delta_{23}(r_2 + r_3) \\ &= 2x_0(x_2 - x_3) + 2y_0(y_2 - y_3) + (a_3^2 - a_2^2) \end{aligned} \tag{7b}$$

$$\begin{aligned} r_1^2 - r_3^2 &= \Delta_{31}(r_3 + r_1) \\ &= 2x_0(x_3 - x_1) + 2y_0(y_3 - y_1) + (a_1^2 - a_3^2) \end{aligned} \tag{7c}$$

Now solving for range sums one gets

$$r_1 + r_2 = 2x_0\frac{x_1 - x_2}{\Delta_{12}} + 2y_0\frac{y_1 - y_2}{\Delta_{12}} + \frac{a_2^2 - a_1^2}{\Delta_{12}} \tag{8a}$$

$$r_2 + r_3 = 2x_0\frac{x_2 - x_3}{\Delta_{23}} + 2y_0\frac{y_2 - y_3}{\Delta_{23}} + \frac{a_3^2 - a_2^2}{\Delta_{23}} \tag{8b}$$

$$r_3 + r_1 = 2x_0\frac{x_3 - x_1}{\Delta_{31}} + 2y_0\frac{y_3 - y_1}{\Delta_{31}} + \frac{a_1^2 - a_3^2}{\Delta_{31}} \tag{8c}$$

which in turn leads to

$$\begin{aligned} (r_2 + r_3) - (r_3 + r_1) =& r_2 - r_1 \\ =& \Delta_{12} \\ =& 2x_0\left[\frac{x_2 - x_3}{\Delta_{23}} - \frac{x_3 - x_1}{\Delta_{31}}\right] \\ &+ 2y_0\left[\frac{y_2 - y_3}{\Delta_{23}} - \frac{y_3 - y_1}{\Delta_{31}}\right] \\ &+ \frac{a_3^2 - a_2^2}{\Delta_{23}} - \frac{a_1^2 - a_3^2}{\Delta_{31}} \end{aligned} \tag{9}$$

Finally, after multiplying (9) with the common denominator and rearranging one obtains

$$\begin{aligned} &[x_1\Delta_{23} + x_2\Delta_{31} + x_3(\Delta_{12} - \Sigma)]x_0 \\ &\quad + [y_1\Delta_{23} + y_2\Delta_{31} + y_3(\Delta_{12} - \Sigma)]y_0 \\ &= \frac{1}{2}[\Delta_{12}\Delta_{23}\Delta_{31} + a_1^2\Delta_{23} + a_2^2\Delta_{31} + a_3^2(\Delta_{12} - \Sigma)] \end{aligned} \tag{10}$$

where

$$\Sigma \triangleq \Delta_{12} + \Delta_{23} + \Delta_{31} \tag{11}$$

Equation (9) can be written in a more compact form as $Ax_0 + By_0 = C$, where A, B and C are constants, determined by the receiver locations R_i for $i = 1, 2, 3$ and the measured range differences Δ_{12}, Δ_{23} and Δ_{31}. Thus it is established that the transmitter $T = [x_0\ y_0]^T$ lies on the straight line given by $Ax + By = C$.
In [1] it is shown that equation $Ax + By = C$ together with the three receiver locations R_i for $i = 1, 2, 3$ uniquely define a conic whose axis (= line, where the foci of the conic lie on) is given by $Ax + By = C$. Moreover the transmitter location $T = [x_0\ y_0]^T$ is given as one of the two foci. This means that once we have determined the coefficients A, B and C from the measured range differences $\Delta_{12}, \Delta_{23}, \Delta_{31}$ and the receiver locations R_i for $i = 1, 2, 3$ for one receiver triad (= set of three receivers) there are two ways to proceed in order to get a fix (= estimate of the transmitter position). If more than three receivers are available we can determine the coefficients A', B' and C' for a second receiver triad and the fix will be given as the intersection of the two conic axes defined by $Ax + By = C$ and $A'x + B'y = C'$.
If on the other hand only three receivers are available three different cases have to be considered. In the first case the conic is an ellipse. Then the focus which does not correspond to the transmitter position T would produce the negative range differences $-\Delta_{12}, -\Delta_{23}$ and $-\Delta_{31}$, if the transmitter had its location there, see figure 8. In the second case the conic is an hyperbola. Then a transmitter being located at either of the foci would produce the same range differences, see figure 6. This means that in the case of the conic being a hyperbola, an ambiguity arises and it is not possible to decide mathematically which of the two foci corresponds to the transmitter position T. In the third case one of the foci has moved to infinity and the conic is a parabola [2].

Overall, only the case of the conic being a hyperbola leads to an ambiguity in determining the transmitter location T.

5. DIVIDE AND CONQUER APPROACH - DAC

So far we have considered two cases. In the first case only three receivers are available and the fix can be estimated by determining the foci of a conic (LOCA). In the second case

see the program

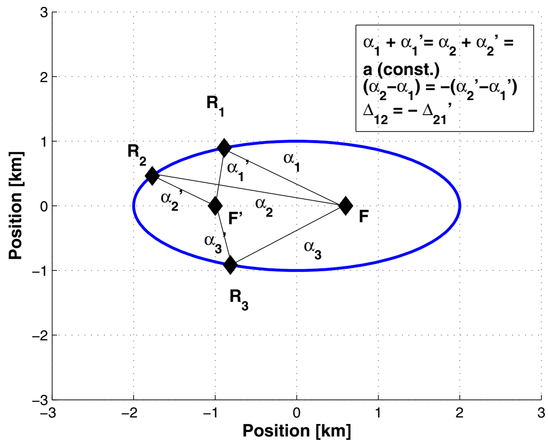


Fig. 5. The solution conic is an ellipse

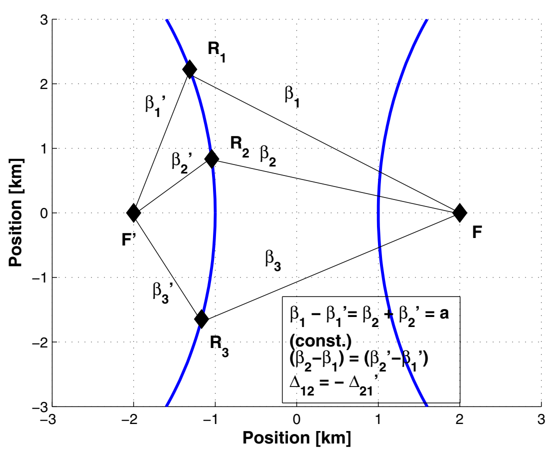


Fig. 6. The solution conic is an hyperbola

more than three receivers are available and the fix can be estimated through cutting the conic axes corresponding to the different receiver triads.

Another approach for the case that $N > 3$ receivers are available would be to determine a fix for each receiver triad by applying LOCA which leads to a set of $\binom{N}{3}$ fixes. In the next step these $\binom{N}{3}$ fixes should be combined in an optimal way in order to produce a reliable estimate of the source location. In this section we will show how the so called divide and conquer approach proposed in [3] can be applied to obtain a maximum likelihood (ML) estimate of the source location by optimally combining the $\binom{N}{3}$ appropriately weighted fixes.

For the mathematical background we refer to [3].

5.1. Formation of the DAC estimate

In order to form the DAC estimate one has to divide the data vector $\mathbf{x}$ in m possibly overlapping sub vectors $\mathbf{x}_i$ for $i = 1, \dots, m$. We have to make sure that each element of $\mathbf{x}$ is represented in at least one $\mathbf{x}_i$. In the next step for each sub vector $\mathbf{x}_i$ an estimate $\hat{\theta}_i$ is determined via maximum likelihood, for example through solving (**??**). In general θ_i can be a sub vector of θ given as

$$\theta_i = \mathbf{S}_i \theta \tag{12}$$

where $\mathbf{S}_i$ is a selection matrix. $\mathbf{S}_i$ can be obtained from an identity matrix by deleting the rows j corresponding to the elements j of θ that one does not want to estimate via $\mathbf{x}_i$. The DAC estimate $\hat{\theta}_{DAC}$ can then be obtained by linearly combining the estimates θ_i according to [4,5]

$$\hat{\theta}_{DAC} = (\mathbf{S}^T \mathbf{W} \mathbf{S})^{-1} \mathbf{S}^T \mathbf{W} \hat{\vartheta} \tag{13}$$

where $\mathbf{S}$ and $\hat{\vartheta}$ are constructed through concatenating the $\hat{\theta}_i$ and $\mathbf{S}_i$ as

$$\hat{\vartheta} \triangleq \begin{bmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_m \end{bmatrix}, \quad \mathbf{S} \triangleq \begin{bmatrix} \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_m \end{bmatrix} \tag{14}$$

and $\mathbf{W}$ is a positive definite weighting matrix, chosen in a way that the mean squared error of the estimate $\hat{\theta}_{DAC}$ is minimized.

5.2. Bias and Variance

In this subsection we will derive bias and covariance of the DAC estimate in (13) for the case that the data vector $\mathbf{x}$ is Gaussian distributed. Lets start with the following definitions

see the program

$$\mu_i \triangleq E\{\mathbf{x}_i\} \quad (15)$$
$$\mathbf{C}_i \triangleq E\{[\mathbf{x}_i - \mu_i][\mathbf{x}_i - \mu_i]^T\} \quad (16)$$
$$\mathbf{C}_{ij} \triangleq E\left\{[\mathbf{x}_i - \mu_i][\mathbf{x}_j - \mu_j]^T\right\} \quad (17)$$
$$\mathbf{J}_i \triangleq \frac{\partial \mu_i^T}{\partial \theta_i}\mathbf{C}_i^{-1}\frac{\partial \mu_i}{\partial \theta_i^T} \quad (18)$$

Assuming the $\hat{\theta}_i$ to be ML estimates and sufficiently large Fisher information $\mathbf{J}_i$ we can apply (19) getting from [3]

$$\hat{\theta}_{ML} \approx \theta + \mathbf{J}^{-1}(\theta)\frac{\partial \ln p(\mathbf{x},\theta)}{\partial \theta} \quad (19)$$

and come up with

$$E\{\hat{\theta}_i\} \approx E\left\{\theta_i + \mathbf{J}_i^{-1}\frac{\partial \ln p(\mathbf{x}_i,\theta_i)}{\partial \theta_i}\right\} \quad (20)$$
$$= E\{\theta_i + \frac{1}{\partial \theta_i}\partial(-\frac{1}{2}\ln\det\{2\pi\mathbf{C}_i\} \quad (21)$$
$$-\frac{1}{2}[\mathbf{x}_i - \mu_i]^T\mathbf{C}_i^{-1}[\mathbf{x}_i - \mu_i])\} \quad (22)$$
$$= E\left\{\theta_i + \frac{\partial(-\frac{1}{2}[\mathbf{x}_i - \mu_i]^T\mathbf{C}_i^{-1}[\mathbf{x}_i - \mu_i])}{\partial \theta_i}\right\} \quad (23)$$
$$= E\left\{\theta_i + \frac{\partial \mu_i^T}{\partial \theta_i}\mathbf{C}_i^{-1}[\mathbf{x}_i - \mu_i]\right\} \quad (24)$$
$$= E\{\theta_i\} + \frac{\partial \mu_i^T}{\partial \theta_i}\mathbf{C}_i^{-1}[E\{\mathbf{x}_i\} - \mu_i] \quad (25)$$
$$= \theta_i \quad (26)$$

where in the step from (21) to (23) it has been used that $\mathbf{C}_i$ does not depend on θ_i. In order to obtain (24) from (23) we applied the chain rule in combination with the matrix equality in (**??**). Finally the linearity of the expectation operator $E\{.\}$ has been used to get (25). Thus it has been established that $\hat{\theta}_i$ is unbiased.

Now we will proceed with the determination of the covariance between $\hat{\theta}_i$ and $\hat{\theta}_j$, denoted as $(\mathbf{C}_{\hat{\vartheta}})_{ij}$. Applying (19) once again leads to

$$\begin{aligned}(\mathbf{C}_{\hat{\vartheta}})_{ij} &\triangleq E\left\{[\hat{\theta}_i - E\{\hat{\theta}_i\}][\hat{\theta}_j - E\{\hat{\theta}_j\}]^T\right\} \\ &\approx E\{\left[\theta_i + \mathbf{J}_i^{-1}\frac{\partial \ln p(\mathbf{x}_i,\theta_i)}{\partial \theta_i} - E\{\hat{\theta}_i\}\right] \\ &\quad \left[\theta_j + \mathbf{J}_j^{-1}\frac{\partial \ln p(\mathbf{x}_j,\theta_j)}{\partial \theta_j} - E\{\hat{\theta}_j\}\right]^T\} \\ &= \mathbf{J}_i^{-1}E\left\{\frac{\partial \ln p(\mathbf{x}_i,\theta_i)}{\partial \theta_i}\frac{\partial \ln p(\mathbf{x}_j,\theta_j)}{\partial \theta_j^T}\right\}\mathbf{J}_j^{-1} \\ &= \mathbf{J}_i^{-1}\frac{\partial \mu_i^T}{\partial \theta_i}\mathbf{C}_i^{-1} \\ &\quad * E\left\{[\mathbf{x}_i - \mu_i][\mathbf{x}_j - \mu_j]^T\right\}\mathbf{C}_j^{-1}\frac{\partial \mu_j}{\partial \theta_j^T}\mathbf{J}_j^{-1} \\ &= \mathbf{J}_i^{-1}\frac{\partial \mu_i^T}{\partial \theta_i}\mathbf{C}_i^{-1}\mathbf{C}_{ij}\mathbf{C}_j^{-1}\frac{\partial \mu_j}{\partial \theta_j^T}\mathbf{J}_j^{-1}\end{aligned} \quad (27)$$

Using (13) and $E\{\hat{\vartheta}\} = \mathbf{S}\theta$ the expectation of the DAC estimate $\hat{\theta}$ can be computed as

$$\begin{aligned}E\{\hat{\theta}_{DAC}\} &= (\mathbf{S}^T\mathbf{W}\mathbf{S})^{-1}\mathbf{S}^T\mathbf{W}E\{\hat{\theta}\} \\ &= \theta\end{aligned} \quad (28)$$

and one finds that $\hat{\theta}_{DAC}$ is unbiased. Finally from (13) and (27) the covariance $\mathbf{C}_{\hat{\theta}_{DAC}}$ of $\hat{\theta}_{DAC}$ can be determined [4]

$$\begin{aligned}\mathbf{C}_{\hat{\theta}_{DAC}} &= (\mathbf{S}^T\mathbf{W}\mathbf{S})^{-1}\mathbf{S}^T\mathbf{W}\mathbf{C}_{\hat{\vartheta}}\mathbf{W}\mathbf{S}(\mathbf{S}^T\mathbf{W}\mathbf{S})^{-1} \\ &\succeq (\mathbf{S}\mathbf{C}_{\hat{\vartheta}}^{-1}\mathbf{S})^{-1}\end{aligned} \quad (29)$$

with $\mathbf{C}_{\hat{\theta}_{DAC}} = (\mathbf{S}\mathbf{C}_{\hat{\vartheta}}^{-1}\mathbf{S})^{-1}$ when $\mathbf{W} = \mathbf{C}_{\hat{\vartheta}}^{-1}$. It should be stressed that all the derivations in this section are based on the approximation (19) and therefore require sufficiently large Fisher information $\mathbf{J}_i$ for $i = 1, \ldots, m$.

In the following we will evaluate $\mathbf{C}_{\hat{\theta}_{DAC}}$ for the case that the sub vectors $\mathbf{x}_i$ and the corresponding parameter vectors θ_i have equal length. Then the matrices $\frac{\partial \mu_i^T}{\partial \theta_i}$ are square and since we are assuming non-singular Fisher information matrices $\mathbf{J}_i$ we can conclude

$$\begin{aligned}\mathbf{J}_i^{-1} &= \left(\frac{\partial \mu_i^T}{\partial \theta_i}\mathbf{C}_i^{-1}\frac{\partial \mu_i}{\partial \theta_i^T}\right)^{-1} \\ &= \left(\frac{\partial \mu_i}{\partial \theta_i^T}\right)^{-1}\mathbf{C}_i\left(\frac{\partial \mu_i^T}{\partial \theta_i}\right)^{-1}\end{aligned} \quad (30)$$

which shows that the matrices $\frac{\partial \mu_i^T}{\partial \theta_i}$ are invertible. Now plugging (30) in (27) we can express $(\mathbf{C}_{\hat{\vartheta}})_{ij}$ as

see the program

$$
\begin{aligned}
(\mathbf{C}_{\hat{\vartheta}})_{ij} &= \mathbf{J}_i^{-1} \frac{\partial \mu_i^T}{\partial \theta_i} \mathbf{C}_i^{-1} \mathbf{C}_{ij} \mathbf{C}_j^{-1} \frac{\partial \mu_j}{\partial \theta_j^T} \mathbf{J}_j^{-1} \\
&= \left(\frac{\partial \mu_i}{\partial \theta_i^T} \right)^{-1} \mathbf{C}_{ij} \left(\frac{\partial \mu_j^T}{\partial \theta_j} \right)^{-1}
\end{aligned} \tag{31}
$$

Defining $\mathbf{\Lambda}_\partial$ as the block diagonal matrix with ith block $\frac{\partial \mu_i}{\partial \theta_i^T}$ (31) can be written as

$$\mathbf{C}_{\hat{\vartheta}} = \mathbf{\Lambda}_\partial^{-1} \underline{\mathbf{C}} \mathbf{\Lambda}_\partial^{-T} \tag{32}$$

where

$$\underline{\mathbf{C}} \triangleq \begin{bmatrix} \mathbf{C}_{11} & \dots & \mathbf{C}_{1m} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{m1} & \dots & \mathbf{C}_{mm} \end{bmatrix} \tag{33}$$

Now let $\mathbf{W} = \mathbf{C}_{\hat{\vartheta}}^{-1}$ and plug (32) into (29) to obtain

$$
\begin{aligned}
\mathbf{C}_{\hat{\theta}_{DAC}} &= \left[\mathbf{S}^T (\mathbf{\Lambda}_\partial^{-1} \underline{\mathbf{C}} \mathbf{\Lambda}_\partial^{-T})^{-1} \mathbf{S} \right]^{-1} \\
&= \left[\mathbf{S}^T \mathbf{\Lambda}_\partial^T \underline{\mathbf{C}}^{-1} \mathbf{\Lambda}_\partial \mathbf{S} \right]^{-1} \\
&= \left[\frac{\partial \underline{\mu}^T}{\partial \theta} \underline{\mathbf{C}}^{-1} \frac{\partial \underline{\mu}}{\partial \theta^T} \right]^{-1}
\end{aligned} \tag{34}
$$

where

$$
\begin{aligned}
\underline{\mu} &\triangleq E\{[\mathbf{x}_1 \dots \mathbf{x}_m]^T\} \\
&= [\mu_i \dots \mu_m]^T
\end{aligned} \tag{35}
$$

is the CRLB. Thus it is shown that assuming sufficiently large Fisher information the DAC estimate approaches the CRLB if the parameter vectors $\hat{\theta}_i$ and the sub vectors $\mathbf{x}_i$ have the same length.

5.3. Application of DAC to Range Difference Location

In order to apply the DAC to the range difference location problem, at first identify $\mathbf{x}$ as the vector of measured range differences

$$\mathbf{x} \triangleq [\tilde{\Delta}_{12} \dots \tilde{\Delta}_{1N} \dots \tilde{\Delta}_{k(k+1)} \dots \tilde{\Delta}_{kN} \dots \tilde{\Delta}_{(N-2)N}]^T \tag{36}$$

In the next step we will divide $\mathbf{x}$ into $M = \binom{N}{3}$ overlapping sub vectors $\mathbf{x}_i \in \mathbb{R}^{2\times 1}$. In order to do so, at first number the receivers arbitrarily from 1 to N. Then define sub vectors as

$$\mathbf{x}_i^{(k)}(l,m) \triangleq \begin{bmatrix} \tilde{\Delta}_{kl} \\ \tilde{\Delta}_{km} \end{bmatrix}_i, \quad \text{for} \quad i = 1, \dots, M \tag{37}$$

where $k \in \{1, \dots, N-2\}$, $l \in \{k+1, \dots, N-1\}$ and $m \in \{l+1, \dots, N\}$. In order to illustrate these rather abstract definitions lets assume the case of $M = 4$ receivers. Then we have

$$\mathbf{x} \triangleq [\tilde{\Delta}_{12}\ \tilde{\Delta}_{13}\ \tilde{\Delta}_{14}\ \tilde{\Delta}_{23}\ \tilde{\Delta}_{24}]^T \tag{38}$$

and

$$\mathbf{x}_1^{(1)} \triangleq \begin{bmatrix} \tilde{\Delta}_{12} \\ \tilde{\Delta}_{13} \end{bmatrix}, \mathbf{x}_2^{(1)} \triangleq \begin{bmatrix} \tilde{\Delta}_{12} \\ \tilde{\Delta}_{14} \end{bmatrix}, \mathbf{x}_3^{(1)} \triangleq \begin{bmatrix} \tilde{\Delta}_{13} \\ \tilde{\Delta}_{14} \end{bmatrix}, \mathbf{x}_4^{(2)} \triangleq \begin{bmatrix} \tilde{\Delta}_{23} \\ \tilde{\Delta}_{24} \end{bmatrix} \tag{39}$$

Now we can collect the vectors $\mathbf{x}_i^{(k)}(l,m)$ from (37) in the vector $\underline{\mathbf{x}}$

$$\underline{\mathbf{x}} \triangleq \left[\mathbf{x}_1^{(1)} \dots \mathbf{x}_{\binom{N-1}{2}}^{(1)}\ \mathbf{x}_{\binom{N-1}{2}+1}^{(2)} \dots \mathbf{x}_{\binom{N-1}{2}+\binom{N-2}{2}}^{(2)} \dots \mathbf{x}_M^{(N-2)} \right]^T \tag{40}$$

and proceed with the identifications

$$
\begin{aligned}
\theta &\triangleq [x_0\ y_0]^T \\
\underline{\mu} &\triangleq E\{\underline{\mathbf{x}}\} \\
\underline{\mathbf{C}} &\triangleq E\{(\underline{\mathbf{x}} - \underline{\mu})(\underline{\mathbf{x}} - \underline{\mu})^T\}
\end{aligned} \tag{41}
$$

To compute the weighting matrix $\mathbf{W}$ in (13) according to section (5.2) the only thing that remains to determine is $\frac{\partial \mu_i^{(k)}}{\partial \theta^T} \triangleq \frac{\partial E\{\mathbf{x}_i^{(k)}\}}{\partial \theta^T}$. In order to do so, remember that

$$
\begin{aligned}
\Delta_{kl} &\triangleq r_l - r_k \\
&= \sqrt{(x_0 - x_l)^2 + (y_0 - y_l)^2} \\
&\quad - \sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2} \\
\Delta_{km} &\triangleq r_m - r_k \\
&= \sqrt{(x_0 - x_m)^2 + (y_0 - y_m)^2} \\
&\quad - \sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}
\end{aligned} \tag{42}
$$

which together with (41) yields

$$
\begin{aligned}
\frac{\partial \mu_i^{(k)}}{\partial \theta^T} &\triangleq \frac{\partial E\{\mathbf{x}_i^{(k)}\}}{\partial \theta^T} \\
&= \begin{bmatrix} \frac{\partial \Delta_{kl}}{\partial x_0} & \frac{\partial \Delta_{kl}}{\partial y_0} \\[6pt] \frac{\partial \Delta_{km}}{\partial x_0} & \frac{\partial \Delta_{km}}{\partial y_0} \end{bmatrix}
\end{aligned} \tag{43}
$$

with

see the program

$$
\begin{aligned}
\frac{\partial \Delta_{kl}}{\partial x_0} =& \frac{x_0 - x_l}{\sqrt{(x_0 - x_l)^2 + (y_0 - y_l)^2}} \\
& - \frac{x_0 - x_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}} \\
\frac{\partial \Delta_{kl}}{\partial y_0} =& \frac{y_0 - y_l}{\sqrt{(x_0 - x_l)^2 + (y_0 - y_l)^2}} \\
& - \frac{y_0 - y_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}} \\
\frac{\partial \Delta_{km}}{\partial x_0} =& \frac{x_0 - x_m}{\sqrt{(x_0 - x_m)^2 + (y_0 - y_m)^2}} \\
& - \frac{x_0 - x_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}} \\
\frac{\partial \Delta_{km}}{\partial y_0} =& \frac{y_0 - y_m}{\sqrt{(x_0 - x_m)^2 + (y_0 - y_m)^2}} \\
& - \frac{y_0 - y_k}{\sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2}}
\end{aligned}
\tag{44}
$$

Finally set

$$
\mathbf{S}_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{for} \quad i = 1, \dots, M \tag{45}
$$

such that we can devise the following algorithm for range difference location

Algorithm 1:

1. compute $\hat{\theta}_i$ based on the sub vectors $\mathbf{x}_i$ for $i = 1, \dots, M$ using LOCA
2. compute μ_i for $i = 1, \dots, M$ according to (43)
3. compute $\mathbf{C}_{\hat{\vartheta}}$ according to (32)
4. set $\mathbf{W} = \mathbf{C}_{\hat{\vartheta}}^{-1}$
5. compute $\hat{\theta}_{DAC} = [\hat{x}_0 \; \hat{y}_0]^T_{DAC}$ according to (13)

Note that $[x_0 \; y_0]^T$ needed in step two for the computation of μ_i for $i = 1, \dots, M$ is unknown, therefore one should use the estimate $[\hat{x}_0 \; \hat{y}_0]^T \triangleq \frac{1}{M} \sum_{i=1}^{M} \hat{\theta}_i$ instead.
The performance of algorithm 1 has been simulated in section (6).

6. SIMULATION RESULTS

In this section we will describe the simulation results. The first and second pictures show the error contour plot of a LOCA system. With increasing the number of receivers we get a better overall RMS error.

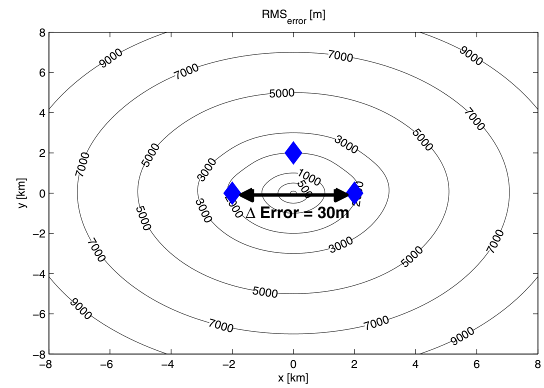


Fig. 7. Error contour plot of LOCA system with three receivers

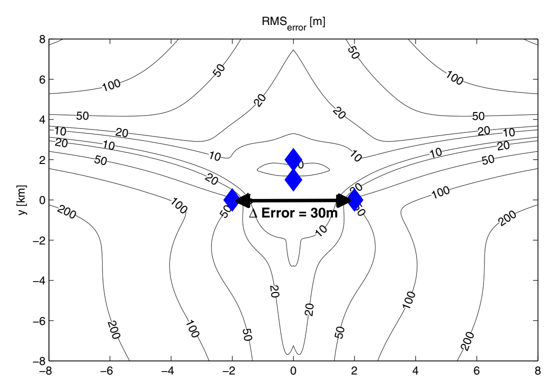


Fig. 8. Error contour plot of LOCA system with four receivers

see the program

7. CONCLUSIONS

In this work we investigated different localization methods based on the time of arrival. The time difference between two receivers was performed by a standard cross-correlation function. The simulation was influenced by COST259 channel model conditions. The main difference between the method based on intersecting hyperbolas and location on the conic axis (LOCA), is the number of necessary receivers. In the case of an ellipse used the LOCA method, it is enough to use three receivers for an ambiguity result. But with increasing the number of receivers we reduce the RMS error over the whole plane. The divide and conquer method gives a good approach to combine fixes.

8. REFERENCES

[1] Luis M. Correia et.al. *Wireless Flexilble Personalised Communications - COST259*. Wiley, 2001.

[2] R. O. Schmidt *A new approach to geometry of range difference location.* In IEEE Trans. Aerosp. Electron. Syst., vol. 8, no. 6, pp. 821-835,Nov. 1972.

[3] J. S. Abel. *A divide and conquer approach to least-squares estimation.* In IEEE Trans. Aerosp. Electron. Syst., vol. 26, no. 2, pp. 423-427

lit.dac

see the program

AN EIGENFILTER APPROACH TO THE DESIGN OF FREQUENCY INVARIANT BEAMFORMERS

Yong Zhao, Wei Liu and Richard Langley

Communications Research Group
Department of Electronic and Electrical Engineering
University of Sheffield, UK

Abstract. An eigenfilter approach to the design of frequency invariant beamformers (FIBs) is proposed and its solution is provided by finding the minimum generalized eigenvector of two matrices. Several design examples are provided with satisfactory frequency invariant property and sidelobe attenuation. The proposed approach is general and can be applied to different array structures, such as linear arrays, circular arrays and rectangular arrays.

1. INTRODUCTION

Broadband beamforming has been studied extensively in the past and amongst them is a class of beamformers with frequency invariant responses [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12], which can form beams pointing to the signal of interest with a constant beamwidth. With the FIB technique, both adaptive broadband beamforming and broadband DOA (direction of arrival) estimation can be simplified greatly [13, 4, 14, 15]. In [4, 9, 10, 11, 12], the design is achieved based on simple multi-dimensional inverse Fourier transforms by exploiting the relationship between the array's spatial and temporal parameters and its beam pattern. More recently, a direct optimization approach was adopted using the convex optimization methods [16, 17, 18].

The least squares approach is a conventional and well-known method for the design of both FIR filters and broadband beamformers [19, 20, 21, 22, 23] and its solutions are normally obtained by matrix inversion based methods and eigenfilter based methods. Compared with the matrix inversion based methods, the eigenfilter based ones have the advantage of improved numerical stability. In this paper, we will extend the traditional eigenfilter based methods to the design of frequency invariant beamformers. The key to its success is to form a cost function enforcing the frequency invariance requirement in the design and one novel unconstrained least squares(ULS) formulation of the problem is proposed, with its solution based on the eigenfilter method and obtained by finding the minimum generalized eigenvector of two related matrices.

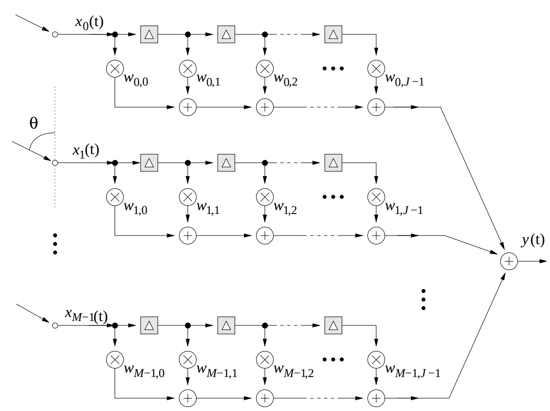


Fig. 1. A general broadband beamforming structure.

This paper is organized as follows. The general broadband beamforming structures are reviewed briefly in Section 2. The ULS formulation for the design of FIBs with its corresponding solution is given in Section 3. Several design examples based on different array structures are provided in Section 4. Conclusions are drawn in Section 5.

2. BROADBAND BEAMFORMING STRUCTURES

A general structure for broadband beamforming with tapped delay-lines (TDLs) or FIR/IIR filters is shown in Fig. 1, in which J is the number of delay elements associated with each

see the program

of the M sensor channels. The beamformer with such a structure samples the propagating wave field in both space and time. Its response as a function of the signal angular frequency ω and the direction of arrival angle θ can be expressed as

$$\tilde{R}(\omega,\theta)=\sum_{m=0}^{M-1}\sum_{k=0}^{J-1}w_{m,k}e^{-j\omega(\tau_m+kT_s)}, \qquad (1)$$

where T_s is the delay between adjacent taps of the TDLs and τ_m is the spatial propagation delay between the $m-th$ sensor and a reference point. We can rewrite the response in a vector form

$$\tilde{R}(\omega,\theta)=\mathbf{w}^T\mathbf{s}(\omega,\theta) \qquad (2)$$

where $\mathbf{w}$ is the coefficient vector defined as

$$\mathbf{w}=[w_{0,0},\cdots w_{M-1,0}\cdots w_{0,J-1}\cdots w_{M-1,J-1}]^T\ , \qquad (3)$$

and $\mathbf{s}(\omega,\theta)$ is the $M\times J$ steering vector given by

$$\mathbf{s}(\omega,\theta)=\mathbf{s}_{T_s}(\omega)\otimes\mathbf{s}_{\tau_m}(\omega,\theta) \qquad (4)$$

where $\otimes$ denotes the Kronecker product, and

$$\mathbf{s}_{T_s}(\omega)=[1,e^{-j\omega T_s},\cdots,e^{-j(J-1)\omega T_s}]^T\ , \qquad (5)$$

$$\mathbf{s}_{\tau_m}(\omega,\theta)=[e^{-j\omega\tau_0},e^{-j\omega\tau_1},\cdots,e^{-j\tau_{M-1}\omega}]^T\ . \qquad (6)$$

For a uniformly spaced linear array with an inter-element spacing d, we have the spatial propagation delay τ_m given by $\tau_m=m\tau_1=m\frac{d}{c}\cos\theta$, with the first sensor position as the phase reference point. With the normalized angular frequency $\Omega=\omega T_s$, $\omega(m\tau_1+kT_s)$ changes to $m\mu\Omega\cos\theta+k\Omega$ with $\mu=\frac{d}{cT_s}$. Then the steering vector changes to

$$\mathbf{s}(\Omega,\theta)=\mathbf{s}_{T_s}(\Omega)\otimes\mathbf{s}_{\tau_m}(\Omega,\theta) \qquad (7)$$

and

$$\mathbf{s}_{T_s}(\Omega)=[1,e^{-j\Omega},\cdots,e^{-j(J-1)\Omega}]^T \qquad (8)$$

$$\mathbf{s}_{\tau_m}(\Omega,\theta)=[1,e^{-j\mu\Omega\cos\theta},\cdots,e^{-j(M-1)\mu\Omega\cos\theta}]^T\ . \qquad (9)$$

Then we obtain the response as a function of Ω and θ

$$R(\Omega,\theta)=\mathbf{w}^T\mathbf{s}(\Omega,\theta)\ . \qquad (10)$$

For a uniform circular array with a circumferential spacing of d, as shown in Fig. 2, we have the spatial delay between the center of the circular array and the $m-th$ sensor, $\tau_m=\frac{r\cos(\theta-\theta_m)}{c}$, where $r=\frac{Md}{2\pi}$ is the radius of the circular array and $\theta_m=m\frac{2\pi}{M}$ is the angle of the $m-th$ sensor. With the normalized angular frequency $\Omega=\omega T_s$, $\omega(\tau_m+kT_s)$ changes to $\Omega\beta\mu\cos(\theta-\theta_m)+k\Omega$ with $\beta=\frac{M}{2\pi}$. Then we obtain $\mathbf{s}_{\tau_m}$ for the uniform circular array case

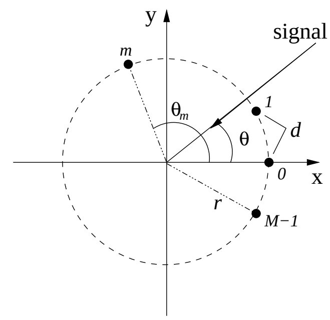


Fig. 2. A uniformly spaced circular array.

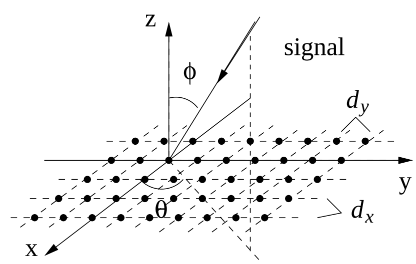


Fig. 3. A uniformly spaced rectangular array, which can be considered as a linear array with SDLs.

$$\mathbf{s}_{\tau_m}(\Omega,\theta)\ =\ [e^{-j\Omega\beta\mu\cos(\theta-\theta_0)},e^{-j\Omega\beta\mu\cos(\theta-\theta_1)},\cdots,$$
$$e^{-j\Omega\beta\mu\cos(\theta-\theta_{M-1})}]^T\ . \qquad (11)$$

When the TDLs in the broadband linear array system shown in Fig. 1 are replaced by sensor delay-lines (SDLs), we will have a rectangular array system without TDLs [24]. Such a structure is shown in Fig. 3, which is an equally spaced rectangular array with $M\times J$ sensors and an inter-element spacing of d_x and d_y, respectively. Without loss of generality, we assume $d_x=d_y=d$ and the signals come from the direction $\varphi=\frac{\pi}{2}$. Since it is a broadband beamforming structure with spatial-only information, we replace kT_s in (1) by δ_k which represents the spatial propagation delay along the y axis and is given by $\delta_k=k\delta_1=m\frac{d}{c}\sin\theta$, with the first sensor position as the phase reference point. With the normalized angular frequency $\Omega=\omega T_s$, $\omega(m\tau_1+k\delta_1)$ changes to $m\mu\Omega\cos\theta + k\mu\Omega\sin\theta$. Then we obtain the steering vector for the rectangular array as

$$\mathbf{s}(\Omega,\theta) = \mathbf{s}_{\delta_k}(\Omega) \otimes \mathbf{s}_{\tau_m}(\Omega,\theta) \tag{12}$$

where

$$\mathbf{s}_{\delta_k}(\Omega,\theta) = [1, e^{-j\mu\Omega\sin\theta}, \cdots, e^{-j(J-1)\mu\Omega\sin\theta}]^T . \tag{13}$$

3. THE EIGENFILTER APPROACH TO THE DESIGN OF FIBS

In the proposed method, the cost function for the design of FIBs consists of three parts. Firstly, the frequency invariance property is formulated by minimizing the Euclidean distance between the response at a fixed reference frequency Ω_r and those at all the other operating frequencies over a range of directions in which frequency invariance is considered. The cost function related to this property is given by

$$J_1 = \int_\Omega \int_{\Theta_{FI}} \left|\mathbf{w}^T\mathbf{s}(\Omega,\theta) - \mathbf{w}^T\mathbf{s}(\Omega_r,\theta)\right|^2 d\Omega d\theta \tag{14}$$

where Θ_{FI} represents a direction range in which frequency invariance is considered. It can be either the main beam direction area or the whole angle range from $0°$ to $180°$ for the linear and planar arrays and from $-180°$ to $180°$ for the circular array. Here we will always consider the full angle range.

Secondly, we also need to minimize the response of the beamformer at the reference frequency Ω_r over the sidelobe region. This part of the cost function is defined as

$$J_2 = \int_{\Theta_s} \left|\mathbf{w}^T\mathbf{s}(\Omega_r,\theta)\right|^2 d\theta \tag{15}$$

where Θ_s denotes the sidelobe region.

Thirdly, we need to maximize the spectrum energy at the reference frequency Ω_r over the mainlobe region.

Then, a complete formulation of the FIB design problem is obtained by combining the three elements together

$$J_{ULS} = \frac{\int_\Omega \int_{\Theta_{FI}} \left|\mathbf{w}^T\mathbf{s}(\Omega,\theta) - \mathbf{w}^T\mathbf{s}(\Omega_r,\theta)\right|^2 d\Omega d\theta + \alpha \int_{\Theta_s} \left|\mathbf{w}^T\mathbf{s}(\Omega_r,\theta)\right|^2 d\theta}{\int_{\Theta_m} \left|\mathbf{w}^T\mathbf{s}(\Omega_r,\theta)\right|^2 d\Omega d\theta} \tag{16}$$

where α is a trade off parameter between the frequency invariant property and the sidelobe attenuation.

Now let (Ω_n, θ_k) be the grid chosen from the continuous frequency and angle ranges above. Then the least squares formulation for the FIB design becomes

$$J_{ULS} = \frac{\sum_{n=0}^{N-1}\sum_{k=0}^{K-1} \left|\mathbf{w}^T\mathbf{s}(\Omega_n,\theta_k) - \mathbf{w}^T\mathbf{s}(\Omega_r,\theta_k)\right|^2 + \alpha \sum_{\theta_k\in\Theta_s} \left|\mathbf{w}^T\mathbf{s}(\Omega_r,\theta_k)\right|^2}{\sum_{\theta_k\in\Theta_m} \left|\mathbf{w}^T\mathbf{s}(\Omega_r,\theta_k)\right|^2} \tag{17}$$

where N and K are the number of samples chosen over the frequency and the angle ranges considered for frequency invariance, respectively. Note that the spectrum energy used here is just the summation of energy over the mainlobe at the reference frequency θ_r, which leads to a much more reduced computational complexity than the conventional methods in the design of broadband beamformers [22]. We can rewrite (17) as

$$J_{ULS} = \frac{\mathbf{w}^T\mathbf{Q}_1\mathbf{w}}{\mathbf{w}^T\mathbf{Q}_2\mathbf{w}} \tag{18}$$

where

$$\begin{aligned}\mathbf{Q}_1 = &\sum_{n=0}^{N-1}\sum_{k=0}^{K-1} |\mathbf{s}(\Omega_n,\theta_k) - \mathbf{s}(\Omega_r,\theta_k)|^2 \\ &+ \alpha \sum_{\theta_k\in\Theta_s} \mathbf{s}(\Omega_r,\theta_k)\mathbf{s}(\Omega_r,\theta_k)^H\end{aligned} \tag{19}$$

and

$$\mathbf{Q}_2 = \sum_{\theta_k\in\Theta_m} \mathbf{s}(\Omega_r,\theta_k)\mathbf{s}(\Omega_r,\theta_k)^H \tag{20}$$

As $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are clearly Hermitian, we can obtain $\mathbf{w}^T\mathbf{Q}_1\mathbf{w} = \mathbf{w}^T\mathbf{Q}_{1R}\mathbf{w}$ and $\mathbf{w}^T\mathbf{Q}_2\mathbf{w} = \mathbf{w}^T\mathbf{Q}_{2R}\mathbf{w}$, where $\mathbf{Q}_{1R}$ and $\mathbf{Q}_{2R}$ are the real symmetric matrices of $\mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively. Thus Equation (18) changes to

$$J_{ULS} = \frac{\mathbf{w}^T\mathbf{Q}_{1R}\mathbf{w}}{\mathbf{w}^T\mathbf{Q}_{2R}\mathbf{w}} \tag{21}$$

Thus the optimal $\mathbf{w}$ can be obtained by finding the generalized eigenvector corresponding to the minimum eigenvalue of $\mathbf{Q}_{1R}$ and $\mathbf{Q}_{2R}$. Note that the optimum weight vector obtained in this way will not necessarily have a unity response at the look direction θ_r. To have the desired unity response, we need to normalize the resultant weight vector $\mathbf{w}$. To avoid this normalisation and also have a direct control of the array response to some specific directions, such as a unity response to the desired signal direction, we can follow the total least squares (TLS) approach in the design of broadband beamformers [22] and obtain the following formulation of the problem

$$J_{CTLS} = \frac{\mathbf{w}^T\mathbf{Q}_{1R}\mathbf{w}}{\mathbf{w}^T\mathbf{Q}_{2R}\mathbf{w}+1} \quad \text{subject to } \mathbf{C}^T\mathbf{w} = \mathbf{f}, \tag{22}$$

see the program

where $\mathbf{C}^T\mathbf{w} = \mathbf{f}$ provides the desired constraints on the array coefficients. This formulation can be transformed to

$$J_{CTLS} = \frac{\hat{\mathbf{w}}^T\hat{\mathbf{Q}}_1\hat{\mathbf{w}}}{\hat{\mathbf{w}}^T\hat{\mathbf{Q}}_2\hat{\mathbf{w}}} \text{ subject to } \hat{\mathbf{C}}\hat{\mathbf{w}} = \mathbf{0} \tag{23}$$

with $\hat{\mathbf{w}}$, $\hat{\mathbf{Q}}_1$, $\hat{\mathbf{Q}}_2$ and $\hat{\mathbf{C}}$ defined as

$$\begin{aligned} \hat{\mathbf{w}} &= [\mathbf{w}^T, -1]^T, \hat{\mathbf{Q}}_1 = \left[\begin{array}{cc} \mathbf{Q}_{1R} & \mathbf{0} \\ \mathbf{0} & 0 \end{array}\right], \\ \hat{\mathbf{Q}}_2 &= \left[\begin{array}{cc} \mathbf{Q}_{2R} & \mathbf{0} \\ \mathbf{0} & 1 \end{array}\right], \hat{\mathbf{C}} = [\mathbf{C}^T, \mathbf{f}] \end{aligned} \tag{24}$$

The solution can be derived in a similar way as in [22].

A problem with this formulation is its high computational complexity and in the design examples, we will only use the formulation given in (21).

4. DESIGN EXAMPLES

To show the effectiveness of the proposed method, we provide several design examples based on a uniformly spaced linear array, a uniformly spaced circular array, and a uniformly spaced rectangular array, respectively.

4.1. Uniformly Spaced Linear Array

We apply the proposed method to a uniformly spaced linear array to design FIBs with a broadside main beam, i.e., a look direction $\theta_r = 90°$. The design is based on a linear array with 14 sensors and each followed by an 18-tap FIR filter. The array spacing is assumed to be half the wavelength corresponding to the maximum normalized signal frequency π so that $\mu = 1$. The frequency range of interest in the design is $[0.3\pi, 0.9\pi]$ and the mainlobe region is $[75°, 105°]$. The fixed reference frequency point is $\Omega_r = 0.7\pi$ and α is set to be 0.05.

The resultant beam pattern with the ULS formulation is shown in Fig. 4 which exhibits a good frequency invariant property over the frequency range $[0.3\pi, 0.9\pi]$ with a satisfactory sidelobe attenuation.

4.2. Uniformly Spaced Circular Array

The proposed least squares approach is general and can be applied to other array structures too. In the following we give an example based on a uniformly spaced circular array with

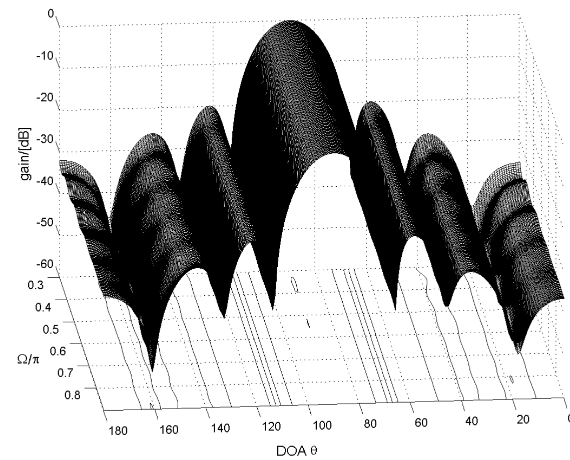


Fig. 4. The resultant beam pattern using the ULS formulation for a main beam direction at $\theta = 90°$.

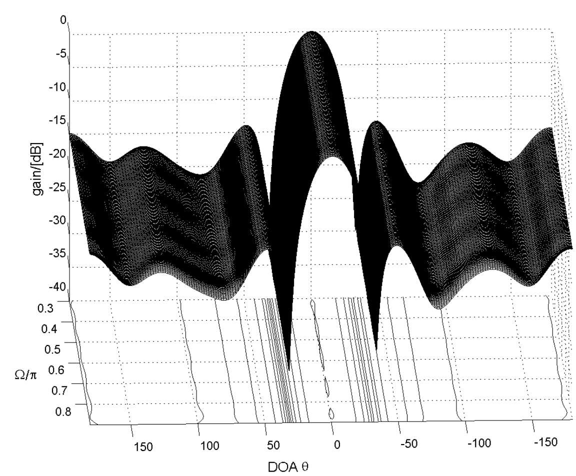


Fig. 5. The resultant beam pattern using the ULS formulation for the uniform circular array.

see the program

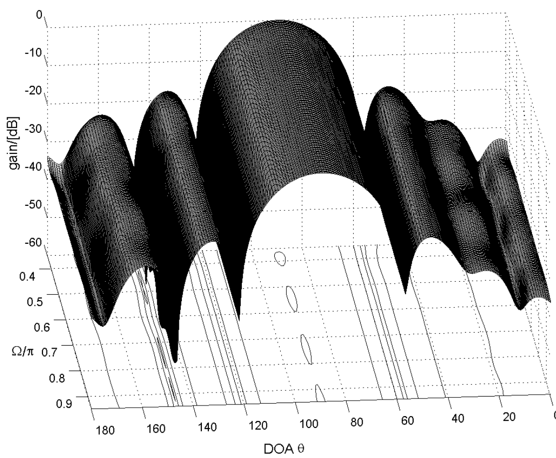


Fig. 6. The resultant beam pattern using the ULS formulation for the uniform rectangular array with a main beam direction at $\theta = 90°$.

14 sensors and each followed by a 20−tap FIR filter. The look direction is $\theta_r = 0°$ and the mainlobe region is $[-30°, 30°]$. All of the other parameters are the same as in the linear array case except that α is set to be 0.1.

The resultant beam pattern is shown in Fig. 5. Again a good frequency invariant property is achieved over the frequency range of interest.

4.3. Uniformly Spaced Rectangular Array

Finally, we apply the proposed method to a uniformly spaced rectangular array without TDLs to design a broadside main beam. Except for the operating frequency range which changes to $[0.35\pi, 0.95\pi]$, all of the other parameters are the same as in the linear array case. The resultant beam pattern is shown in Fig. 6, with a good frequency invariance property and a satisfactory sidelobe performance.

5. CONCLUSION

A novel least squares formulation to the design of frequency invariant beamformers with an eigenfilter based solution has been proposed, which leads to satisfactory design results as demonstrated by the provided design examples. The proposed approach is general and can be applied to different array structures, such as linear arrays, circular arrays and rectangular arrays.

6. REFERENCES

[1] M. M. Goodwin and G. W. Elko, "Constant beamwidth beamforming," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, USA, April 1993, vol. 1, pp. 169–172.

[2] T. Chou, "Frequency-independent beamformer with low response error," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Detroit, USA, May 1995, vol. 5, pp. 2995–2998.

[3] D. B. Ward, R. A. Kennedy, and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," *Journal of the Acoustic Society of America*, vol. 97, no. 2, pp. 1023–1034, February 1995.

[4] T. Sekiguchi and Y. Karasawa, "Wideband beamspace adaptive array utilizing FIR fan filters for multibeam forming," *IEEE Transactions on Signal Processing*, vol. 48, no. 1, pp. 277–284, January 2000.

[5] L. C. Parra, "Steerable frequency-invariant beamforming for arbitrary arrays," *Journal of the Acoustic Society of America*, vol. 119, pp. 3839–3847, June 2006.

[6] H. H. Chen and S. C. Chan, "Adaptive beamforming and DOA estimation using uniform concentric spherical arrays with frequency invariant characteristics," *Journal of VLSI Signal Processing*, vol. 46, no. 1, pp. 15–34, January 2007.

[7] H. H. Chen, S. C. Chan, and K. L. Ho, "Adaptive beamforming using frequency invariant uniform concentric circular arrays," *IEEE Transactions on Circuits & Systems I: Regular Papers*, vol. 54, no. 9, pp. 1938–1949, September 2007.

[8] S. C. Chan and H. H. Chen, "Uniform concentric circular arrays with frequency-invariant characteristics– theory, design, adaptive beamforming and DOA estimation," *IEEE Transactions on Signal Processing*, vol. 55, pp. 165–177, January 2007.

[9] W. Liu, S. Weiss, J. G. McWhirter, and I. K. Proudler, "Frequency invariant beamforming for two-dimensional and three-dimensional arrays," *Signal Processing*, vol. 87, pp. 2535–2543, November 2007.

see the program

[10] W. Liu and S. Weiss, "Design of frequency invarint beamformers for broadband arrays," *IEEE Trans. Signal Processing*, vol. 56, no. 2, pp. 860, February 2008.

[11] W. Liu, D. McLernon, and M. Ghogho, "Design of frequency invariant beamformer without temporal filtering," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 798–802, February 2009.

[12] W. Liu and S. Weiss, "Off-broadside main beam design and subband implementation for a class of frequency invariant beamformers," *Signal Processing*, vol. 89, pp. 913–920, May 2009.

[13] D. B. Ward, Z. Ding, and R. A. Kennedy, "Broadband DOA estimation using frequency invariant beamforming," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1463–1469, May 1998.

[14] W. Liu and D. P. Mandic, "Semi-blind source separation for convolutive mixtures based on frequency invariant transformation," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, USA, March 2005, vol. 5, pp. 285–288.

[15] W. Liu, R. Wu, and R. Langley, "Design and analysis of broadband beamspace adaptive arrays," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 12, pp. 3413–3420, December 2007.

[16] S.F. Yan and Y. L. Ma, "Design of FIR beamformer with frequency invariant patterns via jointly optimizing spatial and frequency responses," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, USA, March 2005, pp. 789–792.

[17] H. Duan, B. P. Ng, C. M. See, and J. Fang, "Applications of the SRV constraint in broadband pattern synthesis," *Signal Processing*, vol. 88, pp. 1035–1045, April 2008.

[18] Y. Zhao, W. Liu, and R. J. Langley, "Efficient design of frequency invariant beamformers with sensor delay-lines," in *Proc. IEEE Workshop on Sensor Array and Multichannel Signal Processing*, Darmstadt, Germany, July 2008, pp. 335–339.

[19] P. Vaidyanathan and T. Q. Nguyen, "Eigenfilters: A new approach to least-squares FIR filter design and applications including Nyquist filters," *IEEE Transactions on Circuits & Systems*, vol. 34, pp. 11–23, January 1987.

[20] S. C. Pei and C. C. Tseng, "A new eigenfilter based on total least squares error criterion," *IEEE Transactions on Circuits & Systems I: Regular Papers*, vol. 48, pp. 699–709, 2001.

[21] A. Tkacenko, P. P. Vaidyanathan, and T. Q. Nguyen, "On the eigenfilter design method and its applications: a tutorial," *IEEE Transactions on Circuits and Systems — II: Analog and Digital Signal Processing*, vol. 50, pp. 497–517, September 2003.

[22] S. Doclo and M. Moonen, "Design of far-field and near-field broadband beamformers using eigenfilters," *Signal Processing*, vol. 83, no. 12, pp. 2641–2673, December 2004.

[23] Y. Zhao, W. Liu, and R. J. Langley, "A least squares approach to the design of frequency invariant beamformers," *submitted to IEEE Trans. Signal Processing*, December 2008.

[24] W. Liu, "Adaptive wideband beamforming with sensor delay-lines," *Signal Processing*, vol. 89, pp. 876–882, May 2009.

see the program

Author Index