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

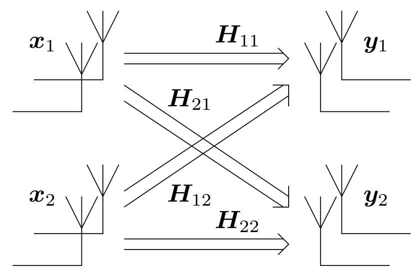

Fig. 1. A MIMO IFC with two Antennas at Each Terminal

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.

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.

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 $\mathrm{E}[\boldsymbol{\eta}_i \boldsymbol{\eta}_i^{\mathrm{H}}] = \sigma^2 \boldsymbol{1}$, 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 $\boldsymbol{R}_i = \boldsymbol{H}_{ij} \boldsymbol{Q}_j \boldsymbol{H}_{ij}^{\mathrm{H}} + \sigma^2 \boldsymbol{1}$, 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]

$$R_i = \log\det\left(\boldsymbol{1} + \boldsymbol{H}_{ii} \boldsymbol{Q}_i \boldsymbol{H}_{ii}^{\mathrm{H}} \boldsymbol{R}_i^{-1}\right) \quad (1)$$

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.

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.

[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 $\log 2$.

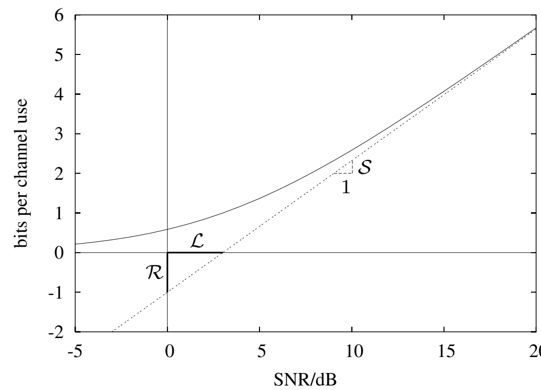

Fig. 2. Graphical Interpretation of the High SNR Slope $\mathcal{S}$, the High SNR Power Offset $\mathcal{L}$ and the High SNR Rate Offset $\mathcal{R}$

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 $\log(\mathrm{SNR})$-versus-rate curve at high SNR and is also referred to as *degrees of freedom* (DOF) or *multiplexing gain*:

$$\mathcal{S} = \lim_{\rho\to\infty} \frac{R(\rho)}{\log\rho}, \quad (2)$$

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 bits / 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(SNR)-versus-rate coordinate system and the point where the high SNR asymptote of the rate curve intersects the log(SNR)-axis. We can see from Fig. 2 that [6]

$$\mathcal{L} = \lim_{\rho\to\infty}\left(\log\rho - \frac{R(\rho)}{\mathcal{S}}\right). \quad (3)$$

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

$$\mathcal{L}_i = -{\mathcal{S}_i}^{-1}\log\lim_{\sigma^2\to 0}\left(\left(\sigma^2\right)^{\mathcal{S}_i} D_i\right) \quad (4)$$

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) $\boldsymbol{I}_i = \boldsymbol{U}_i\boldsymbol{\Lambda}_i\boldsymbol{U}_i^{\mathrm{H}}$:

$$\left(\sigma^2\mathbf{1} + \boldsymbol{I}_i\right)^{-1} = \sigma^{-2}\left(\mathbf{1} - \boldsymbol{I}_i\left(\sigma^2\mathbf{1}{+}\boldsymbol{I}_i\right)^{-1}\right) \quad (5)$$

$$= \sigma^{-2}\left(\mathbf{1} - \boldsymbol{U}_i\boldsymbol{\Lambda}_i\left(\sigma^2\mathbf{1}{+}\boldsymbol{\Lambda}_i\right)^{-1}\boldsymbol{U}_i^{\mathrm{H}}\right) \quad (6)$$

Now we can calculate the behaviour for vanishing noise:

$$\lim_{\sigma^2\to 0}\boldsymbol{U}_i\boldsymbol{\Lambda}_i\left(\sigma^2\mathbf{1} + \boldsymbol{\Lambda}_i\right)^{-1}\boldsymbol{U}_i^{\mathrm{H}} = \boldsymbol{U}_i\boldsymbol{J}_i\boldsymbol{U}_i^{\mathrm{H}} \quad (7)$$

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:

$$\mathcal{R}_i = \log\det\left(\boldsymbol{T}_i^{\mathrm{H}}\boldsymbol{H}_{ii}^{\mathrm{H}}(\mathbf{1} - \boldsymbol{U}_i\boldsymbol{J}_i\boldsymbol{U}_i^{\mathrm{H}})\boldsymbol{H}_{ii}\boldsymbol{T}_i\right) \quad (8)$$

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 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

$$\boldsymbol{H}'_{ii} = \boldsymbol{B}_i^{\mathrm{H}}\boldsymbol{H}_{ii} \qquad \boldsymbol{H}'_{ij} = \boldsymbol{B}_i^{\mathrm{H}}\boldsymbol{H}_{ij}\boldsymbol{B}_j$$
$$\boldsymbol{H}'_{ji} = \boldsymbol{H}_{ji} \qquad \boldsymbol{H}'_{jj} = \boldsymbol{H}_{jj}\boldsymbol{B}_j.$$

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:

$$\boldsymbol{h}_{ii}'^{\mathrm{H}} = \boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ii} \qquad \boldsymbol{h}_{ij}'^{\mathrm{H}} = \boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ij},$$

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

$$\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{12}\boldsymbol{t}_2 = 0 \qquad \boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{21}\boldsymbol{t}_1 = 0.$$

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]:

$$\boldsymbol{t}_i = \frac{\boldsymbol{P}_{ji}^{\perp}\boldsymbol{h}'_{ii}}{\sqrt{\boldsymbol{h}_{ii}'^{\mathrm{H}}\boldsymbol{P}_{ji}^{\perp}\boldsymbol{h}'_{ii}}} \quad \text{with} \quad \boldsymbol{P}_{ji}^{\perp} = \mathbf{1} - \frac{\boldsymbol{h}'_{ji}\boldsymbol{h}_{ji}'^{\mathrm{H}}}{\boldsymbol{h}_{ji}'^{\mathrm{H}}\boldsymbol{h}'_{ji}}. \tag{9}$$

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

$$\mathcal{R}_i = \log\left(\boldsymbol{t}_i^{\mathrm{H}}\boldsymbol{h}'_{ii}\boldsymbol{h}_{ii}'^{\mathrm{H}}\boldsymbol{t}_i\right) = \log\left(\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{H}_{ii}\boldsymbol{P}_{ji}^{\perp}\boldsymbol{H}_{ii}^{\mathrm{H}}\boldsymbol{g}_i\right) \tag{10}$$
$$\text{with} \quad \boldsymbol{P}_{ji}^{\perp} = \mathbf{1} - \frac{\boldsymbol{H}_{ji}^{\mathrm{H}}\boldsymbol{g}_j\boldsymbol{g}_j^{\mathrm{H}}\boldsymbol{H}_{ji}}{\boldsymbol{g}_j^{\mathrm{H}}\boldsymbol{H}_{ji}\boldsymbol{H}_{ji}^{\mathrm{H}}\boldsymbol{g}_j}$$

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 $\boldsymbol{H}_{ii}$.

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

$$\boldsymbol{t}_i = \frac{\boldsymbol{P}\boldsymbol{h}'^*_{ji}}{\sqrt{\boldsymbol{h}'^{\mathrm{H}}_{ji}\boldsymbol{h}'_{ji}}} \qquad \text{with} \qquad \boldsymbol{P} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{11}$$

This allows us to rewrite the high SNR sum rate offset:

$$\begin{aligned}\mathcal{R} =& \log\left(\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{11}\boldsymbol{P}^{\mathrm{H}}\left(\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{g}_2\boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{21}\right)^*\boldsymbol{P}\boldsymbol{H}_{11}^{\mathrm{H}}\boldsymbol{g}_1\right)\\ &+\log\left(\boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{22}\boldsymbol{P}^{\mathrm{H}}\left(\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{12}\right)^*\boldsymbol{P}\boldsymbol{H}_{22}^{\mathrm{H}}\boldsymbol{g}_2\right)\\ &-\log\left(\boldsymbol{g}_2^{\mathrm{H}}\boldsymbol{H}_{21}\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{g}_2\right)-\log\left(\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{H}_{12}\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{g}_1\right)\end{aligned} \tag{12}$$

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

$$\boldsymbol{D}_{ii} = \boldsymbol{M}_{ii}^{\mathrm{H}}\boldsymbol{g}_j^*\boldsymbol{g}_j^{\mathrm{T}}\boldsymbol{M}_{ii} \qquad \boldsymbol{M}_{ii} = \boldsymbol{H}_{ji}^*\boldsymbol{P}\boldsymbol{H}_{ii}^{\mathrm{H}} \tag{15}$$

$$\boldsymbol{D}_{ij} = \boldsymbol{M}_{ij}^{\mathrm{H}}\boldsymbol{g}_j^*\boldsymbol{g}_j^{\mathrm{T}}\boldsymbol{M}_{ij} \qquad \boldsymbol{M}_{ij} = \boldsymbol{H}_{jj}^*\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}_{ij}^{\mathrm{H}} \tag{16}$$

$$\boldsymbol{K}_i = \boldsymbol{H}_{ij}\boldsymbol{H}_{ij}^{\mathrm{H}}.$$

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 $\alpha'_i = \pi - \alpha_i \in [0, \pi]$.

$$0 = \Re\left\{\frac{-|t_i|^2\left([\boldsymbol{M}_{ii}]_{11}^* + [\boldsymbol{M}_{ii}]_{21}^*\, t_j\right) + t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^*\, t_j\right)}{\left([\boldsymbol{M}_{ii}]_{11}^* + [\boldsymbol{M}_{ii}]_{21}^*\, t_j\right) + t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^*\, t_j\right)}\right\} + \Re\left\{\frac{-|t_i|^2\left([\boldsymbol{M}_{ij}]_{11}^* + [\boldsymbol{M}_{ij}]_{21}^*\, t_j\right) + t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^*\, t_j\right)}{\left([\boldsymbol{M}_{ij}]_{11}^* + [\boldsymbol{M}_{ij}]_{21}^*\, t_j\right) + t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^*\, t_j\right)}\right\} - \frac{|t_i|^2(k_{i,22} - k_{i,11}) + (1-|t_i|^2)\Re\left\{k_{i,12}^* t_i\right\}}{\left(k_{i,11} + k_{i,22}|t_i|^2 + 2\Re\left\{k_{i,12}^* t_i\right\}\right)} \quad (21)$$

$$0 = \Re\left\{\frac{\mathrm{j}\, t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^*\, t_j\right)}{\left([\boldsymbol{M}_{ii}]_{11}^* + [\boldsymbol{M}_{ii}]_{21}^*\, t_j\right) + t_i\left([\boldsymbol{M}_{ii}]_{12}^* + [\boldsymbol{M}_{ii}]_{22}^*\, t_j\right)}\right\} + \Re\left\{\frac{\mathrm{j}\, t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^*\, t_j\right)}{\left([\boldsymbol{M}_{ij}]_{11}^* + [\boldsymbol{M}_{ij}]_{21}^*\, t_j\right) + t_i\left([\boldsymbol{M}_{ij}]_{12}^* + [\boldsymbol{M}_{ij}]_{22}^*\, t_j\right)}\right\} - \frac{\Re\left\{\mathrm{j}\, k_{i,12}^* t_i\right\}}{k_{i,11} + k_{i,22}|t_i|^2 + 2\Re\left\{k_{i,12}^* t_i\right\}} \quad (22)$$

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:

$$\max_{\alpha_i, \beta_i} \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{D}_i \boldsymbol{g}_i \quad \text{with} \quad \boldsymbol{D}_i = \boldsymbol{M}_i^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{M}_i, \quad \boldsymbol{M}_i = \boldsymbol{H}_{ji}^* \boldsymbol{P} \boldsymbol{H}_{ii}^{\mathrm{H}}. \quad (23)$$

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):

$$\boldsymbol{D}_1 = \boldsymbol{M}_1^{\mathrm{H}} \boldsymbol{g}_2^* \boldsymbol{g}_2^{\mathrm{T}} \boldsymbol{M}_1 = \mu_2' \boldsymbol{M}_1^{\mathrm{H}} \boldsymbol{D}_2^* \boldsymbol{M}_1 = \mu_1' \mu_2' \boldsymbol{M}_1^{\mathrm{H}} \boldsymbol{M}_2^{\mathrm{T}} \boldsymbol{D}_1 \boldsymbol{M}_2^* \boldsymbol{M}_1 =: \mu \boldsymbol{M}^{\mathrm{H}} \boldsymbol{D}_1 \boldsymbol{M} \quad (24)$$

[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.

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:

$$[\boldsymbol{M}]_{21} d_1^2 + ([\boldsymbol{M}]_{11} - [\boldsymbol{M}]_{22}) d_1 - [\boldsymbol{M}]_{12} = 0 \quad (25)$$

From this quadratic equation we can explicitly calculate two solutions for d_1.[5] For d_2 we then get

$$d_2 = \frac{[\boldsymbol{D}_2]_{12}}{[\boldsymbol{D}_2]_{11}} = \frac{\left[\boldsymbol{M}_2^{\mathrm{H}} \boldsymbol{D}_1^* \boldsymbol{M}_2\right]_{12}}{\left[\boldsymbol{M}_2^{\mathrm{H}} \boldsymbol{D}_1^* \boldsymbol{M}_2\right]_{11}}. \quad (26)$$

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.

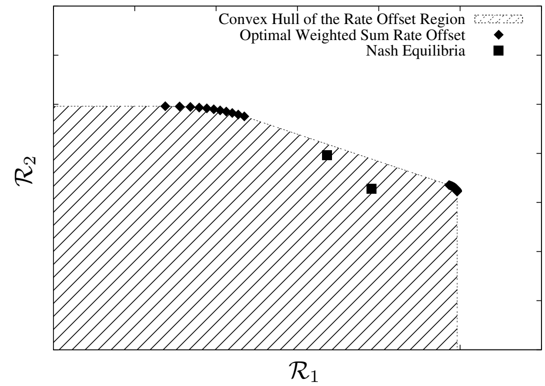

Fig. 3. Example for the Approximation of the Convex Hull of a High SNR Rate Offset Region.

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

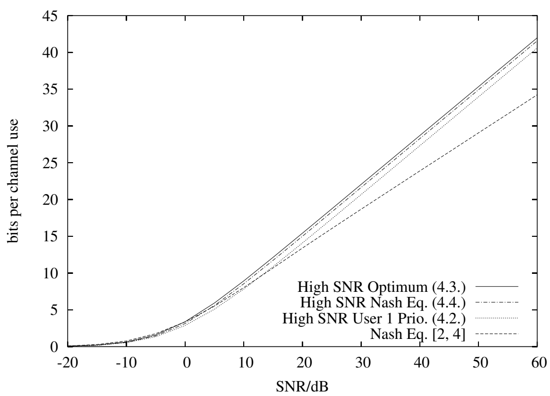

Fig. 4. Sum Rate Performance of the High SNR Solutions in a Two Stream MIMO IFC.

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.