ADAPTIVE CHANNEL DIRECTION QUANTIZATION FOR FREQUENCY SELECTIVE CHANNELS

Stefan Schwarz and Markus Rupp

Institute of Telecommunications, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
Email: {sschwarz, mrupp}@nt.tuwien.ac.at

ABSTRACT

In this work, we consider the calculation of channel state information (CSI) feedback for frequency selective OFDM based multiple input single output (MISO) wireless communication systems. We focus on the quantization and feedback of the normalized MISO channel vector which is required, e.g., for multi-user beamforming systems. Such systems require accurate CSI at the transmitter to avoid interference between users. Our algorithms provide feedback information for a subset of the OFDM subcarriers to reduce the feedback overhead. Residual correlation between such CSI pilot positions in time and frequency is exploited to improve the quantization accuracy. This work is an extension of our previous proposal [1] to frequency selective OFDM systems.

Index Terms— CSI estimation, CSI feedback, channel vector quantization, multi-user MIMO, LTE, OFDMA

1. INTRODUCTION

Current state of the art single-user multiple input multiple output (MIMO) transmit strategies require only little channel state information (CSI) at the transmitter (CSIT) to achieve a considerable multiplexing, diversity and/or beamforming gain [2]. On the other hand, most of the techniques potentially considered for improving the capacity and cell edge performance of future wireless communication systems (e.g., multi-user MIMO, cooperative multi point transmission), necessitate CSIT with much higher accuracy, to avoid/minimize interference between users. While in low mobility time division duplex (TDD) systems accurate CSIT may readily be available due to channel reciprocity, the currently dominating frequency division duplex (FDD) systems (e.g., FDD LTE-A [3], WiMAX 2.0 [4]) require a dedicated feedback channel to obtain CSIT. Due to the limited capacity of the feedback link efficient CSI quantization algorithms are required, targeting a minimal quantization error for a given feedback rate.

In this work, we propose an efficient channel vector quantization algorithm for OFDM systems, which is an extension of our previously presented algorithm [1] to frequency selective channels. The algorithm exploits time and frequency channel correlation to improve the quantization accuracy, by iteratively adapting the quantization codebook to match the temporal evolution of the wireless channel. We focus on quantization of the normalized channel vector (channel direction), because its accuracy significantly determines the performance of interference mitigation/cancellation techniques, while coarse channel magnitude information is typically sufficient in practical systems for the choice of an appropriate modulation and coding scheme. To reduce the feedback overhead, CSIT is only provided for a subset of all OFDM subcarriers, denoted pilot subcarriers. The residual CSI is estimated at the transmitter by means of interpolation. To maintain a small interpolation error the distance between the CSI pilots must be kept moderate in relation to the channel coherence bandwidth [5]. Thus, residual correlation between the pilots exists and can be exploited to improve the quantization accuracy. This is demonstrated by means of simulation results in Section 3. The quantization problem was also considered by other authors, e.g, [6, 7]. A main distinguishing feature between existing algorithms and our algorithm is that we employ adaptive filters to adapt to the statistics (time-frequency correlation) of the wireless channel.

This work has been funded by A1 Telekom Austria AG and the Institute of Telecommunications, Vienna University of Technology.

2. QUANTIZATION ALGORITHM

In this section, we present the considered OFDM system model and given an overview of the proposed CSI feedback scheme. Furthermore, the quantization algorithm and the structure of the codebook adaptation filters are detailed.

2.1. System Model

The currently dominating modulation and multiple access technology of wireless communication systems (e.g., LTE, WiMAX) is orthogonal frequency division multiple access. We thus focus on OFDM systems which convert the broadband frequency selective wireless channel into N orthogonal narrowband frequency flat channels (subcarriers), by means of an FFT and application of a cylcic prefix. The considered system architecture is shown in Fig. 1.

CSIT is provided via a dedicated limited capacity feedback link, by quantizing the channel experienced on a subset $\mathcal{N}_p$ of all subcarriers. The subcarriers in $\mathcal{N}_p$ are denoted pilot sub-

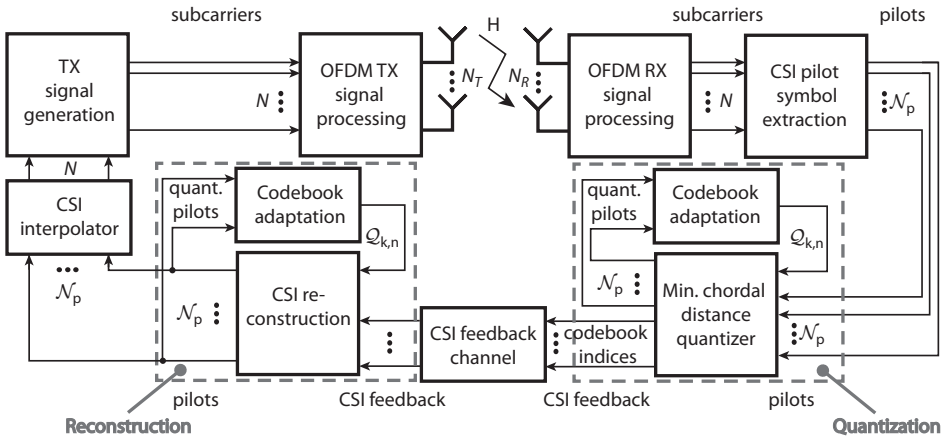

Fig. 1. System architecture of the considered OFDM system together with the proposed CSI feedback scheme.

carriers. A constant distance Δf_s in multiples of subcarriers between pilots is assumed. The size of the pilot set $\mathcal{N}_p$ is denoted N_p. The receiver quantizes normalized channel vectors, called channel directions, $\mathbf{h}_{(k,n)} \in \mathbb{C}^{N_T \times 1}, ||\mathbf{h}_{(k,n)}||_2 = 1, n \in \mathcal{N}_p$, where N_T denotes the number of transmit antennas and k, n are the time and subcarrier indices.

The channel vectors depend on the considered transmit and receive strategies, e.g., for eigen-beamforming a subset of the right singular vectors of the channel matrix is quantized. For simplicity we restrict ourselves to $N_R = 1$ receive antenna, allowing the receiver to directly quantize the normalized channel matrix $\mathbf{h}_{(k,n)} = \frac{\mathbf{H}_{(k,n)}}{||\mathbf{H}_{(k,n)}||_2}$. Here, $\mathbf{H}_{(k,n)} \in \mathbb{C}^{N_T \times N_R}$ denotes the channel matrix on subcarrier n at time instant k. As pointed out in [7], the absolut phase angle of the channel direction is unimportant for multi-user transmission strategies like zero forcing beamforming. Thus all vectors $\mathbf{h}_{(k,n)}e^{j\alpha}, \forall \alpha \in \mathbb{R}$ are equivalent in our quantization problem, and form a Grassmann manifold $\mathcal{G}_{\mathbb{C}}(1, N_T)$ as shown in [7]. In contrast to our previous proposal [1], where we only exploited the unit norm (Stiefel manifold) structure of the quantization proplem, we here additionally make use of the phase independence of the Grassmann manifold.

It was demonstrated in [1] that uniform quantization of the unit sphere with satisfactory quantization accuracy for multi-user beamforming is only possible with a large quantization codebook, causing a large CSI feedback overhead. To reduce the codebook size while still providing good quantization accuracy, it is necessary to drop the insistence on uniform quantization and allow for codebooks that provide high code vector density only in the currently relevant region of the manifold. The quantization accuracy is determined by the precision of the algorithms that determine the relevant quantization region by exploiting time-frequency channel correlation.

2.2. Quantization Codebook Adaptation

The basic idea of our algorithms is to exploit time-frequency channel correlation, by recognizing that the channel direction $\mathbf{h}_{(k,n)}$ does not change arbitrarily fast, neither over time k nor over frequency n. Thus, the next channel direction to be quantized can be predicted from previous observations. We can then employ a quantization codebook that quantizes only a surface area around the predicted vector, whose size depends on the prediction accuracy. To ensure that the transmitter and receiver employ the same codebooks, the prediction and adaptation of the codebook must be based on quantized channel knowledge. The codebook employed at quantization instance (k, n) is denoted $\mathcal{Q}_{k,n} \subset \mathbb{C}^{N_T \times 1}$

We propose a quantization algorithm, that serially quantizes the pilot subcarriers $n \in \mathcal{N}_p$ at each time instant k. Therefore, potentially all quantized pilots from previous time samples $\{k-1, k-2, \ldots\}$ and additionally pilots with subcarrier indices $i < n$ from the current sample k can be utilized for prediction and codebook adaptation. To reduce complexity only neighbouring pilots are used in the simulations.

One difficulty in the prediction of vectors belonging to the Grassmann manifold arises from the unit norm constraint imposed by the normalization of the channel vector. A linear prediction based on vectors from the Grassmann manifold does in general not fulfill this constraint, because the Grassmann manifold is not a linear vector space. Thus, we propose to exploit the differentiable structure of the Grassmann manifold and perform the prediction in the Euclidean tangent space associated with the manifold (see [8] for details). The tangent space $\mathcal{T}_{\mathbf{x}_i}$ to a vector $\mathbf{x}_i \in \mathcal{G}_{\mathbb{C}}(1, N_T)$ is defined as the space of all N_T-vectors $\mathbf{x}_i^{(t)}$ orthogonal to $\mathbf{x}_i : \mathcal{T}_{\mathbf{x}_i} = \{\mathbf{x}_i^{(t)} : \mathbf{x}_i^H \mathbf{x}_i^{(t)} = 0\}$. The tangent space enables us to describe the transition between two vectors $\mathbf{h}_{q,(k-1,n)}$ and $\mathbf{h}_{q,(k,n)}$ via the geodesic curve, that connects these two vectors over the shortest path in the manifold (see [1] for details). Here, $\mathbf{h}_{q,(k,n)} \in \mathcal{G}_{\mathbb{C}}(1, N_T)$ denotes the quantized channel direction obtained at quantization instant (k, n). This is visualized in the upper part of Fig. 2, where the time-frequency grid of channel directions together with the corresponding tangent vectors is shown. The geodesic curve is described by the tangent vector $\mathbf{h}_{q,(k-1,n)}^{(t)} \in \mathcal{T}_{\mathbf{h}_{q,(k-1,n)}}$

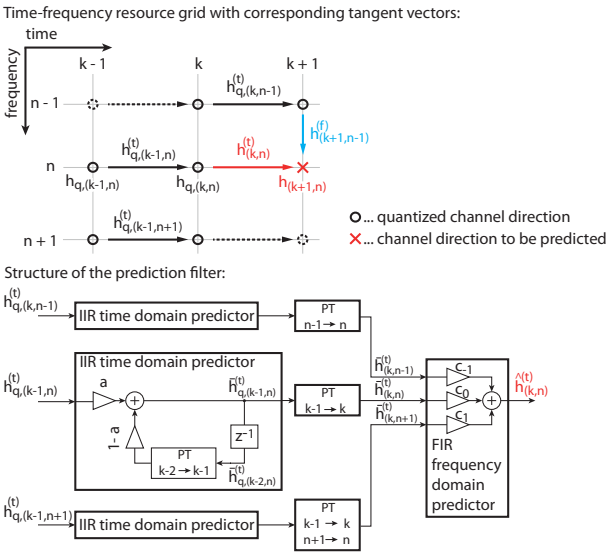

Fig. 2. Time-frequency resource grid with corresponding tangent vectors, and filter structure of the employed predictor.

(see [7])

$$\mathbf{h}^{(t)}_{q,(k-1,n)} = \operatorname{atan}\left(\frac{d}{|\rho|}\right)\frac{\mathbf{h}_{q,(k,n)}/\rho - \mathbf{h}_{q,(k-1,n)}}{d/|\rho|} \quad (1)$$

$$\rho = \mathbf{h}^{H}_{q,(k-1,n)}\mathbf{h}_{q,(k,n)}, \quad d = \sqrt{1-|\rho|^2} \quad (2)$$

Conversely, knowing the tangent vector $\mathbf{h}^{(t)}_{q,(k-1,n)}$ and its originating position $\mathbf{h}_{q,(k-1,n)}$, denoted anchor, one can obtain $\mathbf{h}_{q,(k,n)}$ via the geodesic curve (see Equation (4) in [7]). Hence a direct prediction of the next channel direction to be quantized, $\mathbf{h}_{(k+1,n)}$ in Fig. 2, is equivalent to a prediction of the tangent vector $\mathbf{h}^{(t)}_{(k,n)}$ that describes the transition from $\mathbf{h}_{q,(k,n)}$ to $\mathbf{h}_{(k+1,n)}$.

We obtain the prediction $\hat{\mathbf{h}}^{(t)}_{(k,n)}$ by linear combination of previously observed tangents between quantized channel directions, employing the filter structure shown in the lower part of Fig. 2. To reduce the complexity, a separable filter is employed consisting of a time domain infinite impulse response (IIR) exponential averaging part, followed by a finite impulse response (FIR) frequency domain combiner. The main purpose of the IIR filter is a reduction of the quantization noise in the tangent vectors [1].

A linear combination of tangent vectors is only meaningful if these vectors have the same anchor point. As shown in Fig. 2, this is per se not the case and it is therefore necessary to translate the anchor points of tangent vectors, while keeping their relative orientations (relative to the anchor point) the same. This is enabled by means of parallel transport (PT) [8]. Wherever PT is necessary, there is a box in Fig. 2 denoted PT and labeled with the required PT steps. E.g., in the IIR time domain predictor a short term exponential average of tangent vectors is computed. This is achieved by combining the previous average tangent vector $\bar{\mathbf{h}}^{(t)}_{q,(k-2,n)}$ with the current tangent $\mathbf{h}^{(t)}_{q,(k-1,n)}$. Thus, the anchor of $\bar{\mathbf{h}}^{(t)}_{q,(k-2,n)}$ is transported from $\mathbf{h}_{q,(k-2,n)}$ to $\mathbf{h}_{q,(k-1,n)}$ along the geodesic curve connecting these two channel directions.

Analogous to utilizing tangents between consecutive pilots in the time domain $\mathbf{h}^{(t)}_{q,(k,n)}$, one can exploit frequency domain tangents $\mathbf{h}^{(f)}_{q,(k,n)}$ as well. In this case, the channel direction $\mathbf{h}_{(k+1,n)}$ is predicted, by estimating the tangent $\mathbf{h}^{(f)}_{(k+1,n-1)}$ between $\mathbf{h}_{q,(k+1,n-1)}$ and $\mathbf{h}_{(k+1,n)}$ (see Fig. 2). For reasons of clarity, we did not include the corresponding prediction filters in Fig. 2, but the very same filter structure as for the time domain tangents can be employed in this case as well. Depending on whether time or frequency correlation is dominant in the system, the time or frequency tangent based prediction proofs to be more accurate. In Section 2.4, we show how both predictions $\hat{\mathbf{h}}_{t,(k+1,n)}$ and $\hat{\mathbf{h}}_{f,(k+1,n)}$ can be combined adaptively to deliver the final estimate $\hat{\mathbf{h}}_{(k+1,n)}$.

As soon as $\hat{\mathbf{h}}_{(k+1,n)}$ is calculated, the quantization codebook $\mathcal{Q}_{k+1,n}$, intended for quantizing $\mathbf{h}_{(k+1,n)}$, is obtained by following the same steps as detailed in Section *Codebook Adaptation* of [1].

2.3. Filter Adaptation

The FIR and IIR filters employed for channel direction prediction are both adaptive filters, that are sequentially optimized at each quantization instant. For the adaptation of the FIR filter, we employ the normalized least mean squares (NLMS) algorithm, as summarized below. A frequency band of size $2N_f + 1$ pilots (symmetric around n) is used as input to the FIR filter bank. The filter input vectors are combined in a matrix $\mathbf{H}_{(k,n)} \in \mathbb{C}^{N_T \times 2N_f+1}$

$$\mathbf{H}_{(k,n)} = \left[\bar{\mathbf{h}}^{(t)}_{(k,n-N_f)}, \ldots, \bar{\mathbf{h}}^{(t)}_{(k,n)}, \ldots, \bar{\mathbf{h}}^{(t)}_{(k,n+N_f)}\right]. \quad (3)$$

The vectors $\bar{\mathbf{h}}^{(t)}_{(k,n-i)}$ denote the parallel transported versions of IIR filter outputs, see Fig. 2. The FIR filter output is

$$\hat{\mathbf{h}}^{(t)}_{(k,n)} = \mathbf{H}_{(k,n)}\mathbf{c}_{(k,n)}, \quad \mathbf{c}_{(k,n)} = \left[c_{-N_f}, \ldots, c_{N_f}\right]^{T}_{(k,n)} \quad (4)$$

where $\mathbf{c}_{(k,n)}$ denotes the filter coefficient vector at quantization instant (k,n). Note that the same filter coefficients are employed for all entries of the tangent vectors. This assumes the same statistics of the individual entries of the tangents and neglects a possible correlation between the entries, which could be exploited to further improve the prediction performance. The NLMS update rule for the filter is

$$\mathbf{e}_{(k,n)} = \mathbf{h}^{(t)}_{(k,n)} - \hat{\mathbf{h}}^{(t)}_{(k,n)} = \mathbf{h}^{(t)}_{(k,n)} - \mathbf{H}_{(k,n)}\mathbf{c}_{(k,n)} \quad (5)$$

$$\mathbf{c}_{(k,n+1)} = \mathbf{c}_{(k,n)} + \mu\frac{\mathbf{H}^{H}_{(k,n)}}{||\mathbf{H}_{(k,n)}||^2_F}\mathbf{e}_{(k,n)}. \quad (6)$$

Here, $\mathbf{e}_{(k,n)}$ denotes the prediction error and $||\cdot||_F$ is the Frobenius norm. The step size parameter μ has to satisfy

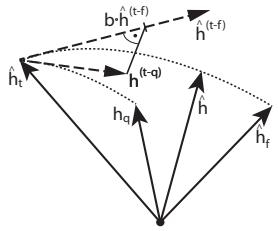

Fig. 3. Visualization of the combination of the time and frequency domain predictions in the tangent space.

$0 < \mu < 2$ for convergence [9] (we use $\mu = 0.1$). As an alternative to the NLMS algorithm, the recursive least squares (RLS) algorithm can be employed as well. We tested both algorithms, but opted for NLMS due to its lower complexity at very similar performance.

For the adaptation of the real-valued IIR filter we impose the constraint $0 \leq a_{(k,n)} \leq 1$ to achieve the desired exponential averaging behavior. Although a formulation of the NLMS algorithm obeying additional constraints exists, we utilize a simple line search to find the optimal parameter $\hat{a}_{(k,n)}$ at each iteration and use the stochastic learning rule

$$a_{(k,n+1)} = \left(1 - \frac{1}{\beta}\right) a_{(k,n)} + \frac{1}{\beta}\hat{a}_{(k,n)}. \tag{7}$$

to obtain an average value ($\beta \geq 1$, we set $\beta = 25$). Similar update rules hold for the frequency domain tangent filters.

At the borders of the system bandwidth, the full pilot band of size $2N_f + 1$ is not available (because there are no pilots outside the system band). Therefore it is necessary to utilize narrowed FIR filter banks at the frequency band borders, that are adapted independently from the full width filter bank.

In Section 3, we examine two distinct quantization orders:

1. Straight order: at each time instant k, the subcarrier quantization order is $n = \{1, 2, \ldots, N\}$.
2. Zigzag order: at alternating time instants, the subcarrier quantization order switches between $n = \{1, \ldots, N\}$ and $n = \{N, N-1, \ldots, 1\}$.

2.4. Prediction Combination in the Tangent Space

As mentioned in Section 2.2, we obtain two different channel direction predictions $\hat{\mathbf{h}}_{t,(k,n)}$ and $\hat{\mathbf{h}}_{f,(k,n)}$, by employing time as well as frequency domain tangent vectors. Averaging over these two predictions, it is possible to reduce the prediction noise such as to obtain a more accurate prediction. A simple arithmetic mean cannot be used for that purpose, because the result has to be an element of $\mathcal{G}_{\mathbb{C}}$. Therefore, we propose to perform the averaging in the tangent space, by calculating the tangent $\hat{\mathbf{h}}_{(k,n)}^{(t-f)}$ between $\hat{\mathbf{h}}_{t,(k,n)}$ and $\hat{\mathbf{h}}_{f,(k,n)}$, and choosing a tangent $b_{(k,n)}\hat{\mathbf{h}}_{(k,n)}^{(t-f)}$, to obtain the final prediction $\hat{\mathbf{h}}_{(k,n)}$ via the corresponding geodesic, as visualized in Fig. 3. As soon as $\mathbf{h}_{(k,n)}$ is quantized, we adapt the scalar $b_{(k,n)}$ such as to minimize the mean squared error (MSE) between the predicted tangent $b_{(k,n)}\hat{\mathbf{h}}_{(k,n)}^{(t-f)}$ and the observation $\mathbf{h}_{(k,n)}^{(t-q)}$ by means of NLMS.

2.5. Channel Direction Quantization

The quantization of the channel direction is based on the chordal distance

$$d_c(\mathbf{h}_1, \mathbf{h}_2) := \sqrt{1 - |\mathbf{h}_1^H \mathbf{h}_2|^2}. \tag{8}$$

The quantizer calculates the chordal distance between each codevector $\mathbf{q}^{(i)} \in \mathcal{Q}_{k,n}$ and the current channel direction $\mathbf{h}_{(k,n)}$ and feeds back the index i of the codevector that achieves the minimal chordal distance [1]. The chordal distance is the natural distance measure on $\mathcal{G}_{\mathbb{C}}$. It also determines the SINR of zero forcing multi-user beamforming systems with limited feedback [10]. Thus minimizing d_c is equivalent to maximizing the SINR in such systems.

3. SIMULATION RESULTS

In this section, we evaluate the chordal distance MSE achieved with our quantization algorithms. An LTE standard compliant OFDM system with 1.4 MHz bandwidth ($N = 72$ subcarriers with a spacing of $f_s = 15\,\text{kHz}$) and an antenna configuration $N_T \times N_R = 4 \times 1$ with uncorrelated transmit antennas is considered. The quantization algorithm is called every subframe (every $T_s = 1\,\text{ms}$) in time and with different pilot distances Δf_s in multiples of subcarriers. The MSE is calculated on the pilot positions only; CSI interpolation is not considered. Two different power delay profile based channel models are evaluated, the 3GPP models PedA with a root mean square delay spread $\tau_{\text{rms}} = 45\,\text{ns}$ and VehA, $\tau_{\text{rms}} = 370\,\text{ns}$. The temporal correlation of the channel is determined by the maximum Doppler frequency f_d according to Jakes' model. The LTE standard defined beamforming codebook is used as our initial quantization codebook $\mathcal{Q}_{0,0}$. This codebook requires 4 bit for indexing of codevectors, and our algorithm requires one additional bit for the codebook adaptation (see [1]). The feedback channel is assumed error free. Although restrictive, this assumption is not unrealistic, because the feedback channel is typically strongly error protected. In case of error a not-acknowledged (NACK) message would be required to re-synchronize the algorithm. The simulations run for a duration of 1 000 subframes and the MSE is calculated over the last 900 subframes (to discard the initial convergence phase of the filters). Furthermore, the results are averaged over 10 independent simulation runs.

Fig. 4 shows the quantization error over the normalized Doppler frequency $\nu_d = f_d T_s$. Different frequency sampling intervals Δf_s are considered. Without exploiting frequency correlation the dashed curve is obtained. It can be seen that the codebook adaptation in this case improves the quantization performance for $\nu_d \leq 0.15$. Otherwise the same performance as with uniform quantization is obtained. The performance can be improved by exploiting frequency correlation. By decreasing Δf_s, the frequency correlation between the pilots is increased and the quantizer performance improves, but also the feedback overhead grows. We adopt the definitions

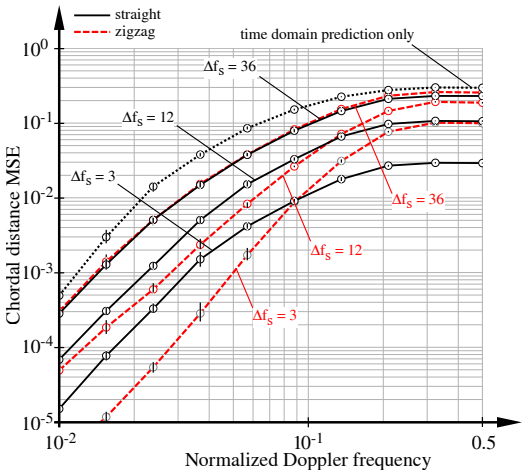

Fig. 4. Chordal distance MSE versus Doppler frequency for an $N_T \times N_R = 4 \times 1$ PedA channel, with a $4+1$ bit codebook.

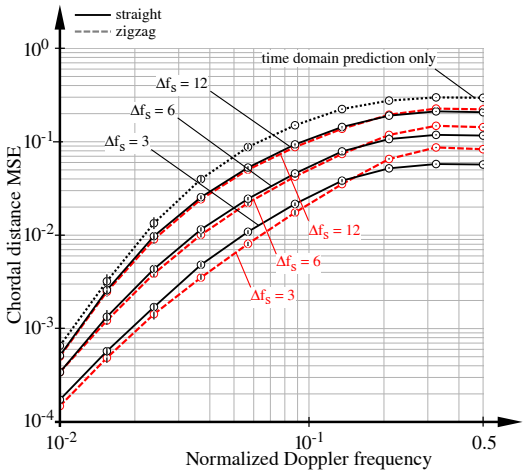

Fig. 5. Chordal distance MSE versus Doppler frequency for an $N_T \times N_R = 4 \times 1$ VehA channel, with a $4+1$ bit codebook.

of coherence time and bandwidth from [11]

$$T_c = \frac{9}{16\pi f_d}, \qquad B_c = \frac{1}{2\pi\tau_{\text{rms}}}. \tag{9}$$

and define their normalized version as $T_c^{(n)} = T_c/T_s$ and $B_c^{(n)} = B_c/(\Delta f_s \cdot f_s)$. The range of these values is $B_c^{(n)} \in [6.5, 78.5]$ and $T_c^{(n)} \in [0.36, 18]$. When temporal correlation is large $\nu_d \leq 0.15$ the zig-zag quantization order outperforms the straight order, and vice-versa for $\nu_d > 0.15$.

With the VehA channel model the normalized coherence bandwidth ranges only over $B_c^{(n)} \in [2.4, 9.5]$, and thus the gain obtained by reducing Δf_s is much smaller than for the PedA channel, see Fig. 5. Comparing both Figures 4 and 5, it can be observed that the slope of the MSE versus f_d is better for PedA than for VehA. We believe that this is caused by the fact that the same prediction filter is employed on all pilots. For the highly correlated PedA channel model that seems to be okay, while independent pilot filters might perform better for the VehA channel. But then also the convergence speed is much decreased, because each filter is only updated once per subframe compared to N_p times. In the VehA channel, the quantization order does not play an important role.

4. CONCLUSION

Accurate CSIT is a key enabler of many candidate technologies considered for boosting the performance of future wireless communication systems, such as interference alignment and multi-user spatial multiplexing. In this work we propose a channel vector quantization algorithm, that is intended for providing CSIT over a limited capacity feedback link. To minimize the feedback overhead, time and frequency channel correlation must be exploited during quantization. We achieve this with an adaptive approach, that provides a quantization codebook that is matched to the channel evolution. The performance of the proposed algorithms is demonstrated by means of simulations. The next step is to incorporate the quantizer in a simulated communication system, to investigate its throughput improvement capabilities.

5. REFERENCES

[1] S. Schwarz and M. Rupp, "Adpative channel direction quantization based on spherical prediction," Ottawa, Canada, June 2012, IEEE International Conference on Communications.

[2] D.J. Love, R.W. Heath Jr., V.K.N. Lau, D. Gesbert, B.D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, Oct. 2008.

[3] 3GPP, "Techn. Spec. Group Radio Access Network; (E-UTRA) and (E-UTRAN); Overall description; Stage 2 (release 10)," Dec. 2010, [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36300.htm.

[4] IEEE and P802.16m, "The draft IEEE 802.16m System Descritpion Document," April 2009, IEEE 802.16m-08/003r8.

[5] S. Schwarz and M. Rupp, "Adaptive channel direction quantization – enabling multi user MIMO gains in practice," Ottawa, Canada, June 2012, IEEE International Conference on Communications.

[6] P. Svedman, E.A. Jorswieck, and B. Ottersten, "Reduced feedback SDMA based on subspace packings," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, March 2009.

[7] T. Inoue and R.W. Heath Jr., "Grassmannian predictive coding for delayed limited feedback MIMO systems," in *47th Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2009.

[8] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.

[9] Simon Haykin, *Adaptive filter theory*, Prentice Hall, Upper Saddle River, NJ, 4th edition, 2002.

[10] M. Trivellato, F. Boccardi, and F. Tosate, "User selection schemes for MIMO broadcast channels with limited feedback," in *65th IEEE Vehicular Technology Conference, Spring 2007*, Dublin, April 2007.

[11] John B. Anderson and Arne Svensson, *Coded Modulation Systems*, Kluwer Academic Publishers, Norwell, MA, USA, 2002.