2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications: Track 1 - PHY and Fundamentals

Orthogonal Precoding with Channel Prediction for High Mobility Massive MIMO

David Löschenbrand, Markus Hofer, Thomas Zemen

Security & Communication Technologies, *AIT Austrian Institute of Technology GmbH*
Vienna, Austria
david.loeschenbrand@ait.ac.at

***Abstract*—We present a strategy to overcome channel aging and to enable massive multiple-input multiple-output (MIMO) in highly time-varying scenarios. Given the self-contained 5G slot structure, we introduce channel prediction methods for the down-link utilizing up-link reference symbols and investigate the prediction quality for synthetic and empirical high mobility channel measurement data. Further, we combine channel prediction with orthogonal precoding to decrease the bit error rate (BER) and allow longer prediction horizons. We show through link-level simulations that this framework exhibits a performance drop of only $1\,\mathrm{dB}$ compared to perfectly known channel state information for a velocity of $160\,\mathrm{km/h}$ at $3.5\,\mathrm{GHz}$. We conclude the paper with a parametric BER study in dependence of the prediction/precoding horizon length.**

***Index Terms*—massive MIMO, orthogonal precoding, channel prediction, BER, high mobility**

I. Introduction

Massive multiple-input multiple-output (MIMO) is a key enabler for the currently rolled out mobile communication standard 5G [1]. Its future advancements like very large aperture arrays, cell-free massive MIMO and intelligent surfaces are already under consideration for the succeeding standard [2]. In favorable propagation scenarios, massive MIMO offers significant improvements in terms of spectral efficiency and it mitigates random wireless channel fluctuations (fading), thus serving both enhanced mobile broadband (eMBB) and ultra-reliable low latency communication (URLLC) use-cases.

To achieve those promises in a time division duplex (TDD) system and to exploit channel reciprocity, timely up-link channel state information (CSI) is of utmost importance. It is used for beam-forming and therefore guarantees constructive interference of impinging paths at the intended user and, depending on the beam-forming method, random or destructive interference at other users participating in the cell. For highly time-varying scenarios, the channel changes significantly between up-link CSI acquisition and down-link beam-forming (an effect called *channel aging*). Channel aging causes an increase of the bit error rate (BER). Mitigation strategies include decreasing the time between up-link CSI acquisition and down-link beam-forming, CSI prediction and orthogonal precoding (OP), which will be detailed in the following.

978-1-7281-4490-0/20/$31.00 © 2020 IEEE

There is vast literature on channel prediction for MIMO systems applying various methods. In [3], a Markov-model based approach with a Kalman-filter tracking the channel statistics is proposed. The authors in [4], [5] directly apply a Kalman-predictor. The work in [6] describes recurrent neural networks as well as special training methods for channel prediction and [7] investigates deep-learning prediction approaches. The authors in [8] investigate mitigation of channel aging via optimal Wiener channel prediction. However, none of the above explicitly investigates prediction quality given the constraints posed by the self-contained slot structure of the 5G standard [9].

Another strategy to mitigate fading induced by channel aging is spreading information in time and frequency as initially introduced in [10], a method we refer to as OP. In [11] the authors investigate the reduction of fading (*channel hardening*) in a massive MIMO system by applying OP with either outdated or perfect channel knowledge and achieved a significant BER reduction in a high mobility scenario.

This work, to the best of the authors knowledge, for the first time combines both channel prediction and OP within the self-contained slot structure of the 5G standard to leverage the full potential of massive MIMO in highly time-varying scenarios.

Our main scientific contributions are the following:

- We compare the CSI prediction quality of various methods in a framework resembling the self-contained slot structure in the 5G standard [9] for synthetic and empirical data from a vehicular measurement campaign.
- We show that with additional utilization of OP, the BER performance of polynomial (Wiener) prediction approaches that of perfectly known CSI to within $2\,\mathrm{dB}$ ($1\,\mathrm{dB}$) for a velocity of $160\,\mathrm{km/h}$ at $3.5\,\mathrm{GHz}$.
- We investigate the dependence of the BER performance on the number of base station (BS) antennas

978-1-7281-4490-0/20/$31.00 ©2020 IEEE

and the prediction/precoding horizon.

II. Signal Model

We consider a massive MIMO down-link system where a BS with a large number of antennas A transmits to single-antenna mobile stations (MSs) in a highly time-varying scenario. We focus on one MS for ease of notation, but the given signal model can readily be extended to several users. The underlying multiplexing scheme is assumed to be TDD, the modulation is orthogonal frequency-division multiplexing (OFDM) and the transmission takes place in a blocked manner, thus M time samples and Q frequency samples are grouped in one block. To obtain the vectorized transmit symbols $\mathbf{d} \in \mathbb{C}^{MQ\times 1}$, the vectorized information symbols $\mathbf{b} \in \mathbb{C}^{MQ\times 1}$ are spread over a time-frequency block of size $M \times Q$, as described in [11], via the orthogonal spreading matrix $\mathbf{S} \in \mathbb{C}^{MQ\times MQ}$, as

$$\mathbf{d} = \mathbf{S}\mathbf{b}. \tag{1}$$

Suitable spreading matrices are derived from orthogonal basis sets such as the discrete symplectic Fourier transform (DSFT) [10], Walsh-Hadamard or discrete prolate spheroidal (DPS) sequences [12]. The rationale for spreading the information symbols over time and frequency is to leverage diversity of the doubly-selective propagation channel to mitigate fading and combat channel aging [11]. As has been shown in the cited work, DSFT spreading performs marginally better than the other mentioned methods and is therefore considered in this work.

The time index is denoted by $m \in \{0, \dots, M-1\}$ and the sub-carrier index is denoted by $q \in \{0, \dots, Q-1\}$. With $\mathrm{vec}(\cdot)$ denoting vectorization, we compactly write the vectorized matrix $\mathbf{t} = \mathrm{vec}(\mathbf{T})$ with elements $(\mathbf{T})_{ij} = t_{i,j}$ as $\mathbf{t} = \mathrm{vec}(t_{i,j})$. The vectorized received samples $\mathbf{y} = \mathrm{vec}(y_{m,q}) \in \mathbb{C}^{MQ\times 1}$ at the MS in a system applying OP are then modeled as

$$\mathbf{y} = \mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{d} + \mathbf{z}, \tag{2}$$

with the massive MIMO channel matrix

$$\mathbf{H} = \begin{bmatrix} \mathrm{diag}(\mathbf{h}_1) \\ \vdots \\ \mathrm{diag}(\mathbf{h}_A) \end{bmatrix} \in \mathbb{C}^{MQA\times MQ} \tag{3}$$

and the (down-link) beam-forming matrix

$$\mathbf{W} = \begin{bmatrix} \mathrm{diag}(\mathbf{w}_1) \\ \vdots \\ \mathrm{diag}(\mathbf{w}_A) \end{bmatrix} \in \mathbb{C}^{MQA\times MQ}. \tag{4}$$

The individual channel vectors $\mathbf{h}_a = \mathrm{vec}(h_{m,q,a}) \in \mathbb{C}^{MQ\times 1}$ contain MQ time-frequency channel coefficients from BS antenna a to the MS. They are diagonalized and stacked to form the channel matrix $\mathbf{H}$. Likewise, the individual beam-forming vectors $\mathbf{w}_a \in \mathbb{C}^{MQ\times 1}$ are diagonalized and stacked to form the beam-forming matrix $\mathbf{W}$. The additive noise is denoted by $\mathbf{z} \sim \mathcal{CN}(0, {\sigma_z}^2\mathbf{I}_{MQ})$, with ${\sigma_z}^2$ denoting the noise variance. We normalize the channel matrix $\mathbb{E}\left(\|\mathbf{H}\|_2^2\right) = MQ$ such that the array gain is not taken into account. The down-link signal to noise ratio (SNR) is than given by $1/{\sigma_z}^2$.

III. Methods to Combat Channel Aging

The prospect of massive MIMO is that, since the number of BS antennas is large, i.e. $A \gg 1$, multiplying the transmit symbols $\mathbf{d}$ with a suitable beam-forming matrix $\mathbf{W}$ leads to coherent superposition of all multipath components at the position of the MS. Thus, fading is (almost) eliminated and the channel becomes deterministic (a phenomenon often referred to as channel hardening [13]) in the case of favorable propagation [1]. Popular methods in literature for choosing $\mathbf{W}$ are maximum ratio transmission (MRT), zero-forcing (ZF) and miminum mean square error (MMSE) beam-forming [14]. For the calculation of the beam-forming matrix, reciprocity and timely up-link CSI is assumed. The CSI becomes outdated quickly in highly time-varying scenarios, an effect called channel aging, and thereby severely degrades reliability and throughput of massive MIMO systems [8], [11]. To prevent channel aging and enable massive MIMO in highly time-varying scenarios, accurate CSI is needed at the BS through either a suitable frame structure, CSI prediction, or both. Additionally, fading as a result of channel aging is successfully mitigated by OP.

These three basic approaches are detailed in the following:

a) Flexible frame structure: Here, the frame structure of the communication link is designed in a way that minimizes the time delay between CSI acquisition in the up-link and down-link beam-forming. In long term evolution (LTE) systems this is possible on a sub-frame basis with time delays exceeding $1\,\mathrm{ms}$. However, 5G will implement a more flexible frame structure that in particular departs from the strict designation of up-link, guard or down-link for each sub-frame to combat channel aging. It foresees hybrid slots, each comprised of 14 symbols, that allow flexible allocation of both up-link and down-link symbols as well as special up-link sounding reference signals (SRSs) [9], see Fig. 1. This effort decreases the time passing between the CSI acquisition and down-link beam-forming.

b) OP: This method does not mitigate the cause of channel aging (i.e. the change of the propagation channel during the time delay between up-link and down-link), but it offers a way to mitigate its effect. By spreading information bits over a time-frequency block, the system is no longer susceptible to errors caused by fading of individual subcarriers over time. The effect of channel hardening is here achieved by exploiting time-frequency diversity rather than spatial diversity [11].

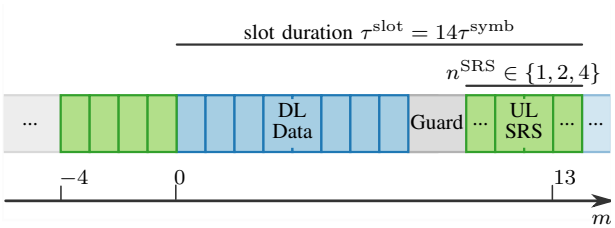

Fig. 1. Self-contained slot structure of the 5G standard. For a given down-link (DL) phase with symbol indices $m \geq 0$, the up-link (UL) SRS symbols with indices $m \in \{-4, -3, -2, -1\}$ provide timely CSI.

c) Channel prediction: Based on acquired CSI in the up-link, the propagation channel is predicted for the time of down-link beam-forming and transmission. The channel prediction can be done in several ways which differ in requirements and complexity and will be detailed in section III-A. With high quality channel prediction, the adverse effects of channel aging are eliminated and massive MIMO in highly time-varying scenarios is enabled. However, the prediction horizon for which this is possible strongly depends on the relative velocities of the MS in the reflecting surrounding as well as on the carrier frequency.

In this work, we consider the flexible frame structure offered by the 5G standard and investigate the impact of OP, channel prediction and their combination on the BER of a single-user massive MIMO system.

A. Channel Prediction Methods

The general framework in which we investigate channel prediction algorithms is based on the 5G new radio (NR) self-contained TDD slot structure. It allows to carry $n^{\text{SRS}} \in \{1, 2, 4\}$ consecutive SRS symbols within the last six symbols of one slot in the up-link. With this measure, timely CSI is acquired at the BS before down-link beam-forming. Figure 1 shows the self-contained slot structure with eight down-link data symbols, two guard symbols and four up-link SRS symbols at the end of the slot. Assuming a subcarrier spacing of 15 kHz, the slot duration equals the sub-frame duration at $\tau^{\text{slot}} = 1\,\text{ms}$ and the symbol duration is assumed to be $\tau^{\text{symb}} = \tau^{\text{slot}}/14 = 71.43\,\mu\text{s}$.

The goal of the channel prediction schemes is to predict the channel for the down-link symbols given the four up-link SRS symbols from the preceding slot. We consider prediction on a per sub-carrier basis and therefore drop the subcarrier index q. The channel vectors $\mathbf{h}_a = [h_{0,a}, \ldots, h_{M-1,a}]^{\text{T}} \in \mathbb{C}^{M\times 1}$ from BS antenna a to the MS reduce to time series vectors with the block length M. Further, the SRS channel coefficients $\hat{h}_{m,a}$ are estimated in the up-link from known SRS channel coefficients d_m^{SRS} by

$$\hat{h}_{m,a} = \frac{h_{m,a} d_m^{\text{SRS}} + w_{m,a}}{d_m^{\text{SRS}}} = h_{m,a} + \hat{w}_{m,a}. \quad (5)$$

The additive white Gaussian noise process $\hat{w}_{m,a} \sim \mathcal{CN}(0, \sigma_{\hat{w}}^2)$ is assumed stationary over one slot duration and independent on a. Time indices $m \in \{-4, \ldots, -1\}$ indicate SRS channel coefficients from the preceding slot.

Given the above notation, we seek a prediction $\tilde{\mathbf{h}}_a = [\tilde{h}_{0,a}, \ldots, \tilde{h}_{M-1,a}]^{\text{T}}$ of the actual channel vector $\mathbf{h}_a$ given the SRS channel coefficient estimates $\hat{h}_{m,a}$ of the preceding slot. The metric to quantify the prediction quality is the relative time-varying mean squared error (MSE)

$$\epsilon\left(\tilde{\mathbf{h}}_a, \mathbf{h}_a\right) = \frac{\left(\tilde{\mathbf{h}}_a - \mathbf{h}_a\right)^{\text{H}} \left(\tilde{\mathbf{h}}_a - \mathbf{h}_a\right)}{\mathbf{h}_a^{\text{H}} \mathbf{h}_a}. \quad (6)$$

To obtain the predicted channel vectors $\tilde{\mathbf{h}}_a$ we investigate four prediction methods which are detailed in the following.

a) Constant prediction: This methods assumes a slowly varying channel with a coherence time much larger than the slot duration τ^{slot}. In this case, the last estimated channel coefficient $\hat{h}_{-1,a}$ is considered a suitable prediction for the following channel coefficients, i.e.

$$\tilde{h}_{m,a} = \hat{h}_{-1,a}. \quad (7)$$

In other words, no prediction is done, but the last known noisy up-link CSI is applied for down-link beam-forming (which depends on the channel vector). Only one up-link SRS channel coefficient is used with this method, so the overhead is minimized.

b) Linear prediction: This method applies a linear extrapolation using the last channel coefficients. The linear prediction is calculated by solving the linear set of equation $\hat{h}_{m,a} = b_0 m + b_1, m \in \{-2, -1\}$ for b_0, b_1 and subsequent calculation of

$$\tilde{h}_{m,a} = b_0 m + b_1 = \hat{h}_{-1,a} + (m+1)\left(\hat{h}_{-1,a} - \hat{h}_{-2,a}\right) \quad (8)$$

for $m \geq 0$.

c) Polynomial prediction: A polynomial of third degree is fitted to the SRS channel coefficients and extrapolated to obtain the channel coefficients for $m \geq 0$. The polynomial prediction is calculated as

$$\tilde{h}_{m,a} = \sum_{m'=-4}^{-1} \hat{h}_{m',a} L_{m'}(m) \quad (9)$$

with the Lagrange basis functions defined as

$$L_{m'}(m) = \prod_{\substack{m''=-4,\\ m'' \neq m'}}^{-1} \frac{m - m''}{m' - m''}. \quad (10)$$

Four SRS channel coefficients are required for the polynomial prediction of third degree.

d) Wiener prediction: Assuming that the fading process of the channel coefficients is described by the Jakes model, the autocorrelation of the channel coefficients in time independent on the BS antenna a is given by

$$r_h(m) = J_0\left(2\pi f_\mathrm{D} \tau^\mathrm{symb} |m|\right) \tag{11}$$

with $J_0(\cdot)$ being the zeroth-order Bessel function of the first kind and f_D the maximum Doppler shift.

Given that the up-link SRS channel coefficients are corrupted by noise according to (5), the Wiener predictor is written as [15]

$$\tilde{h}_{m,a} = \mathbf{r}_h^\mathrm{H}(m) \mathbf{R}_{\hat{h}}^{-1} \hat{\mathbf{h}}_a \tag{12}$$

with

$$\mathbf{R}_{\hat{h}} = \mathbf{R}_h + \sigma_{\hat{w}}^2 \mathbf{I}_4, \tag{13}$$

$$(\mathbf{R}_h)_{m,m'} = r_h(m - m'), \tag{14}$$

$$\mathbf{r}_h(m) = [r_h(m), \ldots, r_h(m-3)]^\mathrm{T} \tag{15}$$

$$\hat{\mathbf{h}}_a = [\hat{h}_{-1,a}, \ldots, \hat{h}_{-4a}]^\mathrm{T} \tag{16}$$

and known noise variance $\sigma_{\hat{w}}^2$.

IV. Results

First, we investigate the prediction algorithms outlined in Section III-A in terms of prediction quality, i.e. the relative MSE from (6), depending on the block length M in time, for a given maximum Doppler frequency f_D. The case of perfectly known channel coefficients, i.e. perfect prediction, is also included for reference when appropriate. We analyze both synthetic data and data from vehicular measurements. In a second approach, we investigate the combination of OP and channel prediction in a link-level simulation using similar parameters as foreseen in the 5G standard documents.

A. Prediction Quality

Simulations are done for a speed of $70\,\mathrm{km/h}$ at a carrier frequency of $3.5\,\mathrm{GHz}$, resulting in a maximum Doppler frequency of $227\,\mathrm{Hz}$. A correlated noiseless Rayleigh fading channel realization according to [16] is assumed which exhibits a Jakes Doppler power spectral density. The number of paths is 40 and no correlation across antennas is considered. The number of channel realizations of length M over which the relative MSE is calculated and averaged is 65000.

Figure 2 shows the relative MSE $\epsilon\left(\tilde{\mathbf{h}}_a, \mathbf{h}_a\right)$ over the block length M. The Wiener predictor is superior to the other approaches as the autocorrelation process and the maximum Doppler shift f_D in (11) are known perfectly. The polynomial prediction performs reasonably well without any assumptions on the underlying channel statistics. The marked lines in Figure 2 show the relative MSE for the Wiener predictor with a correlation matrix obtained by decreasing or increasing the design Doppler frequency f_D in (11) by $25\,\%$. In this case, the prediction

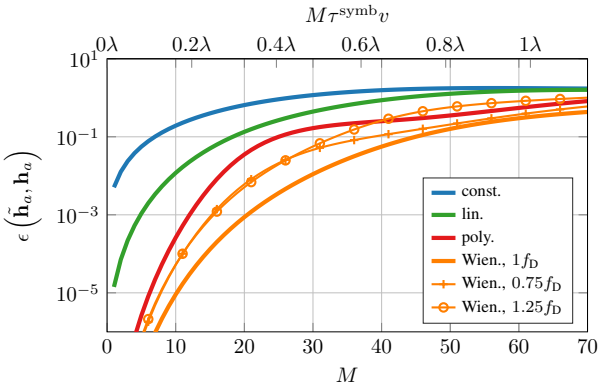

Fig. 2. Relative MSE over the block length M in time for a velocity of $70\,\mathrm{km/h}$ at $3.5\,\mathrm{GHz}$ in a simulated Rayleigh channel. The Wiener predictor with perfect knowledge of the channel statistics exhibits the best prediction quality. For a mismatching maximum Doppler frequency f_D it approaches the quality of the polynomial prediction.

quality deteriorates and approaches the quality of the polynomial predictor.

Simulation models only reproduce the real world to a certain degree. Therefore we use empirical data from vehicular measurements to analyze prediction quality in more realistic conditions. Given the high SNR of the measurements, we consider the gathered empirical data as perfect (i.e. noiseless) representation of the channel. The empirical channel coefficients $\mathbf{h}_{m,a}$ were obtained in a line of sight (LOS) scenario with a MS mounted on a car with a velocity of approximately $60\,\mathrm{km/h}$ at a center frequency of $3.5\,\mathrm{GHz}$. Details on the setup used to gather the empirical data are found in [17]. The repetition rate of the measurement was $1\,\mathrm{ms}$ and a DPS interpolation method with oversampling factor 14 as described in [18] is used to obtain channel realizations with a repetition rate of $\tau^\mathrm{symb} = 71.43\,\mu\mathrm{s}$.

Figure 3 shows the relative MSE $\epsilon\left(\tilde{\mathbf{h}}_a, \mathbf{h}_a\right)$ for the investigated prediction methods over the block length M in time. Constant and linear prediction show similar prediction quality as with simulated data. The maximum Doppler frequency f_D for the Wiener predictor is derived from the position information obtained during the measurement and the autocorrelation process is not known perfectly. Therefore the Wiener predictor shows degraded quality and approaches the performance of the polynomial approach.

B. Link-Level Simulation

To assess the joint performance and trade-offs for massive MIMO beam-forming (in space), OP (in time-frequency) and CSI prediction using the 5G self-contained slot structure, extensive link-level simulations are performed. Similar to before, a correlated Rayleigh fading channel model according to [16] is assumed with a Jakes Doppler power spectral density. The number of paths is 40 and no correlation across antennas is considered. The maximum Doppler frequency f_D =

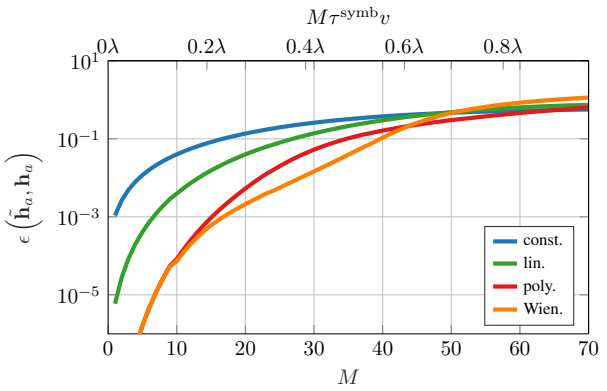

Fig. 3. Relative MSE over the block length M in time for a velocity of 60 km/h at 3.5 GHz for empirical channel data from vehicular measurements.

520 Hz unless noted otherwise, according to a velocity of 160 km/h at 3.5 GHz.

The physical layer is an OFDM system and we simulate $Q = 512$ subcarriers with a spacing of 15 kHz. Of all simulated subcarriers 120 adjacent ones are considered for precoding, corresponding to 10 resource blocks with 12 subcarriers each. As before, the slot duration $\tau^{\text{slot}} = 1$ ms and the symbol duration $\tau^{\text{symb}} = 71.43\,\mu$s. Prediction and precoding are calculated for a block length of $M = 22$ in time. This block length is chosen such that, given 4 SRS symbols are transmitted in the up-link, one whole slot with 14 down-link symbols and one self-contained slot with 8 down-link symbols is predicted and simulated. The remaining 6 symbols of the self-contained second slot are reserved for guard and SRS symbols of the subsequent down-link phase and are therefore not considered (see also Fig. 1).

Perfect up-link channel information is considered, i.e. $\hat{w}_{m,a} = 0$ in (5). The down-link symbols consider an SNR as specified in the plots. An iterative receiver applying a BCJR decoder and soft-symbol feedback as described in [11] is used. The BS is considered to have $A = 64$ antennas and 500 frames are simulated.

Figure 4 shows a bar plot of the simulated BER for all considered prediction methods for four different velocities, three SNR values and no precoding over time and frequency employed, i.e. $\mathbf{d} = \mathbf{b}$ in (1). It is clearly observed that the BER increases for increasing velocities if constant CSI is assumed, and that the BER is only dependent on the SNR if CSI is perfectly known. Linear prediction offers decreased BER for low velocities, but performance quickly deteriorates for higher speeds. If the channel statistics are perfectly known, as is assumed here, Wiener prediction of the CSI significantly increases performance and is largely independent on the velocity as long as the prediction horizon is not larger than approximately 0.6λ.

Figure 5 plots the simulated BERs for the considered prediction methods, and compares no precoding to OP using DSFT. The assumed velocity is 160 km/h and

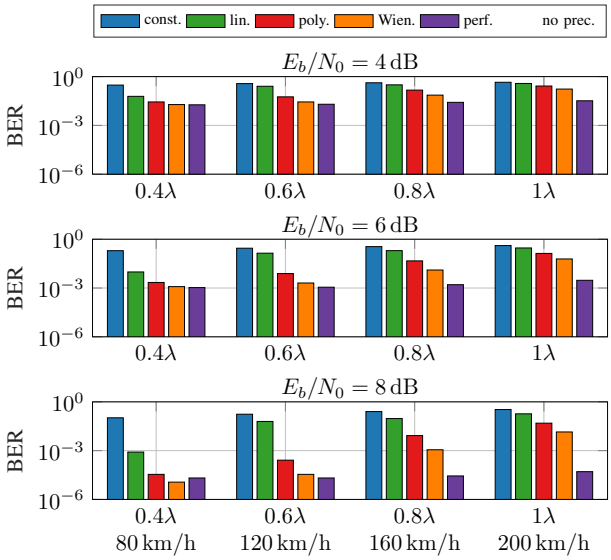

Fig. 4. Bar plot of the simulated BER for a massive MIMO system employing the considered prediction methods without precoding given various velocity and SNR combinations, and the number of BS antennas $A = 64$.

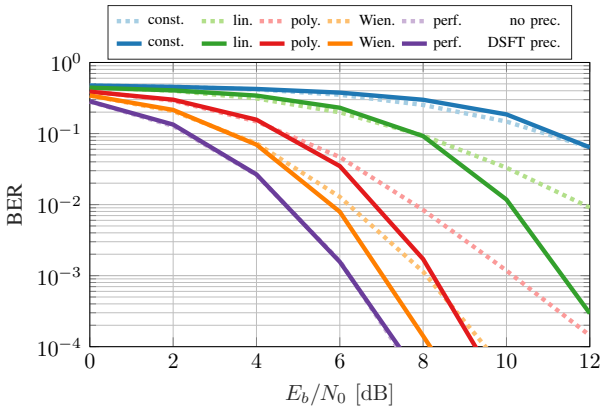

Fig. 5. Simulated BER for a massive MIMO system employing the considered prediction methods without precoding and with DSFT precoding. The prediction horizon is 0.8λ at a velocity of 160 km/h and a frequency of 3.5 GHz, and the number of BS antennas $A = 64$. A significant performance gain is achieved through OP.

channel statistics are known for the Wiener predictor. Precoding in combination with CSI prediction significantly increases performance. The penalty for imperfect channel prediction is roughly 1 dB for the Wiener filter with perfect channel statistic knowledge and roughly 2 dB for the polynomial predictor without any knowledge of the channel. In real-world scenarios, the performance of the Wiener prediction method will be closer to the polynomial one, as the channel statistics are hardly known perfectly and change over time.

It is seen from the dashed lines in Figure 5 that the deployment of polynomial prediction achieves BER performance gains over constant and linear prediction of > 4 dB (Wiener prediction: > 5 dB).

With the same number of overhead SRS symbols, but different SRS symbol repetition rate, Wiener and

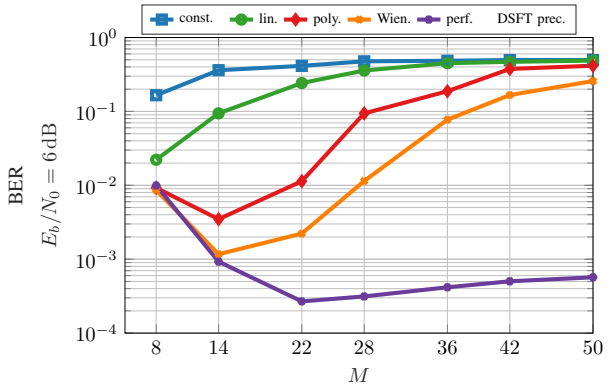

Fig. 6. Simulated BER over block length M in time for $A = 64$ BS antennas, a velocity of $160\,\mathrm{km/h}$ and $E_b/N_0 = 6\,\mathrm{dB}$.

polynomial prediction show lower BER than constant prediction (results not shown).

Figure 6 shows the dependence of the BER on the block length M in time for $A = 64$ BS antennas, a fixed velocity of $160\,\mathrm{km/h}$ and an SNR of $E_b/N_0 = 6\,\mathrm{dB}$ using DSFT OP. We can observe that, for Wiener and polynomial prediction, the lowest BER is not achieved with the smallest prediction horizon $M = 8$ as would be expected without precoding, but with the increased block size $M = 14$. This effect arises due to the increased time-frequency diversity that is harvested by the OP method and that successfully compensates for increased prediction errors towards the end of the slot. We conclude that for a given velocity and the corresponding spatial prediction horizon there exists an optimal block length M that entails small prediction errors and large diversity gains.

V. Conclusion

We investigated the combination of OP and various CSI prediction methods in a massive MIMO single user down-link system. We showed that the prediction quality of simple polynomial prediction is similar to the Wiener filter with empirical data from a vehicular measurement campaign. Link-level simulations showed polynomial and Wiener CSI prediction significantly improves ($> 5\,\mathrm{dB}$) BER performance of our system in time-varying scenarios compared to constant or linear prediction. Additional utilization of OP further decreases the BER and enables approaching the performance of the system with perfect channel knowledge to within $2\,\mathrm{dB}$ ($1\,\mathrm{dB}$) for polynomial (Wiener) prediction for a velocity of $160\,\mathrm{km/h}$. We show there is a trade-off between increased time-frequency diversity utilization and decreased predicted CSI quality for increasing block length in time.

Acknowledgment

This work is funded by the Austrian Research Promotion Agency (FFG) and the Austrian Ministry for Transport, Innovation and Technology (bmvit) within the project MARCONI (861208) of the funding program ICT of the Future.

References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[2] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—What is next?: Five promising research directions for antenna arrays," *Digital Signal Processing: A Review Journal*, vol. 94, pp. 3–20, 2019.

[3] K. Huber and S. Haykin, "Improved Bayesian MIMO channel tracking for wireless communications: Incorporating a dynamical model," *IEEE Transactions on Wireless Communications*, vol. 5, no. 9, pp. 2468–2476, 2006.

[4] A. J. Tenenbaum, R. S. Adve, and Y.-S. Yuk, "Channel prediction and feedback in multiuser broadcast channels," in *2009 11th Canadian Workshop on Information Theory (CWIT)*, 2009, pp. 67–70.

[5] S. Kashyap, C. Mollen, E. Bjornson, and E. G. Larsson, "Performance analysis of (TDD) massive MIMO with Kalman channel prediction," in *International Conference on Acoustics, Speech and Signal Processing*, 2017, pp. 3554–3558.

[6] C. Potter, G. K. Venayagamoorthy, and K. Kosbar, "RNN based MIMO channel prediction," *Signal Processing*, vol. 90, no. 2, pp. 440–450, 2010.

[7] Y. Yang, F. Gao, G. Y. Li, and M. Jian, "Deep Learning-Based Downlink Channel Prediction for FDD Massive MIMO System," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1994–1998, 2019.

[8] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338–351, 2013.

[9] 3GPP, "NR; Physical channels and modulation," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.211, 12 2019, version 15.8.0.

[10] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," in *IEEE Wireless Communications and Networking Conference, WCNC*. IEEE, 2017, pp. 1–6.

[11] T. Zemen, D. Löschenbrand, M. Hofer, C. Pacher, and B. Rainer, "Orthogonally Precoded Massive MIMO for High Mobility Scenarios," *IEEE Access*, vol. 7, pp. 132 979–132 990, 2019.

[12] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case," *Bell System Technical Journal*, vol. 57, no. 5, pp. 1371–1430, 1978.

[13] S. Gunnarsson, J. Flordelis, L. Van Der Perre, and F. Tufvesson, "Channel Hardening in Massive MIMO-A Measurement Based Analysis," in *IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC*, vol. 2018-June. IEEE, 2018, pp. 1–5.

[14] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[15] T. Zemen, C. F. Mecklenbräuker, and B. H. Fleury, "Minimum-energy bandlimited time-variant channel prediction with dynamic subspace selection," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4534–4548, 2007.

[16] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 51, no. 6, pp. 920–928, 2003.

[17] D. Löschenbrand, M. Hofer, L. Bernadó, G. Humer, B. Schrenk, S. Zelenbaba, and T. Zemen, "Distributed Massive MIMO Channel Measurements in Urban Vehicular Scenario," in *13th European Conference on Antennas and Propagation (EuCAP)*, 2019, pp. 1–5.

[18] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3597–3607, 2005.