# Robust Transmit Zero-Forcing Filters

Raphael Hunger, Frank A. Dietrich, Michael Joham, and Wolfgang Utschick

8th ITG Workshop on Smart Antennas (WSA 2004)
Munich, Germany
pp. 130–137, March 18th–19th, 2004



Munich University of Technology
Institute for Circuit Theory and Signal Processing
http://www.nws.ei.tum.de

# Robust Transmit Zero-Forcing Filters

Raphael Hunger, Frank A. Dietrich, Michael Joham, and Wolfgang Utschick
Institute for Circuit Theory and Signal Processing, Munich University of Technology, 80290 Munich, Germany
E-Mail: {Hunger, Dietrich, Joham, Utschick}@nws.ei.tum.de

*Abstract*— **We present linear and non-linear robust transmit *zero-forcing* filters for the downlink of *multi-user multiple-input single-output* (MU-MISO) *time-division-duplex* (TDD) systems which are robust with respect to errors in the *channel state information* (CSI) arising from channel estimation and time lags in mobile communications. Based on a set of estimated CSI of previous uplink slots, we apply a *conditional mean* to the cost function underlying the respective filter for the current downlink slot resulting in channel prediction for the CSI and robust filter structures. Thus, the respective transmit filters are less sensitive to imperfect channel knowledge and show dramatic performance improvements compared to their non-robust counterparts. Our approach can be interpreted as a joint optimization of channel prediction and preequalization. Additionally, we point out the relation of the presented approach to another robust technique named *stochastic programming* and show the analogy to a regularization approach.**

## I. Introduction

Due to the reciprocity of the channel in TDD systems, channel equalization can be transferred from the receiver to the transmitter since uplink and downlink alternate in the same frequency band. Thereby, complexity of the mobile stations can be reduced dramatically, since channel estimation, equalization, and FIR filtering drop out at the receiver side. If both transmitter and receiver are calibrated correctly, instantaneous CSI is available to the *base station* (BS) to a certain degree. Most of the existing transmit filters are based upon perfect CSI, which unfortunately is not available because of channel estimation on the one hand, but mainly due to the movement of the mobile stations on the other hand [1], [2]. Not surprisingly, the performance of these filters rapidly degrades with increasing inaccuracy of the CSI. Our contribution is to develop robust transmit filters which take into account, that the CSI is imperfect. To this end, we modify the cost function underlying the conventional non-robust filter replacing the deterministic channel matrix of the current downlink slot by its respective matrix of random variables and afterwards minimizing the *conditional mean* of the cost function given a set of noisy channel coefficients of all receivers and antennas from previous uplink channel estimations. Interestingly, this robust approach reduces to a *conditional mean prediction* (CMP) of the channel matrix and its Gram. Thereby, the deterministic channel matrix is replaced by its Wiener prediction, whereas the Gram is replaced by the Gram of the Wiener prediction *plus* a regularization term.

This paper is organized as follows: In Section II, we explain the system and channel model, Section III describes the robust transmit filter including the *conditional mean channel prediction*. We show analogies to other robust techniques in Section IV and present simulation results in Section V. The sensitivity of the new approach with respect to errors in the temporal auto-correlation function of the channel is discussed in Section VI. After the evaluation of a similar heuristic robust approach in Section VII, this paper finally concludes with Section VIII.

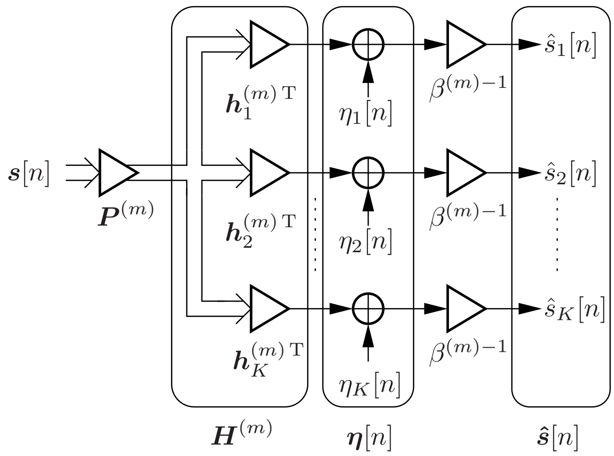


Fig. 1. Downlink for Multi-User System (MU-MISO)

## II. System and Channel Model

Deterministic vectors and matrices are denoted by lower and upper case italic bold letters, whereas the respective random variables are sans serif. The operators $\mathrm{E}[\bullet]$, $\mathrm{E}_{\mathsf{H}}[\bullet]$, $(\bullet)^{\mathrm{T}}$, $(\bullet)^*$, $(\bullet)^{\mathrm{H}}$, and $\mathrm{tr}(\bullet)$ stand for expectation with respect to symbols and noise, expectation with respect to the channel, transpose, complex conjugate, Hermitian transpose, and trace of a matrix, respectively. The scalar element $a_{b,c}$ of the matrix $\boldsymbol{A}$ in row $b$ and column $c$ is denoted by $[\boldsymbol{A}]_{b,c}$, and $\boldsymbol{I}_K$ stands for the $K \times K$ identity matrix. Additionally, we make use of the Kronecker-Delta $\delta_{a,b}$, which vanishes for $a \neq b$ and returns 1, if $a = b$.

Fig. 1 shows the downlink of the frequency flat MU-MISO TDD system. Such a scenario is commonly termed as a system with broadcast channel [3] or as *decentralized receivers* (e.g. [4], [5], [6]). We use $\boldsymbol{s}[n] \in \mathbb{C}^K$, $\boldsymbol{\eta}[n] \in \mathbb{C}^K$, $\beta^{(m)} \in \mathbb{R}_+$, $\boldsymbol{P}^{(m)} \in \mathbb{C}^{N_\mathrm{a} \times K}$, and

$$\boldsymbol{H}^{(m)} = [\boldsymbol{h}_1^{(m)}, \boldsymbol{h}_2^{(m)}, \ldots, \boldsymbol{h}_K^{(m)}]^{\mathrm{T}} \in \mathbb{C}^{K \times N_\mathrm{a}}$$

for the $n$-th symbol vector, the noise vector, the scalar weight for the receivers, the transmit filter, and the channel matrix of time slot $m$, respectively. Each of the $K$ receivers is equipped with a single antenna, whereas $N_\mathrm{a}$ antenna elements are deployed at the BS. Furthermore, we assume the channel to have a temporally correlated block fading, which means that the realization $\boldsymbol{H}^{(m)}$ of the random channel matrix $\mathbf{H}^{(m)}$ for time slot $m$ is assumed to be constant within this slot. To simplify notation, the elements $\mathsf{h}_{k,a}^{(m)} = [\mathbf{H}^{(m)}]_{k,a}$ of $\mathbf{H}^{(m)}$ are zero-mean independent complex Gaussian random variables with

$$\mathsf{h}_{k,a}^{(m)} \sim \mathcal{CN}\left(0; \sigma_{\mathsf{h}_k}^2\right) \;\; \forall a \in \{1, ..., N_\mathrm{a}\} \;\; \forall m \in \mathbb{Z}, \tag{1}$$

the variance $\sigma_{\mathsf{h}_k}^2$ being defined by

$$\sigma_{\mathsf{h}_k}^2 = \mathrm{E}_{\mathbf{H}}\left[|\mathsf{h}_{k,a}^{(m)}|^2\right]. \tag{2}$$

Moreover, we assume a Jakes power density spectrum, leading to a temporal auto-correlation function for receiver $k$, which reads as [7], [8]

$$\mathrm{E}_{\mathbf{H}}\left[\mathsf{h}_{k,a}^{(m-\ell)} \mathsf{h}_{k,a}^{(m)*}\right] = \sigma_{\mathsf{h}_k}^2 \mathrm{J}_0\left(2\pi \frac{f_{\mathrm{D}_k}}{f_\mathrm{slot}} \ell\right), \tag{3}$$

for all antennas $a$, with $\mathrm{J}_0(\bullet)$ denoting the Bessel function of the first kind of order zero, $f_{\mathrm{D}_k}$ being the *maximum* Doppler frequency of receiver $k$, and $f_\mathrm{slot}$ denoting the slot rate depending on the frame format, $f_\mathrm{slot} = 1500$ Hz is assumed for the simulations. We focus on a frame format, where uplink and downlink slots are alternating, i.e. no two successive uplink or downlink slots exist. The BS cannot access the estimated channel coefficients $\hat{\boldsymbol{H}}^{(m-1)}$ of the directly preceding uplink slot since the channel estimation has not been performed yet. Only CSI from slots $m-3$, $m-5$, $m-7$ and so on is available [2], which corresponds to the worst-case processing delay. Due to the alternating slots within one frame, these distances $\ell$ (cf. Eq. 3) stay constant for all $m$.

## III. Robust Transmit Filtering

Before introducing and deriving the robust version (RTxZF) of the linear transmit *zero-forcing* filter and the robust version (RTxZF THP) of the non-linear transmit *zero-forcing* filter with spatial *Tomlinson-Harashima Precoding* (THP), we briefly review the non-robust versions, i.e. the TxZF and the TxZF THP, for a better understanding of the modified cost function.

### A. Conventional Linear Transmit Zero-Forcing Filter

The conventional linear TxZF assumes perfect CSI and results from the minimization of the *mean square error* (MSE)

$$\varepsilon(\boldsymbol{P}, \beta) = \mathrm{E}\left[\|\mathbf{s}[n] - \beta^{-1}(\boldsymbol{H}^{(m)}\boldsymbol{P}\mathbf{s}[n] + \mathbf{n}[n])\|_2^2\right] \tag{4}$$

constrained to the removal of the complete *multiple access interference* (MAI). Moreover, the emerging *average* transmit power $\mathrm{E}[\|\boldsymbol{P}\mathbf{s}[n]\|_2^2] = \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_\mathbf{s}\boldsymbol{P}^\mathrm{H})$ may not exceed the limit $E_\mathrm{tr}$ [9], [10], since only finite transmit power is available:

$$\begin{aligned} &\{\boldsymbol{P}_\mathrm{ZF}^{(m)}, \beta_\mathrm{ZF}^{(m)}\} = \underset{\boldsymbol{P},\beta}{\mathrm{argmin}}\, \varepsilon(\boldsymbol{P}, \beta) \\ &\mathrm{s.t.}: \beta^{-1}\boldsymbol{H}^{(m)}\boldsymbol{P} = \boldsymbol{I}_K \;\text{ and }\; \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_\mathbf{s}\boldsymbol{P}^\mathrm{H}) \le E_\mathrm{tr}, \end{aligned} \tag{5}$$

where $\boldsymbol{R}_\mathbf{s} = \mathrm{E}[\mathbf{s}[n]\mathbf{s}^\mathrm{H}[n]]$. Note that the MSE reduces to $\beta^{-2}\,\mathrm{tr}(\boldsymbol{R}_\mathbf{n})$ when we insert the first constraint of (5) into (4), and $\boldsymbol{R}_\mathbf{n} = \mathrm{E}[\mathbf{n}[n]\mathbf{n}^\mathrm{H}[n]]$. Equivalently, we could minimize $\beta^{-2}$ constrained to the removal of the complete MAI and to the limited average transmit power, since dropping the constant scalar factor $\mathrm{tr}(\boldsymbol{R}_\mathbf{n})$ in the cost function does not change the optimum solution $\boldsymbol{P}_\mathrm{ZF}^{(m)}$ and $\beta_\mathrm{ZF}^{(m)}$. Solving (5) yields

$$\boldsymbol{P}_\mathrm{ZF}^{(m)} = \beta_\mathrm{ZF}^{(m)} \boldsymbol{H}^{(m)\,\mathrm{H}}\left(\boldsymbol{H}^{(m)}\boldsymbol{H}^{(m)\,\mathrm{H}}\right)^{-1} \in \mathbb{C}^{N_\mathrm{a}\times K}, \tag{6}$$

and $\beta_\mathrm{ZF}^{(m)}$ is chosen to fulfill the equality of the power constraint in (5), i.e.

$$\beta_\mathrm{ZF}^{(m)} = \sqrt{\frac{E_\mathrm{tr}}{\mathrm{tr}\left((\boldsymbol{H}^{(m)}\boldsymbol{H}^{(m)\,\mathrm{H}})^{-1}\boldsymbol{R}_\mathbf{s}\right)}} \in \mathbb{R}_+, \tag{7}$$

so the maximum value $E_\mathrm{tr}$ is dissipated on average.

### B. Robust Linear Transmit Zero-Forcing Filter

Since the channel is not perfectly known in practice for the reasons mentioned in Section I, the MAI cannot be suppressed completely and the first constraint in (5) cannot be fulfilled any longer. Therefore, we try to reject the MAI as good as possible which corresponds to the main idea of the *zero-forcing* approach. Unfortunately, we cannot apply $\hat{\boldsymbol{H}}^{(m-1)}$ to the solution in (6) and (7) because the channel estimation has not been performed yet. According to the frame structure, time slot $m-2$ is allocated to the downlink leading to the fact, that no information about the channel is present. Non-robust transmit filters therefore would have to access estimated CSI from time slot $m-3$. But instead of employing the estimated channel coefficients $\hat{\boldsymbol{H}}^{(m-3)}$ for the filter in (6) and (7) as if they perfectly described the transmission over the channel for time slot $m$, we first store $p$ preceding estimated channel coefficients in $\hat{\boldsymbol{\mathcal{H}}}^{(m)}$. Then, we set up a new metric

$$\varepsilon'(\boldsymbol{P}, \beta) = \mathrm{E}_{\mathbf{H}}\left[\mathrm{E}\left[\|\mathbf{s}[n] - \beta^{-1}\mathbf{H}^{(m)}\boldsymbol{P}\mathbf{s}[n]\|_2^2\right] \middle| \hat{\boldsymbol{\mathcal{H}}}^{(m)}\right], \tag{8}$$

which is a measure of interference *on average* and is similar to the *average* deviation of the *zero-forcing* constraint in (5), but now we allow for a bias. In [11], Rey et al. proposed a similar approach for joint transmitter and receiver MMSE optimization in OFDM systems where they exploited the correlations between different subcarriers to reduce uncertainties in CSI. Contrary to the approach of the conventional TxZF in (5), we neglect the noise contributions at the receivers (cf. the definition of $\varepsilon(\boldsymbol{P}, \beta)$ in Eq. 4), but add an expectation with respect to the channel given the set of observed channel coefficients $\hat{\boldsymbol{\mathcal{H}}}^{(m)}$ motivated by the robust *least squares* (LS) solution in [12]. The desired linear robust transmit *zero-forcing* filter can then be found solving

$$\begin{aligned} &\{\boldsymbol{P}_\mathrm{RZF}^{(m)}, \beta_\mathrm{RZF}^{(m)}\} = \underset{\boldsymbol{P},\beta}{\mathrm{argmin}}\, \varepsilon'(\boldsymbol{P}, \beta) \\ &\mathrm{s.t.}: \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_\mathbf{s}\boldsymbol{P}^\mathrm{H}) = E_\mathrm{tr}. \end{aligned} \tag{9}$$

Different from (5), the equality of the power constraint already has to be enforced as part of the optimization, otherwise $\beta_\mathrm{RZF}^{(m)}$ would only be upper bounded but would not have a unique solution (inserting the optimum solution $\boldsymbol{P}_\mathrm{RZF}^{(m)}$ into (8) will finally reveal, that $\varepsilon'(\boldsymbol{P},\beta)$ does not depend on $\beta$). Note that the *conditional mean* in the new cost function $\varepsilon'(\boldsymbol{P},\beta)$ reduces to a CMP for the channel matrix $\mathsf{H}^{(m)}$ and its Gram, which becomes clear, if we express (8) by

$$\begin{aligned}\varepsilon'(\boldsymbol{P},\beta) &= \mathrm{tr}(\boldsymbol{R}_{\mathsf{s}}) \\ &\quad - \beta^{-1}\,\mathrm{tr}\left(\boldsymbol{R}_{\mathsf{s}}\boldsymbol{P}^{\mathrm{H}}\mathrm{E}_{\mathsf{H}}\left[\mathsf{H}^{(m)\,\mathrm{H}}\,\middle|\,\hat{\boldsymbol{\mathcal{H}}}^{(m)}\right]\right) \\ &\quad - \beta^{-1}\,\mathrm{tr}\left(\mathrm{E}_{\mathsf{H}}\left[\mathsf{H}^{(m)}\,\middle|\,\hat{\boldsymbol{\mathcal{H}}}^{(m)}\right]\boldsymbol{P}\boldsymbol{R}_{\mathsf{s}}\right) \\ &\quad + \beta^{-2}\,\mathrm{tr}\left(\mathrm{E}_{\mathsf{H}}\left[\mathsf{H}^{(m)\,\mathrm{H}}\mathsf{H}^{(m)}\,\middle|\,\hat{\boldsymbol{\mathcal{H}}}^{(m)}\right]\boldsymbol{P}\boldsymbol{R}_{\mathsf{s}}\boldsymbol{P}^{\mathrm{H}}\right).\end{aligned} \tag{10}$$

For the computation of the *conditional mean*, we make the simplifying assumption, that all channel coefficients are uncorrelated. Thus, we only have to focus on a single coefficient $\mathsf{h}_{k,a}^{(m)} = [\mathsf{H}^{(m)}]_{k,a}$ from antenna $a$ to receiver $k$ representative for all antennas $a \in \{1,\ldots,N_{\mathrm{a}}\}$ and all receivers $k \in \{1,\ldots,K\}$. We store $p$ preceding estimated channel coefficients in

$$\hat{\boldsymbol{h}}_{\mathrm{T},k,a}^{(m)} = [\hat{h}_{k,a}^{(m-3)}, \hat{h}_{k,a}^{(m-5)}, \ldots, \hat{h}_{k,a}^{(m-2p-1)}]^{\mathrm{T}} \in \mathbb{C}^{p},$$

which is part of $\hat{\boldsymbol{\mathcal{H}}}^{(m)}$. The estimated vector $\hat{\boldsymbol{h}}_{\mathrm{T},k,a}^{(m)}$ can be expressed by the sum of the original channel coefficients plus an error due to imperfect channel estimation (as in a LS channel estimation for example), i.e. $\hat{\boldsymbol{h}}_{\mathrm{T},k,a}^{(m)} = \boldsymbol{h}_{\mathrm{T},k,a}^{(m)} + \boldsymbol{\eta}_{\mathrm{p},k,a}^{(m)}$. We assume a joint multivariate complex Gaussian probability density function for $\mathsf{h}_{k,a}^{(m)}$ and $\hat{\mathbf{h}}_{\mathrm{T},k,a}^{(m)} = \mathbf{h}_{\mathrm{T},k,a}^{(m)} + \mathbf{n}_{\mathrm{p},k,a}^{(m)}$, which reads as (cf. Eq. 3)

$$\begin{bmatrix} \mathsf{h}_{k,a}^{(m)} \\ \hat{\mathbf{h}}_{\mathrm{T},k,a}^{(m)} \end{bmatrix} \sim \mathcal{CN}\left(\mathbf{0}; \sigma_{\mathsf{h}_k}^2 \begin{bmatrix} 1 & \boldsymbol{r}_k^{\mathrm{H}} \\ \boldsymbol{r}_k & \boldsymbol{R}_k \end{bmatrix}\right), \tag{11}$$

where $\boldsymbol{r}_k = \mathrm{E}_{\mathsf{H}}[\mathbf{h}_{\mathrm{T},k,a}^{(m)}\mathsf{h}_{k,a}^{(m)*}]/\sigma_{h_k}^2$ denotes the normalized correlation between the channel coefficient of time slot $m$ and $p$ preceding channel realizations ($\mathsf{h}_{k,a}^{(m)}$ and $\mathbf{n}_{\mathrm{LS},k,a}^{(m)}$ are uncorrelated and zero-mean). Furthermore,

$$\boldsymbol{R}_k = \mathrm{E}_{\mathsf{H}}[(\mathbf{h}_{\mathrm{T},k,a}^{(m)} + \mathbf{n}_{\mathrm{p},k,a}^{(m)})(\mathbf{h}_{\mathrm{T},k,a}^{(m)} + \mathbf{n}_{\mathrm{p},k,a}^{(m)})^{\mathrm{H}}]/\sigma_{h_k}^2 \in \mathbb{C}^{p\times p}$$

denotes the normalized covariance matrix of the (noisy) uplink estimates. Applying the rule of Bayes to (11), the *conditional* distribution of $\mathsf{h}_{k,a}^{(m)}$ given the observed values $\hat{\mathbf{h}}_{\mathrm{T},k,a}^{(m)} = \hat{\boldsymbol{h}}_{\mathrm{T},k,a}^{(m)}$ again has a complex Gaussian probability density function, whose mean value now does not vanish [13]:

$$\mathsf{h}_{k,a}^{(m)}\,\middle|\,\hat{\boldsymbol{h}}_{\mathrm{T},k,a}^{(m)} \sim \mathcal{CN}\left(\overline{h_{k,a}^{(m)}}; \zeta_k\right). \tag{12}$$

The *conditional mean* $\overline{h_{k,a}^{(m)}}$ is defined by

$$\overline{h_{k,a}^{(m)}} = \mathrm{E}_{\mathsf{H}}\left[\mathsf{h}_{k,a}^{(m)}\,\middle|\,\hat{\boldsymbol{h}}_{\mathrm{T},k,a}^{(m)}\right] = \boldsymbol{r}_k^{\mathrm{H}}\boldsymbol{R}_k^{-1}\hat{\boldsymbol{h}}_{\mathrm{T},k,a}, \tag{13}$$

and equals the Wiener prediction for the channel coefficient $h_{k,a}^{(m)}$, since we assumed zero-mean Gaussian pdfs. With (2), the variance $\zeta_k$ of the *conditional* pdf is defined by the Schur complement of $\sigma_{\mathsf{h}_k}^2$ and reads as [13]

$$\zeta_k = \mathrm{E}_{\mathsf{H}}\left[|\mathsf{h}_{k,a}^{(m)} - \overline{h_{k,a}^{(m)}}|^2\,\middle|\,\hat{\boldsymbol{h}}_{\mathrm{T},k,a}^{(m)}\right] = \sigma_{\mathsf{h}_k}^2(1 - \boldsymbol{r}_k^{\mathrm{H}}\boldsymbol{R}_k^{-1}\boldsymbol{r}_k). \tag{14}$$

The CMP for $|\mathsf{h}_{k,a}^{(m)}|^2$ is simply the sum of the conditional variance $\zeta_k$ and the magnitude of the squared conditional mean $|\overline{h_{k,a}^{(m)}}|^2$. Generalizing these results for the multi-user and multi-antenna case, we obtain by means of the matrix $\boldsymbol{\zeta} = \mathrm{diag}\{\zeta_k\}_{k=1}^{K} \in \mathbb{R}^{K\times K}$

$$\mathrm{E}_{\mathsf{H}}\Big[\mathsf{H}^{(m)}\,\Big|\,\hat{\boldsymbol{\mathcal{H}}}^{(m)}\Big] = \overline{\boldsymbol{H}^{(m)}}, \tag{15}$$

$$\begin{aligned}\mathrm{E}_{\mathsf{H}}\Big[\mathsf{H}^{(m)\,\mathrm{H}}\mathsf{H}^{(m)}\,\Big|\,\hat{\boldsymbol{\mathcal{H}}}^{(m)}\Big] &= \overline{\boldsymbol{H}^{(m)\,\mathrm{H}}\boldsymbol{H}^{(m)}} \\ &= \overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}\;\overline{\boldsymbol{H}^{(m)}} + \boldsymbol{I}_{N_{\mathrm{a}}}\mathrm{tr}(\boldsymbol{\zeta}).\end{aligned} \tag{16}$$

The matrix $\overline{\boldsymbol{H}^{(m)}}$ simply consists of the predicted entries of $\boldsymbol{H}^{(m)}$, i.e. $[\overline{\boldsymbol{H}^{(m)}}]_{k,a} = \overline{h_{k,a}^{(m)}}$. The *conditional mean* of the Gram $\mathsf{H}^{(m)\,\mathrm{H}}\mathsf{H}^{(m)}$ of $\mathsf{H}^{(m)}$ leads to the Gram of $\overline{\boldsymbol{H}^{(m)}}$ *plus* a scaled identity matrix, since the main diagonal consists of squared magnitudes of the channel coefficients, i.e.

$$\Big[\mathsf{H}^{(m)\,\mathrm{H}}\mathsf{H}^{(m)}\Big]_{a,a} = \sum_{k=1}^{K}|\mathsf{h}_{k,a}^{(m)}|^2 \quad \forall a \in \{1,\ldots,N_{\mathrm{a}}\}. \tag{17}$$

The error variance in (12) only has to be taken into account if indices $k,l$ and $a,b$ are identical, i.e. $k=l$ and $a=b$, since different channel coefficients are assumed to be uncorrelated in our scenario:

$$\mathrm{E}_{\mathsf{H}}\left[\mathsf{h}_{k,a}^{(m)}\mathsf{h}_{l,b}^{(m)*}\,\middle|\,\hat{\boldsymbol{\mathcal{H}}}^{(m)}\right] = \overline{h_{k,a}^{(m)}}\;\overline{h_{l,b}^{(m)*}} + \delta_{k,l}\delta_{a,b}\zeta_k. \tag{18}$$

Inserting (15) and (16) into (10) and solving (9) for $\boldsymbol{P}$, we find the solution for the *robust* linear transmit *zero-forcing* filter

$$\boldsymbol{P}_{\mathrm{RZF}}^{(m)} = \beta_{\mathrm{RZF}}^{(m)}\Big(\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}\;\overline{\boldsymbol{H}^{(m)}} + \boldsymbol{I}_{N_{\mathrm{a}}}\,\mathrm{tr}(\boldsymbol{\zeta})\Big)^{-1}\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}, \tag{19}$$

where $\beta_{\mathrm{RZF}}^{(m)}$ is again chosen to meet the power constraint. From (19) we can conclude, that the RTxZF filter is similar in structure to the conventional transmit Wiener filter [9], [14]. Hence, the uncertainty in CSI can be regarded as an equivalent noise source.

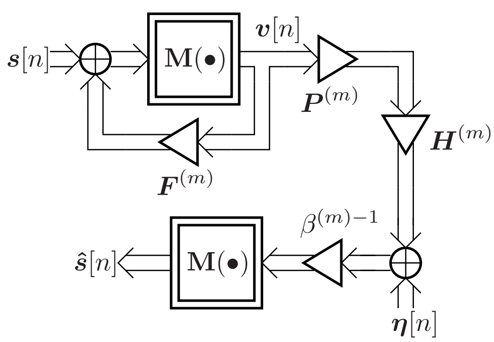


Fig. 2. Downlink for THP MU-MISO System

### C. Conventional Non-Linear Transmit Zero-Forcing Filter with Tomlinson-Harashima Precoding

Different from the originally intended purpose of the *Tomlinson-Harashima* Precoding in [15], [16], where intersymbol-interference in a *single-user single-input single-output* system is eliminated by a non-linear recursive structure at the transmitter, this approach can also be applied to a MU-MISO system in the frequency flat case. For THP systems, both transmitter and receivers have to be equipped with a non-linear *modulo* device $\mathrm{M}(\bullet)$, which on the transmitter side prevents from a power increase and finally revokes the modulo operation of the transmitter at the receivers. Unfortunately, this operation at the receivers leads to an infinite repetition of the signal constellation in the complex plane and implicates the generation of new neighbours [5], [6], [17]. Thus, noise contributions cause an increased *bit error probability* in *zero-forcing* systems for low *signal-to-noise ratios* compared to the linear precoder.

Fig. 2 shows the downlink of the MU-MISO TDD system, when THP is used. In addition to the *modulo* device $\mathrm{M}(\bullet)$, the transmitter is extended by a spatial feedback filter $\boldsymbol{F}^{(m)} \in \mathbb{C}^{K\times K}$, which feeds back the already precoded symbols $v_i[n] = [\boldsymbol{v}[n]]_{i,1} \in \mathbb{C}$. Since only already precoded symbols $v_i[n]$ can be fed back, the structure of $\boldsymbol{F}^{(m)}$ has to be lower triangular with zero main diagonal, as we assume an ordered channel matrix for the derivation. An attractive sub-optimum approach for the sorting order of the precoding can be found in [17] and will be used for the simulations. Because of the *modulo* operation at the transmitter, the statistic of the precoded symbols $v_i[n]$ differs from the statistic of the data symbols $s_k[n]$ (e.g. [5], [6], [17], [18]) and the emerging *average* transmit power reads as $\mathrm{E}[\|\boldsymbol{P}\mathbf{v}[n]\|_2^2] = \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_{\mathbf{v}}\boldsymbol{P}^{\mathrm{H}})$, where $\boldsymbol{R}_{\mathbf{v}} \in \mathbb{R}^{K\times K}$ has diagonal structure and $[\boldsymbol{R}_{\mathbf{v}}]_{1,1} = [\boldsymbol{R}_{\mathbf{s}}]_{1,1}$ and $[\boldsymbol{R}_{\mathbf{v}}]_{i,i} = \frac{M}{M-1}[\boldsymbol{R}_{\mathbf{s}}]_{i,i}\ \forall i \in \{2,\ldots,K\}$. Here, $M$ denotes the cardinality of the modulation alphabet (e.g. $M=4$ for QPSK, $M=16$ for 16-QAM). In order to be able to apply the framework of linear algebra as we did for the linear TxZF, we define the optimization criterion *in front of* the *modulo* operation at the receivers. Thus, the *mean square error* reads as

$$\varepsilon(\boldsymbol{P},\boldsymbol{F},\beta) = \mathrm{E}\left[\|\mathbf{d}[n] - \beta^{-1}(\boldsymbol{H}^{(m)}\boldsymbol{P}\mathbf{v}[n] + \mathbf{n}[n])\|_2^2\right], \quad (20)$$

where we replaced the data symbol vector $\mathbf{s}[n]$ by the desired symbol vector $\mathbf{d}[n] = (\boldsymbol{I}_K - \boldsymbol{F})\mathbf{v}[n]$ and by the precoded symbol vector $\mathbf{v}[n]$, respectively (cf. Fig. 3). The TxZF THP then follows from the following optimization (sorted channel matrix assumed):

$$\begin{aligned}
&\{\boldsymbol{P}_{\mathrm{ZF}}^{(m)}, \boldsymbol{F}_{\mathrm{ZF}}^{(m)}, \beta_{\mathrm{ZF}}^{(m)}\} = \underset{\boldsymbol{P},\boldsymbol{F},\beta}{\mathrm{argmin}}\, \varepsilon(\boldsymbol{P},\boldsymbol{F},\beta) \\
&\mathrm{s.t.:}\ \beta^{-1}\boldsymbol{H}^{(m)}\boldsymbol{P} = \boldsymbol{I}_K - \boldsymbol{F},\ \text{and} \\
&\qquad \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_{\mathbf{v}}\boldsymbol{P}^{\mathrm{H}}) \leq E_{\mathrm{tr}},\ \text{and} \\
&\qquad \boldsymbol{F}\ \text{lower triangular, zero main diagonal.}
\end{aligned} \quad (21)$$

With Lagrangian multipliers, we find the solution

$$\begin{aligned}
\boldsymbol{P}_{\mathrm{ZF}}^{(m)} &= \beta_{\mathrm{ZF}} \sum_{k=1}^{K} \boldsymbol{H}^{(m)}\boldsymbol{S}_k^{\mathrm{T}}\boldsymbol{A}_{k,\mathrm{ZF}}^{(m)-1}\boldsymbol{S}_k\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}, \\
\boldsymbol{F}_{\mathrm{ZF}}^{(m)} &= \sum_{k=1}^{K}(\boldsymbol{e}_k - \boldsymbol{H}^{(m)}\boldsymbol{H}^{(m)\,\mathrm{H}}\boldsymbol{S}_k^{\mathrm{T}}\boldsymbol{A}_{k,\mathrm{ZF}}^{(m)-1}\boldsymbol{S}_k\boldsymbol{e}_k)\boldsymbol{e}_k^{\mathrm{T}},
\end{aligned} \quad (22)$$

where $\beta_{\mathrm{ZF}}$ is chosen to meet the equality of the power constraint and $\boldsymbol{A}_{k,\mathrm{ZF}}^{(m)} = \boldsymbol{S}_k\boldsymbol{H}^{(m)}\boldsymbol{H}^{(m)\,\mathrm{H}}\boldsymbol{S}_k^{\mathrm{T}} \in \mathbb{C}^{k\times k}$. The $k$-th column of the $K\times K$ identity matrix $\boldsymbol{I}_K$ is denoted by $\boldsymbol{e}_k$, and $\boldsymbol{S}_k = [\boldsymbol{I}_k, \boldsymbol{0}] \in \{0;1\}^{k\times K}$ is a selection matrix. For a detailed description, see [17].

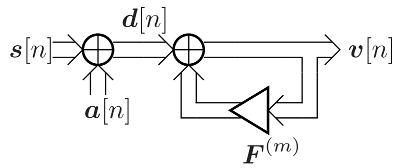


Fig. 3. Linear representation of the *modulo* device: the signal $\boldsymbol{a}[n]$ is added to restrict the real and imaginary part of $\boldsymbol{v}[n]$ to the interval specified by the *modulo* operation

### D. Robust Non-Linear Transmit Zero-Forcing Filter with Tomlinson-Harashima Precoding

For the *robust* version of the *zero-forcing* THP filter, we proceed in the same manner as we did for the linear filter. To this end, we replace the cost function $\varepsilon(\boldsymbol{P},\boldsymbol{F},\beta)$ by our new metric $\varepsilon'(\boldsymbol{P},\boldsymbol{F},\beta)$ (see also Eq. 8):

$$\varepsilon'(\boldsymbol{P},\boldsymbol{F},\beta) = \mathrm{E}_{\mathbf{H}}\Big[\mathrm{E}\Big[\|\mathbf{d}[n] - \beta^{-1}\mathbf{H}^{(m)}\boldsymbol{P}\mathbf{v}[n]\|_2^2\Big|\hat{\boldsymbol{\mathcal{H}}}^{(m)}\Big]\Big] \quad (23)$$

The minimization of this metric has again to be done constrained to a maximum average power emission and the lower triangular structure of $\boldsymbol{F}$ with zero main diagonal:

$$\begin{aligned}
&\{\boldsymbol{P}_{\mathrm{RZF}}^{(m)}, \boldsymbol{F}_{\mathrm{RZF}}^{(m)}, \beta_{\mathrm{RZF}}^{(m)}\} = \underset{\boldsymbol{P},\boldsymbol{F},\beta}{\mathrm{argmin}}\, \varepsilon'(\boldsymbol{P},\boldsymbol{F},\beta) \\
&\mathrm{s.t.:}\ \mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_{\mathbf{v}}\boldsymbol{P}^{\mathrm{H}}) = E_{\mathrm{tr}},\ \text{and} \\
&\qquad \boldsymbol{F}\ \text{lower triangular, zero main diagonal.}
\end{aligned} \quad (24)$$

The solution of (24) reveals, that the *robust* version of the non-linear *zero-forcing* THP filter again resembles the structure of the non-linear Wiener THP filter, where the loading term for the noise at the receivers is again replaced by a loading which depends on the uncertainties of CSI.

$$\begin{aligned}
\boldsymbol{P}_{\mathrm{RZF}}^{(m)} &= \beta_{\mathrm{RZF}} \sum_{k=1}^{K} \boldsymbol{A}_{k,\mathrm{RZF}}^{(m)-1}\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}, \\
\boldsymbol{F}_{\mathrm{RZF}}^{(m)} &= \sum_{k=1}^{K}(\boldsymbol{S}_k^{\mathrm{T}}\boldsymbol{S}_k - \boldsymbol{I}_K)\overline{\boldsymbol{H}^{(m)}}\boldsymbol{A}_{k,\mathrm{RZF}}^{(m)-1}\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}.
\end{aligned} \quad (25)$$

In turn, $\beta_{\mathrm{RZF}}$ has to fulfill the first constraint in (24) and $\boldsymbol{A}_{k,\mathrm{RZF}}^{(m)} = \overline{\boldsymbol{H}^{(m)\,\mathrm{H}}\boldsymbol{S}_k^{\mathrm{T}}\boldsymbol{S}_k\boldsymbol{H}^{(m)}} + \boldsymbol{I}_{N_{\mathrm{a}}}\,\mathrm{tr}(\boldsymbol{\zeta}) \in \mathbb{C}^{N_{\mathrm{a}}\times N_{\mathrm{a}}}$. Since the feedback and the feedforward filter in (25) are different from those in (22), the sub-optimum ordering strategy may also lead to a different precoding order.

## IV. Classification of the Presented Robust Approach and its Relationship to other Robust Techniques

From the point of view of the transmitter the channel $\mathbf{H}^{(m)}$ is a random variable. He has access to its out-dated realization $\boldsymbol{H}^{(m-\ell)}$ via pilot symbols received from the uplink. Modeling the channel as a random variable instead of known deterministic parameter results in a random cost function and stochastic linear equalities as *zero-forcing* constraints (Eq. 5 and 21). Obviously, complete *zero-forcing* cannot be achieved if perfect CSI is not available, and strict "*zero-forcing*" based on imperfect CSI is not desired. Thus, the constraint should be relaxed, i.e. some interference may be allowed. One way to achieve this is via worst case constraints as in [19] or [20] for a deterministic error model $\boldsymbol{H}^{(m)} = \hat{\boldsymbol{H}}^{(m-\ell)} + \boldsymbol{E}^{(m)}$. Alternatively, for a stochastic error model $\mathbf{H}^{(m)} = \hat{\boldsymbol{H}}^{(m-\ell)} + \mathbf{E}^{(m)}$ the principle of chance programming can be employed [21], which allows a certain deviation from the exact equality with an a priori chosen probability. The weakness of both approaches is that a heuristic is required to choose the new free parameters a priori. These approaches typically result in a significantly increased complexity compared to the non-robust design.

Our approach formulates the relaxation of the *zero-forcing* constraint in the cost function $\bar{\varepsilon}(\boldsymbol{P}, \beta, \mathbf{H}^{(m)}) = \mathrm{E}[\|\mathbf{s}[n] - \beta^{-1}\mathbf{H}^{(m)}\boldsymbol{P}\mathbf{s}[n]\|_2^2]$ itself, which is now a random variable in the unknown channel. Thus, we allow for a minimum amount of interference. Solving this optimization problem for a given channel realization $\boldsymbol{H}^{(m)}$ yields an "ill-posed" problem, i.e. the solution is not unique [22]. Ill-posed problems are commonly solved bringing additional side information into the formulation of the problem. This is done in (8) and (23) following the Bayesian philosophy: Interference is minimized on average given the knowledge about the channel from previous channel estimates. The conditional mean in (8) and (23) yields a well-posed problem with a unique solution. It adds a regularization term to the cost function, e.g. as in (8)

$$\varepsilon'(\boldsymbol{P}, \beta) = \bar{\varepsilon}(\boldsymbol{P}, \beta, \overline{\boldsymbol{H}^{(m)}}) + \underbrace{\mathrm{tr}(\boldsymbol{\zeta})\beta^{-2}\mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_{\mathbf{s}}\boldsymbol{P}^{\mathrm{H}})}_{\text{regularization term}}, \quad (26)$$

reducing the Frobenius norm of the solution $\beta^{-1}\boldsymbol{P}$. The second summand in (26) regularizes the solution $\beta^{-1}\boldsymbol{P}$ in the sense of Tikhonov [22], which results in a diagonal loading of the standard zero forcing solution with optimum loading factor $\mathrm{tr}(\boldsymbol{\zeta})$ (regularization parameter) given by the model (cf. Eq. 19). This decreases the sensitivity of the cost function w.r.t. parameter uncertainties [23], [24], i.e. a more robust solution is obtained.

The advantage of our approach is that the error model is given by the second order statistics of the error in (12) and (14). Thus, the robust precoder can be adapted to the current scenario. This way we avoid a performance degradation of the robust design in the nominal case [24], [25].

Given the robust solutions (19) and (25), the problem statement (9) and (24), respectively, can be interpreted as a joint optimization of the predictor and precoder (transmit filter).

Alternatively, we could assume a stochastic error model $\mathbf{H}^{(m)} = \hat{\boldsymbol{H}}^{(m-\ell)} + \mathbf{E}^{(m)}$ and take the expected value w.r.t. the error $\mathbf{E}^{(m)}$ instead of the conditional mean (e.g. Eq. 8), which is a standard method in static stochastic programming [21]. This cost function for linear ZF

$$\begin{aligned}\mathrm{E}_{\mathbf{E}}[\bar{\varepsilon}(\boldsymbol{P}, \beta, \overline{\boldsymbol{H}^{(m)}} + \mathbf{E}^{(m)})] = {} & \bar{\varepsilon}(\boldsymbol{P}, \beta, \overline{\boldsymbol{H}^{(m)}}) \\ & + \beta^{-2}\mathrm{tr}(\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}}\boldsymbol{P}\boldsymbol{R}_{\mathbf{s}}\boldsymbol{P}^{\mathrm{H}})\end{aligned} \quad (27)$$

is identical to (26) if an LMMSE predictor as in (15) is used a priori for prediction. The error covariance matrix is $\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}} = \mathrm{E}_{\mathbf{E}}[\mathbf{E}^{(m)\mathrm{H}}\mathbf{E}^{(m)}] = \boldsymbol{I}_{N_{\mathrm{a}}}\mathrm{tr}(\boldsymbol{\zeta})$ for the simplifying assumptions from Section II, but the equivalence holds for the general case, too.

## V. Simulation Results

Figs. 4 and 5 show the uncoded *bit error ratio* (BER) averaged over 10000 channel realizations and averaged over all receivers versus the *transmit* signal-to-noise ratio $E_{\mathrm{tr}}/\mathrm{E}[\mathbf{n}^{\mathrm{H}}[n]\mathbf{n}[n]]$, which is identical to the ratio $E_{\mathrm{S}}/N_0$ of the *average* transmit symbol energy $E_{\mathrm{S}}$ to the noise power density $N_0$. The BS has $N_{\mathrm{a}} = 4$ antennas, 100 symbols are transmitted per receiver and channel realization. The modulation alphabet is QPSK and the prediction filter is of length $p = 5$. The LS channel estimation is based on 256 pilots at an uplink signal-to-noise ratio of 3 dB. We assume a Jakes power density spectrum and each receiver has a maximum Doppler frequency of $f_{\mathrm{D}_k} = 100$ Hz (app. 54 km/h at 2 GHz). The non-linear *zero-forcing* Tomlinson-Harashima-Precoder (TxZF THP) and the linear *zero-forcing* filter (TxZF) have dashed lines, whereas their respective robust counterparts (RTxZF THP and RTxZF) are represented by solid lines. Both figures show dramatic improvements of the robust filters with respect to their non-robust versions which are also based on the predicted channel (i.e. in (6) and (7), $\boldsymbol{H}^{(m)}$ is replaced by $\overline{\boldsymbol{H}^{(m)}}$), since we do not want to show the effect of channel prediction but the impact of robust filtering. Especially in the $K = 4$ receivers scenario (Fig. 5), the robust filters show superior performance, since there is no degree of freedom left for the TxZF to reduce the noise at the receivers. The *loading* in (19) prevents the RTxZF from removing interference erroneously based on imperfect CSI. Moreover, the non-linear filter experiences a larger gain than the linear filter in both scenarios. But as we mentioned before, the THP filters exhibit a point of intersection with the linear filters and offer larger BERs than their linear counterparts for small $E_{\mathrm{S}}/N_0$ values. Fortunately, this point of intersection has a smaller $E_{\mathrm{S}}/N_0$ value for the robust filters compared to the $E_{\mathrm{S}}/N_0$ of the intercept point for the non-robust filters making robust THP systems very attractive.

## VI. Sensitivity of the Presented Approach with Respect to Errors in the Temporal Auto-Correlation Function of the Channel

One argument, that - at first glance - might speak against the procedure illustrated above, is, that the problem of uncer-

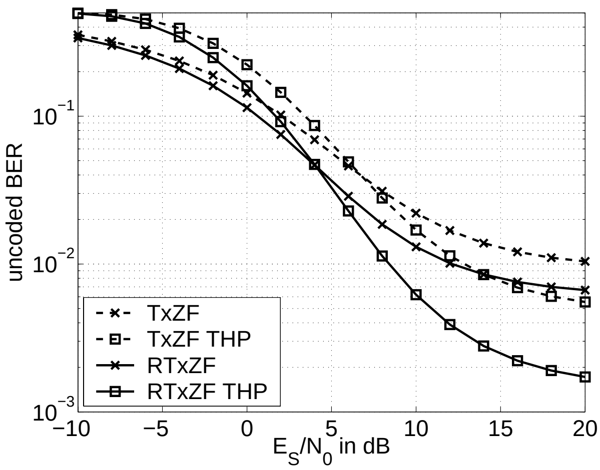


Fig. 4. Robust linear TxZF (RTxZF) and Robust non-linear TxZF THP (RTxZF THP) vs. Conventional TxZF and Conventional TxZF THP with predicted CSI ($K = 3$ receivers, $N_\mathrm{a} = 4$ antennas)

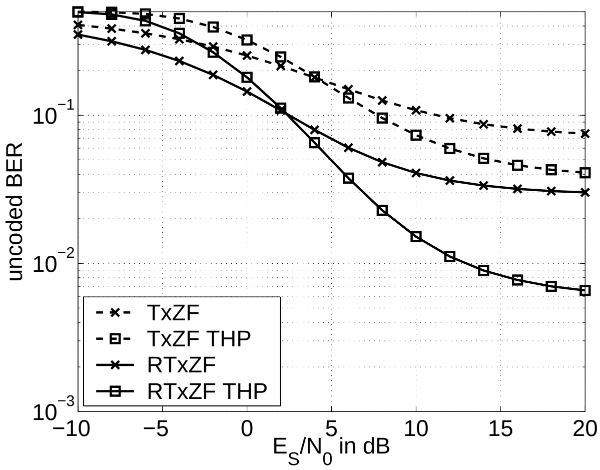


Fig. 5. Robust linear TxZF (RTxZF) and Robust non-linear TxZF THP (RTxZF THP) vs. Conventional TxZF and Conventional TxZF THP with predicted CSI ($K = 4$ receivers, $N_\mathrm{a} = 4$ antennas)

tainties in CSI has been solved by the exact knowledge of the temporal auto-correlation function in (3). As it seems, the performance gains are bought by knowledge, which in practice is inaccurate. On the one hand, this objection is partially true, since errors in the auto-correlation function will definitely degrade the efficiency of the robust filters, but on the other hand, the conventional non-robust filters will degrade *in the same way*, since they are also based on predicted CSI in our simulations and will therefore suffer from imperfect prediction as well. If the auto-correlation function is not perfectly known, then robust transmit filters have to be applied *more than ever*, since the circumstance of imperfect knowledge could be taken into account during the filter design. Due to larger errors, this would lead to a *loading* by a scaled identity matrix, which has a larger (e.g. Frobenius-) norm than the one in (19).

Nevertheless, we investigate the impact of an erroneous temporal auto-correlation function by assuming a constant *power density spectrum* of the fading process within the interval $[-f_{\mathrm{D}_k}; f_{\mathrm{D}_k})$, despite it has Jakes form. This leads to the sinc-function instead of the Bessel function for the auto-correlation function. Fig. 6 shows the influence of the imperfect auto-correlation function on the performance of both the robust and non-robust linear TxZF filter for $K = 3$ receivers and $N_\mathrm{a} = 4$ antennas. We observe, that *both* the robust *and* the non-robust version show a degradation of the BER in the interference dominated region, i.e. for large SNRs. But since there are more antennas than receivers, the impact of the imperfect auto-correlation function is not that severe as it would be in a $K = 4$ receivers scenario.

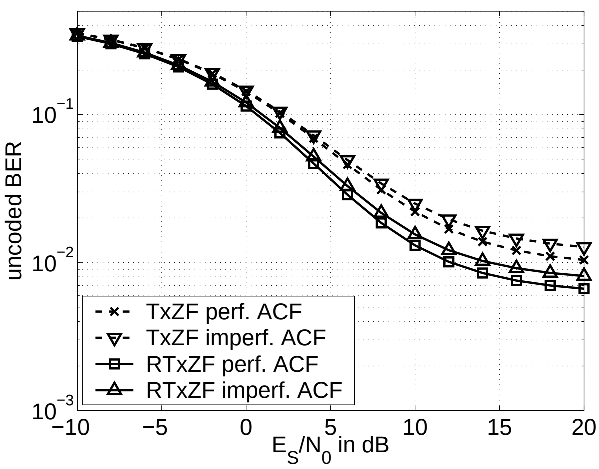


Fig. 6. Robust (RTxZF perf. ACF) and Conventional TxZF (TxZF perf. ACF) with *perfect* auto-correlation knowledge vs. Robust (RTxZF imperf. ACF) and Conventional TxZF (TxZF imperf. ACF) with *imperfect* auto-correlation knowledge ($K = 3$ receivers, $N_\mathrm{a} = 4$ antennas)

## VII. Evaluation of a Similar Heuristic Robust Approach

Another robust approach one might think of results from directly starting with the solution of the *conventional* filter in (6) and applying the *conditional mean* to the individual components instead of starting with the cost function in (8).

### A. Derivation of the Heuristic Robust Transmit Zero-Forcing Filter

For the derivation of the *heuristic* robust TxZF, the deterministic channel matrices in (6) are replaced by their respective random variables and afterwards, the matrix product $\mathbf{H}^{(m)}\mathbf{H}^{(m)\,\mathrm{H}}$ inside the inverse operator and the matrix $\mathbf{H}^{(m)\,\mathrm{H}}$ in front of the inverse are replaced by their *conditional mean* values. Due to the reversed order of the matrix product $\mathbf{H}^{(m)}\mathbf{H}^{(m)\,\mathrm{H}}$ compared to (16), this approach will lead to different results in the majority of the cases. To illustrate this circumstance, we take a look at the main diagonal entries of $\mathbf{H}^{(m)}\mathbf{H}^{(m)\,\mathrm{H}}$ (since only channel coefficients with same indices are correlated in our scenario). Different from (17),

$$\left[\mathbf{H}^{(m)}\mathbf{H}^{(m)\,\mathrm{H}}\right]_{k,k} = \sum_{a=1}^{N_\mathrm{a}} |\mathsf{h}_{k,a}^{(m)}|^2 \quad \forall k \in \{1,\ldots,K\}, \qquad (28)$$

and thus, the *conditional mean* now reads as

$$\begin{aligned}\mathrm{E}_{\mathbf{H}}\left[\mathbf{H}^{(m)}\mathbf{H}^{(m)\,\mathrm{H}}\,\middle|\,\hat{\mathcal{H}}^{(m)}\right] &= \overline{\boldsymbol{H}^{(m)}\boldsymbol{H}^{(m)\,\mathrm{H}}} \\ &= \overline{\boldsymbol{H}^{(m)}}\;\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}} + N_{\mathrm{a}}\boldsymbol{\zeta},\end{aligned} \tag{29}$$

whereas (15) is still valid. The *heuristic* robust filter can then be expressed by

$$\boldsymbol{P}_{\mathrm{HRZF}}^{(m)} = \beta_{\mathrm{HRZF}}^{(m)}\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}\left(\overline{\boldsymbol{H}^{(m)}}\;\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}} + N_{\mathrm{a}}\boldsymbol{\zeta}\right)^{-1}, \tag{30}$$

and $\beta_{\mathrm{HRZF}}^{(m)}$ is again chosen to meet the power constraint. By means of the inversion lemma [13], (19) can be transformed into a similar structure:

$$\boldsymbol{P}_{\mathrm{RZF}}^{(m)} = \beta_{\mathrm{RZF}}^{(m)}\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}\Big(\overline{\boldsymbol{H}^{(m)}}\;\overline{\boldsymbol{H}^{(m)\,\mathrm{H}}}+\boldsymbol{I}_K\,\mathrm{tr}(\boldsymbol{\zeta})\Big)^{-1}. \tag{31}$$

Comparing (31) with (30), equality of both algorithms holds, if $N_{\mathrm{a}}\boldsymbol{\zeta} = \boldsymbol{I}_K\,\mathrm{tr}(\boldsymbol{\zeta})$ which obviously becomes true, if all receivers have the same Doppler frequency $f_{\mathrm{D}_k}$ and the same average channel power $\sigma^2_{\mathsf{h}_k}$. But this configuration is very unlikely in a multi-user communication environment, since the Doppler frequencies are directly related to the speed of the receivers. In cases, where (30) and (31) are different, the *heuristic* approach is outperformed by the algorithm starting with the new cost function in (8) in terms of interference suppression. Under the assumption, that the received signals of all users are perturbed by noise with the same average power, all receivers should eventually exhibit the same raw BER (because of the same scaling factor $\beta^{-1}$ for all receivers), if interference was completely removed. The following simulation results demonstrate, that this goal is achieved best by the robust filter in (19) (or equivalently Eq. 31).

### B. Simulation Results for the Heuristic Approach

In Figs. 7 and 8, the Doppler frequency $f_{\mathrm{D}_1}$ of receiver 1 is reduced from 100 Hz to 20 Hz, whereas receivers 2 and 3 still have Doppler frequencies of $f_{\mathrm{D}_2} = f_{\mathrm{D}_3} = 100$ Hz. All other parameters are left unchanged. From this fact, we can conclude that receiver 1 will finally show a smaller BER than the other receivers due to the reduced uncertainty in CSI.

Fig. 7 shows the BER *averaged over all receivers* versus $E_{\mathrm{S}}/N_0$ for the conventional TxZF, the robust TxZF and the *heuristic* robust TxZF. Compared to Fig. 4, we observe a small performance improvement for both the RTxZF and the TxZF because of the smaller CSI error of receiver 1. For $E_{\mathrm{S}}/N_0$ values smaller than 11 dB, the *heuristic* robust approach outperforms the conventional TxZF, but shows worse performance if $E_{\mathrm{S}}/N_0$ is above 11 dB. In contrast, the RTxZF offers a smaller BER than its *heuristic* counterpart for the complete $E_{\mathrm{S}}/N_0$ range.

Fig. 8 shows the BER versus $E_{\mathrm{S}}/N_0$, where all *receivers are resolved*. We observe, that the *heuristic* robust filter facilitates a BER for receiver 1, which is smaller than the BER of the two other filter types. At the same time, the BERs of receivers 2 and 3 are considerably larger than those of the robust (RTxZF) filter and for large $E_{\mathrm{S}}/N_0$ values even larger than those of the conventional non-robust TxZF. The structure of the RTxZF curves is similar to the structure of the TxZF curves, but smaller BERs are achieved, especially for large $E_{\mathrm{S}}/N_0$ values. For these reasons, the RTxZF meets the desired behaviour of interference removal best, since it offers low BERs and tries to make equal BERs of all receivers possible, even if their uncertainties in CSI are of different size.

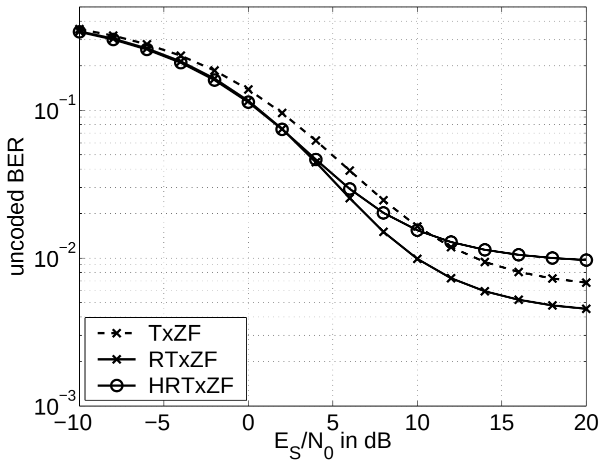


Fig. 7. Robust TxZF (RTxZF) vs. *Heuristic* Robust TxZF (HRTxZF) vs. Conventional TxZF (TxZF) with predicted CSI ($K = 3$ receivers, $N_{\mathrm{a}} = 4$ antennas)

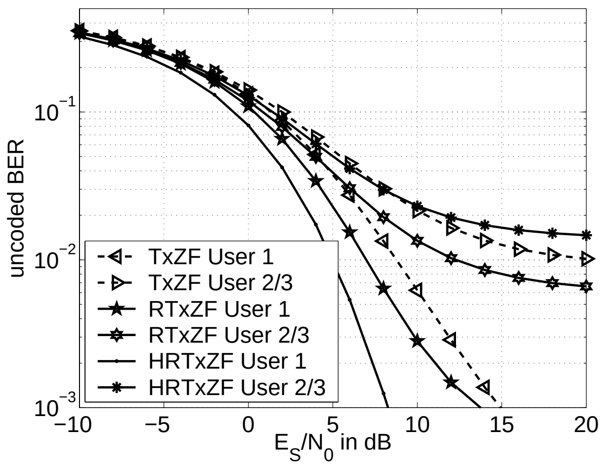


Fig. 8. Robust TxZF (RTxZF) vs. *Heuristic* Robust TxZF (HRTxZF) vs. Conventional TxZF (TxZF) with predicted CSI (*receivers resolved*, $K = 3$ receivers, $N_{\mathrm{a}} = 4$ antennas)

## VIII. Conclusion

In this paper, we addressed the impact of uncertainties in the CSI resulting from channel estimation and time lags in a TDD system with transmit processing and presented a *robust* version of the linear and non-linear THP transmit *zero-forcing* filter for the downlink. The robust versions clearly outperform their non-robust counterparts in terms of bit error ratio without increasing the computational complexity. Even if the temporal correlation of the fading process is not perfectly known, robust filters exhibit superior performance and are therefore essential since complete CSI is never available. For notational simplicity, we restricted ourselves to uncorrelated and frequency flat Rayleigh fading channels during the derivation of the robust filters, but the extension towards correlated, frequency selective, and non-zero-mean channels is straightforward.

## REFERENCES


[1] F. Kowalewski and P. Mangold, "Joint Predistortion and Transmit Diversity," in *Proc. Globecom '00*, pp. 245–249, November 2000.

[2] F. A. Dietrich, R. Hunger, M. Joham, and W. Utschick, "Linear Precoding over Time-Varying Channels in TDD Systems," in *Proc. ICASSP 2003*, vol. V, pp. 117–120, April 2003.

[3] G. Caire and S. Shamai (Shitz), "On the Multiple Antenna Broadcast Channel," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 1188–1193, November 2001.

[4] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "MIMO Precoding for Decentralized Receivers," in *ISIT '02*, p. 496, 2002.

[5] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "Space-Time Transmission using Tomlinson-Harashima Precoding," pp. 139–147, January 2002.

[6] Jia Liu and Alexandra Duel-Hallen, "Tomlinson-Harashima Transmitter Precoding for Synchronous Multiuser Communications," in *Proc. CISS'03*, March 2003.

[7] W. C. Jakes, *Microwave Mobile Communications*, Wiley, 1974.

[8] A. Duel-Hallen, Hu Shengquan, and H. Hallen, "Long-Range Prediction of Fading Signals," *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.

[9] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener Filter for the Downlink of TDD DS-CDMA Systems," in *Proc. ISSSTA 2002*, vol. 1, pp. 9–13, September 2002.

[10] P. W. Baier, M. W. Meurer, T. Weber, and H. Tröger, "Joint Transmission (JT), an alternative rationale for the downlink of time division CDMA using multi-element transmit antennas," in *Proc. ISSSTA'00*, vol. 1, pp. 1–5, Sept. 2000.

[11] F. Rey, M. Lamarca, and G. Vázquez, "Optimal Power Allocation with Partial Channel Knowledge for MIMO Multicarrier Systems," in *Proc. VTC Fall*, vol. 4, pp. 2121–2125, September 2002.

[12] H. A. Hindi and S. P. Boyd, "Robust solutions to $l_1$, $l_2$, and $l_\infty$ uncertain linear approximation problems using convex optimization," in *Proc. of the American Control Conference*, pp. 3487–3491, June 1998.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing*, vol. 1, Prentice Hall, 1993.

[14] R.L. Choi and R.D. Murch, "Transmit MMSE pre-RAKE pre-processing with simplified receivers for the downlink of MISO TDD-CDMA systems," in *Proc. Globecom 2002*, pp. 429–433, November 2002.

[15] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, no. 5/6, pp. 138–139, March 1971.

[16] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, vol. 20, pp. 774–780, August 1972.

[17] M. Joham, J. Brehmer, and W. Utschick, "MMSE Approaches to Multiuser Spatio-Temporal Tomlinson-Harashima Precoding," in *Proc. ITG SCC'04*, pp. 387–394, January 2004.

[18] C. Windpassinger, T. Vencel, and R. Fischer, "Precoding and Loading for BLAST-like Systems," in *Proceedings of IEEE International Conference on Communications (ICC) 2003*, May 2003.

[19] S. Shahbazpanahi, A. B. Gershman, Z.-Q. Luo, and K. M. Wong, "Robust Adaptive Beamforming for General-Rank Signal Models," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2257–2269, Sept. 2003.

[20] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Handbook on Semidefinite Programming*, chapter Robust Semidefinite Programming, pp. 139–162, Kluwer Academic Publishers, 2000.

[21] András Prékopa, *Stochastic Programming*, Kluwer, 1995.

[22] Andrey N. Tikhonov and Vasiliy Y. Arsenin, *Solutions of Ill-Posed Problems*, V. H. Winston & Sons, 1977.

[23] F. A. Dietrich, R. Hunger, M. Joham, and W. Utschick, "Robust Transmit Wiener Filter for Time Division Duplex Systems," in *Proceedings of the International Symposium on Signal Processing and Information Technology (ISSPIT)*, December 2003.

[24] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, January 2004.

[25] A. A. Kassam and H. V. Poor, "Robust Techniques for Signal Processing: A Survey," *Proc. IEEE*, vol. 73(3), pp. 433–481, March 1985.