ROBUST TRANSMIT WIENER FILTER FOR TIME DIVISION DUPLEX SYSTEMS

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

Institute for Circuit Theory and Signal Processing, Munich University of Technology,
Arcisstrasse 21, 80290 Munich, Germany, E-Mail: Dietrich@nws.ei.tum.de

ABSTRACT

Errors in channel state information have a significant impact on the performance of linear precoding (transmit filters) in frequency selective MISO communication links. We present a new robust optimization scheme for the design of the robust transmit Wiener filter, which is based on a stochastic error model. This general paradigm achieves an average robustness as it is relevant for physical layer signal processing.

1. INTRODUCTION

Space-time equalization improves the quality of a wireless communication link at the expense of an increased computational complexity. With linear precoding the complexity of channel equalization is shifted from the receiver to the transmitter, e.g. the base station in the downlink, where power and cost constraints are less stringent. We consider a link with multiple antennas at the transmitter and a single antenna at the receiver. The design of the linear transmit filter considered here relies on *full channel state information* (CSI), i.e. knowledge of all current channel coefficients. 3 main types of linear precoders exist: the *transmit matched filter* (TxMF), *transmit zero forcing* (TxZF), and the *transmit Wiener filter* (TxWF) [1], where the first two are special cases of the TxWF. In time-division duplex (TDD) links the channel estimates from the uplink can be used for linear precoding of the downlink symbols due to reciprocity of the up- and downlink channels. But channel estimates are only available from a previous uplink slot and therefore out-dated, which results in a significant degradation in performance and requires linear prediction to improve the transmitter's channel knowledge [2]. The nominal design methods for standard linear transmit filters do not model the errors in the transmitter's channel knowledge. To our knowledge robust design of linear precoders for frequency selective channels has not been considered in the literature. We present a new robust design of the TxWF based on the paradigm of static stochastic programming [3] for a stochastic model of the parameter errors, which decreases the sensitivity of the TxWF to parameter errors.

After defining the system model and the slot structure, which yields out-dated CSI, the standard TxWF (nominal design) is revisited. In Section 4 we discuss our notion of robust signal processing in the physical layer. For the robust design of the TxWF in Section 5 we define the robust optimization criterion and derive its solution. Interpreting the new cost function and the robust TxWF (RTxWF) we give an explanation for the performance difference of the robust version and its relation to the field of regularization. Finally, it is shown by simulation that the RTxWF decreases the error floor of the uncoded bit error ratio (BER).

Throughout the paper, $\hat{x}$ denotes an estimate of x, '$\otimes$' the Kronecker product, $\mathbf{0}_{M\times N}$ the $M \times N$ zero matrix, and $\mathbf{1}_M$ the $M \times M$ identity matrix. Deterministic variables as well as realizations of a random variable x are written as $\boldsymbol{x}$.

2. SYSTEM MODEL

A single user MISO system is considered. A QPSK symbol sequence $\mathsf{s} \in \mathbb{C}^W$ with correlation matrix $\boldsymbol{R}_\mathsf{s} = \mathrm{E}\left[\mathsf{s}\mathsf{s}^\mathrm{H}\right] \in \mathbb{C}^{W\times W}$ is precoded with $\boldsymbol{P} \in \mathbb{C}^{MW\times W}$ and transmitted over M antenna elements and the received signal in the *downlink* is given by (Fig. 1)

$$\hat{\mathsf{s}} = \boldsymbol{H}_n \boldsymbol{P} \mathsf{s} + \boldsymbol{\eta} \in \mathbb{C}^{W+Q-1}. \quad (1)$$

The model can be generalized to include multiple users and processing at the receiver independent of $\boldsymbol{P}$ [1]. The discrete time frequency selective channel of length Q is constant during slot n and given by the block Toeplitz matrix

$$\boldsymbol{H}_n = \sum_{q=0}^{Q-1} \boldsymbol{S}^\mathrm{T}_{(q,W,Q-1)} \otimes \boldsymbol{h}^\mathrm{T}_{n,q} \in \mathbb{C}^{(W+Q-1)\times MW} \quad (2)$$

with selection matrix

$$\boldsymbol{S}_{(q,M,N)} = \left[\mathbf{0}_{M\times q}, \mathbf{1}_M, \mathbf{0}_{M\times(N-q)}\right] \in \{0,1\}^{M\times(M+N)}$$

and the vector channel coefficient $\boldsymbol{h}_{n,q} \in \mathbb{C}^M$ of tap q. Here, $\boldsymbol{\eta} \sim \mathcal{N}_\mathrm{c}(\mathbf{0}, \boldsymbol{R}_{\boldsymbol{\eta}})$ with $\boldsymbol{R}_{\boldsymbol{\eta}} = \mathrm{E}\left[\boldsymbol{\eta}\boldsymbol{\eta}^\mathrm{H}\right]$ is additive complex Gaussian noise. Signal and noise are uncorrelated, i.e. $\mathrm{E}\left[\mathsf{s}\boldsymbol{\eta}^\mathrm{H}\right] = \mathbf{0}_{W\times(W+Q-1)}$.

The received pilot sequence in the *uplink* is (Fig. 1)

$$\mathsf{y}_\mathrm{u} = \boldsymbol{S}_\mathrm{p} \boldsymbol{h}_{n-\ell} + \boldsymbol{\eta}_\mathrm{u} \in \mathbb{C}^{M(N_\mathrm{P}-Q+1)} \quad (3)$$

with the spatio-temporal channel vector

$$\boldsymbol{h}_{n-\ell} = \left[\boldsymbol{h}^\mathrm{T}_{n-\ell,0}, \cdots, \boldsymbol{h}^\mathrm{T}_{n-\ell,Q-1}\right]^\mathrm{T} \in \mathbb{C}^{MQ}$$

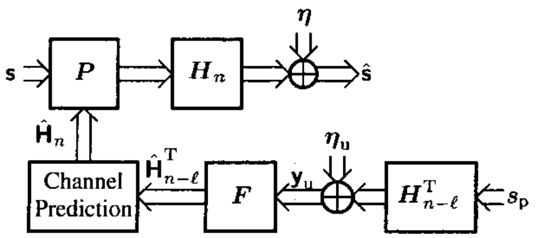

Fig. 1. Downlink transmit processing based on channel estimation and prediction from the uplink.

and matrix $\boldsymbol{S}_{\mathrm{p}} = \boldsymbol{S}'_{\mathrm{p}} \otimes \mathbf{1}_M$, where $\boldsymbol{S}'_{\mathrm{p}}$ is Toeplitz and contains N_{p} pilot symbols.

The *maximum likelihood* (ML) *channel estimator* ($\boldsymbol{R}_{\eta_{\mathrm{u}}} = \sigma_{\mathrm{n}}^2 \mathbf{1}_{M(N_{\mathrm{p}}-Q+1)}$) for uplink slot $n-\ell$ is

$$\hat{\mathbf{h}}_{n-\ell}^{\mathrm{ML}} = \boldsymbol{F}\mathbf{y}_{\mathrm{u}}, \quad \boldsymbol{F} = (\boldsymbol{S}_{\mathrm{p}}^{\mathrm{H}}\boldsymbol{S}_{\mathrm{p}})^{-1}\boldsymbol{S}_{\mathrm{p}}^{\mathrm{H}}. \tag{4}$$

The positions ℓ of available uplink slots depend on the slot assignment: In the TDD mode of UMTS, for example, the signal is divided into frames, which consist of 15 slots [4]. Each slot contains a midamble and a data signal. From various ways of assigning slots to the up- ("↑") or downlink ("↓"), we pick the following 2 proposed frame structures for medium and high data rates [4]:

↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓

Symmetric multiple switching point downlink frame.

↓	↓	↓	↑	↓	↓	↓	↑	↓	↓	↓	↑	↓	↓	↓

Asymmetric multiple switching point downlink frame.

Using the ML channel estimates the linear minimum mean square error (MSE) or Wiener prediction for tap q with $\boldsymbol{w}_q$, the solution of the Wiener-Hopf equation, reads

$$\hat{\mathbf{h}}_{n,q}^{\mathrm{WF}} = \left[\hat{\mathbf{h}}_{n-15p,q}^{\mathrm{ML}}, \cdots, \hat{\mathbf{h}}_{n-1,q}^{\mathrm{ML}}\right] \boldsymbol{T}_k \boldsymbol{w}_q. \tag{5}$$

The diagonal matrix $\boldsymbol{T}_k \in \{0,1\}^{15p\times 15p}$ selects the uplink slots from past $15p$ slots (p frames). A detailed description of the parameters and solution can be found in [2].

In the sequel we assume that 1st and 2nd-order statistics of all random sequences are known perfectly.

3. STANDARD TRANSMIT WIENER FILTER

The *nominal design* of the TxWF $\boldsymbol{P}_{\mathrm{WF}}$ assuming perfect knowledge of $\boldsymbol{H}_n$ minimizes the *modified mean square error* $\sigma_{\varepsilon}^2(\boldsymbol{P}, \beta, \boldsymbol{H}_n, \mathcal{P}_{\mathrm{S}}) = \mathrm{E}\left[\|\boldsymbol{\varepsilon}(\boldsymbol{H}_n)\|_2^2\right]$ of $\boldsymbol{\varepsilon}(\boldsymbol{H}_n) = \boldsymbol{\Psi}\mathbf{s} - \beta^{-1}\hat{\mathbf{s}}$, where $\mathcal{P}_{\mathrm{S}} = \{\boldsymbol{R}_{\mathbf{s}}, \boldsymbol{R}_{\eta}\}$ and $\boldsymbol{\Psi} = [\mathbf{1}_W, \mathbf{0}_{W\times Q-1}]^{\mathrm{T}}$, using the whole available transmit power E_{tr}:

$$[\boldsymbol{P}_{\mathrm{WF}}, \beta_{\mathrm{WF}}] = \arg\min_{\boldsymbol{P},\beta} \mathrm{E}\left[\|\boldsymbol{\Psi}\mathbf{s} - \beta^{-1}\hat{\mathbf{s}}\|_2^2\right] \quad \text{s.t.: } \mathrm{E}\left[\|\boldsymbol{P}\mathbf{s}\|_2^2\right] \leq E_{\mathrm{tr}}, \tag{6}$$

It can be written as [1]

$$\boldsymbol{P}_{\mathrm{WF}} = \beta_{\mathrm{WF}}\left(\boldsymbol{H}_n^{\mathrm{H}}\boldsymbol{H}_n + \xi\mathbf{1}_{MW}\right)^{-1}\boldsymbol{H}_n^{\mathrm{H}}\boldsymbol{\Psi} \tag{7}$$

with $\xi = \mathrm{tr}(\boldsymbol{R}_{\eta})/E_{\mathrm{tr}}$. β_{WF} is used to set the transmit power to E_{tr}. To reduce the necessary complexity we decided to employ block filters which do not process the whole slot at once, but split the slot into smaller windows with W symbols. Therefore, we have to suppress the interference generated for the following window described by $\boldsymbol{\Psi}$.

4. ROBUST SIGNAL PROCESSING FOR COMMUNICATIONS

Dealing with errors in the channel parameters, the true channel matrix is modeled by the transmitter as

$$\mathbf{H}_n = \hat{\boldsymbol{H}}_n + \mathbf{E}_n \text{ with } \mathbf{E}_n = \sum_{q=0}^{Q-1} \boldsymbol{S}_{(q,W,Q-1)}^{\mathrm{T}} \otimes \mathbf{e}_{n,q}^{\mathrm{T}}. \tag{8}$$

Robust optimization methods can be classified based on the underlying assumptions on the model errors:

1) A *deterministic error model* such as a norm bound $\|\boldsymbol{E}_n\|_F \leq \alpha$ (unstructured error) or $\|\boldsymbol{e}_{n,q}\|_2 \leq \alpha_q$ (structured error) is used for *min-max optimization*, minimizing the error σ_{ε}^2 in the worst case [5, 6]. It guarantees a cost not exceeding the worst case, but requires knowledge about the compact set of errors, e.g. the norm bound α or α_q.

2) If the error is modeled as a (stationary) *stochastic process*, e.g. $\mathbf{e}_{n,q} \sim \mathcal{N}_{\mathrm{c}}(\boldsymbol{\mu}_{\mathbf{e}_q}, \boldsymbol{R}_{\mathbf{e}_q})$ with $\boldsymbol{\mu}_{\mathbf{e}_q} = \mathrm{E}[\mathbf{e}_{n,q}]$ and $\boldsymbol{R}_{\mathbf{e}_q} = \mathrm{E}[\mathbf{e}_{n,q}\mathbf{e}_{n,q}^{\mathrm{H}}]$, the cost function is also a stochastic process and the paradigms known from *static stochastic programming* [3] can be applied. The solution depends on the error distribution or, e.g., its first/second order moments. Prékopa in [3] proposes to minimize the expected value of the cost function w.r.t. $\mathbf{e}_{n,q}$:

$$\sigma_{\mathrm{R},\varepsilon}^2(\boldsymbol{P}, \beta, \hat{\boldsymbol{H}}_n, \mathcal{P}_{\mathrm{S}}) = \mathrm{E}_{\mathbf{E}_n}[\sigma_{\varepsilon}^2(\boldsymbol{P}, \beta, \hat{\boldsymbol{H}}_n + \mathbf{E}_n, \mathcal{P}_{\mathrm{S}})]. \tag{9}$$

Thus, it provides a robustness on average.

Application of the min-max approach based on the deterministic error model to physical layer design in communications would have 2 major *disadvantages*:

1) In communications errors in channel parameters are caused by noise, modeled by a (complex) Gaussian distribution, or a delay. The stochastic noise model above implies an unlimited set size α (not compact), which we would need to choose heuristically or based on an additional criterion.

2) Physical layer processing is part of a *large* system: As *higher layers handle worst case* of the physical layer, a useful definition the of worst case at this layer becomes unclear. Moreover, the design goal of physical layer signal processing is to provide good performance on *average* (Average BER, MSE). Thus, *average* robustness appears as a consistent notion of robustness for the physical layer.

5. ROBUST TRANSMIT WIENER FILTER

5.1. Stochastic Error Model

As discussed above, the transmitter views the real channel coefficients $\mathbf{h}_{n,q}$ as a stochastic process centered at the estimated coefficients $\hat{\mathbf{h}}_{n,q}$ corrected by the average error $\boldsymbol{\mu}_{\mathbf{e}_q}$:

$$\mathbf{h}_{n,q} = \hat{\boldsymbol{h}}_{n,q} + \mathbf{e}_{n,q} \text{ with } \mathbf{h}_{n,q} \sim \mathcal{N}_c(\hat{\boldsymbol{h}}_{n,q} + \boldsymbol{\mu}_{\mathbf{e}_q}, \boldsymbol{R}_{\mathbf{e}_q}).$$

For delayed CSI from ML channel estimation improved by linear prediction [2] we set $\hat{\boldsymbol{h}}_{n,q} = \hat{\boldsymbol{h}}_{n,q}^{\mathrm{WF}}$. For Rayleigh fading and spatially correlated coefficients the 1st and 2nd order moments can be derived (in several steps) from (5) as $\boldsymbol{\mu}_{\mathbf{e}_q} = \mathbf{0}$ and

$$\begin{aligned}\boldsymbol{R}_{\mathbf{e}_q} = \boldsymbol{R}_{\mathbf{h}_q} - 2\,\mathrm{Re}\,(\boldsymbol{R}_{\mathrm{c},q}(\mathbf{1}_M \otimes \boldsymbol{T}_k \boldsymbol{w}_q)) +\\ (\mathbf{1}_M \otimes \boldsymbol{w}_q^{\mathrm{H}} \boldsymbol{T}_k) \boldsymbol{R}_{\mathrm{T},q} (\mathbf{1}_M \otimes \boldsymbol{T}_k \boldsymbol{w}_q).\end{aligned} \quad (10)$$

The spatial correlation matrix of tap q is $\boldsymbol{R}_{\mathbf{h}_q} = \mathrm{E}[\boldsymbol{h}_{n,q}\boldsymbol{h}_{n,q}^{\mathrm{H}}]$, spatial-temporal correlations are summarized in

$$\begin{aligned}\boldsymbol{R}_{\mathrm{T},q} = \boldsymbol{R}_{\mathbf{h}_q} \otimes \boldsymbol{R}_{\mathrm{a},q}\\ + \boldsymbol{S}_{(qM,M,(Q-1)M)} \boldsymbol{F} \boldsymbol{R}_{\boldsymbol{\eta}_u} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{S}_{(qM,M,(Q-1)M)}^{\mathrm{T}},\end{aligned}$$

where $\boldsymbol{R}_{\mathrm{a},q}$ is Toeplitz with 1st row $[r_q[0], \cdots, r_q[15p-1]]$ and autocorrelation sequence $r_q[\ell] = \mathrm{E}[\mathbf{h}_{n,q}^{\mathrm{T}} \mathbf{h}_{n-\ell,q}^{*}]/M$ (temporal correlations are spatially invariant for small antenna spacing), and the cross-correlation matrix $\boldsymbol{R}_{\mathrm{c},q} = \boldsymbol{R}_{\mathbf{h}_q} \otimes \boldsymbol{r}_q^{\mathrm{T}}$ with correlation vector $\boldsymbol{r}_q = [r_q[1], \cdots, r_q[15p]]^{\mathrm{T}}$. The situation that only delayed and estimated CSI is available are special cases of (10).

5.2. Robust Optimization

Applying the method of *static stochastic programming* we include knowledge about parameters error in the design. To obtain a more transparent notation we assume $\boldsymbol{\mu}_{\mathbf{e}_q} = \mathbf{0}$. The error w.r.t. $\hat{\mathbf{s}}$ (Eqn. 1) considering errors in $\hat{\boldsymbol{H}}_n$ (Eqn. 8) is

$$\begin{aligned}\boldsymbol{\varepsilon}(\hat{\boldsymbol{H}}_n + \mathbf{E}_n) = \boldsymbol{\Psi s} - \beta^{-1}\left(\left(\hat{\boldsymbol{H}}_n + \mathbf{E}_n\right)\boldsymbol{P}\mathbf{s} + \boldsymbol{\eta}\right)\\ = \boldsymbol{\varepsilon}(\hat{\boldsymbol{H}}_n) - \beta^{-1}\mathbf{E}_n \boldsymbol{P}\mathbf{s}.\end{aligned} \quad (11)$$

According to (9) we take the expected value of $\boldsymbol{\varepsilon}(\hat{\boldsymbol{H}}_n + \mathbf{E}_n)$ w.r.t. $\boldsymbol{\eta}$, $\mathbf{s}$, *and* $\mathbf{E}_n$. As $\mathbf{E}_n$ is statistically independent from the former two random variables we get

$$\begin{aligned}\sigma_{\mathrm{R},\boldsymbol{\varepsilon}}^2(\boldsymbol{P}, \beta, \hat{\boldsymbol{H}}_n, \mathcal{P}_{\mathrm{R}}) = \sigma_{\boldsymbol{\varepsilon}}^2(\boldsymbol{P}, \beta, \hat{\boldsymbol{H}}_n, \mathcal{P}_{\mathrm{S}}) +\\ \beta^{-2}\,\mathrm{tr}\left(\boldsymbol{R}_{\mathbf{s}} \boldsymbol{P}^{\mathrm{H}} \boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}} \boldsymbol{P}\right),\end{aligned} \quad (12)$$

with $\mathcal{P}_{\mathrm{R}} = \mathcal{P}_{\mathrm{S}} \cup \{\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}}\}$ and error covariance matrix

$$\begin{aligned}\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}} = \mathrm{E}_{\mathbf{E}_n}\left[\mathbf{E}_n^{\mathrm{H}} \mathbf{E}_n\right]\\ = \sum_{q=0}^{Q-1} (\boldsymbol{S}_{(q,WF,Q-1)} \boldsymbol{S}_{(q,WF,Q-1)}^{\mathrm{T}}) \otimes \boldsymbol{R}_{\mathbf{e}_q}^{*}.\end{aligned} \quad (13)$$

The new robust optimization goal is minimizing the modified MSE with average robustness constrained by the available transmit power E_{tr}:

$$\begin{aligned}[\boldsymbol{P}_{\mathrm{RWF}}, \beta_{\mathrm{RWF}}] = \arg\min_{\boldsymbol{P},\beta} \sigma_{\mathrm{R},\boldsymbol{\varepsilon}}^2(\boldsymbol{P}, \beta, \hat{\boldsymbol{H}}_n, \mathcal{P}_{\mathrm{R}})\\ \text{s.t. : } \mathrm{E}\left[\|\boldsymbol{P}\mathbf{s}\|_2^2\right] \le E_{\mathrm{tr}}.\end{aligned} \quad (14)$$

The problem remains convex [7]. Its solution is

$$\boldsymbol{P}_{\mathrm{RWF}} = \beta_{\mathrm{RWF}} \left(\hat{\boldsymbol{H}}_n^{\mathrm{H}} \hat{\boldsymbol{H}}_n + \xi \mathbf{1}_{MW} + \boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}}\right)^{-1} \hat{\boldsymbol{H}}_n^{\mathrm{H}} \boldsymbol{\Psi}, \quad (15)$$

and β_{RWF} is chosen to satisfy the constraint with equality.

Comparing (15) with (7) we observe that the regularization term in the robust cost function (12) results in a *"loading"* of the matrix inverse according to the size and structure of the channel parameter errors described by $\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}}$. The loading keeps TxWF closer to the (inherently robust) TxMF and thus achieves an on average optimum trade-off between interference cancellation as performed by the TxZF and serving the user with maximum power as done by the TxMF [2].

The only additional *complexity* for the design of the RTxWF compared to the TxWF are some matrix multiplications and additions to determine $\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}}$ (Eqn. 10 and 13). In case the auto-correlation properties of the channel coefficients are not already needed to design the linear predictor (Section 2), they have to be estimated to compute $\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}}$.

5.3. Interpretation of Regularization and Robustness

The robust criterion in (9) and (12) regularizes the solution by $\Theta(\boldsymbol{P}) = \beta^{-2}\,\mathrm{tr}\left(\boldsymbol{R}_{\mathbf{s}} \boldsymbol{P}^{\mathrm{H}} \boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}} \boldsymbol{P}\right)$. If $\boldsymbol{R}_{\mathbf{s}} = \mathbf{1}_W$ and $\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}} = \sigma_{\mathrm{e}}^2\,\mathbf{1}_{MW}$, additionally to a small ε, regularization tries to keep the Frobenius norm of $\boldsymbol{P}$ small, since $\Theta(\boldsymbol{P}) = \beta^{-2}\,\sigma_{\mathrm{e}}^2\,\|\boldsymbol{P}\|_{\mathrm{F}}^2$. This is known as *Tychonov regularization* with weight given by the variance σ_{e}^2 of the (uncorrelated) error [7]. More generally, for $\boldsymbol{R}_{\mathbf{s}} = \mathbf{1}_W$ and the eigenvalue decomposition of $\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}} = \boldsymbol{U}_{\boldsymbol{E}} \boldsymbol{\Lambda}_{\boldsymbol{E}} \boldsymbol{U}_{\boldsymbol{E}}^{\mathrm{H}}$ we obtain $\Theta(\boldsymbol{P}) = \beta^{-2}\,\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{H}} \boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}} \boldsymbol{P}\right) = \beta^{-2} \sum_{i=1}^{MW} \lambda_{E,i}\, \|\boldsymbol{p}_i'\|_2^2$, where $\boldsymbol{p}_i'$ are the rows of $\boldsymbol{U}_{\boldsymbol{E}}^{\mathrm{H}} \boldsymbol{P}$ and $\lambda_{E,i}$ form the diagonal of $\boldsymbol{\Lambda}_{\boldsymbol{E}}$. Thus, $\boldsymbol{R}_{\mathbf{E}^{\mathrm{H}}}$ determines in which direction the columns of $\boldsymbol{P}$ should be regularized (small norm), i.e. where the error in channel knowledge is large. $\boldsymbol{R}_{\mathbf{s}}$ determines which subspaces of the row space of $\boldsymbol{P}$ should be of small norm, i.e. those used by data symbols for transmission.

We observed that the optimization in (12) prefers solutions $\boldsymbol{P}$ with small $\|\boldsymbol{P}\|_{\mathrm{F}}^2$ or norm of transformed columns $\|\boldsymbol{p}_i'\|_2^2$. It remains to illustrate the impact of regularization on the sensitivity of $\sigma_{\boldsymbol{\varepsilon}}^2(\boldsymbol{P}, \beta, \boldsymbol{H}_n, \mathcal{P}_{\mathrm{S}})$ in (6) to errors in $\boldsymbol{H}_n$: In contrast to the transmitter's channel error model (8), we now assume $\hat{\mathbf{H}}_n = \boldsymbol{H}_n + \mathbf{E}_n$ to evaluate the sensitivity. From Jensen's inequality [7] follows

$$\mathrm{E}_{\hat{\mathbf{H}}_n}\left[\sigma_{\boldsymbol{\varepsilon}}^2(\boldsymbol{P}, \beta, \hat{\mathbf{H}}_n, \mathcal{P}_{\mathrm{S}})\right] \ge \sigma_{\boldsymbol{\varepsilon}}^2(\boldsymbol{P}, \beta, \mathrm{E}_{\hat{\mathbf{H}}_n}[\hat{\mathbf{H}}_n], \mathcal{P}_{\mathrm{S}}),$$

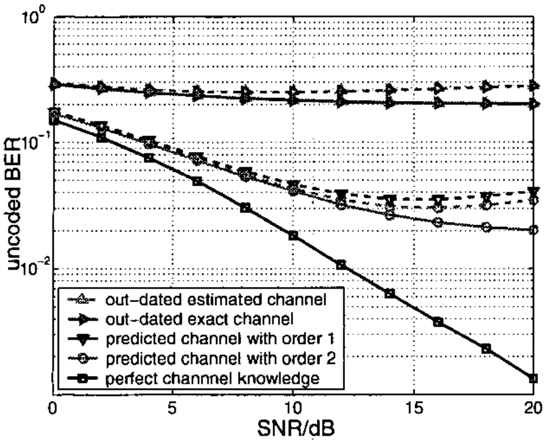

Fig. 2. Symmetric frame: RTxWF (solid) vs. standard TxWF (dashed)

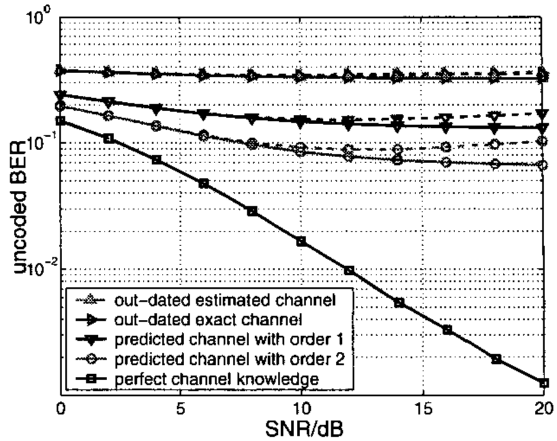

Fig. 3. Asymmetric frame: RTxWF (solid) vs. standard TxWF (dashed)

i.e. variations in $\boldsymbol{H}_n\boldsymbol{P}$ increase the MSE σ_ε^2 and the error $\boldsymbol{E}_n$ is amplified by $\boldsymbol{P}$: The sensitivity of the cost function $\sigma_\varepsilon^2(\boldsymbol{P},\beta,\boldsymbol{H}_n,\mathcal{P}_\mathrm{S})$ to $\boldsymbol{E}_n$ increases with $\|\boldsymbol{P}\|_\mathrm{F}^2$ or $\|\boldsymbol{p}_i'\|_2^2$.

6. BER PERFORMANCE

With a window size $W = 8$, 128 symbols per slot are transmitted over $M = 2$ antenna elements in the downlink. An uplink slot contains $N_\mathrm{p} = 256$ pilot symbols at an SNR of 3 dB. The channel coefficients are i.i.d. complex Gaussian distributed ($Q = 4$ paths) assuming a Jakes power spectrum describing temporal correlations with maximum Doppler frequency $f_\mathrm{dmax} = 100$ Hz (approx. 54 km/h at a carrier frequency of 2 GHz). Moreover, we assume $\boldsymbol{R}_\eta = \sigma_\mathrm{n}^2 \mathbf{1}_{W+Q-1}$ and $\boldsymbol{R}_\mathrm{s} = \mathbf{1}_W$. The SNR is the ratio of transmit power E_tr and noise variance at the receiver σ_n^2.

In Figs. 2 and 3, we compare the uncoded BER of the robust and nominal TxWF for *symmetric* and *asymmetric* frame structure, respectively (out-dated exact or estimated channel: $\hat{\boldsymbol{h}}_{n,q} = \boldsymbol{h}_{n-\ell,q}$ or $\hat{\boldsymbol{h}}_{n,q} = \hat{\boldsymbol{h}}_{n-\ell,q}^{\mathrm{ML}}$; Wiener prediction with order $p \in \{1,2\}$). For increasing SNR, the BERs of the nominal TxWFs (dashed lines) with outdated CSI have a minimum at a *finite* SNR between 10 and 15 dB. This behavior follows from the convergence of the TxWF to the TxZF for high SNR, since the non-robust TxZF tries to remove interference based on outdated CSI and thus, introduces additional interference. On the other hand, the robust TxWF (solid lines) only shows a lower saturation for high SNR due to the loading term which makes the TxWF more similar to the inherently robust TxMF [2]. Therefore, not only CSI prediction is necessary for uncoded BER below 10^{-1} (see Fig. 2), but also a robust design of the transmit filter (cf. Fig. 3).

7. CONCLUSIONS

A *robust* version of the *transmit wiener filter* was derived from an optimization problem tailored to achieve *average robustness* as it is relevant for physical layer design. The solution finds the best trade-off between transmit matched filter and zero-forcing based on 1st and 2nd order moments of the errors in channel state information, i.e. it adapts to the error statistic with little additional complexity. From the interpretation of the new cost function and numerical results we conclude that linear precoding for time-variant channels requires linear prediction and a robust design.

8. REFERENCES

[1] 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*, September 2002, vol. 1, pp. 9–13.

[2] F. Dietrich, R. Hunger, M. Joham, and W. Utschick, "Linear Precoding over Time-Varying Channels in TDD Systems," in *Proceedings of ICASSP 2003*, Hong Kong, April 2003.

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

[4] 3GPP, "Technical Specifications, 3GPP TS 25.221, Version 4.3.0," http://www.3gpp.org/, Dec. 2001.

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

[6] L. El Ghaoui and H. Lebret, "Robust Solutions to Least-Squares Problems with Uncertain Data," *SIAM J. Matrix Anal. Appl.*, vol. 18(4), pp. 1035–1064, Oct. 1997.

[7] S. Boyd and L. Vandenberghe, "Convex Optimization," http://www.stanford.edu/~boyd/cvxbook.html, to be published, Dec. 2002.