Linear Transmit Processing in MIMO Communications Systems

Michael Joham, *Member, IEEE*, Wolfgang Utschick, *Member, IEEE*, and Josef A. Nossek, *Fellow, IEEE*

***Abstract*—We examine and compare the different types of linear transmit processing for multiple input, multiple output systems, where we assume that the receive filter is independent of the transmit filter contrary to the joint optimization of transmit and receive filters. We can identify three filter types similar to receive processing: the transmit matched filter, the transmit zero-forcing filter, and the transmit Wiener filter. We show that the transmit filters are based on similar optimizations as the respective receive filters with an additional constraint for the transmit power. Moreover, the transmit Wiener filter has similar convergence properties as the receive Wiener filter, i.e., it converges to the matched filter and the zero-forcing filter for low and high signal-to-noise ratio, respectively. We give closed-form solutions for all transmit filters and present the fundamental result that their mean-square errors are equal to the errors of the respective receive filters, if the information symbols and the additive noise are uncorrelated. However, our simulations reveal that the bit-error ratio results of the transmit filters differ from the results for the respective receive filters.**

***Index Terms*—Linear transmit processing, linear precoding, multiple input, multiple output systems (MIMO), pre-equalization, prerake, Wiener filtering.**

I. Introduction

THE task of receive equalization filters is to remove the distortion due to the channel and to minimize the effect of the received noise. Three basic receive filter types are well researched and understood [1]: the *receive matched filter* (RxMF), the *receive zero-forcing filter* (RxZF), and the *receive Wiener filter* (RxWF). The RxMF, which is also called *conventional filter* [1] or *rake* [2], maximizes the signal portion of the desired signal, the RxZF [3] removes interference, and the RxWF or linear *minimum mean-square error (MMSE) filter* finds a tradeoff between noise and interference [1], [4].

The major drawback of receive filters is the increased complexity of the receiver, because channel estimation and adaptation of the receive filter is necessary. For example, in the uplink of cellular mobile radio systems, *receive processing* is advantageous, because the complexity resides at the *base station* (BS). On the other hand, in the downlink, receive processing leads to more complex *mobile stations* (MSs).

If the downlink channel impulse response is available at the BS, *transmit processing* becomes possible which equalizes the signal at the receiver with a filter at the transmitter. The main advantage of transmit processing is the possibility to simplify the receivers, i.e., the MSs. The assumption that the downlink channel impulse response is known at the BS is valid in *time division duplex* (TDD) systems, e.g., *TDD-code division multiple access* (TDD–CDMA, [5]) or *time division CDMA* (TD–CDMA, [6]), because the uplink and the downlink share the same frequency band. Thus, all channel parameters are the same for uplink and downlink, if the BS and MSs are calibrated correctly [7]–[9] and the *coherence time* [10] of the channel is large enough so that the channel estimate is still valid, when it is used for the transmit processing algorithm. In *frequency division duplex* (FDD) systems, the two links reside in different frequency bands, and, hence, the channel parameters are different for uplink and downlink. This lack of knowledge can be overcome by exploiting the slowly changing properties of the channel which are independent from frequency (see, e.g., [11]), *viz.*, path delays and average path attenuation, or can be transformed from the uplink to the downlink frequency, i.e., steering vectors (e.g., [12] and [13]). In this article, we assume that the fast changing properties of the *multiple input, multiple output* (MIMO) channel are known at the BS. To this end, feedback from the MSs to the BS is necessary in FDD systems. However, note that spatial temporal transmit processing based on slowly changing channel properties is possible as proposed in [14]–[16].

Contrary to nonlinear approaches, like in [17]–[24], we employ linear transmit processing at the BS to end up with the simplest possible receivers at the MSs which are filters matched to the signal waveform. Thus, no channel estimation is necessary at the MSs, if transmit processing, *pre-equalization*, or *precoding* is used at the BS. Furthermore, we assume that the BS knows *a priori* which type of signal processing is applied at the MSs. This presumption is equivalent to the assumption that the receiver knows the adopted signal processing at the transmitter in the case of receive processing.

Many publications focus on joint optimization of transmit and receive filters, e.g., [25]–[30]. Obviously, transmit processing, as well as receive processing, belong to a constrained category of the joint optimization of transmit and receive filters, i.e., transmit processing and receive processing are suboptimum solutions of the joint optimization. However, besides the advantage of simplifying one side of the link, that is, the receiver for transmit processing and the transmitter for receive processing, the two approaches are of high practical importance in systems with noncooperative transmitters (e.g., uplink of a cellular system), where the transmit signals of the transmitters cannot be cooperatively pretransformed or noncooperative receivers (e.g., downlink of a cellular system

Manuscript received May 16, 2003; revised June 3, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiadong Wang.

The authors are with the Institute for Circuit Theory and Signal Processing, Technical University of Munich, 80290 Munich, Germany (e-mail: joham@nws.ei.tum.de; utschick@nws.ei.tum.de; nossek@nws.ei.tum.de.).

Digital Object Identifier 10.1109/TSP.2005.850331

or *broadcast channel* [31]), where the received signals of the receivers cannot be cooperatively post-transformed. Note that the joint optimization is solely based on the assumptions of fully cooperative receivers and fully cooperative transmitters, that is, the transmit signals and the receive signals can be cooperatively pre- and post-transformed, respectively.

The idea of exploiting the reciprocity of the uplink and downlink channels in TDD systems by applying a linear transmit filter in the downlink was introduced by Henry *et al.* in [32] for flat fading channels and multiple antenna elements at the BS. The approach of [32] is a special case of the *prerake* proposed by Esmailzadeh *et al.* [33], [34] which resulted from the intuitive idea to move the part of the RxMF matched to the channel (rake) from the receiver to the transmitter. The prerake has been researched extensively (e.g., [35]–[41]) and Revés *et al.* [42] reported the implementation of the prerake in the downlink of a DS–CDMA indoor system. Joham *et al.* showed in [43] that the prerake is the *transmit matched filter* (TxMF) which maximizes the power of the desired signal at the respective receiver with a transmit power constraint. In [44], Wang *et al.* optimized the prerake under the assumption that the receiver is equipped with a rake (see also [45] and [46]) and Noll Barreto *et al.* [47] extended the prerake concept with a rake matched to the prerake and the channel at the receiver. Both approaches maximize the *signal-to-noise ratio* (SNR) at the rake output, but lead to an increased complexity at the receiver due to the additional operations necessary for the rake which is not only matched to the channel but also to the transmit filter.

Because the transmitter (BS) has no influence on the noise at the receivers (MSs), the most intuitive approach for transmit processing is a *transmit zero-forcing filter* (TxZF) which removes all *interference* at the MSs. Tang *et al.* presented a *predecorrelating* technique for flat fading scenarios in [48] and Liu *et al.* [49] proposed zero-forcing pre-equalization in a *TDD-time division multiple access* (TDD–TDMA) system for the single user and multiuser case. A TxZF for synchronous and asynchronous CDMA systems over flat fading channels was filed by Weerackody [50]. In [51], Vojčić *et al.* showed that zero-forcing precoding results from the MMSE criterion for the detector signal at the receiver and Brandt–Pearce *et al.* [52] presented a symbol-wise zero-forcing prefilter under the assumption of a small delay spread compared to symbol time. Montalbano *et al.* presented zero-forcing spatiotemporal transmit processing for TDD–CDMA in [53] and also a solution for the TxZF for FDD–CDMA in [14], whereas Forster *et al.* [15] developed the TxZF for FDD–TDMA in frequency domain. In [54], Karimi *et al.* compared transmit processing and receive processing for flat fading MIMO systems and Sampath *et al.* [55] derived the TxZF as matrix FIR filter for frequency selective MIMO systems by exploiting orthogonality properties of steering vectors. The TxZF was applied to TD–CDMA multiuser systems by Baier *et al.* [56] and Joham *et al.* [57], and to TDD–CDMA by Noll Barreto *et al.* [58]. Kowalewski *et al.* [59] examined the influence of channel estimation and change of the channel impulse responses due to the time separation of uplink and downlink (see also [60]), whereas Walke *et al.* [61] compared the TxZF to the RxZF for TD–CDMA systems. Morgado *et al.* [62] developed the TxZF in frequency domain by utilizing the redundancy of nonoverlapping band of a *direct sequence CDMA* (DS–CDMA) signal. In [63], Georgoulis *et al.* compared the TxZF to the TxMF for a TDD–CDMA system. Reynolds *et al.* [64] constructed and analyzed the TxZF for the downlink based on blind channel estimation in the uplink. Guncavdi *et al.* [65] presented a suboptimum TxZF with reduced complexity for synchronous DS–CDMA motivated by a similar receive filter [66] and included a long range fading prediction. Meurer *et al.* [67] proposed to combine the TxZF with a RxZF which would also remove the interference, when the TxZF is not used. However, this approach removes the advantage of transmit processing, i.e., reduced complexity at the receiver.

Contrary to the other two transmit filters, the *transmit Wiener filter* (TxWF) has been proposed and examined only by a few authors, because it cannot be found in such a straightforward way as the TxZF and the TxMF. In [51], Vojčić *et al.* not only noted that the transmit filter minimizing the *mean square error* (MSE) is the TxZF, but also discussed the possibility to include a transmit power constraint. Vojčić *et al.* also reported that this *constrained MMSE transmit filter* (TxCMMSE) outperforms the TxZF for low SNR, but is worse for high SNR, because it is interference limited. Noll Barreto *et al.* in [58] proposed to replace the equality for the transmit power constraint by an inequality, but stated that the resulting optimization has no closed form solution. In [68], Georgoulis *et al.* extended the TxZF of [52] to the TxCMMSE. Joham in [69] reported that the solution for the TxCMMSE can be obtained by finding the positive root of a polynomial and showed that the TxCMMSE is a suboptimum TxWF designed for a fixed SNR. The TxWF was first mentioned by Karimi *et al.* in [54] who obtained the *transmit MMSE filter* by simply adding a weighted identity matrix in the solution of the TxZF in an intuitive way (see also [70]). The necessary optimization for the TxWF was published by Joham *et al.* in [71] and [72] and by Choi *et al.* in [73]. Choi *et al.* compared the TxWF only to the RxMF and the TxMF by *bit-error ratio* (BER) simulations, but did not discuss the convergence of the TxWF to the TxMF for low SNR, whereas Joham *et al.* showed that the TxWF has similar properties as the RxWF, i.e., the TxWF converges to the TxMF for low SNR and to the TxZF for high SNR where the SNR is defined as the ratio of transmit power to noise power at the receiver.

Our contributions are as follows.

1) We show that the transmit filters can be found with similar optimizations as the respective receive filters.
2) We derive the TxWF and discuss its convergence to the other transmit filters for high and low SNR.
3) We compare the different transmit filters and prove analytically that the MSE of the TxZF is lower bounded by the MSE of the TxMF for low SNR but is upper bounded by the MSE of the TxMF for high SNR.
4) We show analytically that the MSEs of the transmit filters and the respective receive filters are the same for uncorrelated symbols and noise.

We will compare the three transmit filters to the respective receive filters in terms of MSE and BER by applying them to a MIMO system under the assumption of perfect channel state information at the transmitter and the receiver. To this end, we explain the system model in Section II and briefly review the RxMF, RxZF, and RxWF in Section III. In Sections IV–VI, we derive the TxMF, TxZF, and TxWF, respectively, and show that their MSEs equal to the MSEs of the respective receive filters in Section VII. Simulation results are given in Section VIII.

A. Notation

Vectors and matrices are denoted by lower-case bold and capital bold letters, respectively. We use $\mathrm{E}[\bullet], \mathrm{Re}(\bullet), \mathrm{tr}(\bullet), (\bullet)^*, (\bullet)^{\mathrm{T}}, (\bullet)^{\mathrm{H}}, \| \bullet \|_2$, and $\| \bullet \|_{\mathrm{F}}$ for expectation, real part of the argument, trace of a matrix, complex conjugation, transposition, conjugate transposition, Euclidian norm, and Frobenius norm, respectively. All random variables are assumed to be zero mean. The covariance matrix of the vector random variable $\boldsymbol{x}$ is denoted by $\boldsymbol{R_x} = \mathrm{E}[\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}]$, whereas the variance of the scalar random variable y is denoted by $\sigma_y^2 = \mathrm{E}[|y|^2]$. The $N \times N$ identity matrix is $\mathbf{1}_N$, the $N \times M$ zero matrix is $\mathbf{0}_{N\times M}$, and the N-dimensional zero vector is denoted by $\mathbf{0}_N$. We use the same definition for the derivative $\partial a(\boldsymbol{B})/\partial \boldsymbol{B}$ of a scalar $a(\boldsymbol{B})$ with respect to the $N \times M$ matrix $\boldsymbol{B}$ as in [74], i.e., each entry of the resulting $N \times M$ matrix is the derivative of the scalar $a(\boldsymbol{B})$ with respect to the respective entry of $\boldsymbol{B}$. Since the cost functions of the investigated optimizations are not analytic, we employ the following derivative (see, e.g., [75])

$$\frac{\partial b(\boldsymbol{A})}{\partial \boldsymbol{A}} = \frac{1}{2}\left(\frac{\partial b(\boldsymbol{A})}{\partial \mathrm{Re}(\boldsymbol{A})} - \mathrm{j}\frac{\partial b(\boldsymbol{A})}{\partial \mathrm{Im}(\boldsymbol{A})}\right)$$

where $b(\boldsymbol{A}) \in \mathbb{C}$ and $\boldsymbol{A} \in \mathbb{C}^{N\times M}$.

II. System Model and Assumptions

We consider a MIMO system as depicted in Fig. 1 which consists of the transmit filter $\boldsymbol{P}$, the channel $\boldsymbol{H}$, and the receive filter $\boldsymbol{G}$. We assume that $\boldsymbol{H} \in \mathbb{C}^{M\times N}$ is tall or square $(M \geq N)$ for receive processing and wide or square $(M \leq N)$ for transmit processing. Moreover, the number of information symbols B does not exceed $\min(M, N)$. If we consider receive processing, the signal processing $\boldsymbol{P} \in \mathbb{C}^{N\times B}$ at the transmitter is *a priori* known to the receiver and the chain $\boldsymbol{HP} \in \mathbb{C}^{M\times B}$ of $\boldsymbol{P}$ and $\boldsymbol{H}$ has full rank, i.e., $\mathrm{rank}(\boldsymbol{HP}) = B$. Accordingly, the filter $\boldsymbol{G} \in \mathbb{C}^{B\times M}$ at the receiver is *a priori* known to the transmitter in the case of transmit processing and $\boldsymbol{GH} \in \mathbb{C}^{B\times N}$ has full rank, that is, $\mathrm{rank}(\boldsymbol{GH}) = B$. With these presumptions, it is possible to design linear transmit filters $\boldsymbol{P}$ and linear receive filters $\boldsymbol{G}$ also with zero-forcing constraints. Note that the filter $\boldsymbol{P}$ for receive processing and the filter $\boldsymbol{G}$ for transmit processing need not be constant over time, i.e., they can depend on the channel realization $\boldsymbol{H}$. Thus, possible examples are the filter $\boldsymbol{H}^{\mathrm{H}}$ matched to the channel or a filter based on the statistics (e.g., covariance matrix) of the channel. Also, note that we do not assume any special structure of the channel matrix $\boldsymbol{H}$. Therefore, the system model, as well as the derivations presented in this paper, are applicable to systems with flat fading and frequency selective fading channels (for the special case of block transmission over FIR channels, $\boldsymbol{H}$ is block Toeplitz, see, e.g., [30]).

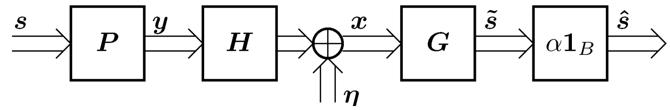


Fig. 1. MIMO system with linear transmit and receive filters.

The transmitted signal $\boldsymbol{y}$ is the desired signal $\boldsymbol{s} \in \mathbb{C}^B$ transformed by the transmit filter (cf. Fig. 1)

$$\boldsymbol{y} = \boldsymbol{Ps} \in \mathbb{C}^N \tag{1}$$

and we assume that the average transmit power is fixed

$$\mathrm{E}\left[\|\boldsymbol{y}\|_2^2\right] = \mathrm{E}\left[\|\boldsymbol{Ps}\|_2^2\right] = \mathrm{tr}(\boldsymbol{PR_sP}^{\mathrm{H}}) = E_{\mathrm{tr}}. \tag{2}$$

After transmission over the channel $\boldsymbol{H}$, the received signal is perturbed by the noise $\boldsymbol{\eta} \in \mathbb{C}^M$ and passed through the receive filter $\boldsymbol{G}$ to obtain the estimate

$$\tilde{\boldsymbol{s}} = \boldsymbol{G}(\boldsymbol{HPs} + \boldsymbol{\eta}) \in \mathbb{C}^B. \tag{3}$$

Note that we assume that the noise is uncorrelated with the symbols, that is, $\mathrm{E}[\boldsymbol{\eta}\boldsymbol{s}^{\mathrm{H}}] = \mathbf{0}_{M\times B}$.

The above MIMO system will be used to compare the different receive and transmit filters by computing the MSE

$$\varepsilon = \mathrm{E}\left[\|\boldsymbol{s} - \hat{\boldsymbol{s}}\|_2^2\right] = \mathrm{E}\left[\|\boldsymbol{s} - \alpha\tilde{\boldsymbol{s}}\|_2^2\right] \in \mathbb{R}_{0,+} \tag{4}$$

where we included a scalar Wiener filter α at the receiver (cf. Section III-C) to be able to give reasonable expressions for the MSEs of all filters, especially the RxMF and the transmit filters. Note that the scalar α can be interpreted as an *automatic gain control* which is necessary in any real MIMO system. Also, note that above MSE is different from the MSE used to find the receive and transmit filters.

We also define the SNR as the transmit power per data stream divided by the noise power per receive antenna element

$$\gamma = \frac{E_{\mathrm{tr}}/B}{\mathrm{tr}(\boldsymbol{R_\eta})/M}. \tag{5}$$

III. Receive Filters

The classical way to deal with the distortions generated by the channel and the perturbation caused by the noise is receive processing, where the receive filter $\boldsymbol{G}$ is designed upon the knowledge of the transmit filter $\boldsymbol{P}$, the channel $\boldsymbol{H}$ (see Fig. 1), and the covariance matrices $\boldsymbol{R_s}$ and $\boldsymbol{R_\eta}$. In this section, we briefly discuss the three receive filters and compute the resulting MSEs.

A. Receive Matched Filter (RxMF)

The RxMF maximizes the SNR at the filter output (e.g., [1]). Therefore, it is optimum for noise limited scenarios. As the RxMF does not regard interference, one way to derive the matrix RxMF $\boldsymbol{G}_{\mathrm{MF}}$ is splitting the vector estimate $\tilde{\boldsymbol{s}}$ into its scalar components, designing row vector RxMFs for the scalar signals, and combining the rows to the matrix filter $\boldsymbol{G}_{\mathrm{MF}}$. Alternatively, we can employ following optimization [76], where we separate the desired signal portion in the estimate from the noise portion by correlation

$$\boldsymbol{G}_{\mathrm{MF}} = \mathrm{argmax}_{\boldsymbol{G}} \frac{|\mathrm{E}[\boldsymbol{s}^{\mathrm{H}}\tilde{\boldsymbol{s}}]|^2}{\mathrm{E}[\|\boldsymbol{G}\boldsymbol{\eta}\|_2^2]} \tag{6}$$

whose solution can be obtained by setting the derivation of the cost function with respect to $\boldsymbol{G}$ to zero and reads as

$$\boldsymbol{G}_{\mathrm{MF}} = \alpha' \boldsymbol{R_s} \boldsymbol{P}^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{R}_{\boldsymbol{\eta}}^{-1} \in \mathbb{C}^{B\times M}. \tag{7}$$

The scalar $\alpha' \in \mathbb{C}$ can be freely chosen and we set $\alpha' = 1$ in the following.

B. Receive Zero-Forcing Filter (RxZF)

Another type of linear receive processing arises from the constraint that $\tilde{\boldsymbol{s}}$ is an interference-free estimate of $\boldsymbol{s}$. Thus, we have to fulfill following equation [see (3)]:

$$\tilde{\boldsymbol{s}}\,|_{\boldsymbol{\eta}=\mathbf{0}_M} = \boldsymbol{GHPs} \equiv \boldsymbol{s}.$$

Since $\boldsymbol{s}$ is arbitrary and unknown to the receiver, the chain of the transmit filter $\boldsymbol{P}$, the channel $\boldsymbol{H}$, and the receive filter $\boldsymbol{G}$ must result in an identity mapping

$$\boldsymbol{GHP} = \mathbf{1}_B. \tag{8}$$

Note that this constraint can be fulfilled, because we assumed $\mathrm{rank}(\boldsymbol{HP}) = B$. With the above constraint and (3), the MSE of the RxZF (without the scalar Wiener filter of Fig. 1) can be shown to be the noise power at the filter output [3]

$$\mathrm{E}\left[\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2^2\right] = \mathrm{E}\left[\|\boldsymbol{G}\boldsymbol{\eta}\|_2^2\right].$$

The RxZF minimizes the above MSE and removes the interference [cf. (8)]

$$\boldsymbol{G}_{\mathrm{ZF}} = \mathrm{argmin}_{\boldsymbol{G}} \mathrm{E}\left[\|\boldsymbol{G}\boldsymbol{\eta}\|_2^2\right] \quad \text{s.t.: } \boldsymbol{GHP} = \mathbf{1}_B. \tag{9}$$

With the Lagrangian multiplier method (e.g., [77]), we obtain the RxZF

$$\boldsymbol{G}_{\mathrm{ZF}} = \left(\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}\boldsymbol{HP}\right)^{-1} \boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1} \in \mathbb{C}^{B\times M}. \tag{10}$$

We see that the RxZF is a RxMF followed by the transformation $(\boldsymbol{R_s}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}\boldsymbol{HP})^{-1}$ for interference suppression.

C. Receive Wiener Filter (RxWF)

The RxWF [78], [1] minimizes the MSE without an additional constraint [see also (3)]

$$\boldsymbol{G}_{\mathrm{WF}} = \mathrm{argmin}_{\boldsymbol{G}} \mathrm{E}\left[\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2^2\right]. \tag{11}$$

After setting the derivative of the MSE to zero, we yield

$$\boldsymbol{G}_{\mathrm{WF}} = \left(\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}\boldsymbol{HP} + \boldsymbol{R}_{\boldsymbol{s}}^{-1}\right)^{-1} \boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1} \tag{12}$$

where we utilized the *matrix inversion lemma* (e.g., [79]). Equation (12) helps to understand the dependence of the RxWF on the SNR. For decreasing SNR, the first summand of the inverse gets smaller compared to the second summand and the RxWF converges to the RxMF [compared to (7)]. On the contrary, the second summand can be neglected for high SNR and the RxWF converges to the RxZF [cf. (10)]. From (12), we also see that the RxWF is the RxMF followed by the interference canceller $(\boldsymbol{R_s}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}\boldsymbol{HP} + \mathbf{1}_B)^{-1}$.

Up until now, we only reviewed different approaches to obtain the matrix filter $\boldsymbol{G}$ in Fig. 1, but we stated in Section II that we need a scalar Wiener filter α at the end of the filter chain to get a reasonable comparison.[1] The scalar Wiener filter α minimizes the MSE $\mathrm{E}[\|\boldsymbol{s} - \alpha\tilde{\boldsymbol{s}}\|_2^2]$ of (4) and is found in a similar way as $\boldsymbol{G}_{\mathrm{WF}}$. We obtain for the scalar Wiener filter

$$\alpha = \frac{\mathrm{tr}\left(\boldsymbol{R_s}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right)}{\mathrm{tr}\left(\boldsymbol{G}\left(\boldsymbol{HPR_s}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_{\boldsymbol{\eta}}\right)\boldsymbol{G}^{\mathrm{H}}\right)} \in \mathbb{C} \tag{13}$$

whose MSE reads as [cf. (4)]

$$\varepsilon = \mathrm{tr}(\boldsymbol{R_s}) - \frac{|\mathrm{tr}(\boldsymbol{R_s}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}})|^2}{\mathrm{tr}(\boldsymbol{G}(\boldsymbol{HPR_s}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_{\boldsymbol{\eta}})\boldsymbol{G}^{\mathrm{H}})}. \tag{14}$$

D. MSEs of the Receive Filters

Due to (14) the RxMF [see (7)], the RxZF [see (10)], and the RxWF [see (12)] applied to the system in Fig. 1 result in the MSEs

$$\varepsilon_{\mathrm{RxMF}} = \mathrm{tr}(\boldsymbol{R_s}) - \frac{\mathrm{tr}^2(\boldsymbol{J}_{\mathrm{rx}}\boldsymbol{R_s})}{\mathrm{tr}\left(\left(\boldsymbol{J}_{\mathrm{rx}}^2 + \boldsymbol{J}_{\mathrm{rx}}\right)\boldsymbol{R_s}\right)} \tag{15}$$

$$\varepsilon_{\mathrm{RxZF}} = \mathrm{tr}(\boldsymbol{R_s}) - \frac{\mathrm{tr}^2(\boldsymbol{R_s})}{\mathrm{tr}(\boldsymbol{R_s}) + \mathrm{tr}\left(\boldsymbol{J}_{\mathrm{rx}}^{-1}\boldsymbol{R_s}\right)} \tag{16}$$

and

$$\varepsilon_{\mathrm{RxWF}} = \mathrm{tr}(\boldsymbol{R_s}) - \mathrm{tr}((\boldsymbol{J}_{\mathrm{rx}} + \mathbf{1}_B)^{-1}\boldsymbol{J}_{\mathrm{rx}}\boldsymbol{R_s}) \tag{17}$$

respectively. Here, we introduced

$$\boldsymbol{J}_{\mathrm{rx}} = \boldsymbol{R}_{\boldsymbol{s}}^{1/2,\mathrm{H}}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}\boldsymbol{HPR}_{\boldsymbol{s}}^{1/2}$$

and we used the matrix inversion lemma for (17). The square-root matrix $\boldsymbol{R}_{\boldsymbol{s}}^{1/2} = \boldsymbol{R}_{\boldsymbol{s}}^{1/2,\mathrm{H}}$ fulfills $\boldsymbol{R_s} = (\boldsymbol{R}_{\boldsymbol{s}}^{1/2})^2$. Note that $\boldsymbol{J}_{\mathrm{rx}} = \gamma B\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{HP}/\mathrm{tr}(\boldsymbol{PP}^{\mathrm{H}})$ for uncorrelated symbols and noise, i.e., $\boldsymbol{R_s} = \sigma_{\boldsymbol{s}}^2\mathbf{1}_B$ and $\boldsymbol{R}_{\boldsymbol{\eta}} = \sigma_{\boldsymbol{\eta}}^2\mathbf{1}_M$.

IV. Transmit Matched Filter (TxMF)

Contrary to the previous section on receive filters, we assume an *a priori* known constant receive filter $\boldsymbol{G}$ in this and the following sections. Thus, the transmitter can design a precoding filter $\boldsymbol{P}$, if the channel matrix $\boldsymbol{H}$ (besides the *a priori* known $\boldsymbol{G}$ and $\boldsymbol{R_s}$) is available at the transmitter. As we presume that the transmitter perfectly knows the instantaneous channel matrix $\boldsymbol{H}$, we can design a transmit filter maximizing the desired signal portion at the receiver.

A. Derivation of TxMF

The TxMF was introduced by Esmailzadeh *et al.* [33] by moving the channel matched filter $\boldsymbol{H}^{\mathrm{H}}$ from the receiver to the transmitter, but they only gave an intuitive explanation. The

[1]For example, the output amplitude of the RxMF depends on the channel realization $\boldsymbol{H}$. The weighting α is necessary to minimize this dependency.

TxMF can be derived by utilizing the same cost function as for the RxMF [cf. (6)], but we have to ensure that only the available transmit power $E_{\rm tr}$ is used

$$\boldsymbol{P}_{\rm MF} = \mathrm{argmax}_{\boldsymbol{P}} \frac{|\mathrm{E}[\boldsymbol{s}^{\rm H}\tilde{\boldsymbol{s}}]|^2}{\mathrm{E}[\|\boldsymbol{G\eta}\|_2^2]} \quad \text{s.t.: } \mathrm{E}\left[\|\boldsymbol{Ps}\|_2^2\right] = E_{\rm tr}. \quad (18)$$

With the Lagrangian multiplier method (e.g., [77]), we end up with the TxMF

$$\boldsymbol{P}_{\rm MF} = \beta_{\rm TxMF}\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H} \in \mathbb{C}^{N\times B}$$

and

$$\beta_{\rm TxMF} = \sqrt{\frac{E_{\rm tr}}{\mathrm{tr}(\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}\boldsymbol{R_s}\boldsymbol{GH})}}. \quad (19)$$

Note that the optimization in (18) also allows a complex valued scaling $\beta_{\rm TxMF}$, but we have chosen $\beta_{\rm TxMF} \in \mathbb{R}_+$. In [33], the same structure with a similar scalar factor was presented. However, the scalar factor was justified by comparing the SNR of the TxMF at the receiver to the one of the RxMF.

When we compare the result for the TxMF in (19) to the RxMF in (7), we can observe that the matched filter at the transmitter is also the conjugate transpose of the subsequent filters. However, the TxMF does not regard the properties of the noise $\boldsymbol{\eta}$, since the noise covariance matrix is not included in (19). This result is not surprising, because the transmitter has no influence on the noise at the receiver. If the receiver incorporates a noise whitening filter together with its matched filter, i.e., $\boldsymbol{G}' = \boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}$, the transmitter is able to adapt to the properties of the noise. But this approach would increase the system complexity due to the necessity to estimate the noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}}$ and to feedback it from the receiver to the transmitter. More surprisingly, the structure of the TxMF is not influenced by the covariance matrix $\boldsymbol{R_s}$ of the desired signal $\boldsymbol{s}$. Our explanation for this result is the inability of the receiver to deal with a transmit filter $\boldsymbol{P}$ which depends upon $\boldsymbol{R_s}$, since the receiver does not consider the correlations of the different entries of $\boldsymbol{s}$ as it uses the *a priori* defined receive filter $\boldsymbol{G}$ which is independent of $\boldsymbol{R_s}$.

B. MSE of TxMF

Due to the transmit power restriction, the amplitude at the output of the receive filter $\boldsymbol{G}$ depends on the transmit power $E_{\rm tr}$, the channel realization $\boldsymbol{H}$, and the choice of the fixed receive filter $\boldsymbol{G}$. We can expect that the output $\tilde{\boldsymbol{s}}$ of $\boldsymbol{G}$ has a wrong amplitude in most cases. Therefore, we need the scalar Wiener filter α of (13) to correct the amplitude, when we apply transmit processing. With (14), the resulting MSE of the TxMF can be written as

$$\varepsilon_{\rm TxMF} = \mathrm{tr}\,(\boldsymbol{R_s}) - \frac{\mathrm{tr}^2(\boldsymbol{J}_{\rm tx}\boldsymbol{R_s})}{\mathrm{tr}((\boldsymbol{J}_{\rm tx}^2 + \boldsymbol{J}_{\rm tx})\boldsymbol{R_s})} \quad (20)$$

where we introduced

$$\boldsymbol{J}_{\rm tx} = \boldsymbol{GHH}^{\rm H}\boldsymbol{G}^{\rm H}\frac{E_{\rm tr}}{\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\rm H})}.$$

Note that $\boldsymbol{J}_{\rm tx} = \gamma B\boldsymbol{GHH}^{\rm H}\boldsymbol{G}^{\rm H}/\mathrm{tr}(\boldsymbol{GG}^{\rm H})$ for uncorrelated symbols and noise. We observe that above expression for the TxMF's MSE has the same form as the MSE of the RxMF in (15)—only $\boldsymbol{J}_{\rm tx}$ has to be replaced by $\boldsymbol{J}_{\rm rx}$, but as $\boldsymbol{J}_{\rm tx}$ is different from $\boldsymbol{J}_{\rm rx}$, the two MSEs are different in general.

We can follow from the SNR definition in (5) that the entries of $\boldsymbol{J}_{\rm tx}$ converge to zero and infinity for $\gamma \to 0$ and $\gamma \to \infty$, respectively. Hence, we get for low and high SNR

$$\gamma \to 0: \quad \varepsilon_{\rm TxMF} \to \mathrm{tr}(\boldsymbol{R_s}) - \mathrm{tr}(\boldsymbol{J}_{\rm tx}\boldsymbol{R_s}) \quad (21)$$

and

$$\gamma \to \infty: \quad \varepsilon_{\rm TxMF} \to \mathrm{tr}(\boldsymbol{R_s}) - \frac{\mathrm{tr}^2(\boldsymbol{J}_{\rm tx}\boldsymbol{R_s})}{\mathrm{tr}(\boldsymbol{J}_{\rm tx}^2\boldsymbol{R_s})} \quad (22)$$

respectively. As expected, the MSE for low SNR $(\gamma \to 0)$ converges to $\mathrm{tr}(\boldsymbol{R_s})$, because the scalar Wiener filter α becomes zero. The MSE for high SNR $(\gamma \to \infty)$ is independent of γ and different from zero, when the eigenvalues of $\boldsymbol{J}_{\rm tx}$ are not identical (see Appendix A). The TxMF is *interference limited* due to this behavior for high SNR, since no noise is present and the error follows from the remaining interference.

V. Transmit Zero-Forcing Filter (TxZF)

If the transmitter knows the channel matrix $\boldsymbol{H}$, the constant signal processing $\boldsymbol{G}$ at the receiver, and the signal covariance matrix $\boldsymbol{R_s}$, not only transmit processing which maximizes the received desired signal as with the TxMF is possible, but also a transmit filter which generates a received signal without interference. We call the transmit filter with this property the TxZF.

A. Derivation of TxZF

To avoid the limitation due to interference caused by the TxMF we design a transmit filter which completely removes the interference. Thus, we force the chain of the transmit filter $\boldsymbol{P}$, channel $\boldsymbol{H}$, and the receive filter $\boldsymbol{G}$ to be an identity mapping as for the RxZF [cf. (8)]

$$\boldsymbol{GHP} = \mathbf{1}_B.$$

Since the transmitter has no influence on the noise at the receiver, this constraint seems to be optimum, because we remove all perturbation caused by the transmitter, namely the interference. However, as will be shown in the next section, it is beneficial to allow some interference at the receiver to increase the received power of the desired signal. We have to minimize the transmit power instead of the receive noise power [compare with (9)] to yield the TxZF

$$\tilde{\boldsymbol{P}}_{\rm ZF} = \mathrm{argmin}_{\boldsymbol{P}}\mathrm{E}\left[\|\boldsymbol{Ps}\|_2^2\right] \quad \text{s.t.: } \boldsymbol{GHP} = \mathbf{1}_B. \quad (23)$$

After setting the derivation of the appropriate Lagrangian function to zero and with the constraint of above optimization, the resulting transmit filter can be written as

$$\tilde{\boldsymbol{P}}_{\rm ZF} = \boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}(\boldsymbol{GHH}^{\rm H}\boldsymbol{G}^{\rm H})^{-1} \in \mathbb{C}^{N\times B}. \quad (24)$$

This result is not satisfactory, because the resulting transmit power $\mathrm{tr}(\tilde{\boldsymbol{P}}_{\rm ZF}\boldsymbol{R_s}\tilde{\boldsymbol{P}}_{\rm ZF}^{\rm H})$ has no predefined value and depends upon the channel $\boldsymbol{H}$. A heuristic approach to deal with this problem is to indroduce a scalar $\beta \in \mathbb{R}_+$ which scales the transmit filter, i.e., $\beta\tilde{\boldsymbol{P}}_{\rm ZF}$, to set the transmit power to a fixed value (e.g., [51], [58])

$$\beta_{\rm TxZF}^2\mathrm{tr}\left(\tilde{\boldsymbol{P}}_{\rm ZF}\boldsymbol{R_s}\tilde{\boldsymbol{P}}_{\rm ZF}^{\rm H}\right) \equiv E_{\rm tr}.$$

Therefore, the TxZF reads as

$$\boldsymbol{P}_{\rm ZF} = \beta_{\rm TxZF}\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H})^{-1} \in \mathbb{C}^{N\times B} \tag{25}$$

with the scaling factor

$$\beta_{\rm TxZF} = \sqrt{\frac{E_{\rm tr}}{{\rm tr}((\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H})^{-1}\boldsymbol{R}_{\boldsymbol{s}})}}. \tag{26}$$

To find a deeper understanding of the TxZF, let us examine following optimization:[2]

$$\begin{aligned}\{\boldsymbol{P}_{\rm ZF}, \beta_{\rm TxZF}\} &= {\rm argmin}_{\{\boldsymbol{P},\beta\}}\beta^{-2}\\ \text{s.t.: } \boldsymbol{G}\boldsymbol{H}\boldsymbol{P} &= \beta\mathbf{1}_B \quad\text{and}\quad {\rm E}\left[\|\boldsymbol{P}\boldsymbol{s}\|_2^2\right] = E_{\rm tr}\end{aligned} \tag{27}$$

where we defined the mean gain of the filter chain $\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}$ for the desired part of the received signal to be equal to

$$\beta = \frac{1}{B}{\rm tr}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}). \tag{28}$$

In general, the gain β is a complex number, but for the optimization in (27), we assume that $\beta \in \mathbb{R}_+$. With the Lagrangian multiplier method, it can be easily shown that (27) leads to the previous result in (25) and (26). Hence, the optimization in (23) combined with the intuitive scaling by a scalar is equivalent to the optimization in (27). We can see that the TxZF leads to a signal at the receive filter output which is free of interference and whose transmit power is constraint to be $E_{\rm tr}$. Moreover, the TxZF maximizes the gain from the transmit filter input to the receive filter output.

If the matrix $\boldsymbol{G}\boldsymbol{H}$ is bad conditioned (e.g., [80]), that is, the ratio of the maximum to the minimum singular value of $\boldsymbol{G}\boldsymbol{H}$ is large, the matrix inversion in (25) leads to a small gain of the TxZF [cf. (26) and (28)]

$$\beta_{\rm TxZF} = \frac{{\rm tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}_{\rm ZF}\right)}{B} = \sqrt{\frac{E_{\rm tr}}{\sigma_{\boldsymbol{s}}^2{\rm tr}\left(\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}\right)^{-1}\right)}}$$

compared to the gain of the TxMF [cf. (19) and (28)][3]

$$\beta'_{\rm TxMF} = \frac{{\rm tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}_{\rm MF}\right)}{B} = \frac{1}{B}\sqrt{\frac{E_{\rm tr}{\rm tr}\left(\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}\right)}{\sigma_{\boldsymbol{s}}^2}} \tag{29}$$

where we assumed uncorrelated symbols, i.e., $\boldsymbol{R}_{\boldsymbol{s}} = \sigma_{\boldsymbol{s}}^2\mathbf{1}_B$. The proof of this statement can be found in Appendix B. A similar property of the RxZF is often called *noise enhancement* (e.g., [1]).

B. MSE of TxZF

The MSE of the scalar Wiener filter in (13) for the TxZF can be expressed as [see (14)]

$$\varepsilon_{\rm TxZF} = {\rm tr}(\boldsymbol{R}_{\boldsymbol{s}}) - \frac{{\rm tr}^2(\boldsymbol{R}_{\boldsymbol{s}})}{{\rm tr}(\boldsymbol{R}_{\boldsymbol{s}}) + {\rm tr}(\boldsymbol{J}_{\rm tx}^{-1}\boldsymbol{R}_{\boldsymbol{s}})} \tag{30}$$

which is different from the MSE of the RxZF, in general (compare to (16)), since $\boldsymbol{J}_{\rm tx} \neq \boldsymbol{J}_{\rm rx}$. Similar to the previous section, we examine the behavior of the TxZF with respect to the SNR γ defined in (5)

$$\gamma \to 0: \quad \varepsilon_{\rm TxZF} \to {\rm tr}(\boldsymbol{R}_{\boldsymbol{s}}) - \frac{{\rm tr}^2(\boldsymbol{R}_{\boldsymbol{s}})}{{\rm tr}\left(\boldsymbol{J}_{\rm tx}^{-1}\boldsymbol{R}_{\boldsymbol{s}}\right)} \tag{31}$$

and

$$\gamma \to \infty: \quad \varepsilon_{\rm TxZF} \to {\rm tr}\left(\boldsymbol{J}_{\rm tx}^{-1}\boldsymbol{R}_{\boldsymbol{s}}\right). \tag{32}$$

The MSE of the TxZF tends toward zero for high SNR ($\gamma \to \infty$), since the entries of $\boldsymbol{J}_{\rm tx}$ converge to infinity. Thus, the TxZF is not interference limited as the TxMF due to the first constraint in (27). For low SNR ($\gamma \to 0$), the MSE converges to the maximum value ${\rm tr}(\boldsymbol{R}_{\boldsymbol{s}})$ like the MSE of the TxMF [cf. (21)] because of the scalar Wiener filter α at the receiver. However, the TxZF is outperformed by the TxMF for low SNR ($\gamma \to 0$) as we show in Appendix C. This gives the motivation to find a transmit filter which is optimum for all SNR values similar to the RxWF.

VI. Transmit Wiener Filter (TxWF)

In the previous sections, we have seen that the TxMF is worse than the TxZF for high SNR, but outperforms the TxZF for low SNR. This dependence on the SNR can also be observed, when we compare the respective receive filters, since $\boldsymbol{J}_{\rm rx}$ has the same convergence properties as $\boldsymbol{J}_{\rm tx}$ depending on the SNR γ. The RxWF finds the optimum tradeoff between the signal maximization of the RxMF and the interference elimination of the RxZF, because the MSE of the RxWF is always smaller than the MSEs of the RxMF and the RxZF. In this section, we will present the TxWF and show its superiority compared to the TxMF and the TxZF. Note that the knowledge of the noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}}$ (besides $\boldsymbol{H}, \boldsymbol{G}$, and $\boldsymbol{R}_{\boldsymbol{s}}$) is necessary for the design of the TxWF contrary to the TxMF and TxZF.

A. Discussion of Transmit Filters Related to the TxWF

As the RxWF was found by minimizing the MSE [cf. (11)], the TxWF might result from following optimization:

$$\boldsymbol{P}_{\rm MMSE} = {\rm argmin}_{\boldsymbol{P}}{\rm E}\left[\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2^2\right]. \tag{33}$$

The minimum MSE can be obtained by setting the derivation with respect to the transmit filter $\boldsymbol{P}$ to zero. Hence, the transmit filter $\boldsymbol{P}_{\rm MMSE}$ has to fulfill following requirement:

$$\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}_{\rm MMSE} = \boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}.$$

Note that the matrix $\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}\boldsymbol{G}\boldsymbol{H}$ is not invertible under the assumptions of Section II, but the above equation is solvable, since the columns of the matrix $\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}$ lie in the span of the matrix $\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}\boldsymbol{G}\boldsymbol{H}$. One possible solution is the TxZF without scalar scaling [see (24)]

$$\boldsymbol{P}_{\rm MMSE} = \boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^{\rm H}\boldsymbol{G}^{\rm H})^{-1} \in \mathbb{C}^{N\times B}.$$

This result was obtained in [51], [53], and [54]. We could conclude that the TxZF minimizes the MSE, but, as we discussed in Section V, the above solution is only valid for unconstrained transmit power and the TxZF is outperformed by the TxMF for low SNR. Consequently, we must include a constraint for the

[2]The special choice for the objective function will be clear with the discussion in Appendix D.

[3]Note that the gain $\beta'_{\rm TxMF}$ is different from $\beta_{\rm TxMF}$ in (19).

transmit power. In [58] and [68], an upper bound for the transmit power was introduced[4]

$$\boldsymbol{P}_{\mathrm{CMMSE}}=\mathrm{argmin}_{\boldsymbol{P}}\mathrm{E}\left[\|\boldsymbol{s}-\tilde{\boldsymbol{s}}\|_2^2\right] \quad \text{s.t.: } \mathrm{E}\left[\|\boldsymbol{P}\boldsymbol{s}\|_2^2\right]\le E_{\mathrm{tr}}. \tag{34}$$

With the Lagrangian multiplier $\lambda \geq 0$ (see, e.g., [77]), we get the *constrained MMSE transmit filter* (TxCMMSE)

$$\boldsymbol{P}_{\mathrm{CMMSE}}=(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}\boldsymbol{H}+\lambda\mathbf{1}_N)^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\in\mathbb{C}^{N\times B} \tag{35}$$

where λ can be computed with the inequality

$$\mathrm{tr}((\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}\boldsymbol{H}+\lambda\mathbf{1}_N)^{-2}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{G}\boldsymbol{H})\le E_{\mathrm{tr}}.$$

The Lagrangian multiplier λ is 0 for large available transmit power E_{tr}, otherwise it is the only positive real root of a polynomial [69]. When the transmitter can use a large transmit power ($E_{\mathrm{tr}}\to\infty$), the resulting filter is equal to the TxZF without scalar scaling of (24) as can be seen from (35) after applying the matrix inversion lemma, because the constraint of (34) is inactive ($\lambda=0$) in this case. For small available transmit power ($E_{\mathrm{tr}}\to 0$), the constraint is active and $\lambda\to\infty$, since the transmit power necessary for the unscaled TxZF is larger than E_{tr}. Hence, the transmit filter in (35) converges to the TxMF.

This behavior with respect to the available transmit power shows the relationship of the TxCMMSE to the RxWF. However, the above transmit filter is independent of the properties of the noise at the receiver. Thus, the TxCMMSE will be like the TxZF, if the available transmit power is large enough, even when the noise power is very large, but we have learned in the last section that the TxMF outperforms the TxZF in this SNR region, since the power of the received signal is larger for the TxMF than for the TxZF.

B. Derivation of TxWF

The amplitude β of the desired portion in the received signal has to be as large as possible to combat the effect of the noise, because the automatic gain control of the receiver will not only scale the desired portion but also the noise portion of the received signal with β^{-1}. The TxWF includes the weighting with β^{-1} in the definition of the MSE and uses the whole available transmit power [72], [73][5]

$$\begin{aligned}\{\boldsymbol{P}_{\mathrm{WF}},\beta_{\mathrm{TxWF}}\}&=\mathrm{argmin}_{\{\boldsymbol{P},\beta\}}\mathrm{E}\left[\|\boldsymbol{s}-\beta^{-1}\tilde{\boldsymbol{s}}\|_2^2\right]\\ &\text{s.t.: } \mathrm{E}\left[\|\boldsymbol{P}\boldsymbol{s}\|_2^2\right]=E_{\mathrm{tr}}.\end{aligned} \tag{36}$$

We can find necessary conditions for the transmit filter $\boldsymbol{P}$ and the weight $\beta\in\mathbb{R}_+$ by constructing the Lagrangian function

$$L(\boldsymbol{P},\beta,\lambda)=\mathrm{E}\left[\|\boldsymbol{s}-\beta^{-1}\tilde{\boldsymbol{s}}\|_2^2\right]+\lambda(\mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{P}^{\mathrm{H}})-E_{\mathrm{tr}})$$

with the Lagrangian multiplier $\lambda\in\mathbb{R}$ and setting its derivations to zero (see, e.g., [77])

$$\begin{aligned}\frac{\partial L(\boldsymbol{P},\beta,\lambda)}{\partial\boldsymbol{P}}&=\beta^{-2}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}^{*}\boldsymbol{H}^{*}\boldsymbol{P}^{*}\boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{T}}+\lambda\boldsymbol{P}^{*}\boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{T}}\\ &\quad-\beta^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{T}}=\mathbf{0}_{N\times B}\end{aligned} \tag{37}$$

and

$$\begin{aligned}\frac{\partial L(\boldsymbol{P},\beta,\lambda)}{\partial\beta}&=\mathrm{tr}(-\boldsymbol{G}(\boldsymbol{H}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{R}_{\boldsymbol{\eta}})\boldsymbol{G}^{\mathrm{H}}\\ &\quad+\beta\mathrm{Re}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{s}}))2\beta^{-3}=0\end{aligned} \tag{38}$$

where we used $\partial\mathrm{tr}(\boldsymbol{A}\boldsymbol{B})/\partial\boldsymbol{A}=\boldsymbol{B}^{\mathrm{T}}$. The structure of the resulting transmit filter follows from (37)

$$\boldsymbol{P}(\lambda\beta^2)=\beta\tilde{\boldsymbol{P}}(\lambda\beta^2)$$

with

$$\tilde{\boldsymbol{P}}(\lambda\beta^2)=(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}\boldsymbol{H}+\lambda\beta^2\mathbf{1}_N)^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}$$

and

$$\beta=\sqrt{\frac{E_{\mathrm{tr}}}{\mathrm{tr}(\tilde{\boldsymbol{P}}(\lambda\beta^2)\boldsymbol{R}_{\boldsymbol{s}}\tilde{\boldsymbol{P}}^{\mathrm{H}}(\lambda\beta^2))}}$$

where we used the constraint of (36). In contrast to the TxMF and the TxZF, the structural part $\tilde{\boldsymbol{P}}(\lambda\beta^2)$ of the TxWF solution depends on the scalar β, too. Thus, the optimum scaling is the solution of an implicit function. Fortunately, by introducing $\xi=\lambda\beta^2\in\mathbb{R}$ and the determination of (38), we find[6]

$$\begin{aligned}&\mathrm{tr}(\beta^2\boldsymbol{G}\boldsymbol{H}\tilde{\boldsymbol{P}}(\xi)\boldsymbol{R}_{\boldsymbol{s}}\tilde{\boldsymbol{P}}^{\mathrm{H}}(\xi)\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}+\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}}\\ &\quad-\beta^2\boldsymbol{G}\boldsymbol{H}\tilde{\boldsymbol{P}}(\xi)\boldsymbol{R}_{\boldsymbol{s}})=0\end{aligned}$$

or

$$\begin{aligned}&\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})-\xi\beta^2\mathrm{tr}(\tilde{\boldsymbol{P}}(\xi)\boldsymbol{R}_{\boldsymbol{s}}\tilde{\boldsymbol{P}}^{\mathrm{H}}(\xi))\\ &\quad=\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})-\xi E_{\mathrm{tr}}=0.\end{aligned}$$

Hence, it follows that

$$\xi=\frac{\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})}{E_{\mathrm{tr}}}.$$

Therefore, we have found a closed form solution for the optimization in (36)

$$\boldsymbol{P}_{\mathrm{WF}}=\beta_{\mathrm{TxWF}}\boldsymbol{F}^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\in\mathbb{C}^{N\times B}$$

and

$$\beta_{\mathrm{TxWF}}=\sqrt{\frac{E_{\mathrm{tr}}}{\mathrm{tr}(\boldsymbol{F}^{-2}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{G}\boldsymbol{H})}} \tag{39}$$

where we defined

$$\boldsymbol{F}=\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}\boldsymbol{H}+\frac{\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})}{E_{\mathrm{tr}}}\mathbf{1}_N\in\mathbb{C}^{N\times N}.$$

From (36), we can see that the TxWF is the transmit filter which minimizes the variance of the difference between the desired signal and the output of the receive filter $\boldsymbol{G}$ weighted by $\beta_{\mathrm{TxWF}}^{-1}$. Consequently, we can interpret β_{TxWF} to be the optimum gain

[4]The discussion for the TxCMMSE with equality constraint [51] is similar.

[5]$\mathrm{E}[\|\boldsymbol{s}-\beta^{-1}\tilde{\boldsymbol{s}}\|_2^2]=\mathrm{tr}(\boldsymbol{R}_{\boldsymbol{s}})-2\beta^{-1}\mathrm{Re}(\mathrm{tr}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{s}}))+\beta^{-2}\mathrm{tr}(\boldsymbol{G}\boldsymbol{H}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}+\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})$.

[6]Note that $\mathrm{tr}(\mathrm{Re}(\boldsymbol{G}\boldsymbol{H}\tilde{\boldsymbol{P}}(\xi)\boldsymbol{R}_{\boldsymbol{s}}))=\mathrm{tr}(\boldsymbol{G}\boldsymbol{H}\tilde{\boldsymbol{P}}(\xi)\boldsymbol{R}_{\boldsymbol{s}})$ and $\mathrm{tr}(\boldsymbol{G}\boldsymbol{H}\tilde{\boldsymbol{P}}(\xi)\boldsymbol{R}_{\boldsymbol{s}})=\mathrm{tr}((\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}\boldsymbol{H}+\xi\mathbf{1}_N)\tilde{\boldsymbol{P}}(\xi)\boldsymbol{R}_{\boldsymbol{s}}\tilde{\boldsymbol{P}}^{\mathrm{H}}(\xi))$.

of the filter chain $\boldsymbol{GHP}$ for a given transmit power and noise power. To support this conjecture, we examine following limits:

$$\frac{E_{\mathrm{tr}}}{\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})} \to 0: \quad \beta_{\mathrm{TxWF}} \to \xi\beta_{\mathrm{TxMF}}$$

and

$$\frac{E_{\mathrm{tr}}}{\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})} \to \infty: \quad \beta_{\mathrm{TxWF}} \to \beta_{\mathrm{TxZF}}$$

where we applied the matrix inversion lemma to (39) for the second limit. If the available transmit power is small compared to the noise power at the receiver, the weight β_{TxWF} becomes large. In other words, it is optimum to maximize the portion in the estimate $\tilde{\boldsymbol{s}}$ due to the desired signal $\boldsymbol{s}$, in this case, like the TxMF does. The second limit, where the transmit power is large compared to the noise power, shows that the weighting of the TxZF is optimum for this scenario. The respective limits of the TxWF confirm above discussion

$$\frac{E_{\mathrm{tr}}}{\mathrm{tr}\left(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}}\right)} \to 0: \quad \boldsymbol{P}_{\mathrm{WF}} \to \boldsymbol{P}_{\mathrm{MF}}$$

and

$$\frac{E_{\mathrm{tr}}}{\mathrm{tr}\left(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}}\right)} \to \infty: \quad \boldsymbol{P}_{\mathrm{WF}} \to \boldsymbol{P}_{\mathrm{ZF}}.$$

We see that the TxWF converges to the TxMF and the TxZF for small and large transmit power compared to the noise power, respectively. Remember that we mentioned a similar convergence property of the RxWF in Section III.

Contrary to the TxCMMSE in (35), the TxWF in (39) always uses the whole available transmit power and its structure depends upon the properties of the noise. However, the noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}}$ is only included inside a trace operator. Therefore, the noise covariance matrix $\boldsymbol{R}_{\boldsymbol{\eta}}$ does not directly influence the TxWF, but the scalar value $\mathrm{tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\mathrm{H}})$, which is the noise power at the receive filter output and is a measured for the ability of the receive filter $\boldsymbol{G}$ to deal with the noise $\boldsymbol{\eta}$. This scalar value can be easily determined by the receiver and has to be fed back from the receiver to transmitter, since the transmitter has no chance to measure this quantity. We could include a noise whitener together with its matched filter at the receiver, i.e., $\boldsymbol{G}' = \boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}^{-1}$, enabling the TxWF to adapt to all properties of the noise represented by the covariance matrix, but this approach increases the system complexity dramatically, because not only one scalar but the whole covariance matrix has to be estimated at the receiver and fed back from the receiver to the transmitter.

The structure of the TxWF, but not the weight β_{TxWF}, is independent from the covariance matrix $\boldsymbol{R}_{\boldsymbol{s}}$ of the transmitted symbols $\boldsymbol{s}$, because the receiver applies the *a priori* defined receive filter $\boldsymbol{G}$. As the receiver cannot adapt to the properties of the transmitted symbols, the TxWF is unable to exploit these properties.

C. MSE of TxWF

Due to (14), the TxWF leads to the MSE

$$\varepsilon_{\mathrm{TxWF}} = \mathrm{tr}(\boldsymbol{R}_{\boldsymbol{s}}) - \mathrm{tr}((\boldsymbol{J}_{\mathrm{tx}} + \mathbf{1}_B)^{-1}\boldsymbol{J}_{\mathrm{tx}}\boldsymbol{R}_{\boldsymbol{s}}) \qquad (40)$$

which has the same form as the MSE of the RxWF in (17), but is not the same, since $\boldsymbol{J}_{\mathrm{tx}} \neq \boldsymbol{J}_{\mathrm{rx}}$ in general. We can follow that the MSE of the TxWF is always smaller than the MSEs of the TxMF and the TxZF, since the RxWF minimizes the MSE and the MSE expressions for all transmit filters can be obtained from the MSEs of the respective receive filters by substituting $\boldsymbol{J}_{\mathrm{rx}}$ with $\boldsymbol{J}_{\mathrm{tx}}$. This result justifies to name the transmit filter obtained from (39) as TxWF.

Alternatively, we can write

$$\varepsilon_{\mathrm{TxWF}} = \mathrm{tr}((\boldsymbol{J}_{\mathrm{tx}} + \mathbf{1}_B)^{-1}\boldsymbol{R}_{\boldsymbol{s}})$$

which helps to compute the second of the following limits:

$$\gamma \to 0: \quad \varepsilon_{\mathrm{TxWF}} \to \mathrm{tr}(\boldsymbol{R}_{\boldsymbol{s}}) - \mathrm{tr}(\boldsymbol{J}_{\mathrm{tx}}\boldsymbol{R}_{\boldsymbol{s}}) \qquad (41)$$

and

$$\gamma \to \infty: \quad \varepsilon_{\mathrm{TxWF}} \to \mathrm{tr}(\boldsymbol{J}_{\mathrm{tx}}^{-1}\boldsymbol{R}_{\boldsymbol{s}}) \qquad (42)$$

as the entries of $\boldsymbol{J}_{\mathrm{tx}}$ tend to zero and infinity for $\gamma \to 0$ and $\gamma \to \infty$. According to the discussion in the previous section, the MSE of the TxWF converges to the MSE of the TxMF for low SNR [$\gamma \to 0$, cf. (21)] and to the MSE of the TxZF for high SNR [$\gamma \to \infty$, cf. (32)].

VII. Equivalence of Receive and Transmit Processing for Uncorrelated Symbols and Noise

In the last sections, we have seen that the MSE expressions for the transmit filters can be obtained from the MSE expressions of the respective receive filters, when $\boldsymbol{J}_{\mathrm{rx}}$ is replaced by $\boldsymbol{J}_{\mathrm{tx}}$. We can follow that the MSEs of the transmit filters equal the MSEs of the respective receive filters, if $\boldsymbol{J}_{\mathrm{rx}} = \boldsymbol{J}_{\mathrm{tx}}$. However, this equality can only be fulfilled in some trivial cases, since the structure of $\boldsymbol{J}_{\mathrm{rx}}$ depends on $\boldsymbol{R}_{\boldsymbol{\eta}}$ and $\boldsymbol{R}_{\boldsymbol{s}}$, whereas the structure of $\boldsymbol{J}_{\mathrm{tx}}$ is independent of these entities.

When we restrict ourselves to the case of uncorrelated symbols and noise, that is, $\boldsymbol{R}_{\boldsymbol{s}} = \sigma_{\boldsymbol{s}}^2\mathbf{1}_B$ and $\boldsymbol{R}_{\boldsymbol{\eta}} = \sigma_{\boldsymbol{\eta}}^2\mathbf{1}_M$, the MSEs of the transmit and receive filters are the same for two important scenarios.

1) The transmit filter and the respective receive filter are applied to the same channel $\boldsymbol{H} \in \mathbb{C}^{B\times B}$, where the receive filter for transmit processing and the transmit filter for receive processing are identity mappings, that is, $\boldsymbol{G}_{\mathrm{tx}} = \mathbf{1}_B$ and $\boldsymbol{P}_{\mathrm{rx}} = \mathbf{1}_B$, respectively. Because $\boldsymbol{J}_{\mathrm{rx}} = \gamma\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}$ and $\boldsymbol{J}_{\mathrm{tx}} = \gamma\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}$, we follow that $\varepsilon_{\mathrm{TxMF}} = \varepsilon_{\mathrm{RxMF}}, \varepsilon_{\mathrm{TxZF}} = \varepsilon_{\mathrm{RxZF}}$, and $\varepsilon_{\mathrm{TxWF}} = \varepsilon_{\mathrm{RxWF}}$.
2) Transmit processing is employed for one link (e.g., downlink) and receive processing for the other link (e.g., uplink), where we have to assume full reciprocity of the channel, that means, $\boldsymbol{H}_{\mathrm{tx}} = \boldsymbol{H}_{\mathrm{rx}}^{\mathrm{T}}$, which is fulfilled in TDD systems. Moreover, we set $\boldsymbol{G}_{\mathrm{tx}} = \boldsymbol{P}_{\mathrm{rx}}^{\mathrm{T}}$, i.e., the receiver for transmit processing applies the transpose of the transmit filter for receive processing. Again, we have $\varepsilon_{\mathrm{TxMF}} = \varepsilon_{\mathrm{RxMF}}, \varepsilon_{\mathrm{TxZF}} = \varepsilon_{\mathrm{RxZF}}$, and $\varepsilon_{\mathrm{TxWF}} = \varepsilon_{\mathrm{RxWF}}$, because $\boldsymbol{J}_{\mathrm{rx}} = \boldsymbol{J}_{\mathrm{tx}}^*$.

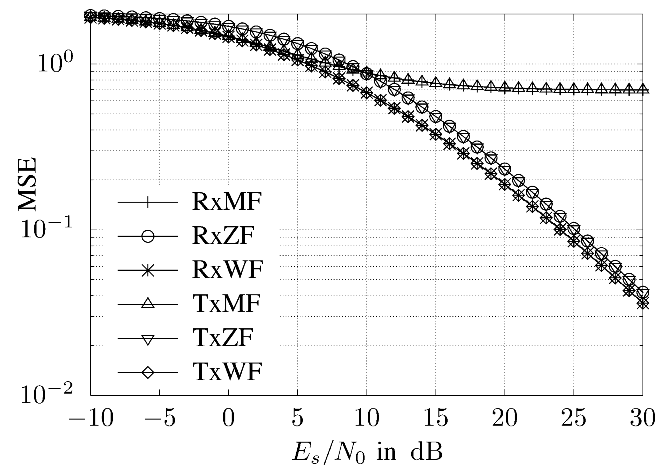


Fig. 2. Receive and transmit filters: MSE ε versus SNR γ for spatially white noise.

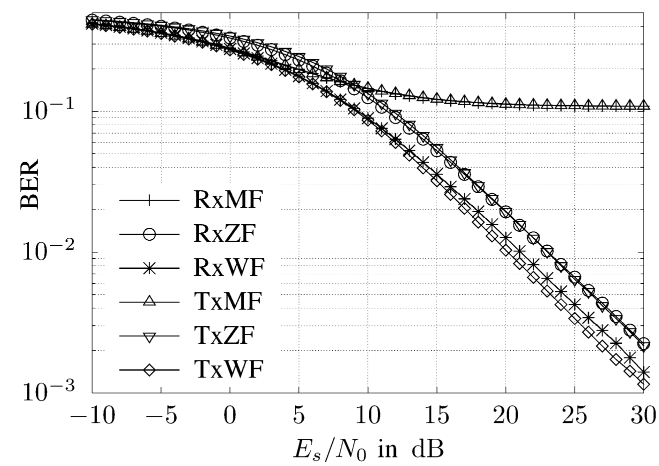


Fig. 3. Receive and transmit filters: BER versus SNR γ for spatially white noise.

VIII. SIMULATION RESULTS

We compare the different receive and transmit filters by applying them to a MIMO system with $M = 2$ antennas at the transmitter and $N = 2$ antennas at the receiver. The transmit filter for the case of receive processing and the receive filter for the case of transmit processing are identity mappings, i.e., $\boldsymbol{P}_{\text{rx}} = \boldsymbol{G}_{\text{tx}} = \mathbf{1}_2$. Per channel realization 100 QPSK symbols for each of the $B = 2$ parallel data streams are transmitted, where we assume uncorrelated data streams and noise, i.e., $\boldsymbol{R_s} = \mathbf{1}_2$ and $\boldsymbol{R_\eta} = \sigma_{\boldsymbol{\eta}}^2\mathbf{1}_2$. We set the transmit power to $E_{\text{tr}} = 2$, that is, unit transmit power is used for one symbol in the average. We assume uncorrelated Rayleigh fading and normalize the channel matrix such that $\mathrm{E}[\|\boldsymbol{H}\|_{\mathrm{F}}^2] = 1$. All results are the mean of 100 000 channel realizations and the transmitter knows the exact instantaneous channel state information.

In Fig. 2, we depict the mean MSEs [(4) averaged over different channel realizations] of the receive and transmit filters. The MSEs of the matched filters saturate for high SNR as shown in Section IV-B, whereas the zero-forcing filters are outperformed by the matched filters for low SNR (see Section V-B). The Wiener filters are always superior compared to the other two filter types as mentioned in Section VI-C. Since we have shown in the previous sections that the MSEs of the transmit filters equal to the MSEs of the respective receive filters for uncorrelated symbols and noise, we are not suprised by the result that this statement is also true in the average. However, the BERs are not the same as can be seen in Fig. 3. This difference between receive and transmit filters is due to the coloring of the noise in the case of receive processing, because the noise is passed through the receive filter. Therefore, the resulting SNRs for the two data streams are different for receive and transmit processing leading to a slight advantage for the receive filters at low SNR and for the transmit filters at high SNR.

The discussed transmit filters are compared in Fig. 4. Besides the results for the already examined TxMF, TxZF, and TxWF, we also show the BER of the TxCMMSE. We observe a strong dependence of the TxCMMSE on the available transmit power. If more transmit power can be used ($E_{\text{tr}} = 20$ instead of $E_{\text{tr}} = 2$), the TxCMMSE saturates at a lower BER for high SNR, but

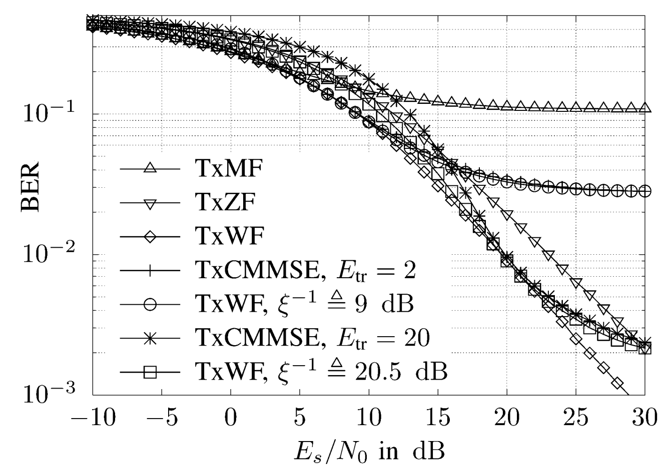


Fig. 4. Transmit filters: BER versus SNR γ for spatially white noise.

is worse for low SNR, since the TxCMMSE does not use the whole transmit power for some channel realizations. We also included the results for the TxWF with constant weighting ξ instead of $\operatorname{tr}(\boldsymbol{G}_{\text{tx}}\boldsymbol{R_\eta}\boldsymbol{G}_{\text{tx}}^{\mathrm{H}})/E_{\text{tr}}$ [see (39)]. The TxWF designed for an SNR of 9 dB is nearly the same as the TxCMMSE with $E_{\text{tr}} = 2$ and the TxWF for an SNR of 20.5 dB is similar to the TxCMMSE with $E_{\text{tr}} = 20$. Thus, the TxCMMSE is equal to or even outperformed by TxWFs designed for constant SNR. Moreover, we see that the weighting $\operatorname{tr}(\boldsymbol{G}_{\text{tx}}\boldsymbol{R_\eta}\boldsymbol{G}_{\text{tx}}^{\mathrm{H}})/E_{\text{tr}}$ of the TxWF obtained with (36) is optimum, as the BERs of the filters with constant weighting only touch the curve of the TxWF for one SNR value (9 and 20.5 dB).

Figs. 5 and 6 depict the results for the receive and transmit filters, when the noise $\boldsymbol{\eta}$ has the covariance matrix

$$\boldsymbol{R_\eta} = \frac{\sigma_{\boldsymbol{\eta}}^2}{11}\begin{bmatrix} 11 & 9 \\ 9 & 11 \end{bmatrix}.$$

Since the transmit filters cannot account for the properties of the noise given by the covariance matrix $\boldsymbol{R_\eta}$, the MSEs of the receive filters are smaller than the MSEs of the transmit filters except for the RxMF which saturates at a higher MSE than the TxMF for high SNR (see Fig. 5). Again, the BER results (cf.

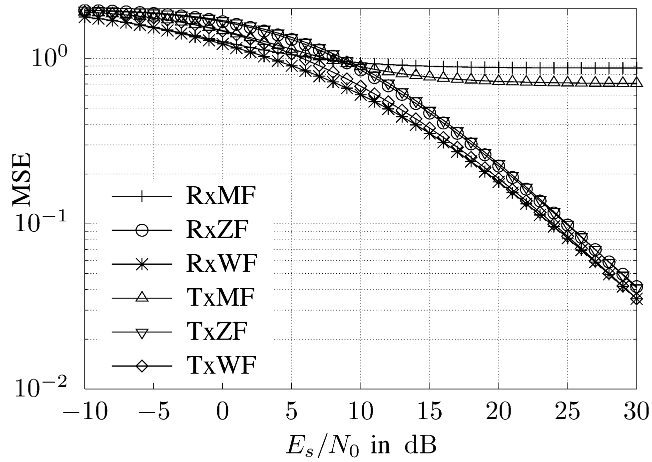


Fig. 5. Receive and transmit filters: MSE ε versus SNR γ for spatially colored noise.

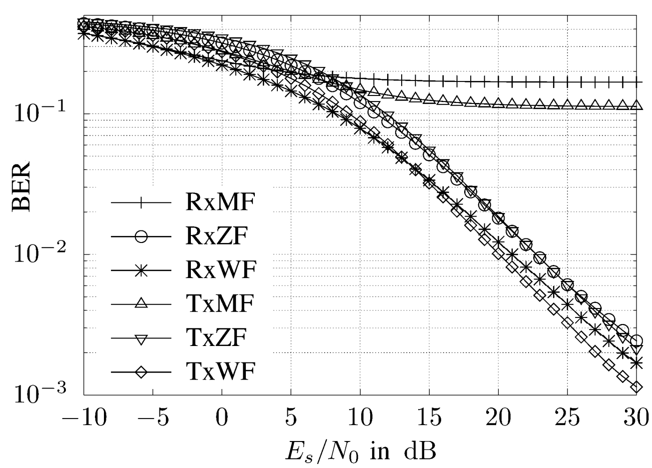


Fig. 6. Receive and transmit filters: BER versus SNR γ for spatially colored noise.

Fig. 6) are different due to the filtering of the noise by the receive filters. We can observe that the receive filters clearly outperform the transmit filters for low SNR, but the transmit filters are slightly superior for high SNR. This can be easily understood, if we use the zero-forcing filters as example. The RxZF and the TxZF lead to interference free estimates, where the amplitude of the two scalar estimates is the same for the noisefree case. As the TxZF does not change the noise, we end up with two data streams with equal SNR. On the other hand, the RxZF filters the noise and leads to different noise powers for the two estimates. Consequently, the SNRs are different in the case of the RxZF leading to a disadvantage compared to the TxZF for high SNR.

IX. Conclusion

We have shown that the different transmit filters can be obtained with the same optimizations as the respective receive filters, where only a transmit power constraint has to be included (see Table I which summarizes the optimizations for the system of Fig. 1). We compared the transmit filters to the respective receive filters in terms of MSE and BER and showed that the

TABLE I
OPTIMIZATIONS FOR RECEIVE AND TRANSMIT FILTERS

RxMF	$\boldsymbol{G}_{\mathrm{MF}} = \underset{\boldsymbol{G}}{\operatorname{argmax}} \left\lvert \mathrm{E}[\boldsymbol{s}^{\mathrm{H}}\tilde{\boldsymbol{s}}]\right\rvert^2 / \mathrm{E}[\lVert \boldsymbol{G}\boldsymbol{\eta}\rVert_2^2]$	
RxZF	$\boldsymbol{G}_{\mathrm{ZF}} = \underset{\boldsymbol{G}}{\operatorname{argmin}} \mathrm{E}\left[\lVert \boldsymbol{s} - \tilde{\boldsymbol{s}}\rVert_2^2\right]$	s.t.: $\tilde{\boldsymbol{s}}\rvert_{\boldsymbol{\eta}=\boldsymbol{0}_M} = \boldsymbol{s}$
RxWF	$\boldsymbol{G}_{\mathrm{WF}} = \underset{\boldsymbol{G}}{\operatorname{argmin}} \mathrm{E}\left[\lVert \boldsymbol{s} - \tilde{\boldsymbol{s}}\rVert_2^2\right]$	
TxMF	$\boldsymbol{P}_{\mathrm{MF}} = \underset{\boldsymbol{P}}{\operatorname{argmax}} \left\lvert \mathrm{E}[\boldsymbol{s}^{\mathrm{H}}\tilde{\boldsymbol{s}}]\right\rvert^2 / \mathrm{E}[\lVert \boldsymbol{G}\boldsymbol{\eta}\rVert_2^2]$	s.t.: $\mathrm{E}\left[\lVert \boldsymbol{P}\boldsymbol{s}\rVert_2^2\right] = E_{\mathrm{tr}}$
TxZF	$\{\boldsymbol{P}_{\mathrm{ZF}}, \beta_{\mathrm{TxZF}}\} = \underset{\{\boldsymbol{P},\beta\}}{\operatorname{argmin}} \mathrm{E}\left[\lVert \boldsymbol{s} - \beta^{-1}\tilde{\boldsymbol{s}}\rVert_2^2\right]$	s.t.: $\tilde{\boldsymbol{s}}\rvert_{\boldsymbol{\eta}=\boldsymbol{0}_M} = \beta\boldsymbol{s}$ and $\mathrm{E}\left[\lVert \boldsymbol{P}\boldsymbol{s}\rVert_2^2\right] = E_{\mathrm{tr}}$
TxWF	$\{\boldsymbol{P}_{\mathrm{WF}}, \beta_{\mathrm{TxWF}}\} = \underset{\{\boldsymbol{P},\beta\}}{\operatorname{argmin}} \mathrm{E}\left[\lVert \boldsymbol{s} - \beta^{-1}\tilde{\boldsymbol{s}}\rVert_2^2\right]$	s.t.: $\mathrm{E}\left[\lVert \boldsymbol{P}\boldsymbol{s}\rVert_2^2\right] = E_{\mathrm{tr}}$

MSEs are the same, if the transmitted data and the noise are uncorrelated. The BER simulations revealed that the receive filters outperform the transmit filters for low SNR especially for colored noise, but the transmit filters show better results for high SNR.

Appendix A
TxMF for High SNR

With the *eigenvalue decomposition* (EVD, e.g., [80]) of

$$\boldsymbol{J}_{\mathrm{tx}} = \boldsymbol{V}\boldsymbol{\Psi}\boldsymbol{V}^{\mathrm{H}} = \sum_{a=1}^{B} \psi_a \boldsymbol{v}_a \boldsymbol{v}_a^{\mathrm{H}} \tag{43}$$

where $\psi_a \in \mathbb{R}_+$ denotes the ath eigenvalue and $\boldsymbol{v}_a \in \mathbb{C}^B$ the respective eigenvector, $\boldsymbol{\Psi} = \mathrm{diag}(\psi_1, \ldots, \psi_B)$, and the modal matrix $\boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_B] \in \mathbb{C}^{B\times B}$, the MSE of the TxMF for high SNR reads as [cf. (22)]

$$\varepsilon_{\mathrm{TxMF}} = \mathrm{tr}\left(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V}\right) - \frac{\mathrm{tr}^2\left(\boldsymbol{\Psi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V}\right)}{\mathrm{tr}\left(\boldsymbol{\Psi}^2\boldsymbol{V}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V}\right)}.$$

Here, we used $\boldsymbol{V}\boldsymbol{V}^{\mathrm{H}} = \mathbf{1}_B$. From the definition of the MSE in (4), we can follow that the MSE is always larger than or equal to zero, i.e., $\varepsilon_{\mathrm{TxMF}} \geq 0$. Our aim is to find a condition, when the MSE is zero, i.e., when the equality holds. To this end, we rewrite above inequality

$$\sum_{a=1}^{B}\sum_{b=1}^{B} \psi_b^2 \rho_a \rho_b - \left(\sum_{b=1}^{B} \psi_b \rho_b\right)^2 \geq 0$$

and

$$\sum_{a=1}^{B}\sum_{b=1}^{B} \left(\psi_b^2 - \psi_a\psi_b\right) \rho_a \rho_b \geq 0$$

where we introduced $\rho_b = \boldsymbol{v}_b^{\mathrm{T}} \boldsymbol{R}_{\boldsymbol{s}} \boldsymbol{v}_b^* \in \mathbb{R}_+$ and note that $\mathrm{tr}(\boldsymbol{\Psi}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V}) = \sum_{b=1}^{B} \psi_b \rho_b$, because $\boldsymbol{\Psi}$ is a diagonal matrix. Since the summand vanishes for $a = b$, we get

$$\sum_{a=1}^{B}\sum_{b=a}^{B} \left(\psi_b^2 - \psi_a\psi_b + \psi_a^2 - \psi_b\psi_a\right) \rho_a \rho_b \geq 0$$

and

$$\sum_{a=1}^{B}\sum_{b=a}^{B} (\psi_a - \psi_b)^2 \rho_a \rho_b \geq 0.$$

Obviously, the MSE is only zero, if $\psi_b = \psi, b = 1, \ldots, B$. In other words, the TxMF exhibits a residual error, when at least one eigenvalue value of $\boldsymbol{J}_{\rm tx}$ is different from the others.

Appendix B
Comparison of the Gains of TxMF and TxZF

The gains of the TxMF and the TxZF [cf. (29) and (26)] can be expressed as

$$\beta'_{\rm TxMF} = \frac{1}{B}\sqrt{\zeta {\rm tr}(\boldsymbol{J}_{\rm tx})}$$

and

$$\beta_{\rm TxZF} = \sqrt{\zeta {\rm tr}^{-1}\left(\boldsymbol{J}_{\rm tx}^{-1}\right)}$$

respectively, where we defined $\zeta = {\rm tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\rm H})/\sigma_{\boldsymbol{s}}^2$ and incorporated the EVD of $\boldsymbol{J}_{\rm tx}$ [see (43)]. We claim that $\beta'_{\rm TxMF} \geq \beta_{\rm TxZF}$ and with the EVD of $\boldsymbol{J}_{\rm tx}$, we obtain

$$\frac{1}{B^2}\sum_{a=1}^{B}\psi_a \geq \left(\sum_{a=1}^{B}\psi_a^{-1}\right)^{-1}$$
$$\sum_{a=1}^{B}\sum_{b=1}^{B}\psi_a\psi_b^{-1} \geq \sum_{a=1}^{B}\sum_{b=1}^{B}1$$

and

$$\sum_{a=1}^{B}\sum_{b=1}^{B}\frac{\psi_a-\psi_b}{\psi_b} \geq 0.$$

Because the summand is zero for $a = b$, it follows that

$$\sum_{a=1}^{B}\sum_{b=a}^{B}\frac{\psi_a-\psi_b}{\psi_b} + \frac{\psi_b-\psi_a}{\psi_a} \geq 0$$

and

$$\sum_{a=1}^{B}\sum_{b=a}^{B}\frac{(\psi_a-\psi_b)^2}{\psi_a\psi_b} \geq 0.$$

Therefore, we have proven that the gain of the TxZF is smaller than or equal to the gain of the TxMF for uncorrelated symbols.

Appendix C
Comparison of TxMF and TxZF for Low SNR

The MSEs of the TxMF and the TxZF for low SNR ($\gamma \rightarrow 0$, cf. (21) and (31)) expressed with the EVD of the matrix $\boldsymbol{J}_{\rm tx}$ [see (43)] can be written as

$$\varepsilon_{\rm TxMF} = {\rm tr}(\boldsymbol{V}^{\rm H}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V}) - {\rm tr}\left(\boldsymbol{\Psi}\boldsymbol{V}^{\rm H}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V}\right)$$

and

$$\varepsilon_{\rm TxZF} = {\rm tr}(\boldsymbol{V}^{\rm H}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V}) - \frac{{\rm tr}^2(\boldsymbol{V}^{\rm H}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V})}{{\rm tr}(\boldsymbol{\Psi}^{-1}\boldsymbol{V}^{\rm H}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{V})}$$

respectively. We will show that the MSE of the TxZF for low SNR is larger than or equal to the respective MSE of the TxMF. Thus, we have to prove that

$$\sum_{a=1}^{B}\sum_{b=1}^{B}\psi_a\psi_b^{-1}\rho_a\rho_b \geq \sum_{a=1}^{B}\sum_{b=1}^{B}\rho_a\rho_b$$
$$\sum_{a=1}^{B}\sum_{b=1}^{B}\frac{\psi_a-\psi_b}{\psi_b}\rho_a\rho_b \geq 0$$
$$\sum_{a=1}^{B}\sum_{b=a}^{B}\left(\frac{\psi_a-\psi_b}{\psi_b} + \frac{\psi_b-\psi_a}{\psi_a}\right)\rho_a\rho_b \geq 0$$

and

$$\sum_{a=1}^{B}\sum_{b=a}^{B}\frac{(\psi_a-\psi_b)^2}{\psi_a\psi_b}\rho_a\rho_b \geq 0.$$

Since the last inequality is always true, we have proven that the TxMF has a lower MSE than the TxZF for low SNR. Additionally, we can see that the TxMF and the TxZF exhibit the same MSE for low SNR, if all eigenvalues of $\boldsymbol{J}_{\rm tx}$ have the same value, i.e., $\psi_b = \psi, b = 1, \ldots, B$.

Appendix D
Optimization for the TxZF

With the understanding of the TxWF (cf. Section VI) we can explain the optimization for the TxZF in (27). Since the TxWF minimizes the error ${\rm E}[\|\boldsymbol{s} - \beta^{-1}\tilde{\boldsymbol{s}}\|_2^2]$ instead of the MSE ${\rm E}[\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2^2]$ as the RxWF does, we also have to include the weighting β^{-1} in the objective function of the RxZF optimization [see (9)] and have to include a transmit power constraint to get from the RxZF to the TxZF optimization[7]

$$\{\boldsymbol{P}_{\rm ZF}, \beta_{\rm TxZF}\} = {\rm argmin}_{\{\boldsymbol{P},\beta\}}{\rm E}\left[\|\boldsymbol{s} - \beta^{-1}\tilde{\boldsymbol{s}}\|_2^2\right]$$
$$\text{s.t.: } \boldsymbol{G}\boldsymbol{H}\boldsymbol{P} = \beta\boldsymbol{1}_B \quad \text{and} \quad {\rm E}\left[\|\boldsymbol{P}\boldsymbol{s}\|_2^2\right] = E_{\rm tr}. \quad (44)$$

Note that the cost function of above optimization reduces to

$${\rm E}\left[\|\boldsymbol{s} - \beta^{-1}\tilde{\boldsymbol{s}}\|_2^2\right] = \beta^{-2}{\rm tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\rm H})$$

due to the first constraint. Consequently, the optimizations in (44) and (27) are equal, since ${\rm tr}(\boldsymbol{G}\boldsymbol{R}_{\boldsymbol{\eta}}\boldsymbol{G}^{\rm H})$ is constant.

References

[1] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[2] R. Price and P. E. Green, "A communication technique for multipath channels," *Proc. IRE*, vol. 46, no. 3, pp. 555–570, Mar. 1958.
[3] D. A. Shnidman, "A generalized nyquist criterion and an optimum linear receiver for a pulse modulation system," *Bell Syst. Tech. J.*, vol. 46, no. 9, pp. 2163–2177, Nov. 1967.
[4] D. A. George, "Matched filters for interfering signals," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, pp. 153–154, Jan. 1965.
[5] R. Esmailzadeh, M. Nakagawa, and E. A. Sourour, "Time-division duplex CDMA communications," *IEEE Personal Commun.*, vol. 4, no. 2, pp. 51–56, Apr. 1997.
[6] M. Haardt, A. Klein, R. Koehn, S. Oestreich, M. Purat, V. Sommer, and T. Ulrich, "The TD–CDMA based UTRA TDD mode," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 8, pp. 1375–1385, Aug. 2000.
[7] G. V. Tsoulos and M. A. Beach, "Calibration and linearity issues for an adaptive antenna system," in *Proc. VTC*, vol. 3, May 1997, pp. 1597–1600.

[7]An alternative formulation for the TxZF optimization can be obtained by replacing the minimization of the error ${\rm E}[\|\boldsymbol{s} - \beta^{-1}\tilde{\boldsymbol{s}}\|_2^2]$ with the maximization of the SNR performed by the TxMF [see (18)].

[8] W. Keusgen, C. M. Walke, and B. Rembold, "A system model considering the influence of front-end imperfections on the reciprocity of up- and downlink system impulse responses," in *Proc. ASST*, Sep. 2001, pp. 243–248.

[9] K. Nishimori, K. Cho, Y. Takatori, and T. Hori, "Automatic calibration method using transmitting signals of an adaptive array for TDD systems," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1636–1640, Nov. 2001.

[10] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[11] L. Bigler, H. P. Lin, S. S. Jeng, and G. Xu, "Experimental direction of arrival and spatial signature measurements at 900 MHz for smart antenna systems," in *Proc. VTC*, vol. 1, Jul. 1995, pp. 55–58.

[12] T. Asté, P. Forster, L. Féty, and S. Mayrargue, "Downlink beamforming avoiding DOA estimation for cellular mobile communications," in *Proc. ICASSP*, vol. VI, May 1998, pp. 3313–3316.

[13] W. Utschick and J. A. Nossek, "Downlink beamforming for FDD mobile radio systems based on spatial covariances," in *Proc. Eur. Wireless*, Oct. 1999, pp. 65–67.

[14] G. Montalbano, I. Ghauri, and D. T. M. Slock, "Spatio-temporal array processing for FDD/CDMA/SDMA downlink transmission," in *Proc. VTC*, vol. 3, Sep. 1999, pp. 1910–1914.

[15] P. Forster, L. Féty, and M. Le Bot, "Spatio-temporal filters for downlink processing in FDD systems," in *Proc. ICASSP*, vol. V, Jun. 2000, pp. 2585–2588.

[16] M. Joham, K. Kusume, W. Utschick, and J. A. Nossek, "Transmit matched filter and transmit wiener filter for the downlink of FDD DS–CDMA systems," in *Proc. PIMRC*, vol. 5, Sep. 2002, pp. 2312–2316.

[17] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electron. Lett.*, vol. 7, no. 5/6, pp. 138–139, Mar. 1971.

[18] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, no. 4, pp. 774–780, Aug. 1972.

[19] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Space-time Transmission using Tomlinson-Harashima precoding," in *Proc. 4th ITG Conf. Source and Channel Coding*, Jan. 2002, pp. 139–147.

[20] ——, "MIMO precoding for decentralized receivers," in *Proc. ISIT*, Jun./Jul. 2002, p. 496.

[21] G. Ginis and J. M. Cioffi, "A multi-user precoding scheme achieving crosstalk cancellation with application to DSL systems," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, vol. 2, Oct. 2000, pp. 1627–1631.

[22] J. Liu and A. Duel-Hallen, "Tomlinson-Harashima transmitter precoding for synchronous multiuser communications," presented at the CISS, Mar. 2003.

[23] M. Joham, J. Brehmer, and W. Utschick, "MMSE approaches to multiuser spatio-temporal Tomlinson-Harashima precoding," in *Proc. ITG SCC*, Jan. 2004, pp. 387–394.

[24] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Linear and nonlinear preequalization/equalization for MIMO systems with long-term channel state information at the transmitter," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 373–378, Mar. 2004.

[25] J. P. Costas, "Coding with linear systems," in *Proc. IRE*, vol. 40, Sep. 1952, pp. 1101–1103.

[26] K. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. COM-24, no. 12, pp. 1283–1290, Dec. 1976.

[27] J. Salz, "Digital transmission over cross-coupled linear channels," *AT&T Tech. J.*, vol. 64, no. 6, pp. 1147–1159, Jul.–Aug. 1985.

[28] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[29] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1988–2005, Jul. 1999.

[30] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-tune linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[31] G. Caire and S. S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[32] P. S. Henry and B. S. Glance, "A new approach to high-capacity digital mobile radio," *Bell Syst. Tech. J.*, vol. 60, no. 8, pp. 1891–1904, Oct. 1981.

[33] R. Esmailzadeh and M. Nakagawa, "Pre-RAKE diversity combination for direct sequence spread spectrum mobile communications systems," *IEICE Trans. Commun.*, vol. E76-B, no. 8, pp. 1008–1015, Aug. 1993.

[34] R. Esmailzadeh, E. Sourour, and M. Nakagawa, "Prerake diversity combining in time-division duplex CDMA mobile communications," *IEEE Trans. Veh. Technol.*, vol. 48, no. 3, pp. 795–801, May 1999.

[35] T. A. Kadous, E. E. Sourour, and S. E. El-Khamy, "Comparison between various diversity techniques of the pre-Rake combining system in TDD/CDMA," in *Proc. VTC*, vol. 3, May 1997, pp. 2210–2214.

[36] I. Jeong and M. Nakagawa, "A novel transmission diversity system in TDD–CDMA," *IEICE Trans. Commun.*, vol. E81-B, no. 7, pp. 1409–1416, Jul. 1998.

[37] A. Noll Barreto and G. Fettweis, "On the downlink capacity of TDD CDMA systems using a pre-rake," in *Proc. Globecom*, vol. 1a, Dec. 1999, pp. 117–121.

[38] S. Zazo, F. Bader, and J. M. Páez-Borrallo, "Pre-rake and multiuser detection techniques in UTRA-TDD systems," in *Proc. VTC*, vol. 3, May 2000, pp. 2069–2072.

[39] R. L. Choi, K. B. Letaief, and R. D. Murch, "CDMA pre-rake diversity system with base station transmit diversity," in *Proc. VTC* , vol. 2, May 2000, pp. 951–955.

[40] ——, "MISO CDMA transmission with simplified receiver for wireless communication handsets," *IEEE Trans. Commun.*, vol. 49, no. 5, pp. 888–898, May 2001.

[41] S. Guncavdi and A. Duel-Hallen, "A novel space-time pre-rake transmitter diversity method for W–CDMA using long range prediction," in *Proc. CISS*, vol. 1, Mar. 2001, pp. 32–37.

[42] X. Revés, A. Gelonch, and F. Casadevall, "FPGA resource allocation of a DS–CDMA indoor system," presented at the 1st Mobile Communications Summit, Sep. 2001.

[43] M. Joham, W. Utschick, and J. A. Nossek, "On the equivalence of prerake and transmit matched filter," in *Proc. ASST*, Sep. 2001, pp. 313–318.

[44] J. Wang, M. Zhao, S. Zhou, and Y. Yao, "A novel multipath transmission diversity' scheme in TDD–CDMA systems," *IEICE Trans. Commun.*, vol. E82-B, no. 10, pp. 1706–1709, Oct. 1999.

[45] R. Irmer, A. Noll Barreto, and G. Fettweis, "Transmitter precoding for spread-spectrum signals in frequency-selective fading channels," in *Proc. 3G Wireless*, May 2001, pp. 939–944.

[46] J.-K. Han, M.-W. Lee, and H.-K. Park, "Principal ratio combining for pre/post-rake diversity," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 234–236, Jun. 2002.

[47] A. Noll Barreto and G. Fettweis, "Performance improvement in DS-spread spectrum CDMA systems using a pre- and a post-rake," in *Proc. Int. Zurich Seminar Broadband Communications*, Feb. 2000, pp. 39–46.

[48] Z. Tang and S. Cheng, "Interference cancellation for DS–CDMA systems over flat fading channels through pre-decorrelating," in *Proc. PIMRC*, vol. II, Sep. 1994, pp. 435–438.

[49] H. Liu and G. Xu, "Multiuser blind channel estimation and spatial channel pre-equalization," in *Proc. ICASSP*, vol. 3, May 1995, pp. 1756–1759.

[50] V. Weerackody, "Proceding of signature sequences for CDMA systems," U.S. Patent 5 461 610, Oct. 1995.

[51] B. R. Vojčić and W. M. Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1346–1355, Oct. 1998.

[52] M. Brandt-Pearce and A. Dharap, "Transmitter-based multiuser interference rejection for the down-link of a wireless CDMA system in a multipath environment," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 407–417, Mar. 2000.

[53] G. Montalbano, I. Ghauri, and D. T. M. Slock, "Spatio-temporal array processing for CDMA/SDMA downlink transmission," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, vol. 2, Nov. 1998, pp. 1337–1341.

[54] H. R. Karimi, M. Sandell, and J. Salz, "Comparison between transmitter and receiver array processing to achieve interference nulling and diversity," in *Proc. PIMRC*, vol. 3, Sep. 1999, pp. 997–1001.

[55] H. Sampath, H. Bölcskei, and A. J. Paulraj, "Pre-equalization for MIMO wireless channels with delay spread," in *Proc. VTC*, vol. 3, Sep. 2000, pp. 1175–1178.

[56] P. W. Baier, M. 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*, vol. 1, Sep. 2000, pp. 1–5.

[57] M. Joham and W. Utschick, "Downlink processing for mitigation of intracell interference in DS–CDMA systems," in *Proc. ISSSTA*, vol. 1, Sep. 2000, pp. 15–19.

[58] A. Noll Barreto and G. Fettweis, "Capacity increase in the downlink of spread spectrum systems through joint signal precoding," in *Proc. ICC*, vol. 4, Jun. 2001, pp. 1142–1146.

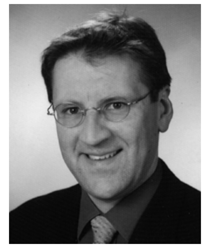
[59] F. Kowalewski and P. Mangold, "Joint predistortion and transmit diversity," in *Proc. Globecom*, vol. 1, Nov. 2000, pp. 245–249.

[60] (1999, Jul.) Tx Diversity with joint predistortion. Bosch, 3GPP. [Online]. Available: http://www.3gpp.org/ftp/tsg-ranIWGl-RLl/TSGRl-06/Docs/Pdfs/rl-99 918.PDF, Tech. Rep. 3GPP TSGR1#6(99)918
[61] C. M. Walke and B. Rembold, "Joint detection and joint pre-distortion techniques for SD/TD/CDMA systems," *Frequenz*, vol. 55, no. 7–8, pp. 204–213, Jul.–Aug. 2001.
[62] A. Morgado, P. Pinho, A. Gameiro, and J. Fernandes, "Pre-equalization technique for interference cancellation in the UMTS-TDD downlink channel," in *Proc. VTC Fall*, vol. 2, Oct. 2001, pp. 878–881.
[63] S. Georgoulis and D. G. M. Cruickshank, "Pre-equalization, transmitter precoding and joint transmission techniques for time division duplex CDMA," in *Proc. Conf. 3G Mobile Communication Technologies*, Mar. 2001, pp. 257–261.
[64] D. Reynolds, A. Høst-Madsen, and X. Wang, "Adaptive transmitter precoding for time division duplex CDMA in fading multipath channels: Strategy and analysis," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 12, pp. 1325–1334, Dec. 2002.
[65] S. Guncavdi and A. Duel-Hallen, "Pre-RAKE multiuser transmitter precoding for DS/CDMA systems," presented at the CISS, Mar. 2003.
[66] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 123–136, Jan. 1989.
[67] M. Meurer, H. Tröger, T. Weber, and P. W. Baier, "Synthesis of joint detection arid joint transmission in CDMA downlinks," *Electron. Lett.*, vol. 37, no. 14, pp. 919–920, Jul. 2001.
[68] S. L. Georgoulis and D. G. M. Cruickshank, "Transmitter based inverse filters for MAI and ISI mitigation in a TDD/CDMA downlink," in *Proc. VTC*, vol. 2, Sep. 2002, pp. 681–685.
[69] M. Joham, "On the constrained minimum mean square error transmit filter and its relation to the transmit Wiener filter," Munich Univ. Technol., Munich, Germany, Tech. Rep. TUM-LNS-TR-03-01, Mar. 2003.
[70] F. P. Wathan, R. Irmer, and G. Fettweis, "On transmitter-based interference mitigation in TDD-downlink with frequency selective fading channel environment," in *Proc. Asia-Pacific Conf. Communications*, Sep. 2002, pp. 125–128.
[71] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter," Munich Univ. Technol., Munich, Germany, Tech. Rep. TUM-LNS-TR-02-1, Jan. 2002.
[72] ——, "Transmit wiener filter for the downlink of TDD DS–CDMA systems," in *Proc. ISSSTA*, vol. 1, Sep. 2002, pp. 9–13.
[73] 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*, vol. 1, Nov. 2002, pp. 429–433.
[74] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, no. 9, pp. 772–781, Sep. 1978.
[75] S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, NJ: Prentice-Hall, 2002.
[76] M. Joham, "On the Equivalence of two optimizations for the receive matched filter," Munich Univ. Technol., Munich, Germany, Tech. Rep. TUM-LNS-TR-02-04, Jul. 2002. In German.
[77] R. Fletcher, *Practical Methods of Optimization*. New York: Wiley, 1987.
[78] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Cambridge, MA: MIT Press, 1949.
[79] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.
[80] G. Golub and C. V. Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

Michael Joham (S'99–M'05) was born in Kufstein, Austria, in 1974. He received the Dipl.-Ing. and Dr.-Ing. degrees (both *summa cum laude*) in electrical engineering from the Munich University of Technology (TUM), Munich, Germany, in 1999 and 2004, respectively.

He has been with the Institute of Circuit Theory and Signal Processing, TUM, since 1999, where he is currently a Senior Researcher. In the summers of 1998 and 2000, he visited the Purdue University, West Lafayette, IN. His main research interests are estimation theory, reduced-rank processing, and precoding in mobile communications.

Dr. Joham received the *VDE Preis* for his diploma thesis in 1999 and the *Texas-Instruments-Preis* for his dissertation in 2004.



Wolfgang Utschick (M'97) was born in Ingolstadt, Germany, in 1964. He completed several industrial education programs before receiving the Dipl.-Ing. and Dr.-Ing. degrees (both with honors) in electrical engineering from the Munich University of Technology (TUM), Munich, Germany, in 1993 and 1998, respectively. During this period, he held a scholarship from the Bavarian Ministry of Education for exceptional students and a scholarship from Siemens AG.

In 1993, he became a part-time lecturer at a technical school for industrial education. From 1998 to 2002, he was head of the Signal Processing Group at the Institute of Circuit Theory and Signal Processing, TUM. He has been a guest researcher at the ETH Zurich, Zurich, Switzerland, in the summer of 2000, and from 2000 to 2001, he was instrumental in the 3rd Generation Partnership Project as an academic consultant in the field of multielement antenna wireless communication systems. Since October 2002, he has been Professor of signal processing with the Department of Electrical Engineering and Information Technology, TUM. He completed the biennial PROLEHRE program for professional academic teaching. He gives national and international undergraduate and graduate courses in the field of signal processing, communication systems, and smart antennas and conducts research on array signal processing, multielement antenna communication systems, x-layer optimization, neural networks, etc. He has authored and co-authored many papers on the field of neural networks and signal processing with regard to communication applications and holds several patents in the field of wireless communications.

Josef A. Nossek (S'72–M'74–SM'81–F'93) was born on December 17, 1947, in Vienna, Austria. He received the Dipl.-Ing. and the Dr. techn. degrees in electrical engineering from the University of Technology, Vienna, in 1974 and 1980, respectively.

In 1974, he joined Siemens AG, Munich, Germany, as a member of the Technical Staff, where he worked on the design of filters for communication systems. In 1978, he became Supervisor of a group working on discrete-time circuits (switched-capacitor and CCD filters), and, from 1980 on, he was Head of the department responsible for electromechanical, microwave, and digital filter design activities. In 1982, he moved into digital microwave radio design, where he was first responsible for high data-rate modems employing multilevel modulation techniques. In 1987, he was promoted to Head of all Radio Systems Design. Since 1989, he has been a Full Professor for Circuit Theory and Signal Processing at the Munich University of Technology (TUM), where he teaches undergraduate and graduate courses on circuits and systems theory and signal processing and leads research on signal processing algorithms for communications systems, theory of linear systems, and VLSI architectures. From 1999 to 2002, he was Dean of the Department of Electrical Engineering and Information Technology, TUM. He was a Guest Professor at the University of Cape Town, Cape Town, South Africa, in 1984, the University of California, Berkeley, in 1992 and 1998, and the University of Technology, Vienna, in 1995.

Prof. Nossek served as Guest Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS in 1993, Associate Editor from 1991 to 1993, and Editor-in-Chief from 1995 to 1997. He is on the Editorial Board of a number of scientific and technical journals. He was program Co-Chairman of the IEEE International Conference on Acoustics, Speech, and Signal Processing in Munich in 1997. He was a member of the Board of Governors of the IEEE Circuits and Systems Society from 1998 to 2000. His awards include the ITG Best Paper Award 1988, the Mannesmann Mobilfunk Innovations award 1998, and the Golden Jubilee Medal of the IEEE Circuits and Systems Society for "Outstanding Contributions to the Society." He was President Elect of the IEEE Circuits and Systems Society in 2001, President in 2002, and Past President in 2003.