A Krylov Subspace Multi-Stage Decomposition of the Transmit Wiener Filter

Guido Dietl, Michael Joham, Philipp Kreuter, Johannes Brehmer, and Wolfgang Utschick

8th ITG Workshop on Smart Antennas (WSA 2004)
Munich, Germany
pp. 247–252, March 18th–19th, 2004

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

A Krylov Subspace Multi-Stage Decomposition of the Transmit Wiener Filter

Guido Dietl, Michael Joham, Philipp Kreuter, Johannes Brehmer, Wolfgang Utschick
Institute for Circuit Theory and Signal Processing,
Munich University of Technology, 80333 Munich, Germany
Email: gudi@nws.ei.tum.de

***Abstract*— Besides methods based on eigensubspaces, the reduced-rank *Multi-Stage Wiener Filter* (MSWF) is a well-known approach for the approximation of the *Wiener Filter*, the optimum linear receive filter in the *minimum mean square error* sense, in a lower-dimensional subspace in order to reduce computational complexity and enhance performance in case of low sample support. Analogous, the *Transmit Wiener Filter* (TxWF) is the optimum linear filter at the transmitter side where the receiver is kept simple since it applies only a scalar weighting.**

In this paper, we use the principles of the MSWF to derive a multi-stage decomposition of the TxWF which we denote *Transmit Multi-Stage Wiener Filter* (TxMSWF). In addition, we will show that the reduced-rank TxMSWF can be seen as an approximation of the TxWF in a Krylov subspace. Simulation results reveal that the TxMSWF achieves near optimum performance for relatively low rank. Thus, it is an interesting alternative to low complexity approximations of the TxWF in eigensubspaces.

I. Introduction

The *Wiener Filter* (WF) [1], [2] performs optimal linear receive processing in the *Minimum Mean Square Error* (MMSE) sense, based only on second order statistics, in order to estimate an unknown signal from an observation vector which is correlated to the desired data signal. Since the computation of the filter coefficients ends up in solving the *Wiener-Hopf* equation, the computational complexity increases with the third order of the dimension of the observation vector. Reduced-rank receive processing is a well-known approach to either reduce computational complexity and/or enhance performance in case of estimation errors due to estimated statistics based on a low number of samples. The first reduced-rank approaches, viz. the *Principal Component* (PC) analysis [3] and the *Cross-Spectral* (CS) method [4], were based on the approximation of the WF in eigensubspaces. Recently, Goldstein et al. introduced the *Multi-Stage WF* (MSWF) in [5] which is an approximation of the WF in the *Krylov* subspace [6], [7], [8] of the autocorrelation matrix of the observation and the cross-correlation vector between the observation and the data signal. Thus, the MSWF can be implemented using algorithms based on Krylov subspaces like the *Lanczos* [9], [8] or the *Conjugate Gradient* method [10], [11], [12]. Simulation results of numerous systems, e. g. an *Enhanced Data rates GSM Evolution* (EDGE) system [8], a *Global Positioning System* (GPS) [13], and a *cdma2000* system [14], have shown that the Krylov subspace based MSWF outperforms the eigensubspace based methods if the computational complexity of both has the same order.

Reduced-rank receive processing is especially advantagous in the downlink where the *Mobile Station* (MS) has to be simple for a low power consumption. The computational complexity of the MS can be further reduced by moving the signal processing except for a scalar weighting from the receiver to the transmitter at the *Base Station* (BS) so that the channel acts as an equalizer for the predistorted transmit signal. Such a method requires normally full *Channel State Information* (CSI) at the transmitter which restricts the application to *Time Division Duplex* (TDD) systems where the channel is slowly varying. Due to the reciprocity of the channel in TDD systems, the channel parameters estimated at the BS in the uplink can be used for transmit processing in the downlink.

One popular transmit strategy is the *Transmit WF* (TxWF) which minimizes the *Mean Square Error* (MSE) between the data signal and the received signal after the scalar weighting (see [15], [16]). The TxWF outperforms not only the *Transmit Matched Filter* (TxMF) [17], maximizing the cross-correlation between the desired data signal and the received signal, but also the *Transmit Zero-Forcing Filter* (TxZF) [18]) which is optimized to completely suppress interference. Again, the calculation of the TxWF is computational intensive and may be an obstacle for the implementation even at the BS.

Motivated by the excellent properties of the reduced-rank MSWF, we propose the *Transmit MSWF* (TxMSWF) which is based on a similar decomposition as the MSWF at the receiver. In accordance with the derivations of the MSWF, we show that the reduced-rank TxMSWF lies in a Krylov subspace which leads to an alternative implementation based on the Lanczos algorithm. Finally, we present simulation results of the reduced-rank TxMSWF as a precoder for a time-dispersive communication channel and compare it to an eigenvector based method motivated by the PC analysis known from receive processing.

In the next section, we introduce the system model and review briefly the computation of the TxWF. After the derivation of the TxMSWF and its Lanczos implementation in Section III, we apply the reduced-rank approaches to a frequency-selective communication system in Section IV.

Throughout the paper, vectors and matrices are denoted by lower case bold and capital bold letters, respectively. The matrix $\mathbf{1}_n$ is the $n \times n$ identity matrix, $\boldsymbol{e}_\nu$ its ν-th column, $\mathbf{0}_{m\times n}$ the $m \times n$ zero matrix, and $\mathbf{0}_n$ the n-dimensional zero vector. The operation '$*$' denotes discrete convolution, '$\otimes$' the

Kronecker product, $\mathrm{E}\{\cdot\}$ expectation, $(\cdot)^*$ conjugate complex, $(\cdot)^{\mathrm{T}}$ transpose, $(\cdot)^{\mathrm{H}}$ Hermitian, i. e. conjugate transpose, $\|\cdot\|_2$ the Euclidian norm, and $\mathrm{O}(\cdot)$ the Landau symbol. We use $\mathrm{span}\{\boldsymbol{A}\}$ as the span of the matrix $\boldsymbol{A}$ and $\mathrm{null}\{\boldsymbol{A}\}$ to denote its null-space. All random processes are assumed to be zero-mean and stationary. The variance of the scalar process $\mathsf{x}[k]$ is denoted by $\sigma_{\mathsf{x}}^2 = \mathrm{E}\left\{|\mathsf{x}[k]|^2\right\}$.

II. System Model and Transmit Wiener Filter

We consider the communication system depicted in Figure 1. The vector *Finite Impulse Response* (FIR) transmit filter

$$\boldsymbol{p}[k] = \sum_{\ell=0}^{L-1} \boldsymbol{p}_\ell \delta[k-\ell] \in \mathbb{C}^N \tag{1}$$

of length L is applied to the data signal $\mathsf{s}[k] \in \mathbb{M}$ with variance σ_{s}^2, to get the transmit signal for the N antenna elements which propagates over the channel

$$\boldsymbol{h}[k] = \sum_{q=0}^{Q-1} \boldsymbol{h}_q \delta[k-q] \in \mathbb{C}^N \tag{2}$$

of length Q and is perturbed by additive white Gaussian noise $\mathsf{n}[k] \in \mathbb{C}$ with variance σ_{n}^2. The set $\mathbb{M}$ denotes the modulation alphabet.

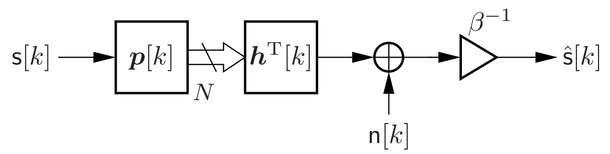

Fig. 1. System with FIR Transmit Filter

The resulting estimate after the multiplication of the received signal with the scalar β^{-1}, can be written as

$$\hat{\mathsf{s}}[k] = \beta^{-1}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{H}\mathbf{s}[k] + \beta^{-1}\mathsf{n}[k] \in \mathbb{C}, \tag{3}$$

where

$$\boldsymbol{p} = \left[\boldsymbol{p}_0^{\mathrm{T}}, \quad \boldsymbol{p}_1^{\mathrm{T}}, \quad \ldots, \quad \boldsymbol{p}_{L-1}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{NL} \tag{4}$$

comprises the transmit filter coefficients,

$$\boldsymbol{H} = \sum_{q=0}^{Q-1} \left[\mathbf{0}_{L\times q}, \quad \mathbf{1}_L, \quad \mathbf{0}_{L\times(Q-q-1)}\right] \otimes \boldsymbol{h}_q \tag{5}$$

is a $NL \times (L+Q-1)$ block Toeplitz matrix representing the channel $\boldsymbol{h}[k]$, and the data symbols effecting $\hat{\mathsf{s}}[k]$ are collected in the data vector $\mathbf{s}[k] \in \mathbb{M}^{L+Q-1}$, i. e.

$$\mathbf{s}[k] = \left[\mathsf{s}[k], \quad \mathsf{s}[k-1], \quad \ldots, \quad \mathsf{s}[k-L-Q+2]\right]^{\mathrm{T}}. \tag{6}$$

The *Transmit Wiener Filter* (TxWF) $\boldsymbol{p}_{\mathrm{WF}}$ and the optimal scalar β_{WF} at the receiver minimize the *Mean Square Error* (MSE) between the data signal $\mathsf{s}[k-\kappa]$ and its estimate $\hat{\mathsf{s}}[n]$ together with a transmit power constraint [15]:

$$\{\boldsymbol{p}_{\mathrm{WF}}, \beta_{\mathrm{WF}}\} = \underset{\{\boldsymbol{p},\beta\}}{\mathrm{argmin}}\, \mathrm{E}\left\{|\mathsf{s}[k-\kappa] - \hat{\mathsf{s}}[k]|^2\right\}$$
$$\text{s. t. } \mathrm{E}\left\{\|\boldsymbol{p}[k] * \mathsf{s}[k]\|_2^2\right\} = E_{\mathrm{tr}}. \tag{7}$$

The latency time $\kappa \in \mathbb{N}$ is chosen adequately but fixed, i. e. not optimized for every channel realization. Without loss of generality, the scalar β can be assumed to be positive real, i. e. $\beta \in \mathbb{R}_+$, since it is only needed to fulfill the transmit power constraint. In the sequel, we assume additionally temporally uncorrelated data, i. e. $\mathrm{E}\{\mathsf{s}[k]\mathsf{s}^*[k+\mu]\} = 0$ for $\mu \neq 0$, and that the data is uncorrelated to the noise, i. e. $\mathrm{E}\{\mathsf{s}[k]\mathsf{n}^*[k+\mu]\} = 0$ for $\mu \in \mathbb{Z}$. Using the Lagrangian function

$$L(\boldsymbol{p},\beta,\lambda) = \sigma_{\mathsf{s}}^2 + \beta^{-2}\sigma_{\mathsf{n}}^2 - 2\beta^{-1}\sigma_{\mathsf{s}}^2\,\mathrm{Re}\left\{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{e}_{\kappa+1}\right\}$$
$$+ \beta^{-2}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{p}^* + \lambda\left(E_{\mathrm{tr}} - \sigma_{\mathsf{s}}^2\boldsymbol{p}^{\mathrm{H}}\boldsymbol{p}\right), \ \lambda \in \mathbb{R}, \tag{8}$$

we get the TxWF

$$\boldsymbol{p}_{\mathrm{WF}} = \beta_{\mathrm{WF}}\boldsymbol{p}_0 \text{ with } \boldsymbol{p}_0 = \boldsymbol{R}_0^{-1}\boldsymbol{r}_0 \in \mathbb{C}^{NL}, \tag{9}$$

where $\boldsymbol{R}_0$ and $\boldsymbol{r}_0$ are defined as

$$\boldsymbol{R}_0 = \boldsymbol{H}^*\boldsymbol{H}^{\mathrm{T}} + \frac{\sigma_{\mathsf{n}}^2}{E_{\mathrm{tr}}}\mathbf{1}_{NL}, \quad \boldsymbol{r}_0 = \boldsymbol{H}^*\boldsymbol{e}_{\kappa+1}, \tag{10}$$

and the optimal scalar weight

$$\beta_{\mathrm{WF}} = \sqrt{\frac{E_{\mathrm{tr}}}{\sigma_{\mathsf{s}}^2\boldsymbol{r}_0^{\mathrm{H}}\boldsymbol{R}_0^{-2}\boldsymbol{r}_0}}. \tag{11}$$

Note that additionally to the full CSI, the transmitter needs information about the noise variance at the receiver which requires a feedback loop. Nevertheless, it was shown in [15] that the performance of the TxWF is quite robust against a wrongly choice of the noise variance if the system does not provide a possibility for feedback.

III. Transmit Multi-Stage Wiener Filter

A. TxWF with Postfiltering

In this section, we combine the transmit filter $\boldsymbol{p} \in \mathbb{C}^{NL}$ with the matrix FIR postfilter

$$\boldsymbol{T}_1[k] = \sum_{\ell=0}^{L-1} \boldsymbol{T}_{1,\ell}\delta[k-\ell] \in \mathbb{C}^{N\times NL} \tag{12}$$

of length L as shown in Figure 2.

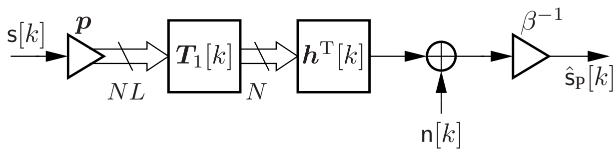

Fig. 2. Transmit Filter with FIR Postfilter

With the quadratic postfilter matrix

$$\boldsymbol{T}_1 = \left[\boldsymbol{T}_{1,0}^{\mathrm{T}}, \quad \boldsymbol{T}_{1,1}^{\mathrm{T}}, \quad \ldots, \quad \boldsymbol{T}_{1,L-1}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{NL\times NL}, \tag{13}$$

the estimate reads as

$$\hat{\mathsf{s}}_{\mathsf{P}}[k] = \beta^{-1}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{T}_1^{\mathrm{T}}\boldsymbol{H}\mathbf{s}[k] + \beta^{-1}\mathsf{n}[k]. \tag{14}$$

With the optimization (cf. Equation 7)

$$\{\boldsymbol{p}_{\mathrm{PWF}}, \beta_{\mathrm{PWF}}\} = \underset{\{\boldsymbol{p},\beta\}}{\operatorname{argmin}} \mathrm{E}\left\{ |\mathsf{s}[k-\kappa] - \hat{\mathsf{s}}_{\mathrm{P}}[k]|^2 \right\} \quad \text{s.t. } \sigma_{\mathsf{s}}^2 \left\| \boldsymbol{T}_1 \boldsymbol{p} \right\|_2^2 = E_{\mathrm{tr}}, \tag{15}$$

we get the postfiltered TxWF

$$\boldsymbol{p}_{\mathrm{PWF}} = \beta_{\mathrm{PWF}} \boldsymbol{p}_{\mathrm{P}} \text{ with } \boldsymbol{p}_{\mathrm{P}} = \left(\boldsymbol{T}_1^{\mathrm{H}} \boldsymbol{R}_0 \boldsymbol{T}_1 \right)^{-1} \boldsymbol{T}_1^{\mathrm{H}} \boldsymbol{r}_0, \tag{16}$$

and the optimal scalar

$$\beta_{\mathrm{PWF}} = \sqrt{ \frac{E_{\mathrm{tr}}}{ \sigma_{\mathsf{s}}^2 \boldsymbol{r}_0^{\mathrm{H}} \left(\boldsymbol{T}_1 \left(\boldsymbol{T}_1^{\mathrm{H}} \boldsymbol{R}_0 \boldsymbol{T}_1 \right)^{-1} \boldsymbol{T}_1^{\mathrm{H}} \right)^2 \boldsymbol{r}_0 } }. \tag{17}$$

Note that if $\boldsymbol{T}_1$ is a full-rank matrix, the estimate $\hat{\mathsf{s}}[k]$ of the TxWF and the estimate $\hat{\mathsf{s}}_{\mathrm{P}}[k]$ of the postfiltered TxWF are the same, i. e. $\hat{\mathsf{s}}_{\mathrm{P}}[k] = \hat{\mathsf{s}}[k]$, because in this case, $\boldsymbol{p}_{\mathrm{WF}} = \boldsymbol{T}_1 \boldsymbol{p}_{\mathrm{PWF}}$ and $\beta_{\mathrm{WF}} = \beta_{\mathrm{PWF}}$. Consequently, $\boldsymbol{p}_0 = \boldsymbol{T}_1 \boldsymbol{p}_{\mathrm{P}}$.

In the following, we consider the full-rank postfilter matrix of the structure

$$\boldsymbol{T}_1 = \begin{bmatrix} \boldsymbol{m}_1, & \boldsymbol{B}_1 \end{bmatrix}. \tag{18}$$

The vector $\boldsymbol{m}_1 \in \mathbb{C}^{NL}$ is chosen to be a normalized *Transmit Matched Filter* (TxMF) maximizing the cross-correlation between the received signal portion $\hat{\mathsf{s}}_1[k, \boldsymbol{m}] = \alpha_1 \boldsymbol{m}^{\mathrm{T}} \boldsymbol{H} \mathbf{s}[k] + \beta_{\mathrm{WF}}^{-1} \mathsf{n}[k]$ due to $\boldsymbol{m}$ and the desired signal $\mathsf{s}[k-\kappa]$, i. e.

$$\begin{aligned} \boldsymbol{m}_1 &= \underset{\boldsymbol{m}}{\operatorname{argmax}} \mathrm{E}\left\{ \mathrm{Re}\left\{ \hat{\mathsf{s}}_1[k, \boldsymbol{m}] \mathbf{s}^{\mathrm{H}}[k] \boldsymbol{e}_{\kappa+1} \right\} \right\} \\ &= \underset{\boldsymbol{m}}{\operatorname{argmax}} \left(\alpha_1 \boldsymbol{m}^{\mathrm{T}} \boldsymbol{r}_0^* + \alpha_1^* \boldsymbol{m}^{\mathrm{H}} \boldsymbol{r}_0 \right) \quad \text{s.t. } \boldsymbol{m}^{\mathrm{H}} \boldsymbol{m} = 1, \end{aligned} \tag{19}$$

where the first element of the filter vector $\boldsymbol{p}_{\mathrm{P}}$ is denoted as α_1. The solution computes as

$$\boldsymbol{m}_1 = \frac{\alpha_1^*}{|\alpha_1|} \frac{\boldsymbol{r}_0}{\|\boldsymbol{r}_0\|_2} = \frac{\boldsymbol{r}_0}{\|\boldsymbol{r}_0\|_2} \in \mathbb{C}^{NL}, \tag{20}$$

if we assume that $\alpha_1 \in \mathbb{R}_+$. This can be done without loss of generality since the product $\alpha_1 \boldsymbol{m}_1^{\mathrm{T}} = |\alpha_1| \, \boldsymbol{r}_0^{\mathrm{T}} / \|\boldsymbol{r}_0\|_2$ which is finally applied to the transmit filter, is independent of the phase of α_1. Due to the normalization of $\boldsymbol{m}_1$, an additional transmit power constraint need not be taken into account. Finally, the columns of the matrix $\boldsymbol{B}_1 \in \mathbb{C}^{NL \times (NL-1)}$ are defined to be orthogonal to $\boldsymbol{m}_1$, i. e.

$$\mathrm{span}\left(\boldsymbol{B}_1 \right) = \mathrm{null}\left(\boldsymbol{m}_1^{\mathrm{H}} \right) \Leftrightarrow \boldsymbol{B}_1^{\mathrm{H}} \boldsymbol{m}_1 = \boldsymbol{0}_{NL \times 1}. \tag{21}$$

Applying the inversion lemma for partitioned matrices [1], we get for the filter vector

$$\boldsymbol{p}_{\mathrm{P}} = \alpha_1 \begin{bmatrix} 1 \\ -\boldsymbol{p}_1 \end{bmatrix} \tag{22}$$

with

$$\boldsymbol{p}_1 = \boldsymbol{R}_1^{-1} \boldsymbol{r}_1 \in \mathbb{C}^{NL-1}, \tag{23}$$

$$\alpha_1 = \|\boldsymbol{r}_0\|_2 \left(\sigma_1^2 - \boldsymbol{r}_1^{\mathrm{H}} \boldsymbol{R}_1^{-1} \boldsymbol{r}_1 \right)^{-1} \in \mathbb{R}_+, \tag{24}$$

where

$$\boldsymbol{R}_1 = \boldsymbol{B}_1^{\mathrm{H}} \boldsymbol{R}_0 \boldsymbol{B}_1 \in \mathbb{C}^{(NL-1)\times(NL-1)}, \tag{25}$$

$$\boldsymbol{r}_1 = \boldsymbol{B}_1^{\mathrm{H}} \boldsymbol{R}_0 \boldsymbol{m}_1 \in \mathbb{C}^{NL-1}, \text{ and} \tag{26}$$

$$\sigma_1^2 = \boldsymbol{m}_1^{\mathrm{H}} \boldsymbol{R}_0 \boldsymbol{m}_1 \in \mathbb{R}_+. \tag{27}$$

Therefore, the filter $\boldsymbol{p}_0$ (cf. Equation 9) can be substituted by $\alpha_1 \left(\boldsymbol{m}_1 - \boldsymbol{B}_1 \boldsymbol{p}_1 \right)$ leading to the first step of the multi-stage decomposition of the TxWF depicted in Figure 3 where the non-zero coefficients of the matrix FIR filter $\boldsymbol{B}_1[k] \in \mathbb{C}^{N \times (NL-1)}$ and the vector FIR filter $\boldsymbol{m}_1[k] \in \mathbb{C}^N$, both of length L, are defined implicitly by

$$\boldsymbol{B}_1 = \left[\boldsymbol{B}_1^{\mathrm{T}}[0], \quad \boldsymbol{B}_1^{\mathrm{T}}[1], \quad \ldots, \quad \boldsymbol{B}_1^{\mathrm{T}}[L-1] \right]^{\mathrm{T}}, \tag{28}$$

$$\boldsymbol{m}_1 = \left[\boldsymbol{m}_1^{\mathrm{T}}[0], \quad \boldsymbol{m}_1^{\mathrm{T}}[1], \quad \ldots, \quad \boldsymbol{m}_1^{\mathrm{T}}[L-1] \right]^{\mathrm{T}}. \tag{29}$$

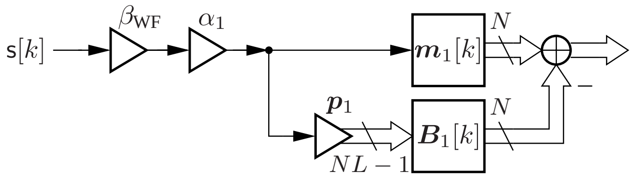

Fig. 3. Transmit Multi-Stage Wiener Filter after First Step

It can be shown that $\boldsymbol{p}_1 \in \mathbb{C}^{NL-1}$ is a reduced-dimension TxWF minimizing the MSE between the received signal portion due to $\boldsymbol{m}_1$ and the signal portion due to $\boldsymbol{B}_1$, i. e.

$$\boldsymbol{p}_1 = \underset{\boldsymbol{p}}{\operatorname{argmin}} \mathrm{E}\left\{ \left| \boldsymbol{m}_1^{\mathrm{T}} \boldsymbol{H} \mathbf{s}[k] - \boldsymbol{p}^{\mathrm{T}} \boldsymbol{B}_1^{\mathrm{T}} \boldsymbol{H} \mathbf{s}[k] \right|^2 \right\} \quad \text{s.t. } \beta_{\mathrm{WF}}^2 \sigma_{\mathsf{s}}^2 \left\| \alpha_1 \left(\boldsymbol{m}_1 - \boldsymbol{B}_1 \boldsymbol{p} \right) \right\|_2^2 = E_{\mathrm{tr}}, \tag{30}$$

and $\alpha_1 \in \mathbb{R}_+$ is a scalar TxWF minimizing the MSE between $\hat{\mathsf{s}}[k]$ and $\mathsf{s}[k-\kappa]$ with the corresponding transmit power constraint and under the assumption of the filter structure given in Figure 3. Remember that the derived structure produces the same estimate $\hat{\mathsf{s}}[k]$ as the TxWF of Equations (9) to (11).

B. Multi-Stage Decomposition of the TxWF

The fundamental idea of the multi-stage decomposition is to repeat the substitution of $\boldsymbol{p}_0 = \alpha_1 \left(\boldsymbol{m}_1 - \boldsymbol{B}_1 \boldsymbol{p}_1 \right) \in \mathbb{C}^{NL}$ for $\boldsymbol{p}_1 \in \mathbb{C}^{NL-1}$, and generally for the $(i-1)$-th TxWF $\boldsymbol{p}_{i-1} \in \mathbb{C}^{NL-i+1}$, $i \in \{1, 2, \ldots, NL\}$, i. e.

$$\boldsymbol{p}_{i-1} = \alpha_i \left(\boldsymbol{m}_i - \boldsymbol{B}_i \boldsymbol{p}_i \right). \tag{31}$$

If we define the received signal portion (cf. Figure 4)

$$\hat{\mathsf{s}}_i[k, \boldsymbol{m}] = (-1)^{i+1} \left(\prod_{j=1}^{i} \alpha_j \right) \boldsymbol{m}^{\mathrm{T}} \left(\prod_{j=i-1}^{1} \boldsymbol{B}_j^{\mathrm{T}} \right) \boldsymbol{H} \mathbf{s}[k] + \beta_{\mathrm{WF}}^{-1} \mathsf{n}[k] \tag{32}$$

due to the filter chain with the vector $\boldsymbol{m} \in \mathbb{C}^{NL-i+1}$, the vector filter $\boldsymbol{m}_i \in \mathbb{C}^{NL-i+1}$ is chosen to maximize the cross-correlation between $\hat{\mathsf{s}}_i[k, \boldsymbol{m}]$ and $\hat{\mathsf{s}}_{i-1}[k, \boldsymbol{m}_{i-1}]$, i. e.

$$\boldsymbol{m}_i = \underset{\boldsymbol{m}}{\operatorname{argmax}} \mathrm{E}\left\{ \mathrm{Re}\left\{ \hat{\mathsf{s}}_i[k, \boldsymbol{m}] \hat{\mathsf{s}}_{i-1}^*[k, \boldsymbol{m}_{i-1}] \right\} \right\} \quad \text{s.t. } \boldsymbol{m}^{\mathrm{H}} \boldsymbol{m} = 1. \tag{33}$$

The solution is given by the normalized TxMF

$$\boldsymbol{m}_i = \frac{\boldsymbol{r}_{i-1}}{\|\boldsymbol{r}_{i-1}\|_2} \in \mathbb{C}^{NL-i+1}, \tag{34}$$

with

$$\boldsymbol{r}_i = \boldsymbol{B}_i^{\mathrm{H}} \boldsymbol{R}_{i-1} \boldsymbol{m}_i \in \mathbb{C}^{NL-i} \text{ and} \tag{35}$$

$$\boldsymbol{R}_i = \boldsymbol{B}_i^{\mathrm{H}} \boldsymbol{R}_{i-1} \boldsymbol{B}_i \in \mathbb{C}^{(NL-i)\times(NL-i)}, \tag{36}$$

if we assume that the columns of the postfilter matrices $\boldsymbol{B}_j$, $j \in \{1, 2, \ldots, i-1\}$, are orthonormal, i. e. $\boldsymbol{B}_j^{\mathrm{H}} \boldsymbol{B}_j = \mathbf{1}_{NL-j}$, and the scalars $\alpha_j \in \mathbb{R}_+$ for $j \in \{1, 2, \ldots, i\}$. Analogous to the first step of the multi-stage decomposition, the columns of the matrix $\boldsymbol{B}_i \in \mathbb{C}^{(NL-i+1)\times(NL-i)}$ span the subspace orthogonal to $\boldsymbol{m}_i$, i. e.

$$\operatorname{span}(\boldsymbol{B}_i) = \operatorname{null}\left(\boldsymbol{m}_i^{\mathrm{H}}\right) \Leftrightarrow \boldsymbol{B}_i^{\mathrm{H}} \boldsymbol{m}_i = \mathbf{0}_{NL-i}. \tag{37}$$

Again, we assume in the following that $\boldsymbol{B}_i^{\mathrm{H}} \boldsymbol{B}_i = \mathbf{1}_{NL-i}$. With the definition of $\boldsymbol{m}_i$ and $\boldsymbol{B}_i$ given by Equation (34) and (37), respectively, the vector filter $\boldsymbol{p}_i$ and the scalar α_i compute as

$$\boldsymbol{p}_i = \boldsymbol{R}_i^{-1} \boldsymbol{r}_i \in \mathbb{C}^{NL-i}, \text{ and} \tag{38}$$

$$\alpha_i = \|\boldsymbol{r}_{i-1}\|_2 \left(\sigma_i^2 - \boldsymbol{r}_i^{\mathrm{H}} \boldsymbol{R}_i^{-1} \boldsymbol{r}_i\right)^{-1} \in \mathbb{R}_+, \tag{39}$$

where $\sigma_i^2 = \boldsymbol{m}_i^{\mathrm{H}} \boldsymbol{R}_{i-1} \boldsymbol{m}_i$. After some computation steps which are not shown in this paper due to space limitations, the vector filter $\boldsymbol{p}_i$ is also the solution of the optimization problem

$$\begin{aligned} \boldsymbol{p}_i = \underset{\boldsymbol{p}}{\operatorname{argmin}} \operatorname{E}\left\{\left|\hat{\mathsf{s}}_i[k, \boldsymbol{m}_i) - \hat{\mathsf{s}}_i[k, \boldsymbol{B}_i \boldsymbol{p})\right|^2\right\} \\ \text{s. t. } \beta_{\mathrm{WF}}^2 \sigma_{\mathsf{s}}^2 \left\| \sum_{j=1}^{i} (-1)^{j+1} \left(\prod_{u=1}^{j-1} \boldsymbol{B}_u\right) \boldsymbol{m}_j \left(\prod_{u=1}^{j} \alpha_u\right) \right. \\ \left. + (-1)^i \left(\prod_{j=1}^{i} \boldsymbol{B}_j\right) \boldsymbol{p} \left(\prod_{j=1}^{i} \alpha_j\right) \right\|_2^2 = E_{\mathrm{tr}}, \end{aligned} \tag{40}$$

i. e. a TxWF minimizing the MSE between the received signal portion $\hat{\mathsf{s}}_i[k, \boldsymbol{m}_i)$ due to the filter chain with $\boldsymbol{m}_i$ and $\hat{\mathsf{s}}_i[k, \boldsymbol{B}_i \boldsymbol{p})$ due to the filter chain with $\boldsymbol{p} \in \mathbb{C}^{NL-i}$. Analogous, α_i is a scalar TxWF minimizing the MSE of two adjacent received signal portions.

After the last substitution of $\boldsymbol{p}_{NL-1}$ by $\alpha_{NL} \boldsymbol{m}_{NL}$, the multi-stage decomposition of the TxWF is complete. The resulting *Transmit Multi-Stage Wiener Filter* (TxMSWF) is depicted in Figure 4. Note that only $\boldsymbol{m}_1$ and $\boldsymbol{B}_1$ comprise coefficients of FIR filters, whereas the following stages only determine how the columns of $\boldsymbol{B}_1$ have to be combined.

C. Reduced-Rank TxMSWF and its Relationship to Krylov Subspace Methods

Again, the transmit filter structure of Figure 4 produces the same output as the full-rank TxWF. The reduced-rank TxMSWF with rank D can be found by stopping the multi-stage decomposition after D steps and replacing $\boldsymbol{p}_{D-1}$ by $\alpha_D \boldsymbol{m}_D$ similar to the MSWF at the receiver (e. g. [5]). In Figure 5, we present a filterbank representation of the reduced-rank TxMSWF motivated by [8]. The vector representation of the i-th FIR filter $\boldsymbol{t}_i[k]$ of length L, i. e.

$$\boldsymbol{t}_i = \left[\boldsymbol{t}_i^{\mathrm{T}}[0], \boldsymbol{t}_i^{\mathrm{T}}[1], \ldots, \boldsymbol{t}_i^{\mathrm{T}}[L-1]\right]^{\mathrm{T}} \in \mathbb{C}^{NL}, \tag{41}$$

is simply the combination of $\boldsymbol{m}_i$ with the succeeding blocking matrices $\boldsymbol{B}_{i-1}$, $\boldsymbol{B}_{i-2}$, ..., and $\boldsymbol{B}_1$, i. e.

$$\boldsymbol{t}_i = \left(\prod_{j=1}^{i-1} \boldsymbol{B}_j\right) \boldsymbol{m}_i. \tag{42}$$

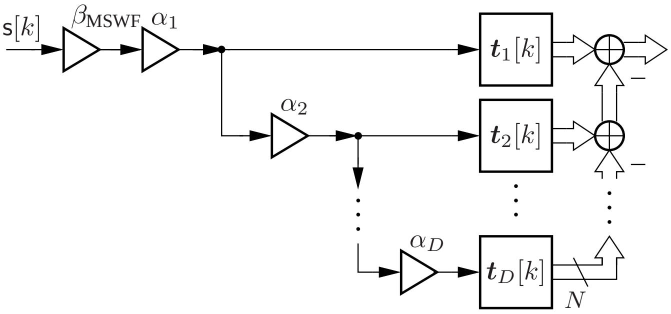

Fig. 5. Rank D TxMSWF as Filterbank

With the definition of the received signal portion

$$\hat{\mathsf{s}}_i^{(D)}[k, \boldsymbol{t}) = (-1)^{i+1} \left(\prod_{j=1}^{i} \alpha_j\right) \boldsymbol{t}^{\mathrm{T}} \boldsymbol{H} \mathsf{s}[k] + \beta_{\mathrm{MSWF}}^{-1} \mathsf{n}[k] \tag{43}$$

and the assumption that the postfilter vectors $\boldsymbol{t}_i$ are mutually orthonormal, they may be computed by solving the optimization problem (cf. Equation 33)

$$\begin{aligned} \boldsymbol{t}_i &= \underset{\boldsymbol{t}}{\operatorname{argmax}} \operatorname{E}\left\{\operatorname{Re}\left\{\hat{\mathsf{s}}_i^{(D)}[k, \boldsymbol{t}) \hat{\mathsf{s}}_{i-1}^{(D),*}[k, \boldsymbol{t}_{i-1})\right\}\right\} \\ &= \underset{\boldsymbol{t}}{\operatorname{argmax}} \left(\boldsymbol{t}^{\mathrm{T}} \boldsymbol{H} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{t}_{i-1}^* + \boldsymbol{t}^{\mathrm{H}} \boldsymbol{H}^* \boldsymbol{H}^{\mathrm{T}} \boldsymbol{t}_{i-1}\right) \\ &\quad \text{s. t. } \boldsymbol{t}^{\mathrm{H}} \boldsymbol{t} = 1 \text{ and} \\ &\quad\quad \boldsymbol{t}^{\mathrm{H}} \boldsymbol{t}_j = 0 \text{ for } j \in \{1, 2, \ldots, i-1\}. \end{aligned} \tag{44}$$

Using the method of Lagrangian multipliers and the fact that $\boldsymbol{P}_{i-1} \boldsymbol{t}_{i-1} = \mathbf{0}_{NL}$, the solution leads to the recursive filter formula

$$\boldsymbol{t}_i = \frac{\boldsymbol{P}_{i-2} \boldsymbol{P}_{i-1} \boldsymbol{R}_0 \boldsymbol{t}_{i-1}}{\|\boldsymbol{P}_{i-2} \boldsymbol{P}_{i-1} \boldsymbol{R}_0 \boldsymbol{t}_{i-1}\|_2}, \tag{45}$$

where $\boldsymbol{P}_i = \mathbf{1}_{NL} - \boldsymbol{t}_i \boldsymbol{t}_i^{\mathrm{H}}$ are projector matrices projecting onto the subspace orthogonal to the one spanned by the vectors $\boldsymbol{t}_i$. The above recursive algorithm is the modified Gram-Schmidt Lanczos algorithm [9], [6]. Thus, we see that the set of filters $\{\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_D\}$ is an orthonormal basis of the D-dimensional Krylov subspace [6]

$$\mathcal{K}^{(D)}(\boldsymbol{R}_0, \boldsymbol{r}_0) = \operatorname{span}\left\{\boldsymbol{r}_0, \boldsymbol{R}_0 \boldsymbol{r}_0, \ldots, \boldsymbol{R}_0^{(D-1)} \boldsymbol{r}_0\right\}. \tag{46}$$

Note that each Lanczos step performs a matrix-vector multiplication of the $NL \times NL$-matrix $\boldsymbol{R}_0$ with a NL-dimensional

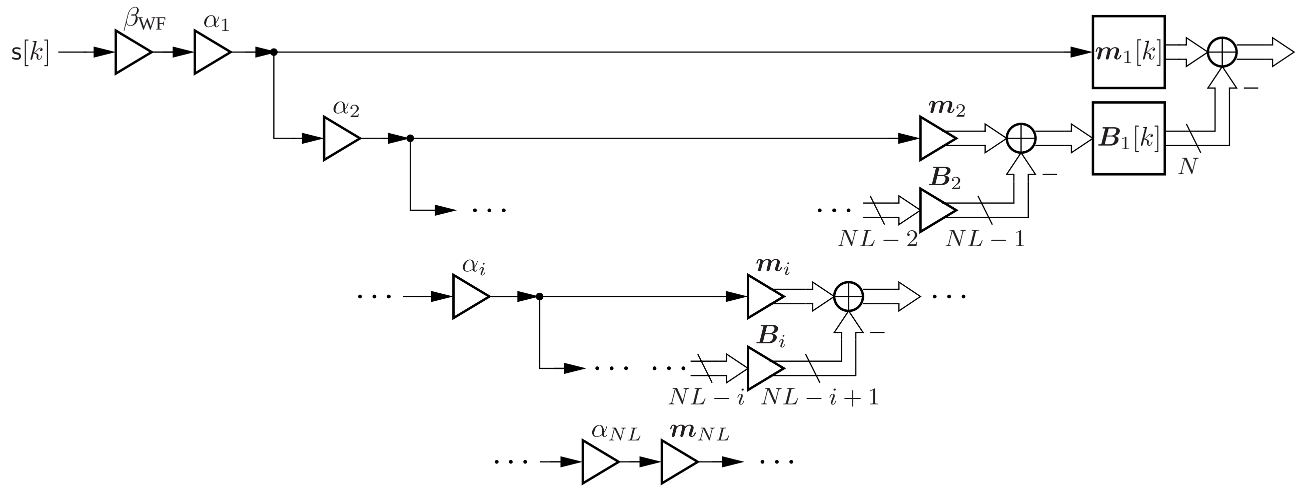

Fig. 4. Transmit Multi-Stage Wiener Filter

vector, which determines the order of the computational complexity to be $\mathrm{O}\left(N^2L^2\right)$. Since the rank D TxMSWF performs the Lanczos step D times, its order is $\mathrm{O}\left(DN^2L^2\right)$ whereas the full-rank TxWF has $\mathrm{O}\left(N^3L^3\right)$. If $D \ll NL$, the TxMSWF leads to an enormous reduction in computational complexity compared to the TxWF.

IV. Simulation Results

We simulate the signal transmission over a single user communication system with QPSK symbol modulation. The BS comprises two transmit antenna elements and the MS one receive antenna element. The channel is assumed to be frequency-selective with three dominant propagation paths, the first one without delay, i. e. the path with line of sight, the second one with delay $Q = 7$ which is equal to the maximum delay spread of the channel, and the third one with an integer delay chosen arbitrarily for every channel realization between the delays of the first two paths. All path weights are independent and identically normal distributed and the sum of their variances per transmit antenna is normalized. The channel impulse response is assumed to be instananously known at the transmitter. This assumption holds in a *Time Division Duplex* (TDD) system with a slow varying channel because of its reciprocity. The perfectly estimated channel of the uplink reception, is used for the calculation of the transmit filter applied to the downlink transmission. The impulse response length of the FIR precoding filter $\boldsymbol{p}[k]$ is set to $L = 9$. Thus, the received signal is obtained by linear combination of $L + Q - 1 = 15$ adjacent input signals $\mathsf{s}[k]$. The latency time of the transmit filter is fixed to $\kappa = 7$. Note that in order to further improve the behavior of the presented algorithms, the latency time can be optimized for every channel realization. Nevertheless, this leads to an increase in computational complexity especially for the Krylov subspace based algorithms and is therefore not shown in this paper since we are interested in computational cheap implementations. The following simulation results are averaged over several thousand channel realizations.

In Figure 6, the uncoded *Bit Error Rates* (BERs) of the different methods are plotted over the *Signal to Noise Ratio* (SNR), i. e. the ratio of the signal power at the transmitter output to the noise power at the receiver input. In the following, the reduced-rank *Transmit Eigensubspace based WF* (TxEWF) denotes the approximation of the TxWF in the subspace spanned by the principal eigenvectors of the matrix $\boldsymbol{R}_0$, i. e. the prefilter vectors $\boldsymbol{t}_i$ are no longer base vectors of the Krylov subspace but eigenvectors corresponding to the largest eigenvalues of $\boldsymbol{R}_0$. It can be seen that the rank $D = 6$ TxEWF is clearly outperformed by the reduced-rank TxMSWF, even if its rank is reduced to $D = 1$ where the TxMSWF is equal to the TxMF. Unlike the TxEWF, the TxMSWF exhibits already for a rank equal to $D = 3$ a far better BER than the TxMF and for a rank of $D = 6$, its performance comes even very close to the one of the full-rank solution of the TxMSWF with $D_{\text{max}} = NL = 18$ which corresponds to the TxWF. Only for SNR values greater than $14\,\text{dB}$, we obtain marginally better results with the TxWF.

Figure 7 depicts the simulation results for the BER depending on the rank D of the proposed filter solutions at a SNR value of $10\,\text{dB}$. Note that the TxWF (full-rank TxMSWF) is plotted over D as an asymptote for the compared methods and again, the TxMF corresponds to the TxMSWF at rank $D = 1$. The TxEWF solution of rank $D = 1$ yields a BER of approximately 0.4 whereas the TxMSWF with the same rank leads to a BER which is by a factor of 10 lower. The TxEWF outbalances the TxMF only for ranks greater or equal than $D = 8$ and converges to the BER of the TxWF for a rank of $D = 15$ which is a rather high value compared to the full rank of $D_{\text{max}} = 18$. Moreover, for every desired BER, the TxEWF needs a greater rank than the TxMSWF, resulting in a higher computational complexity since the TxMSWF and the TxEWF of rank D have the same order of computational complexity, i. e. $\mathrm{O}\left(DN^2L^2\right)$. Contrary to the behavior of the TxEWF, the TxMSWF reaches already BER values comparable to the TxWF for a rank of $D = 5$. Hence, the TxMSWF enables

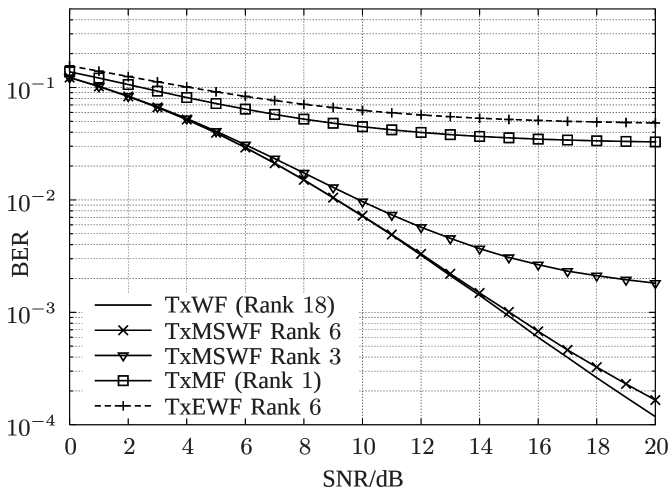

Fig. 6. BER comparison between the reduced-rank TxMSWF, the reduced-rank TxEWF, the TxMF, and the optimal TxWF over SNR

a tremendous reduction in computational complexity without significant performance degradation.

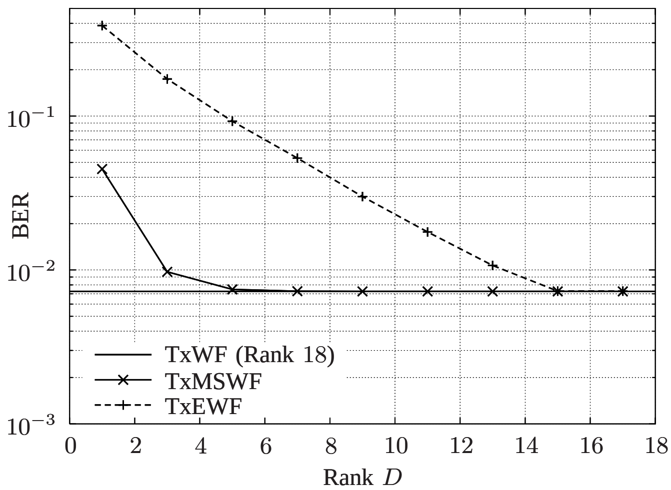

Fig. 7. BER comparison between the TxMSWF and the TxEWF method over the rank D (SNR $= 10\,\mathrm{dB}$)

V. Conclusions

In this paper, we derived the TxMSWF, a multi-stage decomposition of the TxWF. Furthermore, we presented a Lanczos based implementation of the TxMSWF exploiting its relationship to Krylov subspace methods. Simulation results of the application to a time-dispersive communication system showed that the TxMSWF achieves near optimum performance for relatively low rank, thus, tremendously outperforming the eigensubspace based method.

References

[1] L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1991.

[2] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 1999.

[3] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educational Psychology*, vol. 24, no. 6/7, pp. 417–441, 498–520, September/October 1933.

[4] J. S. Goldstein and I. S. Reed, "Reduced-Rank Adaptive Filtering," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 492–496, February 1997.

[5] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2943–2959, November 1998.

[6] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS – *out of print*, 1996.

[7] M. L. Honig and W. Xiao, "Performance of Reduced-Rank Linear Interference Suppression," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1928–1946, July 2001.

[8] M. Joham, Y. Sun, M. D. Zoltowski, M. Honig, and J. S. Goldstein, "A New Backward Recursion for the Multi-Stage Nested Wiener Filter Employing Krylov Subspace Methods," in *Proc. Milcom 2001*, October 2001, vol. 2, pp. 1210–1213.

[9] C. Lanczos, "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," *Journal of Research of the National Bureau of Standards*, vol. 45, no. 4, pp. 255–282, October 1950.

[10] M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, December 1952.

[11] G. Dietl, M. D. Zoltowski, and M. Joham, "Recursive Reduced-Rank Adaptive Equalization for Wireless Communications," in *Proc. SPIE 2001*, April 2001, vol. 4395, pp. 16–27.

[12] L. L. Scharf, L. T. McWhorter, E. K. P. Chong, J. S. Goldstein, and M. D. Zoltowski, "Algebraic Equivalence of Conjugate Directions and Multistage Wiener Filters," in *Proceedings of the Eleventh Annual Workshop on ASAP*, March 2003.

[13] W. L. Myrick, M. D. Zoltowski, and J. S. Goldstein, "Anti-Jam Space-Time Preprocessor for GPS Based on Multistage Nested Wiener Filter," in *Proc. MILCOM '99*, October 1999.

[14] S. Chowdhury, M. D. Zoltowski, and J. S. Goldstein, "Application of Reduced-Rank Chip-Level MMSE Equalization to Forward Link DS-CDMA with Frequency Selective Multipath," in *Proc. 38th Annual Allerton Conference on Communications, Systems, and Computing*, October 2000.

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

[16] R. L. Choi and R. D. Murch, "Transmit MMSE Pre-Rake Pre-processing with Simplified Receivers for the Downlink of MISO TDD-CDMA Systems," in *Proc. Globecom 2002*, November 2002, vol. 1, pp. 429–433.

[17] R. E. McIntosh and S. E. El-Khamy, "Optimum Pulse Transmission Through a Plasma Medium," *IEEE Transactions on Antennas and Propagation*, vol. AP-18, no. 5, pp. 666–671, September 1970.

[18] B. R. Vojčić and W. M. Jang, "Transmitter Precoding in Synchronous Multiuser Communications," *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1346–1355, October 1998.