LOW-COMPLEXITY MMSE RECEIVERS BASED ON WEIGHTED MATRIX POLYNOMIALS IN FREQUENCY-SELECTIVE MIMO SYSTEMS

Guido Dietl, Ingmar Groh, and Wolfgang Utschick

Institute for Circuit Theory and Signal Processing
Munich University of Technology, 80333 Munich, Germany
E-mail: dietl@tum.de

ABSTRACT

Compared to the *Matrix Wiener Filter* (MWF), reduced-rank signal processing in the *Minimum Mean Square Error* (MMSE) sense is a well-known method for the design of low-complexity receivers. In this paper, we reveal the relationship between different reduced-rank receivers based on weighted matrix polynomials approximating the MWF in a *Krylov subspace*, viz., the *Multi-Stage Matrix WF* (MSMWF), the parallel implementation of *Multi-Stage Vector WFs* (MSVWFs), and *Polynomial Expansion* (PE). Besides, we present PE where the weights are approximated based on *Random Matrix* (RM) *theory* assuming the application to a frequency-selective *Multiple-Input Multiple-Output* (MIMO) system. Simulation results show that the MSMWF outperforms the considered reduced-rank methods if their order of computational complexity is the same.

1. INTRODUCTION

The MWF [1] estimates an unknown signal vector from an observation signal vector based on the MMSE criterion. Consequently, the MWF computes as the solution of the *Wiener-Hopf equation* which is computationally intense for observations of high dimensionality. Reduced-rank methods approximate the MWF in a lower dimensional subspace, thus, they can be written as a weighted sum of the corresponding base vectors. If the computation of this basis and the calculation of the optimal weights is less complex than the computation of the full-rank MWF, the reduced-rank approach decreases computational complexity.

The MSMWF introduced by Goldstein et al. [2, 3] approximates the MWF in a *Krylov subspace* composed of the auto-correlation matrix of the observation and the cross-correlation matrix between the observation and the desired signal. Note that the MSMWF is an extension of the MSVWF [4] estimating a scalar signal instead of a vector. Another Krylov subspace based MWF approximation replaces the inverse of the auto-correlation matrix needed to solve the Wiener-Hopf equation, by an aborted *Neumann series* [5]. Note that the Neumann series has unit weights, thus, its polynomial needs in nite degree to provide an exact representation of the MWF. PE [6] is an alternative to the *Neumann series* where the unit weights are replaced by the *Mean Square Error* (MSE) optimal ones. Hence, the degree of the weighted polynomial which is equal to the exact MWF, is the dimension of the observation reduced by one.

RM theory is a well-known tool to analyze analytically the performance of large-scale communication systems (see e.g., [7, 8, 9]). Recently, it has been used for the design of reduced-rank receive lters in *multi-user code division multiple access* systems [10, 11]. Thereby, the PE weights of the MWF approximation which depend generally on the current channel realization, have been replaced by realization independent approximations exploiting the statistical characteristics of the channel. Note that these approximations of the polynomial coef cients have been derived according to the *Signal to Interference and Noise Ratio* (SINR) criterion. In this paper, we present PE where the weights are approximated based on RM theory assuming the application to a frequency-selective *Multiple-Input Multiple-Output* (MIMO) system. Compared to previous contributions, the weights are based on the same criterion as used for the derivation of the MWF, i. e., the MSE criterion.

The next section de nes the channel model and Sec. 3 brie y reviews the derivation of the MWF. Before presenting the simulation results in Sec. 5, we show the fundamental difference between the considered reduced-rank methods and approximate weights based on RM theory in Sec. 4.

Throughout the paper, vectors and matrices are denoted by lower and upper case bold letters, and random variables are written using *sans serif* font. The matrix $\boldsymbol{I}_n$ is the $n \times n$ identity matrix and $\mathbf{0}_{n\times m}$ the $n \times m$ zero matrix. $\boldsymbol{R}_{\mathsf{u}} = \mathrm{E}\{\mathsf{u}\mathsf{u}^H\}$ with the expectation operator $\mathrm{E}\{\cdot\}$, denotes the auto-correlation matrix of the random vector u and $\boldsymbol{R}_{\mathsf{u},\mathsf{v}} = \mathrm{E}\{\mathsf{u}\mathsf{v}^H\}$ the cross-correlation matrix between the random vectors u and v.

2. CHANNEL MODEL

In this paper, we consider the time-dispersive MIMO channel of length L with T inputs and R outputs which can be described by the channel matrix impulse response

$$\boldsymbol{H}[n] = \sum_{\ell=0}^{L-1} \boldsymbol{H}_\ell \delta[n-\ell] \in \mathbb{C}^{R\times T}, \tag{1}$$

with the unit impulse function $\delta[n]$. The weighting matrices of the propagation paths, $\boldsymbol{H}_\ell$, $\ell \in \{0, 1, \ldots, L-1\}$, are realizations of the random matrix $\boldsymbol{H}_\ell$ which has *independent and identically distributed* (i. i. d.) entries with the complex normal distribution $\mathcal{N}_\mathrm{c}(0, 1/L)$. The received signal vector

$$\boldsymbol{r}[n] = \boldsymbol{H}[n] * \boldsymbol{s}[n] + \boldsymbol{\eta}[n] \in \mathbb{C}^R, \tag{2}$$

where '$*$' denotes convolution, is perturbed by additive white Gaussian noise $\boldsymbol{\eta}[n] \in \mathbb{C}^R$ with $\mathcal{N}_\mathrm{c}\left(\mathbf{0}_{R\times 1}, \sigma^2 \boldsymbol{I}_R\right)$. The transmit signal vector $\boldsymbol{s}[n]$ at time index n is composed of T zero-mean i. i. d. symbols with variance one. Note that the total transmit power is $P_\mathrm{Tx} = \operatorname{tr}\{\boldsymbol{R}_{\boldsymbol{s}}\} = T$.

In order to compute the linear equalizer filter of length K (cf. Sec. 3), we derive an alternative matrix-vector model of the time-dispersive MIMO channel. The vector $\tilde{\boldsymbol{r}}[n] = \left[\boldsymbol{r}^\mathrm{T}[n], \boldsymbol{r}^\mathrm{T}[n-1], \ldots, \boldsymbol{r}^\mathrm{T}[n-K+1]\right]^\mathrm{T} \in \mathbb{C}^{KR}$, where $(\cdot)^\mathrm{T}$ denotes transpose, is composed of K adjacent receive signal vectors $\boldsymbol{r}[n]$. Using the block Toeplitz matrix

$$\tilde{\boldsymbol{H}} = \sum_{\ell=0}^{L-1} \boldsymbol{S}_{(\ell,K,L-1)} \otimes \boldsymbol{H}_\ell \in \mathbb{C}^{KR\times(K+L-1)T}, \tag{3}$$

where '$\otimes$' denotes the Kronecker product and $\boldsymbol{S}_{(\ell,K,L-1)} = \left[\mathbf{0}_{K\times\ell}, \boldsymbol{I}_K, \mathbf{0}_{K\times(L-1-\ell)}\right] \in \{0,1\}^{K\times(K+L-1)}$ the selection matrix, Eq. (2) may be rewritten as

$$\tilde{\boldsymbol{r}}[n] = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{s}}[n] + \tilde{\boldsymbol{\eta}}[n] \in \mathbb{C}^{KR}. \tag{4}$$

Analogous to $\tilde{\boldsymbol{r}}[n]$, the vector $\tilde{\boldsymbol{s}}[n] \in \mathbb{C}^{(K+L-1)T}$ is composed of $K+L-1$ adjacent transmit signal vectors $\boldsymbol{s}[n]$ and $\tilde{\boldsymbol{\eta}}[n] \in \mathbb{C}^{KR}$ of K adjacent noise vectors $\boldsymbol{\eta}[n]$.

3. MATRIX WIENER FILTER

Applying the linear matrix filter $\boldsymbol{\mathcal{W}} \in \mathbb{C}^{M\times N}$, $M \in \mathbb{N}$, $N \in \mathbb{N}$, to the observation vector $\boldsymbol{y}[n] \in \mathbb{C}^N$, $N \geq M$, leads to the estimate $\hat{\boldsymbol{x}}[n] = \boldsymbol{\mathcal{W}}\boldsymbol{y}[n]$ of the desired signal vector $\boldsymbol{x}[n] \in \mathbb{C}^M$. The power of the Euclidean norm of the estimation error $\boldsymbol{\varepsilon}[n] = \boldsymbol{x}[n] - \hat{\boldsymbol{x}}[n]$ is the MSE

$$\xi(\boldsymbol{\mathcal{W}}) = \operatorname{tr}\left\{\boldsymbol{R}_{\boldsymbol{x}} - \boldsymbol{\mathcal{W}}\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}} - \boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}^\mathrm{H}\boldsymbol{\mathcal{W}}^\mathrm{H} + \boldsymbol{\mathcal{W}}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{\mathcal{W}}^\mathrm{H}\right\}, \tag{5}$$

where $(\cdot)^\mathrm{H}$ denotes Hermitian, i. e., conjugate transpose. The MWF $\boldsymbol{W}$ minimizes $\xi(\boldsymbol{\mathcal{W}})$ and consequently solves the *Wiener-Hopf equation*

$$\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{W}^\mathrm{H} = \boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}} \quad \Leftrightarrow \quad \boldsymbol{W} = \boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}^\mathrm{H}\boldsymbol{R}_{\boldsymbol{y}}^{-1}. \tag{6}$$

The filter algorithms may be applied to the given MIMO scenario of Sec. 2 if we set $\boldsymbol{y}[n] = \tilde{\boldsymbol{r}}[n]$ and $\boldsymbol{x}[n] = \boldsymbol{s}[n-\nu]$ where ν is the latency time introduced by the equalizer. Consequently, $N = KR$, $M = T$, $\boldsymbol{R}_{\boldsymbol{y}} = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^\mathrm{H} + \sigma^2\boldsymbol{I}_N$, and $\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}} = \tilde{\boldsymbol{H}}\boldsymbol{S}_\nu^\mathrm{T}$ with $\boldsymbol{S}_\nu = \boldsymbol{S}_{(M\nu,M,M(K+L-2))}$.

4. REDUCED-RANK MWF APPROXIMATIONS BASED ON WEIGHTED MATRIX POLYNOMIALS

4.1. Multi-Stage Matrix and Vector Wiener Filter

The rank D MSMWF [2, 3] can be written as

$$\boldsymbol{W}_\mathrm{MSMWF}^{(D)} = \sum_{i=0}^{d-1} \boldsymbol{W}_i \boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}^\mathrm{H}\boldsymbol{R}_{\boldsymbol{y}}^i \in \mathbb{C}^{M\times N}, \tag{7}$$

where $d = D/M$ and $D \in \{1, 2, \ldots, N\}$. The weighting matrices $\boldsymbol{W}_i \in \mathbb{C}^{M\times M}$ are determined by the block Lanczos algorithm [12, 3] which generates an orthonormal basis of the D-dimensional *Krylov subspace* [13, 3]

$$\mathcal{K}^{(D)} = \operatorname{span}\left\{\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}, \boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}, \ldots, \boldsymbol{R}_{\boldsymbol{y}}^{d-1}\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}\right\}, \tag{8}$$

leading to a block tridiagonal auto-correlation matrix of the prefiltered observation vector (e. g., [3]). Thus, the computation of the weighting matrices is negligible compared to the generation of the corresponding orthonormal Krylov basis. Since each of the $d-1$ iterations needs a multiplication of $\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}$ and $\boldsymbol{R}_{\boldsymbol{y}}$ [3], the computational complexity of the MSMWF is of order $\mathrm{O}(dMN^2) = \mathrm{O}(DN^2)$.

An alternative reduced-rank MWF in $\mathcal{K}^{(D')}$ can be obtained by estimating each element of $\boldsymbol{x}[n]$ via a MSVWF [4]. Hence, the resulting matrix filter computes as

$$\left[\boldsymbol{W}_\mathrm{MSVWF}^{(D')}\right]_i = \left[\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}^\mathrm{H}\right]_i \sum_{j=0}^{D'-1} w_{ij}\boldsymbol{R}_{\boldsymbol{y}}^j,\ i \in \{1, 2, \ldots, M\}, \tag{9}$$

where $D' \in \{1, 2, \ldots, N\}$ and $[\cdot]_i$ denotes the i-th row of a matrix. Note that each of the M rows requires a computational complexity of $\mathrm{O}(D'N^2)$ due to the calculation of $D'-1$ matrix-vector products [3]. Again, the complexity of computing the optimal weights w_{ij} is negligible. Consequently, for a fair comparison between MSMWF and the parallel implementation of MSVWFs in Sec. 5, we set $D' = d$.

4.2. Polynomial Expansion

PE [6] is strongly related to the *Neumann series* [5] of the inverse $\boldsymbol{R}_{\boldsymbol{y}}^{-1} = \sum_{i=0}^{\infty}(\boldsymbol{I}_N - \boldsymbol{R}_{\boldsymbol{y}})^i$, $\|\boldsymbol{I}_N - \boldsymbol{R}_{\boldsymbol{y}}\|_\mathrm{F} < 1$, where $\|\cdot\|_\mathrm{F}$ denotes the Frobenius norm of a matrix. However, the inverse $\boldsymbol{R}_{\boldsymbol{y}}^{-1}$ is approximated by a matrix polynomial of finite degree, i. e.,

$$\boldsymbol{W}_\mathrm{PE}^{(D')} = \boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}^\mathrm{H} \sum_{i=0}^{D'-1} w_i \boldsymbol{R}_{\boldsymbol{y}}^i, \tag{10}$$

where $D' \in \{1, 2, \ldots, N\}$ and w_i are the MSE optimal weights. Note that contrary to the parallel implementation of MSVWFs (cf. Eq. 9), the PE weights w_i are equal for all columns of $\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}$, although both are approximations of the MWF in the same subspace $\mathcal{K}^{(D')}$. Again, for a fair performance comparison with respect to computational complexity, we choose $D' = d$.

In order to end up in a system of linear equations for the determination of the optimal weights w_i in Eq. (10) which is also used for the derivation of the RM coef cients in the next section, we de ne the MSE

$$\xi'(\boldsymbol{\mathcal{W}}) = \mathrm{E}\left\{\|\tilde{\boldsymbol{s}}[n] - \boldsymbol{\mathcal{W}}\tilde{\boldsymbol{r}}[n]\|_2^2\right\}, \tag{11}$$

where $\|\cdot\|_2$ denotes the Euclidean norm, and where we restrict the linear matrix lter to be $\boldsymbol{\mathcal{W}} = \tilde{\boldsymbol{H}}^{\mathrm{H}}\sum_{i=0}^{D'-1}\omega_i'\boldsymbol{R}_{\boldsymbol{y}}^i = \tilde{\boldsymbol{H}}^{\mathrm{H}}\sum_{i=0}^{D'-1}\omega_i(\tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^{\mathrm{H}})^i$. Recall that $\boldsymbol{y}[n] = \tilde{\boldsymbol{r}}[n]$. The latter equality holds since the matrices $\boldsymbol{R}_{\boldsymbol{y}}$ and $\tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^{\mathrm{H}}$ differ only in a scaled identity matrix. It can easily be shown that

$$\{w_i\}_{i=0}^{D'-1} = \underset{\{\omega_i\}_{i=0}^{D'-1}}{\mathrm{argmin}}\ \xi'(\boldsymbol{\mathcal{W}}) = \underset{\{\omega_i\}_{i=0}^{D'-1}}{\mathrm{argmin}}\ \xi(\boldsymbol{S}_\nu\boldsymbol{\mathcal{W}}) \tag{12}$$

are the optimal weights needed to nally compute the PE lter matrix (cf. Eq. 10)

$$\boldsymbol{W}_{\mathrm{PE}}^{(D')} = \boldsymbol{S}_\nu\tilde{\boldsymbol{H}}^{\mathrm{H}}\sum_{i=0}^{D'-1} w_i\left(\tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^{\mathrm{H}}\right)^i. \tag{13}$$

Solving Eq. (12) yields the system of linear equations

$$\begin{aligned}&\mathrm{tr}\left\{\sum_{i=0}^{D'-1} w_i\left(\left(\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{H}}\right)^{i+j+2} + \sigma^2\left(\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{H}}\right)^{i+j+1}\right)\right\}\\ &= \mathrm{tr}\left\{\left(\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{H}}\right)^{j+1}\right\},\ j\in\{0,1,\ldots,D'-1\},\end{aligned} \tag{14}$$

which can be used to calculate w_i, $i \in \{0,1,\ldots,D'-1\}$.

4.3. Approximation of Weighting Coefficients Using Random Matrix Theory

In this subsection, we derive the reduced-rank MWF

$$\boldsymbol{W}_{\text{PE-RM}}^{(D')} = \boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}}\sum_{i=0}^{D'-1}\bar{w}_i\boldsymbol{R}_{\boldsymbol{y}}^i, \tag{15}$$

where the polynomial coef cients $\bar{w}_i$ are the approximations of the optimal weights w_i in $\boldsymbol{W}_{\mathrm{PE}}^{(D')}$ (cf. Eq. 10) based on RM theory. Thereby, the frequency-selective MIMO channel matrix $\tilde{\boldsymbol{H}}$ is seen as a realization of the random matrix $\tilde{\boldsymbol{\mathsf{H}}}$.

Assume the *Eigen-Value Decomposition* (EVD) of $\boldsymbol{\mathsf{R}}_1 = \tilde{\boldsymbol{\mathsf{H}}}^{\mathrm{H}}\tilde{\boldsymbol{\mathsf{H}}} = \boldsymbol{\mathsf{Q}}\boldsymbol{\Lambda}\boldsymbol{\mathsf{Q}}^{\mathrm{H}} \in \mathbb{C}^{M'\times M'}$, $M' = (L+K-1)T$, to be given. Since $\boldsymbol{\mathsf{R}}_1$ is positive semide nite and Hermitian, the EVD always exists. Replacing the k-th empirical moment $\mathrm{tr}\{\boldsymbol{\Lambda}^k\}/M'$ of the eigenvalue distribution of $\boldsymbol{\mathsf{R}}_1$ in Eq. (14) by the k-th statistical moment $m_k = \mathrm{E}\{\lambda_{\boldsymbol{\mathsf{R}}_1}^k\}$,[1] yields the system of linear equations

$$\boldsymbol{M}\bar{\boldsymbol{w}} = \boldsymbol{m}, \tag{16}$$

for the determination of the RM approximated PE coef cients $\bar{\boldsymbol{w}} = [\bar{w}_0, \bar{w}_1, \ldots, \bar{w}_{D'-1}]^{\mathrm{T}} \in \mathbb{R}^{D'}$, where $\boldsymbol{m} = [m_1, m_2\ldots, m_{D'}]^{\mathrm{T}} \in \mathbb{R}^{D'}$ and the (i,j)-th element of $\boldsymbol{M}$ is $m_{i+j} + \sigma^2 m_{i+j-1}$, $i,j \in \{1,2,\ldots,D'\}$ [10, 9]. Since statistics do not depend on the explicit channel realization, we have *a priori* computable polynomial coef cients.

The remaining problem is to derive the statistical moments m_k of the eigenvalue distribution of $\boldsymbol{\mathsf{R}}_1$. In order to do so, it is necessary to assume the existence of an asymptotic eigenvalue distribution $F_{\boldsymbol{\mathsf{R}}_1}(\lambda) = \lim_{M'\to\infty}|\{\lambda_{\boldsymbol{\mathsf{R}}_1} : \lambda_{\boldsymbol{\mathsf{R}}_1} < \lambda\}|/M'$ where $|\cdot|$ denotes the cardinality of a set. Moreover, we exploit the fact that the eigenvalue distributions of the random matrices $\boldsymbol{\mathsf{R}}_2 = \tilde{\boldsymbol{\mathsf{H}}}\tilde{\boldsymbol{\mathsf{H}}}^{\mathrm{H}} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{\mathsf{R}}_1 = \tilde{\boldsymbol{\mathsf{H}}}^{\mathrm{H}}\tilde{\boldsymbol{\mathsf{H}}}$ are scaled versions of each other.

In the limiting case $M' \to \infty$, $\tilde{\boldsymbol{\mathsf{H}}}$ can be considered as cyclic [9]. Consequently,

$$\tilde{\boldsymbol{\mathsf{H}}} = \left(\boldsymbol{F}_K^{\mathrm{H}}\otimes\boldsymbol{I}_R\right)\check{\boldsymbol{\mathsf{H}}}\left(\boldsymbol{F}_{K+L-1}\otimes\boldsymbol{I}_T\right) \in \mathbb{C}^{N\times M'}, \tag{17}$$

where $\check{\boldsymbol{\mathsf{H}}} \in \mathbb{C}^{N\times M'}$ is a block-diagonal matrix with the blocks $\check{\boldsymbol{\mathsf{H}}}_i = \sum_{\ell=0}^{L-1}\boldsymbol{\mathsf{H}}_\ell\varphi_L^{i\ell} \in \mathbb{C}^{R\times T}$ on the main diagonal, and where $\boldsymbol{F}_n$ denotes the $n\times n$-dimensional discrete Fourier matrix, i. e., the $(k+1,\ell+1)$-th element of $\boldsymbol{F}_n$ is given by $\varphi_n^{k\ell} = \exp(-\mathrm{j}\,2\pi k\ell/n)$, $k,\ell \in \{0,1,\ldots,n-1\}$. In the following, we will show that the statistical properties of the eigenvalues of $\boldsymbol{\mathsf{R}}_2 = \tilde{\boldsymbol{\mathsf{H}}}\tilde{\boldsymbol{\mathsf{H}}}^{\mathrm{H}}$ will not differ from those of $\boldsymbol{\mathsf{R}}_3 = \check{\boldsymbol{\mathsf{H}}}_i\check{\boldsymbol{\mathsf{H}}}_i^{\mathrm{H}} \in \mathbb{C}^{R\times R}$ independent of the respective index i. First, we rewrite the Gramian matrix $\check{\boldsymbol{\mathsf{H}}}_i\check{\boldsymbol{\mathsf{H}}}_i^{\mathrm{H}}$ as

$$\boldsymbol{\mathsf{R}}_3 = \check{\boldsymbol{\mathsf{H}}}_i\check{\boldsymbol{\mathsf{H}}}_i^{\mathrm{H}} = \bar{\boldsymbol{\mathsf{H}}}\left(\boldsymbol{f}_i\boldsymbol{f}_i^{\mathrm{H}}\otimes\boldsymbol{I}_T\right)\bar{\boldsymbol{\mathsf{H}}}^{\mathrm{H}} = \bar{\boldsymbol{\mathsf{H}}}\boldsymbol{\mathsf{P}}\bar{\boldsymbol{\mathsf{H}}}^{\mathrm{H}}, \tag{18}$$

where $\bar{\boldsymbol{\mathsf{H}}} = [\boldsymbol{\mathsf{H}}_0, \boldsymbol{\mathsf{H}}_1, \ldots, \boldsymbol{\mathsf{H}}_{L-1}] \in \mathbb{C}^{R\times TL}$ and $\boldsymbol{f}_i \in \mathbb{C}^L$ denotes the i-th column of the discrete Fourier matrix $\boldsymbol{F}_L \in \mathbb{C}^{L\times L}$, $i \in \{1,2,\ldots,L\}$. Then, we analyze the matrix $\boldsymbol{\mathsf{P}} = \boldsymbol{f}_i\boldsymbol{f}_i^{\mathrm{H}}\otimes\boldsymbol{I}_T$ which is build by a Kronecker product of the matrices $\boldsymbol{f}_i\boldsymbol{f}_i^{\mathrm{H}}$ and $\boldsymbol{I}_T$. Thus, the eigenvalues of $\boldsymbol{\mathsf{P}}$ are given by $\mathrm{diag}\{[L, \boldsymbol{0}_{1\times(L-1)}]\}\otimes\boldsymbol{I}_T$, i. e., the Kronecker product of the eigenvalues $\mathrm{diag}\{[L, \boldsymbol{0}_{1\times(L-1)}]\}$ of $\boldsymbol{f}_i\boldsymbol{f}_i^{\mathrm{H}}$ and the eigenvalues of the identity matrix $\boldsymbol{I}_T$. Therefore, $\boldsymbol{\mathsf{P}}$ and $\boldsymbol{f}_i\boldsymbol{f}_i^{\mathrm{H}}$ have the same eigenvalue density, namely

$$f_{\boldsymbol{\mathsf{P}}}(\lambda) = f_{\boldsymbol{f}_i\boldsymbol{f}_i^{\mathrm{H}}}(\lambda) = \frac{L-1}{L}\delta(\lambda) + \frac{1}{L}\delta(\lambda - L), \tag{19}$$

with the Dirac delta function $\delta(\lambda)$. We observe in Eq. (19) that the eigenvalue density $f_{\boldsymbol{\mathsf{P}}}(\lambda)$ does not depend on the block index i. Hence, the eigenvalue density of $\boldsymbol{\mathsf{R}}_3$ is also independent of i (cf. Eq. 18). Consequently, if every block in $\check{\boldsymbol{\mathsf{H}}}\check{\boldsymbol{\mathsf{H}}}^{\mathrm{H}} \in \mathbb{C}^{N\times N}$ has the same eigenvalue statistics, this holds for the matrix containing all blocks as well. Moreover, according to Eq. (17), $\check{\boldsymbol{\mathsf{H}}}\check{\boldsymbol{\mathsf{H}}}^{\mathrm{H}}$ and $\boldsymbol{\mathsf{R}}_2 = \tilde{\boldsymbol{\mathsf{H}}}\tilde{\boldsymbol{\mathsf{H}}}^{\mathrm{H}}$ have the same non-zero eigenvalues. So far, we have shown that the eigenvalue density of $\boldsymbol{\mathsf{R}}_2$ is equal to the eigenvalue density of the random matrix $\boldsymbol{\mathsf{R}}_3 = \bar{\boldsymbol{\mathsf{H}}}\boldsymbol{\mathsf{P}}\bar{\boldsymbol{\mathsf{H}}}^{\mathrm{H}} + \boldsymbol{\mathsf{N}} \in \mathbb{C}^{R\times R}$ (cf. Eq. 18) where $\boldsymbol{\mathsf{N}} = \boldsymbol{0}_{R\times R}$. It remains to calculate the statistical moments of the eigenvalue distribution of $\boldsymbol{\mathsf{R}}_3$.

With the *Stieltjes transform*

$$G_{\boldsymbol{\mathsf{R}}_3}(\zeta) = \int\frac{1}{\lambda-\zeta}\mathrm{d}F_{\boldsymbol{\mathsf{R}}_3}(\lambda), \tag{20}$$

[1] Note that each non-zero element of the diagonal matrix $\boldsymbol{\Lambda}$ of eigenvalues is a realization of the random variable $\lambda_{\boldsymbol{\mathsf{R}}_1}$.

the statistical moments $m'_k = \mathrm{E}\{\lambda^k_{\boldsymbol{R}_3}\}$ of the eigenvalue distribution of $\boldsymbol{R}_3$ can be calculated by the series expansion [9]

$$\frac{G_{\boldsymbol{R}_3}(\zeta^{-1})}{-\zeta} = \sum_{k=0}^{\infty} m'_k \zeta^k. \tag{21}$$

We determine the Stieltjes transform $G_{\boldsymbol{R}_3}(\zeta)$ from the following xed-point equation [14]

$$G_{\boldsymbol{R}_3}(\zeta) = G_{\boldsymbol{N}}\left(\zeta - \frac{TL}{R}\int \frac{\lambda f_{\boldsymbol{P}}(\lambda)}{1+\lambda G_{\boldsymbol{R}_3}(\zeta)}\mathrm{d}\lambda\right), \tag{22}$$

where $f_{\boldsymbol{N}}(\lambda) = \delta(\lambda)$ and $G_{\boldsymbol{N}}(\zeta) = -1/\zeta$. Finally, we obtain for $k > 1$,

PSfrag replacements

$$m'_k = \sum_{i=\lceil \frac{k+1}{2} \rceil}^{k+1} \binom{i}{k+1-i}\binom{\frac{1}{2}}{i}\left(L\left(\frac{T}{R}-1\right)\right)^{2(k-i+1)} \cdot \frac{1}{2L}\left(-2L^2\left(\frac{T}{R}+1\right)\right)^{2i-k}, \tag{23}$$

and $m_k = \mathrm{E}\{\lambda^k_{\boldsymbol{R}_1}\} = Nm'_k/M'$ [10, 9] for $k > 1$ whereas $m_0 = 1$, and the polynomial coef cients $\bar{\boldsymbol{w}}$ can be computed according to Eq. (16).

5. SIMULATION RESULTS

We consider QPSK transmission over a MIMO channel of length $L = 2$ with $T = 4$ inputs and $R = 8$ outputs. The MWF length is $K = 3$, thus, the observation vector has dimension $N = KR = 24$. The latency time is set to $\nu = L - 1 = 1$. Fig. 1 depicts the uncoded *Bit Error Rates* (BERs) over *Signal to Noise Ratio* (SNR) $10\log_{10}(P_{\mathrm{Tx}}/\sigma^2)$ of the different reduced-rank MWFs presented in the previous section. We observe that for a comparable order of computational complexity, the MSMWF outperforms both the parallel implementation of MSVWFs and the two reduced-rank MWFs with optimal (PE) and RM approximated polynomial weights (PE-RM). This is due to the fact that the MSMWF exploits a higher dimensional Krylov subspace $\mathcal{K}^{(D)}$ in a computational ef cient way. Nevertheless, for small SNRs, the performance of the PE-RM approximation is close to the one of optimal PE as well as the MSMWF and all of the reduced-rank MWFs are good approximations of the full-rank MWF despite of their reduced computational complexity.

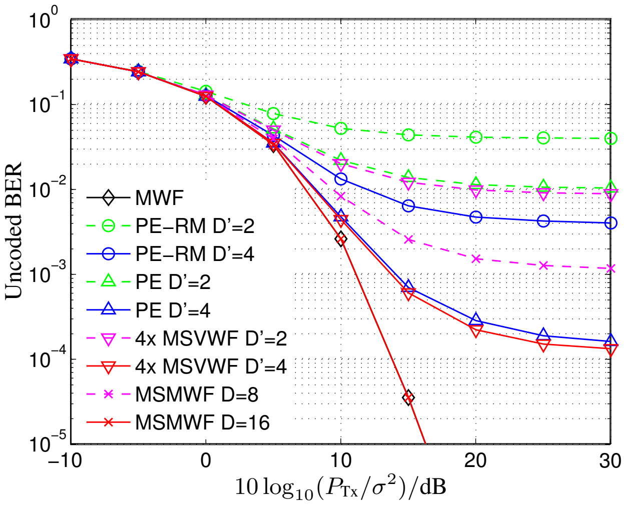

Fig. 1. Uncoded BER vs. SNR for different reduced-rank MWFs where $T = 4$, $R = 8$, $L = 2$, and $K = 3$

6. REFERENCES

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

[2] J. S. Goldstein, I. S. Reed, D. E. Dungeon, and J. R. Guerci, "A Multistage Matrix Wiener Filter for Subspace Detection," *Proc. DECI 1999*, February 1999.

[3] G. Dietl, P. Breun, and W. Utschick, "Block Krylov Methods in Time-Dispersive MIMO Systems," in *Proc. ICC 2004*, June 2004, vol. 5, pp. 2683–2688.

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

[5] I. Gohberg, S. Goldberg, and M. A. Kaashoek, *Basic Classes of Linear Operators*, Birkhäuser, 2003.

[6] G. M. A. Sessler and F. K. Jondral, "Low Complexity Polynomial Expansion Multiuser Dectector with Strong Near-Far Resistance," *Proc. VTC Fall 2002*, vol. 3, pp. 1869–1872, September 2002.

[7] D. N. C. Tse and S. V. Hanly, "Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity," *IEEE Trans. on Inform. Theory*, vol. 45, no. 2, pp. 641–657, March 1999.

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

[9] R. R. Müller, "A Random Matrix Model of Communications Via Antenna Arrays," *IEEE Trans. on Inform. Theory*, vol. 48, no. 9, pp. 2495–2506, September 2002.

[10] R. R. Müller and S. Verdù, "Design and Analysis of Low-Complexity Interference Mitigation Based on Vector Channels," *IEEE J. on Select. Areas Commun.*, vol. 19, no. 4, pp. 1429–1441, August 2001.

[11] L. Li, A. M. Tulino, and S. Verdù, "Design of Reduced-Rank MMSE Multiuser Detectors Using Random Matrix Methods," *IEEE Trans. on Inform. Theory*, vol. 50, no. 6, pp. 986–1008, June 2004.

[12] G. H. Golub and R. Underwood, "The block Lanczos method for computing eigenvalues," in *Mathematical Software III*, J. R. Rice, Ed., pp. 361–377. Academic Press, 1977.

[13] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2003.

[14] J. W. Silverstein and Z. D. Bai, "On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices," *J. Multivariate Anal.*, vol. 54, pp. 175–192, 1995.