MMSE Approaches to Multiuser Spatio-Temporal Tomlinson-Harashima Precoding

Michael Joham, Johannes Brehmer, and Wolfgang Utschick

Institute for Circuit Theory and Signal Processing, Munich University of Technology, 80290 Munich, Germany
Phone: +49 (89) 289-28510, Fax: +49 (89) 289-28504, Email: Michael.Joham@nws.ei.tum.de

Abstract

We investigate *Tomlinson-Harashima precoding* (THP) for multiuser systems. Since we aim at simple receivers, the feedforward and feedback filters are located at the transmitter. The filters are derived based on *minimum mean square error* (MMSE) criteria, with (zero-forcing THP) and without (Wiener THP) additional constraints. For flat fading and frequency selective scenarios, we modify the well known linear transmit filter optimizations to yield the optimizations for the respective THP approaches. The resulting feedback filters directly depend on the feedforward filters which are computed successively. Inspired by this observation, we include a suboptimum ordering strategy for spatial THP similar to *vertical Bell Laboratories space time* (V-BLAST) at the receiver. Moreover, we incorporate a suboptimum latency time optimization for temporal THP. Simulation results reveal that THP is particularly advantageous in conjunction with higher-order modulation schemes and for high *signal to noise ratio* (SNR).

1 Introduction

Motivated by the need for cheap mobiles with low power consumption, we focus on systems where the complex signal processing is performed at the *base station* (BS) as proposed by Gibbard et al. [1], i. e. receive processing in the uplink and transmit processing in the downlink. Similar to the noise enhancement problem of linear receive filters (e. g. [2]), linear transmit filters (e. g. [3], [4], [5], [6], [7], [8], [9], [10], [11]) suffer from the need of increased transmit power. To overcome this problem, THP feeds back already *transmitted symbols* to reduce interference at the receiver [12], [13]. A similar principle at the receiver is the *decision feedback equalization* (DFE, [14]), where already *detected symbols* are fed back to reduce interference.

THP was originally proposed to nonlinearly combat *intersymbol interference* (ISI), since the recursive filter necessary to equalize an FIR channel linearly can be unstable (see [12], [13], and also [15]). Spatial equalization of a flat fading *multiple input multiple output* (MIMO) system with THP was considered by Ginis et al. [16] and Fischer et al. [17]. In [18], Fischer et al. used THP for a DS-CDMA system which was also investigated by Liu et al. [19]. Contrary to the popular assumption of full channel knowledge at the transmitter, Fischer et al. [20] proposed THP based on partial channel state information (see also [21]). However, we assume perfect instantaneous channel state information at the transmitter which is available in *time division duplex* (TDD) systems from the channel estimation during the reception in the other link, if the calibration is good enough (see e. g. [22]) and the channel coherence time (e. g. [23]) is large enough. As channel estimation errors can be neglected compared to the variation of the channel due to Doppler (cf. [24]), prediction increases the applicability of precoding (see [25], [24]). In [26], channel variations are combatted with a robust filter design. In *frequency division duplex* (FDD) systems, feedback of the channel state information from the receivers to the transmitter is necessary.

Most contributions on THP assumed that the feedforward filter is located at the receiver. This assumption was made, because THP was used to overcome the *error propagation problem* of DFE. Obviously, the complexity of the receiver is only reduced slightly in this case, as the feedforward filter still has to be optimized and applied by the receiver. If the feedforward filter of THP is deployed at the transmitter, the receiver can be simplified dramatically. This setup for *decentralized receivers* has only been considered in [1], [16], [27], [18], [28], [19] and will be used in this paper.

First, we focus on multiuser systems with flat fading vector channels (which are equivalent to MIMO systems without cooperation of the receiving antenna elements) and modify the well known optimizations for the linear zero-forcing and the linear Wiener transmit filter to obtain the optimizations for the respective THP filters under the assumption of known ordering as we have done for DFE in [29]. For both THP filter types, the feedback filter directly depends on the feedforward filter and the channel. Since the columns of the resulting feedforward filters are computed successively by simply dropping the respective row in the channel matrix at each step, we employ a successive algorithm to find the suboptimum ordering. This heuristic approach to find the ordering for THP is similar to the V-BLAST ordering for DFE [30]. As our THP filters are based on an optimization that applies the same scalar weight at all receivers, our precoding tries to support the different

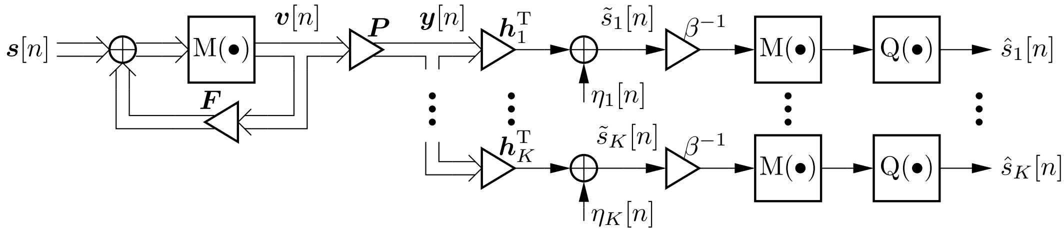

Fig. 1. Transmission over Nondispersive Vector Channels to K Receivers Applying Modulo Operator $\mathrm{M}(\bullet)$ and Quantizer $\mathrm{Q}(\bullet)$

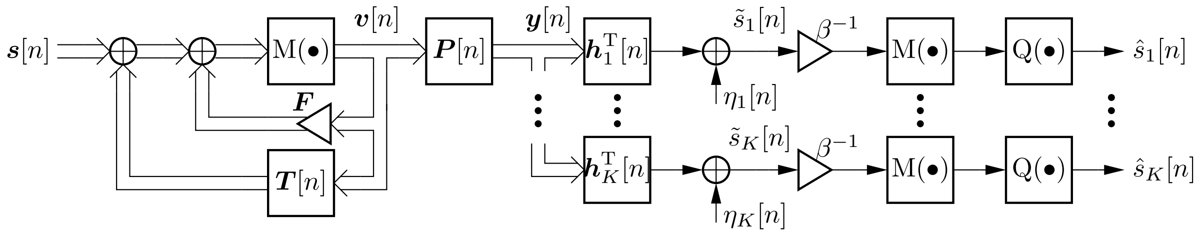

Fig. 2. Transmission over Time Dispersive Vector Channels to K Receivers Applying Modulo Operator $\mathrm{M}(\bullet)$ and Quantizer $\mathrm{Q}(\bullet)$

users equally in contrast to [31] where Windpassinger et al. proposed an ordering which leads to maximally different receive powers and is advandageous in conjunction with adaptive modulation.

Second, we investigate THP for frequency selective fading channels. Starting from the linear transmit filters, we develop spatio-temporal THP which suppresses the ISI of succeeding symbols linearly, the *co-channel interference* (CCI) with spatial THP, and ISI due to already transmitted symbols with temporal THP. We apply the ordering algorithm developed for flat fading channels also to spatio-temporal THP and incorporate a latency time optimization (see e. g. [32]) in THP for frequency selective channels.

We explain the system models for flat fading and frequency selective vector channels in Section 2 and discuss zero-forcing and Wiener THP together with the respective ordering algorithms for flat fading channels in Section 3. After introducing spatio-temporal THP for frequency selective channels in Section 4, we present and analyze the simulation results in Section 5.

1.1 Notation

Vectors and matrices are denoted by lower case bold and capital bold letters, respectively. We use $\mathrm{E}[\bullet]$, '$*$', '$\otimes$', $(\bullet)^*$, $(\bullet)^{\mathrm{T}}$, and $(\bullet)^{\mathrm{H}}$ for expectation, convolution, the Kronecker product, complex conjugation, transposition, and conjugate transposition, respectively. The trace of a matrix is denoted by $\mathrm{tr}(\bullet)$ and the pseudo inverse by $(\bullet)^+$. All random processes are assumed to be zero-mean and stationary. The covariance matrix of the vector process $\boldsymbol{x}[n]$ is denoted by $\boldsymbol{R}_{\boldsymbol{x}} = \mathrm{E}\left[\boldsymbol{x}[n]\boldsymbol{x}^{\mathrm{H}}[n]\right]$, whereas the variance of the scalar process $y[n]$ is denoted by $\sigma_y^2 = \mathrm{E}\left[|y[n]|^2\right]$. The $N \times M$ zero matrix is $\mathbf{0}_{N\times M}$, the M-dimensional zero vector is $\mathbf{0}_M$, and the $N \times N$ identity matrix is $\mathbf{1}_N$, whose n-th column is $\boldsymbol{e}_n$. Throughout the paper, we utilize the selection matrix $\boldsymbol{S}_{(q,M,N)} = [\mathbf{0}_{M\times q}, \mathbf{1}_M, \mathbf{0}_{M\times N-q}] \in \{0,1\}^{M\times M+N}$.

2 System Model

2.1 Flat Fading Vector Channels

The symbols $s_k[n], k = 1,\ldots,K$, for the K receivers are stacked in the vector signal

$$\boldsymbol{s}[n] = \left[s_1[n],\ldots,s_K[n]\right]^{\mathrm{T}} \in \mathbb{C}^K$$

and passed through the precoder consisting of the feedforward filter $\boldsymbol{P} \in \mathbb{C}^{N_\mathrm{a}\times K}$ and the feedback filter $\boldsymbol{F} \in \mathbb{C}^{K\times K}$ (cf. Fig. 1), where N_a denotes the number of antenna elements at the transmitter. For THP, the modulo operation for complex x is defined as

$$\mathrm{M}(x) = x - \left\lfloor \frac{\mathrm{Re}\,(x)}{\tau} + \frac{1}{2} \right\rfloor \tau - \mathrm{j} \left\lfloor \frac{\mathrm{Im}\,(x)}{\tau} + \frac{1}{2} \right\rfloor \tau.$$

The floor operator $\lfloor \bullet \rfloor$ gives the integer number smaller than or equal to the argument and the constant τ depends on the used modulation alphabet (e. g. [17]), e. g. as $s_k[n] \in \{\exp(\mathrm{j}\,\mu\pi/4)\,|\,\mu \in \{-3,-1,+1,+3\}\}$, we set $\tau = 2\sqrt{2}$ for QPSK. For linear transmit processing, the modulo operator reduces to $\mathrm{M}(x) = x$ and the feedback filter is inactive, i. e. $\boldsymbol{F} = \mathbf{0}_{K\times K}$.

The precoded signal $\boldsymbol{y}[n] = \boldsymbol{P}\boldsymbol{v}[n] \in \mathbb{C}^{N_\mathrm{a}}$ propagates over the k-th of the K vector channels $\boldsymbol{h}_k \in \mathbb{C}^{N_\mathrm{a}}$, and leads together with the complex Gaussian noise $\eta_k[n]$ to the received signal $\tilde{s}_k[n]$. After forming vector signals $\tilde{\boldsymbol{s}}[n], \boldsymbol{\eta}[n] \in \mathbb{C}^K$ similar to $\boldsymbol{s}[n]$, the received signals can be concisely expressed as

$$\tilde{\boldsymbol{s}}[n] = \boldsymbol{H}\boldsymbol{y}[n] + \boldsymbol{\eta}[n] = \boldsymbol{H}\boldsymbol{P}\boldsymbol{v}[n] + \boldsymbol{\eta}[n], \qquad (1)$$

where $\boldsymbol{H} = [\boldsymbol{h}_1,\ldots,\boldsymbol{h}_K]^{\mathrm{T}} \in \mathbb{C}^{K\times N_\mathrm{a}}$. All receivers apply the same scalar weighting β^{-1} to correct the amplitude of the desired signal which follows from the transmit power constraint. Note that [18], [19] did not make the restriction of identical weighting, but a unitary feedforward filter $\boldsymbol{P}$ was assumed.

Every receiver must apply a modulo operation $\mathrm{M}(\bullet)$ to remove the effect of the modulo operation at the transmitter. The quantizer $\mathrm{Q}(\bullet)$ maps onto the signal constellation to generate the symbol estimates. Note that the chain of modulo operation and quantizer can be interpreted as a modified quantizer that also maps onto the signal constellation, whose mapping is based on an infinite periodic repetition of the original signal constellation. Due to the additional neighbours in the periodic repetition, the error probability is increased.

2.2 Frequency Selective Vector Channels

Due to the frequency selectivity of the channels

$$\boldsymbol{h}_k[n] = \sum_{q=0}^{Q} \boldsymbol{h}_{k,q}\delta[n-q] \in \mathbb{C}^{N_\mathrm{a}},\ k=1,\ldots,K,$$

with maximum delay Q, the feedforward filter also has to be FIR:

$$\boldsymbol{P}[n] = \sum_{\ell=0}^{L} \boldsymbol{P}_\ell \delta[n-\ell] \in \mathbb{C}^{N_\mathrm{a}\times K}.$$

Additionally to the feedback filter $\boldsymbol{F} \in \mathbb{C}^{K\times K}$ for spatial THP, the feedback filter $\boldsymbol{T}[n] \in \mathbb{C}^{K\times K}$ is necessary for temporal THP. Again, when linear precoding is utilized, $\mathrm{M}(x) = x$ and the feedback filters are inactive, i. e. $\boldsymbol{F} = \boldsymbol{0}_{K\times K}$ and $\boldsymbol{T}[n] = \boldsymbol{0}_{K\times K}$.

The k-th received signal reads as (see Fig. 2)

$$\begin{aligned} \tilde{s}_k[n] &= \boldsymbol{h}_k^\mathrm{T}[n] * \boldsymbol{P}[n] * \boldsymbol{v}[n] + \eta_k[n] \quad \text{or} \\ \tilde{s}_k[n] &= \sum_{i=1}^{K} \boldsymbol{p}_i^\mathrm{T} \boldsymbol{H}_k \boldsymbol{v}_i[n] + \eta_k[n], \end{aligned} \tag{2}$$

where we defined

$$\begin{aligned} \boldsymbol{H}_k &= \sum_{q=0}^{Q} \boldsymbol{S}_{(q,L+1,Q)} \otimes \boldsymbol{h}_{k,q} \in \mathbb{C}^{N_\mathrm{a}(L+1)\times L+Q+1}, \\ \boldsymbol{p}_i &= [\boldsymbol{e}_i^\mathrm{T}\boldsymbol{P}_0^\mathrm{T}, \ldots, \boldsymbol{e}_i^\mathrm{T}\boldsymbol{P}_L^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{a}(L+1)}, \quad \text{and} \\ \boldsymbol{v}_i[n] &= [v_i[n], \ldots, v_i[n-L-Q]]^\mathrm{T} \in \mathbb{C}^{L+Q+1}, \end{aligned}$$

with $v_i[n] = \boldsymbol{e}_i^\mathrm{T}\boldsymbol{v}[n]$. Since we assume uncorrelated outputs of the modulo operator at the transmitter, i. e. $\mathrm{E}\big[\boldsymbol{v}_i[n]\boldsymbol{v}_j^\mathrm{T}[n]\big] = \delta[i-j]\sigma_v^2\boldsymbol{1}_{L+Q+1}$, the transmit power is simply $\mathrm{E}[\|\boldsymbol{y}[n]\|_2^2] = \sigma_v^2 \sum_{i=1}^{K} \|\boldsymbol{p}_i\|_2^2$.

3 THP for Flat Fading Channels

3.1 Zero-Forcing THP (ZF-THP)

Linear zero-forcing precoding (TxZF) results from MSE minimization under the constraint of an unbiased and interference-free estimate $\beta^{-1}\tilde{s}[n]$. The TxZF optimization can be written as (e. g. [8])

$$\begin{aligned} \{\boldsymbol{P}_\mathrm{ZF}', \beta_\mathrm{ZF}'\} &= \underset{\{\boldsymbol{P},\beta\}}{\mathrm{argmin}}\, \mathrm{E}\Big[\big\|\boldsymbol{s}[n] - \beta^{-1}\tilde{\boldsymbol{s}}[n]\big\|_2^2\Big] \\ \text{s. t.:}\ \ \tilde{\boldsymbol{s}}[n]\big|_{\boldsymbol{\eta}[n]=\boldsymbol{0}_K} &= \beta\boldsymbol{s}[n] \text{ and } \mathrm{E}\Big[\|\boldsymbol{y}[n]\|_2^2\Big] = E_\mathrm{tr}, \end{aligned} \tag{3}$$

where E_tr denotes the available transmit power. With

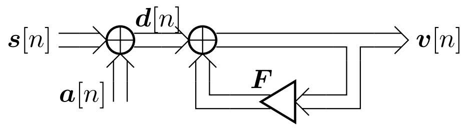

Fig. 3. Linear Representation of the Precoder

Lagrangian multipliers we easily obtain

$$\boldsymbol{P}_\mathrm{ZF}' = \beta_\mathrm{ZF}'\boldsymbol{H}^\mathrm{H}\left(\boldsymbol{H}\boldsymbol{H}^\mathrm{H}\right)^{-1}$$

and β_ZF' is chosen to set the transmit power to E_tr.

Fig. 3 which depicts a linear representation of the precoder used in the original papers [12], [13] suggests how we have to modify the optimization in Eqn. (3) to end up with the optimization for ZF-THP. The auxiliary signal $\boldsymbol{a}[n]$, whose entries have real and imaginary parts which are integer multiples of τ, is chosen to get entries of $\boldsymbol{v}[n]$ whose real and imaginary parts lie in the interval $[-\tau/2, +\tau/2)$. As $\boldsymbol{a}[n]$ is removed by the modulo operators at the receivers, the desired value for $\tilde{\boldsymbol{s}}[n]$ is $\boldsymbol{d}[n] = \boldsymbol{s}[n] + \boldsymbol{a}[n]$ instead of $\boldsymbol{s}[n]$. Moreover, we have to take into account that $\boldsymbol{F}$ can only feed back already precoded entries of $\boldsymbol{v}[n]$. Thus, the feedback filter $\boldsymbol{F}$ has to be lower triangular with zero main diagonal under the assumption of ordered data streams. Thus, we get following optimization for ZF-THP:

$$\begin{aligned} \{\boldsymbol{P}_\mathrm{ZF}, \boldsymbol{F}_\mathrm{ZF}, \beta_\mathrm{ZF}\} &= \underset{\{\boldsymbol{P},\boldsymbol{F},\beta\}}{\mathrm{argmin}}\, \mathrm{E}\Big[\big\|\boldsymbol{d}[n] - \beta^{-1}\tilde{\boldsymbol{s}}[n]\big\|_2^2\Big] \\ \text{s. t.:}\ \ \tilde{\boldsymbol{s}}[n]\big|_{\boldsymbol{\eta}[n]=\boldsymbol{0}_K} &= \beta\boldsymbol{d}[n],\ \mathrm{E}\Big[\|\boldsymbol{y}[n]\|_2^2\Big] = E_\mathrm{tr}, \\ \boldsymbol{F}:&\ \text{lower triangular, zero main diagonal.} \end{aligned} \tag{4}$$

Since $\boldsymbol{d}[n] = \boldsymbol{v}[n] - \boldsymbol{F}\boldsymbol{v}[n]$ as can be concluded from Fig. 3, above optimization simplifies to

$$\begin{aligned} \{\boldsymbol{P}_\mathrm{ZF}, \boldsymbol{F}_\mathrm{ZF}, \beta_\mathrm{ZF}\} &= \underset{\{\boldsymbol{P},\boldsymbol{F},\beta\}}{\mathrm{argmin}}\, \sigma_{\boldsymbol{\eta}}^2\beta^{-2} \\ \text{s. t.:}\ \ \boldsymbol{H}\boldsymbol{P} &= \beta\left(\boldsymbol{1}_K - \boldsymbol{F}\right),\ \mathrm{tr}\left(\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{v}}\boldsymbol{P}^\mathrm{H}\right) = E_\mathrm{tr},\ \text{and} \\ \boldsymbol{S}_i\boldsymbol{F}\boldsymbol{e}_i &= \boldsymbol{0}_i,\ i=1,\ldots,K, \end{aligned} \tag{5}$$

where $\sigma_{\boldsymbol{\eta}}^2 = \sum_{i=1}^{K}\sigma_{\eta_i}^2$. We split the constraint on the structure of $\boldsymbol{F}$ into K constraints, each for one column of $\boldsymbol{F}$. To this end, we introduced the selection matrix $\boldsymbol{S}_i = \boldsymbol{S}_{(0,i,K-i)} \in \{0,1\}^{i\times K}$ to cut out the first i entries of the i-th column of $\boldsymbol{F}$. Furthermore, we assume that the signals at the output of the modulo operator at the transmitter are uncorrelated, that is $\boldsymbol{R}_{\boldsymbol{v}} = \mathrm{E}[\boldsymbol{v}[n]\boldsymbol{v}^\mathrm{H}[n]] = \mathrm{diag}(\sigma_{v_1}^2, \ldots, \sigma_{v_K}^2)$. Then, the solution of Eqn. (5) can be found with Lagrangian multipliers:

$$\boldsymbol{P}_\mathrm{ZF} = \beta_\mathrm{ZF}\sum_{i=1}^{K}\boldsymbol{H}^\mathrm{H}\boldsymbol{S}_i^\mathrm{T}(\boldsymbol{S}_i\boldsymbol{H}\boldsymbol{H}^\mathrm{H}\boldsymbol{S}_i^\mathrm{T})^{-1}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^\mathrm{T},$$

$$\boldsymbol{F}_\mathrm{ZF} = \sum_{i=1}^{K}(\boldsymbol{e}_i - \boldsymbol{H}\boldsymbol{H}^\mathrm{H}\boldsymbol{S}_i^\mathrm{T}(\boldsymbol{S}_i\boldsymbol{H}\boldsymbol{H}^\mathrm{H}\boldsymbol{S}_i^\mathrm{T})^{-1}\boldsymbol{S}_i\boldsymbol{e}_i)\boldsymbol{e}_i^\mathrm{T},$$

and β_ZF is chosen to fulfill the transmit power constraint. Note that $\boldsymbol{P}_\mathrm{ZF}$ and $\boldsymbol{F}_\mathrm{ZF}$ only exist, if $N_\mathrm{a} \geq K$.

1:	$\mathcal{K} \leftarrow \{1,\ldots,K\}$
	$\boldsymbol{G} \leftarrow \boldsymbol{H}$
	for $i = K,\ldots,1$:
4:	$\quad \boldsymbol{P} \leftarrow \boldsymbol{G}^{+}$
5:	$\quad k_i \leftarrow \underset{k\in\mathcal{K}}{\operatorname{argmin}} \\|\boldsymbol{P}\boldsymbol{e}_k\\|_2^2$
6:	$\quad \boldsymbol{p}_i \leftarrow \boldsymbol{P}\boldsymbol{e}_{k_i}$
7:	$\quad \boldsymbol{f}_i \leftarrow \boldsymbol{e}_{k_i} - \boldsymbol{H}\boldsymbol{p}_i$
	$\quad \mathcal{K} \leftarrow \mathcal{K}\backslash\{k_i\}$
	$\quad \boldsymbol{G} \leftarrow (\boldsymbol{1}_K - \boldsymbol{e}_{k_i}\boldsymbol{e}_{k_i}^{\mathrm{T}})\boldsymbol{G}$
10:	$\chi \leftarrow \sigma_s^2\\|\boldsymbol{p}_1\\|_2^2 + \sigma_v^2\sum_{i=2}^{K}\\|\boldsymbol{p}_i\\|_2^2$
	$\beta \leftarrow \sqrt{E_{\mathrm{tr}}/\chi}$
	$\boldsymbol{P} \leftarrow \beta[\boldsymbol{p}_1,\ldots,\boldsymbol{p}_K]$

TABLE I
FILTER AND ORDERING COMPUTATION FOR SPATIAL ZF-THP

The i-th column of $\boldsymbol{P}_{\mathrm{ZF}}$ is the i-th column of the pseudo-inverse of $\boldsymbol{S}_i\boldsymbol{H}$ weighted with β_{ZF}, whereas the i-th column of $\boldsymbol{F}_{\mathrm{ZF}}$ is the respective column of the pseudo-inverse weighted with $\boldsymbol{H}$ and subtracted from $\boldsymbol{e}_i$. Since $\boldsymbol{S}_i\boldsymbol{H} \in \mathbb{C}^{i\times K}$ consists of the first i rows of $\boldsymbol{H}$, we observe that the i-th column of $\boldsymbol{P}_{\mathrm{ZF}}$ only depends on the first i rows of $\boldsymbol{H}$ which suggests to compute the filters $\boldsymbol{P}_{\mathrm{ZF}}$ and $\boldsymbol{F}_{\mathrm{ZF}}$ successively. Therefore, we start with $i = K$ ($\boldsymbol{H}_K = \boldsymbol{H}$), end with $i = 1$, drop the last row of $\boldsymbol{H}_i$ to form the new channel matrix $\boldsymbol{H}_{i-1}$ after step i, set the i-th column of $\boldsymbol{P}_{\mathrm{ZF}}$ equal to the weighted last column of the pseudo inverse of $\boldsymbol{H}_i$, and compute the i-th column of $\boldsymbol{F}_{\mathrm{ZF}}$ accordingly.

Obviously, the order of precoding cannot be noticed by the receivers. Therefore, we can change the ordering to further minimize the cost function. Since the columns of $\boldsymbol{P}_{\mathrm{ZF}}$ are orthogonal, we get

$$\sigma_{\boldsymbol{\eta}}^2\beta_{\mathrm{ZF}}^{-2} = \frac{\sigma_{\boldsymbol{\eta}}^2}{E_{\mathrm{tr}}}\sum_{i=1}^{K}\sigma_{v_{k_i}}^2\left\|\left(\boldsymbol{\Pi}_i\boldsymbol{H}\right)^{+}\boldsymbol{e}_{k_i}\right\|_2^2. \qquad (6)$$

Note that k_i is the index of the receiver which is precoded at the i-th step and the projection matrix

$$\boldsymbol{\Pi}_i = \boldsymbol{1}_K - \sum_{j=i+1}^{K}\boldsymbol{e}_{k_j}\boldsymbol{e}_{k_j}^{\mathrm{T}}$$

sets the rows of the channel matrix belonging to receivers which will be precoded at a later step to zero. As the optimum ordering $\{k_1,\ldots,k_K\}$ can only be found by examining all $K!$ possibilities, we use following suboptimum criterion known from V-BLAST instead of minimizing Eqn. (6):

$$k_i = \underset{k\in\{1,\ldots,K\}\backslash\{k_{i+1},\ldots,k_K\}}{\operatorname{argmin}}\left\|(\boldsymbol{\Pi}_i\boldsymbol{H})^{+}\boldsymbol{e}_k\right\|_2^2,$$

for $i = K,\ldots,1$. Hence, we minimize the i-th summand for fixed $k_{i+1},\ldots,k_K$ instead of the whole sum. The resulting algorithm for joint order and filter computation is listed as pseudo code in Table I. The algorithm for ordered spatial THP is found in Table II. Note that we used $\sigma_s^2 = \mathrm{E}[|s_{k_1}[n]|^2]$ as variance of the first precoded signal $v_1[n]$ in line 10 of Table I, since $v_1[n]$ is simply $s_{k_1}[n]$ (see line 2 in Table II).

1:	for $i = 1,\ldots,K$:
2:	$\quad v_i[n] \leftarrow \mathrm{M}(s_{k_i}[n] + \boldsymbol{e}_{k_i}^{\mathrm{T}}\sum_{j=1}^{i-1}\boldsymbol{f}_j v_j[n])$
	$\boldsymbol{v}[n] = [v_1[n],\ldots,v_K[n]]^{\mathrm{T}}$
	$\boldsymbol{y}[n] = \boldsymbol{P}\boldsymbol{v}[n]$

TABLE II
ORDERED SPATIAL THP

The variance of the other outputs of the modulo operator is approximately equal to $\sigma_v^2 = \mathrm{E}[|v_i[n]|^2] = \tau^2/6$, $i = 2,\ldots,K$, under the assumption of uniformly distributed symbols (see e. g. [17]). When comparing the Tables I and II, we see that the columns of $\boldsymbol{P}_{\mathrm{ZF}}$ are applied in reverse order of their computation, e. g. the K-th column is computed first, but applied last. Because this property is in contrast to V-BLAST for DFE, we do not term the ordering algorithm V-BLAST for THP like in [31].

3.2 Wiener THP (WF-THP)

The optimization of the *linear Wiener transmit filter* (TxWF, e. g. [8], [9]) is found by dropping the zero-forcing constraint in Eqn. (3):

$$\begin{aligned}&\{\boldsymbol{P}'_{\mathrm{WF}},\beta'_{\mathrm{WF}}\} = \underset{\{\boldsymbol{P},\beta\}}{\operatorname{argmin}}\,\mathrm{E}\left[\left\|\boldsymbol{s}[n]-\beta^{-1}\tilde{\boldsymbol{s}}[n]\right\|_2^2\right]\\ &\text{s. t.: } \mathrm{E}\left[\|\boldsymbol{y}[n]\|_2^2\right] = E_{\mathrm{tr}},\end{aligned} \qquad (7)$$

whose solution reads as

$$\boldsymbol{P}'_{\mathrm{WF}} = \beta'_{\mathrm{WF}}\left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \xi_{\mathrm{WF}}\boldsymbol{1}_{N_{\mathrm{a}}}\right)^{-1}\boldsymbol{H}^{\mathrm{H}},$$

where $\xi_{\mathrm{WF}} = \sigma_{\boldsymbol{\eta}}^2/E_{\mathrm{tr}}$ and β'_{WF} is necessary to fulfill the transmit power constraint $\mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{P}^{\mathrm{H}}) = E_{\mathrm{tr}}$.

Similar to ZF-THP, the WF-THP optimization can be obtained from the TxWF optimization by replacing the desired signal $\boldsymbol{s}[n]$ by $\boldsymbol{d}[n] = \boldsymbol{v}[n] - \boldsymbol{F}\boldsymbol{v}[n]$ and constraining the feedback filter $\boldsymbol{F}$ to be lower triangular with zero main diagonal:

$$\begin{aligned}&\{\boldsymbol{P}_{\mathrm{WF}},\boldsymbol{F}_{\mathrm{WF}},\beta_{\mathrm{WF}}\} = \underset{\{\boldsymbol{P},\boldsymbol{F},\beta\}}{\operatorname{argmin}}\,\mathrm{E}\left[\|\boldsymbol{\varepsilon}[n]\|_2^2\right]\\ &\text{s. t.: } \mathrm{tr}\left(\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{v}}\boldsymbol{P}^{\mathrm{H}}\right) = E_{\mathrm{tr}} \text{ and}\\ &\qquad \boldsymbol{S}_i\boldsymbol{F}\boldsymbol{e}_i = \boldsymbol{0}_i,\ i = 1,\ldots,K,\end{aligned} \qquad (8)$$

where the error signal can be written as

$$\boldsymbol{\varepsilon}[n] = \boldsymbol{v}[n] - \boldsymbol{F}\boldsymbol{v}[n] - \beta^{-1}\boldsymbol{H}\boldsymbol{P}\boldsymbol{v}[n] - \beta^{-1}\boldsymbol{\eta}[n].$$

With the Lagrangian multiplier method, the result of above optimization can be found:

$$\boldsymbol{P}_{\mathrm{WF}} = \beta_{\mathrm{WF}}\sum_{i=1}^{K}\boldsymbol{A}_i^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}},$$

$$\boldsymbol{F}_{\mathrm{WF}} = \sum_{i=1}^{K}\left(\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i - \boldsymbol{1}_K\right)\boldsymbol{H}\boldsymbol{A}_i^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}},$$

and β_{WF} is defined by $\mathrm{tr}(\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{v}}\boldsymbol{P}^{\mathrm{H}}) = E_{\mathrm{tr}}$. Here, we used the abbreviation $\boldsymbol{A}_i = \boldsymbol{H}^{\mathrm{H}}\boldsymbol{S}_i^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{H} + \xi_{\mathrm{WF}}\boldsymbol{1}_{N_{\mathrm{a}}}$ and the scalar ξ_{WF} is again $\sigma_{\boldsymbol{\eta}}^2/E_{\mathrm{tr}}$.

1: $\mathcal{K} \leftarrow \{1,\ldots,K\}$
$\boldsymbol{G} \leftarrow \boldsymbol{H}$
for $i = K,\ldots,1$:
4: $\quad \boldsymbol{P} \leftarrow (\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}} + \xi_{\mathrm{WF}}\mathbf{1}_K)^{-1}$
5: $\quad k_i \leftarrow \operatorname*{argmin}_{k\in\mathcal{K}} \boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{P}\boldsymbol{e}_k$
6: $\quad \boldsymbol{p}_i \leftarrow \boldsymbol{G}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{e}_{k_i}$
7: $\quad \boldsymbol{f}_i \leftarrow (\boldsymbol{\Pi}_i - \mathbf{1}_K)\boldsymbol{H}\boldsymbol{p}_i$
$\quad \mathcal{K} \leftarrow \mathcal{K}\backslash\{k_i\}$
$\quad \boldsymbol{G} \leftarrow (\mathbf{1}_K - \boldsymbol{e}_{k_i}\boldsymbol{e}_{k_i}^{\mathrm{T}})\boldsymbol{G}$
$\chi \leftarrow \sigma_s^2\|\boldsymbol{p}_1\|_2^2 + \sigma_v^2\sum_{i=2}^{K}\|\boldsymbol{p}_i\|_2^2$
$\beta \leftarrow \sqrt{E_{\mathrm{tr}}/\chi}$
$\boldsymbol{P} \leftarrow \beta[\boldsymbol{p}_1,\ldots,\boldsymbol{p}_K]$

TABLE III
FILTER AND ORDERING COMPUTATION FOR SPATIAL WF-THP

We make the same observation as for ZF-THP: the i-th column of $\boldsymbol{P}_{\mathrm{WF}}$ only depends on the first i rows of $\boldsymbol{H}$ and the respective column of $\boldsymbol{F}_{\mathrm{WF}}$ directly follows from the i-th column of $\boldsymbol{P}_{\mathrm{WF}}$. Thus, we can use a similar successive algorithm for WF-THP as for ZF-THP. Plugging the optimum $\boldsymbol{P}_{\mathrm{WF}}$, $\boldsymbol{F}_{\mathrm{WF}}$, and β_{WF} into the cost function $\sigma_{\boldsymbol{\varepsilon}}^2 = \mathrm{E}[\|\boldsymbol{\varepsilon}[n]\|_2^2]$ yields:

$$\frac{\sigma_{\boldsymbol{\varepsilon}}^2}{\xi_{\mathrm{WF}}} = \sum_{i=1}^{K} \sigma_{v_{k_i}}^2 \boldsymbol{e}_{k_i}^{\mathrm{T}} \left(\boldsymbol{\Pi}_i\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\Pi}_i + \xi_{\mathrm{WF}}\mathbf{1}_K\right)^{-1} \boldsymbol{e}_{k_i}, \tag{9}$$

where we used the *matrix inversion lemma* (e. g. [33]), employed the already defined projection matrix $\boldsymbol{\Pi}_i$, and included the ordering index k_i. To avoid the examination of all $K!$ possibilities for ordering, we make the heuristic simplification to minimize each summand separately instead of the whole sum in Eqn. (9).

The resulting algorithm for joint WF-THP filter and ordering computation can be found in Table III which we obtained by changing the lines 4, 5, 6, and 7 of Table I. The resulting WF-THP filters are applied in the same manner as their ZF-THP counterparts. Thus, Table II is also valid for WF-THP.

4 Multiuser Space-Time THP

In the following, we focus on the design of *Wiener spatio-temporal THP* (WF-ST-THP) as the design of the zero-forcing variant is straightforward — only the constraint for an unbiased and interference-free estimate $\beta^{-1}\tilde{\boldsymbol{s}}[n]$ has to be included.

The TxWF for frequency selective channels follows from a MSE minimization similar to the TxWF for flat fading channels in Eqn. (7):

$$\{\check{\boldsymbol{P}}_{\mathrm{WF}}, \check{\beta}_{\mathrm{WF}}\} = \operatorname*{argmin}_{\{\boldsymbol{P},\beta\}} \sum_{k=1}^{K} \mathrm{E}\left[|\check{\varepsilon}_k[n]|^2\right] \quad \text{s.t.: } \mathrm{E}\left[\|\boldsymbol{y}[n]\|_2^2\right] = E_{\mathrm{tr}}, \tag{10}$$

where we used $\boldsymbol{P} = [\boldsymbol{p}_1,\ldots,\boldsymbol{p}_K] \in \mathbb{C}^{N_{\mathrm{a}}(L+1)\times K}$ and $\check{\varepsilon}_k[n] = s_k[n-\nu] - \beta^{-1}\tilde{s}_k[n]$. The latency time is

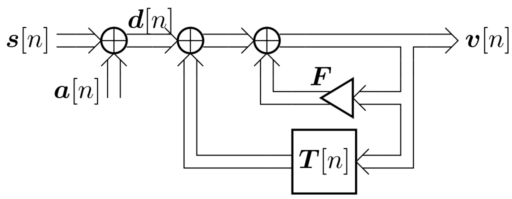

Fig. 4. Linear Representation of the Precoder

denoted by ν. The coefficients of the TxWF filter for the k-th data stream $s_k[n]$ can be expressed as

$$\check{\boldsymbol{p}}_{\mathrm{WF},k}^{\mathrm{T}} = \check{\beta}_{\mathrm{WF}}\boldsymbol{e}_{\nu+1}^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{H}} \left(\check{\boldsymbol{H}}\check{\boldsymbol{H}}^{\mathrm{H}} + \xi_{\mathrm{WF}}\mathbf{1}_{N_{\mathrm{a}}(L+1)}\right)^{-1},$$

with $\check{\boldsymbol{H}} = [\boldsymbol{H}_1,\ldots,\boldsymbol{H}_K] \in \mathbb{C}^{N_{\mathrm{a}}(L+1)\times K(L+Q+1)}$ and $\xi_{\mathrm{WF}} = \sigma_{\boldsymbol{\eta}}^2/E_{\mathrm{tr}}$. Note that the TxZF which is the TxWF after applying the matrix inversion lemma and setting $\xi = 0$ only exists, if $N_{\mathrm{a}}(L+1) \geq K(L+Q+1)$.

Like for flat fading channels, we have to employ the signal $d_k[n-\nu] = \boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{d}[n-\nu]$ (see Fig. 4) as desired signal instead of $s_k[n-\nu]$ to get the THP optimization. Additionally, we have to constrain the structure of the spatial feedback filter $\boldsymbol{F}$:

$$\{\boldsymbol{P}_{\mathrm{WF}}, \boldsymbol{F}_{\mathrm{WF}}, \boldsymbol{T}_{\mathrm{WF}}, \beta_{\mathrm{WF}}\} = \operatorname*{argmin}_{\{\boldsymbol{P},\boldsymbol{F},\boldsymbol{T},\beta\}} \sum_{k=1}^{K} \sigma_{\varepsilon_k}^2 \quad \text{s.t.: } \mathrm{E}\left[\|\boldsymbol{y}[n]\|_2^2\right] = E_{\mathrm{tr}} \text{ and } \boldsymbol{F}: \text{ lower triangular, zero main diagonal.} \tag{11}$$

Here, $\sigma_{\varepsilon_k}^2$ is the variance of the error for the k-th data stream which can be expressed as

$$\varepsilon_k[n] = d_k[n-\nu] - \beta^{-1}\tilde{s}_k[n]$$

and as can be seen in Fig. 4, the signal $d_k[n-\nu]$ is

$$\boldsymbol{e}_{\nu+1}^{\mathrm{T}}\boldsymbol{v}_k[n] - \boldsymbol{e}_k^{\mathrm{T}}\sum_{i=1}^{K}\left(\boldsymbol{F}\boldsymbol{e}_i\boldsymbol{e}_{\nu+1}^{\mathrm{T}} + \boldsymbol{T}_i\boldsymbol{S}^{(\nu)}\right)\boldsymbol{v}_i[n],$$

where we collected the coefficients of the temporal feedback filter $\boldsymbol{T}[n] = \sum_{r=1}^{L+Q-\nu}\tilde{\boldsymbol{T}}_r\delta[n-r]$ for the i-th THP stream $v_i[n]$ in

$$\boldsymbol{T}_i = \left[\tilde{\boldsymbol{T}}_1\boldsymbol{e}_i,\ldots,\tilde{\boldsymbol{T}}_{L+Q-\nu}\boldsymbol{e}_i\right] \in \mathbb{C}^{K\times L+Q-\nu}$$

and in the optimization of Eqn. (11), we utilized the abbreviation $\boldsymbol{T} = [\boldsymbol{T}_1,\ldots,\boldsymbol{T}_K] \in \mathbb{C}^{K\times K(L+Q-\nu)}$. Moreover, we introduced the special selection matrix $\boldsymbol{S}^{(\nu)} = \boldsymbol{S}_{(\nu+1,L+Q-\nu,\nu+1)} \in \{0,1\}^{L+Q+\nu\times L+Q+1}$ to cut out the last $L+Q-\nu$ entries of $\boldsymbol{v}_i[n]$.

To solve the optimization in Eqn. (11) we divide the constraint on the structure of $\boldsymbol{F}$ into K constraints for the K columns of $\boldsymbol{F}$ like in the flat fading case (cf. Eqn. 8). With Lagrangian multipliers we easily find expressions for the feedback filters as functions of the

1:	$\nu \leftarrow \underset{\nu'}{\operatorname{argmin}} \sum_{i=1}^{K} \boldsymbol{e}_i^{(\nu'),\mathrm{T}} \boldsymbol{B}_K^{(\nu'),-1} \boldsymbol{e}_i^{(\nu')}$
	$\mathcal{K} \leftarrow \{1, \ldots, K\}$
	$\boldsymbol{G} \leftarrow \check{\boldsymbol{H}} \boldsymbol{\Pi}_K^{(\nu)}$
	for $i = K, \ldots, 1$:
5:	$\quad \boldsymbol{P} \leftarrow (\boldsymbol{G}^{\mathrm{H}} \boldsymbol{G} + \xi_{\mathrm{WF}} \mathbf{1}_{K(L+Q+1)})^{-1}$
6:	$\quad k_i \leftarrow \underset{k \in \mathcal{K}}{\operatorname{argmin}} \, \boldsymbol{e}_k^{(\nu),\mathrm{T}} \boldsymbol{P} \boldsymbol{e}_k^{(\nu)}$
7:	$\quad \boldsymbol{p}_i^{\mathrm{T}} \leftarrow \boldsymbol{e}_{k_i}^{(\nu),\mathrm{T}} \boldsymbol{P} \boldsymbol{G}^{\mathrm{H}}$
	$\quad \boldsymbol{f}_i \leftarrow (\boldsymbol{\Pi}_i - \mathbf{1}_K) \sum_{j=1}^{K} \boldsymbol{e}_j \boldsymbol{p}_i^{\mathrm{T}} \check{\boldsymbol{H}} \boldsymbol{e}_j^{(\nu)}$
	$\quad \boldsymbol{T}_i \leftarrow -\sum_{j=1}^{K} \boldsymbol{e}_j \boldsymbol{p}_i^{\mathrm{T}} \boldsymbol{H}_j \boldsymbol{S}^{(\nu),\mathrm{T}}$
	$\quad \mathcal{K} \leftarrow \mathcal{K} \backslash \{k_i\}$
	$\quad \boldsymbol{G} \leftarrow \boldsymbol{G}(\mathbf{1}_{K(L+Q+1)} - \boldsymbol{e}_{k_i}^{(\nu)} \boldsymbol{e}_{k_i}^{(\nu),\mathrm{T}})$
	$\chi \leftarrow \sigma_v^2 \sum_{i=1}^{K} \lVert \boldsymbol{p}_i \rVert_2^2$
	$\beta \leftarrow \sqrt{E_{\mathrm{tr}}/\chi}$
	for $i = 0, \ldots, L$:
	$\quad \boldsymbol{P}_i = \beta \boldsymbol{S}_{(N_{\mathrm{a}} i, N_{\mathrm{a}}, N_{\mathrm{a}} L)} [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_K]$
	for $i = 1, \ldots, L + Q - \nu$:
	$\quad \tilde{\boldsymbol{T}}_i = [\boldsymbol{T}_1 \boldsymbol{e}_i, \ldots, \boldsymbol{T}_K \boldsymbol{e}_i]$

TABLE IV

FILTER AND ORDERING COMPUTATION FOR WIENER ST-THP

feedforward filters $\boldsymbol{p}_{\mathrm{WF},1}^{\mathrm{T}}, \ldots, \boldsymbol{p}_{\mathrm{WF},K}^{\mathrm{T}}$:

$$\boldsymbol{T}_{\mathrm{WF},i} = -\sum_{k=1}^{K} \beta_{\mathrm{WF}}^{-1} \boldsymbol{e}_k \boldsymbol{p}_{\mathrm{WF},i}^{\mathrm{T}} \boldsymbol{H}_k \boldsymbol{S}^{(\nu),\mathrm{T}} \quad \text{and}$$

$$\boldsymbol{F}_{\mathrm{WF}} = \sum_{i,k=1}^{K} \left(\boldsymbol{S}_i^{\mathrm{T}} \boldsymbol{S}_i - \mathbf{1}_K \right) \beta_{\mathrm{WF}}^{-1} \boldsymbol{e}_k \boldsymbol{p}_{\mathrm{WF},i}^{\mathrm{T}} \boldsymbol{H}_k \boldsymbol{e}_{\nu+1} \boldsymbol{e}_i^{\mathrm{T}}.$$

With the projector $\boldsymbol{S}_{k,i}^{(\nu)} \boldsymbol{S}_{k,i}^{(\nu),\mathrm{T}}$ which is defined as

$$\mathbf{1}_{L+Q+1} + \left(\lVert \boldsymbol{S}_i \boldsymbol{e}_k \rVert_2^2 - 1 \right) \boldsymbol{e}_{\nu+1} \boldsymbol{e}_{\nu+1}^{\mathrm{T}} - \boldsymbol{S}^{(\nu),\mathrm{T}} \boldsymbol{S}^{(\nu)},$$

and $\boldsymbol{A}_i = \sum_{k=1}^{K} \boldsymbol{H}_k \boldsymbol{S}_{k,i}^{(\nu)} \boldsymbol{S}_{k,i}^{(\nu),\mathrm{T}} \boldsymbol{H}_k^{\mathrm{H}} + \xi_{\mathrm{WF}} \mathbf{1}_{N_{\mathrm{a}}(L+1)}$, the feedforward filters can be written as

$$\boldsymbol{p}_{\mathrm{WF},i}^{\mathrm{T}} = \beta_{\mathrm{WF}} \boldsymbol{e}_{\nu+1}^{\mathrm{T}} \boldsymbol{S}_{i,i}^{(\nu)} \boldsymbol{S}_{i,i}^{(\nu),\mathrm{T}} \boldsymbol{H}_i^{\mathrm{H}} \boldsymbol{A}_i^{-1}, \; i = 1, \ldots, K.$$

The scalar β_{WF} is defined by the transmit power constraint, i. e. $\sigma_v^2 \sum_{i=1}^{K} \lVert \boldsymbol{p}_{\mathrm{WF},i} \rVert_2^2 = E_{\mathrm{tr}}$.

We conclude from above results that the i-th feedforward filter $\boldsymbol{p}_{\mathrm{WF},i}^{\mathrm{T}}$ only depends on the first $\nu + 1$ columns of $\boldsymbol{H}_1, \ldots, \boldsymbol{H}_i$ and the first ν columns of $\boldsymbol{H}_{i+1}, \ldots, \boldsymbol{H}_K$. Moreover, $\boldsymbol{T}_{\mathrm{WF},i}$ and also the i-th column of $\boldsymbol{F}_{\mathrm{WF}}$ only depend on $\boldsymbol{p}_{\mathrm{WF},i}$. This observation suggests to compute the filters successively like in the flat fading case: start with the first $\nu+1$ columns of all channel matrices to compute $\boldsymbol{p}_{\mathrm{WF},K}^{\mathrm{T}}$, drop the $\nu + 1$-th column of $\boldsymbol{H}_i$ after computing $\boldsymbol{p}_{\mathrm{WF},i}^{\mathrm{T}}$, and base $\boldsymbol{p}_{\mathrm{WF},i-1}^{\mathrm{T}}$ on the remaining columns of the channel matrices.

Again, we can change the order of precoding as done for spatial THP, but we also can optimize the latency time to further minimize the MSE. To this end, we plug the filters for fixed ordering $\{k_1, \ldots, k_K\}$ and fixed latency time ν into the MSE $\sigma_\varepsilon^2 = \sum_{i=1}^{K} \sigma_{\varepsilon_i}^2$:

$$\sigma_\varepsilon^2 = \sigma_v^2 \xi_{\mathrm{WF}} \sum_{i=1}^{K} \boldsymbol{e}_{k_i}^{(\nu),\mathrm{T}} \boldsymbol{B}_i^{(\nu),-1} \boldsymbol{e}_{k_i}^{(\nu)}, \qquad (12)$$

where $\boldsymbol{B}_i^{(\nu)} = \boldsymbol{\Pi}_i^{(\nu)} \check{\boldsymbol{H}}^{\mathrm{H}} \check{\boldsymbol{H}} \boldsymbol{\Pi}_i^{(\nu)} + \xi_{\mathrm{WF}} \mathbf{1}_{K(L+Q+1)}$ and $\boldsymbol{e}_k^{(\nu)} = \boldsymbol{e}_{(L+Q+1)(k-1)+\nu+1}$. The k-th block diagonal entry of the projector

$$\boldsymbol{\Pi}_i^{(\nu)} = \operatorname{blockdiag} \left(\boldsymbol{\Pi}_{1,i}^{(\nu)}, \ldots, \boldsymbol{\Pi}_{K,i}^{(\nu)} \right)$$

sets the last $L + Q + 1 - \nu$ or $L + Q - \nu$ columns of the respective channel matrix $\boldsymbol{H}_k$ to zero:

$$\boldsymbol{\Pi}_{k,i}^{(\nu)} = \mathbf{1}_{L+Q+1} - \boldsymbol{S}^{(\nu),\mathrm{T}} \boldsymbol{S}^{(\nu)} - \sum_{j=i+1}^{K} \boldsymbol{e}_{\nu+1} \boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{e}_{k_j} \boldsymbol{e}_{\nu+1}^{\mathrm{T}}.$$

Note that the sum vanishes for $i = K$ and therefore, $\boldsymbol{\Pi}_{k,K}^{(\nu)} = \mathbf{1}_{L+Q+1} - \boldsymbol{S}^{(\nu),\mathrm{T}} \boldsymbol{S}^{(\nu)}$.

Instead of the minimization of the MSE in Eqn. (12) which requires the examination of $K!(L + Q + 1)$ possible ordering and latency time combinations we split the problem into two subproblems. First, we minimize the MSE by choosing the latency time under the assumption that no spatial THP is applied. Second, we use the heuristic successive algorithm which we employed for spatial THP over flat fading channels to find the ordering. The resulting algorithm for joint order and filter computation of WF-ST-THP can be found in Table IV. The ST-THP procedure is outlined in Table V.

Zero-forcing spatio-temporal THP (ZF-ST-THP) can be computed with the same algorithm as in Table IV, only lines 5, 6, and 7 have to be changed (see Table VI).

1:	for $i = 1, \ldots, K$:
	$\quad v_i[n] \leftarrow \mathrm{M}(s_{k_i}[n] + \boldsymbol{e}_{k_i}^{\mathrm{T}} \sum_{j=1}^{i-1} \boldsymbol{f}_j v_j[n])$
	$\quad v_i[n] \leftarrow \mathrm{M}(v_i[n] + \boldsymbol{e}_{k_i}^{\mathrm{T}} \sum_{j=1}^{L+Q-\nu} \tilde{\boldsymbol{T}}_j \boldsymbol{v}[n-j])$
	$\boldsymbol{v}[n] = [v_1[n], \ldots, v_K[n]]^{\mathrm{T}}$
	$\boldsymbol{y}[n] = \sum_{\ell=0}^{L} \boldsymbol{P}_\ell \boldsymbol{v}[n-\ell]$

TABLE V

ORDERED SPATIO-TEMPORAL THP

5:	$\boldsymbol{P} \leftarrow \boldsymbol{G}^{+}$
6:	$k_i \leftarrow \underset{k \in \mathcal{K}}{\operatorname{argmin}} \lVert \boldsymbol{P}^{\mathrm{T}} \boldsymbol{e}_k^{(\nu)} \rVert_2^2$
7:	$\boldsymbol{p}_i^{\mathrm{T}} \leftarrow \boldsymbol{e}_{k_i}^{(\nu),\mathrm{T}} \boldsymbol{P}$

TABLE VI

FILTER AND ORDERING COMPUTATION FOR ZF-ST-THP

5 Simulation Results

We present *uncoded bit error ratio* (uncoded BER) results comparing the linear transmit filters with the different MMSE approaches to THP. We use following (transmit) SNR definition: SNR $= 10 \log_{10}(E_{\mathrm{tr}}/\sigma_{\boldsymbol{\eta}}^2)$. In all scenarios, we assume that the transmitter has full channel state information.

5.1 Flat Fading Vector Channels

In Fig. 5, we present results for a system with $N_{\mathrm{a}} = 4$ antenna elements at the transmitter and $K = 3$ receivers

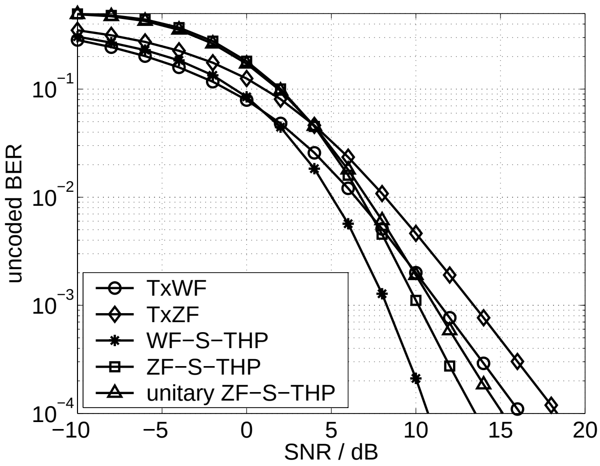

Fig. 5. QPSK Transmission over Frequency Flat Channels, 3 Users, 4 Tx Antennas

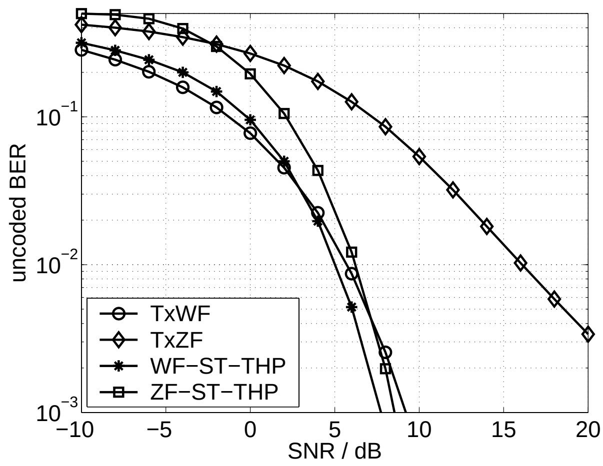

Fig. 6. QPSK Transmission over Frequency Selective Channels with Exponential Power Delay Profile

in a spatially uncorrelated flat Rayleigh fading scenario, using QPSK modulation. The results are the mean of 10000 channel realizations and 20 symbol vectors were transmitted for each realization. We observe that the linear transmit filters outperform their THP counterparts for low SNR values. This behaviour can be explained by the fact that in THP systems, even though QPSK is used, the decision depends on phase and amplitude information, due to the modulo operation at the receiver. For higher SNR values, *spatial THP* (S-THP) provides superior performance. The *Wiener S-THP* (WF-S-THP) solution performs significantly better than both the *zero-forcing S-THP* (ZF-S-THP) presented in this paper and the zero-forcing approach with unitary feedforward filter presented by Fischer et al. in [18]. Moreover, the ZF-S-THP solution derived from the TxZF optimization leads to better results for high SNR than the unitary ZF-S-THP although the unitary ZF-S-THP allows different scalar weighting at the receivers and has therefore potentially more degrees of freedom. We can conclude that the weighting for unitary ZF-S-THP arising from the LQ-factorization as proposed in [18] is suboptimum with respect to overall BER, since we used the same ordering as for ZF-S-THP.

5.2 Frequency Selective Vector Channels

In Fig. 6, we show results for a system with $N_\mathrm{a} = 4$ antenna elements at the transmitter, $K = 3$ receivers, and QPSK modulation. The frequency selective channels have $Q + 1 = 6$ paths and an exponential power delay profile. Again, we assume uncorrelated Rayleigh fading. The results are the mean of 3000 channel realizations, 100 symbol vectors were transmitted per channel realization. The linear transmit filters TxWF and TxZF are of order $L = 15$, but the order of the feedforward filter of *spatio-temporal THP* (ST-THP) is set to $L = 7$ taking into account the coefficients of the THP systems' feedback filter in order to end up with a fair comparison between the different systems.

In Fig. 6, we can observe that the linear TxZF is clearly outperformed by the *zero-forcing ST-THP* (ZF-ST-THP). In contrast to the ZF solutions, in the uncoded BER region between 10^{-1} and 10^{-2}, the *Wiener ST-THP* (WF-ST-THP) and the TxWF show almost identical performance (difference is less than $1\,\mathrm{dB}$). For high SNR values, the WF-ST-THP and the ZF-ST-THP both outperform the TxWF.

For the results in Fig. 7, we only changed the modulation to 16QAM, leaving all other simulation parameters unchanged. The results highlight the impact of the chosen modulation on the THP systems, which was already discussed in Subsection 5.1. Clearly, the higher the order of the modulation alphabet, the better the performance of the THP systems compared to the linear transmit filters. From Fig. 7, it can be seen that in the given scenario, the WF-ST-THP solution provides optimum performance, regardless of the SNR region.

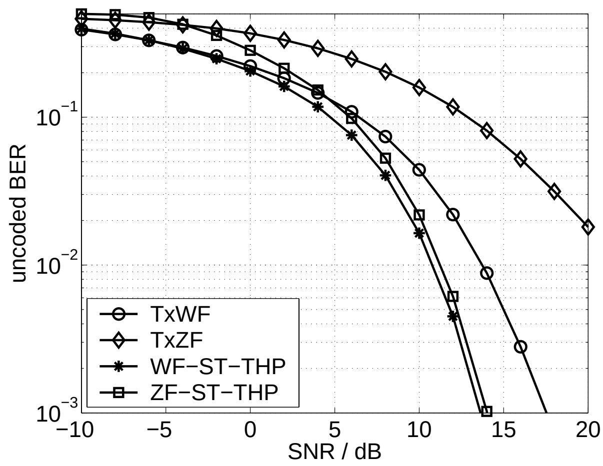

Fig. 7. 16QAM Transmission over Frequency Selective Channels with Exponential Power Delay Profile

Fig. 8 shows results for a channel of order $Q = 4$ with a uniform power delay profile. For the linear

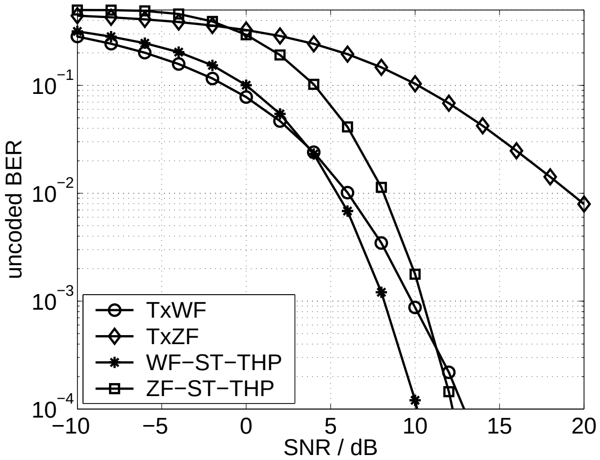

Fig. 8. QPSK Transmission over Frequency Selective Channels with Uniform Power Delay Profile

transmit filters, we set $L = 11$, for the feedforward filters of the THP-systems we chose $L = 5$. It can be observed that in case of a uniform power delay profile, the difference between the ZF-ST-THP solution and the WF-ST-THP solution is larger than for the exponential power delay profile.

References

[1] M. R. Gibbard and A. B. Sesay, "Asymmetric Signal Processing for Indoor Wireless LAN's," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 6, pp. 2053–2064, November 1999.

[2] S. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.

[3] 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.

[4] G. Montalbano, I. Ghauri, and D. T. M. Slock, "Spatio-Temporal Array Processing for CDMA/SDMA Downlink Transmission," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 1337–1341, November 1998.

[5] 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 2000*, vol. 1, pp. 1–5, September 2000.

[6] M. Joham and W. Utschick, "Downlink Processing for Mitigation of Intracell Interference in DS-CDMA Systems," in *Proc. ISSSTA 2000*, vol. 1, pp. 15–19, September 2000.

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

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

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

[10] A. Noll Barreto and G. Fettweis, "Joint Signal Precoding in the Downlink of Spread Spectrum Systems," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 511–518, May 2003.

[11] M. Joham, R. Irmer, S. Berger, G. Fettweis, and W. Utschick, "Linear Precoding Approaches for the TDD DS-CDMA Downlink," in *Proc. WPMC'03*, vol. 3, pp. 323–327, October 2003.

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

[13] H. Harashima and H. Miyakawa, "Matched-Transmission Technique for Channels With Intersymbol Interference," *IEEE Transactions on Communications*, vol. 20, no. 4, pp. 774–780, August 1972.

[14] P. Monsen, "Feedback Equalization for Fading Dispersive Channels," *IEEE Transactions on Information Theory*, vol. IT-17, no. 1, pp. 56–64, January 1971.

[15] G. D. Forney, Jr. and M. V. Eyuboğlu, "Combined Equalization and Coding Using Precoding," *IEEE Communications Magazine*, vol. 29, no. 12, pp. 25–34, December 1991.

[16] G. Ginis and J. M. Cioffi, "A Multi-user Precoding Scheme achieving Crosstalk Cancellation with Application to DSL Systems," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, vol. 2, pp. 1627–1631, October 2000.

[17] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Space-Time Transmission using Tomlinson-Harashima Precoding," in *Proc. 4th ITG Conference on Source and Channel Coding*, pp. 139–147, January 2002.

[18] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "MIMO Precoding for Decentralized Receivers," in *Proc. ISIT 2002*, p. 496, June/July 2002.

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

[20] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Tomlinson-Harashima Precoding in Space-Time Transmission for Low-Rate Backward Channel," in *Proc. 2002 International Zurich Seminar on Broadband Communications*, pp. 7–1–7–6, February 2002.

[21] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Linear and non-linear preequalization/equalization for MIMO systems with long-term channel state information at the transmitter," *Accepted for publication in IEEE Transactions on Wireless Communications*, 2003.

[22] 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 2001*, pp. 243–248, September 2001.

[23] J. G. Proakis, *Digital Communications*, McGraw-Hill, Inc., 1995.

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

[25] 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'01*, vol. 1, pp. 32–37, March 2001.

[26] F. A. Dietrich, R. Hunger, M. Joham, and W. Utschick, "Robust Transmit Wiener Filter for Time Division Duplex Systems," *Accepted for presentation at ISSPIT 2003*, December 2003.

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

[28] M. Schubert and H. Boche, "Joint 'Dirty Paper' Pre-Coding and Downlink Beamforming," in *Proc. ISSSTA 2002*, vol. 2, pp. 536–540, September 2002.

[29] A. Voulgarelis, M. Joham, and W. Utschick, "Space-Time Equalization Based on V-BLAST and DFE for Frequency Selective MIMO Channels," in *Proc. ICASSP 2003*, vol. IV, pp. 381–384, April 2003.

[30] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel," in *Proc. ISSSE 98*, pp. 295–300, September 1998.

[31] C. Windpassinger, T. Vencel, and R. F. H. Fischer, "Precoding and Loading for BLAST-like Systems," in *Proc. ICC'03*, vol. 5, pp. 3061–3065, May 2003.

[32] M. Joham, W. Utschick, and J. A. Nossek, "Latency Time Optimization for FIR and Block Transmit Filters," in *Proc. ISSPA 2003*, vol. 1, pp. 273–276, July 2003.

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