20 Ordered spatial Tomlinson-Harashima precoding

Michael Joham and Wolfgang Utschick

20.1. Introduction

Similar to receive processing (e.g., [1]), linear transmit processing (see, e.g., [2]) can be outperformed by nonlinear transmit processing. Peel et al. [3, 4] recently proposed the *minimization* of the necessary *transmit energy* for every vector symbol of a flat fading MIMO system, when assuming that the receiver applies a *modulo operation*. The necessary search for the minimum length vector of a lattice has exponential complexity (see [5] for a near-optimum variant with $O(B^4)$, where B denotes the number of scalar data streams). Additionally, the transmit signal has to be weighted with a data-dependent scalar to meet the transmit energy constraint [3, 4], if the computation of the average transmit energy (a search of exponential complexity) has to be avoided. Consequently, we do not consider this approach. Another type of nonlinear transmit processing is the *minimization* of the *bit error probability* by choosing the appropriate transmit signal [6, 7]. As the resulting nonconvex optimization can only be solved analytically for special channel matrices [7], we will not investigate this approach due to its prohibitive complexity. In [8], Fischer et al. decomposed the real-valued representation of the channel matrix into the product of a real-valued matrix and an integer-valued matrix motivated by the promising result of Yao et al. [9] who included lattice reduction techniques in the receive filter design and achieved the same diversity order as the maximum likelihood detector. The precoder of [8] which nearly reaches the diversity order of the maximum likelihood detector only equalizes the real-valued matrix of the decomposition, since the modulo operation at the receiver removes the interference caused by the integer-valued matrix. However, the channel matrix decomposition was found by Monte-Carlo search in [8], because an algorithm to obtain the decomposition is an open problem.

An alternative type of nonlinear transmit processing is *Tomlinson-Harashima precoding* (THP) which is closely related to the *decision feedback equalizer* proposed by Austin [10] (see also [11, 12, 13, 14]). Whereas DFE feeds back already *quantized* symbols, the already *transmitted* symbols are fed back in a THP system

and *modulo operations* are applied in the transmitter and the receiver(s). The principle of THP was introduced by Tomlinson [15] and Harashima et al. [16] nearly at the same time. They applied THP to a SISO system without adaptive receive filter in order to suppress the ISI caused by the frequency selectivity of the channel, since the recursive filter necessary to equalize the channel can be unstable (see also [17]). Gibbard et al. [18] proposed an asymmetric transmission, where for the first link, the receiver performs DFE and in the second link, the transmitter applies THP. This approach leads to a simplification and a lower power consumption of the device(s) at one side of the link (e.g., the mobile terminal(s)). Spatial THP without ordering for flat fading MIMO channels was proposed by Ginis et al. in [19] and Fischer et al. in [20]. Whereas Ginis et al. included a feedforward filter at the transmitter and assumed a receive filter which is a diagonal matrix, Fischer et al. investigated a system with the feedforward filter at the receiver. THP was used for a DS-CDMA system by Fischer et al. in [21] and also by Liu et al. in [22]. THP with partial channel state information at the transmitter has only been investigated by Fischer et al. [23] and Simeone et al. [24], all other publications assumed full knowledge of the channel at the transmitter. Joham et al. presented the necessary optimizations for THP with FIR feedforward and feedback filters for frequency selective MIMO channels in [25, 26]. In [27], Fischer et al. designed THP for frequency selective MIMO channels with IIR feedforward filter by applying a spectral factorization of the channel transfer function.

We restrict ourselves to systems with nondispersive channels and noncooperative receivers (e.g., mobiles in the downlink), that is, the signals of the different receivers cannot be jointly transformed. Therefore, the feedforward filter has to be located at the transmitter and the receivers only apply scalar weightings. For simplicity, we make the additional assumption that all receivers use the same scalar weight. The examined THP approaches are based on full channel state information without estimation errors (for a robust design taking into account the estimation errors, see [28]). Note that the channel can only be fully equalized by the transmitter in a system with noncooperative receivers. Thus, receive processing is only a suboptimum approach for such systems.

Contrary to most other contributions on THP, we base the THP filter design on an optimization. Since the THP optimizations are an extension of the well-known optimizations for linear transmit filters, we first review the linear *transmit zero-forcing filter* (TxZF) and the linear *transmit Wiener filter* (TxWF) in Section 20.2. With the linear representation of THP introduced in Section 20.3, we are able to formulate the optimizations for *zero-forcing THP* (ZF-THP) and *Wiener THP* (WF-THP) including not only the THP filters but also the ordering in Section 20.4. Thus, we obtain the algorithms for the optimum orderings for the two THP types and since these algorithms are too complex, we also present suboptimum ordering algorithms closely related to the *vertical Bell Laboratories Layered Space Time* (V-BLAST) algorithm well known for spatial DFE [29]. The simulation results in Section 20.5 reveal that the Wiener THP clearly outperforms the state-of-the-art zero-forcing THP approaches.

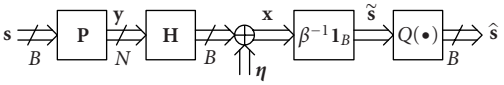

FIGURE 20.1. System with linear transmit filter.

Notation. Vectors and matrices are denoted by lower-case bold and capital bold letters, respectively. We use $\mathrm{E}[\bullet]$, "$*$," $(\bullet)^*$, $(\bullet)^{\mathrm{T}}$, and $(\bullet)^{\mathrm{H}}$ for expectation, convolution, complex conjugation, transposition, and conjugate transposition, respectively. The pseudoinverse is denoted by $(\bullet)^+$. All random variables are assumed to be zero mean and stationary. The variance of the scalar random variable y is denoted by $\sigma_y^2 = \mathrm{E}[|y|^2]$ and the covariance matrix of the vector random variable $\mathbf{x}$ by $\mathbf{R}_{\mathbf{x}} = \mathrm{E}[\mathbf{x}\mathbf{x}^{\mathrm{H}}]$. 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 nth column is $\mathbf{e}_n \in \{0,1\}^N$.

20.2. Linear transmit filters

In a system with linear transmit filter, the data signal $\mathbf{s} = [s_1, \ldots, s_B] \in \mathbb{C}^B$ comprising the symbols for the B noncooperative receivers is passed through the linear precoder $\mathbf{P} \in \mathbb{C}^{N\times B}$ to form the transmit signal of the N transmit antenna elements

$$\mathbf{y} = \mathbf{P}\mathbf{s} \in \mathbb{C}^N. \tag{20.1}$$

The transmit filter has to be designed to satisfy the transmit energy constraint, that is,

$$\mathrm{E}\left[\|\mathbf{P}\mathbf{s}\|_2^2\right] = \mathrm{tr}\left(\mathbf{P}\mathbf{R}_{\mathbf{s}}\mathbf{P}^{\mathrm{H}}\right) = E_{\mathrm{tr}}. \tag{20.2}$$

After propagation over the channel $\mathbf{H} \in \mathbb{C}^{B\times N}$ and perturbation by the noise $\eta \in \mathbb{C}^B$, the received signal $\mathbf{x} \in \mathbb{C}^B$ is weighted with the scalar β^{-1} to form the estimate (see Figure 20.1)

$$\widetilde{\mathbf{s}} = \beta^{-1}\mathbf{H}\mathbf{P}\mathbf{s} + \beta^{-1}\boldsymbol{\eta} \in \mathbb{C}^B. \tag{20.3}$$

Note that the scalar $\beta^{-1} \in \mathbb{R}_+$ at the receiver is necessary to correct the amplitude of the desired signal part in the estimate $\widetilde{\mathbf{s}}$, since the transmitter only has a limited transmit power E_{tr}. The estimate $\widetilde{\mathbf{s}}$ is the input of the nearest-neighbor quantizer $Q(\bullet)$, whose output is denoted by $\widehat{\mathbf{s}} \in \mathbb{C}^B$.

The TxZF minimizes the *mean square error* (MSE) under the transmit energy constraint (20.2) together with the constraint of full interference suppression and unbiasedness (see, e.g., [2]):

$$\{\mathbf{P}_{\mathrm{ZF}}, \beta_{\mathrm{ZF}}\} = \underset{\{\mathbf{P},\beta\}}{\operatorname{argmin}} \mathrm{E}\left[\|\mathbf{s} - \widetilde{\mathbf{s}}\|_2^2\right] \quad \text{s.t. } \mathrm{E}\left[\|\mathbf{P}\mathbf{s}\|_2^2\right] = E_{\mathrm{tr}},\ \widetilde{\mathbf{s}}|_{\boldsymbol{\eta}=\mathbf{0}} = \mathbf{s}. \tag{20.4}$$

The solution of the above optimization can be obtained with the method of Lagrangian multipliers (e.g., [30]) and reads as

$$\begin{aligned}\mathbf{P}_{\mathrm{ZF}} &= \beta_{\mathrm{ZF}}\mathbf{H}^{\mathrm{H}}\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1} \in \mathbb{C}^{N\times B},\\ \beta_{\mathrm{ZF}} &= \sqrt{\frac{E_{\mathrm{tr}}}{\mathrm{tr}\left(\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\mathbf{R}_{\mathbf{s}}\right)}} \in \mathbb{R}_{+}.\end{aligned} \tag{20.5}$$

Note that the MSE of the TxZF is simply the noise portion at the quantizer input due to the second constraint of (20.4), that is, $\mathrm{E}[\|\mathbf{s}-\widetilde{\mathbf{s}}\|_2^2] = \beta_{\mathrm{ZF}}^{-2}\,\mathrm{tr}(\mathbf{R}_{\boldsymbol{\eta}})$. Hence, we can expect that the performance of the TxZF is poor, when the channel matrix $\mathbf{H}$ is ill conditioned (see, e.g., [31]), since the weight β_{ZF} is small in this case.

To reduce this *noise enhancement* of the TxZF for ill-conditioned $\mathbf{H}$, we have to drop the second constraint in (20.4) and end up with the TxWF optimization [32, 33, 2]:

$$\left\{\mathbf{P}_{\mathrm{WF}},\beta_{\mathrm{WF}}\right\} = \underset{\{\mathbf{P},\beta\}}{\mathrm{argmin}}\,\mathrm{E}\left[\|\mathbf{s}-\widetilde{\mathbf{s}}\|_2^2\right] \quad \text{s.t. } \mathrm{E}\left[\|\mathbf{P}\mathbf{s}\|_2^2\right] = E_{\mathrm{tr}}. \tag{20.6}$$

As shown in [2], we get with Lagrangian multipliers:

$$\begin{aligned}\mathbf{P}_{\mathrm{WF}} &= \beta_{\mathrm{WF}}\left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \frac{\mathrm{tr}\left(\mathbf{R}_{\boldsymbol{\eta}}\right)}{E_{\mathrm{tr}}}\mathbf{1}_N\right)^{-1}\mathbf{H}^{\mathrm{H}} \in \mathbb{C}^{N\times B},\\ \beta_{\mathrm{WF}} &= \sqrt{\frac{E_{\mathrm{tr}}}{\mathrm{tr}\left(\left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \left(\mathrm{tr}\left(\mathbf{R}_{\boldsymbol{\eta}}\right)/E_{\mathrm{tr}}\right)\mathbf{1}_N\right)^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{R}_{\mathbf{s}}\mathbf{H}\right)}} \in \mathbb{R}_{+}.\end{aligned} \tag{20.7}$$

Obviously, the TxWF leads to a smaller MSE than the TxZF, since the TxZF minimizes the MSE under an additional constraint. Moreover, the TxZF is independent of the properties of the noise, whereas the TxWF takes into account the mean noise power $\mathrm{tr}(\mathbf{R}_{\boldsymbol{\eta}})$, because it depends on the *signal-to-noise ratio* (SNR) $E_S/N_0 = E_{\mathrm{tr}}/\mathrm{tr}(\mathbf{R}_{\boldsymbol{\eta}})$, that is, the ratio of the average energy per transmitted scalar symbol over the average noise power per receive antenna. Thus, we can expect that the TxZF is outperformed by the TxWF.

20.3. System model for Tomlinson-Harashima precoding

Instead of directly applying the feedforward filter $\mathbf{P} \in \mathbb{C}^{N\times B}$ as in the case of linear transmit processing (see previous subsection), the data signal $\mathbf{s} = [s_1,\ldots,s_B]^{\mathrm{T}} \in \mathbb{M}^B$ is first transformed by the permutation matrix (see Figure 20.2):

$$\boldsymbol{\Pi}^{(\mathcal{O})} = \sum_{i=1}^{B}\mathbf{e}_i\mathbf{e}_{b_i}^{\mathrm{T}} \in \{0,1\}^{B\times B} \quad \text{with } \boldsymbol{\Pi}^{(\mathcal{O}),-1} = \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}, \tag{20.8}$$

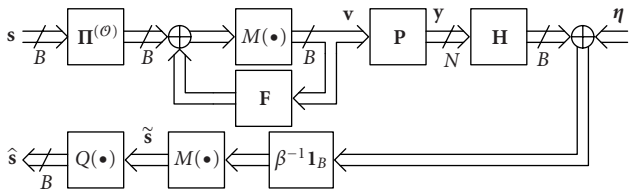

FIGURE 20.2. THP transmission over a nondispersive channel.

that is, the b_ith scalar symbol s_{b_i} is put at the ith position for $i = 1,\dots,B$. Thus, the b_1th data symbol s_{b_1} is precoded first and the b_Bth data symbol s_{b_B} last. For compactness, we collect the indices $b_1,\dots,b_B$ in the B-tupel

$$\mathcal{O} = (b_1,\dots,b_B) \quad \text{with } b_i \in \{1,\dots,B\} \setminus \{b_1,\dots,b_{i-1}\}. \tag{20.9}$$

We will also use the term *ordering* for the B-tupel $\mathcal{O}$. The transmitter can choose the ordering $\mathcal{O}$ freely, since the reordering by the permutation matrix $\mathbf{\Pi}^{(\mathcal{O})}$ cannot be recognized by the receiver. Hence, the ordering $\mathcal{O}$ represents additional degrees of freedom.

After the reordering by $\mathbf{\Pi}^{(\mathcal{O})}$, the signal is passed through the nonlinear feedback loop to get the signal $\mathbf{v} \in \mathbb{M}^B$. The modulo operator in Figure 20.2 is defined element-wise:

$$\begin{aligned} \mathrm{M}(\mathbf{x}) &= \left[\,\mathrm{M}\left(x_1\right),\dots,\mathrm{M}\left(x_B\right)\right]^{\mathrm{T}} \in \mathbb{M}^B \quad \text{with } \mathbf{x} \in \mathbb{C}^B, \\ \mathrm{M}\left(x_i\right) &= x_i - \left\lfloor \frac{\mathrm{Re}\left(x_i\right)}{\tau} + \frac{1}{2} \right\rfloor \tau - \mathrm{j} \left\lfloor \frac{\mathrm{Im}\left(x_i\right)}{\tau} + \frac{1}{2} \right\rfloor \tau, \end{aligned} \tag{20.10}$$

where $x_i \in \mathbb{C}$, $i = 1,\dots,B$, denotes the ith entry of $\mathbf{x}$ and $\lfloor \bullet \rfloor$ denotes the floor operator which gives the largest integer smaller than or equal to the argument. Note that the amplitude of the modulo operator is upper bounded, since $\mathrm{M}(x_i) \in \mathbb{M}$, where

$$\mathbb{M} = \left\{ z \in \mathbb{C} \mid -\frac{\tau}{2} \le \mathrm{Re}(z) < \frac{\tau}{2} \text{ and } -\frac{\tau}{2} \le \mathrm{Im}(z) < \frac{\tau}{2} \right\}. \tag{20.11}$$

The modulo constant $\tau \in \mathbb{R}_+$ is chosen depending on the modulation alphabet (all symbols of the modulation alphabet have to be elements of $\mathbb{M}$, see, for example, [34]). For example, we set $\tau = 2\sqrt{2}$ for QPSK modulation (see Figure 20.3) whose symbols are elements of the set $\{\exp(\mathrm{j}\boldsymbol{\mu}\pi/4) \mid \boldsymbol{\mu} \in \{-3,-1,+1,+3\}\}$.

We assume that the scalar entries $v_1,\dots,v_B \in \mathbb{M}$ of the modulo operator output $\mathbf{v} \in \mathbb{M}^B$ are uncorrelated due to the modulo operation $\mathrm{M}(\bullet)$, that is, $\mathbf{R}_{\mathbf{v}} = \mathrm{diag}(\sigma_{v_1}^2,\dots,\sigma_{v_B}^2) \in \mathbb{R}_+^{B\times B}$. Additionally, we make the popular assumption that the ith output v_i, $i = 2,\dots,B$, of the modulo operator $\mathrm{M}(\bullet)$ at the transmitter

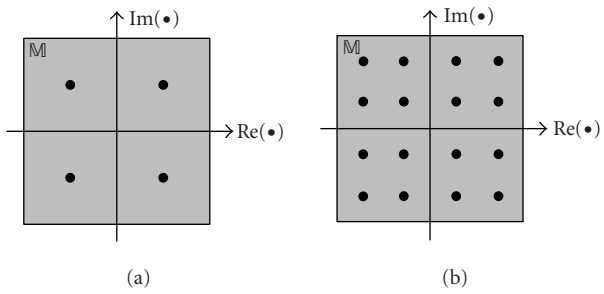

FIGURE 20.3. Modulo operator M(•): set $\mathbb{M}$ of output values, (a) QPSK: $\tau = 2\sqrt{2}$, (b) 16QAM: $\tau = 8/\sqrt{10}$.

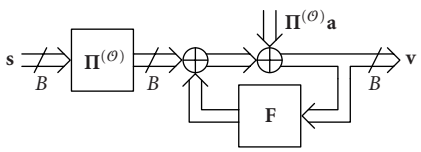

FIGURE 20.4. Linear representation of the modulo operator at the transmitter.

is uniformly distributed over $\mathbb{M}$, which results in the variance

$$\sigma_{\nu_i}^2 = \sigma_{\nu}^2 = \mathrm{E}\Big[\,|\,\nu_i\,|^2\Big] = \frac{\tau^2}{6}, \quad i = 2,\ldots,B. \tag{20.12}$$

Since the first modulo operator output ν_1 is simply s_{b_1} as we will see in the following, its variance is $\sigma_{\nu_1}^2 = \sigma_s^2$, where $\sigma_s^2 = \mathrm{E}[|s_{b_1}|^2]$.

Since M(•) is defined element-wise and we assume an ordering such that the first entry of the permuted data vector $\mathbf{\Pi}^{(\mathcal{O})}\mathbf{s}$ is precoded first and the last entry last, the feedback filter $\mathbf{F} \in \mathbb{C}^{B\times B}$ has to be *lower triangular with zero main diagonal* to ensure the realizability of the feedback loop.[1] This property of $\mathbf{F}$ is often called *spatial causality*, as only data symbols which have already been precoded are fed back.

Due to the definition of M(•) in (20.10), we can follow that the output of the modulo operator M(•) is simply the sum of the input and a term which ensures that all scalar entries of the output are elements of $\mathbb{M}$. When taking this observation into account, we end up with the linear representation of the feedback loop at the transmitter depicted in Figure 20.4. Note that the auxiliary signal $\mathbf{\Pi}^{(\mathcal{O})}\mathbf{a} \in \mathbb{C}^B$,

[1]The scalar signal ν_1 is constructed without feedback (the first row of $\mathbf{F}$ is zero) and is equal to s_{b_1}, whereas the scalar signal ν_B depends on s_{b_B} and $\nu_1,\ldots,\nu_{B-1}$ but not on itself (the last element of the last row of $\mathbf{F}$ is zero). We could also assume alternative orderings, for example, the reverse ordering, that is, s_{b_B} is precoded first and s_{b_1} last. The resulting $\mathbf{F}$ would be upper triangular with zero main diagonal instead.

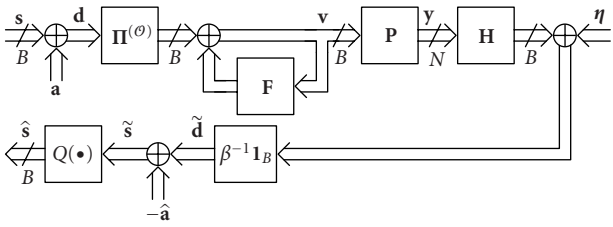

FIGURE 20.5. Linear representation of THP transmission.

whose scalar entries have real and imaginary parts which are integer multiples of the modulo constant τ, can be moved from the inside to the front of the feedback loop.[2] When replacing the modulo operators in Figure 20.2 by the summation with the appropriate auxiliary signals and moving the auxiliary signal $\mathbf{\Pi}^{(\mathcal{O})}\mathbf{a}$ at the transmitter to the front of the feedback loop, we obtain the linear representation in Figure 20.5. As the auxiliary signals $\mathbf{a}$ and $-\tilde{\mathbf{a}}$ which are added at the transmitter and the receiver, respectively, are included automatically by the modulo operators, we use $\mathbf{d}$ as the desired signal and $\tilde{\mathbf{d}}$ as the estimate in the following optimizations. We see in Figure 20.5 that the system whose input and output is $\mathbf{d}$ and $\widetilde{\mathbf{d}}$, respectively, is linear. Consequently, we can apply the optimizations for the linear transmit filters reviewed in the previous subsection also for the design of the THP filters, but we have to take into account the special structure of the feedback filter $\mathbf{F}$ and the statistical properties of the modulo operator output $\mathbf{v}$. We can conclude that the advantage of THP compared to the linear transmit filters is due to the advantageous statistical properties of the signal $\mathbf{v}$, whose amplitude is upper bounded, since $\mathbf{v} \in \mathbb{M}^B$.

The desired signal $\mathbf{d}$ as a function of the modulo operator output $\mathbf{v} \in \mathbb{M}^B$ can be expressed as (see Figure 20.5)

$$\mathbf{d} = \mathbf{\Pi}^{(\mathcal{O}),\mathrm{T}}(\mathbf{1}_B - \mathbf{F})\mathbf{v} \in \mathbb{C}^B. \tag{20.13}$$

The output $\mathbf{v}$ of the modulo operator is passed through the feedforward filter $\mathbf{P} \in \mathbb{C}^{N_a \times B}$, propagates over the channel $\mathbf{H} \in \mathbb{C}^{B \times N_a}$, is perturbed by the noise $\boldsymbol{\eta} \in \mathbb{C}^B$, and is weighted with β^{-1} at the receiver to form the estimate

$$\tilde{\mathbf{d}} = \beta^{-1}\mathbf{HPv} + \beta^{-1}\boldsymbol{\eta} \in \mathbb{C}^B. \tag{20.14}$$

Note that the scalar weight β^{-1} at the receiver is always necessary in a THP system, since the choice of the modulo constant τ is based on the assumption that the

[2] From Figure 20.4, we see that $\mathbf{v} = \mathbf{\Pi}^{(\mathcal{O})}\mathbf{s} + \mathbf{\Pi}^{(\mathcal{O})}\mathbf{a} + \mathbf{Fv}$ or equivalently, we obtain for the output signal of the modulo operator at the transmitter $\mathbf{v} = (\mathbf{1}_B - \mathbf{F})^{-1}\mathbf{\Pi}^{(\mathcal{O})}(\mathbf{s} + \mathbf{a})$. Consequently, $\mathbf{a}$ can be directly added to $\mathbf{s}$ in front of the feedback loop.

amplitude at the modulo operator input is correct. The transmit signal is simply $\mathbf{y} = \mathbf{P}\mathbf{v} \in \mathbb{C}^{N_a}$ and hence, the expression for the transmit energy reads as

$$\mathrm{E}\left[||\mathbf{y}||_2^2\right] = \mathrm{tr}\,(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}). \tag{20.15}$$

20.4. Spatial Tomlinson-Harashima precoding

20.4.1. Zero-forcing spatial Tomlinson-Harashima precoding

The zero-forcing variant of *spatial THP* (S-THP) for nondispersive channels can be found by applying an optimization which is based on the optimization for linear zero-forcing transmit processing (cf. (20.4)). First, we have to replace the desired signal $\mathbf{s}$ by $\mathbf{d}$ and the estimate $\hat{\mathbf{s}}$ by $\tilde{\mathbf{d}}$. Second, we have to include the ordering $\mathcal{O}$ and a constraint that the feedback filter $\mathbf{F} \in \mathbb{C}^{B\times B}$ is spatially causal:

$$\begin{aligned}\left\{\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}, \mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}}, \beta_{\mathrm{ZF}}^{\mathrm{THP}}, \mathcal{O}_{\mathrm{ZF}}^{\mathrm{THP}}\right\} &= \underset{\{\mathbf{P},\mathbf{F},\beta,\mathcal{O}\}}{\operatorname{argmin}} \mathrm{E}\left[||\mathbf{d}-\tilde{\mathbf{d}}||_2^2\right] \\ \text{s.t. : } \tilde{\mathbf{d}}|_{\boldsymbol{\eta}=\mathbf{0}_B} &= \mathbf{d}, \quad \mathrm{E}\left[||\mathbf{y}||_2^2\right] = E_{\mathrm{tr}}, \\ \mathbf{F} &: \text{lower triangular, zero main diagonal.}\end{aligned} \tag{20.16}$$

When plugging (20.13), (20.14), and (20.15) into the above optimization, we get

$$\begin{aligned}\left\{\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}, \mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}}, \beta_{\mathrm{ZF}}^{\mathrm{THP}}, \mathcal{O}_{\mathrm{ZF}}^{\mathrm{THP}}\right\} &= \underset{\{\mathbf{P},\mathbf{F},\beta,\mathcal{O}\}}{\operatorname{argmin}} \beta^{-2}\,\mathrm{tr}\,(\mathbf{R}_{\boldsymbol{\eta}}) \\ \text{s.t. } \beta^{-1}\mathbf{H}\mathbf{P} &= \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}(\mathbf{1}_B - \mathbf{F}), \quad \mathrm{tr}\,(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}) = E_{\mathrm{tr}}, \\ \mathbf{S}_i\mathbf{F}\mathbf{e}_i &= \mathbf{0}_i,\ i = 1,\ldots,B,\end{aligned} \tag{20.17}$$

where we split up the constraint on the spatial causality of $\mathbf{F}$ into B constraints for the columns of $\mathbf{F}$. Here, $\mathbf{e}_i \in \{0,1\}^B$ and we introduced the selection matrix

$$\mathbf{S}_i = \mathbf{S}_{(0,i,B-i)} = \left[\mathbf{1}_i, \mathbf{0}_{i\times B-i}\right] \in \{0,1\}^{i\times B} \tag{20.18}$$

which cuts out the first i elements of a B-dimensional vector.

Obviously, the optimization in (20.17) is not convex due to the first constraint. However, we obtain necessary conditions for the optimum *zero-forcing S-THP* (ZF-S-THP) filters, when setting the derivatives of the Lagrangian function

$$\begin{aligned}&L(\mathbf{P},\mathbf{F},\beta,\mathcal{O},\boldsymbol{\Lambda},\rho,\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\ldots,\boldsymbol{\mu}_B) \\ &\quad= \beta^{-2}\,\mathrm{tr}\,(\mathbf{R}_{\boldsymbol{\eta}}) - \rho(\,\mathrm{tr}\,(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}) - E_{\mathrm{tr}}) \\ &\qquad- 2\,\mathrm{Re}\,(\,\mathrm{tr}\,(\boldsymbol{\Lambda}(\beta^{-1}\mathbf{H}\mathbf{P} - \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}(\mathbf{1}_B - \mathbf{F})))) - \sum_{i=1}^{B} 2\,\mathrm{Re}\,(\boldsymbol{\mu}_i^{\mathrm{T}}\mathbf{S}_i\mathbf{F}\mathbf{e}_i),\end{aligned} \tag{20.19}$$

with the Lagrangian multipliers $\mathbf{\Lambda} \in \mathbb{C}^{B\times B}$, $\rho \in \mathbb{R}$, and $\boldsymbol{\mu}_i \in \mathbb{C}^i$, $i = 1,\ldots,B$, to zero:

$$\begin{aligned}
\frac{\partial L(\mathbf{P},\mathbf{F},\beta,\mathbf{\Lambda},\rho,\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\ldots,\boldsymbol{\mu}_B)}{\partial \mathbf{P}} &= -\beta^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{\Lambda}^{\mathrm{T}} - \rho\mathbf{P}^{*}\mathbf{R}_{\mathbf{v}}^{\mathrm{T}} = \mathbf{0}_{N_{\mathrm{a}}\times B},\\
\frac{\partial L(\mathbf{P},\mathbf{F},\beta,\mathbf{\Lambda},\rho,\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\ldots,\boldsymbol{\mu}_B)}{\partial \mathbf{F}} &= -\mathbf{\Pi}^{(\mathcal{O})}\mathbf{\Lambda}^{\mathrm{T}} - \sum_{i=1}^{B}\mathbf{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i\mathbf{e}_i^{\mathrm{T}} = \mathbf{0}_{B\times B},\\
\frac{\partial L(\mathbf{P},\mathbf{F},\beta,\mathbf{\Lambda},\rho,\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\ldots,\boldsymbol{\mu}_B)}{\partial \beta} &= -2\beta^{-3}\,\mathrm{tr}\,(\mathbf{R}_{\boldsymbol{\eta}}) + \beta^{-2}\,\mathrm{Re}\,(\,\mathrm{tr}(\mathbf{\Lambda}\mathbf{H}\mathbf{P})).
\end{aligned} \tag{20.20}$$

From the derivative with respect to the feedback filter $\mathbf{F}$, it follows that

$$\mathbf{\Lambda}^{\mathrm{H}} = -\mathbf{\Pi}^{(\mathcal{O}),\mathrm{T}}\sum_{i=1}^{B}\mathbf{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^{*}\mathbf{e}_i^{\mathrm{T}}. \tag{20.21}$$

Plugging this result into the derivative of the Lagrangian function with respect to the feedforward filter $\mathbf{P}$ yields

$$\rho\sigma_{v_i}^2\mathbf{P}\mathbf{e}_i = \beta^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{\Pi}^{(\mathcal{O}),\mathrm{T}}\mathbf{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^{*}, \tag{20.22}$$

where we multiplied by $\mathbf{e}_i \in \{0,1\}^B$ and used the assumption that the output $\mathbf{v}$ of the modulo operation at the transmitter is uncorrelated. Due to the first constraint of (20.17), the ith column of the feedback filter can be written as

$$\mathbf{F}\mathbf{e}_i = \mathbf{e}_i - \rho^{-1}\sigma_{v_i}^{-2}\beta^{-2}\mathbf{\Pi}^{(\mathcal{O})}\mathbf{H}\mathbf{H}^{\mathrm{H}}\mathbf{\Pi}^{(\mathcal{O}),\mathrm{T}}\mathbf{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i^{*}. \tag{20.23}$$

Consequently, by employing the last constraint of (20.17), we obtain for the Lagrangian multiplier

$$\boldsymbol{\mu}_i^{*} = \rho\sigma_{v_i}^2\beta^2\left(\mathbf{S}_i\mathbf{\Pi}^{(\mathcal{O})}\mathbf{H}\mathbf{H}^{\mathrm{H}}\mathbf{\Pi}^{(\mathcal{O}),\mathrm{T}}\mathbf{S}_i^{\mathrm{T}}\right)^{-1}\mathbf{S}_i\mathbf{e}_i, \tag{20.24}$$

which leads to following expression for the ith column of the feedforward filter $\mathbf{P}$:

$$\mathbf{P}\mathbf{e}_i = \beta\mathbf{H}^{\mathrm{H}}\mathbf{\Pi}^{(\mathcal{O}),\mathrm{T}}\mathbf{S}_i^{\mathrm{T}}\left(\mathbf{S}_i\mathbf{\Pi}^{(\mathcal{O})}\mathbf{H}\mathbf{H}^{\mathrm{H}}\mathbf{\Pi}^{(\mathcal{O}),\mathrm{T}}\mathbf{S}_i^{\mathrm{T}}\right)^{-1}\mathbf{S}_i\mathbf{e}_i. \tag{20.25}$$

Note that the columns of $\mathbf{P}$ are orthogonal, that is,

$$\mathbf{e}_j^{\mathrm{T}}\mathbf{P}^{\mathrm{H}}\mathbf{P}\mathbf{e}_i = 0, \tag{20.26}$$

for $j \neq i$, since the selection matrix $\mathbf{S}_i = [\mathbf{1}_i, \mathbf{0}_{i\times B-i}] \in \{0,1\}^{i\times B}$ has the following properties:

$$\mathbf{S}_j = \mathbf{S}_j\mathbf{S}_i^{\mathrm{T}}\mathbf{S}_i, \quad \mathbf{S}_j\mathbf{e}_i = \mathbf{0}_j, \quad \text{for } j < i. \tag{20.27}$$

Since the modulo output $\mathbf{v}$ is uncorrelated, we can rewrite the transmit energy constraint:

$$\operatorname{tr}\left(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\right) = \beta^2 \sum_{i=1}^{B} \sigma_{v_i}^2 \mathbf{e}_i^{\mathrm{T}} \mathbf{S}_i^{\mathrm{T}} \left(\mathbf{S}_i \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \mathbf{S}_i^{\mathrm{T}}\right)^{-1} \mathbf{S}_i \mathbf{e}_i = E_{\mathrm{tr}}, \tag{20.28}$$

which leads to the ZF-S-THP solution depending on the ordering $\mathcal{O}$:

$$\begin{aligned}
\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}} &= \beta_{\mathrm{ZF}}^{\mathrm{THP}} \sum_{i=1}^{B} \mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \mathbf{S}_i^{\mathrm{T}} \left(\mathbf{S}_i \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \mathbf{S}_i^{\mathrm{T}}\right)^{-1} \mathbf{S}_i \mathbf{e}_i \mathbf{e}_i^{\mathrm{T}}, \\
\mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}} &= \mathbf{1}_B - \beta_{\mathrm{ZF}}^{\mathrm{THP},-1} \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H} \mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}} \in \mathbb{C}^{B \times B}, \\
\beta_{\mathrm{ZF}}^{\mathrm{THP}} &= \sqrt{\frac{E_{\mathrm{tr}}}{\sum_{i=1}^{B} \sigma_{v_i}^2 \mathbf{e}_i^{\mathrm{T}} \mathbf{S}_i^{\mathrm{T}} \left(\mathbf{S}_i \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \mathbf{S}_i^{\mathrm{T}}\right)^{-1} \mathbf{S}_i \mathbf{e}_i}} \in \mathbb{R}_+.
\end{aligned} \tag{20.29}$$

We observe that the ith column of the ZF-S-THP feedforward filter $\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}$ only depends on the first i rows of the sorted channel matrix $\boldsymbol{\Pi}^{(\mathcal{O})}\mathbf{H}$. Thus, the first column of $\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}$ is simply the weighted *transmit matched filter* (TxMF, see, e.g., [2]) for the first scalar data stream s_{b_1} and the last column is the weighted b_Bth column of the linear TxZF $\mathbf{P}_{\mathrm{ZF}}$. We can conclude that the first scalar data stream s_{b_1} is transmitted without taking into account the interference which is introduced in the other scalar estimates, since this interference is removed by the feedback filter $\mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}}$ and the modulo operations. The second column of the feedforward filter $\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}$ is orthogonal to the b_1th row of the channel $\mathbf{H}$, that is, the second precoded signal v_2 is not interfering with the estimate $\widetilde{s}_{b_1}$ of the first scalar data stream, but as the second column of $\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}$ only depends on the first two rows of the sorted channel matrix $\boldsymbol{\Pi}^{(\mathcal{O})}\mathbf{H}$, the signal v_2 contributes interference to the other estimates, namely, the estimates for $s_{b_3}, \ldots, s_{b_B}$. This interference has to be removed by the feedback filter $\mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}}$ and the modulo operations. The signals $v_1, \ldots, v_{B-1}$ cause interference in the estimate for the last data stream s_{b_B}. Thus, the feedback filter $\mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}}$ has to suppress this interference and the resulting output of the modulo operation v_B does not interfere with the estimates of the other data streams, since the last column of $\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}$ is orthogonal to the first $B-1$ rows of the sorted channel matrix $\boldsymbol{\Pi}^{(\mathcal{O})}\mathbf{H}$.

Note that we can employ the *Cholesky factorization* (e.g., [31, 35]) of the channel Gram with *symmetric permutation*,[3]

$$\boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} = \mathbf{L}\mathbf{L}^{\mathrm{H}}, \tag{20.30}$$

[3] Alternatively, we can use the QR factorization (e.g., [31, 35]) of the Hermitian $\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} = \mathbf{Q}\mathbf{R}$ of the sorted channel matrix. Then, $\mathbf{L} = \mathbf{R}^{\mathrm{H}}$.

with the lower triangular matrix $\mathbf{L} \in \mathbb{C}^{B\times B}$, to find following alternative expressions for the ZF-S-THP filters:

$$\begin{aligned}\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}} &= \beta_{\mathrm{ZF}}^{\mathrm{THP}}\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}\mathbf{L}^{\mathrm{H},-1}\,\mathrm{diag}\,(l_{1,1}^{-1},\ldots,l_{B,B}^{-1}),\\ \mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}} &= \mathbf{1}_B - \mathbf{L}\,\mathrm{diag}\,(l_{1,1}^{-1},\ldots,l_{B,B}^{-1}),\end{aligned} \tag{20.31}$$

and

$$\beta_{\mathrm{ZF}}^{\mathrm{THP}} = \sqrt{\frac{E_{\mathrm{tr}}}{\mathrm{tr}\,(\,\mathrm{diag}\,(l_{1,1}^{-2},\ldots,l_{B,B}^{-2})\mathbf{R}_{\mathbf{v}})}}, \tag{20.32}$$

where $l_{i,i} \in \mathbb{R}_+$ denotes the ith diagonal element of $\mathbf{L}$. Remember that we have obtained this result by solving the optimization in (20.17) and afterwards rewriting the solution with the Cholesky decomposition of the symmetrically permuted channel Gram or the QR factorization of the Hermitian of the sorted channel matrix. Contrarily, *no optimization* was performed and the expressions for the S-THP filters were found intuitively by using the *unsorted* QR factorization in [19, 21], where the weighting at the receiver was assumed to be the diagonal matrix $\mathrm{diag}(l_{1,1}^{-1},\ldots,l_{B,B}^{-1})$ instead of $\beta^{-1}\mathbf{1}_B$, although this choice for the diagonal weighting is suboptimum (see [25]).

When using the projector

$$\boldsymbol{\Pi}_i^{(\mathcal{O})} = \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}\mathbf{S}_i^{\mathrm{T}}\mathbf{S}_i\boldsymbol{\Pi}^{(\mathcal{O})} = \mathbf{1}_B - \sum_{j=i+1}^{B}\mathbf{e}_{b_j}\mathbf{e}_{b_j}^{\mathrm{T}} \in \{0,1\}^{B\times B}, \tag{20.33}$$

we can rewrite the ZF-S-THP solution in (20.29):

$$\begin{aligned}\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}} &= \beta_{\mathrm{ZF}}^{\mathrm{THP}}\sum_{i=1}^{B}\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}_i^{(\mathcal{O})}\left(\boldsymbol{\Pi}_i^{(\mathcal{O})}\mathbf{H}\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}_i^{(\mathcal{O})}\right)^{+}\mathbf{e}_{b_i}\mathbf{e}_i^{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{a}}\times B},\\ \mathbf{F}_{\mathrm{ZF}}^{\mathrm{THP}} &= \mathbf{1}_B - \beta_{\mathrm{ZF}}^{\mathrm{THP},-1}\boldsymbol{\Pi}^{(\mathcal{O})}\mathbf{H}\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}},\\ \beta_{\mathrm{ZF}}^{\mathrm{THP}} &= \sqrt{\frac{E_{\mathrm{tr}}}{\sum_{i=1}^{B}\sigma_{v_i}^2\mathbf{e}_{b_i}^{\mathrm{T}}\left(\boldsymbol{\Pi}_i^{(\mathcal{O})}\mathbf{H}\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}_i^{(\mathcal{O})}\right)^{+}\mathbf{e}_{b_i}}} \in \mathbb{R}_+.\end{aligned} \tag{20.34}$$

The result for the scalar weight $\beta_{\mathrm{ZF}}^{\mathrm{THP}}$ depending on the ordering $\mathcal{O}$ enables us to further minimize the MSE by the appropriate ordering:

$$\mathcal{O}'_{\mathrm{ZF}} = \underset{\mathcal{O}}{\mathrm{argmin}}\,\frac{\mathrm{tr}\,(\mathbf{R}_{\boldsymbol{\eta}})}{E_{\mathrm{tr}}}\sum_{i=1}^{B}\sigma_{v_i}^2\mathbf{e}_{b_i}^{\mathrm{T}}\left(\boldsymbol{\Pi}_i^{(\mathcal{O})}\mathbf{H}\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}_i^{(\mathcal{O})}\right)^{+}\mathbf{e}_{b_i}. \tag{20.35}$$

Thus, the optimal ordering $\mathcal{O}'_{\mathrm{ZF}}$ can only be found by computing the MSEs for all $B!$ possible orderings and choosing the ordering with the minimum MSE. Since the above optimization is very complex, we suggest to use the following suboptimum

approach. Instead of minimizing the sum of MSEs for the B scalar data streams, the ordering is found by successively minimizing each MSE summand under the assumption that the ordering of the *succeeding* MSE summands is fixed:

$$\mathcal{O}_{\mathrm{ZF}} = (b_{\mathrm{ZF},1}, \ldots, b_{\mathrm{ZF},B}) \tag{20.36}$$

with

$$b_{\mathrm{ZF},i} = \underset{b \in \{1,\ldots,B\} \setminus \{b_{\mathrm{ZF},i+1},\ldots,b_{\mathrm{ZF},B}\}}{\operatorname{argmin}} \mathbf{e}_b^{\mathrm{T}} \left(\boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{ZF}})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{ZF}})} \right)^{+} \mathbf{e}_b, \tag{20.37}$$

where $i = B, \ldots, 1$. Note that $\boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{ZF}})}$ only depends on the already computed $b_{\mathrm{ZF},i+1}$, $\ldots, b_{\mathrm{ZF},B}$ and hence, the pseudoinverse of $\boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{ZF}})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{ZF}})}$ is independent of b. Therefore, the complexity is $\mathrm{O}(B^4)$, whereas the optimum ordering of (20.35) has $\mathrm{O}(B!B^3)$. The above suboptimum procedure for ZF-S-THP is similar to the ordering optimization known as V-BLAST for spatial DFE, but the sorting is computed starting with the index for the data stream precoded *last*, whereas V-BLAST starts with the index of the data stream detected *first*.

20.4.2. Wiener spatial Tomlinson-Harashima precoding

The *Wiener S-THP* (WF-S-THP) filters for nondispersive channels results from the minimization of the mean square error together with the transmit energy constraint and the restriction of a spatially causal feedback filter:

$$\begin{aligned} &\left\{ \mathbf{P}_{\mathrm{WF}}^{\mathrm{THP}}, \mathbf{F}_{\mathrm{WF}}^{\mathrm{THP}}, \beta_{\mathrm{WF}}^{\mathrm{THP}}, \mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}} \right\} = \underset{\{\mathbf{P},\mathbf{F},\beta,\mathcal{O}\}}{\operatorname{argmin}} \mathrm{E}\left[||\mathbf{d} - \tilde{\mathbf{d}}||_2^2 \right] \\ &\text{s.t. } \mathrm{E}\left[||\mathbf{y}||_2^2 \right] = E_{\mathrm{tr}}, \quad \mathbf{F} : \text{lower triangular, zero main diagonal.} \end{aligned} \tag{20.38}$$

With (20.13), (20.14), and (20.15), the above optimization can be written as

$$\begin{aligned} &\left\{ \mathbf{P}_{\mathrm{WF}}^{\mathrm{THP}}, \mathbf{F}_{\mathrm{WF}}^{\mathrm{THP}}, \beta_{\mathrm{WF}}^{\mathrm{THP}}, \mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}} \right\} = \underset{\{\mathbf{P},\mathbf{F},\beta,\mathcal{O}\}}{\operatorname{argmin}} \sigma_{\boldsymbol{\varepsilon}}^2(\mathbf{P}, \mathbf{F}, \beta, \mathcal{O}) \\ &\text{s.t. } \mathrm{tr}\left(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\right) = E_{\mathrm{tr}}, \qquad \mathbf{S}_i \mathbf{F} \mathbf{e}_i = \mathbf{0}_i, \; i = 1, \ldots, B, \end{aligned} \tag{20.39}$$

where the MSE $\sigma_{\boldsymbol{\varepsilon}}^2(\mathbf{P}, \mathbf{F}, \beta, \mathcal{O}) = \mathrm{E}[\|\mathbf{d} - \tilde{\mathbf{d}}\|_2^2]$ is defined as

$$\begin{aligned} &\sigma_{\boldsymbol{\varepsilon}}^2(\mathbf{P}, \mathbf{F}, \beta, \mathcal{O}) \\ &\quad = -2\beta^{-1} \mathrm{tr}\left(\mathrm{Re}\left(\boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} (\mathbf{1}_B - \mathbf{F}) \mathbf{R}_{\mathbf{v}} \mathbf{P}^{\mathrm{H}} \mathbf{H}^{\mathrm{H}} \right) \right) \\ &\qquad + \mathrm{tr}\left((\mathbf{1}_B - \mathbf{F}) \mathbf{R}_{\mathbf{v}} (\mathbf{1}_B - \mathbf{F}^{\mathrm{H}}) \right) + \beta^{-2} \mathrm{tr}\left(\mathbf{H}\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} + \mathbf{R}_{\boldsymbol{\eta}} \right). \end{aligned} \tag{20.40}$$

The selection matrix $\mathbf{S}_i = [\mathbf{1}_i, \mathbf{0}_{i \times B-i}] \in \{0,1\}^{i \times B}$ cuts out the first i elements of a B-dimensional column vector and $\mathbf{e}_i \in \{0,1\}^B$. Employing the Lagrangian

multipliers $\rho \in \mathbb{R}$ and $\boldsymbol{\mu}_i \in \mathbb{C}^i$ with $i = 1, \ldots, B$, we form the Lagrangian function

$$\begin{aligned} &L(\mathbf{P}, \mathbf{F}, \beta, \mathcal{O}, \rho, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_B) \\ &\quad = \sigma_{\boldsymbol{\varepsilon}}^2(\mathbf{P}, \mathbf{F}, \beta, \mathcal{O}) - \rho\big(\operatorname{tr}\left(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\right) - E_{\mathrm{tr}}\big) - \sum_{i=1}^{B} 2\operatorname{Re}\left(\boldsymbol{\mu}_i^{\mathrm{T}}\mathbf{S}_i\mathbf{F}\mathbf{e}_i\right), \end{aligned} \tag{20.41}$$

whose derivatives with respect to the feedforward filter $\mathbf{P}$, the feedback filter $\mathbf{F}$, and the scalar weight β must vanish:

$$\begin{aligned} \frac{\partial L(\cdots)}{\partial \mathbf{P}} &= \beta^{-2}\mathbf{H}^{\mathrm{T}}\mathbf{H}^{*}\mathbf{P}^{*}\mathbf{R}_{\mathbf{v}}^{\mathrm{T}} - \beta^{-1}\mathbf{H}^{\mathrm{T}}\boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}(\mathbf{1}_B - \mathbf{F}^{*})\mathbf{R}_{\mathbf{v}}^{\mathrm{T}} \\ &\quad - \rho\mathbf{P}^{*}\mathbf{R}_{\mathbf{v}}^{\mathrm{T}} = \mathbf{0}_{N\times B}, \\ \frac{\partial L(\cdots)}{\partial \mathbf{F}} &= -(\mathbf{1}_B - \mathbf{F}^{*})\mathbf{R}_{\mathbf{v}}^{\mathrm{T}} + \beta^{-1}\boldsymbol{\Pi}^{(\mathcal{O})}\mathbf{H}^{*}\mathbf{P}^{*}\mathbf{R}_{\mathbf{v}}^{\mathrm{T}} - \sum_{i=1}^{B}\mathbf{S}_i^{\mathrm{T}}\boldsymbol{\mu}_i\mathbf{e}_i^{\mathrm{T}} = \mathbf{0}_{B\times B}, \\ \frac{\partial L(\cdots)}{\partial \beta} &= 2\beta^{-2}\operatorname{tr}\left(\operatorname{Re}\left(\boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}(\mathbf{1}_B - \mathbf{F})\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\right)\right) \\ &\quad - 2\beta^{-3}\operatorname{tr}\left(\mathbf{H}\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} + \mathbf{R}_{\boldsymbol{\eta}}\right) = 0. \end{aligned} \tag{20.42}$$

When taking the complex conjugate of the derivative with respect to $\mathbf{P}$, multiplying with $\mathbf{P}^{\mathrm{H}}$ from the right, and applying the trace operator, we find following equality:

$$\begin{aligned} &-\beta^{-1}\operatorname{tr}\left(\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}(\mathbf{1}_B - \mathbf{F})\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\right) \\ &\qquad + \beta^{-2}\operatorname{tr}\left(\mathbf{H}^{\mathrm{H}}\mathbf{H}\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\right) - \rho\operatorname{tr}\left(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}\right) = 0. \end{aligned} \tag{20.43}$$

We can conclude that $\operatorname{tr}(\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}}(\mathbf{1}_B - \mathbf{F})\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}) \in \mathbb{R}$, since all other terms are real valued. Therefore, we can plug the above result into the derivative of the Lagrangian function with respect to the scalar weight β to obtain

$$\rho = -\beta^{-2}\frac{\operatorname{tr}\left(\mathbf{R}_{\boldsymbol{\eta}}\right)}{E_{\mathrm{tr}}}, \tag{20.44}$$

where we used the transmit energy constraint, that is, $\operatorname{tr}(\mathbf{P}\mathbf{R}_{\mathbf{v}}\mathbf{P}^{\mathrm{H}}) = E_{\mathrm{tr}}$. In the following, we use the abbreviation

$$\xi_{\mathrm{WF}} = \frac{\operatorname{tr}\left(\mathbf{R}_{\boldsymbol{\eta}}\right)}{E_{\mathrm{tr}}}. \tag{20.45}$$

Due to the derivative with respect to $\mathbf{P}$ and the above expression for the Lagrangian

multiplier ρ, the feedforward filter can be written as

$$\mathbf{P} = \beta \mathbf{H}^{\mathrm{H}} \left(\mathbf{H}\mathbf{H}^{\mathrm{H}} + \xi_{\mathrm{WF}} \mathbf{1}_B \right)^{-1} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \left(\mathbf{1}_B - \mathbf{F} \right), \tag{20.46}$$

where we also applied the matrix inversion lemma. Plugging this result into the derivative of the Lagrangian function with respect to the feedback filter $\mathbf{F}$ yields

$$\mathbf{F} = \mathbf{1}_B + \xi_{\mathrm{WF}}^{-1} \left(\boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} + \xi_{\mathrm{WF}} \mathbf{1}_B \right) \sum_{i=1}^{B} \sigma_{v_i}^{-2} \mathbf{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \mathbf{e}_i^{\mathrm{T}}, \tag{20.47}$$

which can be used to find the Lagrangian multipliers $\boldsymbol{\mu}_i$, $i = 1, \ldots, B$, with the second constraint of (20.39):

$$\boldsymbol{\mu}_i^* = -\xi_{\mathrm{WF}} \sigma_{v_i}^2 \left(\mathbf{S}_i \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \mathbf{S}_i^{\mathrm{T}} + \xi_{\mathrm{WF}} \mathbf{S}_i \mathbf{S}_i^{\mathrm{T}} \right)^{-1} \mathbf{S}_i \mathbf{e}_i, \tag{20.48}$$

with $i = 1, \ldots, B$. For the last two expressions, we employed the assumption that the modulo operator outputs at the transmitter are uncorrelated, that is, $\mathbf{R}_{\mathbf{v}} = \mathrm{diag}(\sigma_{v_1}^2, \ldots, \sigma_{v_B}^2)$. When we replace the identity matrix $\mathbf{1}_B$ in the expression for the feedback filter $\mathbf{F}$ by

$$\begin{aligned} \mathbf{1}_B &= \sum_{i=1}^{B} \mathbf{e}_i \mathbf{e}_i^{\mathrm{T}} = \sum_{i=1}^{B} \mathbf{S}_i^{\mathrm{T}} \mathbf{S}_i \mathbf{e}_i \mathbf{e}_i^{\mathrm{T}} \\ &= \sum_{i=1}^{B} -\xi_{\mathrm{WF}}^{-1} \sigma_{v_i}^{-2} \mathbf{S}_i^{\mathrm{T}} \mathbf{S}_i \left(\boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} + \xi_{\mathrm{WF}} \mathbf{1}_B \right) \mathbf{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \mathbf{e}_i^{\mathrm{T}}, \end{aligned} \tag{20.49}$$

we obtain for the feedback filter

$$\mathbf{F} = \sum_{i=1}^{B} \xi_{\mathrm{WF}}^{-1} \sigma_{v_i}^{-2} \left(\mathbf{1}_B - \mathbf{S}_i^{\mathrm{T}} \mathbf{S}_i \right) \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \mathbf{S}_i^{\mathrm{T}} \boldsymbol{\mu}_i^* \mathbf{e}_i^{\mathrm{T}}, \tag{20.50}$$

because $\mathbf{S}_i^{\mathrm{T}} \mathbf{S}_i \mathbf{S}_i^{\mathrm{T}} - \mathbf{S}_i^{\mathrm{T}} = \mathbf{0}_{B \times i}$. With the alternative expression for the Lagrangian multipliers

$$\boldsymbol{\mu}_i^* = -\xi_{\mathrm{WF}} \sigma_{v_i}^2 \mathbf{S}_i \boldsymbol{\Pi}^{(\mathcal{O})} \left(\boldsymbol{\Pi}_i^{(\mathcal{O})} \mathbf{H}\mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O})} + \xi_{\mathrm{WF}} \mathbf{1}_B \right)^{-1} \mathbf{e}_{b_i}, \quad i = 1, \ldots, B, \tag{20.51}$$

where we again introduced the projector

$$\boldsymbol{\Pi}_i^{(\mathcal{O})} = \boldsymbol{\Pi}^{(\mathcal{O}),\mathrm{T}} \mathbf{S}_i^{\mathrm{T}} \mathbf{S}_i \boldsymbol{\Pi}^{(\mathcal{O})} = \mathbf{1}_B - \sum_{j=i+1}^{B} \mathbf{e}_{b_j} \mathbf{e}_{b_j}^{\mathrm{T}} \in \{0, 1\}^{B \times B}, \tag{20.52}$$

we finally get the WF-S-THP solution for flat fading channels depending on the ordering $\mathcal{O}$:

$$\mathbf{P}_{\mathrm{WF}}^{\mathrm{THP}} = \beta_{\mathrm{WF}}^{\mathrm{THP}} \sum_{i=1}^{B} \mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O})} \mathbf{A}_{\mathrm{WF},i}^{-1} \mathbf{e}_{b_i} \mathbf{e}_i^{\mathrm{T}} \in \mathbb{C}^{N \times B}, \tag{20.53}$$

$$\mathbf{F}_{\mathrm{WF}}^{\mathrm{THP}} = \sum_{i=1}^{B} \left(\mathbf{S}_i^{\mathrm{T}} \mathbf{S}_i - \mathbf{1}_B \right) \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H} \mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O})} \mathbf{A}_{\mathrm{WF},i}^{-1} \mathbf{e}_{b_i} \mathbf{e}_i^{\mathrm{T}}, \tag{20.54}$$

and

$$\beta_{\mathrm{WF}}^{\mathrm{THP}} = \sqrt{\frac{E_{\mathrm{tr}}}{\sum_{i=1}^{B} \sigma_{v_i}^2 \mathbf{e}_{b_i}^{\mathrm{T}} \mathbf{A}_{\mathrm{WF},i}^{-2} \boldsymbol{\Pi}_i^{(\mathcal{O})} \mathbf{H} \mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O})} \mathbf{e}_{b_i}}} \in \mathbb{R}_+, \tag{20.55}$$

where $\mathbf{A}_{\mathrm{WF},i} = \boldsymbol{\Pi}_i^{(\mathcal{O})} \mathbf{H} \mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O})} + \xi_{\mathrm{WF}} \mathbf{1}_B$. The expression for the scalar $\beta_{\mathrm{WF}}^{\mathrm{THP}}$ was found with the transmit energy constraint. Interestingly, we get the same scalar weight $\xi_{\mathrm{WF}} = \mathrm{tr}(\mathbf{R}_{\boldsymbol{\eta}})/E_{\mathrm{tr}}$ for the identity matrix inside the inverse as for the linear TxWF (cf. (20.7)). Note that we can alternatively write for the feedback filter:

$$\mathbf{F}_{\mathrm{WF}}^{\mathrm{THP}} = \beta_{\mathrm{WF}}^{\mathrm{THP},-1} \sum_{i=1}^{B} \left(\mathbf{S}_i^{\mathrm{T}} \mathbf{S}_i - \mathbf{1}_B \right) \boldsymbol{\Pi}^{(\mathcal{O})} \mathbf{H} \mathbf{P}_{\mathrm{WF}}^{\mathrm{THP}} \mathbf{e}_i \mathbf{e}_i^{\mathrm{T}}, \tag{20.56}$$

that is, the ith column of the feedback filter is constructed by using the ith column of the filter chain $\beta_{\mathrm{WF}}^{\mathrm{THP},-1} \mathbf{H} \mathbf{P}_{\mathrm{WF}}^{\mathrm{THP}}$ and setting the first i elements to zero. The MSE for the WF-S-THP approach in terms of the feedback filter $\mathbf{F}_{\mathrm{WF}}^{\mathrm{THP}}$ can be expressed as

$$\sigma_{\boldsymbol{\varepsilon}}^2 \left(\mathbf{P}_{\mathrm{WF}}^{\mathrm{THP}}, \mathbf{F}_{\mathrm{WF}}^{\mathrm{THP}}, \beta_{\mathrm{WF}}^{\mathrm{THP}}, \mathcal{O} \right) = \xi_{\mathrm{WF}} \, \mathrm{tr} \left(\left(\mathbf{1}_B - \mathbf{F}_{\mathrm{WF}}^{\mathrm{THP}} \right) \mathbf{R}_{\mathbf{v}} \left(\mathbf{1}_B - \mathbf{F}_{\mathrm{WF}}^{\mathrm{THP,H}} \right) \mathbf{A}_{\mathrm{WF},i}^{-1} \right). \tag{20.57}$$

The MSE is further minimized by the choice of the ordering $\mathcal{O}$. With (20.54), we find the WF-S-THP ordering optimization:

$$\mathcal{O}'_{\mathrm{WF}} = \underset{\mathcal{O}}{\mathrm{argmin}} \, \xi_{\mathrm{WF}} \sum_{i=1}^{B} \sigma_{v_i}^2 \mathbf{e}_{b_i}^{\mathrm{T}} \left(\boldsymbol{\Pi}_i^{(\mathcal{O})} \mathbf{H} \mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O})} + \xi_{\mathrm{WF}} \mathbf{1}_B \right)^{-1} \mathbf{e}_{b_i}. \tag{20.58}$$

To avoid the high complexity $\mathrm{O}(B!B^3)$ of this optimization, we suggest to employ the following suboptimum approach instead:

$$\mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}} = \left(b_{\mathrm{WF},1}^{\mathrm{THP}}, \ldots, b_{\mathrm{WF},B}^{\mathrm{THP}} \right) \tag{20.59}$$

with

$$b_{\mathrm{WF},i}^{\mathrm{THP}} = \underset{b \in \mathbb{O}_i}{\mathrm{argmin}} \, \mathbf{e}_b^{\mathrm{T}} \left(\boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}})} \mathbf{H} \mathbf{H}^{\mathrm{H}} \boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}})} + \xi_{\mathrm{WF}} \mathbf{1}_B \right)^{-1} \mathbf{e}_b, \tag{20.60}$$

```
1:   𝕆 ← {1,...,B}
     G ← H
     for i = B,...,1:
4:     P ← (GG^H + ξ_WF 1_B)^{-1}
5:     b_i ← argmin_{b∈𝕆} e_b^T P e_b
6:     p_i ← G^H P e_{b_i}
       𝕆 ← 𝕆\{b_i}
       G ← (1_B − e_{b_i} e_{b_i}^T) G
     for i = 1,...,B:
10:    f_i ← (S_i^T S_i − 1_B) Π^(𝒪) H p_i
11:  χ ← σ_s^2 ‖p_1‖_2^2 + σ_v^2 Σ_{i=2}^B ‖p_i‖_2^2
     β ← √(E_tr/χ)
     P ← β[p_1,...,p_B]
```

Algorithm 20.1. Filter and ordering computation for spatial WF-THP over nondispersive channels.

```
1:   for i = 1,...,B:
2:     ν_i ← M(s_{b_i} + Σ_{j=1}^{i−1} e_i^T f_j ν_j)
     v = [ν_1,...,ν_K]^T
     y = Pv
```

Algorithm 20.2. Ordered spatial THP over nondispersive channels.

where $\mathbb{O}_i = \{1,\ldots,B\} \setminus \{b_{\mathrm{WF},i+1}^{\mathrm{THP}},\ldots,b_{\mathrm{WF},B}^{\mathrm{THP}}\}$ and $i = B,\ldots,1$. Thus, each summand of (20.58) is minimized for fixed *succeeding* indices $b_{\mathrm{WF},i+1}^{\mathrm{THP}},\ldots,b_{\mathrm{WF},B}^{\mathrm{THP}}$ starting from the index of the data stream precoded last and ending with the index of the data stream precoded first. Note that $\boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}})}$ only depends on the succeeding indices $b_{\mathrm{WF},i+1}^{\mathrm{THP}},\ldots,b_{\mathrm{WF},B}^{\mathrm{THP}}$. Thus, the inverse of $\boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}})}\mathbf{H}\mathbf{H}^{\mathrm{H}}\boldsymbol{\Pi}_i^{(\mathcal{O}_{\mathrm{WF}}^{\mathrm{THP}})} + \xi_{\mathrm{WF}}\mathbf{1}_B$ only has to be computed once for each step and the complexity of the above optimization is $\mathrm{O}(B^4)$.

In Algorithm 20.1, we present the resulting algorithm to compute the WF-S-THP filters as pseudocode. S-THP for nondispersive channels is illustrated by Algorithm 20.2. Note that we included the assumption in line 11 that all modulo outputs have variance $\sigma_{\mathbf{v}}^2 = \tau^2/6$ except the first modulo output ν_1 which is equal to s_{b_1} (see line 2 of Algorithm 20.2).

The ZF-S-THP variant can be found with a similar algorithm as the one in Algorithm 20.1, we only have to replace the lines 4–6 and 10 by the respective lines in Algorithm 20.3. Note that this step is equivalent to the limit $\xi_{\mathrm{WF}} \to 0$. However, the replacement is necessary, since we would end up with an inversion of a rank deficient matrix (see line 4 in Algorithm 20.1).

4: $\mathbf{P} \leftarrow \mathbf{G}^{+}$
5: $b_i \leftarrow \operatorname{argmin}_{b \in \mathbb{O}} \|\mathbf{P}\mathbf{e}_b\|_2^2$
6: $\mathbf{p}_i \leftarrow \mathbf{P}\mathbf{e}_{b_i}$
10: $\mathbf{f}_i \leftarrow \mathbf{e}_i - \mathbf{\Pi}^{(\mathcal{O})}\mathbf{H}\mathbf{p}_i$

Algorithm 20.3. Filter and ordering computation for spatial ZF-THP over nondispersive channels.

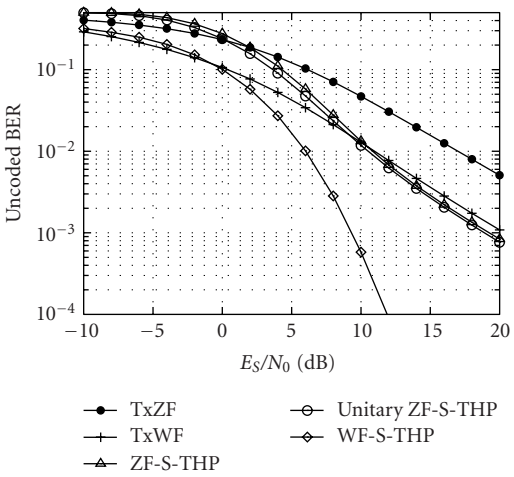

Figure 20.6. QPSK transmission over nondispersive MIMO channel with four transmitting and four receiving antenna elements: BER versus SNR for linear and nonlinear transmit processing.

20.5. Simulation results

We apply the spatial THP approaches discussed in this section to a nondispersive *multiple-input multiple-output* (MIMO) system, where we assume that the entries of the channel matrix are i.i.d. complex Gaussian distributed. The uncoded BER results are the mean of 240 000 channel realizations, where 100 vector symbols are transmitted per realization. The linear transmit filters TxZF and TxWF discussed in Section 20.2 are used for comparison to highlight the capabilities of THP. Additionally, we include the uncoded BER results for *unitary ZF-S-THP* [21] which is a variant of ZF-S-THP with unitary feedforward filter and a weighting with a diagonal matrix at the receiver instead of the scalar weighting employed in this section.

In Figure 20.6, we present the results for a system with four antenna elements deployed at the transmitter and four antenna elements at the receiver. Four QPSK symbols are transmitted per channel use. We observe that the S-THP approaches clearly outperform the respective linear filters, where WF-S-THP needs about 4 dB less SNR than ZF-S-THP for a BER of 10%. The unitary ZF-S-THP is slightly

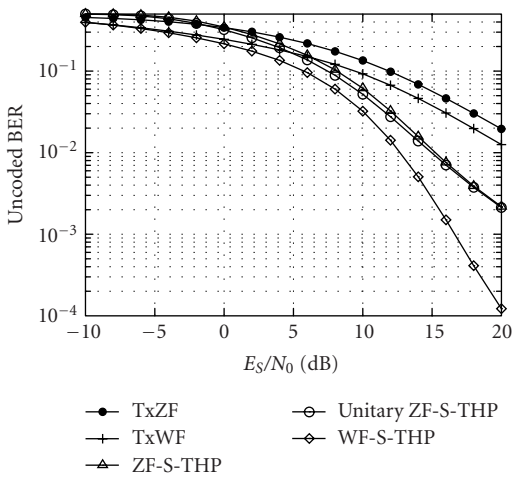

FIGURE 20.7. 16QAM transmission over nondispersive MIMO channel with four transmitting and four receiving antenna elements: BER versus SNR for linear and nonlinear transmit processing.

better than ZF-S-THP with scalar weighting, since the diagonal weighting of unitary ZF-S-THP offers more degrees of freedom. Note that the BERs of the two zero-forcing THP variants have the same slope as the two linear filters for high SNR. This behavior can be explained by the fact that the last column of the feedforward filter $\mathbf{P}_{\mathrm{ZF}}^{\mathrm{THP}}$ is the weighted b_Bth column of the TxZF $\mathbf{P}_{\mathrm{ZF}}$. Thus, the diversity order of the THP data stream precoded last is the same as the diversity order of the TxZF data streams (in our case, diversity order 1, slop is one magnitude of BER per 10 dB SNR). As the smallest diversity order is dominant, the ZF-THP approaches have the same diversity order as the linear transmit filters for high SNR.

The results for 16QAM transmission in Figure 20.7 are similar to the QPSK results in Figure 20.6. WF-S-THP is superior compared to all other depicted approaches and outperforms the linear TxWF even for low SNR. This result illustrates the dependence of THP on the modulation alphabet, since the modulo operation at the receiver introduces additional allowed constellation points. As the number of constellation points for 16QAM is larger than for QPSK, the impact of the modulo operation at the receiver is less pronounced for 16QAM.

When reducing the number of antenna elements at the receiver to three and transmitting three QPSK symbols per channel use, we end up with the results in Figure 20.8. Due to the increased number of freedoms compared to the case with four data streams of Figure 20.6, the linear transmit filters lead to better results than the THP approaches for low and medium SNRs, for example, the TxWF has a lower uncoded BER than WF-S-THP for an SNR below 1 dB. We can also observe that unitary ZF-S-THP is outperformed by ZF-S-THP with scalar weighting at the receiver. Consequently, the intuitively chosen diagonal weighting of unitary ZF-S-THP is suboptimum.

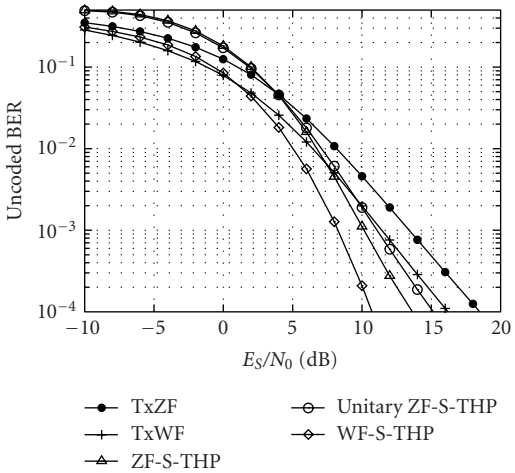

FIGURE 20.8. QPSK transmission over nondispersive MIMO channel with four transmitting and three receiving antenna elements: BER versus SNR for linear and nonlinear transmit processing.

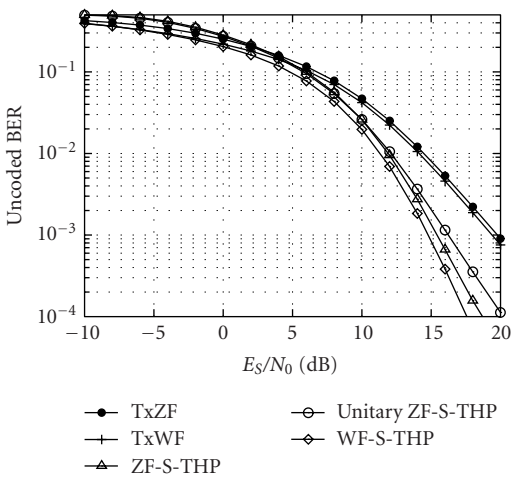

FIGURE 20.9. 16QAM transmission over nondispersive MIMO channel with four transmitting and three receiving antenna elements: BER versus SNR for linear and nonlinear transmit processing.

In Figure 20.9, three 16QAM symbols are transmitted per channel use. As the modulo operation at the receiver is less harmful for 16QAM, WF-S-THP exhibits the best uncoded BER results in the whole depicted SNR region contrary to Figure 20.8. Note how close the BER curves of the zero-forcing and Wiener precoder types lie in Figure 20.9 due to the available degrees of freedom, because the number of receive antenna elements is smaller than the number of transmit antenna elements.

Abbreviations

BER	Bit error rate
DFE	Decision feedback equalizer
DS-CDMA	Direct-sequence code-division multiple access
FIR	Finite impulse response
IIR	Infinite impulse response
MIMO	Multiple-input multiple-output
MSE	Mean square error
QAM	Quadrature amplitude modulation
QPSK	Quaternary phase-shift keying
SISO	Single-input single-output
SNR	Signal-to-noise ratio
S-THP	Spatial Tomlinson-Harashima precoding
THP	Tomlinson-Harashima precoding
TxMF	Transmit matched filter
TxWF	Transmit Wiener filter
TxZF	Transmit zero-forcing filter
V-BLAST	Vertical Bell Laboratories Space Time
WF-S-THP	Wiener spatial Tomlinson-Harashima precoding
WF-THP	Wiener Tomlinson-Harashima precoding
ZF-S-THP	Zero-forcing spatial Tomlinson-Harashima precoding
ZF-THP	Zero-forcing Tomlinson-Harashima precoding

Bibliography

[1] S. Verdú, *Multiuser Detection*, Cambridge University Press, New York, NY, USA, 1998.

[2] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," to appear in *IEEE Trans. Signal Processing*.

[3] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part I: channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, 2005.

[4] C. Peel, B. Hochwald, and L. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multi-user communication—Part II: perturbation," to appear in *IEEE Trans. Commun.*

[5] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-reduction-aided broadcast precoding," in *Proc. 5th International ITG Conference on Source and Channel Coding (SCC '04)*, pp. 403–408, Erlangen, Germany, January 2004.

[6] R. Irmer, R. Habendorf, W. Rave, and G. Fettweis, "Nonlinear multiuser transmission using multiple antennas for TDD-CDMA," in *Proc. IEEE 6th International Symposium on Wireless Personal Multimedia Communications (WPMC '03)*, vol. 3, pp. 251–255, Yokosuka, Japan, October 2003.

[7] T. Weber and M. Meurer, "Optimum joint transmission: potentials and dualities," in *Proc. IEEE 6th International Symposium on Wireless Personal Multimedia Communications (WPMC '03)*, vol. 1, pp. 79–83, Yokosuka, Japan, October 2003.

[8] R. F. H. Fischer and C. Windpassinger, "Improved MIMO precoding for decentralized receivers resembling concepts from lattice reduction," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '03)*, vol. 4, pp. 1852–1856, San Francisco, Calif, USA, December 2003.

[9] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE Global Telecommunications Conference (IEEE GLOBECOM '02)*, vol. 1, pp. 424–428, Taipei, Taiwan, November 2002.

[10] M. E. Austin, "Decision-feedback equalization for digital communication over dispersive channels," Tech. Rep. 437, MIT/Lincoln Laboratory, Lexington, Mass, USA, August 1967.
[11] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Processing*, vol. 48, no. 10, pp. 2921–2936, 2000.
[12] C. A. Belfiore and J. H. Park, "Decision feedback equalization," *Proc. IEEE*, vol. 67, no. 8, pp. 1143–1156, 1979.
[13] A. Duel-Hallen, "Equalizers for multiple input/multiple output channels and PAM systems with cyclostationary input sequences," *IEEE J. Select. Areas Commun.*, vol. 10, no. 3, pp. 630–639, 1992.
[14] P. Monsen, "Feedback equalization for fading dispersive channels," *IEEE Trans. Inform. Theory*, vol. 17, no. 1, pp. 56–64, 1971.
[15] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, no. 5/6, pp. 138–139, 1971.
[16] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, no. 4, pp. 774–780, 1972.
[17] G. D. Forney and M. V. Eyuboğlu, "Combined equalization and coding using precoding," *IEEE Commun. Mag.*, vol. 29, no. 12, pp. 25–34, 1991.
[18] M. R. Gibbard and A. B. Sesay, "Asymmetric signal processing for indoor wireless LAN's," *IEEE Trans. Veh. Technol.*, vol. 48, no. 6, pp. 2053–2064, 1999.
[19] 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, Pacific Grove, Calif, USA, October 2000.
[20] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Space-time transmission using Tomlinson-Harashima precoding," in *Proc. 4th International ITG Conference on Source and Channel Coding (SCC '02)*, pp. 139–147, Erlangen, Germany, January 2002.
[21] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "MIMO precoding for decentralized receivers," in *Proc. IEEE International Symposium on Information Theory (ISIT '02)*, p. 496, Lausanne, Switzerland, June–July 2002.
[22] J. Liu and A. Duel-Hallen, "Tomlinson-Harashima transmitter precoding for synchronous multiuser communications," in *Proc. CISS '03*, Baltimore, Md, USA, March 2003.
[23] 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. International Zurich Seminar on Broadband Communications (IZS '02)*, pp. 7-1–7-6, Zurich, Switzerland, February 2002.
[24] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Linear and nonlinear preequalization/equalization for MIMO systems with long-term channel state information at the transmitter," *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 373–378, 2004.
[25] M. Joham, J. Brehmer, and W. Utschick, "MMSE approaches to multiuser spatio-temporal Tomlinson-Harashima precoding," in *Proc. 5th International ITG Conference on Source and Channel Coding (SCC '04)*, pp. 387–394, Erlangen, Germany, January 2004.
[26] M. Joham, J. Brehmer, A. Voulgarelis, and W. Utschick, "Multiuser spatio-temporal Tomlinson-Harashima precoding for frequency selective vector channels," in *Proc. ITG Workshop on Smart Antennas*, pp. 208–215, Munich, Germany, March 2004.
[27] R. F. H. Fischer, C. Stierstorfer, and J. B. Huber, "Precoding for point-to-multipoint transmission over MIMO ISI channels," in *Proc. International Zurich Seminar on Broadband Communications (IZS '04)*, pp. 208–211, Zurich, Switzerland, February 2004.
[28] R. Hunger, F. Dietrich, M. Joham, and W. Utschick, "Robust transmit zero-forcing filters," in *Proc. ITG Workshop on Smart Antennas*, pp. 130–137, Munich, Germany, March 2004.
[29] 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. International Symposium on Signals, Systems, and Electronics (ISSSE '98)*, pp. 295–300, Pisa, Italy, September 1998.
[30] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, NY, USA, 1987.
[31] G. Golub and C. V. Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Md, USA, 1996.

[32] 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. IEEE Global Telecommunications Conference (GLOBECOM '02)*, vol. 1, pp. 429–433, Taipei, Taiwan, November 2002.

[33] 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. IEEE 7th International Symposiumon Spread Spectrum Techniques and Applications (ISSSTA '02)*, vol. 1, pp. 9–13, Prague, Czech Republic, September 2002.

[34] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*, John Wiley & Sons, New York, NY, USA, 2002.

[35] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 1997.

Michael Joham: Institute for Circuit Theory and Signal Processing, Munich University of Technology, 80 290 München, Germany

Email: joham@tum.de

Wolfgang Utschick: Institute for Circuit Theory and Signal Processing, Munich University of Technology, 80 290 München, Germany

Email: wolfgang.utschick@tum.de