

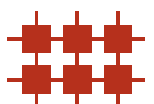
Efficient Tomlinson-Harashima Precoding for Spatial Multiplexing on Flat MIMO Channel

Katsutoshi Kusume, Michael Joham, Wolfgang Utschick, and
Gerhard Bauch

40th IEEE International Conference on Communications
Seoul, Korea

volume 3, pp. 2021–2025, May 16th–20th, 2005

©2005 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.



Efficient Tomlinson-Harashima Precoding for Spatial Multiplexing on Flat MIMO Channel

Katsutoshi Kusume*, Michael Joham†, Wolfgang Utschick†, Gerhard Bauch*

* DoCoMo Communications Laboratories Europe GmbH

Landsbergerstr. 312, 80687 Munich, Germany

Email: {kusume,bauch}@docomolab-euro.com

† Munich University of Technology

Arcisstr. 21, 80290 Munich, Germany

Email: {joham,utschick}@nws.ei.tum.de

Abstract—Nonlinear minimum mean square error Tomlinson-Harashima precoding considered in this paper is an attractive solution for a scenario where a transmitter serves spatially separated receivers and no cooperation among them is possible. Unfortunately, the large performance gain against linear precoding comes along with significantly higher complexity than linear filters in the case of a large number of receivers. We show that superior performance of the nonlinear minimum mean square error Tomlinson-Harashima precoding can be obtained with complexity equivalent to linear precoding. Our proposed algorithm reduces the complexity by a factor of N_R which is the number of receivers.

I. INTRODUCTION

High spectral efficiency is expected in future wireless communication systems. In [1] it was shown that enormous capacity increase is promised on *multiple input multiple output* (MIMO) channels in rich scattering environments. In order to obtain such a capacity advantage on MIMO channels with reasonable complexity, *vertical Bell Labs layered space-time* (V-BLAST) was proposed [2]. V-BLAST can be seen as a block *decision feedback equalizer* (DFE) which iteratively equalizes spatial interference nonlinearly [3].

While V-BLAST suffers from error propagation, a counterpart of DFE, called *spatial Tomlinson-Harashima precoding* (THP) has been proposed in [4] to avoid the error propagation. THP was originally proposed for dispersive *single input single output* (SISO) channels in [5], [6] to avoid intersymbol interference. It moves the feedback filter of the DFE to the transmitter to circumvent the error propagation under the assumption that the channel is known at the transmitter. The same principle can be applied to resolve the spatial interference as proposed in [4]. However, because of the feedforward filter remained at the receiver, all the received signals must be cooperatively processed. That may not be possible in some scenarios, e.g. when the receive antennas belong to spatially separated users.

An interesting approach of spatial THP has been proposed in [7], [8]. This approach moves not only the backward filter but also the forward filter to the transmitter. This architecture enables very simple receivers and more importantly, no signal processing among different receive antennas is necessary. A

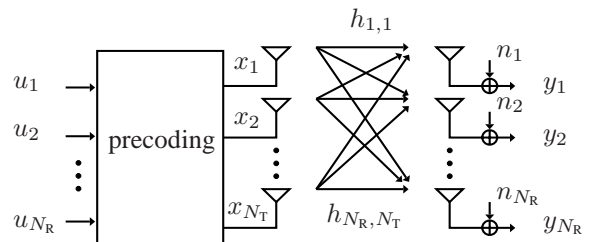


Fig. 1. System model of precoding over flat MIMO channel. There is no cooperation among the receive antennas.

particularly interesting situation is that one transmitter is serving for decentralized receivers or users where no cooperation among receivers is possible. Unfortunately, the complexity of this approach at the transmitter becomes very high for a large number of receivers.

In this paper we propose new computationally efficient *minimum mean square error* (MMSE) THP algorithms. Optimum and suboptimum solutions are presented. Our optimum solution achieves equivalent performance of [7] with significantly lower complexity. Further complexity reduction is possible with our suboptimum solution, whose complexity is equivalent to that of simple linear filters. Although the suboptimum solution shows slight performance degradation, as we will show, it is almost negligible.

This paper is organized as follows. Our system model is introduced in Section II and the MMSE THP is thoroughly reviewed in Section III. Our proposed algorithms are presented in Section IV and the complexity is analyzed in Section V. Section VI shows simulation results, then this paper is summarized in Section VII.

II. SYSTEM MODEL

We consider a system equipped with N_T transmit antennas and N_R receive antennas where $N_T \geq N_R$. We assume narrow band signals, i.e. a non-dispersive fading channel. The discrete time system model in the equivalent complex baseband is illustrated in Fig. 1. The inputs $u_i, i = 1, \dots, N_R$ are complex valued baseband signals and are filtered by the precoding algorithm, which is the main focus of this paper and is explained in the next section. Its output signals $x_i, i = 1, \dots, N_T$ are transmitted from N_T antennas simultaneously.

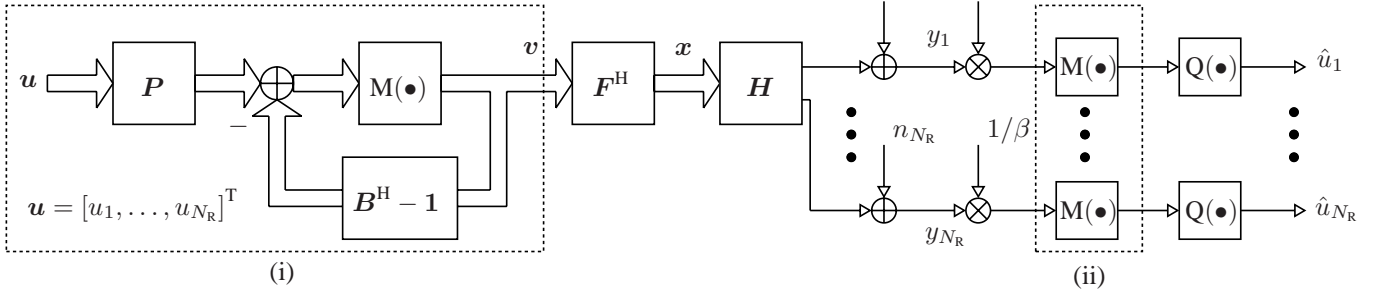


Fig. 2. Block diagram of Tomlinson-Harashima transmission. Also see alternative representations of the sub-blocks (i) and (ii) in Fig. 4.

The channel tap gain from transmit antenna i to receive antenna j is denoted by $h_{j,i}$. These channel taps are assumed to be independent zero mean complex Gaussian variables of equal variance $\mathbb{E}[|h_{j,i}|^2] = 1$. This assumption of independent paths holds if antenna spacing is sufficiently large and the system is surrounded by rich scattering environments.

The signal at receive antenna j can be expressed by

$$y_j = \sum_{i=1}^{N_T} h_{j,i} x_i + n_j, \quad (1)$$

where n_j is additive noise at the receive antenna j . By collecting (1) for N_R receive antennas, the receive signals can be concisely expressed in matrix form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2)$$

where $[\mathbf{H}]_{j,i} = h_{j,i}$, $\mathbf{y} = [y_1, \dots, y_{N_R}]^T$, $\mathbf{x} = [x_1, \dots, x_{N_T}]^T$, $\mathbf{n} = [n_1, \dots, n_{N_R}]^T$, and $(\bullet)^T$ denotes transposition. Our goal is to design a computationally efficient precoding algorithm which maps the input signals $\mathbf{u} = [u_1, \dots, u_{N_R}]^T$ to the transmitted signals \mathbf{x} .

III. MMSE TOMLINSON-HARASHIMA PRECODING

This section reviews the MMSE Tomlinson-Harashima precoding scheme presented in [7]. The overall system structure is illustrated as a block diagram in Fig. 2 where the permutation matrix \mathbf{P} is additionally introduced to the system model in [7]. The input signal vector $\mathbf{u} \in \mathbb{C}^{N_R}$ is firstly reordered in such a way that it can take the best benefit from the filters that follow. The ordering is expressed by the permutation matrix defined as

$$\mathbf{P} = \sum_{i=1}^{N_R} \mathbf{e}_i \mathbf{e}_{k_i}^T \in \{0, 1\}^{N_R \times N_R},$$

where \mathbf{e}_i is the i -th column of the $N_R \times N_R$ identity matrix and $\{k_1, \dots, k_{N_R}\}$ denotes the precoding ordering. The signal vector $\mathbf{v} \in \mathbb{C}^{N_R}$ is iteratively filtered by the backward filter $\mathbf{B}^H \in \mathbb{C}^{N_R \times N_R}$ and by the modulo operator $\mathbf{M}(\bullet)$ where $(\bullet)^H$ denotes Hermitian transpose.

The modulo operator is introduced to reduce the signal power increased by \mathbf{B}^H . The modulo operator for a complex variable c is defined as

$$\mathbf{M}(c) = c - \lfloor \text{Re}(c)/\kappa + 1/2 \rfloor \kappa - j \lfloor \text{Im}(c)/\kappa + 1/2 \rfloor \kappa, \quad (3)$$

where the floor operator $\lfloor \bullet \rfloor$ gives the integer smaller than or equal to the argument and the constant κ is determined

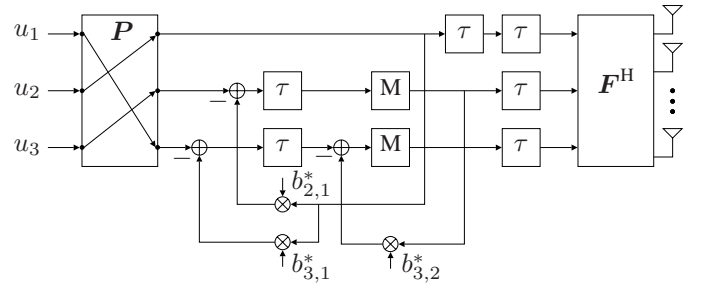


Fig. 3. Example of THP structure for three data streams. τ denotes delay.

TABLE I
ITERATIVE PRECODING PROCEDURE WITH ORDERING.

$$\begin{aligned} \mathbf{v} &= \mathbf{P}\mathbf{u} \\ \text{for } i &= 1, \dots, N_R \\ \mathbf{v}(i) &= \mathbf{M}(\mathbf{v}(i) - \mathbf{B}^H(:, i)\mathbf{v}) \\ \mathbf{x} &= \mathbf{F}^H \mathbf{v} \end{aligned}$$

by the modulation alphabet used, e.g. $\kappa = 2\sqrt{2}$ for QPSK symbols [7], [9]. Consequently, the output signal v_j from $\mathbf{M}(\bullet)$ is an element of the set $\mathbb{M} = \{x + jy \mid x, y \in [-\kappa/2, \kappa/2)\}$ and has the variance of $\sigma_v^2 = \kappa^2/6$ under the assumption that v_j is uniformly distributed both in real and imaginary parts.

The output \mathbf{v} from the feedback section is finally processed by the forward filter $\mathbf{F}^H \in \mathbb{C}^{N_T \times N_R}$ to get the transmit signal \mathbf{x} . The complete precoding procedure is concisely summarized in Table I. Note that there are two constraints to be satisfied by precoding filters. The first one is to limit the total transmit power to certain value E_s . The second constraint is imposed on the backward filter which must be strictly triangular. This can be better understood by the example in Fig. 3. This shows a case of three data streams. The data streams after the permutation \mathbf{P} are precoded from the upper to the lower stream. The filter coefficients can be arranged in the matrix $(\mathbf{B}^H - \mathbf{1})$ which is strictly lower triangular when \mathbf{B}^H is unit lower triangular. The triangular structure ensures the causality of the feedback process.

The receive signal y_j at the j -th antenna is multiplied with the automatic gain control $1/\beta \in \mathbb{R}_+$ which is determined by pilot signals, then the modulo operator $\mathbf{M}(\bullet)$ is applied to get rid of the respective effect of $\mathbf{M}(\bullet)$ at the transmitter. The quantizer $\mathbf{Q}(\bullet)$ generates the estimate \hat{u}_j .

Our goal is to jointly optimize the three tuple of \mathbf{P} , \mathbf{B}^H , and \mathbf{F}^H . In order to formulate the joint optimization using only linear equations, the nonlinear modulo operator in Fig. 2 is interpreted by the linear representation as shown in Fig. 4.

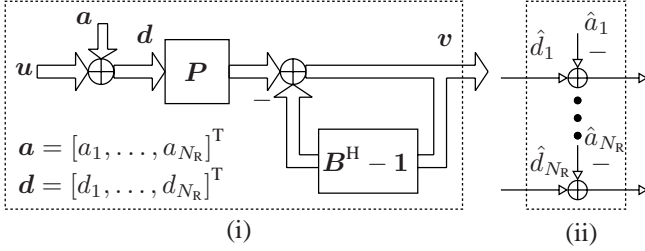


Fig. 4. Alternative linear representation of modulo operators in Fig. 2.

We introduce the signals \mathbf{a} and \mathbf{d} which satisfy \mathbf{v} to be in the set \mathbb{M} . The real and imaginary parts of \mathbf{a} are integer multiples of the constant κ . Indeed, \mathbf{a} and $\hat{\mathbf{a}}$ are implicitly chosen by the modulo operation in (3). Note that the value of \mathbf{a} is not our main interest, but the optimization can be formulated using only linear equations with respect to the desired signal \mathbf{d} .

The signal \mathbf{v} after the backward filtering can be written as $\mathbf{v} = \mathbf{P}\mathbf{d} - (\mathbf{B}^H - \mathbf{1})\mathbf{v}$, which we solve for \mathbf{d} yielding $\mathbf{d} = \mathbf{P}^T \mathbf{B}^H \mathbf{v}$, while the estimated desired signal at the receiver reads as $\hat{\mathbf{d}} = \beta^{-1} \mathbf{H} \mathbf{F}^H \mathbf{v} + \beta^{-1} \mathbf{n}$. By defining the error signal $\boldsymbol{\varepsilon} = \hat{\mathbf{d}} - \mathbf{d}$, the error covariance matrix is calculated as

$$\mathbb{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^H] = \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \beta^{-2} \mathbf{H} \mathbf{F}^H \boldsymbol{\Phi}_{vv} \mathbf{F} \mathbf{H}^H + \mathbf{P}^T \mathbf{B}^H \boldsymbol{\Phi}_{vv} \mathbf{B} \mathbf{P} - 2\beta^{-1} \text{Re}(\mathbf{H} \mathbf{F}^H \boldsymbol{\Phi}_{vv} \mathbf{B} \mathbf{P}) + \beta^{-2} \boldsymbol{\Phi}_{nn}, \quad (4)$$

where we assume $\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \boldsymbol{\Phi}_{nn}$, $\mathbb{E}[\mathbf{v}\mathbf{n}^H] = \mathbf{0}$,

$$\mathbb{E}[\mathbf{v}\mathbf{v}^H] = \boldsymbol{\Phi}_{vv} = \text{diag}(\sigma_{v_1}^2, \dots, \sigma_{v_{N_R}}^2),$$

and $\sigma_{v_j}^2 = \sigma_v^2 = \kappa^2/6$ for $j \neq 1$ and $\sigma_{v_1}^2 = \sigma_u^2 = \mathbb{E}[|u_i|^2] \forall i$. With the error covariance matrix, the MSE is expressed as

$$\varphi = \mathbb{E}[\boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon}] = \text{tr}(\boldsymbol{\Phi}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}}), \quad (5)$$

where ‘tr’ denotes the trace operator. This is the cost function to be minimized by the optimization. With two constraints which we already discussed, the optimization problem is formulated as

$$\begin{aligned} \{\mathbf{F}_{\text{opt}}^H, \mathbf{B}_{\text{opt}}^H, \beta_{\text{opt}}\} &= \underset{\{\mathbf{F}^H, \mathbf{B}^H, \beta\}}{\text{argmin}} \varphi \\ \text{s.t.} & \mathbb{E}[\|\mathbf{x}\|_2^2] = E_s \quad \text{and} \\ \mathbf{S}_i (\mathbf{B}^H - \mathbf{1}) \mathbf{e}_i &= \mathbf{0}_i \quad \text{for } i = 1, \dots, N_R \end{aligned} \quad (6)$$

where \mathbf{S}_i is defined as $\mathbf{S}_i = [\mathbf{1}_i, \mathbf{0}_{i \times N_R - i}] \in \{0, 1\}^{i \times N_R}$. Note that the constraint for lower triangular structure is defined for every column of the backward filter so that its upper triangular part must be zero. Note also that, as seen from (6), the optimization does not take into account \mathbf{P} , therefore its optimality depends on the choice of \mathbf{P} . That will be discussed in the next section. This optimization problem can be solved using Lagrangian multipliers and we get

$$\begin{aligned} \mathbf{B}_{\text{opt}}^H &= \sum_{i=1}^{N_R} \mathbf{P} \boldsymbol{\Phi} \mathbf{P}^T \mathbf{S}_i^T (\mathbf{S}_i \mathbf{P} \boldsymbol{\Phi} \mathbf{P}^T \mathbf{S}_i^T)^{-1} \mathbf{S}_i \mathbf{e}_i \mathbf{e}_i^T, \\ \mathbf{F}_{\text{opt}}^H &= \beta_{\text{opt}} \sum_{i=1}^{N_R} \mathbf{H}^H \mathbf{P}^T \mathbf{S}_i^T (\mathbf{S}_i \mathbf{P} \boldsymbol{\Phi} \mathbf{P}^T \mathbf{S}_i^T)^{-1} \mathbf{S}_i \mathbf{e}_i \mathbf{e}_i^T, \end{aligned} \quad (7)$$

where we define $\boldsymbol{\Phi} = \mathbf{H} \mathbf{H}^H + \gamma^{-1} \mathbf{1}$, and $\gamma = E_s / \text{tr}(\boldsymbol{\Phi}_{nn})$. The optimum scalar β_{opt} can be easily calculated to satisfy the transmit power constraint in (6).

As we can see from (7), the filters are determined column by column, each of which requires one matrix inverse resulting in the total complexity order of $\mathcal{O}(N_R^4)$. That becomes quite complex for large N_R . Detailed complexity analysis is given in Section V. Complete description of the algorithm including ordering strategy based on (7) can be found in [7].

IV. PROPOSED EFFICIENT ALGORITHMS

This section presents our computationally efficient algorithms. One is the suboptimum solution requiring complexity equivalent to only one matrix inverse with slight performance degradation. Our optimum solution performs equivalent to (7) with the complexity slightly higher than the suboptimum one.

A. Suboptimum Ordered Cholesky

One can observe that the solution in (7) has the term involving $\boldsymbol{\Phi}$ which has to be inverted N_R times. Since $\boldsymbol{\Phi}$ is Hermitian and positive definite, the Cholesky factorization with symmetric permutation [10] can be calculated as

$$\mathbf{P} \boldsymbol{\Phi} \mathbf{P}^T = \mathbf{L} \mathbf{D} \mathbf{L}^H, \quad (8)$$

where \mathbf{P} , \mathbf{L} , and $\mathbf{D} = \text{diag}(d_1, \dots, d_{N_R})$ are permutation, unit lower triangular, and diagonal matrices, respectively. With (8), it can be shown that the solution in (7) reduces to

$$\mathbf{B}_{\text{subopt}}^H = \mathbf{L} \quad \text{and} \quad \mathbf{F}_{\text{subopt}}^H = \beta_{\text{subopt}} \mathbf{H}^H \mathbf{P}^T \mathbf{L}^{H,-1} \mathbf{D}^{-1}. \quad (9)$$

This is significant computational reduction, namely one factorization in (8) and the inversion of the triangular matrix in (9) suffice instead of N_R times matrix inversions. This complexity corresponds to only one matrix inversion of the Hermitian positive definite matrix $\boldsymbol{\Phi}$.

The MSE φ in (5) can be also simplified using the results in (9) and (4) as follows

$$\varphi = \gamma^{-1} \sum_{i=1}^{N_R} \sigma_{v_i}^2 d_i^{-1}. \quad (10)$$

This means that the MSE is determined by the diagonal entries of \mathbf{D}^{-1} weighted by $\sigma_{v_i}^2$. The optimality of (9) depends on the way to compute the factorization in (8). A successive algorithm computing (8) finds the diagonal entry starting from d_1 up to d_{N_R} with necessary permutation (cf. [10]). Our proposed pseudo code for the filter calculation is summarized in Table II. Its suboptimality is discussed in the next section.

B. Ordering Strategy for Filter Optimization and Precoding

It was shown in [2] that the *best-first* successive detection strategy (the layer with maximum post SNR is detected first) improves the performance of the weakest layer and leads to the global optimum solution. This strategy can be also applied to the precoding, but there is a difference between the successive detection (V-BLAST) and successive precoding (THP).

In the V-BLAST case, the filter optimization at the later detection stage is less constrained because after each stage one transmit signal less is subject to detection. On the other hand, the precoding filter optimization for the later precoding stage is more constrained since the transmit signal at certain stage

TABLE II
CALCULATION OF THP FILTER WITH SUBOPTIMUM ORDERING.

$$\begin{aligned}
\Phi &= \mathbf{H}\mathbf{H}^H + \frac{N_R}{\gamma} \mathbf{1} \\
\mathbf{P} &= \mathbf{1}_{N_R}, \mathbf{D} = \mathbf{0}_{N_R} \\
\text{for } i &= 1, \dots, N_R \\
q &= \underset{q'=i, \dots, N_R}{\operatorname{argmin}} \sigma_{v_i}^2 \Phi(q', q') \\
\mathbf{P}_i &= \mathbf{1}_{N_R} \text{ whose } i\text{-th and } q\text{-th rows are exchanged} \\
\mathbf{P} &= \mathbf{P}_i \mathbf{P} \\
\Phi &= \mathbf{P}_i \Phi \mathbf{P}_i^T \\
\mathbf{D}(i, i) &= \Phi(i, i) \\
\Phi(i: N_R, i) &= \Phi(i: N_R, i) / \mathbf{D}(i, i) \\
\Phi(i+1: N_R, i+1: N_R) &= \Phi(i+1: N_R, i+1: N_R) \\
&\quad - \Phi(i+1: N_R, i) \Phi(i+1: N_R, i)^H \mathbf{D}(i, i) \\
\mathbf{L} &= \text{lower triangular part of } \Phi \\
\mathbf{B}^H &= \mathbf{L}, \mathbf{F}^H = \mathbf{H}^H \mathbf{P}^T \mathbf{L}^{H,-1} \mathbf{D}^{-1} \\
\chi &= \|\mathbf{F}^H(:, 1)\|_2^2 + \sigma_v^2 \|\mathbf{F}^H(:, 2: N_R)\|_F^2 \\
\beta &= \sqrt{E_s / \chi} \\
\mathbf{F}^H &= \beta \mathbf{F}^H
\end{aligned}$$

should not interfere with the already precoded data streams. Hence, the precoding filter optimization has to be performed in the opposite direction to the precoding ordering. That is why our suboptimum algorithm in Section IV-A does not always lead to the optimum solution. Such optimum solution is possible as presented in the next section.

C. Optimum Ordered Cholesky

In order to have the desired direction of filter optimization as explained in the previous section, the *modified* Cholesky factorization with symmetric permutation is computed for Φ^{-1} instead of Φ (cf. Eq. 8), that is

$$\mathbf{P} \Phi^{-1} \mathbf{P}^T = \mathbf{L}^H \mathbf{D} \mathbf{L}. \quad (11)$$

With this factorization, the backward and forward filters in (7) are calculated as (cf. Eq. 9)

$$\mathbf{B}_{\text{opt}}^H = \mathbf{L}^{-1} \quad \text{and} \quad \mathbf{F}_{\text{opt}}^H = \beta_{\text{opt}} \mathbf{H}^H \mathbf{P}^T \mathbf{L}^H \mathbf{D}. \quad (12)$$

Using (12) and (4), the MSE φ reads as (cf. Eq. 10)

$$\varphi = \gamma^{-1} \sum_{i=1}^{N_R} \sigma_{v_i}^2 d_i. \quad (13)$$

The iterative algorithm computing (11) finds the diagonal entry starting from d_{N_R} down to d_1 with the optimum permutation. This is opposite to the precoding filtering (cf. Table I: for $i=1, \dots, N_R$). This reverse direction is desired as discussed in the previous section. Our optimum algorithm computing the filters is summarized in Table III. In each iteration the algorithm finds the minimum diagonal entry of Φ^{-1} weighted by $\sigma_{v_i}^2$, i.e. $\sigma_{v_i}^2 d_i$ which is the MSE (cf. Eq. 13 and Table III) of the i -th precoded data stream. Therefore, the algorithm iteratively finds the optimum ordering in the MMSE sense.

V. ANALYSIS OF COMPUTATIONAL COMPLEXITY

To compare the complexity, the number of additions and multiplications is separately computed because of their possibly different computational costs for the different processors.

We compare the complexity of our proposed schemes with that of the optimum MMSE THP presented in [7] as

TABLE III
CALCULATION OF THP FILTER WITH OPTIMUM ORDERING.

$$\begin{aligned}
\Phi^{-1} &= (\mathbf{H}\mathbf{H}^H + \frac{N_R}{\gamma} \mathbf{1})^{-1} \\
\mathbf{P} &= \mathbf{1}_{N_R}, \mathbf{D} = \mathbf{0}_{N_R} \\
\text{for } i &= N_R, \dots, 1 \\
q &= \underset{q'=1, \dots, i}{\operatorname{argmin}} \sigma_{v_i}^2 \Phi^{-1}(q', q') \\
\mathbf{P}_i &= \mathbf{1}_{N_R} \text{ whose } i\text{-th and } q\text{-th rows are exchanged} \\
\mathbf{P} &= \mathbf{P}_i \mathbf{P} \\
\Phi^{-1} &= \mathbf{P}_i \Phi^{-1} \mathbf{P}_i^T \\
\mathbf{D}(i, i) &= \Phi^{-1}(i, i) \\
\Phi^{-1}(1:i, i) &= \Phi^{-1}(1:i, i) / \mathbf{D}(i, i) \\
\Phi^{-1}(1:i-1, 1:i-1) &= \Phi^{-1}(1:i-1, 1:i-1) \\
&\quad - \Phi^{-1}(1:i-1, i) \Phi^{-1}(1:i-1, i)^H \mathbf{D}(i, i) \\
\mathbf{L}^H &= \text{upper triangular part of } \Phi^{-1} \\
\mathbf{B}^H &= \mathbf{L}^{-1}, \mathbf{F}^H = \mathbf{H}^H \mathbf{P}^T \mathbf{L}^H \mathbf{D} \\
\chi &= \|\mathbf{F}^H(:, 1)\|_2^2 + \sigma_v^2 \|\mathbf{F}^H(:, 2: N_R)\|_F^2 \\
\beta &= \sqrt{E_s / \chi} \\
\mathbf{F}^H &= \beta \mathbf{F}^H
\end{aligned}$$

TABLE IV
COMPLEXITY OF SYSTEMS WITH $N_R = N_T$ ANTENNAS FOR A PROCESSOR REQUIRING THE SAME OPERATIONS FOR ADDITION AND MULTIPLICATION.

optimum THP in [7]	optimum THP, proposed	suboptimum THP, proposed	linear MMSE
$\frac{13}{6} N_R^4$	$\frac{7}{2} N_R^3$	$\frac{13}{6} N_R^3$	$\frac{13}{6} N_R^3$

well as simple linear MMSE transmit filter. The complexity of the linear transmit filter $\beta \mathbf{H}^H \Phi^{-1}$ can be calculated as $\frac{1}{2} N_R^2 N_T + \frac{1}{3} N_R^3$ additions and $N_R^2 N_T + \frac{1}{3} N_R^3$ multiplications. This complexity is due to the inversion of the Hermitian positive definite matrix Φ . This can be done by (8) and (9) since $\Phi^{-1} = \mathbf{L}^{H,-1} \mathbf{D}^{-1} \mathbf{L}^{-1}$ holds for the special case of no ordering ($\mathbf{P} = \mathbf{1}$). Thus, the complexity of our suboptimum solution is equivalent to that of the simple linear transmit filter. The solution in [7] requires N_R times matrix inversion resulting in $\frac{1}{2} N_R^3 N_T + \frac{1}{3} N_R^4$ additions and $N_R^3 N_T + \frac{1}{3} N_R^4$ multiplications. Our optimum solution needs to compute Φ^{-1} in addition to our suboptimum solution. Its complexity is $\frac{1}{2} N_R^2 N_T + N_R^3$ additions and $N_R^2 N_T + N_R^3$ multiplications.

Note that, although the pseudo codes in Table II and III compute the whole elements of the symmetric matrices Φ and Φ^{-1} for the concise description, the extension to compute only the lower triangular part is easy. The complexity analysis presented above makes use of such triangular structure.

Table IV summarizes the complexities as an example for a system with $N_R = N_T$ antennas using a processor which requires the same number of operations for additions and multiplications. Comparing to the optimum THP in [7], our suboptimum and optimum algorithms reduce the complexity by a factor of N_R and $0.6 N_R$, respectively.

VI. SIMULATION RESULTS

Computer simulations are performed to evaluate the uncoded BER performance over $E_b/N_0 = \frac{E_s}{N_R N \sigma_n^2}$ where N is the number of information bits per channel input and the noise is assumed to be white, i.e. $\Phi_{nn} = \sigma_n^2 \mathbf{1}$. In the following, information bits are QPSK modulated ($N = 2$). Fig. 5 shows the performance of a system with $N_T = N_R = 4$ antennas.

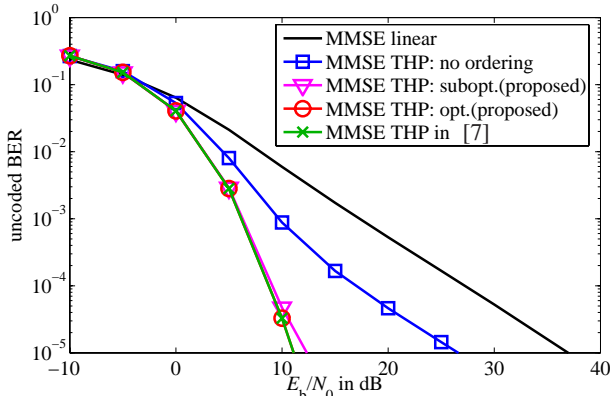


Fig. 5. BER Performance for $N_T = N_R = 4$. Our proposed optimum algorithm performs equivalent to the optimum scheme in [7]. The performance of our suboptimum algorithm is approaching the optimum performance.

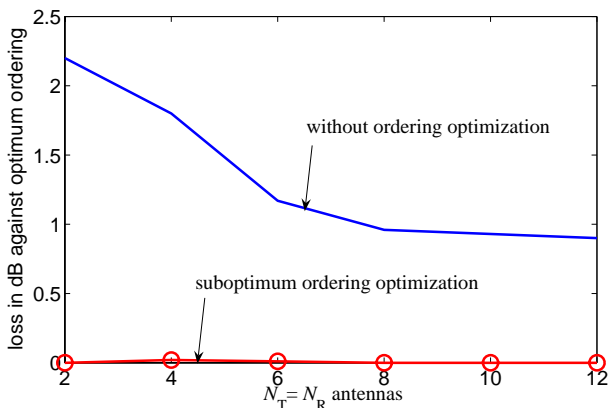


Fig. 6. Performance loss in dB due to the ordering algorithm at an uncoded BER= 10^{-2} . The loss of the proposed suboptimum solution is negligible.

The impact of the ordering optimization can be observed. For comparison, the performance of the MMSE linear filter is also plotted. Significant gain of the nonlinear THP can be seen against the linear filter. Our optimum algorithm shows exactly the same performance of [7]. Our suboptimum algorithm shows slight performance degradation in the low uncoded BER region, but no performance degradation in the uncoded BER, e.g. 10^{-2} which is a practical operating point in coded transmission.

The performance loss due to the suboptimality of the ordering optimization is further investigated. Fig. 6 shows the loss in dB against the optimum ordering at an uncoded BER = 10^{-2} over the number of antenna elements ($N_T = N_R$). We conclude that the performance loss of our suboptimum solution is negligible (below 0.03 dB).

Fig. 7 compares the performance of different transmission schemes. Our proposed algorithm makes it possible to obtain the superior performance of the nonlinear MMSE THP without any computational complexity penalty against simple linear schemes such as zero forcing (ZF) or MMSE linear filters. The THP approach requires a modulo operation, but its complexity is of minor impact on the total complexity.

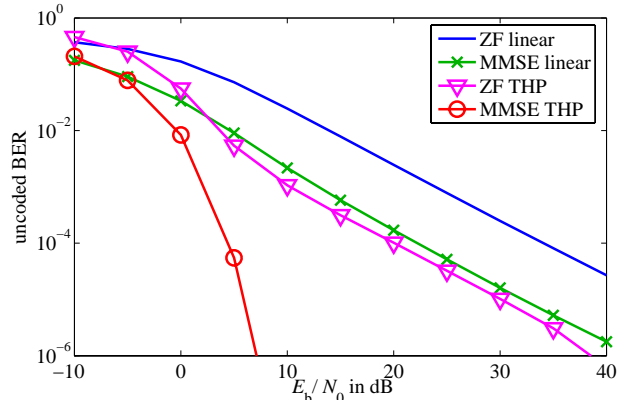


Fig. 7. BER performance of linear and nonlinear transmit processing schemes for $N_T = N_R = 8$. The superior performance of MMSE THP can be obtained with the same complexity of the simple linear scheme by our algorithm.

VII. SUMMARY

We have derived new computationally efficient MMSE THP algorithms for spatial multiplexing on flat MIMO channels. The system considered in this paper is attractive in scenarios where a transmitter serves decentralized receivers which cannot cooperate with each other. Interferences among different receivers are resolved at the transmitter prior to the signal transmission based on the MIMO channel state information. As shown in [7], the performance of the nonlinear MMSE THP is superior to linear variants. However, the complexity is significantly higher for a large number of receivers. We showed that a large performance advantage is possible without any complexity penalty compared to the simple linear filters.

REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, March 1998.
- [2] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified Processing for High Spectral Efficiency Wireless Communication Employing Multi-Element Arrays," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1841–1852, November 1999.
- [3] K. Kusume, M. Joham, and W. Utschick, "MMSE Block Decision-Feedback Equalizer for Spatial Multiplexing with Reduced Complexity," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM 2004)*, vol. 4, Nov/Dec 2004, pp. 2540–2544.
- [4] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "Space-Time Transmission using Tomlinson-Harashima Precoding," in *Proc. of 4. ITG Conference on Source and Channel Coding*, Berlin, January 2002, pp. 139–147.
- [5] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, no. 5/6, pp. 138–139, March 1971.
- [6] 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.
- [7] M. Joham, J. Brehmer, and W. Utschick, "MMSE Approaches to Multiuser Spatio-Temporal Tomlinson-Harashima Precoding," in *Proc. of 5. ITG Conference on Source and Channel Coding*, Erlangen, January 2004, pp. 387–394.
- [8] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "MIMO Precoding for Decentralized Receivers," in *Proc. IEEE International Symposium on Information Theory (ISIT 2002)*, June/July 2002, p. 496.
- [9] J. M. Cioffi and G. P. Dudevoir, "Data Transmission with Mean-Square Partial Response," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM'89)*, vol. 3, November 1989, pp. 1687–1691.
- [10] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, 1996.