Hierarchical Precoding for the Network MIMO Downlink

Sebastian Stern and Robert F.H. Fischer
Institut für Nachrichtentechnik, Universität Ulm, Ulm, Germany
Email: {sebastian.stern, robert.fischer}@uni-ulm.de

***Abstract*— Over the last few years, network MIMO or distributed MIMO has gained enormous interest in research and meanwhile is a candidate for future communication systems. However, existing techniques handling the multi-user interference, e.g., linear preequalization (LPE) or Tomlinson-Harashima precoding (THP), assume full cooperation between the base stations. This might cause significant signaling overhead in the backhaul. In this paper, precoding strategies with reduced coordination effort are presented and assessed. We show how to optimize THP for a setting, where a hierarchy of increasing knowledge over the base stations is expected. A reduced-complexity variant is given. In addition, besides the usual sum power constraint, a per-antenna average power constraint is included in the optimization, which is of relevance in distributed MIMO. The performance of the proposed strategies is covered by means of numerical simulations.**

I. Introduction

In the recent literature, multiuser multiple-input/multiple-output (MIMO) systems have been extended to the principle of *network MIMO*, which is a promising technique for future communication systems, e.g., [7], [8]. In this paper, we study the *network MIMO downlink*, i.e., base stations (BSs) are grouped to supply jointly non-cooperating user equipments (UEs) within a certain service area.

In order to handle the multi-user interference, well-known preequalization or precoding schemes can be employed. Specifically, *linear preequalization* (LPE) and *Tomlinson-Harashima precoding* (THP) have been proposed, e.g., [3], [4], [6], [12]. In the sequel, design criteria, different from the classical ones (zero-forcing (ZF) or minimum mean-squared error (MMSE)), have been applied to incorporate *user balancing* or/and *per-antenna power constraints* for LPE, e.g., [10], [11], [13]. Recently, THP with per-antenna power constraints has been presented in [5].

Inherently, in all the abovementioned work, a central processing unit, having (perfect) channel knowledge and knowing all the data to be transmitted to the UEs, is assumed. In this situation of *full coordination*, the entire user data and the processed/generated transmit symbols have to be communicated via *backhaul links*. In order to decrease the effort of coordination and hence the backhaul traffic, a *decentralized, partial coordination* is of interest. In particular, a successive or hierarchical scheme, where each BS processes only its own data but passes the generated transmit symbols to the next node, seems to be appropriate.

This work was supported by Deutsche Forschungsgemeinschaft (DFG) within the framework COIN under grant Fi 982/4-2.

In this paper, we design preequalization and precoding with only partial coordination in a decentralized setting. Noteworthy, we reduce the effort for communicating the data symbols per discrete time step. Assuming a block fading channel, the initial optimization of the precoding scheme and sorting of the BSs and UEs is done by a central instance. Since, in the optimum, this initial calculation requires high effort, a low-complexity heuristic approximation is proposed. Moreover, a per-antenna power constraint is also included.

The paper is organized as follows: In Sec. II, a geometrical system model for the network MIMO downlink is briefly reviewed. Sec. III presents hierarchical LPE and THP schemes; a reduced-complexity version of determining the (almost) optimum precoding matrices is discussed in Sec. IV. All schemes are assessed by means of numerical simulations. The paper closes with a brief summary and conclusion in Sec. V.

II. System Model

The geometrical network MIMO model, taken from [5], [8], is shown in Fig. 1. $N_{\mathrm{B}} = 3$ cooperating base stations (at inter-site distance of $r_{\mathrm{BS}} = 500$ m) serve $N_{\mathrm{U}} = 3$ non-cooperating (single-antenna) user equipments, uniformly distributed over the dark gray shaded area.[1] UEs located close to a BS (here: $r_{\min} = 125$ m) are excluded from the network MIMO processing, as they anyway have very good channel conditions compared to cell-edge UEs.

The link[2] between BS b and UE u is modeled via the complex channel gain

$$h_{u,b} = 10^{-L_{\mathrm{dB}}(r_{u,b})/20} \cdot 10^{A_{\mathrm{dB}}(\theta_{u,b})/20} \cdot h_{u,b,\mathrm{i.i.d.}} \,. \quad (1)$$

As specified in [1], [2], [5], this model includes the antenna pattern $A_{\mathrm{dB}}(\theta_{u,b}) = \min\left\{12(\theta_{u,b}/70^\circ)^2,\ 20\right\}$ (UE u and BS b at distance $r_{u,b}$ and angle $\theta_{u,b}$, cf. Fig. 1), attenuation, path loss, and shadowing $L_{\mathrm{dB}}(r_{u,b}) = D_{\mathrm{dB}} + 37.6\log_{10}\left(r_{u,b}/[\mathrm{km}]\right) + S_{\mathrm{dB}}$, where S_{dB} is zero-mean Gaussian with 8 dB standard derivation. The fast fading is modeled via the i.i.d. zero-mean unit-variance Gaussian random variable $h_{u,b,\mathrm{i.i.d.}}$.

[1] Assuming the BSs to be part of larger configuration of BSs placed on a hexagonal grid, UEs outside the dark gray area are supplied by a different set of BSs.

[2] The T-spaced complex baseband model includes transmitter-side pulse shaping, the actual channel, and the receiver-side matched filtering. If the channel is not flat, this model can be seen as per carrier in a multi-carrier transmission system.

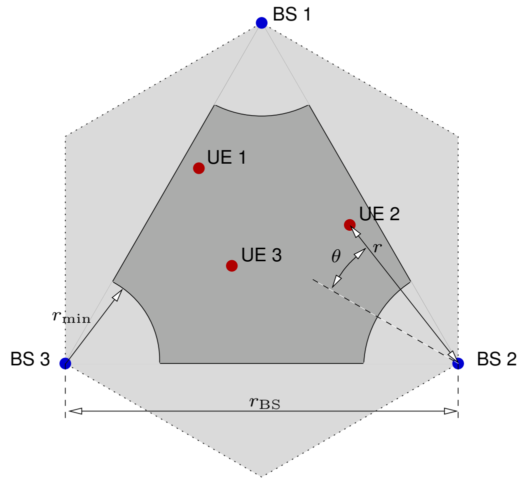

Fig. 1. Geometrical network MIMO system model for $N_{\mathrm{B}} = N_{\mathrm{U}} = 3$, taken from [5].

Combining all channel coefficients into the channel matrix

$$\boldsymbol{H} = \left[\, h_{u,b} \,\right]_{\substack{u=1,\ldots,N_{\mathrm{U}} \\ b=1,\ldots,N_{\mathrm{B}}}} \;, \tag{2}$$

the MIMO system model is given as[3]

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \;, \tag{3}$$

where $\boldsymbol{x} \overset{\text{def}}{=} [x_1, \ldots, x_{N_{\mathrm{B}}}]^{\mathsf{T}}$ denotes the vector of transmit symbols x_b at the BSs $b = 1, \ldots, N_{\mathrm{B}}$, and $\boldsymbol{y} \overset{\text{def}}{=} [y_1, \ldots, y_{N_{\mathrm{U}}}]^{\mathsf{T}}$ the vector of receive symbols y_u of UEs $u = 1, \ldots, N_{\mathrm{U}}$. Finally, $\boldsymbol{n} \overset{\text{def}}{=} [n_1, \ldots, n_{N_{\mathrm{U}}}]^{\mathsf{T}}$ is a vector of i.i.d. complex-valued zero-mean Gaussian noise with covariance matrix $\mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{H}}\} = \sigma_n^2 \boldsymbol{I}$. The variance per component is given by $\sigma_n^2 \overset{\text{def}}{=} N_0/T$, where N_0 denotes the one-sided noise power spectral density.

Given the data symbols a_u, $u = 1, \ldots, N_{\mathrm{U}}$, (drawn from an M-ary QAM constellation with variance σ_a^2) intended for the users, and assuming full channel knowledge, via a preequalization/precoding scheme the vector $\boldsymbol{x}$ of transmit symbols is calculated, cf. Fig 2.

Expecting zero-mean symbols, the average transmit power at each antenna calculates to

$$\sigma_{x_b}^2 \overset{\text{def}}{=} \mathrm{E}\{|x_b|^2\} \;, \quad b = 1, \ldots, N_{\mathrm{B}} \;. \tag{4}$$

In each case, the following *sum power constraint* (SPC) is imposed

$$\sum\nolimits_{b=1}^{N_{\mathrm{B}}} \sigma_{x_b}^2 = \mathrm{E}\{\boldsymbol{x}^{\mathsf{H}}\boldsymbol{x}\} \le P_{\mathrm{sum}} = N_{\mathrm{U}}\sigma_a^2 \;. \tag{5}$$

In addition, a *per-antenna power constraint* (PPC) is meaningful in the network MIMO setting. We demand

$$\mathrm{E}\{|x_b|^2\} \le P_{\mathrm{per}} \;, \quad b = 1, \ldots, N_{\mathrm{B}} \;. \tag{6}$$

Utilizing the maximum sum power, the *signal-to-noise ratio* (SNR), expressed as the ratio of *transmitted energy per information bit* and noise power spectral density reads

$$\frac{E_{\mathrm{b}}}{N_0'} = \frac{\sigma_a^2}{\sigma_n^2 \log_2(M)} \;, \tag{7}$$

as $E_{\mathrm{b}} = N_{\mathrm{U}}\sigma_a^2 T/(N_{\mathrm{U}} \log_2(M))$ and $N_0' = N_0 \cdot 10^{D_{\mathrm{dB}}/10}$ gets rid of the basic attenuation D_{dB}. For details see [5].

At the receivers, only appropriate individual scaling (diagonal scaling matrix $\boldsymbol{G} \overset{\text{def}}{=} \mathbf{diag}(g_1, \ldots, g_{N_{\mathrm{U}}})$) is required. In case of uncoded transmission, a threshold device produces estimates $\hat{a}_u$ of the data symbols.

III. Preequalization and Precoding Schemes for Decentralized Processing

In the network MIMO scenario at hand we have to distinguish two coordination tasks. On the one hand, given the channel, the optimum precoding matrices have to be calculated. In a bursty transmission this is done once, at the beginning of each transmission frame. Throughout the paper we assume a central instance, having perfect channel knowledge, to carry out this optimization task. On the other hand, the transmit symbols x_b, $b = 1, \ldots, N_{\mathrm{B}}$, have to be calculated from the data symbols a_u, $u = 1, \ldots, N_{\mathrm{U}}$, for each time step within the burst. Subsequently, we concentrate ourselves on this coordination/calculation effort.

A. Coordination and Decentralized Processing

Classically, a central unit (CU) would be present, where all user data a_u are communicated to,[4] which calculates the transmit symbols x_b, and communicates them to the BSs, which then radiate these symbols. This strategy is shown on the left of Fig. 3. Clearly, such a procedure requires a large amount of traffic in the backhaul. For our specific scenario, per discrete time step 6 complex numbers have to be communicated.

Omitting the central unit, the processing and distribution of the data can equivalently be done as follows, cf. the middle part of Fig. 3. Each BSs receives only the data symbol of one user. Since THP performs successive encoding [4], BS 1 can calculate $\tilde{x}_1 = a_1$ (cf. Fig. 2). This BS communicates $\tilde{x}_1$ to BSs 2 and 3. Knowing a_2, $\tilde{x}_1$ and the element $b_{2,1}$ of the feedback matrix $\boldsymbol{B} = [b_{u,\nu}]$, BS 2 is able to calculate[5] $\tilde{x}_2 = \mathrm{mod}(a_2 - b_{2,1}\tilde{x}_1)$; this symbol is communicated to BSs 1 and 3. Finally, BS 3 calculates $\tilde{x}_3 = \mathrm{mod}(a_3 - b_{3,1}\tilde{x}_1 - b_{3,2}\tilde{x}_2)$ from a_3 and the knowledge of the already encoded symbols $\tilde{x}_1$ and $\tilde{x}_2$. Additionally, the coefficients $b_{3,1}$ and $b_{3,2}$ have to be known. After BS 3 has communicated $\tilde{x}_3$ back to the other BSs, all BSs know all encoded symbols $\tilde{x}_u$. Via the respective row $\boldsymbol{f}_b^{\mathsf{T}}$ of the feedforward matrix $\boldsymbol{F} = [\boldsymbol{f}_1 \cdots \boldsymbol{f}_{N_{\mathrm{B}}}]^{\mathsf{T}}$, each BSs is now able to calculate its transmit symbol $x_b = \sum_{u=1}^{N_{\mathrm{U}}} f_{b,u}\tilde{x}_u$ individually. For linear preequalization the same holds for $\tilde{\boldsymbol{x}} = \boldsymbol{a}$. As can be seen, in this situation the communication of 6 complex symbols is required, too.

[3]Notation: $\Pr\{\cdot\}$: probability. $\mathrm{E}\{\cdot\}$: expectation. $\boldsymbol{I}$: identity matrix. $\boldsymbol{A}^{\mathsf{T}}$: transpose of matrix $\boldsymbol{A}$. $\boldsymbol{A}^{\mathsf{H}}$: Hermitian of matrix $\boldsymbol{A}$. $\mathbf{diag}(\cdot)$: diagonal matrix with given elements.

[4]We assume that as in the case of no cooperation the data symbols a_u are already present at the base stations.

[5]$\mathrm{mod}(\cdot)$: symmetrical modulo operation [4]; interval according to the respective context (e.g., 16QAM with components at ± 1, ± 3: one-dimensional modulo reduction to $[-4, 4)$ for both components).

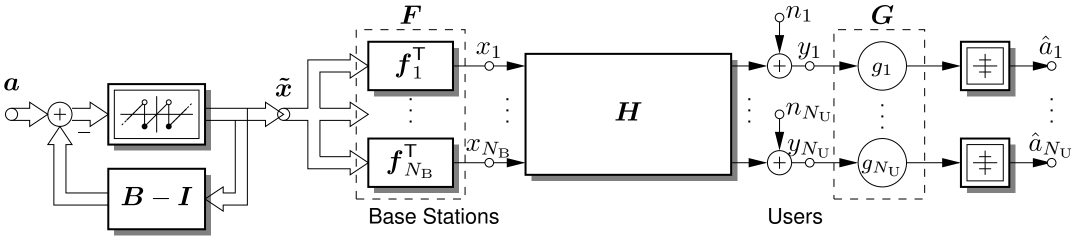

Fig. 2. Block diagram for network MIMO employing Tomlinson-Harashima precoding.

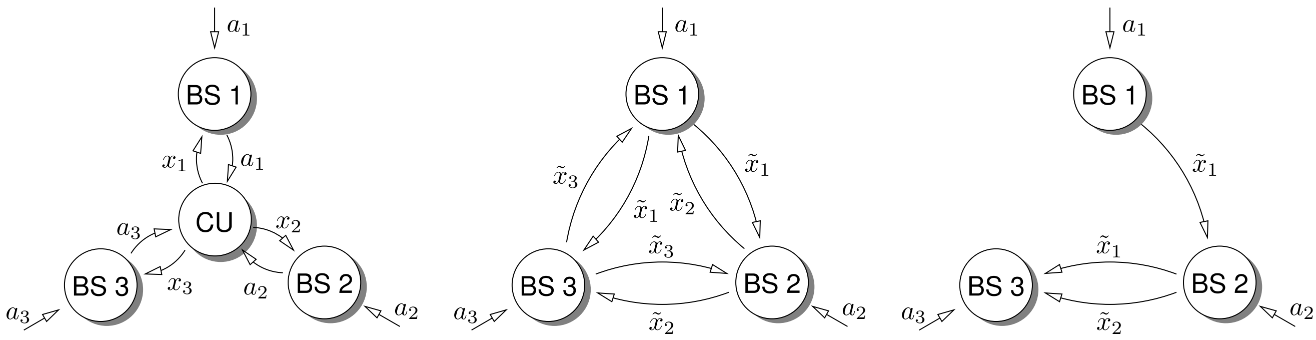

Fig. 3. Network MIMO processing scenarios. Left: Centralized; Middle: Decentralized with full cooperation; Right: Decentralized with hierarchical coordination.

B. Precoding in Network MIMO with Full Cooperation

We now briefly review the optimization of fully-coordinated preequalization/precoding in network MIMO as discussed in [5].

1) Linear Preequalization: The simplest version to handle the interference in the current situation is *linear preequalization*, which is obtained for $\tilde{\boldsymbol{x}} = \boldsymbol{a}$ (i.e., switching off the feedback loop via $\boldsymbol{B} = \boldsymbol{I}$) in Fig. 2. In [10], [11], [13] an optimization according to the *signal-to-interference-plus-noise ratio (SINR)* has been given. Denoting the end-to-end cascade (from data to receive symbols) by $\boldsymbol{C}^{(\mathsf{LPE})} = [c_{u,b}^{(\mathsf{LPE})}] \stackrel{\text{def}}{=} \boldsymbol{HF}$, the SINR of user u is given as [5]

$$\mathrm{SINR}_u^{(\mathsf{LPE})} \stackrel{\text{def}}{=} \frac{|c_{u,u}^{(\mathsf{LPE})}|^2}{\zeta + \sum\limits_{l=1,\ldots,N_\mathrm{U},\, l\neq u} |c_{u,l}^{(\mathsf{LPE})}|^2} , \qquad (8)$$

where $\zeta \stackrel{\text{def}}{=} \sigma_n^2/\sigma_a^2$ denotes the inverse SNR. The optimization task (min SINR) then reads

$$\boldsymbol{F}_{\mathsf{opt}} = \underset{\boldsymbol{F}}{\mathrm{argmax}} \min_{u=1,\ldots,N_\mathrm{U}} \mathrm{SINR}_u^{(\mathsf{LPE})} , \qquad (9)$$

under sum and average power constraints (5) and (6). This optimization can be performed by a *second-order cone program*, particularly applying Algorithm 2 in [10].

2) Tomlinson-Harashima Precoding: Especially under per-antenna constraints, THP provides superior performance compared to LPE [5]. The respective block diagram is shown in Fig. 2.

Since interference of already encoded symbols is perfectly eliminated, the SINR of user u in case of THP is given by [5]

$$\mathrm{SINR}_u^{(\mathsf{THP})} \stackrel{\text{def}}{=} \frac{|c_{u,u}^{(\mathsf{THP})}|^2}{\zeta + \sum\limits_{l=1,\ldots,N_\mathrm{U},\, l> u} |c_{u,l}^{(\mathsf{THP})}|^2} . \qquad (10)$$

Here, the end-to-end cascade is again defined as $\boldsymbol{C}^{(\mathsf{THP})} = [c_{u,b}^{(\mathsf{THP})}] \stackrel{\text{def}}{=} \boldsymbol{HF}$. The optimization task now reads

$$\boldsymbol{F}_{\mathsf{opt}} = \underset{\boldsymbol{F}}{\mathrm{argmax}} \min_{u=1,\ldots,N_\mathrm{U}} \mathrm{SINR}_u^{(\mathsf{THP})} , \qquad (11)$$

obeying sum and average power constraints (5) and (6). Given $\boldsymbol{F}$, the gain matrix $\boldsymbol{G}$ is adjusted such that $\boldsymbol{GHF}$ has unit main diagonal; the feedback matrix $\boldsymbol{B}$ is then the lower triangular part of this matrix.

As THP is a successive procedure, the encoding ordering of the users is of importance. This ordering can equivalently be described by a row permutation (permutation matrix $\boldsymbol{P}_\mathrm{B}$) of the channel matrix, leaving the natural encoding order. Among the $N_\mathrm{U}! = 6$ permutations, that obtained via the BLAST algorithm is (almost) optimum [12]; in case of full coordination we always use this ordering.

C. Hierarchical Network MIMO Processing

In a fully coordinated scheme, all BSs need to know all precoded symbols $\tilde{x}_u$, in order that each BS is able to calculate its transmit symbol x_b. To decrease the backhaul traffic, a *hierarchical distribution of knowledge* is of interest. Here we assume that the degree of knowledge increases from BS to BS, as illustrated in the right part of Fig. 3. BS 1 only knows its allocated data a_1. The precoded symbol $\tilde{x}_1$, generated from a_1 is communicated (via the backhaul) to BS 2. Using this information and its data symbol a_2, BS 2 calculates $\tilde{x}_2$ and forwards both precoded symbols to BS 3. This BS finally generates its symbol $\tilde{x}_3$. Applying this procedure, we obtain half of the coordination effort in the backhaul compared to full cooperation.

However, now BSs 1 and 2 only have partial knowledge on the precoded symbols. The lack of information for the calculation of the transmit symbols can equivalently be modeled as $\boldsymbol{F}$ to have *lower triangular structure*. Especially in case of THP, this approach fits well into the context of successive precoding, as the cancellation of interference is based on the same hierarchical structure (lower triangular feedback matrix $\boldsymbol{B}$). Since, for precoding, all encoded symbols of lower index have to be known anyway, a lower-triangular feedforward matrix does not induce additional coordination effort.

Basically, the optimization of the preequalization/precoding matrices can be done as in (9) and (11), respectively. Though, besides the power constraints (5) and (6), the additional constraint on $\boldsymbol{F}$ being lower triangular has to be taken into account. As above, this optimization can be carried out via a second-order cone program.

The demanded hierarchical approach not only imposes a triangular structure on $\boldsymbol{F}$, but also leads to the fact that the BSs are no longer equivalent. In a fully coordinated scheme, the feedback loop has to be carried out in an optimized way leading to $N_\mathrm{U}!$ possible orderings. W.r.t. performance it is irrelevant, which BS knows which data symbol as here the processing order of data symbols is identical to the processing order of the BSs. After the exchange of the $\tilde{x}_u$, all BSs have full knowledge for the generation of their transmit symbol.

In the hierarchical scheme, this equivalence of the BSs is no longer present, cf. Fig. 3. Now it is additionally of relevance which data symbol (for which user) the BS which starts processing has. This corresponds to the fact that not only *row permutations* (which data symbol is encoded first, second, last) of the channel matrix have to be taken into account, but also *column permutations* (which BS encodes first, second, last). In other words, now the assignment of data symbols to BSs matters, too. Unfortunately, there is no clear strategy which of the $N_\mathrm{U}!N_\mathrm{B}! = 36$ possible permutations leads to the best performance. To obtain reference results we first test all permutations and choose the best one. Subsequently, an heuristic sorting strategy is discussed which lowers the computational complexity significantly.

D. Numerical Results and Comparison

For the numerical evaluation of the abovementioned network MIMO schemes we simulate 25000 channel realizations; each channel is constant over 25000 time steps (burst length). The data symbols are drawn from a 16QAM constellation. The optimization of the min SINR has been performed using the *SeDuMi toolbox* for MATLAB [9]. In case of hierarchical LPE (H-LPE) and hierarchical THP (H-THP) among all possibilities the respectively best user assignment/processing order is chosen. For full cooperation only the ordering according to the BLAST algorithm is used; testing all $N_\mathrm{U}! = 6$ permutations does not give any noticeable gain. For comparison, the conventional MMSE preequalization/precoding schemes [5] are shown as well. Furthermore, the situation without any interference coordination, i.e., $\boldsymbol{F} = \boldsymbol{I}$ and $\boldsymbol{B} = \boldsymbol{I}$ in Fig. 2 (but optimized assignment of users' data symbols to BSs), is also given.

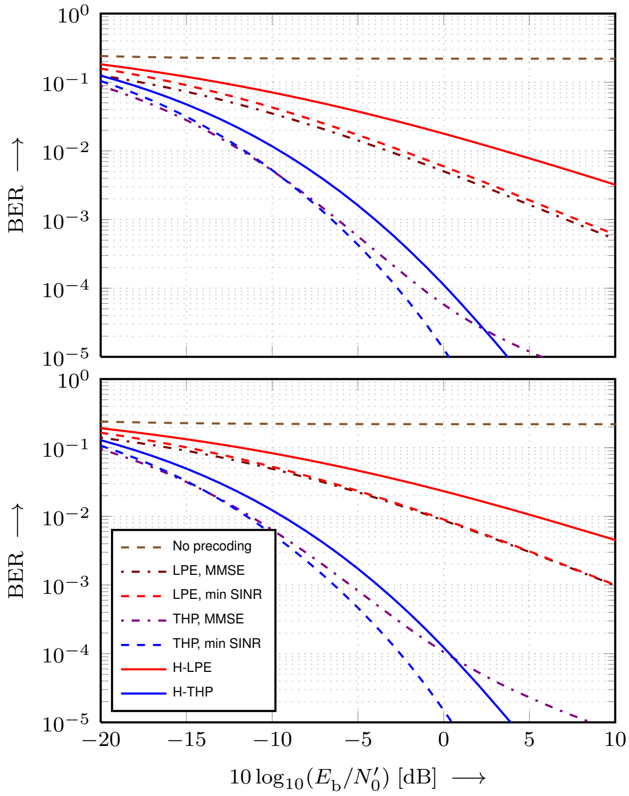

Fig. 4. Bit error rate of network MIMO schemes over the signal-to-noise ratio (in dB) for the system given in Sec. II. 16QAM. Sum power constraint $P_\mathrm{sum} = 3\,\sigma_a^2$; per-antenna power constraint $P_\mathrm{per} \to \infty$ (top), $P_\mathrm{per} = 1.5\,\sigma_a^2$ (bottom).

In Fig. 4 the bit error rate (BER) is plotted over the SNR. In the top only the SPC (5) is active; in the bottom plot additionally a PPC $P_\mathrm{per} = 1.5\,\sigma_a^2$ is imposed. As can be seen, a reasonable network MIMO transmission is not possible without any multiuser processing.

As already known from [5], a PPC only induces some loss in the conventional MMSE schemes. In contrast, using the min SINR optimization the BER curves hardly differ with or without PPC. Of course, the hierarchical schemes do not perform as well as the fully coordinated ones. However, especially in case of H-THP very good performance is possible with heavily reduced communication effort. The additional PPC does almost not degrade the performance.

IV. Preselection Strategy

In the above simulations the min SINR optimization has been carried out for all $N_\mathrm{B}!N_\mathrm{U}!$ permutations of the channel matrix; among them the best solution has been selected. This tremendous computational complexity is now reduced by preselecting a few permutations; only these are fed into the min SINR optimization. In the fully coordinated scheme this preselection was done via the BLAST sorting criterion. Subsequently we restrict ourselves to H-THP, because H-LPE results in too poor performance. Moreover, $N_{\mathrm{B}} = N_{\mathrm{U}} = 3$ is assumed; however, the following considerations can be extended to larger system dimensions.

A. Greedy Preselection Strategy

The main challenge for finding an appropriate preselection strategy is the fact, that the channel matrix $\boldsymbol{H}$ is known before the optimization, but not the end-to-end cascade $\boldsymbol{C}^{(\mathsf{H-THP})}$, as $\boldsymbol{F}$ is calculated during the optimization process. Since $\boldsymbol{C}^{(\mathsf{H-THP})}$ serves as basis for the SINR values, we have to define a substitute criterion to predict the BER performance. A suited measure for that are the SINR values based on $\boldsymbol{H}$, i.e., the SINR for each UE if the feedforward matrix is dropped ($\boldsymbol{F} = \boldsymbol{I}$) but the feedback (interference cancellation) is still active. Hence, in analogy to (10), we define

$$\mathrm{SINR}_u^{(\mathsf{H})} \stackrel{\mathrm{def}}{=} \frac{|h_{u,u}|^2}{\zeta + \sum\limits_{l=1,\ldots,N_{\mathrm{U}},\, l>u} |h_{u,l}|^2} \,. \tag{12}$$

Examining the outcome of the optimization process, it becomes apparent that permutations of the initial channel matrix which have the largest (squared) absolute value $|h_{3,3}|^2$ tend to have the highest minimum SINR value after the optimization process (note from (12) that $\mathrm{SINR}_3^{(\mathsf{H})} = |h_{3,3}|^2/\zeta$). This can be explained as follows: For UE 3, we can eliminate all interference from other users. Since $|h_{3,3}|^2$ is comparatively high, we thus tend to have a high S(I)NR on the subchannel from BS 3 to UE 3, leading to a rather low BER at the receiver. As a consequence, we can spend some transmit power at BS 3 (limited by the PPC P_{per}) in order to supply the other users with inferior links, as all of the precoded symbols $\tilde{\boldsymbol{x}}$ are known at BS 3. Thus, on the one hand, we have a small decrease in SNR for UE 3, which nevertheless remains comparably high, and on the other hand, the additional interference for UE 3 has no negative impact on the BER as we can eliminate it totally. Moreover, since we know all the transmission symbols at BS 3, we have the greatest flexibility in assigning the transmit power to UEs, which gives us the chance to balance up the related SINR values.

The strategy is repeated to assign UE 2 and UE 1. To that end, we use a greedy algorithm, i.e., UE 3 and hence $h_{3,3}$ are determined first (*nine possibilities*) and then kept fixed. This procedure yields the submatrix $\left[\begin{smallmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{smallmatrix}\right]$ where *four permutations* are still possible. For the BER at UE 2, $\mathrm{SINR}_2^{(\mathsf{H})} = |h_{2,2}|^2/(\zeta + |h_{2,3}|^2)$ is the relevant quantity. Following the criterion for UE 3, we choose the permutation which maximizes $\mathrm{SINR}_2^{(\mathsf{H})}$.

The $3! \cdot 3! = 36$ permutations are hence arranged into $9 \cdot 4 = 36$ combinations. The permuted channel matrix can then be written as

$$\boldsymbol{H}_{p_1,p_2} \stackrel{\mathrm{def}}{=} \boldsymbol{P}_{p_1,p_2}^{(\mathsf{row})} \boldsymbol{H} \boldsymbol{P}_{p_1,p_2}^{(\mathsf{col})} \,, \tag{13}$$

with $p_1 = 1, \ldots, 9$ and $p_2 = 1, \ldots, 4$. Here, $\boldsymbol{H}_{p_1,p_2}$ denotes the permutation of $\boldsymbol{H}$ with the p_1^{th} largest value $|h_{3,3}|^2$ (hence $\mathrm{SINR}_3^{(\mathsf{H})}$) and the p_2^{th} largest value $\mathrm{SINR}_2^{(\mathsf{H})}$ under assumption of the p_1^{th} largest value $|h_{3,3}|^2$. The related row and column permutations matrices are $\boldsymbol{P}_{p_1,p_2}^{(\mathsf{row})}$ and $\boldsymbol{P}_{p_1,p_2}^{(\mathsf{col})}$, respectively.

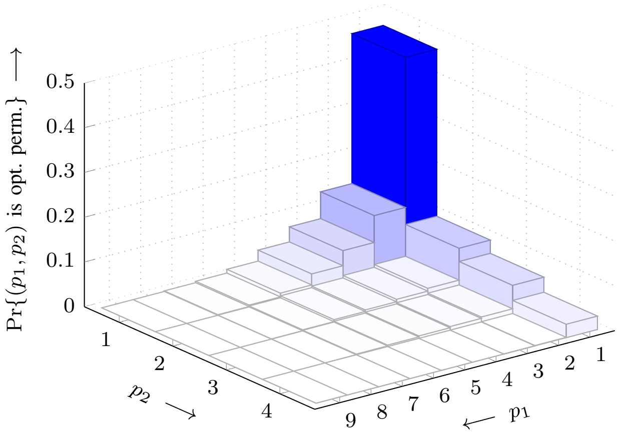

Fig. 5. Simulated probabilities that permutation $\boldsymbol{H}_{p_1,p_2}$ leads to the optimal solution. 25000 channel realizations, SPC $P_{\mathrm{sum}} = 3\,\sigma_a^2$, PPC $P_{\mathrm{per}} = 1.5\,\sigma_a^2$.

Fig. 5 depicts the empirically obtained probabilities that permutation $\boldsymbol{H}_{p_1,p_2}$ gives the highest min SINR after the optimization (11) with $\boldsymbol{F}$ being lower triangular.[6] As clearly visible, in 47% of the channels, the abovementioned strategy of "largest/largest" leads to the best performance. If we want to test only a single permutation, we should preselect this one. However, if we are allowed to test more than one, we should do this in sequence $\boldsymbol{H}_{1,1}$, $\boldsymbol{H}_{2,1}$, $\boldsymbol{H}_{1,2}$, $\boldsymbol{H}_{3,1}$, $\boldsymbol{H}_{1,3}$, $\boldsymbol{H}_{4,1}$, $\boldsymbol{H}_{1,4}$, …(sorted according to $\mathrm{Pr}\{(p_1, p_2)$ is opt. perm.$\}$). If we test $P = 3$, 6, 9, or 12 matrices in the optimized sequence, the optimum permutation (among all 36) is obtained for 69, 86, 93, or 97% of the channels.

B. Numerical Results and Comparison

For the numerical evaluation of the greedy preselection strategy we consider the same setting as above. 25000 different channel realizations and bursts of length 25000 symbols are simulated. 16QAM is used.

In Fig. 6, the BER over the SNR (in dB) is shown for different subset cardinalities P of tested candidate permutations. In the top only the SPC is active; in the bottom plot additionally a PPC $P_{\mathrm{per}} = 1.5\,\sigma_a^2$ is imposed. If we restrict the optimization to one permutation ($P = 1$) selected by the abovementioned greedy algorithm, a loss of about 4 dB for $\mathrm{BER} = 10^{-4}$ is caused in comparison

[6] Strictly speaking, the probability distribution depends on the SNR the system operates. The shown distribution was obtained for $E_{\mathrm{b}}/N_0' \,\widehat{=}\, 0$ dB. However, the actual values do almost not change over a wide range of SNR and hence the preordering of the permutations is universal.

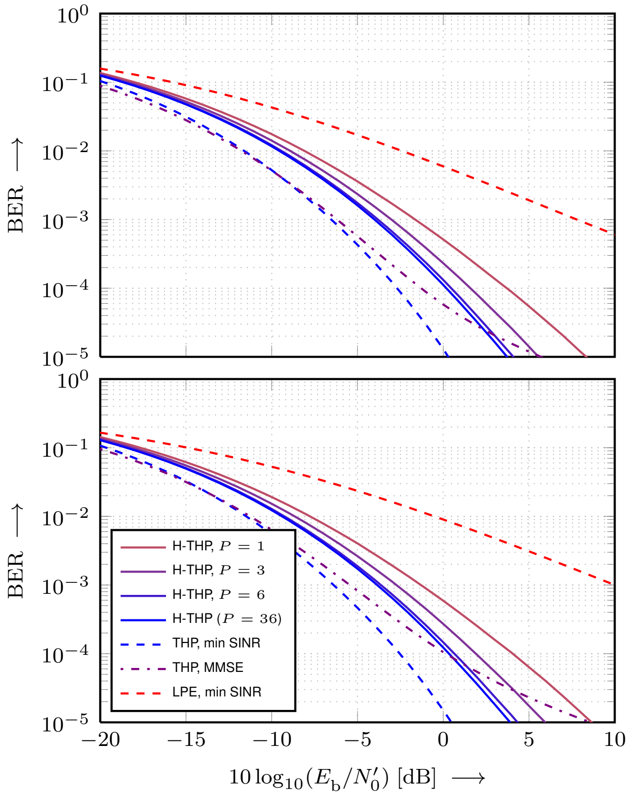

Fig. 6. Bit error rate of network MIMO schemes over the signal-to-noise ratio (in dB) for the system given in Sec. II. 16QAM. Hierarchical THP employing the greedy preselection strategy. Sum power constraint $P_{\text{sum}} = 3\,\sigma_a^2$; per-antenna power constraint $P_{\text{per}} \to \infty$ (top), $P_{\text{per}} = 1.5\,\sigma_a^2$ (bottom).

to the full search, i.e., $P = 36$. Increasing P to 3 reduces the loss to approximately 1.5 dB. For $P \geq 6$ almost no loss over the full search is visible. Hence, testing (running the optimization (11)) of $3 \ldots 6$ permutations is a good compromise between computation effort and performance.

Noteworthy, as already observed in [5], THP optimized according to the min SINR criterion leads to a better performance than classical MMSE THP, where the (unavoidable) flattening to diversity order one occurs earlier. This even holds for the hierarchical scheme. In each case, THP/H-THP clearly outperforms LPE/H-LPE. This fact becomes even more pronounced if the per-antenna power constraint is further lowered. An additional PPC does not degrade performance too much when employing the min SINR optimization.

In summary, we can state that H-THP in combination with the greedy preselection strategy increases the initial computation effort within a tolerable range. The main aim of lowering the coordination effort is achieved; only half of the communication compared to the conventional fully coordinated schemes is required. The loss in BER performance is acceptable; still a very good performance compared to linear schemes is enabled.

V. Summary and Conclusion

In the paper, an approach for decentralized network MIMO processing has been presented. In the initial phase, where all required preequalization or precoding matrices are computed, we still assume full cooperation via some central instance. Moreover, we have studied the computational effort during this phase, specifically, the number of optimizations which have to be carried out.

In order to reduce the signaling effort for communicating the data symbols of the users a hierarchical scheme has been proposed, optimized, and analyzed. The forced hierarchy immediately imposes a lower triangular structure on the transmitter-side feedforward matrix. It has been shown, that choosing a suited permutation of the channel matrix the loss compared to full cooperation is within an acceptable range. The permutation of rows of the channel matrix thereby (as in classical THP) corresponds to the encoding order of the data symbols, the permutation of columns corresponds to the assignment of the data symbols (i.e., users) to the base stations. A heuristic greedy preselection strategy for finding a suited permutation has been proposed.

Current work deals with a simple version how to classify the channel matrices where no coordination is required, hierarchical precoding is sufficient, or full coordination has to be used. By this, only the signaling overhead which is actually needed to guarantee a desired level of performance should be spent.

References

[1] 3GPP. Technical Specification Group Radio Access Network, RF System Scenarios, Technical Report TR 25.996 V3.3.0, 2002.

[2] 3GPP. Technical Specification Group Radio Access Network, Physical Layer Aspects for Evolved Universal Terrestrial Radio Access (UTRA), Technical Report TR 25.814 V7.1.0, 2006.

[3] R.F.H. Fischer, C. Windpassinger, A. Lampe, J.B. Huber. MIMO Precoding for Decentralized Receivers. In *Proc. Int. Symp. on Information Theory (ISIT)*, Lausanne, Switzerland, p. 496, 2002.

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

[5] R.F.H. Fischer, S. Heinrichs. The Network MIMO Downlink with Per-Antenna Power Constraint: Linear Preequalization vs. Tomlinson-Harashima Precoding. In *Proc. 17th Int. ITG Workshop on Smart Antennas (WSA)*, Stuttgart, Germany, Mar. 2013.

[6] M. Joham, J. Brehmer, W. Utschick. MMSE Approaches to Multiuser Spatio-Temporal Tomlinson-Harashima Precoding. In *Proc. 5th Int. ITG Conf. on Source and Channel Coding (SCC)*, pp. 387–394, Erlangen, Germany, Jan. 2004.

[7] M.K. Karakayali, G.J. Foschini, R.A. Valenzuela. Network Coordination for Spectrally Efficient Communication in Cellular Systems. *IEEE Wireless Communications*, pp. 56–61, Aug. 2006.

[8] C. Stierstorfer, C. Siegl, R.F.H. Fischer, T. Wild, C. Hoek. Network MIMO Downlink Transmission. In *Proc. 15th Int. OFDM Workshop (InOWo)*, Hamburg, Germany, Sept. 2010.

[9] J. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software 11–12; Special Issue on Interior Point Methods*, pp. 625–653, 1999.

[10] A. Tölli, M. Codreanu, M. Juntti. Linear Multiuser MIMO Transceiver Design With Quality of Service and Per-Antenna Power Constraints. *IEEE Transactions on Signal Processing*, pp. 3049–3055, July 2008.

[11] A. Wiesel, Y.C. Eldar, S. Shamai (Shitz). Linear Precoding via Conic Optimization for Fixed MIMO Receivers. *IEEE Transactions on Signal Processing*, pp. 161–176, Jan. 2006.

[12] C. Windpassinger. *Detection and Precoding for Multiple Input Multiple Output Channels*. PhD Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, June 2004.

[13] W. Yu, T. Lan. Transmitter Optimization for the Multi-Antenna Downlink With Per-Antenna Power Constraints. *IEEE Transactions on Signal Processing*, pp. 2646–2660, June 2007.