Linear Cooperative Multiuser MIMO Transceiver Design with Per BS Power Constraints

Antti Tölli, Marian Codreanu and Markku Juntti
Centre for Wireless Communications, University of Oulu
P.O. Box 4500, 90014 University of Oulu, Finland. Email: antti.tolli@ee.oulu.fi

Abstract—Joint cooperative processing of transmitted signal from several multiple-input multiple-output (MIMO) base station (BS) antenna heads is considered for users located within a soft handover (SHO) region. Downlink resource allocation problem with different BS power constraints is studied. The mathematical framework for the SHO based MIMO system is derived and the joint design of linear transmit and receive beamformers in a MIMO multiuser transmission according to weighted sum rate maximization criterion and subject to per BS power constraints is considered. The proposed algorithm is shown to provide very efficient solutions despite of the fact that there is no guarantee of achieving the global optimum due to the non-convexity of the problem. Moreover, practical and efficient resource allocation method based on generalized zero forcing transmission is provided.

I. INTRODUCTION

There has been increasing interest to consider network infrastructure based cooperative processing between base stations (BSs) with a cellular system [1]–[4] or fixed relay stations. Recently, [3]–[6] studied the downlink sum rate and spectral efficiency optimization for cooperative multiple-input-multiple-output (MIMO) systems with perfect data cooperation between BSs. Although BS cooperation increases the system complexity, it has significant capacity and coverage benefits making it worth more detailed consideration.

The capacity region of multiuser MIMO downlink (DL) with per antenna or per BS power constraints were recently discovered in [6], [7]. Furthermore, the minimum-power beamformer design for multiple-input single-output (MISO) DL under per antenna or per BS power constraints was investigated in [6], where the original DL problem was transformed into a dual uplink (UL) minimax optimization problem with an uncertain noise covariance. Convex optimization methods [8], such as second-order cone programming [9] and geometric programming [10], are very powerful tools which allow for efficient numerical solution for many signal processing [11]–[13] and optimal transmit and receive beamformer design problems [6], [14]–[17].

The purpose of this paper is to analyze the BS cooperation with linear processing in a more practical scenario. We assume a time division duplex (TDD) system with adaptive MIMO transmission, where the transmission parameters in reciprocal uplink (UL) and downlink (DL) can be adapted according to the channel conditions. Since it may be difficult to attain the channel knowledge between all users and BS antennas in practical cellular network, we consider the case where the joint cooperative processing of transmitted signal from several MIMO BS antenna heads is restricted to an area where the users have comparable signal strengths from adjacent BS antenna heads. Similarly to the soft handover (SHO) feature in (W)CDMA systems [18], SHO region is defined for users with similar received power levels from adjacent distributed BS antenna heads. We assume that the signal processing of the adjacent BSs is concentrated at one controller, and joint beamforming from all the antennas belonging to the "active set" can be performed to the user(s) within the SHO region.

In this paper, we consider transceiver optimization for multiuser MIMO downlink with linear processing. Note that the sum capacity achieving schemes require nonlinear precoding based on dirty paper coding (DPC) [7], [19]. We propose a general method for joint design of the linear transmit and the receive beamformers for weighted sum rate maximization problem subject to per BS power constraints. The method can handle multiple antennas at BSs and mobile users, and any number of data streams is allowed per scheduled user. Furthermore, it can be easily modified to accommodate supplementary constraints, e.g., per antenna power constraints or lower bounds for the SINR values of data streams, and the feasibility of the resulting optimization problems can be easily checked. Unlike the sum capacity achieving scheme, the optimization problems employed in the linear multiuser MIMO transceiver design are not convex in general. Therefore, the problem of finding the global optimum is intrinsically non-tractable. However, by utilizing the recent results on the precoder design via conic optimization [16] and the signomial programming [10], we propose an iterative solution where each step can be efficiently solved by using standard convex optimization tools [20]. Even though each subproblem is optimally solved, the global optimality cannot be guaranteed due to the non-convexity of the original problem. However, the simulation results demonstrate that the achieved locally optimal solutions are very efficient in several practically relevant scenarios. The proposed framework can be extended to other optimization criteria, such as maximization of the weighted minimum SINR per data stream subject to per BS power constraints [21]. In addition, we compare the proposed joint transceiver optimization algorithms with corresponding

This research was supported by Finnish Funding Agency for Technology and Innovation (TEKES), Nokia, the Finnish Defence Forces, Elektrobit, Tauno Tömming Foundation, Nokia Foundation and Infotech Oulu Graduate School.
II. System Model

The cellular MIMO system consists of \( N_B \) base stations, each BS has \( N_T \) transmit antennas and user \( k \) is equipped with \( N_{R_k} \) receive antennas. A user is served by \( M_k \) BSs which define the SHO active set \( S_k \) for the user \( k \). The signal vector \( y_k \in \mathbb{C}^{N_{R_k}} \) received by the user \( k \) can be expressed as

\[
y_k = \sum_{b \in S_k} a_{b,k} \mathbf{H}_{b,k} \left( x_{b,k} + \sum_{r \neq b} x_{r,k} \right) + \sum_{b \notin S_k} a_{b,k} \mathbf{H}_{b,k} x_b + \mathbf{n}_k
\]

where \( x_b \in \mathbb{C}^{N_T} \) is the transmitted signal from the \( b \)th BS to user \( k \), \( x_b \in \mathbb{C}^{N_T} \) denotes the transmitted signal vector from BS transmitter (TX) \( b \), \( \mathbf{n}_k \sim \mathcal{CN}(0, \mathbf{N}_k) \) represents the additive noise sample vector, and \( a_{b,k} \mathbf{H}_{b,k} \in \mathbb{C}^{N_{R_k} \times N_T} \) is the channel matrix from BS \( b \) to user \( k \) with large scale fading coefficient \( a_{b,k} \). The elements of \( \mathbf{H}_{b,k} \) are normalized to have unitary variance.

The signal \( \tilde{x}_k = [x_{S_0(1),k}^T, \ldots, x_{S_0(M_k),k}]^T \in \mathbb{C}^{M_k \times N_T} \) transmitted for user \( k \) is distributed over \( M_k \) base stations being in SHO active set \( S_k \). The global channel matrix \( \tilde{\mathbf{H}}_k \in \mathbb{C}^{N_{R_k} \times M_k \times N_T} \) for user \( k \) from all \( M_k \) BSs is

\[
\tilde{\mathbf{H}}_k = [a_{S_0(1),k} \mathbf{H}_{S_0(1),k}, \ldots, a_{S_0(M_k),k} \mathbf{H}_{S_0(M_k),k}].
\]

The transmitted vector for user \( k \) is generated as \( \tilde{x}_k = \mathbf{M}_k \tilde{d}_k \), where \( \mathbf{M}_k \in \mathbb{C}^{N_{R_k} \times M_k} \) is the pre-coding matrix, \( \tilde{d}_k = [d_{S_0(1),k}, \ldots, d_{S_0(M_k),k}]^T \) is the vector of normalized complex transmitted data symbols, and \( m_k \leq \min(N_T M_k, N_{R_k}) \) denotes the number of active data streams. \( \mathbf{M}_{k,c} \) can be further split into \( \mathbf{M}_{k,c} = \mathbf{V}_{k,c} \mathbf{P}_{k,c}^{1/2} \), where \( \mathbf{V}_{k,c} = [v_{k,c,1}, \ldots, v_{k,c,M_k,c}] \) contains the normalized TX beamformers and \( \mathbf{P}_{k,c} = \text{diag}(\rho_{k,c,1}, \ldots, \rho_{k,c,m_{k,c}}) \) controls the powers allocated to each of \( m_{k,c} \) streams.

The receiver (RX) is assumed to be equipped with a linear minimum mean square error (LMMSE) filter and the decision variables are generated as \( \tilde{d}_k = \mathbf{W}_k^H y_k \). The weight matrix \( \mathbf{W}_k \in \mathbb{C}^{N_{R_k} \times m_k} \) of the LMMSE filter is found by minimizing \( \mathbf{W}_k = \arg \min_{\mathbf{W}_k} \mathbb{E} \left[ \| \tilde{d}_k - \mathbf{W}_k^H y_k \|_2^2 \right] \) and is given as

\[
\mathbf{W}_k^H = \mathbf{M}_k^H \mathbf{H}_k^{-1} \left( \tilde{\mathbf{H}}_k \mathbf{M}_k \mathbf{M}_k^H \mathbf{H}_k^H + \mathbf{Z}_k + \mathbf{R}_k \right)^{-1}
\]

where \( \mathbf{Z}_k \) and \( \mathbf{R}_k \) are the intra- and inter-cell interference covariance matrices, respectively. The matrix \( \mathbf{Z}_k = \sum_{i \neq k} \mathbf{H}_i \mathbf{M}_i \mathbf{M}_i^H \mathbf{H}_i^H \) consists of transmissions to the users \( i \) that have an identical SHO active set composition with user \( k \), \( \mathbf{S}_i = \mathbf{S}_k \). In this paper, we assume \( \mathbf{R}_k = \mathbf{I} \). Ideally, the whitening of colored inter-cell interference can be also contained in \( \mathbf{H}_k \). The practical interference scenarios are considered in the system level study [23].

III. Weighted Sum Rate Maximization with Per BS Constraints

It is possible to serve several users having identical SHO active sets \( S_k \) in the same time-frequency transmission slot using some space division multiple access (SDMA) methods to separate them in space domain. SDMA can be used to improve the utilization of the physical resources (space, time, frequency) by exploiting the available spatial degrees of freedom in downlink multi-user MIMO channel, with an expense of somewhat increased complexity. In the following, we restrict our focus to a single set of users \( A \), where all users \( k \in A \) have identical active set composition, \( \mathbf{S}_k = \mathbf{S}_i, \forall k, i \in A \). We denote by \( M = |S_k| \) the SHO active set size, which is common to all \( k \in A \). Moreover, we focus on linear transmission schemes, where the transmitters send \( S \) independent streams, \( S \leq \min(M N_T, \sum_{k \in A} N_{R_k}) \). Note that it is possible to associate more than one stream to one user, i.e., the cardinality of the set of scheduled users, \( A = \{ k_i | s = 1, \ldots, S \} \), is less than or equal to \( S \).

We consider two general power constraints for cooperative BS processing: a sum power constraint for all \( M \) BSs in the SHO active set \( S_k \) and an individual power constraint for each BS. The total power transmitted by the BS \( n \) is

\[
\text{Tr}(E[\tilde{x}_n^H \tilde{x}_n]) = \text{Tr}(\sum_{k \in A} M_k[n] M_k[n]^H) = \sum_{k \in A} \sum_{i=1}^{M_k[n]} \| v_{k,i} \|_2^2 p_{k,i}
\]

where \( M_k[n] \in \mathbb{C}^{N_{R_k} \times N_T} \) is the pre-coder matrix for user \( k \) that corresponds to \( n \)th BS belonging to \( S_k \), i.e., \( M_k[n] = [M_k[(n-1)N_T+1:n N_T, :]] \), \( n = 1, \ldots, M \). Similarly, \( v_{k,i} \in \mathbb{C}^{N_T} \) is the transmit vector for the \( i \)th stream of user \( k \) from BS \( n \), i.e., \( v_{k,i} = [v_{k,i,(n-1)N_T+1:n N_T}, n = 1, \ldots, M \). Per antenna power constraints can be easily incorporated into (4) by assuming \( n = 1, \ldots, M N_T \) and \( M_k[n] \in \mathbb{C}^{1 \times m_k} \).

A. Joint Design of Linear Tx and Rx Beamformers

In this section, we consider the problem of joint design of the linear transmit and receive beamformers for maximizing the weighted sum of the rates of the individual data streams subject to per BS power constraints. Per data stream processing is considered, where for each data stream \( s, s = 1, \ldots, S \) the central BS’s scheduler unit associates an intended user \( k_s \) with the channel matrix \( \mathbf{H}_{k_s} \in \mathbb{C}^{M_k \times N_{R_k}} \).

Let \( v_s \in \mathbb{C}^{M_k[n] \times 1}, w_s \in \mathbb{C}^{N_{R_k} \times 1} \) and \( p_s \) be arbitrary normalized transmit and receive beamformers and the allocated power for the stream \( s \), respectively. The SINR of the data stream \( s \) can be expressed as

\[
\gamma_s = \frac{p_s \| w_s^H \mathbf{H}_{k_s} v_s \|^2}{1 + \sum_{i=1, i \neq s}^S p_i \| w_i^H \mathbf{H}_{k_s} v_s \|^2}
\]

Similarly to (4), the total power transmitted by \( n \)th BS is given by \( \sum_{s=1}^S p_s \| v_s \|_2^2 \). where \( v_s \in \mathbb{C}^{N_T} \) is the transmit vector for the \( s \)th data stream associated with \( n \)th BS, i.e., \( v_s = [v_{s,1}^T, \ldots, v_{s,M_k[n]}^T] \). Assuming Gaussian codebook [24]
for each data stream, the weighted sum rate is expressed as

$$ R_\beta = \sum_{s=1}^{S} \beta_s r_s = \sum_{s=1}^{S} \beta_s \log(1 + \gamma_s) = \log \prod_{s=1}^{S} (1 + \gamma_s)^{\beta_s} $$

(6)

where $r_s$ and $\gamma_s$ are the rate and the SINR of the $s$'th data stream, respectively. The weight vector, $\beta = [\beta_1, \ldots, \beta_S]^T$, $\beta_s \geq 0$, is used to prioritize differently the data streams and it can be chosen based on different criteria, e.g., states of the queues or buffers in case of cross layer optimization schemes. $\beta = 1$ corresponds to the usual sum rate optimization or best effort. Since $R_\beta$ increases with respect to each $\gamma_s$ and $\log(\cdot)$ is a increasing function, the weighted sum rate maximization problem with per BS power constraints can be formulated as

$$ \text{maximize} \quad \prod_{s=1}^{S} (1 + \gamma_s)^{\beta_s} \]

$$ \text{subject to} \quad \gamma_s \leq \frac{p_s |\mathbf{w}^H \mathbf{H}_k \mathbf{v}_s|^2}{1 + \sum_{i=1, i \neq s}^{S} p_i |\mathbf{w}^H \mathbf{H}_k \mathbf{v}_i|^2}, \quad s = 1, \ldots, S \]

(7)

where the variables are $\gamma_s \in \mathbb{C}^{MNr}$, $\mathbf{w}_s \in \mathbb{C}^{N_{Rs} \times 1}$, $p_s \in \mathbb{R}_+$, $\gamma_s \in \mathbb{R}_+$. It is easy to observe that at the optimal point of (7), the first constraint holds with equality, thus the optimal value of $\gamma_s$ represents the SINR of the $s$'th data stream.

The optimization problem (7) is not convex, and, hence, the problem of finding the global optimum is intrinsically non-tractable. However, (7) can be maximized with respect to different subsets of variables by considering the others fixed. For instance, the normalized maximum SINR receiver

$$ \mathbf{w}_s = \frac{\tilde{\mathbf{w}}_s}{\|\mathbf{w}_s\|_2}, \quad \tilde{\mathbf{w}}_s^H = \sqrt{p_s} \mathbf{v}_s (S \sum_{i=1}^{S} p_i \mathbf{H}_k \mathbf{v}_i \mathbf{v}_s^H + 1)^{-1} $$

(8)

is optimal for any fixed $\mathbf{v}_s$ and $p_s$. Furthermore, by fixing $\mathbf{v}_s$ and $\mathbf{w}_s$, (7) becomes a signomial problem [10], [13], [17]. The problem is not convex as such, but there are efficient methods for approximating the solution by using geometric programming [10], [13]. The procedure consists of searching for a close local maxima by solving a sequence of geometric programs which locally approximate the original problem. This procedure is known to converge fast [10]. Finally, for fixed $\mathbf{w}_s$ and $\gamma_s$ we can find a maximum reduction factor, common for all the per BS power constraints which preserves the SINR values $\gamma_s$ and, implicitly, the rate. This is given by the optimum $\alpha^*$ that solves the problem

$$ \text{minimize} \quad \alpha \]

$$ \text{s. t.} \quad \gamma_s \leq \frac{p_s |\mathbf{w}^H \mathbf{H}_k \mathbf{v}_s|^2}{1 + \sum_{i=1, i \neq s}^{S} p_i |\mathbf{w}^H \mathbf{H}_k \mathbf{v}_i|^2}, \quad s = 1, \ldots, S, \]

$$ \sum_{s=1}^{S} p_s \|\mathbf{v}_s[n]\|^2 \leq \alpha P_n, \quad n = 1, \ldots, M \]

$$ \|\mathbf{v}_s\|_2 = 1, \quad s = 1, \ldots, S \]

(9)

where the variables are $\alpha \in \mathbb{R}_+$, $p_s \in \mathbb{R}_+$, $\mathbf{v}_s \in \mathbb{C}^{MNr}, s = 1, \ldots, S$. The solutions $\mathbf{v}_s^*$ and $p_s^*$ do not directly increase the objective of (7). However, they increase the power margin for a fixed value of the objective, and hence, the saved power can be used to increase the objective. This is realized by updating $\mathbf{v}_s$ and $p_s$ in (7) to the new values $\mathbf{v}_s^*$ and $p_s^*/\alpha^*$ respectively and increasing all $\gamma_s$ until all SINR constraints become tight. Notice that this is an ascent step since $\alpha^* \leq 1$ for any $\mathbf{w}_s$ and $\gamma_s$ that are feasible for (7). The above observations suggest the following iterative optimization algorithm.

**Algorithm 1:** Weighted Sum Rate Maximization Under per BS Power Constraints

1) Initialize $\mathbf{v}_s^{(0)}$ and $p_s^{(0)}$ such that the per BS power constraints are satisfied. Compute the optimal $\mathbf{w}_s^{(0)}$ and $\gamma_s^{(0)}$ according to (8) and (5), respectively. Let $i = 1$ and go to Step 2.

2) Solve the problem (7) for the variables $p_s$ and $\gamma_s$, by fixing $\mathbf{w}_s = \mathbf{w}_s^{(i-1)}$ and $\mathbf{v}_s = \mathbf{v}_s^{(i-1)}$, $s = 1, \ldots, S$. Denote the solutions by $p_s^*$ and $\gamma_s^*$. \Update $p_s^{(i)} = p_s^*/\alpha^*$ and $\gamma_s^{(i)} = \gamma_s^*$. \Step 4) Update $\mathbf{w}_s^{(i)}$ and $\gamma_s^{(i)}$ according to (8) and (5), respectively. Test a stopping criterion. If it is not satisfied, let $i = i + 1$ and go to Step 2, otherwise STOP.

Even though Algorithm 1 increases monotonically the objective of (7), there is no guarantee that the global optimum is found due to non-convexity of the problem. However, the simulations in Sect. IV show that the algorithm converges to a solution, which can be a local optimum, but is still efficient.

Next, we present the algorithm used at Step 2 of Algorithm 1, which solves the problem (7) for fixed $\mathbf{w}_s$ and $\mathbf{v}_s$. The objective of (7), $f_\alpha(\gamma_1, \ldots, \gamma_S) = \prod_{s=1}^{S} (1 + \gamma_s)^{\beta_s}$ is approximated by a monomial function [10] $m(\gamma_1, \ldots, \gamma_S) = c \prod_{s=1}^{S} \gamma_s^{a_s}$, near the point $\hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_S)$, where the parameters $c$ and $a_s$ of the best monomial local approximation are given by [17, Section IV.B, Lemma 1]

$$ a_s^* = \frac{\hat{\gamma}_s}{1 + \hat{\gamma}_s}, \quad c = \frac{f_\alpha(\hat{\gamma}_1, \ldots, \hat{\gamma}_S)}{\prod_{s=1}^{S} \hat{\gamma}_s^{a_s}}. $$

(10)

By using the local approximation in the objective of problem (7), and ignoring the multiplicative constant $c$, we obtain the following iterative algorithm.

**Algorithm 2:** Geometric optimization step

1) Let the initial SINR guess, $\hat{\gamma} = (\gamma_1^{(i-1)}, \ldots, \gamma_S^{(i-1)})$

2) Solve the following geometric program,

$$ \text{maximize} \quad \prod_{s=1}^{S} \gamma_s^{\frac{1}{1+\delta} \frac{s}{s}} \]

$$ \text{s. t.} \quad (1 - \delta) \hat{\gamma}_s \leq \gamma_s \leq (1 + \delta) \hat{\gamma}_s, \quad s = 1, \ldots, S \]

$$ g_{s,x}^{(i)} P_s^{-1} \gamma_s + \sum_{k=1, k \neq s}^{S} g_{x,k} p_k g_{s,x} P_s^{-1} \gamma_s \leq 1, \quad s = 1, \ldots, S \]

$$ \sum_{s=1}^{S} p_s \|\mathbf{v}_s[n]\|^2 \leq P_n, \quad n = 1, \ldots, M \]

(11)

where the variables are $\gamma_s$, $p_s$ and $g_{s,x} = |\mathbf{w}^H \mathbf{H}_k \mathbf{v}_s|^2$ are fixed values. Denote the solution by $p_s^*$ and $\gamma_s^*$. If $\max_s |\gamma_s^* - \hat{\gamma}| > \epsilon$ set $\hat{\gamma} = (\gamma_1^{(i)}, \ldots, \gamma_S^{(i)})$ and go to Step 2, otherwise STOP.
The geometric program (11) approximates the original sigmoidal problem (7) around the point \( \hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_S) \). The first set of inequality constraints of (11) are called trust region constraints [10] and they limit the domain of variables \( \gamma_s \) in a region where the monomial approximation is accurate enough. The constant \( \delta < 1 \) controls the desired approximation accuracy and a typical value is \( \delta = 0.1 \) [10].

Now, we focus on Step 3 of Algorithm 1. First, observe that the change of variable \( \mathbf{m}_s = \sqrt{\mathbf{P}}_s \mathbf{v}_s \) defines a bijective mapping between the sets \( \{ (p_s, \mathbf{v}_s) \mid p_s \in \mathbb{R}_+, \| \mathbf{v}_s \|_2 = 1, \mathbf{v}_s \in \mathcal{O}^{|M|N_T} \} \) and \( \mathbf{m}_s \in \mathcal{O}^{|M|N_T} \). Thus, we can solve (9) for \( \mathbf{m}_s \), and then recover the optimal \( p_s \) and \( \mathbf{v}_s \). Furthermore, by replacing the positive variable \( \alpha \) by \( \rho^2 \) we obtain the following equivalent reformulation of (9)

\[
\begin{align*}
\text{minimize} & \quad \rho^2 \\
\text{s. t.} & \quad \gamma_s \leq \frac{\| w^H_k \mathbf{H}_{ks} \mathbf{m}_s \|^2}{1 + \sum_{s' \neq s} |w^H_{s'k} \mathbf{e}_{s'}^2|}, \quad s = 1, \ldots, S \\
& \quad \sum_{s=1}^S |w^H_k \mathbf{e}_s|^2 \leq \rho^2 P_n, \quad \rho \geq 0, n = 1, \ldots, M
\end{align*}
\]

(12)

Notice that the objective \( \rho^2 \) can be replaced by \( \rho \) since for \( \rho \geq 0 \), minimizing \( \rho^2 \) is equivalent to minimizing \( \rho \). Moreover, the constraints of the problem (12) can be expressed as generalized inequality with respect to the second-order cone [8], [9], [16]. Thus, the problem (12) can be further reformulated as a second-order cone program (SOCP) and it can be efficiently solved numerically by using standard optimization software [20], [25]. By modifying the approach presented in [16, Section IV.B] to accommodate per BS power constraints, we obtain the following equivalent SOCP formulation

\[
\begin{align*}
\text{minimize} & \quad \rho \\
\text{s. t.} & \quad \sqrt{1 + \frac{1}{\rho^2}} w^H_k \mathbf{H}_{ks} \mathbf{m}_s \geq \gamma \mathbf{0}, \quad s = 1, \ldots, S \\
& \quad M \mathbf{H}^H_k \mathbf{w}_s \geq \frac{1}{\rho \sqrt{P_n}} \quad \| \mathbf{vec}(\mathbf{M}[n]) \|_1 \geq \gamma \mathbf{0}, \quad n = 1, \ldots, M
\end{align*}
\]

(13)

where \( \mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_S], \mathbf{M}[n] = [\mathbf{m}_1^n, \ldots, \mathbf{m}_S^n], \) and \( \geq \gamma \) denotes the generalized inequality with respect to the second-order cone [8], [9], i.e., for any \( x \in \mathbb{R} \) and \( y \in \mathbb{C}^m, [x, y^T]^T \geq \gamma \mathbf{0} \) is equivalent to \( x \geq \| y \|_2^2 \). Let us denote the solution of the problem (13) by \( \rho^*, \mathbf{m}_s^*, s = 1, \ldots, S \). The solution of the problem (9) is given by \( \alpha^* = \rho^* \mathbf{v}_s^*/\| \mathbf{m}_s^* \|^2, \mathbf{v}_s^* = \mathbf{m}_s^*/\| \mathbf{m}_s^* \|^2 \), \( p_k^* = \| \mathbf{m}_s^* \|^2, s = 1, \ldots, S \). Note that (13) provides a minimum-power beamformer design under per BS power constraints for MIMO DL with fixed receivers. This is equivalent to the method proposed in [6], where the original DL problem was transformed into a dual UL minimax optimization problem with an uncertain noise covariance.

The optimization problem in (7) can be easily modified to accommodate supplementary constraints, e.g., per antenna or sum power constraints, or minimum SINR values for some of the data streams (\( \gamma_s \geq \gamma_s^{min} \)). The modified problem with minimum SINR constraints can be directly solved using the proposed algorithm, if it is feasible under the initial beamformer configuration obtained at the Step 1 of Algorithm 1, i.e., \( \gamma_s^{(0)} \geq \gamma_s^{min} \). The feasibility of the modified optimization problem with any minimum SINR constraints can be easily checked by iterating between problem (9) with fixed \( \gamma_s = \gamma_s^{min} \) and (8). If the resulting \( \alpha \leq 1 \), then the problem is feasible and the resulting beamformer configuration can be used as a feasible starting point for Algorithm 1.

**B. Multiuser Zero Forcing Solution**

Iterative block diagonalization (BD) of multiple user channels combined with coordinated TX-RX processing and scheduling between users is a simple but efficient ZF method [22], [26]. As a result of the iterative BD processing, e.g., [23, Algorithm 1], the multiple-access interference (MAI) is eliminated between users, i.e., \( M_k^H \mathbf{H}_k \mathbf{M}_k = 0 \) for \( i \neq k \). At the same time, the channel for user \( k \) is diagonalized, \( M_k^H \mathbf{H}_k \mathbf{M}_k = \Lambda_k \mathbf{P}_k \), where the diagonal matrix \( \mathbf{P}_k = \text{diag}(\lambda_k, \ldots, \lambda_k, \mathbf{m}_k) \) controls the powers allocated for each of the \( m_k \) eigenmodes (streams) and the diagonal matrix \( \Lambda_k = \text{diag}(\lambda_k, \ldots, \lambda_k, \mathbf{m}_k) \) includes the first \( m_k \) eigenvalues of of user \( k \). Note that in such a case the receiver in (3) can be reduced to a simple matched filter, \( \mathbf{W}^H_k k = M_k^H \mathbf{H}^H_k \).

With ZF processing the problem of maximizing the weighted sum of rates of the data streams under the per BS power constraint in (7) is reduced to

\[
\begin{align*}
\text{maximize} & \quad \sum_{k \in A} \sum_{i=1}^{m_k} \beta_{ki} \log_2 (1 + \lambda_{ki} p_{k,i}) \\
\text{s. t.} & \quad \sum_{k \in A} \sum_{i=1}^{m_k} \| v_{k,i} \|^2 \leq P_n, n = 1, \ldots, M \\
& \quad p_{k,i} \geq 0, \quad k \in A, \quad i = 1, \ldots, m_k
\end{align*}
\]

(14)

where the variables are \( p_{k,i}, k \in A, i = 1, \ldots, m_k \), and \( P_n \) is the power constraint on the BS \( n \). The weights, \( \beta_{ki}, \forall k, i \), are used to prioritize differently the data streams of different users. It is easy to observe that the objective function of (14) is concave and all the inequality constraints are affine. Thus, the problem (14) is a convex analytic centering problem [8, Chapt. 8.5.3], and it can be efficiently solved numerically by using standard optimization software packages, e.g., CVX [20], SeDuMi [25]. Under the sum power constraint, \( P_{\text{sum}} \) and with equal user priorities, the sum rate is maximized by the well known water-filling power allocation [24], \( p_{k,i} = (\mu - 1/\lambda_{ki})^+, \) where the "water level", \( \mu, \) is chosen such that the sum power constraint holds with equality, i.e., \( \sum_{k \in A} \sum_{i=1}^{m_k} p_{k,i} = P_{\text{sum}} \).

**IV. NUMERICAL RESULTS**

In the simulations, the elements of the channel matrices were modelled as i.i.d. Gaussian random variables and the number of both TX and RX antennas was fixed at 2, \( \{ N_T, N_R \} = \{ 2, 2 \} \). For simplicity, all the BSs are assumed to have equal maximum power limit \( P_t \), i.e., \( P_t = P_t \) \( \forall n \). The impact of the following two power constraints are studied.

- **Sum power constraint**: All \( M \) BSs in \( S_k \) have perfect power and data cooperation. This provides an unrealistic upper bound, where the pooled maximum available power
is always $P_{\text{sum}} = MP_T$, while the antenna array gain from having $MN_T$ TX antennas depends on the RX power differences between BSs.

*Per BS power constraint:* Available power can be increased up to $M$ times depending on the RX power difference between BSs. Also, the gain from having $MN_T$ TX antennas depends on the RX power difference.

The mutual information for 2-branch SHO with different power constraints is studied. We consider a case where two SHO users (labelled as $u = 2$) are served simultaneously by two BSs in a flat fading scenario. Furthermore, we assume that they have identical large scale fading coefficients for two BSs in a flat fading scenario. Furthermore, we assume that they have identical large scale fading coefficients for simplicity, i.e., $a_{S_1,1}, a_{S_1,2}$ and $a_{S_2,1}, a_{S_2,2}$.

First, the impact of beamformer initialization on the ergodic sum rate is studied. Due to non-convexity of the original optimization problem (7), different initialization beamformer configurations from $\{v_1^{(0)}, \ldots, v_S^{(0)}\}$ used in Algorithm 1 may end up in different locally optimal solutions. Fig. 1 depicts the ergodic 2-user sum rate as a function of the number of random beamformer initializations. The best out of $n_{\text{init}}$ random TX beamformer initializations was selected for each channel realization. The received power imbalance $\alpha = a_{S_2}^2/k / a_{S_2}^2$ between BSs is fixed at 0 dB and single link SNRs are 0 dB and 10 dB. $S_k(1)$ is the BS with the strongest reception at the terminal and the single link SNR is defined as $\text{SNR} = P_T a_{S_k(1)} / N_0$. The ergodic sum rate is depicted for the weighted sum rate maximization algorithm (Algorithm 1 labelled as 'lin. max rate') with sum power and per BS power constraints and with weight vector $\beta = 1$. Moreover, the sum capacity with per BS power constraint, computed as in [6], is plotted as the absolute upper bound of the scenario.

A large number of randomly chosen initializations increases the probability to find a solution close to the global optimum for each channel realization. It is seen from Fig. 1 that the impact of the initial beamformer configuration is relatively small and the gain achieved from drawing randomly several initial points saturates rapidly to a fixed value. We studied also a different approach where the initial transmit beamformers were obtained by applying the orthogonal-triangular QR decomposition to the set of dominant right singular vectors of user channels. The singular vectors were ordered in a descending order according to their singular values. The unitary vectors from the resulting Q matrix were used as initial beamformers. As shown in Fig. 1, the QR based initialization method produces a very efficient starting point, e.g., more than two random initializations are required to produce higher ergodic sum rate.

Figs. 2 and 3 illustrate the ergodic mutual information for different power imbalance values, and for 0 dB and 10 dB single link SNRs, respectively. The ergodic 2-user sum rate is depicted for the proposed algorithm with QR based initialization and the corresponding ZF method (Section III-B) with different power constraints. The single user case ($u = 1$) with and without SHO is also included for comparison. The sum capacity bounds shown in Figs. 1, 2 and 3 are not generally achievable with a linear transmission strategy. However, the proposed weighted sum rate maximization algorithm achieve more than 90 percent of the sum capacity with per BS power constraints and a single sum power constraints. Note that the resulting ergodic sum rates for the proposed algorithm can be still slightly improved if a few random initializations for each channel realization are allowed as seen in Fig. 1.

The zero forcing solution, labelled as 'ZF', is depicted for two scenarios: fully loaded case $\{m_1, m_2\} = \{2, 2\}$, labelled as 'FL', and partially loaded case, labelled as 'PL', where the best allocation of $m_k$ among possible combinations $\{m_1, m_2\} = \{2, 2\}, \{2, 1\}, \{1, 2\}, \{1, 1\}$ is selected for each channel realization. It is seen from the figures that the zero forcing with full spatial load ('FL') performs rather bad, especially at low
Fig. 3. Ergodic sum rate of \( \{N_f, N_R_b, M\} = \{2, 2, 2\} \) system at 10 dB single link SNR.

SNR range. Even with large imbalance (−20 dB) both users are intended to be served with two streams. This obviously reduces the achievable rate even below the single link capacity. The zero forcing with partial loading performs reasonably well even at low SNR and approaches the weighted sum rate maximization algorithm (Algorithm 1) at high SNR.

A simple heuristic method which find a suboptimal, but still efficient, power allocation for the problem (14) was introduced in [23]. Similar to the sum power constraint case, a water-filling power allocation is imposed, \( p_{k,i} = (\bar{\mu} - 1/\lambda_{k,i})^+ \), but the water level \( \bar{\mu} \) is increased until one of the BSs reaches its power constraint. The heuristic power loading solution applied to the BD concept (labelled as ‘heur.’) performs almost as good as the optimal method (14).

V. CONCLUSION

The joint cooperative processing of transmitted signal from several MIMO BS antenna heads was considered for users located within a SHO region. The mathematical framework for the SHO based MIMO system was derived and the joint design of linear transmit and receive beamformers in a MIMO multiuser transmission according to weighted sum rate maximization criterion and subject to per BS power constraints was considered. The proposed algorithm was shown to provide very efficient solutions despite of the fact that there is no guarantee of achieving the global optimum due to the non-convexity of the problems. Moreover, practical and efficient resource allocation methods based on generalized zero forcing transmission were provided.

REFERENCES


[11] M. Bengtsson and B. Ottersten, “Optimal and suboptimal transmit filling power allocation is imposed, \( p_{k,i} = (\bar{\mu} - 1/\lambda_{k,i})^+ \), but the water level \( \bar{\mu} \) is increased until one of the BSs reaches its power constraint. The heuristic power loading solution applied to the BD concept (labelled as ‘heur.’) performs almost as good as the optimal method (14).