Joint Transmitter-Receiver Optimization for Downlink Multiuser MIMO Communications

Ali M. Khachan, Adam J. Tenenbaum and Raviraj S. Adve
Dept. of Electrical and Computer Engineering, University of Toronto
10 King’s College Road, Toronto, Ontario, M5S 3G4, Canada
Email: {khachana,adam,rsadve}@comm.utoronto.ca

Abstract

This paper develops the joint optimization of linear precoding and decoding for the downlink in multiuser multiple-input multiple-output (MIMO) systems. The system under investigation consists of a single transmitter (base station) with multiple antennas that communicates with several users, each with one or more receive antennas, possibly receiving multiple data streams. The goal is to jointly optimize the power allocation and transmit-receive filters for all users. Optimization is performed for two criteria: (1) minimizing total transmitted power while satisfying a set of signal-to-interference-plus-noise ratio (SINR) constraints and (2) minimizing the sum mean squared error (SMSE) given a total power budget. The solutions use an uplink/downlink duality, in SINR and MSE sense. These dualities are proven for the general MIMO case. In this regard, this paper generalizes previous multiple-input single-output (MISO) solutions to the MIMO case. Simulation results prove the efficacy of the proposed solutions.

Keywords—Array signal processing, duality, MIMO systems, power control, space division multiplexing.

I. INTRODUCTION

It is now well accepted that exploiting the spatial dimension using an antenna array at the transmitter and/or the receiver can improve reliability and increase the data rate of wireless transmission. More recently, researchers have investigated using such a multiple-input multiple-output (MIMO) system to service multiple users on the same time/frequency channel. In this paper, we study the downlink of a multiuser MIMO system – a single base station communicating with multiple users on a shared channel with each user receiving possibly multiple data streams.

In such a multiuser system, effective communication requires suppression of the resulting multiuser interference (MUI) and inter-stream interference. Given a criterion of optimality, and assuming knowledge of the channel state information (CSI) at the transmitter, the goal of this work is to jointly optimize the power allocation and transmit-receive filters for all users. The optimization is for two criteria: (1) minimizing the total transmit power required to meet Quality of Service (QoS) constraints for each data stream of each user, measured here in terms of signal-to-interference-plus-noise ratio (SINR) constraints for each data stream, and (2) minimizing the sum mean squared error (SMSE) between the transmitted and received signals with a total power constraint at the transmitter. We focus on linear transmit and
receive filters, wherein all signal processing blocks are matrix multiplications, as opposed to non-linear approaches generally based on Tomlinson-Harashima precoding (THP) [1].

The available literature on the above optimization problems can be classified into two different approaches. The first approach uses block diagonalization, or block channel inversion, to force the MUI to zero. This results in effectively single-user communication with each user having interference among its own data streams only. Bourdoux and Khaled [2] suppress inter-stream interference using a combination of transmit and receive filters and minimize MSE over the resulting block diagonal channel. On the other hand, Spencer et al. [3] obtain the optimal transmit matrix to maximize the system throughput resulting in waterfilling for power allocation across data streams. The drawback of these null-space approaches is the requirement that the number of transmit antennas must be equal or greater than the total number of receive antennas. In [4], the authors avoid this problem by extending the single-user SMSE work in [5] to the multiuser domain, and propose an iterative joint optimization algorithm with a per-user power constraint. They also propose a computationally intensive numerical method based on Sequential Quadratic Programming (SQP) that solves the SMSE problem with a sum power constraint.

The second class of solutions employs iterative algorithms that repeatedly switch between power control, optimizing the transmit filter and optimizing the receive filter. The available literature largely focuses on minimizing transmit power with SINR constraints. This problem has been comprehensively investigated for multiple-input single-output (MISO) systems, i.e., for users with only a single receive antenna. Yang and Xu [6] derive the optimal power control policy for fixed transmit beamformers. Rashid-Farrokhi et al. [7] propose an iterative algorithm that converges to transmit weight vectors and downlink power allocations that satisfy a target SINR at each mobile. The algorithm also minimizes the total transmitted power while maintaining the SINR targets. The authors show that finding the optimal power allocation, given transmit and receive filters, is a convex optimization problem. Visotsky and Madhow [8] address feasibility and convergence issues raised in [7]. In [9], Schubert and Boche propose, based on a duality between the uplink and the downlink, the SINR maximization and power minimization solutions of the multiuser downlink given individual SINR constraints. The algorithm solves the problem iteratively in the uplink before transforming the solution to the downlink. Bengtsson and Ottersten [10] argue that although the original power minimization problem is non-linear and non-convex, an efficient solution is possible through semi-definite optimization techniques. The drawback of this approach is its high computational
complexity. Chang et al. [11] solve the downlink problem for MIMO systems with each user receiving a single data stream. However, the solution diverges when some target SINR scenarios are infeasible. In [12], Doostnejad et al. combine this scheme with dirty paper coding, and a non-linear suboptimal solution is presented.

In [13], Serbetli and Yener study the SMSE problem exclusively in the uplink with a per-user power constraint, jointly optimizing the transmit and receive filters. The scheme allows for each user to transmit multiple data streams. Based on the uplink solution in [13], Shi and Schubert [14] solve the downlink SMSE problem in MISO systems, generalizing the scheme in [15] to MIMO systems with a sum power constraint. Their key contributions are extending the uplink-downlink duality to multiple receive antennas and incorporating joint transmit-receive optimization.

This paper solves the most general problem in this area, eliminating most of the constraints placed on this problem in previous works. We jointly optimize the power allocation and transmit and receive filters for both problems of minimizing transmit power with QoS constraints and for minimizing SMSE with a total transmit power constraint for the MIMO case. The system model under consideration allows for an arbitrary number of transmit antennas and an arbitrary number of receive antennas at each user. Each user, in turn, may receive multiple data streams. In this regard, this paper generalizes previously known algorithms for the MISO case incorporating joint transmit/receive processing. The only limitations are resolvability constraints due to the linear precoding/decoding used. Unlike in [11], our proposed power minimization algorithm does not diverge in some scenarios. This allows for an examination of feasibility of a set of target SINRs.

A brief review on the notation used: lower case italics, e.g., \(x\), represent scalars while vectors are represented by lower case boldface type, e.g., \(\mathbf{x}\). Upper case italics, e.g., \(N\), generally represent constants while upper case boldface represents matrices, e.g., \(\mathbf{U}\). The superscripts \(T\) and \(H\) represent the transpose and Hermitian operators respectively. \(\mathbb{E}[\cdot]\) represents the statistical expectation operator while \(\mathbf{I}_N\) represents the \(N \times N\) identity matrix. \(\|x\|_1\) denotes the 1-norm, i.e., the sum of the entries, of a vector \(x\), while \(\text{diag}(x)\) represents the diagonal matrix formed using the entries in vector \(x\).

The paper is organized as follows. Section II presents the system model and states the assumptions made. Section III develops the uplink-downlink duality in the SINR sense for the general MIMO case and presents an iterative solution to the joint optimization problem of power minimization given a set of
SINR constraints. Section IV presents the duality theorem and two algorithms to minimize the sum-MSE with a sum power constraint. In Sections III and IV, numerical examples illustrate the performance of the algorithms developed. Finally, Section V presents our conclusions.

II. SYSTEM MODEL

Consider a single base station equipped with $M$ antennas communicating with $K$ decentralized users. User $k$ is equipped with $N_k$ antennas and $N = \sum_{k=1}^{K} N_k$. In this general setup, user $k$ receives $L_k$ data streams from the base station and $L = \sum_{k=1}^{K} L_k$. Thus we have $M$ transmit antennas transmitting a total of $L$ symbols in every time slot to $K$ users, who have a total of $N$ receive antennas. The symbols of each user are collected in the data vector $\mathbf{x}_k = [x_{k1}, x_{k2}, \ldots, x_{kL_k}]^T$ and the overall data vector is $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_K^T]^T$. We assume the data symbols are drawn from a constellation whose points have unit energy, and that they are independent, i.e., $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_L$.

The $l$th data stream to the $k$th user is processed by transmit filter $\mathbf{u}_{kl}$, i.e., the linear precoder for the $k$th user is the $M \times L_k$ matrix $\mathbf{U}_k = [\mathbf{u}_{k1}, \mathbf{u}_{k2}, \ldots, \mathbf{u}_{kL_k}]$. These individual precoders are collected into the overall $M \times L$ precoder matrix $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_K]$. Similarly, the $l$th data stream of the $k$th user is assigned power $p_{kl}$, i.e., the downlink transmit power vector for user $k$ is $\mathbf{p}_k = [p_{k1}, p_{k2}, \ldots, p_{kL_k}]^T$, with $\mathbf{p} = [\mathbf{p}_1^T, \ldots, \mathbf{p}_K^T]^T$. For convenience, define $\mathbf{P}_k = \text{diag}\{\mathbf{p}_k\}$ and $\mathbf{P} = \text{diag}\{\mathbf{p}\}$. The channel between the transmitter and user $k$ is assumed flat and is represented by the $N_k \times M$ matrix $\mathbf{H}_k^H$. The resulting $N \times M$ channel matrix is $\mathbf{H}$, with $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_K]$.

Based on this model, user $k$ receives a length-$N_k$ vector

$$\mathbf{y}_k = \mathbf{H}_k^H \mathbf{U} \sqrt{\mathbf{P}} \mathbf{x} + \mathbf{n}_k,$$

where $\mathbf{n}_k$ represents the additive white Gaussian noise (AWGN) at the user’s receive antennas with power $\sigma^2$; i.e., $\mathbb{E}[\mathbf{n}_k\mathbf{n}_k^H] = \sigma^2 \mathbf{I}_{N_k}$. To estimate its $l$th data stream, user $k$ processes the received signal using decoder vector $\mathbf{v}_{kl}$ or equivalently, user $k$ decodes its $L_k$ symbols through the individual $N_k \times L_k$ decoder matrix $\mathbf{V}_k^H$, where $\mathbf{V}_k = [\mathbf{v}_{k1}, \mathbf{v}_{k2}, \ldots, \mathbf{v}_{kL_k}]$ resulting in

$$\hat{\mathbf{x}}_{k}^{DL} = \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{U} \sqrt{\mathbf{P}} \mathbf{x} + \mathbf{V}_k^H \mathbf{n}_k,$$

where the superscript $^{DL}$ represents the fact that (2) deals with the downlink. The global receive filter
\( V^H \) is a block diagonal decoder matrix of dimension \( L \times N \):

\[
V^H = \begin{bmatrix}
V_1^H & 0 & \cdots & 0 \\
0 & V_2^H & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & V_K^H \\
\end{bmatrix}.
\]  

(3)

We now construct a virtual uplink system that will prove very useful when exploiting the duality between the downlink and the uplink. In the uplink, user \( k \) transmits \( L_k \) data streams to a single base station receiver through the \( M \times N_k \) channel \( H_k \). The user uses \( V_k \) as its transmit filter and \( U_k^H \) as its receive filter. Let the power allocated to the \( l \)-th data stream of the \( k \)-th user in the uplink be \( q_{kl} \), i.e.,

uplink transmit power vector for user \( k \) is \( q_k = [q_{k1}, q_{k2}, \ldots, q_{kL_k}]^T \), with \( q = [q_1^T, \ldots, q_K^T]^T \). Also, define \( Q_k = \text{diag}\{q_k\} \) and \( Q = \text{diag}\{q\} \). In the uplink, the received vector at the base station, \( y \), and the estimated symbol vector for user \( k \), \( \hat{x}_k \), are

\[
y = \sum_{i=1}^{K} H_i V_i \sqrt{Q_i} x_i + n, \\
\hat{x}_{UL}^k = \sum_{i=1}^{K} U_k^H H_i V_i \sqrt{Q_i} x_i + U_k^H n.
\]  

(4)  

(5)

The noise term, \( n \), is again AWGN with covariance matrix \( \sigma^2 I_M \). Figure 1 illustrates the downlink and the corresponding virtual uplink. The only limitation on the system is that, to ensure resolvability, we require \( L \leq M \) and \( L_k \leq N_k, \forall k \).

III. POWER MINIMIZATION WITH SINR CONSTRAINTS

Consider a set of QoS constraints specified as target SINR \( \gamma_{kl} \) for the \( l \)-th substream of user \( k \). Let the total transmitted power be \( P = \sum_{k=1}^{K} \sum_{l=1}^{L_k} p_{kl} = \|p\|_1 \). The power minimization problem can then formulated as

\[
\begin{align*}
\min_{p, U, V} P &= \sum_{k=1}^{K} \sum_{l=1}^{L_k} p_{kl} = \|p\|_1, \\
\text{subject to} : SINR_{kl} \geq \gamma_{kl}; \quad P &\leq P_{\text{max}},
\end{align*}
\]  

(6)

and where the columns of \( U \) have unit norm.

Using (2), the SINR of the \( l \)-th data stream of the \( k \)-th user in the downlink is

\[
SINR_{DL}^{kl} = p_{kl} \frac{V_k^H S_k^{DL} V_k}{V_k^H T_k^{DL} V_k},
\]  

(7)
where

\[ S_{k\ell}^{DL} = H_k^H u_{k\ell} u_{k\ell}^H H_k, \]  

\[ T_{k\ell}^{DL} = \sum_{j=1,j\neq l}^{L_k} p_{kj} H_k^H u_{kj} u_{kj}^H H_k + \sum_{i=1,i\neq k}^{K} \sum_{m=1}^{L_i} p_{im} H_k^H u_{im} u_{im}^H H_k + \sigma^2 I. \]  

In the virtual uplink, using (5), the equivalent SINR expression is

\[ SINR_{k\ell}^{UL} = q_{k\ell} \frac{u_{k\ell}^H S_{k\ell}^{UL} u_{k\ell}}{u_{k\ell}^H T_{k\ell}^{UL} u_{k\ell}}, \]  

where

\[ S_{k\ell}^{UL} = H_k^H v_{k\ell} v_{k\ell}^H H_k, \]  

\[ T_{k\ell}^{UL} = \sum_{j=1,j\neq l}^{L_k} q_{kj} H_k^H v_{kj} v_{kj}^H H_k + \sum_{i=1,i\neq k}^{K} \sum_{m=1}^{L_i} q_{im} H_k^H v_{im} v_{im}^H H_k + \sigma^2 I. \]  

**A. Power Allocation for Fixed Transmit-Receive Filters**

This section develops the optimal power allocation vectors, \( p \) and \( q \), given fixed transmit/receive filters \( U \) and \( V \). The authors of [9] define the problem

\[ C_{DL} = \max_{1 \leq k \leq K, 1 \leq \ell \leq L_k} \min \frac{SINR_{k\ell}^{DL}}{\gamma_{k\ell}} \text{ subject to: } \|p\|_1 \leq P_{\text{max}}. \]  

If \( C_{DL} \geq 1 \), then the set of SINR targets is feasible. Otherwise, we have infeasible targets and other methods like dropping some users or lowering the SINR targets are necessary. The authors show that for a fixed transmit filter \( U \), and due to \( C_{DL} \) being monotonically increasing in \( P_{\text{max}} \), the optimum power allocation vector, \( p \), achieves active SINR constraints, i.e.,

\[ C_{DL} = \frac{SINR_{k\ell}^{DL}}{\gamma_{k\ell}} \quad \forall k \text{ and } \ell. \]  

It can be easily proven that (14) still holds for the MIMO case, when \( U \) and \( V \) are fixed. Therefore, by collecting the \( L \) equations from (14) we can construct an eigensystem in the downlink and the virtual uplink:

\[ \Upsilon p_{\text{ext}} = \frac{1}{C_{DL}} p_{\text{ext}}, \]  

\[ \Lambda q_{\text{ext}} = \frac{1}{C_{UL}} q_{\text{ext}}, \]
where

\[
\mathbf{Y} = \begin{bmatrix}
\frac{1}{P_{\text{max}}} \mathbf{D} \mathbf{\Psi} \mathbf{D}^T & \frac{1}{P_{\text{max}}} \mathbf{D} \mathbf{\sigma} \\
\end{bmatrix},
\]

(17)

\[
\Lambda = \begin{bmatrix}
\frac{1}{P_{\text{max}}} \mathbf{D} \mathbf{\Psi}^T \mathbf{D}^T & \frac{1}{P_{\text{max}}} \mathbf{D} \mathbf{\sigma}
\end{bmatrix},
\]

(18)

\[
\mathbf{D} = \text{diag}\left(\frac{\gamma_{11}}{|\mathbf{v}_1^H \mathbf{H}_1^H \mathbf{u}_1|}, \ldots, \frac{\gamma_{KL_K}}{|\mathbf{v}_{KL_K}^H \mathbf{H}_{KL_K}^H \mathbf{u}_{KL_K}|}\right).
\]

(19)

Also, \( \mathbf{\Psi} \) is a coupling matrix, defined as

\[
[\mathbf{\Psi}]_{ik} = \begin{cases}
|\mathbf{v}_i^H \mathbf{H}_i^H \mathbf{u}_k|^2 & k \neq i \\
|\mathbf{u}_k^H \mathbf{H}_i \mathbf{v}_i|^2 & k = i.
\end{cases}
\]

(20)

for the case of one data stream per user; in the case of multiple data streams per user, each stream can be treated as a “virtual user” as described in the Appendix. \( \mathbf{p}_{\text{ext}} \) and \( \mathbf{q}_{\text{ext}} \) are extended power allocation vectors, defined as \( \mathbf{p}_{\text{ext}} = [\mathbf{p}^T \mathbf{1}]^T \) and \( \mathbf{q}_{\text{ext}} = [\mathbf{q}^T \mathbf{1}]^T \). Finally, \( \mathbf{\sigma} = \sigma^2 \mathbf{1} \) (where \( \mathbf{1} \) is a vector of ones of appropriate dimension).

Using (15) and (16), finding the optimal power allocation (downlink and uplink) that maximizes \( C_{\text{DL}} \) (or \( C_{\text{UL}} \)) is an eigenvalue problem. The optimal \( \mathbf{p} \) (or \( \mathbf{q} \)) is the dominant eigenvector of \( \mathbf{Y} \) (or \( \Lambda \)) ensuring that \( C_{\text{DL}} \) (or \( C_{\text{UL}} \)) is maximized. If \( C_{\text{DL}} \geq 1 \), then the set of SINR targets is feasible. To minimize the total transmission power (\( P \) in (6)), while achieving the required SINR targets, we therefore change the power allocation policy such that \( SINR_{kl} = \gamma_{kl} \forall \) L streams. The resulting power vectors are

\[
\mathbf{p} = \sigma^2 (\mathbf{D}^{-1} - \mathbf{\Psi})^{-1} \mathbf{1},
\]

(21)

\[
\mathbf{q} = \sigma^2 (\mathbf{D}^{-1} - \mathbf{\Psi}^T)^{-1} \mathbf{1}.
\]

(22)

Several studies have identified an interesting duality between the multiuser uplink and downlink. This duality states that given a fixed transmit filter and a sum power constraint, the same balanced SINR level is achieved in both directions, i.e., \( C_{\text{DL}} = C_{\text{UL}} \). We extend the result to multiuser MIMO systems, with users possibly receiving multiple data streams, in the following theorem:

**Theorem 1:** With linear processing matrices \( \mathbf{U} \) at the base station and \( \mathbf{V} \) over all users, and \( \|\mathbf{p}\|_1 = \|\mathbf{q}\|_1 \), \( C_{\text{DL}} = C_{\text{UL}} \).

**Proof:** See Appendix A.

Consequently, for a fixed \( \mathbf{U}, \mathbf{V} \) and \( P \), there exist downlink/uplink power allocations \( \mathbf{p} \) and \( \mathbf{q} \) such that \( \|\mathbf{p}\|_1 = \|\mathbf{q}\|_1 = P \), and \( SINR_{kl}^{\text{DL}} = SINR_{kl}^{\text{UL}} \) for all \( L \) streams. Note that the theorem makes no
assumption on the optimality of the matrices $U$ and $V$. This remarkable result is very useful since it allows one to solve the optimization problem in either the downlink or the virtual uplink.

B. Transmit-Receive Filters for Fixed Power Allocation

In the previous section, (21) and (22) represent the optimal power allocation policy for a fixed transmit matrix $U$ and collection of receive matrices in $V = \text{diag}(V_1, V_2, \ldots, V_K)$. We now investigate the reverse problem: assuming a fixed power allocation $p$ and $q$, we wish to find the optimal transmit-receive filters $U$ and $V$. By examining (7) and (8), we can observe that in the downlink, and for a fixed $U$, $SINR_{kl}^{DL}$ becomes a function of $v_{kl}$ and does not depend on any other receive beamformers. Hence, the receive beamformers can be optimized independently such that

$$v_{kl}^{opt} = \arg \max_{v_{kl}} \ SINR_{kl}^{DL} = \hat{e}_{\max}(S_{kl}^{DL}, T_{kl}^{DL}),$$

(23)

where $\hat{e}_{\max}(A, B)$ represents the dominant generalized eigenvector of the matrix pair $(A, B)$. Equivalently, in the virtual uplink, for a fixed $V$,

$$u_{kl}^{opt} = \arg \max_{u_{kl}} \ SINR_{kl}^{UL} = \hat{e}_{\max}(S_{kl}^{UL}, T_{kl}^{UL}).$$

(24)

Due to Theorem 1, since the SINR is the same in both the downlink and the virtual uplink, this “receive” matrix in the uplink is also optimal for the downlink. Note that the solutions in (23) and (24) are equivalent to the MMSE filters since $S_{kl}^{UL}$ and $S_{kl}^{DL}$ are rank-1 matrices.

The original problem being addressed attempts to find the optimal transmit matrix, $U$, set of receive matrices, $V^H$, and power allocation policy, $p$, to minimize the transmitted power while meeting QoS (in terms of SINR) constraints for each data stream of each user. The proposed algorithm, summarized in Table I, optimizes each variable by fixing the other variables, and iterating between the uplink and the downlink. Unlike the algorithm in [11], this algorithm does not diverge even if the set of target SINRs is infeasible, such as in the case when $M < K$. Convergence is ensured by using an initial power control policy in (15) that accounts for infeasibility, i.e., when $C^{DL} = C^{UL} < 1$.

C. Numerical Examples

This section presents results of simulations illustrating the performance of the proposed algorithm. The MIMO channels are modelled as being independent and identically distributed using a flat and Rayleigh fading channel model with unit variance. The results are shown for unit variance noise ($\sigma^2 = 1$). The
transmitter is assumed to have full knowledge of the channel matrices and is responsible for all processing, reducing the required complexity at the mobile stations. Define the receive antenna and data stream vectors as $\mathbf{N} = [N_1, N_2, ..., N_K]$ and $\mathbf{L} = [L_1, L_2, ..., L_K]$, respectively.

Figure 2 examines the power minimization algorithm for different scenarios. For simplicity, assume that the target SINRs are equal, i.e., $\gamma = \gamma_{11} = \gamma_{12} = ... = \gamma_{KL}$. It shows the minimal transmitted power $P^*$ required to satisfy the SINR constraint $\gamma$ for $K$ users. The example scenario uses $M = 8$ transmit antennas. Every user is equipped with $N_k = 2$ antennas and is receiving $L_k = 2$ data streams. The maximum allowed transmission power is set at $P_{\text{max}}/\sigma^2 = 13$ dB. The figure illustrates that, as expected, increasing either the SINR targets or the number of users requires an increase in the minimum total power needed to successfully meet the target SINRs. The ceiling on the plot represents the scenarios with $K$ users and SINR targets $\gamma$ where the required $P^*/\sigma^2$ exceeds the power budget of 13 dB. Note that the scenario becomes infeasible when $K \geq M$ and $\gamma > 0$ dB, but the proposed algorithm does not diverge as is the case in [11]. Note that in this example, $K = L$. The fact that the targets become infeasible when $L > M$ indicates a rank-issue with linear precoding.

Figure 3 plots $C^*$ versus the total transmission power $P_{\text{max}}/\sigma^2$, where $C^* = C^{UL} = C^{DL}$ is the optimally balanced ratio of achieved SINR to target SINRs. The scenario includes three users with $N_1 = 2, N_2 = 2, N_3 = 3$ antennas receiving $L_1 = 2, L_2 = 1$ and $L_3 = 3$ data streams respectively. The figure shows $C^*$ for different numbers of transmit antennas $M$. As stated previously, the target SINRs ($\gamma = 1$ dB) can only be achieved for a given $P_{\text{max}}$ if $C^* \geq 1$. From the figure, note that the targets are always feasible for $M > K$; that is, there exists a total power $P_{\text{max}}$ that satisfies $\gamma$. As expected, as the number of users in the network increases, it becomes more difficult to maintain the desired SINR targets for all the users. When the scenario is infeasible, the network has either to drop some users and try to optimize the link again, or the target SINRs have to be relaxed.

Figures 2 and 3 illustrate the validity of the algorithm presented in Table I. While the results are not unexpected, we emphasize that the key contribution here is the development of the algorithm in Table I, where we generalize the problem to the multiuser MIMO case. Significant differences from previously available work are the inclusion of the optimal receive filter and a constraint on the total power used to ensure a convergent solution.

In Figure 4, we compare the performance of linear and non-linear precoding. The simulation uses a
target SINR of $\gamma=10$dB, with $N_k = 6$ receive antennas per user and where each user receives a single data stream. In the simulated scenario, there are $M = 8$ transmit antennas. The figure plots the minimum total transmission power required to satisfy the 10dB target SINR for different numbers of users. An interesting observation is that the proposed linear precoding algorithm performs almost identically to the non-linear algorithm for $K \leq M$. When the number of users exceeds the number of transmit antennas at the base station, the algorithm can not find an acceptable power level to satisfy the targets, and non-linear precoding is required to solve the problem. Consequently, in mobile networks, where $K \leq M$ is expected, the use of low-complexity linear processing techniques is justified.

IV. SMSE MINIMIZATION WITH SUM POWER CONSTRAINT

We now address the popular problem of minimizing the sum mean squared error between the transmitted and received signals over all data streams. This problem is solved in a computationally intensive, brute force manner using SQP in [4]. This section presents a significantly more efficient algorithm based on uplink-downlink duality. Let $E_k^{DL}$ be the $L_k \times L_k$ error covariance matrix of user $k$ in the downlink, where

$$E_k^{DL} = \mathbb{E}[(\hat{x}_k - x_k)(\hat{x}_k - x_k)^H].$$

(25)

The diagonal entries of $E_k^{DL}$ are the MSEs of the $L_k$ substreams of user $k$ and thus $SMSE_k^{DL} = tr[E_k^{DL}]$, where $tr[\cdot]$ is the trace operator. The SMSE minimization problem can be formulated as

$$\min_{p,U,V} \sum_{k=1}^{K} tr[E_k^{DL}]$$

subject to : $\|p\|_1 \leq P_{\text{max}},$

(26)

and where the columns of $U$ have unit norm. The authors of [14] solve the same problem but for MISO systems with single data stream transmission. They prove that under a total power constraint, for a given $U$, the uplink and the downlink have the same normalized MSE achievable region. We extend these results to MIMO systems with possibly multiple data streams per user using the following duality theorem:

Theorem 2: For a given $U$, $V$ and $P_{\text{max}}$, there exist power allocation vectors $p$ and $q$ such that $MSE_{ul}^{kl} = MSE_{dl}^{kl}$, where $MSE_{kl}$ is the MSE of the $l^{th}$ substream of user $k$.

Proof: See Appendix B.

Again, note that the theorem makes no assumptions as to the optimality of $U$ and $V$. In the following we present two different approaches to solving the SMSE minimization problem.
A. Processing in the Uplink

We can solve problem (26) in the virtual uplink and transform the resulting transmit-receive filters and power allocation to the downlink. Using (5) and expanding (25) in the virtual uplink, the resulting error covariance matrix of user $k$ is

$$E_{UL}^k = U_k^H V Q V^H H U_k + \sigma^2 U_k^H U_k + I_{L_k} - U_k^H H_k V_k \sqrt{Q_k} - \sqrt{Q_k} V_k H_k^H U_k.$$  \hspace{1cm} (27)

The optimum MMSE uplink receiver is $U_k^{H}$, where $U_k$ is defined as

$$U_k^{MMSE} = J^{-1} H_k V_k \sqrt{Q_k},$$ \hspace{1cm} (28)

and

$$J = H Q V^H V H^H + \sigma^2 I_M.$$ \hspace{1cm} (29)

$$\Rightarrow E_{UL,MMSE}^k = I_{L_k} - \sqrt{Q_k} V_k^H H_k^H J^{-1} H_k V_k \sqrt{Q_k}.$$ \hspace{1cm} (30)

The sum MSE over all data streams is therefore

$$SMSE = \sum_{k=1}^{K} tr[E_{UL,MMSE}^k] = L - M + \sigma^2 tr[J^{-1}].$$ \hspace{1cm} (31)

The SMSE expression in (31) is a function of two variables: uplink power allocation $Q$ and uplink global transmit filter $V$. We first assume that $V$ is fixed. Analogous to the MISO solution in [14], minimizing SMSE is equivalent to minimizing the trace of $J^{-1}$. Crucially, the resulting optimization problem is convex in $Q$ (see Appendix C), i.e.,

$$Q_{opt} = \arg \min_Q tr[J^{-1}], \ \text{subject to} \ tr[Q] = P_{max},$$ \hspace{1cm} (32)

is a convex optimization problem, allowing for powerful and computationally efficient techniques [16].

Equation (32) finds the optimal uplink power allocation given a fixed transmit matrix (in the uplink) $V$, assuming the optimal receive matrix in (28). The next step is the reverse problem of optimizing $V$ for a fixed power allocation $Q$. The authors of [13] propose a scheme for calculating $v_{kl}$, for the $l^{th}$ data stream of the $k^{th}$ user, given per-user power constraints. We generalize the scheme to allow for a sum power constraint. Using the matrix inversion lemma, $J^{-1}$ can be written as

$$J^{-1} = J^{-1}_{kl} - q_{kl} \frac{J^{-1}_{kl} H_k v_{kl} v_{kl}^H H_k^H J^{-1}_{kl}}{1 + q_{kl} v_{kl}^H H_k^H J^{-1}_{kl} H_k v_{kl}},$$ \hspace{1cm} (33)
where
\[ J_{kl} = J - q_{kl} H_k v_{kl} v_{kl}^H H_k^H. \] (34)

Substituting (33) in (31) simplifies the SMSE expression to
\[ SMSE = W_{kl} - \sigma^2 \frac{v_{kl}^H (H_k^H J_{kl}^{-2} H_k) v_{kl}}{v_{kl}^H (I/q_{kl} + H_k^H J_{kl}^{-1} H_k) v_{kl}}. \] (35)

As indicated in [13], all the terms in \( W_{kl} \) are independent of \( v_{kl} \). Thus the optimal \( v_{kl} \), which minimizes SMSE for a given power allocation when the beamforming vectors of all other streams are fixed, is the solution to a generalized eigenvector problem, i.e.,
\[ v_{kl}^{MSE} = \hat{e}_{max} \left( H_k^H J_{kl}^{-2} H_k, \frac{1}{q_{kl}} + H_k^H J_{kl}^{-1} H_k \right). \] (36)

Table II presents the proposed algorithm which first solves the SMSE minimization problem iteratively in the virtual uplink. By applying duality, we see that the MSE of individual data streams is the same in the uplink or downlink. Thus, these filters are used for our original problem in the downlink. The initialization starts with the uplink power allocation \( Q \) by distributing \( P_{max} \) evenly among all data streams. In the table, \( V_k = \text{SVD}(H_k) \) indicates that \( V_k \) is initialized using the \( L_k \) dominant right singular vectors of \( H_k \). Such a matrix would be optimal in the single-user case. The next step involves optimizing \( V \) and \( Q \) iteratively. Each iteration begins with optimizing \( V \) for the previous power allocation \( Q \), where the optimum \( v_{kl} \) of every stream is found using (36). After the uplink beamforming vectors of all streams are found, power is allocated by solving the convex optimization problem (32). The iterative step is executed repeatedly until SMSE converges to its final value. The algorithm then exploits duality by setting \( \gamma_{kl} = SINR_{UL} \), which leads to \( C^{DL} = C^{UL} = 1 \). The last step finds \( p_{DL} \), the downlink power allocation vector, using (21). Finally, the optimal downlink transmit filter \( U \) is found using the MMSE receiver (28). This algorithm is labelled MIMO1-SMSE.

**B. Iterative SMSE Minimization**

Section IV-A presented an algorithm which iterated between finding the power allocation and the transmit filter entirely in the uplink before transferring to the downlink. This section presents an alternative approach which is closer to the scheme in Section III in philosophy. This approach iteratively optimizes the transmit-receive filters and power allocation by switching between the downlink and the virtual uplink.
We have seen that in the uplink, for fixed transmit filter $V$, the optimal MMSE receiver for user $k$ that minimizes SMSE is (28). Similarly for the downlink, using (2) and expanding (25) in the virtual uplink, the resulting error covariance matrix of user $k$ is

$$E_{DL}^k = V_k^H H_k^H U P U H_k V_k + \sigma^2 V_k^H V_k + I_{L_k} - V_k^H H_k^H U_k \sqrt{P_k} - \sqrt{P_k} U_k H_k V_k. \quad (37)$$

The optimum MMSE downlink receiver for fixed transmit filter $U$ is $V_k^H$, where $V_k$ is defined as

$$V_k^{MMSE} = G_k^{-1} H_k^H U_k \sqrt{P_k}, \quad (38)$$

and

$$G_k = H_k^H U P U H_k + \sigma^2 I_{N_k}. \quad (39)$$

Table III presents a SMSE algorithm which iteratively calculates the receive and transmit matrices by switching between the uplink and downlink from one step to the next. Each iteration of the algorithm begins with finding the optimal power allocation $Q$ in the virtual uplink for the current filter $V$. The virtual uplink receive filter $U^H$ is then calculated for the current $V$ using the MMSE beamformer in (28). Using duality and (21), the downlink power allocation $p$ is found. In the downlink, for a fixed $U$, the MMSE receive filter $V^H$ is calculated using (38). The algorithm switches back to the virtual uplink and repeats the above steps until the SMSE converges to its final value. This algorithm is labelled MIMO2-SMSE.

It is important to note that the three algorithms described here, one to minimize transmit power and two to minimize SMSE, are optimal at each iteration step, but do not guarantee convergence to the global optimal solution. Indeed, this is an issue common to all iterative approaches for the MIMO optimization problem published so far. However, as we will see in the next section, the SMSE algorithms essentially perform as well as the brute force SQP algorithm of [4].

C. Numerical Examples

This section presents the results of simulations illustrating the performance of the algorithms presented in Sections IV-A and IV-B. The simulation scenario makes the same assumptions as in Section III-C. Most importantly, the base station is assumed to know the channels of all $K$ users while the individual users know their own channels.

Figure 5 compares the speed of convergence for the two proposed algorithms that solve the SMSE minimization problem. The scenario investigated is for $K = 3$ users and $M = 6$ transmit antennas. Users
are equipped with $N_1 = 2$, $N_2 = 2$, and $N_3 = 3$ antennas, and are receiving $L_1 = 2$, $L_2 = 1$, and $L_3 = 2$ data streams. The plot shows the number of iterations required by the two algorithms to converge to the minimum SMSE, averaged over many realizations of the channel, versus SNR, where $\text{SNR} = \frac{P_{\text{max}}}{\sigma^2}$. It is clear that MIMO1-SMSE algorithm converges much faster than MIMO2-SMSE algorithm. Furthermore, while the number of iterations for MIMO1-SMSE is almost constant as SNR increases, there is a linear increase in the number of iterations required by MIMO2-SMSE. The better performance of MIMO1-SMSE can be attributed to the fact that this algorithm solves the problem completely in the virtual uplink, while MIMO2-SMSE has to continuously cycle between the downlink and the virtual uplink.

Finally, Figure 6 shows average BER versus SNR for MIMO1-SMSE minimization algorithm. It also shows the same curve for the SQP algorithm. The simulated system has $K = 2$ users, receiving $L_1 = 2$ and $L_2 = 1$ data streams respectively. Both users are equipped with $N_1 = N_2 = 2$ antennas each. The figure shows that the proposed algorithm achieves an almost identical performance in BER to the SQP algorithm which suffers from being computationally intensive.

These examples illustrate the efficacy of the algorithms presented here. The main contribution in this paper is the fact that these algorithms iteratively solve for the jointly optimal transmit-receive filters and power allocation for a multiuser MIMO system.

V. CONCLUSIONS

Multiuser linear precoding has received growing attention recently. In this paper, we have addressed the most general problem of precoding in a multiuser MIMO scenario. We proposed a solution to the problem of joint beamforming and power allocation in the downlink where users have multiple antennas. The investigated scheme also accommodates scenarios where each user receives multiple data streams. The first objective was to find an optimal combination of transmit-receive filters and transmission powers to minimize the total transmitted power while satisfying individual SINR targets for each data stream. We generalized a previously known uplink-downlink duality from the MISO to the MIMO case. The special structure of the SINRs allowed for an iterative approach to optimize beamforming vectors by cycling between the downlink and the virtual uplink.

The second objective was to find this combination that minimizes the sum mean squared error under a total power constraint. We proposed two different algorithms to solve this problem. The first algorithm minimizes the SMSE in the virtual uplink before transforming the solution into the downlink. The second
algorithm iteratively cycles between the downlink and virtual uplink. Simulations showed that the first SMSE minimization algorithm converges significantly faster than the second.

Overall, the three algorithms presented here are effective algorithms with relatively low complexity solving two important optimization problems in multiuser communications.

APPENDIX

Without loss of generality, we simplify the proofs by transforming the system from one having $K$ users with $L_k$ data streams each into a system having $L$ virtual users with single data stream each. The transmit and receive filters become $U = [u_1, u_2, \ldots, u_L]$ and $V = \text{diag}\{v_1, v_2, \ldots, v_L\}$.  

A. Proof of Uplink-Downlink Duality for SINR

Using the coupling matrix $\Psi$ defined in (20), it has been shown in [15] that for the set of SINR constraints to be achievable in the downlink, $\lambda_{\text{max}}(D\Psi) < 1$, where $\lambda_{\text{max}}(.)$ is the maximum eigenvalue operator. Similarly, $\lambda_{\text{max}}(D\Psi^T) < 1$ guarantees achievability in the uplink. The authors in [17] prove that $\lambda_{\text{max}}(D\Psi) = \lambda_{\text{max}}(D\Psi^T)$. Consequently, the SINR achievable region is the same for the uplink and the downlink, with power allocations $q > 0$ and $p > 0$ respectively. Using (22) we write

\[
\|q\|_1 = 1^T q = \sigma^2 1^T (D^{-1} - \Psi^T)^{-1} 1
\]

\[
= \sigma^2 1^T (D^{-1} - \Psi)^{-1} 1 = 1^T p = \|p\|_1.
\]

With the total power in the uplink and the downlink being identical, the SINR targets are achieved and $C_{DL} = C_{UL}$.

B. Proof of Uplink-Downlink Duality for MSE

Along the same lines of the proof presented in [14] for MISO systems, we generalize the MSE duality to systems with multiple antennas at the receiver. In what follows, we adopt the notations $\tilde{v}_i$ and $\tilde{u}_i$ for the MMSE receive beamforming vectors in the downlink and the uplink respectively, where

\[
\tilde{v}_i = (H_i^H U P U^H H_i + \sigma^2 I)^{-1} H_i^H u_i \sqrt{p_i},
\]

(40)

\[
\tilde{u}_i = (H V Q V^H H^H + \sigma^2 I)^{-1} H_i v_i \sqrt{q_i}.
\]

(41)
Also let $v_i = \tilde{v}_i / \|\tilde{v}_i\|$ and $u_i = \tilde{u}_i / \|\tilde{u}_i\|$. We write the MSE expressions for user $i$ in the downlink and the uplink respectively:

$$
\varepsilon_i^{DL} = \tilde{v}_i^H H_i^H U P U^H H_i \tilde{v}_i + \sigma^2 \|\tilde{v}_i\|^2 + 1 - \sqrt{p_i} \tilde{v}_i^H H_i^H u_i - \sqrt{p_i} u_i^H H_i \tilde{v}_i,
$$

(42)

$$
\varepsilon_i^{UL} = \tilde{u}_i^H H V Q V^H H \tilde{u}_i + \sigma^2 \|\tilde{u}_i\|^2 + 1 - \sqrt{q_i} \tilde{v}_i^H H_i^H \tilde{u}_i - \sqrt{q_i} \tilde{u}_i^H H_i \tilde{v}_i.
$$

(43)

Define

$$
B = H V Q V^H H + \sigma^2 I.
$$

(44)

Setting $SINR_i^{DL} = SINR_i^{UL}$ and using some mathematical manipulation, we arrive at the following equation:

$$
\|\tilde{v}_i\|^2 \left( \frac{\tilde{u}_i^H B \tilde{u}_i}{q_i} \right) = \|\tilde{u}_i\|^2 \left( \frac{\tilde{v}_i^H H_i^H U P U^H H_i \tilde{v}_i + \sigma^2 \|\tilde{v}_i\|^2}{p_i} \right)
$$

(45)

Replacing in (42), $\varepsilon_i^{DL}$ becomes

$$
\varepsilon_i^{DL} = \left( \frac{\|\tilde{v}_i\|^2}{\|\tilde{u}_i\|^2} \right) \frac{p_i}{q_i} \tilde{u}_i^H B \tilde{u}_i + 1 - \sqrt{p_i} \tilde{v}_i^H H_i^H u_i - \sqrt{p_i} u_i^H H_i \tilde{v}_i.
$$

(46)

Assuming that $\|\tilde{v}_i\| = \beta / \sqrt{p_i}$ and $\|\tilde{u}_i\| = \beta / \sqrt{q_i}$, where $\beta$ is a constant, we have $\varepsilon_i^{DL} = \varepsilon_i^{UL}$.

C. Proof of Convexity of Power Allocation

The power allocation step in the SMSE minimization algorithms of Section IV involves solving for $[q_1, q_2, ..., q_K]$ in the following problem:

$$
Q^{opt} = \arg \min_Q \text{tr}[J^{-1}]
$$

subject to: \quad $\text{tr}[Q] \leq P_{max}, q_k \geq 0$ for all $k$,

where

$$
J = H V Q V^H H + \sigma^2 I_M.
$$

(48)

Rewrite the positive definite matrix $J$ as

$$
J = Y + tZ,
$$

(49)

where $Y$ and $Z$ are positive definite matrices.
Using eigenvalue decomposition, \( Y^{-1/2} Z Y^{-1/2} = \Lambda \Lambda^H \). Defining a function \( h(t) \), such that \( h(t) = \text{tr}[J^{-1}] \), we write

\[
\begin{align*}
    h(t) &= \text{tr}\left[ (Y + tZ)^{-1} \right] \\
         &= \text{tr}\left[ Y^{-1/2} (I + tY^{-1/2} Z Y^{-1/2})^{-1} Y^{-1/2} \right] \\
         &= \text{tr}\left[ Y^{-1} (I + tY^{-1/2} Z Y^{-1/2})^{-1} \right] \\
         &= \text{tr}\left[ Y^{-1} A (I + tA)^{-1} A^H \right] \\
         &= \text{tr}\left[ A^H Y^{-1} A (I + tA)^{-1} \right] \\
         &= \sum_{i=1}^{K} [A^H Y^{-1} A]_{ii} (1 + t\lambda_i)^{-1}.
\end{align*}
\]

From (50), we see that \( h \) is a positive weighted sum of convex functions and is therefore convex in \( J \). Since \( J \) is a linear function of \( Q \), then problem (47) which has linear constraints is convex in \( Q \).

REFERENCES

Fig. 1. The downlink and virtual uplink in multiuser communications respectively.

Fig. 2. Minimum transmission power vs. $\gamma$ for $M = 8$, $L_k = 2$ and $N_k = 2$. 
TABLE I
POWER MINIMIZATION ALGORITHM

Initialization: $C = 0$, $U = \frac{1}{\sqrt{M}}[1, \ldots, 1]$ and $p = (P_{\text{max}}/L)[1, \ldots, 1]^T$

Iteration:
1- Downlink Receive Beamforming (for $k = 1 : K, l = 1 : L_k$)
   $v_{kl} = \max(S_{kl}^{DL}, T_{kl}^{DL})$
   $v_{kl} = v_{kl}/\|v_{kl}\|$
2- Virtual Uplink Power Allocation
   If $C < 1$
      solve $\Delta q_{eext} = \frac{1}{\lambda_{\text{max}}} q_{eext}$, let $C = 1/\lambda_{\text{max}}$
   else
      $q = \sigma^2(D^{-1} - \Psi)^{-1}1$
3- Virtual Uplink Receive Beamforming (for $k = 1 : K, l = 1 : L_k$)
   $u_{kl} = \max(S_{kl}^{UL}, T_{kl}^{UL})$
   $u_{kl} = u_{kl}/\|u_{kl}\|$
4- Downlink Power Allocation
   If $C < 1$
      solve $\Upsilon p_{eext} = \frac{1}{\lambda_{\text{max}}} p_{eext}$, let $C = 1/\lambda_{\text{max}}$
   else
      $p = \sigma^2(D^{-1} - \Psi)^{-1}1$
      $P = \|p\|_1$
5- Repeat steps 1-4 until convergence

TABLE II
SMSE MINIMIZATION ALGORITHM USING UPLINK

Initialization: $V_k = \text{SVD}(H_k)$ and $q = (P_{\text{max}}/L)[1, \ldots, 1]^T$

Iteration:
1- Virtual Uplink Transmit Beamforming (for $k = 1 : K, l = 1 : L_k$)
   $v_{kl} = \max(H_k^{DL} J_k^2 H_k, I/q_{kl} + H_k^{UL} J_k^{-1} H_k)$
   $v_{kl} = v_{kl}/\|v_{kl}\|$
2- Virtual Uplink Power Allocation
   $q = \arg\min_q \text{tr}[J^{-1}], \text{subject to } q_{kl} \geq 0, \|q\|_1 = P_{\text{max}}$
3- Repeat 1-2 until oldSMSE - newSMSE < $\epsilon$

Update:
4- Downlink Transmit Beamforming (for $k = 1 : K$)
   $U_k = J^{-1} H_k V_k \sqrt{Q_k}$
5- Set target SINR to actual SINR (for $k = 1 : K, l = 1 : L_k$)
   $\gamma_{kl} = SINR_{kl}^{DL}$
6- Downlink Power Allocation
   $p = \sigma^2(D^{-1} - \Psi)^{-1}1$

TABLE III
ITERATIVE SMSE MINIMIZATION ALGORITHM

Initialization: $V_k = \text{SVD}(H_k)$ and $q (P_{\text{max}}/L)[1, \ldots, 1]^T$

Iteration:
1- Virtual Uplink Power Allocation
   $q = \arg\min_q \text{tr}[J^{-1}], \text{subject to } q_{kl} \geq 0, \|q\|_1 = P_{\text{max}}$
2- Virtual Uplink Receive Beamforming (for $k = 1 : K$)
   $U_k = J^{-1} H_k V_k \sqrt{Q_k}$
3- Set target SINR to actual SINR (for $k = 1 : K, l = 1 : L_k$)
   $\gamma_{kl} = SINR_{kl}^{DL}$
4- Downlink Power Allocation
   $p = \sigma^2(D^{-1} - \Psi)^{-1}1$
5- Downlink Receive Beamforming (for $k = 1 : K$)
   $V_k = G_k^{-1} H_k^{UL} U_k \sqrt{P_k}$
6- Repeat 1-5 until oldSMSE - newSMSE < $\epsilon$
Fig. 3. $C$ vs. $P_{\text{max}}$ for $\gamma=1\text{dB}$, $K=3$, $L=[2\ 1\ 2]$ and $N=[2\ 2\ 3]$.

Fig. 4. Minimum required power for linear and non-linear precoding for $M = 8$, $L_k = 1$, $N_k = 6$ and $\gamma = 10\text{dB}$.
Fig. 5. No. of Iterations vs. SNR for SMSE minimization, K=3, M=6, L=[2 1 2] and N=[2 2 3].

Fig. 6. BER vs. SNR for MIMO1-SMSE minimization.