Joint Optimization for Multiuser MIMO Uplink Systems with Imperfect CSI

Minhua Ding and Steven D. Blostein

Dept. of Electrical and Computer Engineering
Queen's University, Kingston, Ontario, K7L3N6, Canada
Email: mding@ee.queensu.ca, steven.blostein@queensu.ca

***Abstract*— In this paper, a joint linear minimum sum mean-squared error transceiver optimization problem is formulated for multiuser MIMO uplink systems under a sum power constraint assuming imperfect channel state information (CSI). Two methods are proposed to solve this problem. One is based on the associated Karush-Kuhn-Tucker conditions. The other is to solve an equivalent problem, approaching the solution by solving a sequence of semi-definite programming problems. After obtaining the solution to the optimization problem, we investigate the effects of channel estimation errors and antenna correlation at the base station on system performance. Simulation results are provided.**

I. Introduction

Due to its low complexity as well as its effectiveness in managing both multiple access interference and inter-stream interference, joint minimum sum mean-squared error (MSMSE) linear precoder-decoder design has been proposed to improve multiuser MIMO spatial multiplexing systems [1]-[7]. Hereafter, we refer to a precoder and decoder pair for each user as a transceiver pair.

Joint MSMSE linear transceiver designs for the MIMO uplink have been studied under both sum power and per-user power constraints [1]-[3]. A separate treatment for the downlink can be found in [4][5]. More recently, an uplink-downlink duality has been found, which implies that with perfect channel state information (CSI), under the same sum power constraint, the achievable signal-to-interference-plus-noise ratio regions or the MSE regions for both links are the same [6]. Based on the duality, the more involved downlink problem has been tackled by forming and solving a dual uplink problem [6]. The same idea has also been adopted in [7].

Most of previous work has assumed *perfect* CSI. However, in practice, CSI is imperfect. Recently, a duality in average sum MSE between the uplink and the dual downlink with imperfect CSI has been shown in [8] and [9] using different approaches. Therefore, in this paper, we focus only on the uplink transceiver design with *imperfect CSI*. Our goal is to jointly optimize the transceiver pairs and then investigate the effects of channel estimation errors and antenna correlation at the base station (BS) on system performance.

After presenting the model of imperfect CSI, we formulate the uplink transceiver optimization problem. Two methods are proposed. One is based on the Karush-Kuhn-Tucker (KKT) conditions associated with the original problem. The other is to solve an equivalent problem, approaching the solution by solving a sequence of semi-definite programming (SDP) problems. The effects of channel estimation errors and channel correlation on system performance are then assessed by simulation.

II. System model and problem formulation

A. Uplink system model

Consider a single cell in cellular communication systems. The BS is equipped with M antennas. There are K mobile stations (MSs, users), each with N_i antennas, $i = 1, \ldots, K$. The uplink channels are denoted by $\mathbf{H}_i$, $i = 1, \ldots, K$.

Suppose that user i has l_i data streams, denoted by the $l_i \times 1$ ($l_i \leq \min(M, N_i)$) vector $\mathbf{x}_i$, $i = 1, \ldots, K$. These data vectors are assumed to be zero-mean, white with $E(\mathbf{x}_i \mathbf{x}_i^H) = \mathbf{I}_{l_i}$, for all i ($\forall i$), and mutually independent among users. Here $\mathbf{I}_m$ denotes the $m \times m$ identity matrix. Before the data streams are sent into the air, a linear precoder is employed for each user, which is denoted by the $N_i \times l_i$ matrix $\mathbf{F}_i$, $i = 1, \ldots, K$. The signal vector received at the BS antennas is given by $\mathbf{y}_{ul} = \sum_{i=1}^{K} \mathbf{H}_i \mathbf{F}_i \mathbf{x}_i + \mathbf{n}_{ul}$. The noise vector $\mathbf{n}_{ul}$ is zero-mean white complex Gaussian, i.e., $\mathcal{N}_c(0, \sigma_n^2 \cdot \mathbf{I}_M)$. The data and the noise are assumed to be statistically independent. At the BS, to recover the data for the user j, a linear decoder, denoted by the $l_j \times M$ matrix $\mathbf{G}_j$, is used. An estimate of the data vector for user j, $j = 1, \ldots, K$, can thus be expressed as $\mathbf{r}_{ul,j} = \mathbf{G}_j \mathbf{y}_{ul} = \mathbf{G}_j \left[\sum_{i=1}^{K} \mathbf{H}_i \mathbf{F}_i \mathbf{x}_i\right] + \mathbf{G}_j \mathbf{n}_{ul}$.

B. Channel model and imperfect channel state information

It is assumed that the antennas at each MS are spatially uncorrelated due to the presence of a large number of local scatterers. Therefore, the uplink channel model is given by [10]: $\mathbf{H}_i = \mathbf{\Sigma}_i^{1/2} \mathbf{H}_{wi}$, where $\mathbf{\Sigma}_i$ (seen by user i) is the normalized BS antenna correlation matrix with unit diagonal entries, $i = 1, \ldots, K$. The entries of $\mathbf{H}_{wi}$ are independent and identically-distributed (i.i.d.) $\mathcal{N}_c(0, 1)$, $\forall i$. In practice, CSI is obtained through channel estimation. The uplink CSI model at the BS can be expressed as [17]: $\mathbf{H}_i = \hat{\mathbf{H}}_i + \mathbf{E}_i$, $i = 1, \ldots, K$, where $\hat{\mathbf{H}}_i = \mathbf{\Sigma}_i^{1/2} \hat{\mathbf{H}}_{wi}$, and $\mathbf{E}_i = \mathbf{\Sigma}_i^{1/2} \mathbf{E}_{wi}$. The entries of $\hat{\mathbf{H}}_{wi}$ and $\mathbf{E}_{wi}$ are i.i.d. $\mathcal{N}_c(0, (1 - \sigma_{Ei}^2))$ and $\mathcal{N}_c(0, \sigma_{Ei}^2)$, respectively, where σ_{Ei}^2 is the channel estimation error variance for user i, $i = 1, \ldots, K$. Furthermore, for each i, the entries of $\hat{\mathbf{H}}_i$ and $\mathbf{E}_i$ are independent. We assume that the channel estimates $\{\hat{\mathbf{H}}_i\}_{i=1}^K$, the channel estimation error variances $\{\sigma_{Ei}^2\}_{i=1}^K$, the noise variance σ_n^2, and the BS antenna correlation matrices $\{\mathbf{\Sigma}_i\}_{i=1}^K$ are available at the BS. Here $\{a_i\}_{i=1}^K$ denotes $\{a_1, \ldots, a_K\}$.

C. Problem formulation

With the above CSI model,

$$\mathbf{y}_{ul} = \sum_{i=1}^K (\hat{\mathbf{H}}_i + \mathbf{E}_i)\mathbf{F}_i\mathbf{x}_i + \mathbf{n}_{ul}.$$

The MSE matrix for user j is given by

$$\begin{aligned} MSE_{ul,j} &= E[(\mathbf{r}_{ul,j} - \mathbf{x}_j)(\mathbf{r}_{ul,j} - \mathbf{x}_j)^H] \\ &= \mathbf{G}_j \left\{ \left[\sum_{i=1}^K \hat{\mathbf{H}}_i \mathbf{F}_i \mathbf{F}_i^H \hat{\mathbf{H}}_i^H \right] + \sigma_n^2 \mathbf{I}_M \right\} \mathbf{G}_j^H \\ &- \mathbf{G}_j \hat{\mathbf{H}}_j \mathbf{F}_j - \mathbf{F}_j^H \hat{\mathbf{H}}_j^H \mathbf{G}_j^H + \mathbf{I}_{l_j} \\ &+ \mathbf{G}_j \left[\sum_{i=1}^K \mathbf{\Sigma}_i^{1/2} E(\mathbf{E}_{wi}\mathbf{F}_i\mathbf{F}_i^H\mathbf{E}_{wi}^H)\mathbf{\Sigma}_i^{1/2} \right] \mathbf{G}_j^H, \forall j, \end{aligned}$$

where $E\{\cdot\}$ denotes statistical expectation taken over the distributions of the channel noise vector, the data vectors and the channel estimation error matrices. Let $tr(\cdot)$ denote the trace operation. Then it can be shown that

$$\begin{aligned} MSE_{ul,j} = \mathbf{G}_j &\left\{ \left[\sum_{i=1}^K \hat{\mathbf{H}}_i \mathbf{F}_i \mathbf{F}_i^H \hat{\mathbf{H}}_i^H \right] + \sigma_n^2 \cdot \mathbf{I}_M \right\} \mathbf{G}_j^H \\ &+ \mathbf{G}_j \left[\sum_{i=1}^K \sigma_{Ei}^2 \cdot tr(\mathbf{F}_i\mathbf{F}_i^H) \cdot \mathbf{\Sigma}_i \right] \mathbf{G}_j^H \\ &- \mathbf{G}_j \hat{\mathbf{H}}_j \mathbf{F}_j - \mathbf{F}_j^H \hat{\mathbf{H}}_j^H \mathbf{G}_j^H + \mathbf{I}_{l_j}, \end{aligned}$$

where $E\{\mathbf{E}_{wi}\mathbf{F}_i\mathbf{F}_i^H\mathbf{E}_{wi}^H\} = \sigma_{Ei}^2 \cdot tr(\mathbf{F}_i\mathbf{F}_i^H) \cdot \mathbf{I}_M$ has been used. The sum MSE from all users is then given by $mse_{ul,t} = \sum_{j=1}^K tr(MSE_{ul,j})$. The uplink problem is to minimize the (average) sum MSE subject to (s.t.) a sum power constraint:

$$\min_{\{(\mathbf{F}_j,\mathbf{G}_j)\}_{j=1}^K} mse_{ul,t} \quad s.t. \quad \sum_{j=1}^K tr(\mathbf{F}_j\mathbf{F}_j^H) \leq P_T. \tag{1}$$

After obtaining the solution to the above problem, we investigate the effect of channel estimation errors and antenna correlation on system performance.

In the following, we assume that the joint optimizations are performed at the BS, and then the optimum filters (i.e., precoders) for the users are sent to the MSs.

III. The joint optimization

A. An iterative algorithm for solving (1) based on the KKT conditions

The problem in (1) is non-convex. However, it can be shown that a global minimum exists for (1) [17]. Furthermore, both the objective and the constraint functions are *continuously differentiable*. Since we only have one inequality constraint, any feasible set $\{\mathbf{F}_i, \mathbf{G}_i\}_{i=1}^K$ is regular. Thus the KKT conditions are necessary for optimality [16]. Similar to [2][5], we now propose an algorithm developed from the KKT conditions. The Lagrangian associated with (1) is: $\mathcal{L}_{ul} = mse_{ul,t} + \mu_{ul}\{[\sum_{j=1}^K tr(\mathbf{F}_j\mathbf{F}_j^H)] - P_T\}$, where μ_{ul} is the Lagrange multiplier associated with the sum power constraint. The associated KKT conditions are given by (2)-(5) (Note: $k = 1, \ldots, K$):

$$\begin{aligned} \mathbf{F}_k^H \hat{\mathbf{H}}_k^H &= \mathbf{G}_k \left\{ \sum_{j=1}^K \hat{\mathbf{H}}_j \mathbf{F}_j \mathbf{F}_j^H \hat{\mathbf{H}}_j^H + \sigma_n^2 \cdot \mathbf{I}_M \right\} \\ &+ \mathbf{G}_k \left[\sum_{j=1}^K \sigma_{Ej}^2 \cdot tr(\mathbf{F}_j\mathbf{F}_j^H) \cdot \mathbf{\Sigma}_j \right], \end{aligned} \tag{2}$$

$$\begin{aligned} \hat{\mathbf{H}}_k^H \mathbf{G}_k^H &= \left\{ \hat{\mathbf{H}}_k^H \left[\sum_{j=1}^K \mathbf{G}_j^H \mathbf{G}_j \right] \hat{\mathbf{H}}_k \right\} \mathbf{F}_k \\ &+ \left\{ \mu_{ul} + \sigma_{Ek}^2 \sum_{j=1}^K tr(\mathbf{G}_j \mathbf{\Sigma}_k \mathbf{G}_j^H) \right\} \mathbf{F}_k, \end{aligned} \tag{3}$$

$$\mu_{ul} \geq 0, \quad \sum_{j=1}^K tr(\mathbf{F}_j\mathbf{F}_j^H) \leq P_T, \tag{4}$$

$$\mu_{ul} \cdot \left[\sum_{j=1}^K tr(\mathbf{F}_j\mathbf{F}_j^H) - P_T \right] = 0. \tag{5}$$

Proposition 1: (Relation between the Lagrange multiplier and the receive filters) For any solutions satisfying the KKT conditions, the following identity holds:

$$\mu_{ul} = (\sigma_n^2/P_T) \cdot \sum_{k=1}^K tr(\mathbf{G}_k\mathbf{G}_k^H). \tag{6}$$

Proof: The proof is solely based on the KKT conditions. It is omitted due to space constraints and can be found in [17].

In Table I, an iterative algorithm is developed based on the KKT conditions [(2)-(5)]. Similar algorithms have been used in [2] and [5] with perfect CSI. However, here we update the Lagrange multiplier using (6), which is simpler and more accurate than the method in [2][5], as the latter requires eigenvalue decompositions and a solution to a non-linear equation for each update of the Lagrange multiplier.

TABLE I

THE KKT-BASED ITERATIVE ALGORITHM

1)	Initialize $\mathbf{F}_k,\ k = 1, \ldots, K$, which are non-zero and satisfy the power constraint with equality.
2)	Update $\mathbf{G}_k$ using (2), $k = 1, \ldots, K$; $\mathbf{G}_k = \mathbf{F}_k^H \hat{\mathbf{H}}_k^H [\sum_{j=1}^K \hat{\mathbf{H}}_j \mathbf{F}_j \mathbf{F}_j^H \hat{\mathbf{H}}_j^H + \sigma_n^2 \cdot \mathbf{I}_M + \sum_{j=1}^K \sigma_{Ej}^2 \cdot tr(\mathbf{F}_j\mathbf{F}_j^H) \cdot \mathbf{\Sigma}_j]^{-1}$;
3)	Update μ_{ul} using (6);
4)	Update $\mathbf{F}_k$ using (3), $k = 1, \ldots, K$; $\mathbf{F}_k = [\hat{\mathbf{H}}_k^H (\sum_{j=1}^K \mathbf{G}_j^H \mathbf{G}_j) \hat{\mathbf{H}}_k + \mu_{ul} \cdot \mathbf{I}_{N_k} + \sigma_{Ek}^2 \cdot \sum_{j=1}^K tr(\mathbf{G}_j \mathbf{\Sigma}_k \mathbf{G}_j^H) \cdot \mathbf{I}_{N_k}]^{-1} \hat{\mathbf{H}}_k^H \mathbf{G}_k^H$;
5)	If the termination condition is met, stop; otherwise, go back to 2).

As in [2][5][7], we cannot show that the iterative algorithm in Table I is guaranteed to achieve the globally optimum solution (except when $K = 1$ [15]), despite the fact that the global minimum exists. This is because the objective function in (1) is not convex in $\{\mathbf{F}_i, \mathbf{G}_i\}_{i=1}^K$ [1], and thus the KKT

[1]Note that in (1), the objective function is non-convex in $\{\mathbf{F}_i, \mathbf{G}_i\}_{i=1}^K$ even when $\sigma_{Ei}^2 = 0, \forall i$.

conditions are not sufficient for global optimality. However, starting from a set of non-zero $\{\mathbf{F}_i\}_{i=1}^K$, the algorithm yields reasonable results as shown by simulation in Section IV. In the next subsection, we seek an alternative solution.

B. Solving (1) by solving an equivalent problem

In this subsection, we assume that the number of data streams is equal to the number of transmit antennas for each user, i.e., $l_i = N_i, i = 1, \ldots, K$. The uplink problem (1) can be equivalently formulated as [12]

$$\min_{\{\mathbf{F}_i\}_{i=1}^K,\ \sum_{i=1}^K tr(\mathbf{F}_i\mathbf{F}_i^H)\le P_T}\ \min_{\{\mathbf{G}_i\}_{i=1}^K} mse_{ul,t}.$$

It turns out that the inner minimization is achieved when (2) is satisfied for all k. Then it can be shown that

$$mse_{ul,t} = tr\left\{\mathbf{X}\left[\sigma_n^2\mathbf{I}_M + \sum_{j=1}^K \sigma_{Ej}^2 tr(\mathbf{Q}_j)\mathbf{\Sigma}_j\right]\right\} + const,$$

where *const* denotes a constant that equals $\sum_{j=1}^K tr\left(\mathbf{I}_{l_j}\right) - tr\left(\mathbf{I}_M\right)$ and we have defined that $\mathbf{Q}_j = \mathbf{F}_j\mathbf{F}_j^H, \forall j$, and

$$\mathbf{X} = \left[\sum_{j=1}^K \hat{\mathbf{H}}_j\mathbf{Q}_j\hat{\mathbf{H}}_j^H + \sum_{j=1}^K \sigma_{Ej}^2 tr(\mathbf{Q}_j)\mathbf{\Sigma}_j + \sigma_n^2\mathbf{I}_M\right]^{-1}. \quad (7)$$

Therefore, instead of solving (1) directly, we attempt to solve the following equivalent formulation [12]:

$$\min_{\mathbf{X};\ \{\mathbf{Q}_j\}_{j=1}^K} tr\left\{\mathbf{X}\left[\sum_{j=1}^K \sigma_{Ej}^2 tr(\mathbf{Q}_j)\mathbf{\Sigma}_j + \sigma_n^2\mathbf{I}_M\right]\right\}, \quad (8)$$

$$s.t. \sum_{j=1}^K tr(\mathbf{Q}_j) \le P_T, \quad (9)$$

$$\mathbf{X} \text{ as given in (7)}, \quad (10)$$

$$\mathbf{Q}_j \succeq 0, j = 1, \ldots, K. \quad (11)$$

Proposition 2: When $\sigma_{Ej}^2 = \sigma_E^2$ and $\mathbf{\Sigma}_j = \mathbf{\Sigma}_{BS}, j = 1, \ldots, K$, (8)-(11) is equivalent to the following SDP problem:

$$\min_{\mathbf{X};\ \{\mathbf{Q}_j\}_{j=1}^K} tr\left\{\mathbf{X}\cdot\left[\sigma_n^2\mathbf{I}_M + \sigma_E^2 P_T\mathbf{\Sigma}_{BS}\right]\right\},$$

$$s.t. \sum_{j=1}^K tr(\mathbf{Q}_j) \le P_T,$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{I}_M \\ \mathbf{I}_M & \sum_{j=1}^K \hat{\mathbf{H}}_j\mathbf{Q}_j\hat{\mathbf{H}}_j^H + \sigma_E^2 P_T\mathbf{\Sigma}_{BS} + \sigma_n^2\mathbf{I}_M \end{pmatrix} \succeq 0,$$

$$\mathbf{Q}_j \succeq 0, j = 1, \ldots, K,$$

where $\mathbf{A} \succeq 0$ means that matrix $\mathbf{A}$ is positive semi-definite.

Proof: The proof is an extension of that in [3]. Details can be found in [17].

A SDP involves a convex optimization [12] and can be solved using the software in [13]. A globally optimum solution is guaranteed. Once the globally optimum solution is obtained, $\{\mathbf{F}_j\}_{j=1}^K$ can be obtained by performing Cholesky factorizations[2] of $\{\mathbf{Q}_j\}_{j=1}^K$, due to the assumption that $l_j = N_j, j = 1, \ldots, K$ [3]. The corresponding $\{\mathbf{G}_j\}_{j=1}^K$ can be obtained using (2).

[2]Here the optimum $\{\mathbf{Q}_j\}_{j=1}^K$ is unique. However, $\{\mathbf{F}_j\}_{j=1}^K$ is not unique. If $\{\mathbf{F}_j\}_{j=1}^K$ is an optimum set, in terms of sum MSE, then $\{\mathbf{F}_j\mathbf{U}_j\}_{j=1}^K$ is also an optimum set, where $\{\mathbf{U}_j\}_{j=1}^K$ is any set of unitary matrices of proper size. We take this into account in our simulations.

Clearly, the result in **Proposition 2** has very limited application, because of the conditions required ($\sigma_{Ej}^2 = \sigma_E^2$ and $\mathbf{\Sigma}_j = \mathbf{\Sigma}_{BS}, \forall j$). In general, the equivalent problem given by (8)-(11) is not a SDP, because the objective function in (8) is not convex. However, **Proposition 2** provides a basis to find a solution to the equivalent problem. Specifically, we have a SDP-based iterative algorithm given in Table II. (See below.)

TABLE II

THE SDP-BASED ITERATIVE ALGORITHM

1)	Initialize $\mathbf{Q}_j = [P_T/(KN_j)]\cdot\mathbf{I}_{N_j}, \forall j$. Calculate the value of the objective function f^{old} using (7) and (8), given $\{\mathbf{Q}_j\}_{j=1}^K$.
2)	Given $\{\mathbf{Q}_j\}_{j=1}^K$, calculate $\mathbf{B} = \sum_{j=1}^K \sigma_{Ej}^2 tr(\mathbf{Q}_j)\mathbf{\Sigma}_j + \sigma_n^2\mathbf{I}_M$.
3)	Solve the SDP problem given by (12)-(15) to obtain a new set of $\{\mathbf{Q}_j\}_{j=1}^K$. Calculate the value of the objective function f^{new}, i.e., the value of (12).
4)	If $\lvert f^{new} - f^{old}\rvert \le \varepsilon$, stop; otherwise, set $f^{old} := f^{new}$, and go back to 2).

During each iteration, the matrix $\mathbf{B}$ is fixed and thus the problem given by (12)-(15) is a SDP problem:

$$\min_{\mathbf{X};\ \{\mathbf{Q}_j\}_{j=1}^K} tr\{\mathbf{XB}\} \quad (12)$$

$$s.t. \sum_{j=1}^K tr(\mathbf{Q}_j) \le P_T, \quad (13)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{I}_M \\ \mathbf{I}_M & \sum_{j=1}^K \hat{\mathbf{H}}_j\mathbf{Q}_j\hat{\mathbf{H}}_j^H + \mathbf{B} \end{pmatrix} \succeq 0, \quad (14)$$

$$\mathbf{Q}_j \succeq 0, j = 1, \ldots, K. \quad (15)$$

Essentially, the algorithm in Table II approaches the solution by solving a sequence of SDP problems which approximate and converge to that given by (8)-(11). After obtaining $\{\mathbf{Q}_j\}_{j=1}^K$, we can obtain $\{\mathbf{F}_j, \mathbf{G}_j\}_{j=1}^K$ as mentioned earlier.

Remark: According to the uplink-downlink duality with imperfect CSI [8][9], when we need to jointly optimize the linear MSMSE transceiver pairs with imperfect CSI for *the downlink*, we can first formulate a dual uplink problem, find the $\{\mathbf{F}_i, \mathbf{G}_i\}_{i=1}^K$ for the uplink using the above two methods, and then translate the transceiver pairs for application in the downlink. (Note that to use duality, we do not need to use the assumption of channel reciprocity. The dual link is virtual.)

IV. NUMERICAL RESULTS

A. Simulation setup

The BS antenna correlations for different users are given by [14][15]: $\mathbf{\Sigma}_{i,pq} = \rho_i^{|p-q|}$, where $0 \le \rho_i < 1, i = 1, \ldots, K$, and $p, q \in \{1, \ldots, M\}$. When simulating the average bit error probability (ABEP) performance, 4-QAM is used in each user's data streams. The sum MSE and the ABEP of User 1 are displayed.

B. Comparison of the KKT-based algorithm and the SDP-based algorithm

To show the equivalence of results obtained from the two algorithms in Tables I and II, we provide an example. Here the error variances are: $\sigma_{E1}^2 = 0.01, \sigma_{E2}^2 = 0.015, \sigma_{E3}^2 = 0.008$. The BS antenna correlation exponents are: $\rho_1 = 0.8, \rho_2 = 0.7, \rho_3 = 0.5$. We set $P_T/\sigma_n^2 = 20$ dB. In addition, let the number of users $K = 3$, the number of BS antennas $M = 6$, and the number of antennas at each user $N_i = l_i = 2, i = 1, \ldots, K$. Using the channel estimation method in Section II-B, we obtain the following channel estimates based on a specific channel realization:

$$\hat{\mathbf{H}}_1 = \begin{pmatrix} -0.3597 - 0.1932i & 0.1346 + 0.0584i \\ 0.3479 - 0.7599i & 0.2668 - 0.8746i \\ 0.4694 - 0.1864i & -0.6294 - 0.1607i \\ 0.1349 + 0.0139i & -0.8027 + 0.4571i \\ 0.8689 - 0.2719i & 0.0640 + 0.2839i \\ 0.5744 - 0.0873i & -1.1839 + 0.0323i \end{pmatrix},$$

$$\hat{\mathbf{H}}_2 = \begin{pmatrix} -0.1467 - 0.7230i & 0.4247 + 0.7612i \\ 0.8363 - 1.0421i & 0.2147 - 0.4386i \\ 0.0837 - 0.5144i & 0.2536 + 0.2494i \\ 0.0150 - 0.0124i & 0.0077 + 1.0058i \\ 0.2739 + 0.0911i & 0.4388 + 0.9669i \\ 0.2869 + 1.0564i & -0.0046 - 0.1422i \end{pmatrix},$$

$$\hat{\mathbf{H}}_3 = \begin{pmatrix} 0.9220 + 0.2587i & -0.0724 - 1.1417i \\ 0.4581 - 0.0645i & 0.4294 - 1.0561i \\ -0.0474 - 0.1073i & 0.0493 + 0.0204i \\ 0.7621 - 0.4112i & -0.8851 + 0.0391i \\ 0.3653 + 0.5164i & -0.3619 + 0.3661i \\ 0.6788 + 0.5998i & -0.8428 + 0.3026i \end{pmatrix}.$$

The comparison of the two algorithms is given in Table III. We find that the two algorithms yield equivalent results. However, the complexity of the SDP-based algorithm is much higher than that of the KKT-based algorithm, where complexity is measured by the computation time required for both algorithms to converge. Both algorithms are run on the same hardware using MATLAB.

TABLE III

A COMPARISON OF THE TWO ALGORITHMS

	SDP-based	KKT-based
achieved sum MSE	0.9384	0.9372
average computation time (seconds)	5.3030	0.1090

Similar comparison results can also be observed for different channel realizations and with different system parameters. Fig. 1 shows more comparisons of these two algorithms with different values of P_T/σ_n^2 and with different channel correlation parameters. Each point on the curve is obtained by averaging the sum MSE from 10,000 channel realizations.

When $\sigma_{Ei}^2 = \sigma_E^2$ and $\rho_i = \rho, \forall i$ (i.e., $\mathbf{\Sigma}_i = \mathbf{\Sigma}_{BS}, \forall i$), the KKT-based algorithm yields the equivalent sum MSE as from solving a SDP problem (see **Proposition 2**). Fig. 2 shows the comparison results, where $\sigma_E^2 = 0.01$, $\rho = 0.5$, $K = 3, M = 6, N_i = l_i = 2, \forall i$.

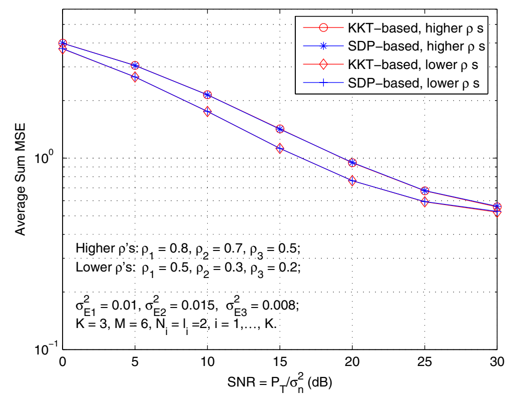

Fig. 1. A comparison between the sum MSEs obtained from the KKT-based and the SDP-based algorithms described in Tables I and II.

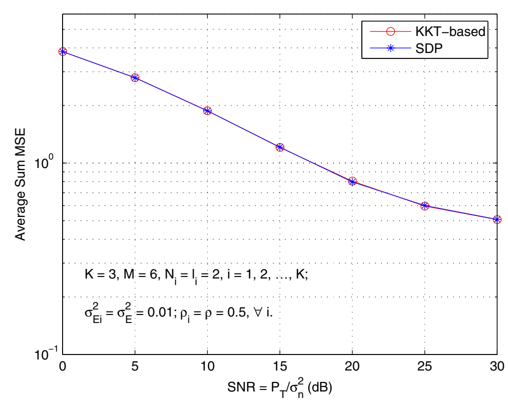

Fig. 2. A comparison between the sum MSEs obtained from the KKT-based algorithm and the SDP described in **Proposition 2**.

We can see that the results obtained using the KKT-based algorithm are consistent with those from the SDP-based algorithm as given by Table II, or from solving a single SDP problem as in the case specified by **Proposition 2**. Therefore, below we investigate the effect of channel estimation error and BS antenna correlation based on the $\{\mathbf{F}_i, \mathbf{G}_i\}_{i=1}^K$ obtained from the algorithm given in Table I or Table II.

C. Effect of channel estimation errors and BS antenna correlation

Fig. 3 shows the effect of channel estimation errors as well as the effect of channel correlation. Comparing Curves 1 and 3 or Curves 2 and 4, we observe that channel estimation errors cause a large performance degradation on the ABEP of User 1. Comparing Curves 1 and 2 or Curves 3 and 4, we can see that BS antenna correlation also has a significant impact on system performance. For example, from Curves 3 and 4, the performance degradation is about 4 dB when $[\rho_1, \rho_2, \rho_3]$ changes from $[0.5, 0.3, 0.2]$ to $[0.8, 0.7, 0.5]$ (with other parameters fixed).

Fig. 4 shows the ABEP results of User 1, when the number of BS antennas, M, increases from 6 to 8. Increasing M implies introducing more antenna diversity. Therefore, from Curves 1 and 3 or Curves 2 and 4 in Fig. 4, it is obvious that the effect of channel estimation errors can be compensated by introducing diversity. Note that one can also introduce diversity by transmitting fewer data streams (i.e., reducing $l_i, i = 1, \ldots, K$).

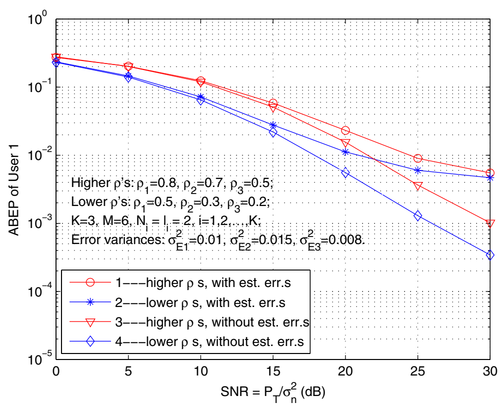

Fig. 3. Comparison of the ABEP of User 1 with or without channel estimation errors and with different amounts of channel correlation.

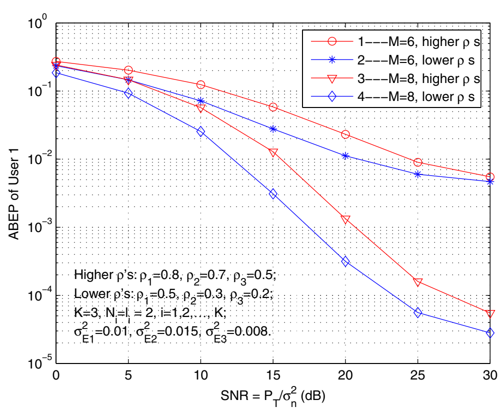

Fig. 4. Comparison of the ABEP of User 1 with channel estimation errors and different amounts of antenna diversity.

V. Summary

A joint linear transceiver optimization problem has been formulated for the multiuser MIMO uplink system. Two algorithms are proposed to obtain the solution. One is based on the KKT conditions, the other is based on solving a sequence of SDP problems. Simulation results have shown the consistency in average sum MSEs obtained from both algorithms. The effect of channel estimation errors and BS antenna correlation has been assessed by simulation.

Acknowledgments

This research has been partially supported by Natural Sciences and Engineering Research Council of Canada Discovery Grant 41731. We would like to thank the reviewers for their helpful comments and for bringing [8] into our attention.

References

[1] E. A. Jorswieck, H. Boche, "Transmission strategies for the MIMO MAC with MMSE receiver: average MSE optimization and achievable individual MSE region," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2872-2881, Nov. 2003.

[2] S. Serbetli, A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Processing*, vol. 52, no. 1, pp. 214-226, Jan. 2004.

[3] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1037-1052, Nov. 2004.

[4] A. Tenenbaum, R. S. Adve, "Joint multiuser transmit-receive optimization using linear processing," in *Proc. IEEE ICC 2004*, Paris, France, June 2004.

[5] J. Zhang, Y. Wu, S. Zhou, J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Commun. Letters*, vol. 9, no. 11, pp. 991-993, Nov. 2005.

[6] M. Schubert, S. Shi, E. A. Jorswieck, H. Boche, "Downlink sum-MSE transceiver optimization for linear multi-user MIMO systems," in *Proc. the 39-th Asilomar Conference on Signals, Systems and Computers, 2005*, pp. 1424-1428, Oct. 28 - Nov. 1, 2005.

[7] A. Khachan, A. Tenenbaum, R. S. Adve, "Joint linear precoding-receiver optimization for multiuser downlink MIMO communications," submitted to *IEEE Trans. on Wireless Commun.*

[8] M. B. Shenouda, T. N. Davidson, "Statistically robust tranceiver design for broadcast channels with uncertainty," in *Proc. Canadian Conf. on Electrical and Computer Engineering (CCECE) 2007*, Apr. 2007.

[9] M. Ding, S. D. Blostein, "Relation between joint optimizations for multiuser MIMO uplink and downlink with imperfect CSI," in *Proc. IEEE ICASSP 2008*, Las Vegas, Mar. 30- Apr. 4, 2008.

[10] D. Shiu, G. J. Foschini, M. J. Gans, J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502-513, Mar. 2000.

[11] S. Serbetli, A. Yener, "MMSE transmitter design for correlated MIMO systems with imperfect channel estimates: power allocation trade-offs," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2295-2304, Aug. 2006.

[12] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[13] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Meth. Software*, vol. 11-12, pp. 625-653, 1999.

[14] A. B. Gershman, N. Sidiropoulos, Editors, *Space-Time Processing for MIMO Communications*, Wiley, 2005.

[15] M. Ding, S. D. Blostein, "MIMO LMMSE transceiver design with imperfect CSI at both ends," in *Proc. IEEE ICC'07*, Glasgow, Jun. 24-28, 2007.

[16] D. P. Bertsekas, *Nonlinear Programming*, second edition, Athena Scientific, 1999.

[17] M. Ding, "Multiple-input multiple-output system designs with imperfect channel knowledge," Ph.D thesis, Queen's University, Kingston, Ontario, Canada, to appear in 2008.