Transceiver Optimization for Multiuser MIMO Systems

Semih Serbetli, *Student Member, IEEE,* and Aylin Yener, *Member, IEEE*

***Abstract*—We consider the uplink of a multiuser system where the transmitters as well as the receiver are equipped with multiple antennas. Each user multiplexes its symbols by a linear precoder through its transmit antennas. We work with the system-wide mean squared error as the performance measure and propose algorithms to find the jointly optimum linear precoders at each transmitter and linear decoders at the receiver. We first work with the case where the number of symbols to be transmitted by each user is given. We then investigate how the symbol rate should be chosen for each user with optimum transmitters and receivers. The convergence analysis of the algorithms is given, and numerical evidence that supports the analysis is presented.**

***Index Terms*—MMSE receivers, multiuser MIMO system, receiver beamforming, transmitter beamforming.**

I. Introduction

DUE to the emerging demand on new multimedia applications, next-generation wireless systems are expected to support higher data rates. The scarcity of wireless bandwidth prompts the need for spectrally efficient methods. Using multiple transmit and receive antennas is an effective means to increase spectral efficiency [1], [2]. Recently, there has been considerable research in exploiting the space dimension through transmit diversity, space-time coding, and spatial multiplexing for multiple input multiple output (MIMO) systems that employ multiple transmit and/or receive antennas [3]–[5]. In particular, spatial multiplexing can be used to transmit multiple data streams that can be separated using receiver signal processing, e.g., [6] and [7].

Performance improvement for MIMO systems can be achieved by exploiting various levels of feedback information available at the transmitter. In the absence of channel state-related feedback, multiantenna structure of the system can be used for spatial multiplexing as in BLAST [6] or for coding as in space-time coding [3], [4]. Antenna selection is a spatial multiplexing technique that assumes limited feedback information from the receiver to the transmitter. The information about which transmitter antennas should be used to achieve a certain data rate is fed back, and the potential high capacity of a MIMO system can be realized with limited increase in complexity at the transmitter. Capacity of MIMO systems with antenna selection has been studied recently [8], [9]. Minimum error rate antenna selection methods are proposed for single-user MIMO systems employing maximum likelihood and linear receivers in the context of spatial multiplexing in [10]. Optimum MIMO transmission schemes with antenna selection are analyzed in [11].

Spatial multiplexing can significantly benefit from transmit precoding when channel information is available at the transmitter side. In such cases, designing the appropriate precoding strategy has been studied under a variety of system objectives [5], [7], [12], [13]. All of these studies, as with most of the MIMO system analysis, have been done for a single-user system that transmits multiple data streams. In the case of a multiuser MIMO system where users' transmissions interfere with each other, the system objectives should be optimized jointly for all users given the channels of all users. Thus, optimal designs of single-user systems are not directly applicable. In this context, optimum or near-optimum transmit strategies that maximize the information theoretic sum capacity of vector multiple access channels have been investigated [14], [15]. A recent reference considers optimum transmit strategies relevant for a multicarrier scenario [16].

Joint transmitter and receiver design is an effective interference management technique for multiuser systems. In particular, signature sequence optimization in CDMA systems, which aims to determine optimum transmitter sets to enhance the performance of the overall system, has been investigated for several channel models. Optimum CDMA signature sequence sets are identified, and iterative algorithms that converge to the optimum signature sequence set are proposed in [17]–[21]. For multipath CDMA systems, jointly optimum transmission schemes are investigated, and iterative algorithms to find the optimum signature sets are proposed in [22]–[24].

Transmit and receiver beamforming for multiuser MIMO systems when each user is transmitting a single data stream have also been studied extensively up to date. Receiver beamforming has been shown to be effective in interference suppression in multiuser systems [25], [26]. Jointly optimum transmit powers and receiver beamformers were found in [27]. Reference [28] proposed an iterative algorithm for determining the downlink powers and transmit beamformers given a signal-to-interference ratio (SIR) target at the single antenna receiver of each user. The optimality of a similar algorithm was shown in [29]. Algorithms that identify transmit and receiver beamforming strategies and the corresponding transmit power assignments are proposed in [30] with the aim of maximizing the minimum achievable SIR or providing each user with its SIR target. The algorithms suggested were numerically shown to enhance system performance but were observed to converge to local optima [30].

Manuscript received December 22, 2002; revised April 22, 2003. This work was supported in part by the National Science Foundation under CAREER Award 02–37727. This paper was presented in part in the 40th Allerton Conference on Communications, Control, and Computing, Monticello, IL, October 2002. The associate editor coordinating the review of this paper and approving it for publication was Dr. Rick S. Blum.

The authors are with the Electrical Engineering Department, The Pennsylvania State University, University Park, PA 16802 USA (e-mail: serbetli@psu.edu; yener@ee.psu.edu).

Digital Object Identifier 10.1109/TSP.2003.819988

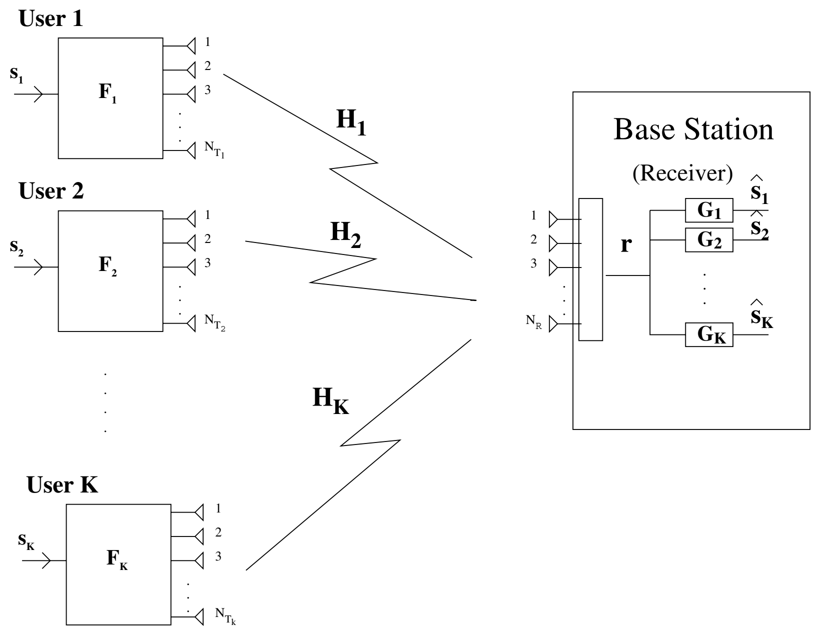

Fig. 1. System model of multiuser MIMO system.

Our aim in this work is to design algorithms that converge to the optimum transmitters (precoders) and receivers (decoders) for all users in a multiuser MIMO system when users transmit possibly multiple data streams. The channels are assumed to be flat and known at the receiver side, and we assume that there exists an error-free and low-delay feedback channel to each user. The transmitters and receivers are assumed to be linear for all users. A multiuser MIMO system can be viewed as a MIMO system with a channel matrix that consists of the channel gains of all transmitter-receiver pairs of all users where each user's symbols can be precoded only by that user's transmitter antennas. We work with a system-wide performance measure for the joint optimization of transmitters and receivers, namely, the system-wide mean squared error (MSE). In contrast to receiver optimization for fixed transmitters, e.g., in [31], optimization of the individual MSEs is not equivalent to total MSE optimization. However, one can construct iterative algorithms for the cases where users transmit single or multiple symbols that monotonically decrease the total MSE under the given system constraints.

For a given number of symbols to be transmitted by each user, we propose two different algorithms that use alternating minimization to find the optimum transmitter and receiver set. The first algorithm enables parallel updates for all users over the transmitters and receivers. The second algorithm allows us to observe that there exists an optimum transmitter set that distributes each user's power equally to each of its symbols. This observation enables the formulation of the multiuser MIMO system with multiple symbols transmitted by each user as a system where multiple virtual users transmitting one symbol each. We then propose an alternative iterative algorithm that optimizes the transmitter vector of each virtual user one at a time for faster convergence. The proposed algorithms are observed to converge to the best transmitter-receiver pairs under the given power constraints and result in enhanced performance for all users. We also observe that when the number of symbols to be transmitted by each user exceeds a certain threshold, the overall system performance degrades, and we suggest guidelines for the number of symbols transmitted per user.

The algorithms proposed employ iterative updates of the transmitters and receivers. It is important to note that the updates depend on the error-free and low-delay feedback channels and that the feedback can be of one of the two forms: Either i) the necessary information at each iteration is fed back to the transmitter side for users to perform the actual update, or ii) the algorithm is run offline, and the resulting transmitter for each user is fed back to the user.

Throughout the derivation of the algorithms, the channels are assumed to be flat and slowly varying. In fact, we need the channels to be varying slowly as compared with the convergence time of the proposed iterative algorithms. Since there is no structural constraint imposed on channel matrices, the algorithms are valid for any multiuser MIMO system having slow channel variations such as multiuser multiantenna systems [30] and OFDMA systems [32] in fixed (or very slow mobility) wireless environments and wireline multiuser systems [33], [34]. The algorithms are derived for synchronous uplink systems such as multiantenna TDMA systems (SDMA/TDMA systems) where time slots are shared by several users, and multiuser detection schemes are performed for each time slot, resulting in improved user capacity [35]–[37]. However, the algorithms can be extended for the asynchronous case simply by increasing the dimensions of the channel, transmitter, and receiver matrices by the time index, and optimum transmission schemes can be found for blocks of symbols for each user.

II. System Model and Performance Metric

We consider the uplink of a single cell synchronous system with K users. The receiver employs N_R antennas. We assume that the ith user multiplexes a fixed number of data streams M_i through its N_{T_i} transmit antennas employing an $N_{T_i} \times M_i$ linear transmitter $\mathbf{F}_i$ in one symbol period (Fig. 1). We assume that the number of symbols for each user is given and fixed first. We will remove this assumption in Section V and determine the number of symbols to be transmitted. Similar to the notation in [7], the received vector is

$$\mathbf{r} = \sum_{j=1}^{K} \mathbf{H}_j \mathbf{F}_j \mathbf{s}_j + \mathbf{n} \tag{1}$$

where $\mathbf{s}_i$ is the $M_i \times 1$ symbol vector, $\mathbf{H}_i$ is the $N_R \times N_{T_i}$ matrix of complex channel gains, $\mathbf{n}$ is the zero mean complex Gaussian noise vector with $E[\mathbf{n}\mathbf{n}^\dagger] = \sigma^2 \mathbf{I}$, and $(\cdot)^\dagger$ denotes the Hermitian of a vector or matrix. Denote the covariance of the transmitted signal vector of user i by $\mathbf{R}_i$. Assuming the symbols of each user are independent, the covariance of the transmitted signal vector is $\mathbf{R}_i = E[\mathbf{F}_i \mathbf{s}_i \mathbf{s}_i^\dagger \mathbf{F}_i^\dagger] = \mathbf{F}_i \mathbf{F}_i^\dagger$. The transmit power constraint for user i is $tr\{\mathbf{F}_i \mathbf{F}_i^\dagger\} = \text{tr}\{\mathbf{R}_i\} \leq p_i$. The $M_i \times N_R$ linear receiver of user i is denoted by $\mathbf{G}_i$. The decision statistic $\mathbf{y}_i$ is given by

$$\mathbf{y}_i = \mathbf{G}_i \left(\sum_{j=1}^{K} \mathbf{H}_j \mathbf{F}_j \mathbf{s}_j + \mathbf{n} \right). \tag{2}$$

In this work, we aim to design transmitter-receiver pairs that minimize the system-wide MSE. The MSE incurred by user i, MSE_i is

$$\begin{aligned} \text{MSE}_i &= E\left[\|\mathbf{y}_i - \mathbf{s}_i\|^2 \right] \\ &= \text{tr}\left\{ \sum_{j=1}^{K} \mathbf{F}_j{}^\dagger \mathbf{H}_j{}^\dagger \mathbf{G}_i{}^\dagger \mathbf{G}_i \mathbf{H}_j \mathbf{F}_j - \mathbf{F}_i{}^\dagger \mathbf{H}_i{}^\dagger \mathbf{G}_i{}^\dagger \right. \\ &\qquad \left. - \mathbf{G}_i \mathbf{H}_i \mathbf{F}_i + \mathbf{I} + \sigma^2 \mathbf{G}_i \mathbf{G}_i{}^\dagger \right\} \end{aligned} \tag{3}$$

where $\text{tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$. The total MSE of all users in the system is given by

$$\begin{aligned} \text{MSE} = \text{tr}\left\{ \sum_{i=1}^{K} \left\{ \sum_{j=1}^{K} \mathbf{F}_j{}^\dagger \mathbf{H}_j{}^\dagger \mathbf{G}_i{}^\dagger \mathbf{G}_i \mathbf{H}_j \mathbf{F}_j - \mathbf{F}_i{}^\dagger \mathbf{H}_i{}^\dagger \mathbf{G}_i{}^\dagger \right. \right. \\ \left. \left. - \mathbf{G}_i \mathbf{H}_i \mathbf{F}_i + \mathbf{I} + \sigma^2 \mathbf{G}_i \mathbf{G}_i{}^\dagger \right\} \right\}. \end{aligned} \tag{4}$$

The performance metric, system-wide total MSE is formulated for synchronous uplink systems such as multiantenna TDMA systems with no intersymbol interference. This assumption is valid, for instance, when a guard time is added to each symbol duration to prevent successive symbol interference and increasing the observation period. We note that the formulation can be adapted for asynchronous case simply by increasing the dimensions of the channel, transmitter, and receiver matrices by the time index for a given time interval.

Consider transmission of a block of L symbols for each user. User i has LM_i symbols to send, and the transmitter and receiver matrices of user i have dimensions of $LN_{T_i} \times LM_i$ and $LM_i \times LN_R$, respectively. The channel matrix of user i for an L symbol period of time can be constructed simply by inserting the channel gains related to the contributing symbol intervals. Defining the appropriate transmitter, receiver, and channel matrices, the system-wide total MSE can be formulated in the same manner as above. Note that for this approach, each symbol is precoded over a time interval of L symbol duration, and the optimum transmission scheme for a block of transmission is found. Keeping in mind that the following development may be extended to address the block transmission scenario, we will use the synchronous model in the sequel for clarity of exposition.

Total MSE minimization by choosing the transmitters and receivers has recently been studied for synchronous CDMA systems with single antennas in the context of CDMA signature optimization [20]. This performance measure is desirable to work with in transmitter optimization, in contrast with each user minimizing its own MSE, as is adapted in receiver optimization [31]. This is because the choice of the transmitter of a user affects the MSE of each user in the system. In the following sections, we pose the problem of minimizing the total MSE in the presence of power constraints and devise iterative algorithms that converge to the solution of the corresponding problems. Note that this problem is a generalized version of what is posed in [16] without any constraints on the number of users or channel structure.

III. MSE Minimization

A. Multiple Symbol Transmission

We now pose the problem of minimizing the MSE subject to a transmit power constraint for each user for a predefined number of symbols to be transmitted for each user $\{M_i\}$s. As is explained in Section II, the transmit power constraint for user i can be expressed as $\text{tr}(\mathbf{F}_i{}^\dagger \mathbf{F}_i) \leq p_i$. Formally, the optimization problem is

$$\min_{\{\mathbf{F}_i, \mathbf{G}_i\}_{i=1,\cdots,K}} \text{MSE} \tag{5}$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{F}_i{}^\dagger \mathbf{F}_i \right) \leq p_i \quad i = 1, \cdots, K \tag{6}$$

where we optimize the MSE over the $N_{T_i} \times M_i$ transmitter matrices $\{\mathbf{F}_i\}$ and the $M_i \times N_R$ receiver matrices $\{\mathbf{G}_i\}$. Notice that the only constraint imposed on the system is the transmit power constraint for each user. Then, the Lagrangian dual objective of this optimization problem is

$$\begin{aligned} &L\left(\{\mathbf{F}_i\}, \{\mathbf{G}_i\}, \{\mu_i\}_{i=1}^K \right) \\ &= \text{tr}\left\{ \sum_{i=1}^{K} \left\{ \sum_{j=1}^{K} \mathbf{F}_j{}^\dagger \mathbf{H}_j{}^\dagger \mathbf{G}_i{}^\dagger \mathbf{G}_i \mathbf{H}_j \mathbf{F}_j - \mathbf{F}_i{}^\dagger \mathbf{H}_i{}^\dagger \mathbf{G}_i{}^\dagger \right. \right. \\ &\qquad \left. \left. - \mathbf{G}_i \mathbf{H}_i \mathbf{F}_i + \mathbf{I} + \sigma^2 \mathbf{G}_i \mathbf{G}_i{}^\dagger \right\} \right\} \\ &\quad + \sum_{i=1}^{K} \mu_i \left[\text{tr}\left\{ \mathbf{F}_i{}^\dagger \mathbf{F}_i - p_i \right\} \right] \end{aligned} \tag{7}$$

where $\mu_i \geq 0$ is the Lagrange multiplier associated with the transmit power constraint of user i. Optimum transmitter and receiver structures should satisfy the first-order optimality conditions for each user. Simply taking the derivative with respect

TABLE I
MULTIPLE SYMBOL ALGORITHM, IMPLEMENTATION 1. SEE APPENDIX A FOR EXPLICIT CALCULATION OF $\mu_k(n+1)$

Step 1 :	Update the receivers in a parallel fashion
	For $k = 1 : K$
	$\mathbf{G}_k(n) = \mathbf{F}_k{}^\dagger(n)\mathbf{H}_k{}^\dagger \left(\sigma^2\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i\mathbf{F}_i(n)\mathbf{F}_i{}^\dagger(n)\mathbf{H}_i{}^\dagger\right)^{-1}$
	End
Step 2 :	Update the transmitters in a parallel fashion
	For $k = 1 : K$
	$\mathbf{X}_k(\bar{\mu}_k) = \left(\bar{\mu}_k\mathbf{I} + \sum_{i=1}^K \mathbf{H}_k{}^\dagger\mathbf{G}_i{}^\dagger(n)\mathbf{G}_i(n)\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{G}_k{}^\dagger(n)$
	$\mu_k(n+1) = \lfloor \bar{\mu}_k \text{ such that } \mathrm{tr}\{\mathbf{X}_k(\bar{\mu}_k)\mathbf{X}_k(\bar{\mu}_k)^\dagger\} = p_k \rfloor_+$
	$\mathbf{F}_k(n+1) = \left(\mu_k(n+1)\mathbf{I} + \sum_{i=1}^K \mathbf{H}_k{}^\dagger\mathbf{G}_i{}^\dagger(n)\mathbf{G}_i(n)\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{G}_k{}^\dagger(n)$
	End

TABLE II
MULTIPLE SYMBOL ALGORITHM, IMPLEMENTATION 2. SEE APPENDIX A FOR EXPLICIT CALCULATION OF $\mu_k(n+1)$

Step 1 :	Update the transmitters in a parallel fashion
	For $k = 1 : K$
	$\mathbf{X}_k(\bar{\mu}_k) = \left(\bar{\mu}_k\mathbf{I} + \mathbf{H}_k{}^\dagger(\mathbf{T}^{-1}(n) - \sigma^2\mathbf{T}^{-2}(n))\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{T}^{-1}(n)\mathbf{H}_k\mathbf{F}_k(n)$
	$\mu_k(n+1) = \lfloor \bar{\mu}_k \text{ such that } \mathrm{tr}\{\mathbf{X}_k(\bar{\mu}_k)\mathbf{X}_k(\bar{\mu}_k)^\dagger\} = p_k \rfloor_+$
	$\mathbf{F}_k(n+1) = \left(\mu_k(n+1)\mathbf{I} + \mathbf{H}_k{}^\dagger(\mathbf{T}^{-1}(n) - \sigma^2\mathbf{T}^{-2}(n))\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{T}^{-1}(n)\mathbf{H}_k\mathbf{F}_k(n)$
	End
Step 2 :	Update T matrix
	$\mathbf{T}(n+1) = \sigma^2\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i\mathbf{F}_i(n)\mathbf{F}_i{}^\dagger(n)\mathbf{H}_i{}^\dagger$

to the transmitter and the receiver of user k and equating it to zero, we arrive at

$$\mathbf{G}_k = \mathbf{F}_k^\dagger\mathbf{H}_k^\dagger\left(\sigma^2\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i\mathbf{F}_i\mathbf{F}_i{}^\dagger\mathbf{H}_i{}^\dagger\right)^{-1} \quad (8)$$

$$\mathbf{F}_k = \left(\mu_k\mathbf{I} + \sum_{i=1}^K \mathbf{H}_k{}^\dagger\mathbf{G}_i{}^\dagger\mathbf{G}_i\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{G}_k{}^\dagger \quad (9)$$

where μ_ks can be found by the power constraint as shown in Appendix A:

$$\mu_k\left(\mathrm{tr}\left\{\mathbf{F}_k{}^\dagger\mathbf{F}_k\right\} - p_k\right) = 0. \quad (10)$$

Note that, as expected, the optimum receivers for a given set of transmitters are in the form of the well-known *MMSE receivers* [31]. Note also that the transmitters are functions of receivers of all users, whereas the receivers are functions of transmitters of all users. To find the joint optimum set of transmitters and receivers, one can devise iterative algorithms that monotonically decrease the total MSE. In particular, alternating minimization where variables are optimized one at a time, keeping all others fixed, proves attractive in the design of such iterative algorithms [38]. Equations (8) and (9) describe the transmitter-receiver updates we can perform. The algorithm starts with a given set of transmitters receivers, and we can update the receivers $\{\mathbf{G}_k\}$ and transmitters $\{\mathbf{F}_k\}$ independently in a parallel fashion using (8) and (9). Note that at each iteration, the Lagrange multiplier μ_k in (9) should be calculated such that the transmit power constraint is satisfied. The algorithm is shown in Table I, where $\lfloor x \rfloor_+ = \max(x, 0)$.

Alternatively, if we assume that each receiver is updated instantaneously when the transmitter is updated, we can reduce the two step iteration given by (8) and (9) to a single iteration. This is accomplished by inserting the resulting receivers of (8) in (9). Let us define

$$\mathbf{T} = \sigma^2\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i\mathbf{F}_i\mathbf{F}_i{}^\dagger\mathbf{H}_i{}^\dagger. \quad (11)$$

Then, following some straightforward algebra, we arrive at the following iteration:

$$\mathbf{F}_k^\star = \left(\mu_k\mathbf{I} + \mathbf{H}_k{}^\dagger(\mathbf{T}^{-1} - \sigma^2\mathbf{T}^{-2})\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{T}^{-1}\mathbf{H}_k\mathbf{F}_k. \quad (12)$$

Note that the iterative algorithm defined by Table II may choose to iterate over each user's transmitter by updating (11) with the newest transmitter found before the next user's iteration for faster convergence. Both the two-step algorithm of (8) and (9) and the algorithm given by (12) decrease the MSE at each step.

The algorithms proposed in Tables I and II can be run online or offline. Online implementation of the algorithms requires the feedback of the necessary information to the users to perform the actual updates, whereas the offline implementation requires the feedback of the resulting optimum transmitter matrix of each user to the user. The algorithm presented in Table II, for example, requires that each user has access to the $\mathbf{T}$ matrix and its channel matrix $\mathbf{H}_k$ to update its transmitter. When the channel is used as time duplex mode, channel state information can be estimated at the transmitter. In such a case, broadcasting the $\mathbf{T}$ matrix after each update will suffice for the implementation of the algorithm.

B. Single Symbol Transmission

In the previous section, we proposed an algorithm to find the optimum transmitter and receiver set using alternating minimization. In this section, we explore a reformulation of the MSE by utilizing the resulting receiver structure (MMSE receivers) so that transmitters and receivers can be optimized jointly. The motivation of this approach is to achieve faster convergence.

First, let us consider a multiuser MIMO system with K users each transmitting a single data stream. In this case, the communication model is the same as the system studied in [30]. The transmitter and receiver matrices become vectors, and we will denote the transmitter and receiver vectors for user i by $\mathbf{f}_i$ and $\mathbf{g}_i$, respectively, where transmitter has a power constraint of p_i. The MMSE receiver for user i is given by

$$\mathbf{g}_i = \mathbf{f}_i{}^\dagger \mathbf{H}_i{}^\dagger \mathbf{T}^{-1} \tag{13}$$

where $\mathbf{T}$ is given by (11), i.e., $\mathbf{T} = \sigma^2\mathbf{I} + \sum_{i=1}^{K} \mathbf{H}_i\mathbf{f}_i\mathbf{f}_i^\dagger\mathbf{H}_i{}^\dagger$. Substituting the above for the receivers, the MSE of user i becomes

$$\text{MSE}_i = 1 - \mathbf{f}_i{}^\dagger\mathbf{H}_i{}^\dagger\mathbf{T}^{-1}\mathbf{H}_i\mathbf{f}_i \tag{14}$$

and the total MSE of all users in the system is

$$\begin{aligned}\text{MSE} &= K - \sum_{i=1}^{K} \mathbf{f}_i{}^\dagger\mathbf{H}_i{}^\dagger\mathbf{T}^{-1}\mathbf{H}_i\mathbf{f}_i \\ &= K - \text{tr}\{\mathbf{I}\} + \sigma^2\text{tr}\{\mathbf{T}^{-1}\}. \end{aligned} \tag{15}$$

Next we note that using the matrix inversion lemma, $\mathbf{T}^{-1}$ can be expressed as

$$\mathbf{T}^{-1} = \mathbf{E}_k{}^{-1} - \frac{\mathbf{E}_k^{-1}\mathbf{H}_k\mathbf{f}_k\mathbf{f}_k{}^\dagger\mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-1}}{1 + \mathbf{f}_k{}^\dagger\mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-1}\mathbf{H}_k\mathbf{f}_k} \tag{16}$$

with $\mathbf{E}_k = \sum_{i\neq k}\mathbf{H}_i\mathbf{f}_i\mathbf{f}_i{}^\dagger\mathbf{H}_i{}^\dagger + \sigma^2\mathbf{I}$. Since $\mathbf{E}_k$ does not depend on $\mathbf{f}_k$, we can easily express the total MSE as

$$\text{MSE} = C_k - \sigma^2\left(\frac{\mathbf{f}_k{}^\dagger\mathbf{H}_k{}^\dagger\mathbf{E}_k^{-2}\mathbf{H}_k\mathbf{f}_k}{1 + \mathbf{f}_k{}^\dagger\mathbf{H}_k{}^\dagger\mathbf{E}_k^{-1}\mathbf{H}_k\mathbf{f}_k}\right) \tag{17}$$

where C_k represents the terms independent of user k. Thus, from the perspective of user k, MSE can be minimized by choosing $\mathbf{f}_k$ to maximize the second term in (17). It is easily shown that we need $\mathbf{f}_k{}^\dagger\mathbf{f}_k = p_k$ to maximize the second term. Thus

$$\text{MSE} = C_k - \sigma^2\left(\frac{\mathbf{f}_k{}^\dagger\mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-2}\mathbf{H}_k\mathbf{f}_k}{\frac{\mathbf{f}_k{}^\dagger\mathbf{f}_k}{p_k} + \mathbf{f}_k{}^\dagger\mathbf{H}_k{}^\dagger\mathbf{E}_k^{-1}\mathbf{H}_k\mathbf{f}_k}\right) \tag{18}$$

$$= C_k - \sigma^2\left(\frac{\mathbf{f}_k{}^\dagger\mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-2}\mathbf{H}_k\mathbf{f}_k}{\mathbf{f}_k{}^\dagger\left(\frac{1}{p_k}\mathbf{I} + \mathbf{H}_k{}^\dagger\mathbf{E}_k^{-1}\mathbf{H}_k\right)\mathbf{f}_k}\right). \tag{19}$$

The maximization of the second term is accomplished simply by choosing $\mathbf{f}_k$ to be the maximum generalized eigenvalued eigenvector of $\mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-2}\mathbf{H}_k$ and $1/p_k\mathbf{I} + \mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-1}\mathbf{H}_k$. An iterative algorithm that minimizes the MSE can be devised as follows: Each user takes turns in optimizing the MSE function from its perspective as explained above, monotonically decreasing the MSE function at each iteration.

For single symbol transmission, the vector form of the transmitters and receivers allow us to optimize each user's transmitter and receiver jointly at each step. Thus, the single symbol algorithm proposed is expected to converge faster than the algorithms proposed in the previous section.

The single symbol algorithm can also be used to devise an algorithm for multisymbol transmission that has faster convergence. Indeed, every multisymbol transmission of each user can be viewed as single-symbol transmissions of each symbol with a linear transmitter of the associated column of linear transmitter matrix. This "conversion" from multisymbol scenario to single-symbol scenario presents us with the opportunity to employ the above transceiver optimization algorithm. However, contrary to the case of the multisymbol algorithm that does not provide specifics regarding the allocation of the total transmit power of a user between its symbols, we now need to constrain the power of each symbol to be able to apply the single symbol algorithm. Hence, we next establish that assuming a certain power distribution, the total transmit power of a user equally distributed between its symbols will not prevent us from achieving the global optimum point.

Recall that when MMSE receivers are used, the MSE of a multisymbol transmission can be reformulated as

$$\begin{aligned}\text{MSE} &= \sum_{i=1}^{K} M_i - \text{tr}\left\{\sum_{j=1}^{K}\mathbf{F}_j{}^\dagger\mathbf{H}_j{}^\dagger\mathbf{T}^{-1}\mathbf{H}_j\mathbf{F}_j\right\} \\ &= \sum_{i=1}^{K} M_i - N_R + \sigma^2\text{tr}\{\mathbf{T}^{-1}\}\end{aligned} \tag{20}$$

where $\mathbf{T} = \sigma^2\mathbf{I} + \sum_{i=1}^{K}\mathbf{H}_i\mathbf{F}_i\mathbf{F}_i{}^\dagger\mathbf{H}_i{}^\dagger$. We observe that any two possible transmitters for user k, $\mathbf{F}_k$, and $\bar{\mathbf{F}}_k$ that satisfy $\mathbf{F}_k\mathbf{F}_k{}^\dagger = \bar{\mathbf{F}}_k\bar{\mathbf{F}}_k{}^\dagger = \mathbf{R}_k$ have the same contribution to the MSE function where $\mathbf{R}_k$ is the $N_{T_k} \times N_{T_k}$ transmitter covariance matrix of user k for $k = 1, \cdots, K$.[1] Notice that there exists an optimum transmitter covariance matrix $\mathbf{R}_k^\star$ that minimizes the MSE in terms of user k, and any optimum transmitter matrix $\mathbf{F}_k^\star$ should satisfy the optimum transmitter covariance matrix constraint $\mathbf{F}_k^\star(\mathbf{F}_k^\star)^\dagger = \mathbf{R}_k^\star$. Our aim is now to impose power constraints on the columns of the optimum transmitter matrix. It is important to note that we cannot pick an arbitrary power distribution at will since a transmitter matrix that satisfies both the given power distribution constraint and the optimum transmitter covariance matrix constraint may not exist. In such a case, we would be confining ourselves to a transmitter search space that is strictly suboptimum. Fortunately, there exists a power distribution that guarantees the existence of at least one transmitter matrix satisfying both the given power distribution constraint and the optimum transmitter covariance matrix constraint for any case. This is the *equal* power distribution, as shown in the following lemma.

Lemma 1: For any $\mathbf{R} = \mathbf{F}\mathbf{F}^\dagger$, there exists a transmitter matrix $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_M]$ such that $\mathbf{f}_i{}^\dagger\mathbf{f}_i = \mathbf{f}_j{}^\dagger\mathbf{f}_j$ for $i, j = 1, 2, \cdots, M$, i.e., each column vector $\mathbf{f}_i$ has equal power.

Proof: Denote r as rank($\mathbf{R}$). λ_i's and $\mathbf{u}_i$'s are the eigenvalues and associated eigenvectors of $\mathbf{R}$. Thus, $\mathbf{R} = \lambda_1\mathbf{u}_1\mathbf{u}_1{}^\dagger + \cdots + \lambda_r\mathbf{u}_r\mathbf{u}_r{}^\dagger$. In addition, in terms of the column vectors of $\mathbf{F}$, $\mathbf{R} = \mathbf{f}_1\mathbf{f}_1{}^\dagger + \cdots + \mathbf{f}_M\mathbf{f}_M{}^\dagger$. Assume that $\mathbf{f}_i = \sum_{k=1}^{r}\sqrt{\lambda_k/M}\mathbf{u}_k e^{j\theta_{ik}}$. Then, a $\{e^{j\theta_{ik}}\}$ set that satisfies $\mathbf{R} = \mathbf{f}_1\mathbf{f}_1{}^\dagger + \cdots + \mathbf{f}_M\mathbf{f}_M{}^\dagger$ remains to be found.

[1]Note that the rank of $\mathbf{R}_k$ cannot exceed the rank of $\mathbf{F}_k$ and, thus, is lower than or equal to M_k.

TABLE III
SINGLE SYMBOL ALGORITHM

Update the transmitters
For $k = 1 : \bar{K}$
$\mathbf{E}_k(n+1) = \sum_{i<k}^{\bar{K}} \mathbf{H}_i\mathbf{f}_i(n+1)\mathbf{f}_i{}^\dagger(n+1)\mathbf{H}_i{}^\dagger + \sum_{i>k}^{\bar{K}} \mathbf{H}_i\mathbf{f}_i(n)\mathbf{f}_i{}^\dagger(n)\mathbf{H}_i{}^\dagger + \sigma^2\mathbf{I}$
$\mathbf{f}_k(n+1) = \sqrt{\bar{p}_k}$ maximum generalized eigenvalued eigenvector of $\mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-2}(n+1)\mathbf{H}_k$ and $1/\bar{p}_k\mathbf{I} + \mathbf{H}_k{}^\dagger\mathbf{E}_k{}^{-1}(n+1)\mathbf{H}_k$
End

Assume that $\mathbf{D}$ is a diagonal matrix with $\sqrt{\lambda_k/M}$ value at the k^{th} entry $\mathbf{U} = [\mathbf{u}_1\mathbf{u}_2\cdots\mathbf{u}_r]$, and Θ is the phase matrix with $e^{j\theta_{ik}}$ at the (i,k)th entry. In terms of $\mathbf{F}$, $\mathbf{R}$ is

$$\mathbf{R} = \mathbf{F}\mathbf{F}^\dagger \tag{21}$$

$$= \mathbf{U}\mathbf{D}\Theta\Theta^\dagger\mathbf{D}\mathbf{U}^\dagger \tag{22}$$

and the eigendecomposition of $\mathbf{R}$ is

$$\mathbf{R} = \mathbf{U}\Delta\mathbf{U}^\dagger. \tag{23}$$

Hence, Θ should satisfy

$$\mathbf{D}\Theta\Theta^\dagger\mathbf{D} = \Delta \tag{24}$$

$$\Theta\Theta^\dagger = \mathbf{D}^{-1}\Delta\mathbf{D}^{-1} \tag{25}$$

$$= M\mathbf{I}. \tag{26}$$

The above shows that Θ should have orthogonal rows. Such a Θ matrix can be formed by assigning $e^{j((2\pi ik)/M)}$ to the (i,k)th entry with the assumption $M \geq r$. □

Lemma 1 shows that an equal power distribution for any given covariance matrix (including the optimum covariance matrix) exists. The proof is by construction and essentially tells us that assuming the symbols as virtual users with equal powers will not *preclude* the single symbol algorithm from achieving the global optimum performance of the system in terms of system-wide total MSE. That is to say that the single symbol algorithm searches for transmitter matrices in a reduced search space that includes the global minimizer of the system-wide MSE.

It follows from the above that any multiple symbol transmission scenario can be viewed as a single symbol transmission scenario, where M_k symbols of user k represent M_k virtual users each with power constraint $\bar{p}_k = p_k/M_k$ and channel matrix $\mathbf{H}_k$. The total number of virtual users is $\bar{K} = \sum_{k=1}^{K} M_k$ and an iterative algorithm that minimizes the MSE can be devised as follows: Each virtual user takes turns in optimizing the MSE function from its perspective as explained for the single symbol transmission case, monotonically decreasing the MSE function at each iteration. The algorithm is shown in Table III. The algorithm can be run both online and offline. For online implementation of the algorithm, each transmitter requires the $\mathbf{E}_k$ matrix and the channel matrix of its own $\mathbf{H}_k$. This algorithm may be preferable for implementation due to its potential for faster convergence than the multiple symbol algorithm.

IV. CONVERGENCE ISSUES

In this section, we investigate the convergence properties of the algorithms we proposed. The algorithms iterate over the users each time decreasing the system-wide MSE. Clearly, the MSE function is bounded below. This implies that the algorithms, which produce decreasing sequences that are lower bounded, are convergent. Unfortunately, although the MSE function is convex over each of the transmitter and receiver matrices, it is not jointly convex on all the variables. Therefore, although each step of each of the the algorithms we propose finds the minimum of the MSE function over the variable over which we optimize, the fixed point of the algorithm is not guaranteed to converge to the global minimum due to the possible multimodality of the MSE function.

At the fixed point of our multi symbol algorithm given in Tables I and II, the set of transmitters remains unchanged when the iteration is performed over all the users:

$$\left(\mu_k\mathbf{I} + \mathbf{H}_k{}^\dagger(\mathbf{T}^{-1} - \sigma^2\mathbf{T}^{-2})\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{T}^{-1}\mathbf{H}_k\mathbf{F}_k = \mathbf{F}_k. \tag{27}$$

Simple linear algebra reveals that at the fixed point, each transmitter needs to satisfy

$$\mathbf{H}_k{}^\dagger\mathbf{T}^{-2}\mathbf{H}_k\mathbf{F}_k = \frac{\mu_k}{\sigma^2}\mathbf{F}_k. \tag{28}$$

Thus, at the fixed point, the columns of transmitter matrix of each user are eigenvectors of the matrix $\mathbf{H}_k{}^\dagger\mathbf{T}^{-2}\mathbf{H}_k$ with the same eigenvalue of μ_k/σ^2.

An observation that can readily be made from (28) is that the set of transmitter matrices at the fixed point is not unique. For example, if $\mathbf{F}_k = [\mathbf{f}_{k1}\mathbf{f}_{k2}\cdots\mathbf{f}_{ki}]$ is the transmitter matrix of user k transmitting i symbols at the fixed point, then phase addition to each column and/or permutation of the column vectors $\mathbf{F}_k^\star = [\mathbf{f}_{k2}e^{j\theta_1}\mathbf{f}_{k1}e^{j\theta_2}\cdots\mathbf{f}_{ki}e^{j\theta_i}]$ will also satisfy the fixed point equation.

It is also easily seen that if all columns of the initial transmitter matrix of any user are in the null space of the channel matrix of that user, then the algorithm will yield the undesirable fixed point of the all zero transmitter matrix. Such undesirable starting points should be avoided while implementing this algorithm, for example, by choosing random starting points.

In general, the nonconvexity of the problem may prevent the convergence of the algorithm to the global optimum. As is common with nonconvex problems, our hope then lies in finding the optimum by choosing multiple random starting points and adopting to the best of the fixed points of multiple runs. A moment's thought reveals, however, that in certain cases, we may be able to construct a mechanism to check if the fixed point we derived is indeed the global optimum.

To see this, let us recall that the MSE function is

$$\begin{aligned}\text{MSE} &= \sum_{i=1}^{K} M_i - \text{tr}\left\{\sum_{j=1}^{K} \mathbf{F}_j^{\dagger}\mathbf{H}_j^{\dagger}\mathbf{T}^{-1}\mathbf{H}_j\mathbf{F}_j\right\} \\ &= \sum_{i=1}^{K} M_i - N_R + \sigma^2 \text{tr}\{\mathbf{T}^{-1}\}\end{aligned} \tag{29}$$

where we have used MMSE receivers for all users. When we substitute $\mathbf{R}_k = \mathbf{F}_k\mathbf{F}_k^{\dagger}$, the total MSE minimization problem can be restated as

$$\min \quad \text{tr}\{\mathbf{T}^{-1}\} \tag{30}$$

$$\text{s.t.} \quad \mathbf{T} \leq \sigma^2\mathbf{I} + \sum_{i=1}^{K} \mathbf{H}_i\mathbf{R}_i\mathbf{H}_i^{\dagger} \tag{31}$$

$$\text{tr}\{\mathbf{R}_i\} \leq p_i; \quad \mathbf{R}_i \geq 0 \quad i = 1, \cdots, K \tag{32}$$

$$\text{rank}(\mathbf{R}_i) \leq \min(N_{T_i}, M_i) \quad i = 1, \cdots, K \tag{33}$$

where $\mathbf{A} \geq 0$ refers to the positive semidefiniteness constraint on $\mathbf{A}$. We will refer to (30)–(33) as the *original transmit covariance optimization problem* in the sequel.

It can be shown that the MSE function is jointly convex in $\{\mathbf{R}_k\}$. In addition, the constraints defined by (31) and (32) are convex constraints. Thus, short of the rank constraint defined by (33), which is a consequence of the fact that each $\mathbf{F}_k$ is $N_{T_k} \times M_k$, we have a convex optimization problem at hand. If we are able to discard this constraint of the original transmit covariance optimization problem, then the convex optimization problem defined by (30)–(32) may prove useful in our quest to test the optimality of the algorithm described in Section III. We will call the resulting problem the *relaxed transmit covariance optimization problem*. Since this optimization problem is convex and Slater's condition is satisfied, the Karush–Kuhn–Tucker (KKT) conditions associated with the relaxed transmit covariance optimization problem are necessary and sufficient to test the optimality of the $\{\mathbf{R}_k\}$ [39]. Constructing the Lagrangian dual problem, one can come up with the following KKT conditions for the relaxed transmit covariance optimization problem as derived in Appendix A:

$$\lambda_k\mathbf{I} = \mathbf{H}_k^{\dagger}\mathbf{T}^{-2}\mathbf{H}_k + \Psi_k \tag{34}$$

$$\text{tr}\{\mathbf{R}_k\} = p_k \tag{35}$$

$$\text{tr}\{\Psi_k\mathbf{R}_k\} = 0 \tag{36}$$

$$\Psi_k, \mathbf{R}_k, \lambda_k \geq 0 \tag{37}$$

where $\{\Psi_k, \lambda_k\}$ are the dual variables.

Notice that anytime we *relax* an optimization problem by relaxing the constraints, the optimum point of the relaxed problem is "better" (has a smaller MSE value) than or as good as the original problem. Hence, if the solution of original problem achieves the cost value for the relaxed problem, then the solution found must optimize both the relaxed and the original problem. If the solution we find for the original nonconvex problem proves to be optimal for the relaxed convex problem, then it must be the optimal for the original problem as well. Because KKT conditions are necessary and sufficient for the relaxed problem, we can check the optimality of the point with respect to the relaxed problem. If the point is optimal, then it is so for both the relaxed and the original problem. Thus, if for given transmitters, which lead to a set of $\{\mathbf{R}_k\}$, the above conditions are satisfied, i.e., the corresponding dual variables are found, then we are at the global minimizer of the MSE function for both the original transmit covariance optimization problem and the relaxed transmit covariance optimization problem. Observe that $\{\mathbf{R}_k\}$ can be readily constructed from the transmitters at the fixed point of our algorithm.

It is important to note that the above serves as an exact optimality check when the rank constraint is redundant for the problem. Observe that the rank constraint for the $N_{T_k} \times N_{T_k}$ matrix $\mathbf{R}_k$ is necessary only for the case when $M_k < N_{T_k}$. Thus, for the cases where all users have $M_k \geq N_{T_k}$, the MSE minimization problem formulated as a function of the transmitters with transmit power constraints, the *relaxed transmit covariance optimization problem* defined by (30)–(32) in terms of $\{\mathbf{R}_k\}$ are equivalent, and the optimality check above is exact. For systems, where $M_k < N_{T_k}$ for at least one user, the KKT conditions strictly correspond to the relaxed transmit covariance optimization problem where the rank constraint is relaxed. In this case, the KKT conditions, which are necessary and sufficient for the relaxed transmit covariance optimization, are sufficient but not necessary for the original transmit covariance optimization problem. Specifically, we may be at the optimum point of our rank constrained problem described by (30)–(33) and not at the optimum point of the *relaxed transmit covariance optimization problem* described by (30)–(32). This may lead to cases where we misjudge the fixed point of our algorithm as a local minimizer.

V. Rate Allocation

So far, we considered transceiver optimization algorithms for a given number of symbols to be transmitted by each user. In this section, we ask the question of how many symbols each user should end up transmitting with the corresponding optimum transmitters and receivers. Note that in this case, the MSE optimization problem consists of (30) to (32), i.e., the rank constraint is no longer present. The following observation is immediate.

Observation 1: User k cannot transmit more than rank$(\mathbf{R}_k^{\star})$ symbols without causing self interference among its symbols where $\mathbf{R}_k^{\star}$ is the optimum transmitter covariance matrix of the relaxed transmit covariance optimization problem.

It is evident that for no self interference among symbols, the transmitter matrix of user k, $\mathbf{F}_k$ should have orthogonal columns, and the number of orthogonal columns rank$(\mathbf{F}_k)$ is equal to rank$(\mathbf{F}_k\mathbf{F}_k^{\dagger})$. For a given power constraint p_k for user k, the maximum rank that the optimum $\mathbf{F}_k\mathbf{F}_k^{\dagger}$ can achieve is rank$(\mathbf{R}_k^{\star})$, where $\mathbf{R}_k^{\star}$ is the optimum transmitter covariance matrix of user k without any rank constraint. Thus, an optimum transmitter $\mathbf{F}_k$ of size $N_T \times$ rank$(\mathbf{R}_k^{\star})$ can be formed, which will result in noninterfering symbols for user k. Note that increasing M_k beyond rank$(\mathbf{R}_k^{\star})$ causes the rank constraint to be redundant and does not change the optimum transmit covariance matrix of the rank-constrained problem.

Observation 2: Increasing the number of symbols of user k beyond rank$(\mathbf{R}_k^{\star})$ increases the system-wide MSE by 1 for each additional symbol.

Noting that the resulting $\mathbf{R}_k^{\star}$ for user k is the same for any $M_k \geq$ rank$(\mathbf{R}_k^{\star})$ and that the transceiver structure of other users remains unchanged, we conclude that increasing the number of symbols of user k only increases MSE_k by $M_k - \text{rank}(\mathbf{R}_k{}^\star)$ as if $M_k - \text{rank}(\mathbf{R}_k{}^\star)$ symbols are not transmitted at all. Obviously, this results in degradation of performance for user k and, hence, the overall system, and is undesirable.

From the two observations, it follows that the the number of symbols transmitted can be chosen to maximize the number of symbols transmitted while avoiding self interference for each user, and this judicious choice of number of transmitted symbols can enhance the overall performance of the system. It is evident that number of symbols transmitted by user k should be $\text{rank}(\mathbf{R}_k{}^\star)$ for the best system performance, in the minimum MSE sense, while maximizing the symbol rate. Finding the optimum $\{\mathbf{R}_k\}$ set in this case aims to find the optimum transmitter, receiver, and the number of symbols sent by each user while maximizing the number of symbols transmitted without causing self interference for each user and minimizing system-wide MSE. Formally, the optimization problem becomes

$$\min_{\{\mathbf{R}_i\}_{i=1,\cdots,K}} \quad \text{MSE} \tag{38}$$

$$\text{s.t.} \quad \text{tr}(\mathbf{R}_i) \le p_i \quad \mathbf{R}_i \ge \mathbf{0} \quad i = 1, \cdots, K. \tag{39}$$

Note that there is no constraint other than the transmit power constraint and positive semidefiniteness of the transmitter covariance matrices $\{\mathbf{R}_k\}$. Specifically, the rank constraints of the transmitter covariance matrices defined by (33) for the MSE minimization with predefined data rates (number of symbols) are no longer present. Rather, the number of symbols to be sent by each user is determined by the rank of the resulting transmit covariance matrices. Notice that (38) and (39) is the same problem as the *relaxed transmit covariance optimization problem*, and the KKT conditions of this problem are those given by (34)–(37) and are necessary and sufficient to check for the optimality of the resulting transceivers.

The multiple symbol algorithms (Tables I and II) or the single symbol algorithm (Table III) with the assumption that each user transmits N_T symbols can be used to find $\{\mathbf{R}_k{}^\star\}$. Using these results, the $\{M_k = \text{rank}(\mathbf{R}_k^\star)\}$ set is determined. Note that the algorithms are run offline to find the transmitters and the corresponding $\{\mathbf{R}_k{}^\star\}$ set. Then, the number of symbols to be transmitted by each user is determined by $\{M_k = \text{rank}(\mathbf{R}_k{}^\star)\}$, and the corresponding optimum transmitter matrices are found by simple factorization. Finally, the resulting transmitter matrices and the number of symbols to be transmitted by each user are fed back to each user.

A final observation before concluding this section is that a similar problem is the information theoretic sum capacity maximization problem for multiuser MIMO systems, which also aims to find the optimum covariance matrix of the transmitter, $\mathbf{R}_k$ [14]:

$$\max \quad \frac{1}{2}\log\left|\sum_{i=1}^{K}\mathbf{H}_i\mathbf{R}_i\mathbf{H}_i{}^\dagger + \sigma^2\mathbf{I}\right| - \frac{1}{2}\sigma^{2N_R} \tag{40}$$

$$\text{s.t.} \quad \text{tr}\{\mathbf{R}_i\} \le p_i; \quad \mathbf{R}_i \ge 0 \qquad i = 1, \cdots, K. \tag{41}$$

Since sum capacity function is jointly convex in $\{\mathbf{R}_k\}$ and the constraints are convex constraints, the information-theoretic sum capacity maximization problem is a convex optimization problem, and the KKT conditions are necessary and sufficient to test the optimality of the $\{\mathbf{R}_k\}$. Although the optimization problem given by (40) and (41) is similar to (30)–(32), the KKT conditions are not identical to (34)–(37) [14]; specifically, compare (34) with (42):

$$\beta_k\mathbf{I} = \mathbf{H}_k{}^\dagger\mathbf{T}^{-1}\mathbf{H}_k + \Theta_k \tag{42}$$

$$\text{tr}\{\mathbf{R}_k\} = p_k \tag{43}$$

$$\text{tr}\{\Theta_k\mathbf{R}_k\} = 0 \tag{44}$$

$$\Theta_k, \mathbf{R}_k, \beta_k \ge 0. \tag{45}$$

Note that $\{\Theta_k, \beta_k\}$ are the dual variables. Observe that the optimal $\{\mathbf{R}_k\}$ set for sum capacity maximization and MSE minimization problems need not to be the same set.

VI. Numerical Results

In this section, we offer numerical results to support our analysis. The simulations are performed for a multiuser MIMO system where each user is equipped with $N_{T_i} = N_T = 4$ transmit antennas, and the receiver has $N_R = 4$ antennas. Each user performs binary modulation and has power constraint $tr\{\mathbf{F}_k\mathbf{F}_k^\dagger\} \le p_k = 1$. The channel realizations used in the simulations are presented in Table IV. The variance of the AWGN noise used in the simulations is 0.8 per complex dimension. Each plot shows the MSE or BER value versus the algorithm iteration index, where an iteration signifies updating all users' transmitters, i.e., K updates. Throughout the simulations, the BER for each iteration is found by averaging the BER of all users for a data sample of 10^7 bits. The BER and MSE values resulting from the simulations are presented in Table V.

First, we consider a $K = 4$ user MIMO system with each user sending a single data stream. Both the algorithm described by (12), dubbed Algorithm 1, and the algorithm described in Section III-B, dubbed Algorithm 2, are simulated. Fig. 2 shows the evolution of both algorithms as they converge to the optimum MSE value. Although both algorithms converge to the optimum value, we observe that the convergence of Algorithm 2 that chooses the maximum generalized eigenvalued eigenvector of $\mathbf{H}_k{}^\dagger\mathbf{E}_k^{-2}\mathbf{H}_k$ and $1/\bar{p}_k\mathbf{I} + \mathbf{H}_k{}^\dagger\mathbf{E}_k^{-1}\mathbf{H}_k$ is faster. This is expected since Algorithm 2 assumes instantaneous receiver updates with each user's transmitter update and jointly optimizes the transmitter and the receiver for each user, whereas Algorithm 1 performs transmitter and receiver optimization for each user in tandem. Fig. 3 shows the evolution of the average BER with transceiver updates. Notice that the average BER is reduced by 29-fold (14.6 dB).

For single-user MIMO systems, precoding with the maximum eigenvalued eigenvector of $\mathbf{H}_k^\dagger\mathbf{H}_k$ is the optimum transmission technique that achieves the minimum BER and the minimum MSE in AWGN. For the system considered in Figs. 2 and 3, we have compared the performance of the case where each user transmits by precoding its symbol by its largest eigenvalued eigenvector with that of the case where each user uses the transceivers defined by the fixed point of our algorithms. Our algorithms resulted in a BER of 0.0020 and an MSE of 0.7192, whereas the eigen-beamforming resulted in a BER of 0.0077 and an MSE of 1.0240. The example demonstrates that the system performance of all users should be jointly optimized and that

TABLE IV
SIMULATED MULTIUSER MIMO SYSTEM MODEL

User	MIMO Channel Model
1	$\mathbf{H}_1 = \begin{bmatrix} -0.2517+0.7565i & 0.8983+0.5839i & 0.9444-0.2283i & -0.1402-0.6712i \\ 1.7573-0.6483i & -1.6711-0.1807i & 0.6351-0.7641i & 0.5614+0.8252i \\ 0.1695-0.3972i & 0.1172-0.1628i & 0.4290+0.0827i & 0.8265-0.1750i \\ -0.7809-0.3602i & -0.1686-0.2386i & -0.1607-0.5321i & -0.9994+0.6477i \end{bmatrix}$
2	$\mathbf{H}_2 = \begin{bmatrix} -0.2244-0.2328i & -0.5066-0.6200i & -0.2329+0.7821i & -0.1352-0.4038i \\ 0.3422+1.0303i & -0.1190+0.1382i & 1.3527+0.2843i & 1.3120+0.3817i \\ 0.3664+0.6793i & -0.8429-1.1067i & -0.5403-1.4129i & -0.4315-0.5916i \\ 0.4498+1.4641i & 0.8073+0.8599i & 1.2230-0.0189i & 1.8329-0.8169i \end{bmatrix}$
3	$\mathbf{H}_3 = \begin{bmatrix} -0.7620+0.3841i & -0.4024-0.1000i & -0.4298+1.4355i & 0.9773-1.0329i \\ -0.3328+1.0721i & 0.7152-0.5951i & 1.3372+0.0034i & 0.3813+0.4625i \\ -0.9295+0.4066i & 0.1483-0.4482i & 0.7943-0.4335i & 0.7506+0.0538i \\ -1.7292+0.2461i & -1.0391+0.9538i & 0.1493-0.8461i & 0.2232+0.3720i \end{bmatrix}$
4	$\mathbf{H}_4 = \begin{bmatrix} -0.2549-0.1365i & 0.2346-0.0957i & 0.2916+0.9616i & -0.1317-0.3497i \\ -0.0489-0.6296i & 0.6176-0.8849i & 0.2128-1.2640i & -0.1904+0.5608i \\ 0.3737+0.4703i & -0.7711-1.8840i & 0.1935+0.5052i & -0.8018+0.2415i \\ 0.4004+0.4881i & -0.4661+0.6581i & -0.2530+0.4960i & -0.2055+0.6228i \end{bmatrix}$
5	$\mathbf{H}_5 = \begin{bmatrix} -0.3629-0.6262i & 1.3243+0.0745i & -1.2232-0.7561i & -1.3743+1.0631i \\ -0.7433-0.4415i & -0.2152+0.2147i & -0.2924+0.6256i & 0.3653+0.4381i \\ -0.7378+0.1222i & -0.4349+0.0302i & -1.2263-0.2358i & 0.1867+0.1017i \\ 0.4321+0.7943i & -0.0841-0.2674i & -0.8899+0.2673i & 1.7397-1.7234i \end{bmatrix}$

TABLE V
BER AND MSE RESULTS FOR THE SIMULATED MULTIUSER MIMO SYSTEM

Example	Starting MSE	MSE at the fixed point	Starting BER	BER of Alg.1	BER of Alg.2
$K=4, M_i=1, p_i \leq 1$	1.8438	0.7192	0.0577	0.002	0.002
$K=5, M_i=1, p_i \leq 1$	2.7366	1.5473	0.1097	0.0248	0.0354
$K=2, M_i=2, p_i \leq 2$	2.3375	1.0267	0.1902	0.0082	0.0081
$K=3, M_i=2, p_i \leq 2$	3.1856	2.590	0.1502	0.1065	0.0991

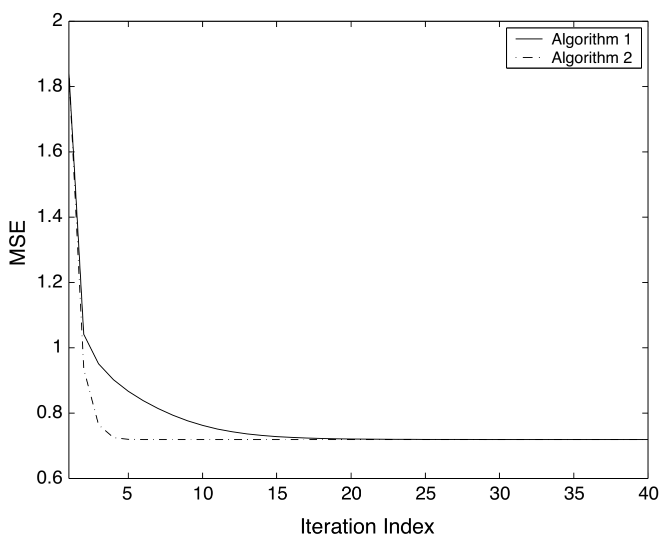

Fig. 2. $K = 4$ user MIMO system with $M_i = 1$ data streams per user. $N_{Ti} = N_R = 4$.

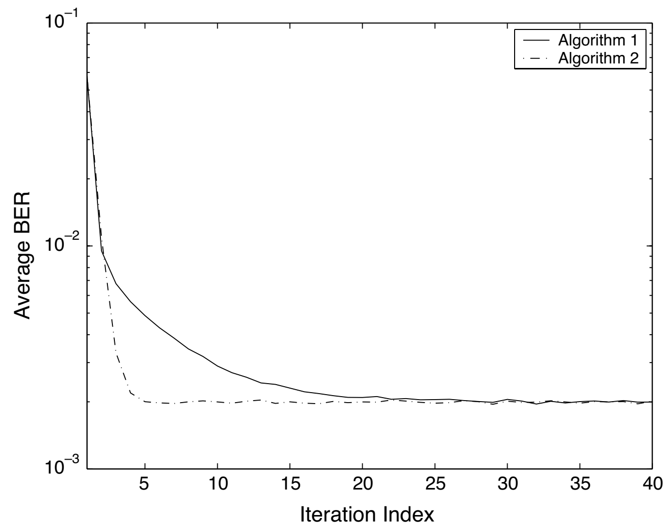

Fig. 3. $K = 4$ user MIMO system with $M_i = 1$ data streams per user. $N_{Ti} = N_R = 4$. BER analysis at each iteration.

the solutions of single-user MIMO systems are not directly applicable to multiuser MIMO systems.

Fig. 4 shows the MSE evolution when a fifth user becomes active in the previous system. The minimum achievable MSE of the system increases as expected due to higher number of users and channel constraints. When the system is overloaded, the number of symbols transmitted is larger than the number of $\min(N_T, N_R)$, and the MSE minimization algorithm does not constrain the distribution of the total MSE to the symbols. Thus, there are many possibilities that achieve the same total MSE but have different BER. Although the resulting transceivers are not necessarily the global optimum in terms of BER, the simulations

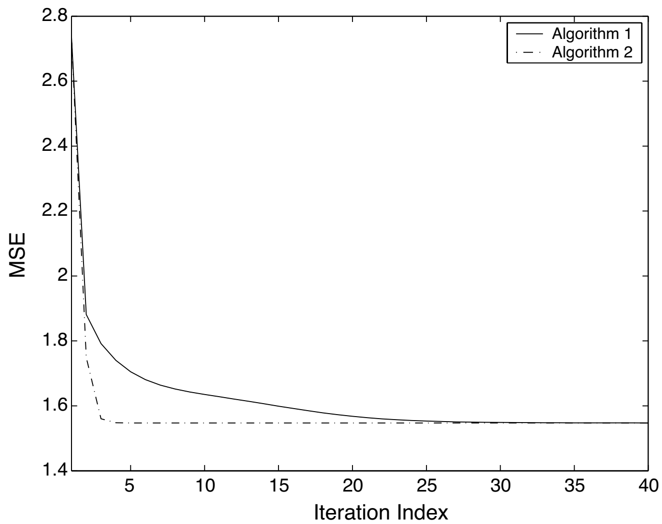

Fig. 4. $K = 5$ user MIMO system with $M_i = 1$ data streams per user. $N_{Ti} = N_R = 4$.

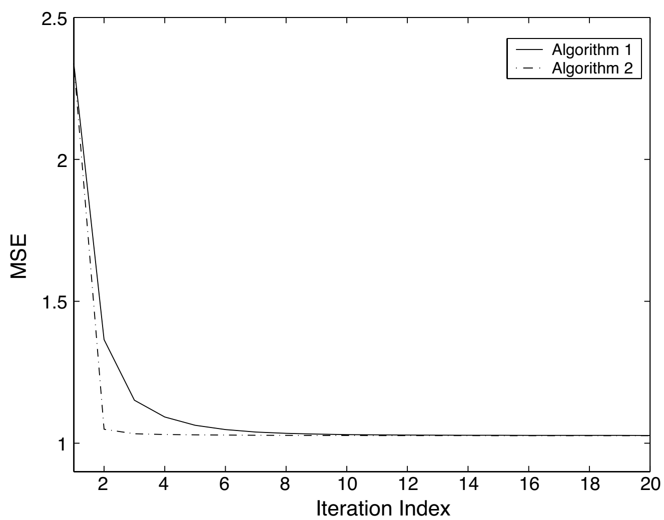

Fig. 5. $K = 2$ user MIMO system with $M_i = 2$ data streams per user. $N_{Ti} = 4, N_R = 4$.

show that the BER values of the resulting transceivers are lower than the BER of the starting transceivers.

The performance comparison of Algorithms 1 and 2 for the case of multiple symbol transmission is shown by Fig. 5 for $K = 2$ users, each equipped with $N_T = 4$ transmitter antennas and transmitting $M_1 = M_2 = 2$ data streams. Total MSE monotonically decreases and converges to its minimum value for both cases. The convergence of Algorithm 2 is much faster than Algorithm 1 for the multiple symbol transmission scenario as well. The evolution of the average BER with the transceiver updates is shown in Fig. 6. The effect of additional users in the system can be observed in Fig. 7. While the performance of the system degrades by the increased number of users, the relative convergence speed of the algorithms remains the same.

Fig. 8 shows the evolution of the Algorithm 1 for a $K = 3$ user system, whereas Fig. 9 shows the evolution of Algorithm 2 for a $K = 2$ user system. Each user transmits 2 data streams. The convergence of the algorithms is observed for five random starting points for both cases. Total MSE monotonically decreases and converges to its minimum value for each starting

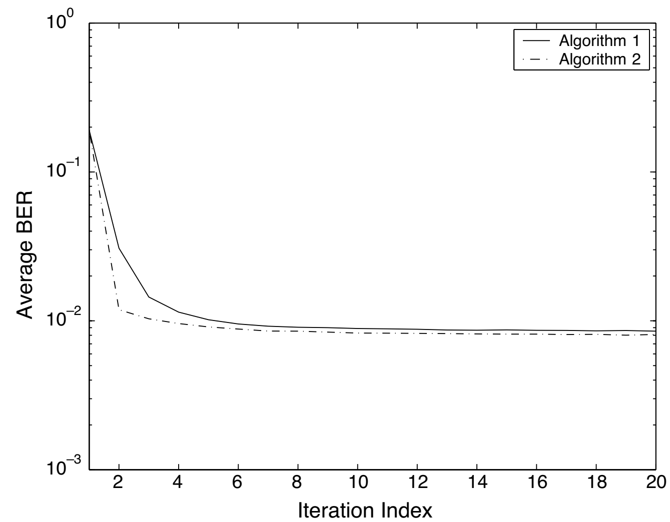

Fig. 6. $K = 2$ user MIMO system with $M_i = 2$ data streams per user. $N_{Ti} = 4, N_R = 4$ BER analysis at each iteration.

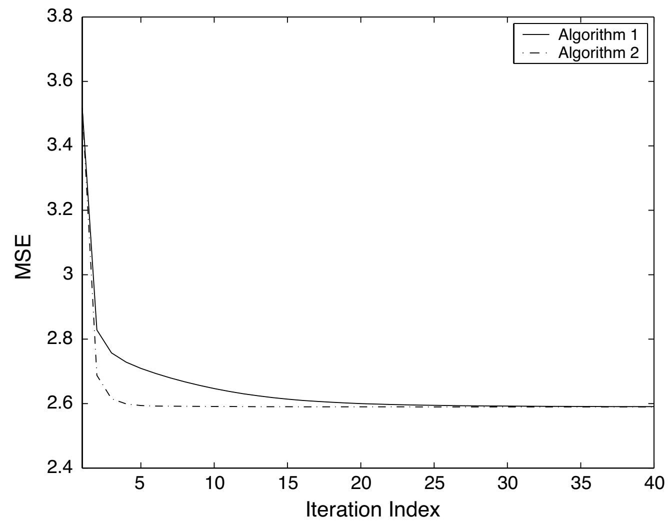

Fig. 7. $K = 3$ user MIMO system with $M_i = 2$ data streams per user. $N_{Ti} = 4, N_R = 4$.

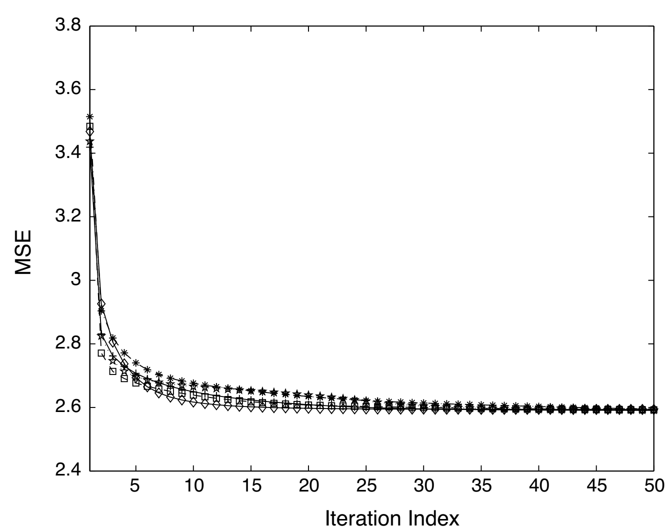

Fig. 8. $K = 3$ user MIMO system with $M_i = 2$ data streams per user. $N_{Ti} = N_R = 4$ performance of Algorithm 1 with 5 different starting transmitter sets.

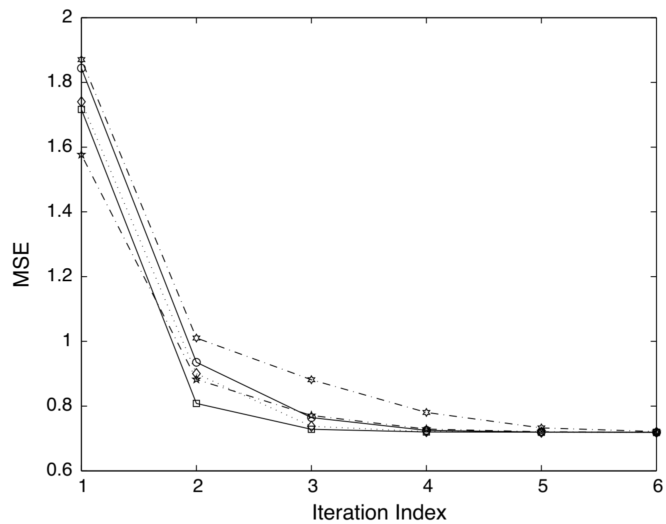

Fig. 9. $K = 2$ user MIMO system with $M_i = 2$ data streams per user. $N_{Ti} = N_R = 4$ performance of Algorithm 2 with 5 different starting transmitter sets.

TABLE VI
PROCESSING TIMES OF THE ITERATIVE ALGORITHMS FOR THE SIMULATED 4 USER 4 × 4 MIMO SYSTEM

Algorithm	Processing Time
SeDuMi based Iterative Algorithm	33 sec.
Algorithm 2 with 1 iteration/turn	1 sec
Algorithm 2 with 2 iteration/turn	1.8 sec
Algorithm 2 with 4 iteration/turn	3.1 sec

transmitter set. Although the resulting transmitter sets are different, the total MSE of the system and $\mathbf{R}_k$ of each user at the fixed point are the same for each starting point. When transmitter matrices are started randomly, we observed they consistently converged to the same total MSE and $\{\mathbf{R}_k\}$ set for a given set of channel matrices $\{\mathbf{H}_k\}$.

In Section IV, a relaxed optimization problem is formed to obtain a global optimality check that is valid for certain cases. In such a formulation, the transmit covariance matrices of the users are jointly convex. Thus, an iterative algorithm can be devised such that each user optimizes its transmit covariance matrix one at a time, and the convergence of such an algorithm to the global minimum of the relaxed problem is guaranteed. Several approaches can be developed to obtain the optimum transmit covariance matrix for a single user. The first one is simply applying Algorithms 1 and 2 to each user where each user transmits as many symbols as its transmit antennas, as mentioned in Section V. For such a case, the rank constraint on the transmit covariance matrix becomes redundant, and the global optimum transmit covariance matrix for the original optimization problem and the relaxed optimization problem become the same. Another approach to solve such a convex optimization problem is semidefinite programming. Such a formulation is done in [40] by using the semidefinite programming software SeDuMi for obtaining optimum transmit covariance matrix of the user when the zero-forcing receiver is employed in a single-user system. We can utilize this software to find the optimum transmit covariance matrices for the multiuser system only when the rank constraint is redundant. In this case, due to the joint convexity of the transmit covariance matrices, an iterative algorithm that runs the SeDuMi software to obtain the optimum transmit covariance matrix of each user at each step can be formed and is guaranteed to converge to the global optimum. For comparison purposes, the SeDuMi-based iterative algorithm is devised as follows: In each iteration, we rewrite the MSE function in terms of the parameters of the associated user and form a semidefinite optimization problem and solve for the optimum transmit covariance matrix of the associated user by using the SeDuMi optimization software. We simulated both Algorithm 2 and the SeDuMi-based iterative algorithm for the same multiuser MIMO system. The simulations showed that both algorithms converge to the same set of transmit covariance matrices, and Table VI summarizes the comparative simulation times of the algorithms. As expected, the processing time for the updates of the SeDuMi-based algorithm is much higher than Algorithm 2.

VII. CONCLUSION

In this paper, we proposed transmitter (and receiver) update algorithms that are geared toward enhancing the system performance by minimizing the total MSE in a multiuser MIMO system. Motivated by the equivalence of the multisymbol multiuser system and the single symbol transmission with multiple virtual users, we proposed an alternative iterative algorithm that is observed to have faster convergence. All the algorithms proposed in this paper can be applied to multiple and single symbol transmission. We investigated the fixed-point properties and identified an optimality check mechanism for the algorithms proposed for the system. We investigated the relationship between the number of symbols each user transmits and the performance of the system. Specifically, we observed that by judicious choice of the number of symbols to be transmitted by each user, the performance of the system can be enhanced by avoiding self interference while maximizing the data rate for each user.

It is important to note that the algorithms we propose here rely on the updates through error-free, low-delay feedback channels. To that end, the effect of the accuracy of the feedback on the performance of the algorithms remain to be investigated.

APPENDIX A
CALCULATION OF LAGRANGE MULTIPLIERS

The algorithms proposed in Section III require the calculation of the Lagrange multipliers (μ_k's) at each step according to the power constraint of each user. The proposed updates for each implementation are as follows:

$$\mathbf{F}_k^\star = \left(\mu_k\mathbf{I}+\sum_{i=1}^{K}\mathbf{H}_k{}^\dagger\mathbf{G}_i{}^\dagger\mathbf{G}_i\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{G}_k{}^\dagger \tag{46}$$

$$\mathbf{F}_k^\star = \left(\mu_k\mathbf{I}+\mathbf{H}_k{}^\dagger(\mathbf{T}^{-1}-\sigma^2\mathbf{T}^{-2})\mathbf{H}_k\right)^{-1}\mathbf{H}_k{}^\dagger\mathbf{T}^{-1}\mathbf{H}_k\mathbf{F}_k \tag{47}$$

where μ_k should satisfy the following condition:

$$\mu_k\left(\operatorname{tr}\left\{(\mathbf{F}_k^\star)^\dagger\,\mathbf{F}_k^\star\right\}-p_k\right)=0, \qquad \mu_k\geq 0. \tag{48}$$

Notice that each update is in the form of

$$\mathbf{F}_k^\star = (\mu_k\mathbf{I}+\mathbf{A}_k)^{-1}\mathbf{B}_k \tag{49}$$

where $\mathbf{A}_k$ and $\mathbf{B}_k$ are known. $\mathbf{A}_k$ can be decomposed as $\mathbf{A}_k = \mathbf{U}_k\mathbf{D}_k\mathbf{U}_k^\dagger$ by singular value decomposition. The value of μ_k satisfying the KKT condition is either the positive value such that the linear transmitter satisfies the power constraint with equality or 0. The value of μ_k satisfying the constraint with equality can be found as

$$\begin{aligned}\mathrm{tr}\left\{\mathbf{F}_k^\star(\mathbf{F}_k^\star)^\dagger\right\} &= p_k &(50)\\ &= \mathrm{tr}\left\{(\mu_k\mathbf{I}+\mathbf{A}_k)^{-1}\times\mathbf{B}_k\mathbf{B}_k^\dagger(\mu_k\mathbf{I}+\mathbf{A}_k)^{-1}\right\} &(51)\\ &= \mathrm{tr}\left\{\left(\mu_k\mathbf{I}+\mathbf{U}_k\mathbf{D}_k\mathbf{U}_k^\dagger\right)^{-1}\times\mathbf{B}_k\mathbf{B}_k^\dagger\left(\mu_k\mathbf{I}+\mathbf{U}_k\mathbf{D}_k\mathbf{U}_k^\dagger\right)^{-1}\right\} &(52)\\ &= \mathrm{tr}\left\{\mathbf{U}_k(\mu_k\mathbf{I}+\mathbf{D}_k)^{-1}\times\mathbf{U}_k^\dagger\mathbf{B}_k\mathbf{B}_k^\dagger\mathbf{U}_k(\mu_k\mathbf{I}+\mathbf{D}_k)^{-1}\mathbf{U}_k^\dagger\right\} &(53)\\ &= \mathrm{tr}\left\{(\mu_k\mathbf{I}+\mathbf{D}_k)^{-2}\mathbf{U}_k^\dagger\mathbf{B}_k\mathbf{B}_k^\dagger\mathbf{U}_k\right\}. &(54)\end{aligned}$$

Defining a new matrix $\mathbf{C}_k = \mathbf{U}_k^\dagger\mathbf{B}_k\mathbf{B}_k^\dagger\mathbf{U}_k$ that has a c_{ij} entry in the ith row and jth column and d_{ij} in the same manner, the equation can be expressed as follows:

$$\sum_{i=1}^{N_{T_k}}\frac{c_{ii}}{(\mu_k+d_{ii})^2} = p_k. \tag{55}$$

Let us define the function $z(\mu_k)=\sum_{i=1}^{N_{T_k}}(c_{ii}/(\mu_k+d_{ii})^2)$. The left-hand side of (55) is a function of the Lagrange multiplier μ_k, and we want to find μ_k, where the function $z(\mu_k)$ gets the value of p_k. Recall that the parameters of the function $z(\mu_k)$, $\{c_{ii}\}$'s, and $\{d_{ii}\}$'s are the diagonal entries of $\mathbf{C}_k$ and eigenvalues of $\mathbf{A}_k$, respectively, and are non-negative due to the positive semidefiniteness of the matrices $\mathbf{C}_k$ and $\mathbf{A}_k$ for both update mechanisms. The derivative of the function $z(\mu_k)$ with respect to the Lagrange multiplier μ_k is

$$\frac{\partial z(\mu_k)}{\partial\mu_k} = \sum_{i=1}^{N_{T_k}}\frac{-2c_{ii}}{(\mu_k+d_{ii})^3} \tag{56}$$

and it is strictly negative due to the non-negativity of c_{ii} and d_{ii} for $\mu_k \geq 0$. The function $z(\mu_k)$ is a monotonically decreasing function of non-negative μ_k, and there exists only one non-negative real value of μ_k that satisfies the (55). Thus, the value of the Lagrange multiplier is the only non-negative real solution of (55) if it exists; otherwise, it is 0.

Appendix B
Derivation of the KKT Conditions for the Relaxed Transmitter Covariance Optimization Problem

The relaxed transmitter covariance optimization problem is

$$\begin{aligned}\min\quad & \mathrm{tr}\{\mathbf{T}^{-1}\} &(57)\\ \text{s.t.}\quad & \mathbf{T} \leq \sigma^2\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i\mathbf{R}_i\mathbf{H}_i{}^\dagger &(58)\\ & \mathrm{tr}\{\mathbf{R}_i\} \leq p_i;\quad \mathbf{R}_i \geq 0\quad i=1,\cdots,K. &(59)\end{aligned}$$

Let Γ, $\{\lambda_i\}$, and $\{\Psi_i\}$ denote the dual variables associated with the constraint on $\mathbf{T}$, the power constraints, and the positive semi-definiteness constraints of the covariance matrices $\{\mathbf{R}_i\}$. The Lagrangian of the optimization problem is

$$\begin{aligned}&L(\{\mathbf{R}_i\},\mathbf{T},\Gamma,\{\lambda_i\},\{\Psi_i\})\\ &= \mathrm{tr}\{\mathbf{T}^{-1}\} + \mathrm{tr}\left\{\Gamma\left(\mathbf{T}-\sigma^2\mathbf{I}-\sum_{i=1}^K\mathbf{H}_i\mathbf{R}_i\mathbf{H}_i{}^\dagger\right)\right\}\\ &\quad + \sum_{i=1}^K\lambda_i(\mathrm{tr}(\mathbf{R}_i)-p_i) - \sum_{i=1}^K\mathrm{tr}\{\Psi_i\mathbf{R}_i\}\\ &= \mathrm{tr}\{\mathbf{T}^{-1}\} + \mathrm{tr}\{\Gamma\mathbf{T}\} - \sigma^2\mathrm{tr}\{\Gamma\} - \sum_{i=1}^K\lambda_i p_i\\ &\quad + \sum_{i=1}^K\mathrm{tr}\left\{\left(\lambda_i\mathbf{I}-\mathbf{H}_i^\dagger\Gamma\mathbf{H}_i-\Psi_i\right)\mathbf{R}_i\right\}.\end{aligned} \tag{60}$$

The objective of the dual program is

$$d(\Gamma,\{\lambda_i\},\{\Psi_i\}) = \min_{\{\mathbf{R}_i\},\mathbf{T}} L(\{\mathbf{R}_i\},\mathbf{T},\Gamma,\{\lambda_i\},\{\Psi_i\}). \tag{61}$$

Taking the gradients and equating to 0, $\mathbf{R}_i$ and $\mathbf{T}$ should satisfy

$$\begin{aligned}\frac{\partial}{\partial\mathbf{T}}\left(\mathrm{tr}\{\mathbf{T}^{-1}\}+\mathrm{tr}\{\Gamma\mathbf{T}\}\right) &= 0 &(62)\\ -\mathbf{T}^{-2}+\Gamma &= 0 &(63)\\ \Gamma &= \mathbf{T}^{-2} &(64)\end{aligned}$$

and

$$\frac{\partial}{\partial\mathbf{R}_i}\left(\mathrm{tr}\left\{\left(\lambda_i\mathbf{I}-\mathbf{H}_i^\dagger\Gamma\mathbf{H}_i-\Psi_i\right)\mathbf{R}_i\right\}\right) = 0 \tag{65}$$

resulting in

$$\Psi_i + \mathbf{H}_i^\dagger\Gamma\mathbf{H}_i = \lambda_i\mathbf{I}\quad i=1,2,\cdots,\quad K. \tag{66}$$

Thus, the dual optimization problem is

$$\begin{aligned}\max\quad & d(\Gamma,\{\lambda_i\},\{\Psi_i\}) = 2\mathrm{tr}\left\{\Gamma^{\frac{1}{2}}\right\} - \sigma^2\mathrm{tr}\{\Gamma\} - \sum_{i=1}^K\lambda_i p_i &(67)\\ \text{subject to}\quad & \lambda_i\mathbf{I} \geq \mathbf{H}_i^\dagger\Gamma\mathbf{H}\quad i=1,2,\cdots,K &(68)\\ & \Gamma \geq 0. &(69)\end{aligned}$$

Since the primal optimization problem is a convex problem, the dual problem achieves its maximum where primal achieves its minimum. Slater's conditions are satisfied, and thus, the KKT conditions are necessary and sufficient for the optimality of the relaxed transmitter covariance problem. Using the complementary slackness condition, we arrive at the following KKT conditions:

$$\begin{aligned}\lambda_k\mathbf{I} &= \mathbf{H}_k{}^\dagger\mathbf{T}^{-2}\mathbf{H}_k + \Psi_k &(70)\\ \mathrm{tr}\{\mathbf{R}_k\} &= p_k &(71)\\ \mathrm{tr}\{\Psi_k\mathbf{R}_k\} &= 0 &(72)\\ \Psi_k,\mathbf{R}_k,\lambda_k &\geq 0. &(73)\end{aligned}$$

The global optimum transmitter covariance matrix should satisfy the KKT conditions derived. Thus, there exists a positive semidefinite matrix Ψ_k and non-negative λ_k for each $\mathbf{R}_k$ if it is global optimum. It is clear that $\Psi_k\mathbf{R}_k = 0$ due to (72) and the positive semidefiniteness of both Ψ_k and $\mathbf{R}_k$. Thus, multiplying both sides of (70) by $\mathbf{R}_k$ results in

$$\lambda_k\mathbf{R}_k = \mathbf{H}_k{}^\dagger\mathbf{T}^{-2}\mathbf{H}_k\mathbf{R}_k. \tag{74}$$

The terms on both sides of the equation are known except for the non-negative slack variable λ_k. If $\mathbf{R}_k$ is the global optimum, such a non-negative λ_k exists, and it can be found by using (74). When λ_k is known, Ψ_k can be calculated by (70) and can be checked for positive semidefiniteness and whether it satisfies the trace condition (72).

REFERENCES

[1] E. Telatar, Capacity of multi-antenna Gaussian channels, in AT&T Bell Labs Memo., Murray Hill, NJ, 1995.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[3] V. Tarokh, N. Seshadri, and R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[4] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[5] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[6] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[7] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.

[8] D. Gore, R. Nabar, and A. Paulraj, "Selection and use of optimal transmit antennas for low rank matrix channel," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 2000, pp. 2785–2788.

[9] A. F. Molisch, M. Z. Win, and J. H. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. Int. Conf. Commun.*, 2001.

[10] R. W. Heath, Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142–144, Apr. 2001.

[11] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 322–324, Aug. 2002.

[12] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, pp. 3221–3231, Dec. 1994.

[13] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May 2002.

[14] W. Rhee, W. Yu, and J. M. Cioffi, "The optimality of beam-forming in uplink multiuser wireless systems," *IEEE Trans. Wireless Commun.*, to be published.

[15] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple access channels," *IEEE Trans. Inform. Theory*, submitted for publication.

[16] Z. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for multiple access through ISI channels," in *Proc. 39th Annu. Allerton Conf. Commun., Contr. Comput.*, Oct. 2001.

[17] M. Rupf and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1261–1266, July 1994.

[18] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1984–1991, Sept. 1999.

[19] S. Ulukus and R. D. Yates, "Iterative construction of optimum signature sequence sets in synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1989–1998, July 2001.

[20] S. Ulukus and A. Yener, "Iterative transmitter and receiver optimization for synchronous CDMA systems," in *Proc. IEEE Int. Symp. Inform. Theory*, June 2002, p. 48.

[21] C. Rose, S. Ulukus, and R. D. Yates, "Wireless systems and interference avoidance," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 415–428, July 2002.

[22] J. I. Concha and S. Ulukus, "Optimization of CDMA signature sequences in multipath channels," in *Proc. IEEE Veh. Technol. Conf.*, 2001, pp. 1978–1982.

[23] W. M. Jang, B. R. Vojcic, and R. L. Pickholtz, "Joint transmitter-receiver optimization in synchronous multiuser communications over multipath channels," *IEEE Trans. Commun.*, vol. 46, pp. 269–278, Feb. 1998.

[24] G. S. Rajappan and M. L. Honig, "Signature sequence adaptation for DS-CDMA with multipath," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 384–395, Feb 2002.

[25] J. Winters, J. Salz, and R. D. Gitlin, "Impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1740–1751, Jan./Feb./Mar. 1994.

[26] A. F. Naguib, A. J. Paulraj, and T. Kailath, "Capacity improvement with base-station antenna arrays in cellular CDMA," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 691–698, Aug. 1994.

[27] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, pp. 247–256, 1998.

[28] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.

[29] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE Veh. Technol. Conf.*, May 1999, pp. 851–856.

[30] J.-H. Chang, F. Rashid-Farrokhi, and L. Tassiulas, "Joint transmitter receiver diversity for efficient space division multiple access," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 16–26, Jan. 2002.

[31] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[32] Z. Wang and G. B. Giannakis, "Linearly precoded or coded OFDM against wireless channel fades?," in *Proc. IEEE Third Workshop Signal Process. Adv. Wireless Commun.*, 2001, pp. 267–270.

[33] S. N. Diggavi, "Multiuser DMT: A multiple access modulation scheme," in *Proc. Global Telecommun. Conf.*, 2001, pp. 1566–1570.

[34] Lecture Notes of EE479—Multiuser Digital Transmission Syst., J. M. Cioffi. (2002). *http://www.stanford.edu/class/ee479/reader.html* [Online]

[35] F. Shad, T. D. Todd, V. Kezys, and J. Litva, "Dynamic slot allocation (DSA) in indoor SDMA/TDMA using a smart antenna basestation," *IEEE/ACM Trans. Networking*, vol. 9, pp. 69–81, Feb. 2001.

[36] J. Fuhl and A. F. Molisch, "Capacity enhancement and BER in a combined SDMA/TDMA system," in *Proc. IEEE 46th Veh. Technol. Conf.*, 1996, pp. 1481–1485.

[37] J.Jia Yunjian, S. Hara, and Y. Hara, "Impact of closed-loop power control on SDMA/TDMA system performance," in *Proc. IEEE 56th Veh. Technol. Conf.*, 2002, pp. 1825–1829.

[38] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[39] D. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.

[40] J. Milanovic, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Design of robust redundant precoding filter banks with zero-forcing equalizers for unknown frequency-selective channels," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 2000, pp. 2741–2764.

Semih Serbetli (S'01) received the B.S. degree in electrical and electronics engineering from Boğaziçi University, Istanbul, Turkey, in 2000. He is currently pursuing the Ph.D. degree and is a graduate research assistant with the Department of Electrical Engineering, Pennsylvania State University, University Park.

His current research interests include multiuser detection and transmitter optimization for wireless communication systems.

Aylin Yener (M'00) received the B.S. degrees in electrical and electronics engineering and in physics, from Boğaziçi University, Istanbul, Turkey, in 1991 and the M.S. and Ph.D. degrees in electrical and computer engineering from Rutgers University, New Brunswick, NJ, in 1994 and 2000, respectively.

During her Ph.D. studies, she was with Wireless Information Network Laboratory (WINLAB), Department of Electrical and Computer Engineering, Rutgers University. Between fall 2000 and fall 2001, she was with the Electrical Engineering and Computer Science Department, Lehigh University, Bethlehem, PA, where she was a P.C. Rossin assistant professor. Currently, she is an assistant professor with the Electrical Engineering Department, Pennsylvania State University, University Park. Her research interests include performance enhancement of multiuser systems, wireless communication theory and wireless networking in general.

Dr. Yener received the NSF CAREER award in 2003. She is an associate editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.