Exploiting Opportunistic Multiuser Detection in Decentralized Multiuser MIMO Systems

Rui Zhang and John M. Cioffi

Abstract

This paper studies the design of a decentralized multiuser multi-antenna (MIMO) system for spectrum sharing over a fixed narrow band, where the coexisting users independently update their transmit covariance matrices for individual transmit-rate maximization via an iterative manner. This design problem was usually investigated in the literature by assuming that each user treats the co-channel interference from all the other users as additional (colored) noise at the receiver, i.e., the conventional single-user decoder (SUD) is applied. This paper proposes a new decoding method for the decentralized multiuser MIMO system, whereby each user opportunistically cancels the co-channel interference from some or all of the other users via applying multiuser detection techniques, thus termed opportunistic multiuser detection (OMD). This paper studies the optimal transmit covariance design for users’ iterative maximization of individual transmit rates with the proposed OMD, and demonstrates the resulting capacity gains in decentralized multiuser MIMO systems against the conventional SUD.

Index Terms

Cognitive radio, decentralized multiuser system, MIMO Gaussian interference channel, multiuser detection.

I. INTRODUCTION

The Gaussian interference channel is a basic mathematical model that characterizes many real-life communication systems with multiple uncoordinated users sharing a common spectrum to transmit independent information at the same time, such as the digital subscriber line (DSL) network [1], the ad-hoc wireless network [2], and the newly emerging cognitive radio (CR) wireless network [3]. From an information-theoretical perspective, the capacity region of the Gaussian interference channel, which constitutes all the simultaneously achievable rates of the users in the system, is still unknown in general [4], while significant progresses have recently been made on approaching this limit [5], [6]. Capacity-approaching techniques usually require certain cooperations among distributed users for their encoding and decoding. A more pragmatic approach that leads to suboptimal achievable rates of the users in the

R. Zhang is with the Institute for Infocomm Research, A*STAR, Singapore (e-mail:rzhang@i2r.a-star.edu.sg).
J. M. Cioffi is with the Department of Electrical Engineering, Stanford University, Stanford, USA (e-mail:cioffi@stanford.edu).
Gaussian interference channel is to restrict the system to operate in a decentralized manner [7], i.e., allowing only single-user encoding and decoding by treating the co-channel interference from the other users as additional Gaussian noise at each user’s receiver. In such a context, decentralized algorithms for users to allocate their transmit resources such as the power, bit-rate, bandwidth, and antenna beam to optimize individual transmission performance and yet to ensure certain fairness among all the users, become most important.

This paper focuses on a multiuser multiple-input multiple-output (MU-MIMO) wireless system, where multiple distributed links, each equipped with multiple transmit and/or receive antennas, share a common narrow band for transmission in a fully decentralized manner. In such a scenario, the system design reduces to finding a set of transmit covariance matrices for the users subject to their co-channel interference resulting from their simultaneous and uncoordinated transmissions. This design problem has been investigated in a vast number of prior works in the literature, e.g., [8]-[16], by treating the co-channel interference as additional colored noise at each user’s receiver, i.e., the conventional single-user decoder (SUD) for the classic point-to-point MIMO channel is applied. In [8], the authors proposed an algorithm, which is in spirit analogous to the iterative water-filling (IWF) algorithm in [7], for each distributed MIMO link to iteratively update transmit covariance matrix to maximize individual transmit rate. Distributed iterative beamforming (the rank of transmit covariance matrix is restricted to be one) algorithms were also studied in [9] for transmit sum-power minimization given individual user’s quality of service (QoS) constraint in terms of the received signal-to-interference-plus-noise ratio (SINR). The throughput of decentralized MU-MIMO systems has been further analyzed in [10] and [11] for the cases of fading channels and large-size systems, respectively. In [12], [13], centralized strategies were proposed where all users’ transmit covariance matrices are jointly searched to maximize their sum-rate, and numerical algorithms were also proposed to converge to a local sum-rate maxima. Analyzing the decentralized MU-MIMO system via a game theoretical approach has recently been done in [14]-[16].

The cited papers on decentralized/centralized designs for the Gaussian MIMO interference channel have all adopted the SUD at each user’s receiver, whereas during the past decade multiuser detection techniques (see, e.g., [17] and references therein) have been thoroughly investigated in the literature, and have been proven in realistic multiuser/MIMO systems to be able to provide substantial performance gains over
the conventional SUD. This motivates our work’s investigation of the following question: Considering a decentralized MU-MIMO system where the users iteratively adapt their transmit covariance matrices for individual rate maximization, “Is applying multiuser detection at each user’s receiver able to enhance the system throughput over the conventional SUD?” Note that because of the randomness of channels among the users, as well as their independent rate assignments, at one particular user’s receiver, multiuser detection can be used to cancel the co-channel interference from some/all of its coexisting users only when their received signals are jointly decodable with this particular user’s own received signal. Thus, we refer to this decoding method as opportunistic multiuser detection (OMD). Also note that the OMD in the context of the decentralized MU-MIMO system is analogous to the “successive group decoder (SGD)” in the fading multiple-access channel (MAC) with unknown channel state information (CSI) at the user transmitters (see, e.g., [18] and references therein). With the proposed OMD, this paper derives the optimal transmit covariance matrix for user’s individual transmit-rate maximization at each iteration of transmit adaptation. By simulation, this paper demonstrates the throughput gains of the converged users’ transmit covariance matrices with the proposed OMD over the conventional SUD.

The rest of this paper is organized as follows. Section II presents the system model of the decentralized MU-MIMO system. Section III studies the optimal design of user transmit covariance matrix with the proposed OMD for the special case with two users in the system. Section IV generalizes the results to the case of more than two users. Section V provides the simulation results to demonstrate the throughput gains with the proposed OMD over the SUD. Finally, Section VI concludes the paper.

Notation: Scalars are denoted by lower-case letters, e.g., $x$, and bold-face lower-case letters are used for vectors, e.g., $\mathbf{x}$, and bold-face upper-case letters for matrices, e.g., $\mathbf{X}$. In addition, $\text{tr}(S)$, $|S|$, $S^{-1}$, and $S^{\frac{1}{2}}$ denote the trace, determinant, inverse, and square-root of a square matrix $S$, respectively, and $S \succeq 0$ means that $S$ is a positive semi-definite matrix [19]. For an arbitrary-size matrix $M$, $M^H$ denotes the conjugate transpose of $M$. $\text{diag}(x_1, \ldots, x_M)$ denotes a $M \times M$ diagonal matrix with $x_1, \ldots, x_M$ as its diagonal elements. $\mathbf{I}$ and $\mathbf{0}$ denote the identity matrix and the all-zero vector, respectively. $\mathbb{E}[\cdot]$ denotes the statistical expectation. The distribution of a circular symmetric complex Gaussian (CSCG) random vector with mean $\mathbf{x}$ and covariance matrix $\Sigma$ is denoted by $\mathcal{CN}(\mathbf{x}, \Sigma)$, and $\sim$ stands for “distributed as”. $\mathbb{C}^{x \times y}$ denotes the space of $x \times y$ matrices with complex-valued elements. $\max(x, y)$ and $\min(x, y)$ denote
the maximum and minimum between two real numbers, $x$ and $y$, respectively, and $(x)^+ = \max(x, 0)$.

II. SYSTEM MODEL

This paper considers a distributed MU-MIMO system where $K$ users transmit independent information to their corresponding receivers simultaneously over a common narrow band. Each user is equipped with multiple transmit and/or receiver antennas, while for user $k$, $k = 1, \ldots, K$, $N_k$ and $M_k$ denote the number of its transmit and receive antennas, respectively. For the time being, it is assumed that perfect time and frequency synchronization with reference to a common clock system have been established for all the users in the system prior to data transmission. We also assume a block-fading model for all the channels involved in the system, and a block-based transmission for all the users over each particular channel fading state. Since the proposed study applies to any channel fading state, for brevity we drop the index of fading state here. The discrete-time baseband signal for the $k$th user transmission is given by

$$y_k = H_{kk}x_k + \sum_{j=1,j\neq k}^{K} H_{jk}x_j + z_k$$

(1)

where $x_k \in \mathbb{C}^{N_k \times 1}$ and $y_k \in \mathbb{C}^{M_k \times 1}$ are the transmitted and received signal vectors for user $k$, respectively, $k \in \{1, \ldots, K\}$; $H_{kk} \in \mathbb{C}^{M_k \times N_k}$ denotes the direct-link channel matrix for user $k$, while $H_{jk} \in \mathbb{C}^{M_k \times N_j}$ denotes the cross-link channel matrix from user $j$ to user $k$, $j \in \{1, \ldots, K\}$, $j \neq k$; and $z_k \in \mathbb{C}^{M_k \times 1}$ is the received noise vector of user $k$.

Without loss of generality, it is assumed that $z_k \sim \mathcal{CN}(0, I)$, $\forall k \in \{1, \ldots, K\}$, and all $z_k$’s are independent. We consider a decentralized multiuser system where the $K$ users independently encode their transmitted messages and thus $x_k$’s are independent over $k$. Since this paper is interested in the information-theoretic limit of each Gaussian MIMO channel involved, it is assumed that $x_k \sim \mathcal{CN}(0, S_k)$, $\forall k \in \{1, \ldots, K\}$, where $S_k = \mathbb{E}[x_kx_k^H]$ is the transmit covariance matrix for user $k$.

This paper considers a similar decentralized operation protocol as in [7], [8], [14]-[16], whereby the users in the system take turns to update their transmit covariance matrices for individual rate maximization, with all the other users’ transmit covariance matrices being fixed, until all users’ transmit covariance matrices and their transmit rates get converged. We consider two types of decoding methods at each user’s receiver. One is the conventional SUD, which has been applied in the above cited papers, where the $k$th user decodes its desired message by treating the co-channel interference from all the other users,
\( j \neq k \), as additional colored Gaussian noise \( \sim CN(0, \sum_{j=1, j \neq k}^{K} H_{jk} S_j H_{jk}^H) \). The other decoding method is the newly proposed OMD, whereby each user opportunistically applies multiuser detection to decode some/all of its coexisting users’ messages so as to cancel their resulted interference, provided that these messages are jointly decodable with this user’s own message. In practice, each user in the system is usually interfered with by all the other users, while due to location-dependent shadowing/fading, only a small group of coexisting users who are closest to one particular user and thus correspond to the strongest cross-link channels to this user, will contribute the most to this user’s received co-channel interference. As a result, this user can effectively estimate the transmit rates as well as the cross-link channels of these “strong” interference users, and employ the proposed OMD to suppress their interference at the receiver. Note that the use of OMD instead of SUD still maintains the fully decentralized property of the existing IWF-like operation protocols given in [7], [8], [14]-[16].

III. TRANSMIT COVARIANCE OPTIMIZATION: THE TWO-USER CASE

In this section, we present the problem formulation as well as the solution to determine the optimal transmit covariance matrix of each user for individual transmit-rate maximization, when the proposed OMD is employed. For the purpose of exposition, we consider the special case where only two users exist in the system. We will address the general case with more than two users in Section IV. For brevity, only user 1’s transmit adaptation is addressed here, while the developed results apply similarly to user 2.

A. Problem Formulation

Note that at one particular iteration of user 1 to update its transmission, user 2’s transmit covariance matrix, \( S_2 \), and transmit rate, denoted by \( r_2 \), are both fixed values. For a given transmit covariance matrix of user 1, \( S_1 \), the resultant maximum transmit rate of user 1 can be expressed as

\[
\log |I + H_{11} S_1 H_{11}^H| \\
\log |I + H_{11} S_1 H_{11}^H + H_{21} S_2 H_{21}^H| - r_2 \\
\log |I + (I + H_{21} S_2 H_{21}^H)^{-1} H_{11} S_1 H_{11}^H| - r_2
\]

where \( r_1(S_1) \) is the resultant maximum transmit rate of user 1 can be expressed as

\[
\begin{align*}
    r_1(S_1) &= \begin{cases} 
    \log |I + H_{11} S_1 H_{11}^H| & r_2 \leq R_2^{(a)} \\
    \log |I + H_{11} S_1 H_{11}^H + H_{21} S_2 H_{21}^H| - r_2 & R_2^{(a)} < r_2 \leq R_2^{(b)} \\
    \log |I + (I + H_{21} S_2 H_{21}^H)^{-1} H_{11} S_1 H_{11}^H| & r_2 > R_2^{(b)}
    \end{cases}
\end{align*}
\]  \hspace{1cm} (2)

where

\[
R_2^{(a)} = \log |I + (I + H_{11} S_1 H_{11}^H)^{-1} H_{21} S_2 H_{21}^H| \hspace{1cm} (3)
\]

\[
R_2^{(b)} = \log |I + H_{21} S_2 H_{21}^H| \hspace{1cm} (4)
\]
The above result is illustrated in the following three cases corresponding to the three expressions of \( r_1 \) in (2) from top to bottom.

- **Strong Interference Case:** In this case, the received signal from user 2 is decodable at user 1’s receiver with the conventional SUD, by treating user 1’s signal as colored Gaussian noise. This is feasible since \( r_2 \leq R_2^{(a)} \) given in (3). After decoding user 2’s message and thereby canceling its associated interference, user 1 can decode its own message with a maximum rate equal to its own channel capacity. The above decoding method is known as **successive decoding** (SD) for the standard Gaussian MAC [20].

- **Moderate Interference Case:** In this case, \( r_2 > R_2^{(a)} \) and thus the received signal from user 2 is not directly decodable by the SUD. However, since \( r_2 \leq R_2^{(b)} \) given in (4), it is still feasible for user 1 to apply **joint decoding** (JD) [20] to decode both users’ messages. In this case, the rate pair of the two users should lie on the 45-degree segment of the corresponding MAC capacity region boundary [20], i.e., \( r_1 + r_2 = \log |I + H_{11}S_1H_{11}^H + H_{21}S_2H_{21}^H| \).

- **Weak Interference Case:** In this case, \( r_2 > R_2^{(b)} \), i.e., the received signal from user 2 is not decodable even without the presence of user 1’s signal. As such, user 1’s receiver has the only option of treating user 2’s signal as colored Gaussian noise and applying the conventional SUD to directly decode user 1’s message, the same as that in the existing IWF-like algorithms (see, e.g., [8], [14]-[16]).

In the above decoding method, multiuser detection is applied in both cases of strong and moderate interferences when \( r_2 \leq R_2^{(b)} \), but not in the case of weak interference when \( r_2 > R_2^{(b)} \). Thus, user 1’s receiver opportunistically applies multiuser detection to decode user 2’s message, either successively (SD) or jointly (JD) with its own message. We thus refer to this decoding method as **opportunistic multiuser detection** (OMD). From (3) and (4), it follows that \( R_2^{(a)} \leq R_2^{(b)} \). Further more, it is easy to verify that \( r_1 \) given in (2) with the proposed OMD is in general larger than the achievable rate with the conventional SUD (given by the third expression of \( r_1 \) in (2) independent of \( r_2 \)), for any given set of \( S_1, S_2 \), and \( r_2 \).

With \( r_1(S_1) \) given in (2) for a fixed \( S_1 \), we can further maximize user 1’s transmit rate by searching

\footnote{Note that SD can also be applied in this case to achieve the same rate for user 1 as JD, if SD is deployed jointly with the “time sharing” [20] or “rate splitting” [21] encoding technique at user 1’s transmitter. Since these techniques require certain cooperations between users, they might not be suitable for the fully decentralized multiuser system considered in this paper.}
over $S_1$. Let $P_1$ denote the transmit power constraint of user 1. This problem can be expressed as

\[
(P1) \quad \max_{S_1} \quad r_1(S_1) \\
\text{s.t.} \quad \text{tr}(S_1) \leq P_1, S_1 \succeq 0
\]

where $r_1(S_1)$ is given in (2). The optimal solution of $S_1$ in (P1) and the corresponding maximum transmit rate of user 1 are denoted by $S_1^{\text{OMD}}$ and $r_1^{\text{OMD}}$, respectively.

**B. Proposed Solution**

In this subsection, we study the solution of (P1) for the optimal transmit covariance matrix of user 1, when the proposed OMD is deployed at user 1’ receiver. Note that although the constraints of (P1) are convex, its objective function is not necessarily concave due to the fact that $R_2^{(a)}$ given in (3) is neither convex nor concave function of $S_1$. As a result, (P1) seems to be non-convex at a first glance. In fact, (P1) is a convex optimization problem after being transformed into a convex form, as will be shown in this subsection. In the following, we will study the solution of (P1) for two cases: $r_2 > R_2^{(b)}$ and $r_2 \leq R_2^{(b)}$, for which the SUD and the multiuser decoding (MD) (in the form of either SD or JD) should be used to achieve $r_1(S_1)$ given in (2), respectively.

1) $r_2 > R_2^{(b)}$: In this case, the SUD should be applied. Note that $R_2^{(b)}$ is a constant unrelated to $S_1$. Thus, the optimal $S_1$ that maximizes the third expression of $r_1(S_1)$ in (2) has the following structure [20]:

\[
S_1^{\text{SUD}} = V \Lambda V^H
\]

where $V \in \mathbb{C}^{N_1 \times T_1}$ with $T_1 = \min(N_1, M_1)$ is obtained from the singular-value decomposition (SVD) of the equivalent channel of user 1 (after the noise whitening) expressed as

\[
(I + H_{21} S_2 H_{21}^H)^{-\frac{1}{2}} H_{11} = U \Sigma V^H
\]

with $U \in \mathbb{C}^{M_1 \times T_1}$, $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_{T_1})$, $\sigma_i \geq 0$, $i = 1, \ldots, T_1$, and $\Lambda = \text{diag}(p_1, \ldots, p_{T_1})$ with $p_i$’s obtained from the standard water-filling solution [20]:

\[
p_i = \left( \mu - \frac{1}{\sigma_i^2} \right)^+, \quad i = 1, \ldots, T_1,
\]
with $\mu$ being a constant to make \[ \sum_{i=1}^{T_i} p_i = P_1. \] The maximum rate of user 1 then becomes
\[ r_1^{\text{SUD}} = \sum_{i=1}^{T_i} \log(1 + \sigma_i^2 p_i). \] (8)

2) $r_2 \leq R_2^{(b)}$: In this case, the MD in the form of either SD or JD should be used. In order to overcome the non-concavity of $r_1(S_1)$ given in (2) due to $R_2^{(b)}$, we re-express the first two expressions of $r_1(S_1)$ in (2) as
\[ r_1^{\text{MD}}(S_1) = \min \left( \log |I + H_{11} S_1 H_{11}^H|, \log |I + H_{11} S_1 H_{11}^H + H_{21} S_2 H_{21}^H| - r_2 \right). \] (9)

Thus, the maximum achievable rate of user 1 can be obtained as
\[ r_1^{\text{MD}} = \max_{S_1: \text{tr}(S_1) \leq P, S_1 \succeq 0} r_1^{\text{MD}}(S_1). \] (10)

The maximization problem in (10) can be explicitly written as
\[
\begin{align*}
\text{(P2)} & \quad \max_{r_1, S_1} r_1 \\
\text{s.t.} & \quad r_1 \leq \log |I + H_{11} S_1 H_{11}^H| \quad (11) \\
& \quad r_1 \leq \log |I + H_{11} S_1 H_{11}^H + H_{21} S_2 H_{21}^H| - r_2 \quad (12) \\
& \quad r_1 \geq 0, \text{tr}(S_1) \leq P, S_1 \succeq 0. \quad (13)
\end{align*}
\]

The optimal solution of $r_1$ in (P2) will be $r_1^{\text{MD}}$. Note that (P2) is a convex optimization problem since its constraints specify a convex set of $(r_1, S_1)$. To solve (P2), we apply the standard Lagrange duality method [19]. First, we introduce two non-negative dual variables, $\mu_1$ and $\mu_2$, associated with the two rate constraints (11) and (12), respectively, and write the associated Lagrangian of (P2) as
\[
\begin{align*}
L(r_1, S_1, \mu_1, \mu_2) & = r_1 - \mu_1 \left( r_1 - \log |I + H_{11} S_1 H_{11}^H| \right) \\
& - \mu_2 \left( r_1 - \log |I + H_{11} S_1 H_{11}^H + H_{21} S_2 H_{21}^H| + r_2 \right) \quad (14)
\end{align*}
\]

By reordering the terms in (14), we obtain
\[
\begin{align*}
L(r_1, S_1, \mu_1, \mu_2) & = (1 - \mu_1 - \mu_2)r_1 + \mu_1 \log |I + H_{11} S_1 H_{11}^H| \\
& + \mu_2 \log |I + H_{11} S_1 H_{11}^H + H_{21} S_2 H_{21}^H| + \mu_2 r_2. \quad (15)
\end{align*}
\]

The Lagrange dual function of (P2) is then defined as
\[ g(\mu_1, \mu_2) = \max_{(r_1, S_1) \in A} L(r_1, S_1, \mu_1, \mu_2) \] (16)
where the set $\mathcal{A}$ specifies the remaining constraints of (P2) given in (13). The dual problem of (P2), of which the optimal value is the same as that of (P2), is defined as

$$
\begin{align*}
(P2-D) \quad \min_{\mu_1 \geq 0, \mu_2 \geq 0} & \quad g(\mu_1, \mu_2). 
\end{align*}
$$

Let $r_1^*$ and $S_1^*$ denote the optimal solutions of (P2). Let $\mu_1^*$ and $\mu_2^*$ denote the optimal dual solutions of the dual problem (P2-D). Next, we will present a key relationship between $\mu_1^*$ and $\mu_2^*$ as follows.

**Lemma 3.1:** In problem (P2-D), the optimal solutions satisfy that $\mu_1^* + \mu_2^* = 1$.

**Proof:** See Appendix I.

Given Lemma 3.1, without loss of generality, we can replace $\mu_2$ by $1 - \mu_1$ in (15). Thus, the maximization problem in (16) can be equivalently rewritten as (by discarding the constant term $\mu_2 r_2$)

$$
\begin{align*}
(P3) \quad \max_{S_1} & \quad \mu_1 \log |I + H_{11} S_1 H_{11}^H| + (1 - \mu_1) \log |I + H_{11} S_1 H_{11}^H + H_{21} S_2 H_{21}^H| \\
\text{s.t.} & \quad \text{tr}(S_1) \leq P_1, S_1 \succeq 0.
\end{align*}
$$

Further more, the dual problem (17) now only needs to minimize $g(\mu_1)$ (since $\mu_2 = 1 - \mu_1$) over $0 \leq \mu_1 \leq 1$. Then, there are the following three cases in which $\mu_1^*$ takes different values.

- **$\mu_1^* = 0$:** In this case, $\mu_2^* = 1$. From the Karush-Kuhn-Tucker (KKT) optimality conditions [19] of (P2), it is known that the constraint (11) is inactive while the constraint (12) is active. This suggests that JD instead of SD is optimal. Furthermore, from (P3), with $\mu_1 = \mu_1^* = 0$, it follows that $S_1^*$, denoted by $S_1^{JD}$, maximizes the sum-rate, $\log |I + H_{11} S_1^{JD} H_{11}^H + H_{21} S_2 H_{21}^H|$, from which we can show that

$$
S_1^{JD} = S_1^{SUD} \tag{19}
$$

where $S_1^{SUD}$ is given in (5), i.e., the optimal transmit covariance matrix is the same for both cases of SUD and JD. However, the optimal $r_1^*$ in this case with JD, denoted by $r_1^{JD}$, is equal to

$$
r_1^{JD} = r_1^{SUD} + R_2^{(b)} - r_2 \tag{20}
$$

where $r_1^{SUD}$ is given in (8). Finally, we need to check the condition under which this case holds. Since the constraint (11) should be inactive, it follows that

$$
r_1^{JD} < \log |I + H_{11} S_1^{JD} H_{11}^H| \tag{21}
$$

It can be easily checked that the Slater’s condition holds for (P2) and thus the duality gap for (P2) is zero [19].
From (20) and (21), it can be shown that the case of interest holds when
\[
 r_2 > \log |I + (I + H_{11}S_1^{\text{JD}}H_{11}^H)^{-1}H_{21}S_2H_{21}^H| \triangleq \bar{R}_2^{(a)}.
\] (22)
Note that \( \bar{R}_2^{(a)} \) can also be obtained from \( R_2^{(a)} \) given in (3) by letting \( S_1 = S_1^{\text{JD}} \).

- \( \mu_1^* = 1 \): In this case, \( \mu_2^* = 0 \). From the KKT optimality conditions of (P2), it is known that the constraint (11) is active while the constraint (12) is inactive. This suggests that SD instead of JD is optimal. Furthermore, from (P3), with \( \mu_1 = \mu_1^* = 1 \), it follows that \( S_1^* \), denoted by \( S_1^{\text{SD}} \), maximizes user 1’s own channel capacity (without the presence of user 2), \( \log |I + H_{11}S_1^*H_{11}^H| \), from which we can easily show that [20]
\[
 S_1^{\text{SD}} = V_1\Lambda_1V_1^H
\] (23)
where \( V_1 \in \mathbb{C}^{N_1 \times T_1} \) is obtained from the SVD of the direct-link channel of user 1 expressed as \( H_{11} = U_1\Gamma V_1^H \), with \( U_1 \in \mathbb{C}^{M_1 \times T_1}, \Gamma_1 = \text{diag}(\gamma_1, \ldots, \gamma_{T_1}), \gamma_i \geq 0, i = 1, \ldots, T_1, \) and \( \Lambda_1 = \text{diag}(q_1, \ldots, q_{T_1}) \) with \( q_i \)'s obtained from the standard water-filling solution [20]:
\[
 q_i = \left( \nu - \frac{1}{\gamma_i^2} \right)^+, \quad i = 1, \ldots, T_1,
\] (24)
with \( \nu \) being a constant to make \( \sum_{i=1}^{T_1} q_i = P_1 \). The optimal \( r_1^* \) in this case with SD, denoted by \( r_1^{\text{SD}} \), then becomes
\[
 r_1^{\text{SD}} = \sum_{i=1}^{T_1} \log(1 + \gamma_i^2 q_i). \] (25)
Similarly like the previous case, we can show that this case holds when
\[
 r_2 < \log |I + (I + H_{11}S_1^{\text{SD}}H_{11}^H)^{-1}H_{21}S_2H_{21}^H| \triangleq \hat{R}_2^{(a)}. \] (26)
At last, we have the following lemma.

**Lemma 3.2:** For \( \bar{R}_2^{(a)} \) defined in (22) and \( \hat{R}_2^{(a)} \) defined in (26), it holds that \( \bar{R}_2^{(a)} \geq \hat{R}_2^{(a)} \).

**Proof:** See Appendix III.

- \( 0 < \mu_1^* < 1 \): In this case, \( 0 < \mu_2^* < 1 \), and from the KKT optimality conditions of (P2), it is known that both the constraints (11) and (12) are active. This suggests that \( r_1^* = \log |I + H_{11}S_1^*H_{11}^H| \), i.e., SD is optimal. However, the optimal solution \( S_1^* \) of (P2), or that of (P3) with \( \mu_1 = \mu_1^* \), denoted by \( \tilde{S}_1^{\text{SD}} \), in general does not have any closed-form expression, and thus needs to be obtained by
a numerical search. Since (P3) is convex, the interior-point method [19] can be used to efficiently obtain its solution for a given \( \mu_1 \). Let \( S_1^*(\mu_1) \) denote the optimal solution of (P3) for a given \( \mu_1 \). Then, \( \mu_1^* \) can be efficiently found by a simple bisection search based upon the sub-gradient [19] of \( g(\mu_1) \), which can be shown from (15) (with \( \mu_2 = 1 - \mu_1 \)) to be

\[
\log \left| I + (I + H_{11}S_1^*(\mu_1)H_{11}^H)^{-1}H_{21}S_2H_{21}^H \right| - r_2. \tag{27}
\]

Once \( \mu_1 \) converges to \( \mu_1^* \), the corresponding \( S_1^*(\mu_1) \) becomes the optimal \( \tilde{S}_1^{SD} \). The optimal \( r_1^* \) in this case with SD, denoted by \( \tilde{r}_1^{SD} \), is then expressed as

\[
\tilde{r}_1^{SD} = \log \left| I + H_{11}\tilde{S}_1^{SD}H_{11}^H \right|. \tag{28}
\]

Similarly like the previous two cases and using Lemma 3.2 we can show that this case holds when

\[
\hat{R}_2^{(a)} \leq r_2 \leq \tilde{R}_2^{(a)}. \tag{29}
\]

3) Combining \( r_2 > R_2^{(b)} \) and \( r_2 \leq R_2^{(b)} \): To summarize, the following theorem is obtained for the optimal solution of (P1).

**Theorem 3.1:** For a given set of \( S_2 \) and \( r_2 \) of user 2, the optimal transmit covariance matrix of user 1 and the maximum transmit rate of user 1 with the proposed OMD are given as follows:

\[
S_1^{OMD} = \begin{cases} 
S_1^{SD}, & 0 < r_2 < \hat{R}_2^{(a)} \\
S_1^{SD}, & \hat{R}_2^{(a)} \leq r_2 \leq \tilde{R}_2^{(a)} \\
S_1^{SD}, & \hat{R}_2^{(a)} < r_2 \leq R_2^{(b)} \\
S_1^{SUD}, & r_2 > R_2^{(b)}, 
\end{cases} \tag{30}
\]

\[
r_1^{OMD} = \begin{cases} 
n_1^{SD}, & 0 < r_2 < \hat{R}_2^{(a)} \\
n_1^{SD}, & \hat{R}_2^{(a)} \leq r_2 \leq \tilde{R}_2^{(a)} \\
n_1^{SD}, & \hat{R}_2^{(a)} < r_2 \leq R_2^{(b)} \\
n_1^{SUD}, & r_2 > R_2^{(b)}. 
\end{cases} \tag{31}
\]

The corresponding optimal decoding methods at user 1’s receiver are (from top to bottom) SD, SD, JD, and SUD, respectively.

In Fig. 1 we show \( r_1^{OMD} \) in (31) as a function of \( r_2 \) for some fixed \( S_2 \). The rate gain of \( r_1^{OMD} \) for OMD over \( r_1^{SUD} \) for SUD is clearly shown when \( r_2 < R_2^{(b)} \). There are three pentagon-shape capacity regions shown in the figure, which are \( C_{MAC}(S_1^{JD}, S_2), C_{MAC}(S_1^{SD}, S_2), \) and \( C_{MAC}(S_1^{SD}, S_2) \), respectively, where...
\( C_{MAC}(S_1, S_2) \) denotes the capacity region of a two-user Gaussian MIMO-MAC with user 1’s and user 2’s transmitters transmitting to user 1’s receiver, and \( S_1, S_2 \) denoting the transmit covariance matrices of user 1 and user 2, respectively. More specifically, \( C_{MAC}(S_1, S_2) \) can be expressed as [20]

\[
C_{MAC}(S_1, S_2) \triangleq \left\{ (r_1, r_2) : \sum_{i \in J} r_i \leq \log \left| I + \sum_{i \in J} H_{ni} S_i H_{ni}^H \right|, \forall J \subseteq \{1, 2\} \right\}.
\]  

(32)

Note that in Fig. 1, the solid line consisting of different rate pairs of \((r_1^{OMD}, r_2)\) constitute the boundary rate pairs of the aforementioned capacity regions. Also note that there is a curved part of this rate-pair line in the case of \( \hat{R}_2^{(a)} < r_2 < \bar{R}_2^{(a)} \), where \( r_1^{OMD} \) is equal to \( \tilde{r}_1^{SD} \) and is achievable by \( \tilde{S}_1^{SD} \), which is the solution of problem (P3) for some given \( \mu_1, 0 < \mu_1 < 1 \).

IV. Extension to More Than Two Users

In this section, we extend the results obtained for the two-user MIMO system to the general MU-MIMO system with more than two users, i.e., \( K > 2 \). Due to the symmetry, we consider only user 1’s transmit optimization over \( S_1 \) to maximize transmit rate \( r_1 \), with all the other users’ transmit rates, \( r_2, \ldots, r_K \), and transmit covariance matrices, \( S_2, \ldots, S_K \), being fixed.

To apply OMD at user 1’s receiver, we need to first identify the group of users whose signals are (jointly or successively) decodable at user 1’s receiver without the presence of user 1’s own received signal. We thus have the following definitions:

**Definition 4.1:** A set \( U_1, U_1 \subseteq \{2, \ldots, K\} \), is called a decodable user set for user 1, if the received signals at user 1’s receiver due to the users in \( U_1 \) are decodable without the presence of user 1’s own received signal, by treating the received signals from the other users in \( \overline{U}_1 \) as colored Gaussian noise, where \( \overline{U}_1 \) denotes the complementary set of \( U_1 \), i.e., \( U_1 \cap \overline{U}_1 = \emptyset \) and \( U_1 \cup \overline{U}_1 = \{2, \ldots, K\} \). More specifically, the transmit rates of users in \( U_1 \) must satisfy [20]

\[
\sum_{i \in J} r_i \leq \log \left| I + \sum_{k \in U_1} H_{k1} S_k H_{k1}^H \right|^{-1} \sum_{i \in J} H_{i1} S_i H_{i1}^H, \forall J \subseteq U_1.
\]  

(33)

**Definition 4.2:** A set \( U_1^* \subseteq \{2, \ldots, K\} \) is called an optimal decodable user set for user 1, if \( U_1^* \) is a decodable user set for user 1, and among all possible decodable user sets for user 1, \( U_1^* \) has the largest size.

Next, we have the following important proposition:
Proposition 4.1: The set $U_1^*$ is unique. Furthermore, for any decodable user set for user 1, $U_1$, it holds that $U_1 \subseteq U_1^*$.

Proof: See Appendix III.

For conciseness, we show the algorithm to find the unique set for user 1, $U_1^*$, in Appendix IV.

From Proposition 4.1, it follows that the optimal decoding strategy for user 1’s receiver is applying OMD to the users in the set $U_1^*$ (it may be possible that $U_1^* = \emptyset$), while taking the users in the set $U_1^C$ as additional colored Gaussian noise. For an arbitrary set $\mathcal{V}$, let $|\mathcal{V}|$ denote the size of $\mathcal{V}$. Note that to make the OMD feasible, the rate of user 1, $r_1$, and the rates of users in $U_1^*$ must be jointly in the capacity region of the corresponding $(|U_1^*| + 1)$-user Gaussian MIMO-MAC for a given set of user transmit covariance matrices and the receiver noise covariance matrix, $\Phi = I + \sum_{k \in U_1^C} H_{k1} S_k H_{k1}^H$, which, similar to (32), can be defined as

$$
C_{MAC}(U_1^*) \triangleq \left\{ (r_1, \{r_i\}_{i \in U_1^*}) : \sum_{i \in \mathcal{J}} r_i \leq \log \left| I + \Phi^{-1} \left( H_{11} S_1 H_{11}^H + \sum_{i \in \mathcal{J}} H_{i1} S_i H_{i1}^H \right) \right|, \forall \mathcal{J} \subseteq \{1\} \cup U_1^* \right\}. \quad (34)
$$

Note that in (34), the rate inequalities involving subsets $\mathcal{J}$’s containing users solely from $U_1^*$ all hold due to the definition of $U_1^*$. Therefore, in order to find the optimal $S_1$ for user 1 to maximize $r_1$, with fixed $r_i$’s and $S_i$’s, $i = 2, \ldots, K$, it is sufficient to consider the following optimization problem:

$$
(P4) \quad \max_{S_1, r_1} \quad r_1
$$

$$
\text{s.t.} \quad r_1 + \sum_{i \in \mathcal{J}} r_i \leq \log \left| I + \Phi^{-1} \left( H_{11} S_1 H_{11}^H + \sum_{i \in \mathcal{J}} H_{i1} S_i H_{i1}^H \right) \right|, \forall \mathcal{J} \subseteq U_1^* \quad (35)
$$

$$
r_1 \geq 0, \text{tr}(S_1) \leq P_1, S_1 \succeq 0 \quad (36)
$$

Problem (P4) is convex in terms of $r_1$ and $S_1$ since its constraints specify a convex set of $(r_1, S_1)$. Similarly like for problem (P2), we introduce a set of non-negative dual variables, $\mu_n$’s, $n = 1, \ldots, 2^{|U_1^*|}$, each associated with one corresponding constraint in (35) for a particular subset $\mathcal{J}$ (including $\mathcal{J} = \emptyset$) denoted by $\mathcal{J}_n$, and obtain an equivalent problem for the optimization over $S_1$ for a given set of fixed $\mu_n$’s, which is expressed as

$$
(P5) \quad \max_{S_1} \sum_{n=1}^{2^{|U_1^*|}} \mu_n \log \left| I + \Phi^{-1} \left( H_{11} S_1 H_{11}^H + \sum_{i \in \mathcal{J}_n} H_{i1} S_i H_{i1}^H \right) \right|
$$

$$
\text{s.t.} \quad \text{tr}(S_1) \leq P_1, S_1 \succeq 0. \quad (37)
$$
It can be shown that problem (P5) is convex, and thus it can be solved via standard convex optimization techniques, e.g., the interior point method [19], while in general, no closed-form solution for (P5) is available, similar to the previous two-user case in Section III. Let the optimal solution of (P5) be denoted by $S_1^*(\{\mu_n\})$. Then, $\mu_n$’s can be updated towards the optimal dual solutions of (P4) via the well-known ellipsoid method [19] subject to an additional constraint, $\sum_n \mu_n = 1$ (similar to Lemma 3.1 in the two-user case). Let the optimal solutions of $\mu_n$’s be denoted by $\mu_n^*$’s. The optimal solution of $S_1$ for (P4) with OMD is then obtained as $S_1^{\text{OMD}} = S_1^*(\{\mu_n^*\})$, and the corresponding maximum achievable rate of user 1, $r_1^{\text{OMD}}$, can be obtained from any active constraint in (35) with equality. The optimal decoding orders/decoding methods for the users in $\mathcal{U}_1^*$ prior to decoding user 1’s message can be obtained according to the optimal non-zero dual solutions, $\mu_n^*$’s, or equivalently, the corresponding active constraints in (35) with equality, via applying the property of polymatroid structure of $\mathcal{C}_\text{MAC}(\mathcal{U}_1^*)$ given in (34) [22].

V. SIMULATION RESULTS

In this section, the performance of the proposed OMD is evaluated in comparison with the conventional SUD in a decentralized MU-MIMO system with $K = 2$ users, where the two users adopt an IWF-like algorithm to successively in turn optimize their transmit covariance matrices for individual rate maximization by deploying OMD or SUD at their receivers. For the purpose of exposition, all the channels involved in the system, including user’s direct-link and cross-link channels, are assumed to have independent Rayleigh-fading distributions, i.e., each element of the channel matrix is independent and identically distributed as zero-mean CSCG random variable. Furthermore, each element of the two users’ direct-link channels is assumed to have the variance $\rho_{11}$ and $\rho_{22}$, for user 1 and 2, respectively; and each element of the two cross-link channels has the variance, $\rho_{12}$ for the channel from user 1 to user 2 and $\rho_{21}$ for the channel from user 2 to user 1, respectively. In total, 5000 independent channel realizations are simulated over which each user’s achievable average rate is computed. For each channel realization, the two users iteratively update their transmit covariance matrices until their rates both get converged. It is assumed that $M_k = N_k = 2, k = 1, 2$.

In Fig. 2 the achievable average sum-rate of the two users is shown for a symmetric system and channel setup, where $P_1 = P_2 = 100$, $\rho_{11} = \rho_{22} = 1$, and $\rho_{12} = \rho_{21} = \rho$. The user sum-rate is plotted
against $\rho$ to investigate the effect of the interference between the two users on their achievable sum-rate. It is observed that the sum-rate with the proposed OMD improves over that with the conventional SUD for all the values of $\rho$, while the rate gains become more substantial in the case of large values of $\rho$, i.e., the “strong” interference case. With SUD, it is observed that the sum-rate first decreases with increasing of $\rho$ (as a result of interference whitening), and then starts to increase with $\rho$ (as a result of interference avoidance), and finally gets converged for large values of $\rho$ (due to the fact that zero-forcing (ZF) -based receive beamforming to completely null the co-channel interference becomes optimal at the high signal-to-noise ratio (SNR) region). However, the sum-rate with the proposed OMD is observed to increase consistently with $\rho$, due to the fact that when the co-channel interference becomes stronger at the receiver, the OMD more easily decodes the interference.

Next, we consider a special scenario of the general system model studied in this paper. In this case, a “cognitive radio (CR)” type of newly emerging wireless system is considered, where user 1 is the so-called primary (non-cognitive) user (PU) who is the legitimate user operating in the frequency band of interest, while user 2 is the secondary (cognitive) user (SU) that transmits simultaneously with the PU over the same spectrum under the constraint that its transmission will not cause the PU’s transmission performance to an unacceptable level [23]. The PU is non-cognitive since it is oblivious to the existence of the SU and applies the conventional SUD at the receiver by treating the interference from the SU as additional noise. While for the SU, it is cognitive in the sense that it is aware of the PU and thus transmits with a much lower average power than that of the PU in order to protect the PU; thus, for this example it is assumed that $P_1 = 10P$ and $P_2 = P$, where $P$ is a given constant. In addition, since the SU is cognitive, it may choose to use the more advanced OMD at the receiver to cope with the interference from the PU. Two cases are thus studied for this example: Case (I) both user 1 and user 2 employ SUD; and Case (II) user 1 employs SUD while user 2 employs OMD. It is assumed that the SU’s link distance is much shorter than that of the PU link, and furthermore the SU transmitter and receiver are both in the vicinity of the PU transmitter while they are both sufficiently far away from the PU receiver. Thus, for this example we assume that $\rho_{11} = 1$, $\rho_{22} = 10$, $\rho_{12} = 10$, and $\rho_{21} = 1$.

In Fig. 3 the achievable user individual rates are shown for different values of $P$ in both Cases I and II. It is observed that the achievable rate of user 2 (the SU) improves significantly in Case II over Case
I, thanks to the use of OMD instead of SUD. This rate gain is substantial because the SU receiver is close to the PU transmitter and thus $\rho_{12}$ is large, i.e., the cross-link channel from PU to SU is a “strong” interference channel, for which the OMD is crucial for the SU to mitigate the PU’s interference. However, it is also observed that the achievable rate of user 1 (the PU) drops slightly in Case II as compared with Case I. This is because that in Case II with OMD, the SU’s transmitted signal has a more spatially spread-out spectrum than that in Case I with SUD, and so does the received SU’s interference at the PU receiver. Nevertheless, due to the small value of $\rho_{21}$ or the weak cross-link channel from SU to PU, the capacity loss of the PU is not significant, which justifies the operation principle of the SU, i.e., the PU transmission should be sufficiently protected.

VI. Conclusion

This paper studied a new decoding method, namely opportunistic multiuser detection (OMD), for the decentralized MU-MIMO system where each user iteratively optimizes transmit covariance matrix for individual rate maximization. In comparison with the conventional single-user detection (SUD), the proposed OMD still allows a fully decentralized processing of each user in the system, while it improves the user’s interference mitigation capability at the receiver, and leads to more optimum spatial spectrum sharing among the users. Simulation results showed that substantial system throughput gains could be achieved by the proposed OMD over the conventional SUD, for certain application scenarios.

APPENDIX I

PROOF OF LEMMA 3.1

We will prove Lemma 3.1 by contradiction. First, suppose that $\mu_1^* + \mu_2^* < 1$. Then, in the maximization problem of (16), from the expression of $L(r_1, S_1, \mu_1, \mu_2)$ in (15), it follows that the optimal $r_1$ that maximizes the Lagrangian is $r_1^* = +\infty$, which contradicts the fact that $r_1$ in (P2) is upper-bounded by finite rate values in the constraints (11) and (12). Second, suppose that $\mu_1^* + \mu_2^* > 1$. Similarly like the previous case, it can shown that $r_1^* = 0$. However, this can not be true since we can easily find a feasible solution set for $(r_1, S_1)$ in (P2) such that $r_1 > 0$. By combining the above two cases, it follows that $\mu_1^* + \mu_2^* = 1$. 

We rewrite $\hat{R}_2^{(a)}$ in (22) and $\hat{R}_2^{(a)}$ in (26) as

$$\hat{R}_2^{(a)} = \log \left| I + H_{11} S_{1}^{JD} H_{11}^H + H_{21} S_{2}^{HD} H_{21}^H \right| - \log \left| I + H_{11} S_{1}^{JD} H_{11}^H \right|$$

(38)

$$\hat{R}_2^{(a)} = \log \left| I + H_{11} S_{1}^{SD} H_{11}^H + H_{21} S_{2}^{HD} H_{21}^H \right| - \log \left| I + H_{11} S_{1}^{SD} H_{11}^H \right|.$$  

(39)

Since $S_{1}^{JD}$ and $S_{1}^{SD}$ are optimal for the sum-capacity (in an equivalent two-user MIMO-MAC) and user’1 channel capacity (without the presence of user 2), respectively, we have

$$\log \left| I + H_{11} S_{1}^{JD} H_{11}^H + H_{21} S_{2}^{HD} H_{21}^H \right| \geq \log \left| I + H_{11} S_{1}^{SD} H_{11}^H + H_{21} S_{2}^{HD} H_{21}^H \right|$$

(40)

$$\log \left| I + H_{11} S_{1}^{JD} H_{11}^H \right| \leq \log \left| I + H_{11} S_{1}^{SD} H_{11}^H \right|.$$  

(41)

Combining the above two inequalities with (38) and (39), it thus follows that $\hat{R}_2^{(a)} \geq \hat{R}_2^{(a)}$.

APPENDIX III

PROOF OF PROPOSITION 4.1

We first prove the former part of Proposition 4.1, i.e., the set $\mathcal{U}_1^*$ is unique, by contradiction. Suppose that there exist two optimal decodable user sets for user 1 with the same size, denoted by $\mathcal{A}_1$ and $\mathcal{B}_1$. Without loss of generality, we let $\mathcal{A}_1 = \{D, C\}$ and $\mathcal{B}_1 = \{E, C\}$, where $C, D$ and $E$ are subsets consisting of completely different user indexes. Then, we can express $\overline{\mathcal{A}}_1 = \{E, F\}$ and $\overline{\mathcal{B}}_1 = \{D, F\}$, where $F = \overline{\mathcal{A}}_1 \cup \overline{\mathcal{B}}_1$. Then, for users in the set $\mathcal{A}_1$, their transmit rates must satisfy [20]

$$\sum_{i \in \mathcal{J} \cup \mathcal{K}} r_i \leq \log \left| I + \left( I + \sum_{k \in \mathcal{A}_1} H_{k1} S_k H_{k1}^H \right)^{-1} \sum_{i \in \mathcal{J} \cup \mathcal{K}} H_{i1} S_i H_{i1}^H \right|, \forall \mathcal{J} \subseteq D, \mathcal{K} \subseteq C.$$  

(42)

Similarly, for users in the subset $E$ of $\mathcal{B}_1$, their transmit rates must satisfy

$$\sum_{i \in \mathcal{I}} r_i \leq \log \left| I + \left( I + \sum_{k \in \mathcal{B}_1} H_{k1} S_k H_{k1}^H \right)^{-1} \sum_{i \in \mathcal{I}} H_{i1} S_i H_{i1}^H \right|, \forall \mathcal{I} \subseteq E.$$  

(43)

Let $\mathcal{J}'$ be an orthogonal set of $\mathcal{J}$, where $\mathcal{J}' \cup \mathcal{J} = D$. Similarly, $\mathcal{I}'$ is defined for $\mathcal{I}$, where $\mathcal{I}' \cup \mathcal{I} = E$. (42) and (43) can thus be further shown as follows:

$$\sum_{i \in \mathcal{J}' \cup \mathcal{K}} r_i \leq \log \left| I + \left( I + \sum_{k \in \mathcal{J} \cup \mathcal{F}} H_{k1} S_k H_{k1}^H \right)^{-1} \sum_{i \in \mathcal{J}' \cup \mathcal{K}} H_{i1} S_i H_{i1}^H \right|.$$  

(44)
From (44) and (45), we obtain

\[ \sum_{i \in I} r_i \leq \log \left| I + \left( I + \sum_{k \in \mathcal{J} \cup \mathcal{F}} H_{k1} S_k H_{k1}^H \right)^{-1} \sum_{i \in I} H_{i1} S_i H_{i1}^H \right| . \]  

(45)

From (44) and (45), we obtain

\[ \sum_{i \in \mathcal{J} \cup \mathcal{K} \cup \mathcal{I}} r_i \leq \log \left| I + \sum_{i \in \mathcal{J} \cup \mathcal{K} \cup \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| + \log \left| I + \sum_{i \in \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| - \log \left| I + \sum_{i \in \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| - \log \left| I + \sum_{i \in \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| \]  

(46)

Since

\[ \log \left| I + \sum_{i \in \mathcal{J} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| - \log \left| I + \sum_{i \in \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| \leq \log \left| I + \sum_{i \in \mathcal{J} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| - \log \left| I + \sum_{i \in \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| \]  

(47)

From (46) and (47), it follows that

\[ \sum_{i \in \mathcal{J} \cup \mathcal{K} \cup \mathcal{I}} r_i \leq \log \left| I + \sum_{i \in \mathcal{J} \cup \mathcal{K} \cup \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| - \log \left| I + \sum_{i \in \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| - \log \left| I + \sum_{i \in \mathcal{I} \cup \mathcal{F}} H_{i1} S_i H_{i1}^H \right| = \log \left| I + \left( I + \sum_{k \in \mathcal{F}} H_{k1} S_k H_{k1}^H \right)^{-1} \sum_{i \in \mathcal{J} \cup \mathcal{K} \cup \mathcal{I}} H_{i1} S_i H_{i1}^H \right| . \]  

(48)

Thus, the set \( \mathcal{J} \cup \mathcal{K} \cup \mathcal{I} \) is a decodable user set for user 1 for any \( \mathcal{J} \subseteq \mathcal{D}, \mathcal{K} \subseteq \mathcal{C}, \text{ and } \mathcal{I} \subseteq \mathcal{E} \), and so is the set \( \mathcal{G}_1 = \mathcal{D} \cup \mathcal{C} \cup \mathcal{E} \). Since the size of \( \mathcal{G}_1 \) is larger than that of \( \mathcal{A}_1 \) or \( \mathcal{B}_1 \), this contradicts the assumption that \( \mathcal{A}_1 \) and \( \mathcal{B}_1 \) are optimal decodable user sets for user 1. The proof of the former part of Proposition 4.1 thus follows.

Next, we prove the latter part of Proposition 4.1 i.e., any decodable user set for user 1, \( \mathcal{U}_1 \), must be a subset of \( \mathcal{U}_1^* \). The proof is also obtained via contradiction. Suppose that there is a set \( \mathcal{U}_1 \) that is not a subset of \( \mathcal{U}_1^* \). Without loss of generality, we can express \( \mathcal{U}_1 = \{ \mathcal{D}, \mathcal{C} \} \) and \( \mathcal{U}_1^* = \{ \mathcal{E}, \mathcal{C} \} \), where \( \mathcal{C}, \mathcal{D} \) and \( \mathcal{E} \) are orthogonal subsets. Based on the proof for the former part of Proposition 4.1, we know that the set \( \mathcal{D} \cup \mathcal{C} \cup \mathcal{E} \) is also a decodable user set for user 1, and apparently, it has a larger size than \( \mathcal{U}_1^* \), which contradicts the fact that \( \mathcal{U}_1^* \) is the optimal decodable user set for user 1. The proof of the latter part of Proposition 4.1 thus follows.
Initialize $V = \{2, \ldots, K\}, \bar{V} = \emptyset$.

While $|V| > 0$ do (1)

Initialze $n = 1$

While $n \leq 2^{|V|} - 1$ do

If $\sum_{i \in V_n} r_i \leq C(V_n)$

Set $n \leftarrow n + 1$

Else

Set $V \leftarrow V - V_n$

Set $\bar{V} \leftarrow \bar{V} \cup V_n$

Go to (1)

End If

End While

Go to (2)

End While

Set $U^*_1 = V$. (2)

### TABLE I

| The Algorithm to Find $U^*_1$. |

#### APPENDIX IV

**Algorithm to Find $U^*_1$**

In this appendix, we present an algorithm to find the optimal decodable user set for user 1, $U^*_1$. First, some notations are given as follows for the convenience of presentation. Let $V_n$ denote a subset of an arbitrary set $V$, $n = 1, \ldots, 2^{|V|} - 1$. Note that here we have excluded the case that $V_n = \emptyset$ for the ease of presentation. The operation $V - V_n$ then stands for removing the subset $V_n$ from $V$.

For a given user set, $V \subseteq \{2, \ldots, K\}$, we know from Definition 4.1 that $V$ is a decodable user set for user 1 if and only if for any subset of $V$, $V_n$, it satisfies that

$$\sum_{i \in V_n} r_i \leq \log \left| I + \left( I + \sum_{k \in V} H_{k1} S_k H_{k1}^H \right)^{-1} \sum_{i \in V_n} H_{i1} S_i H_{i1}^H \right| \triangleq C(V_n). \quad (49)$$

However, if there exists a subset $V_n$ such that $\sum_{i \in V_n} r_i > C(V_n)$, it follows that $V$ should not be a decodable user set for user 1. From the above property, we are able to design an iterative algorithm to find $U^*_1$, which is explained as follows. Initially, we let $V = \{2, \ldots, K\}$. Thus, $\bar{V} = \emptyset$. Then, we will sequentially check for all the subsets of $V$ whether $\sum_{i \in V_n} r_i \leq C(V_n), \forall n$. If this is the case, then we declare that $U^*_1 = V$. However, if we find any $n'$ such that $\sum_{i \in V_{n'}} r_i > C(V_{n'})$, then we conclude that $V$ should not be $U^*_1$ and furthermore $U^*_1 \subseteq V - V_{n'}$. In this case, we will set $V \leftarrow V - V_{n'}, \bar{V} \leftarrow \bar{V} \cup V_{n'}$, and start a new sequence of tests for $\sum_{i \in V_n} r_i \leq C(V_n), \forall n$. The above procedure iterates until we find
a set $\mathcal{V}$ such that $\sum_{i \in \mathcal{V}} r_i \leq C(\mathcal{V}_n), \forall n$ or $\mathcal{V} = \emptyset$. In both cases, we set $\mathcal{U}_i^* = \mathcal{V}$. The above algorithm is summarized in Table I.

REFERENCES

[1] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding digital subscriber line technology*, Englewood Cliffs, NJ: Prentice-Hall, 1999.

[2] J. Andrews et al., “Rethinking information theory for mobile ad hoc networks,” *IEEE Commun. Magazine*, vol. 46, no. 12, pp. 94-101, Dec. 2008.

[3] N. Devroye, M. Vu, and V. Tarokh, “Cognitive radio networks: information theory limits, models and design,” *IEEE Sig. Proces. Magazine*, Nov. 2008.

[4] T. S. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49-60, Jan. 1981.

[5] R. Etkin, D. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,” *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534-5562, Dec. 2008.

[6] V. R. Cadambe and S. A. Jafar, “Interference alignment and the degrees of freedom for the K user interference channel,” *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425-3441, Aug. 2008.

[7] W. Yu, G. Ginis, and J. Cioffi “Distributed multiuser power control for digital subscriber lines”, *IEEE J. Sel. Areas Commun.*, vol. 20, no.5, pp. 1105-1115. Jun. 2002.

[8] M. F. Demirkol and M. A. Ingram, “Power-controlled capacity for interfering MIMO links,” in *Proc. IEEE VTC*, vol. 1 pp. 187-191, 2001.

[9] J. H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, “Joint transmitter receiver diversity for efficient space division multiaccess”, *IEEE Trans. Wireless Commun.*, vol. 1, pp. 16-27, Jan. 2002.

[10] R. S. Blum, “MIMO capacity with interference”, *IEEE J. Sel. Areas Comm.*, vol. 21, pp. 793-801, Jun. 2003.

[11] B. Chen and M. J. Gans, “MIMO communications in ad hoc networks”, *IEEE Trans. Signal Proces.*, vol. 54, pp. 2773-2783, Jul. 2006.

[12] S. Ye and R. S. Blum, “Optimized signaling for MIMO interference systems with feedback,” *IEEE Trans. Sig. Process.*, vol. 51, pp. 2839-2848, Nov. 2003.

[13] E. Larsson and E. Jorswieck, “Competition and collaboration on the MISO interference channel,” in *Proc. Allerton Conference on Communication, Control, and Computing*, 2007.

[14] C. Liang and K. R. Dandekar, “Power management in MIMO ad hoc networks: a game-theoretic approach,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 2866-2882, Apr. 2007.

[15] G. Arslan, M. F. Demirkol, and Y. Song, “Equilibrium efficiency improvement in MIMO interference systems: a decentralized stream control approach,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2984-2993, Aug. 2007.

[16] G. Scutari, D. P. Palomar, and S. Barbarossa, “Competitive design of multiuser MIMO systems based on game theory: a unified view,” *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1089-1103, Sep. 2008.

[17] S. Verdu, *Multi-user detection*, Cambridge University Press, 2003.

[18] N. Prasad, G. Yue, X. Wang, and M. Varanasi, “Optimal successive group decoders for MIMO multiple access channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4821-4846, Nov. 2008.

[19] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[20] T. Cover and J. Thomas, *Elements of information theory*, New York: Wiley, 1991.

[21] B. Rimoldi and R. Urbanke, “A rate-splitting approach to the Gaussian multiple-access channel,” *IEEE Trans. Inf. Theory*, vol. 42, pp. 364-375, Mar. 1996.

[22] D. Tse and S. Hanly, “Multi-access fading channels-Part I: polymatroid structure, optimal resource allocation and throughput capacities,” *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796-2815, Nov. 1998.

[23] R. Zhang and Y. C. Liang, “Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks,” *IEEE J. S. Topics Sig. Proces.*, vol. 2, no. 1, pp. 88-102, Feb. 2008.
Joint Decoding (JD)  
Successive Decoding (SD)  
Single−User Decoding (SUD)

Fig. 1. The maximum achievable rate of user 1 with OMD, $r_1$, as a function of user 2’s rate, $r_2$, for some fixed $S_2$. 
Fig. 2. The achievable sum-rate versus the average cross-link channel power gain, $\rho$, for a MU-MIMO system with $K = 2$, $M_k = N_k = 2$, $k = 1, 2$, and $P_1 = P_2 = 100$.

Fig. 3. The achievable rate versus the average transmit power, $P$, in a MIMO CR system with $M_k = N_k = 2$, $k = 1, 2$, $P_1 = 10P$, and $P_2 = P$, for different decoding methods: Case (I) both PU and SU employ SUD; and Case (II) PU employs SUD and SU employs OMD.