On the Capacity of Multiuser MIMO Networks with Interference

Jia Liu, *Student Member, IEEE,* Y. Thomas Hou, *Senior Member, IEEE,*
Yi Shi, *Member, IEEE,* Hanif D. Sherali, and Sastry Kompella, *Member, IEEE*

***Abstract*— Maximizing the total mutual information of multiuser multiple-input multiple-output (MIMO) systems with interference is a challenging problem. In this paper, we consider the power control problem of finding the maximum sum of mutual information for a multiuser network with mutually interfered MIMO links. We propose a new and powerful global optimization method using a branch-and-bound (BB) framework, coupled with a novel reformulation-linearization technique (RLT). The proposed BB/RLT guarantees finding a global optimum for multiuser MIMO networks with interference. To reduce the complexity of BB/RLT, we propose a modified BB variable selection strategy to accelerate the convergence process. Numerical examples are also given to demonstrate the efficacy of the proposed solution.**

***Index Terms*— Global optimization, multiple input multiple output, multiuser network, power control.**

I. Introduction

MULTIPLE-INPUT Multiple-Output (MIMO) systems have received extensive attention since Telatar [1] and Foschini *et al.* [2] showed the potential of high spectral efficiency provided by multiple antenna systems. The capacity gain by MIMO is achieved at no cost of extra spectrum. Since its introduction, MIMO has penetrated commercial wireless markets and will likely become one of the key underlying transmission technologies for future wireless systems. The invention of the MIMO technology has also brought much interest in a number of research areas, including channel coding [3], [4], MAC [5], [6], routing [7], and, in particular, in information-theoretic studies on the capacity of MIMO systems for both single-user and multiple users [8].

Compared to the research on the capacity of single-user MIMO, for which the "water-filling" solution was found in [1], the capacity limit of *multiuser* MIMO systems is much less studied and some fundamental problems remain unsolved [8]. There is a critical need to extend the MIMO communication concept from single-user to multiuser systems. It has been shown in [9], [10] that the sum-rate capacity of a multiuser MIMO system can be significantly degraded if co-channel interference is not managed carefully. Therefore, the problem of optimal power control and allocation in multiuser MIMO systems is not only of theoretical interest, but is also important in practice.

In this paper, we investigate this problem from a global optimization perspective. In particular, this paper proposes a solution procedure based on *Branch-and-Bound* framework coupled with the *Reformulation Linearization Technique* (BB/RLT). The main contributions of this paper include the mathematical developments of the solution procedure to solve the problem of finding the maximum sum of mutual information (MSMI) for multiuser MIMO systems based on BB/RLT technique, and its convergence speedup techniques. Specifically, we derive tight upper and lower bounds for each potential partitioning variable used for the problem. Each nonlinear term is relaxed with a set of linear constraints based on the bounds we develop to generate a higher dimensional upper-bounding problem. We also utilize a polyhedral outer approximation method to accurately approximate the logarithmic function. During each iteration of the branch-and-bound procedure, we propose a variable selection policy based not only on the relaxation error, but also on the relative significance of the variables in our problem. Our proposed method guarantees the finding of a global optimal solution to the MSMI of multiuser MIMO systems.

The remainder of this paper is organized as follows. We will first briefly review related work in Section II. Section III presents network model and problem formulation. Section IV introduces the BB/RLT framework and key problem-specific components, including factorization, linearization, and convergence speedup techniques. Simulation results are presented in Section V. Section VI concludes this paper.

II. Related Work

In [11], Jorswieck and Boche analyzed the worst-case performance of a multiuser MIMO system with interference. In [5], [6], Demirkol and Ingram introduced an iterative method based on stream control. This algorithm is based on a trial-and-error scheme and only considers simple network configurations, e.g., rectangular or hexagonal network topology. An interesting work was reported in [12], where Chen and Gans analyzed the network spectral efficiency of a MIMO ad hoc network with L simultaneous transmission pairs. They showed that, in the absence of channel state information (CSI) at the transmitters, the network's asymptotic spectral efficiency is limited by n_r nats/s/Hz as $L \to \infty$, and at least $n_t + n_r + 2\sqrt{n_t n_r}$ nats/s/Hz when CSI is available at the transmitters, where n_t and n_r are the numbers of transmitting and receiving antenna elements of each node, respectively. In contrast to such scaling law analysis, which shows trend

Manuscript received September 21, 2006; revised March 10, 2007 and July 24, 2007; accepted August 22, 2007. The associate editor coordinating the review of this paper and approving it for publication was A. Yener. This paper was presented in part at IEEE Globecom'06, San Francisco, California, USA, November 2006.

J. Liu, Y. T. Hou, and Y. Shi are with the Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061 USA (e-mail: {kevinlau, thou, yshi}@vt.edu.)

H. D. Sherali is with the Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061 USA (e-mail: hanifs@vt.edu).

S. Kompella is with Information Technology Division, Naval Research Laboratory, Washington, DC 20375 USA (e-mail: kompella@itd.nrl.navy.mil).

Digital Object Identifier 10.1109/TWC.2008.060732.

for very large networks, in this paper, we are interested in designing an algorithm that can compute exact maximum capacity for a given (finite size) network topology.

In [9], Blum showed that the sum of mutual information of a multiuser MIMO system is neither a convex nor a concave function. Also, a change in the covariance matrix of one user will result in a change of the mutual information of all users. It is thus very difficult to solve the problem analytically. Therefore, most researchers resort to iterative local optimization methods for solving this problem. In [13], Yu *et al.* proposed an "iterative water-filling" technique (IWF). This technique was originally proposed for digital subscriber line (DSL) systems. IWF was applied to MIMO multiple access channel (MIMO-MAC) problems because of its simplicity, and its provable convergence to global optimality due to the convexity of MIMO-MAC channels. However, IWF experiences convergence difficulties in MIMO ad hoc networks with mutually interfered links due to the absence of convexity property. Several variants of IWF were proposed by Jindal *et al.* for MIMO broadcast (MIMO-BC) channels [14]. But they too cannot be directly extended to MIMO ad hoc networks. IWF can be viewed as a noncooperative game and its convergence point is in essence a Nash equilibrium where no user can increase self-utility function by unilaterally deviating from this stationary point. Ye and Blum [15] introduced a "global gradient projection" (GGP) method, which is an extension of the well-known steepest descent method coupled with gradient projection. Unlike the noncooperative game in IWF, all users cooperate in determining covariance matrices by computing the gradients at each iteration. GGP is more sensitive to the choice of a starting point and it is easier to get trapped at a local optimum than IWF. Moreover, GGP exhibits "zigzagging" as it approaches a local optimal solution [16]. Both GGP and IWF can be classified as local optimization techniques, which can quickly find a local optimal solution, but cannot guarantee global optimality for nonconvex optimization problems.

III. Network Model and Problem Formulation

We begin by introducing mathematical notation for matrices, vectors, and complex scalars in this paper. We use boldface to denote matrices and vectors. For a matrix $\mathbf{A}$, $\mathbf{A}^\dagger$ denotes the conjugate transpose. $\mathrm{Tr}\{\mathbf{A}\}$ denotes the trace of $\mathbf{A}$. We let $\mathbf{I}$ denote the identity matrix, whose dimension can be determined from the context. $\mathbf{A} \succeq 0$ represents that $\mathbf{A}$ is Hermitian and positive semidefinite (PSD). $\mathbf{1}$ and $\mathbf{0}$ denote vectors whose elements are all ones and zeros, respectively. The dimensions of $\mathbf{1}$ and $\mathbf{0}$ can be determined from context and thus omitted for brevity. The scalar $a_{(m,n)}$ represents the entry in the m^{th}-row and n^{th}-column of $\mathbf{A}$. For a complex scalar a, $\mathfrak{Re}(a)$ and $\mathfrak{Im}(a)$ represent the real and imaginary parts of a, respectively, $\|a\|$ represents the modulus of a, and $\overline{a}$ represents the conjugate of a. We consider a network consisting of L interfering concurrent MIMO transmission pairs (links), which are indexed by $1, 2, \dots, L$. In this paper, it is assumed that the transmitters have full CSI. Let the matrix $\mathbf{H}_{jl} \in \mathbb{C}^{n_r \times n_t}$ represent the wireless channel gain matrix from the transmitting node of link j to the receiving node of link l, where n_t and n_r are the numbers of transmitting and receiving antenna elements of each node, respectively. Denote ρ_{jl} the signal-to-noise ratio per unit transmit power if $j = l$, or the interference-to-noise ratio per unit transmit power if $j \neq l$. Denote matrix $\mathbf{Q}_l$ the covariance matrix of a zero-mean Gaussian input symbol vector $\mathbf{x}_l$ at link l, i.e., $\mathbf{Q}_l = \mathbb{E}\left\{\mathbf{x}_l \cdot \mathbf{x}_l^\dagger\right\}$. Assume, also, that all nodes in the network are subject to the same maximum transmitting power constraint, i.e., $\mathrm{Tr}\left\{\mathbf{Q}_l\right\} \le P_{\max}$, where $P_{\max}$ is the maximum transmission power. Let $\mathbf{R}_l$ represent the covariance matrix of interference plus noise. Define $\mathcal{I}_l$ as the set of links that can produce interference on link l. The interference-plus-noise is Gaussian distributed and its covariance matrix can be computed as

$$\mathbf{R}_l = \sum_{j \in \mathcal{I}_l} \rho_{jl} \mathbf{H}_{jl} \mathbf{Q}_j \mathbf{H}_{jl}^\dagger + \mathbf{I}. \tag{1}$$

Hence, the mutual information of a MIMO link l with co-channel interference can be computed as $I_l = \log_2 \det\left(\rho_{ll}\mathbf{H}_{ll}\mathbf{Q}_l\mathbf{H}_{ll}^\dagger + \mathbf{R}_l\right) - \log_2 \det \mathbf{R}_l$. Our goal is to maximize the sum of mutual information (MSMI) of this L-link MIMO interference system. Summarizing the previous discussion, this optimization problem can be mathematically formulated as follows:

$$\begin{array}{ll} \max & \sum_{l=1}^{L} I_l \\ \text{s.t.} & I_l = \log_2 \det\left(\rho_{ll}\mathbf{H}_{ll}\mathbf{Q}_l\mathbf{H}_{ll}^\dagger + \mathbf{R}_l\right) - \log_2 \det \mathbf{R}_l \\ & \mathbf{R}_l = \sum_{j \in \mathcal{I}_l} \rho_{jl}\mathbf{H}_{jl}\mathbf{Q}_j\mathbf{H}_{jl}^\dagger + \mathbf{I} \\ & \mathrm{Tr}\{\mathbf{Q}_l\} \le P_{\max}, \mathbf{Q}_l \succeq 0, 1 \le l \le L. \end{array}$$

In this paper, we consider a network where each antenna element in a transmitting node employs equal power allocation. It is necessary to point out that, by saying "equal power allocation", we mean the total power at the same source node is equally allocated to its antenna elements, while different source nodes may have different total transmitting power. The reason behind this approach is that an optimal power allocation, wherein different antenna elements at the same source node have different transmitting power level, puts a high demand of linearity in transmit power amplifiers, which is extremely costly from a practical standpoint [17]. Thus, for low cost hardware implementation, an equal power allocation scheme is more attractive.

Under the equal power allocation approach, the MSMI problem is translated into an optimal power control problem. That is, we are interested in finding an L-dimension power vector $\mathbf{p} = (p_1, p_2, \dots, p_L)^t$, where $0 < p_l \le P_{\max}$, $l = 1, 2, \dots, L$, such that this power vector $\mathbf{p}$ maximizes the sum of mutual information of the links in the network. Mathematically, with equal power allocation to each transmitter at the same node, the input covariance matrix $\mathbf{Q}_l$ becomes an n_t-dimension scaled identity matrix, i.e., $\mathbf{Q}_l = \frac{p_l}{n_t}\mathbf{I}$. Hence, the MSMI problem formulation can be further re-written as follows:

$$\begin{array}{ll} \max & \sum_{l=1}^{L} I_l \\ \text{s.t.} & I_l = \log_2 \det\left(\frac{\rho_{ll}p_l}{n_t}\left(\mathbf{H}_{ll}\mathbf{H}_{ll}^\dagger\right) + \mathbf{R}_l\right) - \log_2 \det \mathbf{R}_l \\ & \mathbf{R}_l = \mathbf{I} + \sum_{j \in \mathcal{I}_l} \frac{\rho_{jl}p_j}{n_t}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right) \end{array} \tag{2}$$

where $0 < p_l \le P_{\max}, 1 \le l \le L$.

IV. Solution Procedure

A. Overview of BB/RLT Method

For a non-convex optimization problem, conventional nonlinear programming methods (e.g., GGP [15]) can at best yield local optimal solutions. On the other hand, the proposed BB/RLT procedure in this paper can find a provably global optimal solution [18]–[20]. The basic idea of BB/RLT is that, by using the RLT technique, we can construct a linear programming (LP) relaxation for the original nonlinear programming (NLP) problem, which can be used to efficiently compute a global upper bound, UB, for the original NLP problem. This relaxation solution is either a feasible solution to the original NLP problem or, if not feasible, can be used as a starting point for a local search algorithm to find a feasible solution to the original NLP problem. This feasible solution will then serve to provide a global lower bound, LB, and an incumbent solution to the original NLP problem. However, it is worth to point out that local search is not necessary in this paper since the LP relaxation for our problem is still subject to the same power constraint $0 \leq p_l \leq P_{\max}$ of the original problem. As a result, solving the LP relaxation of our problem still gives us a feasible solution to the original NLP problem. This will become clearer after we introduce the LP relaxation in Section IV-E. The branch-and-bound process will proceed by tightening UB and LB, and terminates when $LB \geq (1-\epsilon)UB$ is satisfied, where ϵ is the desired approximation error. There is a formal proof that BB/RLT converges to a global optimal solution as long as the partitioning intervals are compact (see [18]–[20] for further details). The general framework of BB/RLT is shown in Algorithm 1.

Algorithm 1 BB/RLT Solution Procedure

Initialization:
1. Let optimal solution $\psi^* = \emptyset$. The initial lower bound $LB = -\infty$.
2. Determine partitioning variables (variables associated with nonlinear terms) and derive their initial bounding intervals.
3. Let the initial problem list contain only the original problem, denoted by P_1.
4. Introduce one new variable for each nonlinear term. Add linear constraints for these variables to build a linear relaxation. Denote the solution to linear relaxation as $\hat{\psi}_1$ and its objective value as the upper bound UB_1.

Main Loop:
1. Select problem P_z that has the largest upper bound among all problems in the problem list.
2. Find, if necessary, a feasible solution ψ_z via a local search algorithm for Problem P_z. Denote the objective value of ψ_z by LB_z.
3. If $LB_z > LB$ then let $\psi^* = \psi_z$ and $LB = LB_z$. If $LB \geq (1-\epsilon)UB$ then stop with the $(1-\epsilon)$-optimal solution ψ^*; else, remove all problems $P_{z'}$ having $(1-\epsilon)UB_{z'} \leq LB$ from the problem list.
4. Compute relaxation error for each nonlinear term.
5. Select a partitioning variable having the maximum relaxation error and divide its bounding interval into two new intervals by partitioning at its value in $\hat{\psi}_z$.
6. Remove the selected problem P_z from the problem list, construct two new problems P_{z1} and P_{z2} based on the two partitioned intervals.
7. Compute two new upper bounds UB_{z1} and UB_{z2} by solving the linear relaxations of P_{z1} and P_{z2}, respectively.
8. If $LB < (1-\epsilon)UB_{z1}$ then add problem P_{z1} to the problem list. If $LB < (1-\epsilon)UB_{z2}$ then add problem P_{z2} to the problem list.
9. If the problem list is empty, stop with the $(1-\epsilon)$-optimal solution ψ^*. Otherwise, go to the Main Loop again.

In the remainder of this section, we develop the key components in the BB/RLT framework, which are problem-specific.

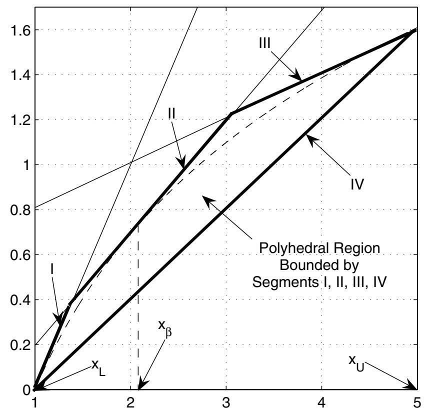

Fig. 1. Polyhedral outer-approximation $y = \ln x$.

B. Factorization and Linearization

Observe that in the MSMI problem formulation, the link mutual information expressions in (2) are nonlinear. To linearize these nonlinear constraints, we first introduce four new variables X_l, Y_l, V_l, and W_l as follows.

$$\begin{cases} X_l = \det\left(\frac{\rho_{ll} p_l}{n_t}\left(\mathbf{H}_{ll}\mathbf{H}_{ll}^{\dagger}\right) + \mathbf{R}_l\right) \triangleq \det \mathbf{D}_l, \\ V_l = \det \mathbf{R}_l, \quad Y_l = \ln X_l, \quad W_l = \ln V_l. \end{cases} \tag{3}$$

The link mutual information constraint in (2) can then be translated into $I_l = \frac{1}{\ln 2}\left(Y_l - W_l\right)$, which is a linear constraint. Also, four groups of new constraints in (3) are added to the problem formulation.

C. Linear Relaxation to Nonlinear Logarithmic Functions

Next, we propose using a polyhedral outer approximation for the curve of logarithmic function. As shown in Fig. 1, the function $y = \ln x$, over an interval defined by suitable upper and lower bounds on x, can be upper-bounded by three tangential segments I, II, and III, which are constructed at $(x_L, \ln x_L)$, $(x_\beta, \ln x_\beta)$, and $(x_U, \ln x_U)$, where x_β is computed as follows:

$$x_\beta = \frac{x_L x_U(\ln x_U - \ln x_L)}{x_U - x_L}. \tag{4}$$

Here x_β is the x-value for the point at the intersection of the extended tangent segments I and III. Segment IV is the chord that joins $(x_L, \ln x_L)$ and $(x_U, \ln x_U)$. The convex region defined by the four segments can be described by the following four linear constraints:

$$\begin{gathered} x_L \cdot y - x \leq x_L(\ln x_L - 1), \\ x_\beta \cdot y - x \leq x_\beta(\ln x_\beta - 1), \\ x_U \cdot y - x \leq x_U(\ln x_U - 1), \\ (x_U - x_L)y + (\ln x_L - \ln x_U)x \geq x_U \cdot \ln x_L - x_L \cdot \ln x_U. \end{gathered}$$

D. Linearizing the Determinants

Substituting the expression for $\mathbf{R}_l$ into that for $\mathbf{D}_l$, and observing the similarity between the expressions for $\mathbf{D}_l$ and $\mathbf{R}_l$, we can write $\mathbf{D}_l$ and $\mathbf{R}_l$ in a more compact form by introducing a notion called the "super interference set" of link l, denoted by $\hat{\mathcal{I}}_l$, with $\hat{\mathcal{I}}_l = \mathcal{I}_l \cup \{l\}$, as follows:

$$\begin{aligned} \mathbf{D}_l &= \textstyle\sum_{j\in\hat{\mathcal{I}}_l} \frac{\rho_{jl}}{n_t}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^{\dagger}\right) p_j + \mathbf{I}, \\ \mathbf{R}_l &= \textstyle\sum_{j\in\mathcal{I}_l} \frac{\rho_{jl}}{n_t}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^{\dagger}\right) p_j + \mathbf{I}. \end{aligned} \quad (5)$$

It is evident from (5) that the determinants of $\mathbf{D}_l$ and $\mathbf{R}_l$ are in essence n_r-order polynomials of the variables $p_1, p_2, \ldots, p_L$. To illustrate how to linearize determinants, let us consider a multiuser MIMO network where every link has two receiving antennas. This means that $\mathbf{D}_l$ and $\mathbf{R}_l$ are 2×2 square matrices. The determinant X_l and V_l are shown in (6) and (7), respectively.

We see that the product terms $p_j p_k$ are the only nonlinear terms, which need to be linearized. To show how linearization works, let us consider a general second-order polynomial term $p_j p_k$, for which we have the following bounding constraints:

$$\begin{aligned} p_j - (p_j)_L \geq 0, \quad (p_j)_U - p_j \geq 0, \\ p_k - (p_k)_L \geq 0, \quad (p_k)_U - p_k \geq 0, \end{aligned} \quad (8)$$

where $(p_j)_L$ and $(p_k)_L$ denote the lower bounds of p_j and p_k, respectively, and $(p_j)_U$ and $(p_k)_U$ denote the upper bounds of p_j and p_k, respectively. Adopting RLT [18], we can derive the following four so-called bounding-factor constraints:

$$\begin{aligned} p_j p_k - (p_k)_L p_j - (p_j)_L p_k &\geq -(p_j)_L (p_k)_L, \\ p_j p_k - (p_k)_U p_j - (p_j)_L p_k &\leq -(p_j)_L (p_k)_U, \\ p_j p_k - (p_k)_L p_j - (p_j)_U p_k &\leq -(p_j)_U (p_k)_L, \\ p_j p_k - (p_k)_U p_j - (p_j)_U p_k &\geq -(p_j)_U (p_k)_U. \end{aligned}$$

In particular, if $j = k$, $p_j p_k$ is given by a general square term p_j^2. Using the following bounding constraints:

$$p_j - (p_j)_L \geq 0 \quad \text{and} \quad (p_j)_U - p_j \geq 0, \quad (9)$$

we can derive the following three bounding-factor constraints:

$$\begin{aligned} p_j^2 - 2(p_j)_L p_j \geq -(p_j)_L^2, \quad p_j^2 - 2(p_j)_U p_j \geq -(p_j)_U^2, \\ p_j^2 - ((p_j)_L + (p_j)_U) p_j \leq -(p_j)_U (p_j)_L. \end{aligned}$$

We now introduce new variables P_{jk} to replace the product terms $p_j p_k$, and P_{jj} to replace the square term p_j^2, respectively. By doing so, the determinant expressions of X_l and V_l become linear constraints. Also, the equality relation $P_{jk} = p_j p_k$ will be relaxed by the above bounding-factor constraint relaxations. All these newly introduced bounding-factor constraints will be appended to the original problem, thus achieving a LP relaxation for the constraints in the original NLP problem.

E. RLT-Based Relaxation for MSMI (R-MSMI)

For convenience, we define the right hand sides of (6) and (7) as $\Phi_{X_l}(\mathbf{p})$ and $\Phi_{V_l}(\mathbf{p})$, respectively. From the discussions in the previous sections, we have the final R-MSMI formulation for a multiuser MIMO system as follows:

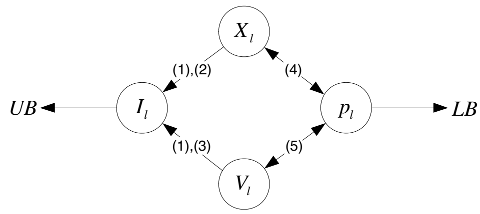

Fig. 2. Relationship among BB variables in MSMI.

R-MSMI

$$\begin{aligned} \max \quad & \textstyle\sum_{l=1}^{L} I_l \\ \text{s.t.} \quad & I_l - \tfrac{1}{\ln 2}(Y_l - W_l) = 0, \quad \forall l. \\ & \text{Three tangential supports for } (Y_l, X_l), \quad \forall l. \\ & \text{Three tangential supports for } (W_l, V_l), \quad \forall l. \\ & X_l - \Phi_{X_l}(\mathbf{p}) = 1, V_l - \Phi_{V_l}(\mathbf{p}) = 1, \quad \forall l. \\ & \text{Bounding constraints for } P_{jk} \text{ and } P_{jj}. \\ & \text{Same power constraints for power variables } p_l, \quad \forall l. \end{aligned}$$

F. Partitioning Variables and Their Upper and Lower Bounds

The partitioning variables in the branch-and-bound process are those that are involved in nonlinear terms, for which we have therefore defined new variables, and whose bounding intervals will need to be partitioned during the RLT-based branch-and-bound algorithm [18]–[20]. In R-MSMI, these partitioning variables include X_l, V_l, and p_l, $l = 1, 2, \ldots, L$. The upper and lower bounds for p_l are $(p_l)_L = 0$, $(p_l)_U = P_{\max}$, for $l = 1, \ldots, L$. From these expressions, we see that the upper bounds for X_l and V_l can, respectively, be computed as $\Phi_{X_l}(P_{\max}\times\mathbf{1})$ and $\Phi_{V_l}(P_{\max}\times\mathbf{1})$, with $\mathbf{1}, \mathbf{0} \in \mathbb{R}^L$. The lower bounds for X_l and V_l are: $(X_l)_L = 1$, and $(V_l)_L = 1$.

G. Convergence Speedup Techniques

In the worst case, BB/RLT has exponential complexity. However, it is possible to exploit certain special structure of the underlying problem to speedup convergence time. For our problem, since the decrease of the global upper bound plays the most critical role in the convergence process, partitioning variables that are able to tighten the upper bound (i.e., X_l and V_l as shown in Fig. 2) should be selected first. Hence, we adopt the following convergence speedup technique, as shown in Algorithm 2. In our numerical results, the threshold value ϵ_1 for $\ln(Z_l^*)_U - \ln(Z_l^*)_L$ are set to 1 since we find that threshold value less than 1 does not further improve the accuracy of the final solution. The relaxation error refers to the difference between the newly introduced variable and its corresponding product terms (e.g., $P_{ij} - p_i p_j$). The threshold value ϵ_2 for testing the relaxation error is 10^{-3} in this paper.

Algorithm 2 Modified BB Variable Selection Strategy

1. Among all X_l and V_l, choose the one, say Z_l^*, having the largest relaxation error.
2. If $(\ln(Z_l^*)_U - \ln(Z_l^*)_L \leq \epsilon_1)$ then
   a) Among all p_l, choose one, say p_l^*, with the largest relaxation error. Denote this relaxation error as E_p;
   b) If $E_p \leq \epsilon_2$, then remove this subproblem; else return p_l^*;

   else return Z_l^*.

$$X_l = 1 + \sum_{j \in \hat{\mathcal{I}}_l} \frac{\rho_{jl}}{n_r} \left[\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(1,1)} + \left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,2)} \right] p_j + \sum_{j \in \hat{\mathcal{I}}_l} \sum_{k \in \hat{\mathcal{I}}_l} \frac{\rho_{jl}\rho_{kl}}{n_r^2} \left[\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(1,1)} \left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,2)} \right.$$
$$\left. - \mathfrak{Re} \left(\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,1)} \right) \mathfrak{Re} \left(\left(\mathbf{H}_{kl} \mathbf{H}_{kl}^\dagger \right)_{(2,1)} \right) - \mathfrak{Im} \left(\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,1)} \right) \mathfrak{Im} \left(\left(\mathbf{H}_{kl} \mathbf{H}_{kl}^\dagger \right)_{(2,1)} \right) \right] p_j p_k. \quad (6)$$

$$V_l = 1 + \sum_{j \in \mathcal{I}_l} \frac{\rho_{jl}}{n_r} \left[\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(1,1)} + \left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,2)} \right] p_j + \sum_{j \in \mathcal{I}_l} \sum_{k \in \mathcal{I}_l} \frac{\rho_{jl}\rho_{kl}}{n_r^2} \left[\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(1,1)} \left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,2)} \right.$$
$$\left. - \mathfrak{Re} \left(\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,1)} \right) \mathfrak{Re} \left(\left(\mathbf{H}_{kl} \mathbf{H}_{kl}^\dagger \right)_{(2,1)} \right) - \mathfrak{Im} \left(\left(\mathbf{H}_{jl} \mathbf{H}_{jl}^\dagger \right)_{(2,1)} \right) \mathfrak{Im} \left(\left(\mathbf{H}_{kl} \mathbf{H}_{kl}^\dagger \right)_{(2,1)} \right) \right] p_j p_k. \quad (7)$$

TABLE I

SNR- AND INR-VALUE FOR A HEAVILY INTERFERED 5-LINK NETWORK (IN DB)

	SNR (in dB)	INR (in dB)					
			L0 Rx	L1 Rx	L2 Rx	L3 Rx	L4 Rx
L0	20.98	L0 Tx	–	13.57	3.79	9.13	2.23
L1	27.04	L1 Tx	18.90	–	6.33	12.38	4.35
L2	20.67	L2 Tx	4.31	6.61	–	7.53	13.39
L3	21.03	L3 Tx	7.39	9.48	9.29	–	4.26
L4	22.57	L4 Tx	4.10	6.19	11.61	5.58	–

TABLE II

SNR- AND INR-VALUES FOR A LESS INTERFERED 5-LINK NETWORK

	SNR (in dB)	INR (in dB)					
			L0 Rx	L1 Rx	L2 Rx	L3 Rx	L4 Rx
L0	20.98	L0 Tx	–	10.98	1.75	0.56	3.48
L1	27.04	L1 Tx	14.84	–	4.02	3.03	7.34
L2	20.67	L2 Tx	1.85	3.41	–	9.02	4.19
L3	21.03	L3 Tx	0.74	2.89	7.93	–	7.74
L4	22.18	L4 Tx	3.17	6.29	4.41	7.51	–

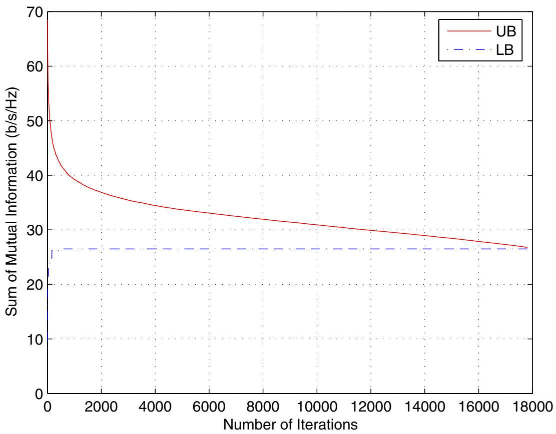

Fig. 3. A 5-link heavily interfered network example.

V. NUMERICAL RESULTS

We first describe our simulation settings. L links are uniformly distributed within a square region. Each node in the network is equipped with two antennas. The maximum transmit power for each node is set to $P_{\max} = 10$ dBm. The network operates in 2.4 GHz ISM band. The channel bandwidth is $W = 30$ MHz. The path loss index is chosen to be $\alpha = 2$.

We use two 5-link network examples to demonstrate the convergence properties of BB/RLT. The first one is an example of a heavily interfered network where each link is interfered by every other link. The desired error bound is chosen to be $\epsilon = 0.01$. That is, the iterative process stops once $LB \geq (1-\epsilon)UB$. The network's SNR- and INR-values are shown in Table I. For example, a cell intersected by row i and column j contains the INR (in dB) from link i to link j. The convergence process is depicted in Fig. 3.

Fig. 3 illustrates the UB and LB in terms of the sum of mutual information (b/s/Hz) at each iteration. In this heavily interfered example, after 17000 iterations, the UB and LB values are both driven to 26.51 b/s/Hz, meaning that the global optimum for the MSMI is 26.51 b/s/Hz. In this example, the optimal power vector is $\mathbf{p} = \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & p_4 \end{bmatrix}^t = \begin{bmatrix} 8.904 & 0.071 & 2.83 & 1.096 & 2.83 \end{bmatrix}^t$ (in mW).

To see how significantly the system's performance could be degraded by interference, we compute the sum of total mutual information for the same network as if there is no interference. In this particular example, the total mutual information for the case of no interference is obtained as 71.2 b/s/Hz. Thus, it can be seen that, even after carefully choosing the optimal power vector $\mathbf{p}$, the spectral efficiency only accounts for approximately 37% of that of the no interference case.

It can also be seen from Fig. 3 that the rate of decrease in UB plays the major role in determining how fast the overall BB/RLT process converges. In this example, UB starts out from 68.47 b/s/Hz, and changes by 41.96 in magnitude by the time the algorithm terminates. Comparatively, LB starts out from 8.37 b/s/Hz. It only changes 18.14 in magnitude by the end of the convergence process. Moreover, the rate of decline of UB becomes slower as it approaches the global optimum.

Next, for comparison, we study another 5-link network having less interference as compared to the previous example. The network's SNR- and INR-values are shown in Table II. The convergence process is depicted in Fig. 4. In this example, BB/RLT converges to the global optimal rather quickly (in about 3810 iterations), yielding an optimum value of 37.65 b/s/Hz. This convergence speed is comparable to many fast local search algorithm, such as GGP [15] and IWF [13]. For this particular example, we also compute the MSMI assuming there is no interference, which gives a value of 74.59 b/s/Hz. That is, in this less interfered network, the spectral efficiency is approximately 50% of that of the no interference case. For this less interfered example, the op-

TABLE III
SNR- AND INR-VALUES FOR A 10-LINK NETWORK

	SNR (in dB)	INR (in dB)										
			L0 Rx	L1 Rx	L2 Rx	L3 Rx	L4 Rx	L5 Rx	L6 Rx	L7 Rx	L8 Rx	L9 Rx
L0	23.19	L0 Tx	–	-0.12	2.29	-2.34	8.01	-1.47	4.73	2.59	-0.28	-3.37
L1	23.99	L1 Tx	0.43	–	8.67	7.54	2.12	0.79	7.98	4.76	-1.53	1.85
L2	20.44	L2 Tx	2.75	8.69	–	2.79	3.58	-0.52	12.74	4.19	-1.77	-0.62
L3	22.41	L3 Tx	-2.20	7.61	1.44	–	-0.22	2.86	2.09	3.32	-1.21	8.34
L4	21.63	L4 Tx	9.11	1.21	1.76	-0.63	–	2.15	5.74	7.88	4.04	-1.08
L5	23.82	L5 Tx	-0.90	1.83	-0.89	3.00	2.18	–	1.31	6.76	5.46	6.86
L6	26.59	L6 Tx	4.99	7.51	9.15	2.81	7.64	1.39	–	8.47	0.27	0.28
L7	23.58	L7 Tx	3.35	4.02	2.27	2.60	9.00	6.76	6.47	–	5.08	2.43
L8	25.23	L8 Tx	0.79	-1.25	-2.18	-1.51	3.90	5.88	0.22	4.44	–	-0.07
L9	25.56	L9 Tx	-3.33	1.97	-1.64	6.03	-1.24	5.56	-0.53	2.09	-0.09	–

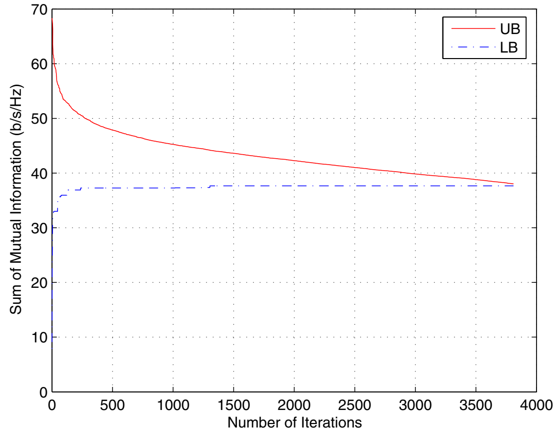

Fig. 4. A 5-link less interfered network example.

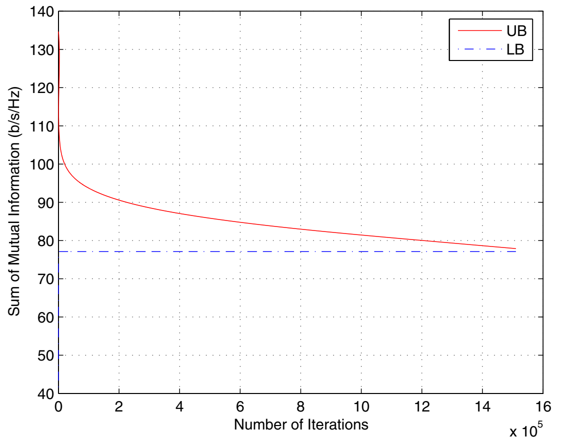

Fig. 5. A 10-link network example.

timal power vector is $\mathbf{p} = \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & p_4 \end{bmatrix}^t = \begin{bmatrix} 1.815 & 0.435 & 4.371 & 2.973 & 5.942 \end{bmatrix}^t$ (in mW). The reason why convergence time is much shorter than the previous heavily interfered network example is because the initial upper bounds of $(X_l)_U$ and $(V_l)_U$ are much smaller. This means that we have relatively a small interval for X_l and V_l variables to partition, which contributes to a faster convergence speed. Also, in a less interfered network, the initial gap between UB and LB is smaller. In this particular example, LB reaches 36.65 b/s/Hz only after 49 iterations, which is already very close to the final optimal value of 37.65 b/s/Hz. At the 49^{th} iteration, UB is at 56.15 b/s/Hz. This gap between LB and UB is about 19.5, which is much smaller than that in the first example. For these reasons, we can see that the less interference in the network, the faster BB/RLT will converge to find the global optimal solution.

To shed light on the huge effect of using the modified partitioning variable selection strategy, we consider the following 10-link network example, whose SNR- and INR-values are shown in Table III. The convergence process is depicted in Fig. 5. The global optimal value for this 10-link network is 77.11 b/s/Hz. It takes approximately 1.5×10^6 iterations to converge with the modified partitioning variable selection strategy (Algorithm 2). On the other hand, if we solve the same 10-link example using BB/RLT without using this strategy, the convergence time is much slower. Via rough estimate, we find that our proposed speedup technique can accelerate the convergence by at least 1000 times for this particular example.

VI. CONCLUSION

In this paper, we studied the maximum sum of mutual information problem for multiuser MIMO network. We proposed a powerful global optimization method using a branch-and-bound framework coupled with the reformulation-linearization technique (BB/RLT). We also proposed a modified branch-and-bound variable selection strategy to accelerate the convergence process. Numerical examples are given to demonstrate the efficacy of the proposed solution.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers for their constructive comments. The work of Y. T. Hou, J. Liu, and Y. Shi has been supported in part by the National Science Foundation (NSF) under Grant CNS-0721421 and the Office of Naval Research (ONR) under Grant N00014-08-1-0084. The work of H.D. Sherali has been supported in part by NSF Grant 0552676. The work of S. Kompella has been supported in part by ONR.

REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.
[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, pp. 311–355, Mar. 1998.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[5] M. F. Demirkol and M. A. Ingram, "Power-controlled capacity for interfering MIMO links," in *Proc. IEEE Veh. Technol. Conf.*, Oct. 2001, pp. 187–191.

[6] ——, "Stream control in network with interfering MIMO links," in *Proc. IEEE Wireless Commun. Networking Conf.*, Mar. 2003, pp. 343–348.

[7] K. Sundaresan and R. Sivakumar, "Routing in ad hoc networks with MIMO links," in *Proc. IEEE Internatinoal Conf. Network Protocols*, Nov. 2005, pp. 85–98.

[8] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 1, pp. 684–702, June 2003.

[9] R. S. Blum, "MIMO capacity with interference," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 793–801, June 2003.

[10] S. Cartreux, L. J. Greenstein, and P. F. Dressen, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Commun. Lett.*, vol. 4, no. 11, pp. 334–336, Nov. 2000.

[11] E. A. Jorswieck and H. Boche, "Performance analysis of capacity of MIMO systems under multiuser interference based on worst-case noise behavior," *EURASIP J. Wireless Commun. Networking*, vol. 2004, no. 2, pp. 273–285, Feb. 2004.

[12] B. Chen and M. J. Gans, "MIMO communications in ad hoc networks," *IEEE Trans. Signal Processing*, vol. 54, no. 7, pp. 2773–2783, July 2006.

[13] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.

[14] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.

[15] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[16] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. New York, NY: John Wiley & Sons Inc., 2006.

[17] J. Jiang, R. M. Buehrer, and W. H. Tranter, "Greedy scheduling performance for a zero-forcing dirty paper coded system," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 789–793, May 2006.

[18] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization-Technique for Solving Discrete and Continuous Nonconvex Problems*. Boston, MA: Kluwer Academic Publishing, 1999.

[19] H. D. Sherali and C. H. Tuncbilek, "A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique," *J. Global Optimization*, vol. 2, no. 1, pp. 101–112, 1992.

[20] H. D. Sherali and H. Wang, "Global optimization of nonconvex factorable programming problems," *Mathematical Programming*, vol. 89, no. 3, pp. 459–478, 1997.