Optimization of Multiuser MIMO Networks with Interference

Jia Liu*, Y. Thomas Hou*, Yi Shi*, and Hanif D. Sherali†
* Department of Electrical and Computer Engineering
† Department of Industrial and Systems Engineering
Virginia Polytechnic Institute and State University, Blacksburg, VA 24061
Email: {kevinlau, thou, yshi, hanifs}@vt.edu

***Abstract*— Maximizing the total mutual information of a multiuser multiple-input multiple-output (MIMO) system with interference is a well-known and challenging problem. In this paper, we consider the power control problem of finding the maximum sum of mutual information for multiuser MIMO systems with equal power allocation at each link. A new and powerful global optimization method using a branch-and-bound framework coupled with the reformulation-linearization technique (BB/RLT) is introduced. The proposed BB/RLT is the first such method that guarantees finding a global optimum for multiuser MIMO systems with interference. In addition, we propose a modified branch-and-bound (BB) variable selection strategy to accelerate the convergence process, and apply the proposed technique to several MIMO systems in order to demonstrate its efficacy.**

I. Introduction

Multiple-Input Multiple-Output (MIMO) system has received extensive attention since Telatar's [1] and Foschini's [2] pioneering works predicting the potential of high capacity provided by multiple antenna systems. Compared to the research on the capacity of single-user MIMO, for which the famous "water-filling" solution has been known for years, the capacity limit of multiuser MIMO system is much less studied and many problems are still unsolved [3]. It has been shown in [4], [5] that the total capacity of a multiuser MIMO system can be degraded significantly if cochannel interference is not managed carefully.

In recent years, several interesting algorithms and results on multiuser MIMO system with mutual interference have been reported in the literature. It is shown in [4] that the sum of mutual information of a multiuser MIMO system is neither convex nor concave function. Thus, it is generally very difficult to solve the problem numerically, let alone analytically. Most current research resort to iterative local optimization methods in solving this problem. In [6], the authors proposed an "iterative water-filling" technique (IWF) for 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 may experience convergence difficulties in MIMO ad hoc network due to the absence of the convexity propoerty. Several variants of IWF found in [7] were proposed for MIMO broadcast (MIMO-BC) channels, but they too cannot be directly extended to MIMO ad hoc networks. The authors of [8] introduced a "global gradient projection" (GGP) method, which is an extension of steepest descent method coupled with gradient projection. GGP and IWF can be classified as local optimization techniques, which can quickly find a local optimal solution, but cannot guarantee the global optimality for nonconvex optimization problems.

Our paper aims to address this challenging optimization problem from a global optimization perspective. In particular, this paper proposes a new and powerful global optimization method using a *Branch-and-Bound* framework coupled with the *Reformulation Linearization Technique* (BB/RLT). It has been mathematically shown in [9] that BB/RLT always converges to a global optimal point under very mild assumptions. The main contributions of this paper include the mathematical developments of the solution procedure to solve the problem of finding the maximal sum of mutual information (MSMI) for multiuser MIMO systems based on BB/RLT technique, and related 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. To the best of our knowledge, the proposed method is the first such method that guarantees the finding of a global optimal solution to the MSMI of multiuser MIMO systems.

The remainder of this paper is organized as follows. Section II discusses the network model and problem formulation. Section III introduces the BB/RLT framework and its key problem-specific components, including factorization, linearization, and the derivation of upper and lower bounds for the partitioning variables. Simulation results are presented in Section IV. A convergence speedup technique for the proposed BB/RLT algorithm is presented in Section V. Section VI concludes this paper.

II. Network Model and Problem Formulation

We begin by introducing the 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), whose indices are denoted by $1, 2, \ldots, L$. In this paper, it is assumed that the transmitters have perfect channel state information (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. Particularly, $\mathbf{H}_{ll}$ represents the channel gain matrix from link l's transmitter to its receiver. We consider a constant channel model in this paper, even though wireless channels in reality are time-varying. This simplification is yet of much interest for the insight it provides and its application in finding the ergodic capacity for block-wise fading channels [3].

Let matrix $\mathbf{Q}_l$ represents the covariance matrix of a zero-mean Gaussian input symbol vector $\mathbf{x}_l$ at link l, i.e., $\mathbf{Q}_l = \mathbb{E}\left\{\mathbf{x}_l \cdot \mathbf{x}_l^{\dagger}\right\}$. It is evident that $\mathbf{Q}_l \succeq 0$. Assume, also, that all nodes in the network are subject to the same maximum transmitting power constraint, i.e., $\mathrm{Tr}\left\{\mathbf{Q}_l\right\} \leq P_{\max}$, 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 cause interference to link l. The interference-plus-noise is Gaussian distributed and its covariance matrix can be computed as [8]

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

Hence, the mutual information of a MIMO link l with co-channel interference can be computed as [8]

$$I_l = \log_2 \det\left(\rho_{ll} \mathbf{H}_{ll} \mathbf{Q}_l \mathbf{H}_{ll}^{\dagger} + \mathbf{R}_l\right) - \log_2 \det \mathbf{R}_l. \quad (2)$$

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:

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

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 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 [10]. Thus, due to the considerations of hardware and realistic 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 $\mathbf{p} = (p_1, p_2, \ldots, p_L)^t$, where $0 < p_l \leq P_{\max}$, $l = 1, 2, \ldots, L$, 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:

$$\begin{array}{ll} \max & \sum_{l=1}^{L} I_l \\ \text{s.t.} & I_l = \log_2 \det\left(\frac{\rho_{ll} p_l}{n_r}\left(\mathbf{H}_{ll}\mathbf{H}_{ll}^{\dagger}\right) + \mathbf{R}_l\right) - \log_2 \det \mathbf{R}_l \\ & \mathbf{R}_l = \mathbf{I} + \sum_{j \in \mathcal{I}_l} \frac{\rho_{jl} p_j}{n_r}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^{\dagger}\right) \end{array} \quad (3)$$

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

III. Solution Procedure

A. Overview of BB/RLT Method

It has been shown in [8] that the objective function of MSMI is neither convex nor concave. For such a non-convex optimization problem, conventional nonlinear programming methods can at best yield local optimal solutions. In this paper, we develop a solution procedure based on the *branch-and-bound* framework coupled with the *Reformulation-Linearization Technique* (BB/RLT) [9], [11], [12]. 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. 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. It has been mathematically proven that BB/RLT converges to a global optimal solution under very mild assumptions [9], [11], [12]. The general framework of BB/RLT is shown in Algorithm 1.

Next, we will develop some important problem-specific components in the general RLT-BB framework to make it work for MSMI.

B. Factorization and Linearization

Observe that in the MSMI problem formulation, the link mutual information expressions in (3) are nonlinear. To linearize these nonlinear constraints, we can introduce four new

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 applied to 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 ϵ-optimal solution ψ^*; else, remove all problems $P_{z'}$ having $(1-\epsilon)UB_{z'} \leq LB$ from the list of problems.
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 ϵ-optimal solution ψ^*. Otherwise, repeat the Main Loop.

variables X_l, Y_l, V_l, and W_l as follows.

$$\begin{cases} X_l = \det\left(\frac{\rho_{ll}p_l}{n_r}\left(\mathbf{H}_{ll}\mathbf{H}_{ll}^\dagger\right) + \mathbf{R}_l\right) \triangleq \det \mathbf{D}_l, \\ V_l = \det \mathbf{R}_l, \quad Y_l = \ln X_l, \quad W_l = \ln V_l. \end{cases} \tag{4}$$

The link mutual information constraint in (3) can then be translated into $I_l = \frac{1}{\ln 2}(Y_l - W_l)$, which is a linear constraint. Also, four groups of new constraints in (4) are added to the problem formulation.

C. Linear Relaxation to Nonlinear Logarithmic Functions

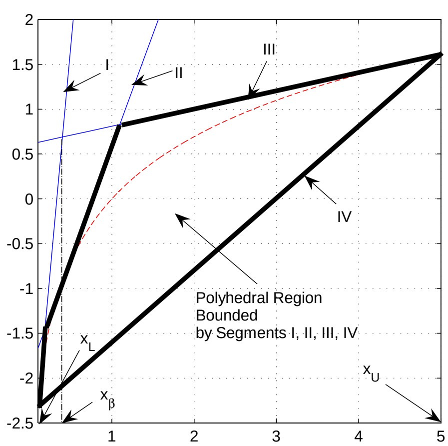

Fig. 1. Polyhedral outer approximation $y = \ln x$.

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:

$$x_\beta = \frac{x_L x_U(\ln x_U - \ln x_L)}{x_U - x_L}. \tag{5}$$

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 which 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:

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

We point out that this polyhedral approximation is a very accurate approximation of $\ln x$. For illustrative purpose, the x_L-value in Fig. 1 is chosen deliberately to be very close to zero to create a significant nonlinearity. Otherwise, segments I, II, III, and IV almost superimpose one another, making the figure hard to discern.

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 consistent form by introducing a notion called the "super interference set" of link l, denoted by $\hat{\mathcal{I}}_l$, where $\hat{\mathcal{I}}_l = \mathcal{I}_l \cup \{l\}$, as follows:

$$\mathbf{D}_l = \textstyle\sum_{j\in\hat{\mathcal{I}}_l} \frac{\rho_{jl}}{n_r}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)p_j + \mathbf{I}, \tag{6}$$

$$\mathbf{R}_l = \textstyle\sum_{j\in\mathcal{I}_l} \frac{\rho_{jl}}{n_r}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)p_j + \mathbf{I}. \tag{7}$$

It is evident from (6) and (7) 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 system where every link has two receiving antennas. This means that $\mathbf{D}_l$ and $\mathbf{R}_l$ are 2×2 square matrices, and their determinants can be computed as in (8) and (9). In (8) and (9), it can be seen that the product terms $p_j p_k$ are the only nonlinear terms, which need to be linearized. To show how the process of linearization works, let us consider a general second-order polynomial term $p_j p_k$, for which we have the following bounding constraints:

$$\begin{aligned} p_j - (p_j)_L \geq 0, \quad (p_j)_U - p_j \geq 0, \\ p_k - (p_k)_L \geq 0, \quad (p_k)_U - p_k \geq 0, \end{aligned} \tag{10}$$

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. From (10), adopting RLT [9], we can derive the following four so-called bounding-factor

$$X_l = 1 + \sum_{j\in\hat{\mathcal{I}}_l} \frac{\rho_{jl}}{n_r}\left[\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(1,1)} + \left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,2)}\right]p_j + \sum_{j\in\hat{\mathcal{I}}_l}\sum_{k\in\hat{\mathcal{I}}_l} \frac{\rho_{jl}\rho_{kl}}{n_r^2}\left[\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(1,1)}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,2)}\right.$$
$$\left. -\mathfrak{Re}\left(\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,1)}\right)\mathfrak{Re}\left(\left(\mathbf{H}_{kl}\mathbf{H}_{kl}^\dagger\right)_{(2,1)}\right) - \mathfrak{Im}\left(\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,1)}\right)\mathfrak{Im}\left(\left(\mathbf{H}_{kl}\mathbf{H}_{kl}^\dagger\right)_{(2,1)}\right)\right] p_j p_k. \quad (8)$$

$$V_l = 1 + \sum_{j\in\mathcal{I}_l} \frac{\rho_{jl}}{n_r}\left[\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(1,1)} + \left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,2)}\right]p_j + \sum_{j\in\mathcal{I}_l}\sum_{k\in\mathcal{I}_l} \frac{\rho_{jl}\rho_{kl}}{n_r^2}\left[\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(1,1)}\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,2)}\right.$$
$$\left. -\mathfrak{Re}\left(\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,1)}\right)\mathfrak{Re}\left(\left(\mathbf{H}_{kl}\mathbf{H}_{kl}^\dagger\right)_{(2,1)}\right) - \mathfrak{Im}\left(\left(\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger\right)_{(2,1)}\right)\mathfrak{Im}\left(\left(\mathbf{H}_{kl}\mathbf{H}_{kl}^\dagger\right)_{(2,1)}\right)\right] p_j p_k. \quad (9)$$

---

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:

$$p_j - (p_j)_L \geq 0 \quad \text{and} \quad (p_j)_U - p_j \geq 0, \qquad (11)$$

we can derive the following three bounding-factor constraints:

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

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, (8) and (9) become linear constraints. The equality relation between $p_j p_k$, p_j and p_k will be replaced by the above bounding-factor constraint relaxations. All these newly introduced bounding-factor constraints will be appended to the original problem, thus achieving a linear programming relaxation for the constraints in the original nonlinear problem.

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

For convenience, we define the right hand sides of (8) and (9) 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 with two transmitting and two receiving antennas (2×2) at each link as follows:

R-MSMI

$$\begin{array}{lll}
\max & \sum_{l=1}^{L} I_l & \\
\text{s.t.} & I_l - \frac{1}{\ln 2}(Y_l - W_l) = 0, \quad \forall l. & (1)\\
& \text{Three tangential supports for } (Y_l, X_l), \quad \forall l. & (2)\\
& \text{Three tangential supports for } (W_l, V_l), \quad \forall l. & (3)\\
& X_l - \Phi_{X_l}(\mathbf{p}) = 1, \quad \forall l. & (4)\\
& V_l - \Phi_{V_l}(\mathbf{p}) = 1, \quad \forall l. & (5)\\
& \text{Bounding constraints for } P_{jk} \text{ and } P_{jj}. & (6)
\end{array}$$

F. Partitioning Variables and Their Upper and Lower Bounds

The partitioning variables in the branch-and-bound (BB) 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 [9], [11], [12] . In R-MSMI, these partitioning variables include X_l, V_l, and p_l, $l = 1, 2, \ldots, L$. For these partitioning variables, we need to derive tight upper and lower bounds. From the definition of MSMI, the upper and lower bounds for p_l are $(p_l)_L = 0$, $(p_l)_U = P_{\max}$, for $l = 1, \ldots, L$. It is evident from these expressions that the upper bounds for X_l and V_l can, respectively, be computed as $\Phi_{X_l}(P_{\max} \times \mathbf{1})$ and $\Phi_{V_l}(P_{\max} \times \mathbf{1})$, where $\mathbf{1}, \mathbf{0} \in \mathbb{R}^L$. It means that all power variables in (8) and (9) take the value $P_{\max}$. On the other hand, noting that the matrix $\mathbf{H}_{jl}\mathbf{H}_{jl}^\dagger$ is PSD and Hermitian, its determinant is greater than or equal to zero. Hence, from (8) and (9), we have lower bounds for X_l and V_l as follows: $(X_l)_L = 1$, and $(V_l)_L = 1$.

IV. Numerical Results

We first describe our simulation settings. A MIMO ad hoc network with L links (transmission pairs) is generated within a square region, whose edge-lengths can be varied. The smaller the area, the more severe interference these links would cause to each other. The L links are uniformly distributed within the square region. Each node in the network is equipped with two transmitting antennas and two receiving antennas. The maximum transmit power for each node is set to $P_{\max} =$ 10dBm. The network operates in 2.4G ISM band. The channel bandwidth is $W = 30$MHz. The path-loss index is chosen to be $\alpha = 2$.

A. Convergence Properties

We use a 5-link network example to demonstrate the convergence properties of BB/RLT. The desired error bound is chosen to be $\epsilon = 0.01$. The network's SNR- and INR-values are shown in Table I. The convergence process is depicted in Fig. 2, where the UB and LB in terms of the sum of mutual information (b/s/Hz) are illustrated at each iteration. It can be seen that, after 17500 iterations, the UB and LB values are both driven to 26.51 b/s/Hz, which means that BB/RLT found a global optimum for the sum of mutual information to be 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). It can also be seen from Fig. 2 that the rate of decrease in UB plays the major role in determining how fast the overall BB/RLT process converges. Moreover, the rate of decline of UB becomes slower as it approaches the global optimal solution value. We will specifically address this convergence speed issue in the next section.

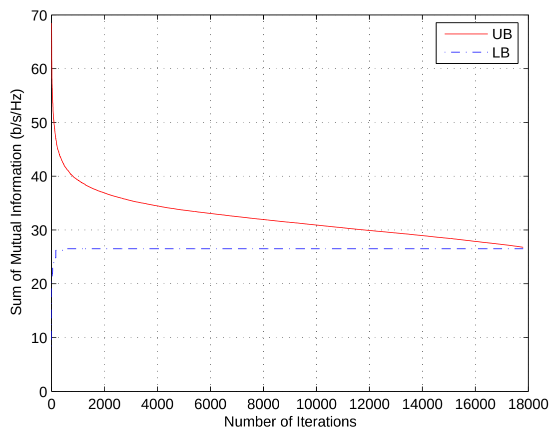

Fig. 2. A 5-link network example.

TABLE I

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

	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	–

V. CONVERGENCE SPEEDUP TECHNIQUES

From the previous section, we can see from Fig. 2 that, although BB/RLT converges quite fast for small-size networks, the decreasing of UB, which plays the major role in affecting the overall convergence speed, exhibits a gradual asymptotic decay as it approaches the global optimal solution. This is because BB is a global optimization technique that is of exponential complexity in general. The performance and convergence speed is always a tradeoff when it comes to choosing global or local optimization techniques [13]. Nonetheless, we can adopt some problem specific strategies that dramatically improve the convergence speed of BB/RLT. Recall that in MSMI, there are three groups of BB variables: X_l, V_l, and p_l. In the original BB/RLT proposed in [9], [11], [12], certain generic discrepancy-based rules are prescribed for selecting the partitioning variables for branching at each stage, which can significantly affect computational time. Upon carefully analyzing X_l, V_l, and p_l for our specific problem, we find that they play very different roles. The relationships among these variables can be visualized in Fig. 3, where the numbers on the link correspond to the constraint number in R-MSMI. For example, an arrow with "(1), (2)" from X_l to I_l means that variables X_l have a direct impact on I_l through constraint (1) and (2) in R-MSMI. From our discussions in the previous section, we know that fast decline of UB plays a major role in accelerating the whole convergence process. Based on this, we conclude that, in order to have a faster convergence speed, X_l and V_l should be of higher priority in partitioning because changing X_l and V_l have a more direct impact on UB. Hence, we adopt the partitioning variable selection strategy described in Algorithm 2.

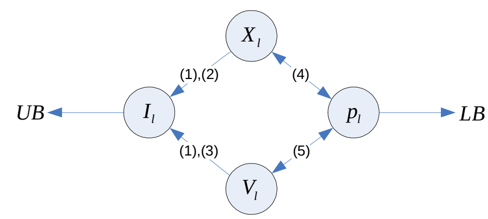

Fig. 3. Relationships among branch-and-bound variables in MSMI

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 $(\ln(Z_l^*)_U - \ln(Z_l^*)_L$ is small enough) then
   a) Among all p_l, choose one, say p_l^*, with the largest relaxation error, denoted as E_p
   b) If E_p is small enough, then skip this subproblem; else return p_l^*;

   else return Z_l^*.

We always select a partitioning variable among the X_l- and V_l-variables first, choosing one that has the largest relaxation error and denoting this variable as Z_l^*. If we find that $\ln(Z_l^*)_U - \ln(Z_l^*)_L$ is already small enough, we must have that the change in $\ln Z_l^*$ by partitioning on Z_l^* would be even smaller than $\ln(Z_l^*)_U - \ln(Z_l^*)_L$ because $\ln\{\cdot\}$ is a monotone function. In addition, since we know that UB is only dependent on the summation of the logarithms of the X_l- and Y_l-variables, we must have that further partitioning on Z_l^* will no longer have much reduction on UB. Therefore, even though the bounding interval of Z_l^* in this subproblem may still be large, we can stop partitioning on Z_l^*. This saves us a huge amount of computation. Next, we can instead switch to partitioning on p_l for this particular subproblem. Partitioning on p_l only indirectly affects UB by tightening the resulting relaxation, but it can help increase LB, which is also beneficial to the convergence process. If, at certain point of the branch-and-bound process, we find that the relaxation error for p_l is small enough, this subproblem needs no further partitioning.

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 II. The convergence process is depicted in Fig. 4. The global optimal value for this 10-link network is

TABLE II

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	–

77.11 b/s/Hz. It takes approximately 1.5×10^6 iterations to converge to the global optimal point after using the modified partitioning variable selection strategy. Although the number of iterations for this 10-link example is seemingly quite large, the BB/RLT procedure manages to take reasonable amount of time to converge, thanks to the availability of fast and robust LP solvers. On the other hand, if we solve the same 10-link example using BB/RLT without using Algorithm 2, the computation time takes so long that the problem basically becomes unsolvable. We roughly estimated that our proposed speedup technique can accelerate the convergence by over 1000 times for this particular example.

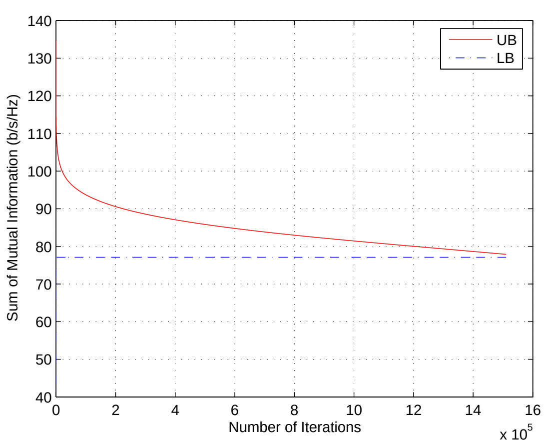

Fig. 4. A 10-link network example.

VI. CONCLUSIONS

In this paper, we considered the power control problem of finding the maximum sum of mutual information for multiuser MIMO systems with equal power allocation for each link. A new and powerful global optimization method using a branch-and-bound framework coupled with the reformulation-linearization technique (BB/RLT) was introduced. Numerical results and detailed discussions on convergence properties using several MIMO systems were provided. This BB/RLT guarantees finding a global optimal solution for multiuser MIMO ad hoc networks. We also proposed a modified branch-and-bound variable selection strategy to accelerate the convergence process, an demonstrated its efficacy.

ACKNOWLEDGEMENT

The work of Y.T. Hou has been supported in part by NSF Grant CNS-0347390 and ONR Grant N00014-05-1-0481. The work of H.D. Sherali has been supported in part by NSF Grant 0552676.

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 Personal Commun.*, vol. 6, pp. 311–355, Mar. 1998.

[3] 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.

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

[5] 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.

[6] 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.

[7] 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. 2004.

[8] 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.

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

[10] 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.

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

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

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.