Optimal Power Allocation for Achieving Perfect Secrecy Capacity in MIMO Wire-Tap Channels

Jia Liu*, Y. Thomas Hou*, and Hanif D. Sherali†
* Bradley Department of Electrical and Computer Engineering
† Grado Department of Industrial and Systems Engineering
Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

***Abstract*—In this paper, we investigate optimal power allocation to achieve perfect secrecy capacity in Gaussian MIMO wire-tap channels. The number of antennas in the MIMO wire-tap channel is arbitrary at the transmitter, the intended receiver, and the eavesdropper. For this challenging non-convex optimization problem, we design a novel global optimization algorithm called branch-and-bound with reformulation and linearization technique (BB/RLT). As opposed to convex programming methods that only yield local optimal solutions, our proposed BB/RLT method guarantees finding a global optimal solution. The main contribution in this paper is that our proposed BB/RLT algorithm is the first method that solves the optimal power allocation problem for achieving perfect secrecy capacity problem in MIMO wire-tap channels. Numerical examples are also given to demonstrate the efficacy of the proposed algorithm.**

I. Introduction

Security and privacy protection is one of the most important issues in wireless communications due to the broadcast nature of wireless channels – anyone within communication range can listen to the transmission in the air and could possibly extract the information. Aside from traditional cryptographic security mechanisms, information-theoretic-based security techniques have gained increasing attention in recent years.

The present information-theoretic security framework was established by Wyner [1], who developed the concept of "wire-tap channel." Wyner showed that, if a wire-tap channel satisfies certain degradedness conditions, perfectly secure communication with non-zero rate between the transmitter and the intended receiver is achievable. Meanwhile, the eavesdropper learns nothing about the secret messages from its observations. The maximum rate of secrecy information from the transmitter to the intended receiver is called *secrecy capacity*. A follow-up work by Leung-Yan-Cheong and Hellman further determined the secrecy capacity of scalar Gaussian wire-tap channel [2]. Wyner's work was later extended by Csiszár and Köner to the case of secret communication over general broadcast channels [3] and scalar Gaussian multiple access channels [4]–[6].

Recently, researchers began to consider exploiting the physical layer characteristics of multiple-input multiple-output (MIMO) channels to further increase the capacity of secrecy communications. The secrecy communication problem for MIMO systems was first studied in [7], where Hero showed that under the restricted case where the eavesdropper is uninformed about its channel, the transmitter can enforce a zero information rate to the eavesdropper while delivering a positive information rate to the intended receiver. In [8], Negi *et al.* proposed a scheme to transmit artificial noise in the intended receiver's null space so that the eavesdropper's channel is degraded. The optimal power allocation strategy for Gaussian multiple-input single-output (MISO) wire-tap channels is studied in [9] and [10]. By somewhat different approaches, both [9] and [10] showed the same result that, under the relatively simpler MISO setting, the optimization problem can be transformed into the well-known Rayleigh quotient problem and thus can be solved analytically.

In a very recent paper, Oggier *et al.* [11] showed that the perfect secrecy capacity of general MIMO wire-tap channels is achieved by Gaussian signaling, thus showing the secrecy capacity inner bound in [9] is tight. However, finding the optimal power allocation under Gaussian signalling for MIMO wire-tap channels is very challenging and remains an open problem. This is because the power allocation problem for MIMO wire-tap channels is non-convex and, unlike simple MISO cases [9], [10], there does not exist a simple way to get around the complex problem structure. In this paper, our major goal is to fill this void and design an algorithm that guarantees finding an optimal solution for such a non-convex optimization problem.

The main contribution of this paper is that we provide an effective solution method to solve the optimal power allocation problem for achieving perfect secrecy capacity in MIMO wire-tap channels. To solve the power allocation problem, we propose a global optimization technique called *branch-and-bound with reformulation linearization technique* (BB/RLT). The basic idea of BB/RLT is that we first use RLT to obtain a linear relaxation for the original problem. By solving the linear relaxation, a global upper bound for the original problem is achieved. We then use this relaxation solution as a starting point to search for a solution that is feasible to the original problem. This feasible solution serves as a global lower bound and an incumbent optimal solution. The branch-and-bound process will then tighten the global upper bound and global lower bound during each iteration, and stop when the gap between global upper and lower bounds is sufficiently small. In this paper, we develop the key problem-specific components in BB/RLT to solve the problem and the related convergence speedup techniques. Specifically, we develop a linearization scheme to generate a higher dimensional upper-bounding problem. We also utilize a polyhedral outer approximation method

978-1-4244-2734-5/09/$25.00 ©2009 IEEE

Authorized licensed use limited to: to IEEExplore provided by Virginia Tech Libraries. Downloaded on June 19, 2009 at 13:23 from IEEE Xplore. Restrictions apply.

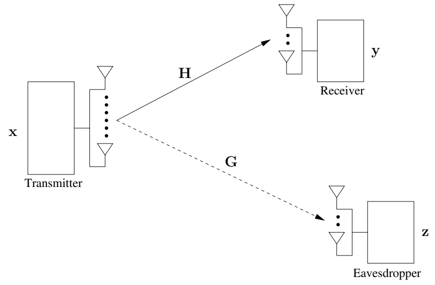

Fig. 1. A MIMO wire-tap channel consists of a transmitter, a receiver, and an eavesdropper, each equipped with multiple antennas.

to accurately approximate the logarithmic function. To speed up the branch-and-bound convergence process, 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, our proposed method is the first one that guarantees finding a global optimal solution for the optimal power allocation problem in MIMO wire-tap channels.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and problem formulation. Section III introduces our proposed BB/RLT solution procedure. A convergence speedup technique for the proposed BB/RLT algorithm is presented in Section IV and a numerical example is given in Section V. Section VI concludes this paper.

II. System Model and Problem Formulation

As shown in Fig. 1, in a Gaussian MIMO wire-tap channel, a transmitter equipped with n_t antennas wishes to reliably communicate a secret message to an intended receiver in the presence of an eavesdropper. The intended receiver has n_r antennas and the eavesdropper has n_e antennas.

A secret message S of rate R is a random integer from the set $\{1, 2, \ldots, 2^{nR}\}$, which is transmitted in n channel uses. The transmitter transmits the coded signal $\mathbf{x} \in \mathbb{C}^{n_t}$ to the receiver, who decodes S based on the output $\mathbf{y} \in \mathbb{C}^{n_r}$. The eavesdropper overhears the output $\mathbf{z} \in \mathbb{C}^{n_e}$. Define the equivocation rate Δ as $\Delta = H(S|\mathbf{z})/H(S)$, i.e., the ratio between the conditional entropy conditioned on the overheard signal and the unconditional entropy. The equivocation rate is a measure of the amount of information the eavesdropper can learn about the message and quantifies the secrecy level [3]. In this paper, we focus on the case $\Delta = 1$, i.e., the eavesdropper learns arbitrary little information regarding message S.

The received signals at the receiver and the eavesdropper can be respectively written as

$$\mathbf{y} = \sqrt{\rho_{\mathbf{H}}}\mathbf{H}\mathbf{x} + \mathbf{n}_1, \quad \mathbf{z} = \sqrt{\rho_{\mathbf{G}}}\mathbf{G}\mathbf{x} + \mathbf{n}_2, \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ and $\mathbf{G} \in \mathbb{C}^{n_e \times n_t}$ are the channel gain matrices from the transmitter to the intended receiver and to the eavesdropper, respectively, $\rho_{\mathbf{H}}$ and $\rho_{\mathbf{G}}$ are the corresponding signal-to-noise ratios (SNR), respectively, and $\mathbf{n}_1 \in \mathbb{C}^{n_r}$, $\mathbf{n}_2 \in \mathbf{C}^{n_e}$ are the corresponding normalized complex circularly symmetric Gaussian random noise terms with identity covariance matrix. In this paper, it is assumed that the transmitter has full knowledge of both channels.

Denote $\mathbf{Q}$ the covariance matrix of the input signal, i.e., $\mathbf{Q} = E[\mathbf{x}\mathbf{x}^{\dagger}]$. Then, the power constraint on $\mathbf{Q}$ is $\mathbf{Q} \succeq 0$ and $\mathrm{Tr}(\mathbf{Q}) \leq 1$. In a very recent paper [11], Oggier *et al.* showed that the perfect secrecy capacity of general MIMO wire-tap channels is achieved by Gaussian signaling. As a result, the perfect secrecy capacity problem of MIMO wire-tap channels can be written as follows:

$$\begin{array}{ll} \text{Maximize} & \log|\mathbf{I} + \rho_{\mathbf{H}}\mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}| - \log|\mathbf{I} + \rho_{\mathbf{G}}\mathbf{G}\mathbf{Q}\mathbf{G}^{\dagger}| \\ \text{subject to} & \mathrm{Tr}(\mathbf{Q}) \leq 1, \ \mathbf{Q} \succeq 0, \ \mathbf{Q} = \mathbf{Q}^{\dagger}. \end{array} \tag{2}$$

Clearly, problem (2) is a challenging non-convex optimization problem and many local maxima may exist. This can be seen by noticing that the objective function is a difference of two concave functions (in the the form of $\log|\cdot|$). As a result, the convexity is not definite. For such a non-convex optimization problem, conventional convex optimization methods do not work well since they yield local optimal solutions at best. In the next section, we will develop an algorithm based on global optimization approach.

III. A Global Optimization Approach

We develop our algorithm by using a state-of-the-art global optimization technique called branch and bound with reformulation-linearization technique (BB/RLT) [12]. The basic idea of BB/RLT is rather simple: First, by using RLT, we obtain a linear relaxation for the original problem. By solving the linear relaxation, we have a relaxation solution that provides a global upper bound UB for the original problem. If this relaxation solution is infeasible to the original problem, we can use it as a starting point to do a local search to determine a feasible solution. Then, this feasible solution will serve as a global lower bound LB and an incumbent optimal solution to the original problem. The branch-and-bound process will then tighten UB and LB during each iteration, and stop when $LB \geq (1 - \epsilon)UB$ is satisfied.

The motivation of using BB/RLT is that BB/RLT guarantees the convergence to a global optimal solution for any non-convex optimization problem as long as the feasible region of the problem is compact (see [12] for further details). This condition is clearly satisfied for problem (2). We summarize the general framework of BB/RLT in Algorithm 1.

We note that the BB/RLT method in Algorithm 1 is a general framework. In order to use Algorithm 1 to solve problem (2), several key problem-specific components remain to be developed and the design is far from trivial. First and foremost, how to construct an RLT linear relaxation for problem (2) is critical to the success of using BB/RLT. Also, the branch-and-bound strategy and the according convergence speed up technique is also very important. In what follows, we will describe in detail how to develop these problem-specific components.

A. Objective Function Linearization

In problem (2), the objective function is nonlinear. To linearize it, we introduce two new variables $A = \ln|\mathbf{I} +$

Authorized licensed use limited to: to IEEExplore provided by Virginia Tech Libraries. Downloaded on June 19, 2009 at 13:23 from IEEE Xplore. Restrictions apply.

Algorithm 1 BB/RLT Algorithmic Framework

```
Initialization:
  1. Let optimal solution $\psi^* = \emptyset$. The initial lower bound $LB = -\infty$.
  2. Determine partition variables (variables associated with nonlinear
     terms) and derive their initial value intervals
  3. Let the initial problem list contains 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.
  5. 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 $\psi_z$ via a local search algorithm
     based on the solution of problem $P_z$. Denote the objective value of
     $\psi_z$ 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 $\epsilon$-optimal solution $\psi^*$;
       else remove all problems $P_{z'}$ with $(1-\epsilon)UB_{z'} \leq LB$.
     endif
  4. Compute relaxation error for each nonlinear term.
  5. Select a partition variable with the maximum relaxation error and
     divide its interval into two new intervals at the point $\hat{\psi}_z$
  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, we stop with the $\epsilon$-optimal solution $\psi^*$.
     Otherwise, repeat step 1.
```

$\rho_{\mathbf{H}}\mathbf{HQH}^\dagger|$ and $B = \ln|\mathbf{I} + \rho_{\mathbf{G}}\mathbf{GQG}^\dagger|$. The objective function can then be linearized to

$$\text{Maximize } \frac{1}{\ln 2}(A - B)\,.$$

For convenience, we let $\mathbf{D} \triangleq \mathbf{I} + \rho_{\mathbf{H}}\mathbf{HQH}^\dagger$ and $\mathbf{R} \triangleq \mathbf{I} + \rho_{\mathbf{G}}\mathbf{GQG}^\dagger$. It is clear that, to evaluate A and B, we need to compute $\ln|\mathbf{D}|$ and $\ln|\mathbf{R}|$.

Generally, determinant computation is cumbersome because a matrix's determinant is in essence a high-order polynomial of all matrix entries. Fortunately, in problem (2), there exist special structures in $\mathbf{D}$ and $\mathbf{R}$ to make determinant computation easier. Noting that $\mathbf{D}$ and $\mathbf{R}$ are positive definite Hermitian matrices, Cholesky decompositions of $\mathbf{D}$ and $\mathbf{R}$ exist. Take $\mathbf{D}$ for example, we have

$$\begin{aligned} A &= \ln|\mathbf{D}| = \ln|\mathbf{D}^{\langle L\rangle}\cdot(\mathbf{D}^{\langle L\rangle})^\dagger| \\ &= \ln(\textstyle\prod_{i=1}^{n_r}\|d_{(i,i)}^{\langle L\rangle}\|^2) = \sum_{i=1}^{n_r}\ln\|d_{(i,i)}^{\langle L\rangle}\|^2 \\ &= \textstyle\sum_{i=1}^{n_r}\left[\ln(\mathfrak{Re}^2(d_{(i,i)}^{\langle L\rangle}) + \mathfrak{Im}^2(d_{(i,i)}^{\langle L\rangle}))\right], \end{aligned}$$

where $\mathbf{D}^{\langle L\rangle}$ is a lower triangular matrix and $d_{(i,i)}^{\langle L\rangle}$ denotes the ith diagonal entry of $\mathbf{D}^{\langle L\rangle}$. Noting that A is a summation of the log values of the square of $d_{(i,i)}^{\langle L\rangle}$'s norm, we can introduce two new groups of variables Y_i and X_i to perform reformulation as follows:

$$\begin{aligned} A &= \textstyle\sum_{i=1}^{n_r} Y_i, \\ Y_i &= \ln X_i, \\ X_i &= \mathfrak{Re}^2(d_{(i,i)}^{\langle L\rangle}) + \mathfrak{Im}^2(d_{(i,i)}^{\langle L\rangle}). \end{aligned}$$

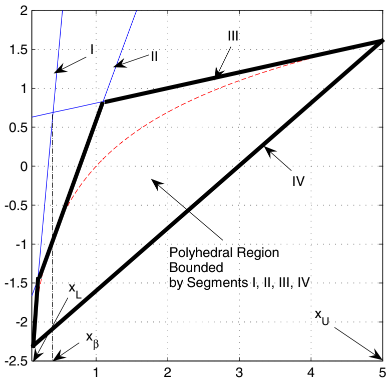

Fig. 2. A convex envelope for $y = \ln x$.

Likewise, after introducing two more groups of variables V_i and W_i, B can be computed using the same method.

Another difficulty in linearizing the objective function lies in handling $\ln\{\cdot\}$. Our basic idea to linearize the curve of the log function is to construct a polyhedral outer approximation. As shown in Fig. 2, the curve $y = \ln x$ over an interval can be relaxed to a *polyhedron* formed 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:

$$x_\beta = \frac{x_L x_U(\ln x_U - \ln x_L)}{x_U - x_L}. \tag{3}$$

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 polyhedron defined by the four line segments can be described by the following four linear constraints:

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

We remark that this polyhedral approximation is a very accurate approximation of the log function. For illustrative purpose, the x_L-value in Fig. 2 is deliberately chosen to be close to zero to generate a significant curvature. Otherwise, segments I, II, III, and IV almost superimpose one another, making the figure hard to discern.

B. Linear Relaxation of Chelosky Decomposition

As indicated earlier, $\mathbf{D}$ and $\mathbf{R}$ can be decomposed into a product form by Cholesky decomposition. Now, it remains to linearize the product terms in the Cholesky decompositions. Take $\mathbf{D}$ for example, we have the following relationship between the entries of $\mathbf{D}$ and the entries of $\mathbf{D}^{\langle L\rangle}$:

$$d_{(i,j)} = \sum_{k=1}^{j} d_{(i,k)}^{\langle L\rangle}\overline{d}_{(j,k)}^{\langle L\rangle} \tag{4}$$

978-1-4244-2734-5/09/$25.00 ©2009 IEEE

Authorized licensed use limited to: to IEEExplore provided by Virginia Tech Libraries. Downloaded on June 19, 2009 at 13:23 from IEEE Xplore. Restrictions apply.

for $1 \leq i \leq n_t$, $1 \leq j \leq i$, and $1 \leq k \leq j$. Expanding the RHS of (4), we have

$$\mathfrak{Re}(d_{(i,j)}) = \sum_{k=1}^{j} \left[\mathfrak{Re}(d_{(i,k)}^{\langle L \rangle}) \mathfrak{Re}(d_{(j,k)}^{\langle L \rangle}) + \mathfrak{Im}(d_{(i,k)}^{\langle L \rangle}) \mathfrak{Im}(d_{(j,k)}^{\langle L \rangle}) \right]$$

$$\mathfrak{Im}(d_{(i,j)}) = \sum_{k=1}^{j} \left[\mathfrak{Im}(d_{(i,k)}^{\langle L \rangle}) \mathfrak{Re}(d_{(j,k)}^{\langle L \rangle}) - \mathfrak{Re}(d_{(i,k)}^{\langle L \rangle}) \mathfrak{Im}(d_{(j,k)}^{\langle L \rangle}) \right]$$

For the diagonal entries of $\mathbf{D}$, we can further simplify the expansions to

$$\mathfrak{Re}(d_{(i,i)}) = \sum_{k=1}^{i} \left[\mathfrak{Re}^2(d_{(i,k)}^{\langle L \rangle}) + \mathfrak{Im}^2(d_{(i,k)}^{\langle L \rangle}) \right], \ \mathfrak{Im}(d_{(i,i)}) = 0.$$

Now, we introduce the so-called RLT variables to linearize these expansion expressions. For convenience, we use a generic term (xy) to represent the terms $\mathfrak{Re}(d_{(i,k)}^{\langle L \rangle})\mathfrak{Re}(d_{(j,k)}^{\langle L \rangle})$, $\mathfrak{Im}(d_{(i,k)}^{\langle L \rangle})\mathfrak{Im}(d_{(j,k)}^{\langle L \rangle})$, $\mathfrak{Im}(d_{(i,k)}^{\langle L \rangle})\mathfrak{Re}(d_{(j,k)}^{\langle L \rangle})$, and $\mathfrak{Re}(d_{(i,k)}^{\langle L \rangle})\mathfrak{Im}(d_{(j,k)}^{\langle L \rangle})$. Also, we use a generic term x^2 to represent the terms $\mathfrak{Re}^2(d_{(i,k)}^{\langle L \rangle})$ and $\mathfrak{Im}^2(d_{(i,k)}^{\langle L \rangle})$.

For a product term xy, we have

$$x - x_L \geq 0, \ x_U - x \geq 0, \ y - y_L \geq 0, \ y_U - y \geq 0, \quad (5)$$

where x_L and y_L denote the lower bounds on x and y, respectively, and x_U and y_U denote the upper bounds on x and y, respectively. From (5), we can derive the following four bounding factor constraints:

$$\begin{array}{ll} (x - x_L) \times (y - y_L) \geq 0, & (x - x_L) \times (y_U - y) \geq 0, \\ (x_U - x) \times (y - y_L) \geq 0, & (x_U - x) \times (y_U - y) \geq 0, \end{array}$$

which can be further expanded to the following four constraints:

$$\begin{array}{c} xy - y_L x - x_L y \geq -x_L y_L, \ -xy + y_U x + x_L y \geq x_L y_U, \\ -xy + y_L x + x_U y \geq x_U y_L, \ xy - y_U x - x_U y \geq -x_U y_U. \end{array} \quad (6)$$

Similarly, for a general square term x^2, we have

$$x - x_L \geq 0 \quad x_U - x \geq 0 \quad (7)$$

From (7), we can derive the following three bounding constraints:

$$(x - x_L)^2 \geq 0, \ (x_U - x)^2 \geq 0, \ (x - x_L) \times (x_U - x) \geq 0$$

which can be expanded to the following three constraints:

$$\begin{array}{c} x^2 - 2x_L x \geq -x_L^2, \quad x^2 - 2x_U x \geq -x_U^2, \\ x^2 - (x_L + x_U)x \leq x_U y_L. \end{array} \quad (8)$$

Now, by treating xy and x^2 as independent variables rather than the products of two terms, all constraints in (6) and (8) become linear constraints.

C. Entrywise Expansions of $\mathbf{D}$ *and* $\mathbf{R}$ *with Respect to* $\mathbf{Q}$

So far, $\mathbf{D}$ and $\mathbf{R}$ are expressed using matrix products of $\mathbf{Q}$, $\mathbf{H}$, and $\mathbf{G}$. However, such matrix product expressions are not convenient to use under BB/RLT framework. Thus, it is necessary to expand these matrices products entry-wise for algorithm design. For example, by expanding $\mathbf{D} = \mathbf{I} + \rho_{\mathbf{H}} \mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}$, the real and imaginary parts of $d_{(i,j)}$, $i \neq j$, can be written as the following linear equalities with respect to $\mathfrak{Re}(q_{(x,y)})$ and $\mathfrak{Im}(q_{(x,y)})$:

$$\begin{aligned} 0 = & \mathfrak{Re}(d_{(i,j)}) - \rho_{\mathbf{H}} \sum_{x=1}^{n_t} \mathfrak{Re}(q_{(x,x)}) \mathfrak{Re}(h_{(i,x)} \overline{h}_{(j,x)}) \\ & -\rho_{\mathbf{H}} \sum_{x=2}^{n_t} \sum_{y=1}^{x-1} \mathfrak{Re}(q_{(x,y)}) \left[\mathfrak{Re}(h_{(i,x)} \overline{h}_{(j,y)}) + \mathfrak{Re}(h_{(i,y)} \overline{h}_{(j,x)}) \right] \\ & -\rho_{\mathbf{H}} \sum_{x=2}^{n_t} \sum_{y=1}^{x-1} \mathfrak{Im}(q_{(x,y)}) \left[\mathfrak{Im}(h_{(i,y)} \overline{h}_{(j,x)}) - \mathfrak{Im}(h_{(i,x)} \overline{h}_{(j,y)}) \right] \end{aligned}$$

$$\begin{aligned} 0 = & \mathfrak{Im}(d_{(i,j)}) - \rho_{\mathbf{H}} \sum_{x=1}^{n_t} \mathfrak{Re}(q_{(x,x)}) \mathfrak{Im}(h_{(i,x)} \overline{h}_{(j,x)}) \\ & -\rho_{\mathbf{H}} \sum_{x=2}^{n_t} \sum_{y=1}^{x-1} \mathfrak{Re}(q_{(x,y)}) \left[\mathfrak{Im}(h_{(i,y)} \overline{h}_{(j,x)}) + \mathfrak{Im}(h_{(i,x)} \overline{h}_{(j,y)}) \right] \\ & -\rho_{\mathbf{H}} \sum_{x=2}^{n_t} \sum_{y=1}^{x-1} \mathfrak{Im}(q_{(x,y)}) \left[\mathfrak{Re}(h_{(i,x)} \overline{h}_{(j,y)}) - \mathfrak{Re}(h_{(i,y)} \overline{h}_{(j,x)}) \right] \end{aligned}$$

For the diagonal elements $d_{(i,i)}$, the expansion can be simplified to

$$\begin{aligned} \mathfrak{Re}(d_{(i,i)}) = & \rho_{\mathbf{H}} \sum_{x=1}^{n_t} \mathfrak{Re}(q_{(x,x)}) \left\| h_{(i,x)} \right\|^2 \\ & +\rho_{\mathbf{H}} \sum_{x=2}^{n_t} \sum_{y=1}^{x-1} \mathfrak{Re}(q_{(x,y)}) \cdot 2\mathfrak{Re}(h_{(i,x)} \overline{h}_{(i,y)}) \\ & +\rho_{\mathbf{H}} \sum_{x=2}^{n_t} \sum_{y=1}^{x-1} \mathfrak{Im}(q_{(x,y)}) \cdot 2\mathfrak{Im}(h_{(i,y)} \overline{h}_{(i,x)}) + 1 \end{aligned} \quad (9)$$

and $\mathfrak{Im}(d_{(i,i)}) = 0$. Since $\mathfrak{Im}(d_{(i,i)}) = 0$, $\forall i$, we can remove these variables from our RLT-based relaxation to reduce the number of variables and constraints without changing the problem. Following the same line, we can derive the expansion of $\mathbf{R} = \mathbf{I} + \rho_{\mathbf{G}} \mathbf{G}\mathbf{Q}\mathbf{G}^{\dagger}$ and the expressions are omitted here for brevity.

D. Linear Relaxation for Power Covariance Matrix $\mathbf{Q}$

In problem (2), the input covariance matrix $\mathbf{Q}$ is subject to $\mathbf{Q} \succeq 0$ and $\mathrm{Tr}(\mathbf{Q}) \leq 1$. The trace constraint on $\mathbf{Q}$ is simply the following linear constraint:

$$\sum_{i=1}^{n_t} \mathfrak{Re}(q_{(i,i)}) \leq 1. \quad (10)$$

However, it is more involved to express the PSD constraint of $\mathbf{Q}$ in a linear form. In what follows, we will uncover the hidden

Authorized licensed use limited to: to IEEExplore provided by Virginia Tech Libraries. Downloaded on June 19, 2009 at 13:23 from IEEE Xplore. Restrictions apply.

entrywise relationship implied by $\mathbf{Q} \succeq 0$. We first introduce the following result, for which the proof can be found in [13].

Theorem 1. *For any Hermitian PSD matrix $\mathbf{Q} \in \mathbb{C}^{n\times n}$, the entry having the largest modulus must appear on the main diagonal of $\mathbf{Q}$. Further, the entries in $\mathbf{Q}$ satisfy the following inequality:*

$$\|q_{(i,j)}\| \leq \frac{q_{(i,i)} + q_{(j,j)}}{2}, \quad 1 \leq i, j \leq n. \tag{11}$$

Since $|\mathfrak{Re}(q_{(i,j)})| \leq \|q_{(i,j)}\|$ and $|\mathfrak{Im}(q_{(i,j)})| \leq \|q_{(i,j)}\|$, by Theorem 1, we have the following relaxation constraints on the entries of $\mathbf{Q}$:

$$\begin{aligned} |\mathfrak{Re}(q_{(i,j)})| &\leq \frac{1}{2}(\mathfrak{Re}(q_{(i,i)}) + \mathfrak{Re}(q_{(j,j)})), \quad \forall i, &(12)\\ |\mathfrak{Im}(q_{(i,j)})| &\leq \frac{1}{2}(\mathfrak{Re}(q_{(i,i)}) + \mathfrak{Re}(q_{(j,j)})), \quad \forall i. &(13)\end{aligned}$$

E. RLT-Based Linear Relaxation

By putting together all new variables and constraints, we have the final RLT linear relaxation of problem (2) as follows:

$$\begin{array}{rl} \text{Max} & \frac{1}{\ln 2}(A - B) \\ \text{s.t.} & A - \sum_{i=1}^{n_r} Y_i = 0 \\ & B - \sum_{i=1}^{n_r} W_i = 0 \\ & \text{Polyhedral approximation for } (Y_i, X_i) \\ & \text{Polyhedral approximation for } (W_i, V_i) \\ & X_i - \mathfrak{Re}^2(d^{\langle L\rangle}_{(i,i)}) - \mathfrak{Im}^2(d^{\langle L\rangle}_{(i,i)}) = 0, \forall i \\ & V_i - \mathfrak{Re}^2(r^{\langle L\rangle}_{(i,i)}) - \mathfrak{Im}^2(r^{\langle L\rangle}_{(i,i)}) = 0, \forall i \\ & \text{Bounding factor constraints for terms in the forms} \\ & \text{of } \mathfrak{Re}(\cdot)\mathfrak{Re}(\cdot),\ \mathfrak{Im}(\cdot)\mathfrak{Im}(\cdot), \text{ and } \mathfrak{Re}(\cdot)\mathfrak{Im}(\cdot). \\ & \text{Entrywise expanssion for } \mathbf{D} \text{ and } \mathbf{R} \text{ w.r.t. } \mathbf{Q} \\ & \sum_{i=1}^{n_t} \mathfrak{Re}(q_{(i,i)}) \leq 1 \\ & |\mathfrak{Re}(q_{(i,j)})| \leq \frac{1}{2}(\mathfrak{Re}(q_{(i,i)}) + \mathfrak{Re}(q_{(j,j)})), \forall i \\ & |\mathfrak{Im}(q_{(i,j)})| \leq \frac{1}{2}(\mathfrak{Re}(q_{(i,i)}) + \mathfrak{Re}(q_{(j,j)})), \forall i. \end{array} \tag{14}$$

F. Partitioning Variables and Their Bounds

In BB/RLT, partitioning variables 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 BB process [12]. In RLT relaxation (14), these BB variables include X_i, V_i, $\mathfrak{Re}(d^{\langle L\rangle}_{(i,j)})$, $\mathfrak{Im}(d^{\langle L\rangle}_{(i,j)})$, $\mathfrak{Re}(r^{\langle L\rangle}_{(i,j)})$, and $\mathfrak{Im}(r^{\langle L\rangle}_{(i,j)})$. For these variables, we need to derive tight upper and lower bounds, which are crucial to the convergence speed of BB/RLT.

1) $\mathfrak{Re}(d^{\langle L\rangle}_{(i,k)})$, and $\mathfrak{Im}(d^{\langle L\rangle}_{(i,k)})$: To derive upper and lower bounds for $\mathfrak{Re}(d^{\langle L\rangle}_{(i,k)})$, and $\mathfrak{Im}(d^{\langle L\rangle}_{(i,k)})$, we need the following result.

Lemma 1. *For any PSD Hermitian matrix $\mathbf{M} \in \mathbb{C}^{n\times n}$, the summation of the modulus of all off-diagonal elements satisfy the following inequality:*

$$\sum_{i=2}^{n}\sum_{j=1}^{i-1} \|m_{(i,j)}\| \leq \frac{n-1}{2}\mathrm{Tr}\{\mathbf{M}\}.$$

Lemma 1 follows directly from Theorem 1. Due to space limitation, we refer readers to [13] for more details. From Lemma 1, we have the following theorem and the proof is also relegated to [13] due to space limitation.

Theorem 2. *The mth diagonal entry of $\mathbf{D}$ can be upper bounded by*

$$\left(d_{(i,i)}\right)_U = \left[\rho_{\mathbf{H}}\left(D^{(1)*}_{(i)} + D^{(2)*}_{(i)}\right)\right] + 1, \tag{15}$$

where $D^{(1)*}_{(i)} = \max\left\|h_{(i,x)}\right\|^2$, for $1 \leq i \leq n_r$, $1 \leq x \leq n_t$, and $D^{(2)*}_{(i)}$ is the optimal objective value of the following linear programming problem:

$$\begin{array}{rl} \max & \sum_{x=2}^{n_t}\sum_{y=1}^{x-1}\left[\mathfrak{Re}(q_{(x,y)})\cdot 2\mathfrak{Re}(h_{(i,x)}\overline{h}_{(i,y)})\right. \\ & \qquad\qquad \left.+\mathfrak{Im}(q_{(x,y)})\cdot 2\mathfrak{Im}(h_{(i,y)}\overline{h}_{(i,x)})\right] \\ \text{s.t.} & -\frac{n_t-1}{2} \leq \sum_{x=2}^{n_t}\sum_{y=1}^{x-1}\mathfrak{Re}(q_{(x,y)}) \leq \frac{n_t-1}{2} \\ & -\frac{n_t-1}{2} \leq \sum_{x=2}^{n_t}\sum_{y=1}^{x-1}\mathfrak{Im}(q_{(x,y)}) \leq \frac{n_t-1}{2} \\ & -1 \leq \mathfrak{Re}(q_{(x,y)}), \mathfrak{Im}(q_{(x,y)}) \leq 1,\ \forall x, y \end{array}$$

With Theorem 2, we are ready to bound $\mathfrak{Re}(d^{\langle L\rangle}_{(i,k)})$ and $\mathfrak{Im}(d^{\langle L\rangle}_{(i,k)})$. Since $d_{(i,i)} = \sum_{k=1}^{i}\left[\mathfrak{Re}^2(d^{\langle L\rangle}_{(i,k)}) + \mathfrak{Im}^2(d^{\langle L\rangle}_{(i,k)})\right]$, we can bound all $\mathfrak{Re}(d^{\langle L\rangle}_{(i,k)})$ and $\mathfrak{Re}(d^{\langle L\rangle}_{(i,k)}), 1 \leq k \leq i, 1 \leq i \leq n_t$, in the the following intervals

$$\left[-\sqrt{[\rho_{\mathbf{H}}(D^{(1)*}_{(i)} + D^{(2)*}_{(i)})] + 1}, \sqrt{[\rho_{\mathbf{H}}(D^{(1)*}_{(i)} + D^{(2)*}_{(i)})] + 1}\right].$$

Since $\mathbf{D}$ and $\mathbf{R}$ are of the identical form, the upper and lower bounds of $\mathfrak{Re}(r^{\langle L\rangle}_{(i,k)})$ and $\mathfrak{Re}(r^{\langle L\rangle}_{(i,k)}), 1 \leq k \leq i, 1 \leq i \leq n_t$ can be derived in exactly the same way and are omitted in here for brevity.

2) X_i, and V_i: Since X_i and V_i are of identical form, we only show how to derive upper and lower bounds for X_i and omit the derivations of V_i for brevity. Recall that $X_i = \mathfrak{Re}^2(d^{\langle L\rangle}_{(i,i)}) + \mathfrak{Im}^2(d^{\langle L\rangle}_{(i,i)})$, we can upper bound X_i as follows:

$$X_i = \|d^{\langle L\rangle}_{(i,i)}\|^2 \leq d_{(i,i)} \leq \left(d_{(i,i)}\right)_U.$$

Since $\det(\mathbf{I} + \mathbf{A}) \geq 1 + \det\mathbf{A}$ for any $\mathbf{A} \succeq 0$ (Minkowski inequality [14]), we have $|\mathbf{D}| = |\mathbf{I} + \rho_{\mathbf{H}}\mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}| = \prod_{i=1}^{n_r}\|d^{\langle L\rangle}_{(i,i)}\|^2 \geq 1$. Hence, X_i can be lower bounded by

$$X_i = \|d^{\langle L\rangle}_{(i,i)}\|^2 \geq \frac{1}{\prod_{\substack{j=1\\ j\neq i}}^{n_r}\left\|d_{(j,j)}\right\|^2} \geq \frac{1}{\prod_{\substack{j=1\\ j\neq i}}^{n_r}\left(d_{(j,j)}\right)_U}.$$

G. Local Search Algorithm

As indicated earlier, the solution to the linear relaxation (14), denoted by $\hat{\mathbf{Q}}$, may not be feasible to the original problem (2). Therefore, during each iteration, we need to obtain a feasible solution based on $\hat{\mathbf{Q}}$ via a local search algorithm if $\hat{\mathbf{Q}}$ is infeasible. The basic idea of our local search is to project $\hat{\mathbf{Q}}$ onto the positive semidefinite cone.

This projection task can be formulated as the following minimization problem: given a Hermitian matrix $\hat{\mathbf{Q}}$, we wish to find a matrix $\mathbf{Q}$ satisfying $\mathbf{Q} \succeq 0$ and $\mathrm{Tr}\{\mathbf{Q}\} \leq 1$ such that $\mathbf{Q}$ minimizes $\|\mathbf{Q} - \hat{\mathbf{Q}}\|_F$, where $\|\cdot\|_F$ represent the Frobenius norm. Mathematically, this can be written as

$$\begin{array}{rl} \text{Minimize} & \frac{1}{2}\|\mathbf{Q} - \hat{\mathbf{Q}}\|_F^2 \\ \text{subject to} & \mathrm{Tr}(\mathbf{Q}) \leq 1,\ \mathbf{Q} \succeq 0. \end{array} \tag{16}$$

Authorized licensed use limited to: to IEEExplore provided by Virginia Tech Libraries. Downloaded on June 19, 2009 at 13:23 from IEEE Xplore. Restrictions apply.

Due to space limitation, we refer reader to [13] for more details about how to solve problem (16).

IV. Convergence Speedup Technique

BB/RLT has exponential complexity due to the NP-hardness of problem (2). However, it is possible to exploit the special structure of problem (2) to significantly speedup its convergence. From our computational experience, we note that the decrease of the global upper bound plays the most important role in the convergence process. Thus, X_i and V_i, should be partitioned with higher priority since they directly impact the global upperbound. We summarize our convergence speedup technique in Algorithm 2.

Algorithm 2 BB Variable Selection Strategy

```
1. Among all $X_i$ and $V_i$, choose the one having the largest relaxation error
   and denoted it as $Z_i^*$.
2. If $(\ln(Z_l^*)_U - \ln(Z_l^*)_L \leq \epsilon_1)$ then
   a) Among all $\mathfrak{Re}(d_{(i,k)}^{\langle L \rangle})$'s, $\mathfrak{Im}(d_{(i,k)}^{\langle L \rangle})$'s, $\mathfrak{Re}(r_{(i,k)}^{\langle L \rangle})$'s, and
      $\mathfrak{Im}(r_{(i,k)}^{\langle L \rangle})$'s, choose one 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 the selected
      variable; else return $Z_l^*$.
```

V. Numerical Example

In this section, We use a numerical example to demonstrate the computational experience of BB/RLT. The simulation setting is as follows: $n_t = n_r = n_e = 4$, $\rho_{\mathbf{H}} = 10$dB, and $\rho_{\mathbf{G}} = 5$dB. We plot the convergence process of BB/RLT in Fig. 3, where the global UB and LB for the secrecy capacity (in b/s/Hz) are illustrated in each iteration. The desired error bound is chosen to be $\epsilon = 0.01$. In this example, after approximately 30000 iterations, the UB and LB values are both driven to 5.45 b/s/Hz, meaning that the secrecy capacity is 5.45 b/s/Hz. Although the number of iterations is seemingly large, the running time of BB/RLT is quite short thanks to the efficiency and robustness of modern day linear programming solvers.

However, we remark that the convergence speedup technique discussed in Section IV is crucial for the success of BB/RLT. For this example, without using the speedup technique, the BB process would easily stall, although it should converge theoretically. This is because without careful consideration in selecting partitioning variables, the algorithm may waste most of its time in partitioning those variables who have very minor effect on closing the gap between global upper and lower bounds. Also, the size of the subproblem list in BB/RLT may become large quickly.

VI. Conclusion

In this paper, we investigated the optimal power allocation problem of Gaussian MIMO wire-tap channels. For this challenging non-convex optimization problem, we designed a global optimization algorithm called branch-and-bound with reformulation and linearization technique (BB/RLT). We developed key problem-specific components for BB/RLT and the related convergence speedup technique for solving the optimal power allocation problem. To the best of knowledge, our proposed BB/RLT algorithm is the first method that guarantees finding the global optimal power allocation to achieve perfect secrecy capacity for Gaussian MIMO wire-tap channels.

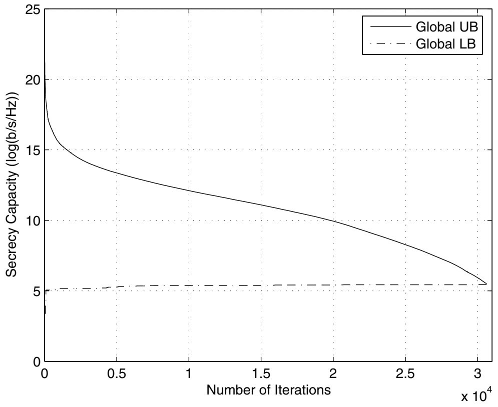

Fig. 3. Convergence process of BB/RLT.

References

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[4] Y. Liang and H. V. Poor, "Multiple access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[5] R. Liu, I. Marić, R. D. Yates, and P. Spasojević, "The discrete memoryless multiple access channel with confidential messages," in *Proc. IEEE ISIT*, Seatle, WA, Jul. 2006, pp. 957–961.

[6] E. Tekin and A. Yener, "Achievable rates for the general gaussian multiple access wire-tap channel with collective secrecy," in *44th Annual Allerton Conference on Communication, Control, and Computing*, Sep. 2006.

[7] A. O. Hero, "Secure space-time communication," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3235–3249, Dec. 2003.

[8] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[9] Z. Li, W. Trappe, and R. Yates, "Secret communication via multi-antenna transmission," in *41st Conference on Information Sciences and Systems*, Baltimore, MD, Mar. 2007, pp. 905–909.

[10] S. Shafiee and S. Ulukus, "Achievable rates in gaussian MISO channels with secrecy constraints," in *Proc. IEEE ISIT*, Nice, France, Jun.24–29 2007, pp. 2466–2470.

[11] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. IEEE ISIT*, Toronto, Canada, Jul.6-11, 2008, pp. 524–528.

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

[13] J. Liu, Y. T. Hou, and H. D. Sherali, "Maximum achievable rates in Gaussian MIMO wire-tap channels with gaussian signaling," *Technical Report, Department of ECE, Virginia Tech*, Jan. 2009. [Online]. Available: http://filebox.vt.edu/users/kevinlau/publications

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1990.

978-1-4244-2734-5/09/$25.00 ©2009 IEEE

Authorized licensed use limited to: to IEEExplore provided by Virginia Tech Libraries. Downloaded on June 19, 2009 at 13:23 from IEEE Xplore. Restrictions apply.