

# Globally Optimal Precoder Design with Finite-Alphabet Inputs for Cognitive Radio Networks

Weiliang Zeng, *Student Member, IEEE*, Chengshan Xiao, *Fellow, IEEE*, Jianhua Lu, *Senior Member, IEEE*, Khaled Ben Letaief, *Fellow, IEEE*

**Abstract**—This paper investigates the linear precoder design for spectrum sharing in multi-antenna cognitive radio networks with finite-alphabet inputs. It formulates the precoding problem by maximizing the constellation-constrained mutual information between the secondary-user transmitter and secondary-user receiver while controlling the interference power to primary-user receivers. This formulation leads to a nonlinear and non-convex problem, presenting a major barrier to obtain optimal solutions. This work proposes a global optimization algorithm, namely Branch-and-bound Aided Mutual Information Optimization (BAMIO), that solves the precoding problem with arbitrary prescribed tolerance. The BAMIO algorithm is designed based on two key observations: First, the precoding problem for spectrum sharing can be reformulated to a problem minimizing a function with bilinear terms over the intersection of multiple co-centered ellipsoids. Second, these bilinear terms can be relaxed by its convex and concave envelopes. In this way, a sequence of relaxed problems is solved over a shrinking feasible region until the tolerance is achieved. The BAMIO algorithm calculates the optimal precoder and the theoretical limit of the transmission rate for spectrum sharing scenarios. By tuning the prescribed tolerance of the solution, it provides a trade-off between desirable performance and computational complexity. Numerical examples show that the BAMIO algorithm offers near global optimal solution with only several iterations. They also verify that the large performance gain in mutual information achieved by the BAMIO algorithm also represents the large gain in the coded bit-error rate.

**Index Terms**—Cognitive radio, spectrum sharing, linear precoding, finite-alphabet inputs, mutual information maximization, multiple-input multiple-output.

Manuscript received 4 January 2012; revised 15 May 2012. The work of W. Zeng and J. Lu was supported in part by National Nature Science Foundation of China under Grants 61101071, 61021001, and 60972021. The work of C. Xiao was supported in part by the U.S. National Science Foundation under Grants CCF-0915846 and ECCS-1231848. The work of K. B. Letaief was supported by QNRF under Project No. NPRP 09-126-2-054 (QUXX01-20L01611/12PN). This work has been carried out while W. Zeng was a research scholar at Missouri University of Science and Technology. Part of the material in this paper has been presented at the IEEE International Conference on Communications (ICC), Ottawa, Canada, 2012.

W. Zeng and J. Lu are with Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: zengwl07@mails.tsinghua.edu.cn; lhh-dee@mail.tsinghua.edu.cn).

C. Xiao is with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA (e-mail: xiaoc@mst.edu).

K. B. Letaief is with the Center for Wireless Information Technology in the Department of Electronic & Computer Engineering, Hong Kong University of Science & Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: eekhaled@ece.ust.hk).

Digital Object Identifier 10.1109/JSAC.2012.121103.

## I. INTRODUCTION

COGNITIVE radio is an emerging technology that has the potential to significantly improve the utilization efficiency of the scarce radio spectrum [1]–[5]. It allows the secondary-user transceiver to access the radio spectrum that is originally allocated to primary users. Since primary users have a higher priority, their quality-of-service should be guaranteed while the secondary-user transceiver shares the spectrum [6]–[8].

The spectrum sharing scheme is allowed provided the secondary-user transceiver is able to control the interference to primary users and thus keeps the performance degradation of each primary user within a tolerable level. Consider a scenario illustrated in Fig. 1, where one secondary-user transmitter (ST), one secondary-user receiver (SR), and  $n$  primary-user receivers (PRs) share the same frequency band at the same time. Suppose the maximum permitted level of interference from ST to the  $i$ -th PR is denoted as  $\gamma_i$  for  $i = 1, 2, \dots, n$ . The investigation of the theoretical limit on the information rate between ST and SR under interference constraints has been an active research topic for recent literature. With Gaussian input, the channel capacity under transmit power and a set of interference power constraints has been determined, and the precoder achieving the capacity has been readily obtained [7]–[9], since the considered problem is convex [10].

Although Gaussian inputs are capacity achieving and thus theoretically optimal, they are rarely realized in practice. Even worse, the precoder designed based on Gaussian inputs often leads to considerable performance degradation when applied to practical systems with finite-alphabet inputs, such as phase shift keying (PSK) modulation, pulse amplitude modulation (PAM), and quadrature amplitude modulation (QAM). Therefore, alternative precoder design approach that maximizes the *constellation-constrained* mutual information (MI) [11] has drawn great research interests in recent years [12]–[16].

However, finding the precoder that maximizes the MI under the transmit and interference power constraints is far from trivial because it implies maximizing a nonconcave function with multiple local maxima [15]. Limited works have considered both interference power constraints and finite alphabet inputs, while several approaches are found in the literature for the special case of point-to-point multiple-input multiple-output (MIMO) precoding problem with only transmit power constraint (i.e., when interference constraints are inactive). For

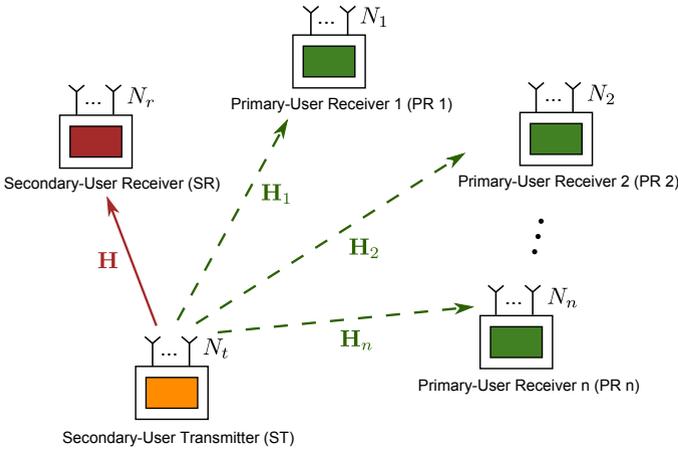


Fig. 1. System model of a transmission scenario, in which the secondary-user transmitter (ST) and the secondary-user receiver (SR) share the same spectrum with multiple primary-user receivers (PRs).

example, the optimal power allocation maximizing the MI is obtained for a diagonal channel matrix [12]. The gradient-ascent method is introduced in [17] for MIMO channels. The structure of the optimal precoder for real-valued MIMO channels is revealed in [18], [19]. A global optimal design is then proposed in [15] for complex-valued MIMO channels. An accurate approximation for MI, which will be used in this work, is recently reported in [20]. It reduces the computational complexity by several orders of magnitude compared to computing the MI directly, and the performance difference between maximizing the approximation and maximizing the MI is negligible.

Existing precoder design algorithms that utilize the optimal precoder structure for point-to-point MIMO channels decompose the precoder into three components: the left singular vector matrix, the diagonal power allocation matrix, and the right singular vector matrix. They first obtain a closed-form solution to the left singular vector matrix, and then different algorithms use different methods to solve the other two components iteratively. This decomposition approach, however, cannot be extended to the case when interference constraints exist because the left singular vector matrix derived for MIMO channels are no longer valid for the precoding problem of spectrum sharing networks. In order to obtain a feasible solution to the more general spectrum sharing problem, a possible precoder satisfying all the constraints may be designed as follows: 1) Design an optimal (or suboptimal) precoder by ignoring the interference constraints with existing algorithms for point-to-point MIMO channels; 2) Rescale the solution to satisfy all interference constraints. This *heuristic* method, however, will lead to significant performance *loss* compared to the best possible solution achieved in this work.

This paper distinguishes itself from other existing methods, not only because of its extension from the point-to-point MIMO channels to the spectrum sharing scenario, but also because of the new developed Branch-and-bound Aided Mutual Information Optimization (BAMIO) algorithm that solves the precoding problem for spectrum sharing with arbitrary prescribed tolerance. The BAMIO algorithm is able to achieve the global optimal precoder for arbitrary signal modulations

and number of transceiver antennas. Similar methods are not reported before, to the best of our knowledge, even for the MIMO channels with only transmit power constraint.

Two key observations enable BAMIO algorithm to efficiently solve the nonconcave problem considered in this work. First, the precoding problem for spectrum sharing can be reformulated to a problem that minimizes a function with bilinear terms over the intersection of multiple co-centered ellipsoids. Second, these bilinear terms can be relaxed by its convex and concave envelopes. In this way, a relaxation of the precoding problem for spectrum sharing is formulated, and it turns out to be a global lower bound for the original problem. Then, the relaxation is used to obtain a solution feasible to the original problem and thus serves as a global upper bound. The global lower and upper bounds are tightened iteratively until the prescribed tolerance is achieved. Two possible methods are considered to speed up the algorithm from different aspects. An improved relaxation is first proposed to decrease the gap between the relaxed problem and the original problem for each iteration. A variable tightening technique is then used to reduce the range of the variables. Both methods cut down the number of iterations and accelerate the convergence. The relationship between the relaxations mentioned in this work and that based on semi-definite relaxation (SDR) is exploited, and the advantages of the proposed algorithm is investigated.

Numerical examples demonstrate the convergence, flexibility, and the performance of the BAMIO algorithm. They show that the algorithm generates very high quality solution with extremely limited computational effort. That is, a near global optimal solution with only several iterations can be used in practice when the computational time is concerned. Numerical examples also verify that the large performance gain in MI achieved by the BAMIO algorithm over the conventional methods represents the large gain in the coded bit-error rate (BER).

The remainder of this paper is organized as follows: Section II introduces the system model and the MI based precoding design. Section III considers the reformulation and relaxation of the precoding problem for spectrum sharing networks. Section IV introduces the BAMIO algorithm achieving the global optimal and the algorithm acceleration and other possible relaxations. Section V presents several numerical examples demonstrating the performance of the proposed algorithm. Finally, Section VI draws conclusions.

*Notation:* Boldface uppercase (lowercase) letters denote matrices (column vectors), and italics denote scalars. The superscripts  $(\cdot)^T$  and  $(\cdot)^H$  stand for transpose and Hermitian operations, respectively. The operator  $\text{Tr}(\cdot)$  denotes the trace of a matrix;  $\text{vec}(\cdot)$  represents the vector obtained by stacking the columns of a matrix;  $\|\cdot\|$  denotes the Frobenius norm of either a matrix or a vector;  $E_{(\cdot)}$  denotes the statistical expectation with respect to its variable;  $\mathbf{I}$  and  $\mathbf{0}$  represent an identity matrix and a zero matrix of appropriate dimensions, respectively;  $\mathbf{A} \otimes \mathbf{B}$  denotes the Kronecker product of two matrices;  $\mathbb{R}$  and  $\mathbb{C}$  denote the real and complex space;  $\Re$  and  $\Im$  are the real and image parts of a complex value;  $\log(\cdot)$  are used for the base two logarithm.

## II. PROBLEM FORMULATION

### A. System Model and Mutual Information

Consider the transmission scenario in Fig. 1. The ST with  $N_t$  antennas attempts to communicate to the SR with  $N_r$  antennas while controlling the power leakage to multiple PRs. The number of PRs is  $n$ , and the number of receive antennas at the  $i$ -th PR is  $N_i$  for  $i = 1, 2, \dots, n$ .

Let  $\mathbf{x} \in \mathbb{C}^{N_t}$  be a transmitted signal with zero mean and identity-matrix covariance (i.e.,  $\mathbb{E}(\mathbf{x}\mathbf{x}^H) = \mathbf{I}$ ). The received signals at the SR and PRs are, respectively, modeled as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{P}\mathbf{x} + \mathbf{n} \\ \mathbf{y}_i &= \mathbf{H}_i\mathbf{P}\mathbf{x} + \mathbf{n}_i, \quad i = 1, 2, \dots, n \end{aligned}$$

where  $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$  and  $\mathbf{H}_i \in \mathbb{C}^{N_i \times N_t}$  are channel matrices from ST to SR and from ST to the  $i$ -th PR, respectively. The vectors  $\mathbf{n} \in \mathbb{C}^{N_r}$  and  $\mathbf{n}_i \in \mathbb{C}^{N_i}$  are independent and identically distributed zero-mean circularly-symmetric Gaussian noises with covariance  $\sigma^2\mathbf{I}$  and  $\sigma_i^2\mathbf{I}$ . The precoding matrix  $\mathbf{P} \in \mathbb{C}^{N_t \times N_t}$  is introduced into the system to improve the transmission rate from ST to SR while restraining the interference to PRs.

The maximal transmit power is  $\gamma$ ; that is,  $\text{Tr}(\mathbf{P}^H\mathbf{P}) \leq \gamma$ . The normalized signal-to-noise ratio (SNR) of the link from ST to SR is thus defined as [21]

$$\text{SNR} \triangleq \frac{\gamma}{N_t} \cdot \frac{\text{Tr}(\mathbf{H}^H\mathbf{H})}{N_r\sigma^2}.$$

The input signal  $\mathbf{x}$  is drawn from equiprobable discrete constellation with cardinality  $M$ . When the channel  $\mathbf{H}$  is known at the receiver, the MI that characterizes the theoretical limit of the constellation-constrained information rate with an arbitrary low probability of error, is given by [11], [21]

$$\mathcal{I}(\mathbf{P}) = K - \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \mathbb{E}_{\mathbf{n}} \left\{ \log \sum_{k=1}^{M^{N_t}} \exp \left( -\frac{\|\mathbf{H}\mathbf{P}\mathbf{e}_{mk} + \mathbf{n}\|^2 - \|\mathbf{n}\|^2}{\sigma^2} \right) \right\} \quad (1)$$

where  $K$ , a constant, equals  $N_t \log M$ , and  $\mathbf{e}_{mk}$  is the difference between  $\mathbf{x}_m$  and  $\mathbf{x}_k$ . Each  $\mathbf{x}_m$  or  $\mathbf{x}_k$  contains  $N_t$  symbols and is taken independently from the  $M$ -ary signal constellation.

The maximization of (1) is difficult for three reasons. First, closed-form expression for MI is challenging, if not impossible, to obtain; see [12] for high SNR approximation of  $\mathcal{I}(\mathbf{P})$  when  $\mathbf{H}$  is diagonal. Second, the expectation over  $\mathbf{n}$  in (1) implies calculation of  $2N_r$  integrals from  $-\infty$  to  $+\infty$ , which can be computationally prohibitive. Third, the gradient of MI involves minimum mean square error matrix [22], which is, theoretically, helpful for maximizing (1), but is even more challenging to calculate. Although Monte Carlo and numerical integral method can be used to estimate MI and its gradient, they can be quite slow and produce a limited accuracy, especially for large input-output dimensions.

To address these difficulties and simplify the precoder design based on constellation-constrained MI, an approximation

of MI is derived in [20]

$$\mathcal{I}_A(\mathbf{P}) = K - \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp \left( -\frac{\mathbf{e}_{mk}^H \mathbf{P}^H \mathbf{H}^H \mathbf{H} \mathbf{P} \mathbf{e}_{mk}}{2\sigma^2} \right) \quad (2)$$

which offers a very accurate approximation to the MI for an arbitrary SNR. It reduces the computational complexity typically by several orders of magnitude compared to calculating the MI directly. For example, the CPU time of evaluating  $\mathcal{I}_A(\mathbf{P})$  for  $2 \times 2$  and  $4 \times 4$  MIMO channels with QPSK inputs is about  $5.9 \times 10^{-5}$  and  $9.6 \times 10^{-6}$ , respectively, times those of evaluating  $\mathcal{I}(\mathbf{P})$ . Besides, the precoder designed based on maximizing  $\mathcal{I}_A(\mathbf{P})$  is proved to be asymptotically optimal, compared to maximizing MI directly, in both low and high SNR regions. Intensive simulations show that for an arbitrary SNR, the performance difference between maximizing MI and maximizing  $\mathcal{I}_A(\mathbf{P})$  is negligible [20].

### B. MI-based Precoding Design

With the assumption of  $\mathbb{E}(\mathbf{x}\mathbf{x}^H) = \mathbf{I}$ , the average transmit signal power at the ST is  $\text{Tr}(\mathbf{P}^H\mathbf{P})$ , while the average receive signal power at the  $i$ -th PR is given by  $\text{Tr}(\mathbf{P}^H\mathbf{H}_i^H\mathbf{H}_i\mathbf{P})$ . Using the approximated MI introduced in (2), the precoding problem, which maximizes MI between ST and SR while restraining the interference to PRs, is formulated as

$$\max. \quad \mathcal{I}_A(\mathbf{P}) \quad (3a)$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{P}^H\mathbf{P}) \leq \gamma \quad (3b)$$

$$\text{Tr}(\mathbf{P}^H\mathbf{H}_i^H\mathbf{H}_i\mathbf{P}) \leq \gamma_i, \quad i = 1, 2, \dots, n \quad (3c)$$

in which (3b) controls the transmit power, and (3c) keeps the interference to  $i$ -th PR below a tolerable level  $\gamma_i$ .

It is interesting to note that problem (3) is related to the point-to-point MIMO precoding problem with finite-alphabet inputs when interference constraints are inactive, e.g.,  $n = 0$  or  $\gamma_i = +\infty$  for  $i = 1, 2, \dots, n$ :

$$\max. \quad \mathcal{I}_A(\mathbf{P})$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{P}^H\mathbf{P}) \leq \gamma. \quad (4)$$

This problem is nonconcave as (3). See more details on how to solve this problem in [20].

The extended problem (3) includes the special case of (4) and thus is even more challenging to solve. The BAMIO algorithm we are proposing differs significantly from existing precoding algorithms for finite-alphabet inputs and Gaussian inputs. It is able to achieve a precoder solution with arbitrary prescribed tolerance for arbitrary signal modulations and arbitrary number of transceiver antennas. Certainly, this algorithm can also be used to solve problem (4).

## III. REFORMULATION AND RELAXATION OF LINEAR PRECODING PROBLEM

This section provides a new reformulation of the precoding problem (3) and also a new relaxation by introducing the convex and concave envelopes.

### A. Reformulation of the Precoding Problem

Since  $\text{Tr}(\mathbf{X}^H \mathbf{Y} \mathbf{X} \mathbf{W}) = \text{vec}(\mathbf{X})^H \cdot (\mathbf{W}^T \otimes \mathbf{Y}) \cdot \text{vec}(\mathbf{X})$  [23], (3a) can be expressed alternatively as

$$\mathcal{I}_A(\mathbf{P}) = K - \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp \left( -\text{vec}(\mathbf{P})^H \cdot (\mathbf{E}_{mk}^T \otimes \mathbf{H}^H \mathbf{H}) \cdot \text{vec}(\mathbf{P}) \right) \quad (5)$$

where  $\mathbf{E}_{mk} = \mathbf{e}_{mk} \mathbf{e}_{mk}^H / (2\sigma^2)$ . By letting

$$\mathbf{p} = \begin{bmatrix} \Re\{\text{vec}(\mathbf{P})\} \\ \Im\{\text{vec}(\mathbf{P})\} \end{bmatrix} \in \mathbb{R}^{2N_t^2} \quad (6)$$

and

$$\mathbf{A}_{mk} = - \begin{bmatrix} \Re\{\mathbf{E}_{mk}^T \otimes \mathbf{H}^H \mathbf{H}\} & -\Im\{\mathbf{E}_{mk}^T \otimes \mathbf{H}^H \mathbf{H}\} \\ \Im\{\mathbf{E}_{mk}^T \otimes \mathbf{H}^H \mathbf{H}\} & \Re\{\mathbf{E}_{mk}^T \otimes \mathbf{H}^H \mathbf{H}\} \end{bmatrix} \quad (7)$$

expression (5) is rewritten as

$$\mathcal{I}_A(\mathbf{P}) = K - \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp(\mathbf{p}^T \mathbf{A}_{mk} \mathbf{p}). \quad (8)$$

Since  $\mathbf{p}^T \mathbf{A}_{mk} \mathbf{p}$  equals  $-\text{Tr}(\|\mathbf{H} \mathbf{P} \mathbf{e}_{mk}\| / 2\sigma^2)$ , non-positive for an arbitrary  $\mathbf{P}$ ,  $\mathbf{A}_{mk}$  is thus negative semi-definite, denoted as  $\mathbf{A}_{mk} \preceq \mathbf{0}$ . Similarly, define

$$\mathbf{B}_i = \begin{bmatrix} \Re\{\mathbf{I} \otimes \mathbf{H}_i^H \mathbf{H}_i\} & -\Im\{\mathbf{I} \otimes \mathbf{H}_i^H \mathbf{H}_i\} \\ \Im\{\mathbf{I} \otimes \mathbf{H}_i^H \mathbf{H}_i\} & \Re\{\mathbf{I} \otimes \mathbf{H}_i^H \mathbf{H}_i\} \end{bmatrix} \quad (9)$$

then  $\mathbf{B}_i$  is positive semi-definite, denoted as  $\mathbf{B}_i \succeq \mathbf{0}$ , and the receive signal power of the  $i$ -th PR is given by  $\mathbf{p}^T \mathbf{B}_i \mathbf{p} \geq 0$ .

Based on the equations and definitions from (5) to (9), we have the following lemma:

*Lemma 1:* The problem (3) can be reduced to a more compact form:

$$\max. \quad K - \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp(\mathbf{p}^T \mathbf{A}_{mk} \mathbf{p}) \quad (10a)$$

$$\text{s.t.} \quad \mathbf{p}^T \mathbf{p} \leq \gamma \quad (10b)$$

$$\mathbf{p}^T \mathbf{B}_i \mathbf{p} \leq \gamma_i, \quad i = 1, 2, \dots, n \quad (10c)$$

where  $\mathbf{p}$ ,  $\mathbf{A}_{mk}$ , and  $\mathbf{B}_i$  are defined in (6), (7), and (9), and  $\gamma$  and  $\gamma_i$  are positive constants controlling the transmit power and interference power, respectively.  $\diamond$

Because  $\mathbf{A}_{mk} \preceq \mathbf{0}$ , (10a) is nonconcave [10]. Geometrically, problem (10) defines a problem that maximizes a nonconcave function over the intersection of  $(n+1)$  co-centered ellipsoids. Since zero vector always satisfies all the constraints, the feasibility of this problem is trivial. However, finding the global maximum can be very difficult, because (10) can be specialized to non-deterministic polynomial-time hard (NP-hard) problems if  $\mathbf{A}_{mk}$  for  $m, k = 1, \dots, M^{N_t}$  are arbitrary negative semi-definite matrix [24].

For convenience of presentation, we further rewrite (10) as a minimization problem with additional bound constraint

$$v^{\text{MI}}(\mathcal{P}) = \min. \quad \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp(\mathbf{p}^T \mathbf{A}_{mk} \mathbf{p}) - K \quad (11a)$$

$$\text{s.t.} \quad \mathbf{p}^T \mathbf{p} \leq \gamma \quad (11b)$$

$$\mathbf{p}^T \mathbf{B}_i \mathbf{p} \leq \gamma_i, \quad i = 1, 2, \dots, n \quad (11c)$$

$$\mathbf{p} \in \mathcal{P}. \quad (11d)$$

The notation  $v^{\text{MI}}(\mathcal{P})$  in (11a) denotes the optimal value of problem (11); the bound  $\mathcal{P}$  in (11d) denotes a box defined as

$$\mathcal{P} = \left\{ \mathbf{p} \in \mathbb{R}^{2N_t^2} \mid \mathbf{l}_p \leq \mathbf{p} \leq \mathbf{u}_p \right\}$$

in which the vector inequalities are element-wise. The lower bound  $\mathbf{l}_p$  and upper bound  $\mathbf{u}_p$  can, respectively, be initialized as  $-\sqrt{\gamma} \cdot \mathbf{1}$  and  $\sqrt{\gamma} \cdot \mathbf{1}$ , where  $\mathbf{1}$  denotes the column vector with all entries being one. This initialization keeps the problems (10) and (11) equivalent because it defines a constraint looser than (11b) and therefore is inactive.

The better initialization for  $\mathbf{l}_p$  and  $\mathbf{u}_p$ , in terms of the size of the defined box, is given by solving  $2N_t^2$  optimization problems. Denote the  $j$ -th element of  $\mathbf{l}_p$  as  $l_{p,j}$ , which is set to be the optimal value of the following convex problem:

$$\begin{aligned} l_{p,j} &= \min. \quad p_j \\ &\text{s.t.} \quad \mathbf{p}^T \mathbf{p} \leq \gamma \\ &\quad \mathbf{p}^T \mathbf{B}_i \mathbf{p} \leq \gamma_i, \quad i = 1, 2, \dots, n \end{aligned} \quad (12)$$

where  $p_j$  is the  $j$ -th element of  $\mathbf{p}$ . Due to the symmetry of the constraints,  $u_{p,j}$ , the  $j$ -th element of  $\mathbf{u}_p$ , is given by  $-l_{p,j}$ .

The sequel proposes an algorithm to iteratively shrink the initial bounded box  $\mathcal{P}$  until a prescribed tolerance is achieved. It starts by introducing the concept of convex and concave envelopes, which is critical to understand the algorithm.

### B. Convex and Concave Envelopes and Bilinear Relaxation

Consider a function  $g : \Omega \rightarrow \mathbb{R}$ , where  $\Omega$  is a region. The *convex envelope* of  $g$  over  $\Omega$ , denoted as  $\text{vex}_\Omega(g)$ , is the pointwise supremum of convex underestimators of  $g$ ; the *concave envelope* of  $g$  over  $\Omega$ , denoted as  $\text{cav}_\Omega(g)$ , is the pointwise infimum of concave overestimators of  $g$ .

A quadratic function is the weighted sum of multiple bilinear functions  $p_i p_j$ :

$$\mathbf{p}^T \mathbf{A}_{mk} \mathbf{p} = \sum_i \sum_j a_{mk,ij} \cdot p_i p_j \quad (13)$$

where  $a_{mk,ij}$  denotes the  $(i, j)$ -th element of  $\mathbf{A}_{mk}$ .

The bilinear function, in general, neither convex nor concave and makes optimization problems difficult to solve [10]. It can be relaxed to a piecewise-linear function based on its convex and concave envelopes. Consider a bilinear function  $g(p_i, p_j) = p_i p_j$  over a rectangular region

$$\Omega = \{(p_i, p_j) \in \mathbb{R}^2 \mid l_{p,i} \leq p_i \leq u_{p,i}, l_{p,j} \leq p_j \leq u_{p,j}\}.$$

Then the convex and concave envelopes over  $\Omega$  are, respectively, given by [25]

$$\begin{aligned} \text{vex}_\Omega(p_i p_j) &= \max \left\{ l_{p,j} p_i + l_{p,i} p_j - l_{p,i} l_{p,j}, \right. \\ &\quad \left. u_{p,j} p_i + u_{p,i} p_j - u_{p,i} u_{p,j} \right\} \\ \text{cav}_\Omega(p_i p_j) &= \min \left\{ u_{p,j} p_i + l_{p,i} p_j - l_{p,i} u_{p,j}, \right. \\ &\quad \left. l_{p,j} p_i + u_{p,i} p_j - u_{p,i} l_{p,j} \right\}. \end{aligned}$$

The relationship between the bilinear function and its envelopes is illustrated in Fig. 2, where  $p_i p_j$ ,  $\text{vex}_\Omega(p_i p_j)$ , and  $\text{cav}_\Omega(p_i p_j)$  are depicted, as an example, over the region  $\Omega = \{(p_i, p_j) \in \mathbb{R}^2 \mid -1 \leq p_i, p_j \leq 1\}$ . Although the bilinear

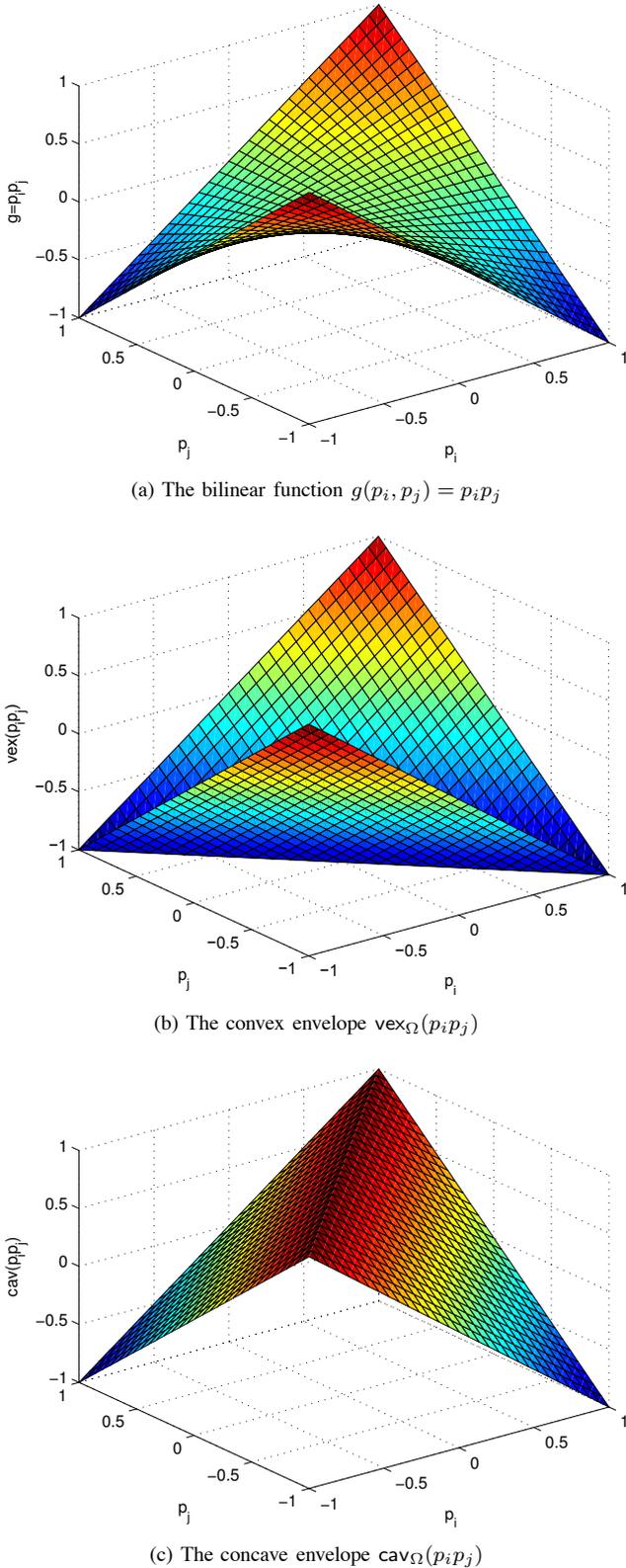


Fig. 2. The bilinear function  $g(p_i, p_j) = p_i p_j$  and its convex and concave envelopes.

function is neither convex nor concave (Fig. 2a), the convex envelope (Fig. 2b) and the concave envelope (Fig. 2c) are piecewise linear and therefore are much easier to use from the viewpoint of optimization.

Using the convex and concave envelopes as the lower and upper estimators of the bilinear function, we introduce an

auxiliary variable  $q_{ij}$  to approximate  $p_i p_j$  by constraining it between the envelopes

$$q_{ij} \geq l_{p,j} p_i + l_{p,i} p_j - l_{p,i} l_{p,j} \quad (14a)$$

$$q_{ij} \geq u_{p,j} p_i + u_{p,i} p_j - u_{p,i} u_{p,j} \quad (14b)$$

$$q_{ij} \leq u_{p,j} p_i + l_{p,i} p_j - l_{p,i} u_{p,j} \quad (14c)$$

$$q_{ij} \leq l_{p,j} p_i + u_{p,i} p_j - u_{p,i} l_{p,j} \quad (14d)$$

Arranging the variables  $q_{ij}$  for  $i, j = 1, \dots, 2N_t^2$  into a matrix  $\mathbf{Q} \in \mathbb{R}^{2N_t^2 \times 2N_t^2}$ , a relaxation of problem (11) over the bounded box  $\mathcal{P}$  can be derived as

$$v^{\text{rlx}}(\mathcal{P}) = \min. \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp \left[ \text{Tr}(\mathbf{A}_{mk} \mathbf{Q}) \right] - K$$

$$\text{s.t. } \text{Tr}(\mathbf{Q}) \leq \gamma \quad (15a)$$

$$\text{Tr}(\mathbf{B}_i \mathbf{Q}) \leq \gamma_i, \quad i = 1, 2, \dots, n \quad (15b)$$

$$\mathbf{Q} - \mathbf{l}_p \cdot \mathbf{p}^T - \mathbf{p} \cdot \mathbf{l}_p^T \geq -\mathbf{l}_p \cdot \mathbf{l}_p^T \quad (15c)$$

$$\mathbf{Q} - \mathbf{u}_p \cdot \mathbf{p}^T - \mathbf{p} \cdot \mathbf{u}_p^T \geq -\mathbf{u}_p \cdot \mathbf{u}_p^T \quad (15d)$$

$$\mathbf{Q} - \mathbf{l}_p \cdot \mathbf{p}^T - \mathbf{p} \cdot \mathbf{u}_p^T \leq -\mathbf{l}_p \cdot \mathbf{u}_p^T \quad (15e)$$

$$\mathbf{Q} = \mathbf{Q}^T, \quad \mathbf{p} \in \mathcal{P} \quad (15f)$$

in which  $\mathbf{l}_p$  and  $\mathbf{u}_p$  are vectors whose  $i$ -th element is  $l_{p,i}$  and  $u_{p,i}$ , respectively; the matrix inequalities of (15c), (15d), and (15e) are element-wise; (15c) and (15d) correspond, respectively, to (14a) and (14b), and (15e) corresponds to (14c) and (14d).

The global optimal value  $v^{\text{rlx}}(\mathcal{P})$  of problem (15) can be readily attained (e.g., by interior-point method [10]) because (15) is a convex problem—all the constraints are linear over variables  $\mathbf{Q}$  and  $\mathbf{p}$ , and the objective function is convex.

Considering the tractability of the relaxed problem, the relationship between  $v^{\text{rlx}}(\mathcal{P})$  and  $v^{\text{MI}}(\mathcal{P})$  is a natural question, which is answered in the following proposition:

*Proposition 1:* The optimal value of the relaxation (15) is a lower bound to the optimal value of (11), i.e.,

$$v^{\text{rlx}}(\mathcal{P}) \leq v^{\text{MI}}(\mathcal{P})$$

and the optimal value of the relaxation (15) converges to that of (11) as the bounded box  $\mathcal{P}$  shrinks down to a point, i.e.,

$$\lim_{\epsilon \rightarrow 0} \left( v^{\text{rlx}}(\mathcal{P}) - v^{\text{MI}}(\mathcal{P}) \right) = 0.$$

where  $\epsilon$  is the diameter of a ball  $\mathcal{B}(\epsilon)$  that includes  $\mathcal{P}$ , i.e.,  $\mathcal{P} \subset \mathcal{B}(\epsilon)$ .

*Proof:* See Appendix A.  $\square$

When  $\epsilon \rightarrow 0$ , the structure of the optimal solution of the relaxed problem (15) is further given in the following corollary:

*Corollary 1:* As the bounded box  $\mathcal{P}$  shrinks down to a point, the solution of the relaxed problem (15) is rank one. That is, solving (15) is equivalent to solving (11).

*Proof:* See Appendix A.  $\square$

Proposition 1 and Corollary 1 help us develop the BAMIO algorithm, which is presented in the next section.

#### IV. THE BAMIO ALGORITHM: ACHIEVING GLOBAL OPTIMUM

Since problem (11) is nonconvex, traditional optimization methods fail to achieve a global optimum. This section proposes the BAMIO algorithm to obtain a provably global optimal solution within a prescribed tolerance  $\delta$  from an arbitrary initial point. At each iteration, BAMIO generates two values  $v^{\text{lb}}(\mathcal{P})$  and  $v^{\text{up}}(\mathcal{P})$  for a bounded box  $\mathcal{P}$ . These values are, respectively, lower and upper bounds of  $v^{\text{MI}}(\mathcal{P})$ :

$$v^{\text{lb}}(\mathcal{P}) \leq v^{\text{MI}}(\mathcal{P}) \leq v^{\text{up}}(\mathcal{P}).$$

As  $\mathcal{P}$  shrinks down, both bounds become tight to ensure the convergence of the algorithm.

##### A. Lower and Upper Bounds Over a Bounded Box

By using the envelopes, the convex relaxation for the original nonconvex problem (11) is formulated in (15). It can be solved efficiently and globally, because of the convexity, and serves as a lower bound  $v^{\text{lb}}(\mathcal{P})$  for the original problem (see Proposition 1).

If the optimal solution of problem (15)  $\mathbf{Q}^*$  is rank one, a feasible and optimal solution for the problem (11)  $\mathbf{p}^*$  that satisfies  $\mathbf{Q}^* = \mathbf{p}^*(\mathbf{p}^*)^T$  is attained. In this case, the lower bound  $v^{\text{lb}}(\mathcal{P})$  is achievable; the upper bound equals the lower bound, and the *relative error*, defined as  $|(v^{\text{up}}(\mathcal{P}) - v^{\text{lb}}(\mathcal{P}))/v^{\text{up}}(\mathcal{P})|$ , is zero, which is within the prescribed tolerance  $\delta$ . Note that the relative error is not necessarily less than one, since the objective function (11a) is less than zero.

If  $\mathbf{Q}^*$  is not rank one, it is used to find a rank-one approximation  $\tilde{\mathbf{Q}}^*$ , which generates a feasible solution  $\mathbf{p}$  and thus serves as an upper bound  $v^{\text{up}}(\mathcal{P})$ . Several methods for rank-one approximation are reported in the recent literature (e.g., using the principal eigenvector of  $\mathbf{Q}^*$  or the Gaussian randomization technique). The rank-one approximation around  $\mathbf{Q}^*$  can be understood as a local search, which is known to provide high-quality solution in different applications (see [26], [27] and references therein).

##### B. Iterative Tightening of Lower and Upper Bounds

The BAMIO algorithm is based on the branch-and-bound framework [28]. It starts by computing the lower and upper bounds over an initialized box  $\mathcal{P}_1$ :

$$L_1 = v^{\text{lb}}(\mathcal{P}_1) \quad \text{and} \quad U_1 = v^{\text{up}}(\mathcal{P}_1).$$

If  $|(U_1 - L_1)/U_1| \leq \delta$ , the algorithm terminates. Otherwise, we divide  $\mathcal{P}_1$  along any of its longest edges into two smaller boxes, with equal size,  $\mathcal{P}_2$  and  $\mathcal{P}_3$  ( $\mathcal{P}_2 \cup \mathcal{P}_3 = \mathcal{P}_1$ ), and compute  $v^{\text{lb}}(\mathcal{P}_i)$  and  $v^{\text{up}}(\mathcal{P}_i)$  for  $i = 2$  and  $3$ . Since  $\mathcal{P}_2$  and  $\mathcal{P}_3$  are smaller boxes, the relaxations over both  $\mathcal{P}_2$  and  $\mathcal{P}_3$  can be expected to be tighter than that over  $\mathcal{P}_1$ . That is, we have

$$\min \{v^{\text{lb}}(\mathcal{P}_2), v^{\text{lb}}(\mathcal{P}_3)\} \geq L_1$$

and

$$\min \{v^{\text{up}}(\mathcal{P}_2), v^{\text{up}}(\mathcal{P}_3)\} \leq U_1.$$

---

#### Algorithm 1 The BAMIO Algorithm

---

##### 1: Initialization

2: Denote the bounded box

$$\mathcal{P}_1 = \{\mathbf{p} \in \mathbb{R}^{2N_t^2} | \mathbf{l}_p \leq \mathbf{p} \leq \mathbf{u}_p\}$$

where  $\mathbf{l}_p$  and  $\mathbf{u}_p$  are either given by  $-\sqrt{\gamma} \cdot \mathbf{1}$  and  $\sqrt{\gamma} \cdot \mathbf{1}$ , respectively, or by solving problems in (12).

3: The set of partitions  $\mathbb{P}$  is initialized as  $\mathcal{P}_1$ .

4: Let  $L_1 = v^{\text{lb}}(\mathcal{P}_1)$ ,  $U_1 = v^{\text{up}}(\mathcal{P}_1)$ , and  $i = 1$ . Set the prescribed tolerance  $\delta$ .

5: **while**  $|(U_i - L_i)/U_i| \geq \delta$  **do**

6: Select  $\mathcal{P}$  from  $\mathbb{P}$  that has the least lower bound.

7: Divide  $\mathcal{P}$  along any of its longest edges into two smaller boxes, with equal size,  $\mathcal{P}_j$  and  $\mathcal{P}_k$ .

8: Remove the selected box  $\mathcal{P}$  from  $\mathbb{P}$ , and add two new boxes  $\mathcal{P}_j$  and  $\mathcal{P}_k$  into  $\mathbb{P}$ .

9: Compute the new lower bound for the  $i$ -th iteration:

$$L_i = \min_{\mathcal{P} \in \mathbb{P}} v^{\text{lb}}(\mathcal{P})$$

and also the new upper bound:

$$U_i = \min_{\mathcal{P} \in \mathbb{P}} v^{\text{up}}(\mathcal{P}).$$

The lower bound  $v^{\text{lb}}(\mathcal{P})$  can be obtained by solving either problem (15) or (17); and the upper bound can be obtained by the rank-one approximation technique (see Sec. IV-A for details).

10: Let  $i = i + 1$ .

11: **end while**

12: **return**  $\delta$ -optimal precoder.

---

The lower and upper bounds are thus refined:

$$\begin{aligned} \min \{v^{\text{lb}}(\mathcal{P}_2), v^{\text{lb}}(\mathcal{P}_3)\} &= L_2 \\ &\leq v^{\text{MI}}(\mathcal{P}_1) \leq U_2 = \min \{v^{\text{up}}(\mathcal{P}_2), v^{\text{up}}(\mathcal{P}_3)\}. \end{aligned} \quad (16)$$

As a result, we have smaller gap between lower and upper bounds after the second iteration. If  $|(U_2 - L_2)/U_2| \leq \delta$ , the algorithm terminates. Otherwise, we divide the box with the minimum lower bound along any of its longest edges into two smaller boxes with equal size and attain new bounds.

For each iteration, one box is divided into two. After  $i$  iterations, the initialized bounded box  $\mathcal{P}_1$  is divided into  $(i + 1)$  smaller boxes. The relaxation on these smaller boxes is closer to the original problem, and a decreased gap between  $L_i$  and  $U_i$  and smaller relative error  $|(U_i - L_i)/U_i|$  can be obtained. Based on this idea, the BAMIO algorithm is given in Algorithm 1.

*Proposition 2:* As the initial bounded box  $\mathcal{P}$  iteratively shrinks, the BAMIO algorithm offers a precoder that globally solves the precoding problem (3). Moreover, the sequences  $\{L_i\}$  and  $\{U_i\}$  for  $i = 1, 2, \dots$  converge to the negative of the maximal objective value of (3).

*Proof:* See Appendix B. □

One benefit of the BAMIO algorithm is the possibility of being able to trade off performance against convergence time by choosing the prescribed tolerance  $\delta$ . A smaller  $\delta$  provides

more accurate solution but leads to longer computing time. This fact will be illustrated in the Simulation section.

### C. Convergence Acceleration for BAMIO Algorithm

This section discusses methods to accelerate the convergence of the BAMIO algorithm. Two possible methods speed up the algorithm from two aspects. For each iteration, we first consider to improve the convex relaxation by decreasing its gap from the original problem. We then consider to reduce the size of each bounded box (i.e., the range of variables). Both methods cut down the number of iterations for a required tolerance and thus accelerate the convergence.

1) *Improving the accuracy of the relaxation*: The convex relaxation in (15) is obtained by raising the variable  $\mathbf{p}$  into the matrix space  $(\mathbf{p}, \mathbf{Q})$ , relaxing the constraint  $\mathbf{Q} = \mathbf{p}\mathbf{p}^T$ , and adding linear inequalities to constrain the range of  $\mathbf{Q}$ . The relaxation serves as a lower bound, which can be tightened by the positive semi-definite (PSD) constraint,  $\mathbf{Q} \succeq \mathbf{p}\mathbf{p}^T$ :

$$v^{\text{imp}}(\mathcal{P}) = \min. \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp \left[ \text{Tr}(\mathbf{A}_{mk} \mathbf{Q}) \right] - K \quad (17a)$$

$$\text{s.t. } \text{Tr}(\mathbf{Q}) \leq \gamma \quad (17b)$$

$$\text{Tr}(\mathbf{B}_i \mathbf{Q}) \leq \gamma_i, \quad i = 1, 2, \dots, n \quad (17b)$$

$$\mathbf{Q} - \mathbf{l}_p \cdot \mathbf{p}^T - \mathbf{p} \cdot \mathbf{l}_p^T \geq -\mathbf{l}_p \cdot \mathbf{l}_p^T \quad (17c)$$

$$\mathbf{Q} - \mathbf{u}_p \cdot \mathbf{p}^T - \mathbf{p} \cdot \mathbf{u}_p^T \geq -\mathbf{u}_p \cdot \mathbf{u}_p^T \quad (17d)$$

$$\mathbf{Q} - \mathbf{l}_p \cdot \mathbf{p}^T - \mathbf{p} \cdot \mathbf{u}_p^T \leq -\mathbf{l}_p \cdot \mathbf{u}_p^T \quad (17e)$$

$$\mathbf{Q} \succeq \mathbf{p}\mathbf{p}^T, \quad \mathbf{Q} = \mathbf{Q}^T, \quad \mathbf{p} \in \mathcal{P} \quad (17f)$$

in which  $v^{\text{imp}}(\mathcal{P})$  denotes the optimal value of problem (17).

The relationship between the solution of problem (15) and that of problem (17) is important to illustrate the tightness. It is given as follows:

*Corollary 2*: The optimal value of the problem (17) is a lower bound to that of the problem (11) and is tighter than that of problem (15), i.e.,

$$v^{\text{rlx}}(\mathcal{P}) \leq v^{\text{imp}}(\mathcal{P}) \leq v^{\text{MI}}(\mathcal{P}).$$

The new relaxation also converges to the original problem as the bounded box shrinks down to a point, i.e.,

$$\lim_{\epsilon \rightarrow 0} (v^{\text{imp}}(\mathcal{P}) - v^{\text{MI}}(\mathcal{P})) = 0$$

where  $\epsilon$  is the diameter of a ball  $\mathcal{B}(\epsilon)$  that includes  $\mathcal{P}$ .

*Proof*: See Appendix B.  $\square$

Since the PSD constraint is convex, (17) is also a convex problem, which can be solved efficiently by established algorithms. Using  $v^{\text{imp}}(\mathcal{P})$  as a lower bound can thus be expected to reduce the number of iterations for a required tolerance  $\delta$ .

2) *Reducing the size of a bounded box*: The BAMIO algorithm starts from solving a convex problem over a bounded box that is initialized as  $\mathcal{P}_1 = \left\{ \mathbf{p} \in \mathbb{R}^{2N_t} \mid \mathbf{l}_p \leq \mathbf{p} \leq \mathbf{u}_p \right\}$ . As the iteration goes, this box is divided along its longest edges into two subboxes with equal size, and the convex problems are solved over each subbox. These smaller boxes define the region of variable  $\mathbf{p}$ , which can be further restricted to accelerate the convergence [29].

Take the first iteration for example. After we solve the convex problem over  $\mathcal{P}_1$ , we obtain the lower bound  $L_1$  and then the upper bound  $U_1$  (see Sec. IV-A). Denote the objective function of (17) as  $f(\mathbf{p})$ , which is less or equal to  $U_1$ . By combining the inequality constraint  $f(\mathbf{p}) \leq U_1$  with the existing ones in (17), we can find, for each element of  $\mathbf{p}$ , the maximal and the minimal values, which define new bounds that can be smaller than  $\mathcal{P}_1$ . Thus, the subboxes divided from the new bounds are also smaller, and the number of iterations required to converge can be less.

### D. Relationship with Semi-Definite Relaxation

It is important to identify the relationship between the SDR developed for quadratic problems [26], [27] and the relaxation (17).

SDR lifts the variable  $\mathbf{p}$  into matrix space  $\mathbf{Q}$  and loses the constraint  $\mathbf{Q} = \mathbf{p}\mathbf{p}^T$ :

$$v^{\text{sdr}} = \min. \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \sum_{k=1}^{M^{N_t}} \exp \left[ \text{Tr}(\mathbf{A}_{mk} \mathbf{Q}) \right] - K \quad (18a)$$

$$\text{s.t. } \text{Tr}(\mathbf{Q}) \leq \gamma \quad (18b)$$

$$\text{Tr}(\mathbf{B}_i \mathbf{Q}) \leq \gamma_i, \quad i = 1, 2, \dots, n \quad (18c)$$

$$\mathbf{Q} \succeq \mathbf{0}, \quad \mathbf{Q} = \mathbf{Q}^T. \quad (18d)$$

After solving the above convex problem, the rank-one approximation technique is used to find a feasible solution. The SDR method has been widely used in different applications (for recent overview, see [27]). On the relationship between the problem (17) and (18), we have the following corollary:

*Corollary 3*: Let  $v^{\text{imp}}(\mathcal{P})$  and  $v^{\text{sdr}}$ , respectively, be the optimal value of problem (17) and (18). Then, we have

$$v^{\text{sdr}} \leq v^{\text{imp}}(\mathcal{P}) \leq v^{\text{MI}}(\mathcal{P}). \quad (19)$$

*Proof*: See Appendix B.  $\square$

The SDR in (18) provides a solution with a fixed relative error. Its solution cannot be improved iteratively because its variable  $\mathbf{Q}$  is irrelevant to  $\mathbf{p}$  thus also irrelevant to  $\mathcal{P}$ . The global optimum cannot be guaranteed by the SDR method since the fixed relative error provided is generally not zero.

## V. SIMULATIONS

This section offers examples to illustrate the trade-off between performance and complexity of the BAMIO algorithm. It also shows the efficacy of different relaxations on the algorithm and the performance comparison with the Gaussian input based design and the heuristic method.

### A. Example 1: Trade-Off Between Performance and Complexity

The proposed BAMIO algorithm converges to the global optimal solution with arbitrary prescribed tolerance. The evolution of the algorithm is examined in this example. Let's first consider the case that the transmit power constraint dominates all other constraints. That is, interference constraints are inactive, and thus we focus on the link between ST and SR, which is depicted by the following matrix as an example

$$\mathbf{H} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \quad (20)$$

which was also used in [15].

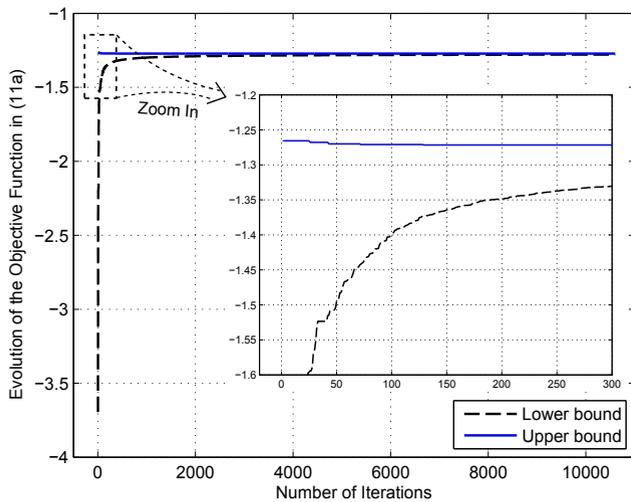


Fig. 3. Typical evolution of objective function in (11a) as the linear precoder is iteratively optimized by the BAMIO algorithm. The signal input is QPSK, and the SNR is -5 dB. The tolerance  $\delta$  is 0.005.

The parameter  $\gamma$  is set to be the same as the number of transmit antennas (in this case, 2), which keeps the transmit power with precoding the same as that without precoding. The input signal  $\mathbf{x}$  is drawn independently from QPSK constellation, and the SNR is -5 dB.

The speedup methods considered in Sec. IV-C is used, and the tolerance is chosen to be  $\delta = 0.005$ . That is, the algorithm terminates until the lower bound  $L_i$  and the upper bound  $U_i$  satisfy  $|(U_i - L_i)/U_i| \leq \delta$ . Figure 3 illustrates the typical evolution of  $L_i$  and  $U_i$ . Note that each iteration of the BAMIO algorithm generates one feasible precoder, which corresponds to the obtained upper bound  $U_i$ . A global optimal solution within tolerance is obtained after 129 iterations; then, more than 10,000 iterations are needed to improve the lower bound  $L_i$  and thus prove that the obtained precoder at 129-th iteration is actually 0.005-optimal.

The trade-off between performance and computational complexity by the tolerance  $\delta$  is illustrated in Fig. 4, where the lower and upper bounds, together with the needed number of iterations versus  $\delta$ , are plotted. The gap between the upper and lower bounds is reduced as the tolerance is smaller. At the same time, the number of iterations increases, and it changes dramatically when  $\delta$  is close to 0.

We are interested in characterizing the relationship between the number of iterations, denoted as  $T$ , and the prescribed tolerance  $\delta$  in Fig. 4 as a mathematical function, which then offers an intuitive way to delineate the relationship. We consider two models to quantify the curve. One model considers an exponential relationship between  $T$  and  $\delta$ ; that is,  $T = a_1 \cdot e^{a_2 \delta}$ . Taking the natural logarithm of both sides results in the equation  $\ln T = \ln a_1 + a_2 \cdot \delta$ . Therefore, a linear relationship between  $\delta$  and  $\ln T$  will indicate the exponential model is appropriate. However, the result plotted in the upper figure of Fig. 5, in which the line is obtained by linear regression, shows the exponential model fails to describe the relationship between  $T$  and  $\delta$ .

The other model considers a polynomial relationship between  $T$  and  $\delta$ ,  $T = a_1 \cdot \delta^{a_2}$ . Taking base-ten logarithm

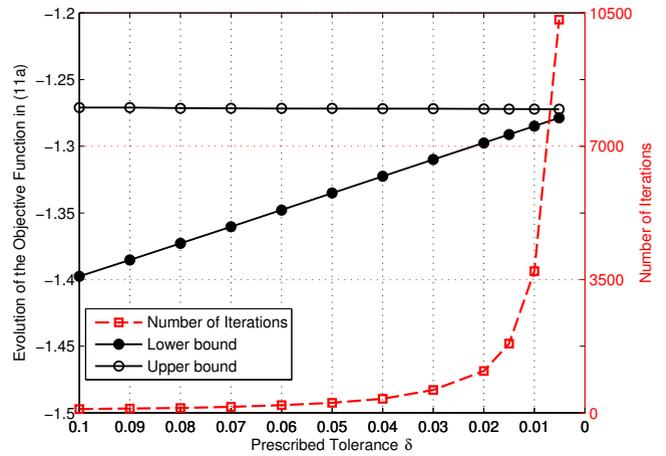


Fig. 4. The evolution of the algorithm and the number of iterations for different prescribed tolerance  $\delta$ . The signal input is QPSK, and the SNR is -5 dB.

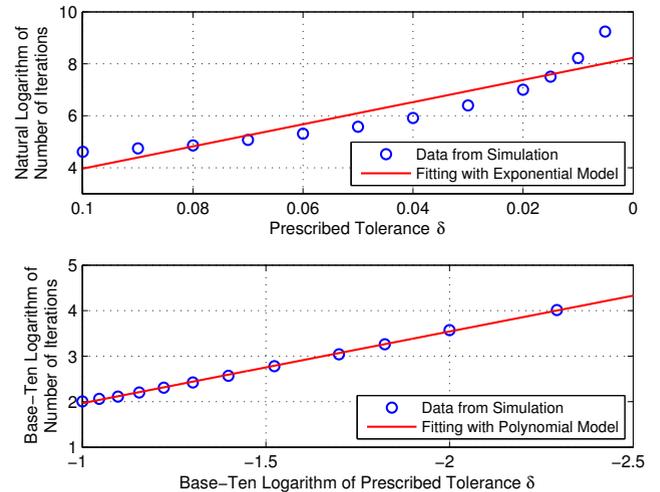


Fig. 5. Regression line for the exponential and polynomial models, which quantify the relationship between the number of iterations and the prescribed tolerance.

of both sides leads to the equation  $\log_{10}(T) = \log_{10}(a_1) + a_2 \cdot \log_{10}(\delta)$ . Therefore, the model is appropriate if a nearly linear relationship exists between  $\log_{10}(T)$  and  $\log_{10}(\delta)$ . We use linear regression to determine  $a_1$  and  $a_2$  and generate the equation  $T = 2.45 \cdot \delta^{-1.58}$ . The curve depicted in lower figure of Fig. 5 shows the polynomial model fits the relationship very well. This fact is also verified by a very high  $R^2$  statistic known as coefficient of determination [30], which is given by 99.9%. Intuitively, this result implies that the number of iterations increases  $10^{a_2} = 38$  times for every 10-time improvement in tolerance (e.g.,  $\delta$  changes from 0.1 to 0.01).

Although the required number of iterations increases faster than the improvement in tolerance, it doesn't necessarily mean a global optimal precoder solution within a prescribed tolerance cannot be obtained by low computational efforts. Observing the convergence curve in Figs. 3 and 4, we have the following remarks:

1) The upper bound converges and stops to decrease after only several iterations, and the lower bound increases gradually to close the gap between both bounds. For example, the

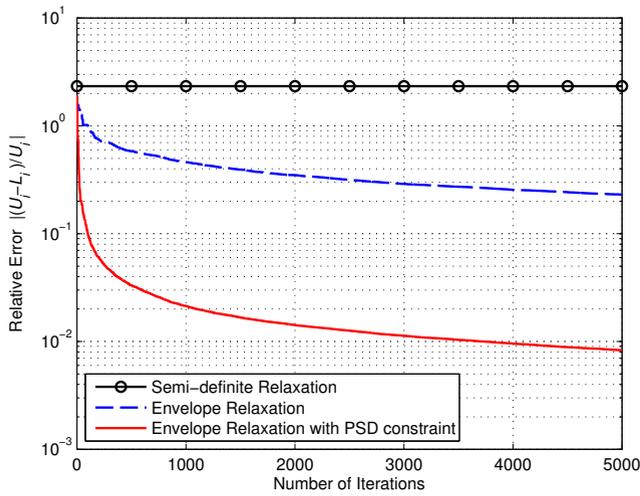


Fig. 6. Comparison of relative error with different relaxations. The signal input is QPSK, and the SNR is -5 dB.

obtained upper bound is -1.272 at the 10,000-th iteration, but up to 99.5% can be achieved at the first iteration. Also, 99.9% can be achieved at the 101-th iteration with  $\delta = 0.1$ .

2) Because the upper bound provides a feasible solution, it represents the achievable objective value. The lower bound, on the other hand, serves as a trusting certificate for the solution.

3) Since the upper bound is insensitive to the iteration, the BAMIO algorithm is able to generate very high quality solution with extremely limited computational effort. That is, a low-complexity BAMIO algorithm with only several iterations can be used in practice when the computational time is concerned and a near optimal solution is permitted.

### B. Example 2: Comparison with Different Relaxations

As shown in Corollary 2 and Corollary 3, different relaxations may have different performance, and the convex-concave envelope relaxation with PSD constraint in (17) dominates both the convex-concave envelope relaxation in (15) and the SDR in (18). This fact is verified in this example, which uses the same simulation setting as those in Example 1. The Gaussian randomization technique (one of possible methods for the rank-one approximation) is used for generating the feasible solution, which is needed by all the three relaxations. As iteration goes, the relative error of these three relaxations are depicted in Fig. 6.

As noted in Sec. IV-D, SDR provides a solution with fixed relative error that cannot be improved iteratively because it introduces new variable  $\mathbf{Q}$  irrelevant to  $\mathbf{p}$  and  $\mathcal{P}$ . However, the solutions based on the envelope relaxation in (15) and envelope relaxation with PSD constraint in (17) are improved iteratively. The latter converges much faster because better lower bound is provided at each iteration (see Corollary 2).

### C. Example 3: Performance Comparison with Different Methods

The BAMIO algorithm allows us to evaluate the optimal precoder and the highest MI with interference constraints. To exemplify the results, we first considers a network with two

PRs, and each PR has two antennas. The channel between ST and SR is the same as (20), while the channels between ST and PRs are depicted by the exponential model:

$$h_{i,jk} = \rho_i^{|j-k|}, \quad i = 1, 2 \quad (21)$$

where  $h_{i,jk}$  is the  $(j, k)$ -th element of  $\mathbf{H}_i$ ,  $j = 1, \dots, N_i$ ,  $k = 1, \dots, N_i$ , and  $\rho_i \in [0, 1)$ . The channel matrices  $\mathbf{H}_1$  and  $\mathbf{H}_2$  are generated by different parameters:  $\rho_1 = 0.95$  and  $\rho_2 = 0.85$ . Meanwhile, the transmit power constraint  $\gamma$  is 2, and different interference power constraints,  $\gamma_i = +\infty, 0.2$ , and 0.02, are considered. For example, when  $\gamma_i$  equals  $+\infty$ , the interference constraints are no longer exist; when  $\gamma_i$  equals 0.2 and 0.02, the maximum receive power permitted at PRs is, respectively, 10 dB and 20 dB less than the transmit power. In this way, the interference power at the PRs can be controlled below a tolerable level.

If we ignore the fact that the input symbols are from finite alphabet and assume deliberately that they are from the ideal Gaussian input, the precoder satisfying all the constraints can be readily obtained, since the problem is maximizing a concave function over a convex set. We name this scheme as *Gaussian based method*. We will show that the precoder designed by the Gaussian based method often leads to considerable performance *degradation* when it is applied to practical systems with finite-alphabet inputs.

Intuitively, if the BAMIO algorithm was not developed, a precoder from the perspective of finite-alphabet inputs and satisfying all the constraints would be designed as follows: 1) Design an optimal precoder by solving (4) (i.e., without interference constraints) using methods developed in existing works, e.g., [15], [20]; 2) Denote the solution as  $\tilde{\mathbf{P}}$  and rescale it to satisfy all constraints:

$$\tilde{\mathbf{P}}^* = \frac{\tilde{\mathbf{P}}}{\sqrt{\max_{i=1, \dots, n} \text{Tr}(\tilde{\mathbf{P}}^H \mathbf{H}_i^H \mathbf{H}_i \tilde{\mathbf{P}}) / \gamma_i}}.$$

The sequel will show that this *heuristic method* may also result in significant performance *loss* compared to the proposed BAMIO algorithm, which finds the global optimal solution.

The performance of different methods are provided in Fig. 7, which shows the Monte Carlo Simulated MI versus SNR for various interference constraints. The capacity with Gaussian inputs, serving as the ultimate upper bound for all possible linear precoders, are also included for comparison purpose.

When interference constraints are inactive (see Fig. 7a), the proposed BAMIO algorithm globally solves the problem (4). It turns out to be an alternative to the heuristic method, which, in this case, solves (4) without rescaling the solution, since the interference constraints are essentially removed. Although the Gaussian based method achieves the capacity when input signal is from Gaussian distribution, it leads to significant performance *loss* if it is directly applied to finite-alphabet inputs. This fact delineates the essential difference between the precoder design with ideal Gaussian inputs and with practical finite-alphabet inputs (see [15], [20] for more details).

When interference constraints are active, the heuristic method satisfies the constraints by reducing the transmit power, which also results in a low performance in terms of MI. Thus, tighter interference constraints lead to larger gap

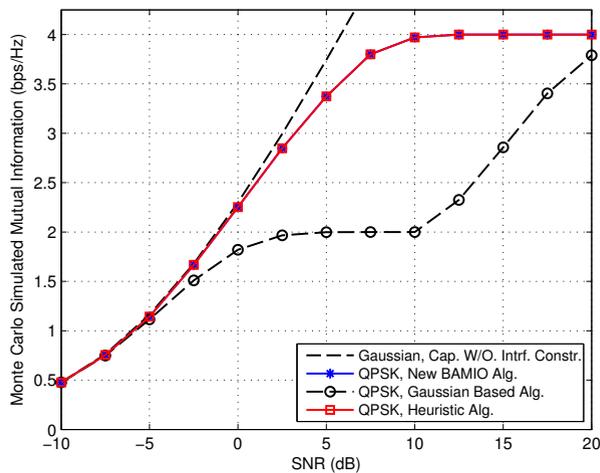
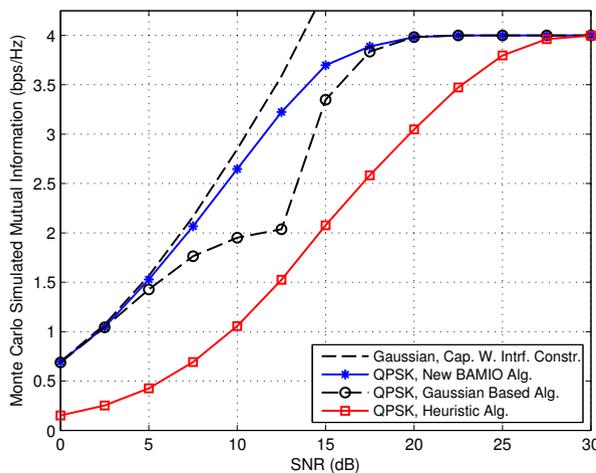
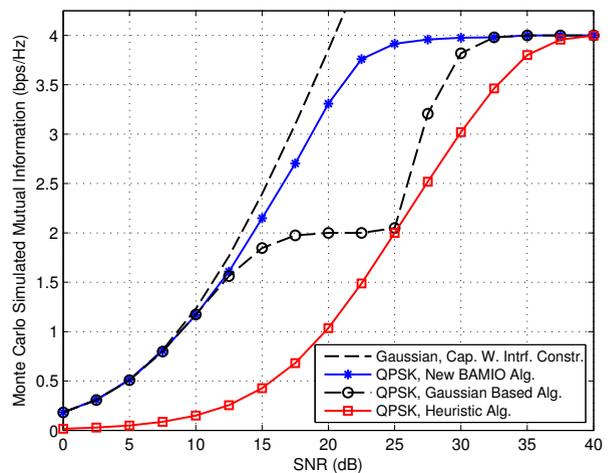
(a) Interference constraints are inactive ( $\gamma_i = +\infty$ ).(b) The maximum receive power permitted at PRs is 10 dB less than the transmit power ( $\gamma_i = 0.2$ ).(c) The maximum receive power permitted at PRs is 20 dB less than the transmit power ( $\gamma_i = 0.02$ ).

Fig. 7. Monte Carlo simulated MI versus SNR for a network with one ST, one SR, and two PRs. Each node has two antennas. The signal input is QPSK; the transmit power constraint  $\gamma$  is 2; the interference constraints  $\gamma_i$  ( $i = 1, 2$ ) are  $+\infty$ , 0.2, and 0.02, respectively.

between the heuristic method and the BAMIO algorithm (see Figs. 7b and 7c). For example, when the channel coding rate is  $3/4$ , the performance of the precoder obtained by the heuristic method is about 8.3 dB and 11.1 dB inferior to that of the precoder obtained by the BAMIO algorithm for  $\gamma_i = 0.2$  and 0.02, respectively.

The Gaussian based method, on the other hand, induces non-smoothness shown in Fig. 7. The reason comes from the signal mismatch between assumptions and reality. Specifically, when the SNR is less than a threshold (10 dB, 12.5 dB, and 25 dB, respectively, for Figs. 7a, 7b, and 7c), the Gaussian based method uses the stronger subchannel and discards the weaker subchannel (similar to the waterfilling method). This method leads to the MI less than 2 bps/Hz. When the SNR is larger than the threshold, the Gaussian based method starts to use the weaker subchannel and therefore makes the MI larger than 2 bps/Hz. This scheme is far away from optimal and results in different amount of performance *loss* depending on different interference constraints. For example, when the channel coding rate is  $3/4$ , the performance of the precoder obtained by the Gaussian based method is about 12.7 dB, 2.9

dB, and 8.1 dB inferior to that of the precoder obtained by the BAMIO algorithm for  $\gamma_i = +\infty$ , 0.2, and 0.02, respectively.

From Figs. 7a, 7b, and 7c, we also conclude that the achievable MI can be expected to be limited if constraints on interference are demanding. For example, when SNR is 5 dB, the achievable MI is 3.74 bps/Hz, 1.53 bps/Hz, and 0.51 bps/Hz for  $\gamma_i = +\infty$ , 0.2, and 0.02, respectively. Therefore, the system operator may consider to relax the specifications of interference and transmit power levels in order to attain a satisfied MI for the link from ST to SR.

The coded BER performance of the aforementioned precoding algorithms is then compared. The transceiver structure in Fig. 8 (see [31] for more details) is realized to capture the benefit of the proposed BAMIO algorithm. Note that the interleaver in [31] is not shown in Fig. 8 because of the usage of the low-density parity-check, or LDPC, codes [32]. The signal sequences at the transmitter are encoded by the LDPC encoder and mapped by the equiprobable discrete constellations. They are then precoded by  $\mathbf{P}$  and transmitted through  $N_t$  antennas. At the receiver, the maximum *a posteriori* (MAP) detector takes channel observations  $\mathbf{y}$  and *a priori* knowledge

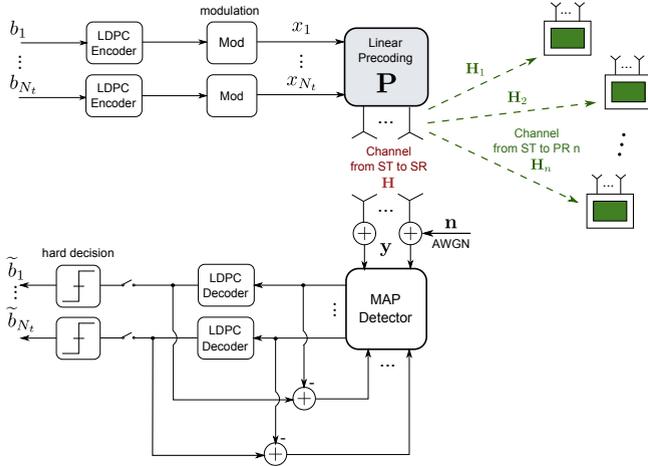


Fig. 8. Block diagram of a multi-antenna spectrum sharing network with channel coding and precoding at the ST and iterative detection and decoding at the SR. The interference from ST to PRs is controlled within the tolerable level.

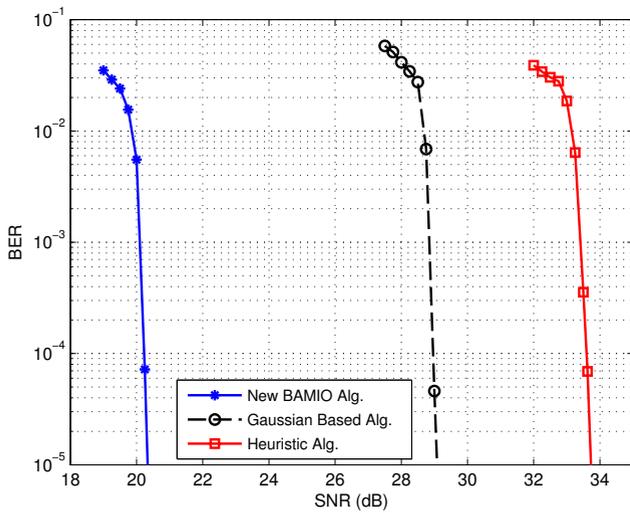


Fig. 9. BER versus SNR for different algorithms. The input signal is drawn from QPSK, the channel coding rate is 3/4, coding block length is 7,200, and the iteration between the MAP detector and the LDPC decoder is 5. The interference constraint  $\gamma_i$  is 0.02.

from the decoder and computes the new information for each coded bit. In this way, the extrinsic information between the MAP detector and the LDPC decoder is exchanged iteratively until the desired performance is obtained.

The LDPC encoder and decoder are derived from [33] with coding rate 3/4. The iteration between the MAP detector and the LDPC decoder is 5. The coding block length of LDPC is 7,200. The obtained coded BER for the corresponding algorithms in Fig. 7c is shown in Fig. 9. Apparently, the proposed BAMIO algorithm outperforms the Gaussian based algorithm and the heuristic algorithm as predicted by MI in Fig. 7c. Based on this result, the large performance gain in MI implies the large performance gain in coded BER. That is, the BAMIO algorithm based on the MI maximization is an excellent approach to provide considerable performance gain for practical spectrum sharing networks.

The performance of the BAMIO algorithm is further investigated when both ST and SR have four antennas and the

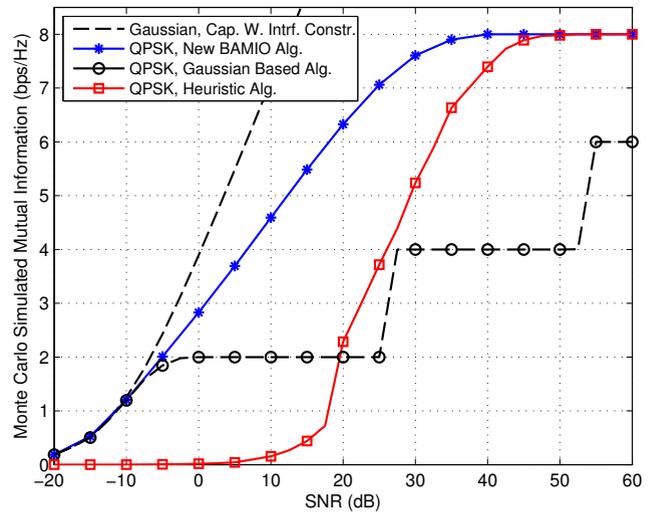


Fig. 10. Monte Carlo simulated MI versus SNR for a network with one ST, one SR, and two PRs. Each node has four antennas. The input signal is drawn from QPSK; the transmit power constraint  $\gamma$  is 4; the interference constraints  $\gamma_i$  ( $i = 1, 2$ ) are 0.04. The  $4 \times 4$  MIMO channel between ST and SR is based on METRA correlation model for the microcell environment.

channel is depicted by

$$\mathbf{H} = \Psi_r \cdot \Psi_t$$

where  $\Psi_r$  and  $\Psi_t$  are, respectively, the receive and transmit correlation matrices for the microcell setting defined by the multiple element transmit-receive antennas (METRA) model [34]. The METRA model is verified by the measured data, and the microcell setting corresponds to a case in which the transmitter is highly correlated.

The two PRs have also four antennas, and the channel between ST and PRs are depicted by the exponential model in (21) with parameters  $\rho_1 = 0.95$  and  $\rho_2 = 0.85$ . The parameter  $\gamma$  is 4, which keeps the transmit power with precoding the same as that without precoding. The interference constraints  $\gamma_i$  ( $i = 1, 2$ ) are 0.04; that is, the maximum receive power permitted at PRs is 20 dB lower than the transmit power.

Figure 10 compares the performance of the BAMIO algorithm with that of the Gaussian based method and the heuristic method. It shows that the curve for the Gaussian based method has a staircase shape. The reason comes from the inefficient usage of the channel. For the  $4 \times 4$  ST-SR channel, the Gaussian based method uses only one subchannel when SNR is less than 25 dB, two subchannels when SNR is more than 25 dB and less than 52.5 dB, and so on. This strategy is capacity achieving when inputs are Gaussian signal; however, it results in significant performance loss when used for a practical system with finite-alphabet inputs. The heuristic method, reducing the transmit power to meet interference constraints, is worse than the Gaussian based method at the low SNR region, and it is better than the Gaussian based method at the high SNR region. Thanks to the BAMIO algorithm, we know both the Gaussian based method and the heuristic method are not good enough because a substantial gain can be achieved by the BAMIO algorithm for a broad range of SNR without violating the interference constraints.

## VI. CONCLUSION

This paper has considered the spectrum sharing problem for multi-antenna cognitive radio networks. It has proposed the precoding algorithm that achieves the theoretical limit on the constellation-constrained information rate between secondary-user transmitter and secondary-user receiver while controlling the interference to multiple primary-user receivers sharing the same frequency band at the same time.

This work has extended the existing precoding study with finite-alphabet inputs, not only because it has considered the more difficult spectrum sharing problem, but also because it has proposed the BAMIO algorithm that solves the open problem of precoding with finite-alphabet inputs. The BAMIO algorithm is guaranteed to converge to a global optimal solution even though the problem is nonconcave. The idea behind the algorithm is to reformulate the precoding design into a problem that minimizes a function with bilinear terms over the intersection of multiple co-centered ellipsoids. A relaxation of the reformulation has been derived by relaxing the bilinear terms based on its convex and concave envelopes. In this way, a sequence of relaxed problems has been solved over shrinking feasible regions. As the iteration goes, the algorithm has provided a solution that eventually converges to the global optimum. Two methods that are able to accelerate the convergence of the algorithm from different aspects have been discussed. The relationship of three possible relaxations has been considered.

Numerical examples have demonstrated the trade-off between the performance and the complexity. They have shown that the BAMIO algorithm generates very high quality solution with extremely limited computational effort. Therefore, a low-complexity algorithm with only several iterations can be used when the computational time is concerned. Numerical examples have also shown the significant performance gain achieved by the proposed BAMIO algorithm for a broad range of SNR compared with the conventional methods. Besides, they verify that the large performance gain in MI achieved by the BAMIO algorithm also represents the large gain in the coded bit-error rate.

## ACKNOWLEDGMENT

The authors would like to thank Prof. Yahong Rosa Zheng, Missouri University of Science and Technology, for her invaluable help throughout this work. They would also like to thank the editor and anonymous reviewers for thorough reviews and constructive suggestions, which have helped to improve the quality of this paper. W. Zeng would like to thank Prof. Nikolaos V. Sahinidis, Carnegie Mellon University, for his useful discussion on the topic of global optimization.

## APPENDIX A

### PROOF OF RELATIONSHIP BETWEEN OPTIMAL VALUES OF PROBLEM (11) AND PROBLEM (15)

*Proof of Proposition 1:* We first prove that the optimal value of the relaxation (15) serves as a lower bound of that of (11). If  $q_{ij}$  for  $i, j = 1, \dots, 2N_t^2$  equals  $p_i p_j$  (i.e.,  $\mathbf{Q} = \mathbf{p} \cdot \mathbf{p}^T$ ), the problem (15) is equivalent to (11). The reason is that the objective function and the first and second constraints

in both problems are the same since  $\text{Tr}(\mathbf{A}\mathbf{Q}) = \text{Tr}(\mathbf{p}^T \mathbf{A}\mathbf{p})$ ; and (15c), (15d), and (15e) become redundant in this case.

If  $q_{ij}$  is relaxed by the convex and concave envelopes, the feasible set of  $q_{ij}$  is enlarged. Therefore, a better solution of variable  $\mathbf{Q}$  with a lower objective value can be found, i.e.,  $v^{\text{rlx}}(\mathcal{P}) \leq v^{\text{MI}}(\mathcal{P})$ .

In order to prove both problems converge to each other as the bounded box shrinks, we first show that the relaxed variable  $q_{ij}$  converges to the bilinear term  $p_i p_j$  when  $\epsilon \rightarrow 0$ . Denote the center of the ball  $\mathcal{B}(\epsilon)$  as  $\bar{\mathbf{p}}$ . Then based on the definition of the concave envelope, we have [25], [35]

$$\begin{aligned} 0 &\leq \lim_{\epsilon \rightarrow 0} \left( \text{cav}_{\mathcal{P}}(\bar{p}_i \bar{p}_j) - \bar{p}_i \bar{p}_j \right) \\ &\leq \lim_{\epsilon \rightarrow 0} \left( \sup_{\mathbf{p} \in \mathcal{P}} (p_i p_j) - \bar{p}_i \bar{p}_j \right) \end{aligned} \quad (22)$$

where  $\bar{p}_i$  and  $\bar{p}_j$  are the  $i$ -th and  $j$ -th elements of  $\bar{\mathbf{p}}$ . Since  $\mathcal{P} \subset \mathcal{B}(\epsilon)$ , we also have

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0} \left( \sup_{\mathbf{p} \in \mathcal{P}} (p_i p_j) - \bar{p}_i \bar{p}_j \right) \\ &\leq \lim_{\epsilon \rightarrow 0} \left( \sup_{\mathbf{p} \in \mathcal{B}(\epsilon)} (p_i p_j) - \bar{p}_i \bar{p}_j \right) = 0 \end{aligned} \quad (23)$$

in which the equality follows from the continuity of  $p_i p_j$ . The expressions in (22) and (23) imply that the concave envelope  $\text{cav}_{\mathcal{P}}(p_i p_j)$  converges to  $p_i p_j$ , that is,

$$\lim_{\epsilon \rightarrow 0} \max_{\mathbf{p} \in \mathcal{P}} \left( \text{cav}_{\mathcal{P}}(p_i p_j) - p_i p_j \right) = 0.$$

Similarly, the convex envelope  $\text{vex}_{\mathcal{P}}(p_i p_j)$  also converges to  $p_i p_j$

$$\lim_{\epsilon \rightarrow 0} \max_{\mathbf{p} \in \mathcal{P}} \left( \text{vex}_{\mathcal{P}}(p_i p_j) - p_i p_j \right) = 0.$$

Combining the convergence of both envelopes, the relaxed variable  $q_{ij}$  constrained between convex and concave envelopes, converges to the bilinear term, i.e.,

$$\lim_{\epsilon \rightarrow 0} q_{ij} = p_i p_j. \quad (24)$$

Now consider the convergence between  $v^{\text{rlx}}(\mathcal{P})$  and  $v^{\text{MI}}(\mathcal{P})$ :

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0} \left( v^{\text{rlx}}(\mathcal{P}) - v^{\text{MI}}(\mathcal{P}) \right) \\ &= \lim_{\epsilon \rightarrow 0} \left( \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \frac{\sum_{k=1}^{M^{N_t}} \exp(\text{Tr}(\mathbf{A}_{mk} \mathbf{Q}))}{\sum_{k=1}^{M^{N_t}} \exp(\mathbf{p}^T \mathbf{A}_{mk} \mathbf{p})} \right) \\ &= \lim_{\epsilon \rightarrow 0} \left( \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \frac{\sum_{k=1}^{M^{N_t}} \exp\left(\sum_{i,j} a_{mk,ij} \cdot q_{ij}\right)}{\sum_{k=1}^{M^{N_t}} \exp\left(\sum_{i,j} a_{mk,ij} \cdot p_i p_j\right)} \right) \\ &= \frac{1}{M^{N_t}} \sum_{m=1}^{M^{N_t}} \log \frac{\sum_{k=1}^{M^{N_t}} \exp\left(\sum_{i,j} a_{mk,ij} \cdot \lim_{\epsilon \rightarrow 0} (q_{ij})\right)}{\sum_{k=1}^{M^{N_t}} \exp\left(\sum_{i,j} a_{mk,ij} \cdot p_i p_j\right)} \\ &= 0 \end{aligned}$$

where the derivation follows from the properties of limits. The proof is now complete.  $\square$

We now provide the proof for Corollary 1:

*Proof of Corollary 1:* From (24), the relaxed variable  $\mathbf{Q}$ , constrained between convex and concave envelopes, converges

to  $\mathbf{pp}^H$  as  $\epsilon \rightarrow 0$ . Consequently, the solution of the relaxed problem is rank one, which ensures all the constraints in (11) is satisfied and that the objective value of (11) is the same as that of problem (15). Therefore, the rank-one solution also solves problem (11). The proof is complete.  $\square$

## APPENDIX B

### PROOF OF CONVERGENCE AND RELATIONSHIP AMONG OPTIMAL VALUES OF PROBLEM (11), (17), AND (18)

We first provide the proof for the convergence of the proposed BAMIO algorithm:

*Proof of Proposition 2:* The convergence for branch-and-bound method is ensured if the difference of lower and upper bounds reduces to zero as the bounded box  $\mathcal{P}$  shrinks down to a point (i.e.,  $\epsilon \rightarrow 0$ ) [36]. This condition holds because of Proposition 1 and Corollary 1.  $\square$

We now prove that the PSD constraint tightens the relaxation based on convex and concave envelopes and therefore provides a better optimal value:

*Proof of Corollary 2:* Since the feasible region defined by problem (17) is tighter than that of (15), the optimal value of the former is larger than the later and thus closer to  $v^{\text{MI}}(\mathcal{P})$ . The second part of Corollary 2 is a direct extension to Proposition 1. It holds because  $v^{\text{imp}}(\mathcal{P})$  is closer to  $v^{\text{MI}}(\mathcal{P})$  than  $v^{\text{rx}}(\mathcal{P})$ .  $\square$

Finally, we prove the relationship between solutions of problem (17) and (18):

*Proof of Corollary 3:* Because of  $\mathbf{pp}^T \succeq \mathbf{0}$  and the linear constraint based on convex and concave envelopes, the constraints in (17) define a region smaller than that of (18). Therefore, the first inequality in (19) holds. The second inequality is from Corollary 2.  $\square$

## REFERENCES

- [1] J. Mitola III and G. Maguire Jr, "Cognitive radio: making software radios more personal," *IEEE Personal Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [3] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.
- [4] J. Ma, Y. G. Li, and B. Juang, "Signal processing in cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.
- [5] Y. Liang, K. Chen, Y. G. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Jul. 2011.
- [6] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May. 2009.
- [7] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007.
- [8] R. Zhang, Y. Liang, and S. Cui, "Dynamic resource allocation in cognitive radio networks," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 102–114, Mar. 2010.
- [9] R. Zhang and Y. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.
- [10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [11] E. Biglieri, *Coding for Wireless Channels*. New York, NY: Springer, 2005.
- [12] A. Lozano, A. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, Jul. 2006.
- [13] G. Caire and K. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proc. IEEE*, vol. 95, no. 12, pp. 2274–2298, Dec. 2007.
- [14] C. Wen and K. Wong, "Asymptotic analysis of spatially correlated MIMO multiple-access channels with arbitrary signaling inputs for joint and separate decoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 252–268, Jan. 2007.
- [15] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3301–3314, Jul. 2011.
- [16] J. Harshan and B. Rajan, "On two-user Gaussian multiple access channels with finite input constellations," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1299–1327, Mar. 2011.
- [17] F. Pérez-Cruz, M. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1070–1084, Mar. 2010.
- [18] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proc. Int. Symp. on Wireless Commun. Syst. (ISWCS)*, 2009, pp. 26–30.
- [19] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Seoul, Korea, 2009, pp. 1085–1089.
- [20] W. Zeng, C. Xiao, and J. Lu, "A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 38–41, Feb. 2012.
- [21] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector Gaussian channels with finite discrete inputs," in *Proc. IEEE Globecom*, New Orleans, LA, 2008.
- [22] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.
- [23] R. Horn and C. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1994.
- [24] A. Nemirovski, C. Roos, and T. Terlaky, "On maximization of quadratic form over intersection of ellipsoids with common center," *Math. Program.*, vol. 86, no. 3, pp. 463–473, 1999.
- [25] F. Al-Khayyal, C. Larsen, and T. Van Voorhis, "A relaxation method for nonconvex quadratically constrained quadratic programs," *J. Global Optim.*, vol. 6, no. 3, pp. 215–230, 1995.
- [26] W. Ma, P. Ching, and Z. Ding, "Semidefinite relaxation based multiuser detection for M-ary PSK multiuser systems," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2862–2872, 2004.
- [27] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Mag.*, vol. 27, no. 3, pp. 20–34, Mar. 2010.
- [28] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, 3rd ed. Berlin, Germany: Springer-Verlag, 1996.
- [29] N. Sahinidis and M. Tawarmalani, "Accelerating branch-and-bound through a modeling language construct for relaxation-specific constraints," *J. Global Optim.*, vol. 32, no. 2, pp. 259–280, 2005.
- [30] R. H. Myers, *Classical and Modern Regression with Applications*. Belmont, CA: Duxbury Press, 1990, vol. 488.
- [31] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.
- [32] B. Lu, X. Wang, and K. Narayanan, "LDPC-based space-time coded OFDM systems over correlated fading channels: Performance analysis and receiver design," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 74–88, Jan. 2002.
- [33] M. Valenti, "Iterative solutions coded modulation library (ISICML)." [Online]. Available: <http://www.iterativesolutions.com/Matlab.htm>
- [34] J. Kermaol, L. Schumacher, K. Pedersen, P. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Jun. 2002.
- [35] X. Bao, N. Sahinidis, and M. Tawarmalani, "Multiterm polyhedral relaxations for nonconvex, quadratically constrained quadratic programs," *Optim. Methods and Softw.*, vol. 24, no. 4, pp. 485–504, 2009.
- [36] R. Horst, "A general class of branch-and-bound methods in global optimization with some new approaches for concave minimization," *J. Optim. Theory Appl.*, vol. 51, no. 2, pp. 271–291, 1986.



**Weiliang Zeng** (S'08) received the B.S. degree in electronic engineering (with the Highest Hons.) from the University of Electronic Sciences and Technology of China (UESTC), Chengdu, China, in 2007. Since September 2007, he has been working towards his Ph.D. degree at the Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, China. During his doctoral studies, he conducted cooperative research at Missouri University of Science and Technology, Rolla, MO.

His research interests include communications and information theory with special emphasis on wireless communications and signal processing. He is the recipient of First-Class Scholarship for graduate students at Tsinghua University.



**Chengshan Xiao** (M'99–SM'02–F'10) received the B.S. degree in electronic engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 1987, the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1989, and the Ph.D. degree in electrical engineering from the University of Sydney, Sydney, Australia, in 1997.

From 1989 to 1993, he was with the Department of Electronic Engineering, Tsinghua University, where he was a member of Research Staff and

then a Lecturer. From 1997 to 1999, he was a Senior Member of Scientific Staff with Nortel, Ottawa, ON, Canada. From 1999 to 2000, he was a Faculty Member with the University of Alberta, Edmonton, AB, Canada. From 2000 to 2007, he was with the University of Missouri, Columbia, where he was an Assistant Professor and then an Associate Professor. He is currently a Professor with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla (formerly University of Missouri, Rolla). His research interests include wireless communications, signal processing, and underwater acoustic communications. He is the holder of three U.S. patents. His algorithms have been implemented into Nortel's base station radios after successful technical field trials and network integration.

Dr. Xiao is the Editor-in-Chief of the IEEE Transactions on Wireless Communications, a Member of the Fellow Evaluation Committee of IEEE Communications Society (ComSoc), a Member at Large of IEEE ComSoc Board of Governors, a Distinguished Lecturer of IEEE ComSoc, and a Distinguished Lecturer of the IEEE Vehicular Technology Society. Previously, he served as the founding Area Editor for Transmission Technology of the IEEE Transactions on Wireless Communications; an Associate Editor of the IEEE Transactions on Vehicular Technology, the IEEE Transactions on Circuits and Systems-I, and the international journal of Multidimensional Systems and Signal Processing. He was the Technical Program Chair of the 2010 IEEE International Conference on Communications (ICC), the Lead Co-chair of the 2008 IEEE ICC Wireless Communications Symposium, and a Phy/MAC Program Co-chair of the 2007 IEEE Wireless Communications and Networking Conference. He served as the founding Chair of the IEEE Technical Committee on Wireless Communications.

**Jianhua Lu** (M'98–SM'07) received B.S.E.E. and M.S.E.E. degrees from Tsinghua University, Beijing, China, in 1986 and 1989, respectively, and his Ph.D. degree in electrical and electronic engineering from Hong Kong University of Science and Technology, Kowloon. Since 1989 he has been with the Department of Electronic Engineering, Tsinghua University, where he is currently a professor. His research interests include broadband wireless communication, multimedia signal processing, satellite communication, and wireless networking. He has

published more than 180 technical papers in international journals and conference proceedings. He has been an active member of professional societies. He was one of the recipients of best paper awards at the International Conference on Communications, Circuits and Systems 2002 and ChinaCom 2006, and was awarded the National Distinguished Young Scholar Fund by the NSF committee of China in 2006. He has served in numerous IEEE conferences as a member of Technical Program Committees and served as Lead Chair of the General Symposium of IEEE ICC 2008, as well as a Program Committee Co-Chair of the 9th IEEE International Conference on Cognitive Informatics. He is a Senior Member of the IEEE Signal Processing Society.



**Khaled B. Letaief** (S'85–M'86–SM'97–F'03) received the BS degree with distinction in Electrical Engineering from Purdue University at West Lafayette, Indiana, USA, in December 1984. He received the MS and Ph.D. Degrees in Electrical Engineering from Purdue University, in August 1986, and May 1990, respectively. From January 1985 and as a Graduate Instructor in the School of Electrical Engineering at Purdue University, he has taught courses in communications and electronics.

From 1990 to 1993, he was a faculty member at the University of Melbourne, Australia. Since 1993, he has been with the Hong Kong University of Science & Technology (HKUST) where he is currently the Dean of Engineering. He is also Chair Professor of Electronic and Computer Engineering as well as the Director of the Hong Kong Telecom Institute of Information Technology. His current research interests include wireless and mobile networks, Broadband wireless access, Cooperative networks, Cognitive radio, and Beyond 3G systems. In these areas, he has over 400 journal and conference papers and given invited keynote talks as well as courses all over the world. He has also 3 granted patents and 10 pending US patents.

Dr. Letaief served as consultants for different organizations and is the founding Editor-in-Chief of the IEEE Transactions on Wireless Communications. He has served on the editorial board of other prestigious journals including the IEEE Journal on Selected Areas in Communications - Wireless Series (as Editor-in-Chief). He has been involved in organizing a number of major international conferences and events. These include serving as the Co-Technical Program Chair of the 2004 IEEE International Conference on Communications, Circuits and Systems, ICCS'04; General Co-Chair of the 2007 IEEE Wireless Communications and Networking Conference, WCNC'07; Technical Program Co-Chair of the 2008 IEEE International Conference on Communication, ICC'08, Vice General Chair of the 2010 IEEE International Conference on Communication, ICC'10, and General Co-Chair of 2011 IEEE Technology Time machine, TTM'11.

He served as an elected member of the IEEE Communications Society Board of Governors, IEEE Distinguished lecturer, and Vice-President for Conferences of the IEEE Communications Society. He also served as the Chair of the IEEE Communications Society Technical Committee on Wireless Communications, Chair of the Steering Committee of the IEEE Transactions on Wireless Communications, and Chair of the 2008 IEEE Technical Activities/Member and Geographic Activities Visits Program. He served as member of the IEEE Communications Society and IEEE Vehicular Technology Society Fellow Evaluation Committees as well as member of the IEEE Technical Activities Board/PSPB Products & Services Committee.

He is the recipient of many distinguished awards including the Michael G. Gale Medal for Distinguished Teaching (Highest university-wide teaching award at HKUST); 2007 IEEE Communications Society Publications Exemplary Award, 2009 IEEE Marconi Prize Award in Wireless Communications, 2010 Outstanding Electrical and Computer Engineer Award by Purdue University, 2011 IEEE Communications Society Harold Sobol Award, 2011 IEEE Communications Society Wireless Communications Committee Recognition Award, and 9 IEEE Best Paper Awards.

Dr. Letaief is a Fellow of IEEE and is currently serving as Member of the IEEE Product Services and Publications Board, and the Treasurer of the IEEE Communications Society. He is also recognized by Thomson Reuters as an ISI Highly Cited Researcher.