

# Generalized Quadratic Matrix Programming: A Unified Approach For Linear Precoder Design

Juening Jin\*, Yahong Rosa Zheng†, Wen Chen\* and Chengshan Xiao†

\*Department of Electronic Engineering, Shanghai Jiao Tong University

Shanghai, 200240, China, Email: jueningjin@gmail.com, wenchen@sjtu.edu.cn

†Department of Electrical and Computer Engineering, Missouri University of Science and Technology

Rolla, MO 65409, USA, Email: zhengyr@mst.edu, xiaoc@mst.edu

**Abstract**—This paper investigates a new class of nonconvex optimization, which provides a unified framework for linear precoder design. The new optimization is called generalized quadratic matrix programming (GQMP). Due to the non-deterministic polynomial time (NP)-hardness of GQMP problems, we provide a polynomial time algorithm that is guaranteed to converge to a Karush-Kuhn-Tucker (KKT) point. In terms of application, we consider the linear precoder design problem for spectrum-sharing secure broadcast channels. We design linear precoders to maximize the average secrecy sum rate with finite-alphabet inputs and statistical channel state information (CSI). The precoder design problem is a GQMP problem and we solve it efficiently by our proposed algorithm. A numerical example is also provided to show the efficacy of our algorithm.

## I. INTRODUCTION

Convex optimization is a broad class of optimization problems that can be solved efficiently both in theory and in practice. In contrast, nonconvex optimization problems are quite challenging to deal with due to the lack of verifiable global optimality conditions. Although solving nonconvex optimization problems is a proven difficult task, much progress has been made for specific classes of nonconvex optimization, such as quadratic constrained quadratic programming (QCQP) [1] and difference of convex functions (DC) programming [2], by means of convex optimization approaches. However, many nonconvex optimization problems with special structures in communications cannot be cast as QCQP or DC problems, thus it is important to design new algorithms that can capture the underlying structure of those problems.

In this paper, we investigate a new class of nonconvex optimization, which provides a unified framework for linear precoder design. The new optimization is a generalization of quadratic matrix programming problems [3], and we call it generalized quadratic matrix programming (GQMP). A GQMP problem is defined as maximizing a generalized quadratic matrix function subject to generalized quadratic matrix inequality constraints, where both objective and constraints are nonconvex functions. Since GQMP problems are NP-hard in general, we develop a numerical algorithm to solve GQMP problems suboptimally with polynomial time complexity. The solution obtained by our proposed algorithm satisfies the KKT optimality conditions. We further analyze the computational complexity of the proposed algorithm and discuss the details of algorithm implementation.

The GQMP algorithm is then applied to design linear precoders in spectrum-sharing secure broadcast channels, where a secondary-user transmitter (ST) communicates with  $I$  secondary-user receivers (SRs) in the presence of  $J$  eavesdroppers (EDs) and subject to interference threshold constraints at  $K$  primary-user receivers (PRs). Each node in the system has multiple antennas. We address the fundamental problem of maximizing the average secrecy sum rate of the secondary users through linear precoding under finite-alphabet inputs and statistical CSI. The precoder design problem is a GQMP problem, and it can be solved by our proposed algorithm. Finally, we present a numerical example to show that our proposed precoding algorithm significantly outperforms the conventional Gaussian precoding method.

**Notations:** Boldface lowercase letters, boldface uppercase letters, and calligraphic letters are used to denote vectors, matrices and sets, respectively. The real and complex number fields are denoted by  $\mathbb{R}$  and  $\mathbb{C}$ , respectively. The space of Hermitian  $n \times n$  matrices is denoted by  $\mathbb{H}^n$ . The superscripts  $(\cdot)'$  and  $(\cdot)^\dagger$  represent transpose and Hermitian operations, respectively.  $\text{tr}(\cdot)$  is the trace of a matrix;  $[a]^+$  denotes  $\max(a, 0)$ ;  $\mathbf{A}^{(+)}$  denotes the positive definite part of a Hermitian matrix  $\mathbf{A}$ , i.e.,  $\mathbf{A}^{(+)} = \sum_{\lambda_i > 0} \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger$ , where  $\lambda_i$  is the  $i$ -th eigenvalue of  $\mathbf{A}$ , and  $\mathbf{u}_i$  is the corresponding eigenvector of  $\mathbf{A}$ ;  $\mathbf{A}^{(-)}$  denotes the negative definite part of  $\mathbf{A}$ , i.e.,  $\mathbf{A}^{(-)} = \sum_{\lambda_i < 0} \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger$ ;  $\|\cdot\|$  denotes the Euclidean norm of a vector;  $E_{\mathbf{x}}(\cdot)$  represents the statistical expectation with respect to  $\mathbf{x}$ ;  $\mathbf{I}$  and  $\mathbf{0}$  denote an identity matrix and a zero matrix, respectively, with appropriate dimensions;  $\mathbf{A} \succeq \mathbf{B}$  represents  $\mathbf{A} - \mathbf{B}$  is positive semidefinite; The symbol  $\mathcal{I}(\cdot)$  represents the mutual information;  $\log(\cdot)$  and  $\ln(\cdot)$  are used for the base two logarithm and natural logarithm, respectively.

## II. GENERALIZED QUADRATIC MATRIX PROGRAMMING

A real-valued function  $h(\mathbf{X})$  is said to be a composite quadratic matrix function if  $h(\mathbf{X})$  can be expressed in the form

$$h(\mathbf{X}) = g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X}), \quad g \in \mathcal{G} \quad (1)$$

where  $\mathbf{X} \in \mathbb{C}^{n \times r}$ ,  $\mathbf{A} \in \mathbb{H}^n$ ,  $g(\mathbf{W}) : \mathbb{H}^r \rightarrow \mathbb{R}$ , and  $\mathcal{G}$  is the family of differentiable convex functions satisfying either matrix-nondecreasing (MND) condition or matrix-

nonincreasing (MNI) condition. The definition of MND and MNI are given respectively as [4, ch. 3.6.1]:

$$\text{MND} : \mathbf{W}_1 \succeq \mathbf{W}_2 \Rightarrow g(\mathbf{W}_1) \geq g(\mathbf{W}_2) \quad (2)$$

$$\text{MNI} : \mathbf{W}_1 \succeq \mathbf{W}_2 \Rightarrow g(\mathbf{W}_1) \leq g(\mathbf{W}_2). \quad (3)$$

A linear combination of composite quadratic matrix functions is called a generalized quadratic matrix function

$$f(\mathbf{X}) = \sum_{k=1}^K \alpha_k g_k(\mathbf{X}^\dagger \mathbf{A}_k \mathbf{X}) \quad (4)$$

where  $\alpha_k \in \mathbb{R}$ ,  $k = 1, 2, \dots, K$ ;  $\mathbf{A}_k \in \mathbb{H}^n$ ,  $k = 1, 2, \dots, K$ ; and  $g_k \in \mathcal{G}$ ,  $k = 1, 2, \dots, K$ .

Programming problems dealing with generalized quadratic matrix functions are called generalized quadratic matrix programming (GQMP) problems. The standard form of a GQMP problem considered in this paper is given by

$$\begin{aligned} & \underset{\mathbf{X}}{\text{maximize}} && f_0(\mathbf{X}) \\ & \text{subject to} && f_j(\mathbf{X}) \geq 0, \quad j = 1, 2, \dots, J \\ & && \mathbf{X} \in \mathcal{X} \end{aligned} \quad (5)$$

where  $\mathbf{X} \in \mathbb{C}^{n \times r}$ ,  $\mathcal{X}$  is a closed convex set, and  $f_j(\mathbf{X})$ ,  $j = 0, 1, \dots, J$  are generalized quadratic matrix functions

$$f_j(\mathbf{X}) = \sum_{k=1}^{K_j} \alpha_{jk} g_{jk}(\mathbf{X}^\dagger \mathbf{A}_{jk} \mathbf{X}). \quad (6)$$

It is worth noting that a GQMP problem can also be expressed in the minimization form, since this is equivalent to the maximization of  $-f_0(\mathbf{X})$ , which is again a generalized quadratic matrix function. Furthermore, when  $\mathcal{X} = \mathbb{C}^{n \times r}$ , and  $g_{jk}(\mathbf{W}) = \text{tr}(\mathbf{W})$  for all  $(j, k)$ , problem (5) is a nonconvex quadratic matrix programming problem [3].

The motivation behind the GQMP formulation in (5) is twofold: First, the throughput with linear precoding for various linear vector Gaussian channels [5]–[7] can be expressed as a linear combination of mutual information. Second, the mutual information in linear vector Gaussian channels is a generalized quadratic matrix function. Therefore, GQMP provides a unified framework for throughput maximization problems with linear precoding strategy. As an illustration, we consider the point-to-point vector Gaussian channel

$$\mathbf{y} = \mathbf{H}\mathbf{P}\mathbf{x} + \mathbf{n} \quad (7)$$

where  $\mathbf{H} \in \mathbb{C}^{n \times r}$  is the complex channel matrix,  $\mathbf{P} \in \mathbb{C}^{r \times r}$  is the linear precoder,  $\mathbf{n} \in \mathbb{C}^{n \times 1}$  is the independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian noise with zero-mean and unit-variance, and  $\mathbf{x} \in \mathbb{C}^{r \times 1}$  is the arbitrarily distributed channel input signal with zero-mean and covariance  $E_{\mathbf{x}}[\mathbf{x}\mathbf{x}^\dagger] = \mathbf{I}$ .

In Theorem 1 of [5], the authors presented three properties of the input-output mutual information  $\mathcal{I}(\mathbf{x}; \mathbf{y})$ :

$$\mathcal{I}(\mathbf{x}; \mathbf{y}) \text{ is a function of } \mathbf{W} = \mathbf{P}^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{P}$$

$$\nabla_{\mathbf{W}} \mathcal{I}(\mathbf{x}; \mathbf{y}) = \Phi \quad (8)$$

$$\mathcal{I}(\mathbf{x}; \mathbf{y}) \text{ is a concave function with respect to } \mathbf{W}$$

where  $\nabla_{\mathbf{W}} \mathcal{I}(\mathbf{x}; \mathbf{y})$  represents the complex gradient of  $\mathcal{I}(\mathbf{x}; \mathbf{y})$  with respect to  $\mathbf{W}$ ;  $\Phi$  is known as the minimum mean square error matrix

$$\Phi = E[(\mathbf{x} - E[\mathbf{x}|\mathbf{y}])(\mathbf{x} - E[\mathbf{x}|\mathbf{y}])^\dagger]. \quad (9)$$

The first property shows that  $\mathcal{I}(\mathbf{x}; \mathbf{y})$  is a function of  $\mathbf{W}$ , thus it can be expressed as  $\mathcal{I}(\mathbf{W})$ . According to [4, ch. 3.6.1], the second property guarantees that  $\mathcal{I}(\mathbf{W})$  is MND because  $\Phi$  is a positive semidefinite matrix. The third property implies that  $-\mathcal{I}(\mathbf{W})$  is a differentiable convex function of  $\mathbf{W}$ . Based on the definitions in (4) and (5),  $\mathcal{I}(\mathbf{P}^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{P})$  is a generalized quadratic matrix function of  $\mathbf{P}$ , and the following throughput maximization problem is a GQMP problem

$$\begin{aligned} & \underset{\mathbf{P}}{\text{maximize}} && \mathcal{I}(\mathbf{P}^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{P}) \\ & \text{subject to} && \text{tr}(\mathbf{P}^\dagger \mathbf{P}) \leq \gamma. \end{aligned} \quad (10)$$

Note that when  $\mathbf{x}$  is Gaussian distributed, the input-output mutual information can be represented as

$$\mathcal{I}(\mathbf{P}^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{P}) = \log \det(\mathbf{I} + \mathbf{P}^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{P}). \quad (11)$$

In this special case,  $-\mathcal{I}(\mathbf{W}) = -\log \det(\mathbf{I} + \mathbf{W})$ , which is indeed a convex and MNI function with respect to  $\mathbf{W}$ .

Although a lot of linear precoder design problems can be brought under the umbrella of GQMP, it is extremely difficult to solve GQMP problems due to the following reasons: first, generalized quadratic matrix functions  $f_j(\mathbf{X})$ ,  $j = 0, 1, \dots, J$  in (6) are neither convex nor concave with respect to  $\mathbf{X}$ , thus problem (5) is a purely nonconvex optimization problem; second, problem (5) is a NP-hard optimization problem because the quadratic matrix programming problem is NP-hard in general. Due to the NP-hardness of problem (5), we develop a polynomial time algorithm that is guaranteed to converge to a KKT point of (5).

### III. ALGORITHM DESIGN

To design a polynomial time algorithm for GQMP problems, we first investigate the underlying structure of composite quadratic matrix functions. For each  $g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X})$ , we provide a concave lower bound and a convex upper bound, which satisfy the following three conditions:

$$\begin{aligned} l(\mathbf{X}; \mathbf{X}_0) &\leq g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X}) \leq u(\mathbf{X}; \mathbf{X}_0) \text{ for all } \mathbf{X} \\ g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X}) &= l(\mathbf{X}; \mathbf{X}_0) = u(\mathbf{X}; \mathbf{X}_0) \text{ when } \mathbf{X} = \mathbf{X}_0 \\ \nabla_{\mathbf{X}} g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X}) &= \nabla_{\mathbf{X}} l(\mathbf{X}; \mathbf{X}_0) = \nabla_{\mathbf{X}} u(\mathbf{X}; \mathbf{X}_0) \text{ when } \mathbf{X} = \mathbf{X}_0 \end{aligned} \quad (12)$$

where  $\mathbf{X}_0 \in \mathbb{C}^{n \times r}$  is an arbitrary matrix;  $l(\mathbf{X}; \mathbf{X}_0)$  and  $u(\mathbf{X}; \mathbf{X}_0)$  serve as the concave lower bound and the convex upper bound of  $g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X})$ , respectively. The conditions in (12) are necessary for us to design an ascent algorithm that converges to a KKT point of problem (5).

*Theorem 1:* The concave lower bound of  $g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X})$  is given by

$$l(\mathbf{X}; \mathbf{X}_0) = \text{tr}(\mathbf{L}^\dagger \mathbf{G}) + g(\mathbf{X}_0^\dagger \mathbf{A} \mathbf{X}_0) - \text{tr}(\mathbf{X}_0^\dagger \mathbf{A} \mathbf{X}_0 \mathbf{G}) \quad (13)$$

with

$$\mathbf{L} = \begin{cases} \mathbf{L}_1, & g(\mathbf{W}) \text{ is MND} \\ \mathbf{L}_2, & g(\mathbf{W}) \text{ is MNI} \end{cases} \quad (14)$$

where  $\mathbf{L}_1 = \mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X} + \mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X}_0 + \mathbf{X}_0^\dagger \mathbf{A}^{(+)} \mathbf{X} - \mathbf{X}_0^\dagger \mathbf{A}^{(+)} \mathbf{X}_0$ ,  $\mathbf{L}_2 = \mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X} + \mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X}_0 + \mathbf{X}_0^\dagger \mathbf{A}^{(-)} \mathbf{X} - \mathbf{X}_0^\dagger \mathbf{A}^{(-)} \mathbf{X}_0$ , and  $\mathbf{G} \in \mathbb{H}^r$  is the complex gradient of  $g(\mathbf{W})$  at  $\mathbf{W} = \mathbf{X}_0^\dagger \mathbf{A} \mathbf{X}_0$ , i.e.,  $\mathbf{G} = \nabla_{\mathbf{W}} g(\mathbf{X}_0^\dagger \mathbf{A} \mathbf{X}_0)$ .

*Proof:* See Appendix A. ■

**Theorem 2:** The convex upper bound of  $g(\mathbf{X}^\dagger \mathbf{A} \mathbf{X})$  is given by

$$u(\mathbf{X}; \mathbf{X}_0) = \begin{cases} g(\mathbf{U}_1), & g(\mathbf{W}) \text{ is MND} \\ g(\mathbf{U}_2), & g(\mathbf{W}) \text{ is MNI} \end{cases} \quad (15)$$

where  $\mathbf{U}_1 = \mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X} + \mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X}_0 + \mathbf{X}_0^\dagger \mathbf{A}^{(-)} \mathbf{X} - \mathbf{X}_0^\dagger \mathbf{A}^{(-)} \mathbf{X}_0$ , and  $\mathbf{U}_2 = \mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X} + \mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X}_0 + \mathbf{X}_0^\dagger \mathbf{A}^{(+)} \mathbf{X} - \mathbf{X}_0^\dagger \mathbf{A}^{(+)} \mathbf{X}_0$ .

*Proof:* The proof is omitted due to space limits. ■

Based on Theorems 1 and 2, the concave lower bound of the generalized quadratic matrix function  $f_j(\mathbf{X})$  in (6) is given by

$$\bar{f}_j(\mathbf{X}; \mathbf{X}_0) = \sum_{\alpha_{jk} > 0} \alpha_{jk} l_{jk}(\mathbf{X}; \mathbf{X}_0) + \sum_{\alpha_{jk} < 0} \alpha_{jk} u_{jk}(\mathbf{X}; \mathbf{X}_0) \quad (16)$$

where  $l_{jk}(\mathbf{X}; \mathbf{X}_0)$  and  $u_{jk}(\mathbf{X}; \mathbf{X}_0)$  represent the concave lower bound and the convex upper bound of  $g_{jk}(\mathbf{X}^\dagger \mathbf{A}_{jk} \mathbf{X})$ , respectively. Since  $l_{jk}(\mathbf{X}; \mathbf{X}_0)$  and  $u_{jk}(\mathbf{X}; \mathbf{X}_0)$  satisfy the conditions in (12), the minorize-maximization (MM) algorithm [8] can be exploited to find a KKT point of problem (5) through solving a sequence of the following concave maximization problems

$$\begin{aligned} & \underset{\mathbf{X}}{\text{maximize}} \quad \bar{f}_0(\mathbf{X}; \mathbf{X}_0) \\ & \text{subject to} \quad \bar{f}_j(\mathbf{X}; \mathbf{X}_0) \geq 0, \quad j = 1, 2, \dots, J \\ & \quad \mathbf{X} \in \mathcal{X}. \end{aligned} \quad (17)$$

In the first iteration, we solve problem (17) at initial  $\mathbf{X}_0$ , and the optimal solution is denoted as  $\bar{\mathbf{X}}_1$ . Then we replace  $\mathbf{X}_0$  in the objective and constraints of problem (17) with  $\bar{\mathbf{X}}_1$ , and solve problem (17) again. At the  $n$ -th iteration, we solve problem (17) by replacing  $\bar{f}_j(\mathbf{X}; \mathbf{X}_0)$  with  $\bar{f}_j(\mathbf{X}; \bar{\mathbf{X}}_{n-1})$ ,  $j = 0, 1, \dots, J$ , where  $\bar{\mathbf{X}}_{n-1}$  is the optimal solution of (17) at the  $(n-1)$ -th iteration. The MM algorithm for solving problem (5) is summarized in Algorithm 1.

The convergence of Algorithm 1 is presented by the following propositions.

**Proposition 1:** The solution obtained by Algorithm 1 satisfies the KKT condition of problem (5).

*Proof:* This lemma follows from Theorem 1 of [9]. ■

**Proposition 2:** The sequence  $\{s_n\}$  generated by Algorithm 1 is monotonically increasing, i.e.,  $s_n \geq s_{n-1}$ .

*Proof:* The proof is omitted due to space limits. ■

The proposed algorithm is an iterative procedure between updating concave lower bounds and solving concave maximization problems (17). Since (17) is a concave maximization problem, it can be solved efficiently by the interior-point

### Algorithm 1 : The MM algorithm

- 1) Initialization: Given tolerance  $\epsilon > 0$ , choose an initial feasible point  $\mathbf{X}_0$ , set  $n = 1$ ,  $s_0 = f_0(\mathbf{X}_0)$ . Let  $\bar{\mathbf{X}}_n$  represent the optimal solution of (17) at the  $n$ -th iteration.
- 2) Stopping criterion: if  $|s_n - s_{n-1}| > \epsilon$  go to the next step, otherwise STOP.
- 3) Concave approximation:
  - a) set  $\mathbf{X}_0 = \bar{\mathbf{X}}_n$  and solve problem (17) to obtain  $\bar{\mathbf{X}}_{n+1}$ .
  - b) set  $s_{n+1} = f_0(\bar{\mathbf{X}}_{n+1})$  and  $\mathbf{X}_{\text{opt}} = \bar{\mathbf{X}}_{n+1}$ .
- 4) Set  $n := n + 1$  and go to step 2).
- 5) Output:  $\mathbf{X}_{\text{opt}}$ .

method. The total number of optimization variables in (17) is  $nr$ , then the complexity order for solving problem (17) is about  $\mathcal{O}((nr)^3)$  [4]. Assuming that Algorithm 1 needs to update the concave lower bounds  $T$  times, the overall complexity is then given by  $\mathcal{O}(T(nr)^3)$ .

### IV. GENERALIZATION AND IMPLEMENTATION

For a multiuser communication system, we often adopt a utility function to measure the overall system performance. In this section, we extend the original GQMP problem (5) to the following optimization problem

$$\begin{aligned} & \underset{\mathbf{X}}{\text{maximize}} \quad \min_{1 \leq j \leq l} f_j(\mathbf{X}) \\ & \text{subject to} \quad f_j(\mathbf{X}) \geq 0, \quad j = l + 1, 2, \dots, J \\ & \quad \mathbf{X} \in \mathcal{X} \end{aligned} \quad (18)$$

where  $\mathbf{X} \in \mathbb{C}^{n \times r}$ ,  $\mathcal{X}$  is a closed convex set, and  $f_j(\mathbf{X})$ ,  $j = 1, 2, \dots, J$  are generalized quadratic matrix functions

$$f_j(\mathbf{X}) = \sum_{k=1}^{K_j} \alpha_{jk} g_{jk}(\mathbf{X}^\dagger \mathbf{A}_{jk} \mathbf{X}). \quad (19)$$

Problem (18) is a GQMP problem with the min-rate utility, and a KKT point of (18) can be obtained through solving a sequence of concave maximization problems:

$$\begin{aligned} R = & \underset{\mathbf{X}}{\text{maximize}} \quad \min_{1 \leq j \leq l} \bar{f}_j(\mathbf{X}; \mathbf{X}_0) \\ & \text{subject to} \quad \bar{f}_j(\mathbf{X}; \mathbf{X}_0) \geq 0, \quad j = l + 1, 2, \dots, J \\ & \quad \mathbf{X} \in \mathcal{X} \end{aligned} \quad (20)$$

where  $R$  is the optimal value of problem (20), and  $\bar{f}_j(\mathbf{X}; \mathbf{X}_0)$  is the concave lower bound of  $f_j(\mathbf{X})$ ,  $j = 1, 2, \dots, J$ . We first replace the pointwise minimum with the log-sum-exp approximation via the following inequality [10]

$$\min_{1 \leq j \leq l} a_j + \frac{1}{\beta} \ln l \leq \frac{1}{\beta} \ln \sum_{j=1}^l \exp(\beta a_j) \leq \min_{1 \leq j \leq l} a_j, \quad \beta < 0. \quad (21)$$

By applying (21) to problem (20), we obtain a smooth concave maximization problem

$$\begin{aligned} \bar{R}(\beta) = & \underset{\mathbf{X} \in \mathcal{X}}{\text{maximize}} \quad \frac{1}{\beta} \ln \sum_{j=1}^l \exp(\beta \bar{f}_j(\mathbf{X}; \mathbf{X}_0)) \\ & \text{subject to} \quad \bar{f}_j(\mathbf{X}; \mathbf{X}_0) \geq 0, \quad j = l + 1, 2, \dots, J \end{aligned} \quad (22)$$

where  $\bar{R}(\beta)$  is the optimal value of problem (22). The relationship between problem (20) and problem (22) is revealed as follows

$$|\bar{R}(\beta) - R| < \frac{1}{|\beta|} \ln l. \quad (23)$$

Since (22) is a convex optimization problem, the optimal value  $\bar{R}(\beta)$  can be readily attained by the interior point method [4, ch. 11.3.1], and the gap between  $\bar{R}(\beta)$  and  $R$  is controlled by  $\beta$ . Subsequently, we can invoke Algorithm 1 to obtain a KKT point of problem (18).

## V. APPLICATION

We consider a spectrum-sharing secure broadcast channel depicted in Fig. 1. A secondary-user transmitter (ST) communicates with  $I$  secondary-user receivers (SRs) in the presence of  $J$  eavesdroppers (EDs) and subject to interference threshold constraints at  $K$  primary-user receivers (PRs). The channel output at the  $i$ -th SR, the  $j$ -th ED and the  $k$ -th PR are, respectively, given by

$$\begin{aligned} \mathbf{y}_i &= \mathbf{H}_i \mathbf{s} + \mathbf{n}_{h_i}, \quad i = 1, 2, \dots, I \\ \mathbf{z}_j &= \mathbf{G}_j \mathbf{s} + \mathbf{n}_{g_j}, \quad j = 1, 2, \dots, J \\ \mathbf{w}_k &= \mathbf{F}_k \mathbf{s} + \mathbf{n}_{f_k}, \quad k = 1, 2, \dots, K \end{aligned} \quad (24)$$

where  $\mathbf{H}_i \in \mathbb{C}^{N_R \times N_T}$ ,  $\mathbf{G}_j \in \mathbb{C}^{N_E \times N_T}$  and  $\mathbf{F}_k \in \mathbb{C}^{N_P \times N_T}$  are complex channel matrices from the ST to the  $i$ -th SR, the  $j$ -th ED, and the  $k$ -th PR, respectively;  $\mathbf{s} \in \mathbb{C}^{N_T \times 1}$  is the channel input at the ST;  $\mathbf{n}_{h_i} \in \mathbb{C}^{N_R \times 1}$ ,  $\mathbf{n}_{g_j} \in \mathbb{C}^{N_E \times 1}$  and  $\mathbf{n}_{f_k} \in \mathbb{C}^{N_P \times 1}$  are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian noises with zero-mean and unit-variance.

The channel input  $\mathbf{s}$  can be represented as

$$\mathbf{s} = \sum_{i=1}^I \mathbf{P}_i \mathbf{x}_i = \mathbf{P} \mathbf{x} \quad (25)$$

where  $\mathbf{P}_i \in \mathbb{C}^{N_T \times N_T}$  is the precoding matrix for the  $i$ -th SR,  $\mathbf{x}_i \in \mathbb{C}^{N_T \times 1}$  is the input data vector for the  $i$ -th SR with zero-mean and covariance  $E_{\mathbf{x}_i} [\mathbf{x}_i \mathbf{x}_i^\dagger] = \mathbf{I}$ ,  $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_I]$  and  $\mathbf{x} = [\mathbf{x}_1^\dagger, \mathbf{x}_2^\dagger, \dots, \mathbf{x}_I^\dagger]^\dagger$ .

The channel matrices considered in this paper are modeled as [11]

$$\begin{aligned} \mathbf{H}_i &= \tilde{\mathbf{H}}_i \Theta_{h_i}^{\frac{1}{2}}, \quad i = 1, 2, \dots, I \\ \mathbf{G}_j &= \tilde{\mathbf{G}}_j \Theta_{g_j}^{\frac{1}{2}}, \quad j = 1, 2, \dots, J \\ \mathbf{F}_k &= \tilde{\mathbf{F}}_k \Theta_{f_k}^{\frac{1}{2}}, \quad k = 1, 2, \dots, K \end{aligned} \quad (26)$$

where  $\tilde{\mathbf{H}}_i \in \mathbb{C}^{N_R \times N_T}$ ,  $\tilde{\mathbf{G}}_j \in \mathbb{C}^{N_E \times N_T}$  and  $\tilde{\mathbf{F}}_k \in \mathbb{C}^{N_P \times N_T}$  are random matrices with i.i.d. zero-mean unit-variance complex Gaussian entries;  $\Theta_{h_i} \in \mathbb{C}^{N_T \times N_T}$ ,  $\Theta_{g_j} \in \mathbb{C}^{N_T \times N_T}$  and  $\Theta_{f_k} \in \mathbb{C}^{N_T \times N_T}$  are transmit correlation matrices of  $\mathbf{H}_i$ ,  $\mathbf{G}_j$  and  $\mathbf{F}_k$ , respectively.

We assume that the  $i$ -th SR knows the instantaneous channel realization of  $\mathbf{H}_i$ , the  $j$ -th ED knows the instantaneous channel realization of  $\mathbf{G}_j$ , and the ST only has statistical channel state information (CSI) of all nodes in the system, i.e., the transmit

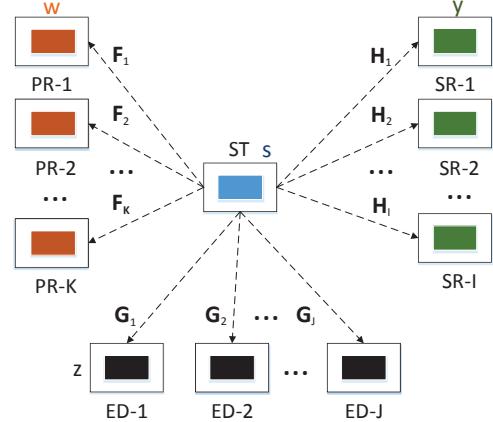


Fig. 1: System model of the spectrum-sharing secure broadcast channel.

correlation matrices  $\{\Theta_{h_i}, \Theta_{g_j}, \Theta_{f_k}, \forall (i, j, k)\}$  as well as the distributions of  $\tilde{\mathbf{H}}_i$ ,  $\tilde{\mathbf{G}}_j$  and  $\tilde{\mathbf{F}}_k$ . In addition, the  $i$ -th SR treats signals of other SRs as interference, and the  $j$ -th ED can at best decode the signal of the  $i$ -th SR while treating signals of other SRs as interference. Under these assumptions, the average secrecy rate for the  $i$ -th SR can be expressed as [12]

$$R_i = \min_{1 \leq j \leq J} [\mathcal{I}(\mathbf{x}_i; \mathbf{y}_i | \mathbf{H}_i) - \mathcal{I}(\mathbf{x}_i; \mathbf{z}_j | \mathbf{G}_j)]^+ \quad (27)$$

$$= \min_{1 \leq j \leq J} [E_{\mathbf{H}_i} \mathcal{I}(\mathbf{x}_i; \mathbf{y}_i | \mathbf{H}_i = \tilde{\mathbf{H}}_i) - E_{\mathbf{G}_j} \mathcal{I}(\mathbf{x}_i; \mathbf{z}_j | \mathbf{G}_j = \tilde{\mathbf{G}}_j)]^+ \quad (28)$$

where  $\tilde{\mathbf{H}}_i$  and  $\tilde{\mathbf{G}}_j$  in (28) are given instantaneous channel realizations of  $\mathbf{H}_i$  and  $\mathbf{G}_j$ , respectively. For notational simplicity, we omit the given channel realization condition in mutual information expressions and the following average secrecy sum rate is achievable

$$R_{BC} = \sum_{i=1}^I \min_{1 \leq j \leq J} [E_{\mathbf{H}_i} \mathcal{I}(\mathbf{x}_i; \mathbf{y}_i) - E_{\mathbf{G}_j} \mathcal{I}(\mathbf{x}_i; \mathbf{z}_j)]^+. \quad (29)$$

We maximize  $R_{BC}$  under the power constraint at the ST and the interference threshold constraints at PRs. The average transmit power at the ST is constrained to  $\gamma_0$

$$E_{\mathbf{x}} \text{tr}(\mathbf{P} \mathbf{x} \mathbf{x}^\dagger \mathbf{P}^\dagger) = \text{tr}(\mathbf{P}^\dagger \mathbf{P}) \leq \gamma_0 \quad (30)$$

and the average interference power at the  $k$ -th PR is limited by  $N_T \cdot \gamma_k$

$$\begin{aligned} &E_{\mathbf{x}, \mathbf{F}_k} \text{tr}(\mathbf{F}_k \mathbf{P} \mathbf{x} \mathbf{x}^\dagger \mathbf{P}^\dagger \mathbf{F}_k^\dagger) \\ &= E_{\tilde{\mathbf{F}}_k} \text{tr}(\mathbf{P}^\dagger (\Theta_{f_k}^{\frac{1}{2}})^\dagger \tilde{\mathbf{F}}_k^\dagger \tilde{\mathbf{F}}_k \Theta_{f_k}^{\frac{1}{2}} \mathbf{P}) \\ &= N_T \cdot \text{tr}(\mathbf{P}^\dagger \Theta_{f_k} \mathbf{P}) \leq N_T \cdot \gamma_k, \quad \forall k. \end{aligned} \quad (31)$$

The second equality in (31) holds because 1) the entries of  $\tilde{\mathbf{F}}_k$  are i.i.d. complex Gaussian variables with zero-mean and unit-variance; 2)  $\tilde{\mathbf{F}}_k$  and  $\mathbf{s}$  are independent. Then the linear precoder design problem can be formulated as

$$\begin{aligned} &\underset{\mathbf{P}}{\text{maximize}} \quad R_{BC} \\ &\text{subject to} \quad \text{tr}(\mathbf{P}^\dagger \mathbf{P}) \leq \gamma_0 \\ &\quad \text{tr}(\mathbf{P}^\dagger \Theta_{f_k} \mathbf{P}) \leq \gamma_k, \quad \forall k. \end{aligned} \quad (32)$$

We solve problem (32) under finite-alphabet inputs by GQMP. Instead of Gaussian inputs, we assume that each symbol of the input data vector  $\mathbf{x}$  is uniformly distributed from a  $Q$ -ary equiprobable discrete constellation set. Then the average constellation-constrained mutual information  $E_{\mathbf{H}_i}\mathcal{I}(\mathbf{x}_i; \mathbf{y}_i)$  and  $E_{\mathbf{G}_j}\mathcal{I}(\mathbf{x}_i; \mathbf{z}_j)$  can be expressed as [7]

$$E_{\mathbf{H}_i}\mathcal{I}(\mathbf{x}_i; \mathbf{y}_i) = \frac{1}{L} \sum_{m=1}^L E_{\mathbf{H}_i, \mathbf{n}_{h_i}} \{ \log(a_m^{ii}) \} \quad (33)$$

$$E_{\mathbf{G}_j}\mathcal{I}(\mathbf{x}_i; \mathbf{z}_j) = \frac{1}{L} \sum_{m=1}^L E_{\mathbf{G}_j, \mathbf{n}_{g_j}} \{ \log(b_m^{ij}) \} \quad (34)$$

with

$$a_m^{ii} = \frac{\sum_{n=1}^L \exp(-\|\mathbf{H}_i \mathbf{P} \mathbf{I}_i \mathbf{e}_{mn} + \mathbf{n}_{h_i}\|^2 + \|\mathbf{n}_{h_i}\|^2)}{\sum_{n=1}^L \exp(-\|\mathbf{H}_i \mathbf{P} \mathbf{e}_{mn} + \mathbf{n}_{h_i}\|^2 + \|\mathbf{n}_{h_i}\|^2)} \quad (35)$$

$$b_m^{ij} = \frac{\sum_{n=1}^L \exp(-\|\mathbf{G}_j \mathbf{P} \mathbf{I}_i \mathbf{e}_{mn} + \mathbf{n}_{g_j}\|^2 + \|\mathbf{n}_{g_j}\|^2)}{\sum_{n=1}^L \exp(-\|\mathbf{G}_j \mathbf{P} \mathbf{e}_{mn} + \mathbf{n}_{g_j}\|^2 + \|\mathbf{n}_{g_j}\|^2)} \quad (36)$$

where the constant  $L$  is equal to  $Q^{N_T I}$ ;  $\mathbf{I}_i$  is a block diagonal matrix formed by replacing the  $i$ -th  $N_T \times N_T$  block diagonal entry of the  $N_T I \times N_T I$  identity matrix  $\mathbf{I}$  with  $\mathbf{0}$ ;  $\mathbf{e}_{mn}$  is the difference between  $\mathbf{x}_m$  and  $\mathbf{x}_n$ , with  $\mathbf{x}_m$  and  $\mathbf{x}_n$  representing two possible input data vectors from  $\mathbf{x}$ .

The computational complexity for evaluating the above average mutual information and its gradient is prohibitively high because both (33) and (34) as well as their gradients have no closed-form expressions. To overcome this difficulty, we adopt an accurate approximation of the average constellation-constrained mutual information [13]. The average secrecy sum rate  $R_{BC}$  can then be approximated as

$$\sum_{i=1}^I \min_{1 \leq j \leq J} [\mathcal{I}_A(\mathbf{x}_i; \mathbf{y}_i) - \mathcal{I}_A(\mathbf{x}_i; \mathbf{z}_j)]^+ \quad (37)$$

where  $\mathcal{I}_A(\mathbf{x}_i; \mathbf{y}_i)$  and  $\mathcal{I}_A(\mathbf{x}_i; \mathbf{z}_j)$  are accurate approximations of  $E_{\mathbf{H}_i}\mathcal{I}(\mathbf{x}_i; \mathbf{y}_i)$  and  $E_{\mathbf{G}_j}\mathcal{I}(\mathbf{x}_i; \mathbf{z}_j)$ , respectively

$$\mathcal{I}_A(\mathbf{x}_i; \mathbf{y}_i) = g(\mathbf{I}_i^\dagger \mathbf{P}^\dagger \Theta_{h_i} \mathbf{P} \mathbf{I}_i, N_R) - g(\mathbf{P}^\dagger \Theta_{h_i} \mathbf{P}, N_R) \quad (38)$$

$$\mathcal{I}_A(\mathbf{x}_i; \mathbf{z}_j) = g(\mathbf{I}_i^\dagger \mathbf{P}^\dagger \Theta_{g_j} \mathbf{P} \mathbf{I}_i, N_E) - g(\mathbf{P}^\dagger \Theta_{g_j} \mathbf{P}, N_E) \quad (39)$$

with

$$g(\mathbf{W}, N) = \frac{1}{L} \sum_{m=1}^L \log \sum_{n=1}^L \left( 1 + \frac{1}{2} \mathbf{e}_{mn}^\dagger \mathbf{W} \mathbf{e}_{mn} \right)^{-N} \quad (40)$$

Using (37) to replace  $R_{BC}$ , problem (32) can be approximated as

$$\underset{\mathbf{P} \in \mathcal{P}}{\text{maximize}} \quad \sum_{i=1}^I \min_{1 \leq j \leq J} [\mathcal{I}_A(\mathbf{x}_i; \mathbf{y}_i) - \mathcal{I}_A(\mathbf{x}_i; \mathbf{z}_j)]^+ \quad (41)$$

where  $\mathcal{P}$  is the set of all feasible precoding matrices

$$\mathcal{P} = \{ \mathbf{P} | \text{tr}(\mathbf{P}^\dagger \mathbf{P}) \leq \gamma_0, \text{tr}(\mathbf{P}^\dagger \Theta_{f_k} \mathbf{P}) \leq \gamma_k, \forall k \}. \quad (42)$$

The operator  $[.]^+$  is a major barrier for solving problem (41). The following proposition shows that the optimal value of (41) can be obtained via solving a sequence of the following optimization problem

$$r_l = \underset{\mathbf{P} \in \mathcal{P}}{\text{maximize}} \quad \sum_{i \in \mathcal{S}_l} \min_{1 \leq j \leq J} [\mathcal{I}_A(\mathbf{x}_i; \mathbf{y}_i) - \mathcal{I}_A(\mathbf{x}_i; \mathbf{z}_j)] \quad (43)$$

where  $r_l$  is the optimal value of (43),  $\mathcal{S}_l$  represents the  $l$ -th non-empty subset of  $\{1, 2, \dots, I\}$ .

*Proposition 3:* Let  $r_{bc}$  denote the optimal value of problem (41). Then we have

$$r_{bc} = \max_{1 \leq l \leq 2^I - 1} r_l \quad (44)$$

*Proof:* The proof is omitted due to space limits. ■

The physical meaning of each subproblem (43) is interpreted as follows. The precoders designed by (43) attempt to guarantee secure transmission for the SRs in  $\mathcal{S}_l$ . The precoded signals of other SRs, with individual average secrecy rates being zeros, are used to confuse the EDs.

Problem (43) is a GQMP problem with the min-rate utility due to the following proposition.

*Proposition 4:*  $g(\mathbf{W}, N)$  is a convex and MNI function with respect to  $\mathbf{W}$ .

*Proof:* The proof is omitted due to space limitations. ■

Based on Proposition 4,  $\mathcal{I}_A(\mathbf{x}_i; \mathbf{y}_i) - \mathcal{I}_A(\mathbf{x}_i; \mathbf{z}_j)$  is a generalized quadratic matrix function for all  $(i, j)$ . Therefore, problem (43) is a GQMP problem with the min-rate utility, and it can be solved efficiently by Algorithm 1.

Finally, we provide a numerical example to demonstrate the efficacy of Algorithm 1. We consider a spectrum-sharing secure broadcast channel with one ST, two SRs, two EDs and two PRs. Each node in the system has two antennas. The transmit correlation matrices are given by

$$\Theta_{h_1} = \mathbf{R}(0.9), \Theta_{h_2} = \mathbf{R}(0.8), \Theta_{g_1} = \mathbf{R}(0.45), \quad (45)$$

$$\Theta_{g_2} = \mathbf{R}(0.55), \Theta_{f_1} = \mathbf{R}(0.2), \Theta_{f_2} = \mathbf{R}(0.6) \quad (46)$$

where  $\mathbf{R}(\rho)$  is described from the exponential correlation matrix model

$$[\mathbf{R}(\rho)]_{ij} = \rho^{|i-j|}, \quad \rho \in [0, 1]. \quad (47)$$

The interference thresholds at PRs are 10dB less than the total transmit power, i.e.,  $\gamma_1 = \gamma_2 = 0.1\gamma_0$ , and the SNR is defined as  $\text{SNR} = \gamma_0$ .

Fig. 2 depicts the comparison results with the conventional Gaussian precoding under BPSK and QPSK modulations. The Gaussian precoding method solves problem (32) under Gaussian signaling by GQMP, and then evaluates the average secrecy sum rate with the corresponding optimal precoders and finite-alphabet inputs. Results in Fig. 2 indicate that our proposed precoding algorithm offers much higher secrecy sum rate than the Gaussian precoding method, especially in the medium and high SNR regimes. This is due to the signal mismatch between ideal Gaussian signals and practical finite-alphabet signals. The Gaussian precoding method only

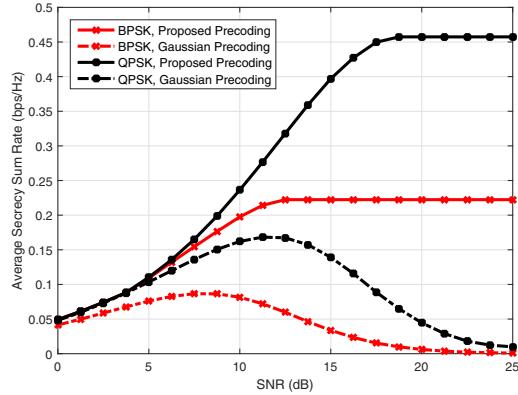


Fig. 2: Proposed precoding versus Gaussian precoding under different modulations.

allocates power to the 1-th SR, and the precoder for the 2-th SR is 0. In contrast, our proposed precoding algorithm allocates power to both SRs, and the precoded signal for the 2-th SR acts as the jamming signal to further confuse EDs. Therefore, our proposed precoding significantly outperforms the Gaussian precoding method.

## VI. CONCLUSION

We introduced a new class of nonconvex optimization called generalized quadratic matrix programming, which unifies linear precoding design in various MIMO Gaussian channels with arbitrary input distributions. In order to solve GQMP problems, we have proposed a polynomial time algorithm which is guaranteed to converge to a KKT point. Then we have considered the application of GQMP on linear precoder design with finite-alphabet inputs in the spectrum-sharing secure broadcast channel. The linear precoder design problem is a GQMP problem, and we solved it efficiently by our proposed algorithm. A numerical result has been shown to demonstrate that our proposed precoding algorithm provides a significant performance gain over the Gaussian precoding method.

## APPENDIX A PROOF OF THEOREM 1

*Proof:* Since  $g(\mathbf{W})$  is a differentiable convex function, it can be lower bounded by its first-order Taylor approximation at any point  $\mathbf{W}_0 = \mathbf{X}_0^\dagger \mathbf{A} \mathbf{X}_0$  [4, ch. 3.1.3]:

$$g(\mathbf{W}) \geq g(\mathbf{W}_0) + \text{tr}[(\mathbf{W} - \mathbf{W}_0)^\dagger \nabla_{\mathbf{W}} g(\mathbf{W}_0)]. \quad (48)$$

By plugging  $\mathbf{W} = \mathbf{X}^\dagger \mathbf{A} \mathbf{X}$  and  $\mathbf{G} = \nabla_{\mathbf{W}} g(\mathbf{W}_0)$  into (48), we have

$$\text{tr}(\mathbf{X}^\dagger \mathbf{A} \mathbf{X}) \geq \text{tr}(\mathbf{X}^\dagger \mathbf{A} \mathbf{X} \mathbf{G}) + g(\mathbf{W}_0) - \text{tr}(\mathbf{W}_0 \mathbf{G}). \quad (49)$$

Note that  $\text{tr}(\mathbf{X}^\dagger \mathbf{A} \mathbf{X} \mathbf{G})$  can be expressed alternatively as

$$\text{tr}(\mathbf{X}^\dagger \mathbf{A} \mathbf{X} \mathbf{G}) = \text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X} \mathbf{G}) + \text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X} \mathbf{G}) \quad (50)$$

$$= \text{vec}(\mathbf{X})^\dagger \cdot [(\mathbf{G}' \otimes \mathbf{A}^{(+)}) + (\mathbf{G}' \otimes \mathbf{A}^{(-)})] \cdot \text{vec}(\mathbf{X}). \quad (51)$$

If  $g(\mathbf{W})$  is MND, the complex gradient  $\mathbf{G} \succeq \mathbf{0}$  for all  $\mathbf{X}_0$  [4, ch. 3.6.1]. Therefore,  $\mathbf{G}' \otimes \mathbf{A}^{(+)} \succeq \mathbf{0}$  and  $\mathbf{G}' \otimes \mathbf{A}^{(-)} \preceq \mathbf{0}$  [14], which imply the convexity of  $\text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X} \mathbf{G})$

and the concavity of  $\text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X} \mathbf{G})$ . The convex function  $\text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X} \mathbf{G})$  can be further lower bounded by its first-order Taylor approximation at  $\mathbf{X}_0$ :

$$\begin{aligned} & \text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X} \mathbf{G}) \\ & \geq \text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X}_0 + \mathbf{X}_0^\dagger \mathbf{A}^{(+)} \mathbf{X} - \mathbf{X}_0^\dagger \mathbf{A}^{(+)} \mathbf{X}_0 \mathbf{G}). \end{aligned} \quad (52)$$

Combining (49), (50) and (52), we obtain a concave lower bound when  $g(\mathbf{W})$  is MND.

If  $g(\mathbf{W})$  is MNI,  $\mathbf{G} \preceq \mathbf{0}$  for all  $\mathbf{X}_0$  [4, ch. 3.6.1]. Therefore,  $\text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(+)} \mathbf{X} \mathbf{G})$  is a concave function of  $\mathbf{X}$  and  $\text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X} \mathbf{G})$  is a convex function of  $\mathbf{X}$ . The convex function  $\text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X} \mathbf{G})$  can be further lower bounded by its first-order Taylor approximation at  $\mathbf{X}_0$ :

$$\begin{aligned} & \text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X} \mathbf{G}) \\ & \geq \text{tr}(\mathbf{X}^\dagger \mathbf{A}^{(-)} \mathbf{X}_0 + \mathbf{X}_0^\dagger \mathbf{A}^{(-)} \mathbf{X} - \mathbf{X}_0^\dagger \mathbf{A}^{(-)} \mathbf{X}_0 \mathbf{G}). \end{aligned} \quad (53)$$

Combining (49), (50) and (53), we obtain a concave lower bound when  $g(\mathbf{W})$  is MNI. ■

## ACKNOWLEDGMENT

This work was supported in part by US National Science Foundation under grants ECCS-1231848, ECCS-1408316 and ECCS-1539316. This work has been carried out while Mr. Juening Jin is visiting Missouri University of Science and Technology.

## REFERENCES

- [1] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, “Semidefinite relaxation of quadratic optimization problems,” *IEEE Signal Process. Mag.*, vol. 27, no. 3, p. 20, May 2010.
- [2] P. D. Tao *et al.*, “The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems,” *Ann. Oper. Res.*, vol. 133, no. 1-4, pp. 23–46, Jan. 2005.
- [3] A. Beck, “Quadratic matrix programming,” *SIAM J. Optim.*, vol. 17, no. 4, pp. 1224–1238, Jan. 2007.
- [4] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [5] 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.
- [6] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, “Linear precoding for finite-alphabet signaling over MIMOME wiretap channels,” *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2599–2612, Jul. 2012.
- [7] Y. Wu, M. Wang, C. Xiao, Z. Ding, and X. Gao, “Linear precoding for MIMO broadcast channels with finite-alphabet constraints,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2906–2920, Aug. 2012.
- [8] D. R. Hunter and K. Lange, “A tutorial on MM algorithms,” *Amer. Statist.*, vol. 58, no. 1, pp. 30–37, Feb. 2004.
- [9] B. R. Marks and G. P. Wright, “A general inner approximation algorithm for nonconvex mathematical programs,” *Oper. Res.*, vol. 26, no. 4, pp. 681–683, Jul. 1978.
- [10] M. Chen, S. C. Liew, Z. Shao, and C. Kai, “Markov approximation for combinatorial network optimization,” *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6301–6327, Oct. 2013.
- [11] C. Xiao, J. Wu, S.-Y. Leong, Y. R. Zheng, and K. Letaief, “A discrete-time model for triply selective MIMO Rayleigh fading channels,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1678–1688, Sep. 2004.
- [12] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai, “Compound wiretap channels,” *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, pp. 1–12, Oct. 2009.
- [13] W. Zeng, C. Xiao, M. Wang, and J. Lu, “Linear precoding for finite-alphabet inputs over MIMO fading channels with statistical CSI,” *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3134–3148, Jun. 2012.
- [14] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1994.