TIGHTNESS OF A NEW AND ENHANCED SEMIDEFINITE RELAXATION FOR MIMO DETECTION[*]

CHENG LU[†], YA-FENG LIU[‡], WEI-QIANG ZHANG[§], AND SHUZHONG ZHANG[¶]

Abstract. In this paper, we consider a fundamental problem in modern digital communications known as multiple-input multiple-output (MIMO) detection, which can be formulated as a complex quadratic programming problem subject to unit-modulus and discrete argument constraints. Various semidefinite-relaxation-based (SDR-based) algorithms have been proposed to solve the problem in the literature. In this paper, we first show that conventional SDR is generally not tight for the problem. Then, we propose a new and enhanced SDR and show its tightness under an easily checkable condition, which essentially requires the level of the noise to be below a certain threshold. The above results have answered an open question posed by So in [*Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'10)*, SIAM, Philadelphia, PA, 2011, pp. 698–711]. Numerical simulation results show that our proposed SDR significantly outperforms the conventional SDR in terms of the relaxation gap.

Key words. complex quadratic programming, semidefinite relaxation, MIMO detection, tight relaxation

AMS subject classifications. 90C22, 90C20, 90C46, 90C27

DOI. 10.1137/17M115075X

1. Introduction. Multiple-input multiple-output (MIMO) detection is a fundamental problem in modern digital communications [46]. Mathematically, the input-output relationship of the MIMO channel can be modeled as

$$(1.1) \qquad r = Hx^* + v,$$

where

- $r \in \mathbb{C}^m$ is the vector of received signals,
- $H \in \mathbb{C}^{m \times n}$ is an $m \times n$ complex channel matrix (for n inputs and m outputs with $m \geq n$),
- $x^* \in \mathbb{C}^n$ is the vector of transmitted symbols, and
- $v \in \mathbb{C}^m$ is an additive white circularly symmetric Gaussian noise.

[†]School of Economics and Management, North China Electric Power University, Beijing, 102206, China (lucheng1983@163.com).

[‡]Corresponding author. State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China (yafliu@lsec.cc.ac.cn).

[§]Department of Electronic Engineering, Tsinghua University, Beijing, 100084, China (wqzhang@tsinghua.edu.cn).

[¶]Department of Industrial and Systems Engineering, University of Minnesota, Minneapolis, MN 55455 (zhangs@umn.edu) and Institute of Data and Decision Analytics, The Chinese University of Hong Kong, Shenzhen, 518172, China (zhangs@cuhk.edu.cn)

Assume that an M-phase-shift keying (M-PSK) modulation scheme with $M \geq 2$ is adopted. Then each entry x_i^* of x^* belongs to a finite set of symbols, i.e.,

$$x_i^* \in \left\{ e^{\mathbf{i}\theta} \mid \theta = 2j\pi/M, \ j = 0, 1, \dots, M-1 \right\}, \ i = 1, 2, \dots, n,$$

where $\mathbf{i}$ is the imaginary unit. The MIMO detection problem is to recover the vector of transmitted symbols x^* from the vector of received signals r based on the knowledge of the channel matrix H. The mathematical formulation of the problem is

$$(\text{P}) \qquad \begin{aligned} &\min_{x \in \mathbb{C}^n} &&\|Hx - r\|_2^2 \\ &\text{s.t.} &&|x_i|^2 = 1, \ \arg(x_i) \in \mathcal{A}, \ i = 1, 2, \dots, n, \end{aligned}$$

where $\|\cdot\|_2$ denotes the Euclidean norm, $\arg(\cdot)$ denotes the argument of the complex number, and

$$(1.2) \qquad \mathcal{A} = \{0, 2\pi/M, \dots, 2(M-1)\pi/M\}.$$

It has been shown that minimizing the square error in the above MIMO detection problem (P) is equivalent to minimizing the probability of vector detection error (see, e.g., [43, Chapter 3]).

There are two problems related to the MIMO detection problem (P). The first one is the following complex quadratic programming problem:

$$(\text{CQP}) \qquad \begin{aligned} &\min_{x \in \mathbb{C}^n} &&x^\dagger Q x + 2\mathrm{Re}(c^\dagger x) \\ &\text{s.t.} &&|x_i|^2 = 1, \ \arg(x_i) \in \mathcal{A}, \ i = 1, 2, \dots, n, \end{aligned}$$

where $Q \in \mathbb{C}^{n \times n}$ is a Hermitian matrix, $c \in \mathbb{C}^n$ is a complex column vector, $(\cdot)^\dagger$ denotes the conjugate transpose, and $\mathrm{Re}(\cdot)$ denotes the elementwise real part of a complex matrix/vector/number. Problem (P) is a special case of problem (CQP), where $Q = H^\dagger H$ and $c = -H^\dagger r$. The second one is the unit-modulus constrained quadratic programming problem

$$(\text{UQP}) \qquad \begin{aligned} &\min_{x \in \mathbb{C}^n} &&x^\dagger Q x + 2\mathrm{Re}(c^\dagger x) \\ &\text{s.t.} &&|x_i|^2 = 1, \ i = 1, 2, \dots, n, \end{aligned}$$

which can be seen as a continuous relaxation of problem (CQP) in the sense that the discrete argument constraints $\arg(x_i) \in \mathcal{A}$ for $i = 1, \dots, n$ in problem (CQP) are dropped.

Both problems (CQP) and (UQP) have been extensively studied due to their broad applications. For instance, the Max-Cut problem [10] and the Max-3-Cut problem [11] are special cases of problem (CQP) with homogeneous objective functions and with $M \in \{2, 3\}$; the classical binary quadratic programming problem [4] is also a special case of problem (CQP) with $M = 2$. Moreover, problems (CQP) and (UQP) have found wide applications in signal processing and wireless communications, including MIMO detection [14, 15], angular synchronization [2, 34], phase retrieval [44], and radar signal processing [28, 31, 38]. However, it is known that both problems (CQP) and (UQP) are NP-hard [47], and the MIMO detection problem, as a special case of problem (CQP), is also NP-hard [42]. Therefore, there is no polynomial-time algorithms which can solve these problems to global optimality in general, unless P = NP.

Returning to the MIMO detection problem (P), various algorithms have been proposed to solve it. The sphere decoder algorithm [6], which can be seen as a special branch-and-bound algorithm [30], has been developed to find the exact solution of problem (P). However, the worst-case complexity of the sphere decoder algorithm is exponential. To overcome the exponential complexity issue, some suboptimal algorithms have been proposed, including the zero-forcing detector [19, 32], the minimum mean-squared error detector [12, 45], and the decision feedback detector [41]. However, the performance of these low-complexity suboptimal algorithms (in terms of the vector detection error rate) is generally poor.

In the last two decades, semidefinite-relaxation-based (SDR-based) algorithms have been widely studied in the signal processing and wireless communication community [26]. For various SDR-based algorithms for problems (CQP) and (UQP) under different signal processing and wireless communication scenarios, we refer the interested reader to [2, 14, 15, 13, 18, 23, 25, 27, 28, 31, 35, 40, 44, 47] and the references therein. The SDR-based algorithms generally perform very well in some signal processing and wireless communication applications, as pointed out in [3]. Similar observations have also been made for MIMO detection [14, 16, 27], asynchronous multisensor data fusion [31], as well as angular synchronization [34]. For the above applications, SDR-based algorithms proved to be impressively effective if the so-called signal-to-noise ratio (SNR) is high. In particular, for problem (P) with $M = 2$, it has been shown in [14, 16] that SDR-based algorithms can achieve the maximum possible diversity order (under the assumption that all entries of H are independent and identically distributed (i.i.d.) Gaussian variables).

Therefore, it has been an important longstanding problem in the field to understand why the performances of the SDR-based algorithms are so remarkably good in practice. One line of research is to analyze the approximation ratios of the SDR-based algorithms. Along this line, the approximation ratios of some SDR-based algorithms have been analyzed in [36, 47] for problems (CQP) and (UQP) with homogeneous positive semidefinite objective functions. Another line of research is to identify conditions under which the SDRs are tight [2, 14, 15, 17, 18, 35]. The phenomenon that some nonconvex problems are equivalent to their convex relaxations under certain conditions can also be regarded as a type of hidden convexity [39]. It is also worth remarking that in some signal processing and wireless communication applications with high SNRs, even first-order algorithms are guaranteed to converge to the global solution of these nonconvex problems [5, 21, 22].

In this paper, we focus on the MIMO detection problem (P). For ease of presentation, we define the *tightness* of an SDR of the MIMO detection problem (P) as follows.

DEFINITION 1.1. *An SDR of problem* (P) *is called tight if the following two conditions hold:*
   − *the gap between the SDR and problem* (P) *is zero, and*
   − *the SDR recovers the true vector of transmitted signals.*

For the MIMO detection problem (P), the tightness of some SDRs has been studied in [14, 15, 17, 18, 35]. In particular, So proved in [35] that, for the case where $M = 2$, there exists a tight SDR (see (BSDP) later on) if the inputs H and v in (1.1) satisfy

$$(1.3) \qquad \lambda_{\min}\left(\mathrm{Re}(H^{\dagger}H)\right) > \|\mathrm{Re}(H^{\dagger}v)\|_{\infty},$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a given matrix and $\|\cdot\|_{\infty}$ denotes

the L_∞-norm. In [35], So also posed the following open question: Is the condition

$$(1.4) \qquad \lambda_{\min}(H^\dagger H) > \|H^\dagger v\|_\infty$$

sufficient for the conventional (complex) SDR being tight for problem (P) with $M \geq 3$?

The main contributions of our paper are twofold. First, we show that the conventional SDR is generally not tight for problem (P) and thus answer an open question posed by So. Second, we propose an enhanced SDR for problem (P), which is much tighter than the conventional SDR. We prove that our proposed enhanced SDR is tight for the case where $M \geq 3$ if the following condition is satisfied:

$$(1.5) \qquad \lambda_{\min}\left(H^\dagger H\right) \sin\left(\frac{\pi}{M}\right) > \left\|H^\dagger v\right\|_\infty.$$

To the best of our knowledge, for the case where $M \geq 3$, our new enhanced SDR is the first one to have a theoretical guarantee of tightness if the SNR of the problem is sufficiently high (or equivalently the noise level of the problem is sufficiently low). Numerical results show that the new enhanced SDR performs significantly better than the conventional SDR in terms of the relaxation gap as well as the ability to recover the vector of transmitted signals.

The rest of this paper is organized as follows. In section 2, we review the conventional SDR for problem (P) and show that it is generally not tight. Then, we propose an enhanced SDR for problem (CQP) in section 3 and prove it to be tight for problem (P)—a special case of problem (CQP)—if condition (1.5) holds in section 4. In section 5, we present some numerical results to show the effectiveness of the newly proposed SDR. Finally, we conclude the paper in section 6.

We adopt the following somewhat standard notation in this paper. For a given complex vector x, we use x_i (or $[x]_i$) to denote its ith entry, $\|x\|_2$ to denote its Euclidean norm, $\|x\|_1$ to denote its L_1-norm, $\|x\|_\infty$ to denote its L_∞-norm, and $\mathrm{Diag}(x)$ to denote the diagonal matrix with the diagonal entries being x. For a given complex Hermitian matrix A, $A \succeq 0$ means A is positive semidefinite, $\mathrm{Trace}(A)$ denotes the trace of A, $A_{i,j}$ denotes the (i,j)th entry of A, and $A^\dagger$ and A^T denote the conjugate transpose and transpose of A, respectively. For two Hermitian matrices A and B, $A \succeq B$ means $A - B \succeq 0$ and $A \bullet B$ means $\mathrm{Re}(\mathrm{Trace}(A^\dagger B))$. For any given matrix $C \in \mathbb{C}^{m \times n}$ (including the scalar case and the vector case), we use $\mathrm{Re}(C)$ and $\mathrm{Im}(C)$ to denote the componentwise real and imaginary parts of C, respectively. For a set $\mathcal{S}$, we use $\mathrm{Conv}(\mathcal{S})$ to denote its convex envelope. Finally, we use $\mathbf{i}$ to denote the imaginary unit, which satisfies the equation $\mathbf{i}^2 = -1$, and use I_n to denote the n-by-n identity matrix. In the remainder of this paper, we will focus on the MIMO detection problem. We define

$$Q = H^\dagger H, \ c = -H^\dagger r, \ \text{and} \ \lambda_{\min} = \lambda_{\min}(Q)$$

unless otherwise specified.

2. Conventional semidefinite relaxations. By introducing $X = xx^\dagger$, relaxing it to $X \succeq xx^\dagger$, and dropping the argument constraints $\arg(x_i) \in \mathcal{A}$ for all $i = 1, \ldots, n$, we get the following conventional SDR for (CQP):

$$
\begin{aligned}
\text{(CSDP)} \qquad \min_{x,\,X} \quad & Q \bullet X + 2\mathrm{Re}(c^\dagger x) \\
\text{s.t.} \quad & X_{i,i} = 1, \ i = 1, 2, \ldots, n, \\
& X \succeq xx^\dagger,
\end{aligned}
$$

where $x \in \mathbb{C}^n$ and $X \in \mathbb{C}^{n \times n}$ are variables and $X \succeq xx^\dagger$ is equivalent to

$$\begin{bmatrix} 1 & x^\dagger \\ x & X \end{bmatrix} \succeq 0.$$

The above (CSDP) is a complex semidefinite program which has been widely used in the literature. Indeed, the approximation algorithms in [2, 14, 15, 18, 27, 28, 36, 35, 44, 47] are all based on (CSDP).

In [35], So studied the tightness of an SDR for the MIMO detection problem. For the special case where $M = 2$, the argument constraints of (CQP) become the binary constraints $x \in \{-1, 1\}^n$. Therefore, (CSDP) can be posed as the following real SDR:

$$
\begin{aligned}
&\min_{x,\,X} && \mathrm{Re}(Q) \bullet X + 2\mathrm{Re}(c)^\mathsf{T} x \\
\text{(BSDP)} \quad &\text{s.t.} && X_{i,i} = 1, \ i = 1, 2, \dots, n, \\
& && X \succeq xx^\mathsf{T},
\end{aligned}
$$

where $x \in \mathbb{R}^n$ and $X \in \mathbb{R}^{n \times n}$. Based on (BSDP), So proved the following theorem.

THEOREM 2.1 (see [35]). *Suppose that $M = 2$. If the inputs H and v in* (1.1) *satisfy* (1.3), *then* (BSDP) *is tight for* (P).

For the case where $M = 2$, Theorem 2.1 proposes a sufficient condition under which (BSDP) is tight for (P). In fact, for the same case, similar sufficient conditions for (BSDP) to be tight have been proposed in the literature. For instance, the condition proposed in [17, Theorem 1] is

$$(2.1) \qquad\qquad \lambda_{\min}\left(\mathrm{Re}(H^\dagger H)\right) > \left\|\mathrm{Re}(H^\dagger v)\right\|_1$$

and the one proposed in [18] is

$$(2.2) \qquad\qquad \lambda_{\min}\left(\mathrm{Re}(H^\dagger H)\right) > \left\|\mathrm{Re}(H^\dagger v)\right\|_2.$$

Since $\|\mathrm{Re}(H^\dagger v)\|_p \geq \|\mathrm{Re}(H^\dagger v)\|_\infty$ for $p \in \{1, 2\}$, condition (1.3) is weaker than the two above.

In addition to the above sufficient conditions, the following sufficient and necessary condition for the case where $M = 2$ is also proposed in [15, Theorem 1] and [14, Theorem 7.1]: Letting $H \in \mathbb{R}^{m \times n}$, $v \in \mathbb{R}^m$, and $x^* \in \{-1, 1\}^n$, (BSDP) is tight if and only if

$$v \in \left\{ u \in \mathbb{R}^m \,|\, H^\mathsf{T} H + [\mathrm{Diag}(x^*)]^{-1}\mathrm{Diag}(H^\mathsf{T} u) \succeq 0 \right\}.$$

The above condition can be extended to the complex case [15], where $H \in \mathbb{C}^{m \times n}$, $v \in \mathbb{C}^m$, and $x^* \in \{-1, 1\}^n$ as follows:

$$v \in \left\{ u \in \mathbb{C}^m \,|\, \mathrm{Re}(H^\dagger H) + [\mathrm{Diag}(x^*)]^{-1}\mathrm{Diag}[\mathrm{Re}(H^\dagger u)] \succeq 0 \right\}.$$

It is simple to verify that condition (1.3), as well as conditions (2.1) and (2.2), can be derived from the above sufficient and necessary condition.

For more general cases where $M \geq 3$, So posed an open question in [35]: Is condition (1.4) sufficient for (CSDP) to be tight for (P)? Next, we show that the answer to this question is negative. More specifically, we show that (CSDP) is not tight for almost all instances of (P) with nonzero random noise in the sense that the probability that there is no gap between (CSDP) and (P) is zero.

To tackle the problem, we first derive a necessary condition for (CSDP) to be tight. Define

$$(2.3) \qquad \hat{Q} = \begin{bmatrix} \mathrm{Re}(Q) & -\mathrm{Im}(Q) \\ \mathrm{Im}(Q) & \mathrm{Re}(Q) \end{bmatrix}, \ \hat{c} = \begin{bmatrix} \mathrm{Re}(c) \\ \mathrm{Im}(c) \end{bmatrix}, \text{ and } y = \begin{bmatrix} \mathrm{Re}(x) \\ \mathrm{Im}(x) \end{bmatrix}.$$

Then (UQP) in its real form can be written as

$$(2.4) \qquad \begin{aligned} \min_{y \in \mathbb{R}^{2n}} \quad & y^{\mathsf{T}} \hat{Q} y + 2\hat{c}^{\mathsf{T}} y \\ \text{s.t.} \quad & y_i^2 + y_{n+i}^2 = 1, \ i = 1, 2, \dots, n. \end{aligned}$$

The Lagrangian function of (2.4) is

$$L(y; \lambda) = y^{\mathsf{T}} \hat{Q} y + 2\hat{c}^{\mathsf{T}} y + \sum_{i=1}^{n} \lambda_i \left(y_i^2 + y_{n+i}^2 - 1 \right),$$

where λ_i is the Lagrange multiplier corresponding to the constraint $y_i^2 + y_{n+i}^2 = 1$. The Lagrangian dual problem of (2.4) is

$$(D) \qquad\qquad\qquad\qquad\qquad \max_{\lambda \in \mathbb{R}^n} \ d(\lambda),$$

where $d(\lambda) := \min_{y \in \mathbb{R}^{2n}} L(y; \lambda)$ is the dual function of problem (2.4). It is well known that problem (D) is equivalent to (CSDP) [20, 33], i.e., the optimal value of problem (D) is equal to that of problem (CSDP).

Next, we derive the KKT optimality condition for problem (2.4). It is simple to verify that the linear independent constraint qualification (LICQ) holds for problem (2.4). Let x^* be an optimal solution of (UQP) and y^* be the corresponding optimal solution of problem (2.4). Then there exist $\{\lambda_i^* \in \mathbb{R}\}_{i=1}^n$ such that

$$\left. \frac{\partial L(y; \lambda)}{\partial y_i} \right|_{y=y^*} = 2 \left[\hat{Q} y^* \right]_i + 2\hat{c}_i + 2\lambda_i y_i^* = 0, \ i = 1, 2, \dots, n,$$

and

$$\left. \frac{\partial L(y; \lambda)}{\partial y_{i+n}} \right|_{y=y^*} = 2 \left[\hat{Q} y^* \right]_{i+n} + 2\hat{c}_{i+n} + 2\lambda_i y_{i+n}^* = 0, \ i = 1, 2, \dots, n.$$

Since

$$2 \left[\hat{Q} y^* \right]_i + 2\hat{c}_i + 2\lambda_i y_i^* = 2\mathrm{Re} \left([Qx^*]_i + c_i \right) + 2\lambda_i \mathrm{Re} \left(x_i^* \right), \ i = 1, 2, \dots, n,$$

and

$$2 \left[\hat{Q} y^* \right]_{i+n} + 2\hat{c}_{i+n} + 2\lambda_i y_{i+n}^* = 2\mathrm{Im} \left([Qx^*]_i + c_i \right) + 2\lambda_i \mathrm{Im} \left(x_i^* \right), \ i = 1, 2, \dots, n,$$

the KKT condition of (UQP) in the complex form becomes

$$[Qx^*]_i + c_i + \lambda_i x_i^* = 0, \ i = 1, 2, \dots, n.$$

Now, we prove the next theorem.

THEOREM 2.2. *Suppose that $M \geq 2$. If* (CSDP) *is tight for* (P), *then there exist* $\{\lambda_i^* \in \mathbb{R}\}_{i=1}^n$ *such that*

$$[H^\dagger v]_i = \lambda_i^* x_i^*, \ i = 1, 2, \ldots, n, \tag{2.5}$$

where H, x^*, and v are given in (1.1).

Proof. Let $\lambda^* \in \mathbb{R}^n$ be the optimal solution of problem (D). If (CSDP) is tight for (P), then the KKT condition of (UQP) is satisfied at x^*, i.e.,

$$[Qx^* + c]_i + \lambda_i^* x_i^* = 0, \ i = 1, 2, \ldots, n.$$

This, together with $Qx^* + c = H^\dagger H x^* - H^\dagger r = -H^\dagger v$, immediately implies the desired result (2.5). $\qquad\square$

The conditions in (2.5) imply that either $[H^\dagger v]_i$ and x_i^* have the same phase or $[H^\dagger v]_i = 0$ for all $i = 1, 2, \ldots, n$. However, this is generally not true in real applications. In particular, the noise vector v is often assumed to follow a circularly symmetric complex Gaussian distribution, and the probability that the event

$$\left\{[H^\dagger v]_i = 0\right\} \bigcup \left\{\arg\left([H^\dagger v]_i\right) = \arg\left(x_i^*\right)\right\}$$

happens is zero for each $i = 1, \ldots, n$. Therefore, the probability that all conditions in (2.5) are simultaneously satisfied is zero. This immediately implies that (CSDP) is generally not tight for (P) (regardless of the condition in (1.4)). This answers the open question posed in [35]. It is also worth noting that (CSDP) and (BSDP) are not equivalent to each other even for the case where $M = 2$.

3. An enhanced semidefinite relaxation. In this section, we propose an enhanced SDR for (CQP). Recall that the (discrete) argument constraints $\arg(x_i) \in \mathcal{A}$ for $i = 1, 2, \ldots, n$ in (CQP) are ignored in its conventional relaxation (CSDP). The idea of designing the enhanced SDR for (CQP) is to better exploit the structure of the argument constraints and develop valid linear constraints for them to tighten (CSDP). Since these valid linear constraints are based on the real form of (CSDP), we first reformulate (CSDP) as the following real SDR:

$$\text{(RSDP)} \quad \begin{aligned} \min_{y, Y} \quad & \hat{Q} \bullet Y + 2\hat{c}^\mathsf{T} y \\ \text{s.t.} \quad & Y_{i,i} + Y_{n+i,n+i} = 1, \ i = 1, 2, \ldots, n, \\ & \begin{bmatrix} 1 & y^\mathsf{T} \\ y & Y \end{bmatrix} \succeq 0, \end{aligned}$$

where $y \in \mathbb{R}^{2n}$ and $Y \in \mathbb{R}^{2n \times 2n}$, and $\hat{Q}$, $\hat{c}$, and y are defined in (2.3).

We remark that our real reformulation (RSDP) of complex (CSDP) is not the same as the ones in [11] and [44]. The dimension of the matrix variable in (RSDP) is $2n + 1$ while the one of the matrix variable in [11] and [44] is $2n + 2$. Hence, the equivalence between (CSDP) and (RSDP) cannot be shown by using the same argument in [11] and [44]. An equivalence proof of (CSDP) and (RSDP) is provided in Appendix A.

Next, we develop some valid linear constraints for the argument constraints based on (RSDP), which leads to an enhanced SDR for (CQP). First, for (RSDP), we define the following 3×3 matrices:

$$\mathcal{Y}(i) := \begin{bmatrix} 1 & y_i & y_{n+i} \\ y_i & Y_{i,i} & Y_{i,n+i} \\ y_{n+i} & Y_{n+i,i} & Y_{n+i,n+i} \end{bmatrix}, \ i = 1, 2, \ldots, n. \tag{3.1}$$

For each $i = 1, \ldots, n$, $\mathcal{Y}(i)$ contains the following five variables in (RSDP) (due to its symmetry):

$$y_i, \ y_{n+i}, \ Y_{i,i}, \ Y_{i,n+i}, \text{ and } Y_{n+i,n+i}.$$

From the definition of y in (2.3), we have $y_i = \mathrm{Re}(x_i)$ and $y_{n+i} = \mathrm{Im}(x_i)$. Since $|x_i|^2 = 1$ and $\arg(x_i) \in \mathcal{A}$, it follows that

$$(y_i, \ y_{n+i}) \in \{(\cos(\theta), \ \sin(\theta)) \mid \theta = 2j\pi/M, \ j = 0, 1, \ldots, M-1\}.$$

Define the following 3×3 real symmetric matrices:

$$(3.2) \qquad P_j = \begin{bmatrix} 1 \\ \cos\left(\frac{2j\pi}{M}\right) \\ \sin\left(\frac{2j\pi}{M}\right) \end{bmatrix} \begin{bmatrix} 1 & \cos\left(\frac{2j\pi}{M}\right) & \sin\left(\frac{2j\pi}{M}\right) \end{bmatrix}, \ j = 0, 1, \ldots, M-1.$$

Then, each $\mathcal{Y}(i)$ must equal one of the matrices P_j with $j = 0, 1, \ldots, M-1$, i.e.,

$$\mathcal{Y}(i) \in \{P_0, \ P_1, \ldots, P_{M-1}\}, \ i = 1, 2, \ldots, n.$$

The convex envelopes of the above constraints are

$$\mathcal{Y}(i) \in \mathrm{Conv}\{P_0, \ P_1, \ldots, P_{M-1}\}, \ i = 1, 2, \ldots, n,$$

which are equivalent to

$$\mathcal{Y}(i) = \sum_{j=0}^{M-1} t_{i,j} P_j, \ \sum_{j=0}^{M-1} t_{i,j} = 1, \ t_{i,j} \geq 0, \ j = 0, 1, \ldots, M-1, \ i = 1, 2, \ldots, n.$$

Due to the symmetry of $\mathcal{Y}(i)$, the constraint $\mathcal{Y}(i) = \sum_{j=0}^{M-1} t_{i,j} P_j$ can be explicitly expressed as the following five linear constraints:

$$y_i = \sum_{j=0}^{M-1} t_{i,j} \cos\left(\frac{2j\pi}{M}\right), \ y_{n+i} = \sum_{j=0}^{M-1} t_{i,j} \sin\left(\frac{2j\pi}{M}\right),$$

$$(3.3) \qquad Y_{i,i} = \sum_{j=0}^{M-1} t_{i,j} \cos^2\left(\frac{2j\pi}{M}\right), \ Y_{n+i,n+i} = \sum_{j=0}^{M-1} t_{i,j} \sin^2\left(\frac{2j\pi}{M}\right),$$

$$Y_{i,n+i} = \sum_{j=0}^{M-1} t_{i,j} \cos\left(\frac{2j\pi}{M}\right) \sin\left(\frac{2j\pi}{M}\right).$$

Obviously, the above equations and $\sum_{j=0}^{M-1} t_{i,j} = 1$ imply $Y_{i,i} + Y_{n+i,n+i} = 1$. By dropping redundant constraints, we get the following enhanced SDR for (CQP):

$$\begin{aligned} \min_{y, \, Y, \, t} \quad & \hat{Q} \bullet Y + 2\hat{c}^{\mathsf{T}} y \\ \text{s.t.} \quad & \mathcal{Y}(i) = \sum_{j=0}^{M-1} t_{i,j} P_j, \ \sum_{j=0}^{M-1} t_{i,j} = 1, \ i = 1, 2, \ldots, n, \\ & t_{i,j} \geq 0, \ i = 1, 2, \ldots, n, \ j = 0, 1, \ldots, M-1, \\ & \begin{bmatrix} 1 & y^{\mathsf{T}} \\ y & Y \end{bmatrix} \succeq 0, \end{aligned}$$

(ERSDP)

where $y \in \mathbb{R}^{2n}$, $Y \in \mathbb{R}^{2n \times 2n}$, $t \in \mathbb{R}^{n \times M}$, $\hat{Q}$ and $\hat{c}$ are defined in (2.3), $\mathcal{Y}(i)$ is defined in (3.1), and P_j is defined in (3.2).

We term the above real SDP "ERSDP," since $\mathcal{Y}(i)$ for all $i = 1, 2, \ldots, n$ in (ERSDP) are constrained in the convex envelope with the Extreme points being $P_0, P_1, \ldots, P_{M-1}$, i.e., the variables y_i, y_{n+i}, $Y_{i,i}$, $Y_{i,n+i}$, and $Y_{n+i,n+i}$ must satisfy the linear constraints given in (3.3), which is the main difference between (ERSDP) and (RSDP). Hence, (ERSDP) is (strictly) tighter than (RSDP) (which is equivalent to (CSDP)).

It is worth noting that the proposed (ERSDP) is not the first SDR for the MIMO detection problem (P) that is stronger/tighter than (CSDP). For instance, both the SDRs proposed in [8, 29] are customized for problem (P). The proposed (ERSDP) is different from those in [8, 29] in the sense that the matrix variables in these relaxations are lifted from different spaces.

4. Tightness of (ERSDP). In this section, we study the tightness of the newly proposed SDP relaxation (ERSDP).

Let us first look at a case where $M = 2$. In this case,

$$P_0 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } P_1 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and the linear constraints in (3.3) reduce to

$$Y_{i,i} = 1, \ y_i = t_{i,1} - t_{i,2}, \ \text{and} \ y_{n+i} = Y_{n+i,i} = Y_{n+i,n+i} = 0.$$

By dropping the zero blocks in the matrix Y, (ERSDP) reduces to (BSDP). Therefore, it follows from Theorem 2.1 that condition (1.3) is sufficient for (ERSDP) to be tight for problem (P) with $M = 2$.

In the remainder of this section, we study the tightness of (ERSDP) for (P) with $M \geq 3$. We prove that condition (1.5) is sufficient for (ERSDP) to be tight for (P). Our proof consists of two main steps: In Step I, we derive a complex SDR called (CSDP2) based on (ERSDP) and show that (ERSDP) is tighter than (CSDP2); in Step II, we show that (CSDP2) is tight for (P) with $M \geq 3$ under a certain condition, and hence (ERSDP) is also tight for (P) under the same condition.

Step I. We derive a complex SDR to be called (CSDP2) from (ERSDP). Recall that, for any feasible solution (y, Y) of (ERSDP), $\mathcal{Y}(i)$ in (3.1) must lie in the convex envelope of $\{P_0, \ldots, P_{M-1}\}$, which implies that

(4.1) $$(y_i, y_{i+n}) \in \text{Conv} \{(\cos(\theta), \sin(\theta)) \mid \theta \in \mathcal{A}\},$$

where $\mathcal{A}$ is defined in (1.2). Note that y_i and y_{i+n} correspond to the real and imaginary parts of the complex variable x_i. Then, one can show that (4.1) is equivalent to $x_i \in \mathcal{F}$, where $\mathcal{F}$ is the convex envelope of the set

$$\mathcal{D} = \{w \mid |w| = 1, \ \arg(w) \in \mathcal{A}\}.$$

For the case where $M \geq 3$, by using a similar argument to that showing Proposition 1 in [24], the set $\mathcal{F}$ can be represented as

$$\left\{ w \in \mathbb{C} \mid \text{Re}(a_j^\dagger w) \leq \cos\left(\frac{\pi}{M}\right), \ j = 0, 1, \ldots, M-1 \right\},$$

where

(4.2)
$$a_j = e^{\mathbf{i}\theta_j} \text{ and } \theta_j = \frac{(2j-1)\pi}{M}.$$

Furthermore, $\mathrm{Re}(a_j^\dagger w) \le \cos(\frac{\pi}{M})$ is equivalent to

$$\cos(\theta_j)\,\mathrm{Re}(w) + \sin(\theta_j)\,\mathrm{Im}(w) \le \cos\left(\frac{\pi}{M}\right).$$

The line associated with the above half-plane connects the two points

$$\cos\left(\frac{2(j-1)\pi}{M}\right) + \sin\left(\frac{2(j-1)\pi}{M}\right)\mathbf{i} \text{ and } \cos\left(\frac{2j\pi}{M}\right) + \sin\left(\frac{2j\pi}{M}\right)\mathbf{i}$$

in the complex domain. See Figure 1 for an illustration of the set $\mathcal{F}$ for the case where $M = 8$.

Based on the above observations, we obtain the following complex SDR:

(CSDP2)
$$\begin{aligned}
\min_{x,\,X} \quad & Q \bullet X + 2\mathrm{Re}(c^\dagger x) \\
\text{s.t.} \quad & X_{ii} = 1,\ i = 1, 2, \ldots, n, \\
& \mathrm{Re}(a_j^\dagger x_i) \le \cos\left(\tfrac{\pi}{M}\right),\ i = 1, 2, \ldots, n,\ j = 0, 1, \ldots, M-1, \\
& \begin{bmatrix} 1 & x^\dagger \\ x & X \end{bmatrix} \succeq 0,
\end{aligned}$$

where $x \in \mathbb{C}^n$ and $X \in \mathbb{C}^{n \times n}$.

Two remarks on the comparison between (CSDP2) and (ERSDP) are in order. First, since the constraints $\mathrm{Re}(a_j^\dagger x_i) \le \cos(\frac{\pi}{M})$ for all $i = 1, 2, \ldots, n$ and $j = 0, 1, \ldots, M-1$ in (CSDP2) are derived from (ERSDP), we know that (ERSDP) is at least as tight as (CSDP2). In fact, we have the following result.

THEOREM 4.1. (ERSDP) *is tighter than* (CSDP2).

A rigorous proof of Theorem 4.1 can be found in Appendix B. Therefore, if (CSDP2) is tight for (P), then (ERSDP) must also be tight. Second, (CSDP2) is tight enough for us to derive our main results in Theorems 4.2 and 4.4. Our analysis is based on a simpler reformulation of (CSDP2) and its dual.

Step II. We show the tightness of (CSDP2) in this part. Our proof is similar to the one that shows Theorem 1 in [15]. The basic idea is to construct a dual feasible solution such that (x^*, X^*) and the constructed dual solution jointly satisfy the KKT optimality conditions of (CSDP2) under some conditions, where $X^* = x^*(x^*)^\dagger$. The dual problem of (CSDP2) is

(4.3)
$$\begin{aligned}
\max_{\lambda,\,\tau,\,\mu} \quad & \tau - \sum_{i=1}^{n} \lambda_i - \cos\left(\frac{\pi}{M}\right) \sum_{i=1}^{n} \sum_{j=0}^{M-1} \mu_{i,j} \\
\text{s.t.} \quad & \begin{bmatrix} -\tau & (c+g)^\dagger \\ c+g & Q + \mathrm{Diag}(\lambda) \end{bmatrix} \succeq 0, \\
& \mu_{i,j} \ge 0,\ i = 1, 2, \ldots, n,\ j = 0, 1, \ldots, M-1,
\end{aligned}$$

where $\lambda \in \mathbb{R}^n$, $\tau \in \mathbb{R}$, $\mu \in \mathbb{R}^{n \times M}$, and $g = [g_1, g_2, \ldots, g_n]^\mathsf{T}$ with

(4.4)
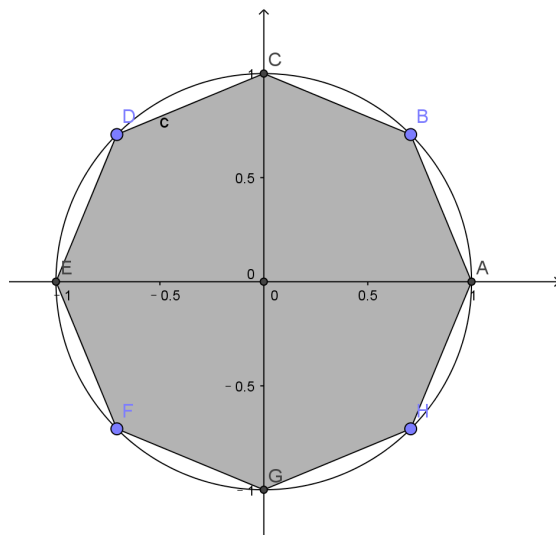$$g_i = \sum_{j=0}^{M-1} \frac{\mu_{i,j}}{2} a_j,\ i = 1, 2, \ldots, n.$$

FIG. 1. *An illustration of the set $\mathcal{F}$ for the case where $M = 8$, which can be represented by 8 linear constraints.*

For completeness, a detailed derivation of the above dual problem is provided in Appendix C. Since both problems (CSDP2) and its dual (4.3) are strictly feasible, it follows that the primal-dual Slater conditions are satisfied. Suppose that (x, X) is an optimal solution of (CSDP2). Then there must exist a feasible solution (λ, τ, μ) (to problem (4.3)) such that (x, X) and (λ, τ, μ) jointly satisfy the following complementarity conditions:

$$(4.5) \qquad \mu_{i,j}\left(\cos\left(\frac{\pi}{M}\right) - \mathrm{Re}\left(a_j^\dagger x_i\right)\right) = 0, \ i = 1, 2, \ldots, n, \ j = 0, 1, \ldots, M-1,$$

and

$$(4.6) \qquad \begin{bmatrix} -\tau & (c+g)^\dagger \\ c+g & Q + \mathrm{Diag}(\lambda) \end{bmatrix} \bullet \begin{bmatrix} 1 & x^\dagger \\ x & X \end{bmatrix} = 0.$$

We are ready to present our main results in this paper.

THEOREM 4.2. *Suppose that $M \geq 3$. Let x^* with*

$$(4.7) \qquad x_i^* = e^{2\pi s_i i/M}, \ s_i \in \{0, 1, \ldots, M-1\}, \ i = 1, 2, \ldots, n,$$

be the vector of transmitted signals. Define

$$(4.8) \qquad t_i = \begin{cases} s_i + 1 & \text{if } s_i < M-1, \\ 0 & \text{if } s_i = M-1, \end{cases} \ i = 1, 2, \ldots, n.$$

Then (x^*, X^*) is an optimal solution to (CSDP2) if and only if there exist

$$(4.9) \qquad \bar{\lambda}_i \in \mathbb{R}, \ \bar{\mu}_{i,s_i} \geq 0, \ \text{and } \bar{\mu}_{i,t_i} \geq 0, \ i = 1, 2, \ldots, n,$$

such that H and v in (1.1) *satisfy*

$$(4.10) \qquad [H^\dagger v]_i = \bar{\lambda}_i x_i^* + \frac{\bar{\mu}_{i,s_i}}{2} a_{s_i} + \frac{\bar{\mu}_{i,t_i}}{2} a_{t_i}, \ i = 1, 2, \ldots, n,$$

and

$$(4.11) \qquad Q + \mathrm{Diag}(\bar{\lambda}) \succeq 0.$$

Furthermore, if $Q + \mathrm{Diag}(\bar{\lambda}) \succ 0$, then (x^*, X^*) is the unique optimal solution to (CSDP2).

Proof. We first prove the necessary direction. Assume that (x^*, X^*) is an optimal solution of (CSDP2) and $(\lambda^*, \tau^*, \mu^*)$ is an optimal solution of dual problem (4.3). Define $\Lambda^* = \mathrm{Diag}(\lambda^*)$. Then, from the optimality conditions, we have

$$(4.12) \qquad \begin{bmatrix} -\tau^* & (c + g^*)^\dagger \\ c + g^* & Q + \Lambda^* \end{bmatrix} \succeq 0,$$

$$(4.13) \qquad \begin{bmatrix} -\tau^* & (c + g^*)^\dagger \\ c + g^* & Q + \Lambda^* \end{bmatrix} \bullet \begin{bmatrix} 1 & (x^*)^\dagger \\ x^* & X^* \end{bmatrix} = \begin{bmatrix} 1 \\ x^* \end{bmatrix}^\dagger \begin{bmatrix} -\tau^* & (c + g^*)^\dagger \\ c + g^* & Q + \Lambda^* \end{bmatrix} \begin{bmatrix} 1 \\ x^* \end{bmatrix} = 0,$$

and

$$(4.14) \qquad \mu_{i,j}^* \left(\cos\left(\frac{\pi}{M}\right) - \mathrm{Re}\left(a_j^\dagger x_i^*\right) \right) = 0, \ i = 1, 2, \ldots, n, \ j = 0, 1, \ldots, M - 1.$$

From (4.12), we immediately get

$$(4.15) \qquad Q + \Lambda^* \succeq 0.$$

Furthermore, it follows from (4.12) and (4.13) that

$$(4.16) \qquad \begin{bmatrix} -\tau^* & (c + g^*)^\dagger \\ c + g^* & Q + \Lambda^* \end{bmatrix} \begin{bmatrix} 1 \\ x^* \end{bmatrix} = 0,$$

which implies $c + g^* = -(Q + \Lambda^*) x^*$. By this and the facts that $Q = H^\dagger H$ and $c = -H^\dagger r = -H^\dagger H x^* - H^\dagger v$, we have

$$(4.17) \qquad H^\dagger v = g^* + \Lambda^* x^*.$$

Moreover, by the complementarity condition (4.14), we obtain $\mu_{i,j}^* = 0$ for $j \notin \{s_i, t_i\}$ and $\mu_{i,j}^* \geq 0$ for $j \in \{s_i, t_i\}$. This, together with the definition of g_i (cf. (4.4)), shows

$$g_i^* = \sum_{j=0}^{M-1} \frac{\mu_{i,j}^*}{2} a_j = \frac{\mu_{i,s_i}^*}{2} a_{s_i} + \frac{\mu_{i,t_i}^*}{2} a_{t_i}, \ i = 1, 2, \ldots, n.$$

Substituting the above into (4.17), we immediately obtain

$$(4.18) \qquad [H^\dagger v]_i = \lambda_i^* x_i^* + \frac{\mu_{i,s_i}^*}{2} a_{s_i} + \frac{\mu_{i,t_i}^*}{2} a_{t_i}, \ i = 1, 2, \ldots, n.$$

Let $\bar{\lambda} = \lambda^*$, $\bar{\mu}_{i,s_i} = \mu_{i,s_i}^*$, and $\bar{\mu}_{i,t_i} = \mu_{i,t_i}^*$ for $i = 1, \ldots, n$. Hence, (4.15) and (4.18) become (4.11) and (4.10), respectively. The proof for the necessity of the condition is completed.

Next, we shall prove that the condition is sufficient too. Under the condition, we construct a feasible solution $(\lambda^*, \tau^*, \mu^*)$ for problem (4.3) as follows:

$$(4.19) \qquad \lambda_i^* = \bar{\lambda}_i, \ i = 1, 2, \ldots, n,$$

and

$$(4.20) \qquad \mu_{i,j}^* = \begin{cases} \bar{\mu}_{i,j} & \text{if } j \in \{s_i, t_i\}, \\ 0 & \text{if } j \notin \{s_i, t_i\}, \end{cases} \quad i = 1, 2, \ldots, n, \ j = 0, 1, \ldots, M-1,$$

where $\bar{\lambda}_i$, $\bar{\mu}_{i,s_i}$, and $\bar{\mu}_{i,t_i}$ satisfy the conditions stipulated in (4.9) and (4.10). Set

$$g_i^* = \sum_{j=0}^{M-1} \frac{\mu_{i,j}^*}{2} a_j, \ i = 1, 2, \ldots, n,$$

and define $\Lambda^* = \text{Diag}(\lambda^*)$. By (4.10), (4.19), and (4.20), we obtain $H^\dagger v = \Lambda^* x^* + g^*$, which, together with the definitions of Q and c, further implies

$$c + g^* = -H^\dagger H x^* - H^\dagger v + g^* = -H^\dagger H x^* - \Lambda^* x^* = -(Q + \Lambda^*) x^*.$$

Furthermore, set

$$(4.21) \qquad \tau^* = (x^*)^\dagger (Q + \Lambda^*) x^* + 2\text{Re}\left((c + g^*)^\dagger x^*\right) = -(x^*)^\dagger (Q + \Lambda^*) x^*.$$

Clearly, under the condition $Q + \Lambda^* \succeq 0$, we have

$$\begin{bmatrix} -\tau^* & (c + g^*)^\dagger \\ c + g^* & Q + \Lambda^* \end{bmatrix} = \begin{bmatrix} (x^*)^\dagger \\ -I_n \end{bmatrix} (Q + \Lambda^*) \begin{bmatrix} x^* & -I_n \end{bmatrix} \succeq 0.$$

Therefore, the above constructed $(\lambda^*, \tau^*, \mu^*)$ is a feasible solution to problem (4.3).

Then, we show that the pairs (x^*, X^*) and $(\lambda^*, \tau^*, \mu^*)$ jointly satisfy the complementarity conditions (4.5) and (4.6). Note that $\mu_{i,j}^* = 0$ if $j \notin \{s_i, t_i\}$ (cf. (4.20)) and $\cos(\frac{\pi}{M}) - \text{Re}(a_j^\dagger x_i^*) = 0$ if $j \in \{s_i, t_i\}$ (from (4.2) and (4.8)). Consequently, (4.5) is true. By simple calculation, we have

$$(4.22)$$
$$\begin{bmatrix} -\tau^* & (c + g^*)^\dagger \\ c + g^* & Q + \Lambda^* \end{bmatrix} \bullet \begin{bmatrix} 1 & (x^*)^\dagger \\ x^* & X^* \end{bmatrix} = (x^*)^\dagger (Q + \Lambda^*) x^* + 2\text{Re}\left((c + g^*)^\dagger x^*\right) - \tau^* = 0.$$

Therefore, both of the complementarity conditions (4.5) and (4.6) hold. From the above analysis, we can conclude that (x^*, X^*) is an optimal solution of (CSDP2), completing the proof that the condition is sufficient.

Finally, if $Q + \bar{\Lambda} \succ 0$, then it follows from [1, Theorem 10] that (x^*, X^*) is the unique optimal solution of (CSDP2). $\qquad \square$

Note that if $\bar{\lambda}_i > -\lambda_{\min}$ for all $i = 1, 2, \ldots, n$, then $Q + \text{Diag}(\bar{\lambda}) \succ 0$. Hence, it follows from Theorem 4.2 that we have the following corollary.

COROLLARY 4.3. *Suppose that $M \geq 3$. Let s_i and t_i be defined as in Theorem* 4.2. *If there exist*

$$(4.23) \qquad \bar{\lambda}_i > -\lambda_{\min}, \ \bar{\mu}_{i,s_i} \geq 0, \ and \ \bar{\mu}_{i,t_i} \geq 0, \ i = 1, 2, \ldots, n,$$

such that H and v in (1.1) satisfy (4.10), then (x^*, X^*) is the unique solution to (CSDP2).

As a direct consequence of Theorem 4.2 and Corollary 4.3, we have the next result.

THEOREM 4.4. *Suppose that $M \geq 3$. If the inputs H and v in (1.1) satisfy (1.5), then both* (CSDP2) *and* (ERSDP) *are tight for* (P).

Proof. Consider the set

$$\mathcal{S}_i = \left\{ \lambda_i x_i^* + \frac{\mu_{i,s_i}}{2} a_{s_i} + \frac{\mu_{i,t_i}}{2} a_{t_i} \in \mathbb{C} \mid \lambda_i \geq -\lambda_{\min},\ \mu_{i,s_i} \geq 0,\ \mu_{i,t_i} \geq 0 \right\}.$$

By (4.2) and (4.7), we have $a_{s_i} = x_i^* e^{-\pi \mathbf{i}/M}$ and $a_{t_i} = x_i^* e^{\pi \mathbf{i}/M}$. Hence, $\mathcal{S}_i$ can be rewritten as $\mathcal{S}_i = x_i^* \hat{\mathcal{S}}_i$, where

$$\hat{\mathcal{S}}_i = \left\{ \lambda_i + \frac{\mu_{i,s_i}}{2} e^{-\pi \mathbf{i}/M} + \frac{\mu_{i,t_i}}{2} e^{\pi \mathbf{i}/M} \in \mathbb{C} \mid \lambda_i \geq -\lambda_{\min},\ \mu_{i,s_i} \geq 0,\ \mu_{i,t_i} \geq 0 \right\}.$$

Note that $\hat{\mathcal{S}}_i$ is a polyhedron in the complex domain with the extreme point being $-\lambda_{\min}$ and two extreme directions being $e^{-\pi \mathbf{i}/M}$ and $e^{\pi \mathbf{i}/M}$, that is, a polyhedron in the two-dimensional real domain with the extreme point being $(-\lambda_{\min}, 0)$ and two extreme directions being $(\cos(\pi/M), -\sin(\pi/M))$ and $(\cos(\pi/M), \sin(\pi/M))$. One can easily verify that (1) the origin lies in the set $\mathcal{S}_i$ and (2) the smallest distance from the origin to the boundary of the set $\mathcal{S}_i$ is $\lambda_{\min} \sin(\frac{\pi}{M})$. Therefore, if H and v in (1.1) satisfy the condition in (1.5), then $[H^\dagger v]_i$ lies in the interior of $\mathcal{S}_i$ for all $i = 1, 2, \ldots, n$, which further implies that the conditions in (4.10) are satisfied. Invoking Theorem 4.2 and Corollary 4.3, we conclude that (CSDP2) is tight for (P) if (1.5) is satisfied. Since (ERSDP) is even tighter than (CSDP2) (cf. Theorem 4.1), it follows that (ERSDP) is also tight for (P) if (1.5) is satisfied. The proof is complete.□

Two remarks on Theorem 4.4 are in order.

First, Theorem 4.4 shows that both (CSDP2) and (ERSDP) are tight for the MIMO detection problem (P) under condition (1.5) and thus answers the open question posed by So in [35]. Moreover, our result in Theorem 4.4 is for arbitrary $M \geq 3$, in contrast to all previous results in [14, 15, 17, 18, 35], which focus on the $M = 2$ case.

Second, Theorem 4.4 states that the true vector of transmitted signals x^* can be recovered by solving appropriate SDPs if condition (1.5) is satisfied. It is worth mentioning a recent related result in [22, Theorem 1] for the general M-PSK modulation case, which shows that x^* can be recovered by using the generalized power method (GPM) if the following conditions are satisfied:

$$(4.24) \qquad \left\| \frac{2\alpha_k}{m} H^\dagger v \right\|_\infty < \frac{\sin(\pi/M)}{2}, \quad \left\| I_n - \frac{2\alpha_k}{m} H^\dagger H \right\|_{\mathrm{op}} < \frac{1}{4},$$

where $\{\alpha_k\}$ are the step sizes used in the GPM and $\| \cdot \|_{\mathrm{op}}$ is the matrix operator norm. One can show that the conditions in (4.24) imply

$$\lambda_{\min}(H^\dagger H) \sin\left(\frac{\pi}{M}\right) > \frac{3}{2} \|H^\dagger v\|_\infty,$$

which is more restrictive than condition (1.5). Another difference between conditions (4.24) in [22] and our condition (1.5) is that the conditions in (4.24) depend on both the algorithm parameters (i.e., $\{\alpha_k\}$) and the problem parameters (i.e., H, v, and M) while our condition (1.5) only depends on the problem parameters (and is independent of the algorithm parameters).

In practice, H and v generally follow a complex Gaussian distribution. The following theorem states how likely it is that condition (1.5) will be satisfied under the assumption that the real and imaginary parts of the entries of H and v are i.i.d. real Gaussian variables. The proof of the theorem is postponed to Appendix B.

THEOREM 4.5. *Consider the MIMO channel model* (1.1), *where the entries of* $\mathrm{Re}(H)$ *and* $\mathrm{Im}(H)$ *are i.i.d. standard real Gaussian random variables and the entries of* $\mathrm{Re}(v)$ *and* $\mathrm{Im}(v)$ *are i.i.d. real Gaussian variables with zero mean and variance* σ^2. *Suppose that* $m > n$ *and*

$$(4.25) \qquad \sigma \leq \frac{(1-\rho)^2 \sin(\pi/M)}{4\sqrt{2}},$$

where $\rho = \sqrt{n/m}$. *Then the probability that condition* (1.5) *is satisfied is at least*

$$1 - e^{-\frac{m(1-\rho)^2}{4}} - 2\sqrt{2/\pi} \cdot n \cdot e^{-m/2} - 8e^{-m/8}.$$

5. Numerical experiments. In this section, we present some numerical results to illustrate the tightness of our proposed enhanced (ERSDP) and the conventional (CSDP) for problem (P). We generate the instances of problem (P) as follows: We generate each entry of the channel matrix $H \in \mathbb{C}^{m \times n}$ according to the complex standard Gaussian distribution (with zero mean and unit variance); for each $i = 1, 2, \ldots, n$, we uniformly choose s_i from $\{0, 1, \ldots, M-1\}$ and set each entry of the vector $x^* \in \mathbb{C}^n$ to be $x_i^* = e^{2s_i\pi\mathbf{i}/M}$; we generate each entry of the noise vector $v \in \mathbb{C}^m$ according to the complex Gaussian distribution with zero mean and with variance σ^2; and finally we compute r as in (1.1). In our numerical experiments, we set $m = 15$ and $n = 10$. In practical digital communications, M is generally 2, 4, 8, ... In our experiments, to study the tightness of (ERSDP) on various different settings, we consider all cases where $M \in \{3, 4, 6, 8\}$.

To compare the numerical performance of (ERSDP) and (CSDP), we generate a total of 40 instances with $M = 3$ and $\sigma^2 \in \{0.01, 0.1, 1.0, 10\}$. Numerical results are summarized in Table 1, where "LBC" and "LBE" denote the lower bounds (i.e., the optimal objective values) returned by the two SDP relaxations (CSDP) and (ERSDP), respectively; "UB" denotes the upper bound (i.e., the objective value at the feasible point) returned by the approximation algorithms in [36] and [47]; "GapC" ($=$ UB $-$ LBC) and "GapE" ($=$ UB $-$ LBE) denote the gaps between the upper bound and the corresponding lower bounds; "ClosedGap" denotes (LBE $-$ LBC)/(UB $-$ LBC), which measures how much of the gap in (CSDP) is closed by (ERSDP); and finally the "Y/N" pair denotes whether (ERSDP) is tight or not and whether condition (1.5) is satisfied or not, with "Y" and "N" denoting "Yes" and "No", respectively. Obviously, the larger the value of ClosedGap is, the larger the portion of the gap in (CSDP) is closed by (ERSDP) and the better the performance of (ERSDP) (compared to (CSDP)).

We can observe from Table 1 that (ERSDP) is generally tight for problem (P) when the noise level σ^2 is low, i.e., when the SNR is high. More specifically, (ERSDP) is tight for all 20 instances with $\sigma^2 \leq 0.1$ and for 5 instances with $\sigma^2 = 1$. In sharp contrast, (CSDP) is not tight for all instances. Furthermore, we can see from Table 1 that LBE is much larger than LBC and GapC is much larger than GapE for instances with high levels of noise, which show that (ERSDP) is much tighter than (CSDP) for instances with high levels of noise. For instance, when $\sigma^2 = 10$, GapE is generally much smaller than GapC and about 70% to 99.8% of the gap in (CSDP) is closed by (ERSDP).

TABLE 1

Numerical results of (ERSDP) *and* (CSDP) *on 40 randomly generated instances of problem* (P) *with different levels of noise and $M = 3$.*

ID	σ^2	LBC	LBE	UB	GapC	GapE	ClosedGap	Y/N
1	0.010	−506.995	−506.883	−506.883	0.111	0.000	100.0%	(Y, N)
2	0.010	−332.651	−332.539	−332.539	0.112	0.000	100.0%	(Y, Y)
3	0.010	−253.831	−253.568	−253.568	0.263	0.000	100.0%	(Y, Y)
4	0.010	−194.345	−194.280	−194.280	0.065	0.000	100.0%	(Y, Y)
5	0.010	−217.328	−217.190	−217.190	0.138	0.000	100.0%	(Y, Y)
6	0.010	−344.753	−344.581	−344.581	0.173	0.000	100.0%	(Y, Y)
7	0.010	−212.575	−212.430	−212.430	0.145	0.000	100.0%	(Y, Y)
8	0.010	−159.017	−158.900	−158.900	0.116	0.000	100.0%	(Y, Y)
9	0.010	−206.139	−206.054	−206.054	0.084	0.000	100.0%	(Y, Y)
10	0.010	−228.058	−227.955	−227.955	0.103	0.000	100.0%	(Y, Y)
11	0.100	−133.312	−132.217	−132.217	1.095	0.000	100.0%	(Y, N)
12	0.100	−398.492	−397.605	−397.605	0.887	0.000	100.0%	(Y, N)
13	0.100	−260.164	−259.290	−259.290	0.875	0.000	100.0%	(Y, N)
14	0.100	−296.578	−295.055	−295.055	1.523	0.000	100.0%	(Y, N)
15	0.100	−202.681	−201.966	−201.966	0.715	0.000	100.0%	(Y, N)
16	0.100	−355.024	−354.248	−354.248	0.775	0.000	100.0%	(Y, N)
17	0.100	−454.615	−453.808	−453.808	0.807	0.000	100.0%	(Y, N)
18	0.100	−302.445	−301.429	−301.429	1.016	0.000	100.0%	(Y, N)
19	0.100	−392.029	−390.804	−390.804	1.225	0.000	100.0%	(Y, N)
20	0.100	−248.856	−247.446	−247.446	1.411	0.000	100.0%	(Y, N)
21	1.000	−183.786	−171.408	−170.839	12.947	0.569	95.6%	(N, N)
22	1.000	−388.290	−374.763	−370.751	17.539	4.012	77.1%	(N, N)
23	1.000	−309.820	−296.907	−296.907	12.913	0.000	100.0%	(Y, N)
24	1.000	−181.890	−157.968	−151.285	30.605	6.683	78.2%	(N, N)
25	1.000	−235.490	−219.367	−217.811	17.679	1.557	91.2%	(N, N)
26	1.000	−263.610	−249.803	−249.803	13.807	0.000	100.0%	(Y, N)
27	1.000	−373.140	−361.259	−361.259	11.882	0.000	100.0%	(Y, N)
28	1.000	−214.068	−201.188	−201.188	12.879	0.000	100.0%	(Y, N)
29	1.000	−404.825	−398.641	−398.564	6.260	0.077	98.8%	(N, N)
30	1.000	−343.993	−339.752	−339.752	4.241	0.000	100.0%	(Y, N)
31	10.000	−390.985	−300.162	−282.554	108.431	17.608	83.8%	(N, N)
32	10.000	−298.119	−221.992	−205.475	92.644	16.517	82.2%	(N, N)
33	10.000	−386.464	−329.595	−328.439	58.025	1.156	98.0%	(N, N)
34	10.000	−356.127	−314.399	−304.091	52.036	10.308	80.2%	(N, N)
35	10.000	−329.525	−243.219	−212.705	116.820	30.514	73.9%	(N, N)
36	10.000	−324.014	−284.622	−269.178	54.836	15.444	71.8%	(N, N)
37	10.000	−548.569	−477.101	−468.082	80.487	9.019	88.8%	(N, N)
38	10.000	−787.389	−737.455	−737.359	50.029	0.095	99.8%	(N, N)
39	10.000	−366.880	−307.478	−305.544	61.336	1.935	96.8%	(N, N)
40	10.000	−377.069	−304.619	−280.803	96.266	23.816	75.3%	(N, N)

To study the empirical probability of (ERSDP) being tight (as a function of σ^2), we generate more problem instances where $M = 3$ and $\sigma^2 \in [0.001, 10]$. For each setting, we randomly generate 1000 instances. Figure 2 illustrates the empirical probability of (ERSDP) being tight and the empirical probability of condition (1.5) being satisfied. We can conclude from Figure 2 that (1) for the case where $m = 15$, $n = 10$, and $M = 3$, the probability of (ERSDP) being tight is very close to one
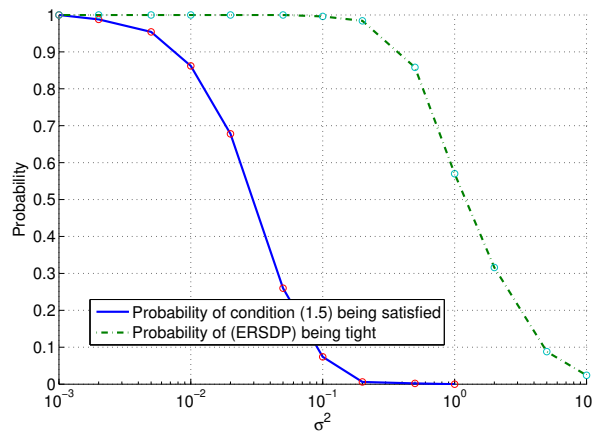
FIG. 2. *Empirical probability of* (ERSDP) *being tight and condition* (1.5) *being satisfied.*

TABLE 2
Numerical results of (ERSDP) *and* (CSDP) *on randomly generated instances of problem* (P) *with different M and different levels of noise.*

M	σ^2	GapC	GapE	ClosedGap	TimeC	TimeE	ProbC	ProbE
4	0.01	0.096	0.000	100.0%	0.05	0.08	0%	100%
6	0.01	0.106	0.000	100.0%	0.05	0.08	0%	100%
8	0.01	0.100	0.000	99.9%	0.05	0.09	0%	99%
4	0.1	0.971	0.000	100.0%	0.05	0.08	0%	100%
6	0.1	0.871	0.015	98.9%	0.05	0.08	0%	80%
8	0.1	0.979	0.056	96.2%	0.05	0.09	0%	59%
4	1.0	9.399	0.934	91.0%	0.05	0.07	0%	32%
6	1.0	9.423	2.646	75.7%	0.05	0.07	0%	7%
8	1.0	9.162	3.739	60.8%	0.05	0.07	0%	1%
4	10	54.251	11.328	78.6%	0.05	0.07	0%	0%
6	10	26.827	8.085	68.2%	0.05	0.07	0%	0%
8	10	14.803	4.617	68.0%	0.05	0.07	0%	0%

when $\sigma^2 \leq 0.1$, and (2) since condition (1.5) is a sufficient condition for (ERSDP) to be tight, the probability of condition (1.5) being satisfied is lower than the probability of (ERSDP) being tight.

To gain more insight towards the numerical performance of (ERSDP), we carry out numerical experiments on more problem instances with different $M \in \{4, 6, 8\}$ and different $\sigma^2 \in \{0.01, 0.1, 1.0, 10\}$. More specifically, we test the performance of (ERSDP) on 12 different settings and for each setting we randomly generate 100 instances. Table 2 summarizes the average results (over the 100 problem instances) and the statistical results (computed based on the 100 problem instances), where "TimeC" and "TimeE" denote the average CPU time for solving (CSDP) and (ERSDP), respectively; "ProbC" and "ProbE" denote the probability that (CSDP) and (ERSDP) are tight, respectively.

We can conclude from Table 2 that (1) the probability that (ERSDP) is tight for (P) is very high if the noise level is low; (2) if the noise level is high, (ERSDP) is much tighter than (CSDP) and (ERSDP) narrows down over 60% of the gap due

to (CSDP); and (3) the CPU time of solving (ERSDP) is not much longer than that of solving (CSDP). Finally, (ERSDP) performs better in terms of the tightness on instances with smaller M (if the noise level is fixed). This is consistent with the sufficient condition in (1.5), which shows that a lower level of the noise is required to guarantee the tightness of (ERSDP) for problem (P) with a larger M.

6. Conclusion. In this paper, we considered the MIMO detection problem (P), an important class of quadratic programming problems with unit-modulus and discrete argument constraints. We showed that the conventional (CSDP) is generally not tight for (P). Moreover, we proposed a new enhanced SDR called (ERSDP) and showed that (ERSDP) is tight for problem (P) under the condition in (1.5). To the best of our knowledge, our proposed (ERSDP) is the first SDR that is theoretically guaranteed to be tight for general cases of problem (P). Our above results answered an open question posed by So in [35]. In addition to enjoying strong theoretical guarantees, our proposed (ERSDP) also performs very well numerically. Our numerical results show that (ERSDP) can return the true vector of transmitted signals with high probability if the level of the noise is low and it can narrow down more than 60% of the gap due to (CSDP) on average if the level of the noise is high. It is interesting to compare the tightness of some existing SDRs (e.g., in [8, 29]) with our proposed (ERSDP) and (CSDP2). If we could show that some existing SDRs are stronger/tighter than (CSDP2), then it follows from this and Theorems 4.2 and 4.4 that condition (1.5) is also a sufficient condition for them to be tight. We believe that our proposed (ERSDP) and related techniques will find further applications such as in the development of quantized precoding for massive multiuser MIMO communications [13, 37].

Appendix A. Equivalence between (CSDP) and (RSDP). We first show that, for any feasible point (x, X) of (CSDP), we can construct a feasible point (y, Y) of (RSDP) that satisfies $\hat{Q} \bullet Y + 2\hat{c}^\mathsf{T} y = Q \bullet X + 2\mathrm{Re}(c^\dagger x)$. Since

$$\begin{bmatrix} 1 & x^\dagger \\ x & X \end{bmatrix} \succeq 0,$$

we assume, without loss of generality, that it has the following decomposition:

$$(A.1) \qquad \begin{bmatrix} 1 & x^\dagger \\ x & X \end{bmatrix} = \sum_{j=1}^{n+1} \begin{bmatrix} t_j \\ v_j \end{bmatrix} \begin{bmatrix} t_j \\ v_j \end{bmatrix}^\dagger,$$

where $t_j \in \mathbb{R}_+$ and $v_j \in \mathbb{C}^n$ for all $j = 1, 2, \ldots, n+1$. Let

$$(A.2) \qquad y = \sum_{j=1}^{n+1} t_j \begin{bmatrix} \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix} \text{ and } Y = \sum_{j=1}^{n+1} \begin{bmatrix} \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix}^\mathsf{T}.$$

It is simple to see that

$$\begin{bmatrix} 1 & y^\mathsf{T} \\ y & Y \end{bmatrix} = \sum_{j=1}^{n+1} \begin{bmatrix} t_j \\ \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix} \begin{bmatrix} t_j \\ \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix}^\mathsf{T} \succeq 0.$$

It follows from (A.1) and (A.2) that, for any $i = 1, 2, \ldots, n$, we have

$$Y_{i,i} + Y_{n+i,n+i} = \sum_{j=1}^{n+1} \left([\mathrm{Re}(v_j)]_i^2 + [\mathrm{Im}(v_j)]_i^2 \right) = \sum_{j=1}^{n+1} \left| [v_j]_i \right|^2 = X_{i,i} = 1.$$

Hence, (y, Y) in (A.2) is a feasible solution to (RSDP). Moreover, we have, from (A.1) and (A.2), that

$$\hat{Q} \bullet Y + 2\hat{c}^{\mathsf{T}} y = \hat{Q} \bullet \left(\sum_{j=1}^{n+1} \begin{bmatrix} \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix}^{\mathsf{T}} \right) + 2\hat{c}^{\mathsf{T}} \left(\sum_{j=1}^{n+1} t_j \begin{bmatrix} \mathrm{Re}(v_j) \\ \mathrm{Im}(v_j) \end{bmatrix} \right)$$

$$= Q \bullet \left(\sum_{j=1}^{n+1} v_j v_j^{\dagger} \right) + 2\mathrm{Re} \left(c^{\dagger} \left(\sum_{j=1}^{n+1} t_j v_j \right) \right)$$

$$= Q \bullet X + 2\mathrm{Re}(c^{\dagger} x).$$

We can conclude from the above that (CSDP) is at least as tight as (RSDP).

Conversely, for any given feasible point (y, Y) of (RSDP), we can construct a feasible point (x, X) of (CSDP) that gives the same objective value, i.e., $Q \bullet X + 2\mathrm{Re}(c^{\dagger} x) = \hat{Q} \bullet Y + 2\hat{c}^{\mathsf{T}} y$. Suppose that

$$y = \begin{bmatrix} a \\ b \end{bmatrix} \text{ and } Y = \begin{bmatrix} A & B \\ B^{\mathsf{T}} & C \end{bmatrix},$$

where $a, b \in \mathbb{R}^n$, $A, B, C \in \mathbb{R}^{n \times n}$, $A = A^{\mathsf{T}}$, and $C = C^{\mathsf{T}}$. Let

(A.3) $$x = a + b\mathbf{i} \text{ and } X = (A + C) + (B^{\mathsf{T}} - B)\mathbf{i}.$$

Next, we first show

(A.4) $$\hat{X} := \begin{bmatrix} 1 & x^{\dagger} \\ x & X \end{bmatrix} \succeq 0.$$

Define

$$M_1 = \begin{bmatrix} 1 & a^{\mathsf{T}} & 0 & b^{\mathsf{T}} \\ a & A & \zeta & B \\ 0 & \zeta^{\mathsf{T}} & 0 & \zeta^{\mathsf{T}} \\ b & B^{\mathsf{T}} & \zeta & C \end{bmatrix} \text{ and } M_2 = \begin{bmatrix} 0 & \zeta^{\mathsf{T}} & 0 & \zeta^{\mathsf{T}} \\ \zeta & C & -b & -B^{\mathsf{T}} \\ 0 & -b^{\mathsf{T}} & 1 & a^{\mathsf{T}} \\ \zeta & -B & a & A \end{bmatrix},$$

where ζ is the all-zero column vector of dimension n. Then, it follows from

$$\begin{bmatrix} 1 & y^{\mathsf{T}} \\ y & Y \end{bmatrix} \succeq 0$$

that

$$M_1 \succeq 0 \text{ and } M_2 = \begin{bmatrix} & I_{n+1} \\ -I_{n+1} & \end{bmatrix} M_1 \begin{bmatrix} & -I_{n+1} \\ I_{n+1} & \end{bmatrix} \succeq 0,$$

which further implies

(A.5) $$M := M_1 + M_2 = \begin{bmatrix} 1 & a^{\mathsf{T}} & 0 & b^{\mathsf{T}} \\ a & A + C & -b & B - B^{\mathsf{T}} \\ 0 & -b^{\mathsf{T}} & 1 & a^{\mathsf{T}} \\ b & B^{\mathsf{T}} - B & a & A + C \end{bmatrix} \succeq 0.$$

Recall that $v \in \mathbb{C}^{n+1}$ is an eigenvector of $\hat{X}$ in (A.4) if and only if

$$\begin{bmatrix} \mathrm{Re}(v) \\ \mathrm{Im}(v) \end{bmatrix} \text{ and } \begin{bmatrix} -\mathrm{Im}(v) \\ \mathrm{Re}(v) \end{bmatrix}$$

are two eigenvectors of M in (A.5) corresponding to the same eigenvalue [11, Page 448]. Then, we have $\hat{X} \geq 0$. The positive semidefiniteness of $\hat{X}$ further implies that all of its diagonal entries are nonnegative. This, together with (A.3), immediately shows that $X_{i,i} = 1$ for all $i = 1, 2, \ldots, n$. Moreover, it is simple to check that x and X in (A.3) satisfy $Q \bullet X + 2\mathrm{Re}(c^\dagger x) = \hat{Q} \bullet Y + 2\hat{c}^\mathsf{T} y$. This shows that (RSDP) is at least as tight as (CSDP).

From the above analysis, we conclude that (CSDP) are (RSDP) are equivalent to each other.

Appendix B. Proof of Theorem 4.1. We first show that, for any feasible point (y, Y) of (ERSDP), we can construct a feasible point (x, X) of (CSDP2) that satisfies $Q \bullet X + 2\mathrm{Re}(c^\dagger x) = \hat{Q} \bullet Y + 2\hat{c}^\mathsf{T} y$. Similar to the proof in Appendix A, we can construct the same x and X in (A.3) and show that $\hat{X}$ in (A.4) satisfies $\hat{X} \succeq 0$, $X_{i,i} = 1$ for all $i = 1, 2, \ldots, n$, and $Q \bullet X + 2\mathrm{Re}(c^\dagger x) = \hat{Q} \bullet Y + 2\hat{c}^\mathsf{T} y$. For each $i = 1, 2, \ldots, n$, since $\mathcal{Y}(i) \in \mathrm{Conv}\{P_0, P_1, \ldots, P_{M-1}\}$ in (ERSDP), it follows that

$$(y_i,\, y_{i+n}) \in \mathrm{Conv}\left(\{(\cos\theta,\, \sin\theta) \mid \theta \in \mathcal{A}\}\right),$$

which is equivalent to

$$\mathrm{Re}(a_j^\dagger x_i) \leq \cos\left(\frac{\pi}{M}\right),\; j = 0, 1, \ldots, M-1,$$

where $x_i = y_i + y_{i+n}\mathbf{i}$ and a_j is given in (4.2). Hence, (x, X) is a feasible solution to (CSDP2). Based on the above analysis, we conclude that (ERSDP) is at least as tight as (CSDP2).

Next, we give an example to illustrate that (ERSDP) is indeed (strictly) tighter than (CSDP2). Let $m = n = 2$ and $M = 3$ (and hence the set $\mathcal{A}$ in (1.2) is $\{0, 2\pi/3, 4\pi/3\}$ in this case); let

$$x^* = \left(\begin{array}{c} \frac{-1-\sqrt{3}\mathbf{i}}{2} \\ 1 \end{array}\right),\; H = \left(\begin{array}{cc} 8-6\mathbf{i} & 8+6\mathbf{i} \\ 3+4\mathbf{i} & -4-3\mathbf{i} \end{array}\right),\; \text{and } v = \left(\begin{array}{c} 5-6\mathbf{i} \\ 4-4\mathbf{i} \end{array}\right);$$

and finally set $r = Hx^* + v$ as in (1.1). Then, we numerically solve (CSDP), (CSDP2), and (ERSDP), and their optimal objective values are

$$-76.3176,\; -45.1273,\; \text{and } -25.4763,$$

respectively. Hence, (CSDP2) is tighter (CSDP) and (ERSDP) is further tighter than (CSDP2).

Appendix C. Derivation of the dual problem of (CSDP2). The canonical form of (CSDP2) is

$$\begin{aligned} \min_{y,\, \hat{X}} \quad & \hat{C} \bullet \hat{X} \\ \text{s.t.} \quad & E_i \bullet \hat{X} = 1,\; i = 1, 2, \ldots, n+1, \\ & \mathrm{Re}(a_j^\dagger \hat{X}_{i+1,1}) + y_{i,j} = \cos\left(\frac{\pi}{M}\right),\; i = 1, 2, \ldots, n,\; j = 0, 1, \ldots, M-1, \\ & y \geq 0,\; \hat{X} \succeq 0, \end{aligned}$$

where $y \in \mathbb{R}^{M \times n}$ and $\hat{X} \in \mathbb{C}^{(n+1) \times (n+1)}$,

$$\hat{C} = \left(\begin{array}{cc} 0 & c^\dagger \\ c & Q \end{array}\right),$$

and $E_i \in \mathbb{R}^{(n+1)\times(n+1)}$ is the all-zero matrix except its ith diagonal entry being 1. Let τ be the Lagrange multiplier corresponding to the constraint $E_1 \bullet \hat{X} = 1$, let $-\lambda_i$ be the Lagrange multiplier corresponding to the constraint $E_{i+1}\bullet\hat{X} = 1$, and let $-\mu_{i,j}$ be the Lagrange multiplier corresponding to the constraint $\text{Re}(a_j^\dagger \hat{X}_{i+1,1})+y_{i,j} = \cos(\frac{\pi}{M})$. Notice that both the nonnegative orthant cone and the semidefinite cone are self-dual. Thus, (4.3) is the dual problem of (CSDP2).

Appendix D. Proof of Theorem 4.5. To prove Theorem 4.5, we first state the following two lemmas.

LEMMA D.1. *Under the same assumptions as in Theorem* 4.5*, the smallest singular value* $s_{\min}$ *of the matrix* $\frac{1}{\sqrt{2m}}H$ *satisfies*

$$(\text{D.1}) \qquad \mathbb{P}\left(s_{\min} \leq 1 - \sqrt{n/m} - t\right) \leq e^{-mt^2}.$$

For the proof of the above lemma, we refer the reader to [7] and [9, Chapter 9]. Substituting $\rho = \sqrt{n/m}$ and $t = (1-\rho)/2$ into (D.1), we get

$$\mathbb{P}\left(s_{\min} \leq (1-\rho)/2\right) \leq e^{-\frac{m(1-\rho)^2}{4}},$$

which, together with the fact $\lambda_{\min}(H^\dagger H) = 2ms_{\min}^2$, further implies

$$(\text{D.2}) \qquad \mathbb{P}\left(\lambda_{\min}(H^\dagger H) \leq m(1-\rho)^2/2\right) \leq e^{-\frac{m(1-\rho)^2}{4}}.$$

The next result is an extension of [35, Proposition 3.6].

LEMMA D.2. *Under the same assumptions as in Theorem* 4.5*, it holds that*

$$(\text{D.3}) \qquad \mathbb{P}\left(\|H^\dagger v\|_\infty > 2\sqrt{2}m\sigma\right) \leq 2\sqrt{2/\pi}\cdot n\cdot e^{-m/2} + 8e^{-m/8}.$$

Proof. Let

$$\hat{H} = \left[\hat{H}_1, \hat{H}_2\right],\ \hat{H}_1 = \left[\begin{array}{c} \text{Re}(H) \\ \text{Im}(H) \end{array}\right],\ \hat{H}_2 = \left[\begin{array}{c} -\text{Im}(H) \\ \text{Re}(H) \end{array}\right],\ \text{and}\ \hat{v} = \left[\begin{array}{c} \text{Re}(v) \\ \text{Im}(v) \end{array}\right].$$

It follows from [35, Proposition 3.6] that

$$\mathbb{P}\left(\|\hat{H}_1^\mathsf{T}\hat{v}\|_\infty \geq 2m\sigma\right) \leq \sqrt{2/\pi}\cdot n\cdot e^{-m/2} + 4e^{-m/8}$$

and

$$\mathbb{P}\left(\|\hat{H}_2^\mathsf{T}\hat{v}\|_\infty \geq 2m\sigma\right) \leq \sqrt{2/\pi}\cdot n\cdot e^{-m/2} + 4e^{-m/8},$$

and thus

$$\mathbb{P}\left(\|\hat{H}^\mathsf{T}\hat{v}\|_\infty \geq 2m\sigma\right) \leq 2\sqrt{2/\pi}\cdot n\cdot e^{-m/2} + 8e^{-m/8}.$$

Since $\|\hat{H}^\mathsf{T}\hat{v}\|_\infty \geq \frac{\sqrt{2}}{2}\|H^\dagger v\|_\infty$, we immediately get the desired result (D.3). $\square$

Now we are ready to prove Theorem 4.5. The proof is similar to the one in [35, Theorem 3.3]. First, combining (4.25) and (D.3), we have

$$(\text{D.4}) \qquad \mathbb{P}\left\{\|H^\dagger v\|_\infty \geq \frac{m(1-\rho)^2\sin(\pi/M)}{2}\right\} \leq 2\sqrt{2/\pi}\cdot n\cdot e^{-m/2} + 8e^{-m/8}.$$

Then, it follows from (D.2) and (D.4) that the probability that the event

$$\left\{ \lambda_{\min}(H^{\dagger}H) \geq \frac{m(1-\rho)^2}{2} \right\} \bigcap \left\{ \|H^{\dagger}v\|_{\infty} \leq \frac{m(1-\rho)^2 \sin(\pi/M)}{2} \right\}$$

happens is at least $1 - e^{-\frac{m(1-\rho)^2}{4}} - 2\sqrt{2/\pi} \cdot n \cdot e^{-m/2} - 8e^{-m/8}$. Finally, it is simple to verify that the probability that the above event happens is greater than or equal to the one that condition (1.5) is satisfied. The proof of Theorem 4.5 is complete.

REFERENCES

[1] F. ALIZADEH, J.-P. A. HAEBERLY, AND M. L. OVERTON, *Complementarity and nondegeneracy in semidefinite programming*, Math. Program., 77 (1997), pp. 111–128.

[2] A. S. BANDEIRA, N. BOUMAL, AND A. SINGER, *Tightness of the maximum likelihood semidefinite relaxation for angular synchronization*, Math. Program., 163 (2017), pp. 145–167.

[3] A. S. BANDEIRA, Y. KHOO, AND A. SINGER, *Open problem: Tightness of maximum likelihood semidefinite relaxations*, in Proceedings of the 27th Conference on Learning Theory, Proc. Mach. Learn. Res. 35, 2014, pp. 1265–1267; available at http://proceedings.mlr.press/v35/bandeira14.html.

[4] A. BECK AND M. TEBOULLE, *Global optimality conditions for quadratic optimization problems with binary constraints*, SIAM J. Optim., 11 (2000), pp. 179–188.

[5] N. BOUMAL, *Nonconvex phase synchronization*, SIAM J. Optim., 26 (2016), pp. 2355–2377.

[6] O. DAMEN, A. CHKEIF, AND J.-C. BELFIORE, *Lattice code decoder for space-time codes*, IEEE Commun. Lett., 4 (2000), pp. 161–163.

[7] K. R. DAVIDSON AND S. J. SZAREK, *Local Operator Theory, Random Matrices and Banach Spaces*, in Handbook of the Geometry of Banach Spaces, Vol. I, North-Holland, Amsterdam, 2001, pp. 317–366.

[8] X. FAN, J. SONG, D. P. PALOMAR, AND O. C. AU, *Universal binary semidefinite relaxation for ML signal detection*, IEEE Trans. Commun., 61 (2013), pp. 4565–4576.

[9] F. S. FOUCART AND H. RAUHUT, *A Mathematical Introduction to Compressive Sensing*, Birkhäuser, Basel, Switzerland, 2013.

[10] M. GOEMANS AND D. WILLIAMSON, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM, 42 (1995), pp. 1115–1145.

[11] M. GOEMANS AND D. WILLIAMSON, *Approximation algorithms for Max-3-Cut and other problems via complex semidefinite programming*, J. Comput. Syst. Sci., 68 (2004), pp. 442–470.

[12] M. HONIG, U. MADHOW, AND S. VERDU, *Blind adaptive multiuser detection*, IEEE Trans. Inf. Theory, 41 (1995), pp. 944–960.

[13] S. JACOBSSON, G. DURISI, M. GOLDSTEIN, AND C. STUDER, *Quantized precoding for massive MU-MIMO*, IEEE Trans. Commun., 65 (2017), pp. 4670–4684.

[14] J. JALDÉN, *Detection for Multiple Input Multiple Output Channels*, Ph.D. Thesis, School of Electrical Engineering, KTH, Stockholm, Sweden, 2006.

[15] J. JALDÉN, C. MARTIN, AND B. OTTERSTEN, *Semidefinite programming for detection in linear systems – Optimality conditions and space-time decoding*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03), IEEE Press, Piscataway, NJ, 2003, pp. 9–12.

[16] J. JALDÉN AND B. OTTERSTEN, *The diversity order of the semidefinite relaxation detector*, IEEE Trans. Inf. Theory, 54 (2008), pp. 1406–1422.

[17] M. KISIALIOU AND Z.-Q. LUO, *Performance analysis of quasi-maximum-likelihood detector based on semi-definite programming*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05), IEEE Press, Piscataway, NJ, 2005, pp. 433–436.

[18] M. KISIALIOU AND Z.-Q. LUO, *Probabilistic analysis of semidefinite relaxation for binary quadratic minimization*, SIAM J. Optim., 20 (2010), pp. 1906–1922.

[19] R. KOHNO, H. IMAI, AND M. HATORI, *Cancellation techniques of co-channel interference in asynchronous spread spectrum multiple access systems*, Trans. IECE Japan, 66-A (1983), pp. 416–423.

[20] C. LEMARECHAL AND F. OUSTRY, *SDP relaxations in combinatorial optimization from a Lagrangian point of view*, in Advances in Convex Analysis and Global Optimization, N. Hadijsavvas and P. M. Paradalos, eds., Kluwer, Norwell, MA, 2001, pp. 119–134.

[21] H. LIU, M.-C. YUE, AND A. M.-C. SO, *On the estimation performance and convergence rate of the generalized power method for phase synchronization*, SIAM J. Optim., 27 (2017), pp. 2426–2446.

[22] H. LIU, M.-C. YUE, AND A. M.-C. SO, *A discrete first-order method for large-scale MIMO detection with provable guarantees*, in Proceedings of the 18th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC'17), IEEE Press, Piscataway, NJ, 2017, pp. 669–673.

[23] Y.-F. LIU, M. HONG, AND Y.-H. DAI, *Max-min fairness linear transceiver design problem for a multi-user SIMO interference channel is polynomial time solvable*, IEEE Signal Process. Lett., 20 (2013), pp. 27–30.

[24] C. LU AND Y.-F. LIU, *An efficient global algorithm for single-group multicast beamforming*, IEEE Trans. Signal Process., 65 (2017), pp. 3761–3774.

[25] C. LU, Y.-F. LIU, AND J. ZHOU, *An efficient global algorithm for nonconvex complex quadratic problems with applications in wireless communications*, in Proceedings of the 6th IEEE/CIC International Conference on Communications in China (ICCC'17), IEEE Press, Piscataway, NJ, 2017, pp. 1–5.

[26] 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., 27 (2010), pp. 20–34.

[27] W.-K. MA, P.-C. CHING, AND Z. DING, *Semidefinite relaxation based multiuser detection for M-ary PSK multiuser systems*, IEEE Trans. Signal Process., 52 (2004), pp. 2862–2872.

[28] A. D. MAIO, S. D. NICOLA, Y. HUANG, Z.-Q. LUO, AND S. ZHANG, *Design of phase codes for radar performance optimization with a similarity constraint*, IEEE Trans. Signal Process., 57 (2009), pp. 610–621.

[29] A. MOBASHER, M. TAHERZADEH, R. SOTIROV, AND A. K. KHANDANI, *A near-maximum-likelihood decoding algorithm for MIMO systems based on semi-definite programming*, IEEE Trans. Inf. Theory, 53 (2007), pp. 3869–3886.

[30] A. D. MURUGAN, H. E. GAMAL, M. O. DAMEN, AND G. CAIRE, *A unified framework for tree search decoding: Rediscovering the sequential decoder*, IEEE Trans. Inf. Theory, 52 (2006), pp. 933–953.

[31] W. PU, Y.-F. LIU, J. YAN, H. LIU, AND Z.-Q. LUO, *Optimal estimation of sensor biases for asynchronous multi-sensor data fusion*, Math. Program. 170 (2018), pp. 357–386.

[32] K. S. SCHNEIDER, *Optimum detection of code division multiplexed signals*, IEEE Trans. Aerosp. Electron. Syst., 15 (1979), pp. 181–185.

[33] N. Z. SHOR AND A. S. DAVYDOV, *Method of obtaining estimates in quadratic extremal problems with Boolean variables*, Cybern. Syst. Anal., 21 (1985), pp. 207–210.

[34] A. SINGER, *Angular synchronization by eigenvectors and semidefinite programming*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 20–36.

[35] A. M.-C. SO, *Probabilistic analysis of the semidefinite relaxation detector in digital communications*, in Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'10), SIAM, Philadelphia, PA, 2011, pp. 698–711.

[36] A. M.-C. SO, J. ZHANG, AND Y. YE , *On approximating complex quadratic optimization problems via semidefinite programs*, Math. Program., 110 (2007), pp. 93–110.

[37] F. SOHRABI, Y.-F. LIU, AND W. YU, *One-bit precoding and constellation range design for massive MIMO with QAM signaling*, IEEE J. Sel. Topics Signal Process., 12 (2018), pp. 557–570.

[38] M. SOLTANALIAN AND P. STOICA, *Designing unimodular codes via quadratic optimization*, IEEE Trans. Signal Process., 62 (2014), pp. 1221–1234.

[39] J. SUN, Q. QU, AND J. WRIGHT, *When are Nonconvex Problems Not Scary?*, preprint, https://arxiv.org/abs/1510.06096, 2016.

[40] P. H. TAN AND L. K. RASMUSSEN, *The application of semidefinite programming for detection in CDMA*, IEEE J. Sel. Areas Commun., 19 (2001), pp. 1442–1449.

[41] M. K. VARANASI, *Decision feedback multiuser detection: A systematic approach*, IEEE Trans. Inf. Theory, 45 (1999), pp. 219–240.

[42] S. VERDÚ, *Computational complexity of optimum multiuser detection*, Algorithmica, 4 (1989), pp. 303–312.

[43] S. VERDÚ, *Multiuser Detection*, Cambridge University Press, New York, 1998.

[44] I. Waldspurger, A. Aspremont, and S. Mallat, *Phase recovery, MaxCut and complex semidefinite programming*, Math. Program., 149 (2015), pp. 47–81.

[45] Z. Xie, R. T. Short, and C. K. Rushforth, *A family of suboptimum detectors for coherent multi-user communications*, IEEE J. Sel. Areas Commun., 8 (1990), pp. 683–690.

[46] S. Yang and L. Hanzo, *Fifty years of MIMO detection: The road to large-scale MIMOs*, IEEE Commun. Surveys Tuts., 17 (2015), pp. 1941–1988.

[47] S. Zhang and Y. Huang, *Complex quadratic optimization and semidefinite programming*, SIAM J. Optim., 16 (2006), pp. 871–890.