Approximation Bounds for Quadratic Optimization with Homogeneous Quadratic Constraints*

Zhi-Quan Luo†, Nicholas D. Sidiropoulos‡, Paul Tseng§ and Shuzhong Zhang¶

June 17, 2006

Abstract

We consider the NP-hard problem of finding a minimum norm vector in n-dimensional real or complex Euclidean space, subject to m concave homogeneous quadratic constraints. We show that a semidefinite programming (SDP) relaxation for this nonconvex quadratically constrained quadratic program (QP) provides an $O(m^2)$ approximation in the real case, and an $O(m)$ approximation in the complex case. Moreover, we show that these bounds are tight up to a constant factor. When the Hessian of each constraint function is of rank 1 (namely, outer products of some given so-called *steering* vectors) and the phase spread of the entries of these steering vectors are bounded away from $\pi/2$, we establish a certain "constant factor" approximation (depending on the phase spread but independent of m and n) for both the SDP relaxation and a convex QP restriction of the original NP-hard problem. Finally, we consider a related problem of finding a maximum norm vector subject to m convex homogeneous quadratic constraints. We show that a SDP relaxation for this nonconvex QP provides an $O(1/\ln(m))$ approximation, which is analogous to a result of Nemirovski, Roos and Terlaky [14] for the real case.

*The first author is supported in part by the National Science Foundation, Grant No. DMS-0312416, and by the Natural Sciences and Engineering Research Council of Canada, Grant No. OPG0090391. The second author is supported in part by the U.S. ARO under ERO, Contract No. N62558-03-C-0012, and the EU under U-BROAD STREP, Grant No. 506790. The third author is supported by the National Science Foundation, Grant No. DMS-0511283. The fourth author is supported by Hong Kong RGC Earmarked Grant CUHK418505.

†Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, U.S.A. (luozq@ece.umn.edu)

‡Department of Electronic and Computer Engineering, Technical University of Crete, 73100 Chania - Crete, Greece. (nikos@telecom.tuc.gr)

§Department of Mathematics, University of Washington, Seattle, Washington 98195, U.S.A. (tseng@math.washington.edu)

¶Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. (zhang@se.cuhk.edu.hk)

1 Introduction

Consider the quadratic optimization problem with concave homogeneous quadratic constraints:

$$
\begin{array}{rll}
v_{\mathrm{qp}} := \min & \|z\|^2 & \\
\text{s.t.} & \displaystyle\sum_{\ell\in\mathcal{I}_i} |h_\ell^H z|^2 \ge 1, \quad i = 1, ..., m, & \\
& z \in \mathbb{F}^{\mathrm{n}}, &
\end{array} \tag{1}
$$

where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$, $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{F}^{\mathrm{n}}$, $m \ge 1$, each h_ℓ is a given vector in $\mathbb{F}^{\mathrm{n}}$, and $\mathcal{I}_1, ..., \mathcal{I}_m$ are nonempty, mutually disjoint index sets satisfying $\mathcal{I}_1 \cup \cdots \cup \mathcal{I}_m = \{1, ..., M\}$. Throughout, the superscript "H" will denote the complex Hermitian transpose, i.e., for $z = x + \mathbf{i}y$, where $x, y \in \mathbb{R}^{\mathrm{n}}$ and $\mathbf{i}^2 = -1$, $z^H = x^T - \mathbf{i}y^T$. Geometrically, the above problem (1) corresponds to finding a least norm vector in a region defined by the intersection of the exteriors of m co-centered ellipsoids. If the vectors $h_1, ..., h_M$ are linearly independent, then M equals the sum of the rank of the matrices defining these m ellipsoids. Notice that the problem (1) is easily solved for the case of $n = 1$, so we assume $n \ge 2$.

We assume that $\sum_{\ell\in\mathcal{I}_i} \|h_\ell\| \neq 0$ for all i, which is clearly a necessary condition for (1) to be feasible. This is also a sufficient condition (since $\bigcup_{i=1}^m \{z \mid \sum_{\ell\in\mathcal{I}_i} |h_\ell^H z|^2 = 0\}$ is a finite union of proper subspaces of $\mathbb{F}^{\mathrm{n}}$, so its complement is nonempty and any point in its complement can be scaled to be feasible for (1)). Thus, the above problem (1) always has an optimal solution (not necessarily unique) since its objective function is coercive, continuous, and its feasible set is nonempty, closed. Notice, however, that the feasible set of (1) is typically nonconvex and disconnected, with an exponential number of connected components exhibiting little symmetry. This is in contrast to the quadratic problems with convex feasible set but nonconvex objective function considered in [13, 14, 22]. Furthermore, unlike the class of quadratic problems studied in [1, 7, 8, 15, 16, 21, 23, 24, 25, 26], the constraint functions in (1) do not depend on $z_1^2, ..., z_n^2$ only.

Our interest in the nonconvex QP (1) is motivated by the transmit beamforming problem for multicasting applications [20] and by the wireless sensor network localization problem [6]. In the transmit beamforming problem, a transmitter utilizes an array of n transmitting antennas to broadcast information within its service area to m radio receivers, with receiver $i \in \{1, ..., m\}$ equipped with $|\mathcal{I}_i|$ receiving antennas. Let h_ℓ, $\ell \in \mathcal{I}_i$, denote the $n \times 1$ complex *steering vector* modelling propagation loss and phase shift from the transmitting antennas to the ℓth receiving antenna of receiver i. Assuming that each receiver performs spatially matched filtering / maximum ratio combining, which is the optimal combining strategy under standard mild assumptions, then the constraint

$$\sum_{\ell \in \mathcal{I}_i} |h_\ell^H z|^2 \geq 1$$

models the requirement that the total received signal power at receiver i must be above a given threshold (normalized to 1). This constraint is also equivalent to a signal-to-noise ratio (SNR) condition commonly used in data communication. Thus, to minimize the total transmit power subject to individual SNR requirements (one at each receiver), we are led to the QP (1). In the special case where each radio receiver is equipped with a single receiving antenna, the problem reduces to [20]:

$$\begin{array}{ll} \min & \|z\|^2 \\ \text{s.t.} & |h_\ell^H z|^2 \geq 1, \quad \ell = 1, ..., m, \\ & z \in \mathbb{F}^{\mathrm{n}}, \end{array} \tag{2}$$

This problem is a special case of (1) whereby each ellipsoid lies in $\mathbb{F}^{\mathrm{n}}$ and the corresponding matrix has rank 1.

In this paper, we first show that the nonconvex QP (2) is NP-hard in either the real or the complex case, which further implies the NP-hardness of the general problem (1). Then, we consider a semidefinite programming (SDP) *relaxation* of (1) and a convex QP *restriction* of (2) and study their worst-case performance. In particular, let υ_{sdp}, υ_{cqp} and υ_{qp} denote the optimal values of the SDP relaxation, the convex QP restriction, and the original QP (1), respectively. We establish a performance ratio of $\upsilon_{\mathrm{qp}}/\upsilon_{\mathrm{sdp}} = O(m^2)$ for the SDP relaxation in the real case, and we give an example showing that this bound is tight up to a constant factor. Similarly, we establish a performance ratio of $\upsilon_{\mathrm{qp}}/\upsilon_{\mathrm{sdp}} = O(m)$ in the complex case, and we give an example showing the tightness of this bound. We further show that, in the case when the phase spread of the entries of $h_1, ..., h_M$ is bounded away from $\pi/2$, the performance ratios $\upsilon_{\mathrm{qp}}/\upsilon_{\mathrm{sdp}}$ and $\upsilon_{\mathrm{cqp}}/\upsilon_{\mathrm{qp}}$ for the SDP relaxation and the convex QP restriction, respectively, are independent of m and n.

In recent years, there have been extensive studies of the performance of SDP relaxations for nonconvex QP. However, to our knowledge, this is the first performance analysis of SDP relaxation for QP with concave quadratic constraints. Our proof techniques also extend to a maximization version of the QP (1) with convex homogeneous quadratic constraints. In particular, we give a simple proof of a result analogous to one of Nemirovski, Roos and Terlaky [14] (also see [13, Theorem 4.7]) for the real case, namely, the SDP relaxation for this nonconvex QP has a performance ratio of $O(1/\ln(m))$.

2 NP-hardness

In this section, we show that the nonconvex QP (1) is NP-hard in general. First, we notice that, by a linear transformation if necessary, the following problem

$$\begin{array}{ll} \text{minimize} & z^H Q z \\ \text{subject to} & |z_\ell| \geq 1, \quad \ell = 1, ..., n, \\ & z \in \mathbb{F}^{\mathrm{n}}, \end{array} \tag{3}$$

is a special case of (1), where $Q \in \mathbb{F}^{\mathrm{n}\times\mathrm{n}}$ is a Hermitian positive definite matrix (i.e., $Q \succ 0$), and z_ℓ denotes the ℓth component of z. Hence, it suffices to establish the NP-hardness of (3). To this end, we consider a reduction from the NP-complete partition problem: Given positive integers $a_1,\ a_2,\ ...,\ a_N$, decide whether there exists a subset $\mathcal{I}$ of $\{1, ..., N\}$ satisfying

$$\sum_{\ell\in\mathcal{I}} a_\ell = \frac{1}{2}\sum_{\ell=1}^{N} a_\ell. \tag{4}$$

Our reductions differ for the real and complex cases. As will be seen, the NP-hardness proof in the complex case[1] is more intricate than in the real case.

2.1 The Real Case

We consider the real case of $\mathbb{F} = \mathbb{R}$. Let $n := N$ and

$$\begin{aligned} a &:= (a_1, \ldots, a_N)^T, \\ Q &:= aa^T + I_n \succ 0, \end{aligned}$$

where I_n denotes the $n \times n$ identity matrix.

We show that a subset $\mathcal{I}$ satisfying (4) exists if and only if the optimization problem (3) has a minimum value of n. Since

$$z^T Q z = |a^T z|^2 + \sum_{\ell=1}^{n} |z_\ell|^2 \geq n \quad \text{whenever} \quad |z_\ell| \geq 1\ \forall\, \ell,\ z \in \mathbb{R}^{\mathrm{n}},$$

we see that (3) has a minimum value of n if and only if there exists a $z \in \mathbb{R}^{\mathrm{n}}$ satisfying

$$a^T z = 0, \quad |z_\ell| = 1\ \ \forall\, \ell.$$

The above condition is equivalent to the existence of a subset $\mathcal{I}$ satisfying (4), with the correspondence $\mathcal{I} = \{\ell \mid z_\ell = 1\}$. This completes the proof.

[1] This NP-hardness proof was first presented in an appendix of [20] and is included here for completeness; also see [26, Proposition 3.5] for a related proof.

2.2 The Complex Case

We consider the complex case of $\mathbb{F} = \mathbb{C}$. Let $n := 2N + 1$ and

$$\begin{array}{rcl} a & := & (a_1, \ldots, a_N)^T, \\ A & := & \left(\begin{array}{ccc} I_N & I_N & -e_N \\ a^T & 0_N^T & -\frac{1}{2} a^T e_N \end{array} \right), \\ Q & := & A^T A + I_n \succ 0, \end{array}$$

where e_N denotes the N-dimensional vector of ones, 0_N denotes the N-dimensional vector of zeros, and I_n and I_N are identity matrices of sizes $n \times n$ and $N \times N$, respectively.

We show that a subset $\mathcal{I}$ satisfying (4) exists if and only if the optimization problem (3) has a minimum value of n. Since

$$z^H Q z = \|Az\|^2 + \sum_{\ell=1}^{n} |z_\ell|^2 \geq n \quad \text{whenever } \; |z_\ell| \geq 1 \; \forall\, \ell, \; z \in \mathbb{C}^{\mathrm{n}},$$

we see that (3) has a minimum value of n if and only if there exists a $z \in \mathbb{C}^{\mathrm{n}}$ satisfying

$$Az = 0, \quad |z_\ell| = 1 \;\; \forall\, \ell.$$

Expanding $Az = 0$ gives the following set of linear equations:

$$0 = z_\ell + z_{N+\ell} - z_n, \quad \ell = 1, ..., N, \tag{5}$$

$$0 = \sum_{\ell=1}^{N} a_\ell z_\ell - \frac{1}{2} \left(\sum_{\ell=1}^{N} a_\ell \right) z_n. \tag{6}$$

For $\ell = 1, ..., 2N$, since $|z_\ell| = |z_n| = 1$ so that $z_\ell / z_n = e^{\mathbf{i}\theta_\ell}$ for some $\theta_\ell \in [0, 2\pi)$, we can rewrite (5) as

$$\begin{array}{rcl} \cos\theta_\ell + \cos\theta_{N+\ell} & = & 1, \\ \sin\theta_\ell + \sin\theta_{N+\ell} & = & 0, \end{array} \quad \ell = 1, ..., N.$$

These equations imply that $\theta_\ell \in \{-\pi/3, \pi/3\}$ for all $\ell \neq n$. In fact, these equations further imply that $\cos\theta_\ell = \cos\theta_{N+\ell} = 1/2$ for $\ell = 1, ..., N$, so that

$$\mathrm{Re}\left(\sum_{\ell=1}^{N} a_\ell \frac{z_\ell}{z_n} - \frac{1}{2} \left(\sum_{\ell=1}^{N} a_\ell \right) \right) = 0.$$

Therefore, (6) is satisfied if and only if

$$\mathrm{Im}\left(\sum_{\ell=1}^{N} a_\ell \frac{z_\ell}{z_n} - \frac{1}{2} \left(\sum_{\ell=1}^{N} a_\ell \right) \right) = \mathrm{Im}\left(\sum_{\ell=1}^{N} a_\ell \frac{z_\ell}{z_n} \right) = 0,$$

which is further equivalent to the existence of a subset $\mathcal{I}$ satisfying (4), with the correspondence $\mathcal{I} = \{\ell \mid \theta_\ell = \pi/3\}$. This completes the proof.

3 Performance analysis of SDP relaxation

In this section, we study the performance of an SDP relaxation of (2). Let

$$H_i := \sum_{\ell \in \mathcal{I}_i} h_\ell h_\ell^H, \quad i = 1, ..., m.$$

The well-known SDP relaxation of (1) [11, 19] is

$$\begin{aligned} v_{\text{sdp}} \;:=\; \min \quad & \mathrm{Tr}(Z) \\ \text{s.t.} \quad & \mathrm{Tr}(H_i Z) \geq 1, \quad i = 1, ..., m, \\ & Z \succeq 0, \quad Z \in \mathbb{F}^{\mathrm{n}\times\mathrm{n}} \text{ is Hermitian.} \end{aligned} \tag{7}$$

An optimal solution of the SDP relaxation (7) can be computed efficiently using, say, interior-point methods; see [18] and references therein.

Clearly $v_{\text{sdp}} \leq v_{\text{qp}}$. We are interested in upper bounds for the relaxation performance of the form

$$v_{\text{qp}} \leq C v_{\text{sdp}},$$

where $C \geq 1$. Since we assume $H_i \neq 0$ for all i, it is easily checked that (7) has an optimal solution, which we denote by Z^*.

3.1 General steering vectors: the real case

We consider the real case of $\mathbb{F} = \mathbb{R}$. Upon obtaining an optimal solution Z^* of (7), we construct a feasible solution of (1) using the following randomization procedure:

1. Generate a random vector $\xi \in \mathbb{R}^{\mathrm{n}}$ from the real-valued normal distribution $N(0, Z^*)$.
2. Let $z^*(\xi) = \xi / \min\limits_{1 \leq i \leq m} \sqrt{\xi^T H_i \xi}$.

We will use $z^*(\xi)$ to analyze the performance of the SDP relaxation. Similar procedures have been used for related problems [1, 3, 4, 5, 14]. First, we need to develop two lemmas. The first lemma estimates the left-tail of the distribution of a convex quadratic form of a Gaussian random vector.

Lemma 1 *Let $H \in \mathbb{R}^{n\times n}$, $Z \in \mathbb{R}^{n\times n}$ be two symmetric positive semidefinite matrices (i.e., $H \succeq 0$, $Z \succeq 0$). Suppose $\xi \in \mathbb{R}^n$ is a random vector generated from the real-valued normal distribution $N(0, Z)$. Then, for any $\gamma > 0$,*

$$\text{Prob}\left(\xi^T H\xi < \gamma E(\xi^T H\xi)\right) \leq \max\left\{\sqrt{\gamma}, \frac{2(\bar{r}-1)\gamma}{\pi - 2}\right\}, \tag{8}$$

where $\bar{r} := \min\{\text{rank}\,(H), \text{rank}\,(Z)\}$.

Proof. Since the covariance matrix $Z \succeq 0$ has rank $r := \text{rank}\,(Z)$, we can write $Z = UU^T$, for some $U \in \mathbb{R}^{n\times r}$ satisfying $U^T ZU = I_r$. Let $\bar{\xi} := Q^T U^T \xi \in \mathbb{R}^r$, where $Q \in \mathbb{R}^{r\times r}$ is an orthogonal matrix corresponding to the eigen-decomposition of the matrix

$$U^T HU = Q\Lambda Q^T,$$

for some diagonal matrix $\Lambda = \text{Diag}\{\lambda_1, \lambda_2, ..., \lambda_r\}$, with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_r \geq 0$. Since $U^T HU$ has rank at most $\bar{r}$, we have $\lambda_i = 0$ for all $i > \bar{r}$. It is readily checked that $\bar{\xi}$ has the normal distribution $N(0, I_r)$. Moreover, ξ is statistically identical to $UQ\bar{\xi}$, so that $\xi^T H\xi$ is statistically identical to

$$\bar{\xi}^T Q^T U^T HUQ\bar{\xi} = \bar{\xi}^T \Lambda\bar{\xi} = \sum_{i=1}^{\bar{r}} \lambda_i |\bar{\xi}_i|^2.$$

Then, we have

$$\begin{aligned}
\text{Prob}\left(\xi^T H\xi < \gamma E(\xi^T H\xi)\right) &= \text{Prob}\left(\sum_{i=1}^{\bar{r}} \lambda_i |\bar{\xi}_i|^2 < \gamma E\left(\sum_{i=1}^{\bar{r}} \lambda_i |\bar{\xi}_i|^2\right)\right) \\
&= \text{Prob}\left(\sum_{i=1}^{\bar{r}} \lambda_i |\bar{\xi}_i|^2 < \gamma \sum_{i=1}^{\bar{r}} \lambda_i\right).
\end{aligned}$$

If $\lambda_1 = 0$, then this probability is zero, which proves (8). Thus, we will assume that $\lambda_1 > 0$. Let $\bar{\lambda}_i := \lambda_i/(\lambda_1 + \cdots + \lambda_{\bar{r}})$, for $i = 1, ..., \bar{r}$. Clearly, we have

$$\bar{\lambda}_1 + \cdots + \bar{\lambda}_{\bar{r}} = 1, \quad \bar{\lambda}_1 \geq \bar{\lambda}_2 \geq \ldots \geq \bar{\lambda}_{\bar{r}} \geq 0.$$

We consider two cases. First, suppose $\bar{\lambda}_1 \geq \alpha$, where $0 < \alpha < 1$. Then, we can bound the above probability as follows:

$$\begin{aligned}
\text{Prob}\left(\xi^T H\xi < \gamma E(\xi^T H\xi)\right) &= \text{Prob}\left(\sum_{i=1}^{\bar{r}} \bar{\lambda}_i |\bar{\xi}_i|^2 < \gamma\right) \\
&\leq \text{Prob}\left(\bar{\lambda}_1 |\bar{\xi}_1|^2 < \gamma\right) \\
&\leq \text{Prob}\left(|\bar{\xi}_1|^2 < \gamma/\alpha\right) \qquad (9) \\
&\leq \sqrt{\frac{2\gamma}{\pi\alpha}},
\end{aligned}$$

where the last step is due to the fact that $\bar{\xi}_1$ is a real-valued zero mean Gaussian random variable with unit variance.

In the second case, we have $\bar{\lambda}_1 < \alpha$, so that

$$\bar{\lambda}_2 + \cdots + \bar{\lambda}_{\bar{r}} = 1 - \bar{\lambda}_1 > 1 - \alpha.$$

This further implies $(\bar{r}-1)\bar{\lambda}_2 \geq \bar{\lambda}_2 + \cdots + \bar{\lambda}_{\bar{r}} > 1 - \alpha$. Hence

$$\bar{\lambda}_1 \geq \bar{\lambda}_2 > \frac{1-\alpha}{\bar{r}-1}.$$

Using this bound, we obtain the following probability estimate:

$$\begin{aligned}
\text{Prob}\left(\xi^T H\xi < \gamma E(\xi^T H\xi)\right) &= \text{Prob}\left(\sum_{i=1}^{\bar{r}} \bar{\lambda}_i|\bar{\xi}_i|^2 < \gamma\right) \\
&\leq \text{Prob}\left(\bar{\lambda}_1|\bar{\xi}_1|^2 < \gamma,\ \bar{\lambda}_2|\bar{\xi}_2|^2 < \gamma\right) \\
&= \text{Prob}\left(\bar{\lambda}_1|\bar{\xi}_1|^2 < \gamma\right) \cdot \text{Prob}\left(\bar{\lambda}_2|\bar{\xi}_2|^2 < \gamma\right) \qquad (10) \\
&\leq \sqrt{\frac{2\gamma}{\pi\bar{\lambda}_1}} \cdot \sqrt{\frac{2\gamma}{\pi\bar{\lambda}_2}} \\
&\leq \frac{2(\bar{r}-1)\gamma}{\pi(1-\alpha)}.
\end{aligned}$$

Combining the estimates for the above two cases and setting $\alpha = 2/\pi$, we immediately obtain the desired bound (8). ■

Lemma 2 *Let* $\mathbb{F} = \mathbb{R}$. *Let* $Z^* \succeq 0$ *be a feasible solution of* (7) *and let* $z^*(\xi)$ *be generated by the randomization procedure described earlier. Then, with probability 1,* $z^*(\xi)$ *is well defined and feasible for* (1). *Moreover, for every* $\gamma > 0$ *and* $\mu > 0$,

$$\text{Prob}\left(\min_{1\leq i\leq m} \xi^T H_i\xi \geq \gamma,\ \|\xi\|^2 \leq \mu \text{Tr}(Z^*)\right) \geq 1 - m \cdot \max\left\{\sqrt{\gamma}, \frac{2(r-1)\gamma}{\pi-2}\right\} - \frac{1}{\mu}, \qquad (11)$$

where $r := \text{rank}\,(Z^*)$.

Proof. Since $Z^* \succeq 0$ is feasible for (7), it follows that $\text{Tr}(H_i Z^*) \geq 1$ for all $i = 1, ..., m$. Since $E(\xi^T H_i \xi) = \text{Tr}(H_i Z^*) \geq 1$ and the density of $\xi^T H_i \xi$ is absolutely continuous, the probability of $\xi^T H_i\xi = 0$ is zero, implying that $z^*(\xi)$ is well defined with probability 1. The feasibility of $z^*(\xi)$ is easily verified.

To prove (11), we first note that $E(\xi\xi^T) = Z^*$. Thus, for any $\gamma > 0$ and $\mu > 0$,

$$
\begin{aligned}
&\text{Prob}\left(\min_{1\le i\le m} \xi^T H_i \xi \ge \gamma,\ \|\xi\|^2 \le \mu \text{Tr}(Z^*)\right) \\
&= \text{Prob}\left(\xi^T H_i \xi \ge \gamma\ \forall\, i = 1, ..., m \text{ and } \|\xi\|^2 \le \mu \text{Tr}(Z^*)\right) \\
&\ge \text{Prob}\left(\xi^T H_i \xi \ge \gamma \text{Tr}(H_i Z^*)\ \forall\, i = 1, ..., m \text{ and } \|\xi\|^2 \le \mu \text{Tr}(Z^*)\right) \\
&= \text{Prob}\left(\xi^T H_i \xi \ge \gamma E(\xi^T H_i \xi)\ \forall\, i = 1, ..., m \text{ and } \|\xi\|^2 \le \mu E(\|\xi\|^2)\right) \\
&= 1 - \text{Prob}\left(\xi^T H_i \xi < \gamma E(\xi^T H_i \xi) \text{ for some } i \text{ or } \|\xi\|^2 > \mu E(\|\xi\|^2)\right) \\
&\ge 1 - \sum_{i=1}^m \text{Prob}\left(\xi^T H_i \xi < \gamma E(\xi^T H_i \xi)\right) - \text{Prob}\left(\|\xi\|^2 > \mu E(\|\xi\|^2)\right) \\
&> 1 - m \cdot \max\left\{\sqrt{\gamma}, \frac{2(r-1)\gamma}{\pi - 2}\right\} - \frac{1}{\mu},
\end{aligned}
$$

where the last step uses Lemma 1 as well as Markov's inequality:

$$\text{Prob}\left(\|\xi\|^2 > \mu E(\|\xi\|^2)\right) \le \frac{1}{\mu}.$$

This completes the proof. ■

We now use Lemma 2 to bound the performance of the SDP relaxation.

Theorem 1 *Let $\mathbb{F} = \mathbb{R}$. For the QP* (1) *and its SDP relaxation* (7)*, we have $\upsilon_{\text{qp}} = \upsilon_{\text{sdp}}$ if $m \le 2$, and otherwise*

$$\upsilon_{\text{qp}} \le \frac{27m^2}{\pi} \upsilon_{\text{sdp}}.$$

Proof. By applying a suitable rank reduction procedure if necessary, we can assume that the rank r of the optimal SDP solution Z^* satisfies $r(r+1)/2 \le m$; see e.g. [17]. Thus $r < \sqrt{2m}$. If $m \le 2$, then $r = 1$, implying that $Z^* = z^*(z^*)^T$ for some $z^* \in \mathbb{R}^n$ and it is readily seen that z^* is an optimal solution of (1), so that $\upsilon_{\text{qp}} = \upsilon_{\text{sdp}}$. Otherwise, we apply the randomization procedure to Z^*. We also choose

$$\mu = 3, \quad \gamma = \frac{\pi}{4m^2}\left(1 - \frac{1}{\mu}\right)^2 = \frac{\pi}{9m^2}.$$

Then, it is easily verified using $r < \sqrt{2m}$ that

$$\sqrt{\gamma} \ge \frac{2(r-1)\gamma}{\pi - 2} \quad \forall\, m = 1, 2, ...$$

Plugging these choices of γ and μ into (11), we see that there is a positive probability (independent of problem size) of at least

$$1 - m\sqrt{\gamma} - \frac{1}{\mu} = 1 - \frac{\sqrt{\pi}}{3} - \frac{1}{3} = 0.0758...$$

that ξ generated by the randomization procedure satisfies

$$\min_{1\le i\le m} \xi^T H_i \xi \ge \frac{\pi}{9m^2} \quad \text{and} \quad \|\xi\|^2 \le 3\,\mathrm{Tr}(Z^*).$$

Let ξ be any vector satisfying these two conditions.[2] Then, $z^*(\xi)$ is feasible for (1), so that

$$v_{\mathrm{qp}} \le \|z^*(\xi)\|^2 = \frac{\|\xi\|^2}{\min_i \xi^T H_i \xi} \le \frac{3\,\mathrm{Tr}(Z^*)}{(\pi/9m^2)} = \frac{27m^2}{\pi} v_{\mathrm{sdp}},$$

where the last equality uses $\mathrm{Tr}(Z^*) = v_{\mathrm{sdp}}$. ■

In the above proof, other choices of μ can also be used, but the resulting bound seems not as sharp. Theorem 1 suggests that the worst-case performance of the SDP relaxation deteriorates quadratically with the number of quadratic constraints. Below we give an example demonstrating that this bound is in fact tight up to a constant factor.

Example 1: For any $m \ge 2$ and $n \ge 2$, consider a special instance of (2), corresponding to (1) with $|\mathcal{I}_i| = 1$ (i.e., each H_i has rank 1), whereby

$$h_\ell = \left(\cos\left(\frac{\ell\pi}{m}\right), \sin\left(\frac{\ell\pi}{m}\right), 0, \ldots, 0\right)^T, \qquad \ell = 1, ..., m.$$

Let $z^* = (z_1^*, \ldots, z_n^*)^T \in \mathbb{R}^n$ be an optimal solution of (2) corresponding to the above choice of steering vectors h_ℓ. We can write

$$(z_1^*, z_2^*) = \rho(\cos\theta, \sin\theta), \quad \text{for some } \theta \in [0, 2\pi).$$

Since $\{\ell\pi/m,\ \ell = 1, ..., m\}$ is uniformly spaced on $[0, \pi)$, there must exist an integer ℓ such that

$$\text{either} \quad \left|\theta - \frac{\ell\pi}{m} - \frac{\pi}{2}\right| \le \frac{\pi}{2m} \quad \text{or} \quad \left|\theta - \frac{\ell\pi}{m} + \frac{\pi}{2}\right| \le \frac{\pi}{2m}.$$

For simplicity, we assume the first case. (The second case can be treated similarly.) Since the last $(n-2)$ entries of h_ℓ are zero, it is readily checked that

$$|h_\ell^T z^*| = \rho \left|\cos\left(\theta - \frac{\ell\pi}{m}\right)\right| = \rho \left|\sin\left(\theta - \frac{\ell\pi}{m} - \frac{\pi}{2}\right)\right| \le \rho \left|\sin\left(\frac{\pi}{2m}\right)\right| \le \frac{\rho\pi}{2m}.$$

[2] The probability that no such ξ is generated after N independent trials is at most $(1-0.0758..)^N$, which for $N = 100$ equals 0.000375.. Thus, such ξ requires relatively few trials to generate.

Since z^* satisfies the constraint $|h_\ell^T z^*| \geq 1$, it follows that

$$\|z^*\| \geq \rho \geq \frac{2m|h_\ell^T z^*|}{\pi} \geq \frac{2m}{\pi},$$

implying

$$\upsilon_{\rm qp} = \|z^*\|^2 \geq \frac{4m^2}{\pi^2}.$$

On the other hand, the positive semidefinite matrix

$$Z^* = \mathrm{Diag}\{1, 1, 0, \ldots, 0\}$$

is feasible for the SDP relaxation (7), and it has an objective value of $\mathrm{Tr}(Z^*) = 2$. Thus, for this instance, we have

$$\upsilon_{\rm qp} \geq \frac{2m^2}{\pi^2}\upsilon_{\rm sdp}.$$

The preceding example and Theorem 1 show that the SDP relaxation (7) can be weak if the number of quadratic constraints is large, especially when the steering vectors h_ℓ are in a certain sense "uniformly distributed" in space.

3.2 General steering vectors: the complex case

We consider the complex case of $\mathbb{F} = \mathbb{C}$. We will show that the performance ratio of the SDP relaxation (7) improves to $O(m)$ in the complex case (as opposed to $O(m^2)$ in the real case). Similar to the real case, upon obtaining an optimal solution Z^* of (7), we construct a feasible solution of (1) using the following randomization procedure:

1. Generate a random vector $\xi \in \mathbb{C}^n$ from the *complex-valued* normal distribution $N_c(0, Z^*)$ [2, 26].
2. Let $z^*(\xi) = \xi / \min\limits_{1\leq i\leq m} \sqrt{\xi^H H_i \xi}$.

Most of the ensuing performance analysis is similar to that of the real case. In particular, we will also need the following two lemmas analogous to Lemmas 1 and 2.

Lemma 3 *Let $H \in \mathbb{C}^{n\times n}$, $Z \in \mathbb{C}^{n\times n}$ be two Hermitian positive semidefinite matrices (i.e., $H \succeq 0$, $Z \succeq 0$). Suppose $\xi \in \mathbb{C}^n$ is a random vector generated from the complex-valued normal distribution $N_c(0, Z)$. Then, for any $\gamma > 0$,*

$$\mathrm{Prob}\left(\xi^H H\xi < \gamma E(\xi^H H\xi)\right) \leq \max\left\{\frac{4}{3}\gamma, 16(\bar{r}-1)^2\gamma^2\right\}, \qquad (12)$$

where $\bar{r} := \min\{\text{rank}\,(H), \text{rank}\,(Z)\}$.

Proof. We follow the same notations and proof as for Lemma 1, except for two blanket changes:

$$\begin{aligned} \text{matrix transpose} &\rightarrow \text{Hermitian transpose}, \\ \text{orthogonal matrix} &\rightarrow \text{unitary matrix}. \end{aligned}$$

Also, $\bar{\xi}$ has the complex-valued normal distribution $N_c(0, I_r)$. With these changes, we consider the same two cases: $\bar{\lambda}_1 \geq \alpha$ and $\bar{\lambda}_1 < \alpha$, where $0 < \alpha < 1$. In the first case, we have similar to (9) that

$$\text{Prob}\left(\xi^H H\xi < \gamma E(\xi^H H\xi)\right) \leq \text{Prob}\left(|\bar{\xi}_1|^2 < \gamma/\alpha\right). \tag{13}$$

Recall that the density function of a complex-valued circular normal random variable $u \sim N_c(0, \sigma^2)$, where σ is the standard deviation, is

$$\frac{1}{\pi\sigma^2} e^{-\frac{|u|^2}{\sigma^2}} \quad \forall\ u \in \mathbb{C}.$$

In polar coordinates, the density function can be written as

$$f(\rho, \theta) = \frac{\rho}{\pi\sigma^2} e^{-\frac{\rho^2}{\sigma^2}} \quad \forall\ \rho \in [0, +\infty),\ \theta \in [0, 2\pi).$$

In fact, a complex-valued normal distribution can be viewed as a joint distribution of its modulus and its argument, with the following particular properties: (1) the modulus and argument are independently distributed; (2) the argument is uniformly distributed over $[0, 2\pi)$; (3) the modulus follows a Weibull distribution with density

$$f(\rho) = \begin{cases} \frac{2\rho}{\sigma^2} e^{-\frac{\rho^2}{\sigma^2}}, & \text{if } \rho \geq 0; \\ 0, & \text{if } \rho < 0, \end{cases}$$

and distribution function

$$\text{Prob}\,\{|u| \leq t\} = 1 - e^{-\frac{t^2}{\sigma^2}}. \tag{14}$$

Since $\bar{\xi}_1 \sim N_c(0, 1)$, substituting this into (13) yields

$$\text{Prob}\left(\xi^H H\xi < \gamma E(\xi^H H\xi)\right) \leq \text{Prob}\left(|\bar{\xi}_1|^2 < \gamma/\alpha\right) \leq 1 - e^{-\gamma/\alpha} \leq \gamma/\alpha,$$

where the last inequality uses the convexity of the exponential function.

In the second case of $\bar{\lambda}_1 < \alpha$, we have similar to (10) that

$$\begin{aligned}
\mathrm{Prob}\left(\xi^H H\xi < \gamma E(\xi^H H\xi)\right) &\le \mathrm{Prob}\left(\bar{\lambda}_1|\bar{\xi}_1|^2 < \gamma\right)\cdot \mathrm{Prob}\left(\bar{\lambda}_2|\bar{\xi}_2|^2 < \gamma\right)\\
&= (1-e^{-\gamma/\bar{\lambda}_1})(1-e^{-\gamma/\bar{\lambda}_2})\\
&\le \frac{\gamma^2}{\bar{\lambda}_1\bar{\lambda}_2}\\
&\le \frac{(\bar{r}-1)^2\gamma^2}{(1-\alpha)^2},
\end{aligned}$$

where last step uses the fact that $\bar{\lambda}_1 \ge \bar{\lambda}_2 \ge (1-\alpha)/(\bar{r}-1)$. Combining the estimates for the above two cases and setting $\alpha = 3/4$, we immediately obtain the desired bound (12). ∎

Lemma 4 *Let* $\mathbb{F} = \mathbb{C}$. *Let* $Z^* \succeq 0$ *be a feasible solution of* (7) *and let* $z^*(\xi)$ *be generated by the randomization procedure described earlier. Then, with probability 1,* $z^*(\xi)$ *is well defined and feasible for* (1). *Moreover, for every* $\gamma > 0$ *and* $\mu > 0$,

$$\mathrm{Prob}\left(\min_{1\le i\le m}\xi^H H_i\xi \ge \gamma,\ \|\xi\|^2 \le \mu \mathrm{Tr}(Z^*)\right) \ge 1 - m\cdot\max\left\{\frac{4}{3}\gamma, 16(r-1)^2\gamma^2\right\} - \frac{1}{\mu},$$

where $r := \mathrm{rank}\,(Z^*)$.

Proof. The proof is mostly the same as that for the real case (see Lemma 2). In particular, for any $\gamma > 0$ and $\mu > 0$, we still have

$$\begin{aligned}
&\mathrm{Prob}\left(\min_{1\le i\le m}\xi^H H_i\xi \ge \gamma,\ \|\xi\|^2 \le \mu \mathrm{Tr}(Z^*)\right)\\
&\ge\ 1 - \sum_{i=1}^m \mathrm{Prob}\left(\xi^H H_i\xi < \gamma E(\xi^H H_i\xi)\right) - \mathrm{Prob}\left(\|\xi\|^2 > \mu E(\|\xi\|^2)\right).
\end{aligned}$$

Therefore, we can invoke Lemma 3 to obtain

$$\begin{aligned}
&\mathrm{Prob}\left(\min_{1\le i\le m}\xi^H H_i\xi \ge \gamma,\ \|\xi\|^2 \le \mu \mathrm{Tr}(Z^*)\right)\\
&\ge\ 1 - m\cdot\max\left\{\frac{4}{3}\gamma, 16(r-1)^2\gamma^2\right\} - \mathrm{Prob}\left(\|\xi\|^2 > \mu E(\|\xi\|^2)\right)\\
&\ge\ 1 - m\cdot\max\left\{\frac{4}{3}\gamma, 16(r-1)^2\gamma^2\right\} - \frac{1}{\mu},
\end{aligned}$$

which completes the proof. ∎

Theorem 2 *Let $\mathbb{F} = \mathbb{C}$. For the QP* (1) *and its SDP relaxation* (7)*, we have $v_{\rm sdp} = v_{\rm qp}$ if $m \le 3$ and otherwise*

$$v_{\rm qp} \le 8m \cdot v_{\rm sdp}.$$

Proof. By applying a suitable rank reduction procedure if necessary, we can assume that the rank r of the optimal SDP solution Z^* satisfies $r = 1$ if $m \le 3$ and $r \le \sqrt{m}$ if $m \ge 4$; see [9, Section 5]. Thus, if $m \le 3$, then $Z^* = z^*(z^*)^H$ for some $z^* \in \mathbb{C}^n$ and it is readily seen that z^* is an optimal solution of (1), so that $v_{\rm sdp} = v_{\rm qp}$. Otherwise, we apply the randomization procedure to Z^*. By choosing $\mu = 2$ and $\gamma = \frac{1}{4m}$, it is easily verified using $r \le \sqrt{m}$ that

$$\frac{4}{3}\gamma \ge 16(r-1)^2\gamma^2 \quad \forall\ m = 1, 2, ...$$

Therefore, it follows from Lemma 4 that

$$\text{Prob}\left\{\min_{1\le i\le m} \xi^H H_i \xi \ge \gamma,\ \|\xi\|^2 \le \mu \text{Tr}(Z^*)\right\} \ge 1 - m\frac{4}{3}\gamma - \frac{1}{\mu} = \frac{1}{6}.$$

Then, similar to the proof of Theorem 1, we obtain that with probability of at least 1/6, $z^*(\xi)$ is a feasible solution of (1) and $v_{\rm qp} \le \|z^*(\xi)\|^2 \le 8m \cdot v_{\rm sdp}$.[3] ■

The proof of Theorem 2 shows that, by repeating the randomization procedure, the probability of generating a feasible solution with a performance ratio no more than $8m$ approaches 1 exponentially fast (independent of problem size). Alternatively, a de-randomization technique from theoretical computer science can perhaps convert the above randomization procedure into a polynomial-time deterministic algorithm [12]; also see [14].

Theorem 2 shows that the worst-case performance of SDP relaxation deteriorates *linearly* with the number of quadratic constraints. This contrasts with the *quadratic* rate of deterioration in the real case (see Theorem 1). Thus, the SDP relaxation can yield better performance in the complex case. This is in the same spirit as the recent results in [26] which showed that the quality of SDP relaxation improves by a constant factor for certain quadratic *maximization* problems when the space is changed from $\mathbb{R}^n$ to $\mathbb{C}^n$. Below we give an example demonstrating that this approximation bound is tight up to a constant factor.

Example 2: For any $m \ge 2$ and $n \ge 2$, let $K = \lceil\sqrt{m}\rceil$ (so $K \ge 2$). Consider a special instance of (2), corresponding to (1) with $|\mathcal{I}_i| = 1$ (i.e., each H_i has rank 1), whereby

$$h_\ell = \left(\cos\frac{j\pi}{K}, \sin\frac{j\pi}{K}e^{\frac{\mathrm{i}2k\pi}{K}}, 0, \ldots, 0\right)^T \quad \text{with} \quad \ell = jK - K + k, \qquad j, k = 1, ..., K.$$

[3] The probability that no such ξ is generated after N independent trials is at most $(5/6)^N$, which for $N = 30$ equals 0.00421.. Thus, such ξ requires relatively few trials to generate.

Hence there are K^2 complex rank-1 constraints. Let $z^* = (z_1^*, \ldots, z_n^*)^T \in \mathbb{C}^n$ be an optimal solution of (2) corresponding to the above choice of $\lceil\sqrt{m}\rceil^2$ steering vectors h_ℓ. By a phase rotation if necessary, we can without loss of generality assume that z_1^* is real and write

$$(z_1^*, z_2^*) = \rho(\cos\theta, \sin\theta e^{\mathbf{i}\psi}), \quad \text{for some } \theta, \psi \in [0, 2\pi).$$

Since $\{2k\pi/K,\ k = 1, ..., K\}$ and $\{j\pi/K,\ j = 1, ..., K\}$ are uniformly spaced in $[0, 2\pi)$ and $[0, \pi)$ respectively, there must exist integers j and k such that

$$\left|\psi - \frac{2k\pi}{K}\right| \le \frac{\pi}{K} \quad \text{and} \quad \text{either } \left|\theta - \frac{j\pi}{K} - \frac{\pi}{2}\right| \le \frac{\pi}{2K} \text{ or } \left|\theta - \frac{j\pi}{K} + \frac{\pi}{2}\right| \le \frac{\pi}{2K}.$$

Without loss of generality, we assume

$$\left|\theta - \frac{j\pi}{K} - \frac{\pi}{2}\right| \le \frac{\pi}{2K}.$$

Since the last $(n-2)$ entries of each h_ℓ are zero, it is readily seen that, for $\ell = jK - K + k$,

$$\begin{aligned}
\left|\mathrm{Re}(h_\ell^H z^*)\right| &= \rho\left|\cos\theta\cos\frac{j\pi}{K} + \sin\theta\sin\frac{j\pi}{K}\cos\left(\psi - \frac{2k\pi}{K}\right)\right| \\
&= \rho\left|\cos\left(\theta - \frac{j\pi}{K}\right) + \sin\theta\sin\frac{j\pi}{K}\left(\cos\left(\psi - \frac{2k\pi}{K}\right) - 1\right)\right| \\
&= \rho\left|\sin\left(\theta - \frac{j\pi}{K} - \frac{\pi}{2}\right) - 2\sin\theta\sin\frac{j\pi}{K}\sin^2\left(\frac{K\psi - 2k\pi}{2K}\right)\right| \\
&\le \rho\left|\sin\frac{\pi}{2K}\right| + 2\rho\sin^2\frac{\pi}{2K} \\
&\le \frac{\rho\pi}{2K} + \frac{\rho\pi^2}{2K^2}.
\end{aligned}$$

In addition, we have

$$\begin{aligned}
\left|\mathrm{Im}(h_\ell^H z^*)\right| &= \rho\left|\sin\theta\sin\frac{j\pi}{K}\sin\left(\psi - \frac{2k\pi}{K}\right)\right| \\
&\le \rho\left|\sin\left(\psi - \frac{2k\pi}{K}\right)\right| \\
&\le \rho\left|\psi - \frac{2k\pi}{K}\right| \le \frac{\rho\pi}{K}.
\end{aligned}$$

Combining the above two bounds, we obtain

$$\left|h_\ell^H z^*\right| \le \left|\mathrm{Re}(h_\ell^H z^*)\right| + \left|\mathrm{Im}(h_\ell^H z^*)\right| \le \frac{3\rho\pi}{2K} + \frac{\rho\pi^2}{2K^2}.$$

Since z^* satisfies the constraint $|h_\ell^H z^*| \ge 1$, it follows that

$$\|z^*\| \ge \rho \ge \frac{2K^2|h_\ell^H z^*|}{\pi(3K + \pi)} \ge \frac{2K^2}{\pi(3K + \pi)},$$

implying

$$v_{\rm qp} = \|z^*\|^2 \geq \frac{4K^4}{\pi^2(3K+\pi)^2} = \frac{4\lceil\sqrt{m}\rceil^4}{\pi^2(3\lceil\sqrt{m}\rceil+\pi)^2}.$$

On the other hand, the positive semidefinite matrix

$$Z^* = \mathrm{Diag}\{1, 1, 0, \ldots, 0\}$$

is feasible for the SDP relaxation (7), and it has an objective value of $\mathrm{Tr}(Z^*) = 2$. Thus, for this instance, we have

$$v_{\rm qp} \geq \frac{2\lceil\sqrt{m}\rceil^4}{\pi^2(3\lceil\sqrt{m}\rceil+\pi)^2}\, v_{\rm sdp} \geq \frac{2m}{\pi^2(3+\pi/2)^2}\, v_{\rm sdp}.$$

The preceding example and Theorem 2 show that the SDP relaxation (7) can be weak if the number of quadratic constraints is large, especially when the steering vectors h_ℓ are in a certain sense "uniformly distributed" in space. In the next subsection, we will tighten the approximation bound in Theorem 2 by considering special cases where the steering vectors are "not too spread out in space".

3.3 Specially configured steering vectors: the complex case

We consider the complex case of $\mathbb{F} = \mathbb{C}$. Let Z^* be any optimal solution of (7). Since Z^* is feasible for (7), $Z^* \neq 0$. Then

$$Z^* = \sum_{k=1}^{r} w_k w_k^H, \tag{15}$$

for some nonzero $w_k \in \mathbb{C}^{\rm n}$, where $r := \mathrm{rank}\,(Z^*) \geq 1$. By decomposing $w_k = u_k + v_k$, with $u_k \in \mathrm{span}\{h_1, ..., h_M\}$ and $v_k \in \mathrm{span}\{h_1, ..., h_M\}^\perp$, it is easily checked that $\tilde{Z} := \sum_{k=1}^r u_k u_k^H$ is feasible for (7) and

$$\langle I, Z^*\rangle = \sum_{k=1}^{r} \|u_k + v_k\|^2 = \sum_{k=1}^{r}(\|u_k\|^2 + \|v_k\|^2) = \langle I, \tilde{Z}\rangle + \sum_{k=1}^{r} \|v_k\|^2.$$

This implies $v_k = 0$ for all k, so that

$$w_k \in \mathrm{span}\{h_1, ..., h_M\}. \tag{16}$$

Below we show that the SDP relaxation (7) provides a *constant factor approximation* to the QP (1) when the phase spread of the entries of h_ℓ is bounded away from $\pi/2$.

Theorem 3 *Suppose that*

$$h_\ell = \sum_{i=1}^{p} \beta_{i\ell} g_i \quad \forall\, \ell = 1, ..., M, \tag{17}$$

for some $p \geq 1$, $\beta_{i\ell} \in \mathbb{C}$ *and* $g_i \in \mathbb{C}^{\mathrm{n}}$ *such that* $\|g_i\| = 1$ *and* $g_i^H g_j = 0$ *for all* $i \neq j$. *Then the following results hold.*

(a) *If* $\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}) > 0$ *whenever* $\beta_{i\ell}^H \beta_{j\ell} \neq 0$, *then* $v_{\mathrm{qp}} \leq C v_{\mathrm{sdp}}$, *where*

$$C := \max_{i,j,\ell \ \mid\ \beta_{i\ell}^H \beta_{j\ell} \neq 0} \left(1 + \frac{|\mathrm{Im}(\beta_{i\ell}^H \beta_{j\ell})|^2}{|\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell})|^2} \right)^{1/2}. \tag{18}$$

(b) *If* $\beta_{i\ell} = |\beta_{i\ell}| e^{\mathbf{i}\phi_{i\ell}}$, *where*

$$\phi_{i\ell} \in [\bar{\phi}_\ell - \phi, \bar{\phi}_\ell + \phi] \quad \forall\, i, \ell, \quad \textit{for some } 0 \leq \phi < \frac{\pi}{4} \textit{ and some } \bar{\phi}_\ell \in \mathbb{R}, \tag{19}$$

then $\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}) > 0$ *whenever* $\beta_{i\ell}^H \beta_{j\ell} \neq 0$, *and* C *given by* (18) *satisfies*

$$C \leq \frac{1}{\cos(2\phi)}. \tag{20}$$

Proof. (a) By (16), we have

$$w_k = \sum_{i=1}^{p} \alpha_{ki} g_i,$$

for some $\alpha_{ki} \in \mathbb{C}$. This together with (15) yields

$$\begin{aligned} \langle I, Z^* \rangle &= \sum_{k=1}^{r} \|w_k\|^2 = \sum_{k=1}^{r} \left\| \sum_{i=1}^{p} \alpha_{ki} g_i \right\|^2 \\ &= \sum_{k=1}^{r} \sum_{i=1}^{p} |\alpha_{ki}|^2 = \sum_{i=1}^{p} \lambda_i^2, \end{aligned}$$

where the third equality uses the orthonormal properties of $g_1, ..., g_p$, and the last equality uses $\lambda_i := \left(\sum_{k=1}^{r} |\alpha_{ki}|^2\right)^{1/2} = \|(\alpha_{ki})_{k=1}^{r}\|$.

Let

$$z^* := \sum_{i=1}^{p} \lambda_i g_i.$$

Then, the orthonormal properties of $g_1, ..., g_p$ yields

$$\|z^*\|^2 = \left\|\sum_{i=1}^{p} \lambda_i g_i\right\|^2 = \sum_{i=1}^{p} \lambda_i^2 = \langle I, Z^* \rangle = v_{\text{sdp}}. \tag{21}$$

Moreover, for each $\ell \in \{1, ..., M\}$, we obtain from (15) that

$$\begin{aligned}
\langle h_\ell h_\ell^H, Z^* \rangle &= \sum_{k=1}^{r} \langle h_\ell h_\ell^H, w_k w_k^H \rangle = \sum_{k=1}^{r} |h_\ell^H w_k|^2 \\
&= \sum_{k=1}^{r} \left| \sum_{i=1}^{p} \alpha_{ki} h_\ell^H g_i \right|^2 = \sum_{k=1}^{r} \left| \sum_{i=1}^{p} \alpha_{ki} \beta_{i\ell} \right|^2 \\
&= \mathrm{Re}\left(\sum_{k=1}^{r} \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_{ki}^H \alpha_{kj} \beta_{i\ell}^H \beta_{j\ell} \right) = \mathrm{Re}\left(\sum_{i=1}^{p} \sum_{j=1}^{p} \beta_{i\ell}^H \beta_{j\ell} \sum_{k=1}^{r} \alpha_{ki}^H \alpha_{kj} \right) \\
&= \sum_{i=1}^{p} \sum_{j=1}^{p} \mathrm{Re}\left(\beta_{i\ell}^H \beta_{j\ell} \sum_{k=1}^{r} \alpha_{ki}^H \alpha_{kj} \right) \\
&\le \sum_{i=1}^{p} \sum_{j=1}^{p} \left| \beta_{i\ell}^H \beta_{j\ell} \right| \left| \sum_{k=1}^{r} \alpha_{ki}^H \alpha_{kj} \right| \le \sum_{i=1}^{p} \sum_{j=1}^{p} \left| \beta_{i\ell}^H \beta_{j\ell} \right| \|(\alpha_{ki})_{k=1}^r\| \|(\alpha_{kj})_{k=1}^r\| \\
&= \sum_{i=1}^{p} \sum_{j=1}^{p} \left| \beta_{i\ell}^H \beta_{j\ell} \right| \lambda_i \lambda_j,
\end{aligned}$$

where the fourth equality uses (17) and the orthonormal properties of $g_1, ..., g_p$; the last inequality is due to the Cauchy-Schwarz inequality. Then, it follows that

$$\begin{aligned}
\langle h_\ell h_\ell^H, Z^* \rangle &\le \sum_{i=1}^{p} \sum_{j=1}^{p} \left(|\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell})|^2 + |\mathrm{Im}(\beta_{i\ell}^H \beta_{j\ell})|^2 \right)^{1/2} \lambda_i \lambda_j \\
&= \sum_{i=1}^{p} \sum_{j=1}^{p} \left| \mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}) \right| \left(1 + \frac{|\mathrm{Im}(\beta_{i\ell}^H \beta_{j\ell})|^2}{|\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell})|^2} \right)^{1/2} \lambda_i \lambda_j \\
&\le \sum_{i=1}^{p} \sum_{j=1}^{p} \left| \mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}) \right| C \lambda_i \lambda_j \\
&= \sum_{i=1}^{p} \sum_{j=1}^{p} \mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}) C \lambda_i \lambda_j,
\end{aligned}$$

where the summation in the second step is taken over i, j with $\beta_{i\ell}^H \beta_{j\ell} \neq 0$, the third step is due to (18), and the last step is due to the assumption that $\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}) > 0$ whenever

$\beta_{i\ell}^H \beta_{j\ell} \neq 0$. Also, we have from (17) and the orthonormal properties of $g_1, ..., g_p$ that

$$|h_\ell^H z^*|^2 = \left\| \sum_{i=1}^p \lambda_i h_\ell^H g_i \right\|^2 = \left\| \sum_{i=1}^p \lambda_i \beta_{i\ell} \right\|^2 = \sum_{i=1}^p \sum_{j=1}^p \lambda_i \lambda_j \mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}).$$

Comparing the above two displayed equations, we see that

$$\langle h_\ell h_\ell^H, Z^* \rangle \leq C |h_\ell^H z^*|^2, \quad \ell = 1, ..., M.$$

Since Z^* is feasible for (7), this shows that $\sqrt{C} z^*$ is feasible for (1), which further implies

$$v_{\mathrm{qp}} \leq \left\| \sqrt{C} z^* \right\|^2 = C \|z^*\|^2 = C v_{\mathrm{sdp}}.$$

This proves the desired result.

(b) The condition (19) implies that $|\phi_{i\ell} - \phi_{j\ell}| \leq 2\phi < \pi/2$. In other words, the phase angle spread of the entries of each $\beta_\ell = (\beta_{1\ell}, \beta_{2\ell}, \ldots, \beta_{n\ell})^T$ is no more than 2ϕ. This further implies that

$$\cos(\phi_{i\ell} - \phi_{j\ell}) \geq \cos(2\phi) \quad \forall\, i, j, \ell. \tag{22}$$

We have

$$\begin{aligned} \beta_{i\ell}^H \beta_{j\ell} &= |\beta_{i\ell}| e^{-\mathbf{i}\phi_{i\ell}} |\beta_{j\ell}| e^{\mathbf{i}\phi_{j\ell}} \\ &= |\beta_{i\ell}||\beta_{j\ell}| e^{\mathbf{i}(\phi_{j\ell} - \phi_{i\ell})} \\ &= |\beta_{i\ell}||\beta_{j\ell}| (\cos(\phi_{j\ell} - \phi_{i\ell}) + \mathbf{i} \sin(\phi_{j\ell} - \phi_{i\ell})). \end{aligned}$$

Since $|\phi_{i\ell} - \phi_{j\ell}| < \pi/2$ so that $\cos(\phi_{j\ell} - \phi_{i\ell}) > 0$, we see that $\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell}) > 0$ whenever $\beta_{i\ell}^H \beta_{j\ell} \neq 0$. Then

$$\left(1 + \frac{|\mathrm{Im}(\beta_{i\ell}^H \beta_{j\ell})|^2}{|\mathrm{Re}(\beta_{i\ell}^H \beta_{j\ell})|^2}\right)^{1/2} \leq \left(1 + \tan^2(\phi_{j\ell} - \phi_{i\ell})\right)^{1/2} = \frac{1}{\cos(\phi_{j\ell} - \phi_{i\ell})} \leq \frac{1}{\cos(2\phi)},$$

where the last step uses (22). Using this in (18) completes the proof. ■

In Theorem 3(b), we can more generally consider $\beta_{i\ell}$ of the form $\beta_{i\ell} = \omega_{i\ell} e^{\mathbf{i}\phi_{i\ell}} (1 + \mathbf{i}\theta_{i\ell})$, where $\omega_{i\ell} \geq 0$, $\alpha_{i\ell}$ satisfies (19), and

$$|\theta_{j\ell} - \theta_{i\ell}| \leq \sigma |1 + \theta_{i\ell}\theta_{j\ell}| \quad \forall\, i, j, \ell, \quad \text{for some } \sigma \geq 0 \text{ with } \tan(2\phi)\sigma < 1. \tag{23}$$

Then the proof of Theorem 3(b) can be extended to show the following upper bound on C given by (18):

$$C \leq \frac{1}{\cos(2\phi)} \cdot \frac{\sqrt{1 + \sigma^2}}{1 - \tan(2\phi)\sigma}. \tag{24}$$

However, this generalization is superficial as we can also derive (24) from (20) by rewriting $\beta_{i\ell}$ as

$$\beta_{i\ell} = |\beta_{i\ell}| e^{\mathbf{i}\tilde{\phi}_{i\ell}} \quad \text{with} \quad \tilde{\phi}_{i\ell} = \phi_{i\ell} + \tan^{-1}(\theta_{i\ell}).$$

Then, applying (20) yields $C \geq \cos(2\tilde{\phi})$, where $\tilde{\phi} = \max_{i,j,\ell} |\tilde{\phi}_{i\ell} - \tilde{\phi}_{j\ell}|/2$. Using trigonometric identity, it can be shown that $\cos(2\tilde{\phi})$ equals the right-hand side of (24) with $\sigma = \max\limits_{i,j,\ell \;|\; \theta_{i\ell}\theta_{j\ell} \neq -1} |\theta_{j\ell} - \theta_{i\ell}|/|1 + \theta_{i\ell}\theta_{j\ell}|$.

Notice that Theorem 3(b) implies that if $\phi = 0$, then the SDP relaxation (7) is tight for the quadratically constrained QP (1) with $\mathbb{F} = \mathbb{C}$. Such is the case when all components of h_ℓ, $\ell = 1, ..., M$, are real and nonnegative.

4 A convex QP restriction

In this subsection, we consider a convex quadratic programming *restriction* of (2) in the complex case of $\mathbb{F} = \mathbb{C}$ and analyze its approximation bound. Let us write h_ℓ (the channel steering vector) as

$$h_\ell = (\ldots, |h_{j\ell}| e^{\mathbf{i}\phi_{j\ell}}, \ldots)^T_{j=1,\ldots,n}.$$

For any $\bar{\phi}_j \in [0, 2\pi)$, $j = 1, ..., n$, and any $\phi \in (0, \pi/2)$, define the four corresponding index subsets:

$$\begin{aligned}
J_\ell^1 &:= \{j \mid \phi_{j\ell} \in [\bar{\phi}_j - \phi, \bar{\phi}_j + \phi]\}, \\
J_\ell^2 &:= \{j \mid \phi_{j\ell} \in [\bar{\phi}_j - \phi + \pi/2, \bar{\phi}_j + \phi + \pi/2]\}, \\
J_\ell^3 &:= \{j \mid \phi_{j\ell} \in [\bar{\phi}_j - \phi + \pi, \bar{\phi}_j + \phi + \pi]\}, \\
J_\ell^4 &:= \{j \mid \phi_{j\ell} \in [\bar{\phi}_j - \phi + 3\pi/2, \bar{\phi}_j + \phi + 3\pi/2]\},
\end{aligned}$$

for $\ell = 1, ..., M$. The above four subsets are pairwise disjoint if and only if $\phi < \pi/4$, and are collectively exhaustive if and only if $\phi \geq \pi/4$. Choose an index subset J with the property that

$$\text{for each } \ell, \text{ at least one of } J_\ell^1,\ J_\ell^2,\ J_\ell^3,\ J_\ell^4 \text{ contains } J.$$

Of course, $J = \emptyset$ is always allowable, but we should choose J maximally since our approximation bound will depend on the ratio $n/|J|$ (see Theorem 4 below). Partition the constraint set index $\{1, ..., M\}$ into four subsets K^1, K^2, K^3, K^4 such that

$$J \subseteq J_\ell^k \quad \forall\, \ell \in K^k,\ k = 1, 2, 3, 4.$$

Consider the following convex QP restriction of (2) corresponding to K^1, K^2, K^3, K^4:

$$\begin{array}{rrrll} \upsilon_{\rm cqp} := & \min & \|z\|^2 & & \\ & \text{s.t.} & \mathrm{Re}(h_\ell^H z) & \geq 1 & \forall\, \ell \in K^1, \\ & & -\mathrm{Im}(h_\ell^H z) & \geq 1 & \forall\, \ell \in K^2, \\ & & -\mathrm{Re}(h_\ell^H z) & \geq 1 & \forall\, \ell \in K^3, \\ & & \mathrm{Im}(h_\ell^H z) & \geq 1 & \forall\, \ell \in K^4. \end{array} \tag{25}$$

The above problem is a restriction of (2) because, for any $z \in \mathbb{C}$,

$$\begin{array}{rcl} |z| & \geq & \max\{|\mathrm{Re}(z)|, |\mathrm{Im}(z)|\} \\ & = & \max\{\mathrm{Re}(z), \mathrm{Im}(z), -\mathrm{Re}(z), -\mathrm{Im}(z)\}. \end{array}$$

If $J \neq \emptyset$ and $(\ldots, h_{j\ell}, \ldots)_{j\in J} \neq 0$ for $\ell = 1, ..., M$, then (25) is feasible, and hence has an optimal solution. Since (25) is a restriction of (2), $\upsilon_{\rm qp} \leq \upsilon_{\rm cqp}$. We have the following approximation bound.

Theorem 4 *Suppose that $J \neq \emptyset$ and (25) is feasible. Then,*

$$\upsilon_{\rm cqp} \leq \upsilon_{\rm qp} \frac{N}{\cos^2\phi} \max_{k=1,\ldots,N} \left(\max_{j\in \hat{J}_k} \frac{\bar{\eta}_j}{\underline{\eta}_{\pi_k(j)}} \right)^2,$$

where $N := \lceil n/|J| \rceil$, $\bar{\eta}_j := \max_\ell |h_{j\ell}|$, $\underline{\eta}_j := \min_{\ell | h_{j\ell}\neq 0} |h_{j\ell}|$, $\hat{J}_1, ..., \hat{J}_N$ is any partition of $\{1, ..., n\}$ satisfying $|\hat{J}_k| \leq |J|$ for $k = 1, ..., N$, and π_k is any injective mapping from $\hat{J}_k$ to J.

Proof. By making the substitution

$$z_j^{\rm new} \leftarrow z_j e^{\mathbf{i}\bar{\phi}_j},$$

we can without loss of generality assume that $\bar{\phi}_j = 0$ for all j and ℓ.

Let z^* denote an optimal solution of (2) and write

$$z^* = (\ldots, r_j e^{\mathbf{i}\beta_j}, \ldots)^T_{j=1,\ldots,n},$$

with $r_j \geq 0$. Then, for any ℓ, we have from $|h_{j\ell}| \leq \bar{\eta}_j$ for all j that

$$1 \leq |h_\ell^H z^*| \leq r := \sum_{j=1}^n r_j \bar{\eta}_j.$$

Also, we have

$$\upsilon_{\rm qp} = \|z^*\|^2 = \sum_{j=1}^{n} r_j^2.$$

Define

$$R_k := \left(\sum_{j \in \hat{J}_k} r_j^2 \right)^{1/2}, \qquad S_k := \sum_{j \in \hat{J}_k} r_j \bar{\eta}_j.$$

Then

$$1 \le r = \sum_{k=1}^{N} S_k, \qquad \upsilon_{\rm qp} = \sum_{k=1}^{N} R_k^2.$$

Without loss of generality, assume that $R_1/S_1 = \min_k R_k/S_k$. Then, using the fact that

$$\min_k \frac{|x_k|}{|y_k|} \le \sqrt{N} \frac{\|x\|_2}{\|y\|_1}$$

for any $x, y \in \mathbb{R}^{\rm N}$ with $y \neq 0$,[4] we see from the above relations that

$$\begin{aligned} \frac{R_1}{S_1} &\le \frac{R_1}{S_1} r \\ &\le \sqrt{N} \frac{\sqrt{\upsilon_{\rm qp}}}{r} r \\ &= \sqrt{N} \sqrt{\upsilon_{\rm qp}}. \end{aligned}$$

Since $|\hat{J}_1| \le |J|$, there is an injective mapping π from $\hat{J}_1$ to J. Let $\omega := \min_{j \in \hat{J}_1} \underline{\eta}_{\pi(j)}/\bar{\eta}_j$. Define the vector $\bar{z} \in \mathbb{C}^{\rm n}$ by

$$\bar{z}_j := \begin{cases} r_{\pi^{-1}(j)}/(S_1 \omega \cos\phi) & \text{if } j \in \pi(\hat{J}_1); \\ 0 & \text{else.} \end{cases}$$

Then,

$$\|\bar{z}\|^2 = \frac{R_1^2}{S_1^2 \omega^2 \cos^2\phi} \le \frac{N \upsilon_{\rm qp}}{\omega^2 \cos^2\phi}.$$

Moreover, for each $\ell \in K^1$, since $\pi(\hat{J}_1) \subseteq J \subseteq J_\ell^1$, we have

$$\mathrm{Re}\left(h_\ell^H \bar{z} \right) = \mathrm{Re}\left(\sum_{j \in \pi(\hat{J}_1)} h_{j\ell}^H \bar{z}_j \right)$$

[4] Proof: Suppose the contrary, so that for some $x, y \in \mathbb{R}^{\rm N}$ with $y \neq 0$, we have $|x_k|/|y_k| > \sqrt{N}\|x\|_2/\|y\|_1$ for all k. Then, multiplying both sides by $|y_k|$ and summing over k yields $\|x\|_1 > \sqrt{N}\|x\|_2$, contradicting properties of 1- and 2-norms.

$$
\begin{aligned}
&= \frac{1}{S_1\omega\cos\phi}\mathrm{Re}\left(\sum_{j\in\pi(\hat{J}_1)} r_{\pi^{-1}(j)}|h_{j\ell}|e^{-\mathbf{i}\phi_{j\ell}}\right)\\
&= \frac{1}{S_1\omega\cos\phi}\sum_{j\in\pi(\hat{J}_1)} r_{\pi^{-1}(j)}|h_{j\ell}|\cos\phi_{j\ell}\\
&\geq \frac{1}{S_1\omega\cos\phi}\sum_{j\in\pi(\hat{J}_1)} r_{\pi^{-1}(j)}\underline{\eta}_j\cos\phi\\
&= \frac{1}{S_1\omega}\sum_{j\in\hat{J}_1} r_j\bar{\eta}_j\,\frac{\underline{\eta}_{\pi(j)}}{\bar{\eta}_j}\\
&\geq \frac{1}{S_1\omega}\sum_{j\in\hat{J}_1} r_j\bar{\eta}_j\cdot\min_{j\in\hat{J}_1}\frac{\underline{\eta}_{\pi(j)}}{\bar{\eta}_j}\\
&= 1,
\end{aligned}
$$

where the first inequality uses $|h_{j\ell}| \geq \underline{\eta}_j$ and $\phi_{j\ell} \in [-\phi, \phi]$ for $j \in J_\ell^1$. Since $\bar{z}_j = 0$ for $j \notin J_\ell^1$, this shows that $\bar{z}$ satisfies the first set of constraints in (25). A similar reasoning shows that $\bar{z}$ satisfies the remaining three sets of constraints in (25). ■

Notice that the $\bar{z}$ constructed in the proof of Theorem 4 is feasible for the further restriction of (25) whereby $z_j = 0$ for all $j \notin J$. This further restricted problem has the same (worst-case) approximation bound specified in Theorem 4.

Let us compare the two approximation bounds in Theorem 3 and Theorem 4. First, the required assumptions are different. On the one hand, the bound in Theorem 3 does not depend on $|h_{j\ell}|$, while the bound in Theorem 4 does. On the other hand, Theorem 3 requires that the bounded angular spread

$$|\phi_{j\ell} - \phi_{i\ell}| \leq 2\phi \quad \forall\, j, \ell, \tag{26}$$

for some $\phi < \pi/4$, while Theorem 4 allows $\phi < \pi/2$ and only requires the condition (26) for all $1 \leq \ell \leq M$ and $j \in J$, where J is a pre-selected index set. Thus, the bounded angular spread condition required in Theorem 3 corresponds exactly to $|J| = n$. Thus, the assumptions required in the two theorems do not imply one another. Second, the two performance ratios are also different. Naturally, the final performance ratio in Theorem 4 depends on the choice of J through the ratio $|J|/n$, so a large J is preferred. In the event that the assumptions of both theorems are satisfied and let us assume for simplicity that $\bar{\eta}_j = \underline{\eta}_j$ for all j, then $|J| = n$ and $\phi < \pi/4$, in which case Theorem 4 gives a performance ratio of $1/\cos^2\phi$ while Theorem 3 gives $1/\cos(2\phi)$. Since $\cos(2\phi) = \cos^2\phi - \sin^2\phi \leq \cos^2\phi$, we have $1/\cos(2\phi) \geq 1/\cos^2\phi$, showing that Theorem 4 gives a tighter approximation bound. However, this does not mean Theorem 4 is stronger than Theorem 3 since the two theorems hold under different assumptions in general.

We can specialize Theorem 4 to a typical situation in transmit beamforming. Consider a uniform linear transmit antenna array consisting of n elements, and let us assume that the M receivers are in a sector area from the far field, and the propagation is line-of-sight. By reciprocity, each steering vector h_ℓ will be Vandermonde with generator $e^{-\mathbf{i}2\pi\frac{d}{\lambda}\sin\theta_\ell}$ (see, e.g., [10]), where d is the inter-antenna spacing, λ is the wavelength, and θ_ℓ is the angle of arrival of the ℓth receiving antenna. In a sector of approximately 60 degrees about the array broadside, we will have $|\theta_\ell| \le \pi/3$. Suppose that $d/\lambda = 1/2$. Then the steering vector corresponding to the ℓth receiving antenna will have the form

$$h_\ell = (\ldots, e^{-\mathbf{i}(j-1)\pi\sin\theta_\ell}, \ldots)^T_{j=1,\ldots,n}.$$

In this case, we have that $\phi_{j\ell} = (j-1)\pi\sin\theta_\ell$ and $|h_{j\ell}| = 1$ for all j and ℓ. We can take, e.g.,

$$\bar{\phi}_j = 0, \quad \phi = \bar{j}\pi\max_\ell|\sin\theta_\ell|, \quad J = \{1, ..., \bar{j}+1\},$$

where $\bar{j} := \lfloor 1/\max_\ell|\sin\theta_\ell|\rfloor$. Thus, the assumptions of Theorem 4 are satisfied. Moreover, since $|\theta_\ell| \le \pi/3$ for all ℓ, it follows that $|J| = \bar{j}+1 \ge 2$. If n is not large, say, $n \le 8$, then Theorem 4 gives a performance ratio of $n/(|J|\cos^2\phi) \le 16$.

More generally, if we can choose the partition $\hat{J}_1, ..., \hat{J}_N$ and the mapping π_k in Theorem 4 such that

$$(\ldots, \bar{\eta}_j, \ldots)_{j\in\hat{J}_k} = (\ldots, \underline{\eta}_{\pi_k(j)}, \ldots)_{j\in J} \quad \forall\, k,$$

then the performance ratio in Theorem 4 simplifies to $N/\cos^2\phi$. In particular, this holds when $|h_{j\ell}| = \eta > 0$ for all j and ℓ or when $J = \{1, ..., n\}$ (so that $N = 1$) and $|h_{j\ell}|$ is independent of ℓ for all j, and more generally, when the channel coefficients periodically repeat their magnitudes. In general, we should choose the partition $\hat{J}_1, ..., \hat{J}_N$ and the mapping π_k to make the performance ratio in Theorem 4 small. For example, if $J = \hat{J}_1 = \{1, 2\}$ and $\bar{\eta}_1 = 100, \bar{\eta}_2 = 10$, $\underline{\eta}_1 = 1, \underline{\eta}_2 = 10$, then $\pi_1(1) = 2, \pi_1(2) = 1$ is the better choice.

5 Homogeneous QP in Maximization Form

Let us now consider the following complex norm maximization problem with convex homogeneous quadratic constraints:

$$\begin{array}{rll} v_{\mathrm{qp}} := \max & \|z\|^2 & \\ \text{s.t.} & \displaystyle\sum_{\ell \in \mathcal{I}_i} |h_\ell^H z|^2 \leq 1, \quad i = 1, ..., m, \\ & z \in \mathbb{C}^{\mathrm{n}}, \end{array} \tag{27}$$

where $h_\ell \in \mathbb{C}^{\mathrm{n}}$.

To motivate this problem, consider the problem of designing an intercept beamformer[5] capable of suppressing signals impinging on the receiving antenna array from irrelevant or hostile emitters, e.g., jammers, whose steering vectors (spatial signatures, or "footprints") have been previously estimated, while achieving as high gain as possible for all other transmissions. The jammer suppression capability is captured in the constraints of (27), and $|\mathcal{I}_i| > 1$ covers the case where a jammer employs more than one transmit antennas. The maximization of the objective $\|z\|^2$ can be motivated as follows. In intercept applications, the steering vector of the emitter of interest, h, is *a priori* unknown, and is naturally modelled as random. A pertinent optimization objective is then the average beamformer output power, measured by $E[|h^H z|^2]$. Under the assumption that the entries of h are uncorrelated and have equal average power, it follows that $E[|h^H z|^2]$ is proportional to $\|z\|^2$, which is often referred to as the beamformer's *white noise gain*.

Similar to (1), we let

$$H_i := \sum_{\ell \in \mathcal{I}_i}^{m} h_\ell h_\ell^H$$

and consider the natural SDP relaxation of (27):

$$\begin{array}{rrl} v_{\mathrm{sdp}} := & \max & \mathrm{Tr}(Z) \\ & \text{s.t.} & \mathrm{Tr}(H_i Z) \leq 1, \quad i = 1, ..., m, \\ & & Z \succeq 0, \quad Z \text{ is complex and Hermitian.} \end{array} \tag{28}$$

We are interested in lower bounds for the relaxation performance of the form

$$v_{\mathrm{qp}} \geq C \, v_{\mathrm{sdp}},$$

where $0 < C \leq 1$. It is easily checked that (28) has an optimal solution.

[5]Note that here we are talking about a receive beamformer, as opposed to our earlier motivating discussion of transmit beamformer design.

Let Z^* be an optimal solution of (28). We will analyze the performance of the SDP relaxation using the following randomization procedure:

1. Generate a random vector $\xi \in \mathbb{C}^n$ from the *complex-valued* normal distribution $N_c(0, Z^*)$.
2. Let $z^*(\xi) = \xi / \max_{1\le i\le m} \sqrt{\xi^H H_i \xi}$.

First, we need the following lemma analogous to Lemmas 1 and 3.

Lemma 5 *Let $H \in \mathbb{C}^{n\times n}$, $Z \in \mathbb{C}^{n\times n}$ be two Hermitian positive semidefinite matrices (i.e., $H \succeq 0$, $Z \succeq 0$). Suppose $\xi \in \mathbb{C}^n$ is a random vector generated from the complex-valued normal distribution $N_c(0, Z)$. Then, for any $\gamma > 0$,*

$$\text{Prob}\left(\xi^H H\xi > \gamma E(\xi^H H\xi)\right) \le \bar{r}\ e^{-\gamma}, \tag{29}$$

where $\bar{r} := \min\{\text{rank}\,(H), \text{rank}\,(Z)\}$.

Proof. If $H = 0$, then (29) is trivially true. Suppose $H \neq 0$. Then, as in the proof of Lemma 1, we have

$$\text{Prob}\left(\xi^H H\xi > \gamma E(\xi^H H\xi)\right) = \text{Prob}\left(\sum_{i=1}^{\bar{r}} \bar{\lambda}_i |\bar{\xi}_i|^2 > \gamma\right),$$

where $\bar{\lambda}_1 \ge \bar{\lambda}_2 \ge \ldots \ge \bar{\lambda}_{\bar{r}} \ge 0$ satisfy $\bar{\lambda}_1 + \cdots + \bar{\lambda}_{\bar{r}} = 1$ and each $\bar{\xi}_i \in \mathbb{C}$ has the complex-valued normal distribution $N_c(0, 1)$. Then

$$\begin{aligned}
\text{Prob}\left(\xi^H H\xi > \gamma E(\xi^H H\xi)\right) &\le \text{Prob}\left(|\bar{\xi}_1|^2 > \gamma \text{ or } |\bar{\xi}_2|^2 > \gamma \text{ or } \cdots \text{ or } |\bar{\xi}_{\bar{r}}|^2 > \gamma\right) \\
&\le \sum_{i=1}^{\bar{r}} \text{Prob}\left(|\bar{\xi}_i|^2 > \gamma\right) \\
&= \bar{r}\ e^{-\gamma},
\end{aligned}$$

where the last step uses (14). ∎

Theorem 5 *For the complex QP (27) and its SDP relaxation (28), we have $v_{\text{sdp}} = v_{\text{qp}}$ if $m \le 3$ and otherwise*

$$v_{\text{qp}} \ge \frac{1}{4\ln(100K)} v_{\text{sdp}},$$

where $K := \sum_{i=1}^m \min\{\text{rank}\,(H_i), \sqrt{m}\}$.

Proof. By applying a suitable rank reduction procedure if necessary, we can assume that the rank r of the optimal SDP solution Z^* satisfies $r = 1$ if $m \le 3$ and $r \le \sqrt{m}$ if $m \ge 4$; see [9, Section 5]. Thus, if $m \le 3$, then $Z^* = z^*(z^*)^H$ for some $z^* \in \mathbb{C}^n$ and it is readily seen that z^* is an optimal solution of (27), so that $v_{\rm sdp} = v_{\rm qp}$. Otherwise, we apply the randomization procedure to Z^*. By using Lemma 5, we have, for any $\gamma > 0$ and $\mu > 0$,

$$\begin{aligned} & \text{Prob}\left(\max_{1\le i\le m} \xi^H H_i \xi \le \gamma,\ \|\xi\|^2 \ge \mu \text{Tr}(Z^*)\right) \\ \ge\ & 1 - \sum_{i=1}^m \text{Prob}\left(\xi^H H_i \xi > \gamma E(\xi^H H_i \xi)\right) - \text{Prob}\left(\|\xi\|^2 < \mu \text{Tr}(Z^*)\right) \\ \ge\ & 1 - K e^{-\gamma} - \text{Prob}\left(\|\xi\|^2 < \mu \text{Tr}(Z^*)\right), \end{aligned} \tag{30}$$

where the last step uses $r \le \sqrt{m}$.

Let

$$\eta_j := \begin{cases} |\xi_j|^2 / Z^*_{jj}, & \text{if } Z^*_{jj} > 0; \\ 0, & \text{if } Z^*_{jj} = 0, \end{cases} \qquad j = 1, ..., n.$$

For simplicity, let us assume that $Z^*_{jj} > 0$ for all $j = 1, ..., n$. Since $\xi_j \sim N_c(0, Z^*_{jj})$, as we discussed in Subsection 3.2, $|\xi_j|$ follows a Weibull distribution with variance Z^*_{jj} (see (14)), and therefore

$$\text{Prob}\,(\eta_j \le t) = 1 - e^{-t} \quad \forall\, t \in [0, \infty).$$

Hence,

$$E(\eta_j) = \int_0^\infty t e^{-t} dt = 1, \quad E(\eta_j^2) = \int_0^\infty t^2 e^{-t} dt = 2, \quad \text{Var}(\eta_j) = 1.$$

Moreover,

$$E(|\eta_j - E(\eta_j)|) = \int_0^1 (1-t) e^{-t} dt + \int_1^\infty (t-1) e^{-t} dt = \frac{2}{e}.$$

Let us denote $\lambda_j = Z^*_{jj}/\text{Tr}(Z^*)$, $j = 1, ..., n$, and $\eta := \sum_{j=1}^n \lambda_j \eta_j$. We have $E(\eta) = 1$ and

$$E(|\eta - E(\eta)|) = E\left(\left|\sum_{j=1}^n \lambda_j (\eta_j - E(\eta_j))\right|\right) \le \sum_{j=1}^n \lambda_j E(|\eta_j - E(\eta_j)|) = \frac{2}{e}.$$

Since, by Markov's inequality,

$$\text{Prob}\,(|\eta - E(\eta)| > \alpha) \le \frac{E(|\eta - E(\eta)|)}{\alpha} \le \frac{2}{\alpha e}, \quad \forall\, \alpha > 0,$$

we have

$$\begin{aligned} \text{Prob}\left(\|\xi\|^2 < \mu \text{Tr}(Z^*)\right) &= \text{Prob}\,(\eta < \mu) \\ &\le \text{Prob}\,(|\eta - E(\eta)| > 1 - \mu) \\ &\le \frac{2}{e(1-\mu)}, \qquad \text{for all } \mu \in (0,1). \end{aligned}$$

Substituting the above inequality into (30), we obtain

$$\text{Prob}\left(\max_{1\le i\le m}\xi^H H_i\xi \le \gamma,\ \|\xi\|^2 \ge \mu\text{Tr}(Z^*)\right) > 1 - Ke^{-\gamma} - \frac{2}{e(1-\mu)}, \quad \forall\, \mu \in (0,1).$$

Setting $\mu = 1/4$ and $\gamma = \ln(100K)$ yields a positive right-hand side of 0.00898.., which then proves the desired bound. ■

The above proof technique also applies to the real case, i.e., $h_\ell \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$. The main difference is that $\xi \sim N(0, Z^*)$, so that $|\bar{\xi}_i|^2$ in the proof of Lemma 5 and η_j in the proof of Theorem 5 both follow a χ^2 distribution with one degree of freedom. Then

$$\text{Prob}\left(|\bar{\xi}_i|^2 > \gamma\right) = \int_{\sqrt{\gamma}}^{\infty}\frac{e^{-t^2/2}}{\sqrt{2\pi}}dt \le \int_{\sqrt{\gamma}}^{\infty}\frac{e^{-\gamma t/2}}{\sqrt{2\pi}}dt = \sqrt{\frac{2}{\pi\gamma}}e^{-\gamma/2}, \quad \forall\,\gamma > 0,$$

$E(\eta_j) = 1$, and

$$\begin{aligned} E|\eta_j - E(\eta_j)| &= \int_0^\infty \frac{e^{-t/2}}{\sqrt{2\pi t}}|t-1|dt \\ &= \frac{1}{\sqrt{2\pi}}\int_0^1 \frac{e^{-t/2}}{\sqrt{t}}dt - \frac{1}{\sqrt{2\pi}}\int_0^1 \sqrt{t}e^{-t/2}dt \\ &\quad + \frac{1}{\sqrt{2\pi}}\int_1^\infty \sqrt{t}e^{-t/2}dt - \frac{1}{\sqrt{2\pi}}\int_1^\infty \frac{e^{-t/2}}{\sqrt{t}}dt \\ &= \frac{4}{\sqrt{2\pi e}} < 0.968, \end{aligned}$$

where in the last step we used integration by parts on the first and the fourth terms. This yields the analogous bound that, for any $\gamma \ge 1$ and $\mu \in (0,1)$,

$$\text{Prob}\left(\max_{1\le i\le m}\xi^T H_i\xi \le \gamma,\ \|\xi\|^2 \ge \mu\text{Tr}(Z^*)\right) > 1 - K\sqrt{\frac{2}{\pi\gamma}}e^{-\gamma/2} - \frac{0.968}{1-\mu} > 1 - Ke^{-\gamma/2} - \frac{0.968}{1-\mu},$$

where $K := \sum_{i=1}^m \min\{\text{rank}\,(H_i), \sqrt{2m}\}$. Setting $\mu = 0.01$ and $\gamma = 2\ln(50K)$ yields a positive right-hand side of 0.0022... This in turn shows that $v_{\text{sdp}} = v_{\text{qp}}$ if $m \le 2$ (see the proof of Theorem 1) and otherwise

$$v_{\text{qp}} \ge \frac{1}{200\ln(50K)}v_{\text{sdp}}.$$

We note that, in the real case, a sharper bound of

$$v_{\text{qp}} \ge \frac{1}{2\ln(2m\mu)}v_{\text{sdp}},$$

where $\mu := \min\{m, \max_i \operatorname{rank}(H_i)\}$, was shown by Nemirovski, Roos and Terlaky [14] (also see [13, Theorem 4.7]), though the above proof seems simpler. Also, an example in [14] shows that the $O(1/\ln m)$ bound is tight (up to a constant factor) in the worst case. This example readily extends to the complex case by identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ and observing that $|h_\ell^H z| \geq |\mathrm{Re}(h_\ell)^T\mathrm{Re}(z) + \mathrm{Im}(h_\ell)^T\mathrm{Im}(z)|$ for any $h_\ell, z \in \mathbb{C}^n$. Thus, in the complex case, the $O(1/\ln m)$ bound is also tight (up to a constant factor).

6 Discussion

In this paper, we have analyzed the worst-case performance of SDP relaxation and convex restriction for a class of NP-hard quadratic optimization problems with homogeneous quadratic constraints. Our analysis is motivated by important emerging applications in transmit beamforming for physical layer multicasting and sensor localization in wireless sensor networks. Our generalization (1) of the basic problem in [20] is useful, for it shows that the same convex approximation approaches and bounds hold in the case where each multicast receiver is equipped with multiple antennas. This scenario is becoming more pertinent with the emergence of small and cheap multi-antenna mobile terminals. Furthermore, our consideration of the related homogeneous QP maximization problem has direct application to the design of jam-resilient intercept beamformers. In addition to these timely topics, more traditional signal processing design problems can be cast in the same mathematical framework; see [20] for further discussions.

While theoretical worst-case analysis is very useful, empirical analysis of the ratio $\frac{v_{\mathrm{qp}}}{v_{\mathrm{sdp}}}$ through simulations with randomly generated steering vectors $\{h_\ell\}$ is often equally important. In the context of transmit beamforming for multicasting [20] for the case $|\mathcal{I}_i| = 1\ \forall i$ (single receiving antenna per subscriber node), simulations have provided the following insights:

- For moderate values of m, n (e.g., $m = 24$, $n = 8$), and independent and identically distributed (i.i.d.) complex-valued circular Gaussian (i.i.d. Rayleigh) entries of the steering vectors $\{h_\ell\}$, the average value of $\frac{v_{\mathrm{qp}}}{v_{\mathrm{sdp}}}$ is under 3 – much lower than the worst-case value predicted by our analysis.
- In all generated instances where all steering vectors have positive real and imaginary parts, the ratio $\frac{v_{\mathrm{qp}}}{v_{\mathrm{sdp}}}$ equals one (with error below 10^{-8}). This is better than what our worst-case analysis predicts for limited phase spread (see Theorem 3).
- In experiments with measured VDSL channel data, for which the steering vectors

follow a correlated log-normal distribution, $\frac{v_{\rm qp}}{v_{\rm sdp}} = 1$ in over 50% of instances.

- Our analysis shows that the worst-case performance ratio $\frac{v_{\rm qp}}{v_{\rm sdp}}$ is smaller in the complex case than in the real case ($O(m)$ versus $O(m^2)$). Moreover, this remains true with high probability when $v_{\rm qp}$ is replaced by its upper bound

$$v_{\rm ubqp} := \min_{k=1,\ldots,N} \|z^*(\xi^k)\|^2,$$

where $\xi^1, ..., \xi^N$ are generated by N independent trials of the randomization procedure (see Subsections 3.1 and 3.2) and N is taken sufficiently large. In our simulation, we used $N = 30nm$. Figure 1 shows our simulation results for the real Gaussian case.[6] It plots $\frac{v_{\rm ubqp}}{v_{\rm sdp}}$ for 300 independent realizations of i.i.d. real-valued Gaussian steering vector entries, for $m = 8$, $n = 4$. Figure 2 plots the corresponding histogram. Figures 3 and 4 show the corresponding results for i.i.d. complex-valued circular Gaussian steering vector entries.[7] Both the mean and the maximum of the upper bound $\frac{v_{\rm ubqp}}{v_{\rm sdp}}$ are lower in the complex case. The simulations indicate that SDP approximation is better in the complex case not only in the worst case but also on average.

The above empirical (worst-case and average-case) analysis complements our theoretical worst-case analysis of the performance of SDP relaxation for the class of problems considered herein.

Finally, we remark that our worst-case analysis of SDP performance is based on the assumption that the homogeneous quadratic constraints are concave (see (1)). Can we extend this analysis to general homogeneous quadratic constraints? The following example in $\mathbb{R}^2$ suggests that this is not possible.

Example 3: For any $L > 0$, consider the quadratic optimization problem with homogeneous quadratic constraints:

$$\begin{array}{rl} \min & \|z\|^2 \\ \text{s.t.} & z_2^2 \geq 1, \quad z_1^2 - Lz_1z_2 \geq 1, \quad z_1^2 + Lz_1z_2 \geq 1, \\ & z \in \mathbb{R}^2. \end{array} \tag{31}$$

The last two constraints imply $z_1^2 \geq L|z_1||z_2| + 1$ which, together with the first constraint $z_2^2 \geq 1$, yield $z_1^2 \geq L|z_1| + 1$ or, equivalently, $|z_1| \geq (L + \sqrt{L^2+4})/2$. So the optimal value

[6]Here the SDP solution is constrained to be real-valued, and real Gaussian randomization is used.

[7]Here the SDP solutions are complex-valued, and complex Gaussian randomization is used.

of (31) is at least $1 + (L + \sqrt{L^2 + 4})^2/4$ (and in fact is equal to this). The natural SDP relaxation of (31) is

$$\begin{array}{ll} \min & Z_{11} + Z_{22} \\ \text{s.t.} & Z_{22} \geq 1, \quad Z_{11} - LZ_{12} \geq 1, \quad Z_{11} + LZ_{12} \geq 1, \\ & Z \succeq 0. \end{array}$$

Clearly, $Z = I_2$ is a feasible solution (and, in fact, an optimal solution) of this SDP, with an objective value of 2. Therefore, the SDP performance ratio for this example is at least $1/2 + (L + \sqrt{L^2 + 4})^2/8$, which can be arbitrarily large.

References

[1] N. Alon and A. Naor, "Approximating the Cut-Norm via Grothendiecks inequality," in *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing,* Chicago, IL, USA, 2004, pp. 72–80.

[2] H.H. Andersen, M. Højbjerre, D. Sørensen, and P.S. Eriksen, "Linear and graphical models for the multivariate complex normal distribution," Lecture Notes in Statistics 101, Springer-Verlag, New York, 1995.

[3] A. Ben-Tal and A. Nemirovski, "On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty," *SIAM Journal on Optimization,* 12 (2002), pp. 811–833.

[4] A. Ben-Tal, A. Nemirovski, and C. Roos, "Extended matrix cube theorems with applications to μ-theory in control," *Mathematics of Operations Research*, 28 (2003), pp. 497–523.

[5] D. Bertsimas and Y. Ye, "Semidefinite relaxations, multivarite normal distribution, and order statistics," In *Handbook of Combinatorial Optimization*, D.Z. Du and P.M. Pardalos, eds., Kluwer Academic Publishers, 3 (1998) pp. 1–19.

[6] B. Biswas and Y. Ye, "Semidefinite programming for ad hoc wireless sensor network localization," Report, Electrical Engineering, Stanford University, Stanford, September 2003. `http://www.stanford.edu/~yyye/`

[7] M.X. Goemans and D.P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the ACM*, 42 (1995), pp. 1115–1145.

[8] M.X. Goemans and D.P. Williamson, "Approximation algorithms for MAX-3-CUT and other problems via complex semidefinite programming," *Journal of Computer and System Sciences,* 68 (2004), pp. 442–470.

[9] Y. Huang and S. Zhang, "Complex matrix decomposition and quadratic programming," Technical Report SEEM2005-02, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, 2005. `http://www.se.cuhk.edu.hk/~zhang/#workingpaper`

[10] D.H. Johnson and D.E. Dugeon, *Array Signal Processing: Concepts and Techniques,* Simon & Schuster, 1992.

[11] L. Lovász and A. Schrijver, "Cones of matrices and set-functions and 0-1 optimization," *SIAM Journal Optimization,* 1 (1991), pp. 166–190.

[12] S. Mahajan and H. Ramesh, "Derandomizing approximation algorithms based on semidefinite programming," *SIAM Journal on Computing*, 28 (1999), pp. 1641–1663.

[13] A. Megretski, "Relaxations of quadratic programs in operator theory and system analysis," *Operator Theory: Advances and Applications*, 129 (2001), pp. 365–392.

[14] A. Nemirovski, C. Roos and T. Terlaky, "On maximization of quadratic form over intersection of ellipsoids with common center," *Mathematical Programming*, 86 (1999), pp. 463–473.

[15] Y. Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization," *Optimization Methods and Softwares,* 9 (1998), pp. 141–160.

[16] Y. Nesterov, H. Wolkowicz, and Y. Ye, "Semidefinite programming relaxations of nonconvex quadratic optimization," in *Handbook of Semidefinite Programming*, edited by H. Wolkowicz, R. Saigal, and L. Vandenberghe, Kluwer, Boston, (2000), pp. 360–419.

[17] G. Pataki, "On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues," *Mathematics of Operations Research*, 23 (1998), pp. 339-358.

[18] G. Pataki, editor, "Computational semidefinite and second order cone programming: the state of the art," *Mathematical Programming*, 95 (2003).

[19] N.Z. Shor, "Quadratic optimization problems," *Soviet J. Comput. Systems Sci.,* 25 (1987), pp. 1–11.

[20] N.D. Sidiropoulos, T.N. Davidson and Z.-Q. Luo, “Transmit beamforming for physical layer multicasting,” *IEEE Transactions on Signal Processing*, 54 (2006), pp. 2239–2251.

[21] A. So, J. Zhang, and Y. Ye, “On approximating complex quadratic optimization problems via semidefinite programming relaxations,” in *Proceedings of the 11th Conference on Integer Programming and Combinatorial Optimization,* Lecture Notes in Computer Science, vol. 3509, M. Junger and V. Kaibel, eds., Springer-Verlag, Berlin, 2005, pp. 125–135.

[22] P. Tseng, “Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation,” *SIAM Journal on Optimization*, 14 (2003), pp. 268–283.

[23] Y. Ye, , “Approximating quadratic programming with bound and quadratic constraints,” *Mathematical Programming,* 84 (1999), pp. 219–226.

[24] Y. Ye, “Approximating global quadratic optimization with convex quadratic constraints,” *Journal of Global Optimization,* 15 (1999), pp. 1–17.

[25] S. Zhang, “Quadratic maximization and semidefinite relaxation,” *Mathematical Programming*, 87 (2000), pp. 453–465.

[26] S. Zhang and Y. Huang, “Complex quadratic optimization and semidefinite programming,” *SIAM Journal on Optimization*, 16 (2006), pp. 871–890.

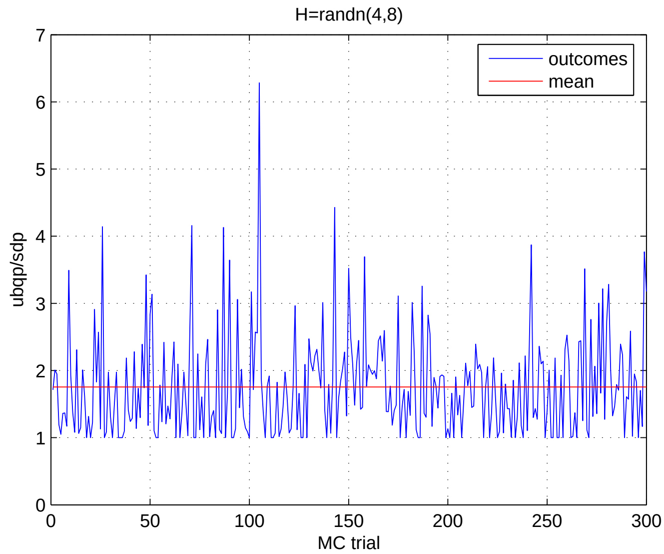

Figure 1: Upper bound on $\frac{v_{\text{qp}}}{v_{\text{sdp}}}$ for $m = 8$, $n = 4$, 300 realizations of real Gaussian i.i.d. steering vector entries, solution constrained to be real.

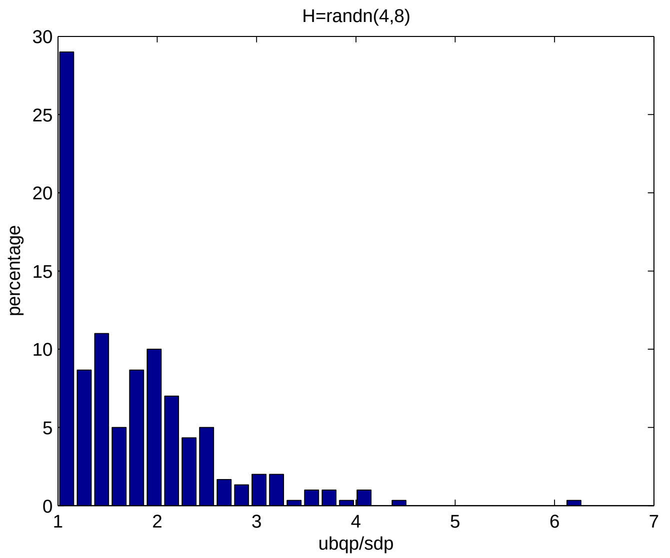

Figure 2: Histogram of the outcomes in Fig. 1.

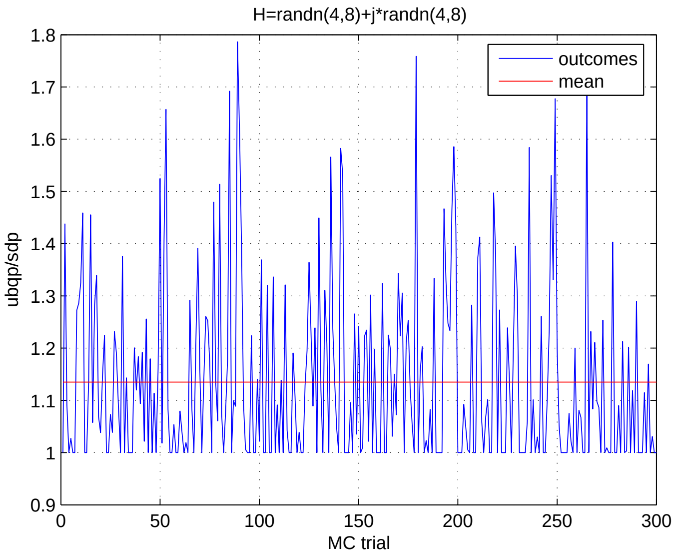

Figure 3: Upper bound on $\frac{v_{\text{qp}}}{v_{\text{sdp}}}$ for $m = 8$, $n = 4$, 300 realizations of complex Gaussian i.i.d. steering vector entries.

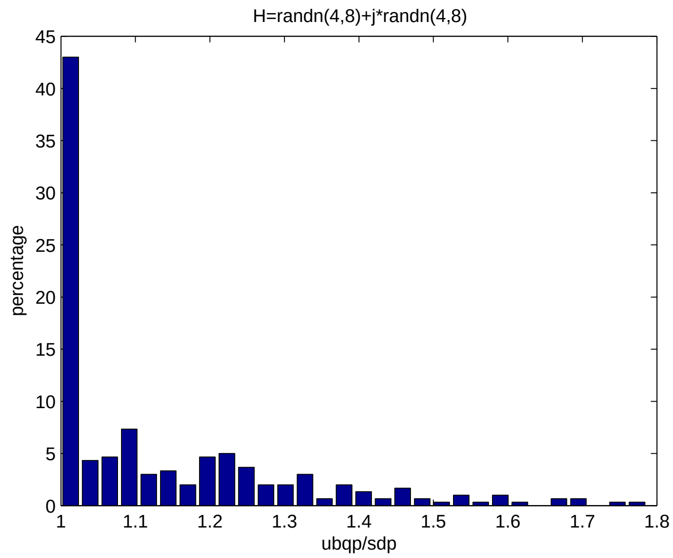

Figure 4: Histogram of the outcomes in Fig. 3.