Math. Program., Ser. B (2007) 110:93–110
DOI 10.1007/s10107-006-0064-6

FULL LENGTH PAPER

On approximating complex quadratic optimization problems via semidefinite programming relaxations

Anthony Man-Cho So · Jiawei Zhang · Yinyu Ye

Received: 26 July 2005 / Accepted: 12 June 2006 /
Published online: 13 December 2006

Abstract In this paper we study semidefinite programming (SDP) models for a class of discrete and continuous quadratic optimization problems in the complex Hermitian form. These problems capture a class of well-known combinatorial optimization problems, as well as problems in control theory. For instance, they include the MAX-3-CUT problem where the Laplacian matrix is positive semidefinite (in particular, some of the edge weights can be negative). We present a generic algorithm and a unified analysis of the SDP relaxations which allow us to obtain good approximation guarantees for our models. Specifically, we give an $(k\sin(\frac{\pi}{k}))^2/(4\pi)$-approximation algorithm for the discrete problem where the decision variables are k-ary and the objective matrix is positive semidefinite. To the best of our knowledge, this is the first known approximation result for this family of problems. For the continuous problem where the objective matrix is positive semidefinite, we obtain the well-known $\pi/4$ result due to Ben-Tal et al.

A preliminary version of this paper has appeared in the Proceedings of the 11th Conference on Integer Programming and Combinatorial Optimization (IPCO XI), 2005.
This research was supported in part by NSF grant DMS-0306611.

A. Man-Cho So (✉)
Department of Computer Science, Stanford University, Stanford, CA 94305, USA
e-mail: manchoso@cs.stanford.edu

J. Zhang
Department of Information, Operations, and Management Sciences, Stern School of Business, New York University, New York, NY 10012, USA
e-mail: jzhang@stern.nyu.edu

Y. Ye
Department of Management Science and Engineering,
and, by courtesy, Electrical Engineering, Stanford University, Stanford, CA 94305, USA
e-mail: yinyu-ye@stanford.edu

[Math Oper Res 28(3):497–523, 2003], and independently, Zhang and Huang [SIAM J Optim 16(3):871–890, 2006]. However, our techniques simplify their analyses and provide a unified framework for treating those problems. In addition, we show for the first time that the gap between the optimal value of the original problem and that of the SDP relaxation can be arbitrarily close to $\pi/4$. We also show that the unified analysis can be used to obtain an $\Omega(1/\log n)$-approximation algorithm for the continuous problem in which the objective matrix is not positive semidefinite.

Keywords Hermitian quadratic functions · Complex semidefinite programming · Grothendieck's inequality

Mathematics Subject Classification (2000) 90C20 · 90C22 · 90C27 · 90C90

1 Introduction

Following the seminal work of Goemans and Williamson [7], there has been an outgrowth in the use of semidefinite programming (SDP) for designing approximation algorithms. Recall that an α-approximation algorithm for a problem $\mathscr{P}$ is a polynomial-time algorithm such that for every instance I of $\mathscr{P}$, it delivers a solution that is within a factor of α of the optimum value [9]. It is well-known that SDPs can be solved in polynomial time (up to any prescribed accuracy) via interior-point algorithms (see, e.g., [14]), and they have been used very successfully in the design of approximation algorithms for a host of NP-hard problems, e.g. graph partitioning, graph coloring, and quadratic optimization [5–8,11,15], just to name a few.

In this paper, we consider a class of discrete and continuous quadratic optimization problems in the complex Hermitian form. Specifically, we consider the following problems:

$$\begin{array}{ll} \text{maximize} & z^H Q z \\ \text{subject to} & z_j \in \left\{1, \omega, \ldots, \omega^{k-1}\right\} \qquad j = 1, 2, \ldots, n \end{array} \tag{1}$$

and

$$\begin{array}{ll} \text{maximize} & z^H Q z \\ \text{subject to} & |z_j| = 1 \qquad j = 1, 2, \ldots, n \\ & z \in C^n \end{array} \tag{2}$$

where $Q \in C^{n\times n}$ is a Hermitian matrix, ω is the principal kth root of unity, and z^H denotes the conjugate transpose of the complex vector $z \in C^n$. The difference between (1) and (2) lies in the values that the decision variables are allowed to take. In Problem (1), we have discrete decision variables, and such variables can be conveniently modelled as roots of unity. On the other hand, in Problem (2), the decision variables are constrained to lie on the unit circle, which is a continuous domain. Such problems arise from many applications.

For instance, the MAX-3-CUT problem where the Laplacian matrix is positive semidefinite can be formulated as an instance of (1). On the other hand, (2) arises from the study of robust optimization as well as control theory [3,13].

It is known that both of these problems are NP-hard, and thus we will settle for approximation algorithms. Previously, various researchers have considered SDP relaxations for (1) and (2). However, approximation guarantee is known only for the continuous problem [3,16], and to the best of our knowledge, no such guarantees are known for the discrete problem prior to our work.

Our main contribution is to present a generic algorithm and a unified treatment of the two seemingly very different problems (1) and (2) using their natural SDP relaxations, and to give the first known approximation result for the discrete problem. Specifically, we are able to achieve an $(k\sin(\frac{\pi}{k}))^2/(4\pi)$-approximation ratio for the discrete problem.[1] As a corollary, we obtain an 0.537-approximation algorithm for the MAX-3-CUT problem where the Laplacian matrix is positive semidefinite. This should be contrasted with the 0.836-approximation algorithm of Goemans and Williamson [8] for MAX-3-CUT with *non-negative* edge weights. For this particular case, our result might be seen as a generalization of Nesterov's result [11] which gives an $2/\pi$-approximation for the MAX-CUT problem where the Laplacian matrix is positive semidefinite.

For the continuous problem, our analysis also achieves the $\pi/4$ guarantee of [3,16]. However, our analysis is simpler than that of [3,16], and it follows the same framework as that of the discrete problem. Moreover, we give a family of examples showing that the gap between the optimal value of Problem (2) and that of the SDP relaxation can be arbitrarily close to $\pi/4$. In addition, we show that the unified analysis can be used to obtain an $\Omega(1/\log n)$-approximation algorithm for the continuous problem in the case where the objective matrix is not positive semidefinite. We remark that the $\Omega(1/\log n)$ bound (as well as the $\Omega(1/\log n)$ bound obtained by Charikar and Wirth [4] for the real analog of the continuous problem) follows directly from a result of Nemirovski, Roos and Terlaky [10]. However, both our algorithm and its analysis are different from those in [10]. In particular, the performance guarantee of our algorithm is a function of the entries of the objective matrix Q, and that function is lower bounded by $\Omega(1/\log n)$. In contrast, it is not clear if the algorithm of [10] has such a feature.

To motivate our approach, we first remark that our rounding schemes are essentially the same as that of Goemans and Williamson [7,8]. However, since we do not make any assumptions on the entries of the objective matrix Q, we cannot use their analysis to establish the desired approximation results. Indeed, one apparent difficulty in analyzing SDP relaxation-based algorithms for Problems (1) and (2) is that the usual Goemans–Williamson analysis [7,8] (and its variants thereof) only provides a term-by-term estimate of the objective

[1] By extending their analysis in [16], Zhang and Huang are also able to obtain the same approximation ratio for the discrete problem. We refer the readers to the journal version of [16] for details.

function and does not provide a global estimate. Although global techniques for analyzing (real) SDP relaxations exist [11], it is not clear how they can be applied to our problems. Our analysis is mainly inspired by a recent result of Alon and Naor [2], who proposed several methods for analyzing (real) SDP relaxations in a global manner using results from functional analysis. One of those methods, which is based on the work of Rietz [12], uses averaging with Gaussian measure and the simple fact that $\sum_{i,j} q_{ij}(v_i \cdot v_j) \geq 0$ if the matrix $Q = (q_{ij})$ is positive semidefinite (here, $v_i \cdot v_j$ is the inner product of two vectors v_i and v_j in some Hilbert space). Our results for (1) and (2) in the case where Q is positive semidefinite are motivated by this method. Although the assumption that Q is positive semidefinite is essential to make the analyses go through, we manage to analyze our algorithm in a unified way for the case where Q is not positive semidefinite as well.

The rest of the paper is organized as follows. In Sect. 2 we give the setting of our problems and a generic algorithm for solving those problems. In Sect. 3 we consider the discrete problem where the objective matrix is positive semidefinite, and give an analysis on the approximation guarantee of our algorithm. In Sect. 4 we consider the continuous problem when the objective matrix is not semidefinite. Finally, in Sect. 5, we summarize our findings and provide some possible future directions.

2 Complex quadratic optimization

Let $Q \in C^{n\times n}$ be a Hermitian matrix, where $n \geq 1$ is an integer. Consider the following discrete quadratic optimization problem:

$$\begin{array}{ll} \text{maximize} & z^H Q z \\ \text{subject to} & z_j \in \{1, \omega, \ldots, \omega^{k-1}\} \qquad j = 1, 2, \ldots, n \end{array} \tag{3}$$

where ω is the principal kth root of unity. We note that as k goes to infinity, the discrete problem (3) becomes a continuous optimization problem:

$$\begin{array}{ll} \text{maximize} & z^H Q z \\ \text{subject to} & |z_j| = 1 \qquad j = 1, 2, \ldots, n \\ & z \in C^n \end{array} \tag{4}$$

Although Problems (3) and (4) are quite different in nature, the following *complex* semidefinite program provides a relaxation for both of them:

$$\begin{array}{ll} \text{maximize} & Q \bullet Z \\ \text{subject to} & Z_{jj} = 1 \qquad j = 1, 2, \ldots, n \\ & Z \succeq 0 \end{array} \tag{5}$$

We use w_{SDP} to denote the optimal value of the SDP relaxation (5).

Our goal is to get a near optimal solution to Problems (3) and (4). Below we present a generic algorithm that can be used to solve both (3) and (4). Our algorithm is quite simple, and it is similar in spirit to the algorithm of Goemans and Williamson [7,8].

Algorithm

Step 1. Solve the SDP relaxation (5) and obtain an optimal solution Z^*. Since Z^* is positive semidefinite, we can obtain a Cholesky decomposition $Z^* = VV^H$, where $V = (v_1, v_2, \ldots, v_n)$.

Step 2. Generate two independent normally distributed random vectors $x \in R^n$ and $y \in R^n$ with mean 0 and covariance matrix $\frac{1}{2}I_n$, where I_n is the $n \times n$ identity matrix. Let $r = x + yi$.

Step 3. For $j = 1, 2, \ldots, n$, let $\hat{z}_j = f(v_j \ldots r)$, where the function $f(\cdot)$ depends on the structure of the problem and will be fixed later. Let $\hat{z} = (\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_n)$ be the resulting solution.

In order to prove the performance guarantee of our algorithm, we are interested in analyzing the quantity:

$$\hat{z}^H Q \hat{z} = Q \bullet \hat{z}\hat{z}^H = \sum_{l,m} Q_{lm} \hat{z}_l \overline{\hat{z}_m} = \sum_{l,m} Q_{lm} f(v_l \cdot r) \overline{f(v_m \cdot r)}$$

Since our algorithm is randomized, we compute the expected objective value of the solution $\hat{z}$. By linearity of expectation, we have:

$$\mathrm{E}\left[\hat{z}^H Q \hat{z}\right] = \sum_{l,m} Q_{lm} \mathrm{E}[f(v_l \cdot r) \overline{f(v_m \cdot r)}]$$

Thus, it would be sufficient to compute the quantity $\mathrm{E}[f(v_l \cdot r)\overline{f(v_m \cdot r)}]$ for any l, m, and this will be the main concern of our analysis. The analysis, of course, depends on the choice of the function $f(\cdot)$. However, the following lemma will be useful and it is independent of the function $f(\cdot)$. Recall that for two vectors $b, c \in C^n$, we have $b \cdot c = \sum_{j=1}^n b_j \overline{c_j}$.

Lemma 1 *For any pair of vectors $b, c \in C^n$, $\mathrm{E}[(b \cdot r)\overline{(c \cdot r)}] = b \cdot c$, where $r = x + yi$ and $x \in R^n$ and $y \in R^n$ are two independent normally distributed random vector with mean 0 and covariance matrix $\frac{1}{2}I_n$.*

Proof This follows from a straightforward computation:

$$\mathrm{E}[(b \cdot r)\overline{(c \cdot r)}] = \mathrm{E}\left[\left(\sum_{j=1}^n b_j \bar{r}_j\right)\left(\sum_{k=1}^n \bar{c}_k r_k\right)\right] = \sum_{j,k=1}^n b_j \bar{c}_k \mathrm{E}[\bar{r}_j r_k] = \sum_{j=1}^n b_j \bar{c}_j$$

where the last equality follows from the fact that the entries of x and y are independent normally distributed with mean 0 and variance 1/2. □

In the sequel, we shall use $r \sim \mathcal{N}_C(0, I_n)$ to indicate that r is an n-dimensional standard complex normal random vector, i.e. $r = x + yi$, where $x, y \in R^n$ are two independent normally distributed random vectors, each with mean 0 and covariance matrix $\frac{1}{2}I_n$.

3 Discrete problems where Q is positive semidefinite

In this section, we assume that Q is Hermitian and positive semidefinite. Consider the discrete complex quadratic optimization problem (3). In this case, we define the function $f(\cdot)$ in the generic algorithm presented in Sect. 2 as follows:

$$f(z) = \begin{cases} 1 & \text{if } \arg(z) \in [-\pi/k, \pi/k) \\ \omega & \text{if } \arg(z) \in [\pi/k, 3\pi/k) \\ \vdots & \vdots \\ \omega^{k-1} & \text{if } \arg(z) \in [(2k-3)\pi/k, (2k-1)\pi/k) \end{cases} \tag{6}$$

By construction, we have $\hat{z}_j \in \{1, \omega, \ldots, \omega^{k-1}\}$ for $j = 1, 2, \ldots, n$, i.e. $\hat{z}$ is a feasible solution to Problem (3). Now, we can establish the following lemma:

Lemma 2 *For any pair of vectors $b, c \in C^n$ and $r \sim \mathcal{N}_C(0, I_n)$, we have:*

$$\mathrm{E}[(b \cdot r)\overline{f(c \cdot r)}] = \frac{k \sin(\pi/k)}{2\sqrt{\pi}}(b \cdot c)$$

Proof By rotation invariance, we may assume without loss of generality that $b = (b_1, b_2, 0, \ldots, 0)$ and $c = (1, 0, \ldots, 0)$. Then, we have

$$\begin{aligned} \mathrm{E}[(b_1\overline{r_1} + b_2\overline{r_2})\overline{f(\overline{r_1})}] &= b_1 \mathrm{E}[\overline{r_1}\overline{f(\overline{r_1})}] \\ &= \frac{b_1}{\pi} \int_{\mathbf{R}} \int_{\mathbf{R}} (x - iy)\overline{f(x - iy)} \exp\{-(x^2 + y^2)\} \, \mathrm{d}x\mathrm{d}y \\ &= \frac{b_1}{\pi} \int_0^{\infty} \int_0^{2\pi} \rho^2 \mathrm{e}^{-i\theta} \overline{f(\rho \mathrm{e}^{-i\theta})} \mathrm{e}^{-\rho^2} \, \mathrm{d}\theta \mathrm{d}\rho \end{aligned}$$

Now, for any $j = 1, \ldots, k$, if $(2j-3)\pi/k < \theta \leq (2j-1)\pi/k$, then $-(2j-1)\pi/k \leq -\theta < -(2j-3)\pi/k$, or

$$\frac{2k - 2j + 1}{k}\pi \leq 2\pi - \theta < \frac{2k - 2j + 3}{k}\pi$$

It then follows from the definition of $f(\cdot)$ that:

$$f\left(\rho \mathrm{e}^{-i\theta}\right) = f\left(\rho \mathrm{e}^{i(2\pi - \theta)}\right) = \omega^{k-j+1}$$

and hence $\overline{f(\rho e^{-i\theta})} = \omega^{j-1}$. Therefore, we have

$$\int_{(2j-3)\pi/k}^{(2j-1)\pi/k} \overline{f(\rho e^{-i\theta})} e^{-i\theta} \, d\theta = \omega^{j-1} \int_{(2j-3)\pi/k}^{(2j-1)\pi/k} e^{-i\theta} \, d\theta = 2\sin(\pi/k)$$

In particular, the above quantity is independent of j. Thus, we conclude that:

$$\int_0^{2\pi} \overline{f(\rho e^{-i\theta})} e^{-i\theta} \, d\theta = 2k\sin(\pi/k)$$

Moreover, since we have:

$$\int_0^{\infty} \rho^2 e^{-\rho^2} \, d\rho = \frac{\sqrt{\pi}}{4}$$

it follows that:

$$E[(b_1\overline{r_1} + b_2\overline{r_2})\overline{f(\overline{r_1})}] = \frac{k\sin(\pi/k)}{2\sqrt{\pi}} b_1 = \frac{k\sin(\pi/k)}{2\sqrt{\pi}} (b \cdot c)$$

as desired. □

We are now ready to prove the main result of this section.

Theorem 1 *Suppose that Q is Hermitian and positive semidefinite. Then, there exists an $\frac{(k\sin(\frac{\pi}{k}))^2}{4\pi}$-approximation algorithm for* (3).

Proof By Lemmas 1 and 2, we have

$$\begin{aligned} &E\left[\left\{(b \cdot r) - \frac{2\sqrt{\pi}}{k\sin(\frac{\pi}{k})} f(b \cdot r)\right\} \overline{\left\{(c \cdot r) - \frac{2\sqrt{\pi}}{k\sin(\frac{\pi}{k})} f(c \cdot r)\right\}}\right] \\ &= -(b \cdot c) + \frac{4\pi}{(k\sin(\frac{\pi}{k}))^2} E[f(b \cdot r)\overline{f(c \cdot r)}] \end{aligned}$$

Now, for each vector $u \in C^n$, let us define the function $h_u : C^n \to C$ by:

$$h_u(r) = u \cdot r - \frac{2\sqrt{\pi}}{k\sin(\frac{\pi}{k})} f(u \cdot r)$$

It then follows that:

$$E[\hat{z}^H Q \hat{z}] = \frac{(k\sin(\frac{\pi}{k}))^2}{4\pi} \left[\sum_{l=1}^{n} \sum_{m=1}^{n} q_{lm}(v_l \cdot v_m) + \sum_{l=1}^{n} \sum_{m=1}^{n} q_{lm} E\left[h_{v_l} \overline{h_{v_m}}\right]\right] \quad (7)$$

Now, we claim that:

$$\sum_{l=1}^{n}\sum_{m=1}^{n} q_{lm}\mathrm{E}\left[h_{v_l}\overline{h_{v_m}}\right] \geq 0 \tag{8}$$

To see this, let G be the standard complex Gaussian measure, i.e.:

$$\mathrm{d}G(r) = \frac{1}{\pi^n}\exp\left(-\|r\|^2\right)\mathrm{d}r$$

where $\|r\|^2 = |r_1|^2 + \cdots + |r_n|^2$ and $\mathrm{d}r$ is the $2n$-dimensional Lebesgue measure. Consider the Hilbert space $L^2(G)$, i.e. the space of all complex-valued measurable functions f on C^n with $\int_{C^n} |f|^2\,\mathrm{d}G < \infty$. Recall that the inner product on $L^2(G)$ is given by:

$$\langle f_u, f_v\rangle \equiv \int\limits_{C^n} f_u(r)\overline{f_v(r)}\,\mathrm{d}G(r) = \mathrm{E}[f_u\overline{f_v}]$$

In particular, since $h_u \in L^2(G)$ for each $u \in C^n$, we see that $\mathrm{E}\left[h_{v_l}\overline{h_{v_m}}\right]$ is an inner product of two vectors in the Hilbert space $L^2(G)$. Moreover, we may consider Q as a positive semidefinite operator defined on the n-dimensional subspace spanned by the vectors $\{h_{v_1}, \ldots, h_{v_n}\}$. These observations allow us to conclude that (8) holds. Finally, upon substituting (8) into (7), we obtain:

$$\mathrm{E}\left[\hat{z}^H Q\hat{z}\right] \geq \frac{(k\sin(\frac{\pi}{k}))^2}{4\pi}\sum_{l=1}^{n}\sum_{m=1}^{n} q_{lm}(v_l \cdot v_m) = \frac{(k\sin(\frac{\pi}{k}))^2}{4\pi} w_{\mathrm{SDP}}$$

i.e. our algorithm gives an $\frac{(k\sin(\frac{\pi}{k}))^2}{4\pi}$-approximation. □

As an application of Theorem 1, we consider the MAX-3-CUT problem, which is defined as follows. We are given an undirected graph $G = (V, E)$ with V being the set of nodes and E being the set of edges. For each edge $(i,j) \in E$, there is a weight w_{ij} that could be positive or negative. For a partition of V into three subsets V_1, V_2 and V_3, we define:

$$\delta(V_1, V_2, V_3) = \{(i,j) \in E : i \in V_k, j \in V_l \quad \text{for } k \neq l\}$$

and

$$w(\delta(V_1, V_2, V_3)) = \sum_{(i,j)\in\delta(V_1,V_2,V_3)} w_{ij}$$

Our goal is to find a tripartition (V_1, V_2, V_3) such that $w(\delta(V_1, V_2, V_3))$ is maximized. Notice that the MAX-3-CUT problem is a generalization of the well-known MAX-CUT problem. In the MAX-CUT problem, we require one of the subsets, say V_3, to be empty.

Goemans and Williamson [8] have given the following complex quadratic programming formulation for the MAX-3-CUT problem:

$$\begin{array}{ll} \text{maximize} & \frac{1}{3}\sum_{(i,j)\in E} w_{ij}\left(2 - z_i \cdot z_j - z_j \cdot z_i\right) \\ \text{subject to} & z_j \in \{1, \omega, \omega^2\} \quad \text{for all} \quad j \in V \end{array} \tag{9}$$

Based on this formulation and its SDP relaxation, Goemans and Willimason [8] are able to give an 0.836-approximation algorithm for the MAX-3-CUT problem when the weights of the edges are nonnegative, i.e. $w_{ij} \geq 0$ for all $(i,j) \in E$. (They also show that their algorithm is actually the same as that of Frieze and Jerrum [5], and thus give a tighter analysis of the algorithm in [5].) However, their analysis does not apply if some of the edges have negative weights.

Notice that since $w_{ij} = w_{ji}$, Problem (9) is equivalent to:

$$\begin{array}{ll} \text{maximize} & \frac{2}{3} z^H L z \\ \text{subject to} & z_j \in \{1, \omega, \omega^2\} \quad \text{for all} \quad j \in V \end{array} \tag{10}$$

where L is the Laplacian matrix of the graph $G = (V, E)$, i.e. $L_{ij} = -w_{ij}$ and $L_{ii} = \sum_{j:(i,j)\in E} w_{ij}$. However, by Theorem 1, Problem (10) can be approximated by a factor of $\frac{(3\sin(\frac{\pi}{3}))^2}{4\pi} \approx 0.537$. Therefore, we obtain the following result:

Corollary 1 *There is a randomized* 0.537*-approximation algorithm for the* MAX-3-CUT *problem when the Laplacian matrix is positive semidefinite.*

Now, let us consider Problem (4) when Q is positive semidefinite. This problem can be seen as a special case of (3) by letting $k \to \infty$. In this case, the function $f(\cdot)$ is defined as follows:

$$f(t) = \begin{cases} \frac{t}{|t|} & \text{if } |t| > 0 \\ 0 & \text{if } t = 0 \end{cases} \tag{11}$$

Note that as $k \to \infty$, we have $\frac{(k\sin(\frac{\pi}{k}))^2}{4\pi} \to \pi/4$. This establishes the following result, which has been proven independently by Ben–Tal, Nemirovski and Roos [3], and Zhang and Huang [16]. However, our proof is quite a bit simpler.

Corollary 2 *Suppose that Q is positive semidefinite and Hermitian. Then, there exists an $\frac{\pi}{4}$-approximation algorithm for* (4).

Next, we show that our analysis is in fact tight for the continuous complex quadratic optimization problem (4). We give a family of examples which shows that the natural SDP relaxation for the above problem has a gap arbitrarily close to $\pi/4$. We begin with a technical lemma.

Lemma 3 *Let u, v be two random, independent vectors on the unit sphere of C^p. Then, we have*

$$\mathrm{E}\left[|u \cdot v|^2\right] = \frac{1}{p}; \qquad \mathrm{E}[|u \cdot v|] = \left(\frac{\sqrt{\pi}}{2} + o(1)\right)\frac{1}{\sqrt{p}}$$

Proof By rotation invariance, we may assume that $v = (1, 0, \ldots, 0)$. Observe that u and $r/\|r\|^2$, where $r \sim \mathcal{N}_C(0, I_p)$, have the same distribution. Thus, the first statement follows from:

$$\mathrm{E}\left[|u \cdot v|^2\right] = \mathrm{E}\left[|u_1|^2\right] = \frac{1}{p}\mathrm{E}\left[|u_1|^2 + \cdots + |u_p|^2\right] = \frac{1}{p}$$

using symmetry and linearity of expectation. For the second statement, observe that:

$$\mathrm{E}\left[|u \cdot v|\right] = \mathrm{E}[|u_1|] = \int_{C^n} \frac{|r_1|}{\sqrt{|r_1|^2 + \cdots + |r_p|^2}}\, \mathrm{d}G(r)$$

Since $|r_1|^2, |r_2|^2, \ldots$ are i.i.d. with $\mathrm{E}\left[|r_1|^2\right] = 1$, we have, by the Strong Law of Large Numbers, that:

$$\sqrt{|r_1|^2 + \cdots + |r_p|^2} \to \sqrt{p} \qquad \text{a.s.}$$

It then follows that:

$$\begin{aligned}\int_{C^n} \frac{|r_1|}{\sqrt{|r_1|^2 + \cdots + |r_p|^2}}\, \mathrm{d}G(r) &= \frac{1 + o(1)}{\sqrt{p}} \cdot \int_C |r_1|\, \mathrm{d}G(r_1) \\ &= \left(\frac{\sqrt{\pi}}{2} + o(1)\right)\frac{1}{\sqrt{p}}\end{aligned}$$

as desired. □

To construct the tight example, let p and $n \gg p$ be fixed. Let $v_1, \ldots, v_n$ be independent random vectors chosen uniformly according to the normalized Haar measure on the unit sphere of C^p. We define $A = (a_{ij})$ by $a_{ij} = \frac{1}{n^2}\overline{(v_i \cdot v_j)}$. By construction, the matrix A is positive semidefinite and Hermitian. Moreover, we have

$$\sum_{i,j} a_{ij}(v_i \cdot v_j) = \frac{1}{n^2}\sum_{i,j} |v_i \cdot v_j|^2$$

By taking $n \to \infty$, the right-hand side converges to the average of the square of the inner product between two random vectors on the unit sphere of C^n.

By Lemma 3, this value is $1/p$, and hence the optimal value of the SDP relaxation is at least $1/p$.

Now, let $z_i \in C$ be such that $|z_i| = 1$. Then, we have

$$z^H A z = \sum_{i,j} a_{ij} \overline{z_i} z_j = \left| \frac{1}{n} \sum_{i=1}^{n} z_i v_i \right|^2$$

Hence, the value of the original SDP is the square of the maximum possible modulus of a vector $\frac{1}{n} \sum_{i=1}^{n} z_i v_i$. If we somehow know that the direction of this optimal vector is given by the unit vector c, then we must set $z_i = \overline{f(v_i \cdot c)}$ in order to maximize the modulus. It then follows that:

$$\frac{1}{n} \sum_{i=1}^{n} z_i v_i \cdot c = \left| \frac{1}{n} \sum_{i=1}^{n} z_i v_i \right|$$

by the Cauchy–Schwarz inequality. Moreover, this quantity converges to the average value of $|v \cdot c|$ as $n \to \infty$. By letting n arbitrarily large and choosing an appropriate ϵ-net of directions on the sphere, we conclude that with high probability, the value of the original SDP is at most $[(\sqrt{\pi}/2 + o(1))/\sqrt{p}]^2 = (\pi/4 + o(1))/p$, which yields the desired result.

4 Continuous problems where Q is not positive semidefinite

In this section, we deal with Problem (4) where the matrix Q is not positive semidefinite. However, for convenience, we assume that $w_{\text{SDP}} > 0$ so that the standard definition of approximation algorithm makes sense for our problem. It is clear that $w_{\text{SDP}} > 0$ as long as all the diagonal entries of Q are zeros.

Assumption 1 *The diagonal entries of Q are all zeros, i.e. $Q_{ii} = 0$ for $i = 1, 2, \ldots, n$.*

In fact, Assumption 1 leads to the following even stronger result.

Lemma 4 *If Q satisfies Assumption 1, then there exists a constant $C > 0$ such that:*

$$w_{\text{SDP}} \geq C \sqrt{\sum_{1 \leq i,j \leq n} |q_{ij}|^2} > 0$$

Proof It is straightforward to show that Problem (4) is equivalent to:

$$\begin{array}{lll} \text{maximize} & (x^{\mathrm{T}}, y^{\mathrm{T}}) \begin{pmatrix} \mathrm{Re}(Q) & \mathrm{Im}(Q) \\ -\mathrm{Im}(Q) & \mathrm{Re}(Q) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} & \\ \text{subject to} & x_j^2 + y_j^2 = 1 & j = 1, 2, \ldots, n \\ & x, y \in R^n & \end{array} \tag{12}$$

Moreover, the objective value of (12) is bounded below by the objective value of the following problem:

$$\begin{array}{lll} \text{maximize} & \left(x^{\mathrm{T}}, y^{\mathrm{T}}\right) \begin{pmatrix} \mathrm{Re}(Q) & \mathrm{Im}(Q) \\ -\mathrm{Im}(Q) & \mathrm{Re}(Q) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} & \\ \text{subject to} & x_j^2 = \frac{1}{2} & j = 1, 2, \ldots, n \\ & y_j^2 = \frac{1}{2} & j = 1, 2, \ldots, n \\ & x, y \in R^n & \end{array} \tag{13}$$

Since Q satisfies Assumption 1, the diagonal entries of

$$\begin{pmatrix} \mathrm{Re}(Q) & \mathrm{Im}(Q) \\ -\mathrm{Im}(Q) & \mathrm{Re}(Q) \end{pmatrix}$$

must also be zeros. It has been shown in [1] that for any real matrix $A = (a_{ij})_{n\times n}$ with a zero diagonal, the optimal objective value of

$$\begin{array}{lll} \text{maximize} & x^{\mathrm{T}} A x & \\ \text{subject to} & x_j^2 = 1 & j = 1, 2, \ldots, n \\ & x \in R^n & \end{array} \tag{14}$$

is bounded below by $C\sqrt{2\sum_{1\le i,j\le n} |a_{ij}|^2}$, for some constant $C > 0$ which is independent of A. This implies that the optimal objective value of Problem (13) is at least

$$\frac{1}{2} C \sqrt{2 \sum_{1\le i,j\le n} 2\left(|\mathrm{Re}(q_{ij})|^2 + |\mathrm{Im}(q_{ij})|^2\right)} \ge C \sqrt{\sum_{1\le i,j\le n} |q_{ij}|^2}$$

which leads to the desired result. □

Again, we use our generic algorithm presented in Sect. 2. In this case, we specify the function $f(\cdot)$ as follows:

$$f(t) = \begin{cases} \frac{t}{T} & \text{if } |t| \le T \\ \frac{t}{|t|} & \text{if } |t| > T \end{cases} \tag{15}$$

where T is a parameter which will be fixed later. If we let $z_j = f(v_j \cdot r)$, then the solution $z = (z_1, \ldots, z_n)$ obtained by this rounding may not be feasible, as the point may not have unit modulus. However, we know that $|z_j| \le 1$. Thus, we can further round the solution as follows:

$$\hat{z} = \begin{cases} z/|z| & \text{with probability } (1 + |z|)/2 \\ -\bar{z}/|z| & \text{with probability } (1 - |z|)/2 \end{cases}$$

The following lemma is a direct consequence of the second randomized rounding.

Lemma 5 *For $i \neq j$, we have* $\mathrm{E}[\hat{z}_i\overline{\hat{z}_j}] = \mathrm{E}[z_i\overline{z_j}]$.

Proof By definition, conditioning on z_i, z_j, we have

$$\begin{aligned}
\mathrm{E}[\hat{z}_i\overline{\hat{z}_j}\,|\,z_i,z_j] &= \Pr\{\hat{z}_i = z_i/|z_i|,\ \hat{z}_j = z_j/|z_j|\}\cdot\frac{z_i\overline{z_j}}{|z_i|\cdot|\overline{z_j}|}\\
&\quad + \Pr\{\hat{z}_i = z_i/|z_i|,\ \hat{z}_j = -z_j/|z_j|\}\cdot\left(-\frac{z_i\overline{z_j}}{|z_i|\cdot|\overline{z_j}|}\right)\\
&\quad + \Pr\{\hat{z}_i = -z_i/|z_i|,\ \hat{z}_j = z_j/|z_j|\}\cdot\left(-\frac{z_i\overline{z_j}}{|z_i|\cdot|\overline{z_j}|}\right)\\
&\quad + \Pr\{\hat{z}_i = -z_i/|z_i|,\ \hat{z}_j = -z_j/|z_j|\}\cdot\frac{z_i\overline{z_j}}{|z_i|\cdot|\overline{z_j}|}\\
&= \frac{1}{2}\left(1+|z_i|\cdot|z_j|\right)\cdot\frac{z_i\overline{z_j}}{|z_i|\cdot|\overline{z_j}|} - \frac{1}{2}\left(1-|z_i|\cdot|z_j|\right)\cdot\frac{z_i\overline{z_j}}{|z_i|\cdot|\overline{z_j}|}\\
&= z_i\overline{z_j}
\end{aligned}$$

The desired result then follows from the tower property of conditional expectation. □

This shows that the expected value of the solution on the circle equals that of the "fractional" solution obtained by applying $f(\cdot)$ to the SDP solution. Therefore, we could still restrict ourselves to the rounding function $f(\cdot)$.

Now, define:

$$g(T) = \frac{1}{T} - \frac{1}{T}\mathrm{e}^{-\mathrm{T}^2} + \sqrt{\pi}(1-\Phi(\sqrt{2}T))$$

where $\Phi(\cdot)$ is the probability distribution function of $\mathcal{N}(0,1)$.

Lemma 6 *For any pair of vectors $b, c \in C^n$ and $r \sim \mathcal{N}_C(0, I_n)$, we have:*

$$\mathrm{E}[(b\cdot r)\overline{f(c\cdot r)}] = g(T)(b\cdot c)$$

Proof Again, without loss of generality, we assume that $c = (1,0,\dots,0)$ and $b = (b_1, b_2, 0, \dots, 0)$. Let $\mathbf{1}_A$ be the indicator function of the set A, i.e. $\mathbf{1}_A(\omega) = 1$

if $\omega \in A$ and $\mathbf{1}_A(\omega) = 0$ otherwise. Then, we have

$$\begin{aligned}
\mathrm{E}[(b \cdot r)\overline{f(c \cdot r)}] &= \mathrm{E}\left[(b_1\bar{r}_1 + b_2\bar{r}_2)\frac{r_1}{T} \cdot \mathbf{1}_{\{|r_1| \le T\}}\right] \\
&\quad + \mathrm{E}\left[(b_1\bar{r}_1 + b_2\bar{r}_2)\frac{r_1}{|r_1|} \cdot \mathbf{1}_{\{|r_1| > T\}}\right] \\
&= \frac{1}{T}\mathrm{E}\left[b_1|r_1|^2 \cdot \mathbf{1}_{\{|r_1| \le T\}}\right] + \mathrm{E}\left[b_1|r_1| \cdot \mathbf{1}_{\{|r_1| > T\}}\right] \\
&= \frac{b_1}{T} \cdot \frac{1}{\pi} \int\limits_{x^2+y^2 \le T^2} \left(x^2 + y^2\right) \ \exp\left(-\left(x^2 + y^2\right)\right) \mathrm{d}x\mathrm{d}y \\
&\quad + \frac{b_1}{\pi} \int\limits_{x^2+y^2 > T^2} \sqrt{x^2 + y^2} \ \exp\left(-\left(x^2 + y^2\right)\right) \mathrm{d}x\mathrm{d}y \\
&= \frac{b_1}{\pi T} \int\limits_0^{2\pi} \int\limits_0^{T} \rho^3 \exp\left(-\rho^2\right) \ \mathrm{d}\rho\mathrm{d}\theta \\
&\quad + \frac{b_1}{\pi} \int\limits_0^{2\pi} \int\limits_T^{\infty} \rho^2 \exp\left(-\rho^2\right) \ \mathrm{d}\rho\mathrm{d}\theta \\
&= g(T)b_1
\end{aligned}$$

where the last equality follows from the facts

$$\int\limits_0^T \rho^3 \exp\left(-\rho^2\right)\mathrm{d}\rho = \frac{1}{2}\left(1 - \left(T^2 + 1\right)\exp\left(-T^2\right)\right)$$

and

$$\int\limits_T^\infty \rho^2 \exp\left(-\rho^2\right)\mathrm{d}\rho = \frac{1}{2}\left(T\exp\left(-T^2\right) + \sqrt{\pi}\left(1 - \Phi\left(\sqrt{2}T\right)\right)\right)$$

This completes the proof. □

Lemma 7 *For any pair of vectors $b, c \in C^n$ and $r \sim \mathcal{N}_C(0, I_n)$, we have:*

$$\mathrm{E}[f(c \cdot r)\overline{f(c \cdot r)}] = \frac{1}{T^2} - \frac{1}{T^2}\exp\left(-T^2\right)$$

Proof The proof is similar to that of Lemma 2. We again assume that $c = (1, 0, \ldots, 0)$. Then, we have

$$
\begin{aligned}
\mathrm{E}[f(c\cdot r)\overline{f(c\cdot r)}] &= \mathrm{E}\left[\frac{\bar{r}_1}{T}\frac{r_1}{T}\cdot \mathbf{1}_{\{|r_1|\le T\}}\right] + \mathrm{E}\left[\frac{\bar{r}_1}{|r_1|}\frac{r_1}{|r_1|}\cdot \mathbf{1}_{\{|r_1|>T\}}\right] \\
&= \frac{1}{T^2}\cdot\frac{1}{\pi}\int\limits_{x^2+y^2\le T^2}\left(x^2+y^2\right)\exp\left(-\left(x^2+y^2\right)\right)\,\mathrm{d}x\mathrm{d}y \\
&\quad + \frac{1}{\pi}\int\limits_{x^2+y^2>T^2}\exp\left(-\left(x^2+y^2\right)\right)\,\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{T^2}\cdot\frac{1}{\pi}\int\limits_0^{2\pi}\int\limits_0^{T}\rho^3\exp\left(-\rho^2\right)\,\mathrm{d}\rho\mathrm{d}\theta \\
&\quad + \frac{1}{\pi}\int\limits_0^{2\pi}\int\limits_T^{\infty}\rho\exp\left(-\rho^2\right)\,\mathrm{d}\rho\mathrm{d}\theta \\
&= \frac{1}{T^2}\left(1-\left(T^2+1\right)\exp\left(-T^2\right)\right)+\exp\left(-T^2\right) \\
&= \frac{1}{T^2}-\frac{1}{T^2}\exp\left(-T^2\right)
\end{aligned}
$$

□

Theorem 2 *If Q satisfies Assumption 1, then there exists a constant $C > 0$ such that*

$$
\mathrm{E}\left[\hat{z}^H Q\hat{z}\right] \ge \frac{1}{3\log(\beta)}\, w_{\mathrm{SDP}}
$$

where $\beta = \max\left\{5, \frac{\sum_{1\le k,m\le n}|q_{km}|}{C\sqrt{\sum_{1\le k,m\le n}|q_{km}|^2}}\right\}$.

Proof By Lemmas 1 and 6, we have

$$
\begin{aligned}
&\mathrm{E}[\{(b\cdot r)-Tf(b\cdot r)\}\overline{\{(c\cdot r)-Tf(c\cdot r)\}}] \\
&= (1-2Tg(T))(b\cdot c)+T^2\mathrm{E}[f(b\cdot r)\overline{f(c\cdot r)}]
\end{aligned}
$$

It follows that

$$\mathrm{E}\left[\hat{z}^H Q \hat{z}\right] = \sum_{k=1}^{n}\sum_{m=1}^{n} \frac{2Tg(T)-1}{T^2} q_{km}(v_k \cdot v_m) + \frac{1}{T^2}\sum_{k=1}^{n}\sum_{m=1}^{n} q_{km}\mathrm{E}[\{(v_k \cdot r) - Tf(v_k \cdot r)\}\overline{\{(v_m \cdot r) - Tf(v_m \cdot r)\}}]$$

Again, the quantity $\mathrm{E}[\{(b \cdot r) - Tf(b \cdot r)\}\overline{\{(c \cdot r) - Tf(c \cdot r)\}}]$ can be seen as an inner product of two vectors in a Hilbert space. Moreover, by letting $b = c$ and using Lemma 7, we know that the norm of an Euclidean unit vector in this Hilbert space is:

$$2 - 2Tg(T) - \exp\left(-T^2\right) = \exp\left(-T^2\right) - 2T\sqrt{\pi}\left(1 - \Phi\left(\sqrt{2}T\right)\right)$$

It follows that

$$\frac{1}{T^2}\sum_{k=1}^{n}\sum_{m=1}^{n} q_{km}\mathrm{E}[\{(v_k \cdot r) - Tf(v_k \cdot r)\} \cdot \overline{\{(v_m \cdot r) - Tf(v_m \cdot r)\}}] \geq -\frac{\exp\left(-T^2\right) - 2T\sqrt{\pi}\left(1 - \Phi\left(\sqrt{2}T\right)\right)}{T^2}\sum_{k=1}^{n}\sum_{m=1}^{n} |q_{km}|$$

On the other hand, by Lemma 4, we have $w_{\mathrm{SDP}} \geq C\sqrt{\sum_{1 \leq k,m \leq n} |q_{km}|^2} > 0$ for some constant $C > 0$. It follows that

$$\begin{aligned}\frac{1}{T^2}\sum_{k=1}^{n}\sum_{m=1}^{n} q_{km}\mathrm{E}[\{(v_k \cdot r) - Tf(v_k \cdot r)\} \cdot \overline{\{(v_m \cdot r) - Tf(v_m \cdot r)\}}] \\ \geq -\frac{\exp\left(-T^2\right) - 2T\sqrt{\pi}\left(1 - \Phi\left(\sqrt{2}T\right)\right)}{T^2} \cdot \frac{\sum_{1 \leq k,m \leq n} |q_{km}|}{C\sqrt{\sum_{1 \leq k,m \leq n} |q_{km}|^2}} \cdot w_{\mathrm{SDP}} \\ \geq -\frac{\exp\left(-T^2\right) - 2T\sqrt{\pi}\left(1 - \Phi\left(\sqrt{2}T\right)\right)}{T^2}\beta \cdot w_{\mathrm{SDP}}\end{aligned}$$

where $\beta = \max\left\{5, \frac{\sum_{1 \leq k,m \leq n} |q_{km}|}{C\sqrt{\sum_{1 \leq k,m \leq n} |q_{km}|^2}}\right\}$. This implies that

$$\begin{aligned}\mathrm{E}\left[\hat{z}^H Q \hat{z}\right] &\geq \left(\frac{2Tg(T)-1}{T^2} - \frac{\exp\left(-T^2\right) - 2T\sqrt{\pi}\left(1 - \Phi\left(\sqrt{2}T\right)\right)}{T^2} n\right) w_{\mathrm{SDP}} \\ &\geq \frac{1 - (2+\beta)\exp\left(-T^2\right)}{T^2}\, w_{\mathrm{SDP}}\end{aligned}$$

By letting $T = \sqrt{2\log\beta}$, we have $\mathrm{E}\left[\hat{z}^H Q \hat{z}\right] \geq \frac{1}{3\log\beta} w_{\mathrm{SDP}}$. □

Notice that by the Cauchy–Schwarz inequality, we have

$$\frac{\sum_{1\le k,m\le n} |q_{km}|}{\sqrt{\sum_{1\le k,m\le n} |q_{km}|^2}} \leq \sqrt{n}$$

This yields the following corollary.

Corollary 3 *If Q satisfies Assumption 1, then* $\mathrm{E}\left[\hat{z}^H Q \hat{z}\right] \geq \Omega\left(\frac{1}{\log n}\right) \cdot w_{\mathrm{SDP}}$.

Remarks Corollary 3 can also be derived from the result of [10]. Indeed, since the diagonal entries of Q are all zeros, Problem (12), which is equivalent to Problem (4), is also equivalent to the following problem:

$$\begin{array}{lll} \text{maximize} & u^{\mathrm{T}} A u & \\ \text{subject to} & u^{\mathrm{T}} A_j u \leq 1 & j = 1, 2, \ldots, n \\ & u \in R^{2n} & \end{array} \tag{16}$$

Here, we have $u = (x, y)$,

$$A = \begin{pmatrix} \mathrm{Re}(Q) & \mathrm{Im}(Q) \\ -\mathrm{Im}(Q) & \mathrm{Re}(Q) \end{pmatrix}, \qquad A_j = \begin{pmatrix} I_j & 0 \\ 0 & I_j \end{pmatrix}$$

where $I_j \in R^{n\times n}$ is the matrix whose jth diagonal entry is 1 and all other entries zero. In particular, we have $\sum_{j=1}^{n} A_j = I \succ 0$ and $\mathrm{rank}(A_j) = 2$ for all $j = 1, \ldots, n$. Hence, we may apply the result of [10] and conclude that we can find a feasible solution $\hat{u}$ to (16) in polynomial time whose value is at least $\Omega(1/\log n)$ times the optimum. This in turn implies an $\Omega(1/\log n)$ approximation to the original problem.

5 Conclusion

We have studied a class of discrete and continuous quadratic optimization problems in the complex Hermitian form and obtained good approximation guarantees for these problems. The techniques we used in analyzing the SDP relaxations suggest that they may have a wider applicability, and it would be interesting to explore those possibilities. In addition, it would be interesting to refine our analysis and obtain better approximation guarantees for the discrete quadratic optimization problem considered in this paper.

Acknowledgments The first two authors would like to thank Shuzhong Zhang for sharing the results in [16], which motivated this work. We also thank him and Yongwei Huang for discussions on this topic, and the referees for their valuable comments.

References

1. Alon, N., Makarychev, K., Makarychev, Y., Naor, A.: Quadratic forms on graphs. In: Proceedings of the 37th annual ACM symposium on theory of computing, pp. 486–493 (2005)
2. Alon, N., Naor, A.: Approximating the Cut–Norm via Grothendieck's inequality. In: Proceedings of the 36th annual ACM symposium on theory of computing, pp. 72–80 (2004)
3. Ben-Tal, A., Nemirovski, A., Roos, C.: Extended matrix cube theorems with applications to μ–theory in control. Math. Oper. Res. **28**(3), 497–523 (2003)
4. Charikar, M., Wirth, A.: Maximizing quadratic programs: extending Grothendieck's inequality. In: Proceedings of the 45th annual IEEE symposium on foundations of computer science, pp. 54–60 (2004)
5. Frieze, A., Jerrum, M.: Improved approximation algorithms for Max *k*–Cut and Max bisection. Algorithmica **18**, 67–81 (1997)
6. Goemans, M.X., Rendl, F.: Combinatorial optimization. In: Wolkowicz, H., Saigal, R., Vandenberghe, L. (eds.) Handbook of semidefinite programming: theory, algorithms and applications, pp. 343–360. Kluwer, Dordrecht (2000)
7. Goemans, M.X., Williamson, D.P.: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J. ACM **42**(6), 1115–1145 (1995)
8. Goemans, M.X., Williamson, D.P.: Approximation algorithms for Max–3–Cut and other problems via complex semidefinite programming. J. Comput. Syst. Sci. **68**(2), 442–470 (2004)
9. Hochbaum, D.S. (ed.): Approximation algorithms for NP-hard problems. PWS Publishing Company, (1997)
10. Nemirovski, A., Roos, C., Terlaky, T.: On maximization of quadratic form over intersection of ellipsoids with common center. Math. Prog. Ser. A **86**, 463–473 (1999)
11. Nesterov, Y.: Global quadratic optimization via conic relaxation, CORE Discussion Paper 9860. Université Catholique de Louvain (1998)
12. Rietz, R.E.: A proof of the Grothendieck inequality. Israel J. Math. **19**, 271–276 (1974)
13. Toker, O., Özbay, H.: On the complexity of purely complex μ computation and related problems in multidimensional systems. IEEE Trans Automat Control **43**(3), 409–414 (1998)
14. Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM Rev **38**(1), 49–95 (1996)
15. Ye, Y.: Approximating quadratic programming with bound and quadratic constraints. Math. Prog. **84**, 219–226 (1999)
16. Zhang, S., Huang, Y.: Complex quadratic optimization and semidefinite programming. Technical Report SEEM 2004–03, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, 2004. SIAM J. Optim. **16**(3), 871–890 (2006)