Math. Program., Ser. A (2009) 120:479–495
DOI 10.1007/s10107-008-0223-z

FULL LENGTH PAPER

On the copositive representation of binary and continuous nonconvex quadratic programs

Samuel Burer

Received: 27 November 2006 / Accepted: 26 March 2008 / Published online: 29 April 2008

Abstract In this paper, we model any nonconvex quadratic program having a mix of binary and continuous variables as a linear program over the dual of the cone of copositive matrices. This result can be viewed as an extension of earlier separate results, which have established the copositive representation of a small collection of NP-hard problems. A simplification, which reduces the dimension of the linear conic program, and an extension to complementarity constraints are established, and computational issues are discussed.

Mathematics Subject Classification (2000) 90C25 · 90C26 · 90C20

1 Introduction

A recent line of research has shown that several NP-hard optimization problems can be expressed as linear programs over the convex cone of copositive matrices (called *copositive programs*, or *COPs*). In particular, as far as we are aware, the following is a complete list of all problems known to have representations as COPs: standard quadratic programming [3], the maximum stable set problem [7], and quadratic programs over transportation matrices satisfying specialized quadratic constraints [15]. (Transportation matrices are element-wise nonnegative with pre-specified row- and column-sums.) The last class of problems includes, for example, a certain minimum-cut graph tri-partitioning problem [17] and the quadratic assignment problem [16]. In each case, the COP has a natural associated dual linear program over the dual cone of completely positive matrices (called a *completely positive program*, or *CPP*).

Research partially supported by NSF Grant CCF-0545514.

S. Burer (✉)
Department of Management Sciences, University of Iowa, Iowa City, IA 52242-1994, USA
e-mail: samuel-burer@uiowa.edu

While expressing a problem as a COP or CPP does not resolve the difficulty of the problem, it does allow one to "package" the difficulty completely inside the convex cones of copositive and/or completely positive matrices. Whatever knowledge is then gained about copositive or completely positive matrices can be applied uniformly to multiple types of NP-hard problems. For example, Parrilo [13] (see also de Klerk and Pasechnik [7]) discusses a hierarchy of linear- and semidefinite-representable cones approximating the copositive cone from the inside. More precisely, letting $\mathcal{C}_q$ denote the cone of $q \times q$ copositive matrices, there exist closed convex cones $\{\mathcal{K}_q^r : r = 0, 1, 2, \ldots\}$ such that $\mathcal{K}_q^r \subset \mathcal{K}_q^{r+1}$ for all $r \geq 0$ and $\overline{\cup_{r\geq 0}\mathcal{K}_q^r} = \mathcal{C}_q$. The dual cones $\{(\mathcal{K}_q^r)^* : r = 0, 1, 2, \ldots\}$ approximate the completely positive cone $\mathcal{C}_q^*$ from the outside, i.e., $(\mathcal{K}_q^r)^* \supset (\mathcal{K}_q^{r+1})^*$ for all $r \geq 0$ and $\cap_{r\geq 0}(\mathcal{K}_q^r)^* = \mathcal{C}_q^*$. Explicit representations of the approximating cones have been worked out by Parrilo [13] and Bomze and de Klerk [2]. For example, $\mathcal{K}_q^0$ is the cone of all symmetric matrices that can be written as the sum of a positive semidefinite matrix and an element-wise nonnegative matrix, and $(\mathcal{K}_q^0)^*$ is the cone of all symmetric matrices that are simultaneously positive semidefinite and element-wise nonnegative. Moreover, using these approximation ideas in the case of standard quadratic programming, Bomze and de Klerk [2] prove bounds on the approximation error as a function of r. This line of reasoning also leads to a polynomial-time approximation scheme for standard quadratic programming (and an extension by de Klerk et al. [6]).

Other than the handful of problems listed above, what types of problems can be represented as COPs or as CPPs? In Sect. 2 of this paper, we consider the extremely large class of nonconvex quadratic programs, having any mixture of binary and continuous variables, and prove that they can be expressed as CPPs. By "quadratic program," we mean any program with a quadratic objective and linear constraints (other than the binary conditions on the discrete variables). In Sect. 3, we show that it is often possible to reduce the dimension of the completely positive representation and also establish an extension to complementarity constraints over bounded variables. Then, in Sect. 4, we compare our approach with earlier results, highlighting key similarities and differences, and in Sect. 5, we discuss some computational issues. In Sect. 6, we conclude with a few final remarks.

1.1 Definitions and notation

The closed, full-dimensional convex cone of $q \times q$ *copositive matrices* is defined as

$$\mathcal{C}_q := \left\{ X \in \Re^{q\times q} : X = X^T,\ v^T X v \geq 0 \;\forall\; v \in \Re^q_+ \right\},$$

and its dual is the closed, full-dimensional convex cone of $q \times q$ *completely positive matrices*:

$$\mathcal{C}_q^* := \left\{ X \in \Re^{q\times q} : X = \sum_{k\in K} z^k (z^k)^T \quad \text{for some finite } \{z^k\}_{k\in K} \subset \Re^q_+ \backslash \{0\} \right\} \cup \{0\}.$$

Note that the decomposition of nonzero $X \in \mathcal{C}_q^*$ defines $z^k \neq 0$ for all $k \in K$. A (*linear*) *copositive program (COP)* is any optimization over $X \in \mathcal{C}_q$ whose objective and constraints are linear in X. Associated with every copositive program is a dual problem over the cone of completely positive matrices, which is a (*linear*) *completely positive program* (*CPP*). In general, satisfaction of Slater's condition in either the primal or dual is needed to guarantee that there is no duality gap between such dual problems.

We use $\mathcal{S}_q$ to denote the cone of $q \times q$ positive semidefinite matrices. The generic notation Feas$(\cdot)$ and opt$(\cdot)$ will be used to denote the feasible set and optimal value of a given problem. Conv$(\cdot)$ will be used to denote the convex hull of a given set, and Cone$(\cdot)$ indicates the conic hull. The notation e_q will be used to denote the q-length vector of all ones; if the dimension q is clear from the context, then we will drop the subscript. The matrix I represents the identity matrix of the appropriate dimension. Also, Diag(v) is the diagonal matrix whose diagonal is v. For a matrix M, vec(M) is the vector gotten by stacking the columns of M in order. Given two conformal matrices A and B, $A \bullet B := \text{trace}(A^T B)$.

2 The problem and its completely positive representation

The general-form problem we consider is

$$\begin{aligned} \min \quad & x^T Q x + 2\, c^T x \\ \text{s.t.} \quad & a_i^T x = b_i \quad \forall i = 1, \ldots, m \\ & x \geq 0 \\ & x_j \in \{0, 1\} \quad \forall j \in B \end{aligned} \tag{P}$$

where $x \in \Re^n_+$ and $B \subseteq \{1, \ldots, n\}$. We assume Feas$(P) \neq \emptyset$. Clearly, any problem as described in the third paragraph of the Introduction ("nonconvex quadratic programs, having any mixture of binary and continuous variables") can be put in the form of (P), and so (P) encompasses a huge class of NP-hard problems, e.g., all 0-1 linear integer programs, all nonconvex (continuous) quadratic programs, and all unconstrained 0-1 quadratic programs. As it will turn out, a distinguishing feature of (P) is that all general linear constraints are represented as equations (as opposed to inequalities).

We make a particular assumption (referred to as the *key assumption*) regarding the linear portion $L := \{x \geq 0 \colon a_i^T x = b_i \ \forall\, i = 1, \ldots, m\}$ of Feas(P):

$$x \in L \implies 0 \leq x_j \leq 1 \quad \forall\, j \in B. \tag{1}$$

If $B = \emptyset$, then the key assumption is vacuous. On the other hand, if $B \neq \emptyset$ and the key assumption is not already implied, then one can achieve it without loss of generality. For example, one could augment (P) by adding the constraint $x_j + s_j = 1$, where $s_j \geq 0$ is a slack variable, for all $j \in B$. Related to the key assumption, we also define the recession cone of L, namely $L_\infty := \left\{d \geq 0 : a_i^T d = 0 \ \forall\, i = 1, \ldots, m\right\}$, which has the following property from (1):

$$d \in L_\infty \implies d_j = 0 \quad \forall j \in B. \tag{2}$$

Standard techniques, including the relaxation of the rank-1 matrix $(1; x)(1; x)^T$ to the symmetric matrix $(1, x^T; x, X)$, yield the following CPP, which is a relaxation of (P):

$$\begin{aligned} \min \quad & Q \bullet X + 2\, c^T x \\ \text{s.t.} \quad & a_i^T x = b_i \quad \forall i = 1, \dots, m \\ & a_i^T X a_i = b_i^2 \quad \forall i = 1, \dots, m \\ & x_j = X_{jj} \quad \forall j \in B \\ & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathcal{C}^*_{1+n}. \end{aligned} \tag{C}$$

In comparison with (P), (C) contains $\mathcal{O}(n^2)$ number of variables but only twice as many general equations. Of course, the difficulty in (C) is the completely positive constraint.

Define the following convex sets of symmetric matrices of size $(1+n) \times (1+n)$:

$$\mathrm{Feas}^+(C) := \left\{ \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} : (x, X) \in \mathrm{Feas}(C) \right\},$$

$$\mathrm{Feas}^+(P) := \mathrm{Conv} \left\{ \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T : x \in \mathrm{Feas}(P) \right\},$$

$$L_\infty^+ := \mathrm{Cone} \left\{ \begin{pmatrix} 0 \\ d \end{pmatrix} \begin{pmatrix} 0 \\ d \end{pmatrix}^T : d \in L_\infty \right\}.$$

We now establish the following fundamental result.

Proposition 2.1 $\mathrm{Feas}^+(P) \subseteq \mathrm{Feas}^+(C) = \mathrm{Feas}^+(P) + L_\infty^+$.

The inclusion $\mathrm{Feas}^+(P) \subseteq \mathrm{Feas}^+(C)$ holds by construction.

Next, we examine L_∞^+ and its relationship to the recession cone of $\mathrm{Feas}^+(C)$, which equals

$$\left\{ \begin{pmatrix} 0 & d^T \\ d & D \end{pmatrix} \in \mathcal{C}^*_{1+n} : \begin{array}{ll} a_i^T d = 0 & i = 1, \dots, m \\ a_i^T D a_i = 0 & i = 1, \dots, m \\ d_j = D_{jj} & j \in B \end{array} \right\}.$$

Because $(0, d^T; d; D) \in \mathcal{C}^*_{1+n}$, it follows that $d = 0$ and so the recession cone simplifies to

$$\left\{ \begin{pmatrix} 0 & 0^T \\ 0 & D \end{pmatrix} \in \mathcal{C}^*_{1+n} : \begin{array}{ll} a_i^T D a_i = 0 & i = 1, \dots, m \\ D_{jj} = 0 & j \in B \end{array} \right\}.$$

From (2), this implies that L_∞^+ is contained in the recession cone, and hence $\text{Feas}^+(C) \supseteq \text{Feas}^+(P) + L_\infty^+$.

To show the remaining "$\subseteq$" inclusion, let $(x, X) \in \text{Feas}(C)$ and consider a completely positive decomposition

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \sum_{k \in K} \begin{pmatrix} \zeta_k \\ z^k \end{pmatrix} \begin{pmatrix} \zeta_k \\ z^k \end{pmatrix}^T, \tag{3}$$

where $(\zeta_k; z^k) \in \Re_+^{1+n}$ for all $k \in K$.

Lemma 2.2 *Let* $(x, X) \in \text{Feas}(C)$. *Given the decomposition* (3), *define* $K_+ := \{k \in K : \zeta_k > 0\}$ *and* $K_0 := \{k \in K : \zeta_k = 0\}$. *Then*:

(i) $z^k/\zeta_k \in L$ *for all* $k \in K_+$;
(ii) $z^k \in L_\infty$ *for all* $k \in K_0$.

Proof Directly from (3), we see

$$\sum_{k \in K} \zeta_k^2 = 1. \tag{4}$$

Moreover, since $a_i^T x = b_i$ and $a_i^T X a_i = b_i^2$, we have

$$b_i = \sum_{k \in K} \zeta_k (a_i^T z^k) \quad \text{and} \quad b_i^2 = \sum_{k \in K} \left(a_i^T z^k\right)^2 \tag{5}$$

for all $i = 1, \ldots, m$. Thus,

$$\left(\sum_{k \in K} \zeta_k \left(a_i^T z^k\right)\right)^2 = \left(\sum_{k \in K} \zeta_k^2\right)\left(\sum_{k \in K} \left(a_i^T z^k\right)^2\right),$$

and so by the equality-case of the Cauchy–Schwarz inequality, it follows that there exist δ_i $(i = 1, \ldots, m)$ such that

$$\delta_i \, \zeta_k = a_i^T z^k \quad \forall k \in K, \; i = 1, \ldots, m. \tag{6}$$

The second statement of the proposition follows directly from (6) and the fact that $\zeta_k = 0$ for all $k \in K_0$. To prove the first statement, it suffices to show $\delta_i = b_i$ for all i. Indeed, the first equality of (5), (6), and (4) imply

$$b_i = \sum_{k \in K} \zeta_k \left(a_i^T z^k\right) = \sum_{k \in K} \zeta_k (\delta_i \, \zeta_k) = \delta_i.$$

$\square$

By taking $\lambda_k := \zeta_k^2$ and $v^k := z^k/\zeta_k$ for all $k \in K_+$, we can write the completely positive decomposition (3) in the more convenient form

$$\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} = \sum_{k \in K_+} \lambda_k \begin{pmatrix} 1 \\ v^k \end{pmatrix} \begin{pmatrix} 1 \\ v^k \end{pmatrix}^T + \sum_{k \in K_0} \begin{pmatrix} 0 \\ z^k \end{pmatrix} \begin{pmatrix} 0 \\ z^k \end{pmatrix}^T, \tag{7}$$

where $\lambda_k > 0$ for all $k \in K_+$, $\sum_{k \in K_+} \lambda_k = 1$, and $v^k \in L$ for all $k \in K_+$. To establish the "$\subseteq$" inclusion, it remains only to show $v_j^k \in \{0, 1\}$ for all $j \in B$.

Lemma 2.3 *Let* $(x, X) \in \text{Feas}(C)$. *Then, in the representation* (7), $v_j^k \in \{0, 1\}$ *for all* $k \in K_+$ *and all* $j \in B$.

Proof Fixing $j \in B$, (1) and (2) imply $0 \le v_j^k \le 1$ for all $k \in K_+$ and $z_j^k = 0$ for all $k \in K_0$. Further, the equality $x_j = X_{jj}$ implies

$$\sum_{k \in K_+} \lambda_k v_j^k = \sum_{k \in K_+} \lambda_k \left(v_j^k\right)^2 + \sum_{k \in K_0} \left(z_j^k\right)^2 \iff \sum_{k \in K_+} \lambda_k \left(v_j^k - \left(v_j^k\right)^2\right) = 0.$$

The last equation expresses that the sum of nonnegative numbers is 0, and so $\lambda_k \left(v_j^k - (v_j^k)^2\right) = 0$ for all $k \in K_+$. Since $\lambda_k > 0$, we conclude $v_j^k = \left(v_j^k\right)^2$. □

This completes the proof of Proposition 2.1.

Because all matrices in the cone L_∞^+ have zeros in the 0-th row and column, Proposition 2.1 also gives rise to the following characterization.

Corollary 2.4 $\text{Conv}(\text{Feas}(P)) = \{x : (x, X) \in \text{Feas}(C)$ *for some* $X\}$.

Said differently, the convex hull of Feas(P) is the projection of Feas(C) onto the coordinates corresponding to x. Moreover, in the case when $L_\infty = \{0\}$, i.e., when Feas(P) is bounded, we can make a slightly stronger statement.

Corollary 2.5 *If* Feas(P) *is bounded, then* $\text{Feas}^+(P) = \text{Feas}^+(C)$.

Note that Feas(P) is bounded if and only if L is bounded due to (1) and (2).

We are now prepared to state and prove the main theorem relating (P) and (C).

Theorem 2.6 (C) *is equivalent to* (P), *i.e.*: (i) $\text{opt}(C) = \text{opt}(P)$; (ii) *if* (x^*, X^*) *is optimal for* (C), *then* x^* *is in the convex hull of optimal solutions for* (P).

Let

$$\ell(Y) := \begin{pmatrix} 0 & c^T \\ c & Q \end{pmatrix} \bullet Y;$$

note that ℓ is linear in Y. Then we can express opt(P) and opt(C) as follows:

$$\text{opt}(P) = \min_{Y \in \text{Feas}^+(P)} \ell(Y) \quad \text{and} \quad \text{opt}(C) = \min_{Y \in \text{Feas}^+(C)} \ell(Y).$$

Hence, the inclusion of Proposition 2.1 immediately implies $\text{opt}(P) \geq \text{opt}(C)$. If $\text{opt}(P) = -\infty$, then $\text{opt}(C) = -\infty$, and item (i) of the theorem follows. In case $\text{opt}(P) \neq -\infty$, to prove $\text{opt}(P) \leq \text{opt}(C)$ and hence item (i), it suffices to use the equation of Proposition 2.1 in conjunction with the following lemma.

Lemma 2.7 *If* $\text{opt}(P) \neq -\infty$*, then* ℓ *is nonnegative over* L_∞^+.

Proof We prove the contrapositive. Suppose ℓ is negative somewhere on L_∞^+. Then, since ℓ is linear, there exists some $0 \neq d \in L_\infty$ such that

$$\ell\left(\begin{pmatrix} 0 \\ d \end{pmatrix} \begin{pmatrix} 0 \\ d \end{pmatrix}^T\right) = d^T Q d < 0.$$

This shows that d is a negative direction of recession for (P), i.e., $\text{opt}(P) = -\infty$. □

Item (ii) of Theorem 2.6 is proved by examining the representation (7) for (x^*, X^*). In such a case, we must have v^k optimal for (P) for all $k \in K_+$; otherwise, $Q \bullet X^* + 2\, c^T x^*$ would not equal $\text{opt}(C) = \text{opt}(P)$. (In fact, we know $(z^k)^T Q z^k = 0$ for all $k \in K_0$.) Since $x^* = \sum_{k \in K_+} \lambda_k v^k$, this proves item (ii).

In addition to Theorem 2.6, we remark that, when $\text{opt}(P) \neq -\infty$, the optimal value of (C) is actually attained because the optimal value of (P) is attained.

3 A simplification and an extension

3.1 Eliminating x from the completely positive representation

We now show that it is often possible to eliminate the variable x in (C) and restrict the completely positive optimization to $\mathcal{C}_n^*$ instead of $\mathcal{C}_{1+n}^*$. The key property required of (P) is the following:

$$\exists \ y \in \Re^m \text{ s.t. } \sum_{i=1}^m y_i a_i \geq 0, \ \sum_{i=1}^m y_i b_i = 1. \tag{8}$$

In this subsection, we assume (8) and define

$$\alpha := \sum_{i=1}^m y_i a_i \geq 0. \tag{9}$$

Note that, by Farkas's lemma, (8) holds if and only if the set $\{z \geq 0 : a_i^T z = -b_i \ \forall i = 1, \ldots, m\}$ is empty. Note also, that if the sum of any subset of variables of $x \in \text{Feas}(P)$ can be bounded explicitly, then (8) can be achieved by adding a slack variable and a new redundant equation to (P). For example, if $x \in \text{Feas}(P)$ implies $x_1 + x_3 \leq 2.5$, then adding the constraint $x_1 + x_3 + s = 2.5$ enables (8) by taking y to have a weight of 0.4 on the new constraint and 0 on all others.

A direct consequence of (8) is that $\alpha^T x = 1$ is redundant for (P). So we consider

$$\begin{array}{ll} \min & x^T Qx + 2\,c^T x \\ \text{s.t.} & x \in \text{Feas}(P) \\ & \alpha^T x = 1 \end{array}$$

as well as its completely positive representation

$$\begin{array}{ll} \min & Q \bullet X + 2\,c^T x \\ \text{s.t.} & (x, X) \in \text{Feas}(C) \\ & \alpha^T x = 1 \\ & \alpha^T X\alpha = 1, \end{array} \tag{10}$$

which is equivalent to (C). The construction of (C) from (P) and Proposition 2.1 also make it clear that the constraint $X\alpha = x$ is redundant for (C) and (10), which means that (10) is equivalent to the reduced problem

$$\begin{array}{ll} \min & Q \bullet X + 2\,c^T X\alpha \\ \text{s.t.} & (X\alpha, X) \in \text{Feas}(C) \\ & \alpha^T X\alpha = 1. \end{array} \tag{11}$$

Finally, the equations $\alpha^T X\alpha = 1$ and

$$\begin{pmatrix} 1 & \alpha^T X \\ X\alpha & X \end{pmatrix} = \begin{pmatrix} \alpha & I \end{pmatrix}^T X \begin{pmatrix} \alpha & I \end{pmatrix}$$

along with $\alpha \geq 0$ imply that

$$X \in \mathcal{C}_n^* \quad \Longrightarrow \quad \begin{pmatrix} 1 & \alpha^T X \\ X\alpha & X \end{pmatrix} \in \mathcal{C}_{1+n}^*.$$

Moreover, the converse holds because principal submatrices of completely positive matrices are completely positive. Hence, (11) is in turn equivalent to

$$\begin{array}{lll} \min & Q \bullet X + 2\,c^T X\alpha & \\ \text{s.t.} & a_i^T X\alpha = b_i & \forall i = 1, \ldots, m \\ & a_i^T X a_i = b_i^2 & \forall i = 1, \ldots, m \\ & [X\alpha]_j = X_{jj} & \forall j \in B \\ & X \in \mathcal{C}_n^* & \\ & \alpha^T X\alpha = 1. & \end{array} \tag{C'}$$

Overall, we conclude that (C) is equivalent to (C'), as stated in the following theorem.

Theorem 3.1 *Suppose* (8) *holds, and define* α *as in* (9)*. Then* (C') *is equivalent to* (P)*, i.e.:* (i) $\text{opt}(C') = \text{opt}(P)$*;* (ii) *if* X^* *is optimal for* (C')*, then* $X^*\alpha$ *is in the convex hull of optimal solutions for* (P)*.*

Besides a more compact representation, an additional benefit of (C') over (C) is that $\text{Feas}(C')$ may have an interior, whereas $\text{Feas}(C)$ never does. This is an important feature of (C') since the existence of an interior is generally a benefit both theoretically and computationally. Let us consider an interior solution $(x, X) \in \text{Feas}(C)$. It satisfies the additional condition that $(1, x^T; x, X)$ is an element of $\text{int}(\mathcal{C}^*_{1+n})$. A fundamental fact, which is clear from the definition of the completely positive matrices, is $\mathcal{C}^*_q \subseteq \mathcal{S}_q$, where $\mathcal{S}_q$ is the set of all $q \times q$ positive semidefinite matrices, and so $\text{int}(\mathcal{C}^*_q) \subseteq \text{int}(\mathcal{S}_q)$. However, $(x, X) \in \text{Feas}(C)$ implies

$$\begin{pmatrix} b_i \\ -a_i \end{pmatrix}^T \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \begin{pmatrix} b_i \\ -a_i \end{pmatrix} = 0 \implies \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \notin \text{int}\,(\mathcal{S}_{1+n}) \supseteq \text{int}\,\left(\mathcal{C}^*_{1+n}\right),$$

establishing that (C) does not have an interior. (This argument also shows that the lack of interior for (C) will also hold for any semidefinite relaxation.) On the other hand, since (C') optimizes over $\mathcal{C}^*_n$ instead of $\mathcal{C}^*_{1+n}$, it certainly may have an interior. For example, consider $\text{Feas}(P) = \{x \geq 0 : e^T x = 1\}$ (see also Sect. 4.1). In this case, (8) holds with $\alpha = e$, and $\text{Feas}(C') = \{X \in \mathcal{C}^*_n : e^T X e = 1\}$ has an interior since $\text{Cone}(\text{Feas}(C')) = \mathcal{C}^*_n$. In particular, $\frac{1}{n} I$ is known to be an element of the interior.

In addition to eliminating x, one may also wonder if it is possible to eliminate the constraints $a_i^T X\alpha = b_i$ in $\text{Feas}(C')$, which are in some sense the last remnants of x. One example when it is possible has just been demonstrated in the previous paragraph. In general, however, it is not possible as the following example shows. (This specific example was first considered by Bomze et al. [3].) Let $\kappa \in (0, 1)$, and consider $\text{Feas}(P) = \{(x_1, x_2) \geq 0 : x_1 = \kappa, x_1 + x_2 = 1\}$ and

$$\text{Feas}(C') = \left\{ \begin{pmatrix} X_{11} & X_{21} \\ X_{21} & X_{22} \end{pmatrix} \in \mathcal{C}^*_2 : \begin{array}{l} X_{11} + X_{21} = \kappa \\ X_{11} = \kappa^2 \\ X_{11} + 2X_{21} + X_{22} = 1 \end{array} \right\}.$$

Clearly, $\text{Feas}(P) = \{(\kappa, 1 - \kappa)\}$, and

$$\text{Feas}(C') = \left\{ \begin{pmatrix} \kappa^2 & \kappa(1-\kappa) \\ \kappa(1-\kappa) & (1-\kappa)^2 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \kappa \\ 1-\kappa \end{pmatrix} \begin{pmatrix} \kappa \\ 1-\kappa \end{pmatrix}^T \right\},$$

as predicted. However, the set

$$\left\{ \begin{pmatrix} X_{11} & X_{21} \\ X_{21} & X_{22} \end{pmatrix} \in \mathcal{C}^*_2 : \begin{array}{l} X_{11} = \kappa^2 \\ X_{11} + 2X_{21} + X_{22} = 1 \end{array} \right\}$$

contains additional points, e.g.,

$$\begin{pmatrix} \kappa^2 & 0 \\ 0 & 1-\kappa^2 \end{pmatrix}.$$

3.2 Complementarity constraints on bounded variables

A natural question is whether the techniques of Sect. 2 can be extended to represent quadratically constrained quadratic programs as CPPs. Indeed, at least one important case, namely that of complementarity constraints on bounded variables, can be handled. We proceed by briefly stating a somewhat general extension and then immediately specializing it to complementarity constraints.

Consider the following extension of (P):

$$\begin{aligned} \min \quad & x^T Q x + 2\, c^T x \\ \text{s.t.} \quad & x \in \text{Feas}(P) \\ & x^T \bar{Q} x + 2\, \bar{c}^T x + \bar{\kappa} = 0, \end{aligned} \qquad (\bar{P})$$

where $\bar{Q}$ is symmetric. Let $\bar{B}$ be the index set of variables x_j involved in the quadratic constraint, i.e., $\bar{B} := \{j : \bar{Q}_{\cdot j} \neq 0 \text{ or } c_j \neq 0\}$. Although a constraint of the form $x_j \in \{0, 1\}$ has already been included in Feas(P), its representation as $x_j - x_j^2 = 0$ serves as a good example for a quadratic equation such as the above. Define L and L_∞ as before and assume

$$x \in L \implies x^T \bar{Q} x + 2\, \bar{c}^T x + \bar{\kappa} \geq 0 \quad \text{and} \quad d \in L_\infty \implies d_j = 0 \quad \forall j \in \bar{B}.$$

In the case of $x_j - x_j^2 = 0$, these assumptions are implied by (1) and (2). In fact, one can easily check that these assumptions are natural generalizations of (1) and (2) and enable the extension of Proposition 2.1, Lemmas 2.3, and 2.7 to $(\bar{P})$ (note that Lemma 2.2 is unchanged), thus enabling the extension of Theorem 2.6 to $(\bar{P})$. The resulting form of the equivalent CPP is

$$\begin{aligned} \min \quad & Q \bullet X + 2\, c^T x \\ \text{s.t.} \quad & (x, X) \in \text{Feas}(C) \\ & \bar{Q} \bullet X + 2\, \bar{c}^T x + \bar{\kappa} = 0, \end{aligned} \qquad (\bar{C})$$

Additional quadratic constraints can be included in $(\bar{P})$ and $(\bar{C})$ as long as each satisfies the two assumptions.

Now we consider complementarity constraints. Let E be a collection of pairs (j, k) indicating the complementarities, e.g., $x_j x_k = 0$ for all $(j, k) \in E$, and let $B^c = \{j : (j, k) \in E \text{ for some } k\}$ be the set of all indices of variables involved in the complementarity constraints. The extension of (P) is as follows:

$$\begin{aligned} \min \quad & x^T Q x + 2\, c^T x \\ \text{s.t.} \quad & x \in \text{Feas}(P) \\ & \sum_{(j,k)\in E} x_j x_k = 0. \end{aligned} \qquad (P^c)$$

The corresponding CPP is

$$\begin{array}{ll} \min & Q \bullet X + 2\, c^T x \\ \text{s.t.} & (x, X) \in \text{Feas}(C) \\ & \displaystyle\sum_{(j,k) \in E} X_{jk} = 0. \end{array} \tag{C^c}$$

Related to the assumptions of the previous paragraph, note that $x \geq 0$ ensures $\sum_{(j,k)\in E} x_j x_k \geq 0$ for all $x \in L$. To guarantee $d \in L_\infty \Longrightarrow d_j = 0$ for all $j \in B^c$, some additional knowledge on L is required. For example, if $x \in L$ implies that x_j is bounded for all $j \in B^c$, then the condition is satisfied. We have the following theorem.

Theorem 3.2 *For each $j \in B^c$, suppose the variable x_j is bounded in L. Then (C^c) is equivalent to (P^c), i.e.*: (i) $\text{opt}(C^c) = \text{opt}(P^c)$; (ii) *if (x^*, X^*) is optimal for (C^c), then x^* is in the convex hull of optimal solutions for (P^c).*

One interesting application of Theorems 2.6 and 3.2 is the following. Consider the case when $B = \emptyset$. Then it is known that (P) can be formulated as a linear program with complementarity constraints via its first-order KKT conditions [9], although care must be taken to bound the complementary variables (i.e., those with $j \in B^c$; see Burer and Vandenbussche [5]). Hence, Theorems 2.6 and 3.2 provide two different CPPs for the same problem.

4 Previous results: description and comparison

4.1 Standard quadratic programming

As far as we are aware, Bomze et al. [3] have established the first copositive representation of an NP-hard problem. In particular, they study quadratic optimization over the simplex (called *standard quadratic programming*), i.e., the special case of (P) with $m = 1$, $B = \emptyset$, and $(a_1, b_1) = (e, 1)$. Without loss of generality, they take $c = 0$ because, if $c \neq 0$, then c can be absorbed into the quadratic part by optimizing with $(\tilde{Q}, \tilde{c}) = (Q + ec^T + ce^T, 0)$, which has the same objective as (Q, c) over the simplex. They also replace $e^T x = 1$ with the equivalent $e^T x x^T e = 1$ and subsequently study $\min\{x^T Q x : e^T x x^T e = 1, x \geq 0\}$. The CPP that Bomze et al. consider is

$$\min \left\{ Q \bullet X : e^T X e = 1,\ X \in \mathcal{C}_n^* \right\}, \tag{12}$$

which is equivalent to (C') in this case. Establishing the special case of Theorem 3.1 then requires the following result: the extreme points of (12) are precisely the rank-1 matrices xx^T with $e^T x = 1$ and $x \geq 0$.

Working in a somewhat different direction, Sturm and Zhang [18] examine the idea of generalized copositive and completely positive cones based on a general domain $D \subseteq \Re^q$:

$$\mathcal{C}_q(D) := \left\{ X \in \Re^{q\times q} : X = X^T,\ v^T X v \geq 0\ \forall v \in D \right\},$$

$$\mathcal{C}_q^*(D) := \left\{ X \in \Re^{q\times q} : X = \sum_{k\in K} z^k (z^k)^T \quad \text{for some finite } \{z^k\}_{k\in K} \subset D\backslash\{0\} \right\} \cup \{0\}.$$

For example, if $D = \Re^q$, then both $\mathcal{C}_q(D)$ and $\mathcal{C}_q^*(D)$ are equal to $\mathcal{S}_q$, and in the case of $D = \Re^q_+$, we have the regular copositive and completely positive cones, $\mathcal{C}_q$ and $\mathcal{C}_q^*$, as defined in this paper. Using these definitions, one implication of their work is that any quadratic minimization over the set $\{x \in D : e^T x = 1\}$ can be represented as a linear program over $\mathcal{C}_q^*(D)$. In this sense, Sturm and Zhang prove a natural extension of the work of Bomze et al. It is also interesting to note that the Sturm–Zhang construction keeps the variable x explicit in the equivalent conic linear program, just as we have kept x explicit in (C).

4.2 The maximum stable set problem

The next example of representing an NP-hard problem as a completely positive program is the maximum stable set problem as established by de Klerk and Pasechnik [7]. One description of their result is as follows. Given a graph $G = (\{1, \ldots, n\}, E)$, it is not difficult to show (see Burer et al. [4]) that the maximum stable set number α for G satisfies

$$\alpha = \max \left\{ x^T e e^T x \ : \ x \geq 0,\ \|x\|^2 = 1,\ x_i x_j = 0 \forall (i, j) \in E \right\}.$$

De Klerk and Pasechnik show that the corresponding CPP also has optimal value α, i.e.,

$$\alpha = \max \left\{ ee^T \bullet X \ : \ X \in \mathcal{C}_n^*,\ I \bullet X = 1,\ X_{ij} = 0 \forall (i, j) \in E \right\}.$$

This program is also closely related to the Lovász-theta-number SDP [10], which is achieved by relaxing $X \in \mathcal{C}_n^*$ to $X \in \mathcal{S}_n$.

In fact, the maximum stable set problem can be formulated as a standard quadratic program [11], and so it is known from Bomze et al. [3] that the maximum stable set problem has a completely positive representation. In this sense, the result of de Klerk and Pasechnik is not completely new. However, in addition to the representation above, de Klerk and Pasechnik show further results, e.g., that an approximation of $\mathcal{C}_n^*$ by $(\mathcal{K}_n^r)^*$ with $r = \alpha^2$ (see the Introduction) yields the optimal value α (after rounding down).

4.3 Quadratic problems over transportation matrices

As mentioned in the Introduction, Povh [15] has shown that a certain class of quadratic programs over transportation matrices can be represented as CPPs. In the following subsections, we investigate two specific cases, which have appeared prior to the generalization established by Povh.

4.3.1 A minimum-cut graph tri-partitioning problem

Povh and Rendl [17] have recently shown that a certain NP-hard graph partitioning problem can be posed as a CPP. Given a graph $G = (\{1, \dots, n\}, E)$ and a length-three integer vector $s = (s_1; s_2; s_3)$ such that $e_3^T s = n$, the problem is to partition $\{1, \dots, n\}$ into sets S_i $(i = 1, 2, 3)$ with $|S_i| = s_i$ such that the number of edges in E crossing from S_1 to S_2 is minimal. Letting $x \in \Re^{3n}$ and $Y \in \Re^{n\times 3}$ be associated by the equation $x = \text{vec}(Y)$, the problem can be formulated as the minimization of a (homogeneous) quadratic objective over the feasible set

$$\left\{ x = \text{vec}(Y) : \ x \geq 0, \ Ye_3 = e_n, \ \ Y^T e_n = s, \ \ x_j \in \{0, 1\} \ \ \forall j = 1, \dots, 3n \right\}. \tag{13}$$

From this representation, Theorem 2.6 thus implies that this partitioning problem can be represented as a CPP.

However, Povh and Rendl take an interesting alternative route. It can be shown that (x, Y) is feasible for (13) if and only if Y is in the set

$$\left\{ Y \geq 0 \ : \ Ye_3 = e_n, \ Y^T e_n = s, \ Y^T Y = \text{Diag}(s) \right\}.$$

Then they multiply linear constraints to include redundant quadratic constraints—ones that ultimately allow them to drop all linear constraints and obtain a completely (homogeneous) quadratic representation of the feasible set:

$$\left\{ Y \geq 0 \ : \ (Ye_3)^2 = e_n, \ \ Ye_3e_n^T Y = e_n s^T, \ \ Y^T e_n e_n^T Y = ss^T, \ Y^T Y = \text{Diag}(s) \right\}.$$

Here $(Ye_3)^2$ indicates the Hadamard (i.e., component-wise) product of Ye_3 with itself. The corresponding CPP is then shown to be equivalent to the graph partitioning problem.

4.3.2 The quadratic assignment problem

In addition to the graph partitioning problem of the previous subsection, Povh and Rendl [16] also study the quadratic assignment problem (QAP), for which a similar approach is possible. Letting $x \in \Re^{n^2}$ and $Y \in \Re^{n\times n}$ be identified by the relationship $x = \text{vec}(Y)$, the QAP can be formulated as minimizing a (homogeneous) quadratic objective over the feasible set

$$\left\{ x = \text{vec}(Y) : \ x \geq 0, \ Ye_n = e_n, \ \ Y^T e_n = e_n, \ \ x_j \in \{0, 1\} \ \ \forall j = 1, \dots, n^2 \right\}, \tag{14}$$

and Theorem 2.6 establishes a completely positive representation.

Here again, however, Povh and Rendl take a different approach. It can be shown that (x, Y) is in the feasible set (14) if and only if Y is in the set

$$\left\{ Y \geq 0 : \; Y^T Y = I \right\}.$$

After adding the redundant constraints $YY^T = I$ and $(e_n^T Y e_n)^2 = n^2$ to yield

$$\left\{ Y \geq 0 : \; Y^T Y = YY^T = I, \; (e_n^T Y e_n)^2 = n^2 \right\},$$

they prove that the resulting CPP represents the QAP exactly.

4.4 Comparison

The results of this paper fit naturally between those of Bomze et al. [3] and Sturm and Zhang [18] (see Sect. 4.1). Indeed, Theorems 2.6 and 3.1 are generalizations of the work of Bomze et al. On the other hand, in contrast to Sturm–Zhang, the theorems make clear that it is not necessary to explicitly utilize all features of the feasible set inside the cone. Instead, the (regular) copositive and completely positive cones suffice. This is an important property since emerging tools are available for approximating $\mathcal{C}_q$ and $\mathcal{C}_q^*$ (see the Introduction)—but not for $\mathcal{C}_q(D)$ and $\mathcal{C}_q^*(D)$ in the case of general $D \subseteq \Re^q$.

From the descriptions above, we see that all previous examples of representing NP-hard problems as CPPs share the following common characteristic: before representation as a CPP, the NP-hard problem is first expressed with homogeneous quadratic objective and homogeneous quadratic equality constraints. We point out that it is not obvious how to extend our approach to more general quadratic constraints (homogeneous or otherwise) beyond the discussion in Sect. 3.2. In this sense, our analysis cannot be said to subsume previous analyses (although the previously studied problems can be formulated differently to fit our setup). On the other hand, our approach does interestingly require that each linear constraint be squared and added as a redundant homogeneous quadratic equation. Overall, it seems that the original equation and its squared version are critical for Theorems 2.6 and 3.1. A related observation is that all previous completely positive representations have been in terms of X only as with (C'), not in terms of both x and X as with (C).

Some additional points of comparison can be made with the works of Povh and Rendl:

- For the graph partitioning problem and the QAP, Povh and Rendl prove the completely positive representation by appealing to the equality case of the Cauchy–Schwarz inequality (in addition to other arguments). In fact, we use Cauchy–Schwarz in a very similar manner as they—namely to prove important properties of the rank-1 matrices in the completely positive decomposition. It is interesting that both approaches share this proof technique.
- For the graph partitioning problem, Povh and Rendl start with a mixed linear and quadratic representation of the feasible set, and then multiply various linear

equations to ultimately convert everything to quadratic equations. In particular, some linear equations are squared, and some are multiplied pair-wise with others. Our approach employs squared equations but does not require pair-wise multiplications.

– Povh and Rendl handle integrality in an implicit manner via the formulations, while we handle it in a more straightforward manner.

5 Computational considerations

The primary purpose of the current paper is theoretical—to demonstrate that a broad class of NP-hard problems can be modeled as a specific class of well-structured convex programs. However, a long-term goal is certainly to provide new practical tools for computationally solving difficult optimization problems. For this reason, we now discuss some computational issues related to the completely positive representations (C) and (C').

It is important to keep in mind that, while both (C) and (C') are convex programming problems, they are themselves NP-hard due to their equivalence with (P). So the results of this paper are not a "magic bullet" for solving NP-hard problems. Indeed, formidable computational challenges still (necessarily) exist. However, the paper does demonstrate that the main obstacle for solving (P) via (C) or (C') is working with the cone of completely positive matrices $\mathcal{C}_q^*$; the only other considerations are linear constraints and a linear objective.

Hence, a natural computational approach is to focus attention on approximating $\mathcal{C}_q^*$ to obtain tractable convex relaxations of (P). This is the idea behind the linear- and semidefinite-representable cones $(\mathcal{K}_q^r)^*$ mentioned in the Introduction. Variations and simplifications of $(\mathcal{K}_q^r)^*$ are explored by Peña et al. [14]. Another approach, which dynamically provides better and better polyhedral approximations of $\mathcal{C}_q^*$ via a type of variable generation scheme, has recently been proposed by Bundfuss and Dür [8]. In each case, the computational burden for improving the quality of the approximation is substantial. For example, the size of the representation of $(\mathcal{K}_q^r)^*$ is approximately $O(n^r)$, which makes the full use of $(\mathcal{K}_q^r)^*$ only feasible for small r.

Complementary to approximating $\mathcal{C}_q^*$, a standard approach for developing convex relaxations of (P) is incorporating valid linear constraints in the space (x, X). For example, one could consider the SDP relaxation

$$\begin{array}{ll} \min & Q \bullet X + 2\,c^T x \\ \text{s.t.} & a_i^T x = b_i, \quad \forall i = 1, \ldots, m \\ & a_i^T X a_k = b_i b_k \quad \forall i = 1, \ldots, m,\ k = i, \ldots, m \\ & x_j = X_{jj} \quad \forall j \in B \\ & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathcal{S}_{1+n}, \end{array}$$

which introduces $O(m^2)$ more linear constraints compared to (C) (though the cone is much simpler). Other variations and types of valid constraints are certainly possible, e.g., there exists an entire class of valid inequalities coming from the binary conditions

$x_j \in \{0, 1\}$ for $j \in B$ [12]. The number and strength of added constraints is of primary computational importance.

In light of the above, an important computational insight provided by the current paper is the following:

> (C) and (C') make clear that the two complementary strategies of approximating $\mathcal{C}_q^*$ and adding valid constraints on (x, X) can actually be combined in a single intuitive framework and applied to a wide class of difficult optimization problems.

For any specific situation, determining the most practical way to combine the two strategies will likely require careful consideration, and simple combination strategies may not be possible. (An added layer of complexity arises when (P) can be formulated in different ways, giving rise to different formulations of (C) and (C').) Nevertheless, we feel the above insight is a critical contribution of this paper and deserves future investigation.

Finally, we mention two success stories arising from the perspective of completely positive programming. (The first is derived directly from this paper, while the second arises from the work of Povh and Rendl). In our opinion, these examples support the viewpoint that the current results can lead to new computational insights for difficult optimization problems.

First, Anstreicher and Burer [1] have recently used the results of this paper to provide a closed-form polyhedral and semidefinite representation of

$$\operatorname{Conv}\left\{ \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix}^T : \; 0 \le x_1 \le 1, \; 0 \le x_2 \le 1 \right\},$$

answering a fundamental open question in the field of global optimization. While this low-dimensional result may seem of limited computational value, it actually provides a complete characterization of all valid constraints for 3×3 leading submatrices in the space $(1; x)(1; x)^T$ for the box-constrained quadratic program $\min\{x^T Q x + 2\, c^T x : 0 \le x \le e\}$. As such, this result is important for the global solution of high-dimensional box-constrained QPs.

Second, by relaxing $\mathcal{C}_q^*$ to $\mathcal{S}_q$ in their completely positive representation of QAP (see Sect. 4.3.2), Povh and Rendl [16] have derived an efficiently computable SDP relaxation of QAP whose quality significantly dominates all previously known SDP relaxations of the same (approximate) size.

6 Conclusion

This paper significantly broadens the application of copositive and completely positive programming. Still, more work is necessary to exploit these formulations in practice as discussed in Sect. 5. Some other questions of interest are:

– Can more general quadratic constraints beyond those in Sect. 3.2 be included in (P)?

– Can we extend the bounds on approximation error proved by Bomze and de Klerk [2] for standard quadratic programming? Even more interesting, can we extend the polynomial-time approximation scheme for standard quadratic programming to more general constraint structures?
– Are there cases in which one can bound r so that (P) is equivalent to optimization over $(\mathcal{K}^r_{1+n})^*$? Positive results along these lines by de Klerk and Pasechnik [7] have been shown for their approach to the maximum stable set problem.

Acknowledgments The author would like to thank: Janez Povh for clarifying and detailing earlier results on copositive representations; Kurt Anstreicher for insightful comments on a draft of this paper; Immanuel Bomze for pointing out an earlier error in Proposition 2.3; and Mirjam Dür for discussions concerning the extension to complementarity constraints. The author is also in debt to the two anonymous referees, who provided many helpful comments and suggestions.

References

1. Anstreicher, K.M., Burer, S.: Computable Representations for Convex Hulls of Low-Dimensional Quadratic Forms. University of Iowa, Manuscript (2007)
2. Bomze, I.M., de Klerk, E.: Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. J. Global Optim. **24**(2), 163–185 (2002). Dedicated to Professor Naum Z. Shor on his 65th birthday
3. Bomze, I.M., Dür, M., de Klerk, E., Roos, C., Quist, A.J., Terlaky, T.: On copositive programming and standard quadratic optimization problems. J. Global Optim. **18**(4), 301–320 (2000). GO'99 Firenze
4. Burer, S., Monteiro, R., Zhang, Y.: Maximum stable set formulations and heuristics based on continuous optimization. Math. Program. **94**, 137–166 (2002)
5. Burer, S., Vandenbussche, D.: A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations. Math. Program. **113**, 259–282 (2008)
6. de Klerk, E., Laurent, M., Parrilo, P.A.: A PTAS for the minimization of polynomials of fixed degree over the simplex. Theor. Comput. Sci. **361**(2–3), 210–225 (2006)
7. de Klerk, E., Pasechnik, D.V.: Approximation of the stability number of a graph via copositive programming. SIAM J. Optim. **12**(4), 875–892 (2002)
8. Bundfuss, S., Dür, M.: An Adaptive Linear Approximation Algorithm for Copositive Programs. Manuscript, Department of Mathematics, Technische Universität Darmstadt, Darmstadt, Germany (2008)
9. Giannessi, F., Tomasin, E.: Nonconvex quadratic programs, linear complementarity problems, and integer Linear Programs. In: Fifth Conference on Optimization Techniques (Rome, 1973), Part I. Lecture Notes in Compute Sciences, Vol. 3, pp. 437–449. Springer, Berlin (1973)
10. Lovász, L.: On the Shannon Capacity of a graph. IEEE Trans. Inf. Theory IT- **25**(1), 1–7 (1979)
11. Motzkin, T.S., Straus, E.G.: Maxima for graphs and a new proof of a theorem of Turán. Can. J. Math. **17**(4), 533–540 (1965)
12. Padberg, M.: The Boolean quadric polytope: some characteristics, facets and relatives. Math. Program. **45**(1, (Ser. B)), 139–172 (1989)
13. Parrilo, P.: Structured Semidefinite Programs and Semi-algebraic Geometry Methods in Robustness and Optimization. Ph.D. thesis, California Institute of Technology (2000)
14. Peña, J., Vera, J., Zuluaga, L.F.: Computing the stability number of a graph via linear and semidefinite programming. SIAM J. Optim. **18**(1), 87–105 (2007)
15. Povh, J.: Application of Semidefinite and Copositive Programming in Combinatorial Optimization. Ph.D. thesis, University of Ljubljana (2006)
16. Povh, J., Rendl, F.: Copostive and semidefinite relaxations of the quadratic assignment problem. Manuscript, Faculty of Logisitics, University in Maribor, Mariborska cesta 3, Celje, Slovenia, July 2006 (submitted)
17. Povh, J., Rendl, F.: A copositive programming approach to graph partitioning. SIAM J. Optim. **18**(1), 223–241 (2007)
18. Sturm, J.F., Zhang, S.: On cones of nonnegative quadratic functions. Math. Oper. Res. **28**(2), 246–267 (2003)