SIAM J. OPTIM.
Vol. 10, No. 3, pp. 750–778

CONES OF MATRICES AND SUCCESSIVE CONVEX RELAXATIONS OF NONCONVEX SETS*

MASAKAZU KOJIMA† AND LEVENT TUNÇEL‡

Abstract. Let F be a compact subset of the n-dimensional Euclidean space R^n represented by (finitely or infinitely many) quadratic inequalities. We propose two methods, one based on successive semidefinite programming (SDP) relaxations and the other on successive linear programming (LP) relaxations. Each of our methods generates a sequence of compact convex subsets C_k $(k = 1, 2, \dots)$ of R^n such that

(a) the convex hull of $F \subseteq C_{k+1} \subseteq C_k$ (monotonicity),
(b) $\cap_{k=1}^{\infty} C_k =$ the convex hull of F (asymptotic convergence).

Our methods are extensions of the corresponding Lovász–Schrijver lift-and-project procedures with the use of SDP or LP relaxation applied to general quadratic optimization problems (QOPs) with infinitely many quadratic inequality constraints. Utilizing descriptions of sets based on cones of matrices and their duals, we establish the exact equivalence of the SDP relaxation and the semi-infinite convex QOP relaxation proposed originally by Fujie and Kojima. Using this equivalence, we investigate some fundamental features of the two methods including (a) and (b) above.

Key words. semidefinite programming, nonconvex quadratic optimization problem, linear matrix inequality, bilinear matrix inequality, semi-infinite programming, global optimization

AMS subject classifications. 15A48, 52A47, 49M39, 90C05, 90C25, 90C26, 90C30, 90C34

PII. S1052623498336450

1. Introduction.

Consider a maximization problem with a linear objective function $\boldsymbol{c}^T\boldsymbol{x}$:

$$\text{maximize } \boldsymbol{c}^T\boldsymbol{x} \quad \text{subject to } \boldsymbol{x} \in F, \tag{1.1}$$

where $\boldsymbol{c}$ denotes a constant vector in the n-dimensional Euclidean space R^n and F a subset of R^n. We can reduce a more general maximization problem with a nonlinear objective function $f(\boldsymbol{x})$ to a maximization problem having a linear objective function represented by a new variable, x_{n+1}, if we replace $f(\boldsymbol{x})$ by x_{n+1} and then add the inequality $f(\boldsymbol{x}) \geq x_{n+1}$ to the constraint. Thus (1.1) covers such a general optimization problem. Throughout the paper we assume that F is compact. Then the problem (1.1) has a global maximizer whenever the feasible region F is nonempty.

For any compact convex set C containing F, the maximization problem

$$\text{maximize } \boldsymbol{c}^T\boldsymbol{x} \quad \text{subject to } \boldsymbol{x} \in C \tag{1.2}$$

serves as a convex relaxation problem, which satisfies the properties that

(i) the maximum objective value ζ of the problem (1.2) gives an upper bound for the maximum objective value ζ^* of the problem (1.1), i.e., $\zeta \geq \zeta^*$, and

*Received by the editors March 31, 1998; accepted for publication (in revised form) July 19, 1999; published electronically March 21, 2000.
http://www.siam.org/journals/siopt/10-3/33645.html

†Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552, Japan (kojima@is.titech.ac.jp).

‡Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada (ltuncel@math.uwaterloo.ca). This work was completed while this author was visiting Tokyo Institute of Technology, Department of Mathematical and Computing Sciences, on a sabbatical leave from University of Waterloo. The research of this author was supported in part by Tokyo Institute of Technology and by a research grant from NSERC of Canada.

(ii) if a maximizer $\hat{\boldsymbol{x}} \in C$ of (1.2) lies in F, it is a maximizer of (1.1).

Since the objective function of (1.1) is linear, we know that if we take the convex hull c.hull(F) (defined as the intersection of all the convex sets containing F) for C in (1.2), then

(i)$'$ $\zeta = \zeta^*$, and

(ii)$'$ the set of the maximizers of (1.2) forms a compact convex set whose extreme points are maximizers of (1.1).

Therefore, if we solve the relaxation problem (1.2) with a convex feasible region C which closely approximates c.hull(F), we can expect to get not only a good upper bound ζ for the maximum objective value ζ^* but also an approximate maximizer of the problem (1.1). We can further prove that for almost every $\boldsymbol{c} \in R^n$ (in the sense of measure), any maximizer $\boldsymbol{x}' \in C = $ c.hull(F) of (1.2) is an extreme point of c.hull(F), which also lies in F; hence $\boldsymbol{x}'$ is a maximizer of (1.1). This follows from a result due to Ewald, Larman, and Rogers [5] for consequences of related results; see also [17]. Furthermore, for many representations of various convex sets C, given $\hat{\boldsymbol{x}} \in C$, we can very efficiently find $\boldsymbol{x}^*$, an extreme point of C, such that $\boldsymbol{c}^T\boldsymbol{x}^* \geq \boldsymbol{c}^T\hat{\boldsymbol{x}}$.

Indeed, the relaxation technique mentioned above has been playing an essential role in practical computational methods for solving various problems in the fields of combinatorial optimization and global optimization. It is often used in hybrid schemes with the branch-and-bound and branch-and-cut techniques in those fields. See, for instance, [2].

The aim of this paper is to present a basic idea on how we can approximate the convex hull of F. This is a quite difficult problem, and also too general. Before making further discussions, we at least need to provide an appropriate (algebraic) representation for the compact feasible region F of the problem (1.1) and the compact convex feasible region C of the relaxation problem (1.2). We employ quadratic inequalities for this purpose.

Let $\mathcal{S}^n$ and $\mathcal{S}^n_+ \subset \mathcal{S}^n$ denote the set of $n \times n$ symmetric matrices and the set of $n \times n$ symmetric positive semidefinite matrices, respectively. Given $\boldsymbol{Q} \in \mathcal{S}^n$, $\boldsymbol{q} \in R^n$, and $\gamma \in R$, we write a quadratic function on R^n with the quadratic term $\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$, the linear term $2\boldsymbol{q}^T\boldsymbol{x}$, and the constant term γ as $p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q})$:

$$p(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) \equiv \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \quad \forall \boldsymbol{x} \in R^n.$$

Then the set $\mathcal{Q}$ of quadratic functions on R^n and the set $\mathcal{Q}_+$ of convex quadratic functions are defined as

$$\mathcal{Q} \equiv \{p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) : \boldsymbol{Q} \in \mathcal{S}^n,\ \boldsymbol{q} \in R^n \text{ and } \gamma \in R\}$$

and

$$\mathcal{Q}_+ \equiv \{p(\cdot;\gamma,\ \boldsymbol{q},\boldsymbol{Q}) : \boldsymbol{Q} \in \mathcal{S}^n_+,\ \boldsymbol{q} \in R^n \text{ and } \gamma \in R\},$$

respectively. We also write $p(\cdot) \in \mathcal{Q}$ (or $\mathcal{Q}_+$) instead of $p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{Q}$ (or $\mathcal{Q}_+$) if $\boldsymbol{Q} \in \mathcal{S}^n$, $\boldsymbol{q} \in R^n$, and $\gamma \in R$ are irrelevant. Throughout the paper, we assume that the feasible region F of the problem (1.1) is represented by a set of quadratic inequalities such that

$$F = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0\ \forall p(\cdot) \in \mathcal{P}_F\},$$

where $\mathcal{P}_F$ denotes a set of quadratic functions, i.e., $\mathcal{P}_F \subseteq \mathcal{Q}$, and we will derive convex relaxations, C, represented by convex quadratic inequalities such that

$$C = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0\ \forall p(\cdot) \in \mathcal{P}_C\},$$

where $\mathcal{P}_C$ denotes a set of convex quadratic functions, i.e., $\mathcal{P}_C \subseteq \mathcal{Q}_+$. We allow cases where $\mathcal{P}_F$ and/or $\mathcal{P}_C$ involve infinitely many quadratic functions. Thus (1.1) or (1.2) (or both) can be a semi-infinite quadratic optimization problem (QOP). Here we use the word "semi-infinite" for optimization problems having a finite number of scalar variables and possibly an infinite number of inequality constraints.

There are some reasons why we have chosen quadratic inequalities for the representation of both problems, the maximization problem (1.1) that we want to solve and its convex relaxation problem (1.2). First, quadratic inequalities form a class of relatively easily manageable nonlinear inequalities, yet they have enough power to describe any compact feasible region F in R^n. Indeed, if F is closed, then its complement $R^n \backslash F$ is open so that it can be represented as the union of the open balls

$$\{\boldsymbol{x} \in R^n : (\boldsymbol{x} - \boldsymbol{x}')^T(\boldsymbol{x} - \boldsymbol{x}') < \epsilon(\boldsymbol{x}')\} \quad \text{with } \exists \epsilon(\boldsymbol{x}') > 0$$

over all $\boldsymbol{x}' \in \mathcal{G}$ for some $\mathcal{G} \subseteq R^n \backslash F$; hence

$$F = \{\boldsymbol{x} \in R^n : (\boldsymbol{x} - \boldsymbol{x}')^T(\boldsymbol{x} - \boldsymbol{x}') \geq \epsilon(\boldsymbol{x}') \ \forall \boldsymbol{x}' \in \mathcal{G}\}.$$

We also know that any single polynomial inequality can be converted into a system of quadratic inequalities; for example,

$$x_1^2 x_2 + 2x_1 x_2^2 - 5 \leq 0$$

can be converted into

$$x_3 - x_1 x_2 \leq 0, \ -x_3 + x_1 x_2 \leq 0 \quad \text{and} \quad x_1 x_3 + 2x_2 x_3 - 5 \leq 0.$$

See [23, 24].

Second, we know that we can solve some classes of maximization problems having linear objective functions and a convex-quadratic-inequality constrained feasible region C efficiently. Among others, we can apply interior-point methods [1, 16] to the problem (1.2) when either $\mathcal{P}_C$ is finite or $\mathcal{P}_C$ is infinite, but its feasible region C is described as the projection of a set characterized by linear matrix inequalities in the space $\mathcal{S}^n$ of $n \times n$ symmetric matrices onto the n-dimensional Euclidean space R^n.

Third, and also most importantly, we can apply the semidefinite programming (SDP) relaxation, which was originally developed for 0-1 integer programming problems by Lovász and Schrijver [12] and later extended to nonconvex quadratic optimization problems [6, 18, 19], to the entire class of maximization problems having a linear objective function and finitely or infinitely many quadratic inequality constraints. See also [1, 8, 9, 13, 15, 23, 24, 29].

In addition to the reasons above, we should mention that the maximization problem with a linear objective function and quadratic inequality constraints involves various optimization problems such as 0-1 integer linear (or quadratic) programming problems which, in principle, include all combinatorial optimization problems [1, 9, 18]. Linear complementarity problems [4], bimatrix games, and bilinear matrix inequalities [14, 20] are also included as special cases.

For some optimization problems, some of the semidefinite programming (SDP) relaxations we provide may be solved in polynomially many iterations (of an interior-point method or an ellipsoid algorithm) approximately. Such conclusion requires, in the case of the ellipsoid method, the existence of a certain polynomial-time separation oracle for the underlying convex cone constraint (see [9]). In the case of interior-point algorithms (whose efficiency in the theory and practice of SDP has been well established), we need to have an efficiently computable self-concordant barrier for the feasible solutions set or at least for the underlying cone constraints (see [16]).

Some of the most exciting activities in combinatorial optimization are currently centered around the applications of SDP to combinatorial optimization problems (see [7]). Such activity in theory and practice is fueled by theoretical results establishing that certain simple SDP relaxations of a combinatorial optimization problem can be effectively utilized in developing polynomial-time approximation algorithms with worst-case approximation-ratio guarantees much better than those previously proven using linear programming or other techniques. (See Goemans [7], Goemans and Williamson [8], Nesterov [15], and Ye [29].) Also outstanding are the results on the stable set problem establishing the fact that SDP techniques can be used in optimizing over a relaxation of the stable set polytope which is contained in the polytope defined by the clique inequalities. (Note that it is NP-hard to optimize over the latter-mentioned polytope, whereas Grötschel, Lovász, and Schrijver [9] and Lovász, and Schrijver [12] were able to utilize polynomial-time methods to achieve a better goal, as far as the proof of approximate optimality of some feasible solutions of the stable set problem is concerned.)

Given an initial approximation C_0 of F, i.e., a compact convex set C_0 containing F, both of the methods, proposed in this paper, generate a sequence of compact convex subsets C_k $(k = 1, 2, \dots)$ of R^n such that

(a) c.hull$(F) \subseteq C_{k+1} \subseteq C_k$ (monotonicity),

(b) $\cap_{k=1}^{\infty} C_k =$ c.hull(F) (asymptotic convergence).

It should be noted that the compactness of each C_k and property (b) imply that

(c) if $F = \emptyset$, then $\cap_{k=1}^{k^*} C_k = \emptyset$ for some finite number k^* (detecting infeasibility).

To generate C_{k+1} at each iteration, the SDP relaxation and the linear programming (LP) relaxation play an essential role, and the entire method may be regarded as an extension of the Lovász–Schrijver lift-and-project procedure for 0-1 integer programming problems to semi-infinite nonconvex quadratic optimization problems, with the use of the SDP relaxation in the first method and the LP relaxation in the second method. The LP relaxation, referred to above, is essentially the same as the reformulation-linearization technique developed for nonconvex quadratic optimization problems by Sherali and Alameddine [21]; see also [2, 22]. However, we should caution the reader that the methods presented here are mostly conceptual in the general settings, because we need to solve a semi-infinite SDP (or a semi-infinite LP) at each iteration. For such a task, an efficient practical algorithm may not be currently available.

In their paper [6], Fujie and Kojima proposed the semi-infinite convex QOP relaxation for nonconvex quadratic optimization problems and showed that the semi-infinite convex QOP relaxation is not stronger than the SDP relaxation in general, but the two relaxations are essentially equivalent under Slater's constraint qualification. We establish the exact equivalence between the two relaxations for semi-infinite nonconvex quadratic optimization problems without any constraint qualification. Using this equivalence, we derive some fundamental features of our methods including (a) and (b) above. One of the common themes in this paper is the usage of cones of matrices (and duality) in our constructions. This was also one of the themes of [12]. The other themes of this paper are the successive applications of SDP relaxations and LP relaxations. We call the related procedures the successive SDP relaxation method and the successive semi-infinite LP relaxation method, respectively.

Section 2 is devoted to preliminaries, where we provide some basic definitions and properties on quadratic inequality representations for closed subsets of R^n, the homogeneous form of quadratic functions, the SDP relaxation, etc. In section 3, we present our first method in detail as well as the main results, including the features (a) and (b). After we present some fundamental characterizations of the SDP relaxation in section 4, we give proofs of the main results in section 5. In section 6, we apply our method to 0-1 semi-infinite nonconvex quadratic optimization problems. Incorporating the basic results on the lift-and-project procedure given by Lovász and Schrijver [12] for 0-1 integer convex optimization problems, we show that our method terminates in at most $(n+1)$ iterations either to generate the convex hull of the feasible region or to detect the emptiness of the feasible region, where n denotes the number of 0-1 variables of the problem. Section 7 contains our second method, which is based on semi-infinite LP relaxations. We establish the same theoretical properties as we do for the successive SDP relaxation method. In section 8, we present two numerical examples showing the worst-case behavior of some of our procedures. In particular, we know from the second example that the best of our procedures requires infinitely many iterations to generate the convex hull of F in the worst case.

2. Preliminaries.

2.1. Semi-infinite quadratic inequality representation. In this subsection, we discuss some representations of a closed subset F of R^n in terms of (possibly infinitely many) quadratic inequalities. If $p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{Q}$, and $p(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) \leq 0$ holds for all $\boldsymbol{x} \in F$, we say that $p(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) \leq 0$ is *a quadratic valid inequality* for F and that $p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q})$ induces a quadratic valid inequality for F. A quadratic valid inequality $p(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) \leq 0$ for F is

linear if $\boldsymbol{Q} = \boldsymbol{O}$,

rank-1 *quadratic* if $p(\boldsymbol{x}) = (\boldsymbol{a}^T\boldsymbol{x}-\alpha)(\boldsymbol{a}^T\boldsymbol{x}-\beta)$ for $\exists \boldsymbol{a} \in R^n$, $\exists \alpha \in R$ and $\exists \beta \in R$ such that $\alpha \leq \boldsymbol{a}^T\boldsymbol{x} \leq \beta \ \forall \boldsymbol{x} \in F$,

rank-2 *quadratic* if $p(\boldsymbol{x}) = -(\boldsymbol{a}^T\boldsymbol{x}-\alpha)(\boldsymbol{b}^T\boldsymbol{x}-\beta)$ for $\exists \boldsymbol{a} \in R^n$, $\exists \boldsymbol{b} \in R^n$, $\exists \alpha \in R$ and $\exists \beta \in R$ such that $\boldsymbol{a}^T\boldsymbol{x} \leq \alpha$ and $\boldsymbol{b}^T\boldsymbol{x} \leq \beta \ \forall \boldsymbol{x} \in F$,

spherical if $p(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{d})^T(\boldsymbol{x}-\boldsymbol{d}) - \rho$ for $\exists \boldsymbol{d} \in R^n$ and $\exists \rho > 0$,

ellipsoidal if $p(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{d})^T\boldsymbol{Q}(\boldsymbol{x}-\boldsymbol{d}) - \rho$ for $\exists \boldsymbol{Q} \in \mathcal{S}^n_{++}$, $\boldsymbol{d} \in R^n$ and $\exists \rho > 0$,

convex quadratic if $\boldsymbol{Q} \in \mathcal{S}^n_{+}$,

respectively. It should be noted that if a quadratic valid inequality $p(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) \leq 0$ for F is rank-2, then the rank of the matrix $\boldsymbol{Q}$ is at most 2 but that the converse is not necessarily true.

We say that F has *a (semi-infinite) quadratic inequality representation* $\mathcal{P} \subseteq \mathcal{Q}$ if

$$F = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) \leq 0 \ \forall p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}\}$$

holds. To designate the underlying representation $\mathcal{P}$ of F, we often write $F(\mathcal{P})$ instead of F. Whenever F is a closed proper subset of R^n, F has infinitely many representations. We allow the cases where $\mathcal{P}$ consists of infinitely many quadratic functions. Hence $p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}$ can be a semi-infinite system of quadratic inequalities. If $\mathcal{P} \subseteq \mathcal{Q}$ is a quadratic inequality representation of F and if $p(\cdot) \in \text{c.cone}(\mathcal{P})$, then $p(\boldsymbol{x}) \leq 0$ is a quadratic valid inequality, where $\text{c.cone}(\mathcal{P})$ denotes the closed convex cone generated by $\mathcal{P}$. Hence if $\mathcal{P} \subseteq \mathcal{P}' \subseteq \text{c.cone}(\mathcal{P})$, then $\mathcal{P}'$ is a quadratic inequality representation of F; $F(\mathcal{P}) = F(\mathcal{P}') = F(\text{c.cone}(\mathcal{P}))$. A quadratic inequality representation $\mathcal{P}$ of F is *finite* if it consists of a finite number of quadratic functions, and *infinite* otherwise. If F is a compact convex subset of R^n, it has a quadratic inequality representation; in fact, the set of all the linear (rank-2 quadratic or spherical) valid inequalities for F forms an inequality representation of F. If, in addition, F is polyhedral, we can take a finite linear inequality representation.

Let C be a compact subset of R^n. We use the following symbols:

$\mathcal{P}^L(C) =$ the set of $p(\cdot)$'s that induce linear valid inequalities for C,

$\mathcal{P}^1(C) =$ the set of $p(\cdot)$'s that induce rank-1 quadratic valid inequalities for C,

$\mathcal{P}^2(C) =$ the set of $p(\cdot)$'s that induce rank-2 quadratic valid inequalities for C,

$\mathcal{P}^S(C) =$ the set of $p(\cdot)$'s that induce spherical valid inequalities for C,

$\mathcal{P}^E(C) =$ the set of $p(\cdot)$'s that induce ellipsoidal valid inequalities for C,

$\mathcal{P}^C(C) =$ the set of $p(\cdot)$'s that induce convex quadratic valid inequalities for C,

$\mathcal{P}^\sharp(C) =$ the set of $p(\cdot)$'s that induce all quadratic valid inequalities for C.

By definition, we see that

$$\begin{array}{l}\left(\mathcal{P}^L(C) \cup \mathcal{P}^1(C) \cup \mathcal{P}^S(C) \cup \mathcal{P}^E(C)\right) \subset \mathcal{P}^C(C) \subset \mathcal{P}^\sharp(C),\\ \mathcal{P}^S(C) \subset \mathcal{P}^E(C) \quad \text{and} \quad \left(\mathcal{P}^L(C) \cup \mathcal{P}^1(C)\right) \subset \mathcal{P}^2(C) \subset \mathcal{P}^\sharp(C).\end{array}$$

Note that if C is convex, then the equality

$$C = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}\}$$

holds with each $\mathcal{P} = \mathcal{P}^L(C), \mathcal{P}^1(C), \mathcal{P}^2(C), \mathcal{P}^S(C), \mathcal{P}^E(C), \mathcal{P}^C(C), \mathcal{P}^\sharp(C)$. Among these, $\mathcal{P}^\sharp(C)$ is *the strongest quadratic inequality representation* of C.

2.2. Homogeneous form of quadratic functions—lifting to the space of symmetric matrices. We introduce a different description of quadratic functions, which we call the homogeneous form. This form leads us to a lifting of a quadratic function defined on the Euclidean space to the space of symmetric matrices and to the SDP relaxation (or to the semi-infinite LP relaxation in section 4.2). For every quadratic function $p(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{Q}$, we connect the variable vector $\boldsymbol{x} \in R^n$ to the $(1+n) \times (1+n)$ rank-1 positive semidefinite matrix

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} (1, \boldsymbol{x}^T) \in \mathcal{S}^{1+n}_+$$

and the triplet of the constant $\gamma \in R$, $\boldsymbol{q} \in R^n$, and $\boldsymbol{Q} \in \mathcal{S}^n$ to the $(1+n) \times (1+n)$ symmetric matrix $\left(\begin{smallmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{smallmatrix}\right) \in \mathcal{S}^{1+n}$. Then we have the identity

$$p(\boldsymbol{x}; \gamma, \boldsymbol{q}, \boldsymbol{Q}) = (1, \boldsymbol{x}^T) \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{pmatrix} \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{pmatrix} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \ \forall \boldsymbol{x} \in R^n.$$

Thus, if $\mathcal{P} \subseteq \mathcal{Q}$ is a quadratic inequality representation of F, then

$$\underline{\mathcal{P}} \equiv \left\{ \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{pmatrix} : p(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P} \right\}$$

provides an equivalent representation of F;

$$F(\mathcal{P}) = \left\{ \boldsymbol{x} \in R^n : \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0 \ \forall \boldsymbol{P} \in \underline{\mathcal{P}} \right\}.$$

Now we have two kinds of description for a quadratic function on R^n: the usual form $p(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) = \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$ and the homogeneous form introduced above.

The former is used in section 5, where we prove our main results, while the latter is suitable for the compact description of the SDP relaxation in section 2.3 and the proof of its equivalence to the semi-infinite convex QOP relaxation in section 4. We will use both forms in parallel, choosing whichever is convenient to us in a given situation. It should be noted that the correspondence

$$p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{Q} \iff \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{pmatrix} \in \mathcal{S}^{1+n}$$

is not only one-to-one but also linear. To save notation, we identify the set $\mathcal{Q}$ of quadratic functions with the set $\mathcal{S}^{1+n}$ of $(1+n)\times(1+n)$ symmetric matrices and any subset of $\mathcal{Q}$ with the corresponding subset of $\mathcal{S}^{1+n}$. Specifically, we write $P = \left(\begin{smallmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{smallmatrix}\right) \in \mathcal{P}$ whenever $p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}$ and identify the set of $(1+n)\times(1+n)$ symmetric matrices

$$\left\{ \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{pmatrix} : \gamma \in R,\ \boldsymbol{q} \in R^n,\ \boldsymbol{Q} \in \mathcal{S}^n \right\}$$

with the set $\mathcal{Q}$ of quadratic functions from R^n to R.

2.3. SDP relaxation. Let $\mathcal{P}$ be a semi-infinite quadratic inequality representation of F:

$$\begin{aligned} F(\mathcal{P}) &= \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \le 0\ \forall p(\cdot) \in \mathcal{P}\} \\ &= \left\{ \boldsymbol{x} \in R^n : \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \le 0\ \forall \boldsymbol{P} \in \mathcal{P} \right\}. \end{aligned}$$

The SDP relaxation $\hat{F}(\mathcal{P})$ of $F(\mathcal{P})$ with the quadratic inequality representation $\mathcal{P}$ is given by

$$\begin{aligned} \hat{F}(\mathcal{P}) &\equiv \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \text{ and} \\ \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \le 0 \quad \forall p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P} \end{array} \right\} \\ &= \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \text{ and} \\ \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \le 0 \quad \forall \boldsymbol{P} \in \mathcal{P} \end{array} \right\}. \end{aligned}$$

If $\boldsymbol{x} \in F(\mathcal{P})$ and $\boldsymbol{P} \in \mathcal{P}$, then $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$ satisfies that

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} (1, \boldsymbol{x}^T) \in \mathcal{S}^{1+n}_+ \text{ and } \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \le 0.$$

This implies that $\boldsymbol{x} \in \hat{F}(\mathcal{P})$ and $F(\mathcal{P}) \subseteq \hat{F}(\mathcal{P})$. We also see that $\hat{F}(\mathcal{P})$ is convex. Hence c.hull$(F(\mathcal{P})) \subseteq \hat{F}(\mathcal{P})$. The SDP relaxation was originally proposed for combinatorial optimization problems and 0-1 integer programming problems [12], and later extended to quadratic optimization problems. See [1, 6, 8, 9, 15, 19, 18, 23, 24, 29].

3. Main results. Now we are ready to describe our method for approximating a quadratic-inequality-constrained compact feasible region F of the minimization problem (1.1). Before running the method, we need to fix a semi-infinite quadratic inequality representation $\mathcal{P}_F$ of F, and choose an initial approximation C_0 of the convex hull of F, i.e., a compact convex set which contains c.hull(F). Starting from C_0, the method generates a sequence of compact convex sets C_k $(k = 0, 1, 2, \dots)$, which we expect to converge to c.hull(F). At each iteration, we choose a semi-infinite quadratic inequality representation $\mathcal{P}_k$ of the kth approximation C_k of c.hull(F). Since c.hull$(F) \subseteq C_k$, the union $(\mathcal{P}_F \cup \mathcal{P}_k)$ forms a semi-infinite quadratic inequality representation of F. We then apply the SDP relaxation to $(\mathcal{P}_F \cup \mathcal{P}_k)$ to generate the next iterate $C_{k+1} = \hat{F}(\mathcal{P}_F \cup \mathcal{P}_k)$. It should be emphasized that during none of the iterations do we modify or strengthen the representation $\mathcal{P}_F$ directly. We only utilize the semi-infinite quadratic inequality representation of the compact convex set C_k that has been computed in the previous iteration.

Successive SDP Relaxation Method.

Step 0: Let $k = 0$.

Step 1: If $C_k = \emptyset$ or $C_k = $ c.hull(F), then stop.

Step 2: Choose a semi-infinite quadratic inequality representation $\mathcal{P}_k$ for C_k.

Step 3: Let

$$C_{k+1} = \hat{F}(\mathcal{P}_F \cup \mathcal{P}_k) = \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}_+^{1+n} \\ \text{and} \\ \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \leq 0 \quad \forall \boldsymbol{P} \in \mathcal{P}_F \cup \mathcal{P}_k \end{array} \right\}. \tag{3.1}$$

Step 4: Let $k = k + 1$, and go to Step 1.

We state two convergence theorems below. We choose the spherical inequality representation $\mathcal{P}^S(C_k)$ for C_k at Step 2 of each iteration in the first theorem, while we choose the rank-2 quadratic inequality representation $\mathcal{P}^2(C_k)$ for C_k at Step 2 of each iteration in the second theorem. Their proofs will be given in section 5.

Theorem 3.1. *Assume that $\mathcal{P}_F$ is a semi-infinite quadratic inequality representation of a compact subset F of R^n, and that $C_0 \supseteq F$ is a compact convex subset of R^n. If we choose $\mathcal{P}_k = \mathcal{P}^S(C_k)$ at Step 2 of each iteration in the successive SDP relaxation method, then the monotonicity property* (a) *and the asymptotic convergence property* (b) *stated in the introduction hold.*

Theorem 3.2. *Under the same assumptions as in Theorem* 3.1*, if we choose $\mathcal{P}_k = \mathcal{P}^2(C_k)$ at Step 2 of each iteration in the successive SDP relaxation method, then* (a) *and* (b) *remain valid.*

We know that if $\mathcal{P} \subset \mathcal{Q}$ and $\mathcal{P}' \subset \mathcal{Q}$ are semi-infinite quadratic inequality representations of C_k and if $\mathcal{P} \subset \mathcal{P}'$, then $\hat{F}(\mathcal{P}') \subseteq \hat{F}(\mathcal{P})$. Hence, even if we replace "$\mathcal{P}_k = \mathcal{P}^S(C_k)$" in Theorem 3.1 by "$\mathcal{P}_k \supseteq \mathcal{P}^S(C_k)$" (or "$\mathcal{P}_k = \mathcal{P}^2(C_k)$" in Theorem 3.2 by "$\mathcal{P}_k \supseteq \mathcal{P}^2(C_k)$"), the properties (a) and (b) remain valid. In particular, (a) and (b) remain valid when we choose any of $\mathcal{P}^E(C_k)$, $\mathcal{P}^C(C_k)$, and $\mathcal{P}^\sharp(C_k)$ for $\mathcal{P}_k$.

If we take the linear representation $\mathcal{P}^L(C_k)$ of C_k at every iteration, then we can prove that

$$C_1 = \tilde{F}(\mathcal{P}_F \cup \mathcal{P}_0) = \tilde{F}(\mathcal{P}_F) \cap C_0 \quad \text{and} \quad C_{k+1} = \tilde{F}(\mathcal{P}_F) \cap C_k = C_1 \quad (k = 1, 2, \dots).$$

(See Lemma 4.1.) Hence (b) does not follow in general.

In section 8, we will give two numerical examples. The first example shows that the rank-1 quadratic inequality representation $\mathcal{P}_k = \mathcal{P}^1(C_k)$ is not strong enough to ensure (b). The second example shows that even when we choose the strongest quadratic inequality representation $\mathcal{P}^\sharp(C_k)$ of C_k for $\mathcal{P}_k$ at every iteration, not only does the convergence "$C_k \to$ c.hull(F)" require infinitely many iterations, but its speed also becomes extremely slow in the worst case.

4. Fundamental characterization of successive convex relaxation.

4.1. Semi-infinite convex QOP relaxation and its equivalence to SDP relaxation.

The semi-infinite convex QOP relaxation of $F(\mathcal{P})$ with the semi-infinite quadratic inequality representation $\mathcal{P}$ is defined as

$$\begin{aligned}\tilde{F}(\mathcal{P}) &\equiv \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \text{c.cone}(\mathcal{P}) \cap \mathcal{Q}_+\} \\ &= \left\{\boldsymbol{x} \in R^n : \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0 \ \forall \boldsymbol{P} \in \text{c.cone}(\mathcal{P}) \cap \mathcal{Q}_+ \right\}.\end{aligned}$$

We observe that

$$F(\mathcal{P}) = \left\{\boldsymbol{x} \in R^n : \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0 \ \forall \boldsymbol{P} \in \text{c.cone}(\mathcal{P}) \right\} \subseteq \tilde{F}(\mathcal{P})$$

and that the set $\tilde{F}(\mathcal{P})$ is a closed convex set. Hence $F(\mathcal{P}) \subseteq \text{c.hull}(F(\mathcal{P})) \subseteq \tilde{F}(\mathcal{P})$.

The semi-infinite convex QOP relaxation was introduced by Fujie and Kojima [6]. It was called the relaxation using convex-quadratic valid inequalities for $F(\mathcal{P})$ in their paper [6]. The following basic properties of the relaxation are essentially due to them.

LEMMA 4.1. *Let $\mathcal{P}_F$ be a semi-infinite quadratic inequality representation of a closed set $F \subset R^n$.*

(i) *Let $\mathcal{P}$ be a set of convex quadratic valid inequalities for F, i.e., $\mathcal{P} \subseteq \mathcal{P}^C(F)$. Then*

$$\tilde{F}(\mathcal{P}_F \cup \mathcal{P}) \subseteq \tilde{F}(\mathcal{P}) = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}\}.$$

(ii) *Let $\mathcal{P}$ be a set of linear valid inequalities for F, i.e., $\mathcal{P} \subseteq \mathcal{P}^L(F)$. Then*

$$\tilde{F}(\mathcal{P}_F \cup \mathcal{P}) = \tilde{F}(\mathcal{P}_F) \cap \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}\}.$$

(iii) *Let $\boldsymbol{x}' \notin$ c.hull(F). Suppose that $p(\boldsymbol{x}'; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \geq 0$ for some $p(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}_F$ with a positive definite $\boldsymbol{Q}$. Then $\boldsymbol{x}' \notin \tilde{F}(\mathcal{P}_F)$.*

Proof. Part (i) follows directly from the definition of the semi-infinite convex QOP relaxation. Now we show (ii). Let $C = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}\}$. Then we see that

$$\tilde{F}(\mathcal{P}_F \cup \mathcal{P}) \subseteq \tilde{F}(\mathcal{P}_F) \cap \tilde{F}(\mathcal{P}) = \tilde{F}(\mathcal{P}_F) \cap C.$$

Hence it suffices to show that $\tilde{F}(\mathcal{P}_F) \cap C \subseteq \tilde{F}(\mathcal{P}_F \cup \mathcal{P})$. Let $p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}) \cap \mathcal{Q}_+$. Then there exist $p(\cdot)_i \in \mathcal{P}_F$ $(i = 1, 2, \ldots, \ell)$, $p(\cdot)_j \in \mathcal{P}$ $(j = \ell + 1, \ldots, m)$, and positive numbers λ_i $(i = 1, 2, \ldots, m)$ such that

$$p(\cdot) = \sum_{i=1}^{\ell} \lambda_i p(\cdot)_i + \sum_{j=\ell+1}^{m} \lambda_i p(\cdot)_i \in \mathcal{Q}_+.$$

Since $p(\cdot)_j \in \mathcal{P}$ $(j = \ell + 1, \ldots, m)$ are linear functions, we see that

$$\sum_{i=1}^{\ell} \lambda_i p(\cdot)_i \in \text{c.cone}(\mathcal{P}_F) \cap \mathcal{Q}_+; \quad \text{hence,} \ \sum_{i=1}^{\ell} \lambda_i p(\boldsymbol{x})_i \leq 0 \quad \forall \boldsymbol{x} \in \tilde{F}(\mathcal{P}_F).$$

Moreover,

$$\sum_{j=\ell+1}^{m} \lambda_i p(\cdot)_i \in \text{c.cone}(\mathcal{P}) \cap \mathcal{Q}_+; \quad \text{hence,} \quad \sum_{j=\ell+1}^{m} \lambda_i p(\boldsymbol{x})_i \leq 0 \quad \forall \boldsymbol{x} \in C.$$

Therefore,

$$p(\boldsymbol{x}) = \sum_{i=1}^{\ell} \lambda_i p_i(\boldsymbol{x}) + \sum_{j=\ell+1}^{m} \lambda_i p_i(\boldsymbol{x}) \leq 0 \quad \forall \boldsymbol{x} \in \tilde{F}(\mathcal{P}_F) \cap C.$$

This proves (ii). Finally we will show (iii). Since $\boldsymbol{x}' \notin F$, there is a $p'(\cdot) \in \mathcal{P}_F$ such that $p'(\boldsymbol{x}') > 0$. Hence, if $\epsilon > 0$ is sufficiently small, we obtain that

$$\epsilon p(\cdot)' + p(\cdot) \in \text{c.cone}(\mathcal{P}_F) \cap \mathcal{Q}_+ \quad \text{and} \quad \epsilon p'(\boldsymbol{x}') + p(\boldsymbol{x}') > 0.$$

This implies $\boldsymbol{x}' \notin \tilde{F}(\mathcal{P}_F)$, and proves (iii). ◻

When $\mathcal{P}$ is finite and $F(\mathcal{P})$ satisfies Slater's constraint qualification, Fujie and Kojima [6] showed that the semi-infinite convex QOP relaxation is essentially equivalent to the SDP relaxation in the sense that $\tilde{F}(\mathcal{P})$ coincides with the closure of $\hat{F}(\mathcal{P})$. The theorem below shows the exact equivalence between them, without any constraint qualification, for more general semi-infinite quadratic inequality representation cases. Since $\tilde{F}(\mathcal{P})$ is closed, one of the consequences of the next theorem is that $\hat{F}(\mathcal{P})$ is always closed. Note that we can assume without loss of generality that $\mathcal{P}$ is a closed convex cone, since every closed set F admits such a representation.

THEOREM 4.2. *Let $\mathcal{P}$ be a closed convex cone, giving a semi-infinite quadratic inequality representation of a closed subset F of R^n; $F(\mathcal{P}) = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}\}$. Then its SDP relaxation and its semi-infinite convex QOP relaxation coincide with each other; $\hat{F}(\mathcal{P}) = \tilde{F}(\mathcal{P})$.*

Proof. Using the dual cone

$$\mathcal{P}^* = \{\boldsymbol{V} \in \mathcal{S} : \boldsymbol{V} \bullet \boldsymbol{U} \geq 0 \ \forall \boldsymbol{U} \in \mathcal{P}\}$$

of $\mathcal{P}$, we can express the sets $\hat{F}(\mathcal{P})$ and $\tilde{F}(\mathcal{P})$ as follows:

$$\hat{F}(\mathcal{P}) = \left\{ \boldsymbol{x} \in R^n : \exists \boldsymbol{X} \in \mathcal{S}^n \ \text{ such that } \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in (-\mathcal{P}^*) \cap \mathcal{S}_+^{1+n} \right\}$$

and

$$\begin{aligned} \tilde{F}(\mathcal{P}) &= \left\{ \boldsymbol{x} \in R^n : \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \in -\left(\mathcal{P} \cap \mathcal{Q}_+\right)^* \right\} \\ &= \left\{ \boldsymbol{x} \in R^n : \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \in -\left[\mathcal{P}^* + \begin{pmatrix} 0 & \boldsymbol{0}^T \\ \boldsymbol{0} & \mathcal{S}_+^n \end{pmatrix} \right] \right\}. \end{aligned}$$

For the last identity above, we have used the fact that for any pair of closed convex cones $\mathcal{K}_1$ and $\mathcal{K}_2$ in R^m, we have $(\mathcal{K}_1 \cap \mathcal{K}_2)^* = \mathcal{K}_1^* + \mathcal{K}_2^*$.

First let $\boldsymbol{x} \in \hat{F}(\mathcal{P})$. Then there exists an $\boldsymbol{X} \in \mathcal{S}^n$ such that

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in (-\mathcal{P}^*) \cap \mathcal{S}_+^{1+n}.$$

Consider the identity

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} = -\left[\begin{pmatrix} -1 & -\boldsymbol{x}^T \\ -\boldsymbol{x} & -\boldsymbol{X} \end{pmatrix} + \begin{pmatrix} 0 & -\boldsymbol{0}^T \\ -\boldsymbol{0} & \boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix}\right].$$

The first matrix on the right-hand side is in $\mathcal{P}^*$ and in the second matrix of the right-hand side, we have $\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^T \in \mathcal{S}^n_+$ since it is the Schur complement of 1 in the symmetric, positive semidefinite matrix $\left(\begin{smallmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{smallmatrix}\right)$. We have proved $\boldsymbol{x} \in \tilde{F}(\mathcal{P})$ and hence $\hat{F}(\mathcal{P}) \subseteq \tilde{F}(\mathcal{P})$.

For the converse, let $\boldsymbol{x} \in \tilde{F}(\mathcal{P})$; that is, there exists some $\boldsymbol{H} \in \mathcal{S}^n_+$ such that

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T + \boldsymbol{H} \end{pmatrix} \in -\mathcal{P}^*.$$

The matrix

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T + \boldsymbol{H} \end{pmatrix}$$

is positive semidefinite if and only if $(\boldsymbol{H} + \boldsymbol{x}\boldsymbol{x}^T - \boldsymbol{x}\boldsymbol{x}^T) = \boldsymbol{H}$ is. But the latter was already established. So,

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T + \boldsymbol{H} \end{pmatrix} \in (-\mathcal{P}^*) \cap \mathcal{S}^{1+n}_+.$$

Therefore $\boldsymbol{x} \in \hat{F}(\mathcal{P})$, and $\tilde{F}(\mathcal{P}) \subseteq \hat{F}(\mathcal{P})$ is proved. □

4.2. Semi-infinite LP relaxation. In section 7, we will also need an analog of the above theorem for our successive semi-infinite LP relaxation method. For every semi-infinite quadratic inequality representation $\mathcal{P}$ of a compact subset F of R^n, let us define

$$\hat{F}^L(\mathcal{P}) \equiv \left\{\boldsymbol{x} \in R^n : \exists X \in \mathcal{S}^n \ \text{ such that } \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \leq 0 \ \ \forall \boldsymbol{P} \in \mathcal{P}\right\}$$

and

$$\tilde{F}^L(\mathcal{P}) \equiv \left\{\boldsymbol{x} \in R^n : \gamma + 2\boldsymbol{q}^T\boldsymbol{x} \leq 0 \ \forall p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \text{c.cone}(\mathcal{P}) \cap \mathcal{L}\right\}$$

of Sherali and Alameddine [21]. Here, $\mathcal{L}$ denotes the set of linear functions on R^n:

$$\mathcal{L} \equiv \{p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{Q} : \boldsymbol{Q} = \boldsymbol{O}\}.$$

The next result can be obtained by following the steps of the proof of Theorem 4.2.

COROLLARY 4.3. *Let $\mathcal{P}$ be a closed convex cone, giving a semi-infinite quadratic inequality representation of a closed subset F of R^n; $F(\mathcal{P}) = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}\}$. Then $\hat{F}^L(\mathcal{P}) = \tilde{F}^L(\mathcal{P})$.*

Proof. We observe that

$$\hat{F}^L(\mathcal{P}) = \left\{\boldsymbol{x} \in R^n : \exists \boldsymbol{X} \in \mathcal{S}^n \ \text{ such that } \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in -\mathcal{P}^*\right\}$$

and

$$\begin{aligned}\tilde{F}^L(\mathcal{P}) &= \left\{ \boldsymbol{x} \in R^n : \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \in -\left(\mathcal{P} \cap \mathcal{L}\right)^* \right\} \\ &= \left\{ \boldsymbol{x} \in R^n : \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \in -\left[\mathcal{P}^* + \begin{pmatrix} 0 & \boldsymbol{0}^T \\ \boldsymbol{0} & \mathcal{S}^n \end{pmatrix} \right] \right\}.\end{aligned}$$

Since it is easy to see that $\exists \boldsymbol{X} \in \mathcal{S}^n$ such that $\left(\begin{smallmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{smallmatrix}\right) \in -\mathcal{P}^*$ if and only if

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \in -\left[\mathcal{P}^* + \begin{pmatrix} 0 & \boldsymbol{0}^T \\ \boldsymbol{0} & \mathcal{S}^n \end{pmatrix} \right],$$

the proof is complete. ◻

4.3. Invariance under one-to-one affine transformation. Let $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ be an arbitrary one-to-one affine transformation on R^n, where $\boldsymbol{A}$ is an $n \times n$ nonsingular matrix and $\boldsymbol{b} \in R^n$.

Then

$$\begin{aligned}\boldsymbol{f}(\hat{F}(\mathcal{P})) = \boldsymbol{f}(\tilde{F}(\mathcal{P})) &= \{\boldsymbol{y} \in R^n : p'(\boldsymbol{y}) \leq 0 \ \forall p'(\cdot) \in \text{c.cone}(\mathcal{P}') \cap \mathcal{Q}_+\}, \\ \boldsymbol{f}(\hat{F}^L(\mathcal{P})) = \boldsymbol{f}(\tilde{F}^L(\mathcal{P})) &= \{\boldsymbol{y} \in R^n : p'(\boldsymbol{y}) \leq 0 \ \forall p'(\cdot) \in \text{c.cone}(\mathcal{P}') \cap \mathcal{L}\},\end{aligned}$$

where $\mathcal{P}' \equiv \{p(\boldsymbol{f}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}\}$ forms a semi-infinite quadratic inequality representation of $\boldsymbol{f}(F(\mathcal{P}))$. This means that the semi-infinite SDP and LP relaxations are invariant under the one-to-one affine transformation $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$.

We also see that

$$\mathcal{P}^U(\boldsymbol{f}(C)) = \{p(\boldsymbol{f}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}^U(C)\}$$

holds, where $U \in \{L,\ 1,\ 2,\ E,\ C,\ \sharp\}$. Therefore, $\mathcal{P}^L(C)$, $\mathcal{P}^1(C)$, $\mathcal{P}^2(C)$, $\mathcal{P}^E(C)$, $\mathcal{P}^C(C)$, and $\mathcal{P}^\sharp(C)$ are invariant under one-to-one affine transformations on R^n. If in addition $\boldsymbol{A}$ is a scalar multiple of an orthogonal matrix, then the above identity also holds for $U = S$; hence $\mathcal{P}^S(C)$ is invariant under such a one-to-one affine transformation on R^n.

At each iteration of the successive SDP relaxation method, we observe that

$$\boldsymbol{f}(C_{k+1}) = \{\boldsymbol{y} \in R^n : p'(\boldsymbol{y}) \leq 0 \ \forall p'(\cdot) \in \text{c.cone}(\mathcal{P}'_F \cup \mathcal{P}'_k) \cap \mathcal{Q}_+\},$$

where $\mathcal{P}'_F \equiv \{p(\boldsymbol{f}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}_F\}$ forms a semi-infinite quadratic inequality representation of $\boldsymbol{f}(F)$ and $\mathcal{P}'_k \equiv \{p(\boldsymbol{f}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}_k\}$ forms a semi-infinite quadratic inequality representation of $\boldsymbol{f}(C_k)$. Furthermore, if we choose one of the invariant semi-infinite quadratic inequality representations $\mathcal{P}^L(C_k)$, $\mathcal{P}^1(C_k)$, $\mathcal{P}^2(C_k)$, $\mathcal{P}^E(C_k)$, $\mathcal{P}^C(C_k)$, and $\mathcal{P}^\sharp(C_k)$ of C_k under any one-to-one affine transformation for $\mathcal{P}_k$, we see that $\mathcal{P}^U(\boldsymbol{f}(C)) = \{p(\boldsymbol{f}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}^U(C)\}$; hence the identity above turns out to be

$$\boldsymbol{f}(C_{k+1}) = \left\{\boldsymbol{y} \in R^n : p'(\boldsymbol{y}) \leq 0 \ \forall p'(\cdot) \in \text{c.cone}(\mathcal{P}'_F \cup \mathcal{P}^U(\boldsymbol{f}(C_k))) \cap \mathcal{Q}_+\right\}.$$

Here $U \in \{L,\ 1,\ 2,\ E,\ C,\ \sharp\}$. Therefore the successive SDP relaxation method is invariant under any one-to-one affine transformation. The same comment applies to the successive semi-infinite LP relaxation method, which we will present in section 7.

5. Proofs of Theorems 3.1 and 3.2. We present three lemmas, Lemma 5.1 in section 5.1, Lemma 5.2 in section 5.2, and Lemma 5.3 in section 5.4. Lemma 5.1 proves the monotonicity property (a) in Theorems 3.1 and 3.2 simultaneously. Lemma 5.2 is used to prove Theorem 3.1 in section 5.3, and Lemma 5.3 to prove Theorem 3.2 in section 5.5.

5.1. Monotonicity. We first establish the monotonicity in general.

LEMMA 5.1. *Let C_0 be a compact convex set containing F. Fix a closed convex cone $\mathcal{S}^{1+n}_+ \subseteq \mathcal{K} \subseteq \mathcal{S}^{1+n}$ and $U \in \{L, 1, 2, S, E, C, \sharp\}$. Define*

$$C_{k+1} \equiv \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{K} \text{ and} \\ \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \leq 0 \ \ \forall \boldsymbol{P} \in \mathcal{P}_F \cup \mathcal{P}^U(C_k) \end{array} \right\}$$

for $k = 1, 2, \ldots$. Assume that

$$C_0 = \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{K} \text{ and} \\ \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \leq 0 \ \ \forall \boldsymbol{P} \in \mathcal{P}^U(C_0) \end{array} \right\}.$$

Then c.hull$(F) \subseteq C_{k+1} \subseteq C_k$ for all $k = 0, 1, 2, \ldots$.

Proof. Since $\mathcal{K} \supseteq \mathcal{S}^{1+n}_+$, it contains all symmetric rank-1 matrices of the form

$$\begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix}.$$

Now, as in the arguments in section 2.3, it follows that c.hull$(F) \subseteq C_k$ for all $k = 0, 1, 2, \ldots$. We will show by induction that $C_{k+1} \subseteq C_k$ for all $k = 0, 1, \ldots$. By the construction of C_1 and the assumption imposed on C_0, we first observe that $C_1 \subseteq C_0$. Now assume that $C_k \subseteq C_{k-1}$ for some $k \geq 1$. Then $\mathcal{P}^U(C_{k-1}) \subseteq \mathcal{P}^U(C_k)$, which implies that $\mathcal{P}_F \cup \mathcal{P}^U(C_{k-1}) \subseteq \mathcal{P}_F \cup \mathcal{P}^U(C_k)$. Therefore, $C_{k+1} \subseteq C_k$, as desired. ∎

5.2. Separating hypersphere. The following lemma easily follows from the separating hyperplane theorem, and the proof is omitted here.

LEMMA 5.2. *Let C be a compact convex subset of R^n and $\boldsymbol{x}' \notin C$. Then there exists a hypersphere $S \equiv \{\boldsymbol{x} \in R^n : \|\boldsymbol{x} - \boldsymbol{d}\| = \eta\}$ which strictly separates the point $\boldsymbol{x}'$ and C such that*

$$\|\boldsymbol{x}' - \boldsymbol{d}\| > \eta > \|\boldsymbol{x} - \boldsymbol{d}\| \quad \forall \boldsymbol{x} \in C, \tag{5.1}$$

where $\boldsymbol{d} \in R^n$ and $\eta > 0$.

5.3. Proof of Theorem 3.1. The monotonicity property (a) follows from Lemma 5.1 by letting $\mathcal{K} \equiv \mathcal{S}^{1+n}_+$ and $U \equiv S$. Let $C \equiv \cap_{k=0}^{\infty} C_k$. We know by (a) that c.hull$(F) \subseteq C \subseteq C_{k+1} \subseteq C_k$ $(k = 0, 1, \ldots)$, and that all the sets c.hull(F), C, and C_k are compact sets. To prove (b), we have the following left to show: $C \subseteq$ c.hull(F). Assume on the contrary that there exists some $\boldsymbol{x}' \in C$ such that $\boldsymbol{x}' \notin$ c.hull(F). Then, by Lemma 5.2, there exists a hypersphere $S \equiv \{\boldsymbol{x} \in R^n : \|\boldsymbol{x} - \boldsymbol{d}\| = \eta\}$ that strictly separates the point $\boldsymbol{x}' \in C$ from c.hull(F) such that

$$\|\boldsymbol{x}' - \boldsymbol{d}\| > \eta > \|\boldsymbol{x} - \boldsymbol{d}\| \quad \forall \boldsymbol{x} \in \text{c.hull}(F),$$

where $\boldsymbol{d} \in R^n$ and $\eta > 0$. Let $\eta^* \equiv \sup\{\|\boldsymbol{x} - \boldsymbol{d}\| : \boldsymbol{x} \in C\}$. Obviously, $\eta < \eta^* = \|\boldsymbol{x}^* - \boldsymbol{d}\|$ for some $\boldsymbol{x}^* \in C$. Since $\boldsymbol{x}^* \notin \text{c.hull}(F)$, there is a quadratic function, $p_1(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}_F$ that cuts off $\boldsymbol{x}^*$; $0 < p_1(\boldsymbol{x}^*; \gamma, \boldsymbol{q}, \boldsymbol{Q})$. Note that if $p_1(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q})$ is such a quadratic function, then so is $\alpha p_1(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q})$ for any $\alpha > 0$. Hence we may assume that the minimum eigenvalue of the matrix $\boldsymbol{Q} \in \mathcal{S}^n$ is at least (-1). Now consider a quadratic function $p_2(\cdot)$ defined by

$$p_2(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{d})^T(\boldsymbol{x} - \boldsymbol{d}) - (\eta^*)^2 - p_1(\boldsymbol{x}^*; \gamma, \boldsymbol{q}, \boldsymbol{Q})/2 \quad \forall \boldsymbol{x} \in R^n.$$

By the definition of η^*, we see that

$$p_2(\boldsymbol{x}) \leq -p_1(\boldsymbol{x}^*)/2 < 0 \quad \forall \boldsymbol{x} \in C.$$

This means that the open ball $B_+ \equiv \{\boldsymbol{x} \in R^n : p_2(\boldsymbol{x}) < 0\}$ with the center $\boldsymbol{d}$ and the radius $\sqrt{(\eta^*)^2 + p_1(\boldsymbol{x}^*; \gamma, \boldsymbol{q}, \boldsymbol{Q})/2}$ forms a neighborhood of the compact set C. On the other hand, the sequence $\{C_k\}$ of compact subsets of R^n satisfies

$$C_{k+1} \subseteq C_k \ (k = 0, 1, 2, \dots) \quad \text{ and } \quad C = \cap_{k=0}^{\infty} C_k.$$

So, we can find a finite positive number ℓ such that the open ball B_+ contains C_ℓ. Hence, $p_2(\boldsymbol{x}) \leq 0$ is a convex quadratic valid inequality for C_ℓ; $p_2(\cdot) \in \mathcal{P}_\ell$. We also see that

$$p_1(\boldsymbol{x}^*; \gamma, \boldsymbol{q}, \boldsymbol{Q}) + p_2(\boldsymbol{x}^*) = p_1(\boldsymbol{x}^*; \gamma, \boldsymbol{q}, \boldsymbol{Q})/2 > 0 \ \text{ and } p_1(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) + p_2(\cdot) \in \mathcal{Q}_+.$$

Thus we have shown that

$$p_1(\boldsymbol{x}^*; \gamma, \boldsymbol{q}, \boldsymbol{Q}) + p_2(\boldsymbol{x}^*) > 0 \ \text{ and } p_1(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) + p_2(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_\ell) \cap \mathcal{Q}_+.$$

Therefore, $\boldsymbol{x}^* \notin C_{\ell+1} = \tilde{F}(\mathcal{P}_F \cup \mathcal{P}_\ell)$, so that $\boldsymbol{x}^* \notin C = \cap_{k=0}^{\infty} C_k$. This is a contradiction. The theorem is proved.

5.4. A family of inequalities of the convex cone of rank-2 quadratic valid inequalities for the unit ball. Let B denote the unit ball $\{\boldsymbol{x} \in R^n : \|\boldsymbol{x}\| \leq 1\}$. Let $\boldsymbol{Q}$ be an arbitrary $n \times n$ symmetric matrix, and let $\boldsymbol{u} \in R^n$ be an arbitrary vector on the boundary of B; $\|\boldsymbol{u}\| = 1$. We will construct a family of quadratic valid inequalities, which lie in the convex cone of rank-2 quadratic valid inequalities, $p_\theta(\boldsymbol{x}) \leq 0$, with a parameter $\theta \in (0, \pi/8)$ for the unit ball B satisfying the properties (i), (ii), and (iii) listed in Lemma 5.3.

We first apply the eigenvalue decomposition to the matrix $\boldsymbol{Q} \in \mathcal{S}^n$. We may assume that the first m eigenvalues are nonnegative and the last $n - m$ eigenvalues are nonpositive for some nonnegative integer $m \leq n$. Then we can write the matrix $\boldsymbol{Q} \in \mathcal{S}^n$ as

$$\boldsymbol{Q} = \sum_{j=1}^{m} \mu_j \boldsymbol{r}_j \boldsymbol{r}_j^T - \sum_{j=m+1}^{n} \mu_j \boldsymbol{r}_j \boldsymbol{r}_j^T,$$

where $\|\boldsymbol{r}_j\| = 1 \ (j = 1, 2, \dots, n)$ and $\mu_j \geq 0 \ (j = 1, 2, \dots, n)$, $\boldsymbol{r}_j \ (j = 1, 2, \dots, n)$ denote eigenvectors of $\boldsymbol{Q}$, which are orthogonal to each other, and $\mu_j \ (j = 1, 2, \dots, m)$ and $-\mu_j \ (j = m + 1, \dots, n)$ denote the eigenvalues corresponding to them.

For each $\theta \in (0, \pi/8)$, we define

$$(5.2)\qquad \begin{cases} \boldsymbol{a}_j(\theta) & \equiv & \boldsymbol{u}\cos\theta + \boldsymbol{r}_j \sin\theta \quad (j = 1, 2, \ldots, n), \\ \bar{\boldsymbol{a}}_j(\theta) & \equiv & \boldsymbol{u}\cos\theta - \boldsymbol{r}_j \sin\theta \quad (j = 1, 2, \ldots, n), \\ \boldsymbol{b}_j & \equiv & +\boldsymbol{r}_j, \ \bar{\boldsymbol{b}}_j \equiv -\boldsymbol{r}_j \quad (j = 1, 2, \ldots, m), \\ \boldsymbol{b}_j & \equiv & -\boldsymbol{r}_j, \ \bar{\boldsymbol{b}}_j \equiv +\boldsymbol{r}_j \quad (j = m+1, \ldots, n), \\ \alpha_j(\theta) & \equiv & \max\{\boldsymbol{a}_j(\theta)^T\boldsymbol{x} : \boldsymbol{x} \in B\} = \|\boldsymbol{a}_j(\theta)\| \quad (j = 1, 2, \ldots, n), \\ \bar{\alpha}_j(\theta) & \equiv & \max\{\bar{\boldsymbol{a}}_j(\theta)^T\boldsymbol{x} : \boldsymbol{x} \in B\} = \|\bar{\boldsymbol{a}}_j(\theta)\| \quad (j = 1, 2, \ldots, n), \\ \beta_j & \equiv & \max\{\boldsymbol{b}_j^T\boldsymbol{x} : \boldsymbol{x} \in B\} = \|\boldsymbol{b}_j\| = 1 \quad (j = 1, 2, \ldots, n), \\ \bar{\beta}_j & \equiv & \max\{\bar{\boldsymbol{b}}_j^T\boldsymbol{x} : \boldsymbol{x} \in B\} = \|\bar{\boldsymbol{b}}_j\| = 1 \quad (j = 1, 2, \ldots, n), \\ \lambda_j(\theta) & \equiv & \dfrac{\mu_j}{2\sin\theta} \geq 0 \quad (j = 1, 2, \ldots, n). \end{cases}$$

Then, $\forall\theta \in (0, \pi/8)$ and $j = 1, 2, \ldots, n$, $\boldsymbol{a}_j(\theta)$, $\bar{\boldsymbol{a}}_j(\theta)$, $\boldsymbol{b}_j(\theta)$, and $\bar{\boldsymbol{b}}_j(\theta)$ are nonzero vectors, and

$$(5.3)\qquad \begin{cases} \boldsymbol{a}_j(\theta)^T\boldsymbol{x} - \alpha_j(\theta) & \leq 0, \quad \boldsymbol{b}_j^T\boldsymbol{x} - \beta_j \leq 0, \\ \bar{\boldsymbol{a}}_j(\theta)^T\boldsymbol{x} - \bar{\alpha}_j(\theta) & \leq 0, \quad \bar{\boldsymbol{b}}_j^T\boldsymbol{x} - \bar{\beta}_j \leq 0 \end{cases}$$

are linear valid inequalities for the unit ball B. For all $\theta \in (0, \pi/8)$, define

$$(5.4)\qquad p_\theta(\boldsymbol{x}) \equiv -\sum_{j=1}^{n} \lambda_j(\theta)\Big((\boldsymbol{a}_j(\theta)^T\boldsymbol{x} - \alpha_j(\theta))(\boldsymbol{b}_j^T\boldsymbol{x} - \beta_j) + (\bar{\boldsymbol{a}}_j(\theta)^T\boldsymbol{x} - \bar{\alpha}_j(\theta))(\bar{\boldsymbol{b}}_j^T\boldsymbol{x} - \bar{\beta}_j)\Big).$$

Then $p_\theta(\cdot) \in \text{c.cone}(\mathcal{P}^2(B))$ for all $\theta \in (0, \pi/8)$. In particular, $p_\theta(\boldsymbol{u}) \leq 0\ \forall\theta \in (0, \pi/8)$.

LEMMA 5.3.
(i) *$p_\theta(\cdot) \in c.cone(\mathcal{P}^2(B))$.*
(ii) *$p_\theta(\boldsymbol{u}) \to 0$ as $\theta \in (0, \pi/8)$ tends to 0.*
(iii) *The Hessian matrix of $p_\theta(\cdot)$ coincides with $-\boldsymbol{Q}$.*

Proof. Part (i) was already shown.

(ii) Let j be fixed. It suffices to show that

$$\epsilon_j(\theta) \equiv \lambda_j(\theta)(\boldsymbol{a}_j(\theta)^T\boldsymbol{u} - \alpha_j(\theta))(\boldsymbol{b}_j^T\boldsymbol{u} - \beta_j) \quad \text{and}$$
$$\bar{\epsilon}_j(\theta) \equiv \lambda_j(\theta)(\bar{\boldsymbol{a}}_j(\theta)^T\boldsymbol{u} - \bar{\alpha}_j(\theta))(\bar{\boldsymbol{b}}_j^T\boldsymbol{u} - \bar{\beta}_j)$$

converge to zero as $\theta \in (0, \pi/8)$ tends to 0. First, we derive that $\epsilon_j(\theta)$ converges to zero as $\theta \in (0, \pi/8)$ tends to 0. We see from (5.2) that

$$(5.5)\qquad \begin{aligned} \epsilon_j(\theta) &= \frac{\mu_j(\cos\theta + \boldsymbol{u}^T\boldsymbol{r}_j\sin\theta - \|\boldsymbol{u}\cos\theta + \boldsymbol{r}_j\sin\theta\|)}{2\sin\theta}(\boldsymbol{b}_j^T\boldsymbol{u} - 1) \\ &= \frac{\mu_j\left(\cos\theta + \boldsymbol{u}^T\boldsymbol{r}_j\sin\theta - (\cos^2\theta + 2\boldsymbol{u}^T\boldsymbol{r}_j\sin\theta\cos\theta + \sin^2\theta)^{\frac{1}{2}}\right)}{2\sin\theta} \\ &\quad \times (\boldsymbol{b}_j^T\boldsymbol{u} - 1). \end{aligned}$$

Since both the numerator and the denominator above converge to zero as $\theta \in (0, \pi/8)$ tends to 0, we calculate their derivatives at $\theta = 0$. The derivative of the numerator turns out to be

$$\mu_j \left(-\sin\theta + \boldsymbol{u}^T\boldsymbol{r}_j \cos\theta + \frac{\boldsymbol{u}^T\boldsymbol{r}_j(\sin^2\theta - \cos^2\theta)}{(2\boldsymbol{u}^T\boldsymbol{r}_j \sin\theta\cos\theta + 1)^{1/2}} \right) (\boldsymbol{b}_j^T\boldsymbol{u} - 1),$$

which vanishes at $\theta = 0$. On the other hand, the derivative "$2\cos\theta$" of the denominator "$2\sin\theta$" in (5.5) does not vanish at $\theta = 0$. Thus, $\epsilon_j(\theta)$ converges to 0 as $\theta \in (0, \pi/8)$ tends to 0. Similarly, we can prove that $\bar{\epsilon}_j(\theta)$ converges to 0 as $\theta \in (0, \pi/8)$ tends to 0.

(iii) It follows from the definitions (5.2) and (5.4) that the Hessian matrix of the quadratic function $p_\theta(\cdot)$

$$\begin{aligned}
&= -\sum_{j=1}^{n} \lambda_j(\theta) \frac{\boldsymbol{a}_j(\theta)\boldsymbol{b}_j^T + \boldsymbol{b}_j\boldsymbol{a}_j^T(\theta) + \bar{\boldsymbol{a}}_j(\theta)\bar{\boldsymbol{b}}_j^T + \bar{\boldsymbol{b}}_j\bar{\boldsymbol{a}}_j(\theta)^T}{2} \\
&= -\sum_{j=1}^{m} \mu_j\boldsymbol{r}_j\boldsymbol{r}_j^T + \sum_{j=m+1}^{n} \mu_j\boldsymbol{r}_j\boldsymbol{r}_j^T \\
&= -\boldsymbol{Q}. \qquad \square
\end{aligned}$$

From the lemma above, we see that the cone $\mathcal{P}^2(B)$ is rich enough to contain rank-2 quadratic functions with any prescribed Hessian, leading to valid inequalities that are tight at any given point on the boundary of B.

5.5. Proof of Theorem 3.2. The monotonicity property (a) follows from Lemma 5.1 by letting $\mathcal{K} \equiv \mathcal{S}_+^{1+n}$ and $U \equiv 2$. To derive (b), it suffices to show that $C \equiv \cap_{k=0}^{\infty} C_k \subseteq \text{c.hull}(F)$ as in the proof of Theorem 3.1. Assume on the contrary that $\boldsymbol{x}' \notin \text{c.hull}(F)$ for some $\boldsymbol{x}' \in C$. By Lemma 5.2, there exists a hypersphere $S \equiv \{\boldsymbol{x} \in R^n : \|\boldsymbol{x} - \boldsymbol{d}\| = \eta\}$ which strictly separates the point $\boldsymbol{x}' \in C$ and c.hull(F) such that

$$\|\boldsymbol{x}' - \boldsymbol{d}\| > \delta > \|\boldsymbol{x} - \boldsymbol{d}\| \quad \forall \boldsymbol{x} \in \text{c.hull}(F),$$

where $\boldsymbol{d} \in R^n$ and $\delta > 0$. Let $\delta^* \equiv \sup\{\|\boldsymbol{x} - \boldsymbol{d}\| : \boldsymbol{x} \in C\}$. Obviously, $\delta^* = \|\boldsymbol{u} - \boldsymbol{d}\| > \delta$ for some $\boldsymbol{u} \in C$. Since the successive SDP relaxation method using the rank-2 quadratic representation for C_k at each iteration is invariant under the affine transformation $(\boldsymbol{x} - \boldsymbol{d})/\delta^* \to \boldsymbol{x}'$, which maps $\boldsymbol{d}$ to the origin and the hypersphere $S \equiv \{\boldsymbol{x} \in R^n : \|\boldsymbol{x} - \boldsymbol{d}\| = \delta^*\}$ onto the unit hypersphere $\{\boldsymbol{x}' \in R^n : \|\boldsymbol{x}'\| = 1\}$, we may assume that $\boldsymbol{d} = \boldsymbol{0}$ and $\delta^* = 1$. Thus, we have obtained that

$$C \subseteq B \equiv \{\boldsymbol{x} \in R^n : \|\boldsymbol{x}\| \leq 1\} \quad \text{and} \quad \boldsymbol{u} \in C,\ \boldsymbol{u} \notin \text{c.hull}(F),\ \|\boldsymbol{u}\| = 1.$$

Since $\boldsymbol{u} \notin F$, there is a quadratic function $p_1(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}_F$ that cuts off $\boldsymbol{u}$; $p_1(\boldsymbol{u}; \gamma, \boldsymbol{q}, \boldsymbol{Q}) > 0$. Now, let $p_\theta(\cdot) \in \mathcal{P}^2(B) \cap \mathcal{Q}_+$ be the quadratic function introduced in section 5.4. See (5.2) and (5.4). By Lemma 5.3, we can choose a $\theta \in (0, \pi/8)$ for which $p_\theta(\boldsymbol{u}) \geq -p_1(\boldsymbol{u}; \gamma, \boldsymbol{q}, \boldsymbol{Q})/3$ holds. Now we define

$$\begin{aligned}
\alpha_j^k &= \max\{\boldsymbol{a}_j(\theta)^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}, \quad \beta_j^k = \max\{\boldsymbol{b}_j(\theta)^T\boldsymbol{x} : \boldsymbol{x} \in C_k\} \quad (1 \leq j \leq n), \\
\bar{\alpha}_j^k &= \max\{\bar{\boldsymbol{a}}_j(\theta)^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}, \quad \bar{\beta}_j^k = \max\{\bar{\boldsymbol{b}}_j(\theta)^T\boldsymbol{x} : \boldsymbol{x} \in C_k\} \quad (1 \leq j \leq n),
\end{aligned}$$

$$\begin{aligned}
p'_k(\boldsymbol{x}) = -\sum_{j=1}^{n} \lambda_j(\theta) \big(&(\boldsymbol{a}_j(\theta)^T\boldsymbol{x} - \alpha_j^k)(\boldsymbol{b}_j(\theta)^T\boldsymbol{x} - \beta_j^k) \\
&+ (\bar{\boldsymbol{a}}_j(\theta)^T\boldsymbol{x} - \bar{\alpha}_j^k)(\bar{\boldsymbol{b}}_j(\theta)^T\boldsymbol{x} - \bar{\beta}_j^k) \big)
\end{aligned}$$

for $k = 0, 1, 2, \dots$. By construction, we know that $p'_k(\cdot) \in \text{c.cone}(\mathcal{P}^2(C_k))$. Since both quadratic functions $p_\theta(\cdot)$ and $p'_k(\cdot)$ have the common Hessian matrix $-\boldsymbol{Q}$,

$$p_1(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) + p'_k(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{L} \subset \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+ \quad \forall k = 0, 1, 2, \dots.$$

We will show that

$$p_1(\boldsymbol{u}) + p'_k(\boldsymbol{u}) \geq p_1(\boldsymbol{u})/3 > 0 \tag{5.6}$$

for every sufficiently large k. Then the above two relations imply $\boldsymbol{u} \notin C_{k+1}$ for such a large k. This contradicts the fact $\boldsymbol{u} \in C = \cap_{k=0}^{\infty} C_k$.

Since the sequence of compact convex subsets C_k $(k = 0, 1, 2, \dots)$ satisfies

$$C_{k+1} \subseteq C_k \ (k = 0, 1, 2, \dots) \ \text{ and } \ \cap_{k=0}^{\infty} C_k = C \subseteq B = \{\boldsymbol{x} : \|\boldsymbol{x}\| \leq 1\},$$

we see that

$$\begin{aligned}
&\alpha_j^k \to \alpha_j^* \equiv \max\{\boldsymbol{a}_j(\theta)^T \boldsymbol{x} : \boldsymbol{x} \in C\} \leq \alpha_j(\theta), \ \beta_j^k \to \beta_j^* \equiv \max\{\boldsymbol{b}_j(\theta)^T \boldsymbol{x} : \boldsymbol{x} \in C\} \leq \beta_j, \\
&\bar{\alpha}_j^k \to \bar{\alpha}_j^* \equiv \max\{\bar{\boldsymbol{a}}_j(\theta)^T \boldsymbol{x} : \boldsymbol{x} \in C\} \leq \bar{\alpha}_j(\theta), \ \bar{\beta}_j^k \to \bar{\beta}_j^* \equiv \max\{\bar{\boldsymbol{b}}_j(\theta)^T \boldsymbol{x} : \boldsymbol{x} \in C\} \leq \bar{\beta}_j
\end{aligned}$$

as $k \to \infty$ $(j = 2, 3, \dots, n)$. By continuity, we see then that for every sufficiently large k

$$\begin{aligned}
p'_k(\boldsymbol{u}) &\geq -p_1(\boldsymbol{u})/3 - \sum_{j=1}^{n} \lambda_j(\theta) \big((\boldsymbol{a}_j(\theta)^T \boldsymbol{u} - \alpha_j^*)(\boldsymbol{b}_j(\theta)^T \boldsymbol{u} - \beta_j^*) \\
&\qquad\qquad + (\bar{\boldsymbol{a}}_j(\theta)^T \boldsymbol{u} - \bar{\alpha}_j^*)(\bar{\boldsymbol{b}}_j(\theta)^T \boldsymbol{u} - \bar{\beta}_j^*) \big) \\
&\geq -p_1(\boldsymbol{u})/3 + p_\theta(\boldsymbol{u}) \\
&\geq -2p_1(\boldsymbol{u})/3 \text{ (since } p_\theta(\boldsymbol{u}) \geq -p_1(\boldsymbol{u})/3).
\end{aligned}$$

Thus we have shown that (5.6) holds for every sufficiently large k. This completes the proof of Theorem 3.2.

6. Application to 0-1 semi-infinite, nonconvex quadratic optimization problems. We briefly recall two of the Lovász–Schrijver procedures for 0-1 integer programming problems, and relate them to our successive SDP relaxation method. Let F be a subset of $\{0,1\}^n$ whose convex hull is to be approximated. In the Lovász–Schrijver procedures, we assume that a compact convex subset C_0 of R^n satisfying $F = C_0 \cap \{0,1\}^n$ is given in advance. We define

$$\mathcal{K}_0 \equiv \{(\lambda, \lambda \boldsymbol{x}^T) \in R^{1+n} : \lambda \geq 0, \text{ and } \boldsymbol{x} \in C_0\}.$$

Let $\mathcal{K}_I$ denote the convex cone spanned by the 0-1 vectors in $\mathcal{K}_0$:

$$\mathcal{K}_I = \{(\lambda, \lambda \boldsymbol{x}^T) \in R^{1+n} : \lambda \geq 0, \text{ and } \boldsymbol{x} \in \text{c.hull}(F)\}.$$

Here the 0th coordinate is special. It is used in homogenizing the sets of interest in R^n. Clearly

$$C_0 = \{\boldsymbol{x} \in R^n : (1, \boldsymbol{x}^T) \in \mathcal{K}_0\} \quad \text{ and } \quad \text{c.hull}(F) = \{\boldsymbol{x} \in R^n : (1, \boldsymbol{x}^T) \in \mathcal{K}_I\}.$$

The closed convex cone $\mathcal{K}_0$ serves as an initial relaxation of $\mathcal{K}_I$. Given the current relaxation $\mathcal{K}_k$ of $\mathcal{K}_I$, first a convex cone $M_+(\mathcal{K}_k, \mathcal{K}_k)$ in the space of $(1+n) \times (1+n)$

symmetric matrices is defined (the lifting operation). Then a projection of this cone gives the next relaxation $N_+(\mathcal{K}_k)$ of $\mathcal{K}_I$.

Now, we define the lifting operation in general. Let $\mathcal{K}$ and $\mathcal{T}$ be closed convex cones in R^{1+n}. A $(1+n) \times (1+n)$ symmetric matrix, $\boldsymbol{Y}$, with real entries is in $M_+(\mathcal{K}, \mathcal{T})$ if

(i) $\boldsymbol{Y} \in \mathcal{S}_+^{1+n}$,

(ii) $\boldsymbol{Y}\boldsymbol{e}_0 = \mathrm{Diag}(\boldsymbol{Y})$,

(iii) $\boldsymbol{u}^T\boldsymbol{Y}\boldsymbol{v} \geq 0 \ \forall\, u \in \mathcal{K}^*,\ v \in \mathcal{T}^*$. (This condition is equivalent to $\boldsymbol{Y}\mathcal{K}^* \subseteq \mathcal{T}$.)

Here, $\boldsymbol{e}_0$ denotes the unit vector with 0th coordinate 1. Item (ii) above serves an important role in Lovász–Schrijver procedures as well as in some of the SDP relaxations used by Goemans and Williamson [8], Nesterov [15], and Ye [29]. This equation is valid simply because for each j for which $x_j \in \{0, 1\}$, the equation $x_j^2 = x_j$ is valid. Indeed, our general framework applies to any compact set in R^n, and the equation $\boldsymbol{Y}\boldsymbol{e}_0 = \mathrm{Diag}(\boldsymbol{Y})$ was not utilized in earlier sections (as it is not valid). In this section, however, the equation is valid and we utilize it. As will be noted in the proof of Theorem 6.3, the inclusion of this equation will be guaranteed by our choice of the initial formulation.

The third condition of Lovász–Schrijver procedures is very interesting. They present a couple of possibilities for the choice of cone $\mathcal{T}$ in 0-1 integer programming. Among them is the cone spanned by all 0-1 vectors with the first component $x_0 = 1$. This choice, since the cone $\mathcal{T}^*$ has a very simple set of generators, allows for the development of polynomial-time algorithms for approximately solving the successive SDP relaxations as long as the number of iterations of the successive procedure is $O(1)$. Their result only assumes that a polynomial-time *weak separation oracle* is available for $\mathcal{K}$. The key is that since $\mathcal{T}^*$ has only $O(n)$ extreme rays, it becomes trivial to check condition (iii) in polynomial time. On the other hand, Lovász and Schrijver [12] note that the choice $\mathcal{T} \equiv \mathcal{K}$ is also possible and leads to at least as good relaxations as the former choice for $\mathcal{T}$. (In many cases the successive relaxations for $\mathcal{T} \equiv \mathcal{K}$ are significantly tighter than the successive relaxations with the simpler choice of $\mathcal{T}$.) In the case of the latter choice, the possibility of polynomial-time solvability of the first few successive relaxations depends on the availability of polynomial-time algorithms to check $\boldsymbol{Y}\mathcal{K}^* \subseteq \mathcal{K}$. Our procedure uses $\mathcal{T} \equiv \mathcal{K}$.

Now, we describe the projection step.

$$N_+(\mathcal{K}) \equiv \{\boldsymbol{Y}\boldsymbol{e}_0 : \boldsymbol{Y} \in M_+(\mathcal{K}, \mathcal{K})\}.$$

We also define the iterated operators $N_+^k(\mathcal{K})$ as follows: $N_+^0(\mathcal{K}) := \mathcal{K}$ and $N_+^k(\mathcal{K}) := N_+(N_+^{k-1}(\mathcal{K}))$ for all integers $k \geq 1$. (We use the notation $N_+(\mathcal{K})$, whereas $N_+(\mathcal{K}, \mathcal{K})$ is used in [12].)

Another procedure studied in [12] uses a weaker relaxation by removing the condition (i) in the lifting procedure. Let $M(\mathcal{K}, \mathcal{K})$ and $N(\mathcal{K})$ denote the related sets for this procedure. We will refer to the first procedure using the lifting $M_+(\mathcal{K}, \mathcal{K})$ (and the projection N_+) as the N_+ procedure. We will call the other (using $M(\mathcal{K}, \mathcal{K})$, and N) the N procedure. Lovász and Schrijver prove the following.

THEOREM 6.1.

$$\mathcal{K} \supseteq N_+(\mathcal{K}) \supseteq N_+^2(\mathcal{K}) \supseteq \cdots \supseteq N_+^n(\mathcal{K}) = \mathcal{K}_I$$

and

$$\mathcal{K} \supseteq N(\mathcal{K}) \supseteq N^2(\mathcal{K}) \supseteq \cdots \supseteq N^n(\mathcal{K}) = \mathcal{K}_I.$$

Let us see how our successive SDP relaxation method applies to 0-1 nonconvex quadratic optimization problems. Consider a 0-1 nonconvex quadratic program:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in F \equiv \{\boldsymbol{x} \in \{0,1\}^n : p(\boldsymbol{x}) \leq 0 \ \forall p(\cdot) \in \mathcal{P}'\}. \end{array} \tag{6.1}$$

We may assume that the set $\mathcal{P}'$ contains the quadratic functions $x_i(x_i-1),\ i = 1,2,\dots,n$. Then we can replace the 0-1 constraint imposed on the variable x_i by the inequality $-x_i(x_i-1) \leq 0$. Thus by adding the quadratic functions $-x_i(x_i-1),\ i = 1,2,\dots,n$, to $\mathcal{P}'$, we obtain a quadratic inequality representation $\mathcal{P}_F$ of the feasible region F. Let $C_0 \equiv [0,1]^n$. Note that $F \neq C_0 \cap \{0,1\}^n = \{0,1\}^n$ in our general setting here. However, $F = C_0 \cap \{0,1\}^n$ has been assumed for some compact convex subset C_0 of R^n in the Lovász–Schrijver procedures discussed above.

LEMMA 6.2. *Suppose that we take $C_0 = [0,1]^n$ and $\mathcal{P}_0 \equiv \{x_i(x_i-1) :\ i = 1,2,\dots,n\} \subset \mathcal{P}^2(C_0)$. Then $F = C_1 \cap \{0,1\}^n$, where*

$$C_1 = \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \ \textit{such that}\ \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \quad \textit{and} \\ \boldsymbol{P} \bullet \boldsymbol{Y} \leq 0 \ \ \forall \boldsymbol{P} \in \mathcal{P}_F \cup \mathcal{P}_0 \end{array} \right\}.$$

Proof. Let C_1' be the semi-infinite convex QOP relaxation of the set F with the quadratic inequality representation $\mathcal{P}_F \cup \mathcal{P}_0$:

$$C_1' \equiv \{\boldsymbol{x} \in R^n : p(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) \leq 0 \ \forall p(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \text{c.cone}\,(\mathcal{P}_F \cup \mathcal{P}_0) \cap \mathcal{Q}_+\}.$$

In view of Theorem 4.2 and Lemma 5.1, we know that

$$F \subseteq \text{c.hull}(F) \subseteq C_1 = C_1' \subseteq C_0.$$

Hence it suffices to show that

$$\{\boldsymbol{x} \in C_1 : x_i = 0 \ \text{ or } 1,\ i = 1,2,\dots,n\} \subseteq F.$$

If F contains all the 0-1 vectors, the inclusion relation above obviously holds. Now assume that $\boldsymbol{x}' \notin F$ is a 0-1 vector. Then there is a quadratic function $p_1(\cdot,\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}_F$ such that

$$p_1(\boldsymbol{x}',\gamma,\boldsymbol{q},\boldsymbol{Q}) > 0.$$

On the other hand, we know that the quadratic function

$$p_2(\boldsymbol{x}) \equiv \sum_{i=1}^n x_i(x_i-1),$$

with the identity matrix as its Hessian matrix, is a member of c.cone$(\mathcal{P}_0)$, and that $p_2(\boldsymbol{x}') = 0$. Hence if $\epsilon > 0$ is sufficiently small, then

$$\begin{array}{l} \epsilon p_1(\cdot,\gamma,\boldsymbol{q},\boldsymbol{Q}) + p_2(\cdot) \in \text{c.cone}\,(\mathcal{P}_F \cup \mathcal{P}_0) \cap \mathcal{Q}_+, \\ \epsilon p_1(\boldsymbol{x}',\gamma,\boldsymbol{q},\boldsymbol{Q}) + p_2(\boldsymbol{x}') > 0. \end{array}$$

This implies that $\boldsymbol{x}' \notin C_1'$. $\square$

As a consequence of the lemma above, we see that the 0-1 nonconvex quadratic optimization problem (6.1) is equivalent to the 0-1 convex quadratic optimization problem

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in F = C_1 \cap \{0,1\}^n. \end{array} \tag{6.2}$$

Using this observation, we can prove that in the case of 0-1 nonconvex quadratic optimization problem (6.1), our successive SDP relaxation method converges in $(1+n)$ iterations.

THEOREM 6.3. *The successive SDP relaxation method, applied to the* 0-1 *nonconvex quadratic optimization problem* (6.1)*, using* $C_0 = [0,1]^n$ *as the initial approximation of c.hull*(F) *and* $\mathcal{P}_k = \mathcal{P}^2(C_k)$ *in each iteration, terminates in at most* $(1+n)$ *iterations with* $C_{1+n} = $ *c.hull*(F).

Proof. We note that by Lemma 6.2, after one iteration of the successive SDP relaxation method, we obtain the 0-1 convex quadratic optimization problem (6.2) that can be used with the original Lovász–Schrijver procedure. We only have to note that the successive SDP relaxation method becomes the Lovász–Schrijver procedure after the first iteration. For this purpose, we compare conditions (i), (ii), and (iii) of the Lovász–Schrijver procedure for $\mathcal{K} = \mathcal{T} = \mathcal{K}_k$ to the conditions used to construct $C_{k+1} = \hat{F}(\mathcal{P}_F \cup \mathcal{P}_k)$ in the successive SDP relaxation method. Here

$$\mathcal{K}_k \equiv \{(\lambda, \lambda\boldsymbol{x}^T) \in R^{1+n} : \lambda \geq 0,\ \boldsymbol{x} \in C_k\}.$$

First, we observe that $\exists \boldsymbol{X'} \in \mathcal{S}^n$ such that $\boldsymbol{Y'} = \left(\begin{smallmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X'} \end{smallmatrix}\right) \in \mathcal{S}^{1+n}_+$ if and only if $\forall\lambda \geq 0$, $\exists \boldsymbol{X} \in \mathcal{S}^n$ such that $\boldsymbol{Y} = \left(\begin{smallmatrix} \lambda & \lambda\boldsymbol{x}^T \\ \lambda\boldsymbol{x} & \boldsymbol{X} \end{smallmatrix}\right) \in \mathcal{S}^{1+n}_+$. Hence (i) is satisfied. For (ii), note that $x_i(x_i - 1) \in \mathcal{P}_F\ \forall\, i$ implies the constraint $\boldsymbol{Y}\boldsymbol{e}_0 \geq \mathrm{Diag}(\boldsymbol{Y})$ and $-x_i(x_i-1) \in \mathcal{P}_F\ \forall\, i$ implies $\boldsymbol{Y}\boldsymbol{e}_0 \leq \mathrm{Diag}(\boldsymbol{Y})$. Finally, for (iii), note that a linear inequality $\boldsymbol{a}^T\boldsymbol{x} \leq \alpha$ is valid for C_k if and only if $(\alpha, -\boldsymbol{a}^T) \in \mathcal{K}_k^*$ (recall $C_k = \{\boldsymbol{x} \in R^n : (1, \boldsymbol{x}^T) \in \mathcal{K}_k\}$). Therefore, we see that

$$\mathcal{P}^2(C_k) = \text{c.cone}\{-\boldsymbol{u}\boldsymbol{v}^T :\ \boldsymbol{u}, \boldsymbol{v} \in \mathcal{K}_k^*\}.$$

Step (3.1) of the successive SDP relaxation method implies that $\boldsymbol{Y} = \left(\begin{smallmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{smallmatrix}\right) \in -(\mathcal{P}^2(C_k))^*$. Thus, we conclude by noting that

$$\boldsymbol{Y} \in -(\mathcal{P}^2(C_k))^* \quad \text{if and only if } \boldsymbol{Y} \bullet \boldsymbol{u}\boldsymbol{v}^T = \boldsymbol{u}^T\boldsymbol{Y}\boldsymbol{v} \geq 0 \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \mathcal{K}^*.$$

Now, Theorem 6.1 implies that n more steps of the procedure is sufficient. ∎

The above discussion and the results show that our successive SDP relaxation method generalizes the Lovász–Schrijver N_+ procedure by ignoring condition (ii), which is no longer valid. Our results in the previous sections already showed that in this full generality, we still have the asymptotic convergence of the method. It is therefore interesting to investigate the same questions about the weaker procedure N:

- What is the generalization of procedure N?
- Does the generalization of procedure N satisfy the same theoretical properties as the successive SDP relaxation method?

We answer both of these questions in the next section. As is shown in [12], in some cases the procedure N_+ is significantly better than N. Procedure N is weaker, but the relaxations given by it are always polyhedral sets (so LP techniques can be employed) and N_+ requires more general techniques. Hence, sometimes procedure N might be more manageable even if the procedure N_+ is not.

We should expect that the generalization of procedure N should be only using condition (iii), $\boldsymbol{Y}\mathcal{K}^* \subseteq \mathcal{K}$, in the definition of the lifting. We would also expect that the generalization should lead to semi-infinite LP (rather than SDP) relaxations. We show in the next section that the above-mentioned generalization of procedure N leads to successive semi-infinite LP relaxations and all the analogs of the theoretical properties established for our successive SDP relaxations can also be established for the successive semi-infinite LP relaxations.

7. Successive semi-infinite LP relaxation.

SUCCESSIVE SEMI-INFINITE LP RELAXATION METHOD.

Step 0: Let $k = 0$.

Step 1: If $C_k = \emptyset$ or $C_k = \text{c.hull}(F)$, then stop.

Step 2: Choose a quadratic inequality representation $\mathcal{P}_k$ for C_k.

Step 3: Let

$$\begin{aligned} C_{k+1} &= \hat{F}^L(\mathcal{P}_F \cup \mathcal{P}_k) \\ &\equiv \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that} \\ \boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \leq 0 \ \ \forall \boldsymbol{P} \in \mathcal{P}_F \cup \mathcal{P}_k \end{array} \right\} \\ &= \tilde{F}^L(\mathcal{P}_F \cup \mathcal{P}_k) \\ &\equiv \left\{ \boldsymbol{x} \in R^n : \gamma + 2\boldsymbol{q}^T\boldsymbol{x} \leq 0 \ \forall p(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in (\text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k)) \cap \mathcal{L} \right\}. \end{aligned}$$

(The equalities above follow from Corollary 4.3.)

Step 4: Let $k = k + 1$, and go to Step 1.

THEOREM 7.1. *Assume that $\mathcal{P}_F$ is a semi-infinite quadratic inequality representation of a compact subset F of R^n, and that $C_0 \supseteq F$ is a compact convex subset of R^n. If we choose $\mathcal{P}_k = \mathcal{P}^2(C_k)$ at Step 2 of each iteration in the successive semi-infinite LP relaxation method, then the monotonicity property* (a) *and the asymptotic convergence property* (b) *stated in the introduction hold.*

Proof. We can apply the same proof as the one given for Theorem 3.2 in section 5.5 to the theorem. $\square$

Note that we can define another semi-infinite LP relaxation based on the semi-infinite convex QOP relaxation. Clearly, if $\boldsymbol{Q} \in \mathcal{S}^n_+$, then

$$\gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \leq 0 \ \text{ implies } \ \gamma + 2\boldsymbol{q}^T\boldsymbol{x} \leq 0 \quad \forall \boldsymbol{x} \in R^n.$$

So, we can define a semi-infinite LP relaxation based on the above observation:

$$\hat{F}^L_+ \equiv \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n, \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{O} \end{pmatrix} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \leq 0, \\ \forall \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{pmatrix} \in \text{c.cone}(\mathcal{P}) \cap \mathcal{Q}_+ \end{array} \right\}$$

and

$$\tilde{F}^L_+ \equiv \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{O} \end{pmatrix} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0, \\ \forall \begin{pmatrix} \gamma & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{Q} \end{pmatrix} \in \text{c.cone}(\mathcal{P}) \cap \mathcal{Q}_+ \end{array} \right\}.$$

In this case, the equivalence $\hat{F}^L_+ = \tilde{F}^L_+$ is evident. The convergence of the successive semi-infinite LP relaxation method using $\hat{F}^L_+$ can be established by following the proofs of Theorems 3.1 and 3.2. Instead, we note $\tilde{F}^L_+ \subseteq \tilde{F}^L$. Therefore, Theorem 7.1 also implies that this particular semi-infinite LP relaxation method has the properties (a) and (b) mentioned in the theorem.

8. Further discussions on successive convex relaxations.

8.1. Conic quadratic inequality representation. The conic quadratic inequality presented below is a generalization of the linear matrix inequality [3, 28] and the bilinear matrix inequality [14, 20]. It will be shown that any conic quadratic inequality can be reduced to a semi-infinite system of standard quadratic inequalities and vice versa.

Let $\mathcal{K}$ and $\mathcal{K}^* = \{\boldsymbol{v} \in R^m : \boldsymbol{v} \cdot \boldsymbol{u} \geq 0 \ \forall \boldsymbol{u} \in \mathcal{K}\}$ be a closed convex cone in R^m and its dual. Here $\boldsymbol{u} \cdot \boldsymbol{v}$ denotes an inner product of $\boldsymbol{u} \in R^m$ and $\boldsymbol{v} \in R^m$. For all $\boldsymbol{u} \in R^m$, we write $\boldsymbol{u} \preceq_K \boldsymbol{0}$ when $-\boldsymbol{u}$ lies in $\mathcal{K}$. Now we introduce a *conic quadratic inequality*:

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T, \quad \sum_{i=0}^{n} \sum_{j=0}^{n} \boldsymbol{g}_{ij} x_i x_j \preceq_K \boldsymbol{0} \quad \text{and} \quad x_0 = 1. \tag{8.1}$$

Here $\boldsymbol{g}_{ij}$, $i = 0, 1, \ldots, n,\ j = 0, 1, \ldots, n$, are constant vectors in R^m. We may assume without loss of generality that $\boldsymbol{g}_{ij} = \boldsymbol{g}_{ji}$. The inequality (8.1) turns out to be a system of m usual quadratic inequalities on R^n if we take the nonnegative orthant R^m_+ of R^m for the cone $\mathcal{K}$. The inequality (8.1) turns out to be a *quadratic matrix inequality*, which is a generalization of linear and bilinear matrix inequalities [3, 28] if we identify the space of $\ell \times \ell$ symmetric matrices with R^m and we take the positive semidefinite cone $\mathcal{S}^\ell_+$ of matrices for the cone $\mathcal{K}$, where $m = \ell \times (\ell + 1)/2$ for some $\ell \geq 1$.

We can rewrite the conic quadratic inequality (8.1) as a semi-infinite system of standard quadratic inequalities in the homogeneous form.

$$\boldsymbol{P} \bullet \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \leq 0 \quad \forall \boldsymbol{P} \in \mathcal{P} \tag{8.2}$$

for some $\mathcal{P} \subseteq \mathcal{Q} = \mathcal{S}^{1+n}$. This means that we can easily include any conic quadratic inequality in the semi-infinite quadratic inequality representation of the feasible region F of the maximization problem (1.1). To see the equivalence between (8.1) and (8.2) for some $\mathcal{P} \subseteq \mathcal{Q} = \mathcal{S}^{1+n}$, we observe that (8.1) can be rewritten as

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T, \quad \left(\sum_{i=0}^{n} \sum_{j=0}^{n} \boldsymbol{g}_{ij} x_i x_j \right) \cdot \boldsymbol{v} \leq 0 \quad \forall \boldsymbol{v} \in \mathcal{K}^* \quad \text{and} \quad x_0 = 1.$$

Therefore, if we define

$$\begin{aligned} \boldsymbol{P}(\boldsymbol{v}) &\equiv \begin{pmatrix} \boldsymbol{g}_{00} \cdot \boldsymbol{v} & \boldsymbol{g}_{01} \cdot \boldsymbol{v} & \cdots & \boldsymbol{g}_{0n} \cdot \boldsymbol{v} \\ \boldsymbol{g}_{10} \cdot \boldsymbol{v} & \boldsymbol{g}_{11} \cdot \boldsymbol{v} & \cdots & \boldsymbol{g}_{1n} \cdot \boldsymbol{v} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{g}_{n0} \cdot \boldsymbol{v} & \boldsymbol{g}_{n2} \cdot \boldsymbol{v} & \cdots & \boldsymbol{g}_{nn} \cdot \boldsymbol{v} \end{pmatrix} \in \mathcal{Q} = \mathcal{S}^{1+n} \quad \forall \boldsymbol{v} \in \mathcal{K}^*, \\ \mathcal{P} &\equiv \{\boldsymbol{P}(\boldsymbol{v}) : \boldsymbol{v} \in \mathcal{K}^*\}, \end{aligned}$$

we obtain the desired semi-infinite system (8.2) of standard quadratic inequalities, which is equivalent to (8.1).

Let $F(\mathcal{P})$ denote the solution set of (8.2) with its quadratic inequality representation $\mathcal{P} \equiv \{\boldsymbol{P}(\boldsymbol{v}) : \boldsymbol{v} \in \mathcal{K}^*\}$. Applying the SDP relaxation to $F(\mathcal{P})$, we obtain that

$$\begin{aligned}
\hat{F}(\mathcal{P}) &\equiv \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \text{ and} \\ \boldsymbol{P}(\boldsymbol{v}) \bullet \boldsymbol{Y} \leq 0 \ \forall \boldsymbol{v} \in \mathcal{K}^* \end{array} \right\} \\
&= \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \text{ and} \\ \left(\sum_{i=0}^n \sum_{j=0}^n \boldsymbol{g}_{ij} Y_{ij} \right) \cdot \boldsymbol{v} \leq 0 \ \forall \boldsymbol{v} \in \mathcal{K}^* \end{array} \right\} \\
&= \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \text{ and} \\ \sum_{i=0}^n \sum_{j=0}^n \boldsymbol{g}_{ij} Y_{ij} \preceq_K \boldsymbol{0} \end{array} \right\}.
\end{aligned}$$

The set in the last line corresponds to the SDP relaxation to the solution set of (8.1). This implies that we can apply the SDP relaxation directly to the conic quadratic inequality (8.1) without converting it into the semi-infinite system (8.2) of standard quadratic inequalities.

Conversely, we can reduce any semi-infinite system of standard quadratic inequalities to a conic quadratic inequality. To show this, consider a semi-infinite system (8.2) of standard quadratic inequalities in the homogeneous form. We may assume without loss of generality that $\mathcal{P} \subseteq \mathcal{S}^{1+n}$ is a closed convex cone. We can rewrite (8.2) as

$$\left(\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} (1, \boldsymbol{x}^T) \right) \preceq_{\mathcal{P}^*} \boldsymbol{O}, \tag{8.3}$$

which is a conic quadratic inequality.

Let F denote the solution set of the conic quadratic inequality (8.3) that we have derived from (8.2) above. Applying the SDP relaxation to F, we obtain that

$$\begin{aligned}
\hat{F} &\equiv \left\{ \boldsymbol{x} \in R^n : \ \exists \boldsymbol{X} \in \mathcal{S}^n \ \text{ such that } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \ \text{ and } \boldsymbol{Y} \preceq_{\mathcal{P}^*} \boldsymbol{O} \right\} \\
&= \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \ \text{ such that } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \ \text{ and} \\ \boldsymbol{P} \bullet \boldsymbol{Y} \leq 0 \ \forall \boldsymbol{P} \in \mathcal{P} \end{array} \right\}.
\end{aligned}$$

Note that the set in the last line corresponds to the SDP relaxation of the solution set of the semi-infinite system (8.2) of standard quadratic inequalities.

In view of the discussions above, we know that the conic quadratic inequality representation is as general as the semi-infinite quadratic inequality representation and that the SDP relaxations to both representations are equivalent. When we deal with the semi-infinite convex QOP relaxation, however, the semi-infinite quadratic inequality representation seems more convenient than the conic quadratic inequality representation.

8.2. A counterexample to the convergence for the rank-1 quadratic inequality representation case. The example below shows that the rank-1 quadratic inequality representation is not strong enough to ensure the convergence of the successive SDP relaxation method. Let

$$\begin{aligned} F &\equiv \{\boldsymbol{x} = (x_1, x_2)^T : p_0(\boldsymbol{x}) \leq 0,\ \|\boldsymbol{x}\|^2 \leq 1\}, \\ S &\equiv \{\boldsymbol{a} \in R^2 : a_1^2 + a_2^2 = 1\}, \\ B &\equiv \{\boldsymbol{x} = (x_1, x_2)^T \in R^2 : x_1^2 + x_2^2 \leq 1\}, \\ C_0 &\equiv B, \\ p_0(\boldsymbol{x}) &\equiv -(x_1 - 1)^2 - (x_2 - 1)^2 + 1, \\ \mathcal{P}_F &\equiv \{p_0(\boldsymbol{x})\} \cup \mathcal{P}^1(B), \end{aligned}$$

where $\mathcal{P}^1(B)$ denotes the rank-1 quadratic inequality representation of the unit ball, which consists of all quadratic functions such that $(\boldsymbol{a}^T\boldsymbol{x} - 1)(\boldsymbol{a}^T\boldsymbol{x} + 1)$ $(\boldsymbol{a} \in \boldsymbol{S})$. We see that

$$\text{c.hull}(F) = \{\boldsymbol{x} = (x_1, x_2)^T \in B : x_1 + x_2 \leq 1\}.$$

THEOREM 8.1. *Suppose that we take $\mathcal{P}_k = \mathcal{P}^1(C_k)$ (the rank-1 quadratic inequality representation of C_k) in the successive SDP relaxation method applied to the example above. Then $C_k = B$ $(k = 0, 1, 2, \dots)$.*

Proof. By definition, $C_0 = B$. We will prove $C_1 = B$, which suffices to establish the theorem. First observe that $C_1 \subseteq B$. Hence it suffices to show $B \subseteq C_1$ or equivalently for all $p(\cdot) \in \text{c.cone}(\mathcal{P}_F) \cap \mathcal{Q}_+$,

$$p(\bar{\boldsymbol{x}}) \leq 0 \quad \forall \bar{\boldsymbol{x}} \in B.$$

Let $p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+$ and $\bar{\boldsymbol{x}} \in B$ be fixed. Then we can choose $\lambda_i \geq 0$ $(i = 0, 1, \dots, \ell)$ and $\boldsymbol{a}_i \in S$ $(i = 1, 2, \dots, \ell)$ such that

$$p(\boldsymbol{x}) = \lambda_0 p_0(\boldsymbol{x}) + \sum_{i=1}^{\ell} \lambda_i (\boldsymbol{a}_i^T\boldsymbol{x} - 1)(\boldsymbol{a}_i^T\boldsymbol{x} + 1) \quad \forall \boldsymbol{x} \in R^n.$$

If $\lambda_0 = 0$, then $p(\bar{\boldsymbol{x}}) \leq 0$. Now assume that $\lambda_0 > 0$. In this case, we may further assume without loss of generality that $\lambda_0 = 1$; hence, for all $\boldsymbol{x} \in R^n$,

$$\begin{aligned} p(\boldsymbol{x}) &= p_0(\boldsymbol{x}) + \sum_{i=1}^{\ell} \lambda_i (\boldsymbol{a}_i^T\boldsymbol{x} - 1)(\boldsymbol{a}_i^T\boldsymbol{x} + 1) \\ &= \boldsymbol{x}^T \left(\sum_{i=1}^{\ell} \lambda_i \boldsymbol{a}_i \boldsymbol{a}_i^T - \boldsymbol{I} \right) \boldsymbol{x} - \sum_{i=1}^{\ell} \lambda_i + 2\boldsymbol{e}^T\boldsymbol{x} - 1. \end{aligned}$$

It follows from $p(\cdot) \in \mathcal{Q}_+$ that the Hessian matrix $\left(\sum_{i=1}^{\ell} \lambda_i \boldsymbol{a}_i \boldsymbol{a}_i^T - \boldsymbol{I}\right)$ is positive semidefinite. Hence if we denote the largest and the smallest eigenvalues of the matrix $\sum_{i=1}^{\ell} \lambda_i \boldsymbol{a}_i \boldsymbol{a}_i^T$ by $\mu_{\max}$ and $\mu_{\min}$, then $1 \leq \mu_{\min} \leq \mu_{\max}$. We also see that

$$\mu_{\max} + \mu_{\min} = \text{trace}\left(\sum_{i=1}^{\ell} \lambda_i \boldsymbol{a}_i \boldsymbol{a}_i^T \right) = \sum_{i=1}^{\ell} \lambda_i \boldsymbol{a}_i^T \boldsymbol{a}_i = \sum_{i=1}^{\ell} \lambda_i.$$

Hence

$$\begin{aligned} p(\bar{\boldsymbol{x}}) &= \bar{\boldsymbol{x}}^T \left(\sum_{i=1}^{\ell} \lambda_i \boldsymbol{a}_i \boldsymbol{a}_i^T - \boldsymbol{I} \right) \bar{\boldsymbol{x}} - \sum_{i=1}^{\ell} \lambda_i + 2\boldsymbol{e}^T \bar{\boldsymbol{x}} - 1 \\ &\le \mu_{\max} - 1 - \sum_{i=1}^{\ell} \lambda_i + 2\boldsymbol{e}^T \bar{\boldsymbol{x}} - 1 \\ &= 2\boldsymbol{e}^T \bar{\boldsymbol{x}} - \mu_{\min} - 2 \\ &\le 2\sqrt{2} - 3 \\ &< 0. \qquad \square \end{aligned}$$

8.3. A counterexample to the finite termination for the strongest quadratic inequality representation case. The example below shows that in the worst case, even when we take the strongest quadratic inequality representation $\mathcal{P}^\sharp(C_k)$ for C_k at every iteration,

- the successive SDP relaxation method requires infinitely many iterations, and
- the convergence is extremely slow.

For every $\boldsymbol{x} = (x_1, x_2)^T \in R^2$, let

$$\begin{aligned} p_1(\boldsymbol{x}) &\equiv x_1^2 + x_2^2 - 4, \\ p_2(\boldsymbol{x}) &\equiv -(x_1 - 1)^2 - (x_2 - 2)^2 + 5, \\ p_3(\boldsymbol{x}) &\equiv p_2(-x_1, x_2) = -(x_1 + 1)^2 - (x_2 - 2)^2 + 5. \end{aligned}$$

Define

$$\begin{aligned} F &\equiv \{\boldsymbol{x} = (x_1, x_2)^T \in R^2 : p_i(\boldsymbol{x}) \le 0 \ (i = 1, 2, 3)\}, \\ \mathcal{P}_F &\equiv \{p_1(\cdot),\ p_2(\cdot),\ p_3(\cdot)\}, \\ C_0 &= \{\boldsymbol{x} = (x_1, x_2)^T \in R^2 : p_1(\boldsymbol{x}) \le 0\}. \end{aligned}$$

Then

$$\begin{aligned} \text{c.hull}(F) &= \{\boldsymbol{x} = (x_1, x_2)^T \in R^2 : p_1(\boldsymbol{x}) \le 0,\ x_2 \le 0\} \\ &= \{\boldsymbol{x} = (x_1, x_2)^T \in R^2 : x_1^2 + x_2^2 \le 4,\ x_2 \le 0\}. \end{aligned}$$

THEOREM 8.2. *Suppose that we take $\mathcal{P}_k = \mathcal{P}^\sharp(C_k)$ (the strongest quadratic inequality representation of C_k) in the successive SDP relaxation method applied to the example above.*

(i) *C_k is symmetric with respect to the x_2 axis:*

$$(x_1, x_2)^T \in C_k \quad \textit{if and only if } (-x_1, x_2)^T \in C_k.$$

(ii) *Let*

$$\xi_k \equiv \max\{x_2 \ : \ (0, x_2)^T \in C_k\}.$$

Then

$$0 < \xi_k \le 2, \tag{8.4}$$

$$0 < \bar{\xi}_{k+1} \equiv \frac{\xi_k}{1 + \xi_k(1 - \xi_k/4)} \le \xi_{k+1}. \tag{8.5}$$

Proof. We will prove (i) and (ii) by induction.

(i) Obviously the assertion is true for $k = 0$. Assume that C_k is symmetric with respect to the x_2 axis. Then we know that

$$p(x_1, x_2) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+ \text{ if and only if } p(-x_1, x_2) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+.$$

This ensures that C_{k+1} is symmetric with respect to the x_2 axis.

(ii) By definition, we know that $\xi_0 = 2$. Hence (8.4) holds for $k = 0$. Assuming that (8.4) holds, we prove that (8.5) holds. We first observe that

$$(2,0)^T \in \text{c.hull}(F) \subseteq C_k, \quad (0, \xi_k)^T \in C_k \quad \text{and} \quad (0, \bar{\xi}_{k+1})^T \in C_k. \tag{8.6}$$

It suffices to show that $(0, \bar{\xi}_{k+1})^T \in C_{k+1}$ or equivalently

$$p(0, \bar{\xi}_{k+1}) \leq 0 \quad \forall p(x_1, x_2) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+.$$

Assume on the contrary that

$$p(0, \bar{\xi}_{k+1}) > 0 \text{ for } \exists p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+.$$

Since $p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+$, we can choose $\lambda_i \geq 0$ $(i = 2, 3)$ and

$$p'(\boldsymbol{x}) \equiv Q_{11}x_1^2 + 2Q_{12}x_1x_2 + Q_{22}x_2^2 + 2q_1x_1 + 2q_2x_2 + \gamma \in \mathcal{P}_k$$

such that

$$p(\boldsymbol{x}) = \sum_{i=2}^{3} \lambda_i p_i(\boldsymbol{x}) + p'(\boldsymbol{x}) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+.$$

Here we remark that $p_1(\cdot)$ can be incorporated into $p'(\cdot)$ since $p_1(\cdot) \in \mathcal{P}_k$. By the symmetry with respect to the x_2 axis, we see that

$$p(-x_1, x_2) = \sum_{i=2}^{3} \lambda_i p_i(-x_1, x_2) + p'(-x_1, x_2) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+.$$

Thus, defining $\tilde{p}(\boldsymbol{x}) = (p(x_1, x_2) + p(-x_1, x_2))/2$, $\mu = \lambda_2 + \lambda_3$, and $p''(x_1, x_2) = (p'(x_1, x_2) + p'(-x_1, x_2))/2$, we obtain that

$$\begin{aligned} \tilde{p}(0, \bar{\xi}_{k+1}) &= p(0, \bar{\xi}_{k+1}) > 0, \\ p''(x_1, x_2) &= Q_{11}x_1^2 + Q_{22}x_2^2 + q_2x_2 + \gamma \in \mathcal{P}_k, \\ \tilde{p}(x_1, x_2) &= \mu(-x_1^2 - (x_2 - 2)^2 + 4) + p''(x_1, x_2) \in \mathcal{Q}_+. \end{aligned} \tag{8.7}$$

It follows from $p''(x_1, x_2) \in \mathcal{P}_k$ and the third inclusion relation of (8.6) that $p''(0, \bar{\xi}_{k+1}) \leq 0$. Hence $\mu > 0$. We may further assume without loss of generality that $\mu = 1$; redefine $p(\boldsymbol{x}) = p(\boldsymbol{x})/\mu$, $p''(\boldsymbol{x}) = p''(\boldsymbol{x})/\mu, \ldots,$ etc.; then all the relations above remain valid. Since $\tilde{p}(x_1, x_2) \in \mathcal{Q}_+$, we see that $Q_{11} \geq 1$ and $Q_{22} \geq 1$. By (8.6) and $p''(x_1, x_2) \in \mathcal{P}_k$,

$$0 \geq p''(2, 0) = 4Q_{11} + \gamma \geq 4 + \gamma \quad \text{and} \quad 0 \geq p''(0, \xi_k);$$

hence

$$\begin{aligned} \tilde{p}(0,0) &= (-0^2 - 2^2 + 4) + p''(0,0) = \gamma \leq -4 \text{ and} \\ \tilde{p}(0, \xi_k) &= (-0^2 - (\xi_k - 2)^2 + 4) + p''(0, \xi_k) \leq (4 - \xi_k)\xi_k. \end{aligned}$$

Therefore, by the convexity of the quadratic function $\tilde{p}(\boldsymbol{x})$, we obtain that

$$\begin{aligned}
\tilde{p}(0, \bar{\xi}_{k+1}) &= \tilde{p}\left(\frac{\xi_k - \bar{\xi}_{k+1}}{\xi_k}(0,0)^T + \frac{\bar{\xi}_{k+1}}{\xi_k}(0,\xi_k)^T\right) \\
&\leq \frac{\xi_k - \bar{\xi}_{k+1}}{\xi_k}\tilde{p}(0,0) + \frac{\bar{\xi}_{k+1}}{\xi_k}\tilde{p}(0,\xi_k) \\
&\leq \frac{\xi_k - \bar{\xi}_{k+1}}{\xi_k}(-4) + \frac{\bar{\xi}_{k+1}}{\xi_k}(4-\xi_k)\xi_k \\
&= \frac{4\bar{\xi}_{k+1}}{\xi_k}\left(1 + (1-\xi_k/4)\xi_k\right) - 4 \\
&= 0.
\end{aligned}$$

This contradicts (8.7). ▯

The above example is simple, yet it illustrates great difficulties for the successive SDP relaxation method. For example, $\xi_{k+1}/\xi_k \to 1$. Therefore, the convergence is slower than linear.

Note that, in any dimension, if we take a pair of ball constraints, one convex (inclusion), the other nonconvex (exclusion), then both of the successive SDP and semi-infinite LP relaxation methods stop in one iteration, returning the convex hull of the intersection. Also, in the above example, if we knew that $p_2(\cdot)$ affects only the definition of F in the region $x_1 \geq 0$ and that $p_3(\cdot)$ is only effective in the region $x_1 \leq 0$, we could do elementary modifications to the method to speed up convergence tremendously. This is a good elementary example to illustrate the fact that for such methods to become more efficient in practice, hybrid approaches including branch-and-bound and branch-and-cut seem necessary. We make further remarks in the next section.

9. Concluding remarks. We propose extensions of two fundamental lift-and-project procedures N and N_+ of Lovász and Schrijver [12]. The original procedures were proposed for 0-1 integer programming problems to compute the convex hull of feasible (integer) solutions. Our procedure applies to any nonconvex region and as a result we do not use the key equations, $\boldsymbol{Y}\boldsymbol{e}_0 = \mathrm{Diag}(\boldsymbol{Y})$, used in N and N_+ procedures. Therefore, our relaxations are based either on two conditions: $\boldsymbol{Y}$ is positive semidefinite and $\boldsymbol{Y}\mathcal{K}^* \subseteq \mathcal{K}$ (successive SDP relaxation method), or on only one condition: $\boldsymbol{Y}\mathcal{K}^* \subseteq \mathcal{K}$ (successive semi-infinite LP relaxation method). In both cases we established the properties (a) monotonicity and (b) asymptotic convergence. The weakest version of our procedures satisfying the properties (a) and (b) uses only rank-2 quadratic valid inequalities. We showed in section 6 that such inequalities ensure the condition $\boldsymbol{Y}\mathcal{K}^* \subseteq \mathcal{K}$. Finally, in section 8 we showed that even the strongest of such relaxation procedures (using all quadratic valid inequalities) uses infinitely many iterations to converge. In the above sense, the strongest positive result is given in section 7 by the successive semi-infinite LP relaxation method based on rank-2 valid inequalities.

On the one hand, theoretically speaking, the best results are given in section 7: the weakest algorithm achieving the strongest results. Moreover, the successive semi-infinite LP relaxation method is more likely to be practical for a given general problem. On the other hand, the relative value of SDP relaxations has been quite impressive so far on some very special problems (e.g., the stable set problem [12]) and less impressive on others (e.g., the matching problem [25]). Therefore, one interesting research direction is to search for interesting classes of nonconvex sets for which the successive SDP relaxation method is significantly better than the successive semi-infinite LP relaxation method. For the same reason, (partial) characterizations of nonconvex sets on which both methods perform comparably are also important.

Our convergence proofs are by contradiction, but the main argument is about cutting off a point using valid inequalities induced by the underlying construction. The strongest convergence result (for the weakest algorithm) uses separating hyperspheres. In the other proofs, for the *bad* points, the separating hyperspheres may have huge radii and converge to hyperplanes. However, for certain points and shapes, the advantage of using more general convex quadratic inequalities is clear. This discussion motivates us to suggest another avenue for research. It would be interesting to find certain invariants and measures of the input of our procedures that lead to nontrivial, *descriptive* convergence rates for our methods, perhaps only for some interesting subclass of problems.

Recently, Kojima and Takeda [11] discussed the computational complexity of the successive SDP and semi-infinite LP relaxation methods. They gave an upper bound on the number of iterations which the methods require to attain a convex relaxation of a quadratically constrained compact set F with a given accuracy $\epsilon > 0$, in terms of ϵ, the diameter of the initial relaxation C_0, the diameter of F, and some other quantities characterizing the Lipschitz continuity and the nonconvexity and nonlinearity of the quadratic inequality representation $\mathcal{P}_F$ of F.

The major difficulty in implementing the idea of the successive SDP (or semi-infinite LP) relaxation method in practice is the solution of a continuum of semi-infinite SDPs (or semi-infinite LPs) to generate a new approximation C_{k+1} of the convex hull of the feasible region F of a nonconvex quadratic program at each iteration. In their succeeding paper [10], the authors propose implementable variants by introducing two new techniques, a discretization technique for approximating continuum of semi-infinite SDPs (or semi-infinite LPs) by a finite number of standard SDPs (or LPs) with a finite number of linear inequality constraints, and a localization technique for generating a convex relaxation of F that is accurate only in certain directions in a neighborhood of the objective direction $\boldsymbol{c}$. They established that, *Given any positive number ϵ, there is an implementable discretized-localized variant of the successive SDP (or semi-infinite LP) relaxation method which generates an upper bound of the objective values within ϵ of their maximum in a finite number of iterations.* See also [27] for a practical implementation of this variant and some numerical results.

REFERENCES

[1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim., 5 (1995), pp. 13–51.

[2] E. Balas, S. Ceria, and G. Cornuéjols, *A lift-and-project cutting plane algorithm for mixed 0-1 programs*, Math. Programming, 58 (1993), pp. 295–323.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Stud. Appl. Math. 15, SIAM, Philadelphia, 1994.

[4] R. W. Cottle, J. S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, New York, 1992.

[5] G. Ewald, D. G. Larman, and C. A. Rogers, *The directions of the line segments and of the r-dimensional balls on the boundary of a convex body in Euclidean space*, Mathematika, 17 (1970), pp. 1–20.

[6] T. Fujie and M. Kojima, *Semidefinite relaxation for nonconvex programs*, J. Global Optim., 10 (1997), pp. 367–380.

[7] M. X. Goemans, *Semidefinite programming in combinatorial optimization*, Math. Programming, 79 (1997), pp. 143–161.

[8] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comput. Mach., 42 (1995), pp. 1115–1145.

[9] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer, New York, 1988.

[10] M. Kojima and L. Tunçel, *Discretization and Localization in Successive Convex Relaxation Methods for Nonconvex Quadratic Optimization Problems*, Technical Report B-341, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, 1998. Also issued as CORR 98-34, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

[11] M. Kojima and A. Takeda, *Complexity Analysis of Conceptual Successive Convex Relaxation of Nonconvex Sets*, Technical Report B-350, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, 1999.

[12] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and* 0-1 *optimization*, SIAM J. Optim., 1 (1991), pp. 166–190.

[13] R. Horst and H. Tuy, *Global Optimization*, 2nd rev. ed., Springer-Verlag, Berlin, 1992.

[14] M. Mesbahi and G. P. Papavassilopoulos, *A cone programming approach to the bilinear matrix inequality problem and its geometry*, Math. Programming, 77 (1997), pp. 247–272.

[15] Yu. E. Nesterov, *Semidefinite relaxation and nonconvex quadratic optimization*, Optim. Methods Softw., 9 (1998), pp. 141–160.

[16] Yu. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, Stud. Appl. Math. 13, SIAM, Philadelphia, 1994.

[17] G. Pataki and L. Tunçel, *On the Generic Properties of Convex Optimization Problems in Conic Form*, Research Report 97–16, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada, 1997, revised 1999.

[18] S. Poljak, F. Rendl, and H. Wolkowicz, *A recipe for semidefinite relaxation for* (0,1)-*quadratic programming*, J. Global Optim., 7 (1995), pp. 51–73.

[19] M. V. Ramana, *An Algorithmic Analysis of Multiquadratic and Semidefinite Programming Problems*, Ph.D. Thesis, Johns Hopkins University, Baltimore, MD, 1993.

[20] M. G. Safonov, K. G. Goh, and J. H. Ly, *Control system synthesis via bilinear matrix inequalities*, in Proceedings of the 1994 American Control Conference, Baltimore, MD, IEEE, 1994, pp. 45–49.

[21] H. D. Sherali and A. Alameddine, *A new reformulation-linearization technique for bilinear programming problems*, J. Global Optim., 2 (1992), pp. 379–410.

[22] H. D. Sherali and C. H. Tuncbilek, *A reformulation-convexification approach for solving nonconvex quadratic programming problems*, J. Global Optim., 7 (1995), pp. 1–31.

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

[24] N. Z. Shor, *Dual quadratic estimates in polynomial and boolean programming*, Ann. Oper. Res., 25 (1990), pp. 163-168.

[25] T. Stephen and L. Tunçel, *On a representation of the matching polytope via semidefinite liftings*, Math. Oper. Res., 24 (1999), pp. 1–7.

[26] J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions* I, Springer-Verlag, Berlin, 1970.

[27] A. Takeda, Y. Dai, M. Fukuda, and M. Kojima, *Towards the implementation of successive convex relaxation method for nonconvex quadratic optimization problem*, in Approximation and Complexity in Numerical Optimization: Continuous and Discrete Problems, P. M. Pardalos, ed., Kluwer Academic Press, Dordrecht, to appear.

[28] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.

[29] Y. Ye, *Approximating quadratic programming with bound and quadratic constraints*, Math. Programming, 84 (1999), pp. 219–226.