Discretization and Localization in Successive Convex Relaxation Methods for Nonconvex Quadratic Optimization

Masakazu Kojima† and Levent Tunçel‡

July 1998, Revised May 2000

Abstract. Based on the authors' previous work which established theoretical foundations of two, conceptual, successive convex relaxation methods, *i.e.*, the SSDP (Successive Semidefinite Programming) Relaxation Method and the SSILP (Successive Semi-Infinite Linear Programming) Relaxation Method, this paper proposes their implementable variants for general quadratic optimization problems. These problems have a linear objective function $\boldsymbol{c}^T\boldsymbol{x}$ to be maximized over a nonconvex compact feasible region F described by a finite number of quadratic inequalities. We introduce two new techniques, "discretization" and "localization," into the SSDP and SSILP Relaxation Methods. The discretization technique makes it possible to approximate an infinite number of semi-infinite SDPs (or semi-infinite LPs) which appeared at each iteration of the original methods by a finite number of standard SDPs (or standard LPs) with a finite number of linear inequality constraints. We establish:

- *Given any open convex set U containing F, there is an implementable discretization of the SSDP (or SSILP) Relaxation Method which generates a compact convex set C such that $F \subseteq C \subseteq U$ in a finite number of iterations.*

The localization technique is for the cases where we are only interested in upper bounds on the optimal objective value (for a fixed objective function vector $\boldsymbol{c}$) but not in a global approximation of the convex hull of F. This technique allows us to generate a convex relaxation of F that is accurate only in certain directions in a neighborhood of the objective direction $\boldsymbol{c}$. This cuts off redundant work to make the convex relaxation accurate in unnecessary directions. We establish:

- *Given any positive number ϵ, there is an implementable localization-discretization of the SSDP (or SSILP) Relaxation Method which generates an upper bound of the objective value within ϵ of its maximum in a finite number of iterations.*

Key words Nonconvex Quadratic Optimization Problem, Semidefinite Programming, Linear Matrix Inequality, Global Optimization, SDP Relaxation, Semi-Infinite LP Relaxation, Interior-Point Method

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

‡ Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. Part of this work was done while this author was visiting Tokyo Institute of Technology, Department of Mathematical and Computing Sciences, on a sabbatical leave from University of Waterloo. Research of this author was supported in part by Tokyo Institute of Technology and by a research grant from NSERC of Canada.
e-mail:*ltuncel@math.uwaterloo.ca*

1 Introduction.

This paper deals with a general quadratic optimization problem (abbreviated by QOP) of the form

$$\text{maximize } \boldsymbol{c}^T\boldsymbol{x} \text{ subject to } \boldsymbol{x} \in F \equiv \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \mathcal{P}_F)\}. \tag{1}$$

Here $\boldsymbol{c}$ denotes a constant vector in the n-dimensional Euclidean space R^n and $\mathcal{P}_F$ a set of (possibly infinitely many) quadratic functions on R^n. We note that QOP (1) covers various important nonconvex optimization problems such as general quadratic programs [12, 16, etc.], mixed 0-1 linear and quadratic integer programs [15, etc.], linear complementarity problems [5, etc.], and bilinear matrix inequalities [13, 17, etc.]. Our major concern is a convex relaxation of QOP (1):

$$\text{maximize } \boldsymbol{c}^T\boldsymbol{x} \text{ subject to } \boldsymbol{x} \in C, \tag{2}$$

where C denotes a convex subset of R^n including F. Throughout the paper we assume that F is compact, and that a compact convex subset C_0 of R^n, which includes F, is known in advance. The set C_0 will serve as an initial relaxation of the convex hull, c.hull(F), of F.

In their previous paper [10], the authors presented two types of successive convex relaxation methods for approximating c.hull(F), which are extensions of the Lovász-Schrijver lift-and-project procedures [11] with the use of SDP (semidefinite programming) and semi-infinite LP (linear programming) relaxations originally developed for combinatorial optimization problems to general nonconvex quadratic optimization problems. We will call the first method the SSDP (Successive Semidefinite Programming) Relaxation Method, and the second the SSILP (Successive Semi-Infinite Linear Programming) Relaxation Method. See also [1, 2, 6, 7, 8, 14, 15, 16, 21, 22, 25, etc.] for SDP and semi-infinite LP relaxations. The semi-infinite LP relaxation is called "a reformulation-convexification" in [21]. In [10], the authors established the following theoretical foundation of their methods: *Both of the methods generate a sequence of compact convex sets C_k of R^n $(k = 0, 1, \dots)$ 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 *if $F = \emptyset$, then $C_{k^*} = \emptyset$ for some finite number k^* (detecting infeasibility).*

The current paper is a continuation of the successive convex relaxation methods proposed in the previous paper [10] toward an implementation of the SSDP and SSILP Relaxation Methods. A major difficulty in implementing these methods is that we are required to solve an infinite number of semi-infinite SDPs (or semi-infinite LPs) to generate a new approximation C_{k+1} of c.hull(F). Here we mean by a semi-infinite SDP (or a semi-infinite LP) a semidefinite program (or a linear program, respectively) with infinitely many inequality constraints. Another difficulty lies in the slow convergence of C_k to c.hull(F). See the numerical example in Section 8.3 of [10]. Among these two major difficulties, the first one is more crucial than the second one; we first need to resolve the first difficulty to implement the methods, and then we will be able to start discussing various practical issues including the global and local convergence of the methods and applications of the methods to special cases.

The purpose of this paper is to propose implementable variants of the SSDP and SSILP Relaxation Methods for QOP (1). We will not address the slow convergence and the possible computational inefficiency of the proposed methods. For our theoretical analyses and results, we allow cases where the representation $\mathcal{P}_F$ of the feasible region F consists of infinitely many quadratic inequalities. To develop implementable variants of the SSDP and SSILP Relaxation Methods for QOP (1), however, we need to assume that $\mathcal{P}_F$ is finite.

We will bring two new techniques, "discretization" and "localization," into the SSDP and SSILP Relaxation Methods. "Discretization" makes it possible to approximate an infinite number of semi-infinite SDPs (or semi-infinite LPs) which appear at each iteration of the method by a finite number of standard SDPs (or standard LPs, respectively) with a finite number of inequalities. We establish the following result in Theorem 3.3 : *For any given open convex subset U of R^n including F, there is a discretization of the SSDP (or SSILP) Relaxation Method which generates a sequence of compact convex sets C_k of R^n $(k = 0, 1, \dots)$ satisfying (a) above and*

(b)' c.hull(F) $\subseteq C_k \subseteq U$ for some finite number k.

The properties (a), (b) and (b)' mentioned above are independent of the objective function $\boldsymbol{c}^T\boldsymbol{x}$ of QOP (1). In other words, the original successive convex relaxation methods and their discretized variant mentioned above do not utilize any information on the objective function $\boldsymbol{c}^T\boldsymbol{x}$ of QOP (1). In many cases, however, we are only interested in an approximate optimal solution of QOP (1) and a good upper bound for the objective values of QOP (1) but not an approximation of the entire c.hull(F). For such cases, it is reasonable to consider a convex relaxation of F that is accurate only in certain directions that lie in a neighborhood of the objective direction $\boldsymbol{c}$ but not necessarily accurate in other directions. The localization technique is developed for generating such a convex relaxation of F. We establish the following result in Theorem 3.5: *For any given $\epsilon > 0$, there is a localization-discretization of the SSDP (or SSILP) Relaxation Method which generates a sequence of compact convex sets C_k of R^n $(k = 0, 1, \dots)$ satisfying (a) and*

(c)' $\zeta_k \equiv \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_k\} \leq \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in F\} + \epsilon$ for some finite number k.

Theoretically we could impose property (c)' on a discretized variant of the SSDP (or SSILP) Relaxation Method without using the localization technique, but we study the localization technique for practical purposes because it eliminates some redundant work to make the convex relaxation accurate in unnecessary directions.

Section 2 is devoted to notation, symbols and conditions used throughout the paper. In Section 3, we present our main results including Theorems 3.3 and 3.5 referred to above. We sketch an outline of the proofs of Theorems 3.3 and 3.5 in Section 4.1, and give a series of fundamental lemmas in Sections 4.2 and 4.3, which are used to prove Theorem 3.3 in Section 5 and Theorem 3.5 in Section 6. Section 7 discusses relations of the SSDP and SSILP Relaxation Methods applied to 0-1 quadratic optimization problems and the Lovász-Schrijver lift-and-project procedures [11].

2 Notation, Symbols and Conditions.

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. We often use the notation $qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q})$ to designate the constant term γ, the linear term $2\boldsymbol{q}^T\boldsymbol{x}$ and the quadratic term $\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$ involved in a quadratic function $p(\cdot)$ defined on R^n;

$$p(\boldsymbol{x}) = qf(\boldsymbol{x}; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \equiv \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x},$$

where $\gamma \in R$, $\boldsymbol{q} \in R^n$ and $\boldsymbol{Q} \in \mathcal{S}^n$. With this convention, we can write the set $\mathcal{Q}$ of quadratic functions on R^n, the set $\mathcal{Q}_+$ of convex quadratic functions on R^n and the set $\mathcal{L}$ of linear functions on R^n as

$$
\begin{aligned}
\mathcal{Q} &\equiv \{qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) : \gamma \in R,\ \boldsymbol{q} \in R^n,\ \boldsymbol{Q} \in \mathcal{S}^n\}, \\
\mathcal{Q}_+ &\equiv \{qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) : \gamma \in R,\ \boldsymbol{q} \in R^n,\ \boldsymbol{Q} \in \mathcal{S}^n_+\}, \\
\mathcal{L} &\equiv \{qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) : \gamma \in R,\ \boldsymbol{q} \in R^n,\ \boldsymbol{Q} = \boldsymbol{O}\},
\end{aligned}
$$

respectively. We say that $\mathcal{P} \subseteq \mathcal{Q}$ ($\mathcal{P} \subseteq \mathcal{Q}_+$ or $\mathcal{P} \subseteq \mathcal{L}$) is *a quadratic (convex quadratic or linear) inequality representation of a closed subset C of R^n* if $C = \{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0\ (\forall p(\cdot) \in \mathcal{P})\}$holds.

We need to impose the following condition on QOP (1) to implement Algorithms 3.1 and 3.2, which are variants of the SSDP and SSILP Relaxation Methods [10], respectively, on a computer.

Condition 2.1

(i) The quadratic inequality representation $\mathcal{P}_F$ of the feasible region F is finite.

(ii) A compact convex set $C_0 \supseteq F$ is known in advance and it has a finite convex quadratic (or linear) inequality representation $\widetilde{\mathcal{P}}_0$.

We remark that all the theoretical analyses and results in Sections 3, 4, 5 and 6 remain valid even when the representations $\mathcal{P}_F$ and $\widetilde{\mathcal{P}}_0$ are infinite.

Let

$$
\begin{aligned}
\text{c.cone}(\mathcal{P}) &\equiv \text{the closure of the convex cone generated by } \mathcal{P} \subseteq \mathcal{Q}, \\
\boldsymbol{Q} \bullet \boldsymbol{X} &\equiv \sum_i \sum_j Q_{ij} X_{ij} \\
&\quad\ \text{(the (trace) inner product of two symmetric matrices } \boldsymbol{Q} \text{ and } \boldsymbol{X}), \\
B(\boldsymbol{x}_c, \rho) &\equiv \{\boldsymbol{x} \in R^n : \|\boldsymbol{x} - \boldsymbol{x}_c\| \leq \rho\}\ (\forall \boldsymbol{x}_c \in R^n \text{ and } \forall \rho > 0) \\
&\quad\ \text{(the closed ball with center } \boldsymbol{x}_c \in R^n \text{ and radius } \rho > 0) \\
B_{\text{int}}(\boldsymbol{x}_c, \rho) &\equiv \{\boldsymbol{x} \in R^n : \|\boldsymbol{x} - \boldsymbol{x}_c\| < \rho\}\ (\forall \boldsymbol{x}_c \in R^n \text{ and } \forall \rho > 0) \\
&\quad\ \text{(the open ball with center } \boldsymbol{x}_c \in R^n \text{ and radius } \rho > 0) \\
\overline{D} &\equiv \{\boldsymbol{d} \in R^n : \|\boldsymbol{d}\| = 1\} \\
&\quad\ \text{(the set of vectors with unit length)}, \\
D(\boldsymbol{w}, \delta) &\equiv \overline{D} \cap B(\boldsymbol{w}, \delta) = \{\boldsymbol{u} \in \overline{D} \ :\ \|\boldsymbol{u} - \boldsymbol{w}\| \leq \delta\}\ (\forall \delta > 0 \text{ and } \forall \boldsymbol{w} \in \overline{D}), \\
D_I &\equiv \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n, -\boldsymbol{e}_1, -\boldsymbol{e}_2, \ldots, -\boldsymbol{e}_n\}.
\end{aligned}
$$

Here $\boldsymbol{e}_i$ denotes the ith unit coordinate vector in R^n $(i = 1, 2, \ldots, n)$.

Let $D \subseteq \overline{D}$. For every $\delta \geq 0$, a subset D' of $\overline{D}$ is *a δ-net of D* if for every $\boldsymbol{d} \in D$ there exists a $\boldsymbol{d}' \in D'$ such that $\|\boldsymbol{d}' - \boldsymbol{d}\| \leq \delta$. By definition, if $\delta > \delta' \geq 0$ and D' is a δ'-net of D then D' is a δ-net of D. In particular, D itself is a δ-net of D for any $\delta \geq 0$. It should be noted that we can take a finite δ-net D' of D whenever $\delta > 0$. Thus, a δ-net D' of D with $\delta > 0$ leads us to a finite discretization of D.

Let C be a compact subset of C_0. For every $\boldsymbol{d}$, $\boldsymbol{d}' \in \overline{D}$ and $\boldsymbol{x} \in R^n$, define

$$
\begin{aligned}
\alpha(C, \boldsymbol{d}) &\equiv \max\{\boldsymbol{d}^T \boldsymbol{x} \ :\ \boldsymbol{x} \in C\}, \\
\ell sf(\boldsymbol{x}; C, \boldsymbol{d}) &\equiv \boldsymbol{d}^T \boldsymbol{x} - \alpha(C, \boldsymbol{d}), \\
r2sf(\boldsymbol{x}; C, \boldsymbol{d}, \boldsymbol{d}') &\equiv -\ell sf(\boldsymbol{x}; C, \boldsymbol{d}) \ell sf(\boldsymbol{x}; C, \boldsymbol{d}'), \\
r1sf(\boldsymbol{x}; C, \boldsymbol{d}) &\equiv -\ell sf(\boldsymbol{x}; C, \boldsymbol{d}) \ell sf(\boldsymbol{x}; C, -\boldsymbol{d}).
\end{aligned}
$$

We call $\ell sf(\cdot; C, \boldsymbol{d})$ *a linear supporting function for* C *with a normal* $\boldsymbol{d}$ *in* $\overline{D}$, $r2sf(\cdot; C, \boldsymbol{d}, \boldsymbol{d}')$ *a rank-2 (quadratic) supporting function for* C *with a pair* $(\boldsymbol{d}, \boldsymbol{d}')$ *of normals in* $D \times \overline{D}$, and $r1sf(\cdot; C, \boldsymbol{d})$ *a rank-1 (quadratic) supporting function for* C *with a normal* $\boldsymbol{d}$ *in* $\overline{D}$. Let

$$\begin{aligned}
\mathcal{P}^L(C, D) &\equiv \{\ell sf(\cdot; C, \boldsymbol{d}) : \boldsymbol{d} \in D\} \ (\forall D \subseteq \overline{D}), \\
\mathcal{P}^2(C, D, D') &\equiv \{r2sf(\cdot; C, \boldsymbol{d}, \boldsymbol{d}') : \boldsymbol{d} \in D,\ \boldsymbol{d}' \in D'\} \ (\forall D,\ D' \subseteq \overline{D}), \\
\mathcal{P}^1(C, D) &\equiv \{r1sf(\cdot; C, \boldsymbol{d}) : \boldsymbol{d} \in D\} \ (\forall D \subseteq \overline{D}).
\end{aligned}$$

For every $\mathcal{P} \subseteq \mathcal{Q}$, let

$$\begin{aligned}
\widehat{G}(C_0, \mathcal{P}) &\equiv \left\{ \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_{+} : \begin{array}{l} \boldsymbol{x} \in C_0 \ \text{ and} \\ \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}) \end{array} \right\}, \\
\widehat{F}(C_0, \mathcal{P}) &\equiv \left\{ \boldsymbol{x} \in R^n : \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \widehat{G}(C_0, \mathcal{P}) \ \text{ for } \exists \boldsymbol{X} \in \mathcal{S}^n \right\} \\
&\equiv \text{an SDP relaxation of} \\
&\quad\ F(C_0, \mathcal{P}) \equiv \{\boldsymbol{x} \in C_0 : qf(\boldsymbol{x}; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P})\}, \\
\widehat{G}^L(C_0, \mathcal{P}) &\equiv \left\{ \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n} : \begin{array}{l} \boldsymbol{x} \in C_0 \ \text{ and} \\ \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}) \end{array} \right\}, \\
\widehat{F}^L(C_0, \mathcal{P}) &\equiv \left\{ \boldsymbol{x} \in R^n : \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \widehat{G}^L(C_0, \mathcal{P}) \ \text{ for } \exists \boldsymbol{X} \in \mathcal{S}^n \right\} \\
&\equiv \text{a semi-infinite LP relaxation of } F(C_0, \mathcal{P}).
\end{aligned}$$

We may regard the SDP relaxation $\widehat{F}(C_0, \mathcal{P})$ (or the semi-infinite LP relaxation $\widehat{F}^L(C_0, \mathcal{P})$) as the projection of $\widehat{G}(C_0, \mathcal{P}) \subset \mathcal{S}^{1+n}$ onto R^n (or the projection of $\widehat{G}^L(C_0, \mathcal{P}) \subset \mathcal{S}^{1+n}$ onto R^n).

3 Main Results.

Algorithms 3.1 and 3.2 below are variants of the SSDP (Successive SDP) Relaxation Method given in Section 3 of [10] and the SSILP (Successive Semi-Infinite LP) Relaxation Method given in Section 7 of [10], respectively.

Algorithm 3.1 (SSDP Relaxation Method)

Step 0: Choose a pair of direction sets $D_0,\ D_1 \subseteq \overline{D}$ such that $D_0 \supseteq D_I$. Let $k = 0$.

Step 1: If $C_k = \emptyset$ then let $\zeta_k = -\infty$ and stop; otherwise, compute an upper bound ζ_k of the maximum objective value of QOP (1) by $\zeta_k = \max\{\boldsymbol{c}^T\boldsymbol{x} \ : \ \boldsymbol{x} \in C_k\}$.

Step 2: Compute $\alpha(C_k, \boldsymbol{d}) = \max\{\boldsymbol{d}^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}$ $(\forall \boldsymbol{d} \in D_0 \cup D_1)$.

Step 3-L: Let $\mathcal{P}_k = \mathcal{P}^L(C_k, D_0) \cup \mathcal{P}^2(C_k, D_0, D_1)$.

Step 4: Let $C_{k+1} = \widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$.

Step 5: Let $k = k + 1$, and go to Step 1.

Algorithm 3.2 (SSILP Relaxation Method)

Step 0, 1, 2 and 3-L: The same as Steps 0, 1, 2 and 3-L of Algorithm 3.1, respectively.

Step 4: Let C_{k+1} be the semi-infinite LP relaxation $\widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$ of F with the quadratic inequality representation $\mathcal{P}_F \cup \mathcal{P}_k$.

Step 5: The same as Step 5 of Algorithm 3.1.

At Step 3-L of both algorithms, we generate a set $\mathcal{P}_k$ of linear and quadratic valid inequalities for F, which consists of the following two subsets, the set $\mathcal{P}^L(C_k, D_0)$ of all linear supporting functions for C_k with normals in D_0, and the set $\mathcal{P}^2(C_k, D_0, D_1)$ of all rank-2 supporting functions for C_k with pairs of normals in $D_0 \times D_1$. Since $F \subseteq C_k$, $\mathcal{P}_k$ together with the original quadratic inequality representation $\mathcal{P}_F$ of F then serves as a new quadratic inequality representation for F. The difference between the two algorithms above lies in their Step 4; Algorithm 3.1 employs the SDP relaxation $\widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$ of F with the quadratic inequality representation $\mathcal{P}_F \cup \mathcal{P}_k$ generated in Step 3-L, while Algorithm 3.2 employs the semi-infinite LP relaxation $\widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$ with the same representation. Evidently, the SDP relaxation is at least as good as the semi-infinite LP relaxation, *i.e.*, $\widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_k) \subseteq \widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$.

Each algorithm above generates a sequence $\{C_k \ (k = 0, 1, 2, \ldots, \ )\}$ of convex subsets of C_0 and a sequence $\{\zeta_k \ (k = 0, 1, 2, \ldots, \ )\}$ of real numbers. If we replaced Step 3-L by

Step 3-R: Choose a quadratic inequality representation $\mathcal{P}_k$ for C_k,

then Algorithms 3.1 and 3.2 would be identical to the original SSDP Relaxation Method in Section 3 of [10] and the original SSILP Relaxation Method given in Section 7 of [10], respectively. In this case, the choice of direction sets $D_0, \ D_1 \subseteq \overline{D}$ in Step 0 and the computation of $\alpha(C_k, \boldsymbol{d})$ $(\forall \boldsymbol{d} \in D_0 \cup D_1)$ in Step 2 would be unnecessary.

In both Algorithms 3.1 and 3.2, $\mathcal{P}_k = \mathcal{P}^L(C_k, D_0) \cup \mathcal{P}^2(C_k, D_0, D_1)$ generated at Step 3-L does not form a quadratic inequality representation of C_k in general, *i.e.*, C_k can be a proper subset of $\{\boldsymbol{x} \in R^n : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \mathcal{P}_k)\}$. But, if we take $D_0 = D_1 = \overline{D}$, then $\mathcal{P}_k = \mathcal{P}^L(C_k, D_0) \cup \mathcal{P}^2(C_k, D_0, D_1)$ forms a quadratic inequality representation of C_k. Thus, we can regard this case as a common special case of Algorithm 3.1 and the original SSDP Relaxation Method given in Section 3 of [10] or a common special case of Algorithm 3.2 and the original SSILP Relaxation Method given in Section 7 of [10]. In this case, we can derive the convergence of the sequence $\{C_k \ (k = 0, 1, 2, \ldots, \ )\}$ to c.hull(F) in the sense of the properties (a) and (b) stated in Introduction. See Theorems 3.2 and 7.1 of [10].

Another common special case is obtained by taking $D_0 = D_I$ and $D_1 = \overline{D}$. In this case, D_0 is finite and $\mathcal{P}_k = \mathcal{P}^L(C_k, D_0) \cup \mathcal{P}^2(C_k, D_0, D_1)$ also forms a quadratic inequality representation of C_k (see (i) of Lemma 4.2 in Section 4). We can derive the properties (a) and (b) (see Corollary 3.4). We should mention, however, that Algorithms 3.1 and 3.2 with this choice of direction sets $D_0, \ D_1 \subseteq \overline{D}$ are still impractical because $D_1 = \overline{D}$ consists of a continuum of direction vectors;

- Step 2 requires to compute a continuum of scalars

$$\alpha(C_k, \boldsymbol{d}) = \max\{\boldsymbol{d}^T \boldsymbol{x} : \boldsymbol{x} \in C_k\} \ (\forall \boldsymbol{d} \in D_I \cup \overline{D}),$$

- Step 4 requires a continuum of linear inequalities

$$\gamma + 2\boldsymbol{q}^T \boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}_F \cup \mathcal{P}_k)$$

  to represent C_{k+1}.

To resolve these difficulties in implementing Algorithms 3.1 and 3.2, we need to choose a finite set for D_1 too in Step 0. In Theorem 3.3 below, we will employ a δ-net of $\overline{D}$, a finite discretization of $\overline{D}$ which uniformly approximates $\overline{D}$, for the direction set D_1. Even with this compromise, we can maintain the monotonicity property (a). But we can not expect the asymptotic convergence property (b) any more, so that we need to introduce an approximation of c.hull(F) in a finite number of iterations: Given a small convex neighborhood U of F (a small open convex set U containing F), we will regard a convex set C_k as a good approximation of c.hull(F), relative to U, if c.hull(F) $\subseteq C_k \subseteq U$ holds.

Now, we are ready to state our main results.

Theorem 3.3 *Let $U \subseteq R^n$ be an arbitrary open convex set containing F. There exists a $\delta > 0$ such that if we take a δ-net D_1 of $\overline{D}$ in Algorithm 3.1 (or Algorithm 3.2), then $\{C_k\ (k = 0, 1, 2, \dots)\}$ generated by Algorithm 3.1 (or Algorithm 3.2, respectively) satisfies the properties (a) and (b)' stated in Introduction.*

We will give a proof of the theorem in Section 5 after preparing several fundamental lemmas in Section 4. As a corollary, we obtain:

Corollary 3.4 *Take $D_1 = \overline{D}$ in Algorithm 3.1 (or Algorithm 3.2). Then $\{C_k\ (k = 0, 1, 2, \dots)\}$ generated by Algorithm 3.1 (or Algorithm 3.2, respectively) satisfies the properties (a) and (b).*

Proof: Since $\overline{D}$ itself is a δ-net of $\overline{D}$ for any $\delta > 0$, the desired result follows from Theorem 3.3. ∎

We will see in Section 7.1 that if we apply Algorithms 3.1 and 3.2 to a 0-1 convex quadratic optimization problem using $D_0 = D_I$ and $D_1 = \overline{D}$ as in the corollary above then they work in a similar way as the Lovász-Schrijver lift-and-project procedures [11].

Since c.hull(F) is compact, given any $\epsilon > 0$, we can choose a convex neighborhood U of F such that if property (b)' holds then so does property (c)' stated in Introduction. Usually property (b)' is unnecessary in practice but property (c)' is enough because we are only interested in an approximate optimal solution of QOP (1) and a good bound for the optimal value but not in an approximation of the entire convex hull, c.hull(F), of the feasible region F of QOP (1). For such cases, we will present *a localization technique* for generating a sequence $\{C_k\ (k = 0, 1, 2, \dots)\}$ of convex relaxations of F that becomes accurate only in certain directions in a neighborhood of the objective direction $\mathbf{c}$ (in the sense of (c)' with a sufficiently small positive ϵ) and cuts off redundant work to make the convex relaxations accurate in unnecessary directions.

Theorem 3.5 *Suppose that $\mathbf{c} \neq \mathbf{0}$ in QOP (1). Let $\kappa > 0$ and $\epsilon > 0$. There exists a $\delta > 0$ such that if we take a δ-net D_1 of $D(\mathbf{c}, \kappa)$ in Algorithm 3.1 (or Algorithm 3.2), then $\{C_k\ (k = 0, 1, 2, \dots)\}$ and $\{\zeta_k\ (k = 0, 1, 2, \dots)\}$ generated by Algorithm 3.1 (or Algorithm 3.2, respectively) satisfy the properties (a) and (c)'.*

We give a proof of Theorem 3.5 in Section 6. As a corollary, we obtain:

Corollary 3.6 *Suppose that $\mathbf{c} \neq \mathbf{0}$ in QOP (1). Let $\kappa > 0$. Take $D_1 = D(\mathbf{c}, \kappa)$ in Algorithm 3.1 (or Algorithm 3.2). Then $\{C_k\ (k = 0, 1, 2, \dots)\}$ and $\{\zeta_k\ (k = 0, 1, 2, \dots)\}$ generated by Algorithm 3.1 (or Algorithm 3.2, respectively) satisfy the properties (a) and*

(c) ζ_k converges to ζ^* as $k \to \infty$.

Proof: Since $D(\boldsymbol{c},\kappa)$ itself is a δ-net of $D(\boldsymbol{c},\kappa)$ for any $\delta > 0$, the desired result follows from Theorem 3.5. ∎

In the remainder of this section, we assume that Condition 2.1 holds, and that we take finite direction sets both for D_0 and D_1 in Algorithms 3.1 and 3.2. Let $\widetilde{\mathcal{P}}_0$ denote the finite convex quadratic inequality representation of the compact convex set C_0 containing F assumed in (ii) of Condition 2.1:

$$C_0 = \{\boldsymbol{x} \in R^n : \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \leq 0 \ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \widetilde{\mathcal{P}}_0)\}.$$

In this case, the maximization of a linear function $\boldsymbol{d}^T\boldsymbol{x}$ over C_0 at Steps 1 and 2 in Algorithms 3.1 turns out to be a convex quadratic optimization problem

$$\begin{array}{ll} \text{maximize} & \boldsymbol{d}^T\boldsymbol{x} \\ \text{subject to} & \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \leq 0 \ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \widetilde{\mathcal{P}}_0), \end{array}$$

which can be further transformed into an SDP

$$\begin{array}{ll} \text{maximize} & \boldsymbol{d}^T\boldsymbol{x} \\ \text{subject to} & \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \widetilde{\mathcal{P}}_0), \\ & \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \end{array}$$

(see Theorem 4.2 of [10]). Let $k \geq 1$. Recall that $C_k = \widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1})$ is the projection of $\widehat{G}(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1}) \subset \mathcal{S}^{1+n}$ onto R^n in Algorithm 3.1. Hence the maximization of a linear function $\boldsymbol{d}^T\boldsymbol{x}$ over C_k at Steps 1 and 2 is carried out by solving an SDP with finitely many inequality constraints:

$$\begin{array}{ll} \text{maximize} & \boldsymbol{d}^T\boldsymbol{x} \\ \text{subject to} & \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \widehat{G}(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1}). \end{array}$$

Here

$$\widehat{G}(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1}) = \left\{ \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ : \begin{array}{l} \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \\ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \widetilde{\mathcal{P}}_0 \cup \mathcal{P}_F \cup \mathcal{P}_{k-1}) \end{array} \right\}.$$

Furthermore, the number of SDPs to be solved in each iteration is finite. Therefore, Algorithm 3.1 is implementable on a computer. In executing Algorithm 3.1, it is not necessary to construct C_k explicitly but to update $\mathcal{P}_k$ in each iteration. It should be also noted that even if $\widehat{G}(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1})$ is represented in terms of a finite number of linear matrix inequalities, its projection $C_k = \widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1})$ onto R^n does not have a finite quadratic inequality representation in general.

Now, we suppose that the compact set C_0 containing F has a finite linear inequality representation $\widetilde{\mathcal{P}}_0$:

$$C_0 = \{\boldsymbol{x} \in R^n : \gamma + 2\boldsymbol{q}^T\boldsymbol{x} \leq 0 \ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{O}) \in \widetilde{\mathcal{P}}_0)\}.$$

Then the maximization of a linear function $\boldsymbol{d}^T\boldsymbol{x}$ over C_0 at Steps 1 and 2 in Algorithm 3.2 forms an LP

$$\begin{array}{ll} \text{maximize} & \boldsymbol{d}^T\boldsymbol{x} \\ \text{subject to} & \gamma + 2\boldsymbol{q}^T\boldsymbol{x} \leq 0 \ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{O}) \in \widetilde{\mathcal{P}}_0). \end{array}$$

Let $k \geq 1$. Since $C_k = \widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1})$ is the projection of $\widehat{G}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_{k-1}) \subset \mathcal{S}^{1+n}$ onto R^n in Algorithm 3.2, the maximization of a linear function $\boldsymbol{d}^T\boldsymbol{x}$ over C_k at Steps 1 and 2 is carried out by solving an LP with finitely many inequality constraints:

$$\begin{array}{ll} \text{maximize} & \boldsymbol{d}^T\boldsymbol{x} \\ \text{subject to} & \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \widetilde{\mathcal{P}}_0 \cup \mathcal{P}_F \cup \mathcal{P}_{k-1}).\mathcal{S}^{1+n} \end{array}$$

As in the case of Algorithm 3.1, it is not necessary to construct C_k explicitly but to update $\mathcal{P}_k$ in each iteration, and the number of LPs to be solved in each iteration is finite; hence Algorithm 3.2 is implementable on a computer.

Remark 3.7 We can view Algorithm 3.2 in the context of the reformulation- linearization technique [19, 20, 21, 18]. Suppose that $F \subseteq C_k$. (This will be shown in Lemma 4.2). Then we can add the redundant inequalities $\boldsymbol{d}^T\boldsymbol{x} - \alpha(\boldsymbol{d}, C_k) \leq 0$ ($\forall \boldsymbol{d} \in D_0 \cup D_1$) to describe F as the intersection of C_0 and the solution set of linear and quadratic inequalities

$$\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C_k) \leq 0 \ (\forall \boldsymbol{d}_0 \in D_0), \tag{3}$$

$$\boldsymbol{d}_1^T\boldsymbol{x} - \alpha(\boldsymbol{d}_1, C_k) \leq 0 \ (\forall \boldsymbol{d}_1 \in D_1), \tag{4}$$

$$qf(\boldsymbol{x}; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}_F). \tag{5}$$

Specifically, the inequalities

$$\boldsymbol{e}_i^T\boldsymbol{x} - \alpha(\boldsymbol{e}_i, C_k) \leq 0 \ (\forall i = 1, 2, \ldots, n)$$

involved in (3) serves as the bound-factor constraints in the reformulation-linearization technique. Then we replace (4) by pairwise products of inequalities in (3) and inequalities in (4):

$$-\left(\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C_k)\right)\left(\boldsymbol{d}_1^T\boldsymbol{x} - \alpha(\boldsymbol{d}_1, C_k)\right) \leq 0 \ (\forall \boldsymbol{d}_0 \in D_0, \ \forall \boldsymbol{d}_1 \in D_1). \tag{6}$$

Thus, the intersection of C_0 and the solution set of (3), (5) and (6) gives us a description of F. This process is a special case of the reformulation technique. Now, we apply the linearization technique to quadratic inequalities in (5) and (6), which replaces every quadratic term x_ix_j by a single variable X_{ij} ($1 \leq i \leq j \leq n$). Note that inequalities in (3) and (6) correspond to $qf(\boldsymbol{x}; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \leq 0$ ($\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}_k$) in our construction. Therefore, the linearization process described above can be written as $F^L(C_0; \mathcal{P}_F \cup \mathcal{P}_k)$ in Step 4.

Remark 3.8 We can replace the set $\mathcal{P}^L(C_k, D_0)$ of all linear supporting functions for C_k with normals in D_0 by the set $\mathcal{P}^1(C_k, D_0)$ of all rank-1 supporting functions for C_k with normals in D_0 in Step 3-L of Algorithms 3.1 and 3.2;

Step 3-1: Let $\mathcal{P}_k = \mathcal{P}^1(C_k, D_0) \cup \mathcal{P}^2(C_k, D_0, D_1)$.

Using Lemma 4.1, we can prove that Step 3-1 works at least as effectively as Step 3-L. Now, suppose that $\{-\boldsymbol{d} \in \overline{D} : \boldsymbol{d} \in D_0\} \subseteq D_1$. (Specifically, this case applies to Corollary 3.4 where we take $D_1 = \overline{D}$). Then we can eliminate $\mathcal{P}^L(C_k, D_0)$ in constructing $\mathcal{P}_k$ in Step 3-L. The resultant step with only $\mathcal{P}^2(C_k, D_0, D_1)$ in the construction of $\mathcal{P}_k$ works equally effectively as Step 3-L.

Remark 3.9 The invariance of the reformulation-linearization technique under any one-to-one affine transformation was shown in the paper [21]. We will see in Lemma 4.3 that Algorithms 3.1 and 3.2 are invariant under any one-to-one affine transformation. This means that the requirement $D_0 \supseteq D_I$ at Step 0 of Algorithms 3.1 and 3.2 can be generalized. Namely we can replace D_I by any $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n, -\boldsymbol{v}_1, -\boldsymbol{v}_2, \ldots, -\boldsymbol{v}_n\}$ whenever the set of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ forms a basis of R^n. All the theorems and corollaries above remain valid with this generalization.

Remark 3.10 Let $\boldsymbol{c} \in R^n$ and $F \subset R^n$. Consider a general nonconvex maximization problem

$$\text{maximize} \;\; \boldsymbol{c}^T\boldsymbol{x} \;\; \text{subject to } \boldsymbol{x} \in F$$

having an optimal value $\boldsymbol{v}^* < \infty$. Given $\epsilon > 0$, we can think of various ϵ-(global-)optimal solutions. The branch-and-bound method has been widely used to compute such ϵ-optimal solutions. Among others, the branch-and-bound method combined with the reformulation linearization technique proposed by Sherali and Tuncbilek [21] attains

- an "ϵ-feasible" solution $\hat{\boldsymbol{x}}$ satisfying $\boldsymbol{c}^T\hat{\boldsymbol{x}} < v^* + \epsilon$

under the assumption that F is a bounded subset of R^n represented by polynomial inequalities. An important feature of our methods is that we do not rely on the branch-and-bound technique. Theorem 3.5 states that with an appropriate choice of δ-net D_1 of $D(\boldsymbol{c}, \kappa)$, both Algorithms 3.1 and 3.2 generate

- a convex relaxation of F such that $\max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C\} < \boldsymbol{v}^* + \epsilon$,

while Theorem 3.3 states that with an appropriate choice of δ-net D_1 of $\overline{D}$, both Algorithms 3.1 and 3.2 generate

- an "ϵ-approximation" of the convex hull of F.

4 Lemmas.

4.1 Outline of the Proofs of Theorems 3.3 and 3.5.

The proofs of our main theorems (Theorems 3.3 and 3.5) are rather lengthy and complicated requiring several lemmas presented in Sections 4.2, 4.3, 5 and 6. Before going into technical details, we briefly state an outline of the proofs and roles of those lemmas.

Among the lemmas, Lemma 4.1 plays an important role not only in the proofs of our main theorems but also in Lemmas 4.2 and 4.3. It gives us an alternative representation of the SDP relaxation $\widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$ of F used in Step 3 of Algorithm 3.1 in terms of convex quadratic inequalities and an alternative representation of the semi-infinite LP relaxation $\widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$ of F used in Step 3 of Algorithm 3.2 in terms of linear inequalities. These representations are originally due to [6] and [3], respectively. By Lemmas 4.1 and 4.2, we know that C_k $(k = 0, 1, 2, \dots,)$ are compact convex subsets of C_0, and that property (a) holds. These facts imply that the intersection $C_\infty \equiv \cap_{k=0}^{\infty} C_k$ is a compact convex set containing F and that $C_\infty = \emptyset$ if and only if $C_k = \emptyset$ for some finite number k. If $C_\infty = \emptyset$ then $F = \emptyset$ and $C_k = \emptyset$ for some finite number k. Therefore, we may restrict ourselves to the case $C_\infty \neq \emptyset$ from then on to prove the properties (b)' and (c)'.

To prove property (b)' in Theorem 3.3, we prepare a (ψ, ρ)-approximation of c.hull(F), denoted by c.hull(F, ψ, ρ), at the beginning of Section 5, where $\psi \geq 0$ and $\rho > 0$ are parameters which determine the accuracy of the approximation. Lemma 5.1 presents the existence of $\psi > 0$ and $\rho > 0$ such that $F \subseteq$ c.hull$(F, \psi, \rho) \subseteq U$ and that c.hull$(F, \psi, \rho) = \emptyset$ if $F = \emptyset$.Thus, property (b)' is reduced to the relation $C_\infty \subseteq$ c.hull(F, ψ, ρ). We prove this relation by contradiction. In other words, we assume that there is a $\hat{\boldsymbol{x}} \in C_\infty \backslash$c.hull$(F, \psi, \rho)$, and show that for every sufficiently large

k, the kth iteration of Algorithms 3.1 and 3.2 "cuts off" $\hat{\boldsymbol{x}}$; $\hat{\boldsymbol{x}} \notin C_{k+1}$. We construct such a linear cut based on Lemmas 4.1 and 4.6.

To prove property (c)' in Theorem 3.5, we will introduce the set

$$C_0' \equiv \left\{ \boldsymbol{x}' \in C_0 : \boldsymbol{c}^T\boldsymbol{x}' \geq \sup\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in F\} + \epsilon \right\}$$

at the beginning of Section 6. Then property (c)' is reduced to the relation $C_\infty \cap C_0' = \emptyset$. We prove this relation by contradiction using Lemmas 6.1, 6.2, 4.1 and 4.6 in a similar way as in the proof of Theorem 3.3.

Lemmas 4.4 and 4.5 are used to derive Lemma 4.6, and Lemma 4.3 is used to simplify the discussions in the proofs of Theorems 3.3 and 3.5.

4.2 Basic Properties of Algorithm 3.1.

Lemma 4.1

(i) $\widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_k) = \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+)\}$.

(ii) $\widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k) = \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{L})\}$.

Proof: See Theorem 4.2 and Corollary 4.3 of [10]. ∎

The next lemma provides us a proof of assertion (a) in Theorems 3.3 and 3.5.

Lemma 4.2 *Let* $\{C_k \ (k = 0, 1, \dots)\}$ *denote the sequence generated by Algorithm 3.1 or 3.2.*

(i) $C_{k+1} \subseteq \{\boldsymbol{x} \in C_0 : \boldsymbol{d}^T\boldsymbol{x} - \alpha(C_k, \boldsymbol{d}) \leq 0 \ (\forall \boldsymbol{d} \in D_0 \cup D_1)\}$.

(ii) $C_{k+1} \subseteq C_k \ (\forall k = 0, 1, 2, \dots)$.

(iii) $F \subseteq C_k \ (\forall k = 0, 1, 2, \dots)$.

Proof: We will only prove (ii) because we can prove (i) in a quite similar way as in the proof of Proposition 1 of [21], and (iii) follows directly from the construction of C_{k+1}. We shall only deal with the case where the sequence $\{C_k \ (k = 0, 1, \dots)\}$ is generated by Algorithm 3.1. We can similarly deal with the other case where the sequence is generated by Algorithm 3.2 if we replace $\mathcal{Q}_+$ by $\mathcal{L}$ below. By definition, we know that $C_1 \subseteq C_0$. Let $k \geq 1$. Assuming that $C_k \subseteq C_{k-1}$, we will show that $C_{k+1} \subseteq C_k$. Let $\boldsymbol{x}' \in C_{k+1}$. It follows from $C_k \subseteq C_{k-1}$ that

$$\alpha(C_k, \boldsymbol{d}) \leq \alpha(C_{k-1}, \boldsymbol{d}) \ (\forall \boldsymbol{d} \in D_0 \cup D_1).$$

In view of (i), we know that

$$\left.\begin{array}{rcl} 0 & \geq & \boldsymbol{d}^T\boldsymbol{x}' - \alpha(C_k, \boldsymbol{d}) \geq \boldsymbol{d}^T\boldsymbol{x}' - \alpha(C_{k-1}, \boldsymbol{d}) \ (\forall \boldsymbol{d} \in D_0 \cup D_1), \\ 0 & \geq & -(\boldsymbol{d}_0^T\boldsymbol{x}' - \alpha(C_k, \boldsymbol{d}_0))(\boldsymbol{d}_1^T\boldsymbol{x}' - \alpha(C_k, \boldsymbol{d}_1)) \\ & \geq & -(\boldsymbol{d}_0^T\boldsymbol{x}' - \alpha(C_{k-1}, \boldsymbol{d}_0))(\boldsymbol{d}_1^T\boldsymbol{x}' - \alpha(C_{k-1}, \boldsymbol{d}_1)) \ (\forall \boldsymbol{d}_0 \in D_0 \ \text{ and } \forall \boldsymbol{d}_1 \in D_1). \end{array}\right\} \quad (7)$$

Now, we will show that if $p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_{k-1}) \cap \mathcal{Q}_+$ then $0 \geq p(\boldsymbol{x}')$; then $\boldsymbol{x}' \in C_k$ follows from Lemma 4.1. We can represent $p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_{k-1}) \cap \mathcal{Q}_+$ as

$$\begin{array}{rcl} p(\boldsymbol{x}) & = & \displaystyle\sum_{i=1}^m \lambda_i p_i(\boldsymbol{x}) + \sum_{\boldsymbol{d}_0 \in D_0} \mu(\boldsymbol{d}_0)(\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(C_{k-1}, \boldsymbol{d}_0)) \\ & & \displaystyle - \sum_{\boldsymbol{d}_0 \in D_0, \boldsymbol{d}_1 \in D_1} \nu(\boldsymbol{d}_0, \boldsymbol{d}_1)(\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(C_{k-1}, \boldsymbol{d}_0))(\boldsymbol{d}_1^T\boldsymbol{x} - \alpha(C_{k-1}, \boldsymbol{d}_1)) \end{array}$$

for some $p_i(\cdot) \in \mathcal{P}_F$ $(i = 1, 2, \cdots, m)$, $\lambda_i \geq 0$ $(i = 1, 2, \cdots, m)$, $\mu(\boldsymbol{d}_0) \geq 0$ $(\boldsymbol{d}_0 \in D_0)$ and $\nu(\boldsymbol{d}_0, \boldsymbol{d}_1) \geq 0$ $(\boldsymbol{d}_0 \in D_0$ and $\boldsymbol{d}_1 \in D_1)$. By (7), we then obtain that

$$\begin{aligned} p(\boldsymbol{x}') &\leq \sum_{i=1}^{m} \lambda_i p_i(\boldsymbol{x}') + \sum_{\boldsymbol{d}_0 \in D_0} \mu(\boldsymbol{d}_0)(\boldsymbol{d}_0^T \boldsymbol{x}' - \alpha(C_k, \boldsymbol{d}_0)) \\ &\quad - \sum_{\boldsymbol{d}_0 \in D_0, \boldsymbol{d}_1 \in D_1} \nu(\boldsymbol{d}_0, \boldsymbol{d}_1)(\boldsymbol{d}_0^T \boldsymbol{x}' - \alpha(C_k, \boldsymbol{d}_1))(\boldsymbol{d}_1^T \boldsymbol{x}' - \alpha(C_k, \boldsymbol{d}_1)) \\ &\leq 0. \end{aligned}$$

Here the last inequality holds because $\boldsymbol{x}' \in C_{k+1}$ and the quadratic function which appeared on the right hand side, belongs to c.cone$(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+$. Thus, we have shown that $\boldsymbol{x}' \in C_k$. ∎

We will show below the fact that Algorithms 3.1 and 3.2 are invariant under any one-to-one affine transformation. Let $\boldsymbol{h} : R^n \to R^n$ be a one-to-one affine transformation of the form $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}_0)$ for $\boldsymbol{x} \in R^n$, where $\boldsymbol{x}_0 \in R^n$ denote a constant vector and $\boldsymbol{A}$ an $n \times n$ constant nonsingular matrix. Let $\{C_k \ (k = 0, 1, \dots)\}$ denote the sequence generated by Algorithm 3.1 or 3.2 with $D_0, \ D_1 \subseteq \overline{D}$ at Step 0. Let

$$\begin{aligned} F' &\equiv \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in F\}, \ \mathcal{P}'_F \equiv \{p(\boldsymbol{h}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}_F\}, \\ C'_0 &\equiv \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_0\}, \\ D'_0 &\equiv \{\boldsymbol{A}^{-T}\boldsymbol{d}_0/\|\boldsymbol{A}^{-T}\boldsymbol{d}_0\| : \boldsymbol{d}_0 \in D_0\}, \ D'_1 \equiv \{\boldsymbol{A}^{-T}\boldsymbol{d}_1/\|\boldsymbol{A}^{-T}\boldsymbol{d}_1\| : \boldsymbol{d}_1 \in D_1\}. \end{aligned}$$

Then $\mathcal{P}'_F$ forms a quadratic inequality representation of F'. Now, applying Algorithm 3.1 or 3.2 to the compact set F' (with its quadratic inequality representation $\mathcal{P}'_F$ of F'), the initial compact convex set $C'_0 \supseteq F'$ and the pair D'_0 and D'_1 of direction sets, we obtain a sequence of compact sets C'_k $(k = 0, 1, \dots)$ such that c.hull$(F') \subseteq C'_{k+1} \subseteq C'_k$ $(k = 0, 1, \dots)$.

Lemma 4.3 $C'_k = \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_k\}$ $(k = 0, 1, \dots)$.

Proof: We shall only deal with the case where the sequence $\{C_k \ (k = 0, 1, \dots)\}$ is generated by Algorithm 3.1. We can similarly deal with the other case where the sequence is generated by Algorithm 3.2 if we replace $\mathcal{Q}_+$ by $\mathcal{L}$ in the last part of the proof. We will prove $C'_k = \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_k\}$ $(k = 0, 1, \dots)$ by induction. By definition, we know that $C'_0 = \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_0\}$. Assuming that $C'_k = \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_k\}$, we will show that $C'_{k+1} = \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_{k+1}\}$. Since $C'_k = \{\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_k\}$ by the induction hypothesis, we see that if $\boldsymbol{d}' = \boldsymbol{A}^{-T}\boldsymbol{d}/\|\boldsymbol{A}^{-T}\boldsymbol{d}\| \in D'_0 \cup D'_1$ for some $\boldsymbol{d} \in D_0 \cup D_1$ then

$$\alpha(C'_k, \boldsymbol{d}') = \max\{(\boldsymbol{A}^{-T}\boldsymbol{d})^T \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}_0)/\|\boldsymbol{A}^{-T}\boldsymbol{d}\| : \boldsymbol{x} \in C_k\} = \frac{\alpha(C_k, \boldsymbol{d}) - \boldsymbol{d}^T \boldsymbol{x}_0}{\|\boldsymbol{A}^{-T}\boldsymbol{d}\|}.$$

Hence we can rewrite each linear supporting function $\ell sf(\cdot; C'_k, \boldsymbol{d}')$ for C'_k with a normal $\boldsymbol{d}'$ in $D'_0 \cup D'_1$ as

$$\ell sf(\boldsymbol{y}; C'_k, \boldsymbol{d}') = \frac{(\boldsymbol{A}^{-T}\boldsymbol{d})^T \boldsymbol{y} - \alpha(C_k, \boldsymbol{d}) + \boldsymbol{d}^T \boldsymbol{x}_0}{\|\boldsymbol{A}^{-T}\boldsymbol{d}\|} = \frac{\ell sf(\boldsymbol{h}^{-1}(\boldsymbol{y}); C_k, \boldsymbol{d})}{\|\boldsymbol{A}^{-T}\boldsymbol{d}\|},$$

where $\boldsymbol{d}' = \boldsymbol{A}^{-T}\boldsymbol{d}/\|\boldsymbol{A}^{-T}\boldsymbol{d}\| \in D'_0 \cup D'_1$ for some $\boldsymbol{d} \in D_0 \cup D_1$, and each rank-2 supporting function $\boldsymbol{r2sf}(\cdot; C'_k, \boldsymbol{d}'_0, \boldsymbol{d}'_1)$ for C'_k in a pair $(\boldsymbol{d}'_0, \boldsymbol{d}'_1)$ of normals in $D'_0 \times D'_1$ as

$$\boldsymbol{r2sf}(\boldsymbol{y}; C'_k, \boldsymbol{d}'_0, \boldsymbol{d}'_1) \quad = \quad -\ell sf(\boldsymbol{y}; C'_k, \boldsymbol{d}'_0)\ell sf(\boldsymbol{y}; C'_k, \boldsymbol{d}'_1) = \frac{\boldsymbol{r2sf}(\boldsymbol{h}^{-1}(\boldsymbol{y}); C_k, \boldsymbol{d}_0, \boldsymbol{d}_1)}{\|\boldsymbol{A}^{-T}\boldsymbol{d}_0\|\|\boldsymbol{A}^{-T}\boldsymbol{d}_1\|},$$

where $\boldsymbol{d}_0' = \boldsymbol{A}^{-T}\boldsymbol{d}_0/\|\boldsymbol{A}^{-T}\boldsymbol{d}_0\| \in D_0'$ for some $\boldsymbol{d}_0 \in D_0$ and $\boldsymbol{d}_1' = \boldsymbol{A}^{-T}\boldsymbol{d}_1/\|\boldsymbol{A}^{-T}\boldsymbol{d}_1\| \in D_1'$ for some $\boldsymbol{d}_1 \in D_1$. Thus, we obtain that

$$\begin{aligned}
\text{c.cone}(\mathcal{P}^L(C_k', D_0')) &= \text{c.cone}\left(\{\ell sf(\cdot; C_k', \boldsymbol{d}_0') : \boldsymbol{d}_0' \in D_0'\}\right) \\
&= \text{c.cone}\left(\left\{\frac{\ell sf(\boldsymbol{h}^{-1}(\cdot); C_k, \boldsymbol{d}_0)}{\|\boldsymbol{A}^{-T}\boldsymbol{d}_0\|} : \boldsymbol{d}_0 \in D_0\right\}\right) \\
&= \text{c.cone}\left(\left\{\ell sf(\boldsymbol{h}^{-1}(\cdot); C_k, \boldsymbol{d}_0) : \boldsymbol{d}_0 \in D_0\right\}\right), \\
\text{c.cone}(\mathcal{P}^2(C_k', D_0', D_1')) &= \text{c.cone}\left(\left\{\boldsymbol{r}2sf(\boldsymbol{h}^{-1}(\cdot); C_k, \boldsymbol{d}_0, \boldsymbol{d}_1) : \boldsymbol{d}_0 \in D_0, \boldsymbol{d}_1 \in D_1\right\}\right), \\
\text{c.cone}(\mathcal{P}_k') &= \text{c.cone}\left(\mathcal{P}^L(C_k'; D_0') \cup \mathcal{P}^2(C_k', D_0', D_1')\right) \\
&= \text{c.cone}\left(\left\{p(\boldsymbol{h}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}_k\right\}\right).
\end{aligned}$$

Also it is easily verified that $p(\cdot) \in \mathcal{Q}_+$ if and only if $p(\boldsymbol{h}^{-1}(\cdot)) \in \mathcal{Q}_+$. Therefore,

$$\text{c.cone}(\mathcal{P}_F' \cup \mathcal{P}_k') = \text{c.cone}\left(\left\{p(\boldsymbol{h}^{-1}(\cdot)) : p(\cdot) \in \mathcal{P}_F \cup \mathcal{P}_k\right\}\right).$$

Applying Lemma 4.1, we consequently obtain that

$$\begin{aligned}
C_{k+1}' &= \{\boldsymbol{y} \in C_0' : p'(\boldsymbol{y}) \leq 0 \ (\forall p'(\cdot) \in \text{c.cone}(\mathcal{P}_F' \cup \mathcal{P}_k') \cap \mathcal{Q}_+)\} \\
&= \{\boldsymbol{y} \in C_0' : p(\boldsymbol{h}^{-1}(\boldsymbol{y})) \leq 0 \ (\forall p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+)\} \\
&= \{\boldsymbol{h}(\boldsymbol{x}) : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{Q}_+)\} \\
&= \{\boldsymbol{h}(\boldsymbol{x}) : \boldsymbol{x} \in C_{k+1}\}.
\end{aligned}$$

∎

4.3 Convex Cones of Rank-2 Supporting Functions for Compact Convex Sets.

Let $\widetilde{\boldsymbol{M}} \in \mathcal{S}^n$. Then we can write $\widetilde{\boldsymbol{M}}$ as

$$\begin{aligned}
\widetilde{\boldsymbol{M}} &= [\tilde{\mu}_{ij}],\ \tilde{\mu}_{ij} = \tilde{\mu}_{ij}^+ - \tilde{\mu}_{ij}^-,\ \tilde{\mu}_{ij}^+ \geq 1,\ \tilde{\mu}_{ij}^- \geq 1, \\
\widetilde{\boldsymbol{M}} &= \widetilde{\boldsymbol{M}}^+ - \widetilde{\boldsymbol{M}}^-,\ \widetilde{\boldsymbol{M}}^+ = [\tilde{\mu}_{ij}^+] \in \mathcal{S}^n,\ \widetilde{\boldsymbol{M}}^- = [\tilde{\mu}_{ij}^-] \in \mathcal{S}^n.
\end{aligned}$$

We will construct a family of quadratic functions $f^+(\cdot; \rho, \theta, \boldsymbol{w})$ and $f^-(\cdot; \rho, \theta, \boldsymbol{w})$ with parameters $\theta \in (0, \pi/8]$ and $\boldsymbol{w} \in \overline{D}$ which satisfies the properties (i), (ii), (iii) and (iv) listed in Lemma 4.4. For every $\theta \in (0, \pi/8]$, $\boldsymbol{w} \in \overline{D}$ and $i = 1, 2, \ldots, n$, define

$$\left.\begin{aligned}
&\nu_i(\theta, \boldsymbol{w}) = \|\boldsymbol{w}\cos\theta + \boldsymbol{e}_i \sin\theta\|, && \bar{\nu}_i(\theta, \boldsymbol{w}) = \|\boldsymbol{w}\cos\theta - \boldsymbol{e}_i \sin\theta\|, \\
&\boldsymbol{b}_i(\theta, \boldsymbol{w}) = \frac{\boldsymbol{w}\cos\theta + \boldsymbol{e}_i \sin\theta}{\nu_i(\theta, \boldsymbol{w})} \in \overline{D}, && \bar{\boldsymbol{b}}_i(\theta, \boldsymbol{w}) = \frac{\boldsymbol{w}\cos\theta - \boldsymbol{e}_i \sin\theta}{\bar{\nu}_i(\theta, \boldsymbol{w})} \in \overline{D}, \\
&\lambda_i(\theta, \boldsymbol{w}) = \frac{\nu_i(\theta, \boldsymbol{w})}{2\sin\theta}, && \bar{\lambda}_i(\theta, \boldsymbol{w}) = \frac{\bar{\nu}_i(\theta, \boldsymbol{w})}{2\sin\theta}.
\end{aligned}\right\} \tag{8}$$

For every $\rho > 0$, $\theta \in (0, \pi/8]$, $\boldsymbol{w} \in \overline{D}$ and $i = 1, 2, \ldots, n$,

$$\boldsymbol{b}_i(\theta, \boldsymbol{w})^T\boldsymbol{x} - \rho \leq 0,\ \bar{\boldsymbol{b}}_i(\theta, \boldsymbol{w})^T\boldsymbol{x} - \rho \leq 0,\ -\boldsymbol{e}_i^T\boldsymbol{x} - \rho \leq 0 \ \text{ and } \boldsymbol{e}_i^T\boldsymbol{x} - \rho \leq 0$$

are linear valid inequalities for the n-dimensional ball $B(\mathbf{0},\rho)$. For every $\boldsymbol{x} \in R^n$, $\rho > 0$, $\theta \in (0,\pi/8]$ and $\boldsymbol{w} \in \overline{D}$, define

$$
\begin{aligned}
f^{+}(\boldsymbol{x};\rho,\theta,\boldsymbol{w}) &\equiv -\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\mu}_{ij}^{+}\Big(\lambda_i(\theta,\boldsymbol{w})(\boldsymbol{b}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho)(-\boldsymbol{e}_j^T\boldsymbol{x}-\rho) \\
&\quad +\bar{\lambda}_i(\theta,\boldsymbol{w})(\bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho)(\boldsymbol{e}_j^T\boldsymbol{x}-\rho)\Big), \qquad (9)\\
f^{-}(\boldsymbol{x};\rho,\theta,\boldsymbol{w}) &\equiv -\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\mu}_{ij}^{-}\Big(\lambda_i(\theta,\boldsymbol{w})(\boldsymbol{b}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho)(\boldsymbol{e}_j^T\boldsymbol{x}-\rho) \\
&\quad +\bar{\lambda}_i(\theta,\boldsymbol{w})(\bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho)(-\boldsymbol{e}_j^T\boldsymbol{x}-\rho)\Big). \qquad (10)
\end{aligned}
$$

The lemma below is a variant of Lemma 5.3 of [10].

Lemma 4.4

(i) Let $\theta \in (0,\pi/8]$, $\boldsymbol{w} \in \overline{D}$ *and* $D' = \{\boldsymbol{b}_j(\theta,\boldsymbol{w}),\ \bar{\boldsymbol{b}}_j(\theta,\boldsymbol{w})\ (\forall j = 1,2,\ldots,n)\}$. *Then*

$$f^{+}(\cdot;\rho,\theta,\boldsymbol{w}),\ f^{-}(\cdot;\rho,\theta,\boldsymbol{w}) \in \text{c.cone}(\mathcal{P}^2(B(\mathbf{0},\rho),D_I,D'))$$

for every $\rho > 0$.

(ii) For every $\boldsymbol{x} \in R^n$, $\rho > 0$, $\theta \in (0,\pi/8]$ *and* $\boldsymbol{w} \in \overline{D}$,

$$f^{+}(\rho\boldsymbol{x};\rho,\theta,\boldsymbol{w}) = \rho^2 f^{+}(\boldsymbol{x};1,\theta,\boldsymbol{w}) \quad \textit{and } f^{-}(\rho\boldsymbol{x};\rho,\theta,\boldsymbol{w}) = \rho^2 f^{-}(\boldsymbol{x};1,\theta,\boldsymbol{w}).$$

(iii) For every $\boldsymbol{w} \in \overline{D}$, $f^{+}(\boldsymbol{w};1,\theta,\boldsymbol{w}) \to 0$ *and* $f^{-}(\boldsymbol{w};1,\theta,\boldsymbol{w}) \to 0$ *as* $\theta \in (0,\pi/8] \to 0$.

(iv) For every $\rho > 0$, $\theta \in (0,\pi/8]$ *and* $\boldsymbol{w} \in \overline{D}$,

$$
\begin{aligned}
&\textit{the Hessian matrix of } f^{+}(\cdot;\rho,\theta,\boldsymbol{w}) = \widetilde{\boldsymbol{M}}^{+},\\
&\textit{the Hessian matrix of } f^{-}(\cdot;\rho,\theta,\boldsymbol{w}) = -\widetilde{\boldsymbol{M}}^{-}.
\end{aligned}
$$

Proof: We will only prove assertion (iii) since the other assertions (i), (ii) and (iv) follow directly from the construction of the quadratic functions $f^{+}(\cdot;\rho,\theta,\boldsymbol{w})$ and $f^{-}(\cdot;\rho,\theta,\boldsymbol{w})$. To prove (iii), let $\boldsymbol{w} \in \overline{D}$ and i be fixed. It suffices to show that

$$\epsilon_i(\theta) \equiv \lambda_i(\theta,\boldsymbol{w})(\boldsymbol{b}_i(\theta,\boldsymbol{w})^T\boldsymbol{w}-1) \ \text{ and } \bar{\epsilon}_i(\theta) \equiv \bar{\lambda}_i(\theta,\boldsymbol{w})(\bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w})^T\boldsymbol{w}-1)$$

converge to zero as $\theta \in (0,\pi/8]$ tends to 0 $(i = 1,2,\ldots,n)$. We will only show the convergence $\epsilon_i(\theta) \to 0$ as $\theta \in (0,\pi/8] \to 0$ below. The convergence $\bar{\epsilon}_i(\theta) \to 0$ as $\theta \in (0,\pi/8] \to 0$ can be derived similarly. We see from (8) that

$$\epsilon_i(\theta) = \frac{\left(\cos\theta + \boldsymbol{w}^T\boldsymbol{e}_i\sin\theta - (2\boldsymbol{w}^T\boldsymbol{e}_i\sin\theta\cos\theta+1)^{1/2}\right)}{2\sin\theta} \qquad (11)$$

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

$$\left(-\sin\theta + \boldsymbol{w}^T\boldsymbol{e}_i\cos\theta + \frac{\boldsymbol{w}^T\boldsymbol{e}_i(\sin^2\theta-\cos^2\theta)}{(2\boldsymbol{w}^T\boldsymbol{e}_i\sin\theta\cos\theta+1)^{1/2}}\right),$$

which vanishes at $\theta = 0$. On the other hand, the derivative "$2\cos\theta$" of the denominator "$2\sin\theta$" in (11) does not vanish at $\theta = 0$. Thus, we see that $\epsilon_i(\theta)$ converges to 0 as $\theta \in (0,\pi/8]$ tends to 0. ∎

Lemma 4.5 *Suppose that*

$$0 < \psi, \ 0 < \kappa \quad \textit{and}\ 0 < \tilde{\rho}. \tag{12}$$

Then there exists a $\delta > 0$ such that every δ-net D_1 of $D(\boldsymbol{c}, \kappa)$ satisfies the following property.

P1*: Assume that*

$$\rho \in (0, \tilde{\rho}] \quad \textit{and}\ \boldsymbol{w} \in D(\boldsymbol{c}, \kappa/2). \tag{13}$$

Then there are quadratic functions $g^+(\cdot)$, $g^-(\cdot) \in$ c.cone$(\mathcal{P}^2(B(\boldsymbol{0}, \rho), D_I, D_1))$ such that

$$\left.\begin{array}{ll} -\psi/4 \leq g^+(\rho\boldsymbol{w}) \leq 0, & \textit{the Hessian matrix of}\ g^+(\cdot) = \widetilde{\boldsymbol{M}}^+, \\ -\psi/4 \leq g^-(\rho\boldsymbol{w}) \leq 0, & \textit{the Hessian matrix of}\ g^-(\cdot) = -\widetilde{\boldsymbol{M}}^-. \end{array}\right\} \tag{14}$$

Proof: We shall only show the existence of $\delta = \delta^+ > 0$ such that every δ^+-net D_1 of $D(\boldsymbol{c}, \kappa)$ satisfies property P1 restricted to the relations involving a quadratic function $g^+(\cdot)$. We can similarly derive the existence of $\delta = \delta^- > 0$ such that every δ^--net D_1 of $D(\boldsymbol{c}, \kappa)$ satisfies property P1 restricted to the relations involving a quadratic function $g^-(\cdot)$. Letting $\delta = \min\{\delta^+, \ \delta^-\}$, we obtain the desired $\delta > 0$.

For every $\rho > 0$, $\theta \in (0, \pi/8]$ and $\boldsymbol{w} \in D(\boldsymbol{c}, \kappa/2)$, we consider the quadratic function $f^+(\cdot)$ of the form (9). By the definitions of $\boldsymbol{b}_i(\theta, \boldsymbol{w})$ and $\bar{\boldsymbol{b}}_i(\theta, \boldsymbol{w})$ in (8) and by Lemma 4.4, for every $\boldsymbol{w}' \in D(\boldsymbol{c}, \kappa/2)$, there is a $\theta(\boldsymbol{w}') > 0$ such that

$$\left.\begin{array}{l} \boldsymbol{b}_i(\theta(\boldsymbol{w}'), \boldsymbol{w}') \in D(\boldsymbol{w}', \kappa/2) \subseteq D(\boldsymbol{c}, \kappa) \ (\forall i = 1, 2, \ldots, n), \\ \bar{\boldsymbol{b}}_i(\theta(\boldsymbol{w}'), \boldsymbol{w}') \in D(\boldsymbol{w}', \kappa/2) \subseteq D(\boldsymbol{c}, \kappa) \ (\forall i = 1, 2, \ldots, n), \\ -\psi/(8\tilde{\rho}^2) \leq f^+(\boldsymbol{w}'; 1, \theta(\boldsymbol{w}'), \boldsymbol{w}') \leq 0, \\ \text{the Hessian matrix of}\ f^+(\cdot; \rho, \theta(\boldsymbol{w}'), \boldsymbol{w}') = \widetilde{\boldsymbol{M}}^+ \ (\forall \rho > 0). \end{array}\right\} \tag{15}$$

For simplicity of notation, we will write

$$\begin{array}{c} \boldsymbol{b}_i(\boldsymbol{w}') \equiv \boldsymbol{b}_i(\theta(\boldsymbol{w}'), \boldsymbol{w}'), \ \bar{\boldsymbol{b}}_i(\boldsymbol{w}') \equiv \bar{\boldsymbol{b}}_i(\theta(\boldsymbol{w}'), \boldsymbol{w}'), \\ \lambda_i(\boldsymbol{w}') \equiv \lambda_i(\theta(\boldsymbol{w}'), \boldsymbol{w}'), \ \bar{\lambda}_i(\boldsymbol{w}') \equiv \bar{\lambda}_i(\theta(\boldsymbol{w}'), \boldsymbol{w}') \end{array}$$

$(\forall i = 1, 2, \ldots, n)$. For every $\boldsymbol{x} \in R^n$, $\rho \in (0, \tilde{\rho}]$, $\boldsymbol{w}' \in \overline{D}$, $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n) \in \overline{D}^n$, $\overline{\boldsymbol{A}} = (\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2, \ldots, \bar{\boldsymbol{a}}_n) \in \overline{D}^n$, $\boldsymbol{M}^+ = [\mu^+_{ij}]$, define

$$\begin{array}{rcl} g^+(\boldsymbol{x}; \rho, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \boldsymbol{M}^+) & \equiv & -\displaystyle\sum_{i=1}^n \sum_{j=1}^n \mu^+_{ij} \Big(\lambda_i(\boldsymbol{w}')(\boldsymbol{a}_i^T \boldsymbol{x} - \rho)(-\boldsymbol{e}_j^T \boldsymbol{x} - \rho) \\ & & +\bar{\lambda}_i(\boldsymbol{w}')(\bar{\boldsymbol{a}}_i^T \boldsymbol{x} - \rho)(\boldsymbol{e}_j^T \boldsymbol{x} - \rho) \Big). \end{array}$$

Then

$$\begin{array}{rcl} g^+(\cdot; \rho, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \boldsymbol{M}^+) & \in & \text{c.cone}(\mathcal{P}^2(B(\boldsymbol{0}, \rho), D_I, \overline{D})) \\ & & (\forall \boldsymbol{w}' \in \overline{D}, \forall \rho \in (0, \tilde{\rho}], \ \forall \boldsymbol{A} \in \overline{D}^n \ \text{ and } \forall \overline{\boldsymbol{A}} \in \overline{D}^n). \end{array}$$

We observe that

$$\begin{array}{rl} & \text{the Hessian matrix of}\ g^+(\cdot; \rho, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \boldsymbol{M}^+) \\ = & -\displaystyle\sum_{i=1}^n \sum_{j=1}^n \mu^+_{ij} \frac{(-\lambda_i(\boldsymbol{w}')\boldsymbol{a}_i\boldsymbol{e}_j^T + \bar{\lambda}_i(\boldsymbol{w}')\bar{\boldsymbol{a}}_i\boldsymbol{e}_j^T) + (-\lambda_i(\boldsymbol{w}')\boldsymbol{a}_i\boldsymbol{e}_j^T + \bar{\lambda}_i(\boldsymbol{w}')\bar{\boldsymbol{a}}_i\boldsymbol{e}_j^T)^T}{2} \\ = & \dfrac{\boldsymbol{H}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})\boldsymbol{M}^+ + \boldsymbol{M}^+\boldsymbol{H}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})^T}{2}. \end{array}$$

Here

$$\begin{aligned}
\boldsymbol{H}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}) &\equiv (\boldsymbol{h}_1(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}), \boldsymbol{h}_2(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}), \ldots, \boldsymbol{h}_n(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})), \\
\boldsymbol{h}_i(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}) &\equiv \lambda_i(\boldsymbol{w}')\boldsymbol{a}_i - \bar{\lambda}_i(\boldsymbol{w}')\bar{\boldsymbol{a}}_i \ (i = 1, 2, \ldots, n).
\end{aligned}$$

Note that if we take $\boldsymbol{a}_i = \boldsymbol{b}_i(\boldsymbol{w}')$ and $\bar{\boldsymbol{a}}_i = \bar{\boldsymbol{b}}_i(\boldsymbol{w}')$ $(i = 1, 2, \ldots, n)$, we know that

$$\begin{aligned}
g^+(\boldsymbol{x}; \rho, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \boldsymbol{M}^+) &= f^+(\boldsymbol{x}; \rho, \theta(\boldsymbol{w}'), \boldsymbol{w}') \ (\forall \rho > 0 \ \text{ and } \forall \boldsymbol{x} \in R^n), \\
\boldsymbol{H}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}) &= \boldsymbol{I}.
\end{aligned}$$

By the continuity with respect to a perturbation to $\boldsymbol{b}_i(\boldsymbol{w}')$, $\bar{\boldsymbol{b}}_i(\boldsymbol{w}')$ at $\rho = 1$ and $\boldsymbol{x} = \boldsymbol{w}'$, there is a $\delta^+(\boldsymbol{w}') > 0$ such that whenever

$$\boldsymbol{u} \in D(\boldsymbol{w}', \delta^+(\boldsymbol{w}')), \ \boldsymbol{a}_i \in D(\boldsymbol{b}_i(\boldsymbol{w}'), \delta^+(\boldsymbol{w}')), \ \bar{\boldsymbol{a}}_i \in D(\bar{\boldsymbol{b}}_i(\boldsymbol{w}'), \delta^+(\boldsymbol{w}')) \ (\forall i = 1, 2, \ldots, n),$$

the Lyapunov type linear equation

$$\frac{\boldsymbol{H}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})\boldsymbol{M}^+ + \boldsymbol{M}^+\boldsymbol{H}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})^T}{2} = \widetilde{\boldsymbol{M}}^+$$

with the Hessian matrix $\boldsymbol{H}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})$ of $g^+(\cdot; \rho, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \boldsymbol{M}^+)$ as its coefficient matrix has a solution $\boldsymbol{M}^+ = \widehat{\boldsymbol{M}}^+(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}) \in \mathcal{S}^n$ such that

$$\begin{aligned}
&\widehat{\boldsymbol{M}}^+(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}) = [\hat{\mu}^+_{ij}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})] \in \mathcal{S}^n; \\
&\qquad\qquad \hat{\mu}^+_{ij}(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}) \geq 1/2 \ (i = 1, 2, \ldots, n, \ j = 1, 2, \ldots, n), \\
&g^+(\cdot; \rho, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})) \in \text{c.cone}(\mathcal{P}^2(B(\boldsymbol{0}, \rho), D_I, \overline{D})), \\
&-\psi/(4\tilde{\rho}^2) \leq g^+(\boldsymbol{u}; 1, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})) \leq 0, \\
&\text{the Hessian matrix of } g^+(\cdot; \rho, \boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}', \boldsymbol{A}, \overline{\boldsymbol{A}})) = \widetilde{\boldsymbol{M}}^+.
\end{aligned}$$

(Note that the Lyapunov type linear equation has a unique solution $\boldsymbol{M}^+$ for every $\widetilde{\boldsymbol{M}}^+$ if $\boldsymbol{H}$ is positive definite.) Since $D(\boldsymbol{c}, \kappa/2)$ is compact, we can find a finite number of unit direction vectors $\boldsymbol{w}^1, \boldsymbol{w}^2, \ldots, \boldsymbol{w}^\ell \in D(\boldsymbol{c}, \kappa/2)$ such that $\cup_{j=1}^{\ell} D(\boldsymbol{w}^j, \delta^+(\boldsymbol{w}^j)) \supseteq D(\boldsymbol{c}, \kappa/2)$. Let $\delta^+ \equiv \min\{\delta^+(\boldsymbol{w}^1), \delta^+(\boldsymbol{w}^2), \ldots, \delta^+(\boldsymbol{w}^\ell)\}$, and let D_1 be a δ^+-net of $D(\boldsymbol{c}, \kappa)$.

Now, assume (13). Then we can find a $j \in \{1, 2, \ldots, \ell\}$ such that $\boldsymbol{w} \in D(\boldsymbol{w}^j, \delta^+(\boldsymbol{w}^j))$, and the relations in (15) hold for $\boldsymbol{w}' = \boldsymbol{w}^j$. Furthermore there exist

$$\boldsymbol{a}_i \in D(\boldsymbol{b}_i(\boldsymbol{w}^j), \delta^+(\boldsymbol{w}^j)) \cap D_1, \quad \bar{\boldsymbol{a}}_i \in D(\bar{\boldsymbol{b}}_i(\boldsymbol{w}^j), \delta^+(\boldsymbol{w}^j)) \cap D_1 \ (\forall i = 1, 2, \ldots, n).$$

Therefore, we can conclude that

$$\begin{aligned}
&g^+(\cdot; \rho, \boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}})) \in \text{c.cone}\mathcal{P}^2(B(\boldsymbol{0}, \rho), D_I, D_1), \\
&-\psi/(4\tilde{\rho}^2) \leq g^+(\boldsymbol{w}; 1, \boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}})) \leq 0, \\
&-\psi/4 \leq \rho^2 g^+(\boldsymbol{w}; 1, \boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}})) = g^+(\rho\boldsymbol{w}; \rho, \boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}})) \leq 0, \\
&\text{the Hessian matrix of } g^+(\cdot; \rho, \boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}}, \widehat{\boldsymbol{M}}^+(\boldsymbol{w}^j, \boldsymbol{A}, \overline{\boldsymbol{A}})) = \widetilde{\boldsymbol{M}}^+.
\end{aligned}$$

∎

Lemma 4.6 *As in Lemma 4.5, suppose that (12) holds. Then there exists a $\delta > 0$ such that every δ-net D_1 of $D(\boldsymbol{c}, \kappa)$ satisfies the following property.*

P2: *In addition to (13) of Lemma 4.5, assume that*

$$\left.\begin{array}{lcl} C_k & : & \textit{a compact convex subset of } R^n;\ C_{k+1} \subseteq C_k \ (k=0,1,\dots), \\ \rho \boldsymbol{w} & \in & C_\infty \equiv \cap_{k=0}^\infty C_k \subseteq B(\boldsymbol{0},\rho), \end{array}\right\} \quad (16)$$

Then there are quadratic functions $f_k^+(\cdot)$, $f_k^-(\cdot)$ *and* $f_k(\cdot) \in$ c.cone$(\mathcal{P}^2(C_k, D_I, D_1))$ $(k = 0,1,\dots)$ *that satisfy*

$$\left.\begin{array}{ll} -\psi/2 \leq f_k^+(\rho\boldsymbol{w}) \leq 0, & \textit{the Hessian matrix of } f_k^+(\cdot) \ = \widetilde{\boldsymbol{M}}^+, \\ -\psi/2 \leq f_k^-(\rho\boldsymbol{w}) \leq 0, & \textit{the Hessian matrix of } f_k^-(\cdot) \ = -\widetilde{\boldsymbol{M}}^-, \\ -\psi \leq f_k(\rho\boldsymbol{w}) \leq 0, & \textit{the Hessian matrix of } f_k(\cdot) \ = \widetilde{\boldsymbol{M}}. \end{array}\right\} \quad (17)$$

for every sufficiently large k.

Proof: By Lemma 4.5, there exists a $\delta > 0$ such that every δ-net D_1 of $D(\boldsymbol{c},\kappa)$ satisfies property P1. Assume that (13) and (16) hold. Then, by property P1, there are quadratic functions $g^+(\cdot)$, $g^-(\cdot) \in$ c.cone$(\mathcal{P}^2(B(\boldsymbol{0},\rho), D_I, D_1))$ for which the relations in (14) hold. In the rest of the proof, modifying $g^+(\cdot) \in$ c.cone$(\mathcal{P}^2(B(\boldsymbol{0},\rho), D_I, D_1))$, we shall only construct quadratic functions $f_k^+(\cdot) \in$ c.cone$(\mathcal{P}^2(C_k, D_I, D_1))$ $(k = 0,1,\dots)$ that satisfy the first two relations in (17) for every sufficiently large k. Modifying $g^-(\cdot) \in$ c.cone$(\mathcal{P}^2(B(\boldsymbol{0},\rho), D_I, D_1))$, we can similarly construct quadratic functions $f_k^-(\cdot) \in$ c.cone$(\mathcal{P}^2(C_k, D_I, D_1))$ $(k = 0,1,\dots)$ that satisfy the middle two relations in (17) for every sufficiently large k. Finally letting $f_k(\cdot) \equiv f_k^+(\cdot) + f_k^-(\cdot)$ $(k = 0,1,2,\dots)$, we obtain quadratic functions $f_k(\cdot) \in$ c.cone$(\mathcal{P}^2(C_k, D_I, D_1))$ $(k = 0,1,2,\dots)$ that satisfy the last two relations in (17) for every sufficiently large k.

We recall that the quadratic function $g^+(\cdot) \in$ c.cone$(\mathcal{P}^2(B(\boldsymbol{0},\rho), D_I, D_1))$ can be written as

$$g^+(\boldsymbol{x}) = -\sum_{i=1}^n \sum_{j=1}^n \mu_{ij}^+ \left(\lambda_i (\boldsymbol{a}_i^T\boldsymbol{x} - \rho)(-\boldsymbol{e}_j^T\boldsymbol{x} - \rho) + \bar{\lambda}_i (\bar{\boldsymbol{a}}_i^T\boldsymbol{x} - \rho)(\boldsymbol{e}_j^T\boldsymbol{x} - \rho) \right),$$

for some $\mu_{ij}^+ \geq 1/2$, $\lambda_i \geq 0$, $\bar{\lambda}_i \geq 0$, $\boldsymbol{a}_i \in D_1$ and $\bar{\boldsymbol{a}}_i \in D_1$ $(i = 1,2,\dots,n,\ j = 1,2,\dots,m)$. For every $k = 0,1,2,\dots$ and $i = 1,2,\dots,m$, let

$$\begin{array}{ll} \alpha_{ik} = \max\{\boldsymbol{a}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}, & \bar{\alpha}_{ik} = \sup\{\bar{\boldsymbol{a}}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}, \\ \beta_{ik} = \max\{\boldsymbol{e}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}, & \bar{\beta}_{ik} = \max\{-\boldsymbol{e}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}. \end{array}$$

Then, for every $k = 0,1,2,\dots$,

$$\boldsymbol{a}_i^T\boldsymbol{x} - \alpha_{ik} \leq 0,\ \bar{\boldsymbol{a}}_i^T\boldsymbol{x} - \bar{\alpha}_{ik} \leq 0,\ \boldsymbol{e}_i^T\boldsymbol{x} - \beta_{ik} \leq 0 \ \text{ and } \ -\boldsymbol{e}_i^T\boldsymbol{x} - \bar{\beta}_{ik} \leq 0$$

$(i = 1,2,\dots,n)$ are all linear valid inequalities for C_k. Since $\rho\boldsymbol{w} \in \cap_{k=0}^\infty C_k \subseteq B(\boldsymbol{0},\rho)$, we know that $\rho\boldsymbol{w}$ satisfies all the inequalities above for every $k = 0,1,2,\dots$. It follows from $C_k \supseteq C_{k+1}$ $(\forall k = 0,1,2,\dots)$ and $C_\infty = \cap_{k=0}^\infty C_k \subseteq B(\boldsymbol{0},\rho)$ that

$$\begin{array}{lcl} \alpha_{ik} & \geq & \alpha_{i(k+1)} \geq \alpha_{i\infty} \equiv \lim_{k\to\infty} \alpha_{ik} = \max\{\boldsymbol{a}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\} \leq \rho, \\ \bar{\alpha}_{ik} & \geq & \bar{\alpha}_{i(k+1)} \geq \bar{\alpha}_{i\infty} \equiv \lim_{k\to\infty} \bar{\alpha}_{ik} = \max\{\bar{\boldsymbol{a}}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\} \leq \rho, \\ \beta_{ik} & \geq & \beta_{i(k+1)} \geq \beta_{i\infty} \equiv \lim_{k\to\infty} \beta_{ik} = \max\{\boldsymbol{e}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\} \leq \rho, \\ \bar{\beta}_{ik} & \geq & \bar{\beta}_{i(k+1)} \geq \bar{\beta}_{i\infty} \equiv \lim_{k\to\infty} \bar{\beta}_{ik} = \max\{-\boldsymbol{e}_i^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\} \leq \rho \end{array}$$

$(i = 1, 2, \dots, n)$. For every $k = 0, 1, 2, \dots$, we define the quadratic function $f_k^+(\cdot)$ by

$$f_k^+(\boldsymbol{x}) \equiv -\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_{ij}^+ \left(\lambda_i(\boldsymbol{a}_i^T\boldsymbol{x} - \alpha_{ik})(-\boldsymbol{e}_j^T\boldsymbol{x} - \bar{\beta}_{jk}) + \bar{\lambda}_i(\bar{\boldsymbol{a}}_i^T\boldsymbol{x} - \bar{\alpha}_{ik})(\boldsymbol{e}_j^T\boldsymbol{x} - \beta_{jk})\right).$$

Then each quadratic function $f_k^+(\cdot)$ belongs to c.cone$(\mathcal{P}^2(C_k, D_I, D_1))$ and has the same Hessian matrix as the quadratic function $g^+(\cdot)$; hence the second identity of (17) holds for every $k = 0, 1, 2, \dots$. We also see that

$$\begin{aligned}
0 &\geq f_{k+1}^+(\rho\boldsymbol{w}) \geq f_k^+(\rho\boldsymbol{w}) \ (\forall k = 0, 1, 2, \dots), \\
0 &\geq \lim_{k\to\infty} f_k^+(\rho\boldsymbol{w}) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_{ij}^+ \left(\lambda_i(\boldsymbol{a}_i^T\rho\boldsymbol{w} - \alpha_{i\infty})(-\boldsymbol{e}_j^T\rho\boldsymbol{w} - \bar{\beta}_{j\infty}) + \bar{\lambda}_i(\bar{\boldsymbol{a}}_i^T\rho\boldsymbol{w} - \bar{\alpha}_{i\infty})(\boldsymbol{e}_j^T\rho\boldsymbol{w} - \beta_{i\infty})\right) \\
&\geq g^+(\rho\boldsymbol{w}).
\end{aligned}$$

Hence, for every sufficiently large k, we have that $-\psi/2 \leq g^+(\rho\boldsymbol{w}) - \psi/4 \leq f_k^+(\rho\boldsymbol{w}) \leq 0$. Thus, we obtain the first relation of (17) for every sufficiently large k. ∎

5 Proof of Theorem 3.3.

For every $\psi \geq 0$ and $\rho > 0$, define

$$\begin{aligned}
F(\psi) &\equiv \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq \psi \ (\forall p(\cdot) \in \mathcal{P}_F)\}, \\
\text{c.hull}(F, \psi, \rho) &\equiv \cap\{B(\boldsymbol{x}_c, \rho') : F(\psi) \subseteq B(\boldsymbol{x}_c, \rho'),\ \boldsymbol{x}_c \in R^n,\ 0 < \rho' \leq \rho\} \\
&= \text{the intersection of all } n\text{-dimensional balls} \\
& \text{ including } F(\psi) \text{ with radii not greater than } \rho.
\end{aligned}$$

By definition, we know that

$$F = F(0) \subseteq F(\psi_1) \subseteq F(\psi_2) \ (0 \leq \forall\psi_1 < \forall\psi_2),$$

$$\left.\begin{aligned}
\text{c.hull}(F, \psi, \rho_1) &\supseteq \text{c.hull}(F, \psi, \rho_2) \\
&\supseteq \cap_{\rho>0}\text{c.hull}(F, \psi, \rho) \\
&= \text{c.hull}(F(\psi))
\end{aligned}\right\} \ (0 \leq \forall\psi,\ 0 < \forall\rho_1 < \forall\rho_2). \tag{18}$$

(For the last identity, see Lemma 5.2 of [10]). Note that $F(\psi)$ depends on the quadratic inequality representation $\mathcal{P}_F$, which is fixed throughout the paper. We call c.hull(F, ψ, ρ) *a* (ψ, ρ)*-approximation of* c.hull(F).

Lemma 5.1 *Let $U \subseteq R^n$ be an open convex set containing F. Then there exist $\psi > 0$ and $\rho > 0$ such that* c.hull(F, ψ, ρ) *is contained in U; $F \subseteq$* c.hull$(F, \psi, \rho) \subseteq U$. *(Note that U can be empty when F is empty).*

Proof: We first show that c.hull$(F(\psi)) \subseteq U$ for some $\psi > 0$. Assume on the contrary that there exist sequences $\{\psi_k > 0\ (k = 1, 2, \dots)\}$ and $\{\boldsymbol{x}_k \in R^n\ (k = 1, 2, \dots)\}$ for which

$$\begin{aligned}
&\psi_k \to 0 \ \text{ as } k \to 0,\ \boldsymbol{x}_k \notin U \ (\forall k = 1, 2, \dots), \\
&\boldsymbol{x}_k \in F(\psi_k),\ \textit{i.e.},\ p(\boldsymbol{x}_k) \leq \psi_k \ (\forall p(\cdot) \in \mathcal{P}_F) \ (\forall k = 1, 2, \dots).
\end{aligned}$$

Since the sequence $\{\boldsymbol{x}_k\ (k=1,2,\dots)\}$ lies in the compact set $C_0\backslash U$, we may assume that it converges to some $\bar{\boldsymbol{x}} \in C_0\backslash U$. On the other hand, we see by the continuity that $p(\bar{\boldsymbol{x}}) = \lim_{k\to\infty} p(\boldsymbol{x}_k) \leq \lim_{k\to\infty}\psi_k = 0$ for every $p(\cdot) \in \mathcal{P}_F$; hence $\bar{\boldsymbol{x}} \in F$. This contradicts the assumption that $F \subseteq U$. Thus, we have shown that c.hull$(F(\psi)) \subseteq U$ for some $\psi > 0$. Now, recall the relations in (18). Since each c.hull(F,ψ,ρ) is a compact subset of C_0, we can conclude that c.hull$(F,\psi,\rho) \subseteq U$ for every sufficiently large $\rho > 0$. ∎

Proof of Theorem 3.3: We will simultaneously deal with the two cases, the one where the sequence $\{C_k\ (k=0,1,2\dots)\}$ is generated by Algorithm 3.1 and the other where the sequence is generated by Algorithm 3.2. By Lemma 4.2, we know that the monotonicity property (a) holds. So we only deal with property (b)'.

(i) By Lemma 5.1, c.hull$(F,\psi,\bar{\rho}) \subseteq U$ for some $\psi > 0$ and $\bar{\rho} > 0$. If F is empty, we can choose $\psi > 0$ and $\bar{\rho} > 0$ such that c.hull$(F,\psi,\bar{\rho})$ is empty. Let $\kappa \geq 4$, $\tilde{\rho} \equiv \bar{\rho} + \eta_0$, where η_0 denotes a positive number larger than the diameter $\sup\{\|\boldsymbol{x}'-\boldsymbol{x}\| : \boldsymbol{x}',\ \boldsymbol{x} \in C_0\}$ of C_0. Then $D(\boldsymbol{c},\kappa/2) = \overline{D}$. By Lemma 4.6, there exists a $\delta > 0$ such that every δ-net of D_1 of $\overline{D}$ satisfies property P2. We will prove this $\delta > 0$ fulfills the requirement of Theorem 3.3. Let a δ-net D_1 of $\overline{D}$ be fixed.

(ii) Let $C_\infty \equiv \cap_{k=0}^{\infty} C_k$. Since each C_k is compact, so is C_∞. If $C_\infty = \emptyset$ then $C_k = \emptyset$ for some finite number k and $F = \emptyset$. Hence property (b)' follows in this case. Therefore, we will only deal with the case $C_\infty \neq \emptyset$ in the rest of the proof.

(iii) Since we know that c.hull$(F,\psi,\bar{\rho}) \subseteq U$, it suffices to show that $C_\infty \subseteq$ c.hull$(F,\psi,\bar{\rho})$. We assume on the contrary that there is an $\bar{\boldsymbol{x}} \in C_\infty\backslash$c.hull$(F,\psi,\bar{\rho})$, and we will derive a contradiction. If $F(\psi) = \emptyset$, let $\boldsymbol{x}_c = \bar{\boldsymbol{x}}$. Otherwise, we can take an n-dimensional ball $B(\boldsymbol{x}_c,\rho')$ with $\rho' \in (0,\bar{\rho}]$ such that $F(\psi) \subseteq B(\boldsymbol{x}_c,\rho')$ and $\bar{\boldsymbol{x}} \notin B(\boldsymbol{x}_c,\rho')$. In both cases, let $\rho \equiv \sup\{\|\boldsymbol{x}-\boldsymbol{x}_c\| : \boldsymbol{x} \in C_\infty\}$, and choose an $\hat{\boldsymbol{x}} \in C_\infty$ such that $\rho \equiv \|\hat{\boldsymbol{x}} - \boldsymbol{x}_c\|$. Then $\hat{\boldsymbol{x}} \notin F(\psi)$. Letting $\tilde{\boldsymbol{x}} = \boldsymbol{x}_c$ if $F(\psi) = \emptyset$ or $\tilde{\boldsymbol{x}} \in F(\psi) \subseteq C_0$, we also see that $\rho = \|\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}} + \tilde{\boldsymbol{x}} - \boldsymbol{x}_c\| \leq \eta_0 + \bar{\rho} = \tilde{\rho}$. Hence we obtain that

$$\rho \in (0,\tilde{\rho}],\ \hat{\boldsymbol{x}} \notin F(\psi) \ \text{ and } \hat{\boldsymbol{x}} \in C_\infty = \cap_{k=0}^{\infty} C_k \subseteq B(\boldsymbol{x}_c,\rho). \tag{19}$$

Let $\boldsymbol{w} \equiv (\hat{\boldsymbol{x}} - \boldsymbol{x}_c)/\rho \in \overline{D}$. We also see from $\hat{\boldsymbol{x}} \notin F(\psi)$ that there is a $p_0(\cdot) \in \mathcal{P}_F$ such that $p_0(\hat{\boldsymbol{x}}) > \psi$.

(iv) Let $\widetilde{\boldsymbol{M}}$ be the negative of the Hessian matrix of the quadratic function $p_0(\cdot)$. By Lemma 4.3, we may assume, without loss of generality, that $\boldsymbol{x}_c = \boldsymbol{0}$ in (19); if necessary, apply the parallel transformation which maps $\boldsymbol{x}_c$ to the origin. Note that such a transformation preserves the Hessian matrix of the quadratic function $p_0(\cdot)$. Then, all the assumptions (13) and (16) in property P2 are satisfied. Therefore, if k is sufficiently large then there is a quadratic function

$$f_k(\cdot) \in \text{c.cone}(\mathcal{P}^2(C_k,D_I,D_1)) \subseteq \text{c.cone}(\mathcal{P}_k)$$

for which (17) holds. We consider the quadratic function $p_0(\cdot) + f_k(\cdot)$. Then, we observe that the Hessian matrix of the function vanishes; hence

$$p_0(\cdot) + f_k(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}^2(C_k,D_I,D_1)) \cap \mathcal{L}.$$

We also see that

$$p_0(\hat{\boldsymbol{x}}) + f_k(\hat{\boldsymbol{x}}) = p_0(\rho\boldsymbol{w}) + f_k(\rho\boldsymbol{w}) \quad > \quad \psi - \psi = 0.$$

Therefore,

$$\begin{aligned} \hat{\boldsymbol{x}} = \rho\boldsymbol{w} \quad &\notin \quad \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0\ (\forall p(\cdot) \in \text{c.cone}(\mathcal{P}_F \cup \mathcal{P}_k) \cap \mathcal{L})\} \\ &= \quad \widehat{F}^L(C_0,\mathcal{P}_F \cup \mathcal{P}_k) \text{ (by Lemma 4.1).} \end{aligned}$$

On the other hand, we know that $C_{k+1} \subseteq \widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$; hence $\hat{\boldsymbol{x}} = \rho \boldsymbol{w} \notin C_{k+1}$. This is a contradiction to $\hat{\boldsymbol{x}} = \rho \boldsymbol{w} \in C_\infty = \cap_{k=0}^\infty C_k$. This completes the proof of the theorem.

6 Proof of Theorem 3.5.

Since $\boldsymbol{c} \neq \boldsymbol{0}$, we may assume, without loss of generality, that $\|\boldsymbol{c}\| = 1$ throughout this section. Theorem 3.5 is proved in a similar way as Theorem 3.3. In particular, the proof relies on Lemma 4.6. Lemmas 6.1 and 6.2 furnish two constants $\tilde{\rho} >$ and $\psi > 0$ which appear in the assumption (12) of Lemma 4.6.

Lemma 6.1 *Let $\epsilon > 0$ and $\eta_0 > \max\{\|\boldsymbol{x} - \boldsymbol{x}'\| : \boldsymbol{x}, \boldsymbol{x}' \in C_0\}$. Suppose that the set*

$$C_0' \equiv \left\{ \boldsymbol{x}' \in C_0 : \boldsymbol{c}^T \boldsymbol{x}' \geq \sup\{\boldsymbol{c}^T \boldsymbol{x} : \boldsymbol{x} \in F\} + \epsilon \right\} \tag{20}$$

is nonempty.

(i) Let $\rho \geq (\eta_0^2 + 2\epsilon^2)/\epsilon$. Then $F \subseteq \mathrm{int}\left(\cap_{\boldsymbol{x}' \in C_0'} B_{\mathrm{int}}(\boldsymbol{x}' - \rho \boldsymbol{c}, \rho)\right)$.

(ii) Let $\kappa > 0$. Suppose that $C_\infty \subseteq C_0$ is a nonempty compact convex set which contains F and intersects with C_0' at an $\bar{\boldsymbol{x}} \in C_\infty \cap C_0'$. Let

$$\left.\begin{array}{l} \bar{\rho} \equiv \max\left\{(\eta_0^2 + 2\epsilon^2)/\epsilon,\ 4\eta_0/\kappa\right\},\ \tilde{\rho} \equiv \bar{\rho} + \eta_0, \\ G \equiv C_0 \backslash \mathrm{int}\left(\cap_{\boldsymbol{x}' \in C_0'} B_{\mathrm{int}}(\boldsymbol{x}' - \bar{\rho} \boldsymbol{c}, \bar{\rho})\right). \end{array}\right\} \tag{21}$$

Then, there exist $\hat{\boldsymbol{x}} \in C_\infty$, $\boldsymbol{w} \in \overline{D}$ and $\rho \in [\bar{\rho}, \tilde{\rho}]$ such that

$$\hat{\boldsymbol{x}} \in G, \text{ (hence } \hat{\boldsymbol{x}} \notin F), \tag{22}$$

$$\boldsymbol{w} \in D(\boldsymbol{c}, \kappa/2) \quad (i.e., \ \boldsymbol{w} \in \overline{D} \text{ and } \|\boldsymbol{w} - \boldsymbol{c}\| < \kappa/2)\ , \tag{23}$$

$$C_\infty \subseteq B(\hat{\boldsymbol{x}} - \rho \boldsymbol{w}, \rho) \ (i.e., \ \|\boldsymbol{x} - (\hat{\boldsymbol{x}} - \rho \boldsymbol{w})\| \leq \rho \text{ for every } \boldsymbol{x} \in C_\infty). \tag{24}$$

(Note that F can be empty).

Proof: (i) If $F = \emptyset$ then the inclusion relation in (i) is obvious. So we only deal with the case that $F \neq \emptyset$. By assumption, we first observe that $\rho^2 \geq (\eta_0^2/\epsilon + 2\epsilon)^2 > \eta_0^2 + 3\epsilon^2$. Let δ be a positive number such that $\delta + \sqrt{\rho^2 - \eta_0^2 - 3\epsilon^2} \leq \rho$. We will show that

$$B(\boldsymbol{x}, \delta) \subseteq B(\boldsymbol{x}' - \rho \boldsymbol{c}, \rho) \ (\forall \boldsymbol{x} \in F \ \text{ and } \forall \boldsymbol{x}' \in C_0'). \tag{25}$$

Then, the inclusion relation in (i) follows. Assume that $\boldsymbol{x} \in F$ and $\boldsymbol{x}' \in C_0'$. Then

$$\rho \geq (\eta_0^2 + 2\epsilon^2)/\epsilon > \eta_0 \ \text{ and } \boldsymbol{c}^T \boldsymbol{x}' - \epsilon \geq \boldsymbol{c}^T \boldsymbol{x} = \boldsymbol{c}^T \boldsymbol{x}' + \boldsymbol{c}^T(\boldsymbol{x} - \boldsymbol{x}') \geq \boldsymbol{c}^T \boldsymbol{x}' - \eta_0. \tag{26}$$

Here, the last inequality follows from the relations $\|\boldsymbol{c}\| = 1$ and $\boldsymbol{x},\ \boldsymbol{x}' \in C_0$. Hence

$$\rho - \epsilon \geq \rho + \boldsymbol{c}^T \boldsymbol{x} - \boldsymbol{c}^T \boldsymbol{x}' \geq \rho - \eta_0 > 0. \tag{27}$$

Let $\boldsymbol{x}'' = \boldsymbol{c}^T(\boldsymbol{x}-\boldsymbol{x}')\boldsymbol{c}+\boldsymbol{x}'$. Note that the first term $\boldsymbol{c}^T(\boldsymbol{x}-\boldsymbol{x}')\boldsymbol{c}$ is the orthogonal projection of the vector $\boldsymbol{x}-\boldsymbol{x}'$ onto the direction $\boldsymbol{c}$. Then, we obtain that

$$\begin{array}{rcl}
\boldsymbol{c}^T(\boldsymbol{x}''-\boldsymbol{x}) & = & \boldsymbol{c}^T\boldsymbol{x}'' - \boldsymbol{c}^T\boldsymbol{x} = \boldsymbol{c}^T(\boldsymbol{x}-\boldsymbol{x}') + \boldsymbol{c}^T\boldsymbol{x}' - \boldsymbol{c}^T\boldsymbol{x} = 0 \text{ (since } \|\boldsymbol{c}\|=1).\\
\eta_0^2 & \geq & \|\boldsymbol{x}-\boldsymbol{x}'\|^2 \text{ (since } \boldsymbol{x}',\ \boldsymbol{x}\in C_0)\\
& = & \|\boldsymbol{x}-\boldsymbol{x}''+\boldsymbol{x}''-\boldsymbol{x}'\|^2\\
& = & \|\boldsymbol{x}-\boldsymbol{x}''\|^2 + \left\|\boldsymbol{c}^T(\boldsymbol{x}-\boldsymbol{x}')\boldsymbol{c}\right\|^2 \text{ (by (28))}\\
& \geq & \|\boldsymbol{x}-\boldsymbol{x}''\|^2, \qquad\qquad (28)\\
\|\boldsymbol{x}-(\boldsymbol{x}'-\rho\boldsymbol{c})\|^2 & = & \|(\boldsymbol{x}-\boldsymbol{x}'')+(\rho\boldsymbol{c}+\boldsymbol{x}''-\boldsymbol{x}')\|^2\\
& = & \left\|(\boldsymbol{x}-\boldsymbol{x}'')+(\rho+\boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{c}^T\boldsymbol{x}')\boldsymbol{c}\right\|^2\\
& & \text{(since } \boldsymbol{x}''=\boldsymbol{c}^T(\boldsymbol{x}-\boldsymbol{x}')\boldsymbol{c}+\boldsymbol{x}')\\
& = & \|\boldsymbol{x}-\boldsymbol{x}''\|^2 + (\rho+\boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{c}^T\boldsymbol{x}')^2 \text{ (by (28) and } \|\boldsymbol{c}\|=1)\\
& \leq & \eta_0^2 + (\rho-\epsilon)^2 \text{ (by (28) and (27))}\\
& = & \eta_0^2+\rho^2-2\epsilon\rho+\epsilon^2\\
& \leq & \eta_0^2+\rho^2+\epsilon^2-2\epsilon\left((\eta_0^2+2\epsilon^2)/\epsilon\right) \text{ (by (26))}\\
& = & \rho^2-\eta_0^2-3\epsilon^2.
\end{array}$$

Hence, we obtain that $\boldsymbol{x}\in B\left(\boldsymbol{x}'-\rho\boldsymbol{c}, \sqrt{\rho^2-\eta_0^0-3\epsilon^2}\right)$. Now, assume that $\tilde{\boldsymbol{x}}\in B(\boldsymbol{x},\delta)$. Then, we see that

$$\|\tilde{\boldsymbol{x}}-(\boldsymbol{x}'-\rho\boldsymbol{c})\| \leq \|\tilde{\boldsymbol{x}}-\boldsymbol{x}\| + \|\boldsymbol{x}-(\boldsymbol{x}'-\rho\boldsymbol{c})\| \leq \delta + \sqrt{\rho^2-\eta_0^0-3\epsilon^2} \leq \rho.$$

Thus, we have shown relation (25).

(ii) Let $\rho\equiv\max\{\|\boldsymbol{x}-(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c})\| : \boldsymbol{x}\in C_\infty\}$, and choose $\hat{\boldsymbol{x}}\in C_\infty$ such that $\rho=\|\hat{\boldsymbol{x}}-(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c})\|$. We then see that

$$\begin{array}{rcl}
C_\infty & \subseteq & B(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c},\rho),\ \bar{\rho}\leq\rho \text{ (since } \bar{\boldsymbol{x}}\in C_\infty),\\
\hat{\boldsymbol{x}} & \in & C_0\backslash B_{\rm int}(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c},\bar{\rho})\subseteq G;\ \text{(hence (22) holds)},\\
\rho & = & \|\hat{\boldsymbol{x}}-(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c})\| \leq \|\hat{\boldsymbol{x}}-\bar{\boldsymbol{x}}\|+\bar{\rho}\leq\eta_0+\bar{\rho}\leq\tilde{\rho};
\end{array}$$

hence $\rho\in[\bar{\rho},\tilde{\rho}]$. Now, we define $\boldsymbol{w}=(\hat{\boldsymbol{x}}-(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c}))/\rho\in\overline{D}$. Then $\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c}=\hat{\boldsymbol{x}}-\rho\boldsymbol{w}$. Hence

$$C_\infty \ \subseteq\ B(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c},\rho)=B(\hat{\boldsymbol{x}}-\rho\boldsymbol{w},\rho).$$

Thus, we have shown (24). We also see that

$$\begin{array}{rcl}
\|\boldsymbol{w}-\boldsymbol{c}\| & = & \|(\hat{\boldsymbol{x}}-(\bar{\boldsymbol{x}}-\bar{\rho}\boldsymbol{c}))/\rho-\boldsymbol{c}\|\\
& = & \|(\hat{\boldsymbol{x}}-\bar{\boldsymbol{x}})/\rho+((\bar{\rho}-\rho)/\rho)\boldsymbol{c}\|\\
& \leq & 2\eta_0/\rho \text{ (since } \bar{\boldsymbol{x}},\ \hat{\boldsymbol{x}}\in C_0,\ \bar{\rho}\leq\rho\leq\bar{\rho}+\eta_0 \text{ and } \|\boldsymbol{c}\|=1)\\
& \leq & 2\eta_0/\bar{\rho} \text{ (since } \rho\in[\bar{\rho},\tilde{\rho}])\\
& \leq & \kappa/2 \text{ (since } 4\eta_0/\kappa\leq\bar{\rho}).
\end{array}$$

Thus, we have shown (23). ∎

Lemma 6.2 *Let G be a compact subset of C_0 such that $F\cap G=\emptyset$. Then,* $\inf_{\boldsymbol{x}\in G}\ \sup_{p(\cdot)\in\mathcal{P}_F} p(\boldsymbol{x})>0$.

Proof: Assume on the contrary that $\inf_{\boldsymbol{x}\in G} \sup_{p(\cdot)\in\mathcal{P}_F} p(\boldsymbol{x}) \leq 0$ Then, there exists a sequence $\{\boldsymbol{x}_k \in G \ (k = 1, 2, \ldots, \ )\}$ such that $\lim_{k\to\infty} \sup_{p(\cdot)\in\mathcal{P}_F} p(\boldsymbol{x}_k) \leq 0$. It follows that for every $p(\cdot) \in \mathcal{P}_F$, $p(\boldsymbol{x}_k) \leq \sup_{p(\cdot)\in\mathcal{P}_F} p(\boldsymbol{x}_k) \ (k = 1, 2, \ldots \ )$. Since G is compact, we may assume without loss of generality that $\boldsymbol{x}_k \to \bar{\boldsymbol{x}} \in G$ as $k \to \infty$. Taking the limit as $k \to \infty$ in the inequality above, we obtain that $p(\bar{\boldsymbol{x}}) \leq 0$ for every $p(\cdot) \in \mathcal{P}_F$, which implies that $\bar{\boldsymbol{x}} \in F \cap G$. This is a contradiction to the assumption $F \cap G = \emptyset$. ∎

Proof of Theorem 3.5: We will simultaneously deal with the two cases, the one where the sequence $\{C_k \ (k = 0, 1, 2 \ldots \ )\}$ is generated by Algorithm 3.1 and the other where the sequence is generated by Algorithm 3.2. By Lemma 4.2, we know that the monotonicity property (a) holds. So we only deal with the property (c)'.

(i) Define C_0' by (20). If $C_0' = \emptyset$, then (c)' holds with $k = 0$. Therefore, we only deal with the case where $C_0' \neq \emptyset$. Define $\bar{\rho} > 0$, $\tilde{\rho} > 0$ and $G \subseteq C_0$ by (21). Then, G forms a compact subset of C_0 and by (i) of Lemma 6.1, we know that $F \cap G = \emptyset$. Hence, by Lemma 6.2, we can take a $\psi > 0$ such that for every $\boldsymbol{x} \in G$ there is a $p(\cdot) \in \mathcal{P}_F$ for which $p(\boldsymbol{x}) > \psi$. Applying Lemma 4.6, we can find a $\delta > 0$ such that every δ-net of D_2 of $D(\boldsymbol{c}, \kappa)$ satisfies property P2. We will prove this δ fulfills all the requirements of Theorem 3.5. Let a δ-net of D_2 of $D(\boldsymbol{c}, \kappa)$ be fixed.

(ii) Let $C_\infty \equiv \cap_{k=0}^{\infty} C_k$. Since each C_k is compact, so is C_∞. Define $\zeta_\infty \equiv \lim_{k\to\infty} \zeta_k$. We will show that

$$\zeta_\infty = \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\}. \tag{29}$$

Since $C_\infty \subseteq C_k \ (k = 0, 1, 2, \ldots \ )$, we know that $\zeta_k \geq \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\} \ (k = 0, 1, 2, \ldots \ )$; hence

$$\zeta_\infty \equiv \lim_{k\to\infty} \zeta_k \geq \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\}.$$

Thus, it remains to show $\zeta_\infty \leq \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\}$. If $C_k = \emptyset$ for some finite number k, then both ζ_∞ and $\max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\}$ becomes $-\infty$. Now, assume that $C_k \neq \emptyset \ (\forall k = 0, 1, \ldots \ )$. Then, by the definition of the sequence $\{\zeta_k \ (k = 0, 1, 2, \ldots \ )\} \subseteq R$, there exists a sequence $\{\boldsymbol{x}_k \ (k = 0, 1, 2, \ldots \ )\} \subseteq R^n$ such that $\boldsymbol{x}_k \in C_k \subseteq C_0$ and $\zeta_k = \boldsymbol{c}^T\boldsymbol{x}_k$. Since C_0 is compact, the sequence $\{\boldsymbol{x}_k \ (k = 0, 1, 2, \ldots)\}$ has a limit point $\boldsymbol{x}_\infty \in C_0$. Since $\boldsymbol{x}_k \in C_j \ (\forall j = 0, 1, 2, \ldots, k, \ \forall k = 0, 1, \ldots \ )$, we see that $\boldsymbol{x}_\infty$ is a common accumulation point of $C_k \ (\forall k = 0, 1, \ldots \ )$. Since C_k is compact, we have that $\boldsymbol{x}_\infty \in C_k \ (\forall k = 0, 1, \ldots \ )$; hence $\boldsymbol{x}_\infty \in C_\infty \equiv \cap_{k=0}^{\infty} C_k$. By the continuity, we obtain that $\max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_\infty\} \geq \boldsymbol{c}^T\boldsymbol{x}_\infty = \lim_{k\to\infty} \zeta_k = \zeta_\infty$. Thus, we have shown the identity (29).

(iii) To prove property (c)', it suffices to show that $\zeta_\infty < \zeta^* + \epsilon$ if $F \neq \emptyset$. Assume on the contrary that $\zeta_\infty \geq \zeta^* + \epsilon$ when $F \neq \emptyset$, and that $C_\infty \neq \emptyset$ when $F = \emptyset$. In both cases, there is an $\bar{\boldsymbol{x}} \in C_\infty \cap C_0'$. By (ii) of Lemma 6.1, there exist $\hat{\boldsymbol{x}} \in C_\infty$, $\boldsymbol{w} \in \overline{D}$ and $\rho \in [\bar{\rho}, \tilde{\rho}]$ satisfying (22), (23) and (24). Define $\boldsymbol{x}_c \equiv \hat{\boldsymbol{x}} - \rho\boldsymbol{w}$. We see from (22) that here is a $p_0(\cdot) \in \mathcal{P}_F$ for which $p_0(\hat{\boldsymbol{x}}) > \psi$.

(iv) The rest of the proof is exactly the same as the argument given in part (iv) of the proof of Theorem 3.3, and it is omitted here.

7 The Successive Convex Relaxation Methods and the Lovász-Schrijver Procedures.

In this section, we connect our Algorithm 3.1 (the SSDP Relaxation Method) and Algorithm 3.2 (the SSILP Relaxation Method) to the Lovász-Schrijver lift-and-project procedures given in the paper [11], and our Corollary 3.4 to the related theoretical results of [11]. To make the connection, we focus on 0-1 programs, and assume throughout this section that F is a subset of $\{0,1\}^n$.

7.1 0-1 Convex Optimization Problems.

We begin with the case where a compact convex subset C_0 of $[0,1]^n$, satisfying $F = C_0 \cap \{0,1\}^n$, is given in advance as in the original Lovász-Schrijver procedures. We will discuss a more general case later in Section 7.2. Let

$$\begin{aligned}
\mathcal{K}_0 &\equiv \text{c.cone}\left(\left\{\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in R^{1+n} : \boldsymbol{x} \in C_0\right\}\right) \\
&\equiv \left\{\lambda \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in R^{1+n} : \lambda \geq 0 \ \text{ and } \boldsymbol{x} \in C_0\right\}, \\
\mathcal{K}_I &\equiv \text{c.cone}\left(\left\{\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in \text{c.hull}(F)\right\}\right) \\
&= \text{the convex cone spanned by the 0-1 vectors in } \mathcal{K}_0.
\end{aligned}$$

Here the 0th coordinate is special. It is used in homogenizing the sets of interest from R^n to R^{1+n}. Clearly,

$$C_0 = \left\{\boldsymbol{x} \in R^n : \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \mathcal{K}_0\right\} \ \text{ and } \text{c.hull}(F) = \left\{\boldsymbol{x} \in R^n : \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \mathcal{K}_I\right\}.$$

The closed convex cone $\mathcal{K}_0$ serves as an initial convex relaxation of $\mathcal{K}_I$ in the Lovász-Schrijver lift-and-project procedures. Given the current convex relaxation $\mathcal{K}_k$ of $\mathcal{K}_I$, first a closed convex cone in the space $\mathcal{S}^{1+n}$ of $(1+n) \times (1+n)$ symmetric matrices is defined (the lifting operation). Then a projection of this cone onto R^{1+n} gives the next convex relaxation $\mathcal{K}_{k+1}$ 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}, and $\mathcal{K}^* \equiv \{\boldsymbol{v} \in R^{1+n} : \boldsymbol{u}^T\boldsymbol{v} \geq 0 \ (\forall \boldsymbol{u} \in \mathcal{K})\}$ and $\mathcal{T}^*$ their duals, respectively. A $(1+n) \times (1+n)$ symmetric matrix $\boldsymbol{Y}$ is in $M_+(\mathcal{K}, \mathcal{T})$ if

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

(ii) $\boldsymbol{Y}\boldsymbol{e}_0 = \text{Diag}(\boldsymbol{Y})$, where $\boldsymbol{e}_0 \in R^{1+n}$ denotes the unit vector with the 0th coordinate 1 and $\text{Diag}(\boldsymbol{Y})$ the $(1+n)$-dimensional vector composed of the diagonal elements of $\boldsymbol{Y}$,

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

Now, we describe the projection operation:

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

The iterated operators $N^k_+(\mathcal{K}, \mathcal{T})$ are defined as follows. $N^0_+(\mathcal{K}, \mathcal{T}) \equiv \mathcal{K}$, and $N^k_+(\mathcal{K}, \mathcal{T}) \equiv N_+(N^{k-1}_+(\mathcal{K}, \mathcal{T}), \mathcal{T})$ $(\forall k = 1, 2, \dots)$.

Another procedure studied in [11], uses a weaker relaxation by removing the condition (i) in the lifting operation. Let $M(\mathcal{K}, \mathcal{T}) \subseteq \mathcal{S}^{1+n}$ and $N(\mathcal{K}, \mathcal{T}) \subseteq R^{1+n}$ denote the corresponding sets for this procedure. We will refer to the first procedure using the lifting $M_+(\mathcal{K}, \mathcal{T})$ and the projection

$N_+(\mathcal{K},\mathcal{T})$ as the N_+ procedure. We will call the other (using $M(\mathcal{K},\mathcal{T})$ and $N(\mathcal{K},\mathcal{T})$) the N procedure.

In Section 6 of the authors' previous paper [10], they used the choice $\mathcal{T} \equiv \mathcal{K}_k$, and related their original SSDP and SSILP Relaxation Methods to the Lovász-Schrijver lift-and-project procedures mentioned above. See Theorem 6.3 and its proof of [10]. In the case of 0-1 integer linear programs, one of the important choices for $\mathcal{T}$ is to fix it to be the convex cone of all 0-1 vectors in R^{1+n} with $x_0 = 1$; we assume in the remainder of this section that

$$\mathcal{T} \equiv \text{c.cone}\{(\boldsymbol{e}_0 + \boldsymbol{e}_1), (\boldsymbol{e}_0 + \boldsymbol{e}_2), \ldots, (\boldsymbol{e}_0 + \boldsymbol{e}_n)\},$$

where $\boldsymbol{e}_i \in R^{1+n}$ denotes the unit vector with the ith coordinate 1. (We used the notation $N_+(\mathcal{K},\mathcal{T})$, whereas $N_+(\mathcal{K})$ is used in [11] for this choice of $\mathcal{T}$.) Lovász and Schrijver proved the following two theorems.

Theorem 7.1 (Theorem 1.4 of [11])

$$\mathcal{K}_0 \supseteq N_+(\mathcal{K}_0,\mathcal{T}) \supseteq N_+^2(\mathcal{K}_0,\mathcal{T}) \supseteq \ldots \supseteq N_+^n(\mathcal{K}_0,\mathcal{T}) = \mathcal{K}_I$$

and

$$\mathcal{K}_0 \supseteq N(\mathcal{K}_0,\mathcal{T}) \supseteq N^2(\mathcal{K}_0,\mathcal{T}) \supseteq \ldots \supseteq N^n(\mathcal{K}_0,\mathcal{T}) = \mathcal{K}_I.$$

Theorem 7.2 (Theorem 1.6 of [11]) *Let $\mathcal{K}$ be a closed convex cone in R^{1+n}. Suppose that there exists a* weak separation oracle *for $\mathcal{K}$. Then the weak separation problem for $N(\mathcal{K},\mathcal{T})$ and for $N_+(\mathcal{K},\mathcal{T})$ can be solved in polynomial time.*

(For the definition of *weak separation oracle*, see [9].) This result hinges upon the facts that conditions (i) and (ii) can be checked in polynomial time and that condition (iii) on Y can also be checked in polynomial time since $\mathcal{T}^*$ turns out to be the polyhedral cone,

$$\text{c.cone}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n, (\boldsymbol{e}_0 - \boldsymbol{e}_1), (\boldsymbol{e}_0 - \boldsymbol{e}_2), \ldots, (\boldsymbol{e}_0 - \boldsymbol{e}_n)\},$$

and a weak separation oracle for $\mathcal{K}$ exists.

Now, let us see how our Algorithm 3.1 (the SSDP Relaxation Method) and Algorithm 3.2 (the SSILP Relaxation Method) apply to 0-1 convex quadratic optimization problems. Let $\mathcal{P}_I$ denote the set of quadratic functions

$$x_i(x_i - 1) \leq 0 \ \text{ and } \ -x_i(x_i - 1) \leq 0 \ (\forall i = 1, 2, \ldots, n). \tag{30}$$

Then we can rewrite the 0-1 quadratic optimization problem under consideration as

$$\left.\begin{array}{ll} \text{maximize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in F \equiv \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \mathcal{P}_I)\}. \end{array}\right\} \tag{31}$$

We will show below that if we take direction sets D_0 and D_1 as in Corollary 3.4, then Algorithm 3.1 and Algorithm 3.2 work in a similar way as the Lovász-Schrijver N_+ and N procedures, respectively.

Lemma 7.3 *Let $\{C_k\ (k = 0, 1, \dots)\}$ be the sequence generated by Algorithm 3.1 (the SSDP Relaxation Method) or Algorithm 3.2 (the SSILP Relaxation Method) applied to* 0-1 *QOP (31), using $D_0 = D_I \equiv \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n, -\boldsymbol{e}_1, -\boldsymbol{e}_2, \dots, -\boldsymbol{e}_n\}$ and $D_1 = \overline{D}$. Then*

$$
\begin{aligned}
N_+^k(\mathcal{K}_0, \mathcal{T}) &\supseteq c.cone\left(\left\{\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in C_k\right\}\right) \text{ when we use Algorithm 3.1,} && (32)\\
N^k(\mathcal{K}_0, \mathcal{T}) &\supseteq c.cone\left(\left\{\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in C_k\right\}\right) \text{ when we use Algorithm 3.2} && (33)
\end{aligned}
$$

$(k = 0, 1, 2, \dots)$.

Proof: We will only prove the relation (32); we could prove the other relation (33) similarly. For simplicity of notation, define

$$
\begin{aligned}
\overline{\mathcal{C}}_k &= \text{c.cone}\left(\left\{\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in C_k\right\}\right) \ (\forall k = 0, 1, 2, \dots),\\
\mathcal{K}_{k+1} &= N_+(\mathcal{K}_k, \mathcal{T}) \ (\forall k = 0, 1, 2, \dots).
\end{aligned}
$$

By definition, $N_+^0(\mathcal{K}_0, \mathcal{T}) = \mathcal{K}_0 = \overline{\mathcal{C}}_0$. Assuming that $\mathcal{K}_k \supseteq \overline{\mathcal{C}}_k$, we will prove that $\mathcal{K}_{k+1} \supseteq \overline{\mathcal{C}}_{k+1}$. If C_k is empty then so are $\overline{\mathcal{C}}_k$ and $\overline{\mathcal{C}}_{k+1}$; hence the desired relation follows. If C_k consists of a single point $\hat{\boldsymbol{x}}$, then we have either $\{\hat{\boldsymbol{x}}\} = F = C_k$ or $F = \emptyset$. We can derive that

$$
\begin{aligned}
\overline{\mathcal{C}}_k &= \overline{\mathcal{C}}_{k+1} = \mathcal{K}_I = \text{c.cone}\left(\left\{\begin{pmatrix} 1 \\ \hat{\boldsymbol{x}} \end{pmatrix}\right\}\right) \text{ if } \{\hat{\boldsymbol{x}}\} = F = C_k,\\
\overline{\mathcal{C}}_{k+1} &= \emptyset \text{ if } F = \emptyset.
\end{aligned}
$$

Thus, the desired relation follows in both cases. Now, we assume that C_k consists of at least two points. Then there is a $j \in \{1, 2, \dots, n\}$ such that

$$
\begin{aligned}
0 &< \max\{\boldsymbol{e}_j^T\boldsymbol{x} : \boldsymbol{x} \in C_k\} - \min\{\boldsymbol{e}_j^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}\\
&= \max\{\boldsymbol{e}_j^T\boldsymbol{x} : \boldsymbol{x} \in C_k\} + \max\{-\boldsymbol{e}_j^T\boldsymbol{x} : \boldsymbol{x} \in C_k\}\\
&= \alpha(C_k, \boldsymbol{e}_j) + \alpha(C_k, -\boldsymbol{e}_j), && (34)
\end{aligned}
$$

where $\boldsymbol{e}_j \in R^n$ denotes the unit vector with the jth coordinate 1. Since $C_k \subseteq [0, 1]^n$, we also see that

$$
-1 \leq \alpha(C_k, -\boldsymbol{e}_i) \leq 0 \leq \alpha(C_k, \boldsymbol{e}_i) \leq 1 \ (\forall i = 1, 2, \dots, n). \tag{35}
$$

On the other hand, we see by Remark 3.8 that

$$
C_{k+1} = \widehat{F}(C_0, \mathcal{P}_I \cup \mathcal{P}^2(C_k, D_I, \overline{D})).
$$

Hence each $\boldsymbol{y} = \lambda \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \overline{\mathcal{C}}_{k+1}$ for $\exists \lambda > 0$ is characterized as $\exists \boldsymbol{X} \in \mathcal{S}^n$;

$$
\begin{aligned}
&\boldsymbol{Y} \equiv \lambda \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}_+^{1+n},\\
&\gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}_I), && (36)\\
&\gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \mathcal{P}^2(C_k, D_I, \overline{D})). && (37)
\end{aligned}
$$

Recall that $\mathcal{P}_I$ consists of the quadratic functions in (30), so that we can rewrite inequality (36) as the equality $\boldsymbol{Y}\boldsymbol{e}_0 = \mathrm{Diag}(\boldsymbol{Y})$ imposed on each $\boldsymbol{Y} \in \mathcal{M}_+(\mathcal{K}_k, \mathcal{T})$, where $\boldsymbol{e}_0 \in R^{1+n}$ denotes the unit vector with the 0th coordinate 1. Letting

$$\begin{aligned} \mathcal{J}_1 &\equiv \text{c.cone}\left(\left\{\begin{pmatrix} \alpha(C_k, \boldsymbol{d}_1) \\ -\boldsymbol{d}_1 \end{pmatrix} : \boldsymbol{d}_1 \in D_I\right\}\right), \\ \mathcal{J}_2 &\equiv \text{c.cone}\left(\left\{\begin{pmatrix} \alpha(C_k, \boldsymbol{d}_2) \\ -\boldsymbol{d}_2 \end{pmatrix} : \boldsymbol{d}_2 \in \overline{D}\right\}\right), \end{aligned}$$

we can rewrite (37) as

$$0 \ \leq \ \boldsymbol{u}^T\boldsymbol{Y}\boldsymbol{v} \ (\forall \boldsymbol{u} \in \mathcal{J}_2 \ \text{ and } \forall \boldsymbol{v} \in \mathcal{J}_1).$$

We then see that

$$\begin{aligned} \mathcal{J}_1 &= \text{c.cone}\left(\left\{\begin{pmatrix} \alpha(C_k, \boldsymbol{e}_i) \\ -\boldsymbol{e}_i \end{pmatrix}, \begin{pmatrix} \alpha(C_k, -\boldsymbol{e}_i) \\ \boldsymbol{e}_i \end{pmatrix} \ (i = 1, 2, \ldots, n)\right\}\right) \\ &\supseteq \text{c.cone}\left(\left\{\begin{pmatrix} 0 \\ \boldsymbol{e}_i \end{pmatrix}, \begin{pmatrix} 1 \\ -\boldsymbol{e}_i \end{pmatrix} \ (i = 1, 2, \ldots, n)\right\}\right) \ \text{(by (34) and (35))} \\ &= \mathcal{T}^*, \\ \mathcal{J}_2 &= \overline{\mathcal{C}}_k^* \supseteq \mathcal{K}_k^* \ (\text{since } \mathcal{K}_k \supseteq \overline{\mathcal{C}}_k). \end{aligned}$$

Hence (37) implies

$$0 \ \leq \ \boldsymbol{u}^T\boldsymbol{Y}\boldsymbol{v} \ (\forall \boldsymbol{u} \in \mathcal{K}_k^* \ \text{ and } \forall \boldsymbol{v} \in \mathcal{T}^*).$$

Thus, we have shown that if $\boldsymbol{y} = \lambda \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \overline{\mathcal{C}}_{k+1}$ for $\exists \lambda > 0$ then $\exists \boldsymbol{X} \in \mathcal{S}^n$;

(i)' $\boldsymbol{Y} = \lambda \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}_+^{1+n}$,

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

(iii)' $\boldsymbol{u}^T\boldsymbol{Y}\boldsymbol{v} \geq 0$ $(\forall \boldsymbol{u} \in \mathcal{K}_k^*$ and $\forall \boldsymbol{v} \in \mathcal{T}^*)$.

These three conditions characterize each $\boldsymbol{y} = \lambda \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \mathcal{K}_{k+1} = N_+(\mathcal{K}_k, \mathcal{T})$. Therefore, we have shown that $\overline{\mathcal{C}}_{k+1} \subseteq \mathcal{K}_{k+1}$. ∎

Remark 7.4 If F does not lie on any of the hyperplanes $x_j = 0$ or $x_j = 1$ $(\forall j = 1, 2, \ldots, n)$ then we have that

$$\alpha(C_k, -\boldsymbol{e}_i) = 0 \ \text{ and } \alpha(C_k, \boldsymbol{e}_i) = 1 \ (\forall i = 1, 2, \ldots, n, \ \forall k = 1, 2, \ldots) \ \text{ and } \mathcal{J}_1 = \mathcal{T}^*.$$

In this case, we see that $\overline{\mathcal{C}}_{k+1} = \mathcal{K}_{k+1}$ and the equality holds in (32) and (33).

Remark 7.5 If $\alpha(C_k, \boldsymbol{e}_i) < 1$ and/or $\alpha(C_k, -\boldsymbol{e}_i) < 0$ in (35) then we know that

$$F \cap \{\boldsymbol{x} \in [0,1]^n : x_i = 1\} = \emptyset \text{ and/or } F \cap \{\boldsymbol{x} \in [0,1]^n : x_i = 0\} = \emptyset.$$

In this case, using Lemma 4.1, we can derive that

$$C_{k+1} \cap \{\boldsymbol{x} \in [0,1]^n : x_i = 1\} = \emptyset \text{ and/or } C_{k+1} \cap \{\boldsymbol{x} \in [0,1]^n : x_i = 0\} = \emptyset.$$

This fact is consistent with Lemma 1.3 and the proof of Theorem 1.6 in [11]. Actually, Lemma 1.3 of [11] combined with our Lemma 7.3 implies

$$C_{k+1} \subseteq \{\boldsymbol{x} \in [0,1]^n : x_i = 0\} \text{ and/or } C_{k+1} \subseteq \{\boldsymbol{x} \in [0,1]^n : x_i = 1\}.$$

Now, we are ready to prove that equality holds in (32) and (33).

Theorem 7.6 *Let $\{C_k \ (k = 0, 1, \dots)\}$ be the sequence generated by Algorithm 3.1 (the SSDP Relaxation Method) or Algorithm 3.2 (the SSILP Relaxation Method) applied to* 0-1 *QOP (31), using $D_0 = D_I \equiv \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n, -\boldsymbol{e}_1, -\boldsymbol{e}_2, \dots, -\boldsymbol{e}_n\}$ and $D_1 = \overline{D}$. Then*

$$N_+^k(\mathcal{K}_0, \mathcal{T}) = c.cone\left(\left\{\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in C_k\right\}\right) \text{ when we use Algorithm 3.1,} \tag{38}$$

$$N^k(\mathcal{K}_0, \mathcal{T}) = c.cone\left(\left\{\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in C_k\right\}\right) \text{ when we use Algorithm 3.2} \tag{39}$$

$(k = 0, 1, 2, \dots)$.

Proof: Consider the proof of Lemma 7.3. To prove the equations (38) and (39), we take the induction hypothesis to be the equality (which holds for $k = 0$) rather than the containment. So, it suffices to prove that $\boldsymbol{Y} \in \mathcal{M}(\mathcal{K}_k, \mathcal{T})$ implies

$$0 \leq \boldsymbol{u}^T \boldsymbol{Y} \boldsymbol{v} \ (\forall \boldsymbol{u} \in \mathcal{J}_2 \text{ and } \forall \boldsymbol{v} \in \mathcal{J}_1).$$

By Remark 7.4, the equations (38) and (39) hold if

$$\alpha(C_k, -\boldsymbol{e}_i) = 0 \text{ and } \alpha(C_k, \boldsymbol{e}_i) = 1 \ (\forall i = 1, 2, \dots, n, \ \forall k = 1, 2, \dots).$$

Now, suppose there exists $r \in \{1, 2, \dots, n\}$ such that $\alpha(C_k, \boldsymbol{e}_r) < 1$. Then by Remark 7.5, $\boldsymbol{Y} \in \mathcal{M}(\mathcal{K}_k, \mathcal{T})$ implies $y_{rr} = y_{r0} = 0$. Let $\boldsymbol{y}_j$ denote the jth column of $\boldsymbol{Y}$. $\boldsymbol{Y} \in \mathcal{M}(\mathcal{K}_k, \mathcal{T})$ also implies $\boldsymbol{y}_0 - \boldsymbol{y}_j \in \mathcal{K}_k$ $(j = 1, 2, \dots, n)$. Since $y_{r0} = 0$ and $\mathcal{K}_k$ lies in the nonnegative orthant, we must have $y_{rj} = 0$ $(j = 1, 2, \dots, n)$. Since $\boldsymbol{Y}$ is symmetric, we also have $y_{ir} = 0$ $(i = 1, 2, \dots, n)$. We proved that

$$\boldsymbol{Y}\begin{pmatrix} \alpha(C_k, \boldsymbol{e}_r) \\ -\boldsymbol{e}_r \end{pmatrix} \in \mathcal{K}_k \text{ for all such } r.$$

We have the same arguments for the indices r for which $\alpha(C_k, -\boldsymbol{e}_r) < 0$ (by the symmetry of the algorithms for 0-1 convex quadratic optimization problems, under the complementation of the variables). We proved that all the extreme rays $\boldsymbol{v}$ of $\mathcal{J}_1$ that are different than those of $\mathcal{T}^*$ also satisfy

$$0 \leq \boldsymbol{u}^T \boldsymbol{Y} \boldsymbol{v} \ (\forall \boldsymbol{u} \in \mathcal{K}_k^*)$$

for every $\boldsymbol{Y} \in \mathcal{M}(\mathcal{K}_k, \mathcal{T})$. By the induction hypothesis, we obtain $\mathcal{J}_2 = \mathcal{K}_k^*$ which completes the proof. ∎

7.2 0-1 Nonconvex Quadratic Optimization Problems.

We now deal with the general case where no compact convex subset C_0 of $[0,1]^n$, satisfying $F = C_0 \cap \{0,1\}^n$, is given in advance. Instead, we consider the 0-1 nonconvex quadratic optimization problem of the form

$$\left.\begin{array}{ll} \text{maximize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in F \equiv \{\boldsymbol{x} \in [0,1]^n : p(\boldsymbol{x}) \leq \ (\forall p(\cdot) \in \mathcal{P}_F)\}. \end{array}\right\} \tag{40}$$

Here $\mathcal{P}_I \subseteq \mathcal{P}_F \subseteq \mathcal{Q}$.

Lemma 7.7 *Let*

$$\begin{array}{rcl} C_0 & = & \widehat{F}([0,1]^n, \mathcal{P}_F \cup \mathcal{P}^2([0,1]^n, D_I, \overline{D})) \\ & = & \left\{ \boldsymbol{x} \in [0,1]^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+ \text{ and} \\ \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}_F \cup \mathcal{P}^2([0,1]^n, D_I, \overline{D})) \end{array} \right\} \end{array}$$

(an SDP relaxation of F).

Then $F = C_0 \cap \{0,1\}^n$.

Proof: It was shown in Lemma 6.2 of [10] that

$$F = \widehat{F}([0,1]^n, \mathcal{P}_F \cup \mathcal{P}^1([0,1]^n, D_I) \cap \{0,1\}^n.$$

Since

$$F \subseteq \widehat{F}([0,1]^n, \mathcal{P}_F \cup \mathcal{P}^2([0,1]^n, D_I, \overline{D})) \subseteq \widehat{F}([0,1]^n, \mathcal{P}_F \cup \mathcal{P}^1([0,1]^n, D_I)),$$

$F = C_0 \cap \{0,1\}^n$ follows. ∎

The conclusion of this lemma stays valid even if we replace $\widehat{F}([0,1]^n, \mathcal{P}_F \cup \mathcal{P}^2([0,1]^n, D_I, \overline{D}))$ by $\widehat{F}^L([0,1]^n, \mathcal{P}_F \cup \mathcal{P}^2([0,1]^n, D_I, \overline{D}))$ (a semi-infinite LP relaxation). (For the proof, we need Lemma 4.4.) As a consequence, we can reduce the 0-1 nonconvex quadratic optimization problem (40) to the 0-1 convex quadratic optimization problem (31) by applying one iteration of the SDP Relaxation Method or one iteration of the Semi-infinite LP relaxation Method. Therefore, in the case of 0-1 nonconvex quadratic optimization problem (40), our Algorithms 3.1 and 3.2 terminate in at most $(1+n)$ iterations if we choose the direction sets D_0 and D_1 as in Corollary 3.4. By using similar arguments to those in the proof of Theorem 1.6 of [11], we can also prove that whenever the representation $\mathcal{P}_F$ of F is finite, we can solve the weak separation problem for C_k in polynomial-time for any $k = O(1)$.

8 Concluding Discussions.

(A) Let $\kappa > 0$ and $\epsilon > 0$. By Theorem 3.5, there exists a $\delta > 0$ such that if we take a δ-net D_1 of $D(\boldsymbol{c}, \kappa)$ in Step 0 of Algorithm 3.1 (or Algorithm 3.2) then we obtain an upper bound on the optimal objective value of the QOP (1) within ϵ in a finite number of iterations. From the proof of Theorem 3.5 given in Section 6, we know that the value of such a δ depends on $\kappa > 0$, $\epsilon > 0$, D_1, the diameter η_0 of C_0 and the quadratic inequality representation $\mathcal{P}_F$ of the feasible region F of the QOP (1). However it seems difficult to characterize the largest δ which works for the purpose at hand. In practice, we might want to avoid guessing δ, but dynamically adjust it. We may start with a large value of δ and decrease it from one iteration to the next as we see appropriate. This idea was implemented in a practical variant of the SSILP Relaxation Method which Takeda, Dai, Fukuda and Kojima [23] recently presented together with some numerical results.

(B) Application to Quadratic Matrix Inequalities. Let $\boldsymbol{A}_{ij} \in \mathcal{S}^n$ $(i, j = 0, 1, 2, \ldots, n)$, and let C_0 be a nonempty compact convex subset of R^n. Consider the maximization problem with a linear objective function $\boldsymbol{c}^T \boldsymbol{x}$ over the feasible region

$$F = \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in C_0 : x_0 = 1, \ \sum_{i=0}^{n} \sum_{j=0}^{n} \boldsymbol{A}_{ij} x_i x_j \in \mathcal{S}^n_+\}.$$

This problem covers bilinear matrix inequalities with bounded variables. We can rewrite F in terms of quadratic inequalities:

$$F = \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in C_0 : x_0 = 1, \ -\boldsymbol{d}^T \left(\sum_{i=0}^{n} \sum_{j=0}^{n} \boldsymbol{A}_{ij} x_i x_j \right) \boldsymbol{d} \leq 0 \ (\forall \boldsymbol{d} \in \bar{D})\}.$$

Therefore, all the theoretical results (Theorem 3.3, Corollary 3.4, Theorem 3.5 and Corollary 3.6) hold for this problem. But, the quadratic inequality representation of F is not finite, so that Condition 2.1 which we have imposed on the quadratic inequality representation of F to implement Algorithms 3.1 and 3.2 does not hold. In this case, however, we can rewrite C_{k+1} to be computed in Step 3 as

$$C_{k+1} = \left\{ \boldsymbol{x} \in R^n : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n \text{ such that } \boldsymbol{Y} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \in \mathcal{S}^{1+n}_+, \\ \displaystyle\sum_{i=0}^{n} \sum_{j=0}^{n} \boldsymbol{A}_{ij} Y_{ij} \in \mathcal{S}^n_+ \text{ and} \\ \gamma + 2\boldsymbol{q}^T \boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0 \ (\forall qf(\cdot; \gamma, \boldsymbol{q}, \boldsymbol{Q}) \in \widetilde{\mathcal{P}}_0 \cup \mathcal{P}_k) \end{array} \right\}.$$

(See Section 8.1 of [10] to derive the representation above). Here $\widetilde{\mathcal{P}}_0$ denotes a convex quadratic inequality representation of C_0. If both $\widetilde{\mathcal{P}}_0$ and $\mathcal{P}_k$ are finite, then the maximization of a linear function $\boldsymbol{a}^T \boldsymbol{x}$ over C_{k+1} turns out to be a standard SDP having a finite number of linear matrix inequality constraints. Therefore, Algorithm 3.1 applied to this problem is implementable.

(C) Finally, our methods provide a canonical way of generating very large scale LP and SDP problems (essentially of arbitrary size) with certain structure. This provides yet more motivation for the study of computational techniques for the solution of such problems, as well as for the study of the mathematical properties of such families of LP and SDP problems.

Acknowledgment: The authors are grateful to Professor Yang Dai, Ms. Akiko Takeda and Mr. Mituhiro Fukuda for their careful reading of the draft of the paper and their valuable comments and suggestions, and also to anonymous referees for many suggestions to improve the presentation of the paper.

References

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

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

[3] Balas, E., Pulleyblank, W. R. (1983): The perfect matching subgraph polytopes of a bipartite graph. *Networks* **13**, 495–516

[4] Boyd, S., Ghaoui, L. E., Feron, E., Balakrishnan, V. (1994): *Linear Matrix Inequalities in System and Control Theory.* SIAM, Philadelphia, PA

[5] Cottle, R. W., Pang, J.-S., Stone, R. E. (1992): *The Linear Complementarity Problem.* Academic Press, New York, NY

[6] Fujie, T., Kojima, M. (1997): Semidefinite relaxation for nonconvex programs. *Journal of Global Optimization* **10**, 367–380

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

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

[9] Grötschel, M., Lovász, L., Schrijver, A. (1988): *Geometric algorithms and combinatorial optimization.* Springer, New York, NY

[10] Kojima, M., Tunçel, L. (2000): Cones of matrices and successive convex relaxations of nonconvex sets. *SIAM J. on Optimization* **10**, 750–778

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

[12] Horst, R., Tuy, H. (1992): *Global Optimization.* Second, Revised Edition, Springer-Verlag, Berlin

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

[14] Nesterov, Yu. E. (1998): Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software* **9**, 141–160

[15] Poljak, S., Rendl, F., Wolkowicz, H. (1995): A recipe for semidefinite relaxation for (0,1)-quadratic programming. *Journal of Global Optimization* **7**, 51–73

[16] Ramana, M. V. (1993): *An algorithmic analysis of multiquadratic and semidefinite programming problems.* PhD thesis, Johns Hopkins University, Baltimore, MD

[17] Safonov, M. G., Goh, K. G., Ly, J. H. (1994): Control system synthesis via bilinear matrix inequalities. In: M. Peshkin ed., *Proceedings of the 1994 American Control Conference*, Baltimore, MD

[18] Sherali, H. D., Adams, W. P. (1990): A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Mathematics* **3**, 411–430

[19] Sherali, H. D., Adams, W. P. (1994): A hierarchy of relaxations and convex hull characterizations for mixed-integer zero-one programming problems. *Discrete Applied Mathematics* **52**, 83–106

[20] Sherali, H. D., Tuncbilek, C. H. (1992): A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique. *Journal of Global Optimization* **2**, 101–112

[21] Sherali, H. D., Tuncbilek, C. H. (1995): A reformulation-convexification approach for solving nonconvex quadratic programming problems. *Journal of Global Optimization* **7**, 1–31

[22] Shor, N. Z. (1990): Dual quadratic estimates in polynomial and boolean programming. *Annals of Operations Research* **25**, 163–168

[23] Takeda, A., Dai, Y., Fukuda, M., Kojima, M. (1999): Towards the implementation of successive convex relaxation method for nonconvex quadratic optimization problem. Technical Report B-347, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan (to appear in *Approximation and Complexity in Numerical Optimization: Continuous and Discrete Problems* (P. M. Pardalos, Editor), Kluwer Academic Press).

[24] Vandenberghe, L., Boyd, S. (1996): Semidefinite programming. *SIAM Review* **38**, 49–95

[25] Ye, Y. (1999): Approximating quadratic programming with bound and quadratic constraints. *Mathematical Programming* **84**, 219–226