Research Reports on Mathematical and Computing Sciences
Series B : Operations Research

Department of Mathematical and Computing Sciences
Tokyo Institute of Technology
2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan

Complexity Analysis of Successive Convex Relaxation Methods for Nonconvex Sets

Masakazu Kojima
e-mail:*kojima@is.titech.ac.jp*

Akiko Takeda
e-mail:*takeda@is.titech.ac.jp*

April 1999, B-350
Revised July 1999

Abstract. This paper discusses computational complexity of conceptual successive convex relaxation methods proposed by Kojima and Tunçel for approximating a convex relaxation of a compact subset $F = \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \mathcal{P}_F)\}$ of the n-dimensional Euclidean space R^n. Here C_0 denotes a nonempty compact convex subset of R^n, and $\mathcal{P}_F$ a set of finitely or infinitely many quadratic functions. We evaluate the number of iterations which the successive convex relaxation methods require to attain a convex relaxation of F with a given accuracy ϵ, in terms of ϵ, the diameter of C_0, the diameter of F, and some other quantities characterizing the Lipschitz continuity, the nonlinearity and the nonconvexity of the set $\mathcal{P}_F$ of quadratic functions.

Keywords: Complexity, Nonconvex Quadratic Program, Semidefinite Programming, Global Optimization, SDP Relaxation, Convex Relaxation, Lift-and-Project Procedure.

1 Introduction.

In their paper [2], Kojima and Tunçel proposed a class of successive convex relaxation methods for a general nonconvex quadratic program:

$$\text{maximize } \boldsymbol{c}^T\boldsymbol{x} \text{ subject to } \boldsymbol{x} \in F, \tag{1}$$

where

$$\left.\begin{array}{lcl} \boldsymbol{c} & = & \text{a constant column vector in the } n\text{-dimensional Euclidean space } R^n, \\ F & = & F(C_0, \mathcal{P}_F) = \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \mathcal{P}_F)\} \\ & & \text{(we often write } F \text{ instead of } F(C_0, \mathcal{P}_F) \text{ for simplicity of notation),} \\ C_0 & = & \text{a nonempty compact convex subset of } R^n, \\ \mathcal{P}_F & = & \text{a set of finitely or infinitely many quadratic functions on } R^n. \end{array}\right\} \tag{2}$$

Their methods generate a sequence $\{C_k\}$ of compact convex sets satisfying

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

(b) if $F = \emptyset$, then $C_k = \emptyset$ for some finite number k (detecting infeasibility),

(c) otherwise, $\cap_{k=0}^{\infty} C_k = \text{c.hull}(F)$ (asymptotic convergence).

Here c.hull(F) denotes the convex hull of F.

In connection with the successive convex relaxation methods, Kojima [4] pointed out that a wide class of nonlinear programs can be reduced to a nonconvex quadratic program of the form (1). More generally, it is known that any closed subset $G \subset R^n$ can be represented as

$$G = \{\boldsymbol{x} \in R^n : g(\boldsymbol{x}) - \|\boldsymbol{x}\|^2 \leq 0\}$$

using some convex function $g(\cdot) : R^n \to R$. See, for example, Corollary 3.5 of [10]. Thus, given any closed subset G of R^n and any compact convex subset C_0 of R^n, we can rewrite maximization of a linear function $\boldsymbol{c}^T\boldsymbol{x}$ over a compact set $G \cap C_0$ as

$$\begin{array}{ll} \text{maximize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & (\boldsymbol{x}, t) \in H = \{(\boldsymbol{x}, t) \in R^{n+1} : \boldsymbol{x} \in C_0,\ g(\boldsymbol{x}) - t \leq 0,\ 0 \leq t \leq \bar{t}\}, \\ & t - \boldsymbol{x}^T\boldsymbol{x} \leq 0, \end{array}$$

where $g(\cdot)$ is a convex function on R^n and $\bar{t} \geq 0$ is a sufficiently large number such that $g(\boldsymbol{x}) \leq \bar{t}$ for $\forall \boldsymbol{x} \in C_0$. Then, H turns out to be compact and convex, and the resultant problem is a special case of the general nonconvex quadratic program (1). Although this construction is not implementable because an explicit algebraic representation of such a convex function $g(\cdot)$ is usually impossible, it certainly shows theoretical potential of the successive convex relaxation methods for quite general nonlinear programs.

The successive convex relaxation methods are extensions of the lift-and-project procedure, which was proposed by Lovász and Schrijver [5] for 0-1 integer programs, to a general quadratic program of the form (1). At each iteration of the methods, we first generate a set $\mathcal{P}_k = \{p(\cdot)\}$ of finitely or infinitely many quadratic functions such that each $p(\boldsymbol{x}) \leq 0$ forms a valid inequality for the kth iterate C_k. Since C_k was chosen to include F in the previous iteration, each $p(\boldsymbol{x}) \leq 0$ serves as a (redundant) valid inequality for F; hence F is represented as

$$F = \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0 \ (\forall p(\cdot) \in \mathcal{P}_F \cup \mathcal{P}_k)\}.$$

We then apply the SDP (semidefinite programming) relaxation or the SILP (semi-infinite linear programming relaxation) to the set F with the above representation in terms of the set $\mathcal{P}_F \cup \mathcal{P}_k$ of quadratic functions to generate the next iterate C_{k+1}. (The latter relaxation is also called the reformulation-convexification approach in the literature [7]). See also [1, 8].

The successive convex relaxation methods outlined above are conceptual in the sense that we need to solve "*infinitely many* semidefinite programs (or linear programs) having *infinitely many* linear inequality constraints" at each iteration. In their succeeding paper [3], Kojima and Tunçel presented two types of techniques which discretize and localize "*infinitely many* semidefinite programs (or linear programs) having *infinitely many* linear inequality constraints" to implement their conceptual methods under a certain assumption on a finite representation of the feasible region F. They established that, for any given $\epsilon > 0$, we can discretize and localize the conceptual methods to generate a sequence $\{C_k\}$ of compact convex sets satisfying the features (a), (b) above and

(c)' $\zeta^* \leq \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in C_k\} \leq \zeta^* + \epsilon$ for some finite k if $F \neq \emptyset$, where $\zeta^* = \max\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in F\}$.

Each iteration of these discretized-localized versions of successive convex relaxation methods requires to solve finitely many semidefinite programs (or linear programs) having finitely many linear inequality constraints, so that these versions are implementable on computer. However, they are still impractical because as a higher accuracy ϵ is required, not only the number of the semidefinite programs (or linear programs) to be solved at each iteration but also their sizes explode quite rapidly.

More recently, Takeda, Dai, Fukuda and Kojima [9] presented practical versions of successive convex relaxation methods by further slimming down the discretized-localized versions to overcome the rapid explosion in the the number of the semidefinite programs (or linear programs) to be solved at each iteration and their sizes. Although these versions no more enjoy the feature (c)', numerical results reported in the paper [9] look promising.

This paper investigates computational complexity of the conceptual successive convex relaxation methods given in the paper [2]. When they are applied to 0-1 integer programs, they work as the Lovász and Schrijver lift-and-project procedure [5], and they terminate in n iterations, where n denotes the number of 0-1 variables. (See Section 6 of [2] and Section 7 of [3].) In the general case where $\mathcal{P}_F$ consists of arbitrary quadratic functions, however, the convergence of $\{C_k\}$ to c.hull(F) is slower than linear in the worst case. (See an example in Section 8.3 of [2].) In this paper, we bound the number k of iterations required to generate an approximation of c.hull(F) with a given "accuracy." To begin with, we need to clarify the following issues.

- Input data of the conceptual successive convex relaxation methods.

- Output of the methods and its quality or "accuracy."
- What are assumed to be possible in the computation?

To discuss these issues in more detail, we need some notation.

$$\left.\begin{array}{rcl} \mathcal{S}^n & = & \text{the set of } n \times n \text{ symmetric matrices.} \\ qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) & = & \text{the quadratic function having the constant term } \gamma, \\ & & \text{the linear term } 2\boldsymbol{q}^T\boldsymbol{x}, \text{ and the quadratic term } \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}; \\ & & qf(\boldsymbol{x};\gamma,\boldsymbol{q},\boldsymbol{Q}) = \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \ \text{ for } \forall \boldsymbol{x} \in R^n, \\ & & \text{where } \gamma \in R, \ \boldsymbol{q} \in R^n, \text{ and } \boldsymbol{Q} \in \mathcal{S}^n. \end{array}\right\} \quad (3)$$

Our input data are a nonempty compact convex subset C_0 of R^n and a set $\mathcal{P}_F$ of finitely or infinitely many quadratic functions. We do not care about how we represent the compact convex set C_0; it may be represented in terms of finitely or infinitely many linear inequalities, nonlinear convex inequalities and/or linear matrix inequalities. Although it seems nonsense to try to define the size of such input data, we extract some quantity and quality from the input data. The diameter $\mathrm{diam}(C_0)$ of C_0 and the diameter $\mathrm{diam}(F)$ of F are relevant on our complexity analysis since C_0 serves as an initial approximation of c.hull(F) which we want to compute. Concerning quality or difficulty of the input data, we introduce

$$\left.\begin{array}{rcl} \nu_{\mathrm{lip}} & = & \nu_{\mathrm{lip}}(\mathcal{P}_F, C_0) = \sup\left\{ \dfrac{|p(\boldsymbol{x}) - p(\boldsymbol{y})|}{\|\boldsymbol{x} - \boldsymbol{y}\|} : \boldsymbol{x}, \ \boldsymbol{y} \in C_0, \boldsymbol{x} \neq \boldsymbol{y}, \ p(\cdot) \in \mathcal{P}_F \right\} \\ & & \text{(a common Lipschitz constant for all } p(\cdot) \in \mathcal{P}_F \text{ on } C_0), \\ \nu_{\mathrm{nc}} & = & \nu_{\mathrm{nc}}(\mathcal{P}_F) = \sup\left\{ -\inf\{\lambda_{\min}(\boldsymbol{Q}) : qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}_F\}, \ 0 \right\}, \\ \nu_{\mathrm{nl}} & = & \nu_{\mathrm{nl}}(\mathcal{P}_F) = \sup\{\sum_{i=1}^n \sum_{j=1}^n |Q_{ij}| : qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}_F\}. \end{array}\right\} \quad (4)$$

Here $\lambda_{\min}(\boldsymbol{Q})$ denotes the minimum eigenvalue of $\boldsymbol{Q} \in \mathcal{S}^n$. Note that all $\nu_{\mathrm{lip}}(\mathcal{P}_F, C_0)$, $\nu_{\mathrm{nc}}(\mathcal{P}_F)$ and $\nu_{\mathrm{nl}}(\mathcal{P}_F)$ take either a finite nonnegative value or $+\infty$. We will assume that they are finite. If $\mathcal{P}_F$ consists of a finite number of quadratic functions, this assumption is satisfied. We may regard $\nu_{\mathrm{nc}}(\mathcal{P}_F)$ as a nonconvexity measure of the set $\mathcal{P}_F$ of quadratic functions; all quadratic functions in $\mathcal{P}_F$ are convex if and only if $\nu_{\mathrm{nc}}(\mathcal{P}_F) = 0$, and $\mathcal{P}_F$ involves more nonconvexity as $\nu_{\mathrm{nc}}(\mathcal{P}_F) \geq 0$ gets larger. On the other hand, we may regard $\nu_{\mathrm{nl}}(\mathcal{P}_F)$ as a nonlinearity measure of the set $\mathcal{P}_F$ of quadratic functions; all functions in $\mathcal{P}_F$ are linear if and only if $\nu_{\mathrm{nl}}(\mathcal{P}_F) = 0$, and $\mathcal{P}_F$ involves more nonlinearity as $\nu_{\mathrm{nl}}(\mathcal{P}_F) \geq 0$ gets larger. $\nu_{\mathrm{lip}}(\mathcal{P}_F, C_0)$, $\nu_{\mathrm{nc}}(\mathcal{P}_F)$ and $\nu_{\mathrm{nl}}(\mathcal{P}_F)$ directly affect the upper bounds that we will derive for the number k of iterations required to generate an approximation of c.hull(F) with a given certain accuracy.

Our output is a compact set $C_k \subset R^n$, which is an approximation of c.hull(F). Again we don't care about its representation and size. In order to evaluate the quality or accuracy of the approximation, we introduce the notation

$$F(\epsilon) = F(\epsilon, C_0, \mathcal{P}_F) = \{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq \epsilon \ (\forall p(\cdot) \in \mathcal{P}_F)\} \quad \text{for } \forall \epsilon \geq 0. \quad (5)$$

By definition, $F = F(0) \subset F(\epsilon_1) \subset F(\epsilon_2) \subset C_0$ whenever $0 \leq \epsilon_1 \leq \epsilon_2$. We say that a compact convex subset C of C_0 is *an ϵ-convex-relaxation* of F if it satisfies

$$F \subset C \subset \text{c.hull}(F(\epsilon)).$$

We set up an ϵ-convex-relaxation of F as our goal of the successive convex relaxation methods.

Now we discuss what are assumed to be possible in the computation. First we assume precise real arithmetic operations. At each iteration, we need to choose a set $\mathcal{P}_k$ of finitely or infinitely many quadratic functions that induce valid inequalities for C_k before performing the SDP or SILP relaxation. Secondly we assume that such a $\mathcal{P}_k$ is available. Kojima and Tunçel [2] proposed and studied several candidates for $\mathcal{P}_k$. In this paper, we focus our attention on the following three models of successive convex relaxation methods.

- Spherical-SDP Model: We take $\mathcal{P}_k$ to be a set of spherical quadratic functions, and we perform the SDP relaxation (Section 3).
- Rank-2-SDP Model: We take $\mathcal{P}_k$ to be a set of rank-2 quadratic functions, and we perform the SDP relaxation (Section 4).
- Rank-2-SILP Model: We take $\mathcal{P}_k$ to be a set of rank-2 quadratic functions, and we perform the semi-infinite LP relaxation (Section 4).

In each model, we assume that a set of infinitely many quadratic functions chosen for $\mathcal{P}_k$ is available.

Our complexity analysis of the Spherical-SDP Model is much simpler than that of the latter two models, and the former analysis helps an easier understanding of the latter two models, which are of more practical interest. The latter two models correspond to extensions of the Lovász-Schrijver lift-and-project procedure with the use of the SDP relaxation and the use of the SILP relaxation, respectively. See Section 6 of [2] and Section 7 of [3]. The Rank-2-SDP and the Rank-2-SILP Models themselves are not implementable yet. Implementable versions of successive convex relaxation methods given in the paper [3] correspond to discretization and localization of these two models. See also [9].

We summarize our main results: For each of the models mentioned above, given an arbitrary positive number ϵ, we bound the number k of iterations which the successive convex relaxation methods require to attain an ϵ-convex-relaxation of $F = F(C_0, \mathcal{P}_F)$ in terms of the quantities $1/\epsilon$, $\mathrm{diam}(C_0)$, $1/\mathrm{diam}(F)$, $\nu_{\mathrm{lip}}(\mathcal{P}_F, C_0)$, $\nu_{\mathrm{nc}}(\mathcal{P}_F)$, $\nu_{\mathrm{nl}}(\mathcal{P}_F)$. The upper bound derived there itself might not be so significant, and might be improved by a more sophisticated analysis. It should be emphasized, however, that the upper bound is polynomial in these quantities, and that this paper provides a new way of analyzing the computational complexity of the successive convex relaxation methods for general nonlinear programs.

2 Preliminaries.

After introducing notation and symbols which we will use throughout the paper, we describe the SSDP and SSILP algorithms in Section 2.2. In Section 2.3, we present another accuracy measure of convex relaxation which will be utilized in establishing our main results.

2.1 Notation and Symbols.

Let

$$
\begin{aligned}
\mathcal{S}^n &= \text{the set of } n \times n \text{ symmetric matrices,} \\
\mathcal{S}^n_+ &= \text{the set of } n \times n \text{ symmetric positive semidefinite matrices,} \\
\boldsymbol{A} \bullet \boldsymbol{B} &= \text{trace of } \boldsymbol{A}^T\boldsymbol{B} = \sum_{i=1}^n \sum_{j=1}^n A_{ij}B_{ij}, \forall \boldsymbol{A},\ \boldsymbol{B} \in \mathcal{S}^n, \\
\mathcal{Q} &= \text{the set of quadratic functions on } R^n \\
&= \{qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) : \gamma \in R,\ \boldsymbol{q} \in R^n,\ \boldsymbol{Q} \in \mathcal{S}^n\} \\
&\quad \text{(see (3) for the definition of } qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q})), \\
\mathcal{Q}_+ &= \text{the set of convex quadratic functions on } R^n \\
&= \{qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) : \gamma \in R,\ \boldsymbol{q} \in R^n,\ \boldsymbol{Q} \in \mathcal{S}^n_+\}, \\
\mathcal{L} &= \text{the set of linear functions on } R^n \\
&= \{qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{O}) : \gamma \in R,\ \boldsymbol{q} \in R^n\}, \\
\text{c.cone}(\mathcal{P}) &= \text{the convex cone generated by } \mathcal{P} \subset \mathcal{Q} \\
&= \{\sum_{i=1}^{\ell} \lambda_i p_i(\cdot) : \lambda_i \geq 0,\ p_i(\cdot) \in \mathcal{P}\ (i = 1,2,\ldots,\ell),\ \ell \geq 0\}, \\
\text{c.hull}(G) &= \text{the convex hull of } G \subset R^n, \\
B(\boldsymbol{\xi},\rho) &= \{\boldsymbol{x} \in R^n : \|\boldsymbol{x} - \boldsymbol{\xi}\| \leq \rho\} \text{ (the (}n\text{-dimensional closed) ball} \\
&\quad \text{with a center } \boldsymbol{\xi} \in R^n \text{ and a radius } \rho > 0), \\
\rho(\boldsymbol{\xi},G) &= \sup\{\|\boldsymbol{x} - \boldsymbol{\xi}\| : \boldsymbol{x} \in G\} \text{ (the radius of the minimum ball} \\
&\quad \text{with a center } \boldsymbol{\xi} \in R^n \text{ that contains } G \subset R^n), \\
\text{diam}(G) &= \text{the diameter of } G \subset R^n\ = \sup\{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{x},\ \boldsymbol{y} \in G\} \text{ for } \forall G \subset R^n.
\end{aligned}
$$

2.2 SSDP and SSILP Relaxation Methods.

Let $C \subset R^n$ and $p(\cdot) \in \mathcal{Q}$. If $p(\boldsymbol{x}) \leq 0$ for $\forall \boldsymbol{x} \in C$, we say that $p(\boldsymbol{x}) \leq 0$ is *a quadratic valid inequality for* C, and that $p(\cdot)$ *induces a quadratic valid inequality for* C.

Let $\mathcal{P}$ be a nonempty subset of quadratic functions, *i.e.*, $\emptyset \neq \mathcal{P} \subset \mathcal{Q}$. (In the algorithms below, we will take $\mathcal{P}$ to be the union of $\mathcal{P}_F$ and a $\mathcal{P}_k$ of quadratic functions which induce quadratic valid inequalities for the kth iterate C_k). We use the notation $\widehat{F}(C_0,\mathcal{P})$ for the SDP relaxation, and the notation $\widehat{F}^L(C_0,\mathcal{P})$ for the SILP relaxation applied to the set $\{\boldsymbol{x} \in C_0 : p(\boldsymbol{x}) \leq 0, p(\cdot) \in \mathcal{P}\}$;

$$
\begin{aligned}
\widehat{F}(C_0,\mathcal{P}) &= \left\{ \boldsymbol{x} \in C_0 : \begin{array}{l} \exists \boldsymbol{X} \in \mathcal{S}^n; \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \text{ is positive semidefinite,} \\ \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0\ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}) \end{array} \right\}, \\
\widehat{F}^L(C_0,\mathcal{P}) &= \left\{ \boldsymbol{x} \in C_0 :\ \exists \boldsymbol{X} \in \mathcal{S}^n; \gamma + 2\boldsymbol{q}^T\boldsymbol{x} + \boldsymbol{Q} \bullet \boldsymbol{X} \leq 0\ (\forall qf(\cdot;\gamma,\boldsymbol{q},\boldsymbol{Q}) \in \mathcal{P}) \right\}.
\end{aligned}
$$

The lemma below provides a fundamental characterization of $\widehat{F}(C_0,\mathcal{P})$ and $\widehat{F}^L(C_0,\mathcal{P})$, which played an essential role in the global convergence analysis of the successive convex relaxation methods in the papers [2, 3]. We will use the lemma in Sections 3 and 4.

Lemma 2.1. *Let* $\emptyset \neq \mathcal{P} \subset \mathcal{Q}$.

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

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

Proof: See Theorem 4.2 and Corollary 4.3 of [2]. ∎

Algorithms 2.2 and 2.3 below are slight variants of the SSDP relaxation method given in Section 3 of [2] and the SSILP relaxation method given in Section 7 of [2], respectively. See also Section 3 of [3].

Algorithm 2.2. (SSDP relaxation method)

Step 0: Let $k = 0$.

Step 1: Choose a set $\mathcal{P}_k \subset \mathcal{Q}$ that induces quadratic valid inequalities for C_k.

Step 2: Let $C_{k+1} = \widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$.

Step 3: Let $k = k+1$ and go to Step 1.

Algorithm 2.3. (SSILP relaxation method)

Step 0: Let $k = 0$.

Step 1: Choose a set $\mathcal{P}_k \subset \mathcal{Q}$ that induces quadratic valid inequalities for C_k.

Step 2: Let $C_{k+1} = \widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$.

Step 3: Let $k = k+1$ and go to Step 1.

In Section 3 where we discuss the complexity of the Spherical-SDP Model, we will take $\mathcal{P}_k$ to be a set of spherical functions, while in Section 4 where we discuss the complexity of the Rank-2-SDP and the Rank-2-SILP Models, we take $\mathcal{P}_k$ to be a set of rank-2 quadratic functions. By definition, $\widehat{F}(C_0, \mathcal{P}_F \cup \mathcal{P}_k) \subset \widehat{F}^L(C_0, \mathcal{P}_F \cup \mathcal{P}_k)$. This inclusion relation will make it possible for us to apply our complexity analysis to the Rank-2-SILP Model, and simultaneously to the Rank-2-SDP Model in Section 4.

2.3 Accuracy Measures of Convex Relaxation.

In Section 1, we have introduced an ϵ-convex-relaxation of $F = F(C_0, \mathcal{P}_F)$ using c.hull$(F(\epsilon))$ $=$ c.hull$(F(\epsilon, C_0, \mathcal{P}_F))$ to measure the quality or accuracy of an approximation C_k of c.hull(F) which we want to compute. In our complexity analysis of Algorithms 2.2 and 2.3, this notion is not easy to manipulate directly. In this section, we introduce another notion (ψ, Ξ)-convex-relaxation of F which is easier to manipulate, and relate it to the ϵ-convex-relaxation of F.

Let $\psi > 0$, and let $\Xi \subset R^n$ be a nonempty compact convex set. We say that a compact convex subset C of C_0 is *an (ψ, Ξ)-convex-relaxation* of $F(C_0, \mathcal{P}_F)$ if

$$F \subset C \subset \text{c.relax}(F(\psi), \Xi) \quad = \quad \bigcap_{\boldsymbol{\xi} \in \Xi} B(\boldsymbol{\xi}, \rho(\boldsymbol{\xi}, F(\psi))).$$

The definition of c.relax$(F(\psi), \Xi)$ is quite similar to that of (ψ, ρ)-approximation of c.hull(F) given in Section 5 of [3]. Note that $B(\boldsymbol{\xi}, \rho(\boldsymbol{\xi}, F(\psi)))$ in the definition of c.relax$(F(\psi), \Xi)$ corresponds to the minimum ball with the center $\boldsymbol{\xi}$ that contains $F(\psi)$. It is easily verified that given an arbitrary open convex set U containing F, if $\psi > 0$ is sufficiently small and if Ξ is a sufficiently large ball with its center in C_0, then $F \subset$ c.relax$(F(\psi), \Xi) \subset U$. See Lemma 5.1 of [3] and its proof. By definition, we also see that

$$\text{c.relax}(F(\psi), \Xi_1) \quad \supset \quad \text{c.relax}(F(\psi), \Xi_2) \quad \text{if } \Xi_1 \subset \Xi_2.$$

We assume in the remainder of this section that there exists a finite common Lipschitz constant for all $p(\cdot) \in \mathcal{P}_F$;

$$\nu_{\text{lip}} \quad = \quad \nu_{\text{lip}}(\mathcal{P}_F, C_0) = \sup \left\{ \frac{|p(\boldsymbol{x}) - p(\boldsymbol{y})|}{\|\boldsymbol{x} - \boldsymbol{y}\|} : \boldsymbol{x},\ \boldsymbol{y} \in C_0, \boldsymbol{x} \neq \boldsymbol{y},\ p(\cdot) \in \mathcal{P}_F \right\} < \infty.$$

If $\mathcal{P}_F$ consists of a finite number of quadratic functions, then this assumption is satisfied.

Lemma 2.4. *Let $\boldsymbol{\xi}_0 \in C_0$ and $\epsilon > 0$. Choose positive numbers ψ, τ and a compact convex set Ξ such that*

$$\psi = \frac{\epsilon}{2}, \ \tau \geq \frac{\text{diam}(C_0)^2}{2\delta} + \frac{\delta}{2} \quad \textit{and } \Xi = B(\boldsymbol{\xi}_0, 3\tau/2), \tag{6}$$

where

$$0 \ < \ \delta \leq \min \left\{ \frac{\epsilon}{2\ \nu_{\text{lip}}}, \ \frac{\text{diam}(C_0)}{4} \right\}. \tag{7}$$

Then c.relax$(F(\psi), \Xi) \subset$ c.hull$(F(\epsilon))$.

Proof: If $F(\psi)$ is empty, then the desired inclusion relation holds with $\emptyset \subset$ c.hull$(F(\epsilon))$. So we will deal with the case that $F(\psi)$ is not empty.

(i) Define a δ-neighborhood G of $F(\psi)$ within C_0 such that

$$\begin{aligned} G &= \{\boldsymbol{x} \in C_0 : B(\boldsymbol{x}, \delta) \cap F(\psi) \neq \emptyset\} \\ &= \{\boldsymbol{x} \in C_0 : \|\boldsymbol{x} - \boldsymbol{y}\| \leq \delta \text{ for } \exists \boldsymbol{y} \in F(\psi)\}. \end{aligned}$$

We show that $\text{c.hull}(G) \subset \text{c.hull}(F(\epsilon))$. Suppose that $\boldsymbol{x} \in G$. Then $\boldsymbol{x} \in C_0$ and there exists a $\boldsymbol{y} \in F(\psi)$ such that $\|\boldsymbol{x} - \boldsymbol{y}\| \leq \delta$; hence

$$\begin{aligned} p(\boldsymbol{x}) &\leq p(\boldsymbol{y}) + |p(\boldsymbol{x}) - p(\boldsymbol{y})| \\ &\leq \psi + \nu_{\text{lip}} \|\boldsymbol{x} - \boldsymbol{y}\| \\ &\quad \text{(by } \boldsymbol{y} \in F(\psi) \text{ and the definition of } \nu_{\text{lip}}) \\ &\leq \epsilon/2 + \nu_{\text{lip}}\, \delta \quad (\text{since } \psi = \epsilon/2 \text{ and } \|\boldsymbol{x} - \boldsymbol{y}\| \leq \delta) \\ &\leq \epsilon \quad (\text{by (7)}). \end{aligned}$$

Thus we have shown that $G \subset F(\epsilon)$, which implies that $\text{c.hull}(G) \subset \text{c.hull}(F(\epsilon))$.

(ii) In view of (i), it suffices to prove that

$$\text{c.relax}(F(\psi), \Xi) \subset \text{c.hull}(G).$$

Assuming that $\bar{\boldsymbol{x}} \notin \text{c.hull}(G)$, we show that $\bar{\boldsymbol{x}} \notin \text{c.relax}(F(\psi), \Xi)$. We will construct a ball $B(\bar{\boldsymbol{\xi}}, \tau) \subset R^n$ such that

$$\bar{\boldsymbol{\xi}} \in \Xi, \ \bar{\boldsymbol{x}} \notin B(\bar{\boldsymbol{\xi}}, \tau) \ \text{ and } F(\psi) \subset B(\bar{\boldsymbol{\xi}}, \tau). \tag{8}$$

Then $\bar{\boldsymbol{x}} \notin \text{c.relax}(F(\psi), \Xi)$ follows from the definition of $\text{c.relax}(F(\psi), \Xi)$. Let $\bar{\boldsymbol{y}} \in \text{c.hull}(F(\psi)) \subset C_0$ be the point that minimize the distance $\|\bar{\boldsymbol{x}} - \boldsymbol{y}\|$ over $\boldsymbol{y} \in \text{c.hull}(F(\psi))$. Let $\bar{\delta} = \|\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}\|$ and $\bar{\boldsymbol{d}} = (\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}})/\bar{\delta}$. Then $\{\boldsymbol{y} \in R^n : \bar{\boldsymbol{d}}^T \boldsymbol{y} = \bar{\boldsymbol{d}}^T \bar{\boldsymbol{y}}\}$ forms a supporting hyperplane of $\text{c.hull}(F(\psi))$ such that

$$\bar{\boldsymbol{d}}^T \boldsymbol{y} \leq \bar{\boldsymbol{d}}^T \bar{\boldsymbol{y}} \ \text{ for } \forall \boldsymbol{y} \in \text{c.hull}(F(\psi)). \tag{9}$$

We will show that $\bar{\delta} > \delta$. Assume on the contrary that $\bar{\delta} \leq \delta$. Since $\bar{\boldsymbol{y}}$ lies in $\text{c.hull}(F(\psi))$, there are $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k \in F(\psi)$ $(1 \leq \exists k \leq n+1)$ such that

$$\bar{\boldsymbol{y}} \in \text{c.hull}\left(\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\}\right).$$

Let

$$\boldsymbol{x}_j = \boldsymbol{y}_j + \bar{\delta}\bar{\boldsymbol{d}} \ (j = 1, 2, \ldots, k).$$

Then we see that

$$\begin{aligned} \|\boldsymbol{x}_j - \boldsymbol{y}_j\| &= \|\bar{\delta}\bar{\boldsymbol{d}}\| \leq \delta \ \text{ and } \boldsymbol{y}_j \in F(\psi) \ (\text{hence } \boldsymbol{x}_j \in G) \text{ for } \forall j = 1, 2, \ldots, k, \\ \bar{\boldsymbol{x}} &= \bar{\boldsymbol{y}} + \bar{\delta}\bar{\boldsymbol{d}} \\ &\in \text{c.hull}\left(\{\boldsymbol{y}_1 + \bar{\delta}\bar{\boldsymbol{d}}, \boldsymbol{y}_2 + \bar{\delta}\bar{\boldsymbol{d}}, \ldots, \boldsymbol{y}_k + \bar{\delta}\bar{\boldsymbol{d}}\}\right) \\ &= \text{c.hull}\left(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}\right). \end{aligned}$$

This implies that $\bar{\boldsymbol{x}} \in \text{c.hull}(G)$, which contradicts to the hypothesis that $\bar{\boldsymbol{x}} \notin \text{c.hull}(G)$. Thus we have shown that $\bar{\delta} > \delta$.

(iii) Now defining $\bar{\boldsymbol{\xi}} = \bar{\boldsymbol{x}} - (\tau + \bar{\delta} - \delta)\bar{\boldsymbol{d}}$, we show that $\bar{\boldsymbol{\xi}}$ and the ball $B(\bar{\boldsymbol{\xi}}, \tau)$ satisfies (8). From the definition of $\bar{\boldsymbol{\xi}}$ above and $\bar{\delta} > \delta$, we first observe that

$$\begin{array}{rcl} \bar{\boldsymbol{x}} & \notin & B(\bar{\boldsymbol{\xi}}, \tau) \;\; \text{(i.e., the 2nd relation of (8))}, \\ \bar{\boldsymbol{y}} & = & \bar{\boldsymbol{x}} - \bar{\delta}\bar{\boldsymbol{d}} = \bar{\boldsymbol{\xi}} + (\tau + \bar{\delta} - \delta)\bar{\boldsymbol{d}} - \bar{\delta}\bar{\boldsymbol{d}} = \bar{\boldsymbol{\xi}} + (\tau - \delta)\bar{\boldsymbol{d}}. \end{array} \tag{10}$$

Next we will show that the third relation of (8), *i.e.*, c.hull$(F(\psi)) \subset B(\bar{\boldsymbol{\xi}}, \tau)$. Suppose that $\boldsymbol{y} \in$ c.hull$(F(\psi))$. Then

$$\begin{array}{rcl} \boldsymbol{y} - \bar{\boldsymbol{\xi}} & = & (\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{\xi}}) + \left(\boldsymbol{I} - \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T\right)(\boldsymbol{y} - \bar{\boldsymbol{\xi}}), \\ \boldsymbol{y} - \bar{\boldsymbol{y}} & = & (\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{y}}) + \left(\boldsymbol{I} - \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T\right)(\boldsymbol{y} - \bar{\boldsymbol{y}}) \\ & = & (\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{y}}) + \left(\boldsymbol{I} - \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T\right)(\boldsymbol{y} - \bar{\boldsymbol{\xi}} - (\tau - \delta)\bar{\boldsymbol{d}}) \;\; \text{(by (10))} \\ & = & (\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{y}}) + \left(\boldsymbol{I} - \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T\right)(\boldsymbol{y} - \bar{\boldsymbol{\xi}}). \end{array}$$

Note that $\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T$ and $\left(\boldsymbol{I} - \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T\right)$ are orthogonal projection matrices onto the one dimensional line $\{\lambda\bar{\boldsymbol{d}} : \lambda \in R\}$ in R^n and its orthogonal complement, respectively. Therefore, from the above equations, we see that $\|\boldsymbol{y} - \bar{\boldsymbol{\xi}}\|^2 = \|(\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{\xi}})\|^2 + \|\left(\boldsymbol{I} - \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T\right)(\boldsymbol{y} - \bar{\boldsymbol{\xi}})\|^2$ and $\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|^2 = \|(\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{y}})\|^2 + \|\left(\boldsymbol{I} - \bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T\right)(\boldsymbol{y} - \bar{\boldsymbol{\xi}})\|^2$, which derive that

$$\begin{array}{rcl} \|\boldsymbol{y} - \bar{\boldsymbol{\xi}}\|^2 & = & \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|^2 + \|(\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{\xi}})\|^2 - \|(\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{y}})\|^2 \\ & \leq & \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|^2 + \|(\bar{\boldsymbol{d}}\bar{\boldsymbol{d}}^T)(\boldsymbol{y} - \bar{\boldsymbol{\xi}})\|^2 \\ & \leq & \mathrm{diam}(C_0)^2 + \left(\bar{\boldsymbol{d}}^T(\boldsymbol{y} - \bar{\boldsymbol{\xi}})\right)^2 \;\; \text{(since both } \boldsymbol{y},\ \bar{\boldsymbol{y}} \in (F(\psi)) \subset C_0). \end{array}$$

Furthermore

$$\begin{array}{rcl} \bar{\boldsymbol{d}}^T(\bar{\boldsymbol{y}} - \bar{\boldsymbol{\xi}}) & \geq & \bar{\boldsymbol{d}}^T(\boldsymbol{y} - \bar{\boldsymbol{\xi}}) \;\; \text{(by (9))} \\ & = & \bar{\boldsymbol{d}}^T(\bar{\boldsymbol{y}} - \bar{\boldsymbol{\xi}} + \boldsymbol{y} - \bar{\boldsymbol{y}}) \\ & = & \bar{\boldsymbol{d}}^T\left((\tau - \delta)\bar{\boldsymbol{d}} + \boldsymbol{y} - \bar{\boldsymbol{y}}\right) \;\; \text{(by (10))} \\ & \geq & \tau - \delta - \|\boldsymbol{y} - \bar{\boldsymbol{y}}\| \;\; \text{(since } \|\bar{\boldsymbol{d}}\| = 1) \\ & \geq & \dfrac{\mathrm{diam}(C_0)^2}{2\delta} + \dfrac{\delta}{2} - \delta - \mathrm{diam}(C_0) \\ & & \text{(by } \boldsymbol{y},\ \bar{\boldsymbol{y}} \in (F(\psi)) \subset C_0 \text{ and (6))} \\ & = & \mathrm{diam}(C_0)\left(\dfrac{\mathrm{diam}(C_0)}{2\delta} - 1\right) - \dfrac{\delta}{2} \\ & \geq & \mathrm{diam}(C_0) - \dfrac{\delta}{2} \;\; \text{(by (7))} \\ & \geq & 4\delta - \dfrac{\delta}{2} \;\; \text{(by (7))} \\ & > & 0. \end{array}$$

Hence

$$\|\boldsymbol{y} - \bar{\boldsymbol{\xi}}\|^2 \;\leq\; \mathrm{diam}(C_0)^2 + \left(\bar{\boldsymbol{d}}^T(\boldsymbol{y} - \bar{\boldsymbol{\xi}})\right)^2$$

$$
\begin{aligned}
&\leq \operatorname{diam}(C_0)^2 + \left(\bar{\boldsymbol{d}}^T(\bar{\boldsymbol{y}} - \bar{\boldsymbol{\xi}})\right)^2 \\
&= \operatorname{diam}(C_0)^2 + \left(\bar{\boldsymbol{d}}^T(\tau - \delta)\bar{\boldsymbol{d}}\right)^2 \quad \text{(by (10))} \\
&= \operatorname{diam}(C_0)^2 + \tau^2 - 2\delta\tau + \delta^2 \quad (\text{since } \|\bar{\boldsymbol{d}}\| = 1) \\
&\leq \operatorname{diam}(C_0)^2 + \tau^2 - 2\delta\left(\frac{\operatorname{diam}(C_0)^2}{2\delta} + \frac{\delta}{2}\right) + \delta^2 \quad \text{(by (6))} \\
&= \tau^2.
\end{aligned}
$$

Thus we have seen that $\text{c.hull}(F(\psi)) \subset B(\bar{\boldsymbol{\xi}}, \tau)$.

(iv) Finally we will show that the first relation of (8), *i.e.*,, $\bar{\boldsymbol{\xi}} \in \Xi$. From the definition of τ, We know form (6) and (7) that

$$
\tau \;\geq\; \frac{\operatorname{diam}(C_0)^2}{2\delta} + \frac{\delta}{2} \geq \frac{\operatorname{diam}(C_0)}{2\delta} \times \operatorname{diam}(C_0) \geq 2\operatorname{diam}(C_0).
$$

Hence if we choose a $\boldsymbol{y} \in \text{c.hull}(F(\psi)) \subset C_0 \cap B(\bar{\boldsymbol{\xi}}, \tau)$, then

$$
\|\bar{\boldsymbol{\xi}} - \boldsymbol{\xi}_0\| \leq \|\bar{\boldsymbol{\xi}} - \boldsymbol{y}\| + \|\boldsymbol{y} - \boldsymbol{\xi}_0\| \leq \tau + \operatorname{diam}(C_0) \leq 3\tau/2.
$$

This implies that $\bar{\boldsymbol{\xi}} \in \Xi = B(\boldsymbol{\xi}_0, 3\tau/2)$. ∎

Corollary 2.5. *Assume that*

$$
\boldsymbol{\xi}_0 \in C_0 \quad \textit{and} \quad 0 < \epsilon \leq \frac{\nu_{\text{lip}} \operatorname{diam}(C_0)}{2}. \tag{11}
$$

Let

$$
\psi = \frac{\epsilon}{2}, \;\; \tau' = \frac{2\nu_{\text{lip}} \operatorname{diam}(C_0)^2}{\epsilon} \;\; \textit{and}\; \Xi = B(\boldsymbol{\xi}_0, \tau'). \tag{12}
$$

Then $\text{c.relax}(F(\psi), \Xi) \subset \text{c.hull}(F(\epsilon))$.

Proof: Let $\delta = \dfrac{\epsilon}{2\,\nu_{\text{lip}}}$. By (11), δ satisfies the inequality (7) of Lemma 2.4. We also see that

$$
\begin{aligned}
\frac{\operatorname{diam}(C_0)^2}{2\delta} + \frac{\delta}{2} &= \frac{\operatorname{diam}(C_0)^2}{2\delta} + \frac{\delta^2}{2\delta} \\
&\leq \frac{\operatorname{diam}(C_0)^2}{2\delta} + \frac{\operatorname{diam}(C_0)^2}{32\delta} \quad \text{(by (7))} \\
&= \frac{17\operatorname{diam}(C_0)^2}{32\delta} \\
&= \frac{17\,\nu_{\text{lip}} \operatorname{diam}(C_0)^2}{16\,\epsilon}.
\end{aligned}
$$

Hence if we take $\tau = \dfrac{4\,\nu_{\text{lip}} \operatorname{diam}(C_0)^2}{3\,\epsilon}$, then τ satisfies (6) and the desired result follows. ∎

3 The Spherical-SDP Model.

Throughout this section, we assume that $\nu_{\rm lip} = \nu_{\rm lip}(\mathcal{P}_F, C_0) < \infty$ and $\nu_{\rm nc} = \nu_{\rm nc}(\mathcal{P}_F) < \infty$. In the Spherical-SDP Model, we take $\mathcal{P}_k = \mathcal{P}^S(C_k)$ (the set of quadratic functions that induce spherical valid inequalities for C_k) at Step 1 of Algorithm 2.2. Here we say that a quadratic valid inequality $p(\boldsymbol{x}) \leq 0$ for C_k is spherical if $p(\cdot) : R^n \rightarrow R$ is of the form

$$p(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\xi})^T(\boldsymbol{x} - \boldsymbol{\xi}) - \rho^2$$

for $\exists \boldsymbol{\xi} \in R^n$ and $\exists \rho > 0$. Let $\{C_k\}$ be the sequence of compact convex sets generated by Algorithm 2.2 with $\mathcal{P}_k = \mathcal{P}^S(C_k)$ at every iteration. Then the sequence $\{C_k\}$ enjoys properties (a) monotonicity, (b) detecting infeasibility, and (c) asymptotic convergence, which we stated in the Introduction. See Theorem 3.1 of [2]. Among these properties, only the monotonicity property is relevant to the succeeding discussions.

Let ψ be an arbitrary positive number, and Ξ an arbitrary nonempty compact convex set containing C_0. In Lemma 3.2, we first present an upper bound $\hat{k}$ for the number k of iterations at which C_k attain an (ψ, Ξ)-convex-relaxation of F, *i.e.*, $C_k \subset \text{c.relax}(F(\psi), \Xi)$ holds for the first time. Then in Theorem 3.3, we will apply Corollary 2.5 to this bound to derive an upper bound k^* for the number k of iterations at which C_k attains an ϵ-convex-relaxation of F, *i.e.*, $C_k \subset \text{c.hull}(F(\epsilon))$ for the first time. Here ϵ denotes an arbitrarily given positive number.

Let η=diam(Ξ). Define

$$\bar{\delta} \;=\; \begin{cases} 0 & \text{if } \nu_{\rm nc}\, \eta^2 \leq \psi, \\ \left(1 - \dfrac{\psi}{\nu_{\rm nc}\, \eta^2}\right)^{\frac{1}{2}} & \text{otherwise.} \end{cases}$$

By definition, we see that $0 \leq \bar{\delta} < 1$.

Lemma 3.1. *Let $k \in \{0, 1, 2, \dots\}$. For $\forall \boldsymbol{\xi} \in \Xi$, define*

$$\rho'(\boldsymbol{\xi}) = \max\left\{\rho\left(\boldsymbol{\xi}, F(\psi)\right),\ \bar{\delta}\rho(\boldsymbol{\xi}, C_k)\right\}.$$

Then

$$C_{k+1} \subset \bigcap_{\boldsymbol{\xi} \in \Xi} B(\boldsymbol{\xi}, \rho'(\boldsymbol{\xi})).$$

Proof: It suffices to show that $C_{k+1} \subset B(\boldsymbol{\xi}, \rho'(\boldsymbol{\xi}))$ for $\forall \boldsymbol{\xi} \in \Xi$. For an arbitrarily fixed $\boldsymbol{\xi} \in \Xi$, let

$$\rho_k = \rho(\boldsymbol{\xi}, C_k) \ \text{ and } \rho' = \max\left\{\rho\left(\boldsymbol{\xi}, F(\psi)\right),\ \bar{\delta}\rho_k\right\}.$$

Since $\boldsymbol{\xi} \in \Xi$ and $C_k \subset C_0 \subset \Xi$, we see that

$$\rho_k \leq \eta, \tag{13}$$

which will be used later. If $\rho(\boldsymbol{\xi}, F(\psi)) \geq \rho_k$ then $\rho' = \rho(\boldsymbol{\xi}, F(\psi)) \geq \rho_k$. In this case the desired result follows from the monotonicity, *i.e.*, $C_{k+1} \subset C_k$. Now suppose that $\rho(\boldsymbol{\xi}, F(\psi)) < \rho_k$. Assuming that $\bar{\boldsymbol{x}} \notin B(\boldsymbol{\xi}, \rho')$, we derive $\bar{\boldsymbol{x}} \notin C_{k+1}$. If $\bar{\boldsymbol{x}} \notin C_k$, we obviously see that $\bar{\boldsymbol{x}} \notin C_{k+1}$ because $C_{k+1} \subset C_k$. Hence we only need to deal with the case that

$$\bar{\boldsymbol{x}} \in C_k \subset C_0, \ \rho(\boldsymbol{\xi}, F(\psi)) \leq \rho' < \|\bar{\boldsymbol{x}} - \boldsymbol{\xi}\| \leq \rho_k. \tag{14}$$

We will show that

$$\begin{aligned} qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) + g(\cdot) &\in \mathcal{Q}_+, & (15)\\ qf(\bar{\boldsymbol{x}}; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) + g(\bar{\boldsymbol{x}}) &> 0. & (16) \end{aligned}$$

for $\exists qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) \in \mathcal{P}_F$ and $\exists g(\cdot) \in \text{c.cone}(\mathcal{P}_k) = \text{c.cone}\left(\mathcal{P}^S(C_k)\right)$ in 3 steps (i), (ii) and (iii) below. Then $\bar{\boldsymbol{x}} \notin C_{k+1}$ follows from Lemma 2.1.

(i) The relations in (14) imply that $\bar{\boldsymbol{x}} \in C_0$ and $\bar{\boldsymbol{x}} \notin F(\psi)$. Hence there exists a quadratic function $qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) \in \mathcal{P}_F$ such that $qf(\bar{\boldsymbol{x}}; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) > \psi$.

(ii) By the definition of ρ_k, we also know that the quadratic function

$$p(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\xi})^T (\boldsymbol{x} - \boldsymbol{\xi}) - \rho_k^2 \ (\forall \boldsymbol{x} \in R^n)$$

is a member of $\mathcal{P}^S(C_k)$. Let $g(\cdot) = \nu_{\text{nc}}\, p(\cdot)$. By the definition of ν_{nc}, all the eigenvalues of the Hessian matrix of the quadratic function $qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) \in \mathcal{P}_F$ are not less than $-\nu_{\text{nc}}$. Hence (15) follows.

(iii) Finally we observe that

$$\begin{aligned} g(\bar{\boldsymbol{x}}) + qf(\bar{\boldsymbol{x}}; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) &= \nu_{\text{nc}}\, p(\bar{\boldsymbol{x}}) + qf(\bar{\boldsymbol{x}}; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) \\ &> \nu_{\text{nc}} \left((\bar{\boldsymbol{x}} - \boldsymbol{\xi})^T (\bar{\boldsymbol{x}} - \boldsymbol{\xi}) - \rho_k^2 \right) + \psi \\ &\geq \nu_{\text{nc}} \left((\bar{\delta}\rho_k)^2 - \rho_k^2 \right) + \psi \ \ (\text{since } \bar{\delta}\rho_k \leq \rho' < \|\bar{\boldsymbol{x}} - \boldsymbol{\xi}\|) \\ &= -\nu_{\text{nc}} \left(1 - \bar{\delta}^2\right) \rho_k^2 + \psi \\ &\geq -\nu_{\text{nc}} \left(1 - \bar{\delta}^2\right) \eta^2 + \psi \ \ (\text{by (13)}) \\ &\geq 0 \ \ (\text{by the definition of } \bar{\delta}). \end{aligned}$$

Thus we have shown (16). ∎

Lemma 3.2. *Suppose that* $\text{diam}(F) > 0$*. Define*

$$\hat{k} = \begin{cases} 1 & \textit{if } \nu_{\text{nc}}\, \eta^2 \leq \psi, \\ \left\lceil \dfrac{2\nu_{\text{nc}}\, \eta^2}{\psi} \ln \dfrac{2\eta}{\text{diam}(F)} \right\rceil & \textit{otherwise.} \end{cases}$$

If $k \geq \hat{k}$*, then* $C_k \subset \text{c.relax}(F(\psi), \Xi)$*.*

Proof: For every $\boldsymbol{\xi} \in \Xi$ and $k = 0, 1, 2, \ldots$, define

$$\rho_k(\boldsymbol{\xi}) = \max \{\rho(\boldsymbol{\xi}, F(\psi)), \rho(\boldsymbol{\xi}, C_k)\}.$$

It suffices to show that if $k \geq \hat{k}$ then

$$\rho_k(\boldsymbol{\xi}) \leq \rho(\boldsymbol{\xi}, F(\psi)) \quad \text{for} \quad \forall \boldsymbol{\xi} \in \Xi. \tag{17}$$

In fact, if (17) holds, then

$$C_k \subset \bigcap_{\boldsymbol{\xi} \in \Xi} B(\boldsymbol{\xi}, \rho_k(\boldsymbol{\xi})) \subset \bigcap_{\boldsymbol{\xi} \in \Xi} B(\boldsymbol{\xi}, \rho(\boldsymbol{\xi}, F(\psi))) = \text{c.relax}(F(\psi), \Xi).$$

By Lemma 3.1,

$$\rho_{k+1}(\boldsymbol{\xi}) \leq \max \left\{\rho(\boldsymbol{\xi}, F(\psi)), \bar{\delta}\rho_k(\boldsymbol{\xi})\right\} \quad \text{for } \forall k = 0, 1, 2, \ldots.$$

This implies that

$$\rho_k(\boldsymbol{\xi}) \leq \max \left\{\rho(\boldsymbol{\xi}, F(\psi)), \bar{\delta}^k \rho_0(\boldsymbol{\xi})\right\} \quad \text{for } \forall \boldsymbol{\xi} \in B(\eta) \quad \text{and } \forall k = 0, 1, 2, \ldots.$$

Hence, for each $\boldsymbol{\xi} \in \Xi$, if k satisfies the inequality

$$\bar{\delta}^k \rho_0(\boldsymbol{\xi}) \leq \rho(\boldsymbol{\xi}, F(\psi)), \tag{18}$$

then $\rho_k(\boldsymbol{\xi}) \leq \rho(\boldsymbol{\xi}, F(\psi))$. When $\nu_{\text{nc}}\, \eta^2 \leq \psi$, we see that $\bar{\delta} = 0$ by definition. Hence (18) holds for $k = 1$. Now assume that $\nu_{\text{nc}}\, \eta^2 > \psi$. Then $\bar{\delta} > 0$. Since

$$F \subset F(\psi) \subset C_0 \subset \Xi,$$

we see that

$$\text{diam}(F)/2 \leq \rho(\boldsymbol{\xi}, F(\psi)) \leq \rho_0(\boldsymbol{\xi}) \leq \eta \quad \text{for } \forall \boldsymbol{\xi} \in \Xi,$$

Hence, if $\bar{\delta}^k \eta \leq \text{diam}(F)/2$, or equivalently, if

$$k(-\ln \bar{\delta}) \geq \ln \frac{2\eta}{\text{diam}(F)},$$

then (18) holds. We also see from the definition of $\bar{\delta}$ that

$$-\ln \bar{\delta} \;\; = \;\; -\frac{1}{2} \ln \left(1 - \frac{\psi}{\nu_{\text{nc}}\, \eta^2}\right) \geq \frac{\psi}{2\nu_{\text{nc}}\, \eta^2} > 0.$$

Therefore if $\dfrac{k\psi}{2\nu_{\text{nc}}\, \eta^2} \geq \ln \dfrac{2\eta}{\text{diam}(F)}$, then (18) holds. Consequently we have shown that if $k \geq \dfrac{2\nu_{\text{nc}}\, \eta^2}{\psi} \ln \dfrac{2\eta}{\text{diam}(F)}$, then (17) holds. ∎

Now we are ready to present our main result in this section.

Theorem 3.3. *Assume (11) as in Corollary 2.5 and* $\mathrm{diam}(F) > 0$. *Define*

$$k^* = \begin{cases} 1 \quad \textit{if } (\nu_{\mathrm{lip}})^2 \, \nu_{\mathrm{nc}} \, \mathrm{diam}(C_0)^4 \le \dfrac{\epsilon^3}{32}, \\ \left\lceil \dfrac{64 \, (\nu_{\mathrm{lip}})^2 \, \nu_{\mathrm{nc}} \, \mathrm{diam}(C_0)^4}{\epsilon^3} \ln \dfrac{8 \, \nu_{\mathrm{lip}} \, \mathrm{diam}(C_0)^2}{\epsilon \, \mathrm{diam}(F)} \right\rceil & \textit{otherwise.} \end{cases}$$

If $k \ge k^*$, *then* $C_k \subset \mathrm{c.hull}(F(\epsilon))$.

It is interesting to note that the bound k^* is proportional to the nonconvexity ν_{nc} of $\mathcal{P}_F$, and $k^* = 1$ when quadratic functions in $\mathcal{P}_F$ are almost convex.

Proof of Theorem 3.3: Choose positive numbers ψ, τ' and a compact convex set $\Xi \subset R^n$ as in (12) of Corollary 2.5. Then $\mathrm{c.relax}(F(\psi), \Xi) \subset \mathrm{c.hull}(F(\epsilon))$. Let $\eta = \mathrm{diam}(\Xi) = 2\tau'$. Now, define $\hat{k}$ as in Lemma 3.2. By Lemma 3.2, if $k \ge \hat{k}$ then

$$C_k \subset \mathrm{c.relax}(F(\psi), \Xi) \subset \mathrm{c.hull}(F(\epsilon)).$$

On the other hand, we see that

$$\begin{aligned} \nu_{\mathrm{nc}} \, \eta^2 &= \nu_{\mathrm{nc}} \, (2\tau')^2 \\ &= \nu_{\mathrm{nc}} \left(\frac{4 \, \nu_{\mathrm{lip}} \, \mathrm{diam}(C_0)^2}{\epsilon} \right)^2 \\ &= \frac{16 \, (\nu_{\mathrm{lip}})^2 \, \nu_{\mathrm{nc}} \, \mathrm{diam}(C_0)^4}{\epsilon^2}. \end{aligned}$$

Hence the inequality $\nu_{\mathrm{nc}} \, \eta^2 \le \psi$ appeared in the definition of $\hat{k}$ can be rewritten as

$$(\nu_{\mathrm{lip}})^2 \, \nu_{\mathrm{nc}} \, \mathrm{diam}(C_0)^4 \le \frac{\epsilon^3}{32}.$$

We also see that

$$\begin{aligned} & \frac{2 \, \nu_{\mathrm{nc}} \, \eta^2}{\psi} \ln \frac{2\eta}{\mathrm{diam}(F)} \\ &= \frac{2 \, \nu_{\mathrm{nc}} \, (2\tau')^2}{\epsilon/2} \ln \frac{4 \, \tau'}{\mathrm{diam}(F)} \\ &= \frac{64 \, (\nu_{\mathrm{lip}})^2 \, \nu_{\mathrm{nc}} \, \mathrm{diam}(C_0)^4}{\epsilon^3} \ln \frac{8 \, \nu_{\mathrm{lip}} \, \mathrm{diam}(C_0)^2}{\epsilon \, \mathrm{diam}(F)}. \end{aligned}$$

Therefore $k^* = \hat{k}$. ∎

4 The Rank-2 Models.

We discuss the Rank-2-SDP and the Rank-2-SILP Models simultaneously. Let

$$\overline{D} \;=\; \{\boldsymbol{d} \in R^n : \|\boldsymbol{d}\| = 1\},$$

$$
\begin{aligned}
D_0 &= \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n, -\boldsymbol{e}_1, -\boldsymbol{e}_2, \ldots, -\boldsymbol{e}_n\}, \text{ where } \boldsymbol{e}_i \text{ denotes} \\
&\quad \text{the } i\text{th unit coordinate vector in } R^n, \\
\alpha(\boldsymbol{d}, C) &= \max\{\boldsymbol{d}^T\boldsymbol{x} : \boldsymbol{x} \in C\} \\
&\quad \text{for } \forall \boldsymbol{d} \in \overline{D} \text{ and } \forall \text{nonempty compact } C \subset R^n, \\
r2f(\boldsymbol{x}; \boldsymbol{d}_0, \boldsymbol{d}, C) &= -(\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C_0))(\boldsymbol{d}^T\boldsymbol{x} - \alpha(\boldsymbol{d}, C)) \quad \text{for } \forall \boldsymbol{x} \in R^n, \\
&\quad \forall \boldsymbol{d}_0 \in D_0, \ \forall \boldsymbol{d} \in \overline{D} \text{ and } \forall \text{nonempty compact } C \subset R^n.
\end{aligned}
$$

We define

$$
\widetilde{\mathcal{P}}^2(C) = \{r2f(\cdot; \boldsymbol{d}_0, \boldsymbol{d}, C) : \boldsymbol{d}_0 \in D_0, \ \boldsymbol{d} \in \overline{D}\}.
$$

Suppose that C is a subset of C_0. Then the two terms $(\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C_0))$ and $(\boldsymbol{d}^T\boldsymbol{x} - \alpha(\boldsymbol{d}, C))$ appeared in the definition of $r2f(\cdot; \boldsymbol{d}_0, \boldsymbol{d}, C)$ are linear supporting functions for $C_0 \subset C$ and C, respectively, and induce linear valid inequalities for $C_0 \subset C$ and C:

$$
\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C_0) \leq 0 \ \text{ and } \boldsymbol{d}^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C) \leq 0,
$$

respectively. Hence the minus of their product forms *a rank-2 quadratic supporting function*, and induces *a rank-2 quadratic valid inequality* for C:

$$
r2f(\boldsymbol{x}; \boldsymbol{d}_0, \boldsymbol{d}, C_k) = -\left(\boldsymbol{d}_0^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C_0)\right)\left(\boldsymbol{d}^T\boldsymbol{x} - \alpha(\boldsymbol{d}_0, C)\right) \leq 0.
$$

In the Rank-2-SDP and the Rank-2-SILP Models of successive convex relaxation methods, we take $\mathcal{P}_k = \widetilde{\mathcal{P}}^2(C_k)$ at Step 1 of Algorithms 2.2 and 2.3, respectively. Let $\{C_k\}$ be the sequence of compact convex sets generated by Algorithm 2.2 or Algorithm 2.3 with $\mathcal{P}_k = \widetilde{\mathcal{P}}^2(C_k)$. Then the sequence $\{C_k\}$ enjoys properties (a) monotonicity, (b) detecting infeasibility, and (c) asymptotic convergence, which we stated in the Introduction. We can prove this fact by similar arguments used in the proof of Theorem 3.3 of [3]. See also Remark 3.8 of [3]. Let ϵ be an arbitrarily given positive number. We will derive an upper bound k^* in Theorem 4.4 at the end of this section for the number k of iterations at which C_k attains an ϵ-convex-relaxation of F. The argument of the derivation of the bound k^* will proceed in a similar way as in the Spherical-SDP Model, although it is more complicated. In the derivation of the bound in the Spherical-SDP Model, it was a key to show the existence of a quadratic function $g(\cdot) \in \text{c.cone}(\mathcal{P}_k) = \text{c.cone}\left(\mathcal{P}^S(C_k)\right)$ satisfying the relations (15) and (16). We need a sophisticated argument, which we develop in Section 4.1, to construct a quadratic function $g(\cdot) \in \text{c.cone}(\mathcal{P}_k) = \text{c.cone}\left(\widetilde{\mathcal{P}}^2(C_k)\right)$ satisfying the corresponding relations (25) and (26) in the Rank-2-SDP and the SILP Models.

4.1 Convex Cones of Rank-2 Quadratic Supporting Functions for n-dimensional Balls.

In this subsection, we are concerned with $\text{c.cone}\left(\widetilde{\mathcal{P}}^2(B(\boldsymbol{\xi}, \rho))\right)$, where $\boldsymbol{\xi} \in R^n$ and $\rho > 0$. Note that if $C_k \subset B(\boldsymbol{\xi}, \rho)$, $r2f(\cdot; \boldsymbol{d}_0, \boldsymbol{d}, B(\boldsymbol{\xi}, \rho)) \in \widetilde{\mathcal{P}}^2(B(\boldsymbol{\xi}, \rho))$ and $r2f(\cdot; \boldsymbol{d}_0, \boldsymbol{d}, C_k) \in \widetilde{\mathcal{P}}^2(C_k)$, then $r2f(\cdot;\boldsymbol{d}_0,\boldsymbol{d},B(\boldsymbol{\xi},\rho))$ induces a weaker valid inequality for C_k than $r2f(\cdot;\boldsymbol{d}_0,\boldsymbol{d},C_k)$ in the sense that

$$r2f(\boldsymbol{x};\boldsymbol{d}_0,\boldsymbol{d},B(\boldsymbol{\xi},\rho)) \leq r2f(\boldsymbol{x};\boldsymbol{d}_0,\boldsymbol{d},C_k) \leq 0 \text{ for } \forall \boldsymbol{x} \in C_k.$$

This fact will be utilized in our complexity analysis of the Rank-2 Models in the next subsection. For simplicity of notation, we only deal with the case that $\boldsymbol{\xi} = \boldsymbol{0}$.

For $\forall\theta \in (0,\pi/8]$, $\forall \boldsymbol{w} \in \bar{D}$, and $\forall i = 1,2,\ldots,n$, define

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

For $\forall \boldsymbol{x} \in R^n$, $\forall\rho > 0$, $\forall\theta \in (0,\pi/8]$, $\forall \boldsymbol{w} \in \bar{D}$, $\forall i = 1,2,\ldots,n$, and $\forall j = 1,2,\ldots,n$, define

$$\left.\begin{array}{rcl} f^+_{ij}(\boldsymbol{x};\rho,\theta,\boldsymbol{w}) & = & \lambda_i(\theta,\boldsymbol{w})\ r2f(\boldsymbol{x};-\boldsymbol{e}_j,\boldsymbol{b}_i(\theta,\boldsymbol{w}),B(\boldsymbol{0},\rho)) \\ & & +\bar{\lambda}_i(\theta,\boldsymbol{w})\ r2f(\boldsymbol{x};\boldsymbol{e}_j,\bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w}),B(\boldsymbol{0},\rho)) \\ & = & -\lambda_i(\theta,\boldsymbol{w})\left(\boldsymbol{b}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho\right)\left(-\boldsymbol{e}_j^T\boldsymbol{x}-\alpha(-\boldsymbol{e}_j,C_0)\right) \\ & & -\bar{\lambda}_i(\theta,\boldsymbol{w})\left(\bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho\right)\left(\boldsymbol{e}_j^T\boldsymbol{x}-\alpha(\boldsymbol{e}_j,C_0)\right), \\ & & \\ f^-_{ij}(\boldsymbol{x};\rho,\theta,\boldsymbol{w}) & = & \lambda_i(\theta,\boldsymbol{w})\ r2f(\boldsymbol{x};\boldsymbol{e}_j,\boldsymbol{b}_i(\theta,\boldsymbol{w}),B(\boldsymbol{0},\rho)) \\ & & +\bar{\lambda}_i(\theta,\boldsymbol{w})\ r2f(\boldsymbol{x};-\boldsymbol{e}_j,\bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w}),B(\boldsymbol{0},\rho)) \\ & = & -\lambda_i(\theta,\boldsymbol{w})\left(\boldsymbol{b}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho\right)\left(\boldsymbol{e}_j^T\boldsymbol{x}-\alpha(\boldsymbol{e}_j,C_0)\right) \\ & & -\bar{\lambda}_i(\theta,\boldsymbol{w})\left(\bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w})^T\boldsymbol{x}-\rho\right)\left(-\boldsymbol{e}_j^T\boldsymbol{x}-\alpha(-\boldsymbol{e}_j,C_0)\right). \end{array}\right\} \quad (19)$$

Essentially the same functions as the quadratic functions $f^+_{ij}(\cdot;\rho,\theta,\boldsymbol{w})$ and $f^-_{ij}(\cdot;\rho,\theta,\boldsymbol{w})$ $(i,j = 1,2,\ldots,n)$ were introduced in the paper [3], and the lemma below is a further elaboration of Lemma 4.4 of [3] on their basic properties.

Lemma 4.1. *Let $\rho > 0$, $\theta \in (0,\pi/8]$, $\boldsymbol{w} \in \overline{D}$, $i \in \{1,2,\ldots,n\}$, and $j \in \{1,2,\ldots,n\}$.*

(i) $f^+_{ij}(\cdot;\rho,\theta,\boldsymbol{w}),\ f^-_{ij}(\cdot;\rho,\theta,\boldsymbol{w}) \in c.cone\left(\widetilde{\mathcal{P}}^2(B(\boldsymbol{0},\rho))\right)$.

(ii)$^+$ The Hessian matrix of the quadratic function $f^+_{ij}(\cdot;\rho,\theta,\boldsymbol{w}) : R^n \to R$ coincides with the $n \times n$ matrix $\dfrac{\boldsymbol{e}_i\boldsymbol{e}_j^T+\boldsymbol{e}_j\boldsymbol{e}_i^T}{2}$.

(ii)$^-$ The Hessian matrix of the quadratic function $f^-_{ij}(\cdot;\rho,\theta,\boldsymbol{w}) : R^n \to R$ coincides with the $n \times n$ matrix $-\dfrac{\boldsymbol{e}_i\boldsymbol{e}_j^T+\boldsymbol{e}_j\boldsymbol{e}_i^T}{2}$.

(iii) Suppose that $\delta \in [0,1]$ and $\delta\rho\boldsymbol{w} \in C_0$. Then

$$f^+_{ij}(\delta\rho\boldsymbol{w};\rho,\theta,\boldsymbol{w}),\ f^-_{ij}(\delta\rho\boldsymbol{w};\rho,\theta,\boldsymbol{w}) \in \left[-\rho\,\mathrm{diam}(C_0)\left(\frac{1-\delta(1-\theta^2/2)}{\theta-\theta^3/6}+1-\delta+\frac{\theta^2}{2}\right),\ 0\right].$$

(iv) Suppose that $\kappa \geq 0$, $1 \geq \delta \geq 1-\kappa\theta \geq 0$ and $\delta\rho\boldsymbol{w} \in C_0$. Then

$$f^+_{ij}(\delta\rho\boldsymbol{w};\rho,\theta,\boldsymbol{w}),\ f^-_{ij}(\delta\rho\boldsymbol{w};\rho,\theta,\boldsymbol{w}) \in [-2\rho\,\mathrm{diam}(C_0)(\kappa+\theta),\ 0]\,.$$

Proof: We will only show the relations on $f^+_{ij}(\cdot;\rho,\theta,\boldsymbol{w})$ because we can derive the corresponding relations on $f^-_{ij}(\cdot;\rho,\theta,\boldsymbol{w})$ quite similarly. First, assertion (i) follows directly from the definitions (19).

(ii)$^+$. The Hessian matrix of the quadratic function $f^+_{ij}(\cdot;\rho,\theta,\boldsymbol{w}) : R^n \to R$ turns out to be the symmetric part of the matrix

$$\begin{aligned}
&-\left(-\lambda_i(\theta,\boldsymbol{w})\ \boldsymbol{b}_i(\theta,\boldsymbol{w})\boldsymbol{e}_j^T + \bar{\lambda}_i(\theta,\boldsymbol{w})\ \bar{\boldsymbol{b}}_i(\theta,\boldsymbol{w})\boldsymbol{e}_j^T\right)\\
&= \quad -\frac{1}{2\sin\theta}\left(-(\boldsymbol{w}\cos\theta+\boldsymbol{e}_i\sin\theta)\boldsymbol{e}_j^T + (\boldsymbol{w}\cos\theta-\boldsymbol{e}_i\sin\theta)\boldsymbol{e}_j^T\right)\\
&= \quad -\frac{1}{2\sin\theta}\left(-2(\sin\theta)(\boldsymbol{e}_i\boldsymbol{e}_j^T)\right)\\
&= \quad \boldsymbol{e}_i\boldsymbol{e}_j^T.
\end{aligned}$$

Thus we have shown (ii)$^+$.

(iii) By definition, we have that

$$\begin{aligned}
&f^+_{ij}(\delta\rho\boldsymbol{w};\rho,\theta,\boldsymbol{w})\\
&\quad = -\frac{1}{2\sin\theta}\Big(\ \rho\left((\boldsymbol{w}\cos\theta+\boldsymbol{e}_i\sin\theta)^T(\delta\boldsymbol{w}) - \nu_i(\theta,\boldsymbol{w})\right)\left(-\boldsymbol{e}_j^T(\delta\rho\boldsymbol{w}) - \alpha(-\boldsymbol{e}_j,C_0)\right)\\
&\qquad\qquad + \ \rho\left((\boldsymbol{w}\cos\theta-\boldsymbol{e}_i\sin\theta)^T(\delta\boldsymbol{w}) - \bar{\nu}_i(\theta,\boldsymbol{w})\right)\left(\boldsymbol{e}_j^T(\delta\rho\boldsymbol{w}) - \alpha(\boldsymbol{e}_j,C_0)\right)\ \Big).
\end{aligned}$$

We will evaluate each term appeared in the right hand side. We first observe that

$$\begin{aligned}
\nu_i(\theta,\boldsymbol{w}) &= \|\boldsymbol{w}\cos\theta+\boldsymbol{e}_i\sin\theta\|\\
&= \left(\|\boldsymbol{w}\cos\theta\|^2 + \|\boldsymbol{e}_i\sin\theta\|^2 + 2\boldsymbol{e}_i^T\boldsymbol{w}\cos\theta\sin\theta\right)^{1/2}\\
&= \left(1+2\boldsymbol{e}_i^T\boldsymbol{w}\cos\theta\sin\theta\right)^{1/2}\\
&\leq 1+\boldsymbol{e}_i^T\boldsymbol{w}\cos\theta\sin\theta,
\end{aligned}$$

where the last inequality follows from

$$0 \leq (1+\xi)^{1/2} \leq 1+\frac{1}{2}\xi \ \text{ for } \forall\xi \geq -1.$$

Hence

$$\begin{aligned}
0 &\geq (\boldsymbol{w}\cos\theta+\boldsymbol{e}_i\sin\theta)^T(\delta\boldsymbol{w}) - \nu_i(\theta,\boldsymbol{w})\\
&\geq \delta(\cos\theta+\boldsymbol{e}_i^T\boldsymbol{w}\sin\theta) - \left(1+\boldsymbol{e}_i^T\boldsymbol{w}\cos\theta\sin\theta\right)\\
&= \delta\cos\theta - 1 + \boldsymbol{e}_i^T\boldsymbol{w}(\sin\theta)(\delta-\cos\theta)\\
&\geq \delta\cos\theta - 1 - (\sin\theta)|\cos\theta-\delta|.
\end{aligned}$$

Similarly

$$
\begin{aligned}
\bar{\nu}_i(\theta, \boldsymbol{w}) &= \|\boldsymbol{w}\cos\theta - \boldsymbol{e}_i \sin\theta\| \\
&= \left(\|\boldsymbol{w}\cos\theta\|^2 + \|\boldsymbol{e}_i \sin\theta\|^2 - 2\boldsymbol{e}_i^T \boldsymbol{w}\cos\theta\sin\theta\right)^{1/2} \\
&= \left(1 - 2\boldsymbol{e}_i^T \boldsymbol{w}\cos\theta\sin\theta\right)^{1/2} \\
&\leq 1 - \boldsymbol{e}_i^T \boldsymbol{w}\cos\theta\sin\theta,
\end{aligned}
$$

$$
\begin{aligned}
0 &\geq (\boldsymbol{w}\cos\theta - \boldsymbol{e}_i \sin\theta)^T(\delta\boldsymbol{w}) - \bar{\nu}_i(\theta, \boldsymbol{w}) \\
&\geq \delta(\cos\theta - \boldsymbol{e}_i^T \boldsymbol{w}\sin\theta) - \left(1 - \boldsymbol{e}_i^T \boldsymbol{w}\cos\theta\sin\theta\right) \\
&= \delta\cos\theta - 1 - \boldsymbol{e}_i^T \boldsymbol{w}(\sin\theta)(\delta - \cos\theta) \\
&\geq \delta\cos\theta - 1 - (\sin\theta)|\cos\theta - \delta|.
\end{aligned}
$$

Since $\delta\rho\boldsymbol{w} \in C_0$, we also see that

$$
\begin{array}{rcccl}
0 & \geq & -\boldsymbol{e}_j^T(\delta\rho\boldsymbol{w}) - \alpha(-\boldsymbol{e}_j, C_0) & \geq & -\mathrm{diam}(C_0), \\
0 & \geq & \boldsymbol{e}_j^T(\delta\rho\boldsymbol{w}) - \alpha(\boldsymbol{e}_j, C_0) & \geq & -\mathrm{diam}(C_0).
\end{array}
$$

Therefore

$$
\begin{aligned}
f_{ij}^+(\delta\rho\boldsymbol{w}; \rho, \theta, \boldsymbol{w}) &\geq -\frac{\rho}{2\sin\theta} \times 2\left(\delta\cos\theta - 1 - (\sin\theta)|\cos\theta - \delta|\right)\left(-\mathrm{diam}(C_0)\right) \\
&= -\frac{\rho\ \mathrm{diam}(C_0)}{\sin\theta}\left(1 - \delta\cos\theta + (\sin\theta)|\cos\theta - \delta|\right) \\
&= -\rho\ \mathrm{diam}(C_0)\left(\frac{1 - \delta\cos\theta}{\sin\theta} + |\cos\theta - \delta|\right) \\
&= -\rho\ \mathrm{diam}(C_0)\left(\frac{1 - \delta(1 - \theta^2/2 + \tilde{\theta}^4/24)}{\theta - \theta^3/6 + \bar{\theta}^5/120} + |1 - \hat{\theta}^2/2 - \delta|\right) \\
&\geq -\rho\ \mathrm{diam}(C_0)\left(\frac{1 - \delta(1 - \theta^2/2)}{\theta - \theta^3/6} + 1 - \delta + \frac{\theta^2}{2}\right) \quad (\text{since } 1 \geq \delta).
\end{aligned}
$$

Here $\tilde{\theta},\ \bar{\theta},\ \hat{\theta} \in [0, \theta]$.

(iv) We see from (iii) that

$$
\begin{aligned}
f_{ij}^+(\delta\rho\boldsymbol{w}; \rho, \theta, \boldsymbol{w}) &\geq -\rho\ \mathrm{diam}(C_0)\left(\frac{1 - \delta(1 - \theta^2/2)}{\theta(1 - \theta^2/6)} + 1 - \delta + \frac{\theta^2}{2}\right) \\
&\geq -\rho\ \mathrm{diam}(C_0)\left(\frac{\kappa\theta + \delta\theta^2/2}{\theta(1 - \theta^2/6)} + \kappa\theta + \frac{\theta^2}{2}\right) \quad (\text{since } 1 \geq \delta \geq 1 - \kappa\theta) \\
&\geq -\rho\ \mathrm{diam}(C_0)\left(\frac{\kappa + \theta/2}{1 - \theta^2/6} + \kappa\theta + \frac{\theta^2}{2}\right) \quad (\text{since } 1 \geq \delta) \\
&\geq -2\rho\ \mathrm{diam}(C_0)(\kappa + \theta) \quad (\text{since } \theta \in (0,\ \pi/8]).
\end{aligned}
$$

∎

4.2 Complexity Analisys.

In the remainder of the section, we assume that $\nu_{\rm lip} = \nu_{\rm lip}(\mathcal{P}_F, C_0) < \infty$ and $\nu_{\rm nl} = \nu_{\rm nl}(\mathcal{P}_F) < \infty$. Let $\{C_k\}$ be the sequence of compact convex sets generated by either Algorithm 2.2 or Algorithm 2.3 with taking $\mathcal{P}_k = \widetilde{\mathcal{P}}^2(C_k)$ at each iteration. Let ψ be an arbitrary positive number, and Ξ an arbitrary compact convex set containing C_0. In Lemma 4.3, we derive an upper bound $\hat{k}$ for the number k of iterations at which C_k attains a (ψ, Ξ)-convex-relaxation of F, *i.e.*, $C_k \subset \text{c.relax}(F(\psi), \Xi)$ holds for the first time. Then we will combine Lemmas 2.4 and 4.3 to derive an upper bound k^* for the number k of iterations at which C_k attains an ϵ-convex-relaxation of F. Here ϵ denotes an arbitrarily given positive number.

Let η=diam(Ξ). Define

$$
\begin{aligned}
\bar{\kappa} &= \begin{cases} \dfrac{8}{\pi} & \text{if } \nu_{\rm nl} \le \dfrac{\pi\,\psi}{32\,\eta\,\text{diam}(C_0)}, \\[2ex] \dfrac{\psi}{4\eta\,\nu_{\rm nl}\,\text{diam}(C_0)} & \text{otherwise,} \end{cases} \\
\bar{\theta} &= \begin{cases} \dfrac{\pi}{8} & \text{if } \nu_{\rm nl} \le \dfrac{2\psi}{\pi\,\eta\,\text{diam}(C_0)}, \\[2ex] \dfrac{\psi}{4\eta\,\nu_{\rm nl}\,\text{diam}(C_0)} & \text{otherwise,} \end{cases} \\
\bar{\delta} &= 1 - \bar{\kappa}\bar{\theta}.
\end{aligned}
$$

It should be noted that $\bar{\kappa}$ and $\bar{\theta}$ are not greater than $\dfrac{8}{\pi}$ and $\dfrac{\pi}{8}$, respectively, and also that $\dfrac{\pi\,\psi}{32\,\eta\,\text{diam}(C_0)} \le \dfrac{2\psi}{\pi\,\eta\,\text{diam}(C_0)}$. By definition, we see that

$$
\left.\begin{aligned}
1 &= \bar{\kappa}\bar{\theta} && \text{if } \nu_{\rm nl} \le \frac{\pi\,\psi}{32\,\eta\,\text{diam}(C_0)}, \\
1 &> \bar{\kappa}\bar{\theta} \ \ge \left(\frac{\psi}{4\eta\,\nu_{\rm nl}\,\text{diam}(C_0)}\right)^2 && \text{otherwise,}
\end{aligned}\right\} \tag{20}
$$

$$
\left.\begin{aligned}
\bar{\delta} &= 0 \quad \text{if } \nu_{\rm nl} \le \frac{\pi\,\psi}{32\,\eta\,\text{diam}(C_0)}, \\
1 > \bar{\delta} &> 0 \quad \text{otherwise,}
\end{aligned}\right\} \tag{21}
$$

$$
\frac{\psi}{2\eta\,\text{diam}(C_0)} \ \ge\ \left(\bar{\kappa} + \bar{\theta}\right)\nu_{\rm nl} \ \ge\ 0. \tag{22}
$$

Lemma 4.2. *Let* $k \in \{0, 1, 2, \dots\}$. *For* $\forall \boldsymbol{\xi} \in \Xi$, *define*

$$
\rho'(\boldsymbol{\xi}) = \max\left\{\rho\left(\boldsymbol{\xi}, F(\psi)\right),\ \bar{\delta}\rho(\boldsymbol{\xi}, C_k)\right\}.
$$

Then

$$
C_{k+1} \subset \bigcap_{\boldsymbol{\xi} \in \Xi} B(\boldsymbol{\xi}, \rho'(\boldsymbol{\xi})).
$$

Proof: It suffices to show that $C_{k+1} \subset B(\boldsymbol{\xi}, \rho'(\boldsymbol{\xi}))$ for $\forall \boldsymbol{\xi} \in \Xi$. For an arbitrarily fixed $\boldsymbol{\xi} \in \Xi$, let

$$\rho_k = \rho(\boldsymbol{\xi}, C_k) \;\text{ and }\; \rho' = \max\left\{\rho\left(\boldsymbol{\xi}, F(\psi)\right),\; \bar{\delta}\rho_k\right\}.$$

We may assume without loss of generality that $\boldsymbol{\xi} = \mathbf{0}$ because Algorithms 2.2 and 2.3 with taking $\mathcal{P}_k = \widetilde{\mathcal{P}}^2(C_k)$ at each iteration are invariant under any parallel transformation. See [3] for more details. Since $\boldsymbol{\xi} = \mathbf{0} \in \Xi$ and $C_k \subset C_0 \subset \Xi$, we see that

$$\rho_k \leq \eta, \tag{23}$$

which will be used later. If $\rho\left(\mathbf{0}, F(\psi)\right) \geq \rho_k$ then $\rho' = \rho\left(\mathbf{0}, F(\psi)\right) \geq \rho_k$. In this case the desired result follows from $C_{k+1} \subset C_k$. Now suppose that $\rho\left(\mathbf{0}, F(\psi)\right) < \rho_k$. Assuming that $\bar{\boldsymbol{x}} \notin B(\mathbf{0}, \rho')$, we will derive that $\bar{\boldsymbol{x}} \notin C_{k+1}$. If $\bar{\boldsymbol{x}} \notin C_k$, we obviously see $\bar{\boldsymbol{x}} \notin C_{k+1}$ because $C_{k+1} \subset C_k$. Hence we only need to deal with the case that

$$\bar{\boldsymbol{x}} \in C_k \subset C_0,\;\; \rho\left(\mathbf{0}, F(\psi)\right) \leq \rho' < \|\bar{\boldsymbol{x}} - \mathbf{0}\| \leq \rho_k. \tag{24}$$

We will show that

$$qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) + g(\cdot) \in \mathcal{L}, \tag{25}$$

$$qf(\bar{\boldsymbol{x}}; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) + g(\bar{\boldsymbol{x}}) > 0. \tag{26}$$

for $\exists qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) \in \mathcal{P}_F$ and $\exists g(\cdot) \in \mathcal{P}_k = \widetilde{\mathcal{P}}^2(C_k)$ in 3 steps (i), (ii) and (iii) below. Since $\mathcal{L} \subset \mathcal{Q}_+$, Lemma 2.1 shows $\bar{\boldsymbol{x}} \notin C_{k+1}$ in both cases of the Rank-2-SDP Model (Algorithm 2.2) and the SILP Model (Algorithm 2.3).

(i) The relations in (24) imply that $\bar{\boldsymbol{x}} \in C_0$ and $\bar{\boldsymbol{x}} \notin F(\psi)$. Hence there exists a quadratic function $qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) \in \mathcal{P}_F$ such that $qf(\bar{\boldsymbol{x}}; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) > \psi$.

(ii) Let $\bar{\boldsymbol{w}} = \bar{\boldsymbol{x}}/\|\bar{\boldsymbol{x}}\|$, and $\delta = \|\bar{\boldsymbol{x}}\|/\rho_k$. Then we see that

$$\begin{array}{rcl} \bar{\boldsymbol{x}} & = & \|\bar{\boldsymbol{x}}\|\bar{\boldsymbol{w}} = \delta\rho_k\bar{\boldsymbol{w}}, \\ 1 & \geq & \delta = \|\bar{\boldsymbol{x}}\|/\rho_k > \rho'/\rho_k \geq \bar{\delta} = 1 - \bar{\kappa}\bar{\theta} \geq 0. \end{array}$$

We will represent the symmetric matrix $\bar{\boldsymbol{Q}}$ as the difference of two $n \times n$ symmetric matrices $\bar{\boldsymbol{Q}}^+$ and $\bar{\boldsymbol{Q}}^-$ with nonnegative elements such that

$$\bar{\boldsymbol{Q}} = \bar{\boldsymbol{Q}}^+ - \bar{\boldsymbol{Q}}^-,\;\; \bar{Q}^+_{ij} \geq 0,\;\; \bar{Q}^-_{ij} \geq 0, \;\;\text{and}\; \bar{Q}^+_{ij}\bar{Q}^-_{ij} = 0 \;(i, j = 1, 2, \ldots, n).$$

For $\forall \boldsymbol{x} \in R^n$, $\forall i = 1, 2, \ldots, n$, and $\forall j = 1, 2, \ldots, n$, we now define

$$\left.\begin{array}{rcl} g_{ij}^{+}(\boldsymbol{x}) & = & \lambda_i(\bar{\theta}, \bar{\boldsymbol{w}})\ r2f(\boldsymbol{x}; -\boldsymbol{e}_j, \boldsymbol{b}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k) \\ & & +\bar{\lambda}_i(\bar{\theta}, \bar{\boldsymbol{w}})\ r2f(\boldsymbol{x}; \boldsymbol{e}_j, \bar{\boldsymbol{b}}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k) \\ & = & -\lambda_i(\bar{\theta}, \bar{\boldsymbol{w}}) \left(\boldsymbol{b}_i(\bar{\theta}, \bar{\boldsymbol{w}})^T \boldsymbol{x} - \alpha(\boldsymbol{b}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k)\right) \left(-\boldsymbol{e}_j^T \boldsymbol{x} - \alpha(-\boldsymbol{e}_j, C_0)\right) \\ & & -\bar{\lambda}_i(\bar{\theta}, \bar{\boldsymbol{w}}) \left(\bar{\boldsymbol{b}}_i(\bar{\theta}, \bar{\boldsymbol{w}})^T \boldsymbol{x} - \alpha(\bar{\boldsymbol{b}}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k)\right) \left(\boldsymbol{e}_j^T \boldsymbol{x} - \alpha(\boldsymbol{e}_j, C_0)\right), \\ & & \\ g_{ij}^{-}(\boldsymbol{x}) & = & \lambda_i(\bar{\theta}, \bar{\boldsymbol{w}})\ r2f(\boldsymbol{x}; \boldsymbol{e}_j, \boldsymbol{b}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k) \\ & & +\bar{\lambda}_i(\bar{\theta}, \bar{\boldsymbol{w}})\ r2f(\boldsymbol{x}; -\boldsymbol{e}_j, \bar{\boldsymbol{b}}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k) \\ & = & -\lambda_i(\bar{\theta}, \bar{\boldsymbol{w}}) \left(\boldsymbol{b}_i(\bar{\theta}, \bar{\boldsymbol{w}})^T \boldsymbol{x} - \alpha(\boldsymbol{b}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k)\right) \left(\boldsymbol{e}_j^T \boldsymbol{x} - \alpha(\boldsymbol{e}_j, C_0)\right) \\ & & -\bar{\lambda}_i(\bar{\theta}, \bar{\boldsymbol{w}}) \left(\bar{\boldsymbol{b}}_i(\bar{\theta}, \bar{\boldsymbol{w}})^T \boldsymbol{x} - \alpha(\bar{\boldsymbol{b}}_i(\bar{\theta}, \bar{\boldsymbol{w}}), C_k)\right) \left(-\boldsymbol{e}_j^T \boldsymbol{x} - \alpha(-\boldsymbol{e}_j, C_0)\right), \\ & & \\ g(\boldsymbol{x}) & = & \sum_{i=1}^n \sum_{j=1}^n \left(\bar{Q}_{ij}^{-} g_{ij}^{+}(\boldsymbol{x}) + \bar{Q}_{ij}^{+} g_{ij}^{-}(\boldsymbol{x})\right), \\ f(\boldsymbol{x}) & = & \sum_{i=1}^n \sum_{j=1}^n \left(\bar{Q}_{ij}^{-} f_{ij}^{+}(\boldsymbol{x}; \rho_k, \bar{\theta}, \bar{\boldsymbol{w}}) + \bar{Q}_{ij}^{+} f_{ij}^{-}(\boldsymbol{x}; \rho_k, \bar{\theta}, \bar{\boldsymbol{w}})\right). \end{array}\right\} \quad (27)$$

Here $f_{ij}^{+}(\cdot; \rho_k, \bar{\theta}, \bar{\boldsymbol{w}})$ and $f_{ij}^{-}(\boldsymbol{x}; \rho_k, \bar{\theta}, \bar{\boldsymbol{w}}))$ are defined as in (19). By construction, $g(\cdot) \in \widetilde{\mathcal{P}}^2(C_k)$. We also see that $g(\cdot)$ has the same Hessian matrix $\bar{\boldsymbol{Q}}^{-} - \bar{\boldsymbol{Q}}^{+} = -\bar{\boldsymbol{Q}}$ as $f(\cdot)$ does. See Lemma 4.1. Hence the Hessian matrix of the quadratic function $g(\cdot) + qf(\cdot; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}})$ vanishes. Thus we have shown (25).

(iii) Finally we observe that

$$\begin{array}{rcl} 0 & \geq & g(\bar{\boldsymbol{x}}) \\ & \geq & f(\bar{\boldsymbol{x}}) \text{ (since } \bar{\boldsymbol{x}} \in C_k \subset B(\boldsymbol{0}, \rho_k)) \\ & = & f(\delta \rho_k \bar{\boldsymbol{w}}) \\ & = & \displaystyle\sum_{i=1}^n \sum_{j=1}^n \left(\bar{Q}_{ij}^{+} f_{ij}^{+}(\delta \rho_k \bar{\boldsymbol{w}}; \rho_k, \bar{\theta}, \bar{\boldsymbol{w}}) + \bar{Q}_{ij}^{-} f_{ij}^{-}(\delta \rho_k \bar{\boldsymbol{w}}; \rho_k, \bar{\theta}, \bar{\boldsymbol{w}})\right) \\ & \geq & \displaystyle\sum_{i=1}^n \sum_{j=1}^n \left(\bar{Q}_{ij}^{+} \left(-2\rho_k \operatorname{diam}(C_0)(\bar{\kappa} + \bar{\theta})\right) + \bar{Q}_{ij}^{-} \left(-2\rho_k \operatorname{diam}(C_0)(\bar{\kappa} + \bar{\theta})\right)\right) \\ & & \text{(by (iv) of Lemma 4.1)} \\ & = & \displaystyle -2\rho_k \operatorname{diam}(C_0)(\bar{\kappa} + \bar{\theta}) \sum_{i=1}^n \sum_{j=1}^n \left(\bar{Q}_{ij}^{+} + \bar{Q}_{ij}^{-}\right) \\ & = & \displaystyle -2\rho_k \operatorname{diam}(C_0)(\bar{\kappa} + \bar{\theta}) \sum_{i=1}^n \sum_{j=1}^n |\bar{Q}_{ij}| \\ & \geq & -2\eta \operatorname{diam}(C_0)(\bar{\kappa} + \bar{\theta})\ \nu_{\mathrm{nl}} \\ & \geq & -\psi \quad \text{(by (22))}. \end{array}$$

Therefore $g(\bar{\boldsymbol{x}}) + qf(\bar{\boldsymbol{x}}; \bar{\gamma}, \bar{\boldsymbol{q}}, \bar{\boldsymbol{Q}}) > 0$. ∎

Lemma 4.3. *Suppose that* $\mathrm{diam}(F) > 0$*. Define*

$$\hat{k} = \begin{cases} 1 & \text{if } \nu_{\mathrm{nl}} \le \dfrac{\pi\ \psi}{32\ \eta\ \mathrm{diam}(C_0)}, \\ \left\lceil \left(\dfrac{4\eta\ \nu_{\mathrm{nl}}\ \mathrm{diam}(C_0)}{\psi} \right)^2 \ln \dfrac{2\eta}{\mathrm{diam}(F)} \right\rceil & \text{otherwise.} \end{cases}$$

If $k \ge \hat{k}$*, then* $C_k \subset \mathrm{c.relax}(F(\psi), \Xi)$*.*

Proof: For every $\boldsymbol{\xi} \in \Xi$ and $k = 0, 1, 2, \dots$, define

$$\rho_k(\boldsymbol{\xi}) = \max\left\{\rho\left(\boldsymbol{\xi}, F(\psi)\right), \rho(\boldsymbol{\xi}, C_k)\right\}.$$

It suffices to show that if $k \ge \hat{k}$ then

$$\rho_k(\boldsymbol{\xi}) \le \rho\left(\boldsymbol{\xi}, F(\psi)\right) \ \text{ for } \forall \boldsymbol{\xi} \in \Xi. \tag{28}$$

In fact, if (28) holds, then

$$C_k \subset \bigcap_{\boldsymbol{\xi}\in\Xi} B(\boldsymbol{\xi}, \rho_k(\boldsymbol{\xi})) \subset \bigcap_{\boldsymbol{\xi}\in\Xi} B\left(\boldsymbol{\xi}, \rho\left(\boldsymbol{\xi}, F(\psi)\right)\right) = \mathrm{c.relax}(F(\psi), \Xi).$$

By Lemma 4.2,

$$\rho_{k+1}(\boldsymbol{\xi}) \le \max\left\{\rho\left(\boldsymbol{\xi}, F(\psi)\right), \bar{\delta}\rho_k(\boldsymbol{\xi})\right\} \ (k = 0, 1, 2, \dots).$$

This implies that

$$\rho_k(\boldsymbol{\xi}) \le \max\left\{\rho\left(\boldsymbol{\xi}, F(\psi)\right), \bar{\delta}^k\rho_0(\boldsymbol{\xi})\right\} \ \text{ for } \forall \boldsymbol{\xi} \in \Xi \ \text{ and } \forall k = 0, 1, 2, \dots .$$

Hence, for each $\boldsymbol{\xi} \in \Xi$, if k satisfies the inequality

$$\bar{\delta}^k\rho_0(\boldsymbol{\xi}) \le \rho\left(\boldsymbol{\xi}, F(\psi)\right), \tag{29}$$

then $\rho_k(\boldsymbol{\xi}) \le \rho\left(\boldsymbol{\xi}, F(\psi)\right)$. When $\nu_{\mathrm{nl}} \le \dfrac{\pi\ \psi}{32\ \eta\ \mathrm{diam}(C_0)}$, we see by (21) that $\bar{\delta} = 0$. Hence (28) holds for $k = 1$. Now assume that $\nu_{\mathrm{nl}} > \dfrac{\pi\ \psi}{32\ \eta\ \mathrm{diam}(C_0)}$. Then $\bar{\delta} > 0$ by (21). Since

$$F \subset F(\psi) \subset C_0 \subset \Xi,$$

we see that

$$\mathrm{diam}(F)/2 \le \rho\left(\boldsymbol{\xi}, F(\psi)\right) \le \rho_0(\boldsymbol{\xi}) \le \eta \ \text{ for } \forall \boldsymbol{\xi} \in \Xi,$$

Hence, if $\bar{\delta}^k\eta \le \mathrm{diam}(F)/2$, or equivalently, if

$$k(-\ln\bar{\delta}) \ge \ln\frac{2\eta}{\mathrm{diam}(F)},$$

then (29) holds. We also see by the definition of $\bar{\delta}$ and (20) that

$$-\ln \bar{\delta} \;=\; -\ln\left(1-\bar{\kappa}\bar{\theta}\right) \geq \bar{\kappa}\bar{\theta} \geq \left(\frac{\psi}{4\eta\ \nu_{\rm nl}\ {\rm diam}(C_0)}\right)^2 > 0.$$

Therefore if $k\left(\frac{\psi}{4\eta\ \nu_{\rm nl}\ {\rm diam}(C_0)}\right)^2 \geq \ln \frac{2\eta}{{\rm diam}(F)}$, then (29) holds. Consequently we have shown that if $k \geq \left(\frac{4\eta\ \nu_{\rm nl}\ {\rm diam}(C_0)}{\psi}\right)^2 \ln \frac{2\eta}{{\rm diam}(F)}$ then (28) holds. ∎

Theorem 4.4. *Assume (11) as in Corollary 2.5 and* ${\rm diam}(F) > 0$*. Define*

$$k^* = \begin{cases} 1 & \textit{if } \nu_{\rm nl} \leq \dfrac{\pi\ \epsilon^2}{256\ \nu_{\rm lip}\ {\rm diam}(C_0)^3}, \\[2ex] \left\lceil \left(\dfrac{32\ \nu_{\rm lip}\ \nu_{\rm nl}\ {\rm diam}(C_0)^3}{\epsilon^2}\right)^2 \ln \dfrac{8\ \nu_{\rm lip}\ {\rm diam}(C_0)^2}{\epsilon\ {\rm diam}(F)} \right\rceil & \textit{otherwise.} \end{cases}$$

If $k \geq k^*$, then $C_k \subset$ c.hull$(F(\epsilon))$.

Note that the bound k^* is proportional to the square of the nonlinearity $\nu_{\rm nl}$ of $\mathcal{P}_F$, and also that $k^* = 1$ when any function in $\mathcal{P}_F$ is almost linear.

Proof of Theorem 4.4: Choose positive numbers ψ, τ' and a compact convex set $\Xi \subset R^n$ as in (12) of Corollary 2.5. Then c.relax$(F(\psi), \Xi) \subset$ c.hull$(F(\epsilon))$. Let $\eta = {\rm diam}(\Xi) = 2\tau'$. Now, define $\hat{k}$ as in Lemma 4.3. By Lemma 4.3, if $k \geq \hat{k}$ then

$$C_k \subset \text{c.relax}(F(\psi), \Xi) \subset \text{c.hull}(F(\epsilon)).$$

On the other hand, we see that

$$\begin{aligned} \frac{\pi\ \psi}{32\ \eta\ {\rm diam}(C_0)} &= \frac{\pi\ \epsilon/2}{32\ (2\tau')\ {\rm diam}(C_0)} \\ &= \frac{\pi\ \epsilon}{128\ {\rm diam}(C_0)}\left(\frac{\epsilon}{2\ \nu_{\rm lip}\ {\rm diam}(C_0)^2}\right) \quad \text{(by (12))} \\ &= \frac{\pi\ \epsilon^2}{256\ \nu_{\rm lip}\ {\rm diam}(C_0)^3}. \end{aligned}$$

Hence the inequality $\nu_{\rm nl} \leq \frac{\pi\ \psi}{32\ \eta\ {\rm diam}(C_0)}$ appeared in the definition of $\hat{k}$ can be rewritten as

$$\nu_{\rm nl} \leq \frac{\pi\ \epsilon^2}{256\ \nu_{\rm lip}\ {\rm diam}(C_0)^3}.$$

We also see that

$$\left(\frac{4\eta\ \nu_{\rm nl}\ {\rm diam}(C_0)}{\psi}\right)^2 \ln \frac{2\eta}{{\rm diam}(F)}$$

$$= \left(\frac{8\tau' \ \nu_{\rm nl} \ {\rm diam}(C_0)}{\epsilon/2}\right)^2 \ln \frac{4\tau'}{{\rm diam}(F)}$$

$$= \left(\frac{16(2 \ \nu_{\rm lip} \ {\rm diam}(C_0)^2) \ \nu_{\rm nl} \ {\rm diam}(C_0)}{\epsilon^2}\right)^2 \ln \frac{4(2 \ \nu_{\rm lip} \ {\rm diam}(C_0)^2)}{\epsilon \ {\rm diam}(F)}$$

$$= \left(\frac{32 \ \nu_{\rm lip} \ \nu_{\rm nl} \ {\rm diam}(C_0)^3}{\epsilon^2}\right)^2 \ln \frac{8 \ \nu_{\rm lip} \ {\rm diam}(C_0)^2}{\epsilon \ {\rm diam}(F)}$$

Therefore $k^* = \hat{k}$. ∎

5 Concluding Discussions.

(A) In the Spherical-SDP Model, we have assumed that every ball with a given center $\boldsymbol{\xi} \in R^n$ that contains the kth iterate C_k is available. This assumption is equivalent to the assumption that for a given $\boldsymbol{\xi} \in R^n$, we can solve a norm maximization problem

$$\text{maximize } \|\boldsymbol{x} - \boldsymbol{\xi}\| \ \text{ subject to } \boldsymbol{x} \in C_k. \tag{30}$$

In fact, if we denote the maximum value of this problem by $\rho(\boldsymbol{\xi}, C_k)$, then we can represent the set of all balls containing C_k as

$$\{B(\boldsymbol{\xi}, \tau) : \tau \geq \rho(\boldsymbol{\xi}, C_k), \ \boldsymbol{\xi} \in R^n\},$$

and $\mathcal{P}_k = \mathcal{P}^S(C_k)$ consists of the spherical quadratic functions $p(\cdot; \boldsymbol{\xi}, \tau)$ ($\boldsymbol{\xi} \in \boldsymbol{R}^n$, $\tau \geq \rho(\boldsymbol{\xi}, C_k)$ such that

$$p(\boldsymbol{x}; \boldsymbol{\xi}, \tau) = (\boldsymbol{x} - \boldsymbol{\xi})^T(\boldsymbol{x} - \boldsymbol{\xi}) - \tau^2 \text{ for } \forall \boldsymbol{x} \in R^n.$$

We can also observe from the argument in Section 3 that among the spherical quadratic functions in $\mathcal{P}^S(C_k)$, only the supporting spherical quadratic functions $p(\cdot; \boldsymbol{\xi}, \rho(\boldsymbol{\xi}, C_k))$ ($\boldsymbol{\xi} \in R^n$) are relevant in constructing the next iterate $C_{k+1} = F(C_k, \mathcal{P}_F \cup \mathcal{P}^S(C_k))$.

In the Rank-2-SDP and the Rank-2-SILP Models, we have assumed that the maximum value $\alpha(\boldsymbol{d}, C_k)$ of the linear function $\boldsymbol{d}^T\boldsymbol{x}$ over $\boldsymbol{x} \in C_k$ is available for every $\boldsymbol{d} \in \overline{D}$. From the practical point of view, the latter assumption on the Rank-2-SDP and the Rank-2-SILP Models looks much weaker than the former assumption on the Spherical-SDP Model because a maximization of a linear function $\boldsymbol{d}^T\boldsymbol{x}$ over C_k is a convex program while the norm maximization problem (30) is a nonconvex program. But the latter assumption still requires to know the maximum value $\alpha(\boldsymbol{d}, C_k)$ for $\forall \boldsymbol{d}$ in the set $\overline{D}$ consisting of infinitely many directions, so that these models are not implementable yet. Kojima and Tunçel [3] proposed discretization and localization techniques to implement the Rank-2-SDP and the Rank-2-SILP Models, and Takeda, Dai, Fukuda and Kojima [9] reported numerical results on practical versions of the Rank-2-SILP Model.

(B) In Section 4, we have focused on the Rank-2-SILP Model and diverted its complexity analysis to the Rank-2-SDP Model since the SDP relaxation is at least as tight as the semi-infinite LP relaxation. But this is somewhat loose. In particular, we should mention one important difference between the upper bounds required to attain an ϵ-convex-relaxation of F in Theorems 4.4 and 3.3: the upper bound in Theorems 4.4 depends on the nonlinearity $\nu_{\mathrm{nl}}(\mathcal{P}_F)$ of $\mathcal{P}_F$, while the upper bound in Theorem 3.3 depends on the nonconvexity $\nu_{\mathrm{nc}}(\mathcal{P}_F)$ but not on the nonlinearity ν_{nl}. This difference is critical when all the quadratic functions are convex but nonlinear. Here we state a complexity analysis on the Rank-2-SDP Model which leads to an upper bound depending on the nonconvexity ν_{nc} of $\mathcal{P}_F$ but not on the nonlinearity ν_{nl} of $\mathcal{P}_F$.

Define $\bar{\kappa}, \bar{\theta}$ and δ as

$$
\begin{aligned}
\bar{\kappa} &= \begin{cases} \dfrac{8}{\pi} & \text{if } \nu_{\mathrm{nc}} \leq \dfrac{\pi\,\psi}{32n\eta\,\mathrm{diam}(C_0)}, \\[2ex] \dfrac{\psi}{4n\eta\,\nu_{\mathrm{nc}}\,\mathrm{diam}(C_0)} & \text{otherwise,} \end{cases} \\
\bar{\theta} &= \begin{cases} \dfrac{\pi}{8} & \text{if } \nu_{\mathrm{nc}} \leq \dfrac{2\psi}{\pi n\eta\,\mathrm{diam}(C_0)}, \\[2ex] \dfrac{\psi}{4n\eta\,\nu_{\mathrm{nc}}\,\mathrm{diam}(C_0)} & \text{otherwise,} \end{cases} \\
\bar{\delta} &= 1 - \bar{\kappa}\bar{\theta}.
\end{aligned}
$$

By definition, we see that $\bar{\kappa} \geq 0$, $1 \geq \bar{\delta} \geq 0$ and $(\bar{\kappa}+\bar{\theta})\nu_{\mathrm{nc}} \leq \dfrac{\psi}{2n\eta\,\mathrm{diam}(C_0)}$. It is easily seen that the assertion of Lemma 4.2 remains valid with the definition above. The proof is quite similar except replacing the quadratic function $g(\cdot)$ in (27) by

$$
g(\boldsymbol{x}) = \nu_{\mathrm{nc}} \sum_{i=1}^{n} g_{ii}^{+}(\boldsymbol{x}) \text{ for } \forall \boldsymbol{x} \in R^n.
$$

By using similar arguments as in Lemma 4.3, Theorem 4.4 and their proofs, we consequently obtain the following result: Assume (11) as in Corollary 2.5 and $\mathrm{diam}(F) > 0$. Define

$$
k^* = \begin{cases} 1 \quad \text{if } \nu_{\mathrm{nc}} \leq \dfrac{\pi\,\epsilon^2}{256\,n\,\nu_{\mathrm{lip}}\,\mathrm{diam}(C_0)^3}, \\[2ex] \left\lceil \left(\dfrac{32\,n\,\nu_{\mathrm{lip}}\,\nu_{\mathrm{nc}}\,\mathrm{diam}(C_0)^3}{\epsilon^2} \right)^2 \ln \dfrac{8\,\nu_{\mathrm{lip}}\,\mathrm{diam}(C_0)^2}{\epsilon\,\mathrm{diam}(F)} \right\rceil & \text{otherwise.} \end{cases}
$$

If $k \geq k^*$, then $C_k \subset \mathrm{c.hull}(F(\epsilon))$.

Note that now the bound k^* is proportional to the square of the nonconvexity ν_{nc} of $\mathcal{P}_F$, and also that $k^* = 1$ when quadratic functions in $\mathcal{P}_F$ are almost convex.

References

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

[2] M. Kojima and L. Tunçel, "Cones of matrices and successive convex relaxations of nonconvex sets," Technical Report B-338, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, March 1998. Revised June 1999. Also issued as CORR 98-6, Dept. of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

[3] M. Kojima and L. Tunçel, "Discretization and localization in successive convex relaxation methods for nonconvex quadratic optimization problems," Technical Report B-341, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, July 1998. Revised July 1999. Also issued as CORR 98-34, Dept. of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

[4] M. Kojima, "Moderate Nonlinearity = Convexity + Quadratic Concavity," Technical Report B-348, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, March 1999.

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

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

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

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

[9] A. Takeda, Y. Dai, M. Fukuda, and M. Kojima, "Towards the Implementation of Successive Convex Relaxation Method for Nonconvex Quadratic Optimization Problems," Technical Report B-347, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo, Japan, March1999.

[10] H. Tuy, *Convex analysis and Global Optimization*, (Kluwer, Dordrecht, 1998).