A Regularized Explicit Exchange Method for Semi-Infinite Programs with an Infinite Number of Conic Constraints*

Takayuki Okuno, Shunsuke Hayashi and Masao Fukushima†

Abstract

The semi-infinite program (SIP) is normally represented with infinitely many inequality constraints, and has been studied extensively so far. However, there have been very few studies on the SIP involving conic constraints, even though it has important applications such as Chebychev-like approximation, filter design, and so on.

In this paper, we focus on the SIP with a convex objective function and infinitely many conic constraints, called an SICP for short. We show that, under the Robinson constraint qualification, an optimum of the SICP satisfies the KKT conditions that can be represented only with a *finite* subset of the conic constraints. We also introduce two exchange type algorithms for solving the SICP. We first provide an explicit exchange method, and show that it has global convergence under the strict convexity assumption on the objective function. We then propose an algorithm combining a regularization method with the explicit exchange method, and establish the global convergence of the hybrid algorithm without the strict convexity assumption.

1 Introduction

In this paper, we focus on the following optimization problem with an infinite number of *conic* constraints:

$$\begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & A(t)^\top x - b(t) \in C \quad \text{for all } t \in T, \end{array} \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable convex function, $A : T \to \mathbb{R}^{n\times m}$ and $b : T \to \mathbb{R}^m$ are continuous functions, $T \subset \mathbb{R}^\ell$ is a given compact set, and $C \subset \mathbb{R}^m$ is a closed convex cone with nonempty interior. We call this problem the semi-infinite conic program, SICP for short. Throughout this paper, we assume that SICP (1.1) has a nonempty solution set.

When $m = 1$ and $C = \mathbb{R}_+ := \{z \in \mathbb{R} \mid z \geq 0\}$, SICP (1.1)) reduces to the classical semi-infinite program (SIP) [8, 15, 12, 18, 19, 22], which has a wide application in engineering, e.g., the air pollution control, the robot trajectory planning, the stress of materials, etc.[12, 18]. So far, many algorithms have been proposed for solving SIPs, such as the discretization method [8], the local reduction based method [9, 16, 21] and the exchange method [15, 10, 22]. The discretization method solves a sequence of relaxed SIPs with T replaced by $T^k \subseteq T$, where T^k is a finite index set such that the distance[1] from T^k to T converges to 0 as k goes to infinity. While this method is comprehensible and easy to implement, the computational cost tends to be huge since the

*Technical report 2011-013, July 28, 2011.

†Department of Applied Mathematics and Phisics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan ({t_okuno, shunhaya, fuku}@amp.i.kyoto-u.ac.jp).

[1]For two sets $X \subseteq Y$, the distance from X to Y is defined as $\operatorname{dist}(X, Y) := \sup_{y\in Y} \inf_{x\in X} \|x - y\|$.

cardinality of T^k grows infinitely large. In the local reduction based method, the infinite number of constraints in the original SIP are rewritten as a finite number of constraints with implicit functions. Although the SIP can be reformulated as a finitely constrained optimization problem by this method, it is not possible in general to evaluate the implicit functions exactly in a direct manner. The exchange method solves a relaxed subproblem with T replaced by a finite subset $T^k \subseteq T$, where T^k is updated so that $T^{k+1} \subseteq T^k \cup \{t_1, t_2, \cdots, t_r\}$ with $\{t_1, t_2, \cdots, t_r\} \subseteq T \setminus T^k$.

A more general choice for C is the symmetric cone such as the second-order cone (SOC) $\mathcal{K}^m := \left\{(z_1, z_2, \ldots, z_m)^\top \in \mathbb{R}^m \mid z_1 \geq \|(z_2, z_3, \ldots, z_m)^\top\|_2\right\}$ and the semi-definite cone $\mathcal{S}^m_+ := \{Z \in \mathbb{R}^{m\times m} \mid Z = Z^\top,\ Z \succeq 0\}$. We note that the algorithm proposed in this paper needs to solve a sequence of subproblems in which T is replaced by a finite subset $\{t_1, t_2, \ldots, t_r\} \subseteq T$. To such a subproblem, we can apply an existing algorithm such as the interior-point method and the smoothing Newton method [1, 6, 11, 17]. There are some important applications of SICP (1.1). For example, when C is an SOC, SICP (1.1) can be used to formulate a Chebychev-like approximation problem involving vector-valued functions. Specifically, let $Y \subseteq \mathbb{R}^n$ be a given compact set, and $\Phi : Y \to \mathbb{R}^m$ and $F : \mathbb{R}^\ell \times Y \to \mathbb{R}^m$ be given functions. Then, we want to determine a parameter $u \in \mathbb{R}^\ell$ such that $\Phi(y) \approx F(u, y)$ for all $y \in Y$. One relevant approach is to solve the following problem:

$$\underset{u}{\text{Minimize}} \quad \max_{y\in Y} \|\Phi(y) - F(u, y)\|_2.$$

By introducing the auxiliary variable $r \in \mathbb{R}$, we can transform the above problem to

$$\begin{array}{ll} \underset{u,r}{\text{Minimize}} & r \\ \text{subject to} & \begin{pmatrix} r \\ \Phi(y) - F(u, y) \end{pmatrix} \in \mathcal{K}^{m+1} \quad \text{for all } y \in Y, \end{array}$$

which is of the form (1.1) when F is affine with respect to u.

The main purpose of the paper is two-fold. First, we study the Karush-Kuhn-Tucker (KKT) conditions for SICP. (We actually focus on the more general SICP of the form (2.1) or (2.6).) Although the original KKT conditions for SICP could be described by means of integration and Borel measure, we show that they can be represented by a *finite* number of elements in T under the Robinson constraint qualification. Second, we provide two algorithms for solving SICP (1.1). Since any closed convex cone can be represented as an intersection of finitely or infinitely many halfspaces, we may reformulate (1.1) as a classical SIP with infinitely many linear inequlity constraints, and solve it by using existing SIP algorithms [12, 18]. However, such a reformulation approach brings more difficulties since the dimension of the index set may become much larger than that of the original SICP (1.1).[2] Therefore, it is more reasonable to deal with the cones directly without losing their special structures. The two algorithms proposed in this paper are based on the exchange method, which solves a sequence of subproblems with *finitely* many conic constraints. The first algorithm is an explicit exchange method, of which we show global convergence under the strict convexity of the objective function. The second algorithm is a regularized explicit exchange method, which is a hybrid of the explicit exchange method and the regularization method. With the help of regularization, global convergence of the algorithm can be established without the strict convexity assumption.

This paper is organized as follows. In Section 2, we discuss the KKT conditions for SICP (1.1). In Section 3, we propose the explicit exchange method for solving SICP (1.1). In Section 4, we combine the explicit exchange method with the regularization method, and show that the hybrid

[2] In the case where $C = \mathcal{K}^m$, since $\mathcal{K}^m = \{z \in \mathbb{R}^m \mid z^\top s \geq 0, \forall s \in S\}$, where $S := \{(1, \bar{s})^\top \in \mathbb{R}^m \mid \|\bar{s}\| = 1\}$, SICP (1.1) can be reformulated as the SIP: min $f(x)$ s.t. $s^\top(A(t)^\top x - b(t)) \geq 0$ for all $(s, t) \in S \times T$. The dimension of $S \times T$ is then equal to $m + \dim T - 1$, where $\dim T$ denotes the dimension of T.

algorithm is globally convergent for SICP (1.1). In Section 5, we give some numerical results to examine the efficiency of the proposed algorithm . In Section 6, we conclude the paper with some remarks.

Throughout the paper, we use the following notations. $\|\cdot\|$ denotes the Euclidean norm defined by $\|z\| := \sqrt{z^\top z}$ for $z \in \mathbb{R}^m$. For $z^i \in \mathbb{R}^{m_i}$ $(i = 1, 2, \dots, p)$, we often write $(z^1, z^2, \dots, z^p)$ for $((z^1)^\top, (z^2)^\top, \dots, (z^p)^\top)^\top \in \mathbb{R}^{m_1+m_2+\cdots+m_p}$. For a given cone $C \subseteq \mathbb{R}^m$, C^d denotes the dual cone defined by $C^d := \{z \in \mathbb{R}^m \mid z^\top w \geq 0,\ \forall w \in C\}$. For vectors $z \in \mathbb{R}^m$ and $w \in \mathbb{R}^m$, the conic complementarity condition, $z^\top w = 0,\ z \in C$ and $w \in C^d$, is also written as $C \ni z \perp w \in C^d$. For a nonemptyset $D \subseteq \mathbb{R}^m$ and a function $h : \mathbb{R}^m \to \mathbb{R}$, $\mathrm{argmin}_{z\in D} h(z)$ denotes the set of minimizers of h over D. In addition, for $z \in \mathbb{R}^m$ and $\delta > 0$, $B(z, \delta) \subseteq \mathbb{R}^m$ denotes the closed ball with center z and radius δ, i.e., $B(z, \delta) := \{w \in \mathbb{R}^m \mid \|w - z\| \leq \delta\}$.

2 KKT conditions for SICP

In this section, we do not assume the convexity of objective function and constraint functions in (1.1). We focus on the following semi-infinite conic program (SICP):

$$\begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && g(x,t) \in C \quad \text{for all } t \in T, \end{aligned} \tag{2.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function, $g : \mathbb{R}^n \times T \to \mathbb{R}^m$ is a continuous function such that $g(\cdot, t)$ is differentiable for each fixed t, $T \subseteq \mathbb{R}^\ell$ is a compact set and $C \subseteq \mathbb{R}^m$ is a closed convex cone with nonempty interior. Notice that SICP (2.1) includes SICP (1.1) as a special case. If f is convex and g is affine with respect to x, then the local optimality analyses in this section hold in the global sense. The goal of this section is to derive the Karush-Kuhn-Tucker (KKT) conditions for SICP (2.1).

When $m = 1$ and $C = \mathbb{R}_+$, SICP (2.1) reduces to the classical semi-infinite program with the KKT conditions given as follows.

Lemma 2.1. [18, Theorem 2] *Let $x^* \in \mathbb{R}^n$ be an arbitrary local optimum of SICP (2.1) with $C := \mathbb{R}_+$. Let $T_{\mathrm{act}}(x)$ be the set of active indices at $x \in \mathbb{R}^n$, i.e., $T_{\mathrm{act}}(x) := \{t \in T \mid g(x, t) = 0\}$. Suppose that the Mangasarian-Fromovitz constraint qualification (MFCQ) holds at x^*, i.e., there exists a vector $d \in \mathbb{R}^n$ such that $\nabla_x g(x^*, t)^\top d > 0$ for any $t \in T_{\mathrm{act}}(x^*)$. Then, there exist p indices $t_1, t_2, \dots, t_p \in T_{\mathrm{act}}(x^*)$ and Lagrange multipliers $\mu_1, \mu_2, \dots, \mu_p \geq 0$ such that $p \leq n$ and*

$$\begin{aligned} &\nabla f(x^*) - \sum_{i=1}^{p} \mu_i \nabla_x g(x^*, t_i) = 0, \\ &\mathbb{R}_+ \ni \mu_i \perp g(x^*, t_i) \in \mathbb{R}_+ \quad (i = 1, 2, \dots, p). \end{aligned}$$

In the above result, the MFCQ plays a key role. However, for SICP (2.1), it is difficult to apply the MFCQ in a straightforward manner. We therefore introduce the Robinson constraint qualification (RCQ), which is defined as follows:

Definition 2.2 (Robinson Constraint Qualification (RCQ))**.** *Let $x \in \mathbb{R}^n$ be a feasible point of SICP (2.1). Then, we say that the Robinson constraint qualification (RCQ) holds at x if there exists a vector $d \in \mathbb{R}^n$ such that*

$$g(x, t) + \nabla_x g(x, t)^\top d \in \operatorname{int} C \quad \textit{for all } t \in T. \tag{2.2}$$

When $m = 1$ and $C = \mathbb{R}_+$, the RCQ reduces to the MFCQ. When g is affine, i.e., $g(x,t) := A(t)^\top x - b(t)$, the RCQ holds at any feasible point if and only if the Slater constraint qualification holds, i.e., there exists $x_0 \in \mathbb{R}^n$ such that $A(t)^\top x_0 - b(t) \in \operatorname{int} C$ for all $t \in T$. For a detail of the RCQ, see [4]. The next proposition states that any closed convex cone is represented as the intersection of finitely or infinitely many halfspaces generated by a certain compact set.

Proposition 2.3. *Let $C \subsetneq \mathbb{R}^m$ be an arbitrary nonempty closed convex cone. Then,* (i) *there exists a nonempty compact set $S \subseteq \{s \in \mathbb{R}^m \mid \|s\| = 1\}$ such that*

$$C = \{x \mid s^\top x \geq 0, \ \ \forall s \in S\}. \tag{2.3}$$

Moreover, (ii) *we have $S \subseteq C^d$.*

Proof. We first show (i). For any $s \in \mathbb{R}^m$ with $s \neq 0$, define the halfspace $H(s) := \{x \in \mathbb{R}^m \mid s^\top x \geq 0\}$. In addition, let $S := \{s \in \mathbb{R}^m \mid \|s\| = 1, \ H(s) \supseteq C\}$. By [20, Corollary 11.7.1], we have $C = \bigcap_{s \in S} H(s)$. Therefore, it suffices to show the compactness of S. Since the boundedness is evident, we only show the closedness. Choose an arbitrary convergent sequence $\{s^k\} \subseteq S$ such that $\lim_{k\to\infty} s^k = s^*$ and let $z \in C$ be an arbitrary vector. Obviously, we have $\|s^k\| = 1$. Moreover, from $C = \bigcap_{s\in S} H(s) \subseteq H(s^k)$, we have $(s^k)^\top z \geq 0$ for all k. Therefore, letting $k \to \infty$, we obtain $\|s^*\| = 1$ and $(s^*)^\top z \geq 0$, which implies $z \in H(s^*)$. Since $z \in C$ was arbitrarily chosen, we have $C \subseteq H(s^*)$. Hence, we have $s^* \in S$.

Next, we show (ii). Choose $s \in S$ arbitrarily. From (2.3), we have $s^\top x \geq 0$ for all $x \in C$, which implies $s \in C^d$. □

By using this proposition, we reformulate SICP (2.1) as a standard semi-infinite program, whereby we can derive the KKT conditions.

Theorem 2.4. *Let $x^* \in \mathbb{R}^n$ be an arbitrary local optimum of SICP* (2.1)*. Suppose that the RCQ holds at x^*. Then, there exist p indices $t_1, t_2, \ldots, t_p \in T$ and Lagrange multipliers $y^1, y^2, \ldots, y^p \in \mathbb{R}^m$ such that $p \leq n$ and*

$$\nabla f(x^*) - \sum_{i=1}^{p} \nabla_x g(x^*, t_i) y^i = 0, \tag{2.4}$$

$$C^d \ni y^i \perp g(x^*, t_i) \in C \quad (i = 1, 2, \ldots, p). \tag{2.5}$$

Proof. By Proposition 2.3, there exists a nonempty compact set $S \subseteq \{s \in \mathbb{R}^m \mid \|s\| = 1\}$ such that SICP (2.1) is equivalent to the following semi-infinite program:

$$\begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && s^\top g(x,t) \geq 0 \ \text{ for all } (s,t) \in S \times T. \end{aligned} \tag{2.6}$$

Let $(S \times T)_{\text{act}}(x^*) := \{(s,t) \in S \times T \mid s^\top g(x^*, t) = 0\}$. If $(S \times T)_{\text{act}}(x^*) = \emptyset$, then we have (2.4) and (2.5) with $y^i = 0$ for all i. Therefore, we suppose $(S \times T)_{\text{act}}(x^*) \neq \emptyset$. We first show that the MFCQ holds for problem (2.6), i.e., there exists a vector $d \in \mathbb{R}^n$ such that

$$(\nabla_x g(x^*, t)s)^\top d > 0 \ \text{ for all } (s,t) \in (S \times T)_{\text{act}}(x^*). \tag{2.7}$$

By assumption, there exists a vector $d \in \mathbb{R}^n$ satisfying RCQ (2.2), i.e., $g(x^*, t) + \nabla_x g(x^*, t)^\top d \in \operatorname{int} C$ for all $t \in T$. By Proposition 2.3, we also have $0 \notin S \subseteq C^d$. Hence, we have $s^\top(g(x^*, t) + \nabla_x g(x^*, t)^\top d) > 0$ for all $(s,t) \in S \times T$, which implies (2.7). Therefore, d is a vector satisfying the MFCQ. Now, applying Lemma 2.1 to (2.6), we have p indices $(s^1, t_1), (s^2, t_2), \dots, (s^p, t_p) \in (S \times T)_{\rm act}(x^*)$ and the Lagrange multipliers $\mu_1, \mu_2, \dots, \mu_p \geq 0$ such that $p \leq n$ and

$$\nabla f(x^*) - \sum_{i=1}^{p} \mu_i \nabla_x g(x^*, t_i) s_i = 0, \tag{2.8}$$

$$\mathbb{R}_+ \ni \mu_i \perp (s^i)^\top g(x^*, t_i) \in \mathbb{R}_+ \quad (i = 1, 2, \dots, p). \tag{2.9}$$

Letting $y^i := \mu_i s^i$ for each i, we have from (2.9) that $0 = \mu_i s_i^\top g(x^*, t_i) = (y^i)^\top g(x^*, t_i)$. We also have $y^i \in C^d$ since $s^i \in S \subseteq C^d$ from Proposition 2.3 and $\mu_i \geq 0$. In addition, we also have $g(x^*, t_i) \in C$ since x^* is feasible to SICP (2.1). Thus, (2.8) and (2.9) yield (2.4) and (2.5), respectively. This completes the proof. □

Before closing this section, we provide a more enhanced theorem available to the case where C has a Cartesian structure, i.e., $C = C^1 \times \cdots \times C^h \subseteq \mathbb{R}^m = \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_h}$. Consider the following problem:

$$\begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & g_j(x, t^j) \in C^j \ \text{ for all } t^j \in T_j, \ \ j = 1, 2, \cdots, h, \end{array} \tag{2.10}$$

where $g_j : \mathbb{R}^n \times T_j \to \mathbb{R}^{m_j}$ is continuous, $g_j(\cdot, t^j)$ is differentiable for each fixed t^j, $T_j \subseteq \mathbb{R}^{\ell_j}$ is compact and $C^j \subseteq \mathbb{R}^{m_j}$ is a closed convex cone with nonempty interior for each j. Then, the following theorem holds.

Theorem 2.5. *Let $x^* \in \mathbb{R}^n$ be an arbitrary local optimum of SICP (2.10). Assume that the RCQ holds at x^*, i.e., there exists a vector $d \in \mathbb{R}^n$ such that*

$$g_j(x^*, t^j) + \nabla_x g_j(x^*, t^j)^\top d \in \operatorname{int} C^j \ \text{ for all } t^j \in T_j, \ j = 1, 2, \dots, h. \tag{2.11}$$

Then, there exist $j_1, j_2, \dots, j_p \in \{1, 2, \dots, h\}$[3] *and $(t_i^{j_i}, y_i^{j_i}) \in T_{j_i} \times \mathbb{R}^{m_{j_i}}$ for $i = 1, 2, \dots, p$ such that $p \leq n$ and*

$$\nabla f(x^*) - \sum_{i=1}^{p} \nabla_x g_{j_i}(x^*, t_i^{j_i}) y_i^{j_i} = 0 \tag{2.12}$$

$$(C^{j_i})^d \ni y_i^{j_i} \perp g_{j_i}(x^*, t_i^{j_i}) \in C^{j_i} \quad (i = 1, 2, \dots, p). \tag{2.13}$$

Proof. For each $j = 1, 2, \dots, h$, let $\tilde{t}^j \in \mathbb{R}^{\ell_j} \setminus T_j$ be an arbitrary point and $\tilde{T}_j$ be defined as $\tilde{T}_j := \{\tilde{t}^1\} \times \cdots \times \{\tilde{t}^{j-1}\} \times T_j \times \{\tilde{t}^{j+1}\} \times \cdots \times \{\tilde{t}^h\} \subseteq \mathbb{R}^{\ell_1 + \ell_2 + \cdots + \ell_h}$. Then we can easily see that $\tilde{T}_j \cap \tilde{T}_{j'} = \emptyset$ for any $j \neq j'$. Let

$$t := (t^1, t^2, \dots, t^h) \in \mathbb{R}^{\ell_1 + \ell_2 + \cdots + \ell_h}, \quad T := \bigcup_{j=1}^{h} \tilde{T}_j \subseteq \mathbb{R}^{\ell_1 + \ell_2 + \cdots + \ell_h}, \tag{2.14}$$

and define $g : \mathbb{R}^n \times T \to \mathbb{R}^{m_1 + m_2 + \cdots + m_h}$ by

$$g(x, t) := (\tilde{g}_1(x, t), \dots, \tilde{g}_h(x, t)), \tag{2.15}$$

where

$$\tilde{g}_j(x, t) := \begin{cases} g_j(x, t^j) & (t \in \tilde{T}_j) \\ \zeta^j & (t \notin \tilde{T}_j) \end{cases} \tag{2.16}$$

[3] Repeated choice of the same index is allowed in the set $\{j_1, j_2, \dots, j_p\}$.

and $\zeta^j \in \operatorname{int} C^j$ is an arbitrary vector. Then, the function g id continuous on $\mathbb{R}^n \times T$ and $g(\cdot, t)$ is differentiable for each $t \in T$. In particular, we have

$$\nabla_x \tilde{g}_j(x,t) := \begin{cases} \nabla_x g_j(x, t^j) & (t \in \tilde{T}_j) \\ 0 & (t \notin \tilde{T}_j) \end{cases}. \tag{2.17}$$

Then, T is nonempty and compact, and SICP (2.10) is equivalent to SICP (2.1) with $C = C^1 \times \cdots \times C^h$ and g defined by (2.15). By letting $d \in \mathbb{R}^n$ satisfy (2.11), we have

$$\tilde{g}_j(x^*, t) + \nabla_x \tilde{g}_j(x^*, t)^\top d = \begin{cases} g_j(x^*, t^j) + \nabla_x g_j(x^*, t^j)^\top d \in \operatorname{int} C^j & (t \in \tilde{T}_j) \\ \zeta^j \in \operatorname{int} C^j & (t \notin \tilde{T}_j) \end{cases}$$

for each $j = 1, 2, \ldots, h$, where the first case follows from (2.11) and the second one follows from (2.16), (2.17) and $\zeta^j \in \operatorname{int} C^j$. Therefore, we have $g(x^*, t) + \nabla g(x^*, t)^\top d \in \operatorname{int} C$ for all $t \in T$, which implies that the RCQ holds at x^* for SICP (2.1). Hence, by Theorem 2.4, there exist $p \le n$, $t_1, t_2, \ldots, t_p \in T$ and $y_1, y_2, \ldots, y_p \in \mathbb{R}^m$ such that

$$\nabla f(x^*) - \sum_{i=1}^{p} \nabla_x g(x^*, t_i) y_i = 0, \tag{2.18}$$

$$C^d \ni y_i \perp g(x^*, t_i) \in C \quad (i = 1, 2, \ldots, p). \tag{2.19}$$

Let $t_i := (t_i^1, t_i^2, \ldots, t_i^h) \in \mathbb{R}^{\ell_1 + \ell_2 + \cdots + \ell_h}$ and $y_i := (y_i^1, y_i^2, \ldots, y_i^h) \in \mathbb{R}^{m_1 + m_2 + \cdots + m_h}$ for $i = 1, 2, \ldots, p$. From (2.14), for each i, there exists $j_i \in \{1, 2, \ldots, h\}$ such that $t_i \in \tilde{T}_{j_i}$, i.e., $t_i^{j_i} \in T_{j_i}$. Then, we have

$$\begin{aligned} \sum_{i=1}^{p} \nabla_x g(x^*, t_i) y_i &= \sum_{i=1}^{p} \Big(\nabla_x \tilde{g}_1(x^*, t_i), \nabla_x \tilde{g}_2(x^*, t_i), \ldots, \nabla_x \tilde{g}_h(x^*, t_i) \Big) \begin{pmatrix} y_i^1 \\ \vdots \\ y_i^h \end{pmatrix} \\ &= \sum_{i=1}^{p} \nabla_x g_{j_i}(x^*, t_i^{j_i}) y_i^{j_i}, \end{aligned}$$

where the second equality follows from (2.16) and (2.17), which together with (2.18) implies (2.12). In the last, we show (2.13). From (2.19) and $C^d = (C^1)^d \times (C^2)^d \times \cdots \times (C^h)^d$, we have $(C^j)^d \ni y_i^j \perp \tilde{g}_j(x^*, t_i) \in C^j$ for $j = 1, 2, \ldots, h$, which together with $\tilde{g}_{j_i}(x^*, t_i) = g_{j_i}(x^*, t_i^{j_i})$ from (2.16) implies (2.13) for $i = 1, 2, \ldots, p$. The proof is complete. □

3 Explicit exchange method for SICP

In this section, we propose an explicit exchange method for solving SICP (1.1), and show its global convergence under the assumption that f is strictly convex.

3.1 Algorithm

The algorithm proposed in this section requires solving conic programs with *finitely* many constraints as subproblems. Let CP(T') be the relaxed problem of SICP (1.1) with T replaced by a finite subset $T' := \{t_1, t_2, \ldots, t_p\} \subseteq T$. Then, CP$(T')$ can be formulated as follows:

$$\mathrm{CP}(T') \qquad \begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & A(t_i)^\top x - b(t_i) \in C \ \ (i = 1, 2, \ldots, p). \end{array}$$

Note that an optimum x^* of $\mathrm{CP}(T')$ satisfies the following KKT conditions:

$$\nabla f(x^*) - \sum_{i=1}^{p} A(t_i) y_{t_i} = 0,$$
$$C^d \ni y_{t_i} \perp A(t_i)^\top x^* - b(t_i) \in C \;\; (i = 1, 2, \ldots, p),$$

where y_{t_i} is the Lagrange multiplier vector corresponding to the constraint $A(t_i)^\top x^* - b(t_i) \in C$ for each i.

Now, we propose the following algorithm.

Algorithm 1 (Explicit exchange method)

Step 0. Let $\{\gamma_k\} \subseteq \mathbb{R}_{++}$ be a positive sequence such that $\lim_{k\to\infty} \gamma_k = 0$. Choose a finite subset $T^0 := \{t_1^0, \ldots, t_\ell^0\} \subseteq T$ with $\ell \geq 0$ chosen arbitrarily[4], and an arbitrary vector $e \in \operatorname{int} C$. Solve $\mathrm{CP}(T^0)$ to obtain an optimum x^0. Set $k := 0$.

Step 1. Obtain x^{k+1} and T^{k+1} by the following steps.

Step 1-0 Set $r := 0$, $E^0 := T^k$, $v^0 := x^k$ and solve $\mathrm{CP}(E^0)$ to obtain an optimum v^0.

Step 1-1 Find a $t_{\mathrm{new}}^r \in T$ such that

$$A(t_{\mathrm{new}}^r)^\top v^r - b(t_{\mathrm{new}}^r) \notin -\gamma_k e + C. \tag{3.1}$$

If such a t_{new}^r does not exist, i.e.,

$$A(t)^\top v^r - b(t) \in -\gamma_k e + C \tag{3.2}$$

for all $t \in T$, then set $x^{k+1} := v^r$, $T^{k+1} := E^r$, and go to Step 2. Otherwise, let

$$\overline{E}^{r+1} := E^r \cup \{t_{\mathrm{new}}^r\},$$

and go to Step 1-2.

Step 1-2 Solve $\mathrm{CP}(\overline{E}^{r+1})$ to obtain an optimum v^{r+1} and the Lagrange multipliers y_t^{r+1} for $t \in \overline{E}^{r+1}$.

Step 1-3 Let $E^{r+1} := \{t \in \overline{E}^{r+1} \mid y_t^{r+1} \neq 0\}$. Set $r := r + 1$ and return to Step 1-1.

Step 2. If γ_k is sufficiently small, then terminate. Otherwise, set $k := k + 1$ and return to Step 1.

When C is a symmetric cone such as an SOC or a semidefinite cone, the most typical choice for the interior vector e is the identity element with respect to Euclidean Jordan algebra [5].[5] Moreover, in Step 1-2, we can apply existing methods such as the primal-dual interior point method, the regularized smoothing method, and so on [1, 6, 11, 14, 17].

Now, denote the optimal values of $\mathrm{CP}(T')$ and SICP (1.1) by $V(T')$ and $V(T)$, respectively. Since E^{r+1} just removes the constraints with zero Lagrange multipliers from $\overline{E}^{r+1}$, and the feasible region of $\mathrm{CP}(E^r)$ is larger than that of $\mathrm{CP}(\overline{E}^{r+1})$, we have

$$V(E^0) \leq V(\overline{E}^1) = V(E^1) \leq \cdots \leq V(E^r) \leq V(\overline{E}^{r+1}) = V(E^{r+1}) \leq \cdots \leq V(T) < \infty. \tag{3.3}$$

In the subsequent convergence analysis, we omit the termination condition in Step 2, so that the algorithm generates an infinite sequence $\{x^k\}$.

[4] $\ell = 0$ means that we do not choose anything from T.

[5] For example, if C is $\mathbb{R}_+$, $\mathcal{K}^m$ and $\mathcal{S}_+^m$, then the identity element is 1, $(1, 0, \ldots, 0)^\top \in \mathbb{R}^m$, and the $m \times m$ identity matrix, respectively.

3.2 Global convergence under strict convexity assumption

In the previous subsection, we proposed the explicit exchange method for solving SICP (1.1). In this subsection, we show that the algorithm generates a sequence converging to the optimal solution under the following assumption.

Assumption A. i) Function f is strictly convex over the feasible region of SICP (1.1). ii) In Step 1-2 of Algorithm 1, $\mathrm{CP}(\overline{E}^{r+1})$ is solvable for each r. iii) A generated sequence $\{v^r\}$ in every Step 1 of Algorithm 1 is bounded.

Notice that all statements i)–iii) automatically hold when f is strongly convex. Under Assumption A, we first show that the inner iterations within Step 1 do not repeat infinitely, which ensures that Algorithm 1 is well-defined. To prove this, we provide the following proposition stating that the distance between v^{r+1} and v^r does not tend to zero during the inner iterations in Step 1.

Proposition 3.1. *Suppose that Assumption A holds. Then, there exists a positive number $N > 0$ such that*

$$\|v^{r+1} - v^r\| \geq N\gamma_k$$

for any $r \geq 0$ and $k \geq 0$.

Proof. Denote $z(v,t) := A(t)^\top v - b(t)$ for simplicity. Due to the continuity of the matrix norm $\|A(t)\| := \max_{\|w\|=1} \|A(t)^\top w\|$ and the compactness of T, there exists a sufficiently large $M > 0$ such that $\|A(t)\| \leq M$ for any $t \in T$. Hence, we have

$$\|z(v^{r+1},t) - z(v^r,t)\| = \|A(t)^\top (v^{r+1} - v^r)\| \leq M\|v^{r+1} - v^r\| \tag{3.4}$$

for any $t \in T$.

We next show that $\|z(v^{r+1}, t^r_{\text{new}}) - z(v^r, t^r_{\text{new}})\|$ is bounded below by some positive number for any $r \geq 0$. Let $e \in \operatorname{int} C$ be the vector chosen in Step 0. Then, there exists a $\delta > 0$ such that $e + B(0,\delta) \subseteq C$. We therefore have

$$\begin{aligned} z(v^{r+1}, t^r_{\text{new}}) + B(0,\delta\gamma_k) &= -\gamma_k e + z(v^{r+1}, t^r_{\text{new}}) + \gamma_k\left(e + B(0,\delta)\right) \\ &\subseteq -\gamma_k e + C, \end{aligned} \tag{3.5}$$

where the inclusion follows since $e + B(0,\delta) \subseteq C$, $\gamma_k > 0$, $z(v^{r+1}, t^r_{\text{new}}) \in C$, and C is a convex cone[6]. From (3.1), we have $z(v^r, t^r_{\text{new}}) \notin -\gamma_k e + C$, which together with (3.5) implies that

$$\|z(v^{r+1}, t^r_{\text{new}}) - z(v^r, t^r_{\text{new}})\| \geq \delta\gamma_k. \tag{3.6}$$

Combining (3.4) and (3.6) with $N := \delta/M$, we obtain

$$\|v^{r+1} - v^r\| \geq \delta\gamma_k/M = N\gamma_k.$$

□

Theorem 3.2. *Suppose that Assumption A holds. Then, Step 1 of Algorithm 1 terminates in a finite number of iterations for each k.*

Proof. Suppose, for a contradiction, that the inner iterations in Step 1 do not terminate finitely for some outer iteration k. (In what follows, k is fixed.) Then, by Assumption A iii), there exist accumulation points v^* and v^{**} of $\{v^r\}$ such that $v^{r_j} \to v^*$ and $v^{r_j+1} \to v^{**}$ as $j \to \infty$. Moreover,

[6] When C is a convex cone, $\alpha x + \beta y \in C$ holds for any $x, y \in C$ and $\alpha, \beta \geq 0$.

we must have $v^* \neq v^{**}$ from Proposition 3.1. Denote $z_t^r := A(t)^\top v^r - b(t)$ for simplicity. Since v^r solves CP($\overline{E}^r$), it satisfies the following KKT conditions:

$$\nabla f(v^r) - \sum_{t \in \overline{E}^r} A(t) y_t^r = 0, \tag{3.7}$$

$$C^d \ni y_t^r \perp z_t^r \in C \ \ (t \in \overline{E}^r), \tag{3.8}$$

where y_t^r are the Lagrange multipliers. From (3.3), we have $f(v^1) \leq f(v^2) \leq \cdots \leq V(T) < +\infty$, which implies

$$\lim_{r \to \infty} \left(f(v^{r+1}) - f(v^r) \right) = 0. \tag{3.9}$$

Let $F_r := f(v^{r+1}) - f(v^r) - \nabla f(v^r)^\top (v^{r+1} - v^r)$. Then, we have

$$\begin{aligned} f(v^{r+1}) - f(v^r) &= F_r + \nabla f(v^r)^\top (v^{r+1} - v^r) \\ &= F_r + \left(\sum_{t \in \overline{E}^r} A(t) y_t^r \right)^\top (v^{r+1} - v^r) && (3.10) \\ &= F_r + \sum_{t \in \overline{E}^r} (y_t^r)^\top z_t^{r+1} - \sum_{t \in \overline{E}^r} (y_t^r)^\top z_t^r && (3.11) \\ &= F_r + \sum_{t \in \overline{E}^r} (y_t^r)^\top z_t^{r+1}, && (3.12) \end{aligned}$$

where (3.10) and (3.12) follow from (3.7) and (3.8), respectively and (3.11) follows from $z_t^r = A(t)^\top v^r - b(t)$ and $z_t^{r+1} = A(t)^\top v^{r+1} - b(t)$. Since f is convex, we have $F_r \geq 0$. In addition, since $y_t^r \in C^d$ and $z_t^{r+1} \in C$, we have $\sum_{t \in \overline{E}^r} (y_t^r)^\top z_t^{r+1} \geq 0$. Therefore, from (3.9) and (3.12), we have

$$0 = \lim_{r \to \infty} F_r = \lim_{j \to \infty} F_{r_j} = f(v^{**}) - f(v^*) - \nabla f(v^*)^\top (v^{**} - v^*). \tag{3.13}$$

However, this contradicts $v^* \neq v^{**}$ and the strict convexity of f. Hence, the inner iterations in Step 1 must terminate finitely. □

The next theorem shows the global convergence of Algorithm 1 under the strict convexity assumption.

Theorem 3.3. *Suppose that SICP (1.1) has a solution and Assumption A holds. Let x^* be the optimum, and $\{x^k\}$ be the sequence generated by Algorithm 1. Then, we have*

$$\lim_{k \to \infty} x^k = x^*.$$

Proof. We first show that $\{x^k\}$ is bounded. Let $X(\gamma) := \{x \in \mathbb{R}^n \mid A(t)^\top x - b(t) + \gamma e \in C, \ \forall t \in T\}$ and $L := \{x \in \mathbb{R}^n \mid f(x) \leq f(x^*)\}$. Since $x^k \in L \cap X(\gamma_k) \subseteq L \cap X(\overline{\gamma})$ with $\overline{\gamma} := \max_{k \geq 0} \gamma_k$, it suffices to show that $L \cap X(\gamma)$ is bounded for any $\gamma > 0$. By Proposition 2.3, there exists a compact set $S \subseteq \mathbb{R}^m$ such that $0 \notin S \subseteq C^d$ and

$$\begin{aligned} X(\gamma) &= \{x \in \mathbb{R}^n \mid s^\top \left(A(t)^\top x - b(t) + \gamma e \right) \geq 0 \, , \forall (s,t) \in S \times T\} \\ &= \{x \in \mathbb{R}^n \mid (e^\top s)^{-1} \left(s^\top b(t) - (A(t)s)^\top x \right) \leq \gamma \, , \forall (s,t) \in S \times T\} \\ &= \left\{ x \in \mathbb{R}^n \mid h(x) := \max_{(s,t) \in S \times T} (e^\top s)^{-1} \left(s^\top b(t) - (A(t)s)^\top x \right) \leq \gamma \right\}, \end{aligned}$$

where the second equality is valid since $e \in \operatorname{int} C$ and $0 \neq s \in S \subseteq C^d$ entail $\min_{s \in S} e^\top s > 0$. Notice that $h(x) < \infty$ from the compactness of $S \times T$ and continuity of $A(\cdot)$ and $b(\cdot)$. Therefore, function h is closed, proper and convex. Now, let $\overline{h} : \mathbb{R}^n \to (-\infty, +\infty]$ be defined as

$$\overline{h}(x) := \begin{cases} h(x) & (x \in L) \\ \infty & (x \notin L) \end{cases}.$$

Then $\overline{h}$ is also closed, proper and convex since L is convex. Notice that

$$L \cap X(\gamma) = \{x \in \mathbb{R}^n \mid \overline{h}(x) \leq \gamma\},$$

i.e., $L \cap X(\gamma)$ is a level set of $\overline{h}$. If a closed proper convex function has at least one compact level set, then its any nonempty level set is also compact [3]. Moreover, we have $L \cap X(0) = \{x^*\}$ since f is strictly convex. Therefore, $L \cap X(\gamma)$ is compact for any $\gamma \geq 0$.

We next show that $\lim_{k\to\infty} x^k = x^*$. Let $\bar{x}$ be an arbitrary accumulation point of $\{x^k\}$. Then, there exists a subsequence $\{x^{k_j}\} \subseteq \{x^k\}$ and $\{\gamma_{k_j}\} \subseteq \{\gamma_k\}$ such that $\lim_{j\to\infty} x^{k_j} = \bar{x}$ and $\lim_{j\to\infty} \gamma_{k_j} = 0$. For all j, we have $A(t)^\top x^{k_j} - b(t) + \gamma_{k_j} e \in C$ $(\forall t \in T)$ and $f(x^{k_j}) \leq f(x^*)$. Hence, letting j tend to ∞, we have

$$A(t)^\top \bar{x} - b(t) \in C \quad (\forall t \in T), \tag{3.14}$$

$$f(\bar{x}) \leq f(x^*) \tag{3.15}$$

from the continuity of f and the closedness of C. From (3.14), we have $f(\bar{x}) \geq f(x^*)$, which together with (3.15) implies $f(\bar{x}) = f(x^*)$. Therefore, $\bar{x}$ solves SICP (1.1). Since f is strictly convex, we must have $\bar{x} = x^*$. We thus have $\lim_{k\to\infty} x^k = x^*$. □

4 Regularized explicit exchange method for SICP

In the previous section, we proposed the explicit exchange method for SICP (1.1) and analyzed the convergence property. However, to ensure the global convergence, we had to assume the strict convexity of the objective function (Assumption A). In this section, we propose a new method combining the regularization technique with the explicit exchange method, and establish the global convergence without assuming the strict convexity.

4.1 Algorithm

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then, function $f_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ defined by $f_\varepsilon(x) := \frac{1}{2}\varepsilon\|x\|^2 + f(x)$ is strongly convex for any $\varepsilon > 0$. So, if we apply Algorithm 1 to the following regularized SICP:

$$\text{RSICP}(\varepsilon) \qquad \begin{array}{ll} \text{Minimize} & f_\varepsilon(x) \\ \text{subject to} & A(t)^\top x - b(t) \in C \ \text{ for all } t \in T, \end{array}$$

then Step 1 terminates in a finite number of (inner) iterations and the sequence generated by Algorithm 1 converges to the unique solution x^*_ε of RSICP(ε).

By introducing a positive sequence $\{\varepsilon_k\}$ converging to 0, we can expect that $x^*_{\varepsilon_k}$ converges to the solution of the original SICP (1.1) as k goes infinity. However, it requires too much computational cost if we solve RSICP(ε_k) exactly for every k. Therefore, in the regularized explicit exchange method, we solve RSICP(ε_k) inexactly by the explicit exchange method. In the inner iterations, we repeatedly solve finitely relaxed regularized problems of the form:

$$\text{CP}(\varepsilon_k, T') \qquad \begin{array}{ll} \text{Minimize} & f_{\varepsilon_k}(x) \\ \text{subject to} & A(t_i)^\top x - b(t_i) \in C \ \ (i = 1, 2, \dots, p), \end{array}$$

where $T' := \{t_1, t_2, \dots, t_p\} \subseteq T$. The detailed steps of the regularized explicit exchange method are described as follows.

Algorithm 2 (Regularized Explicit Exchange Method)

Step 0. Choose positive sequences $\{\gamma_k\} \subseteq \mathbb{R}_{++}$ and $\{\varepsilon_k\} \subseteq \mathbb{R}_{++}$ such that $\lim_{k\to\infty} \gamma_k = \lim_{k\to\infty} \varepsilon_k = 0$. Choose a finite subset $T^0 := \{t_1^0, \ldots, t_l^0\} \subseteq T$ with $\ell \geq 0$ chosen arbitrarily[7]. Moreover, choose $e \in \operatorname{int} C$ arbitrarily. Set $k := 0$.

Step 1. Obtain x^{k+1} and T^{k+1} by the following procedure.

Step 1-0 Set $r := 0$ and $E^0 := T^k$. Solve $\mathrm{CP}(\varepsilon_k, E^0)$ and let v^0 be an optimum.

Step 1-1 Find $t^r_{\mathrm{new}} \in T$ such that

$$A(t^r_{\mathrm{new}})^\top v^r - b(t^r_{\mathrm{new}}) \notin -\gamma_k e + C. \tag{4.1}$$

If such a t^r_{new} does not exist, i.e.,

$$A(t)^\top v^r - b(t) \in -\gamma_k e + C \tag{4.2}$$

for any $t \in T$, then set $x^{k+1} := v^r$ and $T^{k+1} := E^r$, and go to Step 2. Otherwise, let

$$\overline{E}^{r+1} := E^r \cup \{t^r_{\mathrm{new}}\},$$

and go to Step 1-2.

Step 1-2 Solve $\mathrm{CP}(\varepsilon_k, \overline{E}^{r+1})$ to obtain an optimum v^{r+1} and the Lagrange multipliers y_t^{r+1} for $t \in \overline{E}^{r+1}$.

Step 1-3 Let $E^{r+1} := \{t \in \overline{E}^{r+1} \mid y_t^{r+1} \neq 0\}$. Set $r := r + 1$ and return to Step 1-1.

Step 2. If γ_k and ε_k are sufficiently small, then terminate. Otherwise, set $k := k + 1$ and return to Step 1.

In the next convergence analysis, we omit the termination check in Step 2.

4.2 Global convergence without strict convexity assumption

In this section, we show the global convergence of Algorithm 2 for SICP (1.1) without the strict convexity assumption. Indeed, we only need the following assumption for the proof of the convergence.

Assumption B. Function f is convex. Moreover, the Slater constraint qualification (SCQ) holds for SICP (1.1), i.e., there exists an $x_0 \in \mathbb{R}^n$ such that $A(t)^\top x_0 - b(t) \in \operatorname{int} C$ for all $t \in T$.

Notice that, for SICP (1.1), the SCQ holds if and only if any feasible point satisfies the RCQ as studied in Section 2. We first show that Step 1 terminates finitely.

Proposition 4.1. *Suppose that Assumption B holds. Then, Step 1 terminates finitely.*

Proof. By Theorem 3.2, it suffices to show that Assumption A holds when Step 1 of Algorithm 1 is applied to $\mathrm{RSICP}(\varepsilon)$ for any $\varepsilon > 0$. Since Assumptions A i) and ii) hold from the strong convexity of f_ε, we only show Assumption A iii). Let x^*_ε be an optimum of $\mathrm{RSICP}(\varepsilon)$ and $L^*_\varepsilon := \{x \in \mathbb{R}^n \mid f_\varepsilon(x) \leq f_\varepsilon(x^*_\varepsilon)\}$. Then, L^*_ε is compact since f_ε is strongly convex. Moreover, we have $v^r \in L^*_\varepsilon$, i.e., $f_\varepsilon(v^r) \leq f_\varepsilon(x^*_\varepsilon)$ for all r since $\overline{E}^r \subseteq T$. Hence, $\{v^r\}$ is bounded. □

Now, we show that under Assumption B the generated sequence $\{x^k\}$ is bounded, and Algorithm 2 is globally convergent in the sense that the distance from x^k to the solution set of SICP (1.1) tends to 0.

[7] $\ell = 0$ means that we don't choose anything from T.

Theorem 4.2. *Suppose that Assumption B holds. Let $\{x^k\}$ be the sequence generated by Algorithm 2. Then, the following statements hold.*

i) *If $\{\varepsilon_k\}$ and $\{\gamma_k\}$ are chosen in Step 0 so that $\gamma_k = O(\varepsilon_k)$, then $\{x^k\}$ is bounded.*

ii) *Any accumulation point of $\{x^k\}$ solves SICP (1.1).*

Proof. i) Let $x^* \in \mathbb{R}^n$ be an arbitrary solution of SICP (1.1). Since Assumption B holds, Theorem 2.4 holds for SICP (1.1), i.e., there exist $t_1, t_2, \ldots, t_p \in T$ and $y^1, y^2, \ldots, y^p \in \mathbb{R}^m$ such that $p \le n$ and

$$\nabla f(x^*) - \sum_{i=1}^{p} A(t_i) y^i = 0, \tag{4.3}$$

$$C^d \ni y^i \perp A(t_i)^\top x^* - b(t_i) \in C \quad (i = 1, 2, \ldots, p). \tag{4.4}$$

Let $\{x^k\}$ be the sequence generated by Algorithm 2. Since x^k solves $\mathrm{CP}(\varepsilon_{k-1}, T^k)$ and x^* is feasible to $\mathrm{CP}(\varepsilon_{k-1}, T^k)$, we have

$$\frac{1}{2}\varepsilon_{k-1}\|x^k\|^2 + f(x^k) \le \frac{1}{2}\varepsilon_{k-1}\|x^*\|^2 + f(x^*). \tag{4.5}$$

Multiplying both sides of (4.5) by $2/\varepsilon_{k-1}$, we have

$$\begin{aligned} \|x^k\|^2 &\le \|x^*\|^2 - \frac{2}{\varepsilon_{k-1}}(f(x^k) - f(x^*)) \\ &\le \|x^*\|^2 - \frac{2}{\varepsilon_{k-1}} \nabla f(x^*)^\top (x^k - x^*) \\ &= \|x^*\|^2 - \frac{2}{\varepsilon_{k-1}} \left(\sum_{i=1}^{p} A(t_i) y^i \right)^\top (x^k - x^*), \end{aligned} \tag{4.6}$$

where the second inequality holds since f is convex, and the equality follows from (4.3). Moreover, the last term of (4.6) satisfies the following inequalities:

$$\begin{aligned} &- \left(\sum_{i=1}^{p} A(t_i) y^i \right)^\top (x^k - x^*) \\ &= - \sum_{i=1}^{p} (y^i)^\top (A(t_i)^\top x^k - b(t_i) + \gamma_{k-1} e) + \sum_{i=1}^{p} (y^i)^\top (\gamma_{k-1} e) + \sum_{i=1}^{p} (y^i)^\top (A(t_i)^\top x^* - b(t_i)) \\ &\le \sum_{i=1}^{p} (y^i)^\top (\gamma_{k-1} e) \\ &\le p\mu \|e\| \gamma_{k-1}, \end{aligned} \tag{4.7}$$

where $\mu := \max\{\|y^1\|, \|y^2\|, \ldots, \|y^p\|\}$, and the first inequality follows since (4.2) and (4.4) imply $y^i \in C^d$, $A(t_i)^\top x^k - b(t_i) + \gamma_{k-1} e \in C$ and $(y^i)^\top (A(t_i)^\top x^* - b(t_i)) = 0$. Then, by substituting (4.7) into (4.6), we have

$$\|x^k\|^2 \le \|x^*\|^2 + 2p\mu \|e\| \gamma_{k-1}/\varepsilon_{k-1}. \tag{4.8}$$

Since $\gamma_{k-1} = O(\varepsilon_{k-1})$, $\{\gamma_{k-1}/\varepsilon_{k-1}\}$ is bounded, and hence $\{x^k\}$ is also bounded.

ii) Let $\bar{x}$ be an arbitrary accumulation point of $\{x^k\}$. Then, taking a subsequence if necessary, we have

$$x^k \to \bar{x}, \qquad \varepsilon_{k-1} \to 0, \qquad \gamma_{k-1} \to 0 \qquad (k \to \infty).$$

First, we show that $\bar{x}$ is feasible to SICP (1.1). Since x^k is determined as v^r satisfying (4.2) with γ_k replaced by γ_{k-1}, $A(t)^\top x^k - b(t) + \gamma_{k-1} e \in C$ holds for any $t \in T$. Noticing that C is closed, we have $\lim_{k\to\infty} A(t)^\top x^k - b(t) + \gamma_{k-1}e = A(t)^\top \bar{x} - b(t) \in C$ for any $t \in T$. Hence, $\bar{x}$ is feasible to SICP (1.1).

We next show that $\bar{x}$ is optimal to SICP (1.1). Let x^* be an arbitrary optimum of SICP (1.1). Since $\bar{x}$ is feasible to SICP (1.1), we have $f(\bar{x}) \geq f(x^*)$. On the other hand, x^* is feasible to $\mathrm{CP}(\varepsilon_{k-1}, E_k)$ since the feasible region of SICP (1.1) is contained in that of $\mathrm{CP}(\varepsilon_{k-1}, E_k)$. Hence, we have

$$\frac{1}{2}\varepsilon_{k-1}\|x^k\|^2 + f(x^k) \leq \frac{1}{2}\varepsilon_{k-1}\|x^*\|^2 + f(x^*). \tag{4.9}$$

Due to the continuity of f, by letting $k \to \infty$ in (4.9), we have $f(\bar{x}) \leq f(x^*)$. Therefore, we obtain $f(x^*) = f(\bar{x})$, which implies that $\bar{x}$ solves SICP (1.1). □

From the above theorem, we can see that if we choose $\{\varepsilon_k\}$ and $\{\gamma_k\}$ so that $\gamma_k = O(\varepsilon_k)$, then the generated sequence $\{x^k\}$ has an accumulation point and it solves SICP (1.1). Moreover, we can show that, if $\{\varepsilon_k\}$ and $\{\gamma_k\}$ are chosen so that $\gamma_k = o(\varepsilon_k)$, $\{x^k\}$ is actually convergent and its limit point is the least 2-norm solution.

Theorem 4.3. *Suppose that Assumption B holds. Let $\{\varepsilon_k\}$ and $\{\gamma_k\}$ be chosen such that $\gamma_k = o(\varepsilon_k)$, and $\{x^k\}$ be a sequence generated by Algorithm 2. Let $S^* \subseteq \mathbb{R}^n$ denote the nonempty solution set of SICP (1.1) and $x^* \in \mathbb{R}^n$ be the least 2-norm solution, i.e., $x^*_{\min} := \mathrm{argmin}_{x\in S^*}\|x\|$. Then, $\lim_{k\to\infty} x^k = x^*_{\min}$.*

Proof. By Theorem 4.2, $\{x^k\}$ is bounded and every accumulation point belongs to S^*. Moreover, $x^*_{\min}$ can be identified uniquely since S^* is closed and convex. Therefore, it suffices to show that $\|\bar{x}\| = \|x^*_{\min}\|$ for any accumulation point $\bar{x}$ of $\{x^k\}$.

Now, let $\bar{x}$ be an arbitrary accumulation point of $\{x^k\}$. By (4.8) in the proof of Theorem 4.2 ii), we have

$$\|x^k\|^2 \leq \|x^*\|^2 + 2p\mu\|e\|\gamma_{k-1}/\varepsilon_{k-1}. \tag{4.10}$$

Since $\gamma_k = o(\varepsilon_k)$, by letting $k \to \infty$, we obtain $\|\bar{x}\| \leq \|x^*_{\min}\|$. On the other hand, we also have $\|\bar{x}\| \geq \|x^*_{\min}\|$ since $\bar{x} \in S^*$ and $x^*_{\min} = \mathrm{argmin}_{x\in S^*}\|x\|$. We thus have $\|\bar{x}\| = \|x^*_{\min}\|$. □

5 Numerical experiments

In this section, we report some numerical results. The program was coded in Matlab 2008a and run on a machine with an Intel®Core2 Duo E6850 3.00GHz CPU and 4GB RAM. In this experiment, we consider the SICP with a linear objective function and infinitely many second-order cone constraints with respect to a single second-order cone. Actual implementation of Algorithm 2 was carried out as follows. In Step 0, we set $e := (1, 0, \ldots, 0)^\top \in \mathrm{int}\,\mathcal{K}^m$. In Step 1-1, to find t^r_{new} satisfying (4.1), we first check the values of $\lambda\left(A(t)^\top v^r - b(t) + \gamma_k e\right)$ for $t = -1, -0.98, -0.96, \ldots, 0.98, 1$, where $\lambda(\cdot)$ denotes the spectral value of $z \in \mathbb{R}^m$ [6, 11] defined by

$$\lambda(z) := z_1 - \sqrt{z_2^2 + z_3^2 + \cdots + z_m^2}.$$

If we find a $\bar{t} \in \{-1, -0.98, -0.96, \ldots, 0.98, 1\}$ such that $\lambda\left(A(\bar{t})^\top v^r - b(\bar{t}) + \gamma_k e\right) < 0$, then we set $t^r_{\mathrm{new}} := \bar{t}$.[8] Otherwise, we solve

$$\begin{aligned} &\text{Minimize} \;\; \lambda(A(t)^\top v^r - b(t) + \gamma_k e) \\ &\text{subject to} \;\; t \in [-1, 1], \end{aligned} \tag{5.1}$$

[8] Notice that $\lambda\left(A(t)^\top x - b(t) + \gamma e\right) \geq 0$ if and only if $A(t)^\top x - b(t) \in -\gamma e + \mathcal{K}^m$ for any $x \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$.

and check the nonnegativity of its optimal value. To solve (5.1), we choose the initial point $t_0 := \operatorname{argmin}\{\lambda\left(A(t)^\top v^r - b(t) + \gamma_k e\right) \mid t = -1, -0.98, -0.96, \ldots, 0.98, 1\}$ and apply Newton's method combined with the bisection method. In Step 1-2, we solve $\mathrm{CP}(\varepsilon, T')$ by the smoothing method [6, 11]. In Step 1-3, we regard y_t^r as 0 if $\|y_t^r\| \leq 10^{-12}$. In Step 2, we terminate the algorithm if $\max(\varepsilon_k, \gamma_k) \leq 10^{-5}$.

Experiment 1

In the first experiment, we solve the following SICP:

$$\begin{aligned} &\text{Minimize} && c^\top x \\ &\text{subject to} && A(t)^\top x - b(t) \in \mathcal{K}^m \text{ for all } t \in [-1, 1], \end{aligned} \tag{5.2}$$

where $\mathcal{K}^m := \{(x_1, x_2, \ldots, x_m)^\top \in \mathbb{R}^m \mid x_1 \geq \|(x_2, x_3, \ldots, x_m)^\top\|\}$, $c \in \mathbb{R}^n$, $A(t) := (A_{ij}(t)) \in \mathbb{R}^{n\times m}$ with $A_{ij}(t) := \alpha_{ij0} + \alpha_{ij1}t + \alpha_{ij2}t + \alpha_{ij3}t^3$ $(i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, m)$ and $b(t) := (b_j(t)) \in \mathbb{R}^m$ with $b_1(t) := -\sum_{j=2}^m \sum_{\ell=0}^3 |\beta_j^\ell|$ and $b_j(t) := \beta_{j0} + \beta_{j1}t + \beta_{j2}t^2 + \beta_{j3}t^3$ $(j = 2, \ldots, m)$. We choose $\alpha_{ijk}, \beta_{j\ell}$ $(i = 1, 2, \ldots, n,\ j = 2, \ldots, m,\ k = 0, 1, 2, 3,\ \ell = 0, 1, 2, 3)$ and all components of c randomly from $[-1, 1]$. Note that by the choice of $b_1(t)$, feasibility of (5.2) is ensured.[9] In this way, we generate two sets of data $A(t)$, $b(t)$ and c for each of the three pairs $(m, n) = (25, 15)$, $(15, 15)$ and $(10, 15)$, thereby obtaining six problems denoted by Problems $1, 2, \ldots, 6$.

In this experiment, using parameters $\{\varepsilon_k\}$ and $\{\gamma_k\}$ such that $\varepsilon_k = 0.5^k, \gamma_k = 0.3^k$, and the initial set $T^0 := \{-1, 0, 1\}$ in Step 0, we observe the convergence behavior of the algorithm. The results are shown in Table 1, where each column represents the following:

- $\mathrm{ite}_{\mathrm{out}}$: the number of outer iterations,
- $\{\bar{r}_k\}$: the values of $\bar{r}_k$ for $k = 0, 1, \ldots, \mathrm{ite}_{\mathrm{out}} - 1$, where $\bar{r}_k$ denotes the value of r when the inner termination criterion (4.2) is satisfied at the k-th outer iteration,
- $\bar{r}_{\mathrm{sum}}$: the sum of $\bar{r}_k$'s for all $k = 0, 1, 2, \ldots, \mathrm{ite}_{\mathrm{out}} - 1$,
- t_{socp} : the number of times the sub-SOCPs ($\mathrm{CP}(\varepsilon_k, E^0)$ and $\mathrm{CP}(\varepsilon_k, \overline{E}^{r+1})$) are solved,
- T_{fin} : the values of T^k when the algorithm terminates,
- time(sec) : the CPU time in seconds

In the column of $\bar{r}_k$, p^q means that we had $\bar{r}_k = p$ in q consecutive iterations. For example, $0^{10}, 2, 1^4$ means that $\bar{r}_k = 0$ $(k = 0, 1, \ldots, 9), \bar{r}_{10} = 2$ and $\bar{r}_k = 4$ $(k = 11, 12, 13, 14)$. Notice that we always have $t_{\mathrm{socp}} = \mathrm{ite}_{\mathrm{out}} + \bar{r}_{\mathrm{sum}}$, since we solve sub-SOCPs once at Step 1-0 and $\bar{r}_k$ times at Step 1-3, for each k. Although T_{fin} usually represents an approximate active index set at the optimum, the real active index set is $\{-1, 1\}$ for Problems 2 and 3. This is because the inner termination criterion (4.2) was always satisfied with $r = 0$ and therefore the inactive index set $t = 0$ has never been removed at Step 1-3. From the columns of $\bar{r}_k$, we can see that $\bar{r}_k$ was sometimes large for $k \leq 4$, but it was always 0 or 1 for $k = 7, 8, \ldots, 17$. This fact suggests that T_{fin} is usually obtained in the early stage of iterations.

Experiment 2

In the second experiment, we implement the non-regularized exchange method (Algorithm 1) as well as the regularized exchange method (Algorithm 2), and compare their performances. In both

[9] Note that the origin always lies in the interior of the feasible region, since we have $-b(t) \in \operatorname{int}\mathcal{K}^m$ from $-b_1(t) - \|(-b_2(t), \ldots, -b_m(t))^\top\| > 0$ for all $t \in [-1, 1]$.

methods, the initial index set T^0 is set to be $T_a^0 := \{-1, -0.5, 0, 0.5, 1\}, T_b^0 := \{-1, 0, 1\}$, or $T_c^0 = \{-0.5, 0, 0.5\}$. The parameters are chosen as $\gamma_k = 0.5^k$ for Algorithm 1, and $\varepsilon_k = \gamma_k = 0.5^k$ for Algorithm 2. Both methods are applied to the same problems as in Experiment 1.

Table 2 shows the obtained results, where $t_{\rm socp}^a, t_{\rm socp}^b$ and $t_{\rm socp}^c$ denote the values of $t_{\rm socp}$ for the initial index sets T_a^0, T_b^0 and T_c^0, respectively, and "F" means that we failed to solve a problem. From the table, we can observe that $t_{\rm socp}$ for the non-regularized method is much less than $t_{\rm socp}$ for the regularized method. This is due to the fact that the regularized exchange method has to solve the sub-SOCP (CP(ε_k, E^0)) at least once for every outer iteration, whereas the non-regularized exchange method does not need to solve it when the inner termination criterion (3.2) is satisfied for $r = 0$. However, as shown in Sections 3 and 4, the convergence of the non-regularized exchange method is not guaranteed theoretically since the objective function is linear. Indeed, the non-regularized exchange method failed to solve Problems 1, 4 and 5 with $T^0 = T_c^0$ and Problem 6 with $T^0 = T_b^0$ and T_c^0, since CP(T_b^0) for Problems 1, 4, 5, 6 and CP(T_c^0) for Problem 6 have no solutions. On the other hand, the regularized exchange method succeeded in solving all problems for any choice of T^0. This is the main advantage of the regularized exchange method.

Problem	(m, n)	ite$_{\rm out}$	$\{\bar{r}_k\}$	$\bar{r}_{\rm sum}$	$t_{\rm socp}$	$T_{\rm fin}$	time(sec)
1	(25, 15)	18	$0^6, 4, 0^8, 1, 0^2$	5	23	$\{-1, -0.296, 1\}$	5.57
2	(25, 15)	18	0^{18}	0	18	$\{-1, 0, 1\}$	2.41
3	(15, 15)	18	0^{18}	0	18	$\{-1, 0, 1\}$	1.84
4	(15, 15)	18	$0^3, 11, 0^2, 1, 0^{11}$	14	32	$\{-1, -0.2, -0.18, 1\}$	12.49
5	(10, 15)	18	$0^2, 13, 0, 3, 0, 3, 0, 1, 0^9$	20	38	$\{-1, -0.48, -0.46, 1\}$	3.83
6	(10, 15)	18	$0^2, 7, 4, 6, 2^2, 0^5, 1, 0^3, 1, 0$	23	41	$\{-1, -0.387, 0.25, 1\}$	12.91

Table 1: Convergence behavior for Experiment 1

		regularized			non-regularized		
Problem	(m, n)	$t_{\rm socp}^a$	$t_{\rm socp}^b$	$t_{\rm socp}^c$	$t_{\rm socp}^a$	$t_{\rm socp}^b$	$t_{\rm socp}^c$
1	(25, 15)	23	23	34	5	5	F
2	(25, 15)	18	18	20	1	1	11
3	(15, 15)	18	18	25	1	1	15
4	(15, 15)	27	28	44	4	4	F
5	(10, 15)	19	24	29	4	5	F
6	(10, 15)	28	30	46	8	F	F

Table 2: Comparison of regularized and non-regularized exchange methods

Experiment 3

In the third experiment, we apply Algorithm 2 to Chebyshev-like approximation problems for vector-valued functions.

Experiment 3-1 We first focus on the complex Chebyshev approximation, which appears in various fields such as the filter design [13] and so on [2, 7]. Let the complex functions $G : [0, 2\pi] \to \mathbb{C}$ and $g : \mathbb{C}^\ell \times [0, 2\pi] \to \mathbb{C}$ be defined by

$$G(t) := \frac{1}{\cos t - 1 + i(\sin t - 1)}, \quad g(z,t) := \sum_{\nu=1}^{\ell} z_\nu (\cos t + i \sin t)^{\nu-1},$$

respectively, where $i := \sqrt{-1}$ and $z := (z_1, z_2, \dots, z_\ell)^\top \in \mathbb{C}^\ell$. Then, we aim to find a $z \in \mathbb{C}^\ell$ such that $g(z,t)$ approximates $G(t)$ over $[0, 2\pi]$, that is, to solve the following unconstrained minimization problem:

$$\underset{z \in \mathbb{C}^\ell}{\text{Minimize}} \ \max_{t \in [0,2\pi]} |G(t) - g(z,t)|. \tag{5.3}$$

Introducing an auxiliary variable $v \in \mathbb{R}$ and real vectors $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ with $z = x + iy$, we can transform (5.3) into the following SICP with infinitely many three-dimensional second-order cones:

$$\begin{aligned} &\underset{v,x,y}{\text{Minimize}} \quad v \\ &\text{subject to} \quad \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & B_1(t) & B_2(t) & \cdots & B_\ell(t) \end{pmatrix} \begin{pmatrix} v \\ x_1 \\ y_1 \\ \vdots \\ x_\ell \\ y_\ell \end{pmatrix} - \begin{pmatrix} 0 \\ b_1(t) \\ b_2(t) \end{pmatrix} \in \mathcal{K}^3 \\ &\qquad\qquad \text{for all } t \in [0, 2\pi], \end{aligned} \tag{5.4}$$

where

$$\begin{aligned} B_\nu(t) &:= \begin{pmatrix} \cos(\nu-1)t & -\sin(\nu-1)t \\ \sin(\nu-1)t & \cos(\nu-1)t \end{pmatrix} \in \mathbb{R}^{2\times 2}, \quad \nu = 1, 2, \dots, \ell, \\ b_1(t) &:= \frac{\cos t - 1}{(\cos t - 1)^2 + (\sin t - 1)^2} \in \mathbb{R}, \\ b_2(t) &:= \frac{-\sin t + 1}{(\cos t - 1)^2 + (\sin t - 1)^2} \in \mathbb{R}. \end{aligned}$$

We apply Algorithm 2 to SICP (5.3) with $\ell = 3, 5, 7, 9$. In Step 0, we set $T_0 := \{0, \pi\}$ and $\varepsilon_k = \gamma_k := 0.5^k$. The results are shown in Table 3 and Figure 1. Table 3 shows that all problems could be solved in acceptable time, and about two sub-SOCPs were solved on average at each outer iteration k. Figure 1 represents the values of $\log_{10} |v^{\text{exact}} - v^k|$ for each ℓ and k, where v^{exact} denotes the exact optimal value of v for SICP (5.4). In fact, it is known that the value of v^{exact}, which equals the optimal value of (5.3), is explicitly given by $2^{(1-\ell)/2}$ [2]. From the figure, we can observe that v^k gets sufficiently close to v^{exact} much before the termination criterion $\max(\gamma_k, \varepsilon_k) \le 10^{-5}$ is satisfied. Indeed, we have $|v^{\text{exact}} - v^k| \le 10^{-9}$ for all $k \ge 10$ with $\ell = 3, 5, 7, 9$. This fact suggests that the termination criterion employed in the experiment still has a room to be improved.

ℓ	ite_{out}	t_{socp}	time(sec)
3	18	27	1.30
5	18	32	2.24
7	18	37	5.15
9	18	36	8.03

Table 3: Results for Experiment 3-1

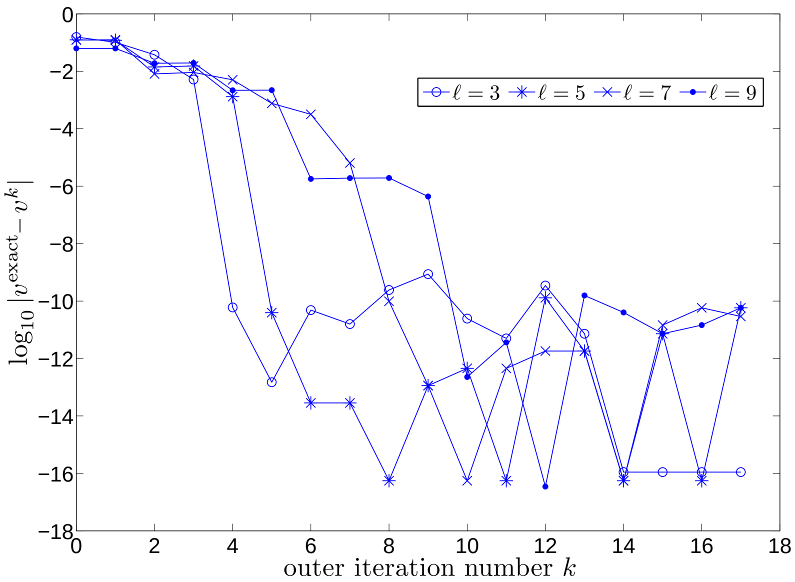

Figure 1: The values of $\log_{10}|v^{\text{exact}} - v^k|$ for each k

Experiment 3-2. We next consider the vector-valued approximation problem with respect to $H : \mathbb{R} \to \mathbb{R}^3$ and $h : \mathbb{R}^8 \times \mathbb{R} \to \mathbb{R}^3$ defined by

$$H(t) = \begin{pmatrix} e^{t^2} \\ 2te^{t^2} \\ (4t^2+2)e^{t^2} \end{pmatrix}, \; h(u,t) := \sum_{\nu=1}^{8} \begin{pmatrix} u_\nu t^{\nu-1} \\ (\nu-1)u_\nu t^{\nu-2} \\ (\nu-1)(\nu-2)u_\nu t^{\nu-3} \end{pmatrix}.$$

In order to find a $u \in \mathbb{R}^8$ such that $h(u,t) \approx H(t)$ over $t \in [-1,1]$, we solve the following problem:

$$\underset{u\in\mathbb{R}^8}{\text{Minimize}} \quad \max_{t\in[-1,1]} \; \|H(t) - h(u,t)\|. \tag{5.5}$$

Introducing an auxiliary variable $v \in \mathbb{R}$, we can reformulate (5.5) as the following SICP with infinitely many second-order cone constraints:

$$\begin{array}{ll} \underset{v,u}{\text{Minimize}} & v \\ \text{subject to} & \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & t & t^2 & \cdots & t^7 \\ 0 & 0 & 1 & 2t & \cdots & 7t^6 \\ 0 & 0 & 0 & 2 & \cdots & 42t^5 \end{pmatrix} \begin{pmatrix} v \\ u \end{pmatrix} - \begin{pmatrix} 0 \\ e^{t^2} \\ 2te^{t^2} \\ (4t^2+2)e^{t^2} \end{pmatrix} \in \mathcal{K}^4 \\ & \text{for all } t \in [-1,1]. \end{array} \tag{5.6}$$

In applying Algorithm 2, we set $T_0 := \{-1, 1\}$ and $\varepsilon_k = \gamma_k := 0.5^k$. Then, the algorithm outputs the solution $v^* = 0.1415$, $u^* = (0.9948, 0.0000, 1.0707, 0.0000, 0.3083, 0.0000, 0.3442, 0.0000)^\top)$ together with $T_{\text{fin}} = \{-1.00, -0.88, -0.52, 0, 0.52, 0.88, 1.00\}$. Notice that we have $u_2^* = u_4^* = u_6^* = u_8^* = 0$. This is reasonable since $H_1(t)$ and $H_3(t)$ are even functions whereas $H_3(t)$ is an odd function. Figure 2 shows the graph of $\|H(t) - h(u^*,t)\|$ over $t \in [-1,1]$. By the graph, we can observe that the values of $\|H(t) - h(u^*,t)\|$ is bounded above by $v^* = 0.1415$, and the bound is attained at multiple points in $[-1,1]$. Actually, those points coincide with $T_{\text{fin}} = \{-1.00, -0.88, -0.52, 0, 0.52, 0.88, 1.00\}$, which correspond to the active constraints at the optimum.

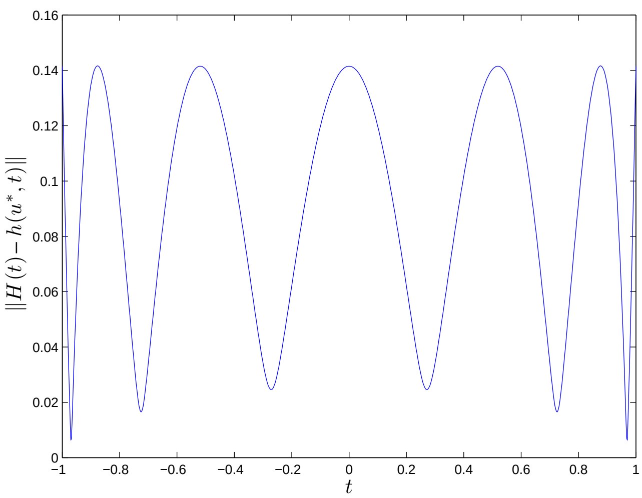

Figure 2: The graph of $\|H(t) - h(u^*,t)\|$ over $t \in [-1,1]$

6 Concluding remarks

For the semi-infinite program with an infinitely many conic constraints (SICP), we have shown that the KKT conditions can be represented with finitely many conic constraints, as long as the Robinson constraint qualification (RCQ) holds. Furthermore, for solving the SICP with a convex objective function and affine conic constrains, we have proposed the regularized explicit exchange method, and established its global convergence under the Slater constraint qualification. Finally, we have conducted numerical experiments with the proposed algorithm and made some observations about its behavior. For the standard semi-infinite program, there have been developed many methods other than the exchange method. It is an interesting future subject of research to extend those methods to the SICP.

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming*, 95:3–51, 2003.

[2] I. Barrodale, L. M. Delves, and J. C. Mason. Linear chebyshev approximation of complex-valued functions. *Math. Comp*, 32(143):853–863, 1978.

[3] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar. *Convex analysis and optimization.* Athena Scientific, 2003.

[4] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer-Verlag, New York, 2000.

[5] J. Faraut and A. Korányi. *Analysis on Symmetric Cones.* Oxford Mathematical Monographs, Oxford University Press, 1994.

[6] M. Fukushima, Z. Q. Luo, and P. Tseng. Smoothing functions for second-order cone complementarity problems. *SIAM Journal on Optimization*, 12:436–460, 2001.

[7] K. Glashoff and K. Roleff. New method for chebyshev approximation of complex-valued functions. *Math. Comput.*, 36(153):233–239, 1981.

[8] M. A. Goberna and M. A. López. *Semi-infinite Programming: Recent Advances.* Kluwer Academic Publishers, Dordrecht, 2001.

[9] G. Gramlich, R. Hettich, and E. W. Sachs. Local convergence of SQP methods in semi-infinite programming. *SIAM Journal on Optimization*, 5:641–658, 1995.

[10] S. Hayashi, T. Yamaguchi, N. Yamashita, and M. Fukushima. A matrix-splitting method for symmetric affine second-order cone complementarity problems. *Journal of Computational and Applied Mathematics*, 175:335–353, 2005.

[11] S. Hayashi, N. Yamashita, and M. Fukushima. A combined smoothing and regularization method for monotone second-order cone complementarity problems. *SIAM Journal on Optimization*, 15:593–615, 2005.

[12] R. Hettich and K. O. Kortanek. Semi-infinite programming: Theory, methods, and applications. *SIAM Review*, 35:380–429, 1993.

[13] L. J. Karam and J. H. McClellan. Complex chebyshev approximation for fir filter design. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, 42(3):207–216, 1995.

[14] H. Kato and M. Fukushima. An SQP-type algorithm for nonlinear second-order cone programs. *Optimization Letters*, 1:129–144, 2007.

[15] H. C. Lai and S.-Y. Wu. On linear semi-infinite programming problems. *Numerical Functional Analysis and Optimization*, 13:287–304, 1992.

[16] C. Ling, L. Qi, G. L. Zhou, and S.-Y. Wu. Global convergence of a robust smoothing SQP method for semi-infinite programming. *Journal of Optimization Theory and Applications*, 129:147–164, 2006.

[17] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284:193–228, 1998.

[18] M. López and G. Still. Semi-infinite programming. *European Journal of Operational Research*, 180:491–518, 2007.

[19] R. Reemtsen and J. Rückmann. *Semi-infinite programming.* Kluwer Academic Publishers, Boston, 1998.

[20] R. T. Rockafellar. *Convex Analysis*. Princeton University, Princeton, 1970.

[21] Y. Tanaka, M. Fukushima, and T. Ibaraki. A globally convergent SQP method for semi-infinite nonlinear optimization. *Journal of Computational and Applied Mathematics*, 23:141–153, 1988.

[22] S.-Y. Wu, D. H. Li, L. Qi, and G. Zhou. An iterative method for solving KKT system of the semi-infinte programming. *Optimization Methods and Software*, 20:629–643, 2005.