

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 Chebyshev-like approximation, filter design, and so on. In this paper, we focus on the SIP with infinitely many conic constraints, called an SICP for short. We show that under the Robinson constraint qualification a local 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 convex 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 global convergence of the hybrid algorithm without the strict convexity assumption. We report some numerical results to examine the effectiveness of the proposed algorithms.

Key words. semi-infinite programming, conic constraints, exchange method

AMS subject classifications. 90C34, 90C90, 65K05

DOI. 10.1137/110839631

1. Introduction. In this paper, we consider the following optimization problem with an infinite number of *conic* constraints:

$$(1.1) \quad \begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{Minimize}} && f(x) \\ & \text{subject to} && g(x, t) \in C \text{ for all } t \in T, \end{aligned}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuously differentiable function, $g : \mathbb{R}^n \times T \rightarrow \mathbb{R}^m$ 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. We call this problem the semi-infinite conic program (SICP).

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) [7, 13, 15, 18, 22], which has a wide application in engineering, e.g., air pollution control, robot trajectory planning, stress of materials, etc. [13, 18]. So far, many algorithms have been proposed for solving SIPs: the discretization method [7, 11, 21], the local reduction based method [8, 12, 24], the exchange method [9, 15, 27], and others [5, 16, 20, 27]. 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 distance¹ 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 high since the cardinality of T^k grows exponentially in the dimension of T . In the local reduction based method, an infinite number of constraints in the

*Received by the editors January 31, 2011; accepted for publication (in revised form) April 16, 2012; published electronically September 11, 2012.

<http://www.siam.org/journals/siopt/22-3/83963.html>

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

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

original SIP are rewritten as a finite number of constraints with implicit functions. Although this method can reformulate the SIP as a finitely constrained optimization problem, it normally works only in a sufficiently small neighborhood of an optimal solution. 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 $T^{k+1} \subseteq T^k \cup \{t_1, t_2, \dots, t_r\}$ with $\{t_1, t_2, \dots, t_r\} \subseteq T \setminus T^k$.

A more general choice for C is the symmetric cone such as the second-order cone (SOC) $\mathcal{K}^m := \{(z_1, z_2, \dots, z_m)^\top \in \mathbb{R}^m \mid z_1 \geq \|(z_2, z_3, \dots, z_m)^\top\|\}$ and the semidefinite 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 $\{t_1, t_2, \dots, t_r\} \subseteq T$. To such a subproblem, we can apply an existing algorithm such as the interior-point method and the smoothing Newton method [1, 6, 10, 17].

In the second half of this paper, we particularly focus on the following special case of SICP (1.1):

$$(1.2) \quad \begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{Minimize}} && f(x) \\ & \text{subject to} && A(t)^\top x - b(t) \in C \text{ for all } t \in T, \end{aligned}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuously differentiable convex function, $A : T \rightarrow \mathbb{R}^{n \times m}$ and $b : T \rightarrow \mathbb{R}^m$ are continuous functions, and T and C are as in SICP (1.1). We will assume that SICP (1.2) has a nonempty solution set.²

There are some important applications of SICP (1.2). For example, when C is an SOC, SICP (1.2) can be used to formulate a Chebyshev-like approximation problem involving vector-valued functions. Specifically, let $X \subseteq \mathbb{R}^\ell$ be a nonempty set, $Y \subseteq \mathbb{R}^n$ be a given compact set, and $\Phi : Y \rightarrow \mathbb{R}^m$ and $F : \mathbb{R}^\ell \times Y \rightarrow \mathbb{R}^m$ be given functions. Then, we want to determine a parameter $u \in X$ such that $\Phi(y) \approx F(u, y)$ for all $y \in Y$. One relevant approach is to solve the following problem:

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

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

$$(1.3) \quad \begin{aligned} & \underset{(u, r) \in X \times \mathbb{R}}{\text{Minimize}} && r \\ & \text{subject to} && \begin{pmatrix} r \\ \Phi(y) - F(u, y) \end{pmatrix} \in \mathcal{K}^{m+1} \text{ for all } y \in Y, \end{aligned}$$

which is of the form (1.2) when F is affine with respect to u .

Another important application for SICP (1.2) is a *finite impulse response* (FIR) filter-design [22, 26]. Generally, the FIR filter-design is to determine a vector $h := (h_0, h_1, \dots, h_{n-1})^\top \in \mathbb{R}^n$ such that the *frequency response* function $H : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{C}$ defined by $H(h, \omega) := \sum_{k=0}^{n-1} h_k e^{-k\omega\sqrt{-1}}$ satisfies some given conditions for all $\omega \in [\omega_1, \omega_2] \subseteq [0, 2\pi]$. The following problem is called the log-Chebyshev approximation FIR filter problem [26]:

$$(1.4) \quad \underset{h \in \mathbb{R}^n}{\text{Minimize}} \quad \sup_{\omega \in [0, \pi]} |\log |H(h, \omega)| - \log D(\omega)|,$$

²One example of SICP with a nonempty feasible set is given by the Chebyshev-like approximation problem (1.3). If, additionally, the constraint set X is compact, then the problem provides us an example of SICP with a nonempty solution set.

where $D : [0, \pi] \rightarrow \mathbb{R}$ is a given *desired frequency magnitude* such that $D(\omega) > 0$ for all $\omega \in [0, \pi]$. By letting $R(h, \omega) := |H(h, \omega)|^2$ and using an auxiliary variable $r \in \mathbb{R}$, problem (1.4) is expressed as

$$\begin{aligned} & \underset{(r, h) \in \mathbb{R} \times \mathbb{R}^n}{\text{Minimize}} && r \\ & \text{subject to} && 1/r \leq R(h, \omega)/D(\omega)^2 \leq r, \\ & && R(h, \omega) \geq 0 \text{ for all } \omega \in [0, \pi], \end{aligned}$$

which can further be rewritten as

$$(1.5) \quad \begin{aligned} & \underset{(r, h) \in \mathbb{R} \times \mathbb{R}^n}{\text{Minimize}} && r \\ & \text{subject to} && rD(\omega)^2 - R(h, \omega) \geq 0, \quad R(h, \omega) \geq 0, \\ & && \begin{pmatrix} R(h, \omega) + r \\ R(h, \omega) - r \\ 2D(\omega) \end{pmatrix} \in \mathcal{K}^3 \text{ for all } \omega \in [0, \pi]. \end{aligned}$$

Since $R(h, \omega) = h_0 + \sum_{k=1}^{n-1} 2h_k \cos k\omega$, the functions involved in problem (1.5) are affine with respect to $(r, h) \in \mathbb{R} \times \mathbb{R}^n$, that is, problem (1.5) is of the form SICP (1.2) with $C = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{K}^3$ and variables $(r, h) \in \mathbb{R} \times \mathbb{R}^n$. In [19, 25], the authors consider other kinds of filter design and show that those design problems can be formulated as SICPs with infinitely many SOC constraints. However, they actually solve such problems via a uniform discretization.

The focus of the paper is twofold. First, we study the Karush–Kuhn–Tucker (KKT) conditions for SICP (1.1). 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 propose two algorithms for solving the convex SICP (1.2). Since any closed convex cone can be represented as an intersection of finitely or infinitely many halfspaces, we may reformulate (1.2) as a classical SIP with infinitely many linear inequality constraints and solve it by using existing SIP algorithms [13, 18]. However, such a reformulation approach brings serious difficulties since the dimension of the index set may become much larger than that of the original SICP (1.2).³ Therefore, it is more reasonable to deal with the cones directly without losing their special structures. The proposed algorithms 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 for which we show global convergence under the strict convexity of the objective function. The second algorithm is a regularized explicit exchange method. With the help of regularization, global convergence of the algorithm is 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 (1.2). In section 4, we combine the explicit exchange method with the regularization method and show that the hybrid algorithm is globally convergent for SICP (1.2). 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.

³In the case where $C = \mathcal{K}^m$, since $\mathcal{K}^m = \{z \in \mathbb{R}^m \mid z^\top s \geq 0 \text{ for all } s \in S\}$, where $S := \{(1, \bar{s})^\top \in \mathbb{R}^m \mid \|\bar{s}\| = 1\}$, SICP (1.2) can be reformulated as the SIP: $\min f(x)$ subject to $s^\top (A(t)^\top x - b(t)) \geq 0$ 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 .

Throughout the paper, we use the following notation. $\|\cdot\|$ 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 $((z^1)^\top, (z^2)^\top, \dots, (z^p)^\top)^\top \in \mathbb{R}^{m_1+m_2+\dots+m_p}$. For a given cone $C \subseteq \mathbb{R}^m$, C^d denotes the dual cone defined by $C^d := \{z \in \mathbb{R}^m \mid z^\top w \geq 0 \text{ for all } w \in C\}$. 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 nonempty set $D \subseteq \mathbb{R}^m$ and a function $h : \mathbb{R}^m \rightarrow \mathbb{R}$, $\operatorname{argmin}_{z \in D} h(z)$ 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., $B(z, \delta) := \{w \in \mathbb{R}^m \mid \|w - z\| \leq \delta\}$. For $z_1 + \sqrt{-1}z_2 \in \mathbb{C}$ with $z_1, z_2 \in \mathbb{R}$, its absolute value is defined by $|z_1 + \sqrt{-1}z_2| := \sqrt{z_1^2 + z_2^2}$.

2. KKT conditions for SICP. In this section, we derive the KKT conditions for SICP (1.1). The result is not only interesting in itself but also provides us with an important key to analyze global convergence of the algorithm proposed in section 4.

When $m = 1$ and $C = \mathbb{R}_+$, SICP (1.1) reduces to the classical semi-infinite program and the KKT conditions are given as follows.

LEMMA 2.1 (see [13, Theorem 3.3]). *Let $x^* \in \mathbb{R}^n$ be a local optimum of SICP (1.1) with $C := \mathbb{R}_+$. Let $T_{\text{act}}(x)$ be the set of active indices at $x \in \mathbb{R}^n$, i.e., $T_{\text{act}}(x) := \{t \in T \mid g(x, t) = 0\}$. 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 $t \in T_{\text{act}}(x^*)$. Then, there exist p indices $t_1, t_2, \dots, t_p \in T_{\text{act}}(x^*)$ 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 lemma, the MFCQ plays a key role. However, for SICP (1.1), 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 (RCQ). *Let $x \in \mathbb{R}^n$ be a feasible point of SICP (1.1). Then, we say that the RCQ holds at x if there exists a vector $d \in \mathbb{R}^n$ such that*

$$(2.1) \quad g(x, t) + \nabla_x g(x, t)^\top d \in \operatorname{int} C \quad \text{for all } t \in T.$$

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 (SCQ) 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 details about the RCQ, see [3]. 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 a nonempty closed convex cone. Then, (i) there exists a nonempty compact set $S \subseteq \{s \in \mathbb{R}^m \mid \|s\| = 1\}$ such that*

$$(2.2) \quad C = \{y \in \mathbb{R}^m \mid s^\top y \geq 0 \text{ for all } s \in S\}.$$

Moreover, we have (ii) $S \subseteq C^d$ and (iii) $\operatorname{int} C \subseteq \{y \in \mathbb{R}^m \mid s^\top y > 0 \text{ for all } s \in S\}$.

Proof. We first show (i). For any $s \in \mathbb{R}^m$ with $s \neq 0$, we define the halfspace $H(s) := \{y \in \mathbb{R}^m \mid s^\top y \geq 0\}$. In addition, let $S := \{s \in \mathbb{R}^m \mid \|s\| = 1, H(s) \supseteq C\}$. By [23, 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 $\lim_{k \rightarrow \infty} s^k = s^*$ 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, by letting $k \rightarrow \infty$, we obtain $\|s^*\| = 1$ and $(s^*)^\top z \geq 0$, which implies $z \in H(s^*)$. Since $z \in C$ was arbitrary, we have $C \subseteq H(s^*)$, and hence $s^* \in S$.

Second, we show (ii). Choose $s \in S$ arbitrarily. From (2.2), we have $s^\top y \geq 0$ for all $y \in C$, which implies $s \in C^d$.

We finally show (iii). Choose $z \in \text{int } C$ arbitrarily. From the compactness of S , there exists $\bar{s} \in S$ such that $\bar{s} \in \arg\min_{s \in S} z^\top s$. To show (iii), we only have to prove $z^\top \bar{s} > 0$. For contradiction, suppose that $z^\top \bar{s} \leq 0$, which together with $\bar{s} \in S \subseteq C^d$ implies $z^\top \bar{s} = 0$. Since $z \in \text{int } C$, we have $z - \delta \bar{s} \in C$ for sufficiently small $\delta > 0$. Then, by using $z^\top \bar{s} = 0$, $z - \delta \bar{s} \in C$, and $\bar{s} \in C^d$, we have $0 \leq (z - \delta \bar{s})^\top \bar{s} = -\delta \|\bar{s}\|^2$, which yields $\bar{s} = 0$. This contradicts the fact $\bar{s} \in S$. $\square$

By using this proposition, we reformulate SICP (1.1) as a standard semi-infinite program, whereby we can derive the KKT conditions.

THEOREM 2.4. *Let $x^* \in \mathbb{R}^n$ be a local optimum of SICP (1.1). Suppose that the RCQ holds at x^* . Then, there exist p indices $t_1, t_2, \dots, t_p \in T$ and Lagrange multipliers $y^1, y^2, \dots, y^p \in \mathbb{R}^m$ such that $p \leq n$ and*

$$(2.3) \quad \nabla f(x^*) - \sum_{i=1}^p \nabla_x g(x^*, t_i) y^i = 0,$$

$$(2.4) \quad C^d \ni y^i \perp g(x^*, t_i) \in C \quad (i = 1, 2, \dots, p).$$

Proof. By Proposition 2.3, there exists a nonempty compact set $S \subseteq \{s \in \mathbb{R}^m \mid \|s\| = 1\}$ such that SICP (1.1) is equivalent to the following semi-infinite program:

$$(2.5) \quad \begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && s^\top g(x, t) \geq 0 \quad \text{for all } (s, t) \in S \times T. \end{aligned}$$

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 (2.3) and (2.4) with $y^i = 0$ for all i . Next, we suppose $(S \times T)_{\text{act}}(x^*) \neq \emptyset$. We first show that the MFCQ holds for problem (2.5), i.e., there exists a vector $d \in \mathbb{R}^n$ such that

$$(2.6) \quad (\nabla_x g(x^*, t) s)^\top d > 0 \quad \text{for all } (s, t) \in (S \times T)_{\text{act}}(x^*).$$

By assumption, there exists a vector $d \in \mathbb{R}^n$ satisfying RCQ (2.1), i.e., $g(x^*, t) + \nabla_x g(x^*, t)^\top d \in \text{int } C$ for all $t \in T$. By Proposition 2.3, we also have $0 \notin S \subseteq C^d$. Hence, from Proposition 2.3(iii), 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 (2.6). Therefore, d satisfies (2.6). Now, applying Lemma 2.1 to problem (2.5), we have p indices $(s^1, t_1), (s^2, t_2), \dots, (s^p, t_p) \in (S \times T)_{\text{act}}(x^*)$ and the Lagrange multipliers $\mu_1, \mu_2, \dots, \mu_p \geq 0$ such that $p \leq n$ and

$$(2.7) \quad \nabla f(x^*) - \sum_{i=1}^p \mu_i \nabla_x g(x^*, t_i) s_i = 0,$$

$$(2.8) \quad \mathbb{R}_+ \ni \mu_i \perp (s^i)^\top g(x^*, t_i) \in \mathbb{R}_+ \quad (i = 1, 2, \dots, p).$$

By letting $y^i := \mu_i s^i$ for each i , we have from (2.8) 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 have $g(x^*, t_i) \in C$ since x^* is feasible for SICP (1.1). Thus, (2.7) and (2.8) yield (2.3) and (2.4), respectively. This completes the proof. $\square$

Before closing this section, we give two theorems. The first one states that the KKT conditions are also sufficient for global optimality when the problem is the convex SICP (1.2).

THEOREM 2.5. *Let $x^* \in \mathbb{R}^n$ be feasible for the convex SICP (1.2). If there exist p indices $t_1, t_2, \dots, t_p \in T$ and p Lagrange multiplier vectors $y^1, y^2, \dots, y^p \in \mathbb{R}^m$ such that*

$$(2.9) \quad \nabla f(x^*) - \sum_{i=1}^p A(t_i) y^i = 0,$$

$$(2.10) \quad C^d \ni y^i \perp A(t_i)^\top x^* - b(t_i) \in C \quad (i = 1, 2, \dots, p),$$

then x^* is a global optimum of SICP (1.2).

Proof. Let $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$ be defined by $\ell(x) := f(x) - \sum_{i=1}^p (y^i)^\top (A(t_i)^\top x - b(t_i))$, and let $\bar{x} \in \mathbb{R}^n$ be an arbitrary feasible point of SICP (1.2). Since ℓ is convex and $\nabla \ell(x^*) = \nabla f(x^*) - \sum_{i=1}^p A(t_i) y^i = 0$ by (2.9), x^* is a global minimum of ℓ , i.e., $\ell(\bar{x}) - \ell(x^*) \geq 0$. Hence, we have $f(\bar{x}) - f(x^*) = \ell(\bar{x}) - \ell(x^*) + \sum_{i=1}^p (y^i)^\top (A(t_i)^\top \bar{x} - b(t_i)) \geq 0$, where the first equality follows from the definition of ℓ and (2.10), and the last inequality follows from $\ell(\bar{x}) - \ell(x^*) \geq 0$, $y^i \in C^d$, and $A(t_i)^\top \bar{x} - b(t_i) \in C$ ($i = 1, 2, \dots, p$). We thus conclude that x^* is a global optimum of SICP (1.2). $\square$

Next, we enhance Theorem 2.4 so that it can elaborate upon the case where C has a Cartesian structure, i.e., $C = C^1 \times \dots \times C^h \subseteq \mathbb{R}^m = \mathbb{R}^{m_1} \times \dots \times \mathbb{R}^{m_h}$. Consider the following problem:

$$(2.11) \quad \begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && g_j(x, t^j) \in C^j \text{ for all } t^j \in T_j, \quad j = 1, 2, \dots, h, \end{aligned}$$

where $g_j : \mathbb{R}^n \times T_j \rightarrow \mathbb{R}^{m_j}$ is continuous, $g_j(\cdot, t^j)$ is differentiable for each fixed t^j , $T_j \subseteq \mathbb{R}^{\ell_j}$ is a nonempty compact set, 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 2.6. *Let $x^* \in \mathbb{R}^n$ be a local optimum of SICP (2.11). Assume that the RCQ holds at x^* , i.e., there exists a vector $d \in \mathbb{R}^n$ such that*

$$(2.12) \quad g_j(x^*, t^j) + \nabla_x g_j(x^*, t^j)^\top d \in \text{int } C^j \text{ for all } t^j \in T_j, \quad j = 1, 2, \dots, h.$$

Then, there exist p indices⁴ $j_1, j_2, \dots, j_p \in \{1, 2, \dots, h\}$ 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

$$(2.13) \quad \nabla f(x^*) - \sum_{i=1}^p \nabla_x g_{j_i}(x^*, t_i^{j_i}) y_i^{j_i} = 0,$$

$$(2.14) \quad (C^{j_i})^d \ni y_i^{j_i} \perp g_{j_i}(x^*, t_i^{j_i}) \in C^{j_i} \quad (i = 1, 2, \dots, p).$$

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 $\tilde{T}_j := \{\tilde{t}^1\} \times \dots \times \{\tilde{t}^{j-1}\} \times T_j \times \{\tilde{t}^{j+1}\} \times \dots \times \{\tilde{t}^h\} \subseteq \mathbb{R}^{\ell_1 + \ell_2 + \dots + \ell_h}$. Then we can easily see that $\tilde{T}_j \cap \tilde{T}_{j'} = \emptyset$ for any $j \neq j'$. Let

⁴Repeated choice of the same index is allowed in the set $\{j_1, j_2, \dots, j_p\}$.

$$(2.15) \quad t := (t^1, t^2, \dots, t^h) \in \mathbb{R}^{\ell_1 + \ell_2 + \dots + \ell_h}, \quad T := \bigcup_{j=1}^h \tilde{T}_j \subseteq \mathbb{R}^{\ell_1 + \ell_2 + \dots + \ell_h},$$

and $g : \mathbb{R}^n \times T \rightarrow \mathbb{R}^{m_1 + m_2 + \dots + m_h}$ be defined by

$$(2.16) \quad g(x, t) := (\tilde{g}_1(x, t), \dots, \tilde{g}_h(x, t)),$$

where

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

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

$$(2.18) \quad \nabla_x \tilde{g}_j(x, t) := \begin{cases} \nabla_x g_j(x, t^j) & (t \in \tilde{T}_j), \\ 0 & (t \notin \tilde{T}_j). \end{cases}$$

Then, T is nonempty and compact, and SICP (2.11) is equivalent to SICP (1.1) with $C = C^1 \times \dots \times C^h$ and g defined by (2.16). By letting $d \in \mathbb{R}^n$ satisfy (2.12), we have

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

for each $j = 1, 2, \dots, h$, where the first case follows from (2.12) and the second one follows from (2.17), (2.18), and $\zeta^j \in \text{int } C^j$. Therefore, we have $g(x^*, t) + \nabla g(x^*, t)^\top d \in \text{int } C$ for all $t \in T$, which implies that the RCQ holds at x^* for SICP (1.1). Hence, by Theorem 2.4, there exist $p \leq n$, $t_1, t_2, \dots, t_p \in T$ and $y_1, y_2, \dots, y_p \in \mathbb{R}^m$ such that

$$(2.19) \quad \nabla f(x^*) - \sum_{i=1}^p \nabla_x g(x^*, t_i) y_i = 0,$$

$$(2.20) \quad C^d \ni y_i \perp g(x^*, t_i) \in C \quad (i = 1, 2, \dots, p).$$

Let $t_i := (t_i^1, t_i^2, \dots, t_i^h) \in \mathbb{R}^{\ell_1 + \ell_2 + \dots + \ell_h}$ and $y_i := (y_i^1, y_i^2, \dots, y_i^h) \in \mathbb{R}^{m_1 + m_2 + \dots + m_h}$ for $i = 1, 2, \dots, p$. From (2.15), for each i , there exists $j_i \in \{1, 2, \dots, h\}$ such that $t_i \in \tilde{T}_{j_i}$, i.e., $t_i^{j_i} \in T_{j_i}$. Then, we have

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

where the second equality follows from (2.17) and (2.18), which together with (2.19) implies (2.13). In the last, we show (2.14). From (2.20) and $C^d = (C^1)^d \times (C^2)^d \times \dots \times (C^h)^d$, we have $(C^j)^d \ni y_i^j \perp \tilde{g}_j(x^*, t_i) \in C^j$ for $j = 1, 2, \dots, h$, which together with $\tilde{g}_{j_i}(x^*, t_i) = g_{j_i}(x^*, t_i^{j_i})$ from (2.17) implies (2.14) for $i = 1, 2, \dots, p$. The proof is complete. $\square$

3. Explicit exchange method for SICP. In this section, we propose an explicit exchange method for solving the convex SICP (1.2) 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 $\text{CP}(T')$ be the relaxed problem of SICP (1.2) with T replaced by a finite subset $T' := \{t_1, t_2, \dots, t_p\} \subseteq T$. Then, $\text{CP}(T')$ can be formulated as follows:

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

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

$$\begin{aligned} \nabla f(x^*) - \sum_{i=1}^p A(t_i) y_{t_i} &= 0, \\ C^d \ni y_{t_i} \perp A(t_i)^\top x^* - b(t_i) &\in C \quad (i = 1, 2, \dots, p), \end{aligned}$$

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 $\lim_{k \rightarrow \infty} \gamma_k = 0$. Choose a finite subset $T^0 := \{t_1^0, \dots, t_\ell^0\} \subseteq T$ for some integer⁵ $\ell \geq 0$ and a vector $e \in \text{int } C$. Set $k := 0$.

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

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

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

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

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

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

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

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

and go to Step 1-2.

Step 1-2. Solve $\text{CP}(\overline{E}^{r+1})$ 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.

⁵We allow $\ell = 0$, which means $T^0 = \emptyset$.

Here, $\gamma_k > 0$ plays the role of a relaxation parameter for the feasible set of SICP (1.2). Let $X(\gamma) := \{x \in \mathbb{R}^n \mid A(t)^\top x - b(t) \in -\gamma e + C \text{ for all } t \in T\}$. Then, $X(0)$ coincides with the feasible set of SICP (1.2), and $X(\gamma)$ expands as γ increases. Note that by the termination criterion (3.2) for the inner loop we have $x^{k+1} \in X(\gamma_k)$ for each k . Hence, we can expect that the distance between x^k and the feasible set of SICP (1.2) tends to 0 as k goes to infinity. Moreover, as will be shown in the next subsection, the positivity of γ_k guarantees the inner loop of Step 1 to terminate in a finite number of iterations for each k .

When C is a symmetric cone such as an SOC or a semidefinite cone, a natural choice for the vector $e \in \text{int } C$ is the identity element with respect to Euclidean Jordan algebra [4].⁶ Moreover, in Step 1-2, we can employ an existing method such as the primal-dual interior point method, the regularized smoothing method, and others [1, 6, 10, 14, 17].

Let us denote the optimal values of $\text{CP}(T')$ and SICP (1.2) by $V(T')$ and $V(T)$, respectively. Since E^{r+1} is obtained by removing the constraints with zero Lagrange multipliers from $\bar{E}^{r+1}$ and the feasible region of $\text{CP}(E^r)$ is larger than that of $\text{CP}(\bar{E}^{r+1})$, we have

$$(3.3) \quad \begin{aligned} V(E^0) &\leq V(\bar{E}^1) = V(E^1) \leq \dots \leq V(E^r) \\ &\leq V(\bar{E}^{r+1}) = V(E^{r+1}) \leq \dots \leq V(T) < \infty. \end{aligned}$$

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

Remark 1. Note that the optimal solution set of $\text{CP}(E^r)$ contains that of $\text{CP}(\bar{E}^r)$ by the construction of E^r in Step 1-3 of Algorithm 1. Therefore, for each $k \geq 1$, we may simply set $v^0 := x^k$ in Step 1-0 without solving $\text{CP}(E^0)$ since $\text{CP}(E^0)$ is identical to $\text{CP}(E_*^r)$ and x^k solves $\text{CP}(\bar{E}_*^r)$, where E_*^r and $\bar{E}_*^r$ are the finite index sets obtained at the end of Step 1 in the previous outer iteration.

3.2. Global convergence under strict convexity assumption. In the previous subsection, we have proposed the explicit exchange method for solving SICP (1.2). 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 (1.2). (ii) In Step 1-2 of Algorithm 1, $\text{CP}(\bar{E}^{r+1})$ is solvable for each r . (iii) A generated sequence $\{v^r\}$ in every Step 1 of Algorithm 1 is bounded.

Notice that 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 $k \geq 0$.

⁶For example, if C is $\mathbb{R}_+$, $\mathcal{K}^m$, and $\mathcal{S}_+^m$, then the identity element is 1, $(1, 0, \dots, 0)^\top \in \mathbb{R}^m$ and the $m \times m$ identity matrix, respectively.

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

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

for any $t \in T$.

We next show that $\|z(v^{r+1}, t_{\text{new}}^r) - z(v^r, t_{\text{new}}^r)\|$ is bounded below by some positive number for any $r \geq 0$. Since $e \in \text{int } C$, there exists a $\delta > 0$ such that $e + B(0, \delta) \subseteq C$. We therefore have

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

where the inclusion holds since $e + B(0, \delta) \subseteq C$, $\gamma_k > 0$, $z(v^{r+1}, t_{\text{new}}^r) \in C$, and C is a convex cone⁷. From (3.1), we have $z(v^r, t_{\text{new}}^r) \notin -\gamma_k e + C$, which together with (3.5) implies that

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

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

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

THEOREM 3.2. *Suppose that Assumption A holds. Then, the inner iterations in Step 1 of Algorithm 1 terminate finitely for each k .*

Proof. Suppose, for contradiction, that the inner iterations in Step 1 do not terminate finitely at 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 $v^{r_j} \rightarrow v^*$ and $v^{r_{j+1}} \rightarrow v^{**}$ as $j \rightarrow \infty$. Moreover, 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 $\text{CP}(\bar{E}^r)$, it satisfies the following KKT conditions:

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

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

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

$$(3.9) \quad \lim_{r \rightarrow \infty} (f(v^{r+1}) - f(v^r)) = 0.$$

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

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

$$(3.11) \quad = F_r + \sum_{t \in \bar{E}^r} (y_t^r)^\top z_t^{r+1} - \sum_{t \in \bar{E}^r} (y_t^r)^\top z_t^r$$

$$(3.12) \quad = F_r + \sum_{t \in \bar{E}^r} (y_t^r)^\top z_t^{r+1},$$

⁷When C is a convex cone, $\alpha x + \beta y \in C$ holds for any $x, y \in C$ and $\alpha, \beta \geq 0$.

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 $\sum_{t \in \bar{E}^r} (y_t^r)^\top z_t^{r+1} \geq 0$. Therefore, from (3.9) and (3.12), we have

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

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

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

THEOREM 3.3. *Suppose that SICP(1.2) 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*

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

Proof. We first show that $\{x^k\}$ is bounded. Let $X(\gamma) := \{x \in \mathbb{R}^n \mid A(t)^\top x - b(t) + \gamma e \in C \text{ for all } t \in T\}$ and $L := \{x \in \mathbb{R}^n \mid f(x) \leq f(x^*)\}$. Since $x^k \in L \cap X(\gamma_k) \subseteq L \cap X(\bar{\gamma})$ with $\bar{\gamma} := \max_{k \geq 0} \gamma_k$, 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

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

where the second equality is valid since $e \in \text{int } C$ and $0 \neq s \in S \subseteq C^d$ entail $\min_{s \in S} e^\top s > 0$ from Proposition 2.3(iii). 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 $\bar{h} : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ be defined as

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

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

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

i.e., $L \cap X(\gamma)$ is a level set of $\bar{h}$. If a closed proper convex function has at least one compact level set, then any nonempty level set is also compact [2]. 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 $\lim_{k \rightarrow \infty} x^k = x^*$. Let $\bar{x}$ be an arbitrary accumulation point of $\{x^k\}$. Then, there exists a subsequence $\{x^{k_j}\} \subseteq \{x^k\}$ such that $\lim_{j \rightarrow \infty} x^{k_j} = \bar{x}$. For all j , we have $A(t)^\top x^{k_j} - b(t) + \gamma_{k_j} e \in C$ for all $t \in T$ and $f(x^{k_j}) \leq f(x^*)$. Hence, by letting j tend to ∞ , we have

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

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

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 (1.2). Since f is strictly convex, we must have $\bar{x} = x^*$. We thus have $\lim_{k \rightarrow \infty} x^k = x^*$. $\square$

4. Regularized explicit exchange method for SICP. In the previous section, we proposed the explicit exchange method for SICP (1.2) and analyzed its convergence property. However, to ensure 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 global convergence without assuming the strict convexity.

4.1. Algorithm. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function. Then, the function $f_\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}$ 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 regularized SICP (RSICP)

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

then Step 1 always 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 (1.2) as k goes infinity. However, since it is computationally prohibitive to solve RSICP(ε_k) exactly for every k , we solve RSICP(ε_k) only approximately by using the explicit exchange method. In the inner iteration of the latter method, we repeatedly solve problems of the form

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

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 $\lim_{k \rightarrow \infty} \gamma_k = \lim_{k \rightarrow \infty} \varepsilon_k = 0$. Choose a finite subset $T^0 := \{t_1^0, \dots, t_\ell^0\} \subseteq T$ for some integer $\ell \geq 0$ and a vector $e \in \text{int } C$. 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 CP(ε_k, E^0) and let v^0 be an optimum.

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

$$(4.1) \quad A(t_{\text{new}}^r)^\top v^r - b(t_{\text{new}}^r) \notin -\gamma_k e + C.$$

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

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

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

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

and go to Step 1-2.

Step 1-2. Solve $\text{CP}(\varepsilon_k, \overline{E}^{r+1})$ 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 $\overline{E}^{r+1} := \{t \in \overline{E}^{r+1} \mid y_t^{r+1} \neq 0\}$. 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.

Algorithm 2 differs from Algorithm 1 only in the choice of $\{\varepsilon_k\}$ in Step 0 and the subproblems $\text{CP}(\varepsilon_k, \overline{E}^{r+1})$ and $\text{CP}(\varepsilon_k, E^0)$ solved in Step 1. But we give a full description of Algorithm 2 for completeness.

In the subsequent convergence analysis, we omit the termination condition in Step 2, so that the algorithm may generate an infinite sequence $\{x^k\}$. Moreover, to ensure convergence, the sequences of $\{\varepsilon_k\}$ and $\{\gamma_k\}$ are required to satisfy the condition $\gamma_k = O(\varepsilon_k)$.

4.2. Global convergence without strict convexity assumption. In this section, we show global convergence of Algorithm 2 for SICP (1.2) without the strict convexity assumption. Indeed, we only need the following assumption for the proof of convergence.

Assumption B. Function f is convex. Moreover, the SCQ holds for SICP (1.2), i.e., there exists an $x_0 \in \mathbb{R}^n$ such that $A(t)^\top x_0 - b(t) \in \text{int } C$ for all $t \in T$. Notice that for SICP (1.2) 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, the inner iterations in Step 1 terminate finitely.*

Proof. By Theorem 3.2, it suffices to show that conditions (i)–(iii) in Assumption A hold when Step 1 of Algorithm 1 is applied to $\text{RSICP}(\varepsilon)$ for any $\varepsilon > 0$. Since conditions (i) and (ii) hold from the strong convexity of f_ε , we only show condition (iii). Let x_ε^* be an optimum of $\text{RSICP}(\varepsilon)$ 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. $\square$

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 (1.2) tends to 0. In the proof, the KKT conditions established in Theorem 2.4 plays a critical role.

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 to satisfy $\gamma_k = O(\varepsilon_k)$, then $\{x^k\}$ is bounded.*
- (ii) *Any accumulation point of $\{x^k\}$ solves SICP(1.2).*

Proof. (i) Let $x^* \in \mathbb{R}^n$ be an optimum of SICP (1.2). Since Assumption B holds, Theorem 2.4 can be applied to SICP (1.2) to ensure that there exist $t_1, t_2, \dots, t_p \in T$ and $y^1, y^2, \dots, y^p \in \mathbb{R}^m$ such that $p \leq n$ and

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

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

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

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

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

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

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 inequalities

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

where $\mu := \max\{\|y^1\|, \|y^2\|, \dots, \|y^p\|\}$, and the first inequality 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

$$(4.8) \quad \|x^k\|^2 \leq \|x^*\|^2 + 2p\mu\|e\|\gamma_{k-1}/\varepsilon_{k-1}.$$

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 accumulation point of $\{x^k\}$. Then, taking a subsequence if necessary, we have

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

First, we show that $\bar{x}$ is feasible to SICP (1.2). 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 $\lim_{k \rightarrow \infty} A(t)^\top x^k - b(t) + \gamma_{k-1}e = A(t)^\top \bar{x} - b(t) \in C$ for any $t \in T$. Hence, $\bar{x}$ is feasible to SICP (1.2).

We next show that $\bar{x}$ is optimal to SICP (1.2). Let x^* be an arbitrary optimum of SICP (1.2). Since $\bar{x}$ is feasible for SICP (1.2), we have $f(\bar{x}) \geq f(x^*)$. On the other hand, x^* is feasible for CP(ε_{k-1}, E_k) since the feasible region of SICP (1.2) is contained in that of CP(ε_{k-1}, E_k). Hence, we have

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

Due to the continuity of f , by letting $k \rightarrow \infty$ 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 (1.2). $\square$

From the above theorem, we can see that if we choose $\{\varepsilon_k\}$ and $\{\gamma_k\}$ such that $\gamma_k = O(\varepsilon_k)$, then the generated sequence $\{x^k\}$ has an accumulation point and it solves SICP (1.2). Moreover, we can show that if $\{\varepsilon_k\}$ and $\{\gamma_k\}$ satisfy $\gamma_k = o(\varepsilon_k)$, then $\{x^k\}$ 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 let $\{x^k\}$ be a sequence generated by Algorithm 2. Let $S^* \subseteq \mathbb{R}^n$ denote the nonempty solution set of SICP(1.2) and $x^* \in \mathbb{R}^n$ be the least 2-norm solution, i.e., $x_{\min}^* := \operatorname{argmin}_{x \in S^*} \|x\|$. Then, we have $\lim_{k \rightarrow \infty} x^k = x_{\min}^*$.*

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 $\{x^k\}$. Note that the inequality (4.8) in the proof of Theorem 4.2. (ii) holds for any $x^* \in S$, in particular for $x_{\min}^*$. Since $\gamma_k = o(\varepsilon_k)$, by letting $k \rightarrow \infty$ in (4.8) we obtain $\|\bar{x}\| \leq \|x_{\min}^*\|$. On the other hand, we have $\|\bar{x}\| \geq \|x_{\min}^*\|$ since $\bar{x} \in S^*$ and $x_{\min}^* = \operatorname{argmin}_{x \in S^*} \|x\|$. We thus have $\|\bar{x}\| = \|x_{\min}^*\|$. $\square$

5. Numerical experiments. In this section, we report some numerical results. The program is coded in MATLAB 2008a and run on a machine with an Intel Core2 Duo E6850 3.00 GHz 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 is carried out as follows. In Step 0, we set $e := (1, 0, \dots, 0)^\top \in \operatorname{int} \mathcal{K}^m$. In Step 1-1, to find t_{new}^r satisfying (4.1) we first choose N grid points $\bar{t}_1, \bar{t}_2, \dots, \bar{t}_N$ from the index set T and compute $\lambda(A(t)^\top v^r - b(t) + \gamma_k e)$ for $t = \bar{t}_1, \bar{t}_2, \dots, \bar{t}_N \in T$, where $\lambda(\cdot)$ denotes the spectral value of $z \in \mathbb{R}^m$ [6, 10] defined by

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

If we find a $\bar{t} \in \{\bar{t}_1, \bar{t}_2, \dots, \bar{t}_N\}$ such that $\lambda(A(\bar{t})^\top v^r - b(\bar{t}) + \gamma_k e) < 0$, then we set $t_{\text{new}}^r := \bar{t}$. Otherwise, we solve

$$(5.1) \quad \begin{aligned} &\text{Minimize } \lambda(A(t)^\top v^r - b(t) + \gamma_k e) \\ &\text{subject to } t \in T \end{aligned}$$

and check the nonnegativity of its optimal value.⁸ To solve (5.1), we apply Newton's method combined with the bisection method when T is one-dimensional and *fmincon* solver in MATLAB Optimization Toolbox when T is multidimensional. For both methods, we set the initial point $\bar{t}_0 \in T$ as $\bar{t}_0 := \operatorname{argmin}\{\lambda(A(t)^\top v^r - b(t) + \gamma_k e) \mid t = \bar{t}_1, \bar{t}_2, \dots, \bar{t}_N\}$. Although there is no theoretical guarantee, in practice we can expect to find a global optimum of (5.1) by taking a sufficiently large N . In Step 1-2, we solve CP(ε, T') by the smoothing method [6, 10]. 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 $\max(\varepsilon_k, \gamma_k) \leq 10^{-5}$. In each experiment reported below, we choose the grid points $\bar{t}_1, \bar{t}_2, \dots, \bar{t}_N \in T$ as in Table 5.1.

Experiment 1. In the first experiment, we solve the following SICP

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

⁸Notice that $\lambda(A(t)^\top x - b(t) + \gamma e) \geq 0$ if and only if $A(t)^\top x - b(t) \in -\gamma e + \mathcal{K}^m$.

TABLE 5.1
Choice of grid points in each experiment.

Experiment	T	N	$\{t_1, t_2, \dots, t_N\}$
1, 2, 3-1	$[-1, 1]$	101	$\{-1 + \frac{1}{50}p\}_{p=0,1,\dots,100}$
3-2	$[0, 1] \times [0, 1]$	2601 (= 51 ²)	$\{\frac{1}{50}(p_1, p_2)\}_{p_1, p_2=0,1,2,\dots,50}$

TABLE 5.2
Convergence behavior for Experiment 1.

Problem	(m, n)	ite _{out}	$\{\bar{r}_k\}$	$\bar{r}_{\text{sum}}$	t_{socp}	T_{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

where $\mathcal{K}^m := \{(x_1, x_2, \dots, x_m)^\top \in \mathbb{R}^m \mid x_1 \geq \|(x_2, x_3, \dots, x_m)^\top\|\}$, $c \in \mathbb{R}^n$, $A(t) := (A_{ij}(t)) \in \mathbb{R}^{n \times m}$ with $A_{ij}(t) := \alpha_{ij0} + \alpha_{ij1}t + \alpha_{ij2}t^2 + \alpha_{ij3}t^3$ ($i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$) 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 $b_j(t) := \beta_{j0} + \beta_{j1}t + \beta_{j2}t^2 + \beta_{j3}t^3$ ($j = 2, \dots, m$). We choose $\alpha_{ijk}, \beta_{j\ell}$ ($i = 1, 2, \dots, n$, $j = 2, \dots, m$, $k = 0, 1, 2, 3$, $\ell = 0, 1, 2, 3$) 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 referred to as problems 1, 2, $\dots$, 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 index set $T^0 := \{-1, 0, 1\}$ in Step 0, we observe the convergence behavior of the algorithm. The results are shown in Table 5.2, where

- ite_{out} : the number of outer iterations,
- $\{\bar{r}_k\}$: the values of r when the inner termination criterion (4.2) is satisfied at the k th outer iteration for $k = 0, 1, \dots, \text{ite}_{\text{out}} - 1$,
- $\bar{r}_{\text{sum}}$: the sum of $\bar{r}_k$'s for all $k = 0, 1, 2, \dots, \text{ite}_{\text{out}} - 1$,
- t_{socp} : the number of times the sub-SOCs (CP(ε_k, E^0) and CP($\varepsilon_k, \bar{E}^{r+1}$)) are solved,
- T_{fin} : the index set T^k upon termination of the algorithm,
- time(sec) : the CPU time in seconds.

In the column of $\bar{r}_k$, p^q means that we have $\bar{r}_k = p$ in q consecutive iterations. For example, $0^{10}, 2, 1^4$ means that $\bar{r}_k = 0$ ($k = 0, 1, \dots, 9$), $\bar{r}_{10} = 2$, and $\bar{r}_k = 1$ ($k = 11, 12, 13, 14$). Notice that we always have $t_{\text{socp}} = \text{ite}_{\text{out}} + \bar{r}_{\text{sum}}$, since we solve sub-SOCs 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) is always satisfied with $r = 0$ and therefore the inactive index $t = 0$ is never removed at Step 1-3. From the columns of $\bar{r}_k$, we can see that $\bar{r}_k$ is sometimes large when $k \leq 4$, but it is always 0 or 1 for $k = 7, 8, \dots, 17$. This fact suggests that T_{fin} is usually obtained in the early stage of iterations.

Experiment 2. In the second experiment, we implement the nonregularized exchange method (Algorithm 1) as well as the regularized exchange method (Algo-

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

TABLE 5.3
Comparison of regularized and nonregularized exchange methods.

Problem	(m, n)	Regularized			Nonregularized		
		t_{socp}^a	t_{socp}^b	t_{socp}^c	t_{socp}^a	t_{socp}^b	t_{socp}^c
1	(25, 15)	23	23	34	5	5	32
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	8	F

rithm 2) and compare their performance. In step 1 of Algorithm 1, for $k \geq 1$ we set $v^0 := x^k$ instead of solving $\text{CP}(E^0)$, as suggested in Remark 1 in section 3. For both 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 those used in Experiment 1.

Table 5.3 shows the obtained results, where t_{socp}^a , t_{socp}^b and t_{socp}^c denote the values of t_{socp} for the initial index sets T_a^0 , T_b^0 , and T_c^0 , respectively, and “F” means that we fail to solve a problem. From the table, we can observe that t_{socp} for the nonregularized method is much less than t_{socp} for the regularized method. This is due to the fact that the regularized exchange method has to solve the sub-SOCP ($\text{CP}(\varepsilon_k, E^0)$) at least once in every outer iteration, whereas the nonregularized exchange method does not need to solve it when the inner termination criterion (3.2) is satisfied for $r = 0$. However, convergence of the nonregularized exchange method is not guaranteed theoretically since the objective function is linear. Indeed, the nonregularized exchange method fails to solve problems 4, 5, and 6 with $T^0 = T_c^0$ since their objective functions are most probably unbounded on the feasible sets.¹⁰ On the other hand, the regularized exchange method succeeds in solving all problems for any choice of T^0 . This is the main advantage of the regularized exchange method.

Experiment 3. In the third experiment, we apply Algorithm 2 to Chebyshev-like approximation problems for vector-valued functions.

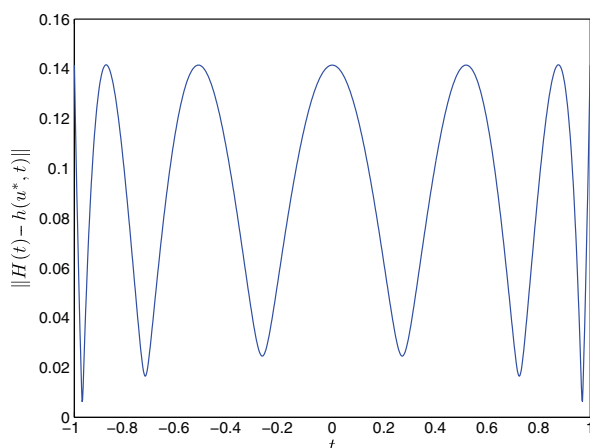
Experiment 3-1. We first consider the vector-valued approximation problem with respect to $H : \mathbb{R} \rightarrow \mathbb{R}^3$ and $h : \mathbb{R}^8 \times \mathbb{R} \rightarrow \mathbb{R}^3$ defined by

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

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:

$$(5.3) \quad \text{Minimize}_{u \in \mathbb{R}^8} \max_{t \in [-1, 1]} \|H(t) - h(u, t)\|.$$

¹⁰In fact, for each $\text{CP}(T_c^0)$ of Problems 4, 5, and 6, we found a feasible point whose objective function value is less than -10^7 .

FIG. 5.1. The graph of $\|H(t) - h(u^*, t)\|$ ($t \in [-1, 1]$) in Experiment 3-1.

Introducing an auxiliary variable $v \in \mathbb{R}$, we can reformulate (5.3) as the following SICP with infinitely many four-dimensional second-order cone constraints:

$$\begin{aligned}
 & \underset{(v, u^T)^T \in \mathbb{R} \times \mathbb{R}^8}{\text{Minimize}} && v \\
 (5.4) \quad & \text{subject to} && \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & t & t^2 & \cdots & t^7 \\ 0 & 0 & 1 & 2t & \cdots & 7t^6 \\ 0 & 0 & 0 & 2 & \cdots & 42t^5 \end{pmatrix} \begin{pmatrix} v \\ u \end{pmatrix} - \begin{pmatrix} 0 \\ e^{t^2} \\ 2te^{t^2} \\ (4t^2 + 2)e^{t^2} \end{pmatrix} \in \mathcal{K}^4 \\
 & && \text{for all } t \in [-1, 1].
 \end{aligned}$$

In applying Algorithm 2, we set $T_0 := \{-1, 1\}$ and $\varepsilon_k = \gamma_k := 0.5^k$. Then, the algorithm outputs the solution $v^* = 0.1415$, $u^* = (0.9948, 0.0000, 1.0707, 0.0000, 0.3083, 0.0000, 0.3442, 0.0000)^T$ 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 $H_5(t)$ is an odd function. Figure 5.1 shows the graph of $\|H(t) - h(u^*, t)\|$ over $t \in [-1, 1]$. From the graph, we can observe that the values of $\|H(t) - h(u^*, t)\|$ are 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.

Experiment 3-2. We next consider a vector-valued approximation problem where T is two-dimensional. Let $\tilde{H} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ and $\tilde{h} : \mathbb{R}^8 \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be defined by

$$\begin{aligned}
 \tilde{H}(t_1, t_2) &:= \begin{pmatrix} \log(t_1 + t_2 + 1) \sin t_1 \\ \sin t_1 / (t_1 + t_2 + 1) + \log(t_1 + t_2 + 1) \cos t_1 \\ \sin t_1 / (t_1 + t_2 + 1) \end{pmatrix}, \\
 \tilde{h}(u, t_1, t_2) &:= \begin{pmatrix} \sum_{\nu=1}^8 u_\nu t_1^{\nu-1} t_2^{8-\nu} \\ \sum_{\nu=2}^8 u_\nu (\nu-1) t_1^{\nu-2} t_2^{8-\nu} \\ \sum_{\nu=2}^7 u_\nu (8-\nu) t_1^{\nu-1} t_2^{7-\nu} \end{pmatrix},
 \end{aligned}$$

TABLE 5.4
Results for Experiment 3-2.

$v^*, (u^*)^\top$	0.9730, (-0.1189, 0.2040, 0.2867, -1.0159, 0.9723, 0.1877, -0.3704, 0.1687)
T_{fin}	$\{(1, 0), (0, 1), (1, 1), (0.64, 0.60), (0.46, 1), (0.60, 0.68), (0.58, 0.74), (0.62, 0.64)\}$
$\{\bar{r}_k\}$	$0^2, 2, 0^2, 3, 1, 4, 0, 1, 5, 3, 12, 6, 1, 2, 0^2$
$\bar{r}_{\text{sum}}$	40
time(sec)	15.40

respectively. In order to find a vector $u := (u_1, u_2, \dots, u_8)^\top \in \mathbb{R}^8$ such that $\tilde{h}(u, t_1, t_2) \approx \tilde{H}(t_1, t_2)$ for $(t_1, t_2) \in [0, 1] \times [0, 1]$, we solve the following problem:

$$(5.5) \quad \underset{u \in \mathbb{R}^8}{\text{Minimize}} \quad \max_{(t_1, t_2) \in [0, 1] \times [0, 1]} \left\| \tilde{H}(t_1, t_2) - \tilde{h}(u, t_1, t_2) \right\|.$$

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

$$(5.6) \quad \begin{aligned} & \underset{(v, u^\top)^\top \in \mathbb{R} \times \mathbb{R}^8}{\text{Minimize}} && v \\ & \text{subject to} && \begin{pmatrix} v \\ \sum_{\nu=1}^8 u_\nu t_1^{\nu-1} t_2^{8-\nu} \\ \sum_{\nu=2}^8 u_\nu (\nu-1) t_1^{\nu-2} t_2^{8-\nu} \\ \sum_{\nu=2}^8 u_\nu (8-\nu) t_1^{\nu-1} t_2^{7-\nu} \\ 0 \\ \log(t_1 + t_2 + 1) \sin t_1 \\ \sin t_1 / (t_1 + t_2 + 1) + \log(t_1 + t_2 + 1) \cos t_1 \\ \sin t_1 / (t_1 + t_2 + 1) \end{pmatrix} \in \mathcal{K}^4 \\ & && \text{for all } (t_1, t_2) \in [0, 1] \times [0, 1]. \end{aligned}$$

In applying Algorithm 2, we set $T_0 := \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ and $\varepsilon_k = \gamma_k := 0.5^k$. The results are shown in Table 5.4, where $(v^*, (u^*)^\top)^\top$ is the computed optimal solution. From the table, we can observe that Algorithm 2 obtained the solution within acceptable time (15.40 seconds). Moreover, the values of $|T_0|$, $|T_{\text{fin}}|$, and $\bar{r}_{\text{sum}}$ indicate that $36 (= |T_0| + \bar{r}_{\text{sum}} - |T_{\text{fin}}|)$ indices are discarded in Step 1-3 in total. Thus, the exchange-scheme in Step 1 worked efficiently to prevent the size of problems $\text{CP}(\varepsilon_k, \bar{E}^{r+1})$ from growing excessively.

6. Concluding remarks. For the semi-infinite program with an infinitely many conic constraints (i.e., SICP), we have shown that the KKT conditions can be represented with finitely many conic constraints, as long as the RCQ holds. Furthermore, for solving the SICP with a convex objective function and affine conic constraints, we have proposed the explicit exchange method and the regularized explicit exchange method and established their global convergence. Finally, we have conducted numerical experiments with the proposed algorithms to examine their effectiveness. For the standard semi-infinite program, there exist many methods other than the exchange method. It is an interesting future subject of research to extend those methods to the SICP.

Acknowledgments. The authors are grateful to the anonymous referees for their helpful comments and suggestions.

REFERENCES

- [1] F. ALIZADEH AND D. GOLDFARB, *Second-order cone programming*, Math. Program., 95 (2003), pp. 3–51.
- [2] D. P. BERTSEKAS, A. NEDIĆ, AND A. E. OZDAGLAR, *Convex Analysis and Optimization*, Athena Scientific, Belmont, MA, 2003.
- [3] J. F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, 2000.
- [4] J. FARAUT AND A. KORÁNYI, *Analysis on Symmetric Cones*, Oxford University Press, New York, 1994.
- [5] C. A. FLOUDAS AND O. STEIN, *The adaptive convexification algorithm: a feasible point method for semi-infinite programming*, SIAM J. Optim., 18 (2007), pp. 1187–1208.
- [6] M. FUKUSHIMA, Z. Q. LUO, AND P. TSENG, *Smoothing functions for second-order cone complementarity problems*, SIAM J. Optim., 12 (2001), pp. 436–460.
- [7] M. A. GOBERNA AND M. A. LÓPEZ, *Semi-Infinite Programming: Recent Advances*, Kluwer Academic Publishers, Dordrecht, 2001.
- [8] G. GRAMLICH, R. HETTICH, AND E. W. SACHS, *Local convergence of SQP methods in semi-infinite programming*, SIAM J. Optim., 5 (1995), pp. 641–658.
- [9] S. HAYASHI AND S.-Y. WU, *An explicit exchange algorithm for linear semi-infinite programming problems with second-order cone constraints*, SIAM J. Optim., 20 (2009), pp. 1527–1546.
- [10] S. HAYASHI, N. YAMASHITA, AND M. FUKUSHIMA, *A combined smoothing and regularization method for monotone second-order cone complementarity problems*, SIAM J. Optim., 15 (2005), pp. 593–615.
- [11] R. HETTICH, *An implementation of a discretization method for semi-infinite programming*, Math. Program., 34 (1986), pp. 354–361.
- [12] R. HETTICH AND W. V. HONSTED, *On quadratically convergent methods for semi-infinite programming*, in *Semi-Infinite Programming*, R. Hettich, ed., Springer, Berlin, 1979, pp. 97–111.
- [13] R. HETTICH AND K. O. KORTANEK, *Semi-infinite programming: Theory, methods, and applications*, SIAM Rev., 35 (1993), pp. 380–429.
- [14] H. KATO AND M. FUKUSHIMA, *An SQP-type algorithm for nonlinear second-order cone programs*, Optim. Lett., 1 (2007), pp. 129–144.
- [15] H. C. LAI AND S.-Y. WU, *On linear semi-infinite programming problems*, Numer. Funct. Anal. Optim., 13 (1992), pp. 287–304.
- [16] D. H. LI, L. QI, J. TAM, AND S.-Y. WU, *A smoothing Newton method for semi-infinite programming*, J. Global Optim., 30 (2004), pp. 169–194.
- [17] M. S. LOBO, L. VANDENBERGHE, S. BOYD, AND H. LEBRET, *Applications of second-order cone programming*, Linear Algebra Appl., 284 (1998), pp. 193–228.
- [18] M. LÓPEZ AND G. STILL, *Semi-infinite programming*, European J. Oper. Res., 180 (2007), pp. 491–518.
- [19] W. S. LU AND T. HINAMOTO, *A second-order cone programming approach for minimax design of 2-D FIR filters with low group delay*, in *proceedings of the IEEE international symposium on Circuits Systems*, 2006, pp. 2521–2524.
- [20] L. QI, S.-Y. WU, AND G. ZHOU, *Semismooth Newton methods for solving semi-infinite programming problems*, J. Global Optim., 27 (2003), pp. 215–232.
- [21] R. REEMTSMA, *Discretization methods for the solution of semi-infinite programming problems*, J. Optim. Theory Appl., 71 (1991), pp. 85–103.
- [22] R. REEMTSMA AND J. J. RÜCKMANN, *Semi-Infinite Programming*, Kluwer Academic Publishers, Boston, 1998.
- [23] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, 1970.
- [24] Y. TANAKA, M. FUKUSHIMA, AND T. IBARAKI, *A globally convergent SQP method for semi-infinite nonlinear optimization*, J. Comput. Appl. Math., 23 (1988), pp. 141–153.
- [25] K. M. TSUI, S. C. CHAN, AND K. S. YEUNG, *Design of FIR digital filters with prescribed flatness and peak error constraints using second-order cone programming*, IEEE Trans. Circuits Systems II, 52 (2005), pp. 601–605.
- [26] S. P. WU, S. BOYD, AND L. VANDENBERGHE, *FIR filter design via spectral factorization and convex optimization*, in *Applied and Computational Control, Signals, and Circuits*, 1999, Birkhäuser, Boston, pp. 215–245.
- [27] S.-Y. WU, D. H. LI, L. QI, AND G. ZHOU, *An iterative method for solving KKT system of the semi-infinite programming*, Optim. Methods Softw., 20 (2005), pp. 629–643.