Title	A new exchange method with a refined subproblem for solving convex semi-infinite programs (Optimization Algorithms : theory, application and implementation)
Author(s)	合元, 健祐; 山下, 信雄; 奥野, 貴之; 林, 俊介
Citation	数理解析研究所講究録 (2015), 1931: 184-196
Issue Date	2015-01
URL	http://hdl.handle.net/2433/223609
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
第 1931 巻 2015 年 184-196

A new exchange method with a refined subproblem for solving convex semi-infinite programs

Kensuke Gomoto*, Nobuo Yamashita†, Takayuki Okuno‡, Shunsuke Hayashi§

Abstract

The semi-infinite programming problem (SIP) is an optimization problem with an infinite number of constraints in a finite dimensional space. The SIP has been studied extensively so far, since a lot of practical problems in various fields such as physics, economics, and engineering can be formulated as the SIPs. The exchange method is one of the most useful algorithms for solving the SIP, and it has been developed by many researchers. In this paper, we focus on the convex SIPs and propose a new exchange method for solving them. While the traditional exchange method solves a sequence of the relaxed problems with finitely many constraints that are selected from the original constraints, our method solves a sequence of semi-infinite programs relaxing the original SIP. These relaxed problems can be solved efficiently by transforming them into certain optimization problems with finitely many constraints. Moreover, under some mild assumptions, they approximate the original SIP more precisely than the finite relaxed problems in the traditional exchange method. We also establish global convergence of the proposed method under strict convexity assumption on the objective function, and examine its efficiency through some numerical experiences.

1 Introduction

The semi-infinite programming problem (SIP) is an optimization problem with a finite dimensional variable $x \in \mathbb{R}^n$ and an infinite number of inequality constraints. The SIP has been studied extensively so far since there are a lot of applications such as Chebyshev approximation in mathematics, optimal control and trajectory control in engineerings, air/water pollution control problem, and production planning, etc. Also, from the 1960s, there have been many theoretical studies such as the optimality condition and duality theorem [12].

In this paper, we focus on the following convex SIP:

$$\begin{aligned} \text{CSIP}: \quad & \underset{x \in X}{\text{minimize}} \quad f(x) \\ & \text{subject to} \quad g(x,t) \leq 0 \quad \forall t \in T, \end{aligned} \tag{1.1}$$

where $X \subseteq \mathbb{R}^n$ is a given convex set, $T \subseteq \mathbb{R}^m$ is a nonempty compact set of the form $T = \{t \in \mathbb{R}^m | At \leq b\}$ with $A \in \mathbb{R}^{l \times m}$ and $b \in \mathbb{R}^l$, $f : \mathbb{R}^n \to \mathbb{R}$ is a function differentiable and convex over X, and $g : \mathbb{R}^n \times T \to \mathbb{R}$ is a function continuously differentiable for any $(x,t) \in X \times T$ and convex with respect to x. Since t plays a role of index in a finitely constrained optimization problem, t and T are called index and index set, respectively.

Many algorithms for solving SIP have been studied so far [6, 8]. Among them, the discretization method and the exchange methods are well known. Let

$$E := \{\bar{t}_1, \bar{t}_2, \ldots, \bar{t}_p\} \subseteq T \tag{1.2}$$

*Fujitsu Corporation (qq336ezd@gmail.com)

†Department of Applied Mathematics and Phisics, Graduate School of Informatics, Kyoto University (nobuo@i.kyoto-u.ac.jp)

‡Department of Management Science, Faculty of Engineering Division I, Tokyo University of Science (t_okuno@ms.kagu.tus.ac.jp)

§Department of Human-Social Information Sciences, Graduate School of Information Sciences, Tohoku University (s_hayashi@plan.civil.tohoku.ac.jp)

be an arbitrary finite subset of T, and let $\mathrm{SP}_{\mathrm{ex}}(E)$ be the finite approximation of CSIP (1.1) with respect to E, that is,

$$\mathrm{SP}_{\mathrm{ex}}(E): \quad \begin{array}{ll} \underset{x\in X}{\text{minimize}} & f(x) \\ \text{subject to} & g(x,\bar{t}_1) \le 0, \\ & \vdots \\ & g(x,\bar{t}_p) \le 0. \end{array} \tag{1.3}$$

The discretization method [10, 13, 14] generates a sequence of index sets $\{T_k\} \subseteq T$ satisfying (i) $|T_k| < \infty$, (ii) $T_0 \subset T_1 \subset T_2 \subset \cdots \subset T$ and (iii) $\lim_{k\to\infty} \mathrm{dist}\,(T_k, T) = 0$[1], so that the optimum x^k of $\mathrm{SP}_{\mathrm{ex}}(T_k)$ converges to the original SIP optimum as k goes infinity. On the other hand, the exchange method [2, 5, 7, 9] generates the sequence converging to the SIP optimum by exchanging an index belonging to T_k by another index belonging to $T \setminus T_k$. Unlike the discretization method, the computational cost for each subproblem does not become very large, since $|T_k|$ is bounded even when $k \to \infty$.

In this paper, we propose an exchange algorithm in which each subproblem is generated by means of the quadratic approximation with respect to t. In the existing exchange algorithms, the iterative point x^k is obtained by solving the finitely approximated subproblem of the form $\mathrm{SP}_{\mathrm{ex}}(E)$. On the other hand, the subproblems of our exchange method is of the form:

$$\mathrm{SP}_{\mathrm{new}}(E) \quad \begin{array}{ll} \underset{x\in X}{\text{minimize}} & f(x) \\ \text{subject to} & g(x,\bar{t}_1) + \nabla_t g(x,\bar{t}_1)^\top (t-\bar{t}_1) - \dfrac{L}{2}\|t-\bar{t}_1\|^2 \le 0 \quad \forall t \in T, \\ & \vdots \\ & g(x,\bar{t}_p) + \nabla_t g(x,\bar{t}_p)^\top (t-\bar{t}_p) - \dfrac{L}{2}\|t-\bar{t}_p\|^2 \le 0 \quad \forall t \in T, \end{array} \tag{1.4}$$

where $L \in \mathbb{R}$ is a positive constant. Although $\mathrm{SP}_{\mathrm{new}}(E)$ is still an SIP, it can be transformed into the problems with a finite number of constraints equivalently. Furthermore, if L is a Lipschiz constant of $\nabla_t g(x,\cdot)$ for any feasible point x of CSIP (1.1), then $\mathrm{SP}_{\mathrm{new}}(E)$ approximates CSIP (1.1) more precisely than $\mathrm{SP}_{\mathrm{ex}}(E)$. For more details, see Proposition 2.4 in Section 2. Consequently, we can expect that our method finds the optimal solution in a lower number of iterations.

This paper is organized as follows. In Section 2, we give some properties of subproblem $\mathrm{SP}_{\mathrm{new}}(E)$ that will be useful in the subsequent analyses. In Section 3, we propose an algorithm and mention some properties. Moreover, we show the global convergence of the algorithm under some assumptions. In Section 4, we provide some techniques how to solve each subproblem and how to treat the constant necessary for the numerical experiments. In Section 5, we give some numerical results relevant to Chebyshev approximation problem. Finally in Section 6, we conclude the paper with some remarks.

2 Some properties of $\mathrm{SP}_{\mathrm{new}}(E)$

In this section, we study some important properties of subproblem $\mathrm{SP}_{\mathrm{new}}(E)$. We first give the following proposition, which shows that the property of a function whose gradient is Lipschitz continuous.

Proposition 2.1 *Let $T \subseteq \mathbb{R}^n$ be a nonempty compact set and $c : T \to \mathbb{R}$ be a differentiable function. Suppose that ∇c is Lipschitz continuous over T with the constant $L_0 \in \mathbb{R}$, i.e.,*

$$\|\nabla c(t') - \nabla c(t'')\| \le L_0\|t' - t''\| \quad \forall (t', t'') \in T \times T. \tag{2.1}$$

Then, for any $(t, \bar{t}) \in T \times T$, we have

$$c(\bar{t}) + \nabla c(\bar{t})^\top (t-\bar{t}) - \frac{L_0}{2}\|t-\bar{t}\|^2 \le c(t).$$

[1]For two sets S and T with $S \subset T$, the distance from S to T is defined as $\mathrm{dist}\,(S,T) = \sup_{t\in T} \inf_{s\in S} \|s-t\|$.

Proof Fix $(t,\bar{t}) \in T \times T$ arbitrarily. Then we have

$$\begin{aligned}
& c(\bar{t}) + \nabla c(\bar{t})^\top (t-\bar{t}) - c(t) \\
&= \nabla c(\bar{t})^\top (t-\bar{t}) - \int_0^1 \nabla c\big(\bar{t} + \lambda(t-\bar{t})\big)^\top (t-\bar{t}) d\lambda \\
&= -\int_0^1 \big(\nabla c(\bar{t} + \lambda(t-\bar{t})) - \nabla c(x,\bar{t})\big)^\top (t-\bar{t}) d\lambda \\
&\le \int_0^1 \|\nabla c(\bar{t} + \lambda(t-\bar{t})) - \nabla c(\bar{t})\| \, \|t-\bar{t}\| d\lambda \\
&\le \int_0^1 L_0 \lambda \|t-\bar{t}\|^2 d\lambda \\
&= L_0 \|t-\bar{t}\|^2 \int_0^1 \lambda d\lambda \;=\; \frac{L_0}{2}\|t-\bar{t}\|^2,
\end{aligned}$$

where the first inequality is due to Cauchy-Schwarz inequality, and the second inequality follows from (2.1) with $t' := \bar{t} + \lambda(t-\bar{t})$ and $t'' := \bar{t}$. This completes the proof. ∎

In what follows, we suppose that the following assumption holds for CSIP (1.1).

Assumption A

(i) *The optimum set of CSIP (1.1) is nonempty and compact.*

(ii) *There exists an $x_0 \in X$ such that $g(x_0,t) < 0$ for any $t \in T$ and $\nabla_t g(x_0,\cdot)$ is locally Lipschitzian. That is, there exists an $L > 0$ such that*

$$\|\nabla_t g(x_0,t') - \nabla_t g(x_0,t'')\| \le L\|t'-t''\| \tag{2.2}$$

for any $(t',t'') \in T \times T$.

Needless to say, Assumption A(i) holds if CSIP (1.1) has a unique optimum. Moreover, as is shown by the following, it implies that X can be assumed to be compact essentially.

Proposition 2.2 [1, Lemma 3.1] *Suppose that Assumption* A(i) *holds. Then, there exists a finite index set $E_{\rm org} \subset T$ such that $|E_{\rm org}| < \infty$ and*

$$\overline{X} := X \cap \{x \in \mathbb{R}^n \mid g(x,\bar{t}) \le 0 \;\; (\bar{t} \in E_{\rm org})\} \cap \{x \in \mathbb{R}^n \mid f(x) \le \alpha\} \tag{2.3}$$

is nonempty and compact, where $\alpha \in \mathbb{R}^n$ is an arbitrary number with $\alpha \ge f(x^*)$.

Due to this proposition, if we can find $E_{\rm org}$ such that the solution set of $\mathrm{SP}_{\rm ex}(E_{\rm org})$ is nonempty and bounded, then we can redefine X as $X := \overline{X}$ without changing the optimum set. Assumption A(ii) holds if $\nabla_t g(\cdot,\cdot)$ is locally Lipschitzian. Assumption A(ii) is the Slater constraint qualification (SCQ) for CSIP (1.1). Moreover, by letting $c(t) := g(x_0,t)$ in Proposition 2.1, it yields

$$g(x_0,\bar{t}) + \nabla_t g(x_0,\bar{t})^\top (\bar{t}-t) - \frac{L}{2}\|\bar{t}-t\|^2 \le g(x_0,t) < 0 \tag{2.4}$$

for any $(\bar{t},t) \in T \times T$. This represents that x_0 is strictly feasible to $\mathrm{SP}_{\rm new}(E)$ for any $E \subset T$.

Now, let $L > 0$ be the Lipschitz constant satisfying (2.2). Moreover, for an arbitrarily fixed $\bar{t} \in T$, let $\hat{g}(\cdot,\bar{t}) : \mathbb{R}^n \to \mathbb{R}$ be defined by

$$\hat{g}(x,\bar{t}) := \max_{t\in T}\left\{ g(x,\bar{t}) + \nabla_t g(x,\bar{t})^\top (t-\bar{t}) - \frac{L}{2}\|t-\bar{t}\|^2 \right\}. \tag{2.5}$$

Then, $\mathrm{SP}_{\rm new}(E)$ can be rewritten equivalently as

$$\mathrm{SP}'_{\rm new}(E) \qquad \begin{array}{ll} \underset{x\in X}{\text{minimize}} & f(x) \\ \text{subject to} & \hat{g}(x,\bar{t}_1) \le 0, \\ & \quad\vdots \\ & \hat{g}(x,\bar{t}_p) \le 0. \end{array}$$

We next give the differentiability property of $\hat{g}(\cdot,\bar{t})$.

Proposition 2.3 *Fix $\bar{t} \in T$ arbitrarily. Then the function $\hat{g}(\cdot,\bar{t}) : \mathbb{R}^n \to \mathbb{R}$ defined by* (2.5) *is differentiable.*

Proof Let $d : \mathbb{R}^n \times T \to \mathbb{R}$ be defined by $d(x,t) := g(x,\bar{t}) + \nabla_t g(x,\bar{t})^\top (t-\bar{t}) - \frac{L}{2}\|t-\bar{t}\|^2$. Then, we have $\hat{g}(x,\bar{t}) = \max_{t\in T} d(x,t)$ for any x.

Now, notice that, for any fixed $x \in \mathbb{R}^n$, $\operatorname{argmax}_{t\in T} d(x,t)$ is a singleton since d is strongly concave with respect to t. Moreover, for any fixed $t \in T$, $d(\cdot,t)$ is continuously differentiable with respect to x. Therefore, by [4, Theorem 10.2.1], $\hat{g}(\cdot,\bar{t})$ is differentiable. ■

Let

$$\mathcal{F}(\mathrm{P}) := \{\text{feasible points of problem P}\}. \tag{2.6}$$

The following proposition provides the inclusion of the feasible sets of CSIP (1.1) and its finitely relaxed subproblems.

Proposition 2.4 *Let $E \subseteq T$ be an arbitrary finite set expressed as* (1.2). *Then, we have*

$$\mathcal{F}(\mathrm{CSIP}) \subseteq \mathcal{F}(\mathrm{SP_{ex}}(E)), \quad \mathcal{F}(\mathrm{SP_{new}}(E)) \subseteq \mathcal{F}(\mathrm{SP_{ex}}(E)).$$

Moreover, suppose that $L > 0$ is the Lipschitz constant of $\nabla_t g$ over not only $\{x_0\} \times T$ but also $X \times T$. Then we have

$$\mathcal{F}(\mathrm{CSIP}) \subseteq \mathcal{F}(\mathrm{SP_{new}}(E)) \subseteq \mathcal{F}(\mathrm{SP_{ex}}(E)). \tag{2.7}$$

Proof We obviously have $\mathcal{F}(\mathrm{CSIP}) \subseteq \mathcal{F}(\mathrm{SP_{ex}}(E))$ since $E \subseteq T$. Also we have $\mathcal{F}(\mathrm{SP_{new}}(E)) \subseteq \mathcal{F}(\mathrm{SP_{ex}}(E))$ since

$$g(x,\bar{t}) = g(x,\bar{t}) + \nabla_t g(x,\bar{t})^\top (\bar{t}-\bar{t}) - \frac{L}{2}\|\bar{t}-\bar{t}\|^2 \le 0$$

for any $x \in \mathcal{F}(\mathrm{SP_{new}}(E))$ and $\bar{t} \in T$.

Moreover, suppose that L is the Lipschitz constant of $\nabla_t g$ over $X \times T$. Let $\bar{t} \in T$ be fixed arbitrarily. Then, Proposition 2.1 with $c(t) := g(x,t)$ yields that

$$g(x,\bar{t}) + \nabla_t g(x,\bar{t})^\top (t-\bar{t}) - \frac{L}{2}\|t-\bar{t}\|^2 \le g(x,t)$$

for any $(x,t) \in X \times T$. Hence we have $\mathcal{F}(\mathrm{CSIP}) \subseteq \mathcal{F}(\mathrm{SP_{new}}(E))$. This completes the proof. ■

This proposition implies that $\mathrm{SP_{new}}(E)$ approximates the original CSIP more precisely than existing exchange methods, and therefore we can expect that our method finds the original CSIP optimum more rapidly than existing exchange methods. As will be stated in the next section, we can redefine $X := \overline{X}$ by Proposition 2.2 without changing the optimum set of CSIP (1.1). In such a case, Assumption A(iv) guarantees that there exists a sufficiently large $L > 0$ satisfying (2.7).

One may think that $\mathrm{SP_{new}}(E)$ is as difficult as (1.1) since $\mathrm{SP_{new}}(E)$ still has an infinite number of inequality constraints. However, $\mathrm{SP_{new}}(E)$ can be transformed into an optimization problem with a finite number of constraints equivalently by using the duality theory for the quadratic programs. We provide the transformation techniques in Section 4.

3 Algorithm

In this section, we propose a new exchange algorithm, and show its convergence property. In what follows, we assume the following.

Assumption B *Function $\hat{g}$ defined by* (2.5) *is convex with respect to x.*

This assumption requires that each subproblem $\mathrm{SP_{new}}(E)$ is a convex optimization problem. Notice that it automatically holds when g is affine with respect to x.

3.1 Approximated algorithm under strict convexity assumption

We first give an approximation algorithm, in which the objective function f is assumed to be strictly convex.

Assumption C

(i) *Function f is strictly convex.*

(ii) *There exists an optimum x^* of CSIP* (1.1) *such that $\nabla_t g(x^*, \cdot)$ is locally Lipschitzian with the constant L, i.e.,*

$$\|\nabla_t g(x^*, t') - \nabla_t g(x^*, t'')\| \leq L\|t' - t''\| \tag{3.1}$$

for any $(t', t'') \in T \times T$.

(iii) *The set X is bounded.* (*Otherwise, we can redefine $X := \overline{X}$ by Proposition* 2.2.)

Assumption C(iii) is necessary to guarantee the boundedness of the sequence generated by the algorithm. Even when the original X is unbounded, Proposition 2.2 and Assumption A(i) guarantee that X can be redefined as a bounded set without changing the optimum set.

Then the details of the algorithm are as follows.

Algorithm 1

Step 0: *Choose a small number $\gamma > 0$ and a finite subset $E_0 \subset T$. Let $L > 0$ be sufficiently large so that* (2.2) *and* (3.1) *hold. Solve* $\mathrm{SP}_{\mathrm{new}}(E_0)$ *to obtain the optimum v^0. Set $r := 0$.*

Step 1: *Find a t^r_{new} such that $g(v^r, t^r_{\mathrm{new}}) > \gamma$. If such a t^r_{new} does not exist, i.e., $g(v^r, t) \leq \gamma$ for any $t \in T$, then output $v^*(\gamma) := v^r$ and $E_{\mathrm{last}} := E_r$. Otherwise, let $\overline{E}_{r+1} := E_r \cup \{P_T(t + \frac{1}{L}\nabla_t g(v^r, t)) \mid t \in E_r\} \cup \{t^r_{\mathrm{new}}\}$ and go to Step* 2.

Step 2: *Solve* $\mathrm{SP}_{\mathrm{new}}(\overline{E}_{r+1})$ *to obtain its optimum v^{r+1} and the corresponding Lagrange multiplier $\lambda^{r+1} := \{\lambda^{r+1}(\bar{t}) \mid \bar{t} \in \overline{E}_{r+1}\}$.*

Step 3: *Let $E_{r+1} := \{\bar{t} \in \overline{E}_{r+1} | \lambda^{r+1}(\bar{t}) \neq 0\}$. Set $r := r + 1$ and return to Step* 1.

In Step 1, $P_T(\cdot)$ denotes the projection onto the index set T, i.e.,

$$P_T(s) := \operatorname*{argmin}_{t \in T} \|s - t\|.$$

Notice that we add not only t^r_{new} but also $\{P_T(t + \frac{1}{L}\nabla_t g(v^r, t)) \mid t \in E_r\}$ to E_r. This is because $P_T(t + \frac{1}{L}\nabla_t g(v^r, t))$ is more desirable than t in the sense that $P_T(t + \frac{1}{L}\nabla_t g(v^r, t))$ is obtained by means of the steepest ascent method with respect to t for a fixed v^r. In Step 3, $\lambda^{r+1}(\bar{t})$ denotes the Lagrange multiplier corresponding to the constraint of the index $\bar{t}$. Here, we remove the inactive indices whose Lagrange multipliers are zero. In the subsequent convergence analysis, we omit the termination condition in Step 2, so that the algorithm may generate an infinite sequence.

The following proposition states that the distance between v^r and v^{r+1} does not tend to zero.

Proposition 3.1 *Suppose that Assumption* C(iii) *holds. Then, there exists a $\delta > 0$ such that*

$$\|v^{r+1} - v^r\| \geq \delta \tag{3.2}$$

for all $r \geq 0$.

Proof Note that the function g is locally Lipschitzian since g is continuously differentiable. Also, T is compact and v^r $(r \geq 0)$ is contained by the compact set X. Then, for any $t \in T$ and $r \geq 0$ there exists some positive number $M > 0$ such that

$$\|g(v^{r+1}, t) - g(v^r, t)\| \leq M\|(v^{r+1}, t) - (v^r, t)\| = M\|(v^{r+1} - v^r\| \tag{3.3}$$

Moreover, we have

$$\|g(v^{r+1}, t^r_{\mathrm{new}}) - g(v^r, t^r_{\mathrm{new}})\| \geq \gamma \tag{3.4}$$

since $g(v^r, t^r_{\mathrm{new}}) > \gamma$ and $g(v^{r+1}, t^r_{\mathrm{new}}) \leq 0$. Thus, (3.3) and (3.4) yield

$$\gamma \leq \|g(v^{r+1}, t^r_{\mathrm{new}}) - g(v^r, t^r_{\mathrm{new}})\| \leq M\|v^{r+1} - v^r\|.$$

Hence we have (3.2) with $\delta := \gamma/M$. ■

Next, we show that the finite termination property of Algorithm 1.

Theorem 3.1 *Suppose that Assumptions* A–C *hold. Let* $\gamma > 0$ *be chosen arbitrarily. Then, Algorithm* 1 *terminates finitely with the outputs* $v^*(\gamma)$ *and* E_{last}. *Moreover, we have* $f(v^*(\gamma)) \leq f^*$, *where* f^* *is the optimal value of CSIP*(1.1).

Proof Note that the Slater constraint qualification holds for $\text{SP}_{\text{new}}(\overline{E}_r)$ from (2.4) under Assumptions A(ii) and B. Then, since $\hat{g}(\cdot, \bar{t})$ is differentiable by Proposition 2.3 and v^r solves $\text{SP}_{\text{new}}(\overline{E}_r)$, we have the following KKT conditions:

$$\begin{aligned} &\nabla f(v^r) + \sum_{\bar{t} \in \overline{E}_r} \lambda^r(\bar{t}) \nabla_x \hat{g}(v^r, \bar{t}) + w^r = 0, \\ &\lambda^r(\bar{t}) \geq 0,\ \hat{g}(v^r, \bar{t}) \leq 0,\ \lambda^r(\bar{t})\hat{g}(v^r, \bar{t}) = 0 \quad (\bar{t} \in \overline{E}_r), \\ &w^r \in \mathcal{N}_X(v^r), \end{aligned} \tag{3.5}$$

where $\{\lambda^r(\bar{t})\}_{\bar{t} \in \overline{E}_r}$ are the Lagrange multipliers and $\mathcal{N}_X(v^r)$ is the normal cone of X at v^r. Let $F_r := f(v^{r+1}) - f(v^r) - \nabla f(v^r)^\top (v^{r+1} - v^r)$. Then, we have

$$\begin{aligned} f(v^{r+1}) - f(v^r) &= F_r + \nabla f(v^r)^\top (v^{r+1} - v^r) \\ &= F_r - \Big(\sum_{\bar{t} \in \overline{E}_r} \lambda^r(\bar{t}) \nabla_x \hat{g}(x, \bar{t}) \Big)^\top (v^{r+1} - v^r) - (w^r)^\top (v^{r+1} - v^r) \\ &\geq F_r - \sum_{\bar{t} \in \overline{E}_r} \{ \lambda^r(\bar{t}) (\hat{g}(v^{r+1}, \bar{t}) - \hat{g}(v^r, \bar{t})) \} \\ &= F_r - \sum_{\bar{t} \in \overline{E}_r} \lambda^r(\bar{t}) \hat{g}(v^{r+1}, \bar{t}) \\ &= F_r - \sum_{\bar{t} \in E_r} \lambda^r(\bar{t}) \hat{g}(v^{r+1}, \bar{t}) - \sum_{\bar{t} \in \overline{E}_r \setminus E_r} \lambda^r(\bar{t}) \hat{g}(v^{r+1}, \bar{t}) \\ &\geq F_r \geq 0, \end{aligned} \tag{3.6}$$

where the first inequality follows from Assumption B, $w^r \in \mathcal{N}_X(v^r)$ and $v^{r+1} \in X$, and the third equality follows from (3.5). In addition, the second inequality holds since $E_r \subseteq \overline{E}_{r+1}$ and $\lambda^r(\bar{t}) = 0$ for any $\bar{t} \in \overline{E}_r \setminus E_r$. Furthermore, for any global optimum x^* of CSIP (1.1), we have $x^* \in \mathcal{F}(\text{SP}_{\text{new}}(E_r))$ from Proposition 2.1 and AssumptionC(ii) together with $\bar{t} \in E_r$. We thus have $f(v^r) \leq f(x^*) = f^*$ for each r. Consequently, we have

$$f(v^1) \leq f(v^2) \leq \cdots \leq f(v^r) \leq f(v^{r+1}) \leq \cdots \leq f^* < \infty, \tag{3.7}$$

which implies

$$\lim_{r \to \infty} \left(f(v^{r+1}) - f(v^r) \right) = 0. \tag{3.8}$$

Now, suppose for contradiction that the algorithm does not terminate finitely. Since $\{v^r\}$ is bounded by Assumption C(iii), there exist accumulation points $\bar{v}$ and $\bar{v}'$ of $\{v^r\}$ such that $v^{r_j} \to \bar{v}$ and $v^{r_j+1} \to \bar{v}'$ as $j \to \infty$. Moreover, we must have $\bar{v} \neq \bar{v}'$ from Proposition 3.1. By (3.6) and (3.8), we have

$$0 = \lim_{r \to \infty} F_r = \lim_{j \to \infty} F_{r_j} = f(\bar{v}') - f(\bar{v}) - \nabla f(\bar{v})^\top (\bar{v}' - \bar{v}),$$

but this contradicts Assumption C (i) since $\bar{v} \neq \bar{v}'$. Thus, Algorithm 1 must terminate finitely for each k. The second part of the theorem can be proved immediately by (3.7). ∎

The next theorem shows that, if a sufficiently small γ is chosen, then $v^*(\gamma)$ obtained by Algorithm 1 sufficiently approximates the original CSIP optimum. Here, we omit the proof since it is quite similar to what is discussed in the next subsection.

Theorem 3.2 *Suppose that Assumptions* A–C *hold. Let* x^* *be the unique optimum of CSIP*(1.1), *and* $v^*(\gamma)$ *be the output of Algorithm* 1. *Then, we have*

$$\lim_{\gamma \to 0} v^*(\gamma) = x^*.$$

3.2 Globally convergent algorithm with parameter controlling scheme

The output obtained by Algorithm 1 is nothing more than an approximation. In this subsection, we introduce an algorithm with the parameter controlling scheme in which Algorithm 1 is employed as a subroutine.

First, we give the following prototype algorithm, and its convergence theorem.

Algorithm 2 (Prototype algorithm)

Step 0: *Let* $\{\gamma_k\}$ *and* $\{\delta_k\}$ *be nonnegative sequences such that* $\lim_{k\to\infty}\gamma_k = 0$ *and* $\lim_{k\to\infty}\delta_k = 0$. *Set* $k := 0$.

Step 1: *Find an* $x^{k+1} \in X$ *such that*

$$g(x^{k+1}, t) \le \gamma_k \quad (\forall t \in T), \tag{3.9}$$

$$f(x^{k+1}) \le f^* + \delta_k, \tag{3.10}$$

where f^* *is the optimal value of CSIP*(1.1).

Step 2: *Terminate if a certain criterion is satisfied. Otherwise, set* $k = k+1$ *and return to Step* 1.

Theorem 3.3 *Suppose that Assumption* A(ii) *holds. Let* $\{\gamma_k\}$ *and* $\{\delta_k\}$ *be arbitrary nonnegative sequences such that* $\lim_{k\to\infty}\gamma_k = 0$ *and* $\lim_{k\to\infty}\delta_k = 0$. *Then, the sequence* $\{x^k\}$ *generated by Algorithm* 2 *is bounded, and its arbitrary accumulation point is the optimum of CSIP*(1.1).

Proof Let $\Omega_g(\gamma) := \{x \in \mathbb{R}^n \mid \max_{t\in T} g(x,t) \le \gamma\}$ and $\Omega_f(\delta) := \{x \in \mathbb{R}^n \mid f(x) \le f^* + \delta\}$. Then, $\Omega_g(\gamma)$ and $\Omega_f(\delta)$ are closed convex sets for any $\gamma \ge 0$ and $\delta \ge 0$, and $X \cap \Omega_g(0) \cap \Omega_f(0)$ coincides with the optimum set of CSIP (1.1). By Assumption A(ii), the set $X \cap \Omega_g(0) \cap \Omega_f(0)$ is nonempty and bounded. Thus, $X \cap \Omega_g(\gamma) \cap \Omega_f(\delta)$ is nonempty and bounded for any $\gamma \ge 0$ and $\delta \ge 0$.

Let $\overline{\gamma} := \max_k \gamma_k$ and $\overline{\delta} := \max_k \delta_k$. Then we have $x^k \in X \cap \Omega_g(\gamma_{k-1}) \cap \Omega_f(\delta_{k-1}) \subseteq X \cap \Omega_g(\overline{\gamma}) \cap \Omega_f(\overline{\delta})$ for any k. Thus $\{x^k\}$ is bounded and any accumulation point solves CSIP (1.1). ∎

Notice that Theorem 3.3 requires only the convexity of f and $g(\cdot, t)$, and Assumption A(ii). In other words, Algorithm 2 and Theorem 3.3 can be applied to the case where f is non-strictly convex. This case will be discussed in the next subsection.

Now, combining Algorithms 1 and 2, we propose the following algorithm.

Algorithm 3

Step 0: *Choose a finite subset* $T_0 \subset T$. *Let* $\{\gamma_k\}$ *be a positive sequence such that* $\lim_{k\to\infty}\gamma_k = 0$. *Let* $L > 0$ *be sufficiently large so that* (2.2) *and* (3.1) *hold. Set* $k := 0$.

Step 1: *Carry out Algorithm* 1 *with* $\gamma := \gamma_k$ *and* $E_0 := T_k$, *and obtain the outputs* $v^*(\gamma_k)$ *and* E_{last}. *Then, let* $x^{k+1} := v^*(\gamma_k)$ *and* $T_{k+1} := E_{\text{last}}$.

Step 2: *Terminate if a certain criterion is satisfied. Otherwise, set* $k = k+1$ *and return to Step* 1.

The convergence result can be given easily by using Theorems 3.1 and 3.3.

Theorem 3.4 *Suppose that Assumptions* A–C *hold. Then, the sequence* $\{x^k\}$ *generated by Algorithm* 3 *converges to the unique optimum of CSIP*(1.1).

Proof Let $\delta_k \equiv 0$. Then, by Theorem 3.1, x^{k+1} satisfies (3.9) and (3.10) for all k. Thus, due to Theorem 3.3, $\{x^k\}$ converges to the unique optimum of CSIP (1.1). ∎

4 Efficient approach to $\mathrm{SP}_{\mathrm{new}}(E)$

In this section, we provide an efficient approach for solving $\mathrm{SP}_{\mathrm{new}}(E)$. Specifically, we reformulate the *semi-infinite* sub-problem $\mathrm{SP}_{\mathrm{new}}(E)$ as two kinds of equivalent problems with finitely many constraints. Let $P_T(\cdot)$ denote the projection onto the index set T. We first consider the case where we have the explicit expression of P_T like the case where T is represented by means of box constraints. Then, for each $i = 1, 2, \dots, p$, we have

$$\operatorname*{argmax}_{t\in T} \left\{ g(x,\bar{t}_i) + \nabla_t g(x,\bar{t}_i)^\top (t-\bar{t}_i) - \frac{L}{2}\|t-\bar{t}_i\|^2 \right\} = P_T\left(\bar{t}_i + \frac{1}{L}\nabla_t g(x,\bar{t}_i)\right).$$

Thus, from (1.4), $\mathrm{SP}_{\mathrm{new}}(E)$ can be cast as the following optimization problem with a finite number of constraints:

$$\begin{aligned}
&\underset{x\in X}{\text{minimize}} && f(x) \\
&\text{subject to} && g(x,\bar{t}_i) + \nabla_t g(x,\bar{t}_i)^\top \left(P_T\left(\bar{t}_i + \frac{1}{L}\nabla_t g(x,\bar{t}_i)\right) - \bar{t}_i\right) \\
&&& -\frac{L}{2}\left\|P_T\left(\bar{t}_i + \frac{1}{L}\nabla_t g(x,\bar{t}_i)\right) - \bar{t}_i\right\|^2 \le 0 \ (i = 1, 2, \dots, p).
\end{aligned} \tag{4.1}$$

and the above problem can be solved by an existing algorithm.

We next consider the case where P_T cannot be calculated explicitly. Fix $i \in \{1, 2, \dots, p\}$ and $x \in X$ arbitrarily. Then, the dual problem of

$$\begin{aligned}
&\text{maximize} && g(x,\bar{t}_i) + \nabla_t g(x,\bar{t}_i)^\top (t-\bar{t}_i) - \frac{L}{2}\|t-\bar{t}_i\|^2 \\
&\text{subject to} && t \in T = \{t \in \mathbb{R}^m \mid At \le b\}
\end{aligned} \tag{4.2}$$

can be represented as

$$\begin{aligned}
&\underset{\eta\in\mathbb{R}^l}{\text{minimize}} && \frac{1}{2L}\|q_{\bar{t}_i}(x,\eta)\|^2 - r_{\bar{t}_i}(x,\eta) \\
&\text{subject to} && \eta \ge 0,
\end{aligned} \tag{4.3}$$

where

$$\begin{aligned}
q_{\bar{t}_i}(x,\eta) &:= -L\bar{t}_i - \nabla_t g(x,\bar{t}_i) + A^\top \eta, \\
r_{\bar{t}_i}(x,\eta) &:= \nabla_t g(x,\bar{t}_i)^\top \bar{t}_i + \frac{L}{2}\|\bar{t}_i\|^2 - g(x,\bar{t}_i) - b^\top \eta
\end{aligned}$$

[3, Section 5.2.4]. Since the strong duality holds between (4.2) and (4.3), $\mathrm{SP}_{\mathrm{new}}(E)$ can be rewritten equivalently as

$$\begin{aligned}
&\underset{x}{\text{minimize}}\ f(x) \\
&\text{subject to}\ \min_{\eta\ge 0} \frac{1}{2L}\|q_{\bar{t}_i}(x,\eta)\|^2 - r_{\bar{t}_i}(x,\eta) \le 0 \ (i = 1, 2, \dots, p),
\end{aligned}$$

which is also equivalent to the following optimization problem with a finite number of constraints:

$$\begin{aligned}
&\underset{x,\eta^1,\dots,\eta^p}{\text{minimize}} && f(x) \\
&\text{subject to} && \frac{1}{2L}\|q_{\bar{t}_i}(x,\eta^i)\|^2 - r_{\bar{t}_i}(x,\eta^i) \le 0,\ \eta^i \ge 0 \ (i = 1, 2, \dots, p).
\end{aligned}$$

This is a convex programming problem with convex quadratic constraints if $g(\cdot, t)$ is affine. Hence, it can be solved effectively by means of the interior point method.

5 Numerical experiments

In this section, we implement Algorithm 1 and report some numerical results. The program is coded in Matlab 7.4.0(R2007a) and run on a machine with an Inter(R) Core(TM)2 Duo E6850 3.00GHz CPU and 3GB RAM. For the sake of comparison, we also implement another exchange-type method named Exchange 2, in which we update the finite index set as $\overline{E}_{r+1} = \overline{E}_r \cup \{t^r_{\mathrm{new}}\}$ and solve a sequence of finitely relaxed problems $\mathrm{SP}_{\mathrm{ex}}(E)$ instead of $\mathrm{SP}_{\mathrm{new}}(E)$ in Step 1.

Experiment 1 (Randomly generated problems)

In the first experiment we consider the following semi-infinite program with a quadratic objective function and infinitely many linear constraint functions:

$$\begin{aligned} &\text{Minimize } \frac{1}{2}x^\top Mx + c^\top x \\ &\text{subject to } a(t)^\top x - b(t) \leq 0 \text{ for all } t \in [-1,1], \end{aligned} \tag{5.1}$$

where $M \in \mathbb{R}^{20\times 20}$, $c \in \mathbb{R}^{20}$, $a(t) := (a_1(t), a_2(t), \ldots, a_{20}(t))^\top \in \mathbb{R}^{20}$ with $a_i(t) := \sum_{j=0}^{5} \alpha_{ij} t^j$ $(i = 1, 2, \ldots, 20)$ and $b(t) := 6 + \sum_{k=1}^{5} \beta_k t^k \in \mathbb{R}$.[2] We choose α_{ij}, β_k $(i = 1, 2, \ldots, 20,\ j = 0, 1, \ldots, 5,\ k = 1, 2, \ldots, 5)$ and all components of c randomly from $[-1, 1]$. Also, we set $M := N^\top N$ where all components of a matrix $N \in \mathbb{R}^{20\times 20}$ are selected from $[-1, 1]$ randomly. The actual implementation of Algorithm 1 and Exchange 2 are carried out as follows. In Step 0, we set the initial index set T^0 as $T^0 = \{-1 + q/10\}_{q=0,1,\ldots,20}$. In Steps 1-0 and 1-2 of Algorithm 1, we solve $\mathrm{SP}_{\mathrm{new}}(E)$ of the form (4.1) with $P_T(s) := \mathrm{med}(-1, s, 1)$. For solving $\mathrm{SP}_{\mathrm{new}}(E)$ and the finite relaxed problem $\mathrm{SP}_{\mathrm{ex}}(E)$, we make use of *fmincon* solver in Matlab Optimization Toolbox. In Step 1-1, we set $t^r_{\mathrm{new}} \in \mathrm{argmax}_{t\in T}\, g(v^r, t)$. For solving $\max_{t\in T} g(v^r, t)$, we first choose grid points $\bar{t}_i := -1 + (i-1)/100$ $(i = 1, \ldots, 201)$ from the index set T and let $t_{\max} \in \mathrm{argmax}_{1\leq i\leq 201}\, g(v^r, \bar{t}_i)$. Furthermore, we run Newton's method starting from $t_{\max}$ and regard the obtained solution as an optimum of $\max_{t\in T} g(v^r, t)$. We apply Algorithm 1 with $L = 30, 40$, and 100 and Exchange 2 to 50 problem instances generated in the above way. The obtained results are shown in Table 1, where each column represents as follows:

From the table, we can observe that Algorithm 1 finds an approximate feasible point v^r such

Algo 1 (M):	Algorithm 1 with $L = M$
ave-time(sec):	the average time in seconds over 50 problems
ave-ite:	the average number of iterations over 50 problems
ave-optv:	the average number of output optimal values over 50 problems

that $\max_{t\in T} g(v^r, t) \leq 10^{-5}$ in a lower number of iterations than Exchange 2. In fact, the average number of iterations for all Algo1s are about one iteration while Exchange 2 takes about 5 iterations. This may represent that a subproblem $\mathrm{SP}_{\mathrm{new}}(E)$ approximates a feasible domain of the original problems more precisely than $\mathrm{SP}_{\mathrm{ex}}(E)$ that is solved in Exchange 2.

Note that Algorithm 1 does not necessarily obtain an optimum of CSIP (5.1) if L is not so large that Assumptions A(i) and C(ii) for the global convergence holds. On the other hand, Exchange 2 finds an approximate optimum of CSIP (5.1) when the stopping condition is satisfied, since it generates iteration points by solving a sequence of the relaxed original problems $\mathrm{SP}_{\mathrm{ex}}(E)$. Hence, we check whether or not Algorithm 1 succeeds in getting an optimum of CSIP (5.1) by comparing the optimal values output by Algorithm 1 and Exchange 2. From the results, Algo 1(30) and Algo 1(40) fail to obtain optima for many problems. Actually, while the average optimal value by Exchange 2 is 1.47, the output values by Algo 1(30) and Algo 1(40) are more than 3. When $L = 100$, we can observe that Algorithm 1 successfully finds optima since the both values of ave-optv for the two methods takes 1.47.

	ave-time(sec)	ave-ite	ave-optv
Algo 1 (30)	1.98	1.04	4.56
Algo 1 (40)	1.89	1.04	3.36
Algo 1 (100)	1.41	1.24	1.47
Exchange 2	1.76	5.24	1.47

Table 1: the results for Experiment 1

[2]Note that the origin is strictly feasible point for (5.1) since we have $-b(t) < 0$ from $\sum_{j=1}^{5} \beta_j t^j$ for all $t \in [-1, 1]$.

Experiment 2 (Chebyshev approximation problem)

In the second experiment, we consider a semi-infinite program derived from the Chebyshev approximation problem. Given a function $h : \mathbb{R} \to \mathbb{R}$, one of the typical Chebyshev approximation problem is to determine the coefficients $(x_1, x_2, \ldots, x_{n-1})^\top \in \mathbb{R}^{n-1}$ such that $\sum_{i=1}^{n-1} x_i t^i \approx h(t)$ over a compact set $T (\subseteq \mathbb{R})$, where t^i denotes the i-th power of $t \in \mathbb{R}$. This can be naturally reformulated as

$$\min_{x\in\mathbb{R}^n} \max_{t\in T} \left| h(t) - \sum_{i=1}^{n-1} x_i t^i \right| .$$

By using an auxiliary variable $x_n \in \mathbb{R}$, the above problem can be transformed into the following semi-infinite program with two linear semi-infinite constrains:

$$\begin{aligned} &\underset{(x_1,x_2,\ldots,x_n)^\top\in\mathbb{R}^n}{\text{minimize}} && x_n \\ &\text{subject to} && g_1(x,t) := x_n - \sum_{i=1}^{n-1} x_i t^{i-1} + h(t) \le 0 \quad (t \in T), \\ & && g_2(x,t) := x_n + \sum_{i=1}^{n-1} x_i t^{i-1} - h(t) \le 0 \quad (t \in T). \end{aligned} \tag{5.2}$$

In the experiment, we actually solve the above SIP with the following specific data:

$$n = 9,\ T = [-5, 5],$$

and

$$h(t) = \begin{cases} t + \frac{5}{6}\pi & (t \le -\frac{5}{6}\pi), \\ \sin(t + \frac{5}{6}\pi) & (-\frac{5}{6}\pi < t \le 0), \\ \frac{1}{2}(1 + \sqrt{3} - \sqrt{3}\exp(t)) & (0 < t \le 2), \\ 5t^2 - \frac{1}{2}(40 + \sqrt{3}\exp(2))t + \frac{1}{2}(1 + \sqrt{3} + \sqrt{3}\exp(2)) & (t > 2). \end{cases}$$

For solving CSIP (5.2) involving two linear semi-infinite constraints, we add some modifications on Algorithm 1 as follows: Let $\{E_1^r\}$ and $\{E_2^r\}$ be sequences of finite index sets corresponding to the constraints g_1 and g_2, respectively. Define $\mathrm{SP_{new}}(E_1, E_2)$ with $E_i := \{\bar{t}_1^i, \bar{t}_2^i, \ldots, \bar{t}_{p_i}^i\} \subset T$ $(i = 1, 2)$ by

$$\begin{aligned} &\underset{x\in X}{\text{minimize}} && f(x) \\ &\text{subject to} && \hat{g}_i(x, \bar{t}_1^i) := g_i(x, \bar{t}_1^i) + \nabla_t g_i(x, \bar{t}_1^i)^\top (t - \bar{t}_1^i) - \frac{L}{2}\|t - \bar{t}_1^i\|^2 \le 0 \quad (t \in T), \\ & && \vdots \\ & && \hat{g}_i(x, \bar{t}_{p_i}^i) := g_i(x, \bar{t}_{p_i}^i) + \nabla_t g_i(x, \bar{t}_{p_i}^i)^\top (t - \bar{t}_{p_i}^i) - \frac{L}{2}\|t - \bar{t}_{p_i}^i\|^2 \le 0 \quad (t \in T), \\ & && (i = 1, 2). \end{aligned}$$

Also we define $\mathrm{SP_{ex}}(E_1, E_2)$ by replacing $\hat{g}_i(x, \bar{t}_j^i) \le 0$ in $\mathrm{SP_{new}}(E_1, E_2)$ with $g(x, \bar{t}_j^i) \le 0$ for $j = 1, 2, \ldots, p_i$.

Modified Algorithm 1

Step 0: Choose a small number $\gamma > 0$ and finite subsets E_0^1, $E_0^2 \subset T$. Solve $\mathrm{SP_{new}}(E_0^1, E_0^2)$ to obtain the optimum v^0. Set $r := 0$.

Step 1: Find a t_{new}^r such that $\max(g_1(v^r, t_{\mathrm{new}}^r), g_2(v^r, t_{\mathrm{new}}^r)) > \gamma$. If such a t_{new}^r does not exist, i.e., $\max(g_1(v^r, t), g_2(v^r, t)) \le \gamma$ for any $t \in T$, then output $v^*(\gamma) := v^r$, $E_{\mathrm{last}}^1 := E_r^1$, and $E_{\mathrm{last}}^2 := E_r^2$. Otherwise, let $\overline{E}_{r+1}^1 := E_r^1 \cup \{P_T(t + \frac{1}{L}\nabla_t g_1(v^r, t)) \mid t \in E_r\} \cup \{t_{\mathrm{new}}^r\}$, $\overline{E}_{r+1}^2 := E_r^2 \cup \{P_T(t + \frac{1}{L}\nabla_t g_2(v^r, t)) \mid t \in E_r\} \cup \{t_{\mathrm{new}}^r\}$ and go to Step 2.

Step 2: Solve $SP_{new}(\overline{E}^1_{r+1}, \overline{E}^2_{r+1})$ to obtain its optimum v^{r+1} and the corresponding Lagrange multiplier $\lambda_1^{r+1} := \{\lambda_1^{r+1}(\bar{t}) \mid \bar{t} \in \overline{E}^1_{r+1}\}$ and $\lambda_2^{r+1} := \{\lambda_2^{r+1}(\bar{t}) \mid \bar{t} \in \overline{E}^2_{r+1}\}$.

Step 3: Let $E^1_{r+1} := \{\bar{t} \in \overline{E}^1_{r+1} | \lambda_1^{r+1}(\bar{t}) \neq 0\}$ and $E^2_{r+1} := \{\bar{t} \in \overline{E}^2_{r+1} | \lambda^{r+1}(\bar{t}) \neq 0\}$. Set $r := r+1$ and return to Step 1.

We also modify Exchange 2 by replacing $SP_{new}(\cdot,\cdot)$ in the modified Algorithm 1 with $SP_{ex}(\cdot,\cdot)$ and updating $\overline{E}^i_{r+1}$ $(i = 1,2)$ as $\overline{E}^i_{r+1} := E^i_r \cup \{t^r_{new}\}$ $(i = 1,2)$ in Step 1. Implementations of Algorithm 1 and Exchange 2 are carried out in the same manner as Experiment 1. The obtained results are shown in Table 2 and Table 3 where .

optval:	the objective functional value of CSIP(5.2) in the final iteration;
$\max g$:	the value of $\max_{t\in T} g(v^r, t)$ in the final iteration;
iter:	the number of iterations;
time(sec):	computational time in seconds;
Algo 1(M):	Algorithm 1 with $L = M$;
T^i_{fin} $(i = 1,2)$:	the index set E^i_r obtained in Step 3 of the final iteration.

We also give Figures 1 and 2 showing how the objective functional value for CSIP(5.2) and $\max_{t\in T} g(v^r, t)$ vary as the iteration proceeds in Exchange 2 and Algo 1(30). From the tables, we can observe that Exchange 2 and Algo 1(30) finds an optimum of CSIP (5.2) successfully. On the other hand, Algo 1(10) fails to attain the optimum although it obtains an accurate feasible point such that $\max_{t\in T} g(v^r, t) = 1.0 \times 10^{-11}$. From Figures 1 and 2, we can also observe that Algorithm 1 with $L = 30$ finds the optimal solution in a lower number of iterations than Exchange 2. However, from the point of view of computational time, Exchange 2 reaches the optimum faster than Algorithm 1. Actually, Exchange 2 takes only 0.98 seconds while Algorithm 1 takes more than 2 seconds. This is due to the fact that a subproblem $SP_{ex}(E)$ solved by Exchange 2 is a just linear programming while $SP_{new}(E)$ solved by Algorithm 1 is a nonlinear programming with a complicated structure. Thus, for the case where SP_{ex} is a nonlinear programming, Algorithm 1 may be superior to Exchange 2 even in the computational time.

	optval	$\max g$	time(sec)	iter
Exchange 2	0.465	0.99×10^{-6}	0.98	16
Algo 1(30)	0.465	-1.0×10^{-10}	3.41	10
Algo 1(10)	0.504	1.0×10^{-11}	2.48	7

Table 2: comparison of the exchange method and Algorithm 1

	T^1_{fin}	T^2_{fin}
Exchange 2	$\{-3.29, 2.41, 0.15, 4.61\}$	$\{-4.56, -1.58, 1.58, 3.59, 5\}$
Algo 1(30)	$\{-3.30, 2.41, 0.15, 4.62\}$	$\{-4.56, -1.58, 1.59, 3.59, 5\}$
Algo 1(10)	$\{-3.53, 0.09, 2.44, 5\}$	$\{-4.85, -1.72, 1.58, 3.69, 5\}$

Table 3: the index sets obtained by the two methods

6 Conclusion

In this paper, we proposed the new algorithm for solving semi-infinite programming problems, and showed its convergence property under some assumptions. We also applied the algorithm to some test problems including a certain Chebyshev approximation problem and observed that the algorithm finds the SIP optimum efficiently. However, there still remain some future works.

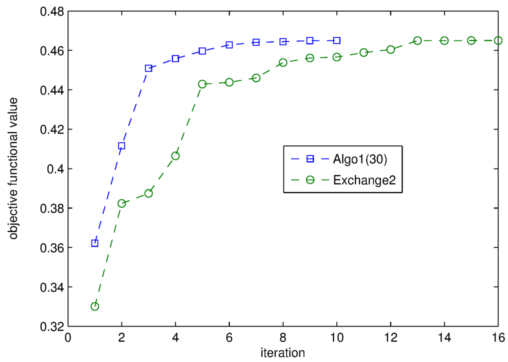

Fig. 1: the optimal values for two methods

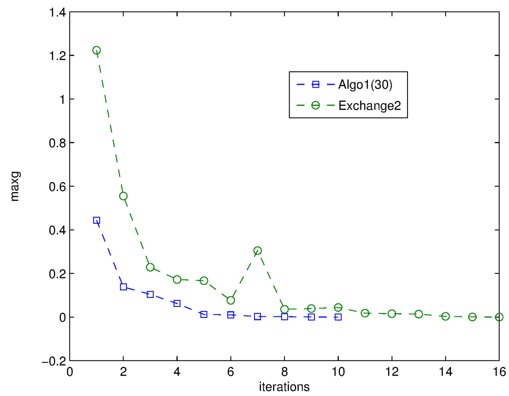

Fig. 2: $\max g$ for two methods

First, it is desired to relax the assumption that were used for the convergence analysis. Also, it is important to consider better techniques of how to choose the constant L when the Lipschitz constant is unknown.

References

[1] A. Auslender, M. A. Goberna, and M. A. López, *Penalty and smoothing methods for convex semi-infinite programming*, Mathematics of Operations Research, 34 (2009), pp. 303–319

[2] B. Betro, *An accelerated central cutting plane algorithm for linear semi-infinite programming*, Mathematical Programming, 101 (2004), pp. 479–495.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004).

[4] F. Facchinel and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vol. 2, Springer, (2003).

[5] S. Hayashi and S. Y. Wu, *An explicit exchange algorithm for linear semi-infinite programming problems with second-order cone constraints*, SIAM Journal on Optimization, 20 (2009), pp. 1527–1546.

[6] R. Hettich and K.O. Kortanek, *Semi-infinite programming: Theory, methods, and applications.* SIAM Review 35(3), 380–429 (1993).

[7] K. O. Kortanek and H. No, *A central cutting plane algorithm for convex semi-infinite programming problems*, SIAM Journal on Optimization, 3 (1993), pp. 901–918.

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

[9] T. Okuno, S. Hayashi and M. Fukushima, *A regularized explicit exchange method for semi-infinite programs with an infinite number of conic constraints*, SIAM Journal on Optimization, 22 (2012), pp. 1009–1028.

[10] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Springer, (1997).

[11] R.T. Rockafellar, *Convex Analysis*, Princeton University Press (1970).

[12] A. Shapiro, *Semi-infinite programming, duality, discretization and optimality conditions*, Optimization, 58 (2009), pp. 133–161.

[13] G. Still, *Discretization in semi-infinite programming: the rate of convergence*, Mathematical Programming, 91 (2001), pp. 53–69.

[14] K. L. Teo, X. Q. Yang, and L. S. Jennings, *Computational discretization algorithms for functional inequality constrained optimization*, Annals of Operations Research, 28 (2000), pp. 215–234.