*Journal of Global Optimization*, **29:** 401–413, 2004.



# A New Quadratic Semi-infinite Programming Algorithm Based on Dual Parametrization

Y. LIU[1], K.L. TEO[2] and S.Y. WU[3]
[1]*Department of Applied Mathematics, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, P.R. China*
[2]*Department of Applied Mathematics, and Center for Multimedia Signal Processing, Department of Electric and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, P.R. China*
[3]*Department of Mathematics, National Cheng Kong University, Tainan, Taiwan*

(Received 1 August 2003; accepted 13 August 2003)

**Abstract.** The so called dual parameterization method for quadratic semi-infinite programming (SIP) problems is developed recently. A dual parameterization algorithm is also proposed for numerical solution of such problems. In this paper, we present and improved adaptive algorithm for quadratic SIP problems with positive definite objective and multiple linear infinite constraints. In each iteration of the new algorithm, only a quadratic programming problem with a limited dimension and a limited number of constraints is required to be solved. Furthermore, convergence result is given. The efficiency of the new algorithm is shown by solving a number of numerical examples.

**Key words.** adaptive method, dual parametrization, global optimization, semi-infinite programming.

## 1. Introduction

Consider the following quadratic semi-infinite program:

**Problem** (P):

$$\min_x \quad f(x)=\frac{1}{2}x^TQx+p^Tx \tag{1}$$

$$\text{s.t.} \quad A(y)x\leqslant b(y) \quad \text{for all} \quad y\in Y, \tag{2}$$

where $x=(x_1,x_2,\ldots,x_n)^T\in R^n$ is the decision vector, $p=(p_1,p_2,\ldots,p_n)^T\in R^n$ is a constant vector, $Q\in R^{n\times n}$ is a positive definite matrix, $Y$ is a compact subset of $R^s$, and $A(y)$: $y\to R^{m\times n}$ and $b(y)$: $y\to R^m$ are continuously differentiable functions defined on $Y\subset R^s$. The vector inequalities (2) are to be understood as component-wise inequalities.

Let $g(y,x)=A(y)x-b(y)$. We will sometimes write the constraint as $g(y,x)\leqslant 0$ for convenience.

The Dorn's dual of (P) can be formulated in the following form:
**Problem** (DP):

$$\min_{x,\nu} \quad \frac{1}{2}x^T Qx + \int_Y b(y)^T \mathrm{d}\nu(y), \tag{3}$$

$$\text{s.t.} \quad Qx + p + \int_Y A(y)^T \mathrm{d}\nu(y) = 0, \tag{4}$$

$$\nu \in M^+(Y), \quad x \in R^n, \tag{5}$$

where $M^+(Y)$ is the set of nonnegative bounded regular Borel measures on $Y$.

Many papers dealt with numerical methods for solving (P). The so-called dual parameterization method [7] and the well-known cutting plane method are two important solution techniques related to this paper. For other methods, see [1]–[5]. The two methods mentioned above are different in concept, and in that the former works on the dual problem and the later deals with the primal problem. Furthermore, the dual parameterization methods is capable of finding the exact solution while the cutting plane method is designed for approximations. But, on the other hand, the two methods have similarities in some technical aspect.

The dual parameterization method is to parameterize the measure $\mu$ in the dual problem (DP) to transform into an equivalent finite dimensional nonlinear programming problem. It is shown in [7] that a global solution of this parameterized dual problem provides a solution to the original problem (P). Method for global solution of the parameterized dual have been developed in [7]. The method combines an approximation method and a local search. An adaptive method was given recently in [9] for finding an approximate solution.

On the other hand, the cutting plane method solves a sequence of regular convex programs in a systematic way. This sequence of convex programs is obtained by a discretization scheme. To be more precise, in the $k$th iteration, the problem is solved by replacing the index set $Y$ with a finite subset $Y_k$. If all the constraints of (DP) are satisfied, up to given tolerance, at the current solution, the algorithm stops and the current solution is taken as the primal solution. Otherwise, find an index point in $Y$ where the constraints violate the most. Add this point to $Y_k$ to form $Y_{k+1}$ and start a new iteration. It is shown that the corresponding solution sequence converges to an optimal solution of problem (P). Wu and Fang [11] developed a relaxed cutting plane method where the set $Y_{k+1}$ is constructed by adding to $Y_k$ a new point at which the constraints are violated but not necessarily violate the most.

The adaptive scheme in the dual parameterization algorithm developed in [9] differs from that of the cutting algorithm in forming the set $Y_{k+1}$. In [9], $Y_{k+1}$ is constructed from $Y_k$ by adding all violated index points of a refined set of grid points to $Y_k$ while dropping all unnecessary point from $Y_k$. The shortcoming of this scheme is that the number of points in $Y_k$ may not be bounded as $k$ increase.

In this paper, we construct $Y_{k+1}$ from $Y_k$ the same way as in [9] except that we add to $Y_k$ only one of the most violated points in a refined set of grid points.

We will show that for each $k$, the number of points in $Y_k$ is limited by $n+2$ where $n$ is the dimension of problem (P). We will also present convergence result under the improved adaptive scheme. Applications and numerical examples demonstrating the effectiveness of the new algorithm are given.

## 2. Dual Parameterization

For completeness, we present in this section the main results in the dual parameterization technique. We assume the following constraint qualification throughout the paper.

ASSUMPTION 1 (Slater condition). There is an $x_0 \in R^n$ such that

$$A(y)x_0 - b(y) < 0 \quad \text{for all} \quad y \in Y. \tag{6}$$

We denote by $C(Y)$ the Banach space of all continuous real functions on $Y$ equipped with the supremum norm, and by $M(Y)$ the space of all signed finite regular Borel measures on $Y$. It is known that $M(Y)$ is the dual space of $C(Y)$. Let $V$ be the cone of $C(Y)$ consisting of all the nonnegative functions in $C(Y)$. The cone in $M(Y)$ associated with $V$, denoted by $V'$, consists of all the nonnegative elements (nonnegative as measure) of $M(Y)$. Thus, $\nu \in V'$ if and only if $\nu(u) \geqslant 0$ for all $u \in V$. We will use the same symbol '$\geqslant$' to denote the partial orders in both $C(Y)$ and $M(Y)$ induced by $V$ and $V'$, respectively. To be more specific, if $u$ and $v$ are two elements in $C(Y)$ (respectively, $M(Y)$), we write $u \geqslant v$ (equivalently, $v \leqslant u$) if and only if $u - v \in V$ (respectively, $V'$).

The main result in the development of the dual parameterization technique is based on the following three results, including the KKT optimality conditions, the Carathéodory's lemma and a classical result in mathematical programming [10]. The first two results are standard and the third one is given in [7]. We state them in the following without proof.

LEMMA 2.1 (KKT conditions). *Let the Slater constraint qualification be satisfied. The minimum of problem* (P) *is achieved at $x^* \in R^n$ if and only if $x^*$ is feasible and there exists a $\nu^* \in M(Y)$ such that*

$$\begin{aligned} &Qx^* + p + \int_Y A(y)^T \mathrm{d}\nu^*(y) = 0, \\ &\int_Y (A(y)^T x^* - b(y)) \mathrm{d}\nu^*(y) = 0, \\ &\nu^* \geqslant 0 \end{aligned} \tag{7}$$

LEMMA 2.2 (Carathéodory). *Let $X$ be a subset of $R^n$. If $x \in \operatorname{cone} X$, i.e., $x$ is a nonnegative linear combination of points in $X$, then there exist $n$ numbers $\alpha_i \geqslant 0$ such that*

$$x = \sum_{i=1}^{n} \alpha_i x^i$$

*for some $x^i \in X$, $i = 1, 2, \ldots, n$, i.e., $x$ can be represented as a nonnegative linear combination of at most $n$ points of $X$.*

LEMMA 2.3. *Let Assumptions* 1 *be satisfied, and assume that the minimum of problem* (P) *is achieved at $x^* \in R^n$. Then $\nu^*$ is a multiplier satisfying the KKT conditions* (7) *if and only if $(x^*, \nu^*)$ is a solutions to the dual problem* (DP).

The dual parameterization method is based on the following result.

THEOREM 2.1. *Let Assumption* 1 *be satisfied, and assume that the minimum of problem* (P) *is achieved at $x^* \in R^n$. Then the solution set of the dual problem* (DP) *contains a solution pair $(x^*, \nu^*)$ of which the measure $\nu^*$ has a finite support of no more than $n$ points.*

The importance of Theorem 2.1 lies in the fact that it allows us to reduce problem (DP) to a finite dimensional problem. In order to solve the primal problem (P), we need only to find a solution pair $(x^*, \nu^*)$ of problem (DP). From Theorem 2.1, we can restrict out search for $\nu^*$ to those nonnegative measures having a finite support of no more that $n$ supporting points. Such a measure $\nu$ is characterized by its $k$ supporting points $y_i \in Y$, $i = 1, 2, \ldots, k$, and the corresponding measures $\mu_i = \nu(\{y_i\}) > 0$, $i = 1, 2, \ldots, k$, at each point.

If we restrict the measure $\nu$ in problem (DP) to the those of finite support of no more than $k$ supporting points which are collectively denoted by $Z$, then problem (DP) is reduced to the following problem ($\mathrm{PDP}_k$).

**Problem** ($\mathrm{PDP}_k$):

$$\begin{aligned} \min_{x,\mu,\tau} \quad & \frac{1}{2} x^T Q x + \sum_{i=1}^{k} b(y_i)^T \mu_i \\ \text{s.t.} \quad & Qx + p + \sum_{i=1}^{k} A(y_i)^T \mu_i = 0 \\ & \mu_i \geqslant 0, \quad i = 1, 2, \ldots, k, \\ & y_i \in Y, \quad i = 1, 2, \ldots, k \end{aligned} \tag{8}$$

where $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ is in the space $R^{m \times k}$ and $\tau = (y_1, y_2, \ldots, y_k)$ is in the space $R^{s \times k}$.

Problem ($\mathrm{PDP}_k$) is called the parameterized dual of problem (P) ([8]) with parameterization number $k$. From the above discussions, we see that once a global solution $(x^*, \mu^*, \tau^*)$ of problem ($\mathrm{PDP}_k$) is obtained, then $x^*$ must be the solution of problem (P) if $k$ is suitably large. Here, suitably large means that $k$ is no less than an integer $k^*$, the minimum parameterization number [8], which is no more than $n$ but not known exactly before solving the problem. Thus, in order to solve problem (P), we need only to deal with problem ($\mathrm{PDP}_k$).

For any finite index set $Z = \{y_1, y_2, \ldots, y_k\} \subset Y$, we define a problem (PDP($Z$)) by

**Problem** (PDP($Z$)):

$$\begin{aligned} \min_{x,\mu} \quad & \frac{1}{2} x^T Q x + \sum_{j=1}^{k} b(y_i)^T \mu_j \\ \text{s.t.} \quad & Qx + p + \sum_{j=1}^{k} A(y_j)^T \mu_j = 0 \\ & \mu_j \geqslant 0, \quad j = 1, 2, \ldots, k, \end{aligned} \tag{9}$$

where $\mu$ is defined as in problem ($\mathrm{PDP}_k$).

In the next section, we present an improved numerical algorithm for solving problem (P).

## 3. Algorithm

From the previous section, in order to solve problem (P), we need to solve a finite dimensional nonlinear programming problem, namely, problem ($\mathrm{PDP}_k$), for a global solution. In [8] and [9], algorithms are proposed to compute a global solution of problem ($\mathrm{PDP}_k$). The smallest number of supporting points, $k$, in the optimal measures is theoretically no more than $n$. However, in the search for an optimal pair $(x^*, \nu^*)$, we technically allow $k$ to be larger than $n$, which only means that we look for an optimal $\nu^*$ in a larger set of measures. The algorithm combines an adaptive scheme for an approximate solution and a local search. The following is an improvement of the adaptive algorithm developed in [9].

For each $i \geqslant 1$, let $Y_i$ be a given subset of $Y$ satisfying

$$d(Y_i, Y) \triangleq \max_{z \in Y} \min_{y \in Y_i} |z - y| \to 0. \tag{10}$$

For any $y \in Y$ and $x \in R^n$, we define

$$g_{\max}(y, x) \triangleq \max_{1 \leqslant j \leqslant m} g_j(y, x) \tag{11}$$

where $g_j(y, x)$ is the $j$th entry of the vector $g(y, x)$.

Now we propose the following algorithm:

**Algorithm 1**

1. Choose an arbitrary $x^0 \in R^n$, a small number $\varepsilon > 0$, a large integer $N$, and a sequence of parameterization sets

$$Y_i = \{y_j^i : j = 1, 2, \ldots, k_i\}, \quad i = 0, 1, \ldots$$

satisfying (10).
2. Let $E^0 = \phi$. Set $i = 0$.
3. Set $i = i + 1$. Find $z_i \in Y_i$ such that

$$g_{\max}(z_i, x^{i-1}) = \max_{y \in Y_i} g_{\max}(y, x^{i-1}).$$

If $g_{\max}(z_i, x^{i-1}) < \varepsilon$,
$Z_i = E_{i-1}$.
If $i \geqslant N$,
goto Step 6.
Else,
set $(x^i, \nu^i) = (x^{i-1}, \nu^{i-1})$, $E_i = E_{i-1}$, repeat Step 3.
End
Else

$$Z_i = \{z_i\} \cup E_{i-1}.$$

End

4. Solve problem (PDP($Z_i$)) to obtain a solution $(x^i, \mu^i)$.
5. Choose a set $E_i \subset Z_i$ with no more than $n+1$ points such that the solution of problem (PDP($E_i$)) is in the form $(x^i, \tilde{\mu}^i)$. Go to step 3.
6. Suppose $Z_i$ has $k$ points $z_1, z_2, \ldots, z_k$ where $z_j$ corresponds to the $j$th entry of $\mu^i$. Find a local minimum $(x^*, \lambda^*, \tau^*)$ for problem (PDP$_k$), starting from $(x^i, \mu^i, \tau^i)$, where $x^i$ and $\mu^i$ are from previously defined and $\tau^i = (z_1, z_2, \ldots, z_k)$ is the $k$ tuple formed by the points in $Z_i$. Then $x^*$ is taken as the solution for problem (P).

The purpose of introducing the integer $N$ is to prevent the algorithm from terminating prematurely. For example, if the iteration number $i$ is small, the subset $Y_{i+1}$ may be relatively sparse in $Y$ and it may happen that the approximate solution $x^i$ as part of the solution $(x^i, \mu^i)$ of problem (PDP($Z_i$)) satisfies all the constraints corresponding to index points in $Y_{i+1}$. Without introducing the integer $N$, the algorithm would terminate the iteration at this stage even if $g_{\max}(y, x^i) > \delta$ at some points $y \in Y \setminus Y_{i+1}$ and goes to the final local search. In this case, $x^i$ may not

be close enough to the primal solution $x^*$ and the subsequent local search in Step 6 does not find a global solution for problem ($\mathrm{PDP}_k$).

An obvious advantage of the above algorithm is that we need only to solve a quadratic programming of dimension not exceeding $n+(n+2)m$ with $n$ linear equality constraints, where $n$ is the dimension of the primal problem and $m$ is the number of its infinite constraints, i.e., the number of rows in the matrix function $A(y)$. The reason is that $Z_i$ contains no more than $n+2$ points for each $i$. Furthermore, at the final local search, we need only to find a local solution of a non-linear programming problem of dimension no more than $n+(n+2)(s+m)$ is solved, where $s$ is the dimension of the index space.

LEMMA 3.1. *Step* 5 *of the above algorithm is numerically feasible.*

*Proof.* Let $Z_i=\{y_1, y_2, \ldots, y_k\}$. Then problem (PDP($Z_i$)) is in the form of (9). Let $(x^i, \mu^i)$ be the solution to this problem obtained in the $i$th iteration. Consider the following linear program:

**Problem** ($\mathrm{LP}_i$):

$$\begin{aligned}
\min_{\mu,\nu} \quad & \sum_{j=1}^{n+1} \nu_j \\
\text{s.t.} \quad & \sum_{j=1}^{k} b(y_j)^T \mu_j + \nu_{n+1} = v^i \\
& \sum_{j=1}^{k} A(y_j)^T \mu_j + \hat{\nu} = c^i \\
& \mu \geqslant 0, \nu \geqslant 0
\end{aligned}$$

where

$$\begin{aligned}
v^i &= \sum_{j=1}^{k} b(y_j)^T \mu_j^i, \\
c^i &= -Qx^i - p, \\
\mu &= (\mu_1^i, \mu_2^i, \ldots, \mu_k^i), \\
\hat{\nu} &= [\nu_1, \nu_2, \ldots, \nu_n]^T,
\end{aligned}$$

and

$$\nu = [\hat{\nu}^T, \nu_{n+1}]^T.$$

In problem ($\mathrm{LP}_i$), which is the phase 1 of a linear program, $\nu_j$, $j=1,2,\ldots,n+1$ are known as artificial variables in the simplex theory of linear programming. Using the simplex method, we obtain a basic feasible solution $(\bar{\mu}^i, \bar{\nu}^i)$ where $\bar{\nu}^i$ is

a zero vector and $\bar{\mu}^i$ contains no more than $n+1$ non-zero components—vectors in $R^m$. Let $\bar{\mu}^i=(\bar{\mu}^i_1,\bar{\mu}^i_2,\ldots,\bar{\mu}^i_k)$ define

$$E_i=\{y_j \mid 1\leqslant j\leqslant k, \bar{\mu}^i_j\neq 0\}.$$

We form a new tuple $\tilde{\mu}^i$ consisting of those column vectors $\bar{u}^i_j$ in $\bar{\mu}^i$ whose corresponding indices $y_j$ are in $E_i$. Then, it is easy to see that $(x^i,\tilde{\mu}^i)$ is a solution to problem (PDP($E_i$)). □

LEMMA 3.2. *Let* $Z=\{y_1,y_2,\ldots,y_k\}\subset Y$ *be any finite subset of index points. Then, problem* (PDP($Z$)) *is the dual* (*in Dorn's form*) *of the following program*: **Problem** (P($Z$)):

$$\min_x \quad (1/2)x^TQx+p^Tx \tag{12}$$

$$\text{s.t.} \quad A(y_i)x-b(y_j)\leqslant 0, \quad j=1,2,\ldots,k. \tag{13}$$

*A vector* $\hat{x}\in R^n$ *is the solution of problem* (P($Z$)) *if and only if there exists some* $\hat{\mu}=(\hat{\mu}_1,\hat{\mu}_2,\ldots,\hat{\mu}_k)\in\prod_{j=1}^k Y$ *such that* $(\hat{x},\hat{\mu})$ *is a solution of problem* (PDP($Z$)). *Furthermore,*

$$V(\text{P}(Z))=-V(\text{PDP}(Z)).$$

*Proof.* The lemma is easy to check and we omit the details. □

In the following, we prove that if the sequence of the parameterization sets $Y_i, i=1,2,\ldots,$ satisfies (10), then the solution sequence $\{x^i\}$ obtained from Algorithm 1 converges to the solution of the problem (P), as shown in the following theorem.

THEOREM 3.1. *If* (10) *is satisfied, then the sequence* $\{x^i\}$ *obtained from Algorithm* 1 *converges to the solution of problem* (P). *Therefore, assuming that problem* ($\text{PDP}_k$) *has only a finite number of local minima for each k, if* $\epsilon$ *and N are suitably chosen, the* $x^*$ *obtained in Step* 6 *is the optimal solution of problem* (P).

*Proof.* From Lemma 3.2, we see that $x^i$ is the solution of problem (P($Z_i$)). According to Step 5, $x^i$ is the solution of problem (P($E_i$)). On the other hand, $x^{i+1}$ is the solution of problem (P($Z_{i+1}$)) of which the constraint index set $Z_{i+1}$ contains $E_i$ as a subset. Thus it is easy to see that

$$f(x^i)\leqslant f(x^{i+1}), \quad i=1,2,\ldots. \tag{14}$$

The existence of a Slater point $x_0$ for problem (P) shows that the sequence $\{f(x^i)\}$ is bounded from above by $f(x_0)$. Thus there exists some constant $f^*$ such that

$$f(x^i)\to f^* \quad (i\to\infty). \tag{15}$$

The strict convexity of the quadratic cost $f(x)$ and the boundedness of $\{f(x^i)\}$ guarantee that the sequence $\{x^i\}$ is bounded. Let $\{x_{i_k}\}$ be any chosen convergent subsequence of $\{x^i\}$ such that

$$x_{i_k} \to \bar{x}, \quad (k \to \infty) \tag{16}$$

for some $\bar{x} \in R^n$. We now show that $\bar{x}$ is a feasible point of problem (P). In fact, if $\bar{x}$ is not a feasible point of problem (P), then there exists $y_0 \in Y$ such that $g_{\max}(y_0, \bar{x}) > 0$. Let

$$\delta = \frac{1}{2} g_{\max}(y_0, \bar{x}). \tag{17}$$

Since $g_{\max}(y, x)$ is continuous, we see that there exists $\eta > 0$ such that

$$|g_{\max}(y, x) - g_{\max}(y_0, \bar{x})| < \delta, \quad \text{for} \quad \|y - y_0\| < \eta, \ \|x - \bar{x}\| < \eta. \tag{18}$$

As a result, we have

$$g_{\max}(y, x) \geqslant \delta, \quad \text{for} \quad \|y - y_0\| < \eta, \ \|x - \bar{x}\| < \eta. \tag{19}$$

From (10) and (16), there exists an integer $K$ such that for $k \geqslant K$, $Y_{i_k}$ and $x^{i_k}$ satisfy

$$\max_{z \in Y} \min_{y \in Y_k} |z - y| < \eta/2 \quad \text{and} \quad \|x^{i_k} - \bar{x}\| < \eta/2.$$

Especially, there exists $y_{i_k} \in Y_{i_k}$ such that $y_{i_k}$ and $x^{i_k}$ satisfy

$$\|y_{i_k} - y_0\| < \eta/2, \quad \|x^{i_k} - \bar{x}\| < \eta/2, \quad \text{for} \quad k \geqslant K. \tag{20}$$

Thus,

$$g_{\max}(y_{i_k}, x^{i_k}) \geqslant \delta, \quad \text{for} \quad k \geqslant K. \tag{21}$$

It is clear from (21) and the definition of $Z_{i_k+1}$ that $y_{i_k}$ is in $Z_{i_k+1}$ and hence

$$g_{\max}(y_{i_k}, x^{i_k+1}) \leqslant 0, \quad \text{for} \quad k \geqslant K. \tag{22}$$

Again, from the definitions of $Z_{i_k+1}$ and $E_{i_k}$, we see that $x^{i_k}$ and $x^{i_k+1}$ are respectively the solution and a feasible point of problem (P($E_{i_k}$)). Hence, from the fact that the feasible set of problem (P($E_{i_k}$)) is convex and its objective function is strictly convex, $f(x)$ is strictly monotone along the segment connecting $x^{i_k}$ and $x^{i_k+1}$. Particularly, we have

$$f(x^{i_k}) < f((x^{i_k} + x^{i_k+1})/2) < f(x^{i_k+1}). \tag{23}$$

Since $\{y_{i_k}\}$ is contained in the compact set $Y$, it has a converging subsequence. Without loss of generality, we suppose $\{y_{i_k}\}$ itself converges to $y' \in Y$. At the same time, we can further suppose that $\{x^{i_k+1}\}$ converges to some limit $\hat{x}$. Letting $k \to \infty$ in (22) and (23), we obtain

$$g_{\max}(y', \hat{x}) \leqslant 0 \tag{24}$$

and

$$f(\bar{x}) \leqslant f((\bar{x}+\hat{x})/2) \leqslant f(\hat{x}). \tag{25}$$

From (24), it follows that

$$|g_{\max}(y', \hat{x}) - g_{\max}(y_0, \bar{x})| \geqslant g_{\max}(y_0, \bar{x}) - g_{\max}(y', \hat{x}) \geqslant 2\delta. \tag{26}$$

According to (20), we have $\|y' - \bar{y}\| \leqslant \eta/2$. Hence, (18) shows that

$$\|\hat{x} - \bar{x}\| \geqslant \eta. \tag{27}$$

From (15), we have

$$f(\bar{x}) = f(\hat{x}). \tag{28}$$

Now we see that (25), (27) and (28) contradict the fact that $f(x)$ is strictly convex. Therefore, $\bar{x}$ is feasible to problem (P).

Next, we show that the whole sequence $\{x^i\}$ converges to the solution $x^*$ of problem (p). Suppose $\{x^i\}$ does not converge. Then there are two subsequences $\{x^{i_k}\}$ and $\{x^{j_k}\}$ converging to $x'$ and $x''$, respectively, where $x' \neq x''$. Then both $x'$ and $x''$ are feasible to problem (P) as we proved above. The point $(x'+x'')/2$ is feasible to problem (P) and hence feasible to problem (P($Z_{i_k}$)) for all $k \geqslant 1$. Therefore,

$$f((x'+x'')/2) < (f(x')+f(x''))/2 = f^*.$$

Since $f(x^{i_k}) \to f^*$ as $k \to \infty$, we have, for sufficiently large $k$,

$$f((x'+x'')/2) < f(x^{i_k}).$$

This contradicts to the fact that $x^{i_k}$ is the solution of problem (P($Z_{i_k}$)). Therefore, $\{x^i\}$ converges to $x^*$. It is clear that $x^*$ is the solution of problem (P).

Finally, if $\epsilon$ is sufficiently small and if $N$ is sufficiently large, the approximation solution $x^i$ found at the termination of the iteration in Step 5 will be so close to the primal solution that the objective value of problem (PDP$_k$) at $(x^i, \mu^i, \tau^i)$ is smaller than the second smallest local minimum value of problem (PDP$_k$). Therefore, the final local search will find the global solution. The proof is complete. □

## 4. Numerical Example

The two numerical examples given in [9] were solved using the new algorithm of this paper. The same numerical results as those of [9] were obtained with much less computing effort. The new algorithm is also successfully applied to a practical filter design problem with 40 decision variables, two infinite constraints, and a two dimensional index set. That example will be presented in a separate paper. Here we present a new example with two dimensional index set.

EXAMPLE 1. Consider the one-sided $L^2$ approximation of the exponential function $e^{1-y_1-y_2}$ on $[0,1]\times[0,1]$ by quadratic (including linear) functions of variables $y_1$ and $y_2$:

$$\begin{aligned}
&\min_x \quad \int\int_Y [x_1y_1^2+x_2y_1y_2+x_3y_2^2+x_4y_1+x_5y_2+x_6-e^{1-y_1-y_2}]^2 dx\,dy\\
&\text{s.t.} \quad y_1^2x_1+y_1y_2x_2+y_2^2x_3+y_1x_4+y_2x_5+x_6\leqslant e^{1-y_1-y_2},\\
&\qquad \text{for } (y_1,y_2)\in[0,1]\times[0,1],
\end{aligned}$$

In other words, the problem is to best approximate the function $e^{1-y_1-y_2}$ from below by polynomials in $y_1, y_2$ of order not exceeding 2. The problem can be transformed into the standard quadratic form:

$$\begin{aligned}
&\min_x \quad \frac{1}{2}x^TQx+p^Tx\\
&\text{s.t.} \quad A(y)x\leqslant b(y) \quad \text{for all } y=(y_1,y_2)\in Y,
\end{aligned}$$

where

$$\begin{aligned}
x&=[x_1,x_2,x_3,x_4,x_5,x_6]^T,\\
Q&=2\int\int_Y A^T(y_1,y_2)A(y_1,y_2)dy_1\,dy_2\\
&=\begin{bmatrix}
2/5 & 1/4 & 2/9 & 1/2 & 1/3 & 2/3\\
1/4 & 2/9 & 1/4 & 1/3 & 1/3 & 1/2\\
2/9 & 1/4 & 2/5 & 1/3 & 1/2 & 2/3\\
1/2 & 1/3 & 1/3 & 2/3 & 1/2 & 1\\
1/3 & 1/3 & 1/2 & 1/2 & 2/3 & 1\\
2/3 & 1/2 & 2/3 & 1 & 1 & 2
\end{bmatrix},\\
p&=-2\int\int_Y A^T(y_1,y_2)e^{1-y_1-y_2}dy_1dy_2,\\
&\approx-[0.5519217\ 0.37959918\ 0.55192174\\
&\qquad 0.90808143\ 0.90808143\ 2.17232254]^T,\\
A(y)&=[y_1^2,y_1y_2,y_2^2,y_1,y_2,1],\\
b(y)&=e^{1-y_1-y_2},
\end{aligned}$$

*Table 1.* Numerical results for Example 1

| | |
|---|---|
| $n$ | 6 |
| $m$ | 1 |
| $s$ | 2 |
| $\varepsilon$ | 1.0e-3 |
| N | 4 |
| Number of iterations for approx. solution | 6 |
| Number of initial index points | 2 |
| Approx. sol. | (0.45800751 0.90296818 0.44626535 −1.97961484 −1.97113439 2.51138762) |
| Approx. index points | (1.000000, 1.000000) (0.062500, 0.687500) (0.437500, 0.312500) (0.468750, 0.218750) (0.531250, 0.156250) |
| Approx. measure at each index point | 0.010801 0.009818 0.010644 0.010005 0.004695 |
| Approx. obj. value | 1.38018 |
| Optimal sol. | 0.45797897 0.90297594 0.44632314 −1.97963318 −1.97108836 2.51138681 |
| Optimal active index points | (1.000000, 1.000000) (0.061988, 0.687005) (0.437095, 0.312078) (0.469177, 0.219176) (0.531462, 0.156459) |
| Optimal measure at each index point | 0.010791 0.009787 0.010602 0.010016 0.004718 |
| Optimal objective | 1.380068 |

and

$$Y = [0, 1] \times [0, 1].$$

The numerical solutions obtained by the present algorithm are shown in Table 1. We note that in the actual computation, the vector $p$ is rounded up to 8 digits.

## 5. Comments

In this paper, we developed an adaptive algorithm for strictly convex quadratic programming problems with multiple linear infinite constraints based on the dual parameterization technique. An obvious advantage of the algorithm is that in each iteration, only no more than $n+2$ index points are chosen to form the parameterized dual problem. Hence, it is only required to solve a standard quadratic program of dimension not exceeding $n+(n+2)m$ with $n$ linear equality constraints. At the termination of the iteration, a local minimization problem of dimension no higher than $n+(n+2)(m+s)$ is to be solved, where $n$, $m$ and $s$ are respectively the dimension of the primal problem, the number of infinite

constraints and the dimension of the index space. Convergence result and its proof are given. In contrast to other algorithms, the algorithm in this paper is capable of finding the exact solution rather than only an approximate solution. Two existing examples are solved by the algorithm and the numerical results show that the algorithm is efficient in computing the solution.

## Acknowledgement

This work was supported by a research grant (PolyU5101/01E) from the Research Grant Council of Hong Kong.

## References

1. Fiacco, A.V. and Ishizuka, Y. (1990), Suggested research topics in sensitivity and stability analysis for semi-infinite programming, *Annals Oper. Res.*, 27, 65–76.
2. Goberna, M.A. and Lopez, M.A. (1998), *Linear Semi-infinite Optimization*, John Wiley & Sons, Chichester.
3. Gustafson, S.-Å. and Kortanek, K.O. (1983), Semi-infinite programming and applications, in Bachen, A. et al. (eds.), *Mathematical Programming*: *the State of the Art*, Springer-Verlag, Berlin, 138–157.
4. Hettich, R. (1983), A review of numerical methods for semi-infinite optimization, in *Semi-infinite Programming and Applications*, Fiacco, A.V. and Kortanek, K.O. (eds.), Springer-verlag, Berlin, 158–178.
5. Hettich, R. and Kortanek, K.O. (1993), Semi-infinite programming: Theory, method, and applications, *SIAM Review*, 35, 380–429.
6. Liu, Y., Teo, L. and Ito, S. (1999), A dual parametrization approach to linear-quadratic semi-infinite programming problems, *Optimization Methods and Software*, 10, 471–491.
7. Ito, S., Liu, Y. and Teo, K.L. (2000), A dual parametrization method for convex semi-infinite programming, *Annals of Operations Research*, 98, 189–213.
8. Liu, Y., Teo, K.L. and Ito, S. (2001), Global optimization in linear-quadratic semi-infinite programming, *Computing*, Supplement 15, 119–132.
9. Liu, Y. and Teo, K.L. (2002), An adaptive dual parameterization algorithm for quadratic semi-infinite programming problems, Journal of Global Optimization, 24, 205–217.
10. Luenberger, D.G. (1969), *Optimization by Vector Space Method*, John Wiley & Sons Inc., New York-London-Sydney-Toronto.
11. Soon-Yi Wu and Fang, S.-C. (1999), Solving convex programs with infinitely many linear constraints by a relaxed cutting plane method, in *Computers and Mathematics with Applications*, Vol. 38.