Journal of Global Optimization **21:** 67–87, 2001.

A Cutting Plane Approach to Solving Quadratic Infinite Programs on Measure Spaces

In memory of my father Chiang-hsiang Wu

S.Y. WU
Department of Mathematics, National Cheng Kung University, Tainan 701, Taiwan (e-mail: soonyi@mail.ncku.edu.tw)

(Received 12 January 2000; accepted in revised form 20 March 2001)

Abstract. We study infinite dimensional quadratic programming (QP) problems of integral type. The decision variable is taken in the space of bounded regular Borel measures on compact Hausdorff spaces. An implicit cutting plane algorithm is developed to obtain an optimal solution of the infinite dimensional QP problem. The major computational tasks in using the implicit cutting plane approach to solve infinite dimensional QP problems lie in finding a global optimizer of a non-linear and non-convex program. We present an explicit scheme to relax this requirement and to get rid of the unnecessary constraints in each iteration in order to reduce the size of the computatioinal programs. A general convergence proof of this approach is also given.

AMS subject classifications: 90C08, 90C48

Key words: Infinite quadratic programming on measure spaces, Cutting plane method, Algorithms, Optimal value, Optimal solution

1. Introduction

Let X and Y be compact metric spaces. We denote by $C(X)$ and $M(X)$ the spaces of continuous real valued functions on X and the set of all regular Borel measures on X, respectively. Let $M^+(X)$ be the subset of $M(X)$ which consists nonnegative Borel measures on X. A function f on $X \times X$ is said to be positive semi-definite if

$$\int_X \int_X f(s,t)d\mu(s)d\mu(t) \geqslant 0 \quad \text{for all } \mu \in M(X).$$

Now we consider the following continuous quadratic programming problem. Let $\phi(s, y)$ be a real-valued continuous function on $X \times Y$, $g(y)$ be a real-valued continuous function on Y, $h(s)$ be a real-valued continuous function on X, and $f(s,t)$ be a positive semi–definite symmetric real-valued continuous function on $X \times X$. Then the continuous quadratic program is as follows:

CQP

$$\begin{aligned}
&\text{Minimize} \quad \frac{1}{2}\int_X\int_X f(s,t)d\mu(s)d\mu(t) + \int_X h(s)d\mu(s)\\
&\text{Subject to} \quad \int_X \phi(s,y)d\mu(s) \geqslant g(y) \ \text{ for each } y \text{ in } Y,\\
&\qquad\qquad \mu \in M^+(X).
\end{aligned} \tag{1}$$

This is an infinite dimensional quadratic programming of integral type. The general capacity problem has been studied by [1, 5, 9, 11, 12, 15, 16 and 17]. The recognition of the general capacity problem as an infinite dimensional linear programming problem, and its study as such, was first made by Yamasaki [17] and Ohtsuka [11]. It is by instinct to consider the quadratic form of the general capacity problem. Thus we study the continuous quadratic program in this paper. The dual problem of CQP has the following form:

DCQP

$$\begin{aligned}
&\text{Maximize} \quad (-\frac{1}{2})\int_X\int_X f(s,t)d\mu(s)d\mu(t) + \int_Y g(y)d\nu(y)\\
&\text{Subject to} \quad \int_Y \phi(s,y)d\nu(y) - \int_X f(t,s)d\mu(t) \leqslant h(s) \ \text{ for each } s \text{ in } X,\\
&\qquad\qquad \nu \in M^+(Y),\ \mu \in M(X).
\end{aligned} \tag{2}$$

We denote by $V(CQP)$ and $V(DCQP)$ the optimal values for CQP and $DCQP$, respectively. Let M_1 and M_2 be the feasible sets for problem CQP and its dual problem $DCQP$, respectively. The author [15] proved the following theorems.

THEOREM 1.1. *Let $\mu = \mu_0$ be an optimal solution to problem CQP and assume that (i) there is a $\overline{\mu} \in M^+(X)$ such that $\int_X \phi(s,y)d\overline{\mu}(s) - g(y) > 0$ for each y in Y, or (ii) $f(s,t) > 0$ for each $s, t \in X \times X$ and $h(s) \geqslant 0$ for each $s \in X$. Then an optimal solution (μ_0, ν_0) exists to problem $DCQP$, and there is no duality gap between CQP and $DCQP$.*

Condition C. We say that $f(s,t)$ satisfies condition C if for every $\mu \in M(X)$ and every sequence $\{\mu_\alpha\}$ such that $\mu_\alpha \xrightarrow{w^*} \mu$, we have

$$\int_X\int_X f(s,t)d\mu_\alpha(s)d\mu_\alpha(t) \to \int_X\int_X f(s,t)d\mu(s)d\mu(t),$$

where $\xrightarrow{w^*}$ denotes weak* convergence.

Condition C$'$. We say that $f(s,t)$ satisfies condition C' if for every $\mu \in M(X)$ and every net $\{\mu_\alpha\}$ $(\alpha \in A)$ such that $\mu_\alpha \xrightarrow[A]{w^*} \mu$, we have

$$\int_X \int_X f(s,t)d\mu_\alpha(s)d\mu_\alpha(t) \to \int_X \int_X f(s,t)d\mu(s)d\mu(t).$$

THEOREM 1.2. *Let $f(s,t)$ satisfy condition C'. Assume that there is a $\nu_0 \in M^+(Y)$ and a positive real number c such that*

$$\int_Y \phi(s,y)d\nu_0(y) + c \leqslant 0 \text{ for each } s \text{ in } X. \tag{3}$$

If the feasible set $M_1 \neq \phi$, then the CQP has an optimal solution.

Proof. Let $\mu \in M_1$. By assumption (3), we have

$$\int_X \int_Y \phi(s,y)d\nu_0(y)d\mu(s) + \int_X cd\mu(s) \leqslant 0. \tag{4}$$

Since μ is a feasible solution of CQP, we have

$$\int_X \phi(s,y)d\mu(s) \geqslant g(y),$$

which implies

$$\int_Y \int_X \phi(s,y)d\mu(s)d\nu_0(y) \geqslant \int_Y g(y)d\nu_0(y). \tag{5}$$

From (4) and (5), we get

$$-c\mu(X) \geqslant \int_Y g(y)d\nu_0(y)$$

and so

$$\mu(X) \leqslant \frac{1}{c}[-\int_Y g(y)d\nu_0(y)].$$

Hence M_1 is bounded in the norm of $M(X) = C(X)^*$. Evidently, M_1 is weak* closed, that is, closed in $\sigma(M(X), C(X))$ topology. By applying the Banach–Alaoglu theorem, we have that M_1 is weak* compact. Let

$$F(\mu) = \int_X \int_X f(s,t)d\mu(s)d\mu(t) + \int_X h(s)d\mu(s) \text{ for any } \mu \in M(X). \tag{6}$$

Since (by C') F is a continuous function on the weak* compact set M_1, it attains its minimum at a point in M_1. □

From Theorem 1.2, we know under some conditions there exists an optimal solution for CQP. In Section 2 we generalize the cutting plane method to develop an implicit algorithm for solving CQP. The major computational tasks in using the implicit algorithm to solve CQP lie in finding a global optimizer of a non-linear and non-convex function. In Section 3, we intend to construct an explicit algorithm to relax this requirement and to get rid of the unnecessary constraints in each iteratin in order to reduce the size of the computational programs. A convergence proof of this approach is given.

2. An implicit algorithm for CQP

In this section we generalize the cutting plane method for solving CQP. Before describing the method, we first introduce the following semi-infinite quadratic programming problem $(QSIP_{T_k})$, which is a 'discretized' version of CQP. Let $T_k = \{s_1, \ldots, s_k\}$.

$(QSIP_{T_k})$

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\sum_{j=1}^{k}\sum_{i=1}^{k} f(s_i, s_j)u_i u_j + \sum_{i=1}^{k} h(s_i)u_i \\ &\text{Subject to} \quad \sum_{i=1}^{k} \phi(s_i, y)u_i \geqslant g(y) \text{ for each } y \text{ in } Y, \\ &\qquad u_i \geqslant 0. \end{aligned} \tag{7}$$

We have the following dual program for the $(QSIP_{T_k})$.

$(DQSIP_{T_k})$.

$$\begin{aligned} &\text{Maximize} \quad (-\frac{1}{2})\sum_{j=1}^{k}\sum_{i=1}^{k} f(s_i, s_j)u_i u_j + \int_Y g(y)d\nu(y) \\ &\text{Subject to} \quad \int_Y \phi(s_j, y)d\nu(y) \leqslant \sum_{i=1}^{k} f(s_i, s_j)u_i + h(s_j) \text{ for } j = 1, 2, \ldots, k \\ &\qquad \text{where} \nu \in M^+(Y), \quad u_i \in \mathbb{R}(i = 1, \ldots, k). \end{aligned} \tag{8}$$

We assume that $(QSIP_{T_k})$ is solvable with an optimal value denoted by $V(QSIP_{T_k})$, and $(DQSIP_{T_k})$ is also solvable with an optimal value denoted by $V(DQSIP_{T_k})$ and $V(QSIP_{T_k}) = V(DQSIP_{T_k})$. Some basic concept and methods for solving $(QSIP_{T_k})$ and $(DQSIP_k)$ can be found in [2, 4, 7, 13 and 14].

Then the method works according to the following scheme.

Algorithm 1:

Step 1. Set $k = 1$, choose any $s_1 \in X$, and set $T_1 = \{s_1\}$.
Step 2. Solve $(QSIP_{T_k})$ with an optimal solution $u^k = (u_1^k, \dots, u_k^k)$, and $(DQSIP_{T_k})$ with an optimal solution (u^k, ν^k).
Step 3. Find a maximizer s_{k+1} of $\phi_{u^k,\nu^k}(s)$ over X where

$$\phi_{u^k,\nu^k}(s) = \int_Y \phi(s, y) d\nu^k(y) - \sum_{i=1}^{k} f(s_i, s) u_i^k - h(s). \tag{9}$$

Step 4. If $\phi_{u^k,\nu^k}(s_{k+1}) \leqslant 0$, then stop. In this case, (u^k, ν^k) is optimal for $(DCQP)$. Otherwise, set $T_{k+1} = T_k \cup \{s_{k+1}\}$, increment $k \leftarrow k + 1$, and go to Step 2.

Let

$u^k = \{u_1^k, \dots, u_k^k\}$ be the optimal solution for $(QSIP_{T_k})$.

We define a discrete measure μ^k on X by letting

$$\mu^k(s) = \begin{cases} u_i^k \ (\geqslant 0) & \text{if } s = s_i \in T_k, \\ 0 & \text{if } s \in X \big| T_k. \end{cases}$$

In this way, $\mu^k(s) \geqslant 0$, for each $s \in X$. If u^k and (u^k, ν^k) are optimal solutions for $(QSIP_{T_k})$ and $(DQSIP_{T_k})$, then u^k, ν^k satisfy the following Kuhn–Tucker conditions.

$$\int_Y \left[\sum_{i=1}^{k} \phi(s_i, y) u_i^k - g(y) \right] d\nu^k(y) = 0 \tag{10}$$

$$\sum_{j=1}^{k} \left[\int_Y \phi(s_j, y) d\nu^k(y) - \sum_{i=1}^{k} f(s_i, s_j) u_i^k - h(s_j) \right] u_j^k = 0 \tag{11}$$

$$\sum_{i=1}^{k} \phi(s_i, y) u_i^k - g(y) \geqslant 0 \qquad \text{for each } y \text{ in } Y \tag{12}$$

$$\int_Y \phi(s_j, y) d\nu^k(y) - \sum_{i=1}^{k} f(s_i, s_j) u_i^k - h(s_j) \leqslant 0 \quad \forall\, j = 1, 2, \dots, k \tag{13}$$

THEOREM 2.1. *If there exists a $M \in \mathbb{R}$ such that $\mu^k(X) + \nu^k(Y) \leqslant M$ and $f(s, t)$ satisfies condition C, then $V(DQSIP_{T_k}) \to V(DCQP)$.*

Proof: Since $V(DQSIP_{T_1}) \geqslant V(DQSIP_{T_2}) \geqslant \cdots \geqslant V(DCQP) = d$, three cases are conceivable, namely:

Case A. The process stops after a finite number of iterations.
Case B. $\lim_{r\to\infty} V(DQSIP_{T_r}) = d + \eta$, where $\eta > 0$.
Case C. $\lim_{r\to\infty} V(DQSIP_{T_r}) = d$.

When Case A or C occurs then we obtain an optimal value for $(DCQP)$. Now we want to show that Case B is not possible. Since $\|\mu^k\| + \|\nu^k\| \leqslant M$ for each k, the sequence $\{\mu^k\} \subset \mathcal{K}$, a weak* compact subset in $M(X)$ and the sequence $\{\nu^k\} \subset \mathcal{K}'$, a weak* compact subset in $M(Y)$. Since $C(X)$ and $C(Y)$ are separable, there exists a subsequence $\{\mu^{i_k}\}$ of $\{\mu^k\}$ and a subsequence $\{\nu^{i_k}\}$ of $\{\nu^k\}$ such that $\{\mu^{i_k}\}$ is weak* convergent to sume μ^*, and ν^{i_k} is weak* convergent to some ν^*. Then we have $\mu^* \in M^+(X)$ and $\nu^* \in M^+(Y)$. Since $f(s,t)$ satisfies the condition C, we have

$$-\frac{1}{2}\int_X\int_X f(s,t)d\mu^*(s)d\mu^*(t) + \int_Y g(y)d\nu^*(y) = d + \eta.$$

Therefore (μ^*, ν^*) does not belong to the feasible set of $(DCQP)$. Since if (μ^*, ν^*) belongs to the feasible set of $(DCQP)$, then

$$-\frac{1}{2}\int_X\int_X f(s,t)d\mu^*(s)\mu^*(t) + \int_Y g(y)d\mu^*(y) = d + \eta \leqslant d.$$

We get a contradiction. Let

$$\phi_{\mu^*,\nu^*}(s) = \int_Y \phi(s,y)d\nu^*(y) - \int_X f(t,s)d\mu^*(t) - h(s), \tag{14}$$

and

$\overline{s}$ be an element in X that maximizes $\phi_{\mu^*,\nu^*}(s)$.

From the definition of (μ^*, ν^*), we have

$$\phi_{\mu^*,\nu^*}(s_i) \leqslant 0, \qquad i = 1, 2, \ldots \tag{15}$$

Let $\{(\mu^{j_i}, \nu^{j_i})\}$ be a subsequence of $\{(\mu^{k_i}, \nu^{k_i})\}$ such that s_{j_i+1} tends towards a limit point s^*. Because of the choice of s_{j_i+1} in the algorithm we find, for each i

$$\phi_{u^{j_i},\nu^{j_i}}(\overline{s}) \leqslant \phi_{u^{j_i},\nu^{j_i}}(s_{j_i+1}).$$

Letting $i \to \infty$, we have $\phi_{\mu^*,\nu^*}(\overline{s}) \leqslant \phi_{\mu^*,\nu^*}(s^*)$. But by (15) we have $\phi_{\mu^*,\nu^*}(s^*) \leqslant 0$. This contradicts the assumption that $\phi_{\mu^*,\nu^*}(\overline{s}) > 0$ and hence case B is not possible. And hence the result is proved. □

3. An explicit algorithm for CQP

In the above algorithm, we add one constraint at a time and the computational bottleneck falls either in solving the quadratic semi-infinite program $(QSIP_{T_k})$ or in finding a global maximizer s_{k+1} of $\phi_{u^k,v^k}(s)$.

To reduce the computational efforts in solving each $(QSIP_{T_k})$, we propose an explicit algorithm which allows to drop some redundant points in T_k. Note that finding a maximizer s_{k+1} of $\phi_{u^k,v^k}(s)$ in Step 3 usually causes a lot of computational problems. Here our new explicit algorithm explores a new idea whereby the subproblem $(QSIP_{T_{k+1}})$ is constructed by only choosing a point s'_{k+1} at which the infinite constraints are violated, i.e., $\phi_{u^k,v^k}(s'_{k+1}) > 0$, rather than the point s_{k+1} where the violation is maximized, i.e. $\phi_{u^k,v^k}(s)$ is maximized. This idea can be easily incorporated into the existing cutting plane algorithm and potentially reduces the computational burden. In the following, we want to derive an algorithm for CQP by iterative processes.

For the purpose of easy description of the proposed approach, we assume that $(QSIP_{T_k})$ and $(DQSIP_{T_k})$ are solvable, and $V(QSIP_{T_k}) = V(DQSIP_{T_k})$.

Algorithm 2:

Step 0. Let $\delta > 0$ be a sufficiently small number (up to machine accuracy, smaller than 10^{-7}).

Step 1. Set $k \leftarrow 1$, choose any $x_1^1 \in X$, and set $T_1 = \{x_1^1\}$ such that the problem $(QSIP_{T_1})$ has an optimal solution $\{u_1^1\}$ where $u_1^1 > 0$. Solve $(DQSIP_{T_1})$ with an optimal solution (u_1^1, v_1^1). Define $\phi_{u_1^1,v_1^1}$ according to (9).

Step 2. Find $x_2^1 \in X$ such that $\phi_{u_1^1,v_1^1}(x_2^1) > \delta$. If x_2^1 does not exist, stop and output (u_1^1, v_1^1) as a solution. Otherwise, set $T_{k+1} = T_k \cup \{x_2^1\}$, go to Step 6.

Step 3. Solve $(QSIP_{T_k})$ with an optimal solution $u^k = (u_1^k, \ldots, u_{m_{k-1}+1}^k)$, and $(DQSIP_{T_k})$ with an optimal solution (u^k, v^k).

Step 4. Define a discrete measure μ^k on X by letting

$$\mu^k(s) = \begin{cases} u_i^k \ (\geqslant 0) & \text{if } s = x_i^{k-1} \in T_k \\ 0 & \text{if } s \in X \big| T_k. \end{cases}$$

Let $E_k = \{s \in T_k, \mu^k(s) > 0\} = \{x_1^k, x_2^k, \ldots, x_{m_k}^k\}$.

Step 5. Find any $x_{m_k+1}^k \in X$ such that $\phi_{u^k,v^k}(x_{m_k+1}^k) > \delta$. If such $x_{m_k+1}^k$ does not exist, stop and output (u^k, v^k) as the solution. Otherwise, set $T_{k+1} = E_k \cup \{x_{m_k+1}^k\}$.

Step 6. Update $k \leftarrow k + 1$ and go to Step 3.

From the above algorithm, we have

$\{u_1^k, \ldots, u_{m_{k-1}+1}^k\}$ is an optimal solution for $(QSIP_{T_k})$.

Remember

$$E_k = \{s \in T_k,\ \mu^k(s) > 0\} = \{x_1^k, x_2^k, \ldots, x_{m_k}^k\}.$$

Let ν^k be a discrete measure concentrated at $y_1^k, y_2^k, \ldots, y_{n_k}^k$.

Since $V(QSIP_{T_k}) = V(DQSIP_{T_k})$, we have

$$\sum_{j=1}^{m_k}\left[\sum_{i=1}^{n_k}\phi(x_j^k, y_i^k)\nu^k(y_i^k) - \sum_{i=1}^{m_k} f(x_i^k, x_j^k)\mu^k(x_i^k) - h(x_j^k)\right]\mu^k(x_j^k) = 0$$

and

$$\sum_{j=1}^{n_k}\left[\sum_{i=1}^{m_k}\phi(x_i^k, y_j^k)\mu^k(x_i^k) - g(y_j^k)\right]\nu^k(y_j^k) = 0. \tag{16}$$

We define H_k to be an $n_k \times m_k$ matrix with rth row being

$$\left(\phi(x_1^k, y_r^k), \phi(x_2^k, y_r^k), \ldots, \phi(x_{m_k}^k, y_r^k)\right) \qquad r = 1, 2, \ldots, n_k.$$

Since $\nu^k(y_j^k) > 0$ for $j = 1, 2, \ldots, n_k$, from (16), we have

$$H_k \underline{u}_k = g^k, \tag{17}$$

where $\underline{u}_k = (\mu^k(x_1^k), \ldots, \mu^k(x_{m_k}^k))$ and $g^k = (g(y_1^k), \ldots, g(y_{n_k}^k))$. Let ϕ_{μ^k,ν^k} be as in (9), i.e., $\phi_{\mu^k,\nu^k}(s) = \int_Y \phi(s,y)d\nu^k(y) - \int_X f(t,s)d\mu^k(t) - h(s)$. For the primal problem, we define

$$\overline{\phi}_{\mu^k}(y) = \int_X \phi(s,y)d\mu^k(s) - g(y). \tag{18}$$

Note that $\overline{\phi}_{\mu^k}(y) \geqslant 0\ \forall\, y \in Y$, and we have the following result:

THEOREM 3.1. *Let μ^k, μ^{k+1}, ν^k and ν^{k+1} be generated in the kth step of the algorithm given as above and $\Delta^k = \mu^{k+1} - \mu^k$. Then*

$$\begin{aligned} &V(DQSIP_{T_{k+1}}) - V(DQSIP_{T_k}) \\ &= \sum_{i=1}^{m_k}\phi_{\mu^{k+1},\nu^{k+1}}(x_i^k)\mu^k(x_i^k) - \sum_{j=1}^{n_{k+1}}\overline{\phi}_{\mu^k}(y_j^{k+1})\nu^{k+1}(y_j^{k+1}) \\ &\quad - \frac{1}{2}\int_X\int_X f(s,t)d\Delta^k(s)\, d\Delta^k(t) \end{aligned} \tag{19}$$

Proof: By the definition of $\overline{\phi}_{\mu^k}(y)$, we have

$$
\begin{aligned}
&\sum_{j=1}^{n_{k+1}} \overline{\phi}_{\mu^k}(y_j^{k+1})\nu^{k+1}(y_j^{k+1}) \\
&= \sum_{j=1}^{n_{k+1}} \left(\sum_{i=1}^{m_k} \phi(x_i^k, y_j^{k+1})\mu^k(x_i^k) - g(y_j^{k+1}) \right) \nu^{k+1}(y_j^{k+1}) \\
&= \sum_{j=1}^{n_{k+1}} \sum_{i=1}^{m_k} \phi(x_i^k, y_j^{k+1})\mu^k(x_i^k)\nu^{k+1}(y_j^{k+1}) - \sum_{j=1}^{n_{k+1}} g(y_j^{k+1})\nu^{k+1}(y_j^{k+1}) \\
&= \sum_{i=1}^{m_k} \left(\sum_{j=1}^{n_{k+1}} \phi(x_i^k, y_j^{k+1})\nu^{k+1}(y_j^{k+1}) \right) \mu^k(x_i^k) - \sum_{j=1}^{n_{k+1}} g(y_j^{k+1})\nu^{k+1}(y_j^{k+1}) \\
&= \sum_{i=1}^{m_k} \left(\phi_{\mu^{k+1},\nu^{k+1}}(x_i^k) + \sum_{j=1}^{m_{k+1}} f(x_j^{k+1}, x_i^k)\mu^{k+1}(x_j^{k+1}) + h(x_i^k) \right) \mu^k(x_i^k) \\
&\quad - \sum_{j=1}^{n_{k+1}} g(y_j^{k+1})\nu^{k+1}(y_j^{k+1}) \\
&= \sum_{i=1}^{m_k} \phi_{\mu^{k+1},\nu^{k+1}}(x_i^k)\mu^k(x_i^k) + \sum_{i=1}^{m_k} \sum_{j=1}^{m_{k+1}} f(x_j^{k+1}, x_i^k)\mu^{k+1}(x_j^{k+1})\mu^k(x_i^k) \\
&\quad + \sum_{i=1}^{m_k} h(x_i^k)\mu^k(x_i^k) - \sum_{j=1}^{n_{k+1}} g(y_j^{k+1})\nu^{k+1}(y_j^{k+1}).
\end{aligned}
\tag{20}
$$

It is obvious that

$$
\sum_{i=1}^{m_k} \sum_{j=1}^{m_{k+1}} f(x_j^{k+1}, x_i^k)\mu^{k+1}(x_j^{k+1})\mu^k(x_i^k) = \int_X \int_X f(s,t)d\mu^{k+1}(s)d\mu^k(t).
$$

Now $\Delta^k = \mu^{k+1} - \mu^k$. Thus

$$
\begin{aligned}
&\int_X \int_X f(s,t)d\mu^{k+1}(s)d\mu^k(t) \\
&= \frac{1}{2} \int_X \int_X f(s,t)d\mu^{k+1}(s)d\mu^k(t) + \frac{1}{2} \int_X \int_X f(s,t)d\mu^{k+1}(s)d\mu^k(t) \\
&= \frac{1}{2} \int_X \int_X f(s,t)d\mu^k(s)d\mu^k(t) + \frac{1}{2} \int_X \int_X f(s,t)d\Delta^k(s)d\mu^k(t) \\
&\quad + \frac{1}{2} \int_X \int_X f(s,t)d\mu^{k+1}(s)d\mu^{k+1}(t) - \frac{1}{2} \int_X \int_X f(s,t)d\mu^{k+1}(s)d\Delta^k(t) \\
&= \frac{1}{2} \int_X \int_X f(s,t)d\mu^k(s)d\mu^k(t) + \frac{1}{2} \int_X \int_X f(s,t)d\mu^{k+1}(s)d\mu^{k+1}(t) \\
&\quad + \frac{1}{2} \int_X \int_X f(s,t)d\Delta^k(s)d\mu^k(t) - \frac{1}{2} \int_X \int_X f(s,t)d\mu^{k+1}(s)d\Delta^k(t).
\end{aligned}
$$

Since f is a symmetric function, we have

$$\int_X \int_X f(s,t)d\Delta^k(s)d\mu^k(t) = \int_X \int_X f(t,s)d\Delta^k(s)d\mu^k(t)$$
$$= \int_X \int_X f(s,t)d\Delta^k(t)d\mu^k(s).$$

Therefore

$$\frac{1}{2}\int_X \int_X f(s,t)d\Delta^k(s)d\mu^k(t) - \frac{1}{2}\int_X \int_X f(s,t)d\mu^{k+1}(s)d\Delta^k(t)$$
$$= -\frac{1}{2}\int_X \int_X f(s,t)d(\mu^{k+1}-\mu^k)(s)d\Delta^k(t)$$
$$= -\frac{1}{2}\int_X \int_X f(s,t)d\Delta^k(s)d\Delta^k(t).$$

It follows that

$$(20) = \sum_{i=1}^{m_k} \phi_{\mu^{k+1},\nu^{k+1}}(x_i^k)\mu^k(x_i^k) + \frac{1}{2}\int_X \int_X f(s,t)d\mu^k(s)d\mu^k(t)$$
$$+ \frac{1}{2}\int_X \int_X f(s,t)d\mu^{k+1}(s)d\mu^{k+1}(t) - \frac{1}{2}\int_X \int_X f(s,t)d\Delta^k(s)d\Delta^k(t)$$
$$+ \sum_{i=1}^{m_k} h(x_i^k)\mu^k(x_i^k) - \sum_{j=1}^{n_{k+1}} g(y_j^{k+1})\nu^{k+1}(y_j^{k+1})$$
$$= \sum_{i=1}^{m_k} \phi_{\mu^{k+1},\nu^{k+1}}(x_i^k)\mu^k(x_i^k) + V(QSIP_{T_k}) - V(DQSIP_{T_{k+1}})$$
$$- \frac{1}{2}\int_X \int_X f(s,t)d\Delta^k(s)d\Delta^k(t).$$

Thus

$$V(DQSIP_{T_{k+1}}) - V(DQSIP_{T_k})$$
$$= \sum_{i=1}^{m_k} \phi_{\mu^{k+1},\nu^{k+1}}(x_i^k)\mu^k(x_i^k) - \sum_{j=1}^{n_{k+1}} \overline{\phi}_{\mu^k}(y_j^{k+1})\nu^{k+1}(y_j^{k+1})$$
$$- \frac{1}{2}\int_X \int_X f(s,t)d\Delta^k(s)d\Delta^k(t).$$

□

Since $\mu^k(x_i^k) > 0$ for $i = 1, \ldots, m_k$, $\nu^{k+1}(y_j^{k+1}) > 0$ for $j = 1, \ldots, n_{k+1}$, and f is a positive semi-definite function, from (19) we know that $V(DQSIP_{T_{k+1}}) < V(DQSIP_{T_k})$ if and only if $\exists\, i \in \{1, 2, \ldots, m_k\}$ such that $\phi_{\mu^{k+1},\nu^{k+1}}(x_i^k) \neq 0$

or $\exists\, j \in \{1, 2, \ldots, n_{k+1}\}$ such that $\overline{\phi}_{\mu^k}(y_j^{k+1}) \neq 0$. Thus we have the following theorem.

THEOREM 3.2. *Let μ^k, μ^{k+1}, ν^k, and ν^{k+1} be generated in the kth and $k+1$ th steps of the algorithm given as above. Then*

$V(QSIP_{T_{k+1}}) < V(QSIP_{T_k})$ if and only if $\phi_{\mu^{k+1},\nu^{k+1}}$ or $\overline{\phi}_{\mu^k}$ satisfies one of the following conditions:

(i) $\exists\, i \in \{1, 2, \ldots, m_k\}$ such that $\phi_{\mu^{k+1},\nu^{k+1}}(x_i^k) \neq 0$.
(ii) $\exists\, j \in \{1, 2, \ldots, n_{k+1}\}$ such that $\overline{\phi}_{\mu^k}(y_j^{k+1}) \neq 0$.

In general, since (μ^k, ν^k) and (μ^{k+1}, ν^{k+1}) are different, $\phi_{\mu^{k+1},\nu^{k+1}}$ or $\overline{\phi}_{\mu^k}$ the satisfaction of condition (i) or (ii) in Theorem 3.2 is not a problem. If $V(DQSIP_{T_{k+1}}) < V(DQSIP_{T_k})$, then we must have $x_{m_k+1}^k \in E_{k+1}$. If $x_{m_k+1}^k \notin E_{k+1}$, then $\mu^{k+1}(x_{m_k+1}^k) = 0$ and μ^{k+1} has nonzero measure only at these points in E^k. Thus we have $V(DQSIP_{T_{k+1}}) = V(DQSIP_{T_k})$ which contradicte to $V(DQSIP_{T_{k+1}}) < V(DQSIP_{T_k})$. Now we let $x_{m_k+1}^k = x_{m_{k+1}}^{k+1}$.

Let $X = [a, b]$ and $Y = [c, d]$. From now on, let $h \in C^\infty([a, b])$, $g \in C^\infty([c, d])$, $\phi \in C^\infty([a, b] \times [c, d])$, and $f \in C^\infty([a, b] \times [a, b])$. Let $M^* > 0$ be a sufficient large such that $|\phi|$, $|f|$, $|g|$, and $|h|$ are bounded by M^*. Let $\delta^* > 0$ be small enough. Consider a function $c \in C^\infty([a, b])$ with $c(t) \leqslant 0$. We will assume that c satisfies the following regularity assumptions (RA):

(i) $c(t) = 0$ only at $t_k \in [c, d]$, $k = 1, \ldots, m$, and $\forall\, k$ there exists an i_k such that the jth derivative $c^{(j)}(t_k) = 0$, $j = 0, 1, \ldots, i_k - 1$, but $c^{(i_k)}(t_k) \neq 0$, $i_k < M^*$.
(ii) (a) For each k, if $t \in N_{\delta^*}(t_k)$ the δ^*–neighborhood of t_k, then $c^{(i_k)}(t)$ has the same sign as $c^{(i_k)}(t_k)$, and $M^* \geqslant \left|c^{(i_k)}(t)\right| \geqslant \delta^*$.
(b) If $\overline{t}$ is a local maximum or minimum of c in $[a, b]$ and $c(\overline{t}) \neq 0$, then $c(\overline{t}) \leqslant -\delta^*$.

The following lemma was proved by Lai and Wu [9].

LEMMA 3.3. *Let $c \in C^\infty([a, b])$ with $c(t) \leqslant 0$ satisfy the condition (RA). Suppose that, for $t \in [a, b]$, $|c(t)| < \varepsilon$ for a sufficient small $\varepsilon > 0$. Then there exists an $\ell \in \{1, 2, \ldots, m\}$ such that $|t - t_\ell| < r(\varepsilon)$ with $r(\varepsilon) \to 0$ whenever $\varepsilon \to 0$.*

We define B_r^k to be a matrix having row vectors

$$\left(\frac{\partial^j}{\partial y^j}\phi(x_1^k, y_r^k), \frac{\partial^j}{\partial y^j}\phi(x_2^k, y_r^k), \ldots, \frac{\partial^j}{\partial y^j}\phi(x_{m_k}^k, y_r^k)\right) \qquad 0 \leqslant j \leqslant 1$$

and

$$B^k = \begin{pmatrix} B_1^k \\ B_2^k \\ \vdots \\ B_{n_k}^k \end{pmatrix}.$$

Since $\overline{\phi}_{\mu^k}(y) \in C^\infty([c,d])$ and $\overline{\phi}_{\mu^k}(y_r^k) = 0$ for $r = 1, \ldots, n_k$, we have $B^k \underline{u}_k = (\overline{g}^k)$ where $\underline{u}_k = (\mu^k(x_1^k), \ldots, \mu^k(x_{m_k}^k))$ and $\overline{g}^k = (g(y_1^k), g'(y_1^k), \ldots, g(y_{n_k}^k), g'(y_{n_k}^k))$.

For μ^k and ν^k in the kth step of algorithm given in Theorem 2.1, we assume that μ^k and ν^k satisfy the following conditions (A):

(A1) $V(DQSIP_{T_{k+1}}) < V(DQSIP_{T_k})$;
(A2) $\|\mu^k\| + \|\nu^k\| \leqslant M^*$;
(A3) $\nu^k(y_j^k) \geqslant \delta$ for $j = 1, \ldots, n_k$,
(A4) There exists a $\overline{K}$ such that $\forall\, k \geqslant \overline{K}$ we have a square submatrix A^k of B^k having rank m_k such that $|\mathrm{Det}(A^k)| > \delta$, and a square submatrix D^k of $(H_k)^T$ having rank n_k such that $|\mathrm{Det}(D^k)| > \delta$;
(A5) $\exists\, \overline{N}$ such that $\forall\, k \geqslant \overline{N}$ we have $x_{m_k}^k$ belongs to one of $N_{\delta_{k-1}}(x_i^{k-1})$, $i = 1, 2, \ldots, m_{k-1}$, where $\delta_{k-1} \to 0$ as $k \to \infty$.

Since $T_{k+1} = E^k \cup \{x_{m_k+1}^k\}$ and $x_{m_k+1}^k = x_{m_{k+1}}^{k+1} \in E^{k+1}$, we have $x_i^{k+1} \in E^k$ for $i = 1, 2, \ldots, m_{k+1}-1$. Without loss of genernality, we may assume that $x_i^{k+1} = x_i^k$ for $i = 1, 2, \ldots, m_{k+1} - 1$. From (A4) and (A5), we can choose $k > \overline{N}'$ large enough such that $x_{m_{k+1}}^{k+1} \notin \bigcup_{i=1}^{m_{k+1}-1} N_{\delta_k}(x_i^k)$. Thus $x_{m_{k+1}}^{k+1}$ belongs to one of $N_{\delta_k}(x_i^k)$ $i = m_{k+1}, \ldots, m_k$. Likewise, we may assume that

$$x_{m_{k+1}}^{k+1} \in N_{\delta_k}(x_{m_{k+1}}^k). \tag{21}$$

Therefore, we immediately have $m_k \geqslant m_{k+1}$.

Under the assumptions (A), we next prove that the above algorithm stops in a finite number of iterations.

THEOREM 3.4. *Given any $\delta > 0$, in each iteration, assume that:*

(a) the measures μ^k and ν^k satisfy the conditions (A1) to (A5), and $\overline{\phi}_{\mu^k}(y)$ satisfies the (RA) condition;
(b) $\phi_{\mu^{k+1}\nu^{k+1}}(s) \leqslant -\delta$ if $s \in T_{k+1} - E_{k+1} - \{x_{m_{k+1}}^k\}$.

Then the algorithm stops in a finite number of iterations.

Proof. Suppose that the algorithm does not stop in a finite number of iterations. From (a) we have

$$V(DQSIP_{T_1}) > V(DQSIP_{T_2}) > V(DQSIP_{T_3}) > \cdots \geqslant V(DCQP) \quad (22)$$

Thus $\lim_{r\to\infty} V(DQSIP_{T_r}) = \alpha \geqslant V(DCQP)$. We claim that $\alpha > V(DCQP)$ is impossible.

Since $\|\mu^k\| + \|\nu^k\| \leqslant M^*$ for each k, the sequence $\{\mu^k\} \subset \mathcal{K}$, a weak* compact subset in $M(X)$, and the sequence $\{\nu^k\} \subset \mathcal{K}'$, a weak* compact subset in $M(Y)$. Since $C(X)$ and $C(Y)$ are separable, there exist subsequences $\{\mu^{i_k}\}$ of $\{\mu^k\}$ and $\{\nu^{i_k}\}$ of $\{\nu^k\}$ such that μ^{i_k} is weak* convergent to some μ^*, ν^{i_k} is weak* convergent to some ν^*, and $\{x^{i_k}_{m_{i_k}+1}\}$ converges to some point x_* as $k \to \infty$. Then we have $\mu^* \in M^+(X)$ and $\nu^* \in M^+(Y)$.

Now let

$$\phi_{\mu^*,\nu^*}(s) = \int_Y \phi(s,y)d\nu^*(y) - \int_X f(t,s)d\mu^*(t) - h(s). \quad (23)$$

Then $\phi_{\mu^{i_k},\nu^{i_k}}(x^{i_k}_{m_{i_k}+1})$ converges to $\phi_{\mu^*,\nu^*}(x_*)$. Since $\phi_{\mu^{i_k},\nu^{i_k}}(x^{i_k}_{m_{i_k+1}}) > \delta$ for each k, we have $\phi_{\mu^*,\nu^*}(x_*) \geqslant \delta$.

Now let ε in $(0,\delta)$ be arbitrary. We can find a large integer $N \in \{i_k\}_{k=1}^{\infty}$ such that

$$|V(DQSIP_{T_N}) - \alpha| < \varepsilon^2 \text{ and } |\phi_{\mu^N,\nu^N}(x^N_{m_N+1}) - \phi_{\nu^*,\nu^*}(x_*)| < \varepsilon^2. \quad (24)$$

From Theorem 3.1 and (22), we have

$$\begin{aligned} &|V(DQSIP_{T_{N+1}}) - V(DQSIP_{T_N})| \\ =&\Big|\sum_{i=1}^{m_N} \phi_{\mu^{N+1},\nu^{N+1}}(x_i^N)\mu^N(x_i^N) - \sum_{j=1}^{n_{N+1}} \overline{\phi}_{\mu^N}(y_j^{N+1})\nu^{N+1}(y_j^{N+1}) \\ &- \frac{1}{2}\int_X\int_X f(s,t)d\Delta^N(s)d\Delta^N(t)\Big| < \varepsilon^2. \end{aligned} \quad (25)$$

Since $\phi_{\mu^{N+1},\nu^{N+1}}(x_i^N) \leqslant 0$ for $i = 1,\ldots,m_N$ and $\overline{\phi}_{\mu^N}(y_j^{N+1}) \geqslant 0$ for $j = 1,\ldots,n_{N+1}$, each term in (25) is nonpositive, and is $\geqslant -\varepsilon^2$. We can get

$$\left|\overline{\phi}_{\mu^N}(y_j^{N+1})\right| < \frac{\varepsilon^2}{\nu^{N+1}(y_j^{N+1})} \qquad \text{for } j = 1,\ldots,n_{N+1}.$$

Since $\nu^{N+1}(y_j^{N+1}) \geqslant \delta$ for $j = 1,2,\ldots,n_{N+1}$, we have

$$\left|\overline{\phi}_{\mu^N}(y_j^{N+1})\right| < \varepsilon \qquad \text{for } j = 1,\ldots,n_{N+1}.$$

Thus, from assumption (a) and Lemma 3.3, we can choose ε so small and find $r_j(\varepsilon)$ such that $y_j^{N+1} \in N_{r_j(\varepsilon)}(y_{r_j}^N)$, and y_i^{N+1}, y_j^{N+1}, for $i \neq j$ are respectively in the disjoint neighborhoods $N_{r_i(\varepsilon)}(y_{r_i}^N)$, $N_{r_j(\varepsilon)}(y_{r_j}^N)$.

From (25), we have that if $N > \overline{N}$ and $m_N > m_{N+1}$, then

$$\left|\phi_{\mu^{N+1},\nu^{N+1}}(x_i^N)\mu^N(x_i^N)\right| < \varepsilon^2 \qquad \text{for } i = m_{N+1}+1, \ldots, m_N.$$

From (b), we have

$$\left|\phi_{\mu^{N+1},\nu^{N+1}}(x_i^N)\right| \geqslant \delta \qquad \text{for } i = m_{N+1}+1, \ldots, m_N,$$

thus

$$\mu^N(x_i^N) < \varepsilon \qquad \text{for } i = m_{N+1}+1, \ldots, m_N. \tag{26}$$

Now we assume that $N > \overline{N}$. Since $|y_i^{N+1} - y_{r_i}^N| < r_i(\varepsilon)$ for $i = 1, 2, \ldots, n_{N+1}$, $x_i^{N+1} = x_i^N$ for $i = 1, 2, \ldots, m_{N+1}-1$, $x_{m_{N+1}}^{N+1} \in N_{\delta_N}(x_{m_{N+1}}^N)$, and $\phi \in C^\infty([a,b]\times[c,d])$, thus as $N \to \infty$, we have

$$\begin{aligned} &\phi(x_i^{N+1}, y_j^{N+1}) - \phi(x_i^N, y_{r_j}^N) \to 0 \\ &\frac{\partial}{\partial y}\phi(x_i^{N+1}, y_j^{N+1}) - \frac{\partial}{\partial y}\phi(x_i^N, y_{r_j}^N) \to 0 \\ &\qquad \text{for } i = 1, 2, \ldots, m_{N+1}, \quad j = 1, 2, \ldots, n_{N+1}. \end{aligned} \tag{27}$$

From the definition of $\overline{\phi}_{\mu^{N+1}}(y)$ in (18), we have

$$\sum_{i=1}^{m_{N+1}} \phi(x_i^{N+1}, y_j^{N+1})\mu^{N+1}(x_i) = g(y_j^{N+1}) \qquad \text{for } j = 1, 2, \ldots, n_{N+1} \tag{28}$$

and

$$\sum_{i=1}^{m_{N+1}} \frac{\partial}{\partial y}\phi(x_i^{N+1}, y_j^{N+1})\mu^{N+1}(x_i) = g'(y_j^{N+1}) \quad \text{for } j = 1, 2, \ldots, n_{N+1}. \tag{29}$$

Let W_N be a matrix with row vectors

$$\big(\phi(x_1^{N+1}, y_j^{N+1}), \phi(x_2^{N+1}, y_j^{N+1}), \ldots, \phi(x_{m_{N+1}}^{N+1}, y_j^{N+1})\big), \quad j = 1, 2, \ldots, n_{N+1}$$

and $\overline{W_N}$ be a matrix with row vectors

$$\left(\frac{\partial}{\partial y}\phi(x_1^{N+1}, y_j^{N+1}), \frac{\partial}{\partial y}\phi(x_2^{N+1}, y_j^{N+1}), \ldots, \frac{\partial}{\partial y}\phi(x_{m_{N+1}}^{N+1}, y_j^{N+1})\right), \quad j = 1, 2, \ldots, n_{N+1}.$$

Therefore we can express (28) and (29) as

$$\begin{pmatrix} W_N \\ \overline{W}_N \end{pmatrix} (\underline{u}_{N+1}) = \begin{pmatrix} g_{N+1} \\ g'_{N+1} \end{pmatrix} \tag{30}$$

where

$$\underline{u}_{N+1} = \left(\mu^{N+1}(x_1^{N+1}), \mu^{N+1}(x_2^{N+1}), \dots, \mu^{N+1}(x_{m_{N+1}}^{N+1})\right),$$
$$g_{N+1} = \left(g(y_1^{N+1}), \dots, g(y_{n_{N+1}}^{N+1})\right), \text{ and } g'_{N+1} = \left(g'(y_1^{N+1}), \dots, g'(y_{n_{N+1}}^{N+1})\right).$$

From the definition of $\overline{\phi}_{\mu^N}(y)$ in (18), we have

$$\sum_{i=1}^{m_{N+1}} \phi(x_i^N, y_{r_j}^N)\mu^N(x_i^N) + \sum_{i=m_{N+1}+1}^{m_N} \phi(x_i^N, y_{r_j}^N)\mu^N(x_i^N) = g(y_{r_j}^N)$$
$$\text{for } j = 1, \dots, n_{N+1}, \tag{31}$$

and

$$\sum_{i=1}^{m_{N+1}} \frac{\partial}{\partial y}\phi(x_i^N, y_{r_j}^N)\mu^N(x_i^N) + \sum_{i=m_{N+1}+1}^{m_N} \frac{\partial}{\partial y}\phi(x_i^N, y_{r_j}^N)\mu^N(x_i^N) = g'(y_{r_j}^N)$$
$$\text{for } j = 1, \dots, n_{N+1}. \tag{32}$$

From (26), we have

$$\sum_{i=1}^{m_{N+1}} \phi(x_i^N, y_{r_j}^N)\mu^N(x_i^N) = g(y_{r_j}^N) + \varepsilon_j(\varepsilon), \tag{33}$$

where $\varepsilon_j(\varepsilon) \to 0$ as $\varepsilon \to 0$ for $j = 1, 2, \dots, n_{N+1}$, and

$$\sum_{i=1}^{m_{N+1}} \frac{\partial}{\partial y}\phi(x_i^N, y_{r_j}^N)\mu^N(x_i^N) = g'(y_{r_j}^N) + \varepsilon'_j(\varepsilon), \tag{34}$$

where $\varepsilon'_j(\varepsilon) \to 0$ as $\varepsilon \to 0$ for $j = 1, 2, \dots, n_{N+1}$.

Let G_N be a matrix with row vectors

$$\left(\phi(x_1^N, y_{r_j}^N), \phi(x_2^N, y_{r_j}^N), \dots, \phi(x_{m_{N+1}}^N, y_{r_j}^N)\right), \quad j = 1, 2, \dots, n_{N+1},$$

and $\overline{G}_N$ be a matrix with row vectors

$$\left(\frac{\partial}{\partial y}\phi(x_1^N, y_{r_j}^N), \frac{\partial}{\partial y}\phi(x_2^N, y_{r_j}^N), \dots, \frac{\partial}{\partial y}\phi(x_{m_{N+1}}^N, y_{r_j}^N)\right),$$
$$j = 1, 2, \dots, n_{N+1}.$$

Therefore, we can express (33) and (34) as

$$\begin{pmatrix} G_N \\ \overline{G}_N \end{pmatrix} (\underline{u}'_N) = \begin{pmatrix} \overline{g}_N \\ \overline{g}'_N \end{pmatrix}, \tag{35}$$

where

$$\underline{u}'_N = \big(\mu^N(x_1^N), \ldots, \mu^N(x_{m_{N+1}}^N)\big),$$
$$\overline{g}_N = \big(g(y_{r_1}^N) + \varepsilon_1(\varepsilon), \ldots, g(y_{r_{n_{N+1}}}^N) + \varepsilon_{n_{N+1}}(\varepsilon)\big),$$

and

$$\overline{g}'_N = \big(g(y_{r_1}^N) + \varepsilon'_1(\varepsilon), \ldots, g(y_{r_{n_{N+1}}}^N) + \varepsilon'_{n_{N+1}}(\varepsilon)\big).$$

Since $\begin{pmatrix} W_N \\ \overline{W}_N \end{pmatrix}$ has rank m_{N+1}, from (27) and assumption (a), we see that $\begin{pmatrix} G_N \\ \overline{G}_N \end{pmatrix}$ has rank m_{N+1} and

$$\big|\mu^{N+1}(x_i^{N+1}) - \mu^N(x_i^N)\big| < \overline{\varepsilon}_i(\varepsilon), \qquad \text{for } i = 1, \ldots, m_{N+1}, \tag{36}$$

where $\overline{\varepsilon}_i(\varepsilon) \to 0$ for $i = 1, 2, \ldots m_{N+1}$ whenever $\varepsilon \to 0$.

From (16), we have

$$\sum_{j=1}^{n_N} \phi(x_i^N, y_j^N)\nu^N(y_j^N) = \sum_{j=1}^{m_N} f(x_j^N, x_i^N)\mu^N(x_j^N) + h(x_i^N),$$
$$i = 1, \ldots, m_{N+1} \tag{37}$$

and

$$\sum_{j=1}^{n_{N+1}} \phi(x_i^{N+1}, y_j^{N+1})\nu^{N+1}(y_j^{N+1})$$
$$= \sum_{j=1}^{m_{N+1}} f(x_j^{N+1}, x_i^{N+1})\mu^{N+1}(x_j^{N+1}) + h(x_i^{N+1}), \qquad i = 1, \ldots, m_{N+1} \tag{38}$$

From (37), it follows that

$$\sum_{j=1}^{n_{N+1}} \phi(x_i^N, y_{r_j}^N)\nu^N(y_{r_j}^N) + \sum_{k=1}^{n_N - n_{N+1}} \phi(x_i^N, y_{t_k}^N)\nu^N(y_{t_k}^N)$$
$$= \sum_{j=1}^{m_N} f(x_j^N, x_i^N)\mu^N(x_j^N) + h(x_i^N), \qquad i = 1, 2, \ldots, m_{N+1}. \tag{39}$$

Let Q_N be a matrix with row vectors

$$\big(\phi(x_i^N, y_{r_1}^N), \phi(x_i^N, y_{r_2}^N), \ldots, \phi(x_i^N, y_{r_{n_{N+1}}}^N)\big), \qquad i = 1, 2, \ldots, m_{N+1},$$

and $\overline{Q}_N$ be a matrix with row vectors

$$\big(\phi(x_i^N, y_{t_1}^N), \phi(x_i^N, y_{t_2}^N), \ldots, \phi(x_i^N, y_{t_{n_N - n_{N+1}}}^N)\big), \qquad i = 1, 2, \ldots, m_{N+1}.$$

Then we can express (39) as

$$(Q_N \;\vdots\; \overline{Q}_N)(\underline{v}_N) = (\ell_N),$$

where

$$\underline{v}_N = \big(\nu^N(y_{r_1}^N), \ldots, \nu^N(y_{r_{n_{N+1}}}^N), \nu^N(y_{t_1}^N), \ldots, \nu^N(y_{t_{n_N - n_{N+1}}}^N)\big)$$

and

$$\ell_N = \Big(\sum_{j=1}^{m_N} f(x_j^N, x_1^N)\mu^N(x_j^N) + h(x_1^N), \ldots, \\ \sum_{j=1}^{m_N} f(x_j^N, x_{m_{N+1}}^N)\mu^N(x_j^N) + h(x_{m_{N+1}}^N)\Big).$$

Let R_N be a matrix with row vectors

$$\big(\phi(x_i^{N+1}, y_1^{N+1}), \phi(x_i^{N+1}, y_2^{N+1}), \ldots, \phi(x_i^{N+1}, y_{n_{N+1}}^{N+1})\big), \\ i = 1, 2, \ldots, m_{N+1}.$$

Then we can express (38) as

$$R_N \underline{v}'_{N+1} = \ell'_{N+1}$$

where

$$\underline{v}'_{N+1} = \big(\nu^{N+1}(y_1^{N+1}), \ldots, \nu^{N+1}(y_{n_{N+1}}^{N+1})\big)$$

and

$$\ell'_{N+1} = \left(\sum_{j=1}^{m_{N+1}} f(x_j^{N+1}, x_1^{N+1})\mu^{N+1}(x_j^{N+1}) + h(x_1^{N+1}), \ldots \right. \\ \left. \ldots, \sum_{j=1}^{m_{N+1}} f(x_j^{N+1}, x_{m_{N+1}}^{N+1})\mu^{N+1}(x_j^{N+1}) + h(x_{m_{N+1}}^{N+1})\right).$$

Thus

$$(R_N \vdots \overline{Q}_N)\begin{pmatrix} \underline{\nu}'_{N+1} \\ O \end{pmatrix} = \ell'_{N+1}.$$

Combining (27), the assumption (a) and the fact that $(Q_N \vdots \overline{Q}_N)$ has rank n_N, and also taking into account of $x^{N+1}_{m_{N+1}} \in N_{\delta_N}(x^N_{m_{N+1}})$, (26) and (36), we see that $(R_N \vdots \overline{Q}_N)$ has rank n_N and

$$\left|\nu^N(y^N_{r_i}) - \nu^{N+1}(y^{N+1}_i)\right| < \overset{\Delta}{\varepsilon}_i\,(\varepsilon), \quad \text{for } i = 1, \ldots, n_{N+1},$$
$$\left|\nu^N(y^N_{t_i})\right| < \varepsilon^*_i(\varepsilon), \quad \text{for } i = 1, \ldots, n_N - n_{N+1},$$

where

$$\overset{\Delta}{\varepsilon}_i\,(\varepsilon) \to 0 \text{ and } \varepsilon^*_i(\varepsilon) \to 0, \text{ as } N \to \infty. \tag{40}$$

Since $\phi_{\mu^{N+1},\nu^{N+1}}(x^{N+1}_{m_{N+1}}) = 0$, from (26), (36) and (40), we have

$$\phi_{\mu^N,\nu^N}(x^{N+1}_{m_{N+1}}) \to 0 \text{ as } N \to \infty. \tag{41}$$

But

$$\phi_{\mu^N,\nu^N}(x^{N+1}_{m_{N+1}}) = \phi_{\mu^N,\nu^N}(x^N_{m_{N+1}}) \to \phi_{\mu^*,\nu^*}(x_*) \neq 0 \text{ as } N \to \infty.$$

Therefore (41) cannot hold, and we have a contradiction. Therefore our claim is valid and the proof is now completed. □

Note that (A2) is commonly assumed in infinite linear programming to simplify proofs. It can be relaxed by using bounded level sets. (A1) is also a technical condition commonly used in linear programming. Moreover, when $\delta > 0$ is chosen to be sufficiently small, (A3), (A4), (A5), and (b) in Theorem 3.4 in general can be satisfied without much difficulty. The violation of any of these four assumptions will lead to some rare instances of degeneracy or inconsistency.

Theorem 3.4 assures that the proposed scheme terminates in finitely many iterations, say k^* iterations, with an optimal solution (μ^{k^*}, ν^{k^*}). Recall that μ^{k^*} is a discrete measure concentrated at the points $x^{k^*}_1, x^{k^*}_2, \ldots, x^{k^*}_{m_{k^*}}$, and ν^{k^*} is a discrete measure concentrated at the points $y^{k^*}_1, y^{k^*}_2, \ldots, y^{k^*}_{n_{k^*}}$. In this case (μ^{k^*}, ν^{k^*}) is, of course, feasible for $(DQSIP_{T_{k^*}})$, and from Theorem 3.5, it can be viewed that (μ^{k^*}, ν^{k^*}) is an approximate solution of $(DCQP)$, if δ is chosen small. It is important to know how ‘good’ such an approximate solution is. The following theorem concerns this issue.

THEOREM 3.5. *For any given* $\delta > 0$, *if there exists* $\overline{\mu} \in M(X)$ *and* $\overline{\nu} \in M^+(Y)$ *with* $\overline{\nu}(y_i^{k^*}) \geqslant -\dfrac{\nu^{k^*}(y_i^{k^*})}{\delta}$ *for* $i = 1, 2, \ldots, n_{k^*}$, *and* $\overline{\nu}(B) \geqslant 0$ *for all Borel set* $B \subset Y - \{y_1^{k^*}, y_2^{k^*}, \ldots, y_{n_{k^*}}^{k^*}\}$, *and such that*

$$\int_Y \phi(x, y) d\overline{\nu}(y) - \int_X f(x, t) d\overline{\mu}(t) \leqslant -1, \tag{42}$$

then

$$\begin{aligned} &|V(DQSIP_{T_{k^*}}) - V(DCQP)| \\ &\leqslant \delta \Big| \int_X \int_X f(s, t) d\mu^{k^*}(s) d\overline{\mu}(t) + \frac{1}{2}\delta \int_X \int_X f(s, t) d\overline{\mu}(s) d\overline{\mu}(t) \\ &\quad - \int_Y g(y) d\overline{\nu}(y) \Big|. \end{aligned} \tag{43}$$

Proof. It is obvious that $\nu^{k^*} + \delta\overline{\nu} \in M^+(X)$, $\mu^{k^*} + \delta\overline{\mu} \in M(X)$, and

$$\begin{aligned} &\int_Y \phi(s, y) d\nu^{k^*}(y) + \delta d\overline{\nu}(y) - \int_X f(t, s) d\mu^{k^*}(t) + \delta d\overline{\mu}(t) - h(s) \\ &= \int_Y \phi(s, y) d\nu^{k^*}(y) - \int_X f(t, s) d\mu^{k^*}(t) - h(s) \\ &\quad + \delta \left(\int_Y \phi(s, y) d\overline{\nu}(y) + \int_X f(t, s) d\overline{\mu}(t) \right) \leqslant \delta + (-\delta) = 0 \quad \text{(from (42))}. \end{aligned}$$

Thus $(\nu^{k^*}, \mu^{k^*}) + \delta(\overline{\nu}, \overline{\mu})$ is feasible for $DCQP$.

From (21), we know that

$$\begin{aligned} &|V(DQSIP_{T_{k^*}}) - V(DCQP)| \\ &\leqslant \Big| -\frac{1}{2} \int_X \int_X f(s, t) d\mu^{k^*}(s) d\mu^{k^*}(t) + \int_Y g(y) d\nu^{k^*}(y) \\ &\quad + \frac{1}{2} \int_X \int_X f(s, t) d(\mu^{k^*} + \delta\overline{\mu})(s) d(\mu^{k^*} + \delta\overline{\mu})(t) \\ &\quad - \int_Y g(y) d(\nu^{k^*} + \delta\overline{\nu})(y) \Big| \\ &= \delta \Big| \int_X \int_X f(s, t) d\mu^{k^*}(s) d\overline{\mu}(t) + \frac{1}{2}\delta \int_X \int_X f(s, t) d\overline{\mu}(s) d\overline{\mu}(t) \\ &\quad - \int_Y g(y) d\overline{\nu}(y) \Big|. \end{aligned}$$

□

Finally, we want to prove that the approximate solution (μ^{k^*}, ν^{k^*}) converges to the optimal solution of CQP as $\delta \to 0$. We use $\overline{\delta}_i$ instead of δ in Algorithm 2 for the stopping criterion. In this case, we assume that the Algorithm 2 terminates in finitely many iterations, say k_i^* iterations, with an optimal solution $(\overline{\mu}^{k_i^*}, \overline{\nu}^{k_i^*})$ in the program $(DQSIP_{T_{k_i^*}})$.

THEOREM 3.6. *If $\overline{\delta}_i$ decreases to* 0 *as* $i \to \infty$, $\overline{\mu}^{k_i^*}(X) + \overline{\nu}^{k_i^*}(Y) \leqslant M$, *and* $f(s,t)$ *satisfies condition C then*

$$V(DQSIP_{T_{k_i^*}}) \to V(DCQP) \quad as \quad i \to \infty.$$

Proof. It is obvious that

$$V(DQSIP_{T_{k_1^*}}) \geqslant V(DQSIP_{T_{k_2^*}}) \geqslant \cdots \geqslant V(DCQP).$$

Then

$$\lim_{r\to\infty} V(DQSIP_{T_{k_r^*}}) = \alpha \geqslant V(DCQP). \tag{44}$$

Since

$$\phi_{\overline{\mu}^{k_i^*}, \overline{\nu}^{k_i^*}}(s) \leqslant \overline{\delta}_i \quad \text{for } i = 1, 2, \ldots, \quad \text{for each } s \in X,$$

where $\overline{\delta}_i \to 0$ as $i \to \infty$, and there exists a subsequence $(\overline{\mu}^{k_{i_j}^*}, \overline{\nu}^{k_{i_j}^*})$ of $(\overline{\mu}^{k_i^*}, \overline{\nu}^{k_i^*})$ weak* convergent to (μ^*, ν^*), we have

$$\phi_{\mu^*,\nu^*}(s) \leqslant 0, \quad \text{for each } s \in X.$$

From (2), we know that (μ^*, ν^*) is feasible for $(DCQP)$. Therefore

$$V(DCQP) \geqslant \int_X \int_X f(s,t)d\mu^*(s)d\mu^*(t) + \int_Y g(y)d\nu^*(y).$$

From the assumption and (44), it follows that

$$\lim_{r\to\infty} V(DQSIP_{T_{k_r^*}}) = \alpha = \int_X \int_X f(s,t)d\mu^*(s)d\mu^*(t) + \int_Y g(y)d\nu^*(y) \geqslant V(DCQP).$$

Thus

$$\int_X \int_X f(s,t)d\mu^*(s)d\mu^*(t) + \int_Y g(y)\nu^*(y) = \alpha = V(DCQP).$$

□

References

1. Choquet, G. (1954), Theory of capacities, *Annales de l'Institut Fourier* 5: 131–295.
2. Fang, S.C., Lin, C.J. and Wu, S.Y. Relaxations of the cutting plane method for quadratic semi-infinite programming. (submitted).
3. Fang, S.C. and Puthenpura, S.C. (1993), *Linear Optimization and Extensions: Theory and Algorithms*, Prentice Hall, Englewood Cliffs, New Jersey.

4. Fang, S.C., Lin, C.J. and Wu, S.Y. (1994), On solving convex quadratic semi–infinite programming problems, *Optimization* 31: 107–125.
5. Fugled, B. (1960), On the theory of potentials in locally compact spaces. *Acta Mathematica*, 103: 139–215.
6. Hernandez–Lerma, O. and Lasserre, J.B. (1998), Approximation schemes for infinite linear programs, *SIAM Journal on Optimization* 8: 973–988.
7. Hettich, R. and Kortanek, K. (1993), Semi-infinite programming: theory, method and application, *SIAM Review* 35: 380–429.
8. Karlin, S. (1959), *Mathematical Methods and Theory in Games, Programming and Economics*, Pergamon, London.
9. Lai, H.C. and Wu, S.Y. (1992), Extremal points and optimal solutions for general capacity problems, *Mathematical Programming*, 54: 87–113.
10. Luenberger, D.G. (1984), *Linear and Nonlinear Programming*, Addison–Wesley, Reading, Massachusetts.
11. Ohtsuka, M. (1966), Generalized capacity and duality theorem in linear programming. Journal of Science, Hiroshima University Series A–I, 30: 45–56.
12. Ohtsuka, M. (1966), A generalization of duality theorem in the theory of linear programming. Journal of Science, Hiroshima University Series A–I, 30: 31–39.
13. Reemtsen, R. and Görner, S. (1998), Numerical methods for semi-infinite programming: a survey. In: R. Reemtsen and J-J. Rueckmann (eds), *Semi-Infinite Programming*, Kluwer Academic Publishers, Boston, pp. 195–275.
14. Tichatschke, R. and Nebeling, V. (1988), A cutting plane algorithm for solving quadratic semi–infinite programs. *Optimization*, 19: 803–817.
15. Wu, S.Y. (1985), Linear programming in measure space. Ph.D. Dissertation, Cambridge University.
16. Wu, S.Y., (1985), The general capacity problem. In W. Oettli et al. (eds.), Methods of Operations Research, Oelgeschlager, Gunn and Hain.
17. Yamasaki, M., (1966), On a capacity problem raised in connection with linear programming. *Journal of Science*, Hiroshima University Series A–I, 30: 57–73.
18. Yosida, M. (1966), Some examples related to duality theorem in linear programming. Journal of Science, Hiroshima University Series A–I, 30: 41–43.