RAIRO Operations Research
RAIRO Oper. Res. **39** (2005) 13–33
DOI: 10.1051/ro:2005002

AN INTERIOR POINT ALGORITHM FOR CONVEX QUADRATIC PROGRAMMING WITH STRICT EQUILIBRIUM CONSTRAINTS

Rachid Benouahboun[1] and Abdelatif Mansouri[1]

Abstract. We describe an interior point algorithm for convex quadratic problem with a strict complementarity constraints. We show that under some assumptions the approach requires a total of $O(\sqrt{n}L)$ number of iterations, where L is the input size of the problem. The algorithm generates a sequence of problems, each of which is approximately solved by Newton's method.

Keywords. Convex quadratic programming with a strict equilibrium constraints, interior point algorithm, Newton's method.

1. Introduction

A mathematical programming with equilibrium constraints (MPEC) is a constrained optimization problem in which the essential constraints are defined by a complementarity system or variational inequality. This field of mathematical programming has been of much interest in the recent years. This is mainly because of its practical usage in many engineering design [3, 10], economic equilibrium [6], multi-level game-theoretic and machine learning problems [4, 9]. The monograph [12] presents a comprehensive study of this important mathematical programming problem.

Definition 1. Let M be a $(n \times n)$ real matrix. We say that a matrix M is copositive if for all $x \geq 0$, $Mx \geq 0$.

Received February 27, 2002. Accepted November 19, 2004.

[1] Département de Mathématiques, Faculté des Sciences Semlalia, Marrakech, Maroc; mansouri@ucam.ac.ma

It is easily seen that if M is co-positive, then for $x \geq 0$, we have $\|XMX\| \leq x^T Mx$, and $\|XMx\| \leq x^T Mx$, where $X = diag(x_1, ..., x_n)$.

In this paper, we consider the following convex quadratic program with a strict linear complementarity constraint:

$$(CQPEC) \begin{cases} Minimize\ \frac{1}{2}x^T Gx + d^T x \\ s.t. \\ Ax = b, \\ x^T Qx = 0, \\ x \geq 0, \end{cases}$$

where x and d are n-vectors, b is an m-vectors, A is an $m \times n$ matrix with $rank(A) = m < n$, G is a symmetric positive semidefinite $n \times n$ matrix, Q is a symmetric co-positive $n \times n$ matrix and the superscript T denotes transposition. Our purpose is to construct an interior point algorithm to solve the problem $(CQPEC)$. Using Newton's direction the algorithm generates a sequence of interior points which under some conditions converges to a solution (stationary point) of $(CQPEC)$ in a polynomial time. The algorithm has a complexity of $O(n^{3.5}L)$ where L is the input length for the problem.

The paper is organized as follows. In Section 2, we give some applications of the problem $(CQPEC)$ in mathematical programming. In Section 3, we present some theoretical background. In Section 4, we present the algorithm and we prove some results related to the convergence properties of the algorithm. Finally, in Section 6 we discuss the initialization of the algorithm. To illustrate our approach we conclude the paper with some numerical results in Section 7.

2. Some applications of *(CQPEC)* in mathematical programming

The problem $(CQPEC)$ have a wide range of applications for example in economic equilibrium, multi-level game and machine learning problems. In this section, we show some applications in mathematical programming. The problems that we consider in this section, are generally NP-Complete. Nevertheless, the approach presented in this paper solves a particular cases of these problems in polynomial times.

2.1. Optimization over the efficient set

Consider the multiobjective linear program

$$Minimize\ \left\{\widetilde{C}\widetilde{x}\ \ s.t.\ \widetilde{x} \in \chi = \left\{\widetilde{A}\widetilde{x} = \widetilde{b},\ \widetilde{x} \geq 0\right\}\right\} \tag{1}$$

where $\widetilde{C}$ is an $k \times n$ matrix. Recall that a point $\overline{x} \in \chi$ is an efficient solution of (1) if and only if there exists no $x \in \chi$ such that $Cx \leq C\overline{x}$ and $Cx \neq C\overline{x}$. Let E

denote the set of the efficient solutions. Consider the problem

$$Minimise \left\{ \frac{1}{2}\widetilde{x}^T \widetilde{G}\widetilde{x} + \widetilde{d}^T\widetilde{x} \;\; s.t \;\; \widetilde{x} \in E \right\} \tag{2}$$

where $\widetilde{G}$ is a symmetric positive semidefinite $n \times n$ matrix and $\widetilde{d}$ is an n-vector. In the linear case this problem has several applications in multiobjective programming (see [1, 2]). (2) can be written in the following form

$$\begin{cases} Minimize \; \frac{1}{2}\widetilde{x}^T \widetilde{G}\widetilde{x} + \widetilde{d}^T\widetilde{x} \\ s.t \\ \widetilde{A}\widetilde{x} = \widetilde{b}, \\ e^T\widetilde{\lambda} = 1, \\ \widetilde{A}^T\widetilde{y} + \widetilde{z} - \widetilde{C}^T\widetilde{\lambda} = 0, \\ \widetilde{x}^T\widetilde{z} = 0, \\ \widetilde{x} \geq 0, \; \widetilde{z} \geq 0, \; \widetilde{\lambda} \geq 0. \end{cases} \tag{3}$$

By taking $\widetilde{y} = \widetilde{y}_1 - \widetilde{y}_2$, with $\widetilde{y}_1 \geq 0$, and $\widetilde{y}_2 \geq 0$,

$$x = \begin{pmatrix} \widetilde{x} \\ \widetilde{z} \\ \widetilde{y}_1 \\ \widetilde{y}_2 \\ \widetilde{\lambda} \end{pmatrix}, \; A = \begin{pmatrix} \widetilde{A} & 0 & 0 & 0 & 0 \\ 0 & I_n & \widetilde{A}^T & -\widetilde{A}^T & -\widetilde{C}^T \end{pmatrix},$$

$$Q = \begin{pmatrix} 0 & I_n & 0 & 0 & 0 \\ I_n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} \widetilde{b} \\ 1 \\ 0 \end{pmatrix}$$

(3) is equivalent to $(CQPEC)$.

2.2. Goal programming

Consider the multiobjective problem (1). To solve this problem, several approaches have been developed. One class of a very used methods is the Distance-Based Methods. These methods assume that the decision-maker can select a point, at each iteration, which can be considered as an "ideally best" point from his viewpoint. If the ideal point is not feasible, the process gives the closer efficient solution to the ideal point as possible. The research of such a solution can be summarized by the mathematical programming problem:

$$Minimize \left\{ d(f^*, \widetilde{C}\widetilde{x}) \; s.t. \; \widetilde{x} \in \chi \right\} \tag{4}$$

where f^* is the ideal point and $d(.,.)$ is a distance of point in $\mathbb{R}^k$. In the case of the L_1-distance and by introducing the following notation

$$d_i^+ = \begin{cases} f_i^* - \widetilde{c}_i^T \widetilde{x}, \text{ if } f_i^* \geq c_i^T \widetilde{x} \\ 0 \quad \text{otherwise} \end{cases}$$
$$d_i^- = \begin{cases} \widetilde{c}_i^T \widetilde{x} - f_i^*, \text{ if } f_i^* < \widetilde{c}_i^T \widetilde{x} \\ 0 \quad \text{otherwise} \end{cases}$$

the problem (4) can be written as follows:

$$\begin{cases} Minimize \sum_{i=1}^{k} w_i (d_i^+ + d_i^-) \\ s.t. \\ \widetilde{A}\widetilde{x} = \widetilde{b} \\ \widetilde{C}\widetilde{x} - d^- + d^+ = f^*, \\ \left(d^+\right)^T d^- = 0, \\ d^+ \geq 0,\ d^- \geq 0, \widetilde{x} \geq 0 \end{cases} \tag{5}$$

where w is a vector of weights. This problem is often called goal programming. In many applications, a mixture of the above process and the method of sequential optimization is referred as a goal programming approach or model.

Obviously the problem (5) is equivalent to $(CQPEC)$.

3. Preliminaries

The algorithm that we consider in this paper is motivated by the application of the mixed penalty technique to problem $(CQPEC)$. The mixed penalty consists of examining the family of problems

$$(P_\mu)\ Minimize \left\{ f_\mu(x) = \frac{1}{2} x^T G x + d^T x + \frac{\left(x^T Q x\right)^2}{4\mu} - \mu \sum_{i=1}^{n} \log(x_i) :\ s.t\ x \in S \right\}$$

where $S = \{x :\ Ax = b,\ x > 0\}$, all $x \in S$ is called an interior point of the problem $(CQPEC)$, and $\mu > 0$ is the penalty parameter. This technique is well-known in the context of general constrained optimization problems. One solves the penalized problem for several values of the parameter μ, with μ decreasing to zero, and the result is a sequence of interior points, called centers, converging to a stationary solution of the original problem [5].

Our aim is to construct a polynomial algorithm. Since, generally, it is impossible to solve (P_μ) exactly in a finite time, we must renounce to the determination of centers, and work in their neighborhood. In other words, we can generate a sequence of interior points $\left\{x^k\right\}$ close to the path of centers such that $\left(x^k\right)^T Q x^k \longrightarrow 0$ and converges to a stationary solution of the problem $(CQPEC)$.

The KKT conditions applied to $(CQPEC)$ implies that if x is a stationary, then there exist $y \in \mathbb{R}^m$, $\beta \in \mathbb{R}$ and $z \in \mathbb{R}^n$, such that:

$$\begin{cases} A^T y - Gx - \beta Qx + z = d, \\ Ax = b, \\ x^T Qx = 0, \\ x^T z = 0, \\ x \geq 0,\ z \geq 0. \end{cases}$$

Let $T = \left\{(x, y, z, \beta) :\ Ax = b,\ A^T y - Gx - \beta Qx + z = d,\ x \geq 0,\ z \geq 0,\ \beta \geq 0\right\}$, and consider the merit function defined for all $(x, y, z, \beta) \in T$, by

$$\Delta(x, y, z, \beta) = x^T Qx + x^T z.$$

With this merit function, the above stationary condition system becomes

$$\begin{cases} (x, y, z, \beta) \in T, \\ \Delta(x, y, z, \beta) = 0. \end{cases}$$

For the time being, our objective is to construct a sequence $\left\{(x^k, y^k, z^k, \beta_k)\right\} \subset T$ such that x^k is an interior point for all k. with an upper bound for the merit function $\Delta(x^k, y^k, z^k, \beta_k) = \left(x^k\right)^T Qx^k + (x^k)^T z^k$, is driven to zero at a fixed rate of $1 - \frac{\sigma}{\sqrt{n}}$, where σ is a given constant. For this, we make the following assumptions.

Assumptions.

(a) There exists $w \in \mathbb{R}^m$ such that $A^T w > 0$ and $b^T w > 0$.

(b) If x is feasible for $(CQPEC)$, then we have for all $i \in \{1, ..., n\}$,

$$x_i > 0 \text{ if and only if } [Qx]_i = 0.$$

(c) For all $(x, y, z, \beta) \in T$, we have $\beta \leq \rho$, where ρ is a given constant.

Remark 1.

(i) The assumption (a) is used, in section 6, to transform the problem $(CQPEC)$ on an equivalent problem for which an initial interior point $x^0 \in S$ is known in advance. As in [7], we show that there exists y^0 such that: $X_0 A^T y^0 = e$.

(ii) The assumption (b) implies that the condition of Constraint Qualification holds for the problem $(CQPEC)$ (see [12]). Remark also that this assumption is satisfied for the problems described in Section 2.

(iii) We give an upper bound for ρ in the next section.

Throughout this paper, we use the following notation.

When x is a lower case letter denotes a vector $x = (x_1, ..., x_n)^T$, then a capital letter will denote the diagonal matrix with the components of the vector on the diagonal, *i.e.*, $X = diag(x_1, ..., x_n)$, and $\|.\|$ denotes the Euclidean norm.

4. The algorithm

Given an $(CQPEC)$ problem in standard form, in the Section 6, we discuss how to transform this problem into a form for which an initial interior point $x^0 \in S$ such that there exists $y^0 \in \mathbb{R}^m$ satisfying

$$X_0 A^T y^0 = e, \tag{6}$$

is given. For the moment we can suppose that a such initial point is known in advance.

For a current iterate $x \in S$, let $\widehat{x}$ denote the next iterate. The direction d_x, chosen to generate $\widehat{x}$ is defined as the Newton direction associated with the penalized problem (P_μ). The Newton direction at the point x is the optimal solution of the quadratic problem:

$$(QP_\mu) \begin{cases} Minimize\ \frac{1}{2} d_x^T \nabla^2 f_\mu(x) d_x + \nabla f_\mu(x)^T d_x \\ s.t. \\ Ad_x = 0. \end{cases}$$

Let M denote the Hessian of the penalty function f_μ at the point x. The following lemma show that if $x^T Q x < \mu$, then the matrix M is positive definite and by consequence the problem (QP_μ) has an unique solution.

Lemma 1. *If $x^T Q x < \mu$, then the matrix $M = G + \left(\frac{x^T Qx}{\mu}\right) Q + \frac{1}{\mu} Q x x^T Q + \mu X^{-2}$ is positive definite.*

Proof. We have $M = \mu X^{-1} \left(XGX + \left(\frac{x^T Qx}{\mu^2}\right) XQX + \frac{1}{\mu^2} XQxx^T QX + I \right) X^{-1}$. On the other hand we have $\left\| \left(\frac{x^T Qx}{\mu^2}\right) XQX \right\| \leq \left(\frac{x^T Qx}{\mu}\right)^2 < 1$, and so $I + \left(\frac{x^T Qx}{\mu^2}\right) XQX$ is positive definite. Thus M is also positive definite. □

By the KKT condition, d_x is determined by the following system of linear equations:

$$\begin{cases} \left[G + \left(\frac{x^T Qx}{\mu}\right) Q + \frac{1}{\mu} Q x x^T Q + \mu X^{-2} \right] d_x + Gx + d + \\ \left(\frac{x^T Qx}{\mu}\right) Qx - \mu X^{-1} e = A^T \widehat{y}, \\ Ad_x = 0. \end{cases} \tag{7}$$

Via a simple calculation, we obtain for d_x, $\widehat{x}$, $\widehat{y}$, $\widehat{z}$ and $\widehat{\beta}$:

$$\begin{aligned} d_x &= \left[M^{-1} - M^{-1}A^T\left(AM^{-1}A^T\right)^{-1}AM^{-1}\right]\vartheta \\ \widehat{x} &= x + d_x \\ \widehat{y} &= -\left(AM^{-1}A^T\right)^{-1}AM^{-1}\vartheta \\ \widehat{\beta} &= \frac{x^TQx}{\mu} \\ \widehat{z} &= G\widehat{x} + d + \widehat{\beta}Q\widehat{x} - A^T\widehat{y} \end{aligned} \tag{8}$$

where $\vartheta = \mu X^{-1}e - Gx - d - \left(\frac{x^TQx}{\mu}\right)Qx$.

We are now ready to describe the algorithm. Let δ, η, ρ and σ be constants satisfying:

$$\eta < \delta\left(\frac{1-\delta}{1+\delta}\right),\ \rho < Min\left\{\frac{\delta\left(\frac{1-\delta}{1+\delta}\right) - \eta}{\delta\theta_2 + \theta_1},\ \frac{\sqrt{\eta}}{\sqrt{\delta}\theta_1\theta_2}\right\} \tag{9}$$

$$\frac{\sigma}{\sqrt{n}} < Min\left\{\theta_1\left[\rho^2 - \rho\left(\frac{\theta_1}{\theta_2} + \delta\right) + \frac{\delta\left(\frac{1-\delta}{1+\delta}\right) - \eta}{\theta_2}\right], 1 - \frac{\rho\sqrt{\delta}\theta_1\theta_2}{\sqrt{\eta}}\right\} \tag{10}$$

where $\theta_1 = 1+\delta$ and $\theta_2 = \delta + \sqrt{n}$. The following lemma gives some results, related to theses constants, that will be useful to prove Lemma 6.

Lemma 2. *Let δ, η, ρ and σ be constants satisfying (9) and (10). Then:*

(i) $\dfrac{(\rho\theta_1\theta_2)^2}{1 - \frac{\sigma}{\sqrt{n}}} < \rho\theta_1\theta_2.$

(ii) $\alpha = 1 - \frac{\sigma}{\sqrt{n}} - \rho\theta_1\theta_2 > 0.$

(iii) $\eta\theta_1 + \delta^2 + \rho^2\theta_1\theta_2\left[\frac{\theta_1}{\rho\theta_2} - 1\right] + \sigma \leq \delta\alpha.$

(iv) $\left(\dfrac{\rho\theta_1\theta_2}{1 - \frac{\sigma}{\sqrt{n}}}\right)^2 \delta \leq \eta.$

Proof.

(i) From (10) we obtain: $\frac{\sigma}{\sqrt{n}} < 1 - \rho\sqrt{\frac{\delta}{\eta}}\theta_1\theta_2 < 1 - \rho\theta_1\theta_2$, which implies

$$\frac{(\rho\theta_1\theta_2)^2}{1 - \frac{\sigma}{\sqrt{n}}} \leq \rho\theta_1\theta_2. \tag{11}$$

(ii) From (11) it follows that $1 - \frac{\sigma}{\sqrt{n}} > \rho\theta_1\theta_2$. Thus $\alpha = 1 - \frac{\sigma}{\sqrt{n}} - \rho\theta_1\theta_2 > 0.$

(iii) By (10) we have

$$\begin{aligned}
&\tfrac{\sigma}{\sqrt{n}} \leq \theta_1 \left[\rho^2 - \rho\left(\tfrac{\theta_1}{\theta_2} + \delta\right) + \tfrac{\delta\left(\frac{1-\delta}{1+\delta}\right)-\eta}{\theta_2}\right] \\
&\tfrac{\sigma}{\sqrt{n}}\theta_2 \leq -\rho^2\theta_1\theta_2\left[\tfrac{\theta_1}{\rho\theta_2} - 1\right] - \rho\theta_1\theta_2\delta + \theta_1\left[\delta\left(\tfrac{1-\delta}{1+\delta}\right) - \eta\right] \\
&\tfrac{\sigma}{\sqrt{n}}\left(\delta + \sqrt{n}\right) \leq -\rho^2\theta_1\theta_2\left[\tfrac{\theta_1}{\rho\theta_2} - 1\right] + \delta - \delta^2 - \eta\theta_1 - \rho\theta_1\theta_2\delta \\
&\tfrac{\sigma}{\sqrt{n}}\delta + \sigma + \delta^2 + \eta\theta_1 + \rho^2\theta_1\theta_2\left[\tfrac{\theta_1}{\rho\theta_2} - 1\right] \leq \delta - \rho\theta_1\theta_2\delta \\
&\delta^2 + \eta\theta_1 + \rho^2\theta_1\theta_2\left[\tfrac{\theta_1}{\rho\theta_2} - 1\right] + \sigma \leq \delta\left(1 - \tfrac{\sigma}{\sqrt{n}} - \rho\theta_1\theta_2\right) \\
&\delta^2 + \eta\theta_1 + \rho^2\theta_1\theta_2\left[\tfrac{\theta_1}{\rho\theta_2} - 1\right] + \sigma \leq \delta\alpha.
\end{aligned}$$

The third inequality follows from the fact that $\theta_1 = 1+\delta$ and $\theta_2 = \delta+\sqrt{n}$.

(iv) By (10) we have $\frac{\sigma}{\sqrt{n}} \leq 1 - \dfrac{\rho\sqrt{\delta}\theta_1\theta_2}{\sqrt{\eta}}$, then we obtain $\dfrac{\rho\theta_1\theta_2}{1-\frac{\sigma}{\sqrt{n}}} \leq \sqrt{\dfrac{\eta}{\delta}}$. Hence

$$\left(\frac{\rho\theta_1\theta_2}{1-\frac{\sigma}{\sqrt{n}}}\right)^2 \delta \leq \eta. \qquad \square$$

We now state the algorithm.

ALGORITHM

- Step.0: let $x^0 \in S$ be a given interior point which satisfies (6), for an initial penalty parameter $\mu_0 > 0$, and $\rho,\ \delta,\ \eta$ and σ a given positive constants satisfying (9) and (10). Let $\varepsilon > 0$ be a tolerance for the merit function Δ.
  Set $k := 0$.
- Step.1: compute $d_x^k,\ y^{k+1},\ \beta_{k+1}$ and z^{k+1} by using (8).
- Step.2: set $x^{p+1} := x^p + d_x^k$ and $\mu_{k+1} := \mu_k\left(1 - \frac{\sigma}{\sqrt{n}}\right)$.
  Set $k := k+1$.
- Step.3: compute the merit function:

$$\Delta(x^k, y^k, z^k, \beta_k) := \left(x^k\right)^T Q x^k + (x^k)^T z^k.$$

  If $\Delta(x^k, y^k, z^k, \beta_k) \leq \varepsilon$, stop.

In the next section, we prove that all points generated by the algorithm lie in the set T and that they remain close to the central path T_c. We also show that the algorithm terminates in at most $O(\sqrt{n}L)$ iterations. This fact will enable us to show that the algorithm performs no more than $O(n^{3.5}L)$ arithmetic operations until its termination.

5. Convergence

We begin this section by stating the following main results.

Theorem 3. *Let ρ, δ, η and σ be positive constants satisfying relations (9) and (10). Assume that x and d_x satisfy $\|X^{-1}d_x\| \leq \delta$ and $\|XQxx^TQd_x\| \leq \eta\mu^2$. Let $\widehat{\mu} = \mu(1-\sigma/\sqrt{n})$ and consider the point $\left(\widehat{x},\widehat{y},\widehat{z},\widehat{\beta}\right) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ given by (8). Then, we have:*

(a) $\left(\widehat{x},\widehat{y},\widehat{z},\widehat{\beta}\right) \in T$;

(b) $\widehat{x}^T\widehat{z} \leq \mu\theta_2(\delta+\eta+\sqrt{n})$ *and* $\dfrac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}} \leq \rho\dfrac{\theta_2^2}{1-\sigma/\sqrt{n}} < 1$;

(c) $\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\| \leq \delta$ *and* $\left\|\widehat{X}Q\widehat{x}\widehat{x}^Td_{\widehat{x}}\right\| \leq \eta\widehat{\mu}^2$.

This theorem show that if, the current iterate is close to the central path, so it is for the next iterate.

Proof. To prove this theorem, we use the following Lemmas 4, 5 and 6 (see the annex for the proofs of those lemmas). □

Lemma 4. *Under the assumptions of the above theorem, we have*

(i) $\left(\widehat{x},\widehat{y},\widehat{z},\widehat{\beta}\right) \in T$ *and* $x^TQx \leq \mu\rho$;

(ii) $\widehat{x}^T\widehat{z} \leq \mu\theta_2(\delta+\eta+\sqrt{n})$ *and* $\dfrac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}} \leq \rho\dfrac{\theta_2^2}{1-\sigma/\sqrt{n}}$.

The following lemma will be used in the proof of Lemma 6.

Lemma 5. *Let x, d_x, and $\widehat{\mu}$ be as above. If $\|XQxx^TQd_x\| \leq \eta\mu^2$ and $\|X^{-1}d_x\| \leq \delta$, then, we have*

(i) $\left\|\widehat{X}X^{-1}-I\right\| \leq \delta$,*and* $\left\|\widehat{X}Q\widehat{x}\right\| \leq \mu\rho\theta_1\theta_2$;

(ii) $\left|\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}} - \frac{x^TQx}{\mu}\right| \leq \rho\left[\frac{\theta_1^2}{1-\sigma/\sqrt{n}}-1\right]$;

(iii) $\left|\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}d_{\widehat{x}}{}^TQd_{\widehat{x}}\right| \leq \mu\rho\theta_1\theta_2\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|^2$;

(iv) $\left\|\frac{1}{\mu^2}\widehat{X}Qxx^TQd_x\right\| \leq \eta\theta_1$.

Now, we prove that if $\|X^{-1}d_x\| \leq \delta$ and $\|XQxx^Td_x\| \leq \eta\mu^2$ are small, then they are also for the next iteration.

Lemma 6. *Let δ, η and σ be as above. If $\|X^{-1}d_x\| \leq \delta$, and $\|XQxx^Td_x\| \leq \eta\mu^2$, then we have*

$$\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\| \leq \delta \quad \text{and} \quad \left\|\widehat{X}Q\widehat{x}\widehat{x}^Td_{\widehat{x}}\right\| \leq \eta\widehat{\mu}^2.$$

Let $x^0 \in S$ be an initial interior point which satisfies (6). The following lemma shows how to choose $\mu_0 > 0$, such that the Newton step d_x^0, at x^0, satisfies $\left\|X_0 Q x^0 (x^0)^T Q d_x^0\right\| \leq \eta \mu_0^2$ and $\left\|X_0^{-1} d_x^0\right\| \leq \delta$.

Lemma 7. *Let δ, η, μ_0 and ρ satisfy (9), (10):*

$$\left\|X_0(Gx^0 + d)\right\| \leq \mu_0 \left[\delta - \rho^2 (1 + \delta)\right] \tag{12}$$

and

$$(x^0)^T Q x^0 \leq \mu_0 \rho. \tag{13}$$

Then

$$\left\|X_0^{-1} d_x^0\right\| \leq \delta \quad \textit{and} \quad \left\|X_0 Q x^0 (x^0)^T Q d_x^0\right\| \leq \eta \mu_0^2.$$

Proof. From (22) we have

$$\begin{aligned}\mu_0 \left\|X_0^{-1} d_x^0\right\|^2 \leq -(d_x^0)^T \left(Gx^0 + d\right) - \frac{(x^0)^T Q x^0}{\mu_0} (d_x^0)^T Q x^0 \\ - \frac{(x^0)^T Q x^0}{\mu_0} (d_x^0)^T Q d_x^0 - \mu_0 (d_x^0)^T X_0^{-1} e. \end{aligned} \tag{14}$$

On the other hand, we have

$$\begin{aligned}\left|\frac{(x^0)^T Q x^0}{\mu_0} (d_x^0)^T Q d_x^0\right| &\leq \frac{(x^0)^T Q x^0}{\mu_0} \left\|X_0 Q X_0\right\| \left\|X_0^{-1} d_x^0\right\|^2 \\ &\leq \frac{\left((x^0)^T Q x^0\right)^2}{\mu_0} \left\|X_0^{-1} d_x^0\right\|^2 \\ &\leq \mu_0 \rho^2 \left\|X_0^{-1} d_x^0\right\|^2.\end{aligned}$$

Then (14) becomes

$$\mu_0 \left(1 - \rho^2\right) \left\|X_0^{-1} d_x^0\right\|^2 \leq -(d_x^0)^T \left(Gx^0 + d\right) - \frac{(x^0)^T Q x^0}{\mu_0} (d_x^0)^T Q x^0 - \mu_0 (d_x^0)^T X_0^{-1} e.$$

From $A d_x^0 = 0$, we have $\left(d_x^0\right)^T A^T y^0 = 0$, then

$$\begin{aligned}\mu_0 \left(1 - \rho^2\right) \left\|X_0^{-1} d_x^0\right\|^2 \leq -(d_x^0)^T \left(Gx^0 + d\right) - \frac{(x^0)^T Q x^0}{\mu_0} (d_x^0)^T Q X^0 \\ - (d_x^0)^T X_0^{-1} \left[e - X_0 A^T y^0\right].\end{aligned}$$

Using (6), we conclude that

$$\begin{aligned}
\mu_0\left(1-\rho^2\right)\left\|X_0^{-1}d_x^0\right\|^2 &\leq (d_x^0)^T X_0^{-1}\left[X_0\left(Gx^0+d\right)+\frac{(x^0)^TQx^0}{\mu_0}X_0Qx^0\right]\\
&\leq \left\|X_0^{-1}d_x^0\right\|\left[\left\|X_0\left(Gx^0+d\right)\right\|+\frac{(x^0)^TQx^0}{\mu_0}\left\|X_0Qx^0\right\|\right]\\
&\leq \left\|X_0^{-1}d_x^0\right\|\left[\left\|X_0\left(Gx^0+d\right)\right\|+\frac{\left((x^0)^TQx^0\right)^2}{\mu_0}\right]\\
&\leq \mu_0\left\|X_0^{-1}d_x^0\right\|\left[\left[\delta-\rho^2\left(1+\delta\right)\right]+\rho^2\right]\\
&\leq \mu_0\left\|X_0^{-1}d_x^0\right\|\left(1-\rho^2\right)\delta.
\end{aligned}$$

Hence

$$\mu_0\left(1-\rho^2\right)\left\|X_0^{-1}d_x^0\right\|^2 \leq \mu_0\left\|X_0^{-1}d_x^0\right\|\left(1-\rho^2\right)\delta$$

which implies

$$\left\|X_0^{-1}d_x^0\right\| \leq \delta.$$

For the last inequality of the lemma, we have

$$\begin{aligned}
\left\|X_0Qx^0(x^0)^TQd_x^0\right\| &\leq \frac{1}{\mu_0^2}\left\|X_0Qx^0\right\|^2\left\|X_0^{-1}d_x^0\right\|\\
&\leq \left(\frac{(x^0)^TQx^0}{\mu_0}\right)^2\left\|X_0^{-1}d_x^0\right\|\\
&\leq \rho^2\delta.
\end{aligned}$$

By (10) we have $\rho^2 \leq \dfrac{\eta}{\delta}$ and hence $\left\|X_0Qx^0(x^0)^TQd_x^0\right\| \leq \mu_0^2\dfrac{\eta}{\delta}\delta = \eta\mu_0^2$. □

As a consequence of Theorem 1, we have the following result:

Corollary 8. *Let $x^0 \in S$ be an initial point which satisfies (6), μ_0 be an initial penalty parameter, and δ, η, σ and ρ constants satisfying the conditions of Lemmas 2 and 5. Then the sequence $\left\{\left(x^k, y^k, z^k, \beta_k\right)\right\}$ generated by the algorithm satisfies for all $k \geq 1$*

(i) $\left\|X_kQx^k(x^k)^TQd_x^k\right\| \leq \eta\mu_k^2$ *and* $\left\|X_k^{-1}d_x^k\right\| \leq \delta$;

(ii) $\left(x^k, y^k, z^k, \beta_k\right) \in T$ *with* $(x^k)^Tz^k \leq \mu_{k-1}\left(\delta+\sqrt{n}\right)\left(\delta+\eta+\sqrt{n}\right)$;

(iii) $(x^k)^TQx^k \leq \mu_k\rho$;

where $\mu_k = \left(1-\sigma/\sqrt{n}\right)^k\mu_0$.

Proof. This result follows trivially from Theorem 1. □

We now derive an upper bound for the total number of iterations performed by the algorithm.

Proposition 9. *The total number of iterations performed by the algorithm is no greater than* $\widehat{k} = \frac{\sqrt{n}}{\sigma} \log\left[\frac{\mu_0}{\varepsilon}\left[(1-\sigma/\sqrt{n})\,\rho + (\delta+\sqrt{n})(\delta+\eta+\sqrt{n})\right]\right]$, *where* $\varepsilon > 0$ *denotes the tolerance for the merit* Δ *and* μ_0 *is the initial penalty parameter.*

Proof. The algorithm terminates whenever

$$\mu_k\left[\left(\delta+\sqrt{n}\right)\left(\delta+\eta+\sqrt{n}\right) + \left(1-\sigma/\sqrt{n}\right)\rho\right] \leq \varepsilon.$$

Thus, it is enough to show that $\widehat{k}$ satisfies this inequality. By the definition of $\widehat{k}$, we have

$$\begin{aligned}
\log\varepsilon &= -\frac{\widehat{k}\sigma}{\sqrt{n}} + \log\left[\mu_0\left[\left(1-\sigma/\sqrt{n}\right)\rho + (\delta+\sqrt{n})(\delta+\eta+\sqrt{n})\right]\right] \\
&\geq \widehat{k}\log\left(1-\sigma/\sqrt{n}\right) + \log\left[\mu_0\left[\left(1-\sigma/\sqrt{n}\right)\rho + (\delta+\sqrt{n})(\delta+\eta+\sqrt{n})\right]\right] \\
&\geq \log\left[\mu_0\left(1-\sigma/\sqrt{n}\right)^{\widehat{k}}\left[\left(1-\sigma/\sqrt{n}\right)\rho + (\delta+\sqrt{n})(\delta+\eta+\sqrt{n})\right]\right] \\
&\geq \log\left[\mu_{\widehat{k}}\left[\left(1-\sigma/\sqrt{n}\right)\rho + (\delta+\sqrt{n})(\delta+\eta+\sqrt{n})\right]\right].
\end{aligned}$$

The second inequality is due the fact that $log(1-x) \leq -x$, for all $x < 1$. Therefore $\widehat{k}$ satisfies $\mu_{\widehat{k}}\left[(\delta+\sqrt{n})\,(\delta+\eta+\sqrt{n}) + (1-\sigma/\sqrt{n})\,\rho\right] \leq \varepsilon$. □

If we define L to be the number of bits necessary to encode the data of problem $(CQPEC)$, then the following corollary clearly holds.

Corollary 10. *If* $\log_2(\mu_0) = O(L)$, $\log_2(\varepsilon) = -O(L)$ *and the hypothesis of the last theorem holds, then the algorithm terminates in at most* $O(\sqrt{n}L)$ *iterations and with* $O(n^{3.5}L)$ *arithmetic operations.*

Proof. From $\log_2(\mu_0) = O(L)$ and $\log_2(\varepsilon) = -O(L)$, it is clear that we have $\widehat{k} = O(\sqrt{n}L)$. Since for each iteration we have to solve one linear system, which requires $O(n^3)$ arithmetic operations, the algorithm converges in at most $O(n^{3.5}L)$. □

6. Initialization of the algorithm

Consider the convex quadratic problem with a strict equilibrium constraint

$$(\widetilde{CQPEC}) \begin{cases} Minimize\ \frac{1}{2}\widetilde{x}^T\widetilde{G}\widetilde{x} + \widetilde{d}^T\widetilde{x} \\ \widetilde{A}\widetilde{x} = b, \\ \widetilde{x}^T\widetilde{Q}\widetilde{x} = 0, \\ \widetilde{x} \geq 0, \end{cases}$$

where $\widetilde{x}$ and $\widetilde{d}$ are vectors of length $\widetilde{n}$, $\widetilde{b}$ is a vector of length $\widetilde{m}$, $\widetilde{A}$ is an $\widetilde{m} \times \widetilde{n}$ matrix, $\widetilde{G}$ is a symmetric positive semidefinite $\widetilde{n} \times \widetilde{n}$ matrix, $\widetilde{Q}$ is a symmetric co-positive $\widetilde{n} \times \widetilde{n}$ matrix. We assume that

Assumption.

($\mathbf{A}_1$) There exists $w^0 \in \mathbb{R}^{\widetilde{m}}$ such that: $\widetilde{A}^T w^0 > 0$ and $\widetilde{b}^T w^0 > 0$.

($\mathbf{A}_2$) For all feasible point $\widetilde{x}$, we have: $\widetilde{x}_i > 0$ if and only if $\left[\widetilde{Q}\widetilde{x}\right]_i = 0$.

($\mathbf{A}_3$) For all $(\widetilde{x}, v, w, \beta) \in \digamma$ we have $\beta \leq \rho$, where

$$\digamma = \left\{(u, w, v, \beta) : \ \widetilde{A}^T w - \widetilde{G}\widetilde{x} - \beta\widetilde{Q}\widetilde{x} + v = \widetilde{d}, \widetilde{x} \geq 0, \ v \geq 0, \ \beta \geq 0\right\}$$

with ρ satisfies (9).

The aim of this section is to transform this problem into a convex quadratic problem with equilibrium constraints that satisfies assumptions (a), (b) and (c) and has an interior point x^0 which satisfies (6). The approach that we propose is one that has been suggested by numerous authors for transforming linear and convex quadratic programs into a form suitable for interior point algorithms.

Let N be a large positive constant. For all i we tack

$$\widetilde{x}_i^0 = \frac{1}{\left[\widetilde{A}^T w^0\right]_i} \quad \text{and} \quad \lambda = \frac{\left(\widetilde{x}^0\right)^T \widetilde{A}^T w^0}{\widetilde{b}^T w^0} = \frac{\widetilde{n}}{\widetilde{b}^T w^0}. \tag{15}$$

As in [8,11] it is clear that the following convex quadratic problem with equilibrium constraint

$$(\widehat{CQPEC}) \begin{cases} Minimize \ \frac{1}{2\lambda}\widehat{x}^T\widetilde{G}\widehat{x} + \widetilde{d}^T\widehat{x} + \alpha_1 N \\ \widetilde{A}\widehat{x} + (\lambda\widetilde{b} - \widetilde{A}\widetilde{x}^0)\alpha_1 = \lambda\widetilde{b}, \\ \alpha_1 + \alpha_2 = 2 \\ \alpha_1\alpha_2 = 0 \\ \frac{1}{\lambda}\widehat{x}^T\widetilde{Q}\widehat{x} = 0, \\ \widehat{x} \geq 0, \alpha_1 \geq 0, \alpha_2 \geq 0, \end{cases}$$

where $\widehat{x} = \lambda\tilde{x}$, can be written in the form of $(CQPE)$ with

$$\widehat{x} = \begin{bmatrix} \lambda\widetilde{x} \\ \alpha_1 \\ \alpha_2 \end{bmatrix}, \ A = \begin{bmatrix} \widetilde{A} & \lambda\widetilde{b} - \widetilde{A}\widetilde{x}^0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \ b = \begin{bmatrix} \lambda\widetilde{b} \\ 2 \end{bmatrix},$$

$$G = \begin{bmatrix} \frac{1}{\lambda}\widetilde{G} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ Q = \begin{bmatrix} \frac{1}{\lambda}\widetilde{Q} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \ d = \begin{bmatrix} \widetilde{d} \\ N \\ 0 \end{bmatrix}.$$

Let $x^0 = \begin{bmatrix} \widetilde{x}^0 \\ 1 \\ 1 \end{bmatrix}$ and $y^0 = \begin{bmatrix} w^0 \\ 1 \end{bmatrix}$. It is clear that x^0 is a feasible point for the problem $(\widehat{CQPEC})$.

Using (15) we obtain

$$\begin{aligned} X_0 A^T y^0 &= X_0 \begin{pmatrix} \widetilde{A}^T w^0 \\ \lambda b^T w^0 - \left(\widetilde{x}^0\right)^T \widetilde{A}^T w^0 + 1 \\ 1 \end{pmatrix} \\ &= X_0 \begin{pmatrix} \widetilde{A}^T w^0 \\ 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \widetilde{X}_0 \widetilde{A}^T w^0 \\ 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \widetilde{e} \\ 1 \\ 1 \end{pmatrix} \\ &= e. \end{aligned}$$

Then x^0 is an initial interior point for the problem $(\widehat{CQPEC})$ which satisfies (6).

On the other hand, we have $A^T y^0 > 0$ and $b^T y^0 > 0$. Thus assumption (a) is verified for $(\widehat{CQPEC})$.

The augmented set T is of the form

$$\begin{aligned} T = \Big\{ (\widehat{x}, y, z, \beta, \tau_1, \tau_2) :\ & \widetilde{A}\widehat{x} + (\lambda\widetilde{b} - \widetilde{A}\widetilde{x}^0)\alpha_1 = \lambda\widetilde{b},\ \alpha_1 + \alpha_2 = 2, \\ \widetilde{A}^T y - \frac{1}{\lambda}\widetilde{G}\widehat{x} - \beta\frac{1}{\lambda}\widetilde{Q}\widehat{x} + z = \widetilde{d}, & \left(\lambda\widetilde{b} - \widetilde{A}\widetilde{x}^0\right)^T y + u - \tau\alpha_2 + v_1 = N, \\ u - \tau\alpha_1 + v_2 = 0,\ \widehat{x} \geq 0, \alpha_1 \geq 0, \alpha_2 \geq 0,\ & z \geq 0, \\ v_1 \geq 0, v_2 \geq 0,\ \beta \geq 0,\ \tau \geq 0 \Big\}. & \end{aligned}$$

It is easily seen that if $(\widehat{x}, \alpha_1, \alpha_2, y, u, z, v_1, v_2, \beta, \tau) \in T$, then we have $\left(\frac{\widehat{x}}{\lambda}, y, z, \beta\right) \in F$ and by the assumption $(\mathbf{A_3})$ we obtain $\beta \leq \rho$. Thus the augmented problem $(\widehat{CQPEC})$ satisfies assumption (c) and by $(\mathbf{A_2})$ this problem satisfies also the assumption (b).

Then we can apply the algorithm to $(\widehat{CQPEC})$ for a large enough N.

If $x = (\widehat{x}, \alpha_1, \alpha_2)$ is an optimal solution of $(\widehat{CQPEC})$, then we have:

- If $\alpha_1 = 0$, then $\widetilde{x} = \dfrac{\widehat{x}}{\lambda}$ is an optimal solution of $(\widetilde{CQPEC})$.
- Else, $(\widetilde{CQPEC})$ has no solution.

7. Numerical example

To illustrate the result of the paper, the progress of the algorithm is presented on a numerical example. The example is a 0-1 convex quadratic problem.

$$\begin{cases} Minimize\ x_1^2 + 2x_2^2 - 3x_1 - x_2 \\ \qquad s.t. \\ -0.25x_1 + 0.25x_2 \leq 0.1875 \\ -0.125x_1 + 0.375x_2 \leq 0.34375 \\ 0.125x_1 + 0.5x_2 \leq 0.640625 \\ 0.25x_1 + 0.125x_2 \leq 0.40625 \\ 0.5x_1 + 0.125x_2 \leq 0.71875 \\ 0.25x_1 - 0.25x_2 \leq 0.28125 \\ 0.25x_1 + 0.75x_2 \geq 0.21875 \\ 0.5x_1 + 0.125x_2 \geq 0.09375 \\ \qquad x_i \in \{0,1\}\,. \end{cases}$$

Adding the necessary surplus variables, this problem becomes

$$\begin{cases} Minimize\ \frac{1}{2}x^T Gx + d^T x \\ s.t \\ Ax + z = b \\ x_i \in \{0,1\} \\ z \geq 0. \end{cases}$$

Converting this system of constraints to the form required by our proposed approach, results in augmented system given by

$$\begin{cases} Minimize\ \frac{1}{2}\widetilde{x}^T \widetilde{G}\widetilde{x} + \widetilde{d}^T \widetilde{x} \\ s.t \\ \widetilde{A}\widetilde{x} = \widetilde{b} \\ \widetilde{x}^T Q\widetilde{x} = 0 \\ \widetilde{x} \geq 0, \end{cases}$$

where

$$\widetilde{x} = \begin{pmatrix} x \\ z \\ e - x \end{pmatrix} \in \mathbb{R}^{12},\ \widetilde{A} = \begin{pmatrix} A & I & 0 \\ I & 0 & I \end{pmatrix},\ \widetilde{G} = \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix}$$

$$Q = \begin{pmatrix} 0 & 0 & I \\ 0 & 0 & 0 \\ I & 0 & 0 \end{pmatrix},\ \widetilde{b} = \begin{pmatrix} b \\ e \end{pmatrix}.$$

The augmented system shown above has the initial interior point:

$$\widetilde{x}_0 = (0.875\ \ 0.875\ \ 0.1875\ \ 0.125\ \ 0.09375\ \ 0.078125\ \ 0.171875\ \ 0.28125\ \ 0.65625\ \ 0.453125)^T.$$

TABLE 1.

iteration	x_1	x_2	merit	$d^T x$
0	0.87500	0.87500	− − −	− − −
1	0.81523	0.70744	4.60204	−1.20179
5	0.99510	0.10504	1.68374	−212514
10	0.99495	0.01498	0.21609	−2.00987
15	0.99954	0.00109	0.03836	−2.00130
20	0.99989	0.00030	0.00479	−2.00017
25	0.99999	0.00002	0.00084	−2.00003
30	1.00000	0.00000	0.00008	−2.00000

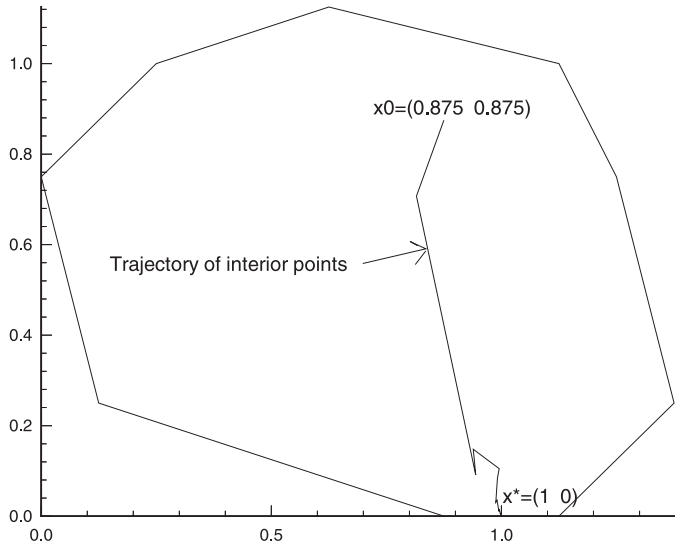

FIGURE 1.

By starting the algorithm with this initial solution (with $\sigma = 1.1$ and $\mu_0 = 0.5$), the process terminates at the solution:

$$\overline{x} = (1 \;\; 0 \;\; 0.43750 \;\; 0.51562 \;\; 0.15625 \;\; 0.21875 \;\; 0.03125 \;\; 0.03125 \;\; 0.40625 \;\; 0 \;\; 1)^T.$$

A summary of the first 30 iterations is given in Table 1.

The trace of the interior solution trajectory is shown in Figure 1.

8. CONCLUSION

We have established polynomial method for convex quadratic programming with strict equilibrium constraints. To our knowledge this is the first time an interior

point algorithm solves the (MPEC) in the convex quadratic case at polynomial time. The complexity obtained here coincides with the bound obtained for linear programming by the most of interior point methods.

Annex

Proof of Lemma 4.

(i) By the definition of $\widehat{x}$ and $\widehat{z}$ we have:

$$\widehat{x} = x + d_x = X\left(e + X^{-1}d_x\right) > 0 \tag{16}$$

and

$$\begin{aligned}\widehat{z} &= \mu X^{-1}e - \frac{1}{\mu}Qxx^TQd_x - \mu X^{-2}d_x \\ &= \mu X^{-1}\left(e - \frac{1}{\mu^2}XQxx^TQd_x - X^{-1}d_x\right) > 0.\end{aligned} \tag{17}$$

This means that $\left(\widehat{x}, \widehat{y}, \widehat{z}, \widehat{\beta}\right) \in T$. On the other hand, we have $\widehat{\beta} = \dfrac{x^TQx}{\mu}$ and from assumption (c) we obtain $x^TQx \leq \mu\rho$.

(ii) From (16) and (17) we have,

$$\begin{aligned}\widehat{x}^T\widehat{z} &\leq \mu\left\|e + X^{-1}d_x\right\|\left\|e - \tfrac{1}{\mu^2}XQxx^TQd_x - X^{-1}d_x\right\| \\ &\leq \mu\theta_2\left(\delta + \eta + \sqrt{n}\right).\end{aligned} \tag{18}$$

Thus, using the fact that $\|XQX\| \leq x^TQx \leq \mu\rho$ we obtain

$$\widehat{x}^TQ\widehat{x} \leq \left\|X^{-1}\widehat{x}\right\|^2\|XQX\| \leq x^TQx\theta_2^2 \leq \mu\rho\theta_2^2.$$

So we have $\dfrac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}} \leq \rho\frac{\theta_2^2}{1-\sigma/\sqrt{n}}$. □

Proof of Lemma 5.

(i) We have $\left\|\widehat{X}X^{-1} - I\right\| = \max\limits_{1\leq i\leq n}\left\{\dfrac{\widehat{x}_i}{x_i} - 1\right\} = \max\limits_{1\leq i\leq n}\left\{\dfrac{h_x^i}{x_i}\right\} \leq \delta$. To prove the last inequality of (i), we have

$$\begin{aligned}\left\|\widehat{X}Q\widehat{x}\right\| &\leq \left\|\widehat{X}X^{-1}\right\|\|XQX\|\left\|X^{-1}\widehat{x}\right\| \\ &\leq \rho\mu\left(1+\delta\right)\left(\delta + \sqrt{n}\right) \\ &= \mu\rho\theta_1\theta_2.\end{aligned}$$

The second inequality follows from the fact that $\|XQX\| \leq x^TQx \leq \mu\rho$ and $\left\|X^{-1}\widehat{x}\right\| = \left\|e + X^{-1}h_x\right\| \leq \delta + \sqrt{n}$.

(ii) Form (i) of the last lemma, it follows that

$$\begin{aligned}\left|\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}-\frac{x^TQx}{\mu}\right| &= \left|\left(\frac{1}{\widehat{\mu}}-\frac{1}{\mu}\right)\frac{x^TQx}{\mu}+2\frac{x^TQd_x}{\widehat{\mu}}+\frac{d_x^TQd_x}{\widehat{\mu}}\right| \\ &\leq \mu\rho\left(\frac{1}{\widehat{\mu}}-\frac{1}{\mu}\right)+2\rho\delta\frac{\mu}{\widehat{\mu}}+\rho\delta^2\frac{\mu}{\widehat{\mu}} \\ &\leq \rho\left[\frac{\mu}{\widehat{\mu}}\left(1+\delta\right)^2-1\right].\end{aligned}$$

Thus

$$\left|\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}-\frac{x^TQx}{\mu}\right| \leq \rho\left[\frac{\theta_1^2}{1-\sigma/\sqrt{n}}-1\right]. \tag{19}$$

(iii) From (ii) of the above Lemma, we conclude

$$\begin{aligned}\left|\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}{d_{\widehat{x}}}^TQd_{\widehat{x}}\right| &\leq \rho\frac{\theta_2^2}{1-\sigma/\sqrt{n}}\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|^2\left\|\widehat{X}X^{-1}\right\|^2\|XQX\| \\ &\leq \mu\rho^2\frac{\theta_2^2\theta_1^2}{1-\sigma/\sqrt{n}}\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|^2 \\ &\leq \mu\rho\theta_2\theta_1\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|^2.\end{aligned}$$

The last inequality is obtained by (i) of Lemma 2.

(iv) We have

$$\begin{aligned}\left\|\frac{1}{\mu^2}\widehat{X}Qxx^TQd_x\right\| &\leq \left\|\widehat{X}X^{-1}\right\|\left\|\frac{1}{\mu^2}XQxx^TQd_x\right\| \\ &\leq \eta\theta_1.\end{aligned}$$

The last inequality follows from the fact that $\left\|\frac{1}{\mu^2}XQxx^TQd_x\right\| \leq \eta$. □

Proof of Lemma 6. From (7) d_x and $d_{\widehat{x}}$ satisfy:

$$\begin{aligned}\left[G+\left(\frac{x^TQx}{\mu}\right)Q+\frac{1}{\mu}Qxx^TQ+\mu X^{-2}\right]d_x+Gx+d& \\ +\left(\frac{x^TQx}{\mu}\right)Qx-\mu X^{-1}e=A^T\widehat{y}&\end{aligned} \tag{20}$$

and

$$\begin{aligned}\left[G+\left(\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}\right)Q+\frac{1}{\widehat{\mu}}Q\widehat{x}\widehat{x}^TQ+\widehat{\mu}\widehat{X}^{-2}\right]d_{\widehat{x}}+G\widehat{x}+d+\left(\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}\right)Q\widehat{x}& \\ -\widehat{\mu}\widehat{X}^{-1}e=A^T\widetilde{y}.&\end{aligned} \tag{21}$$

Using the fact that $Ad_{\widehat{x}} = 0$ and $\widehat{x} = x + d_x$, and multiplying (20) and (21) by $d_{\widehat{x}}^T$ we conclude that:

$$d_{\widehat{x}}^T \left[\frac{1}{\mu} Q x x^T Q + \mu X^{-2}\right] d_x + d_{\widehat{x}}^T G\widehat{x} + d_{\widehat{x}}^T d + \left(\frac{x^T Q x}{\mu}\right) d_{\widehat{x}}^T Q\widehat{x} - \mu d_{\widehat{x}}^T X^{-1} e = 0$$

and

$$d_{\widehat{x}}^T \left[G + \left(\frac{\widehat{x}^T Q \widehat{x}}{\widehat{\mu}}\right) Q + \frac{1}{\widehat{\mu}} Q\widehat{x}\widehat{x}^T Q + \widehat{\mu}\widehat{X}^{-2}\right] d_{\widehat{x}} + d_{\widehat{x}}^T G\widehat{x} + d_{\widehat{x}}^T d + \left(\frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}}\right) d_{\widehat{x}}^T Q\widehat{x} - \widehat{\mu} d_{\widehat{x}}^T \widehat{X}^{-1} e = 0. \quad (22)$$

So, this implies that

$$d_{\widehat{x}}^T \left[G + \frac{1}{\widehat{\mu}} Q\widehat{x}\widehat{x}^T Q + \widehat{\mu}\widehat{X}^{-2}\right] d_{\widehat{x}} = d_{\widehat{x}}^T \left[\frac{1}{\mu} Q x x^T Q + \mu X^{-2}\right] d_x + \left(\frac{x^T Q x}{\mu} - \frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}}\right) d_{\widehat{x}}^T Q\widehat{x} - \left(\frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}}\right) d_{\widehat{x}}^T Q d_{\widehat{x}} + \widehat{\mu} d_{\widehat{x}}^T \widehat{X}^{-1} e - \mu d_{\widehat{x}}^T X^{-1} e.$$

We know that $d_{\widehat{x}}^T Q\widehat{x}\widehat{x}^T Q d_{\widehat{x}} = \left(\widehat{x}^T Q d_{\widehat{x}}\right)^2 \geq 0$ and $d_{\widehat{x}}^T G d_{\widehat{x}} \geq 0$, then we have

$$\widehat{\mu}\left\|\widehat{X}^{-1} d_{\widehat{x}}\right\|^2 + \left(\frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}}\right) d_{\widehat{x}}^T Q d_{\widehat{x}} \leq \mu d_{\widehat{x}}^T \widehat{X}^{-1} \left[\left(\frac{1}{\mu^2}\widehat{X} Q x x^T Q + \widehat{X} X^{-2}\right) d_x + \frac{1}{\mu}\left(\frac{x^T Q x}{\mu} - \frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}}\right)\widehat{X} Q x + \frac{\widehat{\mu}}{\mu} e - \widehat{X} X^{-1} e\right].$$

Using $\frac{\widehat{\mu}}{\mu} e = e - \frac{\sigma}{\sqrt{n}} e$ and the identity $e - \widehat{X} X^{-1} e = -X^{-1} d_x$, we can write the above expression as:

$$\widehat{\mu}\left\|\widehat{X}^{-1} d_{\widehat{x}}\right\|^2 + \left(\frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}}\right) d_{\widehat{x}}^T Q d_{\widehat{x}} \leq \mu d_{\widehat{x}}^T \widehat{X}^{-1} \left[\frac{1}{\mu^2}\widehat{X} Q x x^T Q d_x + \left(\widehat{X} X^{-1} - I\right) X^{-1} d_x + \frac{1}{\mu}\left(\frac{x^T Q x}{\mu} - \frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}}\right)\widehat{X} Q x - \frac{\sigma}{\sqrt{n}} e\right].$$

Now by (ii) of Lemma 5 we have:

$$\left|\frac{\widehat{x}^T Q\widehat{x}}{\widehat{\mu}} d_{\widehat{x}}{}^T Q d_{\widehat{x}}\right| \leq \mu\rho\theta_1\theta_2 \left\|\widehat{X}^{-1} d_{\widehat{x}}\right\|^2.$$

Thus

$$\widehat{\mu}\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|^2-\mu\rho\theta_1\theta_2\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|^2\leq\mu\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|\left[\frac{1}{\mu^2}\left\|\widehat{X}Qxx^TQd_x\right\|\right.$$
$$\left.+\left\|\widehat{X}X^{-1}-I\right\|\left\|X^{-1}d_x\right\|+\frac{1}{\mu}\left|\frac{x^TQx}{\mu}-\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}\right|\left\|\widehat{X}Qx\right\|+\sigma\right].$$

Hence

$$\alpha\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|\leq\frac{1}{\mu^2}\left\|\widehat{X}Qxx^TQd_x\right\|+\left\|\widehat{X}X^{-1}-I\right\|\left\|X^{-1}d_x\right\|$$
$$+\frac{1}{\mu}\left|\frac{x^TQx}{\mu}-\frac{\widehat{x}^TQ\widehat{x}}{\widehat{\mu}}\right|\left\|\widehat{X}Qx\right\|+\sigma,$$

with $\alpha=\frac{\widehat{\mu}}{\mu}-\rho\theta_1\theta_2$.

By using (19), (i) and (iii) of the above lemma, we obtain:

$$\alpha\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|\leq\eta\theta_1+\delta^2+\rho^2\theta_1\theta_2\left[\frac{\theta_1^2}{1-\sigma/\sqrt{n}}-1\right]+\sigma.$$

From (i) of Lemma 2, it follows that

$$\frac{1}{1-\sigma/\sqrt{n}}\leq\frac{1}{\rho\,(1+\delta)\,(\delta+\sqrt{n})}.\tag{23}$$

This implies

$$\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|\leq\frac{1}{\alpha}\left[\eta\theta_1+\delta^2+\rho^2\theta_1\theta_2\left[\frac{\theta_1}{\rho\theta_2}-1\right]+\sigma\right].\tag{24}$$

On the other hand, from (iii) of Lemma 2 we conclude

$$\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|\leq\delta.$$

Now we prove that $\left\|\widehat{X}Q\widehat{x}\widehat{x}^TQd_{\widehat{x}}\right\|\leq\widehat{\mu}^2\eta$. We have

$$\left\|\widehat{X}Q\widehat{x}\widehat{x}^TQd_{\widehat{x}}\right\|\leq\left\|\widehat{X}Q\widehat{x}\right\|^2\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\|,$$

Using (i) of Lemma 5, and $\left\|\widehat{X}^{-1}d_{\widehat{x}}\right\| \leq \delta$, we obtain

$$\left\|\widehat{X}Q\widehat{x}\widehat{x}^{T}Qd_{\widehat{x}}\right\| \leq \widehat{\mu}^{2}\left(\frac{\mu}{\widehat{\mu}}\right)^{2}(\rho\theta_{1}\theta_{2})^{2}\,\delta.$$

By (iv) of Lemma 2, we conclude $\left(\frac{\mu}{\widehat{\mu}}\right)^{2}(\rho\theta_{1}\theta_{2})^{2}\,\delta \leq \eta$.

Therefore $\left\|\widehat{X}Q\widehat{x}\widehat{x}^{T}Qd_{\widehat{x}}\right\| \leq \widehat{\mu}^{2}\eta$. □

References

[1] H.P. Benson, Optimization over the efficient set. *J. Math. Anal. Appl.* **98** (1984) 562–580.

[2] H.P. Benson, An algorithm for optimizing over the weakly-efficient set. *European J. Oper. Res.* **25** (1986) 192–199.

[3] Y. Chen and M. Florian, O-D *demand adjustment problem with cojestion: Part I. Model analysis and optimality conditions*. Publication CRT-94-56.

[4] C. Cinquini and R. Contro, Optimal design of beams discretized by elastic plastic finite element. *Comput. Structures* **20** (1985) 475–585.

[5] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley, New York (1968).

[6] D. Gale, *The theory of Linear Economic*. McGraw-Hill Book Company, New York (1960).

[7] C. Gonzaga, Path-following methods for linear programming. *SIAM Rev.* **34** (1992) 167–274.

[8] D. Goldfarb and S. Liu, An $O(n^3L)$ primal interior point algorithm for convex quadratic programming. *Math. Prog.* **49** (1990/91) 325–340.

[9] M. Kočvara and J.V. Outrata, On the solution of optimum design problems with variational inequalities, in *Recent Advances in Nonsmooth Optimization*, edited by D.Z. Du, L. Qi and R. Womersly, World Sciences (November 1995).

[10] G. Maier, A quadratic programming approach for certain classes of nonlinear structural problems. *Meccanica* **3** (1968) 121–130.

[11] D.C. Monteiro and I. Adler, Interior path following primal-dual algorithms. Part I: Linear programming. *Math. Prog.* **44** (1989) 27–41.

[12] Z.Q. Luo, J.S Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press (1996).