Mathematical Programming 44 (1989) 43-66
North-Holland

INTERIOR PATH FOLLOWING PRIMAL-DUAL ALGORITHMS. PART II: CONVEX QUADRATIC PROGRAMMING

Renato D.C. MONTEIRO and Ilan ADLER

Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, USA

Received 21 September 1987
Revised manuscript received 18 April 1988

We describe a primal-dual interior point algorithm for convex quadratic programming problems which requires a total of $O(\sqrt{n}L)$ number of iterations, where L is the input size. Each iteration updates a penalty parameter and finds an approximate Newton direction associated with the Karush-Kuhn-Tucker system of equations which characterizes a solution of the logarithmic barrier function problem. The algorithm is based on the path following idea. The total number of arithmetic operations is shown to be of the order of $O(n^3L)$.

Key Words: Interior-point methods, convex quadratic programming, Karmarkar's algorithm, polynomial-time algorithms, logarithmic barrier function, path following.

1. Introduction

In Part I of this paper, we introduce an interior path following primal-dual algorithm for linear programming problems which requires a total of $O(\sqrt{n}L)$ interations, where L is the input size, and each iteration can be executed in $O(n^3)$ arithmetic operations. The purpose of the second part is to extend these results in two directions. First, we modify the algorithm in Part I in order to solve convex quadratic programming problems. Second, we reduce the work per iteration to an amortized complexity of $O(n^{2.5})$ arithmetic operations. As a consequence, we obtain an algorithm for convex quadratic programming problems (and hence for linear programming problems) whose total complexity is $O(n^3L)$ arithmetic operations.

Quadratic programming (QP) problems share many of the combinatorial properties of linear programming (LP) problems. Based on these properties, algorithms extending the simplex method have been devised to solve QP problems. However, in the worst case, these algorithms may converge in an exponential number of steps.

Polynomial-time algorithms for convex quadratic programming problems based on the ellipsoid method were presented by Kozlov, Tarasov and Khachiyan [5].

Recently, with the advent of the new interior point algorithm by Karmarkar [3] for solving LP problems, some attention has been devoted to study classes of problems that can be solved by interior point algorithms in polynomial time. Kapoor and Vaidya [2] and Ye and Tse [10] present interior point algorithms for solving convex QP problems based on Karmarkar's projective transformation. Their algorithms are shown to converge in at most $O(nL)$ iterations, with total complexities of $O((\log n)(\log L)n^{3.67}L)$ and $O(n^4L^3)$ arithmetic operations respectively.

This part is organized as follows. In Section 2, we base our discussion on the theoretical background for the interior path following algorithm provided in Part I. We briefly add the necessary extension for convex quadratic programming problems. In Section 3, we describe the algorithm. We motivate and introduce the use of approximate directions to reduce the average effort per iteration. In Section 4, we prove the convergence of the algorithm presented in Section 3. Although the general ideas of the proofs are similar to those presented in Part I, their details are more involved due to the approximation scheme and the introduction of a quadratic component in the objective function of the problem. We also show how to compute an exact solution once we have found a sufficiently accurate feasible solution. In Section 5, we show how the approximation scheme reduces the average effort per iteration. In Section 6, we present an initialization similar to the one presented in Part I, but with detailed proofs which are omitted in Part I. We finally conclude with some remarks.

Since Part I and II share the same basic ideas, we found it necessary, for the sake of completeness and simplicity, to include a certain level of redundancy in the arguments.

2. Theoretical background

In this section, we introduce the problem which will be the object of our study and then briefly review some duality results as well as the extensions of the results presented in Section 2 of Part I to the present context.

We consider the convex quadratic programming problem as follows. Let

$$\text{(P)} \qquad \begin{aligned} \min \quad & c^{\mathrm{T}}x + \tfrac{1}{2}x^{\mathrm{T}}Qx \\ \text{s.t.} \quad & Ax = b, \\ & x \geqslant 0, \end{aligned}$$

where c, x are n-vectors, b is an m-vector, A is an $m \times n$ matrix and Q is a symmetric positive semi-definite $n \times n$ matrix. We assume that the entries of the vectors b, c and the matrices A and Q are integral. As for linear programming problems, we have

Proposition 2.1. *If problem* (P) *does not have an optimal solution then it must be either unbounded or infeasible.*

The Lagrangian dual problem corresponding to problem (P) is another quadratic programming problem given by

$$\text{(D)} \qquad \begin{aligned} \max \quad & -\tfrac{1}{2}v^{\mathrm{T}}Qv + b^{\mathrm{T}}y \\ \text{s.t.} \quad & -Qv + A^{\mathrm{T}}y + z = c, \\ & z \geqslant 0, \end{aligned}$$

where ν and z are n-vectors and y is an m-vector. The relationship between problems (P) and (D) is provided by the following result known as the duality theorem for convex quadratic programming.

Proposition 2.2. (a) *If problem* (P) *is unbounded then problem* (D) *is infeasible. If problem* (D) *is unbounded then problem* (P) *is infeasible.*

(b) *If problem* (P) *has an optimal solution* x^0 *then there exist* y^0 *and* z^0 *such that the point* $(\nu, y, z) = (x^0, y^0, z^0)$ *is an optimal solution of problem* (D). *Conversely, if problem* (D) *has an optimal solution then problem* (P) *has an optimal solution. Moreover, the optimal values of both problems are identical.*

The complementary slackness condition for convex quadratic programming problems is as follows.

Proposition 2.3. *If* x^0 *and* (ν^0, y^0, z^0) *are optimal solutions for problems* (P) *and* (D) *respectively then* $(x^0)^{\mathrm{T}}z^0 = 0$. *Conversely, if* $(\nu, y, z) = x^0, y^0, z^0)$ *is a feasible solution of* (D) *such that* x^0 *is feasible for* (P) *and* $(x^0)^{\mathrm{T}}z^0 = 0$, *then* x^0 *and* (x^0, y^0, z^0) *are optimal solutions of problems* (P) *and* (D) *respectively.*

We impose the following assumptions on the problems (P) and (D) (cf. Part I).

Assumption 2.1. (a) The set $S \equiv \{x \in \mathbb{R}^n;\ Ax = b, x > 0\}$ is non-empty.

(b) The set $T \equiv \{(\nu, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n;\ -Q\nu + A^{\mathrm{T}}y + z = c, z > 0\}$ is non-empty.

(c) $\operatorname{rank}(A) = m$.

We say that points in the sets S and T are interior feasible solutions of problems (P) and (D) respectively.

We will briefly recall some notation already introduced in Part I and also introduce some new notation which is necessary for the present context. If $x = (x_1, \ldots, x_n)^{\mathrm{T}}$ is an n-vector, then the corresponding capital letter X denotes the diagonal matrix $\operatorname{diag}(x_1, \ldots, x_n)$. The lower case letters w and s will be used to denote points $(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ and $(x, z) \in \mathbb{R}^n \times \mathbb{R}^n$ respectively. Also $\mathbb{R}^n_+$ will denote the set of real n-vectors with all components strictly positive. For a real number $a > 0$, we denote its logarithm to the natural base and to the base 2 by $\ln a$ and $\log a$ respectively. If $w \equiv (x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ then $f(w)$ denotes the n-vector XZe where e denotes the vector of all ones. The definition of the set W and the duality gap $g(w)$, for $w \in W$ needs to be reformulated in the present context as follows.

$$W \equiv \{(x, y, z);\ x \in S, (x, y, z) \in T\},$$

$$g(w) \equiv c^{\mathrm{T}}x + x^{\mathrm{T}}Qx - b^{\mathrm{T}}y.$$

Observe that W is the set consisting of the interior feasible solutions of problem (D) such that x is an interior feasible solution of problem (P). Note also that the duality gap $g(w)$ is the value of the objective function of (P) at x minus the value of the objective function of (D) at (x, y, z).

As in Part I, the algorithm considered in this paper is motivated by the application of the logarithmic barrier function method to problem (P). The logarithmic barrier function method consists of examining the family of problems

$$(\mathrm{P}_\mu) \qquad \min \quad c^{\mathrm{T}}x+\tfrac{1}{2}x^{\mathrm{T}}Qx-\mu\sum_{j=1}^{n}\ln x_j$$
$$\text{s.t.} \quad Ax=b$$
$$x>0$$

where $\mu>0$ is the barrier penalty parameter. Similar to relation (2.1) of Part I, one can show that the global solution x of problem (P_μ), if it exists, is completely characterized by the following Karush–Kuhn–Tucker stationary condition.

$$\begin{aligned} &ZXe-\mu e=0,\\ &Ax-b=0, \quad x>0, \\ &-Qx+A^{\mathrm{T}}y+z-c=0. \end{aligned} \tag{2.1}$$

Propositions 2.1 and 2.2 and Corollary 2.1 hold in the present context exactly as they are stated in Part I. We denote the unique solution of system (2.1) (see Corollary 2.1 of Part I) by $w(\mu)\equiv(x(\mu), y(\mu), z(\mu))$, in order to indicate its dependence on the penalty parameter $\mu>0$. Also, we denote by Γ the set (path) of solutions $w(\mu)$, for $\mu>0$. We refer to Γ as the central path associated with the convex QP problem (P).

We conclude this section with the following observations which are easily shown (cf. Section 2 of Part I).

(1) $w(\mu)\in W$, for all $\mu>0$. Or in order words, the central path is entirely contained in the set W. Indeed, $w\in\Gamma$ if, and only, if $w\in W$ and $f(w)=\mu e$, for some $\mu>0$.

(2) $g(w)=x^{\mathrm{T}}z$ for $w\equiv(x, y, z)\in W$. In view of this relation, we will always refer to the duality gap as the quantity $x^{\mathrm{T}}z$.

(3) $g(w(\mu))=n\mu$ for all $\mu>0$. Hence, the duality gap depends linearly on the parameter μ for points in the central path Γ.

Finally, we mention that Proposition 2.3 of Part I also holds for convex QP problem, if we replace $(y(\mu), z(\mu))$ by $(x(\mu), y(\mu), z(\mu))$ [7].

3. The algorithm

In this section, we describe an algorithm to solve the convex QP problem (P). The discussion in this section closely parallels the one presented in Section 3 of Part I. We start the description by discussing how the directions generated by the algorithm are computed.

Assume that a point $w=(x, y, z)\in W$, which we call the current iterate, is given. In order to determine the next iterate, a vector of direction $\Delta w=(\Delta x, \Delta y, \Delta z)\in \mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n$ needs to be generated. As in Part I, we could define Δw to be the Newton direction associated with the Karush–Kuhn–Tucker system of equations (2.1). However, with the objective of improving the worst-case complexity on the number of arithmetic operations, we consider a slight variation of the direction used in Part I. If we denote the left hand side of the system of equations (2.1) by $H(w)\equiv H(x, y, z)$, the Newton direction Δw at $w\in W$ is defined by the system of linear equations

$$D_w H(w)\Delta w = H(w)$$

where $\Delta w=(\Delta x, \Delta y, \Delta z)\in\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n$ and $D_w H(w)$ denotes the Jacobian of H at $w\equiv(x, y, z)$. Note that $D_w H(x, y, z)$ does not depend on the argument $y\in\mathbb{R}^m$. Indeed, the Jacobian of H at $w\equiv(x, y, z)$ is given by

$$J(x, z)\equiv D_w H(w)=\begin{bmatrix} Z & 0 & X \\ A & 0 & 0 \\ -Q & A^{\mathrm{T}} & I \end{bmatrix}.$$

The direction Δw that we are going to consider is defined by the following system of linear equations

$$J(\tilde{x}, \tilde{z})\Delta w = h(x, y, z)$$

where the points $\tilde{x}\in\mathbb{R}^n_+$ and $\tilde{z}\in\mathscr{R}^n_+$ will be chosen to approximate $x\in\mathbb{R}^n$ and $z\in\mathbb{R}^n$ respectively in a manner which will be specified later. More specifically, $\Delta w=(\Delta x, \Delta y, \Delta z)$ is defined by the following system of linear equations

$$\tilde{Z}\Delta x+\tilde{X}\Delta z = XZe-\hat{\mu}e, \tag{3.1.a}$$

$$A\Delta x = 0, \tag{3.1.b}$$

$$-Q\Delta x + A^{\mathrm{T}}\Delta y + \Delta z = 0, \tag{3.1.c}$$

where $\hat{\mu}>0$ is some prespecified penalty parameter. Note that the solution $\Delta w=(\Delta x, \Delta y, \Delta z)$ of the system of equations (3.1) clearly depends on the current iterate $w=(x, y, z)$, on the Jacobian of H at the "approximation" $\tilde{s}=(\tilde{x}, \tilde{z})$ of $s=(x, z)$ and on the penalty parameter $\hat{\mu}>0$. In order to indicate this dependence, we denote the solution Δw of system (3.1) by $\Delta w(w, \tilde{s}, \hat{\mu})$.

By simple calculation, we obtain the following expressions for Δx and Δy:

$$\Delta x = (\tilde{Z}+\tilde{X}Q)^{-1}[I-\tilde{X}A^{\mathrm{T}}(A(\tilde{Z}+\tilde{X}Q)^{-1}\tilde{X}A^{\mathrm{T}})^{-1}A(\tilde{Z}+\tilde{X}Q)^{-1}](XZe-\hat{\mu}e),$$

$$\Delta y = -[(A(\tilde{Z}+\tilde{X}Q)^{-1}\tilde{X}A^{\mathrm{T}})^{-1}A(\tilde{Z}+\tilde{X}Q)^{-1}](XZe-\hat{\mu}e),$$

$$\Delta z = Q\Delta x - A^{\mathrm{T}}\Delta y.$$

Therefore, to calculate the direction $\Delta w\equiv(\Delta x, \Delta y, \Delta z)$, the inverse of the matrix $A(\tilde{Z}+\tilde{X}Q)^{-1}\tilde{X}A^{\mathrm{T}} = A(\tilde{X}^{-1}\tilde{Z}+Q)^{-1}A^{\mathrm{T}}$ needs to be calculated. Note that the inverse of this matrix exists due to the fact that the matrix Q is positive semidefinite and $\tilde{X}$ and $\tilde{Z}$ are positive definite. The main motivation for considering an approximation $\tilde{s}=(\tilde{x}, \tilde{z})$ of $s=(x, z)$ is so that we do not need to invert this matrix from scratch at every iteration. If the current diagonal matrix $\tilde{X}^{-1}\tilde{Z}$ differs from the previous one by exactly l diagonal elements then, as described in Section 5, by performing $2l$ rank-one updates, we are able to compute the inverse of the matrix $A(\tilde{X}^{-1}\tilde{Z}+Q)^{-1}A^{\mathrm{T}}$ in $\mathrm{O}(n^2 l)$ arithmetic operations. Note that all the other operations involved in the computation of $\Delta w \equiv \Delta w(w, \tilde{s}, \hat{\mu})$ are of the order of $\mathrm{O}(n^2)$ arithmetic operations.

Having calculated the Newton direction $\Delta w(w, \tilde{s}, \hat{\mu})$ at the current iterate w, we find the next iterate $\hat{w} \equiv (\hat{x}, \hat{y}, \hat{z})$, by setting $\hat{x}=x-\Delta x$, $\hat{y}=y-\Delta y$ and $\hat{z}=z-\Delta z$, or in more compact notation, $\hat{w}=w-\Delta w$.

We are now ready to describe the algorithm. The algorithm will generate a sequence of points $w^k \in W$, $k=1, 2, 3, \ldots$, where the initial point w^0 is provided as input to the algorithm. We assume that the initial point $w^0 \equiv (x^0, y^0, z^0) \in W$ satisfies the following criterion of closeness with respect to the path Γ:

$$\|f(w^0)-\mu_0 e\| \leq \theta\mu_0 \tag{3.2}$$

where $\|\cdot\|$ denotes the Euclidean norm, μ_0 is a positive constant and $\theta=0.1$. Given a QP problem in standard form, in section 6 we show how to construct an augmented QP problem which immediately yields a solution for the original problem, if there exists one. We also show that this augmented problem satisfies Assumption 2.1 of Section 2 and that an initial point $w^0 \in W$ satisfying the criterion of closeness is readily available.

We now state the algorithm.

Algorithm 3.1. *Step 0*: Let $w^0 \in W$ and $\mu_0 > 0$ satisfy (3.2). Let ε be a given tolerance for the duality gap. Let $\delta \equiv 0.1$ and $\gamma \equiv 0.1$. Set $k := 0$.

Step 1: If $g(w^k)=x^{k\mathrm{T}}z^k \leq \varepsilon$, stop.

Step 2: Choose $\tilde{s}=(\tilde{x}, \tilde{z}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$ satisfying

$$\frac{|x_i^k-\tilde{x}_i|}{\tilde{x}_i} \leq \gamma, \quad i=1, \ldots, n,$$

$$\frac{|z_i^k-\tilde{z}_i|}{\tilde{z}_i} \leq \gamma, \quad i=1, \ldots, n.$$

Step 3: Set $\mu_{k+1} := \mu_k(1-\delta/\sqrt{n})$. Calculate $\Delta w^k \equiv \Delta w(w^k, \tilde{s}, \mu_{k+1})$.

Step 4: Set $w^{k+1} := w^k - \Delta w^k$. Set $k := k+1$ and go to Step 1.

4. Convergence results

In this section, we present convergence results for the algorithm described in Section 3. We start by stating the main result. We first need to introduce some notation.

Given two vectors $x \in \mathbb{R}_+^n$ and $\bar{x} \in \mathbb{R}^n$, we denote the Euclidean norm of the vector $X^{-1}(\bar{x}-x)$ by $\|\bar{x}-x\|_x$, i.e.,

$$\|\bar{x}-x\|_x = \left[\sum_{i=1}^{n} \left(\frac{\bar{x}_i - x_i}{x_i}\right)^2\right]^{1/2}.$$

The main result is

Theorem 4.1. *Let $\theta = \delta = \gamma = 0.1$. Let $w = (x, y, z) \in W$ and $\mu > 0$ satisfy*

$$\|f(w) - \mu e\| \leqslant \theta\mu. \tag{4.1}$$

Let $\tilde{s} = (\tilde{x}, \tilde{z}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n$ satisfy

$$\frac{|x_i - \tilde{x}_i|}{\tilde{x}_i} \leqslant \gamma, \quad i = 1, \ldots, n, \tag{4.2}$$

$$\frac{|z_i - \tilde{z}_i|}{\tilde{z}_i} \leqslant \gamma, \quad i = 1, \ldots, n. \tag{4.3}$$

Let $\hat{\mu} > 0$ be defined as

$$\hat{\mu} = \mu(1 - \delta/\sqrt{n}). \tag{4.4}$$

Consider the point $\hat{w} \equiv (\hat{x}, \hat{y}, \hat{z}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ defined by $\hat{w} \equiv w - \Delta w$ where $\Delta w \equiv \Delta w(w, \tilde{s}, \hat{\mu})$. Then, we have

(a) $\hat{w} \in W$ *and*

$$\|\hat{x} - x\|_x \leqslant 0.28, \tag{4.5}$$

$$\|\hat{z} - z\|_z \leqslant 0.28, \tag{4.6}$$

(b) $\|f(\hat{w}) - \hat{\mu} e\| \leqslant \theta\hat{\mu}$,

(c) $g(\hat{w}) \equiv \hat{x}^{\mathrm{T}}\hat{z} \leqslant 1.1 n\hat{\mu}$.

Theorem 4.1 describes the local behavior of Algorithm 3.1. If w is the current iterate then the next iterate $\hat{w}$, obtained by taking the Newton step $\Delta w(w, \tilde{s}, \hat{\mu})$, is guaranteed to be feasible and to satisfy the criterion of closeness with respect to the reduced penalty parameter $\hat{\mu}$. The proof of Theorem 4.1 requires some technical preliminary results.

Let $w = (x, y, z) \in W$, $\tilde{s} = (\tilde{x}, \tilde{z}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n$ and $\hat{\mu} > 0$. Let $\Delta w = (\Delta x, \Delta y, \Delta z)$ be the direction $\Delta w(w, \tilde{s}, \hat{\mu})$. Consider the point defined by $\hat{w} = w - \Delta w$. The next result provides expressions for the product of complementary variables $f_i(\hat{w}) \equiv \hat{x}_i\hat{z}_i$, $i = 1, \ldots, n$.

Lemma 4.1. *Let w, $\tilde{s}$, Δw and $\hat{w}$ be as above. Then, we have*

$$f_i(\hat{w}) = \hat{\mu} + \Delta x_i \Delta z_i + (\tilde{x}_i - x_i)\Delta z_i + (\tilde{z}_i - z_i)\Delta x_i, \tag{4.7}$$

$$(\Delta x)^{\mathrm{T}}(\Delta z) \geqslant 0. \tag{4.8}$$

Proof. Expression (4.7) can be easily proved using the definition of $f_i(\hat{w})$ and expression (3.1.a). Multiplying expression (3.1.c) on the left by $(\Delta x)^{\mathrm{T}}$, we obtain

$$(A\Delta x)^{\mathrm{T}}\Delta y+(\Delta x)^{\mathrm{T}}\Delta z-(\Delta x)^{\mathrm{T}}Q\Delta x=0 \tag{4.9}$$

which immediately implies (4.8) upon noting expression (3.1.b) and using the fact that the matrix Q is positive semidefinite. This completes the proof of the lemma. □

The next result was proved in Part I of this paper. Its proof is straightforward.

Lemma 4.2. *Let r, s and t be real n-vectors satisfying $r+s=t$ and $r^{\mathrm{T}}s\geqslant 0$. Then, we have*:

$$\max\{\|r\|, \|s\|\}\leqslant \|t\|,$$

$$\|RSe\|\leqslant\frac{\|t\|^2}{2}$$

where R and S denote the diagonal matrices corresponding to the vectors r and s respectively.

The next result is an immediate consequence of the previous lemma. It provides bounds necessary to show that the points generated by Algorithm 3.1 are feasible and remain close to the path Γ. Let $w=(x, y, z)\in W$, $\tilde{s}=(\tilde{x}, \tilde{z})\in\mathbb{R}^n_+\times\mathbb{R}^n_+$ and $\hat{\mu}>0$. Let $\Delta w=(\Delta x, \Delta y, \Delta z)$ be the direction $\Delta w(w, \tilde{s}, \hat{\mu})$. Let $\Delta f=(\Delta X)(\Delta Z)e$, where ΔX and ΔZ are the diagonal matrices corresponding to the vectors Δx and Δz respectively. An upper bound on the Euclidean norm of the vector Δf is given by the following result.

Lemma 4.3. *Let Δf be defined as above. Then, we have*

$$\|\Delta f\|\leqslant\frac{\|f(w)-\hat{\mu}e\|^2}{2\tilde{f}_{\min}}$$

where $\tilde{f}_{\min}\equiv\min\{\tilde{x}_i\tilde{z}_i;\ i=1,\ldots,n\}$. Furthermore, we have

$$\|\tilde{D}\Delta z\|^2\leqslant\frac{\|f(w)-\hat{\mu}e\|^2}{\tilde{f}_{\min}},$$

$$\|\tilde{D}^{-1}\Delta x\|^2\leqslant\frac{\|f(w)-\hat{\mu}e\|^2}{\tilde{f}_{\min}},$$

where $\tilde{D}$ is the diagonal matrix defined by $\tilde{D}=(\tilde{Z}^{-1}\tilde{X})^{1/2}$.

Proof. It follows from (3.1.a) and (4.8) that

$$\tilde{D}^{-1}\Delta x+\tilde{D}\Delta z=(\tilde{X}\tilde{Z})^{-1/2}(XZ-\hat{\mu}e),$$

$$(\tilde{D}^{-1}\Delta x)^{\mathrm{T}}(\tilde{D}\Delta z)\geqslant 0.$$

The lemma now follows by applying Lemma 4.2 and noting that

$$\|(\tilde{X}\tilde{Z})^{-1/2}(XZe-\hat{\mu}e)\|^2 \leqslant \frac{\|f(w)-\hat{\mu}e\|^2}{\tilde{f}_{\min}}.$$

This completes the proof of the lemma. $\square$

The next lemma provides some additional relations that will be useful in the proof of the main theorem.

Lemma 4.4. *Let $0\leqslant\theta<1$, $0\leqslant\gamma<1$ and $0\leqslant\delta\leqslant\sqrt{n}$ be given. Let $w=(x,y,z)\in W$ and $\mu>0$ satisfying (4.1), $\tilde{s}=(\tilde{x},\tilde{z})\in\mathbb{R}^n_+\times\mathbb{R}^n_+$ satisfying (4.2) and (4.3) be given. Define $\hat{\mu}$ as in (4.4). Let $\Delta w\equiv(\Delta x,\Delta y,\Delta z)$ be the direction $\Delta w(w,\tilde{s},\hat{\mu})$. Let p and q be defined by*

$$p=(1-\theta)(1+\gamma)^{-2}, \tag{4.10}$$

$$q=(1+\theta)(1-\gamma)^{-2}. \tag{4.11}$$

Let $\tilde{D}=(\tilde{Z}^{-1}\tilde{X})^{1/2}$. Then, we have the following relations:

$$p\mu\leqslant\tilde{x}_i\tilde{z}_i\leqslant q\mu, \quad i=1,\ldots,n, \tag{4.12}$$

$$\|\Delta f\|\leqslant\frac{(\theta+\delta)^2\mu}{2p}, \tag{4.13}$$

$$\|\tilde{D}^{-1}\Delta x\|^2\leqslant\frac{(\theta+\delta)^2\mu}{p}, \tag{4.14}$$

$$\|\tilde{D}\Delta z\|^2\leqslant\frac{(\theta+\delta)^2\mu}{p}. \tag{4.15}$$

Proof. From (4.2) and (4.3) it follows that

$$0<1-\gamma\leqslant\frac{x_i}{\tilde{x}_i}\leqslant 1+\gamma, \quad i=1,\ldots,n,$$

$$0<1-\gamma\leqslant\frac{z_i}{\tilde{z}_i}\leqslant 1+\gamma, \quad i=1,\ldots,n,$$

which implies,

$$(1-\gamma)^2\leqslant\frac{x_iz_i}{\tilde{x}_i\tilde{z}_i}\leqslant(1+\gamma)^2, \quad i=1,\ldots,n,$$

or equivalently, for all $i=1,\ldots,n$,

$$(1+\gamma)^{-2}x_iz_i\leqslant\tilde{x}_i\tilde{z}_i\leqslant(1-\gamma)^{-2}x_iz_i. \tag{4.16}$$

Using (4.1), we obtain

$$(1-\theta)\mu \le x_i z_i \le (1+\theta)\mu, \quad i=1,\ldots,n. \tag{4.17}$$

Combining (4.16) and (4.17), it follows that for all $i=1,\ldots,n$

$$(1+\gamma)^{-2}(1-\theta)\mu \le \tilde{x}_i\tilde{z}_i \le (1-\gamma)^{-2}(1+\theta)\mu$$

which is exactly (4.12). From (4.1), (4.4) and the fact that $\|e\|=\sqrt{n}$, we have

$$\begin{aligned}\|f(w)-\hat{\mu}e\|^2 &\le (\|f(w)-\mu e\| + \|\mu e-\hat{\mu}e\|)^2 \le (\theta\mu + |\mu-\hat{\mu}|\,\|e\|)^2\\ &= (\theta\mu+\delta\mu)^2 = (\theta+\delta)^2\mu^2.\end{aligned} \tag{4.18}$$

Using Lemma 4.3, relations (4.12) and (4.18), we immediately obtain (4.13), (4.14) and (4.15). This completes the proof of the lemma. □

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. (a) From (3.1.b), (3.1.c) and the fact that $w\in W$, it follows that $\hat{w}\equiv(\hat{x},\hat{y},\hat{z})$ satisfies $A\hat{x}=b$, $A^{\mathrm{T}}\hat{y}+\hat{z}=c$. It is enough to show that $\hat{x}$ and $\hat{z}$ are strictly positive vectors to conclude that $\hat{w}\in W$. We note that (4.5) and (4.6) immediately imply that $\hat{x}>0$ and $\hat{z}>0$ since $x>0$ and $z>0$ by assumption. We will now show that (4.5) and (4.6) hold. From the definition of $\hat{x}$, we have

$$\|\hat{x}-x\|_x^2 = \sum_{i=1}^n\left(\frac{\Delta x_i}{x_i}\right)^2 = \sum_{i=1}^n\left(\frac{\tilde{D}_{ii}}{x_i}\right)^2(\tilde{D}_{ii}^{-1}\Delta x_i)^2 \tag{4.19}$$

where $\tilde{D}_{ii}$ denotes the ith diagonal element of the matrix $\tilde{D}=(\tilde{Z}^{-1}\tilde{X})^{1/2}$, that is, $\tilde{D}_{ii}=(\tilde{x}_i/\tilde{z}_i)^{1/2}$. Using (4.1), (4.2), (4.3) and the definition of $\tilde{D}_{ii}$, we obtain, for all $i=1,\ldots,n$,

$$\left(\frac{\tilde{D}_{ii}}{x_i}\right)^2 = \frac{\tilde{x}_i}{\tilde{z}_i x_i^2} = \left(\frac{z_i}{\tilde{z}_i}\right)\left(\frac{\tilde{x}_i}{x_i}\right)\left(\frac{1}{x_i z_i}\right) \le \left(\frac{1+\gamma}{1-\gamma}\right)\left(\frac{1}{(1-\theta)\mu}\right). \tag{4.20}$$

The last inequality together with (4.19) implies

$$\|\hat{x}-x\|_x^2 \le \left(\frac{1+\gamma}{1-\gamma}\right)\left[\frac{1}{(1-\theta)\mu}\right]\|\tilde{D}^{-1}\Delta x\|^2.$$

The last relation and inequality (4.14) of Lemma 4.8 then imply

$$\|\hat{x}-x\|_x \le \left[\frac{(1+\gamma)}{(1-\gamma)(1-\theta)p}\right]^{1/2}(\theta+\delta).$$

Using the definition of p given by (4.10) and the fact that $\theta=\delta=\gamma=0.1$, we obtain (4.5). In a similar way, one can prove (4.6). This completes the proof of (a).

(b) It follows from expression (4.7) and the properties of norms that

$$\begin{aligned}\|f(\hat{w})-\hat{\mu}e\| &= \|\Delta f + (\tilde{X}-X)\Delta z + (\tilde{Z}-Z)\Delta x\|\\ &\le \|\Delta f\| + \|(\tilde{X}-X)\Delta z\| + \|(\tilde{Z}-Z)\Delta x\|.\end{aligned} \tag{4.21}$$

Using the definition of the Euclidean norm, we obtain

$$\|(\tilde{X}-X)\Delta z\|^2 = \sum_{i=1}^{n} (\tilde{x}_i - x_i)^2(\Delta z_i)^2$$

$$= \sum_{i=1}^{n} [\tilde{D}_{ii}^{-1}(\tilde{x}_i - x_i)]^2(\tilde{D}_{ii}\Delta z_i)^2 \tag{4.22}$$

where $\tilde{D}_{ii}$ denotes the ith diagonal element of the matrix $\tilde{D} = (\tilde{Z}^{-1}\tilde{X})^{1/2}$. Using (4.2) and (4.12), it follows that, for all $i = 1, \ldots, n$,

$$[\tilde{D}_{ii}^{-1}(\tilde{x}_i - x_i)]^2 = (\tilde{x}_i\tilde{z}_i)\left(\frac{\tilde{x}_i - x_i}{\tilde{x}_i}\right)^2 \leq q\mu\gamma^2. \tag{4.23}$$

Relation (4.15) of Lemma 4.4 and relations (4.22) and (4.23) imply that

$$\|(\tilde{X}-X)\Delta z\|^2 \leq \frac{q\gamma^2(\theta+\delta)^2\mu^2}{p}. \tag{4.24}$$

In a similar way, one can prove that

$$\|(\tilde{Z}-Z)\Delta x\|^2 \leq \frac{q\gamma^2(\theta+\delta)^2\mu^2}{p}. \tag{4.25}$$

From (4.21), (4.13), (4.24) and (4.25) it follows that

$$\|f(\hat{w}) - \hat{\mu}e\| \leq \left[\frac{(\theta+\delta)^2}{2p} + 2\gamma(\theta+\delta)\left(\frac{q}{p}\right)^{1/2}\right]\mu.$$

From the expression of $\hat{\mu}$ in (4.4), we obtain

$$\mu \leq \frac{\hat{\mu}}{(1-\delta)}$$

which implies

$$\|f(\hat{w}) - \hat{\mu}e\| \leq \frac{1}{(1-\delta)}\left[\frac{(\theta+\delta)^2}{2p} + 2\gamma(\theta+\delta)\left(\frac{q}{p}\right)^{1/2}\right]\hat{\mu}.$$

Using the definition of p and q given by (4.10) and (4.11) and the fact that $\theta = \delta = \gamma = 0.1$, it follows that $\|f(\hat{w}) - \hat{\mu}e\| \leq \theta\hat{\mu}$ and this completes the proof of (b).

(c) From statement (b), we have $\|f(\hat{w}) - \hat{\mu}e\| \leq 0.1\hat{\mu}$ which implies that $\hat{x}_i\hat{z}_i \leq 1.1\hat{\mu}$. Summing this inequality over all $i = 1, \ldots, n$, we obtain

$$g(\hat{w}) = \sum_{i=1}^{n} \hat{x}_i\hat{z}_i \leq 1.1n\hat{\mu}.$$

This completes the proof of (c). □

We will now describe the consequences of Theorem 4.1.

Corollary 4.1. *The sequence of points* (w^k) *generated by Algorithm* 3.1 *satisfies*

(a) $w^k \in W$, *for all* $k = 1, 2, \ldots,$ *and*

$$\|x^{k+1} - x^k\|_{x^k} \leq 0.28, \qquad \|z^{k+1} - z^k\|_{z^k} \leq 0.28,$$

(b) $\|f(w^k) - \mu_k e\| \leq \theta\mu_k$, *for all* $k = 1, 2, \ldots,$

(c) $g(w^k) \equiv x^{k\mathrm{T}} z^k \leq 1.1 n\mu_k$, *for all* $k = 1, 2, \ldots,$

where $\mu_k = \mu_0(1 - \delta/\sqrt{n})^k$ *for* $k = 1, 2, \ldots.$

Proof. This result follows trivially by arguing inductively and using Theorem 4.1. □

We now derive an upper bound on the total number of iterations performed by Algorithm 3.1. The following result follows easily from Corollary 4.1 and is proved in Section 4 of Part I of this paper.

Proposition 4.1. *The total number of iterations performed by Algorithm* 3.1 *is no greater than* $k^* \equiv \lceil \ln(1.1 n\varepsilon^{-1}\mu_0)\sqrt{n}/\delta \rceil$ *where* $\varepsilon > 0$ *denotes the tolerance for the duality gap and* μ_0 *is the initial penalty parameter.*

We define the size $L(A, Q, b, c)$ of a quadratic programming problem in the standard form (P) as

$$L(A, Q, b, c) = \left\lceil \log\left(\begin{matrix} \text{largest absolute value of the determinant} \\ \text{of any square submatrix of } M \end{matrix} + 1 \right) \right\rceil$$

$$+ \lceil \log(1 + \max_j |c_j|) \rceil + \lceil \log(1 + \max_i |b_i|) \rceil$$

$$+ \lceil \log(m + n) \rceil \tag{4.26}$$

where M is the matrix given by

$$M = \begin{bmatrix} Q & A \\ A^{\mathrm{T}} & 0 \end{bmatrix}. \tag{4.27}$$

In a similar way, we define the size $L(D, d)$ of a system of linear inequalities $Dv \leq d$, where D is an integer $p \times q$ matrix and d an integer p-vector, as follows.

$$L(D, d) = \left\lceil \log\left(\begin{matrix} \text{largest absolute value of the determinant} \\ \text{of any square submatrix of } D \end{matrix} + 1 \right) \right\rceil$$

$$+ \lceil \log(1 + \max_i |d_i|) \rceil + \lceil \log(p + q) \rceil.$$

It is straightforward to verify that the constant $L = L(A, Q, b, c)$ is less than two times the number of bits necessary to represent the data of the QP problem (P). The following result claims that we can find optimal solutions for problems (P) and (D) in $\mathrm{O}(n^3)$ arithmetic operations once the duality gap at a point w^k generated by Algorithm 3.1 becomes sufficiently small.

Proposition 4.2. *Let the convex QP problem in the standard form* (P) *be given. Assume we have a point* $\bar{w}=(\bar{x}, \bar{y}, \bar{z})$ *in the set* W *satisfying*

$$\bar{x}^{\mathrm{T}}\bar{z}\leqslant 2^{-2(L+3)} \tag{4.28}$$

where $L=L(A, Q, b, c)$. *Then, from* $\bar{w}$, *we can find a point* $w^*=(x^*, y^*, z^*)$ *in no more than* $\mathrm{O}(n^3)$ *arithmetic operations, such that* x^* *and* $w^*=(x^*, y^*, z^*)$ *solve problems* (P) *and* (D) *respectively.*

The proof of this proposition follows immediately from the following result. (See Papadimitriou and Steiglitz [8, pp. 173–174, Lemma 8.7].)

Lemma 4.5. *Let* D *be an* $p\times q$ *integer matrix and let* d *be an integer* p*-vector. Let* $0<\zeta\leqslant 2^{-\hat{L}}$ *be given, where* $\hat{L}=L(D, d)$. *Assume that a solution* $\bar{\nu}$ *for the system of linear inequalities* $D\nu<d+\zeta e$ *is known where* e *denotes the vector of all ones. Then, we can find a solution* ν^* *for the system of linear inequalities* $D\nu\leqslant d$ *in* $\mathrm{O}(p^2q)$ *arithmetic operations.*

Using this result, we can prove Proposition 4.2 as follows.

Proof of Proposition 4.2. Let $\zeta=2^{-(L+3)}$. Let $I=\{i;\, \bar{x}_i<\zeta\}$ and $J=\{j;\, \bar{z}_j<\zeta\}$. By (4.28), it follows that $I\cup J=\{1,2,\ldots,n\}$. Consider the system defined by the following linear inequalities:

$$Ax<b+\zeta e,\quad -Ax<-b+\zeta e,\quad -Qx+A^{\mathrm{T}}y<c+\zeta e,$$

$$-(-Qx+A^{\mathrm{T}}y)_j<-c_j+\zeta,\ j\in J,\quad -x<\zeta e \quad\text{and}\quad x_i<\zeta,\ i\in I.$$

We can write this system as $D\nu<d+\zeta e$ where D and d are the appropriate $p\times q$ matrix and p-vector suggested by the above definition, with $p=2(m+n)+|I|+|J|$, $q=m+n$ and ν is the q-vector (x, y). Obviously, $\bar{\nu}=(\bar{x}, \bar{y})$ is a solution for the system $D\nu<d+\zeta e$. Let $\hat{L}=L(D, d)$. One can easily show that $\hat{L}\leqslant L+3$. Hence $\zeta\leqslant 2^{-\hat{L}}$ and therefore, by Lemma 4.5, we can find a solution $\nu^*=(x^*, y^*)$ of the system $D\nu\leqslant d$ in at most $\mathrm{O}(p^2q)=\mathrm{O}(n^3)$ arithmetic operations. Letting $z^*=c+Qx^*-A^{\mathrm{T}}y^*$, it follows from the definition of $D\nu\leqslant d$ that $w^*=(x^*, y^*, z^*)$ is in the set W and satisfies $x_i^*=0$, $i\in I$ and $z_j^*=0$, $j\in J$. Since $I\cup J=\{1,2,\ldots,n\}$, we have $(x^*)^{\mathrm{T}}z^*=0$. The result now follows once we note the converse part of Proposition 2.3. □

Corollary 4.2. *If the initial penalty parameter* μ_0 *satisfies* $\mu_0=2^{\mathrm{O}(L)}$ *then Algorithm* 3.1 *solves problem* (P) *in at most* $\mathrm{O}(\sqrt{n}L)$ *iterations.*

Proof. Using the previous proposition, we can set $\varepsilon=2^{-(L+3)}$ as the tolerance for the duality gap in Algorithm 3.1. From Proposition 4.1, we immediately conclude the validity of this corollary. □

In Section 5, we will see that the initial penalty parameter μ_0 can be chosen to satisfy $\mu_0 = 2^{L(L)}$. One possible choice for the approximation $\tilde{s} \equiv (\tilde{x}, \tilde{z})$ in step 2 of Algorithm 3.1 is to use exact data, that is, to set $\tilde{s}$, on the kth iteration, equal to s^k. With this choice of $\tilde{s}$, we have the following result whose proof is immediate.

Corollary 4.3. *Algorithm* 3.1 *solves problem* (P) *in no more than* $O(n^{3.5}L)$ *arithmetic operations.*

In the next section, we present an alternative choice for the approximation that makes it possible to reduce the complexity of Algorithm 3.1 to $O(n^3L)$ arithmetic operations.

5. A good choice for $\tilde{x}$ and $\tilde{z}$

In this section, we show that the complexity of Algorithm 3.1 can be reduced to $O(n^3L)$ arithmetic operations. We should point out that this idea for reduction of the complexity was first presented in Karmarkar [3] and subsequently in Gonzaga [1] and Vaidya [9]. The reduction basically consists of using a direction that approximates the "exact" direction calculated from using "exact" data, that is, the current iterate. In our case, an approximate direction is implicit in the choice of the approximation $\tilde{s}$. In this section, we show that by choosing the approximation $\tilde{s}$ judiciously, a reduction in the average work per iteration is obtained. The choice of the approximation $\tilde{s}$ is made by an updating scheme as follows. (In the procedure below, k stands for the iteration count.)

Updating scheme 5.1.
For $k := 0$, set $\tilde{x} := x^0$ and $\tilde{z} := z^0$.
For $k > 0$ do
For $i = 1, \ldots, n$ do
If one of the following holds:

(a) $$\frac{|x_i^k - \tilde{x}_i|}{\tilde{x}_i} > \gamma,$$

(b) $$\frac{|z_i^k - \tilde{z}_i|}{\tilde{z}_i} > \gamma,$$

then set $\tilde{x}_i := x_i^k$ and $\tilde{z}_i := z_i^k$.
end of scheme.

In order to calculate the directions Δx, Δy and Δz determined by system (3.1), we need to calculate the inverse of the matrix

$$B \equiv A(\tilde{Z} + \tilde{X}Q)^{-1}\tilde{X}A^{\mathrm{T}} = A(\tilde{X}^{-1}\tilde{Z} + Q)^{-1}A^{\mathrm{T}} \tag{5.1}$$

where $\tilde{s} \equiv (\tilde{x}, \tilde{z})$ represents the approximation for the current iteration. Let $\tilde{s}^k$ and B_k denote the approximation $\tilde{s}$ and the matrix given by (5.1) respectively at the kth iteration of Algorithm 3.1. Also let D_k denote the matrix $(\tilde{X}^k)^{-1}\tilde{Z}^k$. We show next that if the matrix D_k differs from the matrix D_{k-1} by exactly l diagonal elements then the computation of B_k^{-1} can be carried out in $O(n^2 l)$ arithmetic operations by means of $2l$ rank-one updates.

Let $E = Q + D_{k-1}$ and $H = D_k - D_{k-1}$. Then we obtain

$$B_{k-1} = AE^{-1}A^{\mathrm{T}} \tag{5.2}$$

and

$$B_k = A(E+H)^{-1}A^{\mathrm{T}}. \tag{5.3}$$

Obviously, H is a diagonal matrix. Denote the i^{th} diagonal element of the matrix H by h_i. By assumption, exactly l diagonal elements h_i are non-zero. For simplicity of notation, we assume that these elements are the first l diagonal entries of the matrix H. Then the matrix H can be written as

$$H = \sum_{i=1}^{l} h_i u^i (u^i)^{\mathrm{T}}$$

where u^i denotes the n-vector where all components are zero except the ith component which equals one. Let E_j be defined as

$$\begin{aligned} &E_0 = E, \\ &E_j = E_{j-1} + h_j u^j (u^j)^{\mathrm{T}}, \quad j = 1, \ldots, l. \end{aligned} \tag{5.4}$$

Note that $E_l = E + H$. Note also that, by definition, $E_j - Q$ has the first j diagonal elements equal to the corresponding diagonal elements of the matrix $E + H - Q = D_k$ and the others equal to the corresponding diagonal elements of the matrix $E - Q = D_{k-1}$. Therefore the matrices $E_0, E_1, \ldots, E_l$ are positive definite, and hence invertible. Applying the well-known Sherman–Morrison formula of linear algebra to the matrix E_j as given in expression (5.4), we obtain, for $j = 1, \ldots, l$,

$$E_j^{-1} = E_{j-1}^{-1} - \left(\frac{h_j}{1 + h_j (u^j)^{\mathrm{T}} E_{j-1}^{-1} u^j} \right) E_{j-1}^{-1} u^j (u^j)^{\mathrm{T}} E_{j-1}^{-1}.$$

Thus, using the above expression, we can obtain E_l^{-1} as follows.

$$E_l^{-1} = E_0^{-1} - \sum_{i=1}^{l} g_i v^i (v^i)^{\mathrm{T}} \tag{5.5}$$

where the scalars g_i and the n-vectors v^i, $i = 1, \ldots, l$, are generated recursively by the following iterative procedure.

Procedure 5.1. Given E_0^{-1} then:

For $j = 1, \ldots, l$ do

$v^j = E_{j-1}^{-1} u^j,$

$g_j = h_j/(1 + h_j (u^j)^{\mathrm{T}} v^j),$

$E_j^{-1} = E_{j-1}^{-1} - g_j v^j (v^j)^{\mathrm{T}}.$

end of procedure.

Since $E_l = E + H$ and using expressions (5.2), (5.3) and (5.5), we obtain

$$B_k = B_{k-1} - \sum_{j=1}^{l} g_j (A v^j)(A v^j)^{\mathrm{T}}. \tag{5.6}$$

We can also use the same process described above to find the inverse of the matrix B_k using expression (5.6) and the matrix B_{k-1}^{-1} already calculated in the previous iteration of the algorithm. We note that the procedure above involves $\mathrm{O}(n^2 l)$ arithmetic operations. We summarize the discussion above in the following result.

Proposition 5.1. *Let B_k denote the matrix given in* (5.1) *at the kth iteration of Algorithm* 3.1 *and let $D_k = (\tilde{X}^k)^{-1} \tilde{Z}^k$. Assume that the matrices $(D_{k-1} + Q)^{-1}$ and B_{k-1}^{-1} are given. If D_{k-1} differs from D_k by exactly l diagonal elements, then the matrices $(D_k + Q)^{-1}$ and B_k^{-1} can be found in at most $\mathrm{O}(n^2 l)$ arithmetic operations.*

Next we provide an upper bound on the number of diagonal element changes that occur on the matrix $\tilde{X}^{-1}\tilde{Z}$ during K steps of Algorithm 3.1. Note that the ith diagonal element of the matrix $(\tilde{X}^{-1}\tilde{Z})$ changes only when either inequality (a) or (b) of the Updating scheme 5.1 is satisfied.

The following result is due to Gonzaga [1]. Since we state it here in a more general form than is presented in [1] and for the sake of completeness, a proof of this result is given in Appendix A.

Proposition 5.2. *Let $(v^k)_{k=0}^{K}$ be a sequence of n-vectors with all components positive and satisfying*

$$\| v^{k+1} - v^k \|_{v^k} \leq \rho, \quad k = 0, 1, \ldots, K-1 \tag{5.7}$$

where ρ is a positive constant less than one. Define the sequence $(\tilde{v}^k)_{k=0}^{K}$ recursively as follows. Set $\tilde{v}^0 := v^0$ and for $k \geq 1$ and $i = 1, \ldots, n$ let

$$\tilde{v}_i^k := \begin{cases} v_i^k & \text{if } \dfrac{|v_i^k - \tilde{v}_i^{k-1}|}{|\tilde{v}_i^{k-1}|} > \gamma, \\ \tilde{v}_i^{k-1} & \text{otherwise,} \end{cases}$$

where γ is a positive constant less than one. Let V_i^K be the set of indices k defined as

$$V_i^K = \left\{ k;\ \frac{|v_i^k - \tilde{v}_i^{k-1}|}{|\tilde{v}_i^{k-1}|} > \gamma,\ 1 \leq k \leq K \right\}$$

and let $|V_i^K|$ *denote its cardinality, that is, the number of times the ith component of the sequence* $(\tilde{v}^k)_{k=0}^K$ *changes. Then, we have*

$$\sum_{i=1}^n |V_i^K| \leq -\frac{\rho K\sqrt{n}}{(1-\rho)\ln(1-\gamma)}.$$

As a consequence of this result, we have the following corollary.

Corollary 5.1. *Let* $(x^k)_{k=0}^K$ *and* $(z^k)_{k=0}^K$ *be the sequences generated by Algorithm* 3.1 *and let* $\tilde{s}^k = (\tilde{x}^k, \tilde{z}^k)$ *denote the value of the approximation* $\tilde{s} = (\tilde{x}, \tilde{z})$ *at the end of the kth iteration of the Updating scheme* 5.1. *Consider the following two sets*:

$$S_i^K = \left\{k;\ \frac{|x_i^k - \tilde{x}_i^{k-1}|}{\tilde{x}_i^{k-1}} > \gamma,\ 1 \leq k \leq K\right\},$$

$$T_i^K = \left\{k;\ \frac{|z_i^k - \tilde{z}_i^{k-1}|}{\tilde{z}_i^{k-1}} > \gamma,\ 1 \leq k \leq K\right\}.$$

Then, we have the following inequalities.

$$\sum_{i=1}^n |S_i^K| \leq 3.7\sqrt{n}K, \qquad \sum_{i=1}^n |T_i^K| \leq 3.7\sqrt{n}K.$$

Proof. This result follows immediately by using relations (4.5), (4.6) and Proposition 5.2 with $\rho = 0.28$ and $\gamma = 0.1$. $\square$

Thus, the total number of rank-one updates that occur during K steps of Algorithm 3.1 is on the order of $O(\sqrt{n}K)$. As a consequence of this result, we have

Corollary 5.2. *Algorithm* 3.1 *coupled with the Updating scheme* 5.1 *solves problem* (P) *in no more than* $O(n^3L)$ *arithmetic operations.*

Proof. From Corollary 4.2, we know that Algorithm 3.1 finds an optimal solution to Problem (P) in $O(\sqrt{n}L)$ iterations. Corollary 5.1 implies that the total number of rank-one updates is then of the order of $O(nL)$. Since each rank-one update involves $O(n^2)$ arithmetic operations, the total number of arithmetic operations is then of the order of $O(n^3L)$. This completes the proof of the corollary. $\square$

6. Initialization of the algorithm

In this section, we show how to initialize Algorithm 3.1, in order to solve any convex quadratic programming problem. With this aim, we introduce an augmented problem that has an initial start for the algorithm, whose size is of the same order as the original problem, and whose solution immediately yields a solution of the original problem.

Consider the convex quadratic programming problem, which we call the orignal problem, stated as follows.

$$(\tilde{\mathrm{P}}) \qquad \begin{aligned} \min \quad & \tilde{c}^{\mathrm{T}}\tilde{x}+\tfrac{1}{2}\tilde{x}^{\mathrm{T}}\tilde{Q}\tilde{x} \\ \text{s.t.} \quad & \tilde{A}\tilde{x}=\tilde{b}, \\ & \tilde{x}\geqslant 0, \end{aligned}$$

where $\tilde{A}$ is an $\tilde{m}\times\tilde{n}$ matrix which has full row rank, $\tilde{Q}$ is an $\tilde{n}\times\tilde{n}$ symmetric positive semi-definite matrix and $\tilde{b}$, $\tilde{c}$ are vectors of length $\tilde{m}$ and $\tilde{n}$ respectively. We assume that the entries of the vectors $\tilde{b}$, $\tilde{c}$ and the matrices $\tilde{A}$ and $\tilde{Q}$ are integral.

Observe that we might not be able to apply Algorithm 3.1 to solve problem $(\tilde{\mathrm{P}})$ directly for the following reasons. First, problem $(\tilde{\mathrm{P}})$ might not satisfy conditions (a) and (b) of Assumption 2.1. Even if those conditions are satisfied, an initial point satisfying the criterion of closeness (3.2) might not be known a priori. In order to circumvent these apparent difficulties, we introduce an augmented problem which will play an important role in the solution of problem $(\tilde{\mathrm{P}})$.

Before introducing the augmented problem, we need to define some quantities. Let $\tilde{L}=L(\tilde{A},\tilde{Q},\tilde{b},\tilde{c})$ denote the size of $(\tilde{\mathrm{P}})$. Let $n=\tilde{n}+2$ and $m=\tilde{m}+1$. Let $\alpha=2^{4\tilde{L}}$ and $\lambda=2^{2\tilde{L}}$. Let K_b and K_c be constants defined as follows:

$$K_b=\alpha\lambda(\tilde{n}+1)-\lambda\tilde{c}^{\mathrm{T}}e-\lambda^2e^{\mathrm{T}}\tilde{Q}e, \quad K_c=\alpha\lambda. \tag{6.1}$$

The augmented problem can be stated as follows.

$$(\mathrm{P}) \qquad \begin{aligned} \min \quad & \tilde{c}^{\mathrm{T}}\tilde{x}+\tfrac{1}{2}\tilde{x}^{\mathrm{T}}\tilde{Q}\tilde{x}+K_c\tilde{x}_n \\ \text{s.t.} \quad & \tilde{A}\tilde{x}+(\tilde{b}-\lambda\tilde{A}e)\tilde{x}_n=\tilde{b}, \\ & (\alpha e-\lambda\tilde{Q}e-\tilde{c})^{\mathrm{T}}\tilde{x}+\alpha\tilde{x}_{n-1}=K_b, \\ & \tilde{x}\geqslant 0,\ \tilde{x}_{n-1}\geqslant 0,\ \tilde{x}_n\geqslant 0, \end{aligned}$$

where $\tilde{x}=(\tilde{x}_1,\ldots,\tilde{x}_{n-2})^{\mathrm{T}}$ is an $(n-2)$-vector and $\tilde{x}_{n-1}$ and $\tilde{x}_n$ are scalars. The dual problem corresponding to problem (P) is the problem stated as follows.

$$(\mathrm{D}) \qquad \begin{aligned} \max \quad & -\tfrac{1}{2}\tilde{\nu}^{\mathrm{T}}\tilde{Q}\tilde{\nu}+\tilde{b}^{\mathrm{T}}\tilde{y}+K_b\tilde{y}_m \\ \text{s.t.} \quad & -\tilde{Q}\tilde{\nu}+\tilde{A}^{\mathrm{T}}\tilde{y}+(\alpha e-\lambda\tilde{Q}e-\tilde{c})\tilde{y}_m+\tilde{z}=\tilde{c}, \\ & \alpha\tilde{y}_m+\tilde{z}_{n-1}=0, \\ & (\tilde{b}-\lambda\tilde{A}e)^{\mathrm{T}}\tilde{y}+\tilde{z}_n=K_c, \\ & \tilde{z}\geqslant 0,\ \tilde{z}_{n-1}\geqslant 0\ \ \tilde{z}_n\geqslant 0, \end{aligned}$$

where $\tilde{y}$ is an $(m-1)$-vector, $\tilde{\nu}$ and $\tilde{z}$ are $(n-2)$-vectors, $\tilde{y}_m$, $\tilde{z}_{n-1}$ and $\tilde{z}_n$ are scalars. These problems can be recast in the notation of problems (P) and (D) of Section 2 as follows. Let $x=(\tilde{x}^{\mathrm{T}},\tilde{x}_{n-1},\tilde{x}_n)^{\mathrm{T}}\in\mathbb{R}^n$, $y=(\tilde{y}^{\mathrm{T}},\tilde{y}_m)^{\mathrm{T}}\in\mathbb{R}^m$, $z=(\tilde{z}^{\mathrm{T}},\tilde{z}_{n-1},\tilde{z}_n)^{\mathrm{T}}\in\mathbb{R}^n$, and $\nu=(\tilde{\nu}^{\mathrm{T}},\tilde{\nu}_{n-1},\tilde{\nu}_n)^{\mathrm{T}}\in\mathbb{R}^n$. Define $b\in\mathbb{R}^m$, $c\in\mathbb{R}^n$ and $A\in\mathbb{R}^{m\times n}$ as follows:

$$b=\begin{pmatrix}\tilde{b}\\ K_b\end{pmatrix}, \qquad c=\begin{pmatrix}\tilde{c}\\ 0\\ K_c\end{pmatrix}, \qquad A=\begin{bmatrix}\tilde{A} & 0 & \tilde{b}-\lambda\tilde{A}e\\ (\alpha e-\lambda\tilde{Q}e-\tilde{c})^{\mathrm{T}} & \alpha & 0\end{bmatrix}. \tag{6.2}$$

Let $Q \in \mathbb{R}^{n \times n}$ denote the block diagonal matrix as follows:

$$Q = \mathrm{diag}(\tilde{Q}, 0, 0). \tag{6.3}$$

Let $L = L(A, Q, b, c)$ denote the size of (P). With this notation, we can then rewrite problems (P) and (D) as in Section 2. We refer to thse two formats interchangeably.

In the following, we intend to show some results related to the augmented problem (P). First, we show that Algorithm 3.1 can be directly applied to problem (P). We will see that an initial point w^0 belonging to the central path of this augmented problem is readily available. Second, we show that the sizes of problems (P) and $(\tilde{\mathrm{P}})$ are of the same order. Finally, we describe how a solution of problem $(\tilde{\mathrm{P}})$ can be obtained from a solution of problem (P).

We start by verifying that problem (P) satisfies Assumption 2.1 of Section 2. Assumption (c) is obviously satisfied since $\tilde{A}$ was assumed to have full row rank. We verify assumptions (a) and (b) jointly by exhibiting a point $w^0 = (x^0, y^0, z^0)$ which is in the set W defined in Section 2 and satifying the criterion of closeness (3.2). Let $x^0 \equiv (\lambda, \ldots, \lambda, 1)^{\mathrm{T}} \in \mathbb{R}^n$, $y^0 = (0, \ldots, 0, -1)^{\mathrm{T}} \in \mathbb{R}^m$ and $z^0 \equiv (\alpha, \ldots, \alpha, \alpha\lambda)^{\mathrm{T}} \in \mathbb{R}^n$. Using (6.1), one can easily verify that $Ax^0 = b$ and $-Qx^0 + A^{T}y^0 + z^0 = c$. Hence $w^0 \in W$. Moreover, $f(w^0) = \alpha\lambda e$, which implies that not only does w^0 satisfy the criterion of closeness (3.2) with $\mu_0 = \alpha\lambda$, but also that it lies on the central path Γ.

We now show that the sizes of problems (P) and $(\tilde{\mathrm{P}})$ are of the same order. The following observations are easily shown.

(1) From the definition of Q, A, b and c given by expressions (6.2) and (6.3), it follows that $\tilde{L} \leqslant L$. Since $\mu_0 = \alpha\lambda = 2^{6\tilde{L}}$, this implies that $\mu_0 = 2^{\mathrm{O}(L)}$. This fact shows that the assumption of Corollary 4.2 is satisfied.

(2) The largest absolute value of the determinant of any square submatrix of M is at most $2^{23\tilde{L}}$.

(3) $\max_i |b_i| < 2^{7\tilde{L}}$ and $\max_j |c_j| < 2^{6\tilde{L}}$.

(4) Statements (2) and (3) imply that $L \leqslant 36\tilde{L}$. Statement (1) then implies that L and $\tilde{L}$ are of the same order.

We now concentrate our effort towards showing the relationship between the optimal solutions of the augmented problem (P) with the optimal solutions of the original problem $(\tilde{\mathrm{P}})$. We start by pointing out some facts which are important for our purposes.

The Karush-Kuhn-Tucker necessary and sufficient condition for $\tilde{x} \in \mathbb{R}^{\tilde{n}}$ to be a solution of problem $(\tilde{\mathrm{P}})$ is that there exist vectors $\tilde{y} \in \mathbb{R}^{\tilde{m}}$ and $\tilde{z} \in \mathbb{R}^{\tilde{n}}$ such that the following is satisfied:

$$\begin{pmatrix} \tilde{z} \\ 0 \end{pmatrix} - \begin{bmatrix} \tilde{Q} & -\tilde{A}^{\mathrm{T}} \\ \tilde{A} & 0 \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \tilde{c} \\ -\tilde{b} \end{pmatrix}, \tag{6.4.a}$$

$$\tilde{x} \geqslant 0, \quad \tilde{z} \geqslant 0, \tag{6.4.b}$$

$$\tilde{x}^{\mathrm{T}} \tilde{z} = 0. \tag{6.4.c}$$

A well-known result from linear complementary theory is that if the system (6.4) has some solution then it has a solution which is a vertex of the polyhedron given by (6.4.a) and (6.4.b) [6]. The following lemma is a well-known result whose proof follows from an immediate application of Cramer's rule.

Lemma 6.1. *Let $\tilde{w} = (\tilde{x}, \tilde{y}, \tilde{z})$ be a vertex of the polyhedron given by* (6.4.a) *and* (6.4.b). *Then the coordinates of $\tilde{w}$ are rational numbers with numerator and denominator less than or equal to $2^{\tilde{L}}$.*

The next lemma provides an estimate useful for our purposes.

Lemma 6.2. *Let $\tilde{w} = (\tilde{x}, \tilde{y}, \tilde{z})$ be a vertex of the polyhedron given by* (6.4.a) *and* (6.4.b). *Then the right hand side coefficient K_b and the cost coefficient K_c of problem* (P) *satisfy*

$$K_b > (\tilde{b} - \lambda\tilde{A}e)^{\mathrm{T}}\tilde{y},$$

$$K_c > (\alpha e - \lambda\tilde{Q}e - \tilde{c})^{\mathrm{T}}\tilde{x}.$$

The proof of Lemma 6.2 follows from Lemma 6.1 and the definition of the constants K_b and K_c given by relation (6.1). Note that, by construction, problems (P) and (D) have feasible solutions and therefore, they have optimal solutions. We are now ready to state the relationship between the optimal solutions for the original problem with those for the augmented problem.

Lemma 6.3. *Let $x = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $(x, y, z) = ((x_1, \ldots, x_n)^{\mathrm{T}}, (y_1, \ldots, y_m)^{\mathrm{T}}, (z_1, \ldots, z_n)^{\mathrm{T}})$ be optimal solutions for problems* (P) *and* (D) *respectively. Then,* $(\tilde{\mathrm{P}})$ *has an optimal solution if, and only if, $x_n = 0$ and $z_{n-1} = 0$ (and consequently $y_m = 0$). In this case, if we let $\tilde{x} \equiv (x_1, \ldots, x_{n-2})^{\mathrm{T}}$, $\tilde{y} \equiv (y_1, \ldots, y_{m-1})^{\mathrm{T}}$ and $\tilde{z} \equiv (z_1, \ldots, z_{n-2})^{\mathrm{T}}$ then $\tilde{x}$ and $(\tilde{x}, \tilde{y}, \tilde{z})$ are optimal solutions for* $(\tilde{\mathrm{P}})$ *and* $(\tilde{\mathrm{D}})$ *respectively.*

Proof. We first prove the only if part. Assume that problem $(\tilde{\mathrm{P}})$ has an optimal solution. Then, by the observation following (6.4), there exists a vertex of the polyhedron given by (6.4.a) and (6.4.b), say $(\tilde{x}_*, \tilde{y}_*, \tilde{z}_*)$, such that $(\tilde{x}_*)^{\mathrm{T}}\tilde{z}_* = 0$. Consider the vectors $x_* \in \mathbb{R}^n$, $y_* \in \mathbb{R}^m$ and $z_* \in \mathbb{R}^n$ defined as follows:

$$x_* = (\tilde{x}_*^{\mathrm{T}}, K_b - (\alpha e - \lambda\tilde{Q}e - \tilde{c})^{\mathrm{T}}\tilde{x}_*, 0)^{\mathrm{T}},$$

$$y_* = (\tilde{y}_*^{\mathrm{T}}, 0)^{\mathrm{T}},$$

$$z_* = (\tilde{z}_*^{\mathrm{T}}, 0, K_c - (b - \lambda\tilde{A}e)^{\mathrm{T}}\tilde{y}_*)^{\mathrm{T}}.$$

Note that by Lemma 6.2, $(x_*)_{n-1} > 0$ and $(z_*)_n > 0$. Using this fact, one can easily verify that x_* and (x_*, y_*, z_*) are feasible solutions for problems (P) and (D) respectively. Moreover, since $(\tilde{x}_*)^{\mathrm{T}}\tilde{z}_* = 0$, we have $x_*^{\mathrm{T}}z_* = 0$. Hence, by complementary slackness (cf. Proposition 2.3), it follows that x_* and (x_*, y_*, z_*) are optimal solutions for (P) and (D) respectively.

Consider now any optimal solutions x and (x, y, z) of (P) and (D). Then, by complementary slackness and the fact that $(x_*)_{n-1} > 0$, it follows that $z_{n-1} = 0$, and therefore $y_m = 0$. Similarly, since $(z_*)_n > 0$, we have $x_n = 0$. This proves the only if part. Moreover, the above implies that $\tilde{x}$ and $(\tilde{x}, \tilde{y}, \tilde{z})$ are feasible solutions for $(\tilde{\mathrm{P}})$ and $(\tilde{\mathrm{D}})$ respectively and that $\tilde{x}^{\mathrm{T}}\tilde{z}=0$ since $x^{\mathrm{T}}z=0$. Thus, $\tilde{x}$ and $(\tilde{x}, \tilde{y}, \tilde{z})$ are optimal solutions for $(\tilde{\mathrm{P}})$ and $(\tilde{\mathrm{D}})$ respectively.

To see the if part, assume that $x_n=0$ and $z_{n-1}=0$ (and hence, $y_m=0$). This implies that $\tilde{x}$ and $(\tilde{x}, \tilde{y}, \tilde{z})$ are feasible solutions for $(\tilde{\mathrm{P}})$ and $(\tilde{\mathrm{D}})$ respectively. Hence, using Propositions 2.1 and 2.2, we conclude that $(\tilde{\mathrm{P}})$ has an optimal solution. □

To summarize, we now state the main result of this section.

Proposition 6.1. *Problem* (P) *can be solved in at most* $\mathrm{O}(\tilde{n}^3\tilde{L})$ *arithmetic operations.*

Proof. Applying Algorithm 3.1 to problem (P), we obtain vectors $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_m)$ and $z=(z_1,\ldots,z_n)$ such that x and (x, y, z) are optimal solutions for problems (P) and (D) respectively. Let $\tilde{x}\equiv(x_1,\ldots,x_{n-2})^{\mathrm{T}}$, $\tilde{y}\equiv(y_1,\ldots,y_{m-1})^{\mathrm{T}}$ and $\tilde{z}\equiv(z_1,\ldots,z_{n-2})^{\mathrm{T}}$. Consider the following possible cases.

(i) If $x_nz_{n-1}=0$ then it follows from Lemma 6.3 that

(a) If $x_n=0$ and $z_{n-1}=0$ then $\tilde{x}$ and $(\tilde{x}, \tilde{y}, \tilde{z})$ are optimal solutions for problems $(\tilde{\mathrm{P}})$ and $(\tilde{\mathrm{D}})$ respectively.

(b) If $x_n\neq 0$ then $(\tilde{\mathrm{P}})$ is infeasible.

(c) If $z_{n-1}\neq 0$ then $(\tilde{\mathrm{P}})$ is unbounded.

(ii) If $x_nz_{n-1}\neq 0$ then Lemma 6.3 implies that $(\tilde{\mathrm{P}})$ is either unbounded or infeasible. In this case, we solve the LP problem obtained by replacing the objective function of problem (P) by the linear function $K_c\bar{x}_n$. If the resulting optimal solution $\bar{x}$ of this problem satisfies $\bar{x}_n=0$ then $(\tilde{\mathrm{P}})$ is unbounded. Otherwise, $(\tilde{\mathrm{P}})$ is infeasible.

By Corollary 5.2, the computation above can be carried out in at most $\mathrm{O}(n^3L)$ arithmetic operations. Since $\tilde{n}=n-2$ and L and $\tilde{L}$ are of the same order, it follows that the total number of arithmetic operations spent to solve problem $(\tilde{\mathrm{P}})$ is on the order of $\mathrm{O}(\tilde{n}^3\tilde{L})$. □

7. Remarks

The purpose of this paper is to present a theoretical result. Thus in order to simplify the presentation, we constructed $\hat{\mu}=\mu(1-\delta/\sqrt{n})$. Obviously, one can use $\hat{\mu}$ which is less than or equal to the above value, but still ensures statement (b) of Theorem 4.1 and relations (4.5) and (4.6). In this way, one can accelerate the convergence of the algorithm.

We should mention that, throughout the paper, we assumed that $m=\mathrm{O}(n)$. One can easily show that the complexity achieved in Part II of this paper, for linear programming problems, expressed in terms of m and n is $\mathrm{O}((nm^2+n^{1.5}m)L)$. The first term is due to the arithmetic operations spent in the rank-one updates and the other term is due to the remaining operations. The same considerations do not apply for the general convex quadratic programming problem, except for the case when the matrix Q is diagonal.

It is well known that a linear complementarity problem with positive semi-definite matrix can be reduced to an equivalent convex quadratic programming problem and vice-versa (cf. [7]). Thus, the algorithm presented in this paper can be used to solve linear complementarity problems with positive semi-definite matrices. At the time of writing of this paper, we were informed of a recent paper by Kojima et al. [4] which present an algorithm for solving linear complementarity problems with positive semi-definite matrices. They obtained the same complexity as the one achieved in this paper.

Appendix A

In this appendix, we give the proof of Proposition 5.2.

Proof of Proposition 5.2. We first prove that for all $i = 1, \ldots, n$, we have

$$|V_i^K| \leq -[(1-\rho)\ln(1-\gamma)]^{-1} \sum_{k=0}^{K-1} \frac{|\Delta\nu_i^k|}{\nu_i^k} \tag{A.1}$$

where $\Delta\nu_i^k \equiv \nu_i^{k+1} - \nu_i^k$. Inequality (5.7) implies that for all $k = 0, \ldots, K-1$,

$$\frac{|\Delta\nu_i^k|}{\nu_i^k} = \frac{|\nu_i^{k+1} - \nu_i^k|}{\nu_i^k} \leq \rho$$

and therefore, for all $k = 0, \ldots, K-1$, we have

$$\frac{\nu_i^{k+1}}{\nu_i^k} \geq 1 - \rho.$$

The last inequality then implies that, for all $k = 0, \ldots, K-1$,

$$\frac{|\Delta\nu_i^k|}{\nu_i^{k+1}} \leq (1-\rho)^{-1} \frac{|\Delta\nu_i^k|}{\nu_i^k}. \tag{A.2}$$

Assume now that $\hat{k}$ and $\tilde{k}$ are two consecutive indices in the set V_i^K, with $\hat{k} < \tilde{k}$. Then, by the way the set V_i^K and the sequence $(\tilde{\nu}^k)_{k=0}^K$ were defined, we have

$$|\ln \nu_i^{\tilde{k}} - \ln \nu_i^{\hat{k}}| > \min\{\ln(1+\gamma), -\ln(1-\gamma)\} = -\ln(1-\gamma). \tag{A.3}$$

On the other hand, we have

$$\begin{aligned} |\ln \nu_i^{\tilde{k}} - \ln \nu_i^{\hat{k}}| &= \left| \int_{\nu_i^{\hat{k}}}^{\nu_i^{\tilde{k}}} \frac{1}{x}\, dx \right| \\ &\leq \sum_{k=\hat{k}}^{\tilde{k}-1} \max\left\{ \frac{|\Delta\nu_i^k|}{\nu_i^k}, \frac{|\Delta\nu_i^k|}{\nu_i^{k+1}} \right\} \\ &\leq (1-\rho)^{-1} \sum_{k=\hat{k}}^{\tilde{k}-1} \frac{|\Delta\nu_i^k|}{\nu_i^k} \end{aligned} \tag{A.4}$$

where the last inequality follows from (A.2). Combining (A.3) and (A.4), we obtain

$$-\ln(1-\gamma) \leqslant (1-\rho)^{-1} \sum_{k=\hat{k}}^{\tilde{k}-1} \frac{|\Delta \nu_i^k|}{\nu_i^k}. \tag{A.5}$$

Summing the last inequality over all pairs of consecutive indices in the set V_i^K and rearranging, we obtain inequality (A.1). Now, summing inequality (A.1) over all coordinate indices $i = 1, \ldots, n$, we obtain

$$\begin{aligned}
\sum_{i=1}^{n} |V_i^K| &\leqslant -[(1-\rho)\ln(1-\gamma)]^{-1} \sum_{i=1}^{n} \sum_{k=0}^{K-1} \frac{|\Delta \nu_i^k|}{\nu_i^k} \\
&= -[(1-\rho)\ln(1-\gamma)]^{-1} \sum_{k=0}^{K-1} \sum_{i=1}^{n} \frac{|\Delta \nu_i^k|}{\nu_i^k} \\
&\leqslant -[(1-\rho)\ln(1-\gamma)]^{-1} \sum_{k=0}^{K-1} \sqrt{n} \|\nu^{k+1} - \nu^k\|_{\nu^k} \\
&\leqslant -\rho[(1-\rho)\ln(1-\gamma)]^{-1} \sqrt{n} K
\end{aligned}$$

where the last inequality follows from (5.7). This completes the proof of the proposition. □

Acknowledgement

The authors wish to thank Michael Todd and two anonymous referees for many useful comments. This research was partially funded by the Brazilian Post-Graduate Education Federal Agency (CAPES) and by the United States Navy Office of Naval Research under contract N00014-87-K-0202. Their financial support is gratefully acknowledged.

References

[1] C.C. Gonzaga, "An algorithm for solving linear programming problems in $O(n^3 L)$ operations," Memorandum Number UCB/ERL M87/10, Electronics Research Laboratory, University of California (Berkeley, CA, March, 1987).

[2] S. Kappor and P.M. Vaidya, "Fast algorithms for convex quadratic programming and multicommodity flows," in: *Proceedings of the 18th Annual ACM Symposium on Theory of Computing* (Berkeley, California, May 1986) pp. 147–159.

[3] N. Karmarkar, "A new polynomial time algorithm for linear programming," *Combinatorica* 4 (1984) 373–395.

[4] M. Kojima, S. Mizuno and A. Yoshise, "A polynomial-time algorithm for a class of linear complemtarity problems," *Mathematical Programming* 44 (1989) 1–26, this issue.

[5] M.K. Kozlov, S.P. Tarasov and L.G. Khachiyan, "Polynomial solvability of convex quadratic programming," *Doklady Akademiia Nauk SSSR* 5 (1979) 1051–1053. [Translated in: *Soviet Mathematics Doklady* 20 (1979) 1108–1111.]

[6] C.E. Lemke, "On complementary pivot theory," in: G.B. Dantzig and A.F. Vienott, eds., *Mathematics of the Decision Sciences, Part I* (American Mathematical Society, Providence, RI, 1968) pp. 95-114.
[7] N. Megiddo, "Pathways to the optimal set in linear programming," Research Report, IBM Almaden Research Center (San Jose, CA, 1986).
[8] C.R. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity* (Prentice-Hall, Englewood Cliffs, New Jersey, 1982).
[9] P.M. Vaidya, "An algorithm for linear programming which requires $O(((m+n)n^2+(m+n)^{1.5}n)L)$ arithmetic operations," Preprint, AT&T Bell Laboratories (Murray Hill, NJ, 1987).
[10] Y. Ye and E. Tse, "A polynomial algorithm for convex quadratic programming," Working Paper, Department of Engineering-Economic Systems, Stanford University (Stanford, CA, 1986).