

# An Iterative Solver-Based Infeasible Primal-Dual Path-Following Algorithm for Convex QP

Zhaosong Lu <sup>\*</sup>    Renato D.C. Monteiro<sup>†</sup>    Jerome W. O'Neal <sup>‡</sup>

May 3, 2004

## Abstract

In this paper we develop an interior-point primal-dual long-step path-following algorithm for convex quadratic programming (CQP) whose search directions are computed by means of an iterative (linear system) solver. We propose a new linear system, which we refer to as the *augmented normal equation* (ANE), to determine the primal-dual search directions. Since the condition number of the matrix associated with the ANE may become large for degenerate CQP problems, we use a maximum weight basis preconditioner introduced in [16, 14] to better condition this matrix. Using a result obtained in [13], we establish a uniform bound, depending only on CQP data, on the number of iterations required for the iterative solver to obtain a sufficiently accurate solution to the ANE. Since the iterative solver can only generate an approximate solution to the ANE, this solution does not yield a primal-dual search direction satisfying all equations of the primal-dual Newton system. We propose a way to compute an inexact primal-dual search direction so that the Newton equation corresponding to the primal residual is satisfied exactly, while the one corresponding to the dual residual contains a manageable error which allows us to establish a polynomial bound on the number of iterations of our method.

**Keywords:** Convex quadratic programming, iterative solver, inexact search directions, polynomial convergence.

**AMS 2000 subject classification:** 65F10, 65F35, 90C20, 90C25, 90C51.

---

<sup>\*</sup>School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0205. (email: zhaosong@isye.gatech.edu). This author was supported in part by NSF Grant CCR-0203113 and ONR Grant N00014-03-1-0401.

<sup>†</sup>School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0205. (email: monteiro@isye.gatech.edu). This author was supported in part by NSF Grants CCR-0203113 and INT-9910084 and ONR Grant N00014-03-1-0401.

<sup>‡</sup>School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205 (email: joneal@isye.gatech.edu). This author was supported in part by the NDSEG Fellowship Program sponsored by the Department of Defense.

# 1 Introduction

In this paper we develop an interior-point primal-dual long-step path-following algorithm for convex quadratic programming (CQP) whose search directions are computed by means of an iterative (linear system) solver. More specifically, the problem we consider is:

$$\min_x \left\{ \frac{1}{2} x^T Q x + c^T x : Ax = b, x \geq 0 \right\}, \quad (1)$$

where the data are  $Q \in \mathbb{R}^{n \times n}$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ , and  $c \in \mathbb{R}^n$ , and the decision vector is  $x \in \mathbb{R}^n$ . We also assume that  $Q$  is positive semidefinite, and that a factorization  $Q = VV^T$  is explicitly given, where  $V \in \mathbb{R}^{n \times l}$ .

A similar algorithm for solving the special case of linear programming (LP), i.e. problem (1) with  $Q = 0$ , was developed by Monteiro and O’Neal in [12]. The algorithm studied in [12] is essentially the long-step path-following infeasible algorithm studied in [6, 20], the only difference being that the search directions are computed by means of an iterative solver. We refer to the iterations of the iterative solver as the *inner iterations* and to the ones performed by the interior-point method itself as the *outer iterations*. The main step of the algorithm studied in [6, 12, 20] is the computation of the primal-dual search direction  $(\Delta x, \Delta s, \Delta y)$ , whose  $\Delta y$  component can be found by solving a system of the form  $AD^2A^T \Delta y = g$ , referred to as the *normal equation*, where  $g \in \mathbb{R}^m$  and the positive diagonal matrix  $D$  depends on the current primal-dual iterate. In contrast to [6, 20], the algorithm studied in [12] uses an iterative solver to obtain an approximate solution to the normal equation. Since the condition number of the normal matrix  $AD^2A^T$  may become excessively large on degenerate LP problems (see e.g. [9]), the maximum weight basis preconditioner  $T$  introduced in [16, 14] is used to better condition this matrix and an approximate solution of the resulting equivalent system with coefficient matrix  $TAD^2A^T T^T$  is then computed. By using a result obtained in [13], which establishes that the condition number of  $TAD^2A^T T^T$  is uniformly bounded by a quantity depending on  $A$  only, Monteiro and O’Neal [12] shows that the number of inner iterations of the algorithm in [12] can be uniformly bounded by a constant depending on  $n$  and  $A$ .

In the case of CQP, the standard normal equation takes the form

$$A(Q + X^{-1}S)^{-1}A^T \Delta y = g, \quad (2)$$

for some vector  $g$ . When  $Q$  is not diagonal, the matrix  $(Q + X^{-1}S)^{-1}$  is not diagonal, and hence the coefficient matrix of (2) does not have the form required for the result of [13] to hold. To remedy this difficulty, we develop in this paper a new linear system, referred to as the *augmented normal equation* (ANE), to determine a portion of the primal-dual search direction. This equation has the form  $\tilde{A}\tilde{D}^2\tilde{A}^T u = w$ , where  $w \in \mathbb{R}^{m+l}$ ,  $\tilde{D}$  is an  $(n+l) \times (n+l)$  positive diagonal matrix and  $\tilde{A}$  is a  $2 \times 2$  block matrix of dimension  $(m+l) \times (n+l)$  whose blocks consist of  $A$ ,  $V^T$ , the zero matrix and the identity matrix (see equation (21)). As was done in [12], a maximum weight basis preconditioner  $\tilde{T}$  for the ANE is computed and an approximate solution of the resulting preconditioned equation with coefficient matrix

$\tilde{T}\tilde{A}\tilde{D}^2\tilde{A}^T\tilde{T}^T$  is generated using an iterative solver. Using the result of [13], which claims that the condition number of  $\tilde{T}\tilde{A}\tilde{D}^2\tilde{A}^T\tilde{T}^T$  is uniformly bounded regardless of  $\tilde{D}$ , we obtain a uniform bound (depending only on  $\tilde{A}$ ) on the number of inner iterations performed by the iterative solver to find a desirable approximate solution for the ANE (see Theorem 2.6).

Since the iterative solver can only generate an approximate solution to the ANE, it is clear that not all equations of the Newton system, which determines the primal-dual search direction, can be satisfied simultaneously. In the context of LP, Monteiro and O’Neal [12] proposed a recipe to compute an inexact primal-dual search direction so that the Newton equations corresponding to the primal and dual residuals were both satisfied. In the context of CQP, such an approach is no longer possible. Instead, we propose a way to compute an inexact primal-dual search direction so that the Newton equation corresponding to the primal residual is satisfied exactly, while the one corresponding to the dual residual contains a manageable error which allows us to establish a polynomial bound on the number of outer iterations of our method. Interestingly, the presence of this error on the dual residual Newton equation implies that the primal and dual residuals go to zero at different rates. This is a unique feature of our algorithm and stands in contrast with the other interior-point primal-dual path-following algorithms, where the primal and dual residuals go to zero at the same rate.

The use of inexact search directions in interior-point methods has been extensively studied in the context of LP (see e.g. [1, 5, 8, 11, 21]), but little research has been done regarding the use of inexact search directions in interior-point methods for CQP. Moreover, the use of iterative solvers to compute the primal-dual Newton search directions of interior-point path following algorithms in the context of LP have been investigated in [2, 5, 8, 14, 15, 16]; again, little research has been done regarding the use of iterative solvers in the context of CQP. In the context of CQP, Bergamaschi et al. [3] have proposed a preconditioner for the augmented system of equations and have developed an interior-point algorithm which uses an iterative solver on the preconditioned augmented system in order to obtain the search directions. However, it should be noted that this paper does not touch upon the issue of complexity bounds either from the inner or outer iteration points of view. To our knowledge, no one has used the ANE system to obtain the primal-dual search direction in the context of CQP, nor has anyone presented an iterative solver-based CQP algorithm with strong complexity bounds on both the inner and outer iterations.

Our paper is organized as follows. In Subsection 1.1, we give the terminology and notation which will be used throughout our paper. Section 2 describes the algorithm which will be the main focus of our study in this paper and presents both the inner and outer iteration-complexity results obtained for it. In Section 3, we prove the two convergence results stated Section 2. Finally, we present some concluding remarks in Section 4.

## 1.1 Terminology and Notation

Throughout this paper, upper-case Roman letters denote matrices, lower-case Roman letters denote vectors, and lower-case Greek letters denote scalars. We let  $\mathbb{R}^n$ ,  $\mathbb{R}_+^n$  and  $\mathbb{R}_{++}^n$  denote

the set of  $n$ -vectors having real, nonnegative real, and positive real components, respectively. Also, we let  $\mathbb{R}^{m \times n}$  denote the set of  $m \times n$  matrices with real entries. For a vector  $v \in \mathbb{R}^n$ , we let  $|v|$  denote the vector whose  $i$ th component is  $|v_i|$ , for every  $i = 1, \dots, n$ , and we let  $\text{Diag}(v)$  denote the diagonal matrix whose  $i$ -th diagonal element is  $v_i$ , for every  $i = 1, \dots, n$ .

Certain matrices bear special notation, namely the matrices  $X$ ,  $\Delta X$ ,  $S$ ,  $D$ , and  $\tilde{D}$ . These matrices are the diagonal matrices corresponding to the vectors  $x$ ,  $\Delta x$ ,  $s$ ,  $d$ , and  $\tilde{d}$ , respectively, as described in the previous paragraph. The symbol  $0$  will be used to denote a scalar, vector, or matrix of all zeroes; its dimensions should be clear from the context. Also, we denote by  $e$  the vector of all 1's, and by  $I$  the identity matrix; their dimensions should be clear from the context.

For a symmetric positive definite matrix  $W$ , we denote by  $\kappa(W)$  its condition number, i.e. its maximum eigenvalue divided by its minimum eigenvalue. We will denote sets by upper-case script Roman letters (e.g.  $\mathcal{B}$ ,  $\mathcal{N}$ ). For a finite set  $\mathcal{B}$ , we denote its cardinality by  $|\mathcal{B}|$ . Given a matrix  $A \in \mathbb{R}^{m \times n}$  and an ordered set  $\mathcal{B} \subseteq \{1, \dots, n\}$ , we let  $A_{\mathcal{B}}$  denote the submatrix whose columns are  $\{A_i : i \in \mathcal{B}\}$  arranged in the same order as  $\mathcal{B}$ . Similarly, given a vector  $v \in \mathbb{R}^n$  and an ordered set  $\mathcal{B} \subseteq \{1, \dots, n\}$ , we let  $v_{\mathcal{B}}$  denote the subvector consisting of the elements  $\{v_i : i \in \mathcal{B}\}$  arranged in the same order as  $\mathcal{B}$ .

We will use several different norms throughout the paper. For a vector  $z \in \mathbb{R}^n$ ,  $\|z\| = \sqrt{z^T z}$  is the Euclidian norm,  $\|z\|_1 = \sum_{i=1}^n |z_i|$  is the “1-norm”, and  $\|z\|_{\infty} = \max_{i=1, \dots, n} |z_i|$  is the “infinity norm”. For a matrix  $V \in \mathbb{R}^{m \times n}$ ,  $\|V\|$  denotes the operator norm associated with the Euclidian norm:  $\|V\| = \max_{z: \|z\|=1} \|Vz\|$ . Finally,  $\|V\|_F$  denotes the Frobenius norm:  $\|V\|_F = (\sum_{i=1}^m \sum_{j=1}^n V_{ij}^2)^{1/2}$ .

## 2 Main Results

This section, which presents the main results of our paper, is divided into four subsections. In Subsection 2.1, we will review a well-known polynomially convergent primal-dual long-step path-following algorithm for CQP whose search directions are computed by means of an exact solver. Also, we will introduce the “augmented normal equation” (ANE) approach for computing the primal-dual search directions for the above algorithm. In the next three subsections, we discuss a variant of the above algorithm where the corresponding search directions are (inexactly) computed by means of an iterative solver. In Subsection 2.2, we will discuss the use of the ANE in connection with the preconditioning results developed in Monteiro, O’Neal and Tsuchiya [13] and derive a bound on the number of iterations required for the iterative solver to obtain a sufficiently accurate solution to the ANE. In Subsection 2.3, we will discuss a suitable way to define the primal-dual search direction using an inexact solution to the ANE. Finally, in Subsection 2.4 we will present our main algorithm and the corresponding (inner and outer) iteration-complexity results.

## 2.1 Preliminaries and Motivation

In this subsection, we discuss a primal-dual long-step path-following CQP algorithm, Algorithm CQP-exact, which will serve as the basis for Algorithm CQP-IS given in Section 2.4. We will also give the complexity results which have been obtained for Algorithm CQP-exact. In particular, the complexity results obtained in this section follow as an immediate corollary from Theorem 2.5.

Consider the following primal-dual pair of CQP problems:

$$\min_x \left\{ \frac{1}{2} \|V^T x\|^2 + c^T x : Ax = b, x \geq 0 \right\}, \quad (3)$$

$$\max_{(\hat{x}, s, y)} \left\{ -\frac{1}{2} \|V^T \hat{x}\|^2 + b^T y : A^T y + s - VV^T \hat{x} = c, s \geq 0 \right\}, \quad (4)$$

where the data are  $V \in \mathbb{R}^{n \times l}$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  and  $c \in \mathbb{R}^n$ , and the decision variables are  $x \in \mathbb{R}^n$  and  $(\hat{x}, s, y) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ .

It is well-known that if  $x^*$  is an optimal solution for (3) and  $(\hat{x}^*, s^*, y^*)$  is an optimal solution for (4), then  $(x^*, s^*, y^*)$  is also an optimal solution for (4). Now, let  $\mathcal{S}$  denote the set of all vectors  $w := (x, s, y, z) \in \mathbb{R}^{2n+m+l}$  satisfying

$$Ax = b, x \geq 0, \quad (5)$$

$$A^T y + s + Vz = c, s \geq 0, \quad (6)$$

$$Xs = 0, \quad (7)$$

$$V^T x + z = 0. \quad (8)$$

It is clear that  $w \in \mathcal{S}$  if and only if  $x$  is optimal for (3),  $(x, s, y)$  is optimal for (4), and  $z = -V^T x$ . (Throughout this paper, the symbol  $w$  will always denote the quadruple  $(x, s, y, z)$ , where the vectors lie in the appropriate dimensions; similarly,  $\Delta w = (\Delta x, \Delta s, \Delta y, \Delta z)$ ,  $w^k = (x^k, s^k, y^k, z^k)$ ,  $\bar{w} = (\bar{x}, \bar{s}, \bar{y}, \bar{z})$ , etc.)

We observe that the presentation of the primal-dual path-following algorithm based on an exact solver in this subsection differs from the classical way of presenting it in that we introduce an additional variable  $z$  as above. Clearly, it is easy to see that the variable  $z$  is completely redundant and can be eliminated, thereby reducing the method described below to the usual way of presenting it. However, we observe that the introduction of the variable  $z$  plays a crucial role in the version of the algorithm below based on an iterative solver (see Algorithm CQP-IS in Subsection 2.4), where the search directions are computed only approximately.

We will make the following two assumptions throughout the paper:

**Assumption 1** *A has full row rank.*

**Assumption 2** *The set  $\mathcal{S}$  is nonempty.*

For a point  $w \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$ , let us define

$$\mu := \mu(w) = x^T s / n, \quad (9)$$

$$r_p := r_p(w) = Ax - b, \quad (10)$$

$$r_d := r_d(w) = A^T y + s + Vz - c, \quad (11)$$

$$r_V := r_V(w) = V^T x + z, \quad (12)$$

$$r := r(w) = (r_p(w), r_d(w), r_V(w)). \quad (13)$$

Moreover, given  $\gamma \in (0, 1)$  and an initial point  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$ , we define the following neighborhood of the central path:

$$\mathcal{N}_{w^0}(\gamma) := \left\{ w \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l} : Xs \geq (1 - \gamma)\mu e, \frac{\|r\|}{\|r^0\|} \leq \frac{\mu}{\mu_0} \right\}, \quad (14)$$

where  $r := r(w)$ ,  $r^0 := r(w^0)$ ,  $\mu := \mu(w)$ , and  $\mu_0 := \mu(w^0)$ . (Here, we use the convention that  $\nu/0$  is equal to 0 if  $\nu = 0$  and  $\infty$  if  $\nu > 0$ .)

The infeasible primal-dual algorithm which will serve as the basis for our iterative CQP method is as follows:

**Algorithm CQP-exact**

1. **Start:** Let  $\epsilon > 0$ ,  $\gamma \in (0, 1)$ ,  $w^0 \in \mathcal{N}_{w^0}(\gamma)$  and  $0 < \underline{\sigma} < \bar{\sigma} < 1$  be given. Set  $k = 0$ .
2. **While**  $\mu_k := \mu(w^k) > \epsilon$  **do**

(a) Let  $w := w^k$  and  $\mu := \mu_k$ ; choose  $\sigma := \sigma_k \in [\underline{\sigma}, \bar{\sigma}]$ .

(b) Let  $\Delta w = (\Delta x, \Delta s, \Delta y, \Delta z)$  denote the solution of the linear system

$$A\Delta x = -r_p, \quad (15)$$

$$A^T \Delta y + \Delta s + V\Delta z = -r_d, \quad (16)$$

$$X\Delta s + S\Delta x = -Xs + \sigma\mu e, \quad (17)$$

$$V^T \Delta x + \Delta z = -r_V. \quad (18)$$

(c) Let  $\tilde{\alpha} = \operatorname{argmax} \{ \alpha \in [0, 1] : w + \alpha' \Delta w \in \mathcal{N}_{w^0}(\gamma), \forall \alpha' \in [0, \alpha] \}$ .

(d) Let  $\bar{\alpha} = \operatorname{argmin} \{ (x + \alpha \Delta x)^T (s + \alpha \Delta s) : \alpha \in [0, \tilde{\alpha}] \}$ .

(e) Let  $w^{k+1} = w + \bar{\alpha} \Delta w$ , and set  $k \leftarrow k + 1$ .

**End** (while)

(Note: we titled this algorithm ‘‘CQP-exact’’ to distinguish it from our main algorithm in Section 2.4, where system (15)-(18) is solved inexactly using an iterative solver.)

The following result, which establishes an iteration-complexity bound for Algorithm CQP-exact, is well-known.

**Theorem 2.1** *Assume that the constants  $\gamma$ ,  $\underline{\sigma}$ , and  $\bar{\sigma}$  are such that*

$$\max \{ \gamma^{-1}, (1 - \gamma)^{-1}, \underline{\sigma}^{-1}, (1 - \bar{\sigma})^{-1} \} = \mathcal{O}(1),$$

*and that the initial point  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$  satisfies  $(x^0, s^0) \geq (x^*, s^*)$  for some  $w^* \in \mathcal{S}$ . Then, Algorithm CQP-exact finds an iterate  $w^k \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$  satisfying  $\mu_k \leq \epsilon \mu_0$  and  $\|r^k\| \leq \epsilon \|r^0\|$  within  $\mathcal{O}(n^2 \log(1/\epsilon))$  iterations.*

A few approaches have been suggested in the literature for computing the Newton search direction (15)-(18). Instead of pursuing one of them, we will discuss below a new approach, referred to in this paper as the augmented normal equation (ANE) approach, that we believe to be suitable not only for direct solvers but especially for iterative solvers as we will see in the next subsection.

Let us begin by defining the following matrices:

$$D := X^{1/2}S^{-1/2}, \quad (19)$$

$$\tilde{D} := \begin{pmatrix} D & 0 \\ 0 & I \end{pmatrix} \in \mathbb{R}^{(n+l) \times (n+l)}, \quad (20)$$

$$\tilde{A} := \begin{pmatrix} A & 0 \\ V^T & I \end{pmatrix} \in \mathbb{R}^{(m+l) \times (n+l)}. \quad (21)$$

Suppose that we first solve the following system of equations for  $(\Delta y, \Delta z)$ :

$$\tilde{A}\tilde{D}^2\tilde{A}^T \begin{pmatrix} \Delta y \\ \Delta z \end{pmatrix} = \tilde{A} \begin{pmatrix} x - \sigma\mu S^{-1}e - D^2r_d \\ 0 \end{pmatrix} + \begin{pmatrix} -r_p \\ -r_V \end{pmatrix} =: h. \quad (22)$$

This system is what we refer to as the ANE. Next, we obtain  $\Delta s$  and  $\Delta x$  according to:

$$\Delta s = -r_d - A^T \Delta y - V \Delta z, \quad (23)$$

$$\Delta x = -D^2 \Delta s - x + \sigma\mu S^{-1}e. \quad (24)$$

Clearly, the search direction  $\Delta w = (\Delta x, \Delta s, \Delta y, \Delta z)$  computed as above satisfies (16) and (17) in view of (23) and (24). Moreover, it also satisfies (15) and (18) due to the fact that by (20), (21), (22), (23) and (24), we have that

$$\begin{aligned} \tilde{A} \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix} &= \tilde{A} \begin{pmatrix} -D^2 \Delta s - x + \sigma\mu S^{-1}e \\ \Delta z \end{pmatrix} \\ &= \tilde{A} \begin{pmatrix} D^2 r_d + D^2 A^T \Delta y + D^2 V \Delta z - x + \sigma\mu S^{-1}e \\ \Delta z \end{pmatrix} \\ &= \tilde{A} \begin{pmatrix} D^2 A^T \Delta y + D^2 V \Delta z \\ \Delta z \end{pmatrix} + \tilde{A} \begin{pmatrix} D^2 r_d - x + \sigma\mu S^{-1}e \\ 0 \end{pmatrix} \\ &= \tilde{A}\tilde{D}^2\tilde{A}^T \begin{pmatrix} \Delta y \\ \Delta z \end{pmatrix} + \tilde{A} \begin{pmatrix} D^2 r_d - x + \sigma\mu S^{-1}e \\ 0 \end{pmatrix} = \begin{pmatrix} -r_p \\ -r_V \end{pmatrix}, \quad (25) \end{aligned}$$

or equivalently, (15) and (18) hold.

Theorem 2.1 assumes that we can solve  $\Delta w$  exactly. This requires (22) to be computed exactly, which is normally done by computing the Cholesky factorization of (22). In this paper, we will consider a variant of Algorithm CQP-exact which computes an approximate solution to (22) by means of an iterative solver. We then modify equations (17) and (18) to obtain the other search directions in such a manner as to ensure that the number of outer iterations is still polynomially bounded.

## 2.2 Iterative approaches for solving the ANE

In this subsection we discuss an approach for solving the ANE based on iterative solvers. Since the condition number of ANE matrix  $\tilde{A}\tilde{D}^2\tilde{A}^T$  may “blow up” for points  $w$  near an optimal solution, the direct application of a generic iterative solver for solving the ANE (22) without first preconditioning it is generally an ineffective approach. We discuss a natural remedy to this problem which consists of using a preconditioner  $\tilde{T}$ , namely the *maximum weight basis* preconditioner, such that  $\kappa(\tilde{T}\tilde{A}\tilde{D}^2\tilde{A}^T\tilde{T}^T)$  remains uniformly bounded regardless of the iterate  $w$ . Finally, we analyze the complexity of the resulting approach to obtain a suitable approximate solution to the ANE.

We start by describing the maximum weight basis preconditioner. Its construction essentially consists of building a basis  $B$  of  $\tilde{A}$  which gives higher priority to the columns of  $\tilde{A}$  corresponding to larger diagonal elements of  $\tilde{D}$ . More specifically, the maximum weight basis preconditioner is determined by the following algorithm:

### Maximum Weight Basis Algorithm

**Start:** Given  $\tilde{d} \in \mathbb{R}_{++}^{(n+l)}$ , and  $\tilde{A} \in \mathbb{R}^{(m+l) \times (n+l)}$  such that  $\text{rank}(\tilde{A}) = m + l$ ,

1. Order the elements of  $\tilde{d}$  so that  $\tilde{d}_1 \geq \dots \geq \tilde{d}_{n+l}$ ; order the columns of  $\tilde{A}$  accordingly.
2. Let  $\mathcal{B} = \emptyset$ ,  $j = 1$ .
3. **While**  $|\mathcal{B}| < m + l$  **do**
  - (a) If  $\tilde{A}_j$  is linearly independent of  $\{\tilde{A}_i : i \in \mathcal{B}\}$ , set  $\mathcal{B} \leftarrow \mathcal{B} \cup \{j\}$ .
  - (b)  $j \leftarrow j + 1$ .
4. Return to the original ordering of  $\tilde{A}$  and  $\tilde{d}$ ; determine the set  $\mathcal{B}$  according to this ordering and set  $\mathcal{N} := \{1, \dots, n + l\} \setminus \mathcal{B}$ .
5. Set  $B := \tilde{A}_{\mathcal{B}}$  and  $\tilde{D}_{\mathcal{B}} := \text{Diag}(\tilde{d}_{\mathcal{B}})$ .
6. Let  $\tilde{T} = \tilde{T}(\tilde{A}, \tilde{d}) := \tilde{D}_{\mathcal{B}}^{-1} B^{-1}$ .

**end**



Note that the above algorithm can be applied to the matrix  $\tilde{A}$  defined in (21) since this matrix has full row rank due to Assumption 1. The maximum weight basis preconditioner was originally proposed by Resende and Veiga in [16] in the context of the minimum cost network flow problem. In this case,  $\tilde{A} = A$  is the node-arc incidence matrix of a connected digraph (with one row deleted to ensure that  $\tilde{A}$  has full row rank), the entries of  $\tilde{d}$  are weights on the edges of the graph, and the set  $\mathcal{B}$  generated by the above algorithm defines a maximum spanning tree on the digraph. Oliveira and Sorensen [14] later proposed the use of this preconditioner for general matrices  $\tilde{A}$ .

For the purpose of stating the next result, we now introduce some notation. Let us define

$$\varphi_{\tilde{A}} := \max\{\|B^{-1}\tilde{A}\|_F : B \text{ is a basis of } \tilde{A}\}. \quad (26)$$

It is easy to show that  $\varphi_{\tilde{A}} \leq (n+l)^{1/2} \bar{\chi}_{\tilde{A}}$ , where  $\bar{\chi}_{\tilde{A}}$  is a well-known condition number (see [18]) defined as

$$\bar{\chi}_{\tilde{A}} := \sup\{\|\tilde{A}^T(\tilde{A}E\tilde{A}^T)^{-1}\tilde{A}E\| : E \in \text{Diag}(\mathbb{R}_{++}^{(n+l)})\}.$$

Indeed, this follows from the fact that  $\|C\|_F \leq (n+l)^{1/2} \|C\|$  for any matrix  $C \in \mathbb{R}^{(m+l) \times (n+l)}$  and that an equivalent characterization of  $\bar{\chi}_{\tilde{A}}$  is

$$\bar{\chi}_{\tilde{A}} := \max\{\|B^{-1}\tilde{A}\| : B \text{ is a basis of } \tilde{A}\},$$

as shown in [17] and [18].

Lemmas 2.1 and 2.2 of Monteiro, O'Neal and Tsuchiya [13] imply the following result.

**Proposition 2.2** *Let  $\tilde{T} = \tilde{T}(\tilde{A}, \tilde{d})$  be the preconditioner determined according to the Maximum Weight Basis Algorithm, and define  $W := \tilde{T}\tilde{A}\tilde{D}^2\tilde{A}^T\tilde{T}^T$ . Then,  $\|\tilde{T}\tilde{A}\tilde{D}\| \leq \varphi_{\tilde{A}}$  and  $\kappa(W) \leq \varphi_{\tilde{A}}^2$ .*

Note that the bound  $\varphi_{\tilde{A}}^2$  on  $\kappa(W)$  is independent of the diagonal matrix  $\tilde{D}$  and depends only on  $\tilde{A}$ . This will allow us to obtain a uniform bound on the number of iterations needed by an iterative solver to obtain a suitable approximate solution to (22). This topic is the subject of the remainder of this subsection.

Instead of dealing directly with (22), we consider the application of an iterative solver to the preconditioned ANE:

$$Wu = v, \quad (27)$$

where

$$W := \tilde{T}\tilde{A}\tilde{D}^2\tilde{A}^T\tilde{T}^T, \quad v := \tilde{T}h. \quad (28)$$

For the purpose of our analysis below, the only thing we will assume regarding the iterative solver when applied to (27) is that it generates a sequence of iterates  $\{u^j\}$  such that

$$\|v - Wu^j\| \leq c(\kappa) \left[1 - \frac{1}{f(\kappa)}\right]^j \|v - Wu^0\|, \quad \forall j = 0, 1, 2, \dots, \quad (29)$$

where  $c$  and  $f$  are positive functions of  $\kappa \equiv \kappa(W)$ .

Examples of solvers which satisfy (29) include the steepest descent (SD) and conjugate gradient (CG) methods, with the following values for  $c(\kappa)$  and  $f(\kappa)$ :

Solver	$c(\kappa)$	$f(\kappa)$
SD	$\sqrt{\kappa}$	$(\kappa + 1)/2$
CG	$2\sqrt{\kappa}$	$(\sqrt{\kappa} + 1)/2$

Table 2.2

The justification for the table above follows from Section 7.6 and Exercise 10 of Section 8.8 of [10].

The following result gives an upper bound on the number of iterations that the generic iterative linear solver needs to perform to obtain an iterate  $u^j$  satisfying  $\|v - Wu^j\| \leq \xi\sqrt{\mu}$  for some constant  $\xi > 0$ :

**Proposition 2.3** *Let  $u^0$  be an arbitrary starting point. Then, a generic iterative solver with a convergence rate given by (29) generates an iterate  $u^j$  satisfying  $\|v - Wu^j\| \leq \xi\sqrt{\mu}$  in*

$$\mathcal{O}\left(f(\kappa) \log\left(\frac{c(\kappa)\|v - Wu^0\|}{\xi\sqrt{\mu}}\right)\right) \quad (30)$$

iterations, where  $\kappa := \kappa(W)$ .

**Proof:** The proof is similar to Theorem 2.4 of [12]. ■

We will refer to an iterate  $u^j$  satisfying  $\|v - Wu^j\| \leq \xi\sqrt{\mu}$  as a  $\xi$ -approximate solution of (27). From Proposition 2.3, it is clear that different choices of the initial point  $u^0$  lead to different bounds on the number of iterations performed by the iterative solver. In Subsection 2.4, we will describe a suitable way of selecting  $u^0$  and  $\xi$  so that the resulting number of iterations performed by the iterative solver can be uniformly bounded by a universal constant depending on the quantities  $f(\kappa)$ ,  $c(\kappa)$ ,  $n$  and  $\varphi_{\tilde{A}}$  (see the bound (51)).

### 2.3 Computation of the Inexact Search Direction

In this subsection, we will discuss the impact an inexact solution to (27) has on equations (15)–(18). More specifically, we will discuss one way to define the overall search direction so that its residual with respect to system (15)–(18) is appropriate for the development of a polynomially convergent algorithm based on that search direction.

Suppose that we solve (27) inexactly according to Section 2.2. Then our final solution  $u^j$  satisfies  $Wu^j - v = \tilde{f}$  for some vector  $\tilde{f}$  such that  $\|\tilde{f}\| \leq \xi\sqrt{\mu}$ . Letting

$$\begin{pmatrix} \Delta y \\ \Delta z \end{pmatrix} = \tilde{T}^T u^j, \quad (31)$$

we see from (28) that

$$\tilde{A}\tilde{D}^2\tilde{A}^T \begin{pmatrix} \Delta y \\ \Delta z \end{pmatrix} = h + \tilde{T}^{-1}\tilde{f}. \quad (32)$$

The component  $\Delta s$  of the search direction is still obtained as before, i.e. as in (23). However, the component  $\Delta x$  is computed as

$$\Delta x = -D^2\Delta s - x + \sigma\mu S^{-1}e - S^{-1}p, \quad (33)$$

where  $p \in \mathbb{R}^n$  will be specified below. To motivate the choice of  $p$ , we first note that an argument similar to (25) implies that

$$\begin{aligned} & \tilde{A} \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix} + \begin{pmatrix} r_p \\ r_V \end{pmatrix} \\ &= \tilde{A} \begin{pmatrix} -D^2\Delta s - x + \sigma\mu S^{-1}e - S^{-1}p \\ \Delta z \end{pmatrix} + \begin{pmatrix} r_p \\ r_V \end{pmatrix} \\ &= \tilde{A} \begin{pmatrix} D^2r_d + D^2A^T\Delta y + D^2V\Delta z - x + \sigma\mu S^{-1}e - S^{-1}p \\ \Delta z \end{pmatrix} + \begin{pmatrix} r_p \\ r_V \end{pmatrix} \\ &= \tilde{A}\tilde{D}^2 \begin{pmatrix} A^T\Delta y + V\Delta z \\ \Delta z \end{pmatrix} + \tilde{A} \begin{pmatrix} D^2r_d - x + \sigma\mu S^{-1}e \\ 0 \end{pmatrix} - \tilde{A} \begin{pmatrix} S^{-1}p \\ 0 \end{pmatrix} + \begin{pmatrix} r_p \\ r_V \end{pmatrix} \\ &= \tilde{A}\tilde{D}^2\tilde{A}^T \begin{pmatrix} \Delta y \\ \Delta z \end{pmatrix} + \tilde{A} \begin{pmatrix} D^2r_d - x + \sigma\mu S^{-1}e \\ 0 \end{pmatrix} - \tilde{A} \begin{pmatrix} S^{-1}p \\ 0 \end{pmatrix} + \begin{pmatrix} r_p \\ r_V \end{pmatrix} \\ &= \tilde{T}^{-1}\tilde{f} - \tilde{A} \begin{pmatrix} S^{-1}p \\ 0 \end{pmatrix}. \end{aligned} \quad (34)$$

Based on the above equation, one is naturally tempted to choose  $p$  so that the right hand side of (34) is zero, and consequently (15) and (18) are satisfied exactly. However, the existence of such  $p$  cannot be guaranteed and, even if it exists, its magnitude may be so large as to yield a search direction which is not suitable for the development of a polynomially convergent algorithm.

Instead, we consider an alternative approach where  $p$  is chosen so that the first component of (34) is zero and the second component is small enough. More specifically, we choose  $(p, q) \in \mathbb{R}^n \times \mathbb{R}^l$  such that

$$0 = \tilde{f} - \tilde{T}\tilde{A} \begin{pmatrix} S^{-1}p \\ q \end{pmatrix} = \tilde{f} - \tilde{T}\tilde{A}\tilde{D} \begin{pmatrix} (XS)^{-1/2}p \\ q \end{pmatrix}. \quad (35)$$

Then, using (34) and (35), we conclude that

$$\tilde{A} \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix} + \begin{pmatrix} r_p \\ r_V \end{pmatrix} = \tilde{T}^{-1}\tilde{f} - \tilde{A} \begin{pmatrix} S^{-1}p \\ 0 \end{pmatrix} \quad (36)$$

from which we see that the first component of (34) is set to 0 and the second component is exactly  $q$ . We now discuss a convenient way to choose  $(p, q)$  satisfying (35). (Note that the

coefficient matrix of system (35) has full row rank, implying that (35) has multiple solutions.) First, let  $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_{m+l})$  be the ordered set of basic indices computed by the Maximum Weight Basis algorithm and note that, by step 6 of this algorithm, the  $\mathcal{B}_i$ -th column of  $\tilde{T}\tilde{A}\tilde{D}$  is the  $i$ th unit vector for every  $i = 1, \dots, m+l$ . Then, the vector  $u \in \mathbb{R}^{n+l}$  defined as  $u_{\mathcal{B}_i} = \tilde{f}_i$  for  $i = 1, \dots, m+l$  and  $u_j = 0$  for every  $j \notin \{\mathcal{B}_1, \dots, \mathcal{B}_{m+l}\}$  clearly satisfies

$$\tilde{f} = \tilde{T}\tilde{A}\tilde{D}u. \quad (37)$$

We then obtain a pair  $(p, q) \in \mathbb{R}^n \times \mathbb{R}^l$  satisfying (35) by defining

$$\begin{pmatrix} p \\ q \end{pmatrix} := \begin{pmatrix} (XS)^{1/2} & 0 \\ 0 & I \end{pmatrix} u. \quad (38)$$

We summarize the results of this subsection in the following proposition:

**Proposition 2.4** *The search direction  $\Delta w$  derived in this subsection satisfies*

$$A\Delta x = -r_p, \quad (39)$$

$$A^T\Delta y + \Delta s + V\Delta z = -r_d, \quad (40)$$

$$X\Delta s + S\Delta x = -Xs + \sigma\mu e - p, \quad (41)$$

$$V^T\Delta x + \Delta z = -r_V + q, \quad (42)$$

where the vectors  $p$  and  $q$  satisfy:

$$\|p\| \leq \|XS\|^{1/2}\|\tilde{f}\|, \quad (43)$$

$$\|q\| \leq \|\tilde{f}\|. \quad (44)$$

**Proof:** Equations (39) and (42) immediately follow from (21) and (36). Equations (40) and (41) follow from (23) and (33), respectively. The bounds on  $\|p\|$  and  $\|q\|$  follow from (38) and the fact  $\|u\| = \|\tilde{f}\|$ .  $\blacksquare$

## 2.4 The Algorithm

In this subsection, we describe the primal-dual long-step path-following algorithm based on an iterative solver that will be the main focus of the analysis of this paper. We also state the main convergence results for this algorithm, whose proofs are given in the next section. The first result essentially shows that the number of outer iterations is bounded in the same way as for its exact counterpart discussed in Subsection 2.1. The second result shows that the number of inner iterations performed by the iterative solver to compute an approximate solution of (27) is uniformly bounded by a universal constant depending only on the quantities  $f(\kappa)$ ,  $c(\kappa)$ ,  $n$  and  $\varphi_{\bar{A}}$ .

First, given a starting point  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$  and the parameters  $\eta \geq 0$ ,  $\gamma \in [0, 1]$ , and  $\theta > 0$ , we will define the following set:

$$\mathcal{N}_{w^0}(\eta, \gamma, \theta) = \left\{ w \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l} : \begin{array}{l} Xs \geq (1 - \gamma)\mu e, \quad (r_p, r_d) = \eta(r_p^0, r_d^0), \\ \|r_V - \eta r_V^0\| \leq \theta\sqrt{\mu}, \quad \eta \leq \mu/\mu_0 \end{array} \right\}, \quad (45)$$

where  $\mu = \mu(w)$ ,  $\mu_0 = \mu(w^0)$ ,  $r = r(w)$  and  $r^0 = r(w^0)$ . Next, we define the neighborhood

$$\mathcal{N}_{w^0}(\gamma, \theta) = \bigcup_{\eta \in [0, 1]} \mathcal{N}_{w^0}(\eta, \gamma, \theta). \quad (46)$$

We now present the algorithm which will be the focus of the analysis of this paper.

### Algorithm CQP-IS:

1. **Start:** Let  $\epsilon > 0$ ,  $\gamma \in (0, 1)$ ,  $\theta > 0$ ,  $w^0 \in \mathcal{N}_{w^0}(\gamma, \theta)$ , and  $0 < \underline{\sigma} < \bar{\sigma} < 4/5$  be given. Determine  $\bar{w}$  which satisfies equations (49). Set  $k = 0$ .
2. **While**  $\mu_k := \mu(w^k) > \epsilon$  **do**
  - (a) Let  $w := w^k$  and  $\mu := \mu_k$ ; choose  $\sigma \in [\underline{\sigma}, \bar{\sigma}]$ .
  - (b) Define  $D$ ,  $\tilde{A}$  and  $\tilde{D}$  according to (19), (21) and (20). Set  $r_p = Ax - b$ ,  $r_d = A^T y + s + Vz - c$ ,  $r_V = V^T x + z$ , and  $\eta = \|r_p\|/\|r_p^0\|$ .
  - (c) Build the preconditioner  $\tilde{T} = \tilde{T}(\tilde{A}, \tilde{d})$  using the Maximum Weight Basis Algorithm, and define  $W$ ,  $v$ , and  $u^0$  according to (28) and (50).
  - (d) Using  $u^0$  as the start point for an iterative solver, find an approximate solution  $u^j$  of  $Wu = v$  such that  $\tilde{f} = Wu^j - v$  satisfies  $\|\tilde{f}\| \leq \xi\sqrt{\mu}$ , where

$$\xi := \min \left\{ \frac{\gamma\sigma}{4\sqrt{n}}, \left[ \sqrt{1 + \left(1 - \frac{\gamma}{2}\right)\sigma} - 1 \right] \theta \right\}. \quad (47)$$

- (e) Solve for  $\Delta y$  and  $\Delta z$  using equation (31). Let  $(p, q)$  be computed according to (38). Compute  $\Delta s$  and  $\Delta x$  by (23) and (33).
- (f) Compute  $\tilde{\alpha} := \operatorname{argmax}\{\alpha \in [0, 1] : w + \alpha'\Delta w \in \mathcal{N}_{w^0}(\gamma, \theta), \forall \alpha' \in [0, \alpha]\}$ .
- (g) Compute  $\bar{\alpha} := \operatorname{argmin}\{(x + \alpha\Delta x)^T(s + \alpha\Delta s) : \alpha \in [0, \tilde{\alpha}]\}$ .
- (h) Let  $w^{k+1} = w + \bar{\alpha}\Delta w$ , and set  $k \leftarrow k + 1$ .

**End** (while)

Next, we present the main convergence results we have obtained for Algorithm CQP-IS. The first result gives a bound on the number of ‘‘outer’’ iterations needed by the algorithm to obtain an  $\epsilon$ -solution to the KKT conditions (5)–(8).

**Theorem 2.5** Assume that the constants  $\gamma$ ,  $\underline{\sigma}$ ,  $\bar{\sigma}$  and  $\theta$  are such that

$$\max \left\{ \gamma^{-1}, (1 - \gamma)^{-1}, \underline{\sigma}^{-1}, \left(1 - \frac{5}{4}\bar{\sigma}\right)^{-1} \right\} = \mathcal{O}(1), \quad \theta = \mathcal{O}(\sqrt{n}), \quad (48)$$

and that the initial point  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$  satisfies  $(x^0, s^0) \geq (x^*, s^*)$  for some  $w^* \in \mathcal{S}$ . Then, Algorithm CQP-IS generates an iterate  $w^k \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$  satisfying  $\mu_k \leq \epsilon^2 \mu_0$ ,  $\|(r_p^k, r_d^k)\| \leq \epsilon^2 \|(r_p^0, r_d^0)\|$ , and  $\|r_V^k\| \leq \epsilon^2 \|r_V^0\| + \epsilon \theta \sqrt{\mu_0}$  within  $\mathcal{O}(n^2 \log(1/\epsilon))$  iterations.

We will prove Theorem 2.5 in Section 3.2.

We next describe a suitable way of selecting  $u^0$  so that the resulting number of iterations performed by the iterative solver can be uniformly bounded by a universal constant depending only on the quantities  $f(\kappa)$ ,  $c(\kappa)$ ,  $n$  and  $\varphi_{\tilde{A}}$ . First, compute a point  $\bar{w} = (\bar{x}, \bar{s}, \bar{y}, \bar{z})$  such that

$$\tilde{A} \begin{pmatrix} \bar{x} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}, \quad A^T \bar{y} + \bar{s} + V \bar{z} = c. \quad (49)$$

Note that vectors  $\bar{x}$  and  $\bar{z}$  satisfying the first equation in (49) can be easily computed once a basis of  $\tilde{A}$  is available (e.g., the one computed by the Maximum Weight Basis Algorithm in the first outer iteration of Algorithm CQP-IS). Once  $\bar{y}$  is arbitrarily chosen, a vector  $\bar{s}$  satisfying the second equation of (49) is immediately available. We then define

$$u^0 = -\eta \tilde{T}^{-T} \begin{pmatrix} y^0 - \bar{y} \\ z^0 - \bar{z} \end{pmatrix}. \quad (50)$$

For the specific starting point  $u^0$  defined in (50), the following result gives an upper bound on the number of iterations that the generic iterative linear solver needs to perform to obtain an iterate  $u^j$  satisfying  $\|v - Wu^j\| \leq \xi \sqrt{\mu}$ , where  $\xi$  is given in (47). We will prove this result in Subsection 3.1.

**Theorem 2.6** Assume that  $\xi$  is defined in (47), where  $\sigma, \gamma, \theta$  are such that

$$\max\{\sigma^{-1}, \gamma^{-1}, (1 - \gamma)^{-1}, \theta, \theta^{-1}\}$$

is bounded by a polynomial in  $n$ . Assume also that  $w^0$  and  $\bar{w}$  are such that  $(x^0, s^0) > |(\bar{x}, \bar{s})|$  and  $(x^0, s^0) \geq (x^*, s^*)$  for some  $w^* \in \mathcal{S}$ . Then, a generic iterative solver with a convergence rate given by (29) generates an approximate solution  $u^j$  as in step (d) of Algorithm CQP-IS in

$$\mathcal{O}(f(\kappa) \log(c(\kappa)n\varphi_{\tilde{A}})) \quad (51)$$

iterations, where  $\kappa := \kappa(W)$ . As a consequence, the SD and CG methods generate this approximate solution  $u^j$  in  $\mathcal{O}(\varphi_{\tilde{A}}^2 \log(n\varphi_{\tilde{A}}))$  and  $\mathcal{O}(\varphi_{\tilde{A}} \log(n\varphi_{\tilde{A}}))$  iterations, respectively.

### 3 Technical Results

In this section, we will prove Theorems 2.5 and 2.6. Section 3.1 is devoted to the proof of Theorem 2.6, and Section 3.2 will give the proof of Theorem 2.5.

#### 3.1 “Inner” Iteration Results – Proof of Theorem 2.6

In this subsection, we will provide the proof of Theorem 2.6.

We begin by establishing three technical lemmas.

**Lemma 3.1** *Suppose that  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$ ,  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$  for some  $\eta \in [0, 1]$ ,  $\gamma \in [0, 1]$  and  $\theta > 0$ , and  $w^* \in \mathcal{S}$ . Then*

$$(x - \eta x^0 - (1 - \eta)x^*)^T (s - \eta s^0 - (1 - \eta)s^*) \geq -\frac{\theta^2}{4}\mu. \quad (52)$$

**Proof:** Let us define  $\tilde{w} := w - \eta w^0 - (1 - \eta)w^*$ . Using the definitions of  $\mathcal{N}_{w^0}(\eta, \gamma, \theta)$ ,  $r$ , and  $\mathcal{S}$ , we have that

$$\begin{aligned} A\tilde{x} &= 0 \\ A^T\tilde{y} + \tilde{s} + V\tilde{z} &= 0 \\ V^T\tilde{x} + \tilde{z} &= r_V - \eta r_V^0. \end{aligned}$$

Multiplying the second relation by  $\tilde{x}^T$  on the left, and using the first and third relations along with the fact that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$ , we see that

$$\begin{aligned} \tilde{x}^T \tilde{s} &= -\tilde{x}^T V \tilde{z} = [\tilde{z} - (r_V - \eta r_V^0)]^T \tilde{z} = \|\tilde{z}\|^2 - \tilde{z}^T (r_V - \eta r_V^0) \\ &\geq \|\tilde{z}\|^2 - \|\tilde{z}\| \|r_V - \eta r_V^0\| = \left( \|\tilde{z}\| - \frac{\|r_V - \eta r_V^0\|}{2} \right)^2 - \frac{\|r_V - \eta r_V^0\|^2}{4} \\ &\geq -\frac{\|r_V - \eta r_V^0\|^2}{4} \geq -\frac{\theta^2}{4}\mu. \end{aligned}$$

■

**Lemma 3.2** *Suppose that  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$  such that  $(x^0, s^0) \geq (x^*, s^*)$  for some  $w^* \in \mathcal{S}$ . Then, for any  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$  with  $\eta \in [0, 1]$ ,  $\gamma \in [0, 1]$  and  $\theta > 0$ , we have*

$$\eta(x^T s^0 + s^T x^0) \leq \left(3n + \frac{\theta^2}{4}\right)\mu. \quad (53)$$

**Proof:** Using the fact  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$  and (52), we obtain

$$\begin{aligned} x^T s - \eta(x^T s^0 + s^T x^0) + \eta^2 x^{0T} s^0 - (1 - \eta)(x^T s^* + s^T x^*) \\ + \eta(1 - \eta)(x^{*T} s^0 + s^{*T} x^0) + (1 - \eta)^2 x^{*T} s^* \geq -\frac{\theta^2}{4}\mu. \end{aligned}$$

Rearranging the terms in this equation, and using the facts that  $\eta \leq x^T s / x^{0T} s^0$ ,  $x^{*T} s^* = 0$ ,  $(x, s) \geq 0$ ,  $(x^*, s^*) \geq 0$ ,  $(x^0, s^0) > 0$ ,  $\eta \in [0, 1]$ ,  $x^* \leq x^0$ , and  $s^* \leq s^0$ , we conclude that

$$\begin{aligned}
\eta(x^T s^0 + s^T x^0) &\leq \eta^2 x^{0T} s^0 + x^T s + \eta(1 - \eta)(x^{*T} s^0 + s^{*T} x^0) + \frac{\theta^2}{4} \mu \\
&\leq \eta^2 x^{0T} s^0 + x^T s + 2\eta(1 - \eta)x^{0T} s^0 + \frac{\theta^2}{4} \mu \\
&\leq 2\eta x^{0T} s^0 + x^T s + \frac{\theta^2}{4} \mu \\
&\leq 3x^T s + \frac{\theta^2}{4} \mu = \left(3n + \frac{\theta^2}{4}\right) \mu.
\end{aligned}$$

■

**Lemma 3.3** Suppose  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$ ,  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$  for some  $\eta \in [0, 1]$ ,  $\gamma \in [0, 1]$  and  $\theta > 0$ , and  $\bar{w}$  satisfies (49). Let  $W$ ,  $v$  and  $u^0$  be given by (28) and (50), respectively. Then,

$$Wu^0 - v = \tilde{T}\tilde{A} \begin{pmatrix} -x + \sigma\mu S^{-1}e + \eta(x^0 - \bar{x}) + \eta D^2(s^0 - \bar{s}) \\ r_V - \eta r_V^0 \end{pmatrix}. \quad (54)$$

**Proof:** Using the fact that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$  along with (21), (45) and (49), we easily obtain that

$$\begin{pmatrix} r_p \\ r_V \end{pmatrix} = \begin{pmatrix} \eta r_p^0 \\ \eta r_V^0 + (r_V - \eta r_V^0) \end{pmatrix} = \eta \tilde{A} \begin{pmatrix} x^0 - \bar{x} \\ z^0 - \bar{z} \end{pmatrix} + \tilde{A} \begin{pmatrix} 0 \\ r_V - \eta r_V^0 \end{pmatrix} \quad (55)$$

$$s^0 - \bar{s} = -A^T(y^0 - \bar{y}) - V(z^0 - \bar{z}) + r_d^0. \quad (56)$$

Using relations (20), (21), (28), (45), (50), (55) and (56), we obtain

$$\begin{aligned}
Wu^0 - v &= \tilde{T}\tilde{A}\tilde{D}^2\tilde{A}^T\tilde{T}^T u^0 - \tilde{T}\tilde{A} \begin{pmatrix} x - \sigma\mu S^{-1}e - D^2 r_d \\ 0 \end{pmatrix} + \tilde{T} \begin{pmatrix} r_p \\ r_V \end{pmatrix} \\
&= -\eta\tilde{T}\tilde{A}\tilde{D}^2\tilde{A}^T \begin{pmatrix} y^0 - \bar{y} \\ z^0 - \bar{z} \end{pmatrix} - \tilde{T}\tilde{A} \begin{pmatrix} x - \sigma\mu S^{-1}e - \eta D^2 r_d^0 \\ 0 \end{pmatrix} + \tilde{T} \begin{pmatrix} r_p \\ r_V \end{pmatrix} \\
&= -\eta\tilde{T}\tilde{A} \begin{pmatrix} D^2 A^T(y^0 - \bar{y}) + D^2 V(z^0 - \bar{z}) - D^2 r_d^0 \\ z^0 - \bar{z} \end{pmatrix} \\
&\quad - \tilde{T}\tilde{A} \begin{pmatrix} x - \sigma\mu S^{-1}e \\ 0 \end{pmatrix} + \tilde{T} \begin{pmatrix} r_p \\ r_V \end{pmatrix}, \\
&= -\eta\tilde{T}\tilde{A} \begin{pmatrix} -D^2(s^0 - \bar{s}) \\ z^0 - \bar{z} \end{pmatrix} - \tilde{T}\tilde{A} \begin{pmatrix} x - \sigma\mu S^{-1}e \\ 0 \end{pmatrix} \\
&\quad + \eta\tilde{T}\tilde{A} \begin{pmatrix} x^0 - \bar{x} \\ z^0 - \bar{z} \end{pmatrix} + \tilde{T}\tilde{A} \begin{pmatrix} 0 \\ r_V - \eta r_V^0 \end{pmatrix},
\end{aligned}$$



which yields equation (54), as desired.  $\blacksquare$

The next result shows that the initial residual of the iterative solver satisfies  $\|v - Wu^0\| \leq \Psi\sqrt{\mu}$ , where  $\Psi > 0$  is a universal constant independent of the current iterate of Algorithm CQP-IS.

**Lemma 3.4** *Assume that  $\tilde{T} = \tilde{T}(\tilde{A}, \tilde{d})$  is given and that  $w^0$  and  $\bar{w}$  are such that  $(x^0, s^0) > |(\bar{x}, \bar{s})|$  and  $(x^0, s^0) \geq (x^*, s^*)$  for some  $w^* \in \mathcal{S}$ . Further, assume that  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$  for some  $\gamma \in [0, 1]$  and  $\theta > 0$ , and that  $W$ ,  $v$  and  $u^0$  are given by (28) and (50), respectively. Then, the initial residual in (29) satisfies  $\|v - Wu^0\| \leq \Psi\sqrt{\mu}$ , where*

$$\Psi := \left[ \frac{7n + \theta^2/2}{\sqrt{1 - \gamma}} + \theta \right] \varphi_{\tilde{A}}. \quad (57)$$

**Proof:** Since  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$ , we have that  $x_i s_i \geq (1 - \gamma)\mu$  for all  $i$ , which implies

$$\|(XS)^{-1/2}\| \leq \frac{1}{\sqrt{(1 - \gamma)\mu}}. \quad (58)$$

Note that  $\|Xs - \sigma\mu e\|$ , when viewed as a function of  $\sigma \in [0, 1]$ , is convex. Hence, it is maximized at one of its endpoints, which, together with the facts  $\|Xs - \mu e\| < \|Xs\|$  and  $\sigma \in [\underline{\sigma}, \bar{\sigma}] \subset [0, 1]$ , immediately implies that

$$\|Xs - \sigma\mu e\| \leq \|Xs\| \leq \|Xs\|_1 = x^T s = n\mu. \quad (59)$$

Using the fact that  $(x^0, s^0) > |(\bar{x}, \bar{s})|$  together with Lemma 3.2, we obtain that

$$\begin{aligned} \eta\|S(x^0 - \bar{x}) + X(s^0 - \bar{s})\| &\leq \eta\{\|S(x^0 - \bar{x})\| + \|X(s^0 - \bar{s})\|\} \leq 2\eta\{\|Sx^0\| + \|Xs^0\|\} \\ &\leq 2\eta(x^T s^0 + x^T \bar{s}^0) \leq \left(6n + \frac{\theta^2}{2}\right)\mu. \end{aligned} \quad (60)$$

Since  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$ , there exists  $\eta \in [0, 1]$  such that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$ . It is clear that the requirements of Lemma 3.3 are met, so equation (54) holds. By (19), (20) and (54), we see that

$$\begin{aligned} \|v - Wu^0\| &= \left\| \tilde{T}\tilde{A}\tilde{D} \begin{pmatrix} (XS)^{-1/2}\{Xs - \sigma\mu e - \eta[S(x^0 - \bar{x}) + X(s^0 - \bar{s})]\} \\ r_V - \eta r_V^0 \end{pmatrix} \right\| \\ &\leq \|\tilde{T}\tilde{A}\tilde{D}\| \left\{ \|(XS)^{-1/2}\| \left[ \|Xs - \sigma\mu e\| + \eta\|X(s^0 - \bar{s}) + S(x^0 - \bar{x})\| \right] \right. \\ &\quad \left. + \|r_V - \eta r_V^0\| \right\}, \\ &\leq \varphi_{\tilde{A}} \left\{ \frac{1}{\sqrt{(1 - \gamma)\mu}} \left[ n\mu + \left(6n + \frac{\theta^2}{2}\right)\mu \right] + \theta\sqrt{\mu} \right\} = \Psi\sqrt{\mu}, \end{aligned}$$

where the last inequality follows from Proposition 2.2, relations (58), (59), (60), and the assumption that  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$ .  $\blacksquare$

The proof of the first part of Theorem 2.6 immediately follows from Proposition 2.3 and Lemma 3.4. The proof of the second part of Theorem 2.6 follows immediately from Table 2.2 and Proposition 2.2.

### 3.2 “Outer” Iteration Results – Proof of Theorem 2.5

In this subsection, we will present the proof of Theorem 2.5. Specifically, we will show that Algorithm CQP-IS obtains an  $\epsilon$ -approximate solution to (5)–(8) in  $\mathcal{O}(n^2 \log(1/\epsilon))$  outer iterations.

Throughout this section, we use the following notation:

$$w(\alpha) := w + \alpha \Delta w, \quad \mu(\alpha) := \mu(w(\alpha)), \quad r(\alpha) := r(w(\alpha)).$$

**Lemma 3.5** *Assume that  $\Delta w$  satisfies (39)–(42) for some  $\sigma \in \mathbb{R}$ ,  $w \in \mathbb{R}^{2n+m+l}$  and  $(p, q) \in \mathbb{R}^n \times \mathbb{R}^l$ . Then, for every  $\alpha \in \mathbb{R}$ , we have:*

- (a)  $X(\alpha)s(\alpha) = (1 - \alpha)Xs + \alpha\sigma\mu e - \alpha p + \alpha^2\Delta X\Delta s$ ;
- (b)  $\mu(\alpha) = [1 - \alpha(1 - \sigma)]\mu - \alpha p^T e/n + \alpha^2\Delta x^T \Delta s/n$ ;
- (c)  $(r_p(\alpha), r_d(\alpha)) = (1 - \alpha)(r_p, r_d)$ ;
- (d)  $r_V(\alpha) = (1 - \alpha)r_V + \alpha q$ .

**Proof:** Using (41), we obtain

$$\begin{aligned} X(\alpha)s(\alpha) &= (X + \alpha\Delta X)(s + \alpha\Delta s) \\ &= Xs + \alpha(X\Delta s + S\Delta x) + \alpha^2\Delta X\Delta s \\ &= Xs + \alpha(-Xs + \sigma\mu e - p) + \alpha^2\Delta X\Delta s \\ &= (1 - \alpha)Xs + \alpha\sigma\mu e - \alpha p + \alpha^2\Delta X\Delta s, \end{aligned}$$

thereby showing that (a) holds. Left multiplying the above equality by  $e^T$  and dividing the resulting expression by  $n$ , we easily conclude that (b) holds. Statement (c) can be easily verified by means of (39) and (40), while statement (d) follows from (42).  $\blacksquare$

**Lemma 3.6** *Assume that  $\Delta w$  satisfies (39)–(42) for some  $\sigma \in \mathbb{R}$ ,  $w \in \mathbb{R}^{2n+m+l}$  and  $(p, q) \in \mathbb{R}^n \times \mathbb{R}^l$  such that  $p^T e/n \leq \sigma\mu/2$ . Then, for every scalar  $\alpha$  satisfying*

$$0 \leq \alpha \leq \min \left\{ 1, \frac{\sigma\mu}{2\|\Delta X\Delta s\|_\infty} \right\}, \quad (61)$$

we have

$$\frac{\mu(\alpha)}{\mu} \geq 1 - \alpha. \quad (62)$$

**Proof:** Using Lemma 3.5(b) and the assumption that  $p^T e/n \leq \sigma\mu/2$ , we conclude for every  $\alpha$  satisfying (61) that

$$\begin{aligned} \mu(\alpha) &= [1 - \alpha(1 - \sigma)]\mu - \alpha p^T e/n + \alpha^2 \Delta x^T \Delta s/n \\ &\geq [1 - \alpha(1 - \sigma)]\mu - \frac{1}{2} \alpha \sigma \mu + \alpha^2 \Delta x^T \Delta s/n \\ &\geq (1 - \alpha)\mu + \frac{1}{2} \alpha \sigma \mu - \alpha^2 \|\Delta X \Delta s\|_\infty \\ &\geq (1 - \alpha)\mu. \end{aligned}$$

■

**Lemma 3.7** Assume that  $\Delta w$  satisfies (39)-(42) for some  $\sigma > 0$ ,  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$  with  $\gamma \in [0, 1]$ ,  $\theta \geq 0$ , and  $(p, q) \in \mathbb{R}^n \times \mathbb{R}^l$  such that

$$\|p\|_\infty \leq \gamma \sigma \mu / 4 \quad (63)$$

and

$$\|q\| \leq \left[ \sqrt{1 + \left(1 - \frac{\gamma}{2}\right) \sigma} - 1 \right] \theta \sqrt{\mu}. \quad (64)$$

Then,  $w(\alpha) \in \mathcal{N}_{w^0}(\gamma, \theta)$  for every scalar  $\alpha$  satisfying

$$0 \leq \alpha \leq \min \left\{ 1, \frac{\gamma \sigma \mu}{4 \|\Delta X \Delta s\|_\infty} \right\}. \quad (65)$$

**Proof:** Since  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$ , there exists  $\eta \in [0, 1]$  such that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$ . We will show that  $w(\alpha) \in \mathcal{N}_{w^0}((1 - \alpha)\eta, \gamma, \theta) \subseteq \mathcal{N}_{w^0}(\gamma, \theta)$  for every  $\alpha$  satisfying (65).

First, we note that  $(r_p(\alpha), r_d(\alpha)) = (1 - \alpha)\eta(r_p^0, r_d^0)$  by Lemma 3.5(c) and the definition of  $\mathcal{N}_{w^0}(\eta, \gamma, \theta)$ . Next, since the assumption that  $\gamma \in [0, 1]$  and equation (63) imply that  $p^T e/n \leq \sigma\mu/2$ , it follows from Lemma 3.6 that (62) holds for every  $\alpha$  satisfying (61), and hence (65). Thus, for every  $\alpha$  satisfying (65), we have

$$(1 - \alpha)\eta \leq \frac{\mu(\alpha)}{\mu} \eta \leq \frac{\mu(\alpha)}{\mu} \frac{\mu}{\mu_0} = \frac{\mu(\alpha)}{\mu_0}. \quad (66)$$

Now, it is easy to see that for every  $u \in \mathbb{R}^n$  and  $\tau \in [0, n]$ , there holds  $\|u - \tau(u^T e/n)e\|_\infty \leq (1 + \tau)\|u\|_\infty$ . Using this inequality twice, the fact that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$ , relation (63) and

statements (a) and (b) of Lemma 3.5, we conclude for every  $\alpha$  satisfying (65) that

$$\begin{aligned}
& X(\alpha)s(\alpha) - (1 - \gamma)\mu(\alpha)e \\
&= (1 - \alpha)[Xs - (1 - \gamma)\mu e] + \alpha\gamma\sigma\mu e - \alpha \left[ p - (1 - \gamma) \left( \frac{p^T e}{n} \right) e \right] \\
&\quad + \alpha^2 \left[ \Delta X \Delta s - (1 - \gamma) \left( \frac{\Delta x^T \Delta s}{n} \right) e \right] \\
&\geq \alpha \left[ \gamma\sigma\mu - \left\| p - (1 - \gamma) \frac{p^T e}{n} e \right\|_\infty - \alpha \left\| \Delta X \Delta s - (1 - \gamma) \frac{\Delta x^T \Delta s}{n} e \right\|_\infty \right] e \\
&\geq \alpha \left( \gamma\sigma\mu - 2\|p\|_\infty - 2\alpha\|\Delta X \Delta s\|_\infty \right) e \geq \alpha \left( \gamma\sigma\mu - \frac{1}{2}\gamma\sigma\mu - \frac{1}{2}\gamma\sigma\mu \right) e = 0.
\end{aligned}$$

Next, by Lemma 3.5(d), we have that

$$r_V(\alpha) = (1 - \alpha)r_V + \alpha q = (1 - \alpha)\eta r_V^0 + \hat{a},$$

where  $\hat{a} = (1 - \alpha)(r_V - \eta r_V^0) + \alpha q$ . To complete the proof, it suffices to show that  $\|\hat{a}\| \leq \theta\sqrt{\mu(\alpha)}$  for every  $\alpha$  satisfying (65). By using equation (64) and Lemma 3.5(b) along with the facts that  $\|r_V - \eta r_V^0\| \leq \theta\sqrt{\mu}$  and  $\alpha \in [0, 1]$ , we have

$$\begin{aligned}
\|\hat{a}\|^2 - \theta^2\mu(\alpha) &= (1 - \alpha)^2\|r_V - \eta r_V^0\|^2 + 2\alpha(1 - \alpha)(r_V - \eta r_V^0)^T q + \alpha^2\|q\|^2 - \theta^2\mu(\alpha) \\
&\leq (1 - \alpha)^2\theta^2\mu + 2\alpha(1 - \alpha)\theta\sqrt{\mu}\|q\| + \alpha^2\|q\|^2 \\
&\quad - \theta^2 \left\{ [1 - \alpha(1 - \sigma)]\mu - \alpha \frac{p^T e}{n} + \alpha^2 \frac{\Delta x^T \Delta s}{n} \right\} \\
&\leq \alpha^2\|q\|^2 + 2\alpha\theta\sqrt{\mu}\|q\| - \alpha\theta^2\sigma\mu + \alpha\theta^2 \frac{p^T e}{n} - \alpha^2\theta^2 \frac{\Delta x^T \Delta s}{n} \\
&\leq \alpha \left[ \|q\|^2 + 2\theta\sqrt{\mu}\|q\| - \left(1 - \frac{\gamma}{4}\right) \theta^2\sigma\mu + \theta^2\alpha\|\Delta X \Delta s\|_\infty \right] \\
&\leq \alpha \left[ \|q\|^2 + 2\theta\sqrt{\mu}\|q\| - \left(1 - \frac{\gamma}{2}\right) \theta^2\sigma\mu \right] \leq 0,
\end{aligned}$$

where the last inequality follows from the quadratic formula and equation (64).  $\blacksquare$

Next, we consider the minimum step length allowed under our algorithm:

**Lemma 3.8** *In every iteration of Algorithm CQP-IS, the step length  $\bar{\alpha}$  satisfies*

$$\bar{\alpha} \geq \min \left\{ 1, \frac{\min\{\gamma\sigma, 1 - \frac{5}{4}\sigma\}\mu}{4\|\Delta X \Delta s\|_\infty} \right\} \quad (67)$$

and

$$\mu(\bar{\alpha}) \leq \left[ 1 - \left(1 - \frac{5}{4}\sigma\right) \frac{\bar{\alpha}}{2} \right] \mu. \quad (68)$$

**Proof:** By equations (43)–(44) and the bound on  $\|\tilde{f}\|$  in step (d) of Algorithm CQP-IS, we conclude that

$$\|p\|_\infty \leq \|p\| \leq \|XS\|^{1/2}\|\tilde{f}\| \leq \sqrt{n\mu} \frac{\gamma\sigma}{4\sqrt{n}}\sqrt{\mu} = \frac{\gamma\sigma\mu}{4}, \quad (69)$$

$$\|q\| \leq \|\tilde{f}\| \leq \left[ \sqrt{1 + \left(1 - \frac{\gamma}{2}\right)\sigma} - 1 \right] \theta\sqrt{\mu}, \quad (70)$$

so (63) and (64) are satisfied. Hence, by Lemma 3.7, the quantity  $\tilde{\alpha}$  computed in step (g) of Algorithm CQP-IS satisfies

$$\tilde{\alpha} \geq \min \left\{ 1, \frac{\gamma\sigma\mu}{4\|\Delta X \Delta s\|_\infty} \right\}. \quad (71)$$

Moreover, by (69), it follows that the coefficient of  $\alpha$  in the expression for  $\mu(\alpha)$  in Lemma 3.5(b) satisfies

$$-(1-\sigma)\mu - \frac{p^T e}{n} \leq -(1-\sigma)\mu + \|p\|_\infty \leq -(1-\sigma)\mu + \frac{1}{4}\gamma\sigma\mu = -\left(1 - \frac{5}{4}\sigma\right)\mu < 0, \quad (72)$$

since  $\sigma \in (0, \frac{4}{5})$ . Hence, if  $\Delta x^T \Delta s \leq 0$ , it is easy to see that  $\bar{\alpha} = \tilde{\alpha}$ , and hence that (67) holds in view of (71). Moreover, by Lemma 3.5(b) and (72), we have

$$\mu(\bar{\alpha}) \leq [1 - \bar{\alpha}(1-\sigma)]\mu - \bar{\alpha}\frac{p^T e}{n} \leq \left[1 - \left(1 - \frac{5}{4}\sigma\right)\bar{\alpha}\right]\mu \leq \left[1 - \left(1 - \frac{5}{4}\sigma\right)\frac{\bar{\alpha}}{2}\right]\mu,$$

showing that (68) also holds. We now consider the case where  $\Delta x^T \Delta s > 0$ . In this case, we have  $\bar{\alpha} = \min\{\alpha_{\min}, \tilde{\alpha}\}$ , where  $\alpha_{\min}$  is the unconstrained minimum of  $\mu(\alpha)$ . It is easy to see that

$$\alpha_{\min} = \frac{n\mu(1-\sigma) + p^T e}{2\Delta x^T \Delta s} \geq \frac{n[\mu(1-\sigma) - \frac{1}{4}\sigma\mu]}{2\Delta x^T \Delta s} \geq \frac{\mu(1 - \frac{5}{4}\sigma)}{2\|\Delta X \Delta s\|_\infty}.$$

The last two observations together with (71) imply that (67) holds in this case too. Moreover, since the function  $\mu(\alpha)$  is convex, it must lie below the function  $\phi(\alpha)$  over the interval  $[0, \alpha_{\min}]$ , where  $\phi(\alpha)$  is the affine function interpolating  $\mu(\alpha)$  at  $\alpha = 0$  and  $\alpha = \alpha_{\min}$ . Hence,

$$\mu(\bar{\alpha}) \leq \phi(\bar{\alpha}) = \left[1 - (1-\sigma)\frac{\bar{\alpha}}{2}\right]\mu - \bar{\alpha}\frac{p^T e}{2n} \leq \left[1 - \left(1 - \frac{5}{4}\sigma\right)\frac{\bar{\alpha}}{2}\right]\mu, \quad (73)$$

where the second inequality follows from (72). We have thus shown that  $\bar{\alpha}$  satisfies (68). ■

Our next task will be to show that the stepsize  $\bar{\alpha}$  remains bounded away from zero. In view of (67), it suffices to show that the quantity  $\|\Delta X \Delta s\|_\infty$  remains bounded. The next lemma addresses this issue.

**Lemma 3.9** Let  $w^0 \in \mathbb{R}_{++}^{2n} \times \mathbb{R}^{m+l}$  be such that  $(x^0, s^0) \geq (x^*, s^*)$  for some  $w^* \in \mathcal{S}$ , and let  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$  for some  $\gamma \geq 0$  and  $\theta \geq 0$ . Then, the search direction  $\Delta w$  generated by Algorithm CQP-IS satisfies

$$\begin{aligned} \max(\|D^{-1}\Delta x\|, \|D\Delta s\|) &\leq \left(1 - 2\sigma + \frac{\sigma^2}{1 - \gamma}\right)^{1/2} \sqrt{n\mu} \\ &+ \frac{1}{\sqrt{1 - \gamma}} \left(\frac{\gamma\sigma}{4} + 6n + \frac{\theta^2}{2}\right) \sqrt{\mu} + \theta\sqrt{\mu}. \end{aligned} \quad (74)$$

**Proof:** Since  $w \in \mathcal{N}_{w^0}(\gamma, \theta)$ , there exists  $\eta \in [0, 1]$  such that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$ . Let  $\widetilde{\Delta w} := \Delta w + \eta(w^0 - w^*)$ . Using relations (39), (40), (42), and the fact that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$ , we easily see that

$$A\widetilde{\Delta x} = 0 \quad (75)$$

$$A^T\widetilde{\Delta y} + \widetilde{\Delta s} + V\widetilde{\Delta z} = 0, \quad (76)$$

$$V^T\widetilde{\Delta x} + \widetilde{\Delta z} = q - r_V + \eta r_V^0. \quad (77)$$

Pre-multiplying (76) by  $\widetilde{\Delta x}^T$  and using (75) and (77), we obtain

$$\begin{aligned} \widetilde{\Delta x}^T \widetilde{\Delta s} &= -\widetilde{\Delta x}^T V\widetilde{\Delta z} = [\widetilde{\Delta z} - (q - r_V + \eta r_V^0)]^T \widetilde{\Delta z} = \|\widetilde{\Delta z}\|^2 - (q - r_V + \eta r_V^0)^T \widetilde{\Delta z} \\ &\geq \|\widetilde{\Delta z}\|^2 - \|q - r_V + \eta r_V^0\| \|\widetilde{\Delta z}\| \geq -\frac{\|q - r_V + \eta r_V^0\|^2}{4}. \end{aligned} \quad (78)$$

Next, we multiply equation (41) by  $(XS)^{-1/2}$  to obtain  $D^{-1}\Delta x + D\Delta s = H(\sigma) - (XS)^{-1/2}p$ , where  $H(\sigma) := -(XS)^{1/2}e + \sigma\mu(XS)^{-1/2}e$ . Equivalently, we have that

$$D^{-1}\widetilde{\Delta x} + D\widetilde{\Delta s} = H(\sigma) - (XS)^{-1/2}p + \eta [D(s^0 - s^*) + D^{-1}(x^0 - x^*)] =: g.$$

Taking the squared norm of both sides of the above equation and using (78), we obtain

$$\begin{aligned} \|D^{-1}\widetilde{\Delta x}\|^2 + \|D\widetilde{\Delta s}\|^2 &= \|g\|^2 - 2\widetilde{\Delta x}^T \widetilde{\Delta s} \leq \|g\|^2 + \frac{\|q - r_V + \eta r_V^0\|^2}{2} \\ &\leq \left(\|g\| + \frac{\|q - r_V + \eta r_V^0\|}{\sqrt{2}}\right)^2 \leq (\|g\| + \theta\sqrt{\mu})^2, \end{aligned}$$

since  $\|q\| + \|r_V - \eta r_V^0\| \leq [\sqrt{2} - 1] \theta\sqrt{\mu} + \theta\sqrt{\mu} = \sqrt{2}\theta\sqrt{\mu}$  by (45), (70), and the fact that  $1 + (1 - \gamma/2)\sigma \leq 2$ . Thus, we have

$$\begin{aligned} \max(\|D^{-1}\widetilde{\Delta x}\|, \|D\widetilde{\Delta s}\|) &\leq \|g\| + \theta\sqrt{\mu} \\ &\leq \|H(\sigma)\| + \|(XS)^{-1/2}\| \|p\| + \eta [\|D(s^0 - s^*)\| + \|D^{-1}(x^0 - x^*)\|] + \theta\sqrt{\mu}. \end{aligned}$$

This, together with the triangle inequality, the definitions of  $D$  and  $\widetilde{\Delta w}$ , and the fact that  $w \in \mathcal{N}_{w^0}(\eta, \gamma, \theta)$ , imply that

$$\begin{aligned}
& \max(\|D^{-1}\Delta x\|, \|D\Delta s\|) \\
& \leq \|H(\sigma)\| + \|(XS)^{-1/2}\| \|p\| + 2\eta [\|D(s^0 - s^*)\| + \|D^{-1}(x^0 - x^*)\|] + \theta\sqrt{\mu} \\
& \leq \|H(\sigma)\| + \|(XS)^{-1/2}\| \|p\| + 2\eta\|(XS)^{-1/2}\| [\|X(s^0 - s^*)\| + \|S(x^0 - x^*)\|] + \theta\sqrt{\mu} \\
& \leq \|H(\sigma)\| + \frac{1}{\sqrt{(1-\gamma)\mu}} [\|p\| + 2\eta(\|X(s^0 - s^*)\| + \|S(x^0 - x^*)\|)] + \theta\sqrt{\mu}. \tag{79}
\end{aligned}$$

It is well-known (see e.g. [7]) that

$$\|H(\sigma)\| \leq \left(1 - 2\sigma + \frac{\sigma^2}{1-\gamma}\right)^{1/2} \sqrt{n\mu}. \tag{80}$$

Moreover, using the fact that  $s^* \leq s^0$  and  $x^* \leq x^0$  along with Lemma 3.2, we obtain

$$\eta(\|X(s^0 - s^*)\| + \|S(x^0 - x^*)\|) \leq \eta(s^T x^0 + x^T s^0) \leq \left(3n + \frac{\theta^2}{4}\right) \mu. \tag{81}$$

Finally, it is clear from (69) that

$$\|p\| \leq \frac{\gamma\sigma\mu}{4}. \tag{82}$$

The result now follows by incorporating inequalities (80)–(82) into (79).  $\blacksquare$

We are now ready to prove Theorem 2.5.

**Proof of Theorem 2.5:** Let  $\Delta w^k$  denote the search direction, and let  $r^k = r(w^k)$  and  $\mu_k = \mu(w^k)$ , at the  $k$ -th iteration of Algorithm CQP-IS. Clearly,  $w^k \in \mathcal{N}_{w^0}(\gamma, \theta)$ . Hence, using Lemma 3.9, assumption (48) and the inequality

$$\|\Delta X^k \Delta s^k\|_\infty \leq \|\Delta X^k \Delta s^k\| \leq \|(D^k)^{-1} \Delta x^k\| \|D^k \Delta s^k\|,$$

we easily see that  $\|\Delta X^k \Delta s^k\|_\infty = \mathcal{O}(n^2)\mu_k$ . Using this conclusion together with assumption (48) and Lemma 3.8, we see that, for some universal constant  $\beta > 0$ , we have

$$\mu_{k+1} \leq \left(1 - \frac{\beta}{n^2}\right) \mu_k, \quad \forall k \geq 0.$$

Using this observation and some standard arguments (see, for example, Theorem 3.2 of [19]), we easily see that the Algorithm CQP-IS generates an iterate  $w^k \in \mathcal{N}_{w^0}(\gamma, \theta)$  satisfying  $\mu_k/\mu_0 \leq \epsilon^2$  within  $\mathcal{O}(n^2 \log(1/\epsilon))$  iterations. The theorem now follows from this conclusion and the definition of  $\mathcal{N}_{w^0}(\gamma, \theta)$ .  $\blacksquare$

## 4 Concluding Remarks

We have shown that the primal-dual long-step path-following LP algorithm based on an iterative solver presented in [12] can be extended to the CQP case. This was not immediately obvious at first, since the standard normal equation for CQP does not fit into the mold required for the results in [13] to hold. By creating the ANE, we were able to use the results about the maximum weight basis preconditioner developed in [13] in the context of CQP. Another difficulty we encountered was the proper choice of the starting iterate  $u^0$  for the iterative solver. By choosing  $u^0 = 0$  as in the LP case, we obtain  $\|v - Wu^0\| = \|v\|$ , which can only be shown to be  $\mathcal{O}(\max\{\mu, \sqrt{\mu}\})$ . In this case, for every  $\mu > 1$ , Proposition 2.3 would guarantee that the number of inner iterations of the iterative solver is

$$\mathcal{O}(f(\kappa) \max\{\log(c(\kappa)n\varphi_{\bar{A}}), \log \mu\}),$$

a bound which depends on the logarithm of the current duality gap. On the other hand, Theorem 2.6 shows that choosing  $u^0$  as in (50) results in a bound on the number of inner iterations of the iterative solver that does not depend on the current duality gap.

The usual way of defining the dual residual is as the quantity

$$R_d := A^T y + s - VV^T x - c,$$

which, in view of (11) and (12), can be written in terms of the residuals  $r_d$  and  $r_V$  as

$$R_d = r_d - Vr_V. \tag{83}$$

Note that along the iterates generated by Algorithm CQP-IS, we have  $r_d = \mathcal{O}(\mu)$  and  $r_V = \mathcal{O}(\sqrt{\mu})$ , implying that  $R_d = \mathcal{O}(\sqrt{\mu})$ . Hence, while the usual primal residual converges to 0 according to  $\mathcal{O}(\mu)$ , the usual dual residual does so according to  $\mathcal{O}(\sqrt{\mu})$ . This is an unusual feature of the algorithm presented in this paper, and one which contrasts it with other infeasible interior-point path-following algorithms, which require that the primal and dual residuals go to zero at an  $\mathcal{O}(\mu)$ -rate. It should be noted however that this feature is possible due to the specific form of the  $\mathcal{O}(\sqrt{\mu})$ -term present in (83), i.e. one that lies in the range space of  $V$ .

CQP problems where  $V$  is explicitly available arise frequently in the literature. One important example arises in portfolio optimization (see [4]), where the rank of  $V$  is often small. In such problems,  $l$  represents the number of observation periods used to estimate the data for the problem. We believe that Algorithm CQP-IS could be of particular use for this type of application.

## References

- [1] V. Baryamureeba and T. Steihaug. On the convergence of an inexact primal-dual interior point method for linear programming. Technical Report 188, Department of Informatics, University of Bergen, 2000.



- [2] V. Baryamureeba, T. Steihaug, and Y. Zhang. Properties of a class of preconditioners for weighted least squares problems. Technical Report 16, Department of Computational and Applied Mathematics, Rice University, 1999.
- [3] L. Bergamaschi, J. Gondzio, and G. Zilli. Preconditioning indefinite systems in interior point methods for optimization. Technical Report MS-02-002, Department of Mathematical Methods for Applied Sciences, University of Padua, Italy, 2002.
- [4] T.J. Carpenter and R.J. Vanderbei. Symmetric indefinite systems for interior-point methods. *Mathematical Programming*, 58:1–32, 1993.
- [5] R.W. Freund, F. Jarre, and S. Mizuno. Convergence of a class of inexact interior-point algorithms for linear programs. *Mathematics of Operations Research*, 24(1):50–71, 1999.
- [6] M. Kojima, N. Megiddo, and S. Mizuno. A primal-dual infeasible-interior-point algorithm for linear programming. *Mathematical Programming*, 61(3):263–280, 1993.
- [7] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point algorithm for linear programming. in *Progress in Mathematical Programming: Interior-Point and Related Methods*, ch. 2, pages 29–47, 1989.
- [8] J. Korczak. Convergence analysis of inexact infeasible-interior-point algorithms for solving linear programming problems. *SIAM Journal on Optimization*, 11(1):133–148, 2000.
- [9] V.V. Kovacevic-Vujcic and M.D. Asic. Stabilization of interior-point methods for linear programming. *Computational Optimization and Applications*, 14:331–346, 1999.
- [10] D.G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, 1984.
- [11] S. Mizuno and F. Jarre. Global and polynomial-time convergence of an infeasible-interior-point algorithm using inexact computation. *Mathematical Programming*, 84:357–373, 1999.
- [12] R.D.C. Monteiro and J.W. O’Neal. Convergence analysis of a long-step primal-dual infeasible interior-point lp algorithm based on iterative linear solvers. Technical report, Georgia Institute of Technology, 2003. Submitted to *Mathematical Programming*.
- [13] R.D.C. Monteiro, J.W. O’Neal, and T. Tsuchiya. Uniform boundedness of a preconditioned normal matrix used in interior point methods. Technical report, Georgia Institute of Technology, 2003. Submitted to *SIAM Journal on Optimization*.
- [14] A.R.L. Oliveira and D.C. Sorensen. Computational experience with a preconditioner for interior point methods for linear programming. Technical Report 28, Department of Computational and Applied Mathematics, Rice University, 1997.
- [15] L.F. Portugal, M.G.C. Resende, G. Veiga, and J.J. Judice. A truncated primal-infeasible dual-feasible network interior point method. *Networks*, 35:91–108, 2000.

- [16] M.G.C. Resende and G. Veiga. An implementation of the dual affine scaling algorithm for minimum cost flow on bipartite uncapacitated networks. *SIAM Journal on Optimization*, 3:516–537, 1993.
- [17] M.J. Todd, L. Tunçel, and Y. Ye. Probabilistic analysis of two complexity measures for linear programming problems. *Mathematical Programming A*, 90:59–69, 2001.
- [18] S.A. Vavasis and Y. Ye. A primal-dual interior point method whose running time depends only on the constraint matrix. *Mathematical Programming A*, 74:79–120, 1996.
- [19] S.J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, 1997.
- [20] Y. Zhang. On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. *SIAM Journal on Optimization*, 4(1):208–227, 1994.
- [21] G. Zhou and K.-C. Toh. Polynomiality of an inexact infeasible interior point algorithm for semidefinite programming. *Mathematical Programming*. to appear.