Convergence Analysis of an Inexact Feasible Interior Point Method for Convex Quadratic Programming*

Jacek Gondzio†

School of Mathematics
and Maxwell Institute for Mathematical Sciences
The University of Edinburgh
Mayfield Road, Edinburgh EH9 3JZ
United Kingdom.

Abstract

In this paper we will discuss two variants of an inexact feasible interior point algorithm for convex quadratic programming. We will consider two different neighbourhoods: a (small) one induced by the use of the Euclidean norm which yields a short-step algorithm and a symmetric one induced by the use of the infinity norm which yields a (practical) long-step algorithm. Both algorithms allow for the Newton equation system to be solved inexactly. For both algorithms we will provide conditions for the level of error acceptable in the Newton equation and establish the worst-case complexity results.

Keywords: Inexact Newton Method, Interior Point Algorithms, Linear Programming, Quadratic Programming, Worst-case Complexity Analysis, Matrix-Free Methods.

---

†Email: `J.Gondzio@ed.ac.uk`, URL: `http://maths.ed.ac.uk/~gondzio/`
‡For other papers in this series see `http://www.maths.ed.ac.uk/ERGO/`

1 Introduction

It is broadly accepted that interior point methods (IPMs) provide very efficient solution techniques for linear and convex quadratic programming problems [11, 21]. Interior point algorithms for these classes of problems enjoy excellent worst-case complexity bounds: indeed the best known algorithms find the ε-accurate optimal solutions to the problem with n variables in $\mathcal{O}(\sqrt{n}\log(1/\varepsilon))$ or $\mathcal{O}(n\log(1/\varepsilon))$ iterations, depending on how aggressive steps to optimality are allowed. Computational experience provides evidence that the algorithm which uses a more aggressive strategy (the so-called long-step method) solves linear and quadratic programming problems in a number of iterations which may be expressed as $\mathcal{O}(\log n \log(1/\varepsilon))$ [11].

The small number of iterations does not always guarantee the efficiency of the method because occasionally IPMs struggle with a high per-iteration cost of the linear algebra operations. In the most adverse case the cost of solving a dense optimization problem employing a direct linear algebra method to solve the Newton equation system may reach $\mathcal{O}(n^3)$ flops per iteration. The effort of a single IPM iteration is usually significantly lower than this upper bound. However if problems are very large then, although they may display reasonable sparsity features, the use of direct sparsity-exploiting linear algebra techniques may still run into trouble due to excessive memory requirements or unacceptably long CPU time. Iterative methods for linear equations such as conjugate gradients or other approaches from the Krylov subspace family may offer a viable alternative to direct methods in such cases.

The interest in the use of iterative methods to solve the Newton equation system in IPMs has been growing over the last decade: see for example the recent survey of D'Apuzzo et al. [8]. Iterative methods offer several advantages. In particular they are often memory efficient and hence allow much larger problems to be solved. However to take full advantage of iterative methods one usually has to relax the accuracy requirements in the solution of the Newton equation system. Consequently, instead of using an *exact* Newton direction, the resulting IPM employs an *inexact* one. This opens up all sorts of theoretical questions. We will state and answer some of these in this paper. In particular, we will discuss the key issue concerning an acceptable level of error in the inexact Newton method used by an IPM.

The use of an inexact Newton method [9] is well established in the context of solving nonlinear equations [13] and in nonlinear optimization [18]. Bellavia [5] applied an inexact interior point method to solve monotone nonlinear complementarity problems and proved global convergence and local superlinear convergence of the method. Armand et al. [3] studied the behaviour of a primal-dual inexact interior point method for solving nonlinear programming problems. Several successful attempts were also made to shed light on the application of an inexact Newton method in IPMs for linear and convex quadratic programming problems.

Freund, Jarre and Mizuno [10] and Mizuno and Jarre [17] extended a very popular globally convergent infeasible path-following method for linear programming of Kojima, Megiddo and Mizuno [14] to accommodate the inexact solution of Newton systems. In particular Mizuno and Jarre [17] proved that an inexact variant of this algorithm has $\mathcal{O}(n^2\log(1/\varepsilon))$ complexity. Baryamureeba and Steihaug [4] provided another extension of the method of Kojima et al., allowing for inexactness in both primal and dual Newton steps. All these analyses considered a general case in which no assumption was made about *how* the Newton system is solved. The only assumptions made were concerned with the absolute or relative error in the inexact Newton direction.

Several authors tried to specialize their analyses using a better understanding of the specific iterative methods of linear algebra employed to solve the Newton equation system inexactly. Indeed, if a particular preconditioner is used in a given Krylov subspace method applied in this context, it is often possible to control the residuals in specific linear subspaces and design special variants of an inexact interior point algorithm. Al-Jeiroudi and Gondzio [1] considered the case in which an indefinite preconditioner based on a guess of basic-nonbasic partition inspired by the work of Oliveira and Sorensen [19] is used to precondition the indefinite augmented system (the reduced KKT system). They designed an inexact infeasible primal-dual IPM for linear programming and established its $\mathcal{O}(n^2 \log(1/\varepsilon))$ complexity. Lu, Monteiro and O'Neal [15] analysed the case of quadratic programming in which the matrix of the quadratic objective term has a known factorization and proposed an interesting specialized preconditioner for the Newton system in this case. They showed that the resulting inexact IPM converges in $\mathcal{O}(n^2 \log(1/\varepsilon))$ iterations. Cafieri et al. [7] performed an analysis of an inexact potential reduction algorithm for convex quadratic programming.

In this paper we consider a *feasible* primal-dual path-following method for convex quadratic programming and analyse it in a situation when the solutions of the Newton systems admit a certain level of inaccuracy. We prove worst-case iteration complexity results for two variants of such a method. The first variant requires the iterates to stay in a small neighbourhood of the central path induced by the use of the Euclidean norm to control the error in the perturbed complementarity conditions. Such a method has the best known iteration complexity result $\mathcal{O}(\sqrt{n} \log(1/\varepsilon))$ and we prove that the inexact variant preserves the same complexity. However, this method is only of theoretical interest because its implementation demonstrates the behaviour predicted by the worst-case analysis. Therefore, this algorithm is not used in practice. The second variant allows the iterates to stay in a wide symmetric neighbourhood of the central path induced by the use of the infinity norm to control the error in the perturbed complementarity conditions. We show that this method reaches an ε-optimal solution in $\mathcal{O}(n \log(1/\varepsilon))$ iterations. Since the exact method operating in the wide neighbourhood is an attractive approach in practice [11] we analyze its inexact variant because it may provide the basis of an implementable method.

To simplify the analysis and to allow the reader to concentrate on the essential consequences of inexactness in the Newton direction we will analyse *feasible* algorithms. It is worth mentioning that this does not limit the applicability of methods analysed in this paper. We will provide an explanation how any quadratic program may be transformed to another one for which the initial feasible solution is known and therefore a *feasible* IPM may be applied to it. An alternative would be to use the homogeneous and self-dual emdedding [22] which also allows to transform the linear and quadratic programs into new models for which the primal and dual initial feasible solutions are known and therefore primal and dual feasibility can be maintained throughout the computations. The homogeneous and self-dual emdedding [22] was initially used in the context of LPs and implemented by Andersen and Andersen [2] in their Mosek software.

Finally, let us comment that the key motivation for this work is the need to better understand how much inaccuracy is admissible in the Newton systems and provide the foundations for cases in which interior point algorithms rely on iterative methods to solve the underlying linear algebra problems. There has recently been a shift of interest in the IPM community towards the application of iterative methods to solve the reduced KKT systems [8]. There exists a rich body of literature (cf. Benzi et al. [6]) which deals with very similar saddle point problem arising in the discretisations of partial differential equations and we expect that the coming years will bring many interesting developments in iterative methods applicable to IPMs. In particular

there is a clear need to develop alternative preconditioners which are compatible with the spirit of the matrix-free interior point method [11, 12].

The paper is organised as follows. In Section 2 we will introduce the quadratic optimization problem, define the notation used in the paper and point out an essential difference between the *exact* and *inexact* interior point methods. In Section 3 we will perform the worst-case analysis of two variants of an inexact feasible interior point algorithm for convex quadratic programming. First we will analyse the algorithm operating in a small neighbourhood of the central path induced by the 2-norm. Such a method yields the best complexity result known to date but it has only a theoretical importance. Next we will analyse the feasible algorithm operating in a symmetric neighbourhood of the central path induced by the infinity norm. This method has a practical meaning as an implementable algorithm. It provides a theoretical basis for the recently developed *matrix-free* variant of the interior point method [12]. Our analysis will follow that of Wright [21] and will generalize it from linear programming to quadratic programming and from exact to inexact method. In Section 4 we will comment on some practical aspects related to the implementation of inexact interior point method and finally in Section 5 we will give our conclusions.

2 Interior point methods: background

We are concerned in this paper with the theory of interior point methods for solving convex quadratic programming (QP) problems. We consider the following general primal-dual pair of QPs

$$
\begin{array}{ll}
\text{Primal} & \text{Dual} \\
\begin{array}{ll}
\min & c^T x + \frac{1}{2} x^T Q x \\
\text{s.t.} & A x = b, \\
& x \geq 0,
\end{array}
&
\begin{array}{ll}
\max & b^T y - \frac{1}{2} x^T Q x \\
\text{s.t.} & A^T y + s - Q x = c, \\
& x, y \text{ free}, \ s \geq 0,
\end{array}
\end{array}
\tag{1}
$$

where $A \in \mathcal{R}^{m \times n}$ has full row rank $m \leq n$, $Q \in \mathcal{R}^{n \times n}$ is a positive semidefinite matrix, $x, s, c \in \mathcal{R}^n$ and $y, b \in \mathcal{R}^m$. Setting $Q = 0$ yields the special case of the linear programming (LP) primal-dual pair.

To derive a primal-dual interior point method [21] we first introduce the logarithmic barrier function $\mu \sum_{j=1}^{n} \log x_j$ to "replace" the inequality constraint $x \geq 0$ in the primal and then, by using Lagrangian duality theory, write down the first-order optimality conditions

$$
\begin{array}{rcl}
A x & = & b, \\
A^T y + s - Q x & = & c, \\
X S e & = & \mu e, \\
(x, s) & \geq & 0,
\end{array}
\tag{2}
$$

where X and S are diagonal matrices in $\mathcal{R}^{n \times n}$ with elements of vectors x and s spread across the diagonal, respectively and $e \in \mathcal{R}^n$ is the vector of ones. This system of equations has a unique solution $(x(\mu), y(\mu), s(\mu))$, $x(\mu) > 0$, $s(\mu) > 0$ for any $\mu > 0$. The corresponding point is called a μ-centre. A family of these points for all positive values of μ determines a continuous curve $\{(x(\mu), y(\mu), s(\mu)) : \mu > 0\}$ which is called the primal-dual *central trajectory* or *central path*.

The first two equations in (2) are the primal and dual feasibility conditions, respectively. The third equation is a perturbed complementarity condition; the parameter μ associated with the logarithmic barrier function controls the level of perturbation. The convergence to optimality is forced by taking the barrier parameter μ to zero [21, 11].

In this paper we will assume that we work with the *feasible* interior point method and therefore all iterates (x, y, s) satisfy the first two equations in (2). Consequently, the duality gap is equal to the complementarity gap

$$c^T x + \frac{1}{2}x^T Q x - (b^T y - \frac{1}{2}x^T Q x) = x^T s = n\mu, \tag{3}$$

and by reducing the barrier parameter μ, IPM achieves convergence to optimality. The standard interior point algorithm applies the Newton method to (2), that is, computes the Newton direction and makes a step in this direction followed by a reduction of the barrier parameter μ. The reduction of the barrier term is enforced by the use of the parameter $\sigma \in (0, 1)$ and setting $\mu^{new} = \sigma\mu$. Due to the feasibility of (x, y, s) the residual in (2) takes the following form

$$(b - Ax, \ c - A^T y - s + Qx, \ \sigma\mu e - XSe) = (0, 0, \xi). \tag{4}$$

Hence the Newton direction $(\Delta x, \Delta y, \Delta s)$ is obtained by solving the following system of linear equations

$$\begin{bmatrix} A & 0 & 0 \\ -Q & A^T & I \\ S & 0 & X \end{bmatrix} \cdot \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \xi \end{bmatrix}, \tag{5}$$

where I denotes the identity matrix of dimension n.

In this paper we will analyse the method which allows the system (5) to be solved *inexactly*. To be precise, we will assume that all the iterates remain primal and dual feasible and an *inexact* Newton direction $(\Delta x, \Delta y, \Delta s)$ satisfies the following system of linear equations

$$\begin{bmatrix} A & 0 & 0 \\ -Q & A^T & I \\ S & 0 & X \end{bmatrix} \cdot \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \xi + r \end{bmatrix} \tag{6}$$

which admits an error r in the third equation. Let us observe that any step in such a primal-dual inexact Newton direction preserves primal and dual feasibility.

By using the positive semidefiniteness of Q and exploiting the first two equations in the appropriate Newton system, it is easy to demonstrate that for both the exact (5) and the inexact (6) Newton direction $(\Delta x, \Delta y, \Delta s)$, the following property holds:

$$\Delta x^T \Delta s = \Delta x^T Q\, \Delta x \geq 0. \tag{7}$$

The third equation in the Newton system plays a crucial role in the convergence analysis of an interior point algorithm. For the inexact Newton direction (6) this equation takes the following form

$$S\Delta x + X\Delta s = \xi + r = \sigma\mu e - XSe + r,$$

and by using $e^T e = n$ and $x^T s = n\mu$ we get

$$s^T \Delta x + x^T \Delta s = \sigma \mu e^T e - x^T s + e^T r = (\sigma - 1)x^T s + e^T r.$$

Hence the complementarity gap at the new point $(x(\alpha), y(\alpha), s(\alpha)) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s)$ becomes

$$
\begin{aligned}
x(\alpha)^T s(\alpha) &= (x + \alpha \Delta x)^T (s + \alpha \Delta s) \\
&= x^T s + \alpha(s^T \Delta x + x^T \Delta s) + \alpha^2 \Delta x^T \Delta s \\
&= (1 - \alpha(1 - \sigma))x^T s + \alpha e^T r + \alpha^2 \Delta x^T \Delta s
\end{aligned}
\tag{8}
$$

and the corresponding average complementarity gap is

$$\mu(\alpha) = x(\alpha)^T s(\alpha)/n = (1 - \alpha(1 - \sigma))\mu + \alpha e^T r/n + \alpha^2 \Delta x^T \Delta s/n. \tag{9}$$

Under the condition that the error term $e^T r$ and the second order term $\Delta x^T \Delta s$ are kept small enough in comparison with $\alpha(1 - \sigma)x^T s$, the complementarity gap at the new point is reduced compared with that at the previous iteration, thus guaranteeing progress of the algorithm.

The convergence analysis of an interior point algorithm relies on imposing uniform progress in reducing the error in the complementarity conditions which technically is translated into a requirement that the error is small and bounded with $\mathcal{O}(\mu)$. To achieve it we will restrict the iterates to remain in the neighbourhood of the central path $\{(x(\mu), y(\mu), s(\mu)), \mu > 0\}$ and we will control the barrier reduction parameter σ and the stepsize α so that $x(\alpha)^T s(\alpha)$ in (8) is noticeably smaller than $x^T s$.

We will consider two different ways of controlling the proximity to the central path and in both cases we will prove the convergence of the inexact interior point method and derive the worst-case complexity result. In both cases we will consider the *feasible* interior point algorithm and, hence, we will assume that all primal-dual iterates belong to the primal-dual strictly feasible set $\mathcal{F}^0 = \{(x, y, s) \,|\, Ax = b, \, A^T y + s - Qx = c, \, (x, s) > 0\}$. All iterates are confined to a neighbourhood of the central path which translates to a requirement that the error in the perturbed complementarity condition (2) is small. Depending on the norm used to measure this error, we will consider two different neighbourhoods:

- a small neighbourhood induced by the use of the Euclidean norm for some $\theta \in (0, 1)$

$$N_2(\theta) = \{(x, y, s) \in \mathcal{F}^0 \,|\, \|XSe - \mu e\| \leq \theta \mu\}, \tag{10}$$

  which yields a short-step algorithm, and

- a symmetric neighbourhood induced by the use of the infinity norm for some $\gamma \in (0, 1)$

$$N_S(\gamma) = \{(x, y, s) \in \mathcal{F}^0 \,|\, \gamma \mu \leq x_j s_j \leq \frac{1}{\gamma}\mu, \, \forall j\}, \tag{11}$$

  which yields a long-step algorithm.

The former has a theoretical importance as it leads to the algorithm with the best complexity result known to date. However, it is not an implementable method because it leads to poor performance in practice. Indeed, its behaviour reproduces the worst-case analysis. The latter neighbourhood has a practical meaning and leads to an efficient algorithm in practice.

It is worth adding that the symmetric neighbourhood (11) in which both too small and too large complementarity products are forbidden has advantages over a standard (one sided) infinity neighbourhood in which the complementarity products only have to satisfy $x_j s_j \geq \gamma\mu$. In the one sided infinity neighbourhood the largest complementarity product may reach the value of $(n - \gamma(n-1))\mu = \mathcal{O}(n\mu)$ while in the symmetric neighbourhood every product has to remain $\mathcal{O}(\mu)$. This has an important theoretical advantage and an even more important computational advantage because it guarantees that ξ in (5) is well scaled, namely, every entry of it is $\mathcal{O}(\mu)$. The symmetric neighbourhood is very handy when multiple centrality correctors are used (see [11] and the references therein).

Following the general theory of the inexact Newton Method [9, 13], we will assume that the residual r in (6) satisfies

$$\|r\|_p \leq \delta\|\xi\|_p, \tag{12}$$

for some $\delta \in (0,1)$ and an appropriate p-norm. According to the type of the neighbourhood used, (10) or (11), this inequality will use either $p = 2$ or $p = \infty$, respectively. Further in the paper we will omit a subscript 2 for the Euclidean norm unless an expression involves different norms at the same time and such an omission could lead to a confusion.

Let us observe that parameter δ measures the *relative* error in the third equation of (6) and it is independent of μ. Whatever the neighbourhood of the central path is used ((10) or (11)), for any well-centered iterate the right hand side term ξ in (5) and (6) is always $\mathcal{O}(\mu)$. Hence condition (12) implies that the residual r is also $\mathcal{O}(\mu)$. In fact condition (12) is a *standard* stopping criteria for an iterative method which allows the method to terminate when the error drops below a certain fraction δ of the right hand side term in the equation system (5). Consequently, the setting used in this paper is equivalent to an alternative one in which the residual r is directly linked to the barrier term μ and the convergence of the inexact IPM is driven by the barrier term going to zero. In this alternative setting, the condition $r \leq \bar{\delta}\mu$ would involve the forcing term $\bar{\delta}$ as it is usually the case in the analysis of inexact Newton method for nonlinear equations [9, 13]. Such a setting was applied by Bellavia [5] and Armand et al. [3] to analyse the inexact IPMs.

In the next section we will prove that the *feasible* interior point algorithm using an *inexact* Newton direction (6) and applied to a convex quadratic program converges to an ε-accurate solution in $\mathcal{O}(\sqrt{n}\ln(1/\varepsilon))$ or $\mathcal{O}(n\ln(1/\varepsilon))$ iterations if it operates in the $N_2(\theta)$ or $N_S(\gamma)$ neighbourhood, respectively. Our analysis will follow the general scheme used by Wright [21].

3 Worst-case complexity results

The analysis for two different neighbourhoods will share certain common features. The algorithm makes a step in the Newton direction obtained by solving (6). When a step in the Newton direction $(\Delta x, \Delta y, \Delta s)$ is made, the new complementarity product for component j is given by

$$
\begin{aligned}
x_j(\alpha)\,s_j(\alpha) &= (x_j + \alpha\Delta x_j)(s_j + \alpha\Delta s_j) \\
&= x_j s_j + \alpha(s_j\Delta x_j + x_j\Delta s_j) + \alpha^2\Delta x_j\Delta s_j.
\end{aligned} \tag{13}
$$

The third equation in (6) is a local linearization of the complementarity condition and controls the middle term $s_j\Delta x_j + x_j\Delta s_j = \xi_j + r_j$ in the above equation. The error in the approximation

of complementarity products is determined by the second-order term $\Delta x_j \Delta s_j$ in (13). We will provide a bound on these products, namely, we will bound the vector of the second-order error terms $\|\Delta X \Delta S e\|$.

Having multiplied the third equation in the Newton system (6) by $(XS)^{-1/2}$, we obtain

$$X^{-1/2} S^{1/2} \Delta x + X^{1/2} S^{-1/2} \Delta s = (XS)^{-1/2} (\xi + r). \tag{14}$$

Defining $u = X^{-1/2} S^{1/2} \Delta x$ and $v = X^{1/2} S^{-1/2} \Delta s$ and using (7) we obtain $u^T v = \Delta x^T \Delta s \geq 0$. Let us partition all products $u_j v_j$ into positive and negative ones: $\mathcal{P} = \{j \mid u_j v_j \geq 0\}$ and $\mathcal{M} = \{j \mid u_j v_j < 0\}$ and observe that

$$0 \leq u^T v = \sum_{j \in \mathcal{P}} u_j v_j + \sum_{j \in \mathcal{M}} u_j v_j = \sum_{j \in \mathcal{P}} |u_j v_j| - \sum_{j \in \mathcal{M}} |u_j v_j|. \tag{15}$$

Next, let us write equation (14) component-wise as $u_j + v_j = (x_j s_j)^{-1/2} (\xi_j + r_j)$ for every $j \in \{1, 2, \ldots, n\}$ and take the sum of squared equations for components $j \in \mathcal{P}$:

$$0 \leq \sum_{j \in \mathcal{P}} (u_j + v_j)^2 = \sum_{j \in \mathcal{P}} (u_j^2 + v_j^2) + 2 \sum_{j \in \mathcal{P}} u_j v_j = \sum_{j \in \mathcal{P}} (x_j s_j)^{-1} (\xi_j + r_j)^2,$$

to get

$$2 \sum_{j \in \mathcal{P}} |u_j v_j| = 2 \sum_{j \in \mathcal{P}} u_j v_j \leq \sum_{j \in \mathcal{P}} (x_j s_j)^{-1} (\xi_j + r_j)^2.$$

Inequality (15) implies $\sum_{j \in \mathcal{M}} |u_j v_j| \leq \sum_{j \in \mathcal{P}} |u_j v_j|$ and, hence, we can write

$$
\begin{aligned}
\|\Delta X \Delta S e\|_1 &= \sum_{j \in \mathcal{P}} |u_j v_j| + \sum_{j \in \mathcal{M}} |u_j v_j| \leq 2 \sum_{j \in \mathcal{P}} |u_j v_j| \\
&\leq \sum_{j \in \mathcal{P}} (x_j s_j)^{-1} (\xi_j + r_j)^2 \leq \sum_{j=1}^{n} (x_j s_j)^{-1} (\xi_j + r_j)^2.
\end{aligned}
\tag{16}
$$

We will now consider two different algorithms: the short-step method in which the iterates are confined to $N_2(\theta)$ neighbourhood (10) and the long-step method in which the iterates are confined to $N_S(\gamma)$ neighbourhood (11). The names "short-step" and "long-step" describe the steps to optimality and are related to the choice of barrier reduction parameter σ (and should not be confused with the stepsizes taken in the Newton direction). In the short-step method, σ is very close to one and therefore the algorithm makes only a short step to optimality while in the long step method, σ is usually a small number satisfying $\sigma \ll 1$.

3.1 Analysis of the short-step method

In this section we will assume that $(x, y, s) \in N_2(\theta)$ for some $\theta \in (0, 1)$ and inequality (12) holds for the 2-norm: $\|r\|_2 \leq \delta \|\xi\|_2$. It is easy to deduce that since $\|XSe - \mu e\|_\infty \leq \|XSe - \mu e\|_2 \leq \theta \mu$, the complementarity products satisfy the following inequality

$$(1 - \theta)\mu \leq x_j s_j \leq (1 + \theta)\mu \quad \forall j. \tag{17}$$

The barrier reduction parameter for the short-step algorithm is defined as

$$\sigma = 1 - \frac{\beta}{\sqrt{n}}, \tag{18}$$

for some $\beta \in (0, 1)$. Such a definition implies that

$$(1 - \sigma)^2 n = \beta^2 \tag{19}$$

and therefore, using $e^T(XSe - \mu e) = 0$, the norm of the term $\xi = \sigma\mu e - XSe$ in (4) satisfies

$$
\begin{aligned}
\|XSe - \sigma\mu e\|^2 &= \|(XSe - \mu e) + (1 - \sigma)\mu e\|^2 \\
&= \|XSe - \mu e\|^2 + 2(1-\sigma)\mu e^T(XSe - \mu e) + (1-\sigma)^2\mu^2 e^T e \\
&\leq \theta^2\mu^2 + (1 - \sigma)^2 n\mu^2 \\
&= (\theta^2 + \beta^2)\mu^2.
\end{aligned} \tag{20}
$$

We are now ready to derive a bound on $\|\Delta X \Delta Se\|$.

Lemma 3.1 *Let $\theta \in (0, 1)$. If $(x, y, s) \in N_2(\theta)$ then the inexact Newton direction $(\Delta x, \Delta y, \Delta s)$ obtained by solving (6) satisfies*

$$\|\Delta X \Delta Se\| \leq \frac{(1 + \delta)^2(\theta^2 + \beta^2)}{(1 - \theta)}\mu. \tag{21}$$

Proof: Inequality (17) provides a bound $\min_j\{x_j s_j\} \geq (1 - \theta)\mu$. We use it to rewrite (16):

$$\|\Delta X \Delta Se\|_2 \leq \|\Delta X \Delta Se\|_1 \leq \frac{1}{\min_j\{x_j s_j\}}\sum_{j=1}^{n}(\xi_j + r_j)^2 \leq \frac{1}{(1 - \theta)\mu}\|\xi + r\|^2. \tag{22}$$

Using (12) (with $p = 2$) and (20) we write

$$\|\xi + r\|^2 \leq (\|\xi\| + \|r\|)^2 \leq (1 + \delta)^2\|\xi\|^2 \leq (1 + \delta)^2(\theta^2 + \beta^2)\mu^2,$$

and after substituting this expression into (22) obtain the required inequality (21). $\qquad\square$

Next we will show that for appropriately chosen constants θ, β and δ, a full Newton step is feasible and the new iterate $(\bar{x}, \bar{y}, \bar{s}) = (x, y, s) + (\Delta x, \Delta y, \Delta s)$ also belongs to the $N_2(\theta)$ neighbourhood of the central path. We will prove an even stronger result which is that for any step $\alpha \in (0, 1]$ in the Newton direction the following point

$$(x(\alpha), y(\alpha), s(\alpha)) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s) \tag{23}$$

is primal-dual feasible and belongs to the $N_2(\theta)$ neighbourhood.

Using (13) and (9) and $s_j\Delta x_j + x_j\Delta s_j = \sigma\mu - x_j s_j + r_j$, the deviation of the jth complementarity product from the average becomes

$$
\begin{aligned}
x_j(\alpha)s_j(\alpha) - \mu(\alpha) &= (x_j + \alpha\Delta x_j)(s_j + \alpha\Delta s_j) - \mu(\alpha) \\
&= x_j s_j + \alpha(s_j\Delta x_j + x_j\Delta s_j) + \alpha^2\Delta x_j\Delta s_j - \mu(\alpha) \\
&= (1-\alpha)x_j s_j + \alpha\sigma\mu + \alpha r_j + \alpha^2\Delta x_j\Delta s_j - (1-\alpha)\mu - \alpha\sigma\mu - \alpha e^T r/n - \alpha^2\Delta x^T\Delta s/n \\
&= (1 - \alpha)(x_j s_j - \mu) + \alpha(r_j - e^T r/n) + \alpha^2(\Delta x_j\Delta s_j - \Delta x^T\Delta s/n).
\end{aligned}
$$

Consequently, the proximity measure for the point $(x(\alpha), y(\alpha), s(\alpha))$ becomes

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq (1-\alpha)\|XSe - \mu e\| + \alpha\|r - \frac{e^T r}{n} e\| + \alpha^2 \|\Delta X \Delta S e - \frac{\Delta x^T \Delta s}{n} e\|. \qquad (24)$$

The Lemma below sets three parameters: the proximity constant $\theta \in (0,1)$ in (10), $\beta \in (0,1)$ in the barrier reduction parameter $\sigma = 1 - \beta/\sqrt{n}$, and the level of error $\delta \in (0,1)$ allowed in the inexact Newton method (12).

Lemma 3.2 *Let (x, y, s) be the current iterate in $N_2(\theta)$ neighbourhood and $(\Delta x, \Delta y, \Delta s)$ be the inexact Newton direction which solves equation system (6). Let $\theta = \beta = 0.1$. If $\delta = 0.3$ then for any $\alpha \in (0,1]$ the new iterate (23) after a step α in this direction satisfies*

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq \theta\mu(\alpha). \qquad (25)$$

Proof: Expanding the square and using $e^T e = n$ we write

$$\|r - \frac{e^T r}{n} e\|^2 = \|r\|^2 + \frac{1}{n^2}(e^T r)^2 e^T e - \frac{2}{n}(e^T r)^2$$

$$= \|r\|^2 - \frac{1}{n}(e^T r)^2 \leq \|r\|^2. \qquad (26)$$

Similarly, expanding the square, using $e^T e = n$ again and $(\Delta X \Delta S e)^T e = \Delta x^T \Delta s$, we write

$$\|\Delta X \Delta S e - \frac{\Delta x^T \Delta s}{n} e\|^2 = \|\Delta X \Delta S e\|^2 + \frac{1}{n^2}(\Delta x^T \Delta s)^2 e^T e - \frac{2\Delta x^T \Delta s}{n}(\Delta X \Delta S e)^T e$$

$$= \|\Delta X \Delta S e\|^2 - \frac{1}{n}(\Delta x^T \Delta s)^2 \leq \|\Delta X \Delta S e\|^2. \qquad (27)$$

Next, using (24), the definition of $N_2(\theta)$ and inequalities (26), (12) and (27), we write

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq (1-\alpha)\|XSe - \mu e\| + \alpha\|r - \frac{e^T r}{n} e\| + \alpha^2 \|\Delta X \Delta S e - \frac{\Delta x^T \Delta s}{n} e\|$$

$$\leq (1-\alpha)\theta\mu + \alpha\delta\|\xi\| + \alpha^2 \|\Delta X \Delta S e\|.$$

Inequalities (20) and (21) (Lemma 3.1) provide bounds for the last two terms in the above inequality. We use them to write

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq (1-\alpha)\theta\mu + \alpha\delta\sqrt{\theta^2 + \beta^2}\mu + \alpha^2 \frac{(1+\delta)^2(\theta^2 + \beta^2)}{(1-\theta)}\mu.$$

The choice of $\theta = \beta = 0.1$ guarantees that $\sqrt{\theta^2 + \beta^2} \leq \sqrt{2}\theta$ and $\frac{(1+\delta)^2(\theta^2 + \beta^2)}{(1-\theta)} = \frac{2(1+\delta)^2}{9}\theta$ hence

$$\|X(\alpha)S(\alpha)e - \mu(\alpha)e\| \leq (1-\alpha)\theta\mu + \sqrt{2}\alpha\delta\theta\mu + \alpha^2 \frac{2(1+\delta)^2}{9}\theta\mu.$$

Using equality (9) we observe that this lemma will be proved (inequality (25) will be satisfied) if the following holds

$$(1-\alpha)\theta\mu + \sqrt{2}\alpha\delta\theta\mu + \alpha^2 \frac{2(1+\delta)^2}{9}\theta\mu \leq \theta\left((1 - \alpha(1-\sigma))\mu + \alpha\frac{e^T r}{n} + \alpha^2 \frac{\Delta x^T \Delta s}{n}\right).$$

We further simplify this inequality by removing the same terms present on both sides of it and then dividing it by $\alpha\theta$. Inequality (7) guarantees that $\Delta x^T \Delta s$ is nonnegative, hence we conclude that (25) will be satisfied if

$$\sqrt{2}\delta\mu + \alpha\frac{2(1+\delta)^2}{9}\mu \;\leq\; \sigma\mu + \frac{e^T r}{n}.$$

We observe that $|e^T r| \leq \|e\|\|r\| = \sqrt{n}\delta\|\xi\|$ and hence, using (20), we write

$$|\frac{e^T r}{n}| \;\leq\; \frac{\delta}{\sqrt{n}}\sqrt{\theta^2 + \beta^2}\mu \;\leq\; \frac{\sqrt{2}\delta\theta}{\sqrt{n}}\mu. \tag{28}$$

Therefore to guarantee that (25) holds for any $\alpha \in (0,1]$, it suffices to choose δ such that

$$\sqrt{2}\delta\mu + \frac{2(1+\delta)^2}{9}\mu \;\leq\; \sigma\mu - \frac{\sqrt{2}\delta\theta}{\sqrt{n}}\mu$$

and this simplifies to

$$\sqrt{2}\delta(1 + \frac{\theta}{\sqrt{n}}) + \frac{2(1+\delta)^2}{9} \;\leq\; \sigma = 1 - \frac{\beta}{\sqrt{n}}.$$

The left hand side of this inequality is an increasing function of δ and we can easily check that the choice $\delta = 0.3$ gives $0.3\sqrt{2}(1+\theta/\sqrt{n}) + 3.38/9 \leq 1 - \beta/\sqrt{n}$ which holds for $\theta = \beta = 0.1$ and any $n \geq 2$. $\qquad\square$

Lemma 3.2 guarantees that for any $\alpha \in (0,1]$ the new iterate (23) also belongs to the $N_2(\theta)$ neighbourhood of the central path. We will set $\alpha = 1$ and take the full step in the Newton direction. Observe that the inexact Newton direction (6) allows the error r to appear only in the third equation which means that the direction $(\Delta x, \Delta y, \Delta s)$ preserves the feasibility of primal and dual equality constraints. The new iterate is defined as $(\bar{x}, \bar{y}, \bar{s}) = (x, y, s) + (\Delta x, \Delta y, \Delta s)$ and setting $\alpha = 1$ in (9) gives

$$\bar{\mu} = \mu(\alpha) = \sigma\mu + \frac{e^T r}{n} + \frac{\Delta x^T \Delta s}{n}. \tag{29}$$

With $\theta = \beta = 0.1$ and $\delta = 0.3$ the right hand side term in inequality (21) may be simplified to give $\|\Delta X \Delta S e\| \leq 0.3756\beta\mu$ and, using the Cauchy-Schwartz inequality, we get the bound

$$\frac{\Delta x^T \Delta s}{n} = \frac{(\Delta X \Delta S e)^T e}{n} \leq \frac{1}{n}\|\Delta X \Delta S e\|\|e\| \leq \frac{0.3756}{\sqrt{n}}\beta\mu.$$

Using this inequality, our choice $\theta = \beta$, (28) and $\sigma = 1 - \beta/\sqrt{n}$, we obtain the following bound on $\bar{\mu}$ in (29)

$$\bar{\mu} \leq (1 - \frac{\beta}{\sqrt{n}})\mu + \frac{\sqrt{2}\delta\beta}{\sqrt{n}}\mu + \frac{0.3756\beta}{\sqrt{n}}\mu \;\leq\; (1 - \frac{\eta}{\sqrt{n}})\mu, \tag{30}$$

where $\beta(1 - \sqrt{2}\delta - 0.3756) \geq 0.02$, hence we may set $\eta = 0.02$ in (30).

We are now ready to state the complexity result for the inexact short-step feasible interior point method operating in a $N_2(0.1)$ neighbourhood.

Theorem 3.1 *Given $\epsilon > 0$, suppose that a feasible starting point $(x^0, y^0, s^0) \in N_2(0.1)$ satisfies $(x^0)^T s^0 = n\mu^0$, where $\mu^0 \leq 1/\epsilon^\kappa$, for some positive constant κ. Then there exists an index L with $L = \mathcal{O}(\sqrt{n}\ln(1/\epsilon))$ such that $\mu^l \leq \epsilon$, $\forall l \geq L$.*

Proof: is a straightforward application of Theorem 3.2 in Wright [21, Ch. 3]. □

3.2 Analysis of the long-step method

In this section we will assume that $(x, y, s) \in N_S(\gamma)$ for some $\gamma \in (0, 1)$ and inequality (12) holds for the infinity norm: $\|r\|_\infty \leq \delta\|\xi\|_\infty$. We will ask for an aggressive reduction of the duality gap from one iteration to another and set the barrier reduction parameter $\sigma \in (0, 1)$ to be significantly smaller than 1. Therefore it will not be possible in general to make the full step in the Newton direction.

Using definition (11) of the symmetric neighbourhood $N_S(\gamma)$ and observing that $1/\gamma - 1 > 1 - \gamma$ we derive the following bound for the term $\xi = \sigma\mu e - XSe$ in (4)

$$
\begin{aligned}
\|XSe - \sigma\mu e\|_\infty &= \|(XSe - \mu e) + (1 - \sigma)\mu e\|_\infty \\
&\leq \|XSe - \mu e\|_\infty + (1 - \sigma)\mu \\
&\leq \max\{1 - \gamma, \frac{1}{\gamma} - 1\}\mu + (1 - \sigma)\mu \\
&= (\frac{1}{\gamma} - \sigma)\mu.
\end{aligned}
\tag{31}
$$

We are now ready to derive a bound on $\|\Delta X \Delta S e\|_\infty$.

Lemma 3.3 *Let $\gamma \in (0, 1)$. If $(x, y, s) \in N_S(\gamma)$ then the inexact Newton direction $(\Delta x, \Delta y, \Delta s)$ obtained by solving (6) satisfies*

$$
\|\Delta X \Delta S e\|_\infty \leq \|\Delta X \Delta S e\|_1 \leq n\frac{(1 + \delta)^2}{\gamma}(\frac{1}{\gamma} - \sigma)^2\mu
\tag{32}
$$

and

$$
\Delta x_j \Delta s_j \leq \frac{(1 + \delta)^2}{\gamma}(\frac{1}{\gamma} - \sigma)^2\mu, \quad \forall j.
\tag{33}
$$

Proof: The definition of the $N_S(\gamma)$ neighbourhood provides a bound $\min_j\{x_j s_j\} \geq \gamma\mu$. We use it to rewrite (16):

$$
\|\Delta X \Delta S e\|_\infty \leq \|\Delta X \Delta S e\|_1 \leq \frac{1}{\min_j\{x_j s_j\}} \sum_{j=1}^{n}(\xi_j + r_j)^2 \leq \frac{1}{\gamma\mu}\|\xi + r\|_2^2.
\tag{34}
$$

Using (12) (for the infinity norm) and (31) we write

$$
\|\xi + r\|_2^2 \leq n\|\xi + r\|_\infty^2 \leq n(1 + \delta)^2\|\xi\|_\infty^2 \leq n(1 + \delta)^2(\frac{1}{\gamma} - \sigma)^2\mu^2,
$$

and, after substituting this expression into (34), obtain the required inequality (32). We observe that (32) implies that

$$-n\frac{(1+\delta)^2}{\gamma}(\frac{1}{\gamma}-\sigma)^2\mu \le \Delta x_j\Delta s_j \le n\frac{(1+\delta)^2}{\gamma}(\frac{1}{\gamma}-\sigma)^2\mu, \quad \forall j,$$

but we can obtain a tighter upper bound for this component-wise error term. For this we write equation (14) for component j and square both sides of it to get

$$2\Delta x_j\Delta s_j \le \frac{(\xi_j+r_j)^2}{x_js_j} \le \frac{(1+\delta)^2\|\xi\|_\infty^2}{\gamma\mu} \le \frac{(1+\delta)^2}{\gamma}(\frac{1}{\gamma}-\sigma)^2\mu, \quad \forall j,$$

which implies (33) and completes the proof. $\qquad\square$

Next we will show that for appropriately chosen constants σ, γ and δ, a (small) step $\alpha = \mathcal{O}(1/n)$ in the inexact Newton direction is feasible and the new iterate $(x(\alpha), y(\alpha), s(\alpha)) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s)$ remains in the $N_S(\gamma)$ neighbourhood of the central path.

The Lemma below provides conditions which have to be met by three parameters: the proximity constant $\gamma \in (0,1)$ in (11), the barrier reduction parameter $\sigma \in (0,1)$, and the level of error $\delta \in (0,1)$ allowed in the inexact Newton method (12).

Lemma 3.4 *Let (x, y, s) be the current iterate in the $N_S(\gamma)$ neighbourhood and $(\Delta x, \Delta y, \Delta s)$ be the inexact Newton direction which solves equation system (6). If the stepsize $\alpha \in (0,1]$ satisfies the following conditions:*

$$\alpha(\gamma+n)\frac{(1+\delta)^2}{\gamma}(\frac{1}{\gamma}-\sigma)^2 \le \sigma(1-\gamma)-\delta(1+\gamma)(\frac{1}{\gamma}-\sigma) \tag{35}$$

$$\delta(1+\frac{1}{\gamma})(\frac{1}{\gamma}-\sigma)+\alpha\frac{(1+\delta)^2}{\gamma}(\frac{1}{\gamma}-\sigma)^2 \le (\frac{1}{\gamma}-1)\sigma \tag{36}$$

then the new iterate $(x(\alpha), y(\alpha), s(\alpha))$ belongs to the $N_S(\gamma)$ neighbourhood, that is:

$$\gamma\mu(\alpha) \le x_j(\alpha)s_j(\alpha) \le \frac{1}{\gamma}\mu(\alpha), \quad \forall j. \tag{37}$$

Proof: Using the average complementarity gap (9) at the new iterate $(x(\alpha), y(\alpha), s(\alpha))$ together with expression (13) and the third equation in (6), we deduce that the left inequality in (37) will hold for any j if

$$\gamma\left((1-\alpha)\mu+\alpha\sigma\mu+\alpha\frac{e^Tr}{n}+\alpha^2\frac{\Delta x^T\Delta s}{n}\right) \le (1-\alpha)x_js_j+\alpha\sigma\mu+\alpha r_j+\alpha^2\Delta x_j\Delta s_j.$$

Since $(x, y, s) \in N_S(\gamma)$ we know that $\gamma(1-\alpha)\mu \le (1-\alpha)x_js_j$ and therefore the above inequality will hold if we satisfy a tighter version of it:

$$\gamma\left((1-\alpha)\mu+\alpha\sigma\mu+\alpha\frac{e^Tr}{n}+\alpha^2\frac{\Delta x^T\Delta s}{n}\right) \le \gamma(1-\alpha)\mu+\alpha\sigma\mu+\alpha r_j+\alpha^2\Delta x_j\Delta s_j.$$

After removing identical terms from both sides and dividing both sides by α, we conclude that the inequality will hold if

$$\alpha\left(\gamma\frac{\Delta x^T\Delta s}{n}-\Delta x_j\Delta s_j\right) \le \sigma(1-\gamma)\mu+r_j-\gamma\frac{e^Tr}{n}. \tag{38}$$

Using Lemma 3.3 we deduce that

$$\frac{\Delta x^T \Delta s}{n} = \frac{(\Delta X \Delta S e)^T e}{n} \leq \frac{1}{n} \|\Delta X \Delta S e\|_1 \|e\|_\infty \leq \frac{(1+\delta)^2}{\gamma} (\frac{1}{\gamma} - \sigma)^2 \mu \tag{39}$$

hence

$$\gamma \frac{\Delta x^T \Delta s}{n} - \Delta x_j \Delta s_j \leq (\gamma + n) \frac{(1+\delta)^2}{\gamma} (\frac{1}{\gamma} - \sigma)^2 \mu. \tag{40}$$

Using (12) (for the infinity norm) and (31) we deduce that $\|r\|_\infty \leq \delta \|\xi\|_\infty \leq \delta(\frac{1}{\gamma} - \sigma)\mu$ and

$$|e^T r| \leq \|e\|_1 \|r\|_\infty \leq n\delta(\frac{1}{\gamma} - \sigma)\mu \tag{41}$$

hence

$$|r_j - \gamma \frac{e^T r}{n}| \leq |r_j| + \gamma |\frac{e^T r}{n}| \leq \delta(1 + \gamma)(\frac{1}{\gamma} - \sigma)\mu. \tag{42}$$

The inequalities (40) and (42) allow us to determine the most adverse conditions in (38), namely when its left-hand-side is the largest possible and the right-hand-side is the smallest possible:

$$\alpha(\gamma + n) \frac{(1+\delta)^2}{\gamma} (\frac{1}{\gamma} - \sigma)^2 \mu \quad \leq \quad \sigma(1 - \gamma)\mu - \delta(1 + \gamma)(\frac{1}{\gamma} - \sigma)\mu.$$

We then conclude that (35) implies (38) and therefore the left inequality in (37) holds. This completes the first part of the proof.

The second part deals with the right inequality in (37). Again, using (9), (13) and the third equation in (6), we write the required inequality which should be satisfied for any $j \in \{1, 2, \ldots, n\}$

$$(1-\alpha)x_j s_j + \alpha\sigma\mu + \alpha r_j + \alpha^2 \Delta x_j \Delta s_j \leq \frac{1}{\gamma}\left((1-\alpha)\mu + \alpha\sigma\mu + \alpha\frac{e^T r}{n} + \alpha^2 \frac{\Delta x^T \Delta s}{n}\right),$$

and determine the condition under which it holds. We use similar arguments as in the earlier part of the proof: for example, $(x, y, s) \in N_S(\gamma)$ implies $(1-\alpha)x_j s_j \leq \frac{1}{\gamma}(1-\alpha)\mu$. Hence we simplify this inequality by removing identical terms and then divide both sides of it by α. The inequality we need to satisfy becomes

$$r_j - \frac{e^T r}{\gamma n} + \alpha \left(\Delta x_j \Delta s_j - \frac{\Delta x^T \Delta s}{\gamma n}\right) \leq (\frac{1}{\gamma} - 1)\sigma\mu. \tag{43}$$

Using (33) and (7) we write

$$\Delta x_j \Delta s_j - \frac{\Delta x^T \Delta s}{\gamma n} \leq \frac{(1+\delta)^2}{\gamma} (\frac{1}{\gamma} - \sigma)^2 \mu \tag{44}$$

and using (12) and (31) and similar arguments to those which led to (42) we write

$$|r_j - \frac{e^T r}{\gamma n}| \leq |r_j| + |\frac{e^T r}{\gamma n}| \leq \delta(1 + \frac{1}{\gamma})(\frac{1}{\gamma} - \sigma)\mu. \tag{45}$$

By (44) and (45), to satisfy (43) we need to choose α such that:

$$\delta(1 + \frac{1}{\gamma})(\frac{1}{\gamma} - \sigma)\mu + \alpha\frac{(1+\delta)^2}{\gamma}(\frac{1}{\gamma} - \sigma)^2 \mu \leq (\frac{1}{\gamma} - 1)\sigma\mu$$

which simplifies to (36) and completes the second part of the proof.   □

Lemma 3.4 provides conditions which the stepsize $\alpha \in (0, 1]$ needs to satisfy so that the new iterate remains in the $N_S(\gamma)$ neighbourhood of the central path. We still need to demonstrate that after a step is made a sufficient reduction of duality gap is achieved.

Lemma 3.5 *Let $(x, y, s) \in N_S(\gamma)$ be given and let $(\Delta x, \Delta y, \Delta s)$ be the inexact Newton direction which solves equation system (6). If the stepsize $\alpha \in (0, 1]$ satisfies the inequality*

$$\sigma + \delta(\frac{1}{\gamma} - \sigma) + \alpha \frac{(1 + \delta)^2}{\gamma}(\frac{1}{\gamma} - \sigma)^2 \leq 0.9, \tag{46}$$

then the duality gap at the new iterate $(x(\alpha), y(\alpha), s(\alpha))$ satisfies:

$$\mu(\alpha) \leq (1 - 0.1\alpha)\mu. \tag{47}$$

Proof: By substituting (9) into (47), cancelling similar terms and dividing the resulting inequality by α, we replace (47) with a new condition that the stepsize α has to satisfy:

$$\sigma\mu + \frac{e^T r}{n} + \alpha \frac{\Delta x^T \Delta s}{n} \leq 0.9\mu.$$

Using the bounds (41) and (39) derived earlier we conclude that the above inequality will hold if

$$\sigma\mu + \delta(\frac{1}{\gamma} - \sigma)\mu + \alpha \frac{(1 + \delta)^2}{\gamma}(\frac{1}{\gamma} - \sigma)^2\mu \leq 0.9\mu,$$

which is equivalent to (46).   □

It remains to consider the three conditions (35), (36) and (46) and to demonstrate that an appropriate choice of parameters γ, σ and δ guarantees that all these conditions hold for some $\alpha = \mathcal{O}(\frac{1}{n})$.

We set the proximity constant $\gamma = 0.5$ in (11), the barrier reduction parameter $\sigma = 0.5$ and $\delta = 0.05$ as the level of error allowed in the inexact Newton method (12). Indeed, with these parameter settings $\frac{(1+\delta)^2}{\gamma}(\frac{1}{\gamma} - \sigma)^2 = 4.96125$, and we verify that all three conditions (35), (36) and (46) are satisfied by $\hat{\alpha} = \frac{1}{50n}$ for any $n \geq 2$. Substituting such an $\hat{\alpha}$ into (47) gives

$$\bar{\mu} = \mu(\hat{\alpha}) \leq (1 - \frac{\eta}{n})\mu,$$

where $\eta = 0.002$, and allows us to conclude this section with the following complexity result for the long-step inexact feasible interior point method operating in a $N_S(0.5)$ neighbourhood.

Theorem 3.2 *Given $\epsilon > 0$, suppose that a feasible starting point $(x^0, y^0, s^0) \in N_S(0.5)$ satisfies $(x^0)^T s^0 = n\mu^0$, where $\mu^0 \leq 1/\epsilon^\kappa$, for some positive constant κ. Then there exists an index L with $L = \mathcal{O}(n \ln(1/\epsilon))$ such that $\mu^l \leq \epsilon$, $\forall l \geq L$.*

Proof: is a straightforward application of Theorem 3.2 in Wright [21, Ch. 3].   □

4   Practical aspects of the inexact IPM

Linear system (6) is a modification of (5) which admits the error r in its third equation. The analysis in the previous section provides guarantees that when the Newton equation system (6) is solved inexactly and error r satisfies condition (12) the interior point algorithms retain their good complexity results. In this section we will comment on practical ways of computing the Newton direction which meets condition (12).

Let us observe that after eliminating Δs, (5) is reduced to the augmented form

$$\begin{bmatrix} -Q - \Theta^{-1} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} -X^{-1}\xi \\ 0 \end{bmatrix}, \tag{48}$$

where $\Theta = XS^{-1} \in \mathcal{R}^{n \times n}$. Any IPM has to solve at least one such system at each iteration [8, 11]. Numerous attempts have been made during the last decade to employ an iterative method for this task. Iterative methods for linear algebra are particularly attractive if they can be used to find only an approximate solution of the linear system, that is, when their run can be truncated to merely a few iterations. It is common to interrupt the iterative process once the required (loose) accuracy of the solution is obtained. Clearly such a solution is inexact and in the context of (48) this translates to dealing with an inexact Newton direction $(\Delta \tilde{x}, \Delta \tilde{y})$ which satisfies

$$\underbrace{\begin{bmatrix} -Q - \Theta^{-1} & A^T \\ A & 0 \end{bmatrix}}_{\mathcal{K}} \begin{bmatrix} \Delta \tilde{x} \\ \Delta \tilde{y} \end{bmatrix} = \begin{bmatrix} -X^{-1}\xi + r_x \\ r_y \end{bmatrix}, \tag{49}$$

where the errors $r_x \in \mathcal{R}^n$ and $r_y \in \mathcal{R}^m$ determine the level of inexactness.

The analysis presented in this paper applies to the situation when $r_y = 0$. The other error, r_x may take a nonzero value and indeed, (49) becomes equivalent to (6) if $r_y = 0$ and

$$-X^{-1}\xi + r_x = -X^{-1}(\xi + r). \tag{50}$$

This equation combined with condition (12) determines the practical stopping criteria set for an iterative solution method applied to (48):

$$r_y = 0 \quad \text{and} \quad \|r\| = \|X r_x\| \leq \delta \|\xi\|. \tag{51}$$

Let us observe that the condition $r_y = 0$ imposes a certain structure in the solution of (49) and therefore both the IPM and the iterative method have to satisfy the following two conditions, respectively:

C1 A feasible IPM has to be used to provide the correct structure of the right hand side in (5) which will then yield (48).

C2 An iterative method applied to solve equation (48) has to guarantee that the residual r_y in (49) is zero.

Arguments that both these conditions may be satisfied are provided in the discussion below.

In the general case of an arbitrary quadratic optimization problem, we may embed the problem into another one for which the initial primal and dual strictly feasible solutions can be constructed. We replace the primal problem in the primal-dual pair (1) with the following new problem:

$$\begin{aligned}
\min \quad & c^T x + M e_m^T x_+ + M e_m^T x_- + M \zeta_+ + M \zeta_- + \tfrac{1}{2} x^T Q x \\
\text{s.t.} \quad & A x + x_+ - x_- = b, && (52) \\
& e_n^T x + \zeta_+ - \zeta_- = M && (53) \\
& x, x_+, x_-, \zeta_+, \zeta_- \geq 0,
\end{aligned}$$

where M is a large enough number (the "big M"), $x_+, x_- \in \mathcal{R}^m$, $\zeta_+, \zeta_- \in \mathcal{R}$, and $e_m \in \mathcal{R}^m$, $e_n \in \mathcal{R}^n$ denote the vectors of ones. For this problem we define the following primal feasible solution: $x = e_n$, $x_+ = \max\{b - Ae_n, 0\} + e_m$, $x_- = \max\{-(b - Ae_n), 0\} + e_m$, $\zeta_+ = M + 1$ and $\zeta_- = e_n^T e_n + 1 = n + 1$. Let y and y_0 be the Lagrange multipliers associated with equality constraints (52) and (53), respectively and let $s_+, s_-, s_{\zeta_+}, s_{\zeta_-}$ be the Langrange multipliers associated with the nonnegativity constraints $x_+, x_-, \zeta_+, \zeta_- \geq 0$, respectively. The dual feasibility constraints for this new quadratic problem have the following form

$$\begin{aligned}
-Qx + A^T y + e_n y_0 + s &= c \\
y + s_+ &= M e_m \\
y_0 + s_{\zeta_+} &= M && (54) \\
-y + s_- &= M e_m \\
-y_0 + s_{\zeta_-} &= M.
\end{aligned}$$

We use $x = e_n$ as defined before and set $y = 0$. Then we set y_0 to a sufficiently large negative number to guarantee that $s = c + Qe_n - e_n y_0$ is strictly positive. To satisfy it, we need $-y_0 > \|c + Qe_n\|_\infty$ and additionally we request $-y_0 < M$. Finally, we define $s_+ = s_- = M e_m$ and $s_{\zeta_+} = M - y_0$ and $s_{\zeta_-} = M + y_0$, which completes the construction of an initial dual strictly feasible solution.

In conclusion, we can embed the original quadratic optimization problem (1) into a new one for which the initial strictly feasible primal-dual solution may be defined and therefore the feasible interior point algorithm can be applied to it; condition C1 is satisfied.

The requirement $r_y = 0$ in (49) implies that $\tilde{\Delta}x$ lies in the null space of A. It can be met for example if equation (49) is solved by an appropriate Krylov-subspace method with an appropriate preconditioner. Since (49) involves the symmetric indefinite matrix $\mathcal{K}$, two possible Krylov-subspace algorithms which can be used are the minimum residual method (MINRES) for symmetric and possibly indefinite matrices and the generalized minimum residual method (GMRES) for symmetric matrices. Although in exact arithmetic these methods should converge in $n + m$ iterations, in practice one needs to use preconditioning to improve the spectral properties of the matrix, ideally achieving a good clustering of the eigenvalues of the preconditioned matrix [13]. The suitable preconditioners include nonsingular matrices of the form

$$\mathcal{P} = \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix}, \tag{55}$$

where $G \in \mathcal{R}^{n \times n}$ is a nonsingular approximation of the (1,1) block in $\mathcal{K}$. Such preconditioners have been used for example by Lukšan and Vlček [16] any many followers. They are sometimes

called *constraint preconditioners* because they respect the constraints of the optimization problem. Observe that MINRES or GMRES applied to (49), using preconditioner $\mathcal{P}$ of (55) and starting from an initial solution such that $\tilde{\Delta}x = 0$ will initialize r_y to zero and will maintain $r_y = 0$ throughout the whole iterative process.

Since $\mathcal{K}$ is indefinite, the standard conjugate gradient algorithm [13] cannot be applied to (49). However, under certain conditions, such as the use of an appropriate preconditioner which implicitly represents the null space of constraints [16], the preconditioned conjugate gradients can be applied to this system. In a very interesting analysis, Rozložník and Simoncini [20] provided conditions under which the preconditioned conjugate gradients may be applied to an indefinite system. These require the preconditioned matrix $\mathcal{K}\mathcal{P}^{-1}$ to be J-symmetric with $J = \mathcal{P}^{-1}$ and the residual in conjugate gradients to maintain a zero block in the second equation of (49). Such conditions are met for example when (49) is preconditioned with (55) and CG starts from an initial solution $\tilde{\Delta}x = 0$.

In conclusion, there exist Krylov-subspace methods which, with an appropriate preconditioner and an appropriate initial solution, can deliver a solution to (49) such that $r_y = 0$; condition C2 is therefore satisfied.

We summarize this section with an observation which follows from two recent surveys [8, 11] on interior point methods. There is a growing interest in the development of interior point methods which use iterative solvers to compute Newton directions and therefore need new theory and a better understanding of situations when inexact Newton directions are employed. Several approaches using iterative solvers and suitable preconditioners have already been developed and some of them have been implemented, see [8, 11] and the references therein. The author's own experience in this area has been reported for example in [1, 11, 12]. The implemented approaches often differ in details from those which follow the rigorous convergence and worst-case complexity analyses. In this paper we present the theory of two *implementable* variants of the inexact feasible interior point method. Such developments reduce the gap between the theory and the computational practice of IPMs.

5 Conclusions

The analysis presented in this paper provides the proofs of $\mathcal{O}(\sqrt{n}\log(1/\varepsilon))$ and $\mathcal{O}(n\log(1/\varepsilon))$ iteration complexity of the, respectively, short-step and long-step inexact feasible primal-dual algorithms for quadratic programming. The analysis allows for considerable relative errors in the Newton direction. Indeed, δ in (12) may take values 0.3 and 0.05 for the short-step and long-step algorithms, respectively. This shows, somewhat surprisingly, that the inexactness in the solution of the Newton equation system (6) may be quite considerable without adversely affecting the best known worst-case iteration complexity results of these algorithms. It is an encouraging result for researchers who design preconditioners for iterative methods and wish to apply them to solve the reduced Newton equation systems arising in the context of interior point methods.

whose comments helped to improve the paper.

References

[1] G. Al-Jeiroudi and J. Gondzio, *Convergence analysis of inexact infeasible interior point method for linear optimization*, Journal of Optimization Theory and Applications, 141 (2009), pp. 231–247.

[2] E. D. Andersen and K. D. Andersen, *The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm*, in High Performance Optimization, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., Kluwer Academic Publishers, 2000, pp. 197–232.

[3] P. Armand, J. Benoist, and J.-P. Dussault, *Local path-following property of inexact interior methods in nonlinear programming*, Computational Optimization and Applications, 52 (2012), pp. 209–238.

[4] V. Baryamureeba and T. Steihaug, *On the convergence of an inexact primal-dual interior point method for linear programming*, in LSSC 2005, I. Lirkov, S. Margenov, and J. Wasniewski, eds., vol. 3743 of Lecture Notes in Computer Science, Springer-Verlag, Berlin, 2006, pp. 629–637.

[5] S. Bellavia, *Inexact interior point method*, Journal of Optimization Theory and Applications, 96 (1998), pp. 109–121.

[6] M. Benzi, G.H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1–137.

[7] S. Cafieri, M. D'Apuzzo, V. De Simone, D. di Serafino, and G. Toraldo, *Convergence analysis of an inexact potential reduction method for convex quadratic programming*, Journal of Optimization Theory and Applications, 135 (2007), pp. 355–366.

[8] M. D'Apuzzo, V. De Simone, and D. di Serafino, *On mutual impact of numerical linear algebra and large-scale optimization with focus on interior point methods*, Computational Optimization and Applications, 45 (2010), pp. 283–310.

[9] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM Journal on Numerical Analysis, 19 (1982), pp. 400–408.

[10] 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 (1999), pp. 50–71.

[11] J. Gondzio, *Interior point methods 25 years later*, European Journal of Operational Research, 218 (2012), pp. 587–601.

[12] ——, *Matrix-free interior point method*, Computational Optimization and Applications, 51 (2012), pp. 457–480.

[13] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, vol. 16 of Frontiers in Applied Mathematics, SIAM, Philadelphia, 1995.

[14] M. Kojima, N. Megiddo, and S. Mizuno, *A primal-dual infeasible-interior-point algorithm for linear programming*, Mathematical Programming, 61 (1993), pp. 263–280.

[15] Z. Lu, R. D. S Monteiro, and J. W. O'Neal, *An iterative solver-based infeasible primal-dual path-following algorithm for convex quadratic programming*, SIAM Journal on Optimization, 17 (2006), pp. 287–310.

[16] L. Lukšan and J. Vlček, *Indefinitely preconditioned inexact Newton method for large sparse equality constrained nonlinear programming problems*, Numerical Linear Algebra with Applications, 5 (1998), pp. 219–247.

[17] S. Mizuno and F. Jarre, *Global and polynomial-time convergence of an infeasible-interior point algorithm using inexact computation*, Mathematical Programming, 84 (1999), pp. 105–122.

[18] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag Telos, Berlin, 1999.

[19] A. R. L. Oliveira and D. C. Sorensen, *A new class of preconditioners for large-scale linear systems from interior point methods for linear programming*, Linear Algebra and its Applications, 394 (2005), pp. 1–24.

[20] M. Rozloznik and V. Simoncini, *Krylov subspace methods for saddle point problems with indefinite preconditioning*, SIAM Journal of Matrix Analysis and Applications, 24 (2002), pp. 368–391.

[21] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.

[22] Y. Ye, M.J. Todd, and S. Mizuno, *An $o(\sqrt{n}L)$ - iteration homogeneous and self-dual linear programming algorithm*, Mathematics of Operations Research, 19 (1994), pp. 53–67.