A primal-dual interior point algorithm with an exact and differentiable merit function for general nonlinear programming problems*

Ioannis Akrotirianakis[1] and Berç Rustem[2]

Abstract

A primal-dual interior point algorithm for solving general nonlinear programming problems is presented. The algorithm solves the perturbed optimality conditions by applying a quasi-Newton method, where the Hessian of the Lagrangian is replaced by a positive definite approximation. An approximation of Fletcher's exact and differentiable merit function together with line-search procedures are incorporated into the algorithm. The line-search procedures are used to modify the length of the step so that the value of the merit function is always reduced. Different step-sizes are used for the primal and dual variables. The search directions are ensured to be descent for the merit function, which is thus used to guide the algorithm to an optimum solution of the constrained optimisation problem. The monotonic decrease of the merit function at each iteration, ensures the global convergence of the algorithm. Finally, preliminary numerical results demonstrate the efficient performance of the algorithm for a variety of problems.

Key Words: Non-linear programming, constrained optimisation, primal-dual interior point methods, global convergence, non-convex programming.

February, 1999,
Technical Report 98-09,
Department of Computing,
Imperial College of Science, Technology and Medicine,
180 Queen's Gate, London SW7 2BZ, U.K.

1. e-mail: ia4@doc.ic.ac.uk
2. e-mail: br@doc.ic.ac.uk

*Constructive discussions with Professor H. Yabe are gratefully acknowledged by the authors.

1 Introduction

Interior point methods are a successful and efficient class of techniques for solving large-scale Linear Programming (LP) problems. After the announcement of the first interior point method with polynomial complexity by Karmarkar [20], there has been a wide interest in the application of these methods in LP. Among different interior point approaches, primal-dual algorithms have attracted most of the interest. Computational experiments (eg, [24], [3]) and theoretical developments (eg, [5], [29]) have shown that they perform much better than other interior point algorithms and outperform the simplex method in many large-scale LP problems.

The computational success of primal-dual interior point methods in linear programming has motivated substantial interest in their application in Nonlinear Programming (NLP). Most of the effort has been focused on convex Quadratic Programming (QP) (see for example [13], [23]) and convex NLP problems (see for example, [17], [16], [14]), demonstrating that primal-dual interior point methods can solve those problems efficiently. However, only recently general (non-convex) NLP problems have been the subject of research in this area. El-Bakry *et al.* [27], McCormick and Falk [8], and Yamashita [9] have developed globally convergent primal-dual algorithms for that class of problems. Also Lasdon *et al.* [12] have considered various primal-dual formulations of those problems and reported their computational experience.

In this paper, we discuss a primal-dual interior point algorithm for general (non-convex) NLP problems. We are interested in problems with both equality and inequality constraints, and for notational simplicity we consider the NLP problem

$$\begin{array}{ll} min & f(x) \\ ST & g(x) = 0, \quad x \geq 0 \end{array} \tag{1}$$

where $x = (x^1, ..., x^n)^T$, and $f : \Re^n \to \Re$, $g : \Re^n \to \Re^q$ are given smooth functions. The Lagrangian function of problem (1) is

$$L(x, y, z) = f(x) - y^T g(x) - z^T x \tag{2}$$

where $y \in \Re^q$ and $z \in \Re^n$ are the Lagrange multiplier vectors of the equality and inequality constraints, respectively. The first order necessary conditions for a local minimiser x_* of problem (1) are

$$\begin{array}{rcl} \nabla f(x_*) - \nabla g(x_*)^T y_* - z_* & = & 0, \\ g(x_*) & = & 0, \\ X_* Z_* e & = & 0, \quad x_*, z_* \geq 0 \end{array} \tag{3}$$

where $X_* = diag\{x_*^1, ..., x_*^n\}$, $Z_* = diag\{z_*^1, ..., z_*^n\}$, and ∇g denotes the Jacobian matrix of the equality constraints.

Interior point methods solve problem (1) by solving the parametrised equality constrained problems

$$\begin{array}{ll} min & B(x;\mu) = f(x) - \mu \sum_{i=1}^{n} \log(x^i) \\ ST & g(x) = 0 \end{array} \tag{4}$$

for a decreasing sequence of positive barrier parameters μ, converging to zero. The objective function of the parametrised problems (4) is the classical logarithmic barrier function, first introduced by Frisch [6]. Since the logarithm is not defined for non-positive values, the objective function of (4) is defined only in the interior of the feasible region. Fiacco and McCormick [1] have shown under certain assumptions that, if $x(\mu)$ is the exact solution of problem (4) with μ fixed, then the sequence $\{x(\mu)\}$ generated as $\mu \to 0$, converges to an optimum solution x_* of the initial problem (1). The solution of problem (4) is determined by solving its first order optimality conditions. The solution of these conditions is found by using Newton's method, which is efficient for convex programming.

Our approach for solving problem (1) is to use the primal-dual interior point framework [15] to handle the inequality constraints and a Sequential Quadratic Programming (SQP) framework [4] to handle the equality constraints. The motivation of our work is based on the observation that the solution of the first order optimality conditions of any NLP problem, which is the core of interior point algorithms, is not sufficient to guarantee the convergence to an optimum solution, unless the problem is convex. In other words, the algorithm applied, for example, on a minimisation problem, may converge to a local maximum or even worse to a saddle point, since the first order optimality conditions are also satisfied at those points.

To avoid such situations, a merit function is incorporated within the primal-dual interior point algorithm. The adopted merit function has the property that its unconstrained minimisers are solutions of the initial problem (1). The purpose of the merit function is to guide the iterates of the algorithm to a minimiser of the initial problem. This is achieved by ensuring that the merit function is decreased sufficiently at each iteration of the algorithm.

Merit functions have been used extensively in SQP algorithms to achieve global convergence. Recently, some merit functions have been used in interior point methods. One example is the merit function which uses the logarithmic barrier function for the inequality constraints and the classical quadratic penalty function for the equality constraints. That merit function was proposed independently by Vanderbei and Shanno [25], and Akrotirianakis and Rustem [11] and derives mainly from the merit function proposed and studied by Rustem [2] in the context of SQP methods. Other merit functions have been proposed and analysed by Yamashita [9], Gay *et al.* [7], and Gajulapalli and Lasdon [26].

This paper is organised as follows. In section 2 we present the basic framework of primal-dual interior point methods for NLP problems. In section 3 we analyse the adopted merit function and line search rules. In Section 4 we discuss the global convergence of the algorithm. Finally, in Section 5 we report our computational experience.

2 Primal-Dual Interior Point Methods

Primal-dual interior point algorithms find the solution of the initial problem (1), by solving barrier problems (4) for a sequence of strictly positive barrier parameters. They consist of two types of iterations: inner and outer. The inner iterations are associated with the solution of the barrier problem (4) for a fixed value of the barrier parameter μ, whereas the outer iterations are associated with the reduction of μ.

For μ fixed, the Lagrangian function of problem (4) is

$$L_B(x, y; \mu) = f(x) - \mu \sum_{i=1}^{n} \log(x^i) - y^T g(x) \tag{5}$$

and its first order optimality conditions are given by the system of equations

$$\begin{aligned} \nabla f(x) - \nabla g(x)^T y - \mu X^{-1} e &= 0 \\ g(x) &= 0. \end{aligned} \tag{6}$$

where $X = diag\{x^{(1)}, ..., x^{(n)}\}$ and $e = (1, ..., , 1)^T$.

Introducing the non-linear transformation $z = \mu X^{-1} e$, the above system of equations can be written as

$$\begin{aligned} \nabla f(x) - \nabla g(x)^T y - z &= 0 \\ g(x) &= 0 \\ XZe - \mu e &= 0, \end{aligned} \tag{7}$$

where $Z = diag\{z^{(1)}, ..., z^{(n)}\}$ and $x, z > 0$. The introduction of this transformation is essential to the numerical success of the method since system (7) is more stable than system (6).

Also, equations (7) differ from the optimality conditions (3) of the initial problem only in the third equation, known as the complementarity condition. Due to this difference, system (7) is referred to as the perturbed optimality conditions of the initial problem. For positive values of μ, the solutions $(x(\mu), y(\mu), z(\mu))$ of the perturbed optimality conditions, lie in the interior of the feasible region, and form the so-called central path [15]. As $\mu \rightarrow 0$, equations (7) approximate (3) with increasing accuracy. Consequently, the central points $(x(\mu), y(\mu), z(\mu))$ converge to a point satisfying the optimality conditions (3) of the initial problem.

Primal-dual algorithms use Newton's method to solve the perturbed optimality conditions (7). At the k-th iteration and for μ fixed, the first order change of system (6) yields the linear system of equations

$$\begin{bmatrix} H_k & -\nabla g(x_k)^T & -I \\ \nabla g(x_k) & 0 & 0 \\ Z_k & 0 & X_k \end{bmatrix} \begin{pmatrix} \Delta x_k \\ \Delta y_k \\ \Delta z_k \end{pmatrix} = - \begin{pmatrix} \nabla_x L(x_k, y_k, z_k) \\ g(x_k) \\ X_k Z_k e - \mu e \end{pmatrix} \tag{8}$$

where H_k is the Hessian matrix of the Lagrangian function (2), or an approximation to it. If we define $w_k = (x_k, y_k, z_k)$ and $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k)$, system (8) can be written in a more convenient and concise form as

$$J(w_k)\Delta w_k = -F(w_k; \mu), \tag{9}$$

where $F(w_k; \mu)$ is the vector containing the perturbed optimality conditions, and $J(w_k)$ is the corresponding Jacobian matrix.

If we solve the third equation of the Newton system (8) for Δz_k we obtain

$$\Delta z_k = -X_k^{-1} Z_k \Delta x_k - z_k + \mu X_k^{-1} e. \tag{10}$$

Substituting (10) into the first equation of (8) we obtain the reduced system of optimality conditions

$$\begin{bmatrix} H_k + X_k^{-1} Z_k & -\nabla g(x_k)^T \\ \nabla g(x_k) & 0 \end{bmatrix} \begin{pmatrix} \Delta x_k \\ \Delta y_k \end{pmatrix} = - \begin{pmatrix} \nabla_x L_B(x_k, y_k; \mu) \\ g(x_k) \end{pmatrix} \tag{11}$$

The Newton direction Δw_k is then used to find the next iterate

$$w_{k+1} = w_k + A_k \Delta w_k,$$

where $A_k = diag\{\alpha_{x_k} I_n, \alpha_{y_k} I_q, \alpha_{z_k} I_n\}$ and I_n, I_q are the n-th and q-th order identity matrices respectively. The step-lengths α_{x_k}, α_{y_k}, and α_{z_k} are in the interval $(0, 1]$ and may all be equal to or different from each other. Furthermore, $\alpha_{x_k} \leq \alpha_{x_k}^{max}$ and $\alpha_{z_k} \leq \alpha_{z_k}^{max}$, where

$$\alpha_{x_k}^{max} = \gamma \max_{1 \leq j \leq n} \{-\frac{x_k^{(j)}}{\Delta x_k^{(j)}} : \Delta x_k^{(j)} < 0\} \tag{12}$$

and

$$\alpha_{z_k}^{max} = \gamma \max_{1 \leq j \leq n} \{-\frac{z_k^{(j)}}{\Delta z_k^{(j)}} : \Delta z_k^{(j)} < 0\} \tag{13}$$

represent the maximum allowable step sizes, which guarantee that the iterates x_k and z_k

always remain strictly feasible, for some $\gamma \in (0, 1)$.

The distance of the current point from the central path is measured by the Euclidean norm of the perturbed optimality conditions, i.e., $\| F(w_k; \mu) \|$. Once this measure is less than a certain threshold value, the barrier parameter is reduced and the whole process is repeated until the barrier parameter becomes zero.

3 Merit function and step-size rules

The aim of the merit function is to provide a measure of progress towards an optimum solution of the barrier problems (4) as well as the initial problem (1). This is achieved by ensuring that it decreases at each iteration of the algorithm. A procedure for adjusting the step lengths of the variables is used in order to guarantee that the merit function decreases at each iteration. Interior point methods use the logarithmic barrier function to eliminate the inequality constraints of the initial problems. However, they do not provide any means to eliminate the equality constraints, which are carried over without any transformation to the barrier problems. Therefore the merit function is a combination of the barrier function and the equality constraints.

Throughout the paper the following assumptions hold.

Assumptions:

A1: The second order derivatives of the objective function f and the constraints g are continuous.

A2: The columns of the matrix $[\nabla g(x), e_j : j \in I_x^0]$ are linearly independent, where $I_x^0 = \{j : \liminf_{k\to\infty} x_k^{(j)} = 0,\ j = 1, 2, ..., n\}$ and e_j represents the j-th column of the $n \times n$ identity matrix. Also the sequence $\{x_k\}$ is bounded.

A3: Strict complementarity of the solution $w_* = (x_*, y_*, z_*)$ is satisfied, that is if $z_*^i > 0$ then $x_*^i = 0$, for $i = 1, 2, ..., n$ and vice versa.

A4: The second order sufficiency condition for optimality is satisfied at the solution point, i.e., if for all vectors $0 \neq v \in \Re^n$ such that $\nabla g^{(j)}(x_*)^T v = 0$, $j = 1, 2, ..., q$, and $e_j^T v = 0$, for $j \in I_x^0$, then $v^T \nabla_{xx}^2 L(x, y, z) v > 0$. Also, the approximation matrix H_k is such that

$$\frac{1}{M_1} \| v \|^2 \leq v^T H_k v \leq M_1 \| v \|^2 \tag{14}$$

where M_1 is a positive constant and the matrix $H_k + X_k^{-1} Z_k$ is non-singular.

Furthermore we use the notation $f_k = f(x_k)$, $g_k = g(x_k)$, to denote the values of the objective and constraint functions at the k-th iteration. The Euclidean norm is denoted by $\| . \|$.

3.1 Description of the merit function

The merit function used by our algorithm is based on the exact and differentiable penalty function developed by Fletcher [22]. For the barrier problem (4) it has the form

$$\phi(x) = B(x;\mu) - g(x)^T \hat{y}(x) + \frac{1}{2} c \parallel g(x) \parallel^2, \tag{15}$$

where $c \geq 0$ is the penalty parameter and

$$\hat{y}(x) = \left[\nabla g(x)^T \nabla g(x) \right]^{-1} \nabla g(x)^T \nabla B(x;\mu). \tag{16}$$

Note that the multipliers $\hat{y}(x)$ in the merit function $\phi(x)$ are the least squares estimates of the optimal Lagrange multipliers of the barrier problem (4). Furthermore, since they are continuous functions of x, they provide increasingly more accurate approximations of the optimal Lagrange multipliers as the algorithm proceeds.

The function $\phi(x)$ has the essential property of a merit function [21], that is, for μ fixed, $x(\mu)$ is a local minimum of ϕ if and only if it is a local minimum of the barrier problem (4). However, its major disadvantage is that its gradient depends on the second derivatives of the objective and constraint functions, due to the function $\hat{y}(x)$. Hence if ϕ is to be used in a line search procedure of the form

$$\phi(x_k + \alpha_{x_k} \Delta x_k) \leq \phi(x_k) + \rho \alpha_{x_k} \nabla \phi(x_k)^T \Delta x_k, \tag{17}$$

where $\rho \in (0,1)$, the computational effort required for the calculation of $\nabla \phi(x_k)$ may damage the performance of the algorithm. To overcome this difficulty, we use an approximation of the merit function ϕ defined by Powell and Yuan [19]. That approximation has the useful property that its derivative does not depend on the second derivatives of the objective or constraint functions.

Suppose that at the k-th iteration, the Newton direction $(\Delta x_k, \Delta y_k, \Delta z_k)$ has been determined and a trial step $\alpha_{x_k,i}$ is available. Since $\hat{y}(x)$ is a continuous function, we have

$$\hat{y}(x_k + \tau \alpha_{x_k,i} \Delta x_k) \approx \hat{y}_k + \tau \left[\hat{y}(x_k + \alpha_{x_k,i} \Delta x_k) - \hat{y}_k \right],$$

for $\tau \in [0,1]$, where $\hat{y}_k = \hat{y}(x_k)$. Similarly, the values of the merit function ϕ between the points x_k and $x_k + \alpha_{x_k,i} \Delta x_k$ can be approximated by the univariate function

$$\begin{aligned} \Phi_k(\tau \alpha_{x_k,i}) \quad = \quad & B(x_k + \tau \alpha_{x_k,i} \Delta x_k; \mu) \\ & - \left[\hat{y}_k + \tau (\hat{y}(x_k + \alpha_{x_k,i} \Delta x_k) - \hat{y}_k) \right]^T g(x_k + \tau \alpha_{x_k,i} \Delta x_k) \\ & + \frac{1}{2} c_{k,i} \parallel g(x_k + \tau \alpha_{x_k,i} \Delta x_k) \parallel^2, \end{aligned} \tag{18}$$

The strategies that determine the step sizes $\alpha_{x_k,i}$ and the penalty parameters $c_{k,i}$ for different i, are described in the following sections.

By direct substitution we can see that the approximate merit function Φ_k has the property

$$\Phi_k(0) = \phi(x_k) \text{ and } \Phi_k(\alpha_{x_k,i}) = \phi(x_k + \alpha_{x_k,i}\Delta x_k) \tag{19}$$

This property suggests that the line search procedure (17) can be replaced by

$$\Phi_k(\alpha_{x_k,i}) \leq \Phi_k(0) + \rho\alpha_{x_k,i}\Phi_k'(0) \tag{20}$$

where $\Phi_k'(0)$ is the first derivative of Φ_k with respect to τ at $\tau = 0$, defined as

$$\begin{aligned} \Phi_k'(0) &= \nabla B(x_k)^T \Delta x_k - \frac{1}{\alpha_{x_k,i}} \left[\hat{y}(x_k + \alpha_{x_k,i}\Delta x_k) - \hat{y}_k\right]^T g_k \\ &\quad -\hat{y}_k^T \nabla g_k \Delta x_k - c_{k,i} \parallel g_k \parallel^2 . \end{aligned} \tag{21}$$

Since $\Phi_k'(0)$ does not require the calculation of second derivatives of the objective and constraint functions, the approximate merit function Φ_k is preferable to the initial exact penalty function ϕ.

3.2 Penalty parameter selection and step size rules

In this section we describe the mechanism which ensures that Δx_k is a descent direction for the approximate merit function Φ_k, and the procedure for adjusting the step size $\alpha_{x_k,i}$ to guarantee reduction of Φ_k at each iteration. We also describe the strategy that determines the common step size α_z for the dual variables y, z. The analysis assumes that the barrier parameter μ is fixed. By noting that the decrease of the approximate merit function and equations (19) guarantee the decrease of the initial exact merit function ϕ, we show that the algorithm converges to a central point satisfying system (7).

In order to guarantee that Φ_k is reduced at each iteration k, the derivative $\Phi_k'(0)$ must be negative. From the definition of $\Phi_k'(0)$ we can see that if the penalty parameter $c_{k,i}$ is large enough then $\Phi_k'(0)$ can be negative. At the current iteration k, we select the value of the penalty parameter $c_{k,i}$ such that the descent condition

$$\Phi_k'(0) \leq -\frac{1}{2}\left[\Delta_k^T(H_k + X_k^{-1}Z_k)\Delta_k + c_{k,i} \parallel g_k \parallel^2\right] \leq -\frac{1}{4}c_{k,i} \parallel g_k \parallel^2 \leq 0 \tag{22}$$

is satisfied. The index i represents the number of times the descent condition has been checked at the k-th iteration.

Condition (22) is a modification of the corresponding condition defined by Powell and Yuan [19]. The matrix $H_k + X_k^{-1}Z_k$ represents an approximation of the Hessian matrix

$$\nabla_{xx}^2 L_B(x_k, y_k; \mu) = H_k + \mu X_k^{-2}$$

of the Lagrangian (5) of the barrier problem. The use of $H_k + X_k^{-1}Z_k$ instead of $\nabla_{xx}^2 L_B(x_k, y_k; \mu)$ is justified by the fact that our algorithm determines the Newton direction by solving the *primal-dual* system (8).

If (22) is satisfied, the penalty parameter does not increase, i.e., $c_{k,i+1} = c_{k,i}$. Otherwise, the penalty parameter is determined by

$$\begin{aligned} c_{k,i+1} &= \max\{2c_{k,i},\ -2\frac{\Delta_k^T(H_k + X_k^{-1}Z_k)\Delta_k}{\| g_k \|^2}, \\ & \frac{2}{\| g_k \|^2} [\ \frac{1}{2}\Delta_k^T(H_k + X_k^{-1}Z_k)\Delta_k + \nabla g_k \Delta x_k \\ & -\frac{1}{\alpha_{k,i}} [\ \hat{y}(x_k + \alpha_{x_k,i}\Delta x_k) - \hat{y}_k\ ]^T g_k - \hat{y}_k^T \nabla g_k \Delta x_k\ ]\ \}. \end{aligned} \tag{23}$$

The new value $c_{k,i+1}$ of the penalty parameter guarantees that $\Phi'_k(0) \leq 0$, since (22) is satisfied. We then check if inequality (20) is satisfied for $c_{k,i+1}$ and the current value of the step size $\alpha_{x_k,i}$. If it is not satisfied we reduce the step-size by choosing the new one, $\alpha_{x_k,i+1}$, from the interval $[\beta_1\alpha_{x_k,i}, \beta_2\alpha_{x_k,i}]$, for some $\beta_1, \beta_2 \in (0,1)$ and $\beta_1 \leq \beta_2$. This process is repeated until (20) is satisfied for the corresponding values of $c_{k,i+1}$ and $\alpha_{x_k,i+1}$. If on the other hand, (22) is satisfied for $c_{k,i+1}$ and $\alpha_{x_k,i}$ we set

$$x_{k+1} = x_k + \alpha_{x_k,i}\Delta x_k \ \text{ and } \ c_{k+1,0} = c_{k,i+1} \tag{24}$$

For the calculation of the new iterate of the dual variables z we use the information provided by the new primal iterate x_{k+1}. This is a modification of the strategy suggested by Yamashita [9] and Yamashita and Yabe [10].

While the barrier parameter μ is fixed, we determine a step $\alpha_{z_k}^{(j)}$ along the direction $\Delta z_k^{(j)}$, for each dual variable $z_k^{(j)}$, $j = 1, 2, ...n$, such that the box constraints

$$\alpha_{z_k}^{(j)} = \max\{\alpha > 0 :\ LB_k^{(j)} \leq (x_k^{(j)} + \alpha_{x_k}\Delta x_k^{(j)})(z_k^{(j)} + \alpha\Delta z_k^{(j)}) \leq UB_k^{(j)}\}. \tag{25}$$

are satisfied. The lower bounds $LB_k^{(j)}$ and upper bounds $UB_k^{(j)}$, $j = 1, 2, ..., n$ are defined as

$$LB_k^{(j)} = \min\{\frac{1}{2}m\mu, (x_k^{(j)} + \alpha_{x_k}\Delta x_k^{(j)})z_k^{(j)}\}, \tag{26}$$

and

$$UB_k^{(j)} = \max\{2M\mu, (x_k^{(j)} + \alpha_{x_k}\Delta x_k^{(j)})z_k^{(j)}\}, \tag{27}$$

where the parameters m and M are chosen such that

$$0 < m \leq \min\{1, \frac{(1-\gamma)(1-\frac{\gamma}{(M_0)^\mu})\min_i\{x_k^{(j)}z_k^{(j)}\}}{\mu}\}, \tag{28}$$

and

$$M \geq \max\{1, \frac{\max_i\{x_k^{(j)}z_k^{(j)}\}}{\mu}\} > 0, \tag{29}$$

with $\gamma \in (0,1)$ and M_0 a positive large number. These two parameters are always fixed to constants which satisfy (28) and (29), while μ is fixed. The values of m and M change when the barrier parameter μ is decreased.

The common dual step length α_{z_k} is the minimum of all individual step lengths $\alpha_{z_k}^{(j)}$ with the restriction of being always not more that one, namely

$$\alpha_{z_k} = min\{1, \min_{1 \le j \le n} \{\alpha_{z_k}^{(j)}\}\}. \tag{30}$$

The step-size for the dual variables y can be either $\alpha_{yk} = 1$ or $\alpha_{yk} = \alpha_{z_k}$.

The difference between the present step-size rule and the one proposed in [10] lies in the definition of the lower bounds $LB_k^{(j)}$, $j = 1, 2, ..., n$ of the box constraints (25). In particular, the term $1 - \gamma/(M_0)^\mu \in (0, 1)$ in the definition of the parameter m, given by (28), results in the lower bounds $LB_k^{(j)}$ being smaller than the corresponding bounds defined in [10]. Consequently, the step lengths α_{z_k} are larger than those in [10]. Also by noting that

$$\lim_{\mu \to 0} (1 - \gamma/(M_0)^\mu) = 1 - \gamma$$

and

$$z_k^{(j)} + \alpha_{z_k} \Delta z_k^{(j)} \ge (1 - \gamma) z_k^{(j)} > 0, \text{ for all } j = 1, 2, ...n$$

it can be shown that asymptotically the algorithm accepts the maximum allowable step for the dual variables.

A summary of the procedure we use to find the new penalty parameter and different step-sizes for the primal and dual variables of the problem is described bellow

Algorithm 1 Solution of the Barrier problems (4)

At the beginning of the k-th iteration the following items are available:
A point (x_k, y_k, z_k) such that $x_k, z_k > 0$, and parameters $\mu > 0$, $c_{k,0} \ge 0$,
$\eta, \rho > 0$, $\beta_1, \beta_2 \in (0, 1)$ with $\beta_1 \le \beta_2$

Repeat until $\| F(x_k, y_k, z_k; \mu) \| \le \eta\mu$
  Compute Newton direction $(\Delta x_k, \Delta y_k, \Delta z_k)$, by solving system (8)
  Set $i = 0$ and $\alpha_{k,i} = \alpha_{x_k}^{max}$
  If (22) is satisfied then set $c_{k,i+1} = c_{k,i}$
  Else compute $c_{k,i+1}$ from (23)

  Repeat until (20) is satisfied
    Choose $\alpha_{k,i+1} \in [\beta_1 \alpha_{x_k,i}, \beta_2 \alpha_{x_k,i}]$.
    Set $i = i + 1$
  End

  Set $i_k = i$, $x_{k+1} = x_k + \alpha_{x_k,i_k} \Delta x_k$ and $c_{k+1,0} = c_{k,i_k}$
  Compute the dual step-size α_{z_k} from (30)
  Set $z_{k+1} = z_k + \alpha_{z_k} \Delta z_k$ and $y_{k+1} = y_k + \alpha_{z_k} \Delta y_k$
  Set $k = k + 1$

End

From the above discussion, we can derive that the values of the penalty parameter are non-decreasing and either one of the two cases may happen

$$c_{k,i+1} = c_{k,i} \quad \text{or} \quad c_{k,i+1} \geq 2c_{k,i} \tag{31}$$

A critical issue with equation (23) is that a division by zero may occur when the current point satisfies the equality constraints. The following proposition shows that this does not occur as long as the Hessian matrix H_k (or its approximation) is positive definite.

Lemma 1 *Let* $\| (\Delta x_k, \Delta y_k, \Delta z_k) \| > 0$*, the descent condition (22) not be satisfied and the matrix* H_k *be positive definite. Then we have* $\| g_k \| \neq 0$*.*

Proof Assume on the contrary that $\| g_k \| = 0$. Then from the violation of the descent condition (22) we have

$$\nabla f_k^T \Delta x_k - \mu e^T X_k^{-1} \Delta x_k > -\frac{1}{2} \Delta x_k^T (H_k + X_k^{-1} Z_k) \Delta x_k \tag{32}$$

Pre-multiplying the first equation of system (11) by Δx_k^T we have

$$\Delta x_k^T \nabla f_k^T - \mu \Delta x_k^T X_k^{-1} e = -\Delta x_k^T (H_k + X_k^{-1} Z_k) \Delta x_k \tag{33}$$

Substituting (33) into (32) yields

$$-\Delta x_k^T (H_k + X_k^{-1} Z_k) \Delta x_k > -\frac{1}{2} \Delta x_k^T (H_k + X_k^{-1} Z_k) \Delta x_k \tag{34}$$

Recalling that the elements of the diagonal matrix $X_k^{-1} Z_k$ are positive for every iteration, from (34) we can obtain

$$-\Delta x_k^T H_k \Delta x_k > 0 \tag{35}$$

which contradicts the assumption that the H_k is positive definite. Hence $\| g_k \| \neq 0$. •

An immediate consequence of Lemma 1 is that the descent condition (22) is satisfied when $\| g_k \| = 0$, i.e., when feasibility of the equality constraints has been achieved.

Corollary 1 *Let* $\| (\Delta x_k, \Delta y_k, \Delta z_k) \| > 0$*, the matrix* H_k *be positive definite and* $\| g_k \| = 0$*. Then the descent condition (22) is satisfied for the current value* $c_{k,i}$ *of the penalty parameter, and therefore* $c_{k,i+1} = c_{k,i}$*.*

Proof It suffices to show that

$$\Phi_k'(0) + \frac{1}{2} \left[\Delta_k^T (H_k + X_k^{-1} Z_k) \Delta_k + c_{k,i} \| g_k \|^2 \right] < 0 \tag{36}$$

From (21), and the assumption that $\| g_k \| = 0$ we have

$$\Phi_k'(0) = \nabla f_k^T \Delta x_k - \mu e^T X_k^{-1} \Delta x_k$$

Using (33) the above equation becomes

$$\Phi_k'(0) = -\Delta x_k^T (H_k + X_k^{-1} Z_k) \Delta x_k \tag{37}$$

From (37) and the assumptions that H_k is positive definite and $\| g_k \| = 0$, it is clear that (36) is satisfied. •

Furthermore each component of the sequence $\{x_k\}$ is bounded above by Assumption (A2), and away from zero by the existence of the logarithmic barrier term. The next lemma, proved by Yamashita [9], states that the sequence $\{z_k\}$ has similar properties.

Lemma 2 *While μ is fixed, the lower bounds LB_k^i and the upper bounds UB_k^i, $i = 1, 2, ..., n$, of the box constraints in the dual step-size rule, are bounded away from zero and bounded from above respectively, if the corresponding components x_k^i, of the iterates x_k are also bounded above and away from zero.*

Proof The proof can be found in [9]. •

A result of the above lemma is that the elements of the diagonal matrix $X_k^{-1} Z_k$ are bounded above and away from zero. Also from (14) there exists a positive constant M_2 such that

$$\frac{1}{M_2} \| v \|^2 \leq v^T (H_k + X_k^{-1} Z_k) v \leq M_2 \| v \|^2 \tag{38}$$

Based on the above property of the matrix $H_k + X_k^{-1} Z_k$, we can derive a useful upper bound on the derivative $\Phi_k'(0)$ of the approximate merit function.

Lemma 3 *The descent condition (22), Assumptions (A1)-(A4) and (38) imply the inequality*

$$\Phi_k'(0) \leq -\zeta \| \Delta x_k \|^2, \tag{39}$$

where $\zeta > 0$.

Proof The proof is similar to that in [19]. For all iterations $k \geq 0$, define $\nu \in (0, \infty)$ such that

$$2\nu \| H_k + X_k^{-1} Z_k \| + \nu^2 \| H_k + X_k^{-1} Z_k \| \leq \frac{1}{2} \frac{1}{M_2} \tag{40}$$

where M_2 is defined in (38). Furthermore the direction Δx_k can be written as

$$\Delta x_k = \Delta x_k^{(1)} + \Delta x_k^{(2)} \tag{41}$$

where $\Delta x_k^{(1)}$ and $\Delta x_k^{(2)}$ are the projections of Δx_k at the range space of ∇g_k^T and the null space of ∇g_k, respectively. Hence the vectors $\Delta x_k^{(1)}$ and $\Delta x_k^{(2)}$ have the properties

$$\nabla g_k \Delta x_k^{(2)} = 0 \quad \text{and} \quad (\Delta x_k^{(1)})^T \Delta x_k^{(2)} = 0 \tag{42}$$

Substituting (41) into the second equation of system (11), and using the first equation of (42) we have

$$\nabla g_k \Delta x_k^{(1)} = -g_k$$

Using Assumption (A2), the above equation yields

$$\| \Delta x_k^{(1)} \| \leq M_3 \| g_k \| \tag{43}$$

where M_3 is a positive constant. Also from (38) we have

$$(\Delta x_k^{(2)})^T (H_k + X_k^{-1} Z_k) \Delta x_k^{(2)} \leq M_2 \| \Delta x_k^{(2)} \|^2 \tag{44}$$

We distinguish two cases. If $\| \Delta x_k^{(1)} \| \geq \nu \| \Delta x_k^{(2)} \|$, then from (43), using (40), (41) and the second part of (42) yields

$$\| g_k \|^2 \geq \frac{1}{M_3^2} \| \Delta x_k^{(1)} \|^2 \geq \frac{1}{M_3^2} \frac{(\| \Delta x_k^{(1)} \| + \| \Delta x_k^{(2)} \|)^2}{(1+\nu^{-1})^2} = \frac{\| \Delta x_k \|^2}{M_3^2 (1+\nu^{-1})^2} \tag{45}$$

If $\| \Delta x_k^{(1)} \| < \nu \| \Delta x_k^{(2)} \|$, then from (40) and (44) we have

$$\begin{aligned}
\Delta x_k^T (H_k + X_k^{-1} Z_k) \Delta x_k &= (\Delta x_k^{(1)} + \Delta x_k^{(2)})^T (H_k + X_k^{-1} Z_k)(\Delta x_k^{(1)} + \Delta x_k^{(2)}) \\
&\geq \| \Delta x_k^{(2)} \|^2 \, (\frac{1}{M_2} - 2\nu \| H_k + X_k^{-1} Z_k \| \\
&\qquad -\nu^2 \| H_k + X_k^{-1} Z_k \|) \\
&\geq \frac{1}{2} \frac{1}{M_2} \| \Delta x_k^{(2)} \|^2 \\
&> \frac{1}{2} \frac{1}{M_2} \frac{(\| \Delta x_k^{(1)} \| + \| \Delta x_k^{(2)} \|)^2}{(1+\nu)^2} \\
&= \frac{1}{2} \frac{1}{M_2} \frac{\| \Delta x_k \|^2}{(1+\nu)^2}
\end{aligned} \tag{46}$$

Substituting (45) into the last inequality of (22) yields

$$\Phi_k'(0) \leq -\frac{1}{4} c_{k,i} \frac{\| \Delta x_k \|^2}{M_3^2 (1+\nu^{-1})^2} \leq -\frac{1}{4} c_{1,0} \frac{\| \Delta x_k \|^2}{M_3^2 (1+\nu^{-1})^2} \tag{47}$$

where $c_{1,0}$ is the initial value of the penalty parameter. Similarly substituting (46) into the middle inequality of (22) yields

$$\Phi_k'(0) \leq -\frac{1}{4}\frac{1}{M_2}\frac{\| \Delta x_k \|^2}{(1+\nu)^2} \tag{48}$$

Finally from (47) and (48) we can derive that if

$$\zeta = \frac{1}{4}\min\{c_{1,0}\frac{1}{M_3^2(1+\nu^{-1})^2}, \frac{1}{M_2}\frac{1}{(1+\nu)^2}\}$$

then (39) holds. •

We next show that i_k in Algorithm 1, is always finite. This means that the step size $\alpha_{x_k,i}$ is bounded away from zero.

Lemma 4 *Let* $\| (\Delta x_k, \Delta y_k, \Delta z_k) \| > 0$ *and the descent condition (22) hold. Then* i_k *is finite.*

Proof The proof is similar to that in [19]. Assume on the contrary that i_k becomes infinity. Then for all $i = 0, 1, 2, ...,$ we have

$$\Phi_k(\alpha_{x_k,i}) - \Phi_k(0) > \rho\alpha_{x_k,i}\Phi_k'(0) \tag{49}$$

From (31) we can derive that either

$$\lim_{i\to\infty} c_{k,i} = \infty \tag{50}$$

or

$$c_{k,i} = c_{k,i_*} = \overline{c_k}, \quad \text{for all } i \geq i_* \tag{51}$$

First we assume that (50) holds. This means that the descent condition (22) does not hold and the penalty parameter needs to increase infinitely many times. From the violation of (22) we have

$$\begin{aligned} \nabla B(x_k;\mu)^T\Delta x_k - \hat{y}_k^T\nabla g_k\Delta x_k + \frac{1}{2}\Delta x_k^T(H_k + X_K^{-1}Z_k)\Delta x_k \\ -\frac{1}{\alpha_{x_k,i}}\left[\hat{y}(x_k + \alpha_{x_k,i}\Delta x_k) - \hat{y}_k\right]^T g_k \ \geq \ \frac{1}{2}c_{k,i} \| g_k \|^2 \end{aligned} \tag{52}$$

Moreover from Lemma 1 we have that $\| g_k \| \neq 0$. Hence, from (50), we can derive that the left hand side of inequality (52) becomes unbounded, as $i \to \infty$. However only the term

$$\frac{1}{\alpha_{x_k,i}}\left[\hat{y}(x_k + \alpha_{x_k,i}\Delta x_k) - \hat{y}_k\right]^T g_k$$

depends on i and due to the continuity of the second derivatives and the full rank of ∇g_k^T, (52) remains finite. This shows that (50) is impossible.

Assume that (51) holds. Then for all $i \geq i_*$, from the first order Taylor expansion of Φ_k we have

$$\Phi_k(\alpha_{x_k,i}) - \Phi_k(0) - \rho\alpha_{x_k,i}\Phi'_k(0) = (1-\rho)\alpha_{x_k,i}\Phi'_k(0) + o(\alpha_{x_k,i}) \tag{53}$$

From Lemma 3 we deduce that the right hand side of (53) becomes negative as $i \to \infty$, which contradicts (49). •

The next lemma shows that the sequence of penalty parameters $\{c_{k,i}\}$ is bounded, i.e., the penalty parameter does not increase infinitely.

Lemma 5 *Let Assumptions (A1)-(A4) hold and the barrier parameter μ be fixed. Then there exists an iterate k_* and a constant $\overline{c} \in [0, \infty)$, such that*

$$c_{k,i} = c_{k_*,0} = \overline{c} \tag{54}$$

for all $k \geq k_*$ and $0 \leq i \leq i_k$.

Proof Since $c_{k,i}$ satisfies only one of the conditions in (31) it is sufficient to show that there exists a constant $\overline{c}$ such that the descent condition (22) is satisfied if $c_{k,i} \geq \overline{c}$. We first need to show that

$$\nabla B(x_k;\mu)^T \Delta x_k + \Delta x_k^T (H_k + X_k^{-1} Z_k)\Delta x_k - \hat{y}_k^T \nabla g_k \Delta x_k = O(\| \Delta x_k \| \| g_k \|) \tag{55}$$

Pre-multiplying by Δx_k the first equation of system (11) we have

$$\Delta x_k^T(\nabla f_k - \mu X_k^{-1} e) + \Delta x_k^T (H_k + X_k^{-1} Z_k)\Delta x_k = -\Delta x_k^T \nabla g_k^T \overline{y}_k \tag{56}$$

where $\overline{y}_k = y_k + \Delta y_k$. Adding $\Delta x_k^T \nabla g_k^T \hat{y}_k$ to both sides of (56) and using the second equation of (11) yields the bound

$$\begin{aligned} |\Delta x_k^T(\nabla f_k - \mu X_k^{-1} e) + \Delta x_k^T (H_k + X_k^{-1} Z_k)\Delta x_k| &= |\Delta x_k^T \nabla g_k^T(\overline{y}_k - \hat{y}_k)| \\ &\leq \| g_k \| \| \overline{y}_k - \hat{y}_k \| \end{aligned} \tag{57}$$

Moreover, from Assumption (A2) we have the condition

$$\begin{aligned} \| \overline{y}_k - \hat{y}_k \| &= O(\| \nabla g_k^T \overline{y}_k - \nabla g_k^T \hat{y}_k \|) \\ &= O(\| \nabla g_k^T \overline{y}_k + \nabla f_k - \mu X_k^{-1} e \| + \| \nabla g_k^T \hat{y}_k + \nabla f_k - \mu X_k^{-1} e \|) \\ &= O(\| (H_k + X_k^{-1} Z_k)\Delta x_k \|) \end{aligned} \tag{58}$$

where the last equation derives from the definition of $\hat{y}_k$ as the least squares approximation of the Lagrange multipliers, which gives the inequality

$$\| \nabla g_k^T \hat{y}_k + \nabla f_k - \mu X_k^{-1} e \| \leq \| \nabla g_k^T \overline{y}_k + \nabla f_k - \mu X_k^{-1} e \|$$

Therefore (55) follows from the fact that the matrices $H_k + X_k^{-1} Z_k$ are bounded (see Assumption A4 and (38)). The rest of the proof is similar to Lemma 3.4 in [19]. •

Having established that the sequences of iterates $\{x_k\}$ and $\{z_k\}$ are bounded above and away from zero, we show that the iterates $\{y_k\}$, $k \geq 0$ are also bounded. In particular Lemma 7 shows that if at each iteration of the algorithm we take a unit step along the direction Δy_k, then the resulting sequence $\{y_k + \Delta y_k\}$ is bounded. In addition to this, Lemma 7 also shows that the Newton direction $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k)$ is bounded, for fixed μ. We first establish the following technical result.

Lemma 6 *Let w_k is a sequence of vectors generated by Algorithm 1 for fixed μ. Then the matrix sequence $\{\Theta_k^{-1}\}$ is bounded, where*

$$\Theta_k = \begin{pmatrix} 0 & \nabla g_k \\ -\nabla g_k^T & H_k + X_k^{-1} Z_k \end{pmatrix}.$$

Proof The inverse of the partitioned matrix Θ_k is

$$\Theta_k^{-1} = \begin{pmatrix} [\nabla g_k \Omega_k \nabla g_k^T]^{-1} & -[\nabla g_k \Omega_k \nabla g_k^T]^{-1} \nabla g_k \Omega_k \\ \Omega_k \nabla g_k^T [\nabla g_k \Omega_k \nabla g_k^T]^{-1} & \Omega_k - \Omega_k \nabla g_k^T [\nabla g_k \Omega_k \nabla g_k^T]^{-1} \nabla g_k \Omega_k \end{pmatrix},$$

where $\Omega_k = (H_k + X_k^{-1} Z_k)^{-1}$. According to Assumption (A4), and Lemma 2, the matrices Ω_k and $[\nabla g_k \Omega_k \nabla g_k^T]^{-1}$ exist and are bounded. Hence the matrix Θ_k^{-1} is bounded, since all matrices involved in it are bounded. •

Lemma 7 *Let w_k be a sequence of vectors generated by Algorithm 1 for fixed μ. Then the sequence of vectors $\{(\Delta x_k, y_k + \Delta y_k, \Delta z_k)\}$ is bounded.*

Proof Re-arranging the system (11) yields

$$\begin{pmatrix} 0 & \nabla g_k \\ -\nabla g_k^T & H_k + X_k^{-1} Z_k \end{pmatrix} \begin{pmatrix} y_k' \\ \Delta x_k \end{pmatrix} = - \begin{pmatrix} g_k \\ \nabla f_k - \mu X_k^{-1} e \end{pmatrix} \tag{59}$$

where $y_k' = y_k + \Delta y_k$. From the previous lemma we have that the inverse of the matrix in the left side of (59) exists and is bounded. Hence the sequences $\{\Delta x_k\}$ and $\{y_k'\}$ are also bounded. Considering now (10), we deduce that the sequence $\{\Delta z_k\}$ is bounded. •

We then show that the direction Δx_k becomes small.

Lemma 8 *Let Assumptions (A1)-(A4) hold. Then we have*

$$\lim_{k\to\infty} \Delta x_k = 0 \tag{60}$$

Proof The proof is similar to that in Powell and Yuan [19]. From Lemma 5 we know that there exists an iteration k_* such that for every iteration $k \geq k_*$ the penalty parameter does not increase more than a constant $\overline{c}$. From (19) and (20), the exact merit function ϕ is monotonically decreasing for all $k \geq k_*$, and therefore the sequence $\{\phi(x_k) : k \geq k_*\}$ is convergent. We then show that for μ fixed and $k \geq k_*$, if $\| \Delta x_k \| > 0$, i.e., x_k is not on the central path, the Algorithm 1 will find a new primal iterate $x_{k+1} = x_k + \alpha_{x_k,i}\Delta x_k$ such that

$$\phi(x_{k+1}; \overline{c}) < \phi(x_k; \overline{c}). \tag{61}$$

Hence Δx_k is a descent direction for the exact merit function ϕ.

From the first order Taylor expansion of the approximate merit function Φ_k, for $k \geq k_*$, we have

$$\Phi_k(\alpha_{x_k,i}) - \Phi_k(0) - \rho\alpha_{x_k,i}\Phi_k'(0) = (1-\rho)\alpha_{k,i}\Phi_k'(0) + \psi_{k,i}, \tag{62}$$

where

$$\begin{aligned} \psi_{k,i} &= \Phi_k(\alpha_{x_k,i}) - \Phi_k(0) - \alpha_{x_k,i}\Phi_k'(0) \\ &= \phi(x_k + \alpha_{x_k,i}\Delta x_k; \overline{c}) - \phi(x_k; \overline{c}) - \alpha_{x_k,i}\nabla\phi(x_k; \overline{c})^T \Delta x_k \\ & \quad [\hat{y}(x_k + \alpha_{x_k,i}\Delta x_k) - \hat{y}_k]^T g_k - \alpha_{x_k,i}\Delta x_k^T \nabla \hat{y}_k^T g_k. \end{aligned} \tag{63}$$

From (63) we can see that

$$|\psi_{k,i}| = o(\alpha_{x_k,i}). \tag{64}$$

Furthermore, from (64) we can see that there exists a constant $\hat{\epsilon} > 0$, such that, if $\alpha_{x_k,i} < \hat{\epsilon}$, then the right hand side of (62) can be negative i.e.,

$$\psi_{k,i} < -(1-\rho)\alpha_{x_k,i}\Phi_k'(0) \leq (1-\rho)\alpha_{x_k,i}\zeta\epsilon^2 \tag{65}$$

Hence at the k-th iteration the step length is at least $\beta_1\hat{\epsilon} > 0$. From (62), using (19) and (65), we have

$$\begin{aligned} \phi(x_k; \overline{c}) - \phi(x_k + \alpha_{x_k,i}\Delta x_k; \overline{c}) &= \Phi_k(0) - \Phi_k(\alpha_{x_k,i}) \ \geq -\rho\alpha_{x_k,i}\Phi_k'(0) \\ &\geq \rho\beta_1\hat{\epsilon}\zeta\epsilon^2 \ > 0 \end{aligned}$$

which shows that (61) is true. •

The next theorem shows that, while the barrier parameter μ is fixed, the iterates (x_k, y_k, z_k) converge to a point satisfying the perturbed optimality conditions (7).

Theorem 1 *Let the Assumptions (A1)-(A4) hold and let μ be fixed. Then Algorithm 1 terminates at a point, satisfying the perturbed optimality conditions (7).*

Proof We first prove that for k sufficiently large, the dual step, α_{z_k}, becomes unity, by showing that

$$\lim_{k\to\infty} \| z_k + \Delta z_k - \mu X_{k+1}^{-1} e \| = 0. \tag{66}$$

By adding $-\mu X_{k+1}^{-1} e$ to both sides of (10) yields

$$\| z_k + \Delta z_k - \mu X_{k+1}^{-1} e \| \;\; \leq \;\; \| -X_k^{-1} Z_k \| \;\; \| \Delta x_k \| + \mu \| X_k^{-1} - X_{k+1}^{-1} \| \;\; \| e \| \tag{67}$$

Moreover

$$\begin{aligned} \| X_k^{-1} - X_{k+1}^{-1} \|^2 \;\; &\leq \;\; n \max_{1\leq j\leq n} \{ (\frac{1}{x_k^{(j)}} - \frac{1}{x_{k+1}^{(j)}})^2 \} \\ &= \;\; n \max_{1\leq j\leq n} \{ \frac{(\alpha_{x_k})^2 (\Delta x_k^{(j)})^2}{(x_k^{(j)})^2 \; (x_{k+1}^{(j)})^2} \} \end{aligned}$$

Since we always have $\alpha_{x_k} \in (0,1]$, $(\Delta x_k^{(j)})^2 \leq \| \Delta x_k \|^2$ and the sequence $\{x_k\}$ is bounded away from zero, from the above inequality and (60) we can show that

$$\lim_{k\to\infty} \| X_k^{-1} - X_{k+1}^{-1} \|^2 \;\; \leq \;\; n \lim_{k\to\infty} \max_{1\leq j\leq n} \{ \frac{\| \Delta x_k \|^2}{(x_k^{(j)})^2 \; (x_{k+1}^{(j)})^2} \} = 0 \tag{68}$$

Hence letting $k \to \infty$ in (67), and using (60) and (68) it can be deduced that (66) holds. Consequently, $z_{k+1} = z_k + \Delta z_k$, for k sufficiently large.

Furthermore, using (10) and for k sufficiently large, the complementarity condition becomes

$$X_{k+1} z_{k+1} = X_{k+1}(z_k + \Delta z_k) = X_{k+1} X_k^{-1} (-Z_k \Delta x_k + \mu e) \tag{69}$$

From (60) and the fact that the elements of the diagonal matrix $X_{k+1} X_k^{-1}$ can be written as

$$\frac{x_{k+1}^{(j)}}{x_k^{(j)}} = 1 + \alpha_{x_k} \frac{\Delta x_k^{(j)}}{x_k^{(j)}}, \qquad \text{for all } j = 1, 2, ..., n,$$

we can derive that

$$\lim_{k\to\infty} X_{k+1} X_k^{-1} = I_n \tag{70}$$

where I_n is the $n \times n$ identity matrix. Letting $k \to \infty$ in (69), and using (60) and (70) yields

$$\lim_{k\to\infty} X_{k+1} z_{k+1} = X_*(\mu) z_*(\mu) = \mu e \tag{71}$$

Also for $k \to \infty$, the second equation of system (11) and (60) yield

$$\lim_{k\to\infty} (\nabla g_k \Delta x_k) = g(x_*(\mu)) = 0 \tag{72}$$

The first equation of the system (11) can be written as

$$\nabla f_k - \nabla g_k^T y_{k+1} + c_* \nabla g_k^T g_k - \mu X_k^{-1} e = -(H_k + X_k^{-1} Z_k) \Delta x_k$$

where $y_{k+1} = y_k + \Delta y_k$. Letting $k \to \infty$, and using (60) the above equation yields

$$\lim_{k\to\infty} \| \nabla f_k - \nabla g_k^T y_{k+1} + c_* \nabla g_k^T g_k - \mu X_k^{-1} e \| = 0 \tag{73}$$

From the assumptions that the functions f and g have continuous gradients and ∇g_k^T has full column rank and using (68), equation (73) yields

$$\lim_{k\to\infty} \| \nabla f_{k+1} - \nabla g_{k+1}^T y_{k+1} + c_* \nabla g_{k+1}^T g_{k+1} - \mu X_{k+1}^{-1} e \| = 0$$

or equivalently

$$\nabla f(x_*(\mu)) - \nabla g(x_*(\mu))^T y_*(\mu) + c_* \nabla g(x_*(\mu))^T g(x_*(\mu)) - \mu X_*(\mu)^{-1} e = 0 \tag{74}$$

From (74), (72) and (71) we can conclude that the vector $(x_*(\mu), y_*(\mu), z_*(\mu))$ is a solution of the perturbed optimality conditions (8). •

4 Global convergence

In this section we discuss the convergence of the algorithm for a decreasing sequence of positive barrier parameters $\{\mu^l : l = 0, 1, 2, ...\}$. Theorem 1 guarantees that if $\mu = \mu^l$, for some $l \geq 0$, there is an iteration $\tilde{k}$ such that for all $k \geq \tilde{k}$

$$\| F(x_k, y_k, z_k; \mu^l) \| \leq \eta \mu^l \tag{75}$$

for some $\eta > 0$. At this point the barrier parameter is reduced, i.e., $\mu = \mu^{l+1} < \mu^l$ and the iterations proceed. If we define

$$(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) = (x_{\tilde{k}}, y_{\tilde{k}}, z_{\tilde{k}})$$

then the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) : l = 0, 1, 2, ...\}$ approximates the sequence of central points $\{(x(\mu^l), y(\mu^l), z(\mu^l)) : l = 0, 1, 2, ...\}$.

Algorithm 2, below, provides the a description of the main steps performed when the barrier parameter changes.

Algorithm 2

Initialisation: Choose a starting point (x_0, y_0, z_0) such that $x_0, z_0 > 0$.
Choose initial values for the penalty $c_{0,0} \geq 0$ and barrier parameter $\mu^0 > 0$.
Select the parameters $\eta > 0, \epsilon_0, \theta \in (0, 1)$
Set $k = 0$, $l = 0$ and $(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) = (x_k, y_k, z_k)$.

Repeat until $\| F(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l; 0) \| \leq \epsilon_0$

Apply Algorithm 1 to find a point $(x_{\tilde{k}}, y_{\tilde{k}}, z_{\tilde{k}})$, such that

$$\| F(x_{\tilde{k}}, y_{\tilde{k}}, z_{\tilde{k}}; \mu^l) \| \leq \eta \mu^l$$

Set $(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) = (x_{\tilde{k}}, y_{\tilde{k}}, z_{\tilde{k}})$, $\mu^{l+1} = \theta \mu^l$, $l = l + 1$.

End

Note that the index l is used to count the outer iterations (i.e., number of times the barrier parameter is reduced), whereas the index k, which changes within Algorithm 1, is used to count the total number of iterations needed to find an optimal solution of problem (1). In the sequel we show that the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) : l = 0, 1, 2, ...\}$ generated by Algorithm 2, converges to a point $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ which is an optimal solution of the original problem (1).

For a given $\epsilon \geq 0$, consider the set of all the approximate central points, generated by Algorithm 1

$$S(\epsilon) = \{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) : \epsilon \leq \| F(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l; \mu^l) \| \leq \| F(\tilde{x}^0, \tilde{y}^0, \tilde{z}^0; , \mu^0) \|, \ \forall \mu^l < \mu^0\}.$$

If $\epsilon > 0$ then the step-size rules, described in section 3 guarantee that $\tilde{x}^l, \tilde{z}^l \in S(\epsilon)$ are bounded away from zero, for $l \geq 0$. Consequently $(\tilde{x}^l)^T \tilde{z}^l$ is also bounded away from zero in $S(\epsilon)$. The following lemma shows that the sequence $\{\tilde{y}^l\}$ is bounded if the sequence $\{\tilde{z}^l\}$ is also bounded.

Lemma 9 *Let the columns of $\nabla g(\tilde{x}^l)$ be linearly independent and the iterates $\tilde{x}^l$ be in a compact set for $l \geq 0$. Then there exists a constant $M_4 > 0$ such that*

$$\| \tilde{y}^l \| \leq M_4 (1 + \| \tilde{z}^l \|).$$

Proof By defining $r^l = \nabla f(\tilde{x}^l) - \tilde{z}^l - \nabla g(\tilde{x}^l)^T \tilde{y}^l$ and solving for $\nabla g(\tilde{x}^l)^T \tilde{y}^l$ we obtain

$$\nabla g(\tilde{x}^l)^T \tilde{y}^l = \nabla f(\tilde{x}^l) - \tilde{z}^l - r^l.$$

From our assumptions the above equation can be written as

$$\begin{aligned}\tilde{y}^l &= [\nabla g(\tilde{x}^l)\nabla g(\tilde{x}^l)^T]^{-1}\nabla g(\tilde{x}^l)\,(\nabla f(\tilde{x}^l) - r^l) \\ &\quad -[\nabla g(\tilde{x}^l)\nabla g(\tilde{x}^l)^T]^{-1}\nabla g(\tilde{x}^l)\,\tilde{z}^l.\end{aligned}$$

Taking norms in both sides of the above equation yields

$$\begin{aligned}\| \tilde{y}^l \| &\leq \| [\nabla g(\tilde{x}^l)\nabla g(\tilde{x}^l)^T]^{-1}\nabla g(\tilde{x}^l) \| \| \nabla f(\tilde{x}^l) - r^l \| \\ &\quad + \| [\nabla g(\tilde{x}^l)\nabla g(\tilde{x}^l)^T]^{-1}\nabla g(\tilde{x}^l) \| \| \tilde{z}^l \| \\ &\leq M_4(1+ \| \tilde{z}^l \|).\end{aligned}$$

where the constant M_4 is defined as

$$\begin{aligned}M_4 &\geq \max\{ \| [\nabla g(\tilde{x}^l)\nabla g(\tilde{x}^l)^T]^{-1}\nabla g(\tilde{x}^l) \| \| \nabla f(\tilde{x}^l) - r^l \|, \\ &\quad \| [\nabla g(\tilde{x}^l)\nabla g(\tilde{x}^l)^T]^{-1}\nabla g(\tilde{x}^l) \| \}.\end{aligned}$$

and is finite according to our assumptions. •

Lemma 10 *Let $(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) \in S(\epsilon)$ for all $l \geq 0$. Then the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ is bounded above.*

Proof From Lemma 9, it suffices to show that the sequences $\{\tilde{x}^l\}$ and $\{\tilde{z}^l\}$ are bounded from above. By assumption (A2), the sequence $\{\tilde{x}^l\}$ is bounded. Assume that there exists a non-empty set I_z^∞, which contains the indices j of those elements, $(\tilde{z}^l)^{(j)}$, of the vector $\tilde{z}^l$, for which $\lim_{l\to\infty}(\tilde{z}^l)^{(j)} = \infty$. From the boundedness of the sequences $\{(\tilde{x}^l)^{(j)}(\tilde{z}^l)^{(j)}\}$, $j = 1, 2, ..., n$, we obtain $\liminf_{l\to\infty}(\tilde{x}^l)^{(j)} = 0$, for those indices $j \in I_z^\infty$. Furthermore from the definition of the set I_x^0, in Assumption (A4), it is evident that $I_z^\infty \subseteq I_x^0$.

From (75) and the fact that $\{\mu^l\} \to 0$ we have that the sequence

$$\{\| \nabla f(\tilde{x}^l) - \tilde{z}^l - \nabla g(\tilde{x}^l)^T\tilde{y}^l \|\}$$

is bounded. Using this and the fact that $\{\| \nabla f(\tilde{x}^l) \|\}$ is bounded, we conclude that $\{\| -\tilde{z}^l - \nabla g(\tilde{x}^l)^T\tilde{y}^l \|\}$ is also bounded. Hence, we have

$$\frac{\| \tilde{z}^l + \nabla g(\tilde{x}^l)^T\tilde{y}^l \|}{\| (\tilde{y}^l, \tilde{z}^l) \|} \to 0 \tag{76}$$

By setting $\tilde{u}^l = (\tilde{y}^l, \tilde{z}^l)/ \| (\tilde{y}^l, \tilde{z}^l) \|$, we have $\{\tilde{u}^l\}$ bounded and $\{\tilde{u}^l\} \to \tilde{u}^*$. It is clear that $\| \tilde{u}^* \|= 1$ and the components of $\tilde{u}^*$, corresponding to those indices $j \notin I_z^\infty$, i.e., $\{(\tilde{z}^l)^{(j)}\} < \infty$, are zero. If $\hat{u}^*$ is the vector consisting of the components of $\tilde{u}^*$ which correspond to the indices $j \in I_z^\infty$, then $\| \hat{u}^* \|=\| \tilde{u}^* \|= 1$. Furthermore, from (76) we have

$$\frac{\nabla g(\tilde{x}^l)^T \tilde{y}^l + \tilde{z}^l}{\| (\tilde{y}^l, \tilde{z}^l) \|} = \frac{[\nabla g(\tilde{x}^l)^T,\ I_n]\,(\tilde{y}^l, \tilde{z}^l)}{\| (\tilde{y}^l, \tilde{z}^l) \|} = [\nabla g(\tilde{x}^l)^T,\ e_j : j \in I_x^0]\,\hat{u}^* \to 0.$$

However, this result contradicts Assumption (A2). Hence, the set I_z^∞ is empty, or for all indices $j = 1, 2, ..., n$, the sequences $\{(\tilde{z}^l)^{(j)}\}$ are bounded. Consequently, $\{\tilde{z}^l\}$ is also bounded. •

The following theorem shows that the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ converges to an optimum point $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ of the initial constrained optimisation problem (1).

Theorem 2 *Let $\{\mu^l\}$ is a positive monotonically decreasing sequence of barrier parameters with $\{\mu^l\} \to 0$, and let $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) : l \geq 0\}$ be a sequence of approximate central points satisfying (75). Then the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ is bounded and its limit point $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ satisfies the first order optimality conditions of problem (1).*

Proof From Lemma 9 the sequence $\{(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l)\}$ is bounded. Then it is convergent and let $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ be its limit point. From (75) and the fact that $\mu^l \to 0$ we easily obtain that $\lim_{l\to\infty} \| F(\tilde{x}^l, \tilde{y}^l, \tilde{z}^l) \| = 0$. Therefore,

$$\begin{aligned} \nabla f(\tilde{x}^*) - \tilde{z}^* - \nabla g(\tilde{x}^*)^T \tilde{y}^* &= 0 \\ g(\tilde{x}^*) &= 0 \\ \tilde{X}^* \tilde{Z}^* e &= 0. \end{aligned}$$

Clearly from the above equations we may derive that $(\tilde{x}^*, \tilde{y}^*, \tilde{z}^*)$ is an optimum point of the initial constrained optimisation problem (1). •

5 Numerical Results

The algorithm described in the previous sections has been implemented using standard C, on a Dual processor Sun UltraSparc-2, 167 MHz, with 256 megabytes of RAM, running Solaris (release 5.5.1). The test-problems were drawn from the Hock and Schittkowski collection [28]. For most of the problems we used the starting points recommended by in [28].

The various parameters used in Algorithms 1 and 2 are selected as follows: $c_{0,0} = 0$, $\eta = 1000$, $\beta_1 = 0.05$, $\beta_2 = 0.5$, $\rho = 10^{-3}$, $\gamma = 0.995$ $\theta = 0.95$ and $\epsilon_0 = 10^{-8}$. Furthermore, since the values $\Phi_k(0)$, $\Phi_k(\alpha_{x_k,i})$ and $\Phi_k'(0)$ are available, they can be used to provide a quadratic interpolation of the approximate merit function $\Phi_k(\alpha)$, with $\alpha \in [\beta_1 \alpha_{x_k,i}, \beta_2 \alpha_{x_k,i}]$. Thus the new step-size $\alpha_{x_k,i+1}$ is given by

$$\alpha_{x_k,i+1} = \max\{\, \beta_1,\ \min\{\, \beta_2,\ \beta_{x_k,i} \,\}\,\}\, \alpha_{x_k,i}$$

where

$$\beta_{x_k,i} = \frac{-\alpha_{x_k,i}\Phi_k'(0)}{2\left[\Phi_k(\alpha_{x_k,i}) - \Phi_k(0) - \alpha_{x_k,i}\Phi_k'(0)\right]}$$

Moreover, the matrices H_k, which approximate the Lagrangian Hessian, are calculated by using Powell's technique [18] for preserving positive definiteness. Hence H_k is updating according to

$$H_{k+1} = H_k - \frac{H_k p_k p_k^T H_k}{p_k^T H_k p_k} + \frac{r_k r_k^T}{p_k^T r_k}$$

where

$$\begin{aligned} p_k &= x_{k+1} - x_k \\ q_k &= \nabla L_B(x_{k+1}, y_{k+1}; \mu) - \nabla L_B(x_k, y_{k+1}; \mu) \\ r_k &= \omega_k q_k + (1-\omega_k) H_k p_k \end{aligned}$$

and

$$\omega_k = \begin{cases} 1, & \text{if } p_k^T q_k \geq 0.2 p_k^T H_k p_k \\ 0.8 p_k^T H_k p_k / (p_k^T H_k p_k - p_k^T r_k), & \text{if } p_k^T q_k < 0.2 p_k^T H_k p_k \end{cases}$$

The scalar ω_k is introduced to assure that $p_k^T r_k$ is positive for all k, so that positive definiteness of the sequence of matrices $\{H_k\}$ is satisfied. The identity matrix was used as the initial matrix of the sequence $\{H_k\}$.

In Algorithm 2, the value of the barrier parameter is reduced by a constant factor $\theta = 0.95$. This constant reduction is sufficient to guarantee the convergence of our method to an optimum solution of the initial problem (1). However, the efficiency of any primal-dual interior point algorithm heavily depends on the speed by which μ approaches zero.

To accelerate our algorithm we used a strategy where the barrier parameter does not decrease by a constant factor. The reduction strategy derives from two other strategies, presented by Lasdon *et al.* [12] and Gay *et al.* [7]. The basic characteristic of our strategy is that, it determines the new value of μ, by taking into consideration the distance of the current point (x_k, y_k, z_k) from the central path and the optimum solution of the initial problem. The barrier reduction strategy is shown in bellow

Barrier reduction strategy

If $\| F(x_k, y_k, z_k; \mu^l) \|_2 \leq \eta\mu^l$ or $k > 5$ then

  $\mu^{l+1} = \min\{0.95\mu^l, 0.01(0.95)^k \| F(x_k, y_k, z_k; 0) \|_2\}$

  If $\| F(x_k, y_k, z_k; \mu^l) \|_2 \leq 0.1\eta\mu^l$ then

    If $\mu^l < 10^{-4}$ then

      $\mu^{l+1} = \min\{0.85\mu^l, 0.01(0.85)^{k+2\sigma} \| F(x_k, y_k, z_k; 0) \|_2\}$

    Else

$$\mu^{l+1} = \min\{0.85\mu^l, 0.01(0.85)^{k+\sigma} \parallel F(x_k, y_k, z_k; 0) \parallel_2\}$$

The vectors $F(x_k, y_k, z_k; \mu^l)$ and $F(x_k, y_k, z_k; 0)$ represent the perturbed and unperturbed optimality conditions. The threshold determining whether the barrier parameter is going to decrease is initially checked. If the current point is close enough to the central path (i.e., if $\parallel F(x_k, y_k, z_k; \mu^l) \parallel_2 \leq 0.1\eta\mu^l$) and the optimum solution (i.e., if $\mu^l < 10^{-4}$), then the barrier parameter is reduced faster, since it is multiplied by the factor $(0.85)^{2\sigma}$, where $\sigma > 0$. If it is only close to the central path and not close to the optimum solution then the barrier parameter is still reduced but not as fast as before, since it is now multiplied by the larger factor $(0.85)^{\sigma}$. Hence, σ can be thought of as a parameter which accelerates the decrease of μ at appropriate points. In our numerical tests, this barrier reduction rule has performed very effectively. All the numerical results have been obtained by using the above strategy, with $\sigma = 5$.

Table 1 summarise the numerical results, where we use the following abbreviations

Prob: The problem number given in the Hock and Schittkowski collection [28].

Iter: The total number of iterations required to find an optimum solution of the initial constrained problem (1).

$c_{0,0}$: The initial value of the penalty parameter.

c_*: The final value of the penalty parameter.

k_*: The iteration after which the penalty parameter was unchanged.

The algorithm described in the previous sections solved successfully all the problems to the desired accuracy. For all the problems the starting points recommended in [28] were used. The behaviour of the penalty parameter was quite stable. In most of the problems its initial value was set to zero. Its final value remained relatively low and became constant in early iterations.

Finally it should be mentioned that the incorporation of the merit function into the primal-dual framework prevented the algorithm from converging to a local maximum or other stationary points. These results are quite encouraging and indicate that merit functions a very important role in the design of primal-dual interior point algorithms for general nonlinear programming problems.

Prob.	Iter.	$c_{0,0}$	c_*	k_*	Prob.	Iter.	$c_{0,0}$	c_*	k_*
5	8	0	0	1	73	7	0	8.2	1
10	14	0	305.4	10	76	8	0	22	3
11	8	0	99.16	4	83	7	0	0	1
12	11	0	0	1	84	16	0	0	1
14	10	10^7	2×10^7	9	93	14	0	7.2×10^9	9
22	9	0	0	1	95	14	0	0	1
24	9	0	0	1	96	13	0	0	1
27	20	0	5.7×10^4	15	97	20	0	0	1
32	14	0	0	1	98	17	0	0	1
33	12	0	0	1	100	18	0	4.2×10^5	11
34	10	0	1.2	6	104	16	0	9.8×10^4	14
35	7	0	27.03	2	105	36	0	0	1
43	15	0	17.5	4	108	18	0	32	4
57	15	1000	1000	1	110	8	0	0	1
59	13	0	0	1	112	23	0	0	1
64	18	0	2.6×10^{10}	16	113	12	0	23	3
65	10	0	0	1	114	17	0	3.5×10^3	6
66	11	0	2.9×10^3	8	117	22	200	1.4×10^8	17
71	10	0	0.7	2	118	12	0	0	1
72	11	0	6.6×10^9	8	119	19	0	4.7×10^3	4

Table 1: Numerical Results.

References

[1] Fiacco A.V. and McCormick G.P. *Nonlinear Programming: Sequential Unconstrained minimization Techniques.* John Wiley and Sons, New York, 1968.

[2] Rustem B. Equality and inequality constrained optimization algorithms with convergent stepsizes. *JOTA*, 76(3):429–453, 1993.

[3] McShane K.A. Monma C.L. and Shanno D.F. An implementation of a primal-dual interior point method for linear programming. *ORSA Journal on Computing*, 1(2):70–83, 1989.

[4] Bertsekas D. *Nonlinear Programming.* Athena Scientific Publishing, 1995.

[5] Potra F.A. An infeasible interior-point predictor-corrector algorithm for linear programming. Technical Report 26, Department of Mathematics, University of Iowa, Iowa City, IA, 1992.

[6] K. R. Frisch. The logarithmic potential method for convex programming. In stitute of Economics, University of Oslo, Oslo, Norway., 1955.

[7] Overton M.L. Gay D.M. and Wright M.H. A primal-dual interior method for non-convex nonlinesr programming. Technical Report 97-4-08, Bell Laboratories, Murray Hill, N.J., 1997.

[8] McCormick G.P. and Falk J.E. The superlinear convergence of a nonlinear primal-dual algorithm. Technical Report GWU/OR/Serial T-550/91, Department of Operations Research, The Ceorge Washington University, Washington, DC, 1991.

[9] Yamashita H. A globally convergent primal-dual interior point method for constrained optimization. Technical report, Mathematical Systems Institute Inc, 2-5-3 Shinjuku, Shinjuku-ku, Tokyo, Japan, May 1995.

[10] Yamashita H. and Yabe H. Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization. *Mathematical Programming*, 75:377–397, 1996.

[11] Akrotirianakis I. and Rustem B. A globally convergent interior point algorithm for general non-linear programming problems. Technical Report 97-14, Department of Computing, Imperial College, London, UK, November 1997.

[12] Lasdon L.S. Plummer J. and Gang Y. Primal-dual and primal interior point algorithms for general nonlinear programms. *ORSA J. on Computing*, 7(3):321–332, 1995.

[13] Vanderbei R. J. and Carpenter T. J. Symetric indefinite systems for interior point methods. *Mathematical Programming*, (58):1–32, 1991.

[14] Wright S. J. An interior point algorithm for linear constrained optimization. *SIAM Journal on Optimization*, 2:450–473, 1992.

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

[16] Vial J.P. Computational experience with a primal-dual interior point method for smooth convex programming. *Optimization Methods and Software*, 3:285–310, 1994.

[17] Anstreicher K.M. and Vial J.P. On the convergence of an infeasible primal-dual interior point method for convex programming. *Optimization Methods and Software*, 3:273–283, 1994.

[18] Powell M.J.D. A fast algorithm for nonlinearly constrained optimization calculations. In Watson G. A., editor, *Numerical Analysis Proceedings, Biennial Conference, Dundee*, Lecture Notes in Mathematics (630), pages 144–157, Berlin, 1978. Springer Verlag.

[19] Powell M.J.D. and Yuan Y. A recursive quadratic programming algorithm that uses differenciable exact penalty function. *Mathematical Programming*, 35:265–278, 1986.

[20] Karmarkar N.K. A new polynomial time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[21] Boggs P.T. and Tolle J.W. Sequential quadratic programming. *Acta Numerica*, 4, 1995.

[22] Fletcher R. A class of methods for nonlinear programming with termination and convergence properties. In J. Abadie, editor, *Integer and Nonlinear Programming.* North Holland, Amsterdam, 1970.

[23] Monteiro R. and Adler I. Interior path following primal-dual algorithms-part 2: Convex quadratic programming. *Mathematical Programming*, 44:43–66, 1989.

[24] Lusting I.J. Marsten R.E. and Shanno D.F. Computational experience with a primal-dual interior point method for linear programming. *Linear Algebra and Applications*, 152:191–222, 1991.

[25] Vanderbei R.J. and Shanno D.F. An interior point algorithm for nonconvex nonlinear programming. Technical Report SOR-97-21, CEOR, Princeton University, Princeton, NJ, 1997.

[26] Gajulapalli R.S. and Lasdon L.S. Computational experience with a safeguarded barrier algorithm for sparse nonlinear programming. Technical report, University of Texas at Austin, Austin, TX, 1997.

[27] El Bakry A.S. Tapia R.A. Tsuchiya T. and Zhang Y. On the formulation and theory of the newton interior point method for nonlinear programming. *JOTA*, 89(3):507–541, 1996.

[28] Hock W. and Schittkowski K. *Test Examples for Nonlinear Programming Codes.* Springer Verlag, New York, NY, 1981.

[29] Zhang Y. and Tapia R.A. A superlinearly convergent polynomial primal dual interior point algorithm for linear programming. *SIAM J. Optimization*, 3(1):118–133, February 1993.