Proceedings of the International Conference
on Computational and Mathematical Methods
in Science and Engineering, CMMSE 2008
13–17 June 2008.

A Three-D Filter Line Search Method within an Interior Point Framework

M. Fernanda P. Costa[1] and Edite M.G.P. Fernandes[2]

[1] *Mathematics for Science and Technology Department, University of Minho, 4800-058 Guimarães, Portugal*

[2] *Production and Systems Department, University of Minho, 4710-057 Braga, Portugal*

emails: mfc@mct.uminho.pt, emgpf@dps.uminho.pt

Abstract

Here we present a primal-dual interior point three-dimensional filter line search method for nonlinear programming. The three components of the filter aim to measure adequacy of feasibility, centrality and optimality of trial iterates. The algorithm also relies on a monotonic barrier parameter reduction and it includes a feasibility/centrality restoration phase. Numerical experiments with a set of well-known problems are carried out and a comparison with a previous implementation that differs on the optimality measure is presented.

Key words: Nonlinear optimization, interior point, filter method
MSC 2000: 90C51, 90C30

1 Introduction

The filter technique of Fletcher and Leyffer [6] is used to globalize the primal-dual interior point method for solving nonlinear constrained optimization problems. This technique incorporates the concept of nondominance to build a filter that is able to reject poor trial iterates and enforce global convergence from arbitrary starting points. The filter replaces the use of merit functions, avoiding therefore the update of penalty parameters that are associated with the penalization of the constraints in merit functions.

The filter technique has already been adapted to interior point methods. In Ulbrich, Ulbrich and Vicente [12], a filter trust-region strategy based on two components is proposed. The two components combine the three criteria of the first-order optimality conditions: the first component is a measure of quasi-centrality and the second is an optimality measure combining complementarity and criticality. Global convergence to first-order critical points is also proved. In [1, 14, 15, 16], a filter line search strategy that defines two components for each entry in the filter is used. The components are the barrier objective function and the constraints violation. The global convergence is analyzed in [14]. Numerical experiments with a three-dimensional filter based line search strategy are shown in [2, 3]. The three components of the filter measure feasibility, centrality and optimality and are present in the first-order KKT conditions of the barrier problem. The optimality measure relies on the norm of the gradient of the Lagrangian function. Convergence to stationary points has been proved, although convergence to a local minimizer is not guaranteed [4].

The algorithm herein presented is a primal-dual interior point method with a three-dimensional filter line search approach that considers the barrier objective function as the optimality measure. The algorithm also incorporates a restoration phase that aims to improve either feasibility or centrality. In the paper, a performance evaluation is also carried out using a benchmarking tool, known as performance profiles [5], to compare different practical details.

The paper is organized as follows. Section 2 briefly describes the interior point method and Section 3 is devoted to introduce the 3-D filter line search method. Section 4 describes the numerical experiments that were carried out in order to analyze the performance of the new algorithm and to compare its behavior with a previous implementation that differs on the optimality measure. Conclusions are made in Section 5.

2 The interior point method

The formulation of the nonlinear constrained optimization problem that is considered in the paper is the following:

$$\begin{array}{ll} \min_{x\in \mathbb{R}^n} & F(x) \\ \text{s.t.} & h(x) \geq 0 \end{array} \tag{1}$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ and $F : \mathbb{R}^n \to \mathbb{R}$ are nonlinear and twice continuously differentiable functions.

In this interior point paradigm, problem (1) is transformed into an equality constrained problem by using nonnegative slack variables w, as follows:

$$\begin{array}{rl} \min_{x\in \mathbb{R}^n, w\in \mathbb{R}^m} & \varphi_\mu(x,w) \equiv F(x) - \mu \sum_{i=1}^{m} \log(w_i) \\ \text{s.t.} & h(x) - w = 0 \\ & w \geq 0, \end{array} \tag{2}$$

where $\varphi_\mu(x,w)$ is the barrier function and μ is a positive barrier parameter [11, 13]. This is the barrier problem associated with (1). Under acceptable assumptions, the sequence of solutions of the barrier problem converges to the solution of the problem (1) when $\mu \searrow 0$. Thus, primal-dual interior point methods aim to solve a sequence of barrier problems for a positive decreasing sequence of μ values. The first-order KKT conditions for a minimum of (2) define a nonlinear system of $n+2m$ equations in $n+2m$

unknowns

$$\left\{\begin{array}{l}\nabla F(x)-A^T y=0\\-\mu W^{-1}e+y=0\\h(x)-w=0\end{array}\right. \quad (3)$$

where ∇F is the gradient vector of F, A is the Jacobian matrix of the constraints h, y is the vector of dual variables, $W = diag(w_i)$ is a diagonal matrix, and e is an m vector of all ones. Applying the Newton's method to solve (3), the following system, after symmetrization, appears

$$\left[\begin{array}{ccc}-H & 0 & A^T\\0 & -\mu W^{-2} & -I\\A & -I & 0\end{array}\right]\left[\begin{array}{c}\triangle x\\\triangle w\\\triangle y\end{array}\right]=\left[\begin{array}{c}\sigma\\-\gamma_\mu\\\rho\end{array}\right] \quad (4)$$

where

$$H=\nabla^2 F\left(x\right)-\sum_{i=1}^{m} y_i \nabla^2 h_i(x)$$

is the Hessian matrix of the Lagrangian function ($\mathcal{L} = \varphi_\mu(x,w) - y^T(h(x)-w)$) and

$$\sigma=\nabla_x\mathcal{L}=\nabla F(x)-A^T y,\ \ \gamma_\mu=\mu W^{-1}e-y\ \text{ and }\ \rho=w-h(x).$$

Since the second equation in (4) can be used to eliminate Δw without producing any off-diagonal fill-in in the remaining system, one obtains

$$\Delta w=\mu^{-1}W^2\left(\gamma_\mu-\Delta y\right), \quad (5)$$

and the resulting reduced KKT system

$$\left[\begin{array}{cc}-H & A^T\\A & \mu^{-1}W^2\end{array}\right]\left[\begin{array}{c}\Delta x\\\Delta y\end{array}\right]=\left[\begin{array}{c}\sigma\\\pi\end{array}\right] \quad (6)$$

where $\pi = \rho + \mu^{-1}W^2\gamma_\mu$, to compute the search directions Δx, Δw, Δy. Given initial approximations to the primal, slack and dual variables x_0, $w_0 > 0$ and $y_0 > 0$, this interior point method implements a line search procedure that chooses iteratively a step size α_k, at each iteration, and defines a new approximation by

$$\begin{array}{l}x_{k+1}=x_k+\alpha_k\Delta x_k\\w_{k+1}=w_k+\alpha_k\Delta w_k\\y_{k+1}=y_k+\alpha_k\Delta y_k.\end{array}$$

The choice of the step size α_k is a very important issue in nonconvex optimization and in the interior point context aims:

1. to ensure the nonnegativity of the slack and dual variables;

2. to enforce progress towards feasibility, centrality and optimality.

To decide which trial step size is accepted, at each iteration, a backtracking line search framework combined with a three-D filter method is used. This is the subject of the next section.

3 Three-D filter line search method

The methodology of a filter as outline in [6] is adapted to this interior point method. We use a three-dimensional filter. In the sequel, we use the vectors:

$$u = (x, w, y),\ u^1 = (x, w),\ u^2 = (w, y),$$
$$\Delta = (\Delta x, \Delta w, \Delta y),\ \Delta^1 = (\Delta x, \Delta w),\ \Delta^2 = (\Delta w, \Delta y).$$

To define the three components of the filter, we make use of the first-order optimality conditions (3) and the barrier objective function. The first component of the filter measures feasibility, the second measures centrality and the third represents optimality, and they are defined as follows:

$$\theta_f(u^1) = \|\rho\|_2\,,\ \theta_c(u^2) = \|\gamma_\mu\|_2 \text{ and } \varphi_\mu(u^1).$$

We remark that our previous work [2, 3] considered the norm of the gradient of the Lagrangian function in the optimality measure, therein denoted by $\theta_{op} = \frac{1}{2}\|\nabla_x \mathcal{L}\|^2$. While promoting convergence to stationary points [4], the algorithm did not enforce a sufficient decrease in the barrier function. Nonetheless, the practical implementation of the algorithm has shown convergence to minimizers even when saddle points and maximizers are present.

At each iteration k, a backtracking line search framework generates a decreasing sequence of step sizes

$$\alpha_{k,l} \in (0, \alpha_k^{\max}]\,, l = 0, 1, ...,$$

with $\lim_l \alpha_{k,l} = 0$, until a set of acceptance conditions are satisfied. Here, l denotes the iteration counter for the inner loop. $\alpha_k^{\max}$ is the longest step size that can be taken along the search directions to ensure the nonnegativity condition $u_k^2 \geq 0$. Assuming that the initial approximation satisfies $u_0^2 > 0$, the maximal step size $\alpha_k^{\max} \in (0, 1]$ is defined by

$$\alpha_k^{\max} = \max\{\alpha \in (0, 1] : u_k^2 + \alpha\Delta_k^2 \geq (1 - \varepsilon)u_k^2\} \tag{7}$$

for a fixed parameter $\varepsilon \in (0, 1)$.

In this interior point context, the trial iterate $u_k(\alpha_{k,l}) = u_k + \alpha_{k,l}\Delta_k$ is acceptable by the filter, if it leads to sufficient progress in one of the three measures compared to the current iterate,

$$\begin{gathered}\theta_f(u_k^1(\alpha_{k,l})) \leq \left(1 - \gamma_{\theta_f}\right)\theta_f(u_k^1) \quad \text{or} \quad \theta_c(u_k^2(\alpha_{k,l})) \leq (1 - \gamma_{\theta_c})\,\theta_c(u_k^2) \\ \text{or} \quad \varphi_\mu(u_k^1(\alpha_{k,l})) \leq \varphi_\mu(u_k^1) - \gamma_\varphi\theta_f(u_k^1)\end{gathered} \tag{8}$$

where $\gamma_{\theta_f}, \gamma_{\theta_c}, \gamma_\varphi \in (0, 1)$ are fixed constants. However, to prevent convergence to a feasible but nonoptimal point, and whenever for the trial step size $\alpha_{k,l}$, the following switching conditions

$$\begin{gathered}m_k(\alpha_{k,l}) < 0 \quad \text{and} \quad [-m_k(\alpha_{k,l})]^{s_o}\,[\alpha_{k,l}]^{1-s_o} > \delta\left[\theta_f(u_k^1)\right]^{s_f} \\ \text{and} \quad [-m_k(\alpha_{k,l})]^{s_o}\,[\alpha_{k,l}]^{1-s_o} > \delta\left[\theta_c(u_k^2)\right]^{s_c}\end{gathered} \tag{9}$$

hold, with fixed constants $\delta > 0$, $s_f > 1, s_c > 1$, $s_o \geq 1$, where

$$m_k(\alpha) = \alpha \nabla \varphi_\mu(u_k^1)^T \Delta_k^1,$$

then the trial iterate must satisfy the Armijo condition

$$\varphi_\mu(u_k^1(\alpha_{k,l})) \leq \varphi_\mu(u_k^1) + \eta_o m_k(\alpha_{k,l}), \tag{10}$$

instead of (8), to be acceptable. Here, $\eta_o \in (0, 0.5)$ is a constant. A trial step size $\alpha_{k,l}$ is called a φ-step if (10) holds. Similarly, if a φ-step is accepted as the final step size α_k in iteration k, then k is referred to as a φ-type iteration (see also [14]).

To prevent cycling between iterates that improve either the feasibility, or the centrality, or the optimality, at each iteration k, the algorithm maintains a filter that is a set $\overline{F}_k$ that contains values of θ_f, θ_c and φ_μ, that are prohibited for a successful trial iterate in iteration k [12, 14, 15, 16]. Thus, a trial iterate $u_k(\alpha_{k,l})$ is acceptable, if

$$\left(\theta_f(u_k^1(\alpha_{k,l})), \theta_c(u_k^2(\alpha_{k,l})), \varphi_\mu(u_k^1(\alpha_{k,l}))\right) \notin \overline{F}_k.$$

The filter is initialized to

$$\overline{F}_0 \subseteq \left\{(\theta_f, \theta_c, \varphi_\mu) \in \mathbb{R}^3 : \theta_f \geq \theta_f^{\max}, \theta_c \geq \theta_c^{\max}, \varphi_\mu \geq \varphi_\mu^{\max}\right\}, \tag{11}$$

for the nonnegative constants $\theta_f^{\max}$, $\theta_c^{\max}$ and $\varphi_\mu^{\max}$; and is updated whenever the accepted step size satisfies (8) by

$$\begin{aligned}\overline{F}_{k+1} = \overline{F}_k \cup \{(\theta_f, \theta_c, \varphi_\mu) \in \mathbb{R}^3 :\ & \theta_f \geq \left(1 - \gamma_{\theta_f}\right) \theta_f(u_k^1) \text{ and } \theta_c \geq \left(1 - \gamma_{\theta_c}\right) \theta_c(u_k^2) \\ & \text{and } \varphi_\mu \geq \varphi_\mu(u_k^1) - \gamma_\varphi \theta_f(u_k^1)\}.\end{aligned} \tag{12}$$

We remark that the filter remains unchanged whenever (9) and (10) hold for the accepted step size.

Finally, when the backtracking line search cannot find a trial step size $\alpha_{k,l}$ that satisfies the above criteria, we define a minimum desired step size $\alpha_k^{\min}$, using linear models of the involved functions,

$$\alpha_k^{\min} = \gamma_\alpha \begin{cases} \min\left\{\gamma_{\theta_f}, \frac{\gamma_\varphi \theta_f(u_k^1)}{-m_k(\alpha_{k,l})}, \frac{\delta\left[\theta_f(u_k^1)\right]^{s_f}}{\left[-m_k(\alpha_{k,l})\right]^{s_o}}, \frac{\delta\left[\theta_c(u_k^2)\right]^{s_c}}{\left[-m_k(\alpha_{k,l})\right]^{s_o}}\right\}, & \text{if } m_k(\alpha_{k,l}) < 0 \\ & \text{and } (\theta_f(u_k^1) \leq \theta_f^{\min} \text{ or } \theta_c(u_k^2) \leq \theta_c^{\min}) \\ \min\left\{\gamma_{\theta_f}, \frac{\gamma_\varphi \theta_f(u_k^1)}{-m_k(\alpha_{k,l})}\right\}, & \text{if } m_k(\alpha_{k,l}) < 0 \\ & \text{and } (\theta_f(u_k^1) > \theta_f^{\min} \text{ and } \theta_c(u_k^2) > \theta_c^{\min}) \\ \gamma_{\theta_f}, & \text{otherwise} \end{cases} \tag{13}$$

for positive constants $\theta_f^{\min}, \theta_c^{\min}$ and a safety factor $\gamma_\alpha \in (0, 1]$. Whenever the backtracking line search finds a trial step size $\alpha_{k,l} < \alpha_k^{\min}$, the algorithm reverts to a restoration phase. Here, the algorithm tries to find a new iterate u_{k+1} that is acceptable to the current filter, *i.e.*, (8) holds, by reducing either the constraints violation or the centrality within an iterative process.

3.1 Restoration phase

The task of the restoration phase is to compute a new iterate acceptable to the filter by decreasing either the feasibility or the centrality, whenever the backtracking line search procedure cannot make sufficient progress and the step size becomes too small. Thus, the restoration algorithm works with the new functions

$$\theta_{2,f}(u^1) = \frac{1}{2}\|\rho\|_2^2 \text{ and } \theta_{2,c}(u^2) = \frac{1}{2}\|\gamma_\mu\|_2^2$$

and the steps Δ^1 and Δ^2 that are descent directions for $\theta_{2,f}(u^1)$ and $\theta_{2,c}(u^2)$, respectively (as shown in Theorem 2 below).

3.2 Descent properties

While the search directions are computed from solving the reduced KKT system (6), we need for subsequent analysis the explicit formulas for Δx and Δw. Let $N(u) = H + \mu A^T W^{-2} A$ denote the dual normal matrix.

Theorem 1 *If N is nonsingular, then (4) has a unique solution. In particular,*

$$\begin{aligned}\Delta x &= -N^{-1}\nabla F(x) + \mu N^{-1}A^T W^{-1}e + \mu N^{-1}A^T W^{-2}\rho \\ \Delta w &= -AN^{-1}\nabla F(x) + \mu AN^{-1}A^T W^{-1}e - \left(I - \mu AN^{-1}A^T W^{-2}\right)\rho.\end{aligned}$$

Proof. Solving the second block of equations in (6) for Δy and eliminating Δy from first block of equations yields a system involving only Δx whose solution is

$$\begin{aligned}\Delta x &= N^{-1}\left(-\sigma + A^T(\mu W^{-2}\rho + \gamma_\mu)\right) \\ &= N^{-1}\left(-\nabla F(x) + A^T y + A^T(\mu W^{-2}\rho + \mu W^{-1}e - y)\right) \\ &= -N^{-1}\nabla F(x) + N^{-1}A^T y + \mu N^{-1}A^T W^{-2}\rho + \mu N^{-1}A^T W^{-1}e - N^{-1}A^T y \\ &= -N^{-1}\nabla F(x) + \mu N^{-1}A^T W^{-2}\rho + \mu N^{-1}A^T W^{-1}e\end{aligned}$$

where we used the definitions of σ and γ_μ. Using this formula of Δx, we can then solve for Δy and finally for Δw.

The resulting formula for Δw is:

$$\begin{aligned}\Delta w &= \mu^{-1}W^2(\gamma_\mu - \Delta y) \\ &= \mu^{-1}W^2\gamma_\mu - \mu^{-1}W^2\left(\mu W^{-2}\rho + \gamma_\mu - \mu W^{-2}A\Delta x\right) \\ &= \mu^{-1}W^2\gamma_\mu - \rho - \mu^{-1}W^2\gamma_\mu + A\Delta x \\ &= -\rho + AN^{-1}\left(-\sigma + A^T(\mu W^{-2}\rho + \gamma_\mu)\right) \\ &= -\rho - AN^{-1}\sigma + AN^{-1}A^T(\mu W^{-2}\rho + \gamma_\mu) \\ &= -\rho + \mu AN^{-1}A^T W^{-2}\rho - AN^{-1}\sigma + AN^{-1}A^T\gamma_\mu \\ &= -\left(I - \mu AN^{-1}A^T W^{-2}\right)\rho - AN^{-1}\left(\nabla F(x) - A^T y\right) + AN^{-1}A^T\left(\mu W^{-1}e - y\right) \\ &= -AN^{-1}\nabla F(x) + \mu AN^{-1}A^T W^{-1}e - \left(I - \mu AN^{-1}A^T W^{-2}\right)\rho.\end{aligned}$$

■

Theorem 2 *The search directions have the following properties: (i) If the dual matrix N is positive definite and $\rho = 0$, then*

$$\nabla \varphi_\mu^T \Delta^1 \leq 0.$$

ii) Furthermore

$$\nabla \theta_{2,f}^T \Delta^1 \leq 0 \text{ and } \nabla \theta_{2,c}^T \Delta^2 \leq 0.$$

Proof. First we prove (i). It is easy to see that $\nabla_x \varphi_\mu = \nabla F$ and $\nabla_w \varphi_\mu = -\mu W^{-1} e$. Let $y = \mu W^{-1} e$ and $\sigma = \nabla F - A^T y$. From the expressions for Δx and Δw given in Theorem 1, and assuming that $\rho = 0$, we get

$$\begin{aligned}
\begin{pmatrix} \nabla F \\ -y \end{pmatrix}^T \begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix} &= \nabla F^T \Delta x - y^T \Delta w \\
&= \nabla F^T \left(-N^{-1} \nabla F(x) + \mu N^{-1} A^T W^{-1} e\right) - \\
&\quad - y^T \left(-A N^{-1} \nabla F(x) + \mu A N^{-1} A^T W^{-1} e\right) \\
&= \nabla F^T \left(-N^{-1} \nabla F(x) + N^{-1} A^T y\right) - \\
&\quad - y^T \left(-A N^{-1} \nabla F(x) + A N^{-1} A^T y\right) \\
&= \nabla F^T \left(-N^{-1} (\nabla F(x) - A^T y)\right) - \\
&\quad - y^T A \left(-N^{-1} (\nabla F(x) - A^T y)\right) \\
&= \nabla F^T \left(-N^{-1} \sigma\right) - y^T A (-N^{-1} \sigma) \\
&= \left(\nabla F^T - y^T A\right) \left(-N^{-1} \sigma\right) \\
&= -\sigma^T N^{-1} \sigma \leq 0,
\end{aligned}$$

which completes the proof of the first property. To prove (ii), we star by addressing the the feasibility measure $\theta_{2,f}$. It is easy to see that $\nabla_x \theta_{2,f} = -A^T \rho$ and $\nabla_w \theta_{2,f} = \rho$, and from (4) we get

$$\begin{aligned}
\begin{pmatrix} \nabla_x \theta_{2,f} \\ \nabla_w \theta_{2,f} \end{pmatrix}^T \begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix} &= \begin{pmatrix} -A^T \rho \\ \rho \end{pmatrix}^T \begin{pmatrix} \Delta x \\ \Delta w \end{pmatrix} \\
&= \left(-\rho^T A\right) \Delta x + \left(\rho^T\right) \Delta w \\
&= -\rho^T (A \Delta x - \Delta w) \\
&= -\rho^T \rho \leq 0.
\end{aligned}$$

We now address the centrality measure $\theta_{2,c}$. It is easy to see that $\nabla_w \theta_{2,c} = -\mu W^{-2} \gamma_\mu$ and $\nabla_y \theta_{2,c} = -\gamma_\mu$, and from (4) we get

$$\begin{aligned}
\begin{pmatrix} \nabla_w \theta_{2,c} \\ \nabla_y \theta_{2,c} \end{pmatrix}^T \begin{pmatrix} \Delta w \\ \Delta y \end{pmatrix} &= \begin{pmatrix} -\mu W^{-2} \gamma_\mu \\ -\gamma_\mu \end{pmatrix}^T \begin{pmatrix} \Delta w \\ \Delta y \end{pmatrix} \\
&= \gamma_\mu^T (-\mu W^{-2}) \Delta x - \gamma_\mu^T \Delta y \\
&= \gamma_\mu^T \left(-\mu W^{-2} \Delta x - \Delta y\right) \\
&= \gamma_\mu^T \left(-\gamma_\mu\right) \\
&= -\gamma_\mu^T \gamma_\mu \leq 0.
\end{aligned}$$

■

3.3 The algorithm

Next, we present the proposed primal-dual interior point 3-D filter line search algorithm for solving constrained optimization problems.

Algorithm 1 (Interior Point 3-D Filter Line Search Algorithm)

1. Given*: Starting point* x_0*,* $u_0^2 > 0$*;*
   constants $\theta_f^{\max} \in (\theta_f(u_0^1), \infty]$*;* $\theta_f^{\min} \in (0, \theta_f(u_0^1)]$*;* $\theta_c^{\max} \in (\theta_c(u_0^2), \infty]$*;* $\theta_c^{\min} \in (0, \theta_c(u_0^2)]$*;* $\varphi_\mu^{\max} \in (\varphi_\mu(u_0^1), \infty]$*;* $\gamma_{\theta_f}, \gamma_{\theta_c}, \gamma_\varphi \in (0,1)$*;* $\delta > 0$*;* $s_f > 1$*;* $s_c > 1$*;* $s_o \geq 1$*;* $\eta_o, \eta_{\theta_{2,f}}, \eta_{\theta_{2,c}} \in (0, 0.5]$*;* $\varepsilon_{tol} \ll 1$*;* $\varepsilon \in (0,1)$*;* $\delta_\mu, \kappa_\mu \in [0,1)$*;* $\epsilon \in (0,1)$*;*
   compute $\mu_0 > 0$ *using (14).*
2. *Initialize the filter using (11) and set* $k \leftarrow 0$*.*
3. *Stop if termination criterion is satisfied (see (15)).*
4. *If* $k \neq 0$ *compute* μ_k *using (14).*
5. *Compute the search direction* Δ_k *from the linear system (6), and (5).*
6. *6.1 Compute the longest step size* $\alpha_k^{\max}$ *using (7) to ensure positivity of slack and dual variables. Set* $\alpha_{k,l} = \alpha_k^{\max}$*,* $l \leftarrow 0$*.*
   6.2 If $\alpha_{k,l} < \alpha_k^{\min}$*, go to restoration phase in step 10. Otherwise, compute the trial iterate* $u_k(\alpha_{k,l})$*.*
   6.3 If $\left(\theta_f(u_k^1(\alpha_{k,l})), \theta_c(u_k^2(\alpha_{k,l})), \varphi_\mu(u_k^1(\alpha_{k,l}))\right) \in \overline{F}_k$*, reject the trial step size and go to step 6.6.*
   6.4 If $\alpha_{k,l}$ *is a* φ*-step size ((9) holds) and the Armijo condition (10) for the* φ_μ *function holds, accept the trial step and go to step 7.*
   6.5 If (8) holds, accept the trial step and go to step 7. Otherwise go to step 6.6.
   6.6 Set $\alpha_{k,l+1} = \alpha_{k,l}/2$*,* $l \leftarrow l+1$*, and go back to step 6.2.*
7. *Set* $\alpha_k \leftarrow \alpha_{k,l}$ *and* $u_{k+1} \leftarrow u_k(\alpha_k)$*.*
8. *If* k *is not a* φ*-type iteration, augment the filter using (12). Otherwise, leave the filter unchanged.*
9. *Set* $k \leftarrow k+1$ *and go back to step 3.*
10. *Use the* Restoration Algorithm *to produce a point* u_{k+1} *that is acceptable to the filter, i.e.,* $\left(\theta_f(u_{k+1}^1), \theta_c(u_{k+1}^2), \varphi_\mu(u_{k+1}^1)\right) \notin \overline{F}_k$*. Augment the filter using (12) and continue with the regular iteration in step 9.*

In the restoration phase, a sufficient reduction in one of the measures $\theta_{2,f}$ and $\theta_{2,c}$ is required for a trial step size to be acceptable. The Restoration Algorithm is as follows.

Algorithm 2 (Restoration Algorithm)

1. *Set* $\alpha_{k,0}^{\max} = \alpha_k^{\max}$, $u_{k,0} = u_k$, $l = 0$ *and start with step 5.*

2. *If* $u_{k,l}$ *is acceptable to the filter then set* $u_{k+1} = u_{k,l}$ *and stop.*

3. *Compute* $\Delta_{k,l}$ *from the linear system (6), and (5) (with* $u_k = u_{k,l}$*)*

4. *(Define the vectors* $\Delta_{k,l}^1$, $\Delta_{k,l}^2$ *which are used as search directions for the variables* $u_{k,l}^1$, $u_{k,l}^2$*.) Compute* $\alpha_{k,l}^{\max}$

5. *Set* $\alpha_k = \alpha_{k,l}^{\max}$.

6. *Compute the trial iterate* $u_{k,l}(\alpha_k)$,
   If $\theta_{2,f}(u_{k,l}^1(\alpha_k)) \leq \theta_{2,f}(u_{k,l}^1) + \alpha_k \eta_{\theta_{2,f}} \nabla \theta_{2,f}(u_{k,l}^1)^T \Delta_{k,l}^1$ *or*
   $\theta_{2,c}(u_{k,l}^2(\alpha_k)) \leq \theta_{2,c}(u_{k,l}^2) + \alpha_k \eta_{\theta_{2,c}} \nabla \theta_{2,c}(u_{k,l}^2)^T \Delta_{k,l}^2$
   then set $u_{k,l+1} = u_{k,l}(\alpha_k)$, $l = l + 1$, *and return to step 2. Otherwise* $\alpha_k \leftarrow \alpha_k/2$, *and repeat step 6.*

4 Numerical experiments

To analyze the performance of the proposed interior point 3-D filter line search method, as well as to compare with our previous implementation of the algorithm [2], we used 111 constrained problems from the Hock and Schittkowski test set [8]. The tests were done in double precision arithmetic with a Pentium 4. The algorithm is coded in the C programming language and includes an interface to AMPL to read the problems that are coded in the AMPL modeling language [7].

4.1 Implementation details

Next, we report some computational details that were undertaken during our numerical experimentation, such as, for example, the initialization of the variables, the barrier parameter evaluation and the termination criterion.

Initial quasi-Newton approximation Our algorithm is a quasi-Newton based method in the sense that a symmetric positive definite quasi-Newton BFGS approximation, B_k, is used to approximate the Hessian of the Lagrangian H, at each iteration k [9]. In the first iteration, we may set $B_0 = I$ or $B_0 =$ positive definite modification of $\nabla^2 F(x_0)$, depending on the characteristics of the problem to be solved.

Monotonic reduction of the barrier parameter To guarantee a positive decreasing sequence of μ values, the barrier parameter is updated by a formula that couples the theoretical requirement defined on the first-order KKT conditions (3) with a simple heuristic. Thus, μ is updated by

$$\mu_{k+1} = \max\left\{\epsilon, \min\left\{\kappa_\mu \mu_k, \delta_\mu \frac{w_{k+1}^T y_{k+1}}{m}\right\}\right\} \tag{14}$$

where the constants $\kappa_\mu, \delta_\mu \in (0,1)$ and the tolerance ϵ is used to prevent μ from becoming too small so avoiding numerical difficulties at the end of the iterative process.

Reevaluation of centrality and optimality measures in the filter We further remark that each time the barrier parameter is updated, the θ_c component, as well as the barrier objective function value, of points in the filter may be reevaluated using the new μ so that a fair comparison of the current point with points in the filter is made. In practice, only $\theta_c^{\max}$ and $\varphi_\mu^{\max}$ need to be reevaluated.

Initialization of variables Two possible ways to initialize the primal and dual variables consider:

1. the usual published initial x_0, and the dual variables are initialized to one;
2. the published x_0 to define the dual variables and modified primal variables by solving the simplified reduced system:

$$\begin{bmatrix} -(B_0+I) & A^T(x_0) \\ A(x_0) & I \end{bmatrix} \begin{bmatrix} \tilde{x}_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} \nabla F(x_0) \\ 0 \end{bmatrix}.$$

   Further, if $\|y_0\|_\infty > 10^3$ then we set (component-wise) $y_0 = 1$. Similarly, if $\|\tilde{x}_0\|_\infty > 10^3 \|x_0\|_\infty$, we set $\tilde{x}_0 = x_0$.

The nonnegativity of the initial slack variables are ensured by computing $w_0 = \max\{|h(x_0)|, \epsilon_w\}$, for the previously defined x_0, and a fixed positive constant ϵ_w.

Termination criterion The termination criterion considers dual and primal feasibility and centrality measures

$$\max\left\{\frac{\|\sigma\|_\infty}{s}, \|\rho\|_\infty, \frac{\|\gamma_\mu\|_\infty}{s}\right\} \le \varepsilon_{tol}, \tag{15}$$

where

$$s = \max\left\{1, 0.01\frac{\|y\|_1}{m}\right\}$$

and $\varepsilon_{tol} > 0$ is the error tolerance.

4.2 Parameter settings

The chosen values for the parameters involved in the Algorithms 1 and 2 are:
$\theta_f^{\max} = 10^4 \max\left\{1, \theta_f(u_0^1)\right\}$, $\theta_f^{\min} = 10^{-4} \max\left\{1, \theta_f(u_0^1)\right\}$, $\theta_c^{\max} = 10^4 \max\left\{1, \theta_c(u_0^2)\right\}$, $\theta_c^{\min} = 10^{-4} \max\left\{1, \theta_c(u_0^2)\right\}$, $\varphi_\mu^{\max} = 10^4 \max\left\{0, \varphi_\mu(u_0^1)\right\}$, $\gamma_{\theta_f} = \gamma_{\theta_c} = \gamma_\varphi = 10^{-5}$,

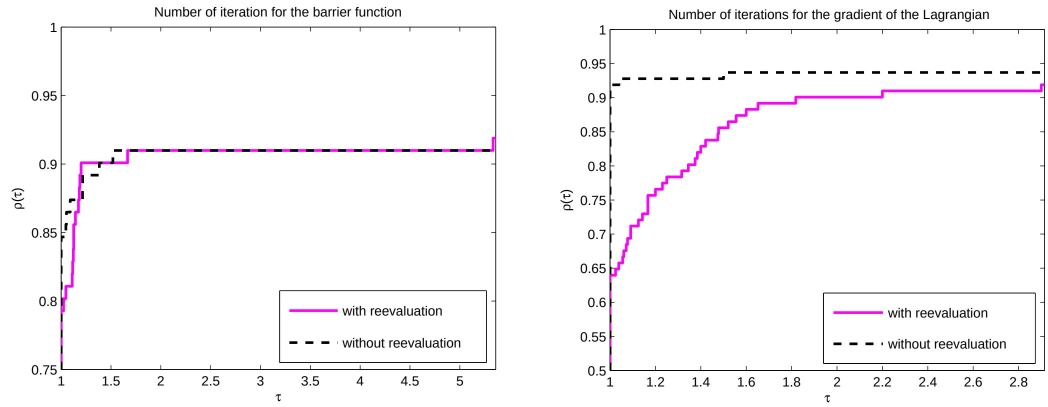

Figure 1: Profiles on the number of iterations: using φ_μ (on the left); using θ_{op} (on the right)

$\delta = 1$, $s_f = 1.1$, $s_c = 1.1$, $s_o = 2.3$, $\eta_o = \eta_{\theta_{2,f}} = \eta_{\theta_{2,c}} = 10^{-4}$, $\varepsilon = 0.95$, , $\delta_\mu = \kappa_\mu = 0.1$, $\epsilon = 10^{-9}$, $\epsilon_w = 0.01$ and $\varepsilon_{tol} = 10^{-4}$.

We carried out a set of experiments considering the two alternatives for setting the initial B_0 (as previously described) and the two ways of primal and dual variables initialization. For the subsequent analysis and comparisons we combined the overall results for each problem and selected the one which yields the smallest number of iterations.

4.3 Dolan-Moré performance profiles

To compare the performance of the reevaluation of the centrality and optimality measures in the filter for each updated μ value, we use the performance profiles as outline in [5]. These profiles represent the cumulative distribution function of a performance ratio, computed from a predefined metric. For this analysis we choose the number of iterations required to achieve the desired accuracy, as reported in (15). A brief explanation of the Dolan-Moré performance profiles follows.

Let $\mathcal{P}$ be the set of problems and $\mathcal{C}$ the set of codes used in the comparative study. Let $t_{p,c}$ be the performance metric - number of iterations required to solve problem p by code c. Then, the comparison relies on the performance ratios

$$r_{p,c} = \frac{t_{p,c}}{\min\{t_{p,c}, c \in \mathcal{C}\}}, p \in \mathcal{P}, c \in \mathcal{C}$$

and the overall assessment of the performance of a code c is given by $\rho_c(\tau) = \frac{n_{P_\tau}}{n_P}$, where n_P is the number of problems in the set $\mathcal{P}$ and n_{P_τ} is the number of problems in the set such that the performance ratio $r_{p,c}$ is less than or equal to $\tau \in \mathbb{R}$ for code $c \in \mathcal{C}$. Thus, $\rho_c(\tau)$ gives the probability (for code c) that $r_{p,c}$ is within a factor τ of the best possible ratio. The function ρ_c is the cumulative distribution function for the performance ratio.

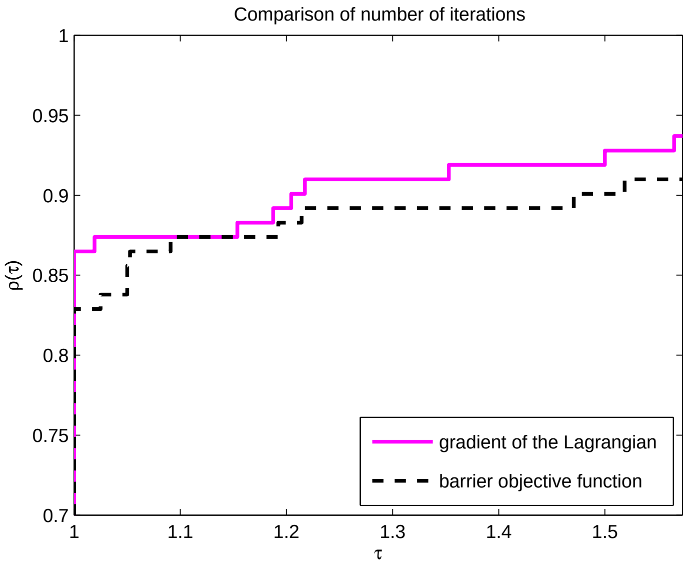

Figure 2: Profiles on the number of iterations: comparison between θ_{op} and φ_μ

First, we examine the practical performance of the reevaluation of the filter for each new μ value within the herein proposed algorithm - where φ_μ is used as the optimality measure. The performance plots on the left of Figure 1 show that the version that does not implement the reevaluation of the filter is the most efficient on 85% of the problems (see the corresponding value of $\rho(1)$). Next, the filter reevaluation process was implemented in our previous implementation of the algorithm - when θ_{op} was used as the optimality measure. The performance plots on the right of Figure 1 definitely show that the filter reevaluation yields the worst performance.

From the previous analysis, we decided to disable the reevaluation filter process from both algorithms and plot the performance profiles together. Figure 2 represents the performance profiles of the number of iterations. The use of the barrier function to measure the trial iterate optimality adequacy did not improve the performance of this interior point based method, at least when the number of iterations is the metric used in these performance profiles.

Finally, to further compare the convergence of both interior point 3-D filter line search algorithms we include Table 1 that records the objective function values at the found solutions. Only the problems that were solved at least by one of the versions in comparison are listed. While the previous implementation did not converge to the required solution on 3 problems (hs046, hs105, hs111), within 100 iterations, the new algorithm did not reach the solution on the following problems: hs064, hs083, hs101, hs106 and hs118. In all the other problems, both algorithms reach the same solution with the desired accuracy.

Table 1: Objective function values at the solution

Prob	with θ_{op}	with φ_μ	Prob	with θ_{op}	with φ_μ	Prob	with θ_{op}	with φ_μ
hs001	8.9525e-13	6.5934e-12	hs038	1.0359e-11	6.6804e-14	hs077	2.4151e-01	2.4151e-01
hs002	5.0426e-02	5.0426e-02	hs039	-1.0000e00	-1.0000e00	hs078	-2.9197e00	-2.9197e00
hs003	1.0000e-04	1.0000e-04	hs040	-2.5000e-01	-2.5000e-01	hs079	7.8777e-02	7.8777e-02
hs004	2.6667e00	2.6667e00	hs041	1.9999e00	1.9999e00	hs080	5.3949e-02	5.3949e-02
hs005	-1.9132e00	-1.9132e00	hs042	1.3858e01	1.3858e01	hs081	5.3950e-02	5.3950e-02
hs006	1.6997e-12	1.6997e-12	hs043	-4.4000e01	-4.4000e01	hs083	-3.0666e04	-
hs007	-1.7321e00	-1.7321e00	hs044	-1.5000e01	-1.5000e01	hs086	-3.2349e01	-3.2349e01
hs008	-1.0000e00	-1.0000e00	hs045	1.0000e00	1.0000e00	hs087	8.8276e03	8.8276e03
hs009	-4.9999e-01	-4.9999e-01	hs046	2.1089e-02	6.2703e-07	hs088	1.3626e00	1.3626e00
hs010	-1.0000e00	-1.0000e00	hs047	1.1658e-07	1.1658e-07	hs089	1.3626e00	1.3626e00
hs011	-8.4985e00	-8.4985e00	hs048	1.6555e-12	1.6555e-12	hs090	1.3626e00	1.3626e00
hs012	-3.0000e01	-3.0000e01	hs049	1.17559e-06	1.1756e-06	hs091	1.3626e00	1.3626e00
hs014	1.3934e00	1.3934e00	hs050	3.1993e-10	3.1993e-10	hs092	1.3627e00	1.3627e00
hs015	3.0650e02	3.0650e02	hs051	3.5090e-10	3.5090e-10	hs093	1.3508e02	1.3508e02
hs016	2.5000e-01	2.5000e-01	hs052	5.3266e00	5.3266e00	hs095	1.5620e-02	1.5620e-02
hs017	1.0000e00	1.0000e00	hs053	4.0930e00	4.0930e00	hs096	1.5620e-02	1.5620e-02
hs018	4.9999e00	5.0000e00	hs054	1.9286e-01	1.9286e-01	hs097	3.1358e00	3.1358e00
hs019	-6.9618e03	-6.9618e03	hs055	6.6667e00	6.6667e00	hs098	4.0712e00	4.0712e00
hs020	3.8199e01	3.8199e01	hs056	-1.0788e-10	-1.0788e-10	hs100	6.8063e02	6.8063e02
hs021	-9.9960e01	-9.9960e01	hs057	3.0648e-02	3.0648e-02	hs101	1.8098e03	-
hs022	1.0000e00	1.0000e00	hs059	-7.8028e00	-7.8028e00	hs102	9.1188e02	9.1188e02
hs023	2.0000e00	2.0000e00	hs060	3.2568e-02	3.2568e-02	hs103	5.4367e02	5.4367e02
hs024	-1.0000e00	-1.0000e00	hs061	-1.4365e02	-1.4365e02	hs104	3.9512e00	3.9512e00
hs025	1.8361e-10	2.7269e-11	hs062	-2.6273e04	-2.6273e04	hs105	-	1.1363e03
hs026	7.6064e-07	1.9872e-07	hs063	9.6172e02	9.6172e02	hs106	7.0492e03	-
hs027	3.9999e-02	3.9999e-02	hs064	6.2998e03	-	hs108	-5.0000e-01	-5.0000e-01
hs028	1.0270e-09	1.0270e-09	hs065	9.5354e-01	9.5354e-01	hs110	-4.5778e01	-4.5778e01
hs029	-2.2627e01	-2.2627e01	hs066	5.1816e-01	5.1816e-01	hs111	-	-4.7761e01
hs030	1.0002e00	1.0002e00	hs067	-1.1620e03	-1.1620e03	hs112	-4.7761e01	-4.7761e01
hs031	5.9999e00	6.0000e00	hs070	2.7971e-01	2.7971e-01	hs113	2.4306e01	2.4306e01
hs032	1.0000e00	1.0000e00	hs071	1.7014e01	1.7014e01	hs114	-1.7688e03	-1.7688e03
hs033	-4.5858e00	-4.5858e00	hs072	7.2760e02	7.2767e02	hs117	3.2349e01	3.2349e01
hs034	-8.3403e-01	-8.3403e-01	hs073	2.9894e01	2.9894e01	hs118	6.6482e02	-
hs035	1.1116e-01	1.1116e-01	hs074	5.1265e03	5.1265e03	hs119	2.4490e02	2.4490e02
hs036	-3.3000e03	-3.3000e03	hs075	5.1744e03	5.1744e03			
hs037	-3.4560e03	-3.4560e03	hs076	-4.6818e00	-4.6818e00			

5 Conclusions

A primal-dual interior point method based on a filter line search approach is presented. The novelty here is that each entry in the filter has three components that represent the feasibility, centrality and optimality of the iterate. Using the barrier objective function as the optimality measure, the algorithm is able to enforce a sufficient decrease of the barrier function and converge to stationary points that are minimizers. The new algorithm is tested with a set of well-known problems and compared with our previous implementation of an interior point three-dimensional filter line search [2, 3], using a benchmarking tool with performance profiles. The numerical results show that both algorithms have similar practical behaviors.

We would like to remark that the performance profiles reflect only the performance of the tested codes on the data being used. Definitive conclusions could be made if different test sets, including larger academic problems and real engineering problems [10], were used. This will be a matter of future research.

References

[1] H.Y. Benson, R.J. Vanderbei and D.F. Shanno, *Interior-point methods for nonconvex nonlinear programming: filter methods and merit functions*, Computational Optimization and Applications, **23** (2002) 257–272.

[2] M.F.P. Costa and E.M.G.P. Fernandes, *Comparison of interior point filter line search strategies for constrained optimization by performance profiles*, International Journal of Mathematics Models and Methods in Applied Sciences, **1** (2007) 111–116.

[3] M.F.P. Costa and E.M.G.P. Fernandes, *Practical implementation of an interior point nonmonotone line search filter method*, International Journal of Computer Mathematics **85** (2008) 397–409.

[4] M.F.P. Costa and E.M.G.P. Fernandes, *A globally convergent interior point filter line search method for nonlinear programming*, submitted to Journal of Numerical Analysis, Industrial and Applied Mathematics 2007.

[5] E.D. Dolan and J.J. Moré, *Benchmarking optimization software with performance porfiles*, Mathematical Programming A **91** (2002) 201–213.

[6] R. Fletcher and S. Leyffer, *Nonlinear programming without a penalty function*, Mathematical Programming **91** (2002) 239–269.

[7] R. Fourer, D.M. Gay and B. Kernighan, *A modeling language for mathematical programming*, Management Science **36** (1990) 519–554.

[8] W. Hock and K. Schittkowski, *Test Examples for Nonlinear Programming*, Springer-Verlag, 1981.

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

[10] K. E. Parsopoulos and M. N. Vrahatis, *Unified particle swarm optimization for solving constrained engineering optimization problems*, Vol 3612 of LNCS, 582–591. Springer-Verlag, ICNC 2005 edition, 2005.

[11] D.F. Shanno and R.J. Vanderbei, *Interior-point methods for nonconvex nonlinear programming: orderings and higher-order methods*, Mathematical Programming B **87** (2000) 303–316.

[12] M. Ulbrich, S. Ulbrich and L.N. Vicente, *A globally convergent primal-dual interior-point filter method for nonlinear programming*, Mathematical Programming **100** (2004) 379–410.

[13] R.J. Vanderbei and D.F. Shanno, *An interior-point algorithm for nonconvex nonlinear programming*, Computational Optimization and Applications **13** (1999) 231–252.

[14] A. Wächter and L.T. Biegler, *Line search filter methods for nonlinear programming: motivation and global convergence*, SIAM Journal on Optimization **16** (2005) 1–31.

[15] A. Wächter and L.T. Biegler, *Line search filter methods for nonlinear programming: local convergence*, SIAM Journal on Optimization **16** (2005) 32–48.

[16] A. Wächter and L.T. Biegler, *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming*, Mathematical Programming **106** (2007) 25–57.