

Research Article

An Improved Nonmonotone Filter Trust Region Method for Equality Constrained Optimization

Zhong Jin

Department of Mathematics, Shanghai Maritime University, Shanghai 201306, China

Correspondence should be addressed to Zhong Jin; asdjz1234@163.com

Received 22 October 2012; Accepted 11 January 2013

Academic Editor: Nikolaos Papageorgiou

Copyright © 2013 Zhong Jin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Motivated by the method of Su and Pu (2009), we present an improved nonmonotone filter trust region algorithm for solving nonlinear equality constrained optimization. In our algorithm a modified nonmonotone filter technique is proposed and the restoration phase is not needed. At every iteration, in common with the composite-step SQP methods, the step is viewed as the sum of two distinct components, a quasinormal step and a tangential step. A more relaxed accepted condition for trial step is given and a crucial criterion is weakened. Under some suitable conditions, the global convergence is established. In the end, numerical results show our method is effective.

1. Introduction

We consider the problem of minimizing a nonlinear function subject to a set of nonlinear equality constraints as follows:

$$\begin{aligned} (P) \quad & \min f(x) \\ & \text{s.t. } c_i(x) = 0, \quad i \in I = \{1, 2, \dots, m\}, \end{aligned} \quad (1)$$

where the objective function $f: R^n \rightarrow R$ and the equality constraints $c_i (i \in I): R^n \rightarrow R$ are all twice continuously differentiable.

The algorithm that we discuss belongs to the class of trust region methods and, more specifically, to that filter methods suggested first by Fletcher and Leyffer [1], in which the use of a penalty function, a common feature of the large majority of the algorithms for constrained optimization, is replaced by the technique so-called “filter.” Subsequently, global convergences of the trust region filter SQP methods were established by Fletcher et al. [2, 3], and Ulbrich [4] got its superlinear local convergence. Especially, the step framework in [3] is similar in spirit to the composite-step SQP methods pioneered by Vardi [5], Byrd et al. [6], and Omojokun [7]. Consequently, filter method has been actually applied in many optimization techniques, for instance, the pattern search method [8], the SLP method [9], the interior method [10], the bundle approaches [11, 12], the system of nonlinear equations and

nonlinear least squares [13], multidimensional filter method [14], line search filter methods [15, 16], and so on.

In fact, filter method exhibits a certain degree of nonmonotonicity. The idea of nonmonotone technique can be traced back to Grippo et al. [17] in 1986, combined with the line search strategy. Due to its excellent numerical exhibition, over the last decades, the nonmonotone technique has been used in trust region method to deal with unconstrained and constrained optimization problems [18–25]. More recently, the nonmonotone trust region method without penalty function has also been developed for constrained optimization [26–29]. Especially, in [29] the nonmonotone idea is employed to the filter technique and restoration phase, a common feature of the large majority of filter methods, is not needed.

Based on [29], motivated by above ideas and methods, in this paper we present an improved filter algorithm that combines the nonmonotone and trust region techniques for solving nonlinear equality constrained optimization. Our method improves previous results and gives a more flexible mechanism, weakens some needed conditions, and has the merits as follows.

- (i) An improved nonmonotone filter technique is presented. Filter technique is viewed as a biojective

optimization problem that minimizes objective function f and violation function h . In [29], the same nonmonotone measure is utilized for both f and h at the k th iteration. In the paper we improve it and define a new measure for h , which may make the discussion and consideration of the nonmonotone properties of f and h more freely.

- (ii) The restoration phase is not needed. The restoration procedure has to be considered in most of filter methods. We employ the nonmonotone idea to the filter technique, so as [29] the restoration phase is not needed.
- (iii) A more relaxed accepted condition for trial step is considered. Compared to general trust region methods, in [29] the condition for the accepted trial step is relaxed. In the paper we improve it and then the accepted condition for trial step is more relaxed.
- (iv) A crucial criterion is weakened in the algorithm. By introducing a new parameter $\eta \in (0, 1]$, we improve the crucial criterion for implementation of the method in [29]. It is obvious that our criterion becomes same with [29] if $\eta = 1$, but new criterion can be easier to satisfy, so this crucial criterion has been weakened when setting $\eta < 1$ in the initialization of our algorithm.

The presentation is organized as follows. Section 2 introduces some preliminary results and improvements on filter technique and some conditions. Section 3 develops a modified nonmonotone filter trust region algorithm, whose global convergence is shown in Section 4. The results of numerical experience with the proposed method are discussed in the last section.

2. Preliminary and Improvements

In this section, we first recall some definitions and preliminary results about the techniques of composite-step SQP type trust region method and fraction of Cauchy decrease condition. And then the improvements on filter technique and some conditions are given.

2.1. Fraction of Cauchy Decrease Condition. To introduce the corresponding results, we consider the unconstrained optimization problem: $\min\{f(x) \mid x \in R^n\}$, where $f : R^n \rightarrow R$ is continuously differentiable. At iteration point x_k , it obtains a trial step d_k by solving the following quadratic program subproblem:

$$\begin{aligned} \min \quad & q(d) = \nabla f(x_k)d + \frac{1}{2}d^T H_k d \\ \text{s.t.} \quad & \|d\| \leq \Delta_k, \end{aligned} \quad (2)$$

where H_k is a symmetric matrix which is either the Hessian matrix of f at x_k or an approximation to it, and $\Delta_k > 0$ is a trust region radius.

To assure the global convergence, the step is only required to satisfy a fraction of Cauchy decrease condition. It means d

must predict via the quadratic model function $q(d)$ at least as much as a fraction of the decreased given by the Cauchy step on $q(d)$, that is, there exists a constant $\sigma > 0$ fixed across all iterations, such that

$$q(0) - q(d) \geq \sigma (q(0) - q(d^{\text{cp}})), \quad (3)$$

where d^{cp} is the steepest descent step for $q(d)$ inside the trust region. And we have the following lemma.

Lemma 1. *If the trial step d satisfies a fraction of Cauchy decrease condition, then*

$$q(0) - q(d) \geq \frac{\sigma}{2} \|\nabla f(x)\| \min \left\{ \Delta, \frac{\|\nabla f(x)\|}{\|H\|} \right\}. \quad (4)$$

Proof. See Powell [30] for the proof. \square

2.2. The Subproblems of Composite-Step SQP Type Trust Region Method. In the composite-step SQP type trust region methods, at current iterate x_k we obtain the trial step $d_k = d_k^n + d_k^t$ by computing a quasi-normal step d_k^n and a tangential step d_k^t . The purpose of the quasi-normal step d_k^n is to improve feasibility and d_k^t in the tangential space of the linearized constraints can provide sufficient decrease for a quadratic model of the objective function $f(x)$ to improve optimality.

For problem (P), d_k^n is the solution to the subproblem

$$\begin{aligned} \min \quad & \frac{1}{2} \|c_k + A_k^T d^n\|^2 \\ \text{s.t.} \quad & \|d^n\| \leq \Delta_k, \end{aligned} \quad (5)$$

where Δ_k is a trust region radius,

$$c(x) = \begin{pmatrix} c_1(x) \\ c_2(x) \\ \vdots \\ c_m(x) \end{pmatrix}, \quad (6)$$

$A(x) = \nabla c(x)$, $c_k = c(x_k)$ and $A_k = A(x_k) \in R^{n \times m}$. In order to improve the value of the objective function, d_k^t can be obtained by the following subproblem

$$\begin{aligned} \min \quad & q_k(d_k^n + d_k^t) \\ \text{s.t.} \quad & A_k^T d_k^t = 0 \\ & \|d_k^t\| \leq \Delta_k, \end{aligned} \quad (7)$$

where $q_k(d) = g_k^T d + (1/2)d^T H_k d$ and $g_k = \nabla f(x_k)$. Then we get the current trial step $d_k = d_k^n + d_k^t$.

In usual ways that impose a trust region in step-decomposition methods, the quasi-normal step d_k^n and the tangential step d_k^t are required to satisfy

$$\|d_k^n\| \leq 0.8\Delta_k, \quad \|d_k^n + d_k^t\| \leq \Delta_k. \quad (8)$$

Here, to simplify the proof, we only impose a trust region on $\|d_k^n\| \leq \Delta_k$ and $\|d_k^t\| \leq \Delta_k$, which is natural.

2.3. *The Improved Accepted Condition for d_k .* Borrowed from the usual trust region idea, we also need to define the following predicted reduction for the violation function $h(x) = \|c(x)\|^2$

$$\text{pred}_k^c = h(x_k) - \|c_k + A_k^T d_k\|^2 \tag{9}$$

and the actual reduction

$$\text{ared}_k^c = h(x_k) - h(x_k + d_k) = \|c_k\|^2 - \|c(x_k + d_k)\|^2. \tag{10}$$

To evaluate the descent properties of the step for the objective function, we use the predicted reduction of $f(x)$

$$\text{pred}_k^f = q_k(0) - q_k(d_k) = -q_k(d_k) \tag{11}$$

and the actual reduction of $f(x)$

$$\text{ared}_k^f = f(x_k) - f(x_k + d_k). \tag{12}$$

In general trust region method, the step d_k will be accepted if

$$\text{ared}_k^f \geq \rho \text{pred}_k^f, \tag{13}$$

where $\rho \in (0, 1)$ is a fixed constant.

In [29], considering nonmonotone technique, the condition (13) is replaced by

$$\text{rared}_k^f \geq \rho \text{pred}_k^f, \tag{14}$$

where rared_k^f is the relaxed actual reduction of $f(x)$, that is,

$$\text{rared}_k^f = \max \left\{ f(x_k), \sum_{r=0}^{m(k)-1} \lambda_{kr} f(x_{k-r}) \right\} - f(x_k + d_k). \tag{15}$$

But in this paper, we consider more relaxed accepted condition by increasing ared_k^f to rared_k^f and reducing pred_k^f to $\min\{\text{pred}_k^f, \text{pred}_k^c + \xi_0\}$ simultaneously, then the condition (14) is replaced by

$$\text{rared}_k^f \geq \rho \min \{ \text{pred}_k^f, \text{pred}_k^c + \xi_0 \}, \tag{16}$$

where ξ_0 is a small positive number.

2.4. *The Improved Nonmonotone Filter Technique.* In order to obtain next iterate, it needs to determine the step; this procedure that decides which trial step is accepted is “filter method.” For the optimization problem with equality constraints

$$\min \{ f(x) \mid c(x) = 0 \}, \tag{17}$$

a promising trial step should either reduce the constraint violation h or the objective function value f . Since $h(x) = \|c(x)\|^2$, it is easy to see that $h(x) = 0$ if and only if x is

a feasible point. So in the traditional filter method, a point x is called acceptance to the filter if and only if

$$h(x) \leq \beta h_j \quad \text{or} \quad f(x) \leq f_j - \gamma h_j \quad \forall (h_j, f_j) \in \mathcal{F}, \tag{18}$$

where $0 < \gamma < \beta < 1$, \mathcal{F} denotes the filter set.

Different from above criteria of filter idea, with nonmonotone technique, in [29] a point x is called to be acceptable to the filter if and only if

$$\begin{aligned} &\text{either } h(x) \leq \beta \max_{0 \leq r \leq m(k)-1} h_{k-r} \\ &\text{or } f(x) \leq \max \left[f_k, \sum_{r=0}^{m(k)-1} \lambda_{kr} f_{k-r} \right] - \gamma h(x), \end{aligned} \tag{19}$$

where $(h_{k-r}, f_{k-r}) \in \mathcal{F}_k$ for $0 \leq r \leq m(k) - 1$, and $1 \leq m(k) \leq \min\{m(k) - 1 + 1, M\}$, $M \geq 1$ is a given positive integer, $\sum_{r=0}^{m(k)-1} \lambda_{kr} = 1$, $\lambda_{kr} \in (0, 1)$, and there exists a positive constant λ such that $\lambda_{kr} \geq \lambda$.

Observe that $m(k)$ is used in both conditions (19), while in this paper we wish to reduce the relationship of the nonmonotone properties of f and h . We consider the nonmonotone properties of f and h , respectively, and call a trial point x is acceptable to the filter if and only if

$$\begin{aligned} &\text{either } h(x) \leq \beta \max_{0 \leq j \leq m(p)-1} \theta_{M+p-j} \\ &\text{or } f(x) \leq \max \left[f_k, \sum_{r=0}^{m(k)-1} \lambda_{kr} f_{k-r} \right] - \gamma h(x), \end{aligned} \tag{20}$$

where θ_{M+p-j} comes from $\{\theta_i\}_{i=1}^{+\infty}$ such that $\theta_1 = \theta_2 = \dots = \theta_{M-1} = 0$ and $\{\theta_i\}_{i=M}^{+\infty}$ is a subsequence of $\{h(x_k)\}$, and $1 \leq m(p) \leq \min\{m(p) - 1 + 1, M\}$.

Similar to the traditional filter methods, if (20) is satisfied, it is called an h -type iteration and the filter set \mathcal{F}_k needs to be updated at each successful h -type iteration.

2.5. *The Weakened Criterion for Implementation of Algorithm.*

We replace the crucial criterion $\text{pred}_k^f \geq \text{pred}_k^c$ in [29] by

$$\frac{\text{pred}_k^f}{\text{pred}_k^c} \geq \eta, \tag{21}$$

where $\eta \in (0, 1]$. It is obvious that this new criterion becomes the same with [29] if $\eta = 1$, but it can be easier to satisfy, so the criterion has been weakened when setting $\eta < 1$ in the initialization of our algorithm.

3. Description of the Algorithm

It will be convenient to introduce the reduced gradient $\hat{g}(x) = W(x)^T g(x)$ where $g(x) = \nabla f(x)$ and $W(x)$ denotes a matrix whose columns are from a basis of the null space of $A(x)^T$. The first order necessary optimality conditions (Karush-Kuhn-Tucher or KKT conditions) at a local solution $\bar{x} \in R^n$ of (P) can be written as

$$c(\bar{x}) = 0, \quad \hat{g}(\bar{x}) = 0. \tag{22}$$

For brevity, at the current iterate x_k , we use $g_k, c_k, A_k, f_k, h_k, W_k, H_k$ to denote $\nabla f(x_k), c(x_k), A(x_k), f(x_k), h(x_k), W(x_k), H(x_k)$. We are now in a position to present a formal statement of our algorithm.

Algorithm 2. Step 0. Initialization: choose an initial guess x_0 , a symmetric matrix $H_0 \in R^{n \times n}$, an initial region radius $\Delta_0 \geq \Delta_{\min} > 0$, $0 < \gamma < \beta < 1$, $0 < \lambda \leq 1$ and $\mathcal{F}_0 = \{(h_0, f_0)\}$. Set $\eta \in (0, 1]$, $\xi_0 > 0$, $0 < \rho < 1$, $0 < \gamma_0 < \gamma_1 < 1 < \gamma_2$ and $M > 0$. Let $m(0) = 1$, $\theta_1 = \theta_2 = \dots = \theta_{M-1} = 0$, $\theta_M = h(x_0)$, $p = 0$ and $k = 0$.

Step 1. Compute $f_k, g_k, c_k, h_k, A_k, W_k$. If $\|\hat{g}_k\| + h_k = 0$, then stop.

Step 2. Compute (5) and (7) to obtain d_k^n and d_k^t , and set $d_k = d_k^n + d_k^t$.

Step 3. If $x_k + d_k$ is acceptable to the filter, go to Step 4, otherwise go to Step 5.

Step 4. If $\text{rared}_k^f < \rho \min\{\text{pred}_k^f, \text{pred}_k^c + \xi_0\}$ and $\text{pred}_k^f / \text{pred}_k^c \geq \eta$, then go to Step 5, otherwise go to Step 6.

Step 5. $\Delta_k \in [\gamma_0 \Delta_k, \gamma_1 \Delta_k]$, go to Step 2.

Step 6. $x_{k+1} = x_k + d_k$, $\Delta_{k+1} \in [\Delta_k, \gamma_2 \Delta_k] \geq \Delta_{\min}$. Update H_k to H_{k+1} , $m(k+1) = \min\{m(k)+1, M\}$. If (20) holds, update the filter set; let $\theta_{M+p+1} = h(x_{k+1})$, $m(p+1) = \min\{m(p)+1, M\}$ and $p = p + 1$. Let $k = k + 1$ and go to Step 1.

Remark 3. At the beginning of each iteration, we always set $\Delta_k \geq \Delta_{\min}$, which will avoid too small trust region radius.

Remark 4. If h -type iteration is satisfied, then θ_{M+p+1} is generated, so if the number of h -type iterations is infinite, we obtain a sequence $\{\theta_i\}_{i=1}^{+\infty}$ such that $\theta_1 = \theta_2 = \dots = \theta_{M-1} = 0$ and $\{\theta_i\}_{i=M}^{+\infty}$ is a subsequence of $\{h(x_k)\}$. Specially, (20) always holds on the whole obtained sequence $\{\theta_i\}_{i=1}^{+\infty}$.

Remark 5. In the above algorithm, let M be a positive integer. For each k , let $m(k)$ satisfying

$$m(0) = 1, \quad 1 \leq m(k) \leq \min\{m(k-1) + 1, M\}, \quad (23)$$

for $k \geq 1$,

and we also find $m(p)$ satisfy

$$m(0) = 1, \quad 1 \leq m(p) \leq \min\{m(p-1) + 1, M\}, \quad (24)$$

for $p \geq 1$,

so the nonmonotonicity is showed as $M > 1$.

4. Global Convergence of Algorithm

In this section, we will prove the global convergence properties of Algorithm 2 under the following assumptions.

(A₁) The objective $f(x)$ and the constraint functions $c_i(x)$, $i = 1, 2, \dots, m$ are all twice continuously differentiable on an open set containing X .

(A₂) All points x_k that are sampled by algorithm lie in a nonempty closed and bounded set X .

(A₃) The matrix sequence $\{H_k\}$ is uniformly bounded.

(A₄) $A(x) = \nabla c(x)$, $(A^T(x)A(x))^{-1}$, $W(x)$ and $(W^T(x)W(x))^{-1}$ are uniformly bounded on X .

From the above assumptions, it is easier to suppose that there exist five constants v_1, v_2, v_3, v_4, v_5 such that $\|f(x)\| \leq v_1$, $\|\nabla f(x)\| \leq v_1$, $\|\nabla^2 f(x)\| \leq v_1$, $\|c(x)\| \leq v_2$, $\|A(x)\| \leq v_2$, $\|\nabla^2 c(x)\| \leq v_2$, $\|(A^T(x)A(x))^{-1}\| \leq v_3$, $\|W(x)\| \leq v_4$, $\|W^T(x)H(x)\| \leq v_5$; this can make our following results convenient to prove.

Lemma 6. *At the current iterate x_k , let the trial point component d_k^n actually be normal to the tangential space. Under the above assumptions, there exists a constant $\alpha_1 > 0$ independent of the iterates such that*

$$\|d_k^n\| \leq \alpha_1 \|c_k\|. \quad (25)$$

Proof. Since d_k^n is actually normal to the tangential space, it follows that

$$\begin{aligned} \|d_k^n\| &= \|A_k(A_k^T A_k)^{-1} A_k^T d_k\| \\ &= \|A_k(A_k^T A_k)^{-1} (c_k + A_k^T d_k - c_k)\| \\ &= \|A_k(A_k^T A_k)^{-1}\| + (\|c_k + A_k^T d_k\| + \|c_k\|), \end{aligned} \quad (26)$$

together with the fact that $\|c_k + A_k^T d_k\| \leq \|c_k\|$, we have

$$\|d_k^n\| \leq 2 \|A_k(A_k^T A_k)^{-1}\| \|c_k\| \triangleq \alpha_1 \|c_k\|. \quad (27)$$

□

Lemma 7. *Suppose that the assumptions hold, there exist positive α_2, α_3 independent of the iterates such that*

$$\|c_k\|^2 - \|c_k + A_k^T d_k\|^2 \geq \alpha_2 \|c_k\| \min\{\|c_k\|, \Delta_k\}, \quad (28)$$

$$\begin{aligned} q_k(d_k^n) - q_k(d_k) \\ \geq \alpha_3 \|W_k^T \nabla q_k(d_k^n)\| \min\{\|W_k^T \nabla q_k(d_k^n)\|, \Delta_k\}. \end{aligned} \quad (29)$$

Proof. The statement is established by an application of Lemma 1 to two subproblems (5) and (7). □

Lemma 8. *Under problem assumptions, Algorithm 2 is well defined.*

Proof. We will show that there exists $\delta > 0$ such that step d_k is accepted whenever $\Delta_k \leq \delta$. In fact,

$$\begin{aligned} & \left| \text{ared}_k^f - \text{pred}_k^f \right| \\ &= \left\| f(x_k) - f(x_k + d_k) + g_k^T d_k + \frac{1}{2} d_k^T H_k d_k \right\| \\ &= \left\| -g_k^T d_k - \frac{1}{2} d_k^T \nabla^2 f(y_k) d_k + g_k^T d_k + \frac{1}{2} d_k^T H_k d_k \right\| \\ &= 4n \Delta_k^2 \frac{1}{2} \left\| \nabla^2 f(y_k) - H_k \right\| \leq 4n \Delta_k^2 b, \end{aligned} \quad (30)$$

where $y_k = x_k + \xi d_k$, $\xi \in (0, 1)$ denotes some point on the line segment from x_k to $x_k + d_k$, $b = (1/2)(\sup \|H_k\| + \max_{x \in S} \|\nabla^2 f(x)\|)$, $\|d_k^f\| \leq \sqrt{n}\Delta_k$, $\|d_k^n\| \leq \sqrt{n}\Delta_k$ and $\|d_k\| \leq 2\sqrt{n}\Delta_k$. We consider two cases in the following.

Case 1 ($h(x_k) \neq 0$). To prove the implementation of Algorithm 2, we need only to show that if $\text{pred}_k^f/\text{pred}_k^c \geq \eta > 0$, it follows $\text{ared}_k^f \geq \rho \min\{\text{pred}_k^f, \text{pred}_k^c + \xi_0\}$. Without loss of generality, we can assume that $\|c_k\| \geq \epsilon$. Then we start with $\delta \in (0, \epsilon]$ such that the closed δ -ball about x_k lies in S . It is obvious from $\delta \leq \epsilon$ that $\Delta_k \leq \delta \leq \epsilon$. And from (28), we have $\text{pred}_k^c \geq \alpha_2 \epsilon \Delta_k$. Observe that

$$\begin{aligned} \left| \frac{\text{ared}_k^f - \text{pred}_k^f}{\text{pred}_k^f} \right| \cdot \left| \frac{\text{pred}_k^f}{\text{pred}_k^c} \right| &= \left| \frac{\text{ared}_k^f - \text{pred}_k^f}{\text{pred}_k^c} \right| \\ &\leq \frac{4n\Delta_k^2 b}{\alpha_2 \epsilon \Delta_k} \rightarrow 0 \quad \text{as } \Delta_k \rightarrow 0, \\ \frac{\text{pred}_k^f}{\text{pred}_k^c} &\geq \eta > 0. \end{aligned} \tag{31}$$

Hence we have

$$\left| \frac{\text{ared}_k^f - \text{pred}_k^f}{\text{pred}_k^f} \right| \rightarrow 0 \quad \text{as } \Delta_k \rightarrow 0, \tag{32}$$

which implies $\text{ared}_k^f \geq \rho \text{pred}_k^f \geq 0$, then $\text{ared}_k^f \geq \text{ared}_k^f \geq \rho \text{pred}_k^f \geq \rho \min\{\text{pred}_k^f, \text{pred}_k^c + \xi_0\}$ for some $\rho \in (0, 1)$.

Case 2 ($h(x_k) = 0$). If $\|\hat{g}_k\| = 0$, then x_k is a KKT-point of (1) by Algorithm 2. Thus, we assume that there exists a constant ϵ such that $\|\hat{g}_k\| \geq \epsilon$. Then

$$\begin{aligned} \|W_k^T \nabla q_k(d_k^n)\| &= \|W_k^T (g_k + H_k d_k^n)\| \\ &\geq \|W_k^T g_k\| - \|W_k^T H_k d_k^n\| \\ &\geq \epsilon - v_5 \Delta_k. \end{aligned} \tag{33}$$

Next, if we reduce δ such that $\delta \leq (1/2v_5)\epsilon$, then for all $\Delta_k \leq \delta$, we have $\Delta_k \leq (1/2v_5)\epsilon$. Therefore

$$\|W_k^T \nabla q_k(d_k^n)\| \geq \frac{\epsilon}{2}. \tag{34}$$

And it holds from $h(x_k) = 0$ that $d_k^n = 0$, then, combining with (29), we find that

$$\text{pred}_k^f = -q_k(d_k) \geq \frac{\alpha}{2} \epsilon \min\left\{\frac{\epsilon}{2}, \Delta_k\right\} = \frac{\alpha}{2} \epsilon \Delta_k. \tag{35}$$

Hence,

$$\left| \frac{\text{ared}_k^f - \text{pred}_k^f}{\text{pred}_k^f} \right| \leq \frac{4n\Delta_k^2 b}{(\alpha/2)\epsilon \Delta_k} \rightarrow 0 \quad \text{as } \Delta_k \rightarrow 0, \tag{36}$$

which also implies $\text{ared}_k^f \geq \text{ared}_k^f \geq \rho \text{pred}_k^f \geq \rho \min\{\text{pred}_k^f, \text{pred}_k^c + \xi_0\}$ for some $\rho \in (0, 1)$. Thus, the trial step is accepted for all $\Delta_k \leq \delta$. \square

Lemma 8 has provided the implementation of Algorithm 2. By mechanism of Algorithm 2, it is obvious that there exists a constant $\bar{\Delta} > 0$, such that $\Delta_k \geq \bar{\Delta}$ for sufficiently large k .

In the remainder of this paper we denote the set of indices of those iterations in which the filter has been augmented by \mathcal{A} , that is to say, if the k th iteration is h -type iteration, we have $k + 1 \in \mathcal{A}$. By this definition, if (20) holds for infinite iterations, then $|\mathcal{A}| = \infty$; otherwise, we can say $|\mathcal{A}| \neq \infty$.

Lemma 9. *Let $\{x_k\}$ be an infinite sequence generated by Algorithm 2. If $|\mathcal{A}| \neq \infty$, then $\lim_{k \rightarrow \infty} h(x_k) = 0$.*

Proof. By mechanism of Algorithm 2, we can assume that for all k , it follows

$$f(x_{k+1}) \leq \max\left\{f(x_k), \sum_{r=0}^{m(k)-1} \lambda_{kr} f(x_{k-r})\right\} - \gamma h(x_k). \tag{37}$$

Then we first show that for all k , it holds

$$\begin{aligned} f(x_k) &\leq f(x_0) - \lambda\gamma \sum_{r=0}^{k-2} h(x_r) - \gamma h(x_{k-1}) \\ &\leq f(x_0) - \lambda\gamma \sum_{r=0}^{k-1} h(x_r). \end{aligned} \tag{38}$$

Next we prove (38) by induction.

If $k = 1$, we have $f(x_1) \leq f(x_0) - \gamma h(x_0) \leq f(x_0) - \lambda\gamma h(x_0)$.

Assume that (38) holds for $1, 2, \dots, k$, then we consider that (38) holds for $k + 1$ in the following two cases.

Case 1. If $\max\{f(x_k), \sum_{r=0}^{m(k)-1} \lambda_{kr} f(x_{k-r})\} = f(x_k)$, then

$$\begin{aligned} f(x_{k+1}) &\leq f(x_k) - \gamma h(x_k) \leq f(x_0) - \lambda\gamma \sum_{r=0}^{k-1} h(x_r) - \gamma h(x_k) \\ &\leq f(x_0) - \lambda\gamma \sum_{r=0}^k h(x_r). \end{aligned} \tag{39}$$

Case 2. If $\max\{f(x_k), \sum_{r=0}^{m(k)-1} \lambda_{kr} f(x_{k-r})\} = \sum_{r=0}^{m(k)-1} \lambda_{kr} f(x_{k-r})$, let $q = m(k) - 1$, then

$$\begin{aligned} f(x_{k+1}) &\leq \sum_{t=0}^q \lambda_{kt} f(x_{k-t}) - \gamma h(x_k) \\ &\leq \sum_{t=0}^q \lambda_{kt} \left(f(x_0) - \lambda\gamma \sum_{r=0}^{k-t-2} h(x_r) \right. \\ &\quad \left. - \gamma h(x_{k-t-1}) \right) - \gamma h(x_k) \end{aligned}$$

$$\begin{aligned}
&= \lambda_{k_0} \left(f(x_0) - \lambda\gamma \sum_{r=0}^{k-q-2} h(x_r) - \lambda\gamma \right. \\
&\quad \times \left. \sum_{r=k-q-1}^{k-2} h(x_r) - \gamma h(x_{k-1}) \right) \\
&+ \lambda_{k_1} \left(f(x_0) - \lambda\gamma \sum_{r=0}^{k-q-2} h(x_r) \right. \\
&\quad \left. - \lambda\gamma \sum_{r=k-q-1}^{k-3} h(x_r) - \gamma h(x_{k-2}) \right) \\
&+ \cdots + \lambda_{k_p} \left(f_0(x_0) - \lambda\gamma \sum_{r=0}^{k-q-2} h(x_r) \right. \\
&\quad \left. - \gamma h(x_{k-q-1}) \right) - \gamma h(x_k) \\
&\leq \sum_{t=0}^q \lambda_{kt} f(x_0) - \lambda\gamma \sum_{r=0}^{k-q-2} \left(\sum_{t=0}^q \lambda_{kt} \right) h(x_r) \\
&\quad - \sum_{t=0}^q \lambda_{kt} \gamma h(x_{k-t-1}) - \gamma h(x_k).
\end{aligned} \tag{40}$$

By the fact that $\sum_{t=0}^q \lambda_{kt} = 1$, $\lambda_{kt} \geq \lambda$ and $h(x_r) \geq 0$, we have

$$\begin{aligned}
f(x_{k+1}) &\leq f(x_0) - \lambda\gamma \sum_{r=0}^{k-q-2} h(x_r) \\
&\quad - \lambda\gamma \sum_{r=k-q-1}^{k-1} h(x_r) - \gamma h(x_k) \\
&= f(x_0) - \lambda\gamma \sum_{r=0}^{k-1} h(x_r) - \gamma h(x_k) \\
&\leq f(x_0) - \lambda\gamma \sum_{r=0}^k h(x_r).
\end{aligned} \tag{41}$$

Then for all k , (38) holds.

Moreover, since $\{f(x_k)\}$ is bounded below, let $k \rightarrow \infty$, we can get that

$$\lambda\gamma \sum_{k=0}^{\infty} h(x_k) < \infty, \tag{42}$$

which implies $h(x_k) \rightarrow 0$. Hence the result follows. \square

Lemma 10. Suppose that the assumptions hold. If Algorithm 2 does not terminate finitely, let $\{x_k\}$ be an infinite sequence generated by Algorithm 2; if $|\mathcal{A}| \neq \infty$, then $\lim_{k \rightarrow \infty} \inf \|\hat{g}_k\| = 0$.

Proof. Suppose to contrary that there exist constant $\epsilon > 0$ and $\bar{k} > 0$ such that $\|\hat{g}_k\| > 2\epsilon$ for all $k > \bar{k}$. Then similar to the proof of Lemma 8, we have

$$\begin{aligned}
\|W_k^T \nabla q_k(d_k^n)\| &\geq \|\hat{g}_k\| - \|W_k^T H_k d_k^n\| \\
&\geq 2\epsilon - v_5 \|d_k^n\| \geq 2\epsilon - v_5 \alpha_1 \|c_k\|,
\end{aligned} \tag{43}$$

for all $k > \bar{k}$. Since $h(x_k) \rightarrow 0$, then $\|c_k\| \rightarrow 0$; there exists $\hat{k} > \bar{k}$ such that

$$\|W_k^T \nabla q_k(d_k^n)\| \geq \epsilon \quad \forall k > \hat{k}. \tag{44}$$

From $h(x_k) \rightarrow 0$, we also have $d_k^n \rightarrow 0$, then $q_k(d_k^n) \rightarrow 0$. Hence

$$\begin{aligned}
\text{pred}_k^f &= -q_k(d_k) = -q_k(d_k^n) + q_k(d_k^n) - q_k(d_k) \\
&\geq -q_k(d_k^n) + \alpha\epsilon \min\{\Delta_k, \epsilon\} \geq \frac{\alpha}{2}\epsilon \min\{\bar{\Delta}, \epsilon\}.
\end{aligned} \tag{45}$$

As in the proof of Lemma 8, there exists $\rho \in (0, 1)$ such that $\text{rared}_k^f \geq \rho \min\{\text{pred}_k^f, \text{pred}_k^c + \xi_0\}$. That is,

$$\begin{aligned}
f(x_{k+1}) &\leq \max \left\{ f(x_k), \sum_{r=0}^{m(k)-1} \lambda_{kr} f(x_{k-r}) \right\} \\
&\quad - \rho \min \{ \text{pred}_k^f, \text{pred}_k^c + \xi_0 \}.
\end{aligned} \tag{46}$$

In common with the proof of Lemma 9, we have

$$\rho \sum_{k=0}^{\infty} \min \{ \text{pred}_k^f, \text{pred}_k^c + \xi_0 \} \rightarrow 0, \tag{47}$$

which implies $\text{pred}_k^f \rightarrow 0$. It contradicts (45). Hence the result follows. \square

Lemma 11. Let $\{x_k\}$ be an infinite sequence generated by Algorithm 2. If $|\mathcal{A}| = \infty$, then $\lim_{k \in \mathcal{A}, k \rightarrow \infty} h(x_k) = 0$.

Proof. In view of convenience, denote

$$\theta_{l(p)} = \max_{0 \leq j \leq m(p)-1} \theta_{M+p-j}, \tag{48}$$

where $M + p - m(p) + 1 \leq l(p) \leq M + p$. From the algorithm, we know $\theta_{M+p+1} = h(x_{k+1})$ and it holds

$$\theta_{M+p+1} \leq \beta \theta_{l(p)}. \tag{49}$$

Since $m(p+1) \leq m(p) + 1$, we have

$$\begin{aligned}
\theta_{l(p+1)} &= \max_{0 \leq j \leq m(p+1)-1} \theta_{M+p+1-j} \\
&\leq \max_{0 \leq j \leq m(p)} \theta_{M+p+1-j} \\
&= \max \{ \theta_{M+p+1}, \theta_{l(p)} \} \\
&= \theta_{l(p)}.
\end{aligned} \tag{50}$$

This implies that $\{\theta_{l(p)}\}$ converges. Moreover, by $\theta_{M+p+1} \leq \beta\theta_{l(p)}$, we obtain

$$\theta_{l(p)} \leq \beta\theta_{l(p)-M-1}. \tag{51}$$

And since $\beta \in (0, 1)$, eventually $\theta_{l(p)} \rightarrow 0$ ($p \rightarrow \infty$).

Thus

$$\theta_{M+p+1} \leq \beta\theta_{l(p)} \rightarrow 0 \tag{52}$$

holds by Algorithm 2, then, with the facts $\theta_{M+p+1} = h(x_{k+1})$ and $k + 1 \in \mathcal{A}$, we have $\lim_{k \in \mathcal{A}, k \rightarrow \infty} h(x_k) = 0$. \square

Lemma 12. *Suppose that the assumptions hold. If Algorithm 2 does not terminate finitely, let $\{x_k\}$ be an infinite sequence generated by Algorithm 2; if $|\mathcal{A}| = \infty$, then $\lim_{k \in \mathcal{A}, k \rightarrow \infty} \inf \|\widehat{g}_k\| = 0$.*

Proof. Suppose to the contrary that there exist constant $\epsilon > 0$ and $\bar{k} > 0$ such that $\|\widehat{g}_k\| > 2\epsilon$ for all $k > \bar{k}$ and $k \in \mathcal{A}$. Then similar to the proof of Lemma 8, we have

$$\begin{aligned} \|W_k^T \nabla q_k(d_k^n)\| &\geq \|\widehat{g}_k\| - \|W_k^T H_k d_k^n\| \\ &\geq 2\epsilon - v_5 \|d_k^n\| \geq 2\epsilon - v_5 \alpha_1 \|c_k\|, \end{aligned} \tag{53}$$

for all $k > \bar{k}$ and $k \in \mathcal{A}$. Since $\lim_{k \in \mathcal{A}, k \rightarrow \infty} h(x_k) \rightarrow 0$, that is, $\|c_k\| \rightarrow 0$ for $k \in \mathcal{A}$, then there exists $\widehat{k} > \bar{k}$ such that

$$\|W_k^T \nabla q_k(d_k^n)\| \geq \epsilon \quad \forall k > \widehat{k}, k \in \mathcal{A}. \tag{54}$$

From $\lim_{k \in \mathcal{A}, k \rightarrow \infty} h(x_k) = 0$, we also have $d_k^n \rightarrow 0$ for $k \in \mathcal{A}$, then $q_k(d_k^n) \rightarrow 0$ for $k \in \mathcal{A}$. Thus, for $k \in \mathcal{A}$

$$\begin{aligned} \text{pred}_k^f &= -q_k(d_k) = -q_k(d_k^n) + q_k(d_k^n) - q_k(d_k) \\ &\geq -q_k(d_k^n) + \alpha\epsilon \min\{\Delta_k, \epsilon\} \geq \frac{\alpha}{2}\epsilon \min\{\overline{\Delta}, \epsilon\}. \end{aligned} \tag{55}$$

Similar to the proof of Lemma 10, it is obtained that

$$\rho \sum_{k=0}^{\infty} \min\{\text{pred}_k^f, \text{pred}_k^c + \xi_0\} \rightarrow 0, \tag{56}$$

which implies $\text{pred}_k^f \rightarrow 0$ for all k . But $\text{pred}_k^f \rightarrow 0$ for $k \in \mathcal{A}$ is also true, which contradicts (55). Hence the result follows. \square

Theorem 13. *Suppose $\{x_k\}$ is an infinite sequence generated by Algorithm 2, one has*

$$\liminf_{k \rightarrow \infty} [\|c_k\| + \|\widehat{g}_k\|] = 0. \tag{57}$$

Namely, there exists at least one cluster point of $\{x_k\}$ that is a KKT point of problem (P).

Proof. The conclusion follows immediately by Lemmas 9, 10, 11, and 12. Thus, the whole proof is completed. \square

5. Numerical Experiments

In this section, we carry out some numerical experiments based on the algorithm over a set of problems from [31, 32]. The program is written in MATLAB and we test and give the performance of Algorithm 2 on these problems. For each test example, we choose the initial matrix $H_0 = I$. As to those constants, we set $\gamma = 0.02$, $\beta = 0.98$, $\rho = 0.5$, $\eta = 0.75$, $\xi_0 = 0.05$, $\gamma_0 = 0.1$, $\gamma_1 = 0.5$, $\gamma_2 = 2$, $\Delta_{\min} = 10^{-6}$, $\Delta_0 = 1$, $M = 3$. We use 10^{-6} as the stopping criterion.

During the numerical experiments, updating of H_k is done by

$$H_{k+1} = H_k - \frac{H_k s_k (s_k)^T H_k}{(s_k)^T H_k s_k} + \frac{r_k (r_k)^T}{(s_k)^T r_k}, \tag{58}$$

where $s_k = x_{k+1} - x_k$, $r_k = \theta_k y_k + (1 - \theta_k) H_k s_k$, $y_k = g_{k+1} - g_k$ and

$$\theta_k = \begin{cases} 1 & \text{if } (s_k)^T y_k \\ & \geq 0.2 (s_k)^T H_k s_k, \\ 0.8 \frac{(s_k)^T H_k s_k}{(s_k)^T H_k s_k - (s_k)^T y_k} & \text{otherwise.} \end{cases} \tag{59}$$

In the following tables, the notations mean as follows:

- (i) SC: the problems from Schittkowski [32];
- (ii) HS: the problems from Hock and Schittkowski [31];
- (iii) n : the number of variables;
- (iv) m : the number of inequality constraints;
- (v) NIT: the number of iterations;
- (vi) NF: the number of evaluations of the objective functions;
- (vii) NG: the number of evaluations of scalar constraint functions;
- (viii) Algorithm 1: our method in this paper;
- (ix) Algorithm 2: the method proposed in [29];
- (x) Algorithm 3: the method proposed in [33].

The detailed numerical results are summarized in Tables 1 and 2. Now we give a brief analysis for numerical test results. From Table 1, we can see that our algorithm executes well for these typical problems taken from [31, 32]. From the computation efficiency in Table 2, by solving some problems in [31, 32] we should point out our algorithm is competitive with the some existed nonmonotone filter type methods to solve equality constrained optimization, for example, [29, 33]. However, in our method an improved nonmonotone filter technique is proposed, which may make the discussion and consideration of the nonmonotone properties of f and h more freely. Furthermore, we consider a more relaxed accepted condition for trial step and a weakened crucial criterion is presented. It is easier to meet those new criteria, so our method is much more flexible. All results summarized show that our algorithm is promising and numerically effective.

TABLE 1: Detailed numerical results.

Problem	n	m	NIT	NF	NG
SC216	2	1	5	16	7
SC219	4	2	28	52	36
SC235	3	1	20	48	30
SC252	3	1	22	46	36
SC254	3	2	18	45	37
SC265	4	2	3	7	3
SC269	5	3	5	10	8
SC344	3	1	12	24	16
HS006	2	1	4	7	5
HS007	2	1	8	14	8
HS008	2	2	3	11	5
HS009	2	1	5	12	6
HS026	3	1	35	87	45
HS028	3	1	7	29	11
HS042	4	2	8	31	10
HS047	5	3	25	86	72
HS050	5	3	12	39	14
HS051	5	3	7	18	9
HS077	5	2	12	22	16
HS079	5	3	11	23	13

TABLE 2: Comparison among three algorithms for same problems in [31, 32].

Problem	(n, m)	NG-NF (Algorithm 1)	NG-NF (Algorithm 2)	NG-NF (Algorithm 3)
SC216	(2, 1)	7-16	6-15	11-13
SC235	(3, 1)	30-48	39-50	39-50
SC252	(3, 1)	36-46	99-122	26-35
SC265	(4, 2)	3-7	3-3	3-3
HS006	(2, 1)	5-7	7-7	9-9
HS007	(2, 1)	8-14	10-13	10-13
HS047	(5, 3)	72-86	95-95	95-95
HS050	(5, 3)	14-39	15-15	15-15
HS051	(5, 3)	9-18	16-16	16-16
HS077	(5, 2)	16-22	22-28	22-28

Acknowledgments

The author would like to thank the anonymous referees for the careful reading and helpful comments, which led to an improved version of the original paper. The research is supported by the Science & Technology Program of Shanghai Maritime University (no. 20120060).

References

- [1] R. Fletcher and S. Leyffer, "Nonlinear programming without a penalty function," *Mathematical Programming*, vol. 91, no. 2, pp. 239–269, 2002.
- [2] R. Fletcher, S. Leyffer, and P. L. Toint, "On the global convergence of a filter-SQP algorithm," *SIAM Journal on Optimization*, vol. 13, no. 1, pp. 44–59, 2002.
- [3] R. Fletcher, N. I. M. Gould, S. Leyffer, P. L. Toint, and A. Wächter, "Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming," *SIAM Journal on Optimization*, vol. 13, no. 3, pp. 635–659, 2002.
- [4] S. Ulbrich, "On the superlinear local convergence of a filter-SQP method," *Mathematical Programming*, vol. 100, no. 1, pp. 217–245, 2004.
- [5] A. Vardi, "A trust region algorithm for equality constrained minimization: convergence properties and implementation," *SIAM Journal on Numerical Analysis*, vol. 22, no. 3, pp. 575–591, 1985.
- [6] R. H. Byrd, R. B. Schnabel, and G. A. Shultz, "A trust region algorithm for nonlinearly constrained optimization," *SIAM Journal on Numerical Analysis*, vol. 24, no. 5, pp. 1152–1170, 1987.
- [7] E. O. Omojokun, *Trust region algorithms for optimization with nonlinear equality and inequality constraints [Ph.D. thesis]*, University of Colorado, Boulder, Colo, USA, 1989.
- [8] C. Audet and J. E. Dennis, "A pattern search filter method for nonlinear programming without derivatives," *SIAM Journal on Optimization*, vol. 14, no. 4, pp. 980–1010, 2004.

- [9] C. M. Chin and R. Fletcher, "On the global convergence of an SLP-filter algorithm that takes EQP steps," *Mathematical Programming*, vol. 96, no. 1, pp. 161–177, 2003.
- [10] M. Ulbrich, S. Ulbrich, and L. N. Vicente, "A globally convergent primal-dual interior-point filter method for nonlinear programming," *Mathematical Programming*, vol. 100, no. 2, pp. 379–410, 2004.
- [11] R. Fletcher and S. Leyffer, "A bundle filter method for nonsmooth nonlinear optimization," Tech. Rep. NA/195, Department of Mathematics, University of Dundee, Dundee, Scotland, 1999.
- [12] E. Karas, A. Ribeiro, C. Sagastizábal, and M. Solodov, "A bundle-filter method for nonsmooth convex constrained optimization," *Mathematical Programming*, vol. 116, no. 1-2, pp. 297–320, 2009.
- [13] N. I. M. Gould, S. Leyffer, and P. L. Toint, "A multidimensional filter algorithm for nonlinear equations and nonlinear least-squares," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 17–38, 2004.
- [14] N. I. M. Gould, C. Sainvitu, and P. L. Toint, "A filter-trust-region method for unconstrained optimization," *SIAM Journal on Optimization*, vol. 16, no. 2, pp. 341–357, 2006.
- [15] A. Wächter and L. T. Biegler, "Line search filter methods for nonlinear programming: motivation and global convergence," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 1–31, 2005.
- [16] A. Wächter and L. T. Biegler, "Line search filter methods for nonlinear programming: local convergence," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 32–48, 2005.
- [17] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search technique for Newton's method," *SIAM Journal on Numerical Analysis*, vol. 23, no. 4, pp. 707–716, 1986.
- [18] Z. W. Chen and X. S. Zhang, "A nonmonotone trust-region algorithm with nonmonotone penalty parameters for constrained optimization," *Journal of Computational and Applied Mathematics*, vol. 172, no. 1, pp. 7–39, 2004.
- [19] N. Y. Deng, Y. Xiao, and F. J. Zhou, "Nonmonotonic trust region algorithm," *Journal of Optimization Theory and Applications*, vol. 76, no. 2, pp. 259–285, 1993.
- [20] P. L. Toint, "Non-monotone trust-region algorithms for nonlinear optimization subject to convex constraints," *Mathematical Programming*, vol. 77, no. 3, pp. 69–94, 1997.
- [21] M. Ulbrich, "Nonmonotone trust-region methods for bound-constrained semismooth equations with applications to nonlinear mixed complementarity problems," *SIAM Journal on Optimization*, vol. 11, no. 4, pp. 889–917, 2001.
- [22] D. C. Xu, J. Y. Han, and Z. W. Chen, "Nonmonotone trust-region method for nonlinear programming with general constraints and simple bounds," *Journal of Optimization Theory and Applications*, vol. 122, no. 1, pp. 185–206, 2004.
- [23] D. T. Zhu, "A nonmonotonic trust region technique for nonlinear constrained optimization," *Journal of Computational Mathematics*, vol. 13, no. 1, pp. 20–31, 1995.
- [24] Z. S. Yu and D. G. Pu, "A new nonmonotone line search technique for unconstrained optimization," *Journal of Computational and Applied Mathematics*, vol. 219, no. 1, pp. 134–144, 2008.
- [25] Z. S. Yu, C. X. He, and Y. Tian, "Global and local convergence of a nonmonotone trust region algorithm for equality constrained optimization," *Applied Mathematical Modelling*, vol. 34, no. 5, pp. 1194–1202, 2010.
- [26] M. Ulbrich and S. Ulbrich, "Non-monotone trust region methods for nonlinear equality constrained optimization without a penalty function," *Mathematical Programming*, vol. 95, no. 1, pp. 103–135, 2003.
- [27] Z. W. Chen, "A penalty-free-type nonmonotone trust-region method for nonlinear constrained optimization," *Applied Mathematics and Computation*, vol. 173, no. 2, pp. 1014–1046, 2006.
- [28] N. Gould and P. L. Toint, "Global convergence of a nonmonotone trust-region filter algorithm for nonlinear programming," in *Proceedings of Gainesville Conference on Multilevel Optimization*, W. Hager, Ed., pp. 129–154, Kluwer, Dordrecht, The Netherlands, 2005.
- [29] K. Su and D. G. Pu, "A nonmonotone filter trust region method for nonlinear constrained optimization," *Journal of Computational and Applied Mathematics*, vol. 223, no. 1, pp. 230–239, 2009.
- [30] M. J. D. Powell, "Convergence properties of a class of minimization algorithm," in *Nonlinear Programming 2*, O. Margasarian, R. Meyer, and S. Robinson, Eds., pp. 1–27, Academic Press, New York, NY, USA, 1975.
- [31] W. Hock and K. Schittkowski, *Test Examples for Nonlinear Programming Codes*, vol. 187 of *Lecture Notes in Economics and Mathematical Systems*, Springer, 1981.
- [32] K. Schittkowski, *More Test Examples for Nonlinear Programming Codes*, vol. 282 of *Lecture Notes in Economics and Mathematical Systems*, Springer, Berlin, Germany, 1987.
- [33] K. Su and H. An, "Global convergence of a nonmonotone filter method for equality constrained optimization," *Applied Mathematics and Computation*, vol. 218, no. 18, pp. 9396–9404, 2012.