

An Inexact Newton-like conditional gradient method for constrained nonlinear systems

M.L.N. Gonçalves * F.R. Oliveira *

May 22, 2017

Abstract

In this paper, we propose an inexact Newton-like conditional gradient method for solving constrained systems of nonlinear equations. The local convergence of the new method as well as results on its rate are established by using a general majorant condition. Two applications of such condition are provided: one is for functions whose the derivative satisfies Hölder-like condition and the other is for functions that satisfies a Smale condition, which includes a substantial class of analytic functions. Some preliminaries numerical experiments illustrating the applicability of the proposed method for medium and large problems are also presented.

Keywords: constrained nonlinear systems; inexact Newton-like method; conditional gradient method; local convergence.

1 Introduction

Let $\Omega \subset \mathbb{R}^n$ be an open set, and $F : \Omega \rightarrow \mathbb{R}^n$ be a continuously differentiable nonlinear function. Consider the following constrained system of nonlinear equations

$$F(x) = 0, \quad x \in C, \tag{1}$$

where $C \subset \Omega$ is a nonempty convex compact set. Constrained nonlinear systems such as (1) appear frequently in many important areas, for instance, engineering, chemistry and economy. Due to this fact, the numerical solutions of problem (1) have been the object of intense research in the last years and, consequently, different methods have been proposed in the literature. Many of them are combinations of Newton methods for solving the unconstrained systems with some strategies

*Instituto de Matemática e Estatística, Universidade Federal de Goiás, Campus II- Caixa Postal 131, CEP 74001-970, Goiânia-GO, Brazil. (E-mails: maxlmg@ufg.br and fabriciaro@gmail.com). The work of these authors was supported in part by CAPES, CNPq Grants 406250/2013-8, 444134/2014-0 and 309370/2014-0.

taking into account the constraint set. Strategies based on projections, trust region, active set and gradient methods have been used; see, e.g., [1, 3, 9, 19, 20, 22, 23, 30, 31, 34, 38, 39, 40].

Recently, paper [15] proposed and established a local convergence analysis of a Newton conditional gradient (Newton-CondG) method for solving (1). Basically, this method consists of computing the Newton step and after applying a conditional gradient (CondG) procedure in order to get the Newton iterative back to the feasible set. It is important to point out that the CondG method, also known as the Frank-Wolfe method, is historically known as one of the earliest first methods for solving convex constrained optimization problems, see [8, 11]. The CondG method and its variants require, at each iteration, to minimize a linear function over the constraint set, which, in general, is significantly simpler than the projection step arising in many proximal-gradient methods. Indeed, projection problems can be computationally hard for some high-dimensional problems. For instance, in large-scale semidefinite programming, each projection subproblem of the proximal-gradient methods requires to obtain the complete eigenvalue decomposition of a large matrix while each subproblem of the CondG methods requires to compute the leading singular vector of such a matrix. The latter requirement is less computationally expensive (see, for example, [18] for more details). Moreover, depending on the application, linear optimization oracles may provide solutions with specific characteristics leading to important properties such as sparsity and low-rank; see, e.g., [12, 18] for a discussion on this subject. Due to these advantages and others, the CondG method has again received much attention, see for instances [12, 16, 18, 25, 29].

It is well-known that implementations of the Newton method for medium- or large-scale problems may be expensive and difficult due to the necessity to compute all the elements of the Jacobian matrix of F , as well as, the exact solution of a linear system for each iteration. For this reason, the main goal of this work is to present an extension of the Newton-CondG method in which the inexact Newton-like method is considered instead of standard Newton method. In each step of this new method, the solution of the linear system and Jacobian matrix can be computed in approximate way; see the INL-CondG method in Section 2 and comments following it. From the theoretical viewpoint, we present a local convergence analysis of the proposed method under a majorant condition. The advantage of using a general condition such as majorant condition in the analyses of Newton methods lies in the fact that it allows to study them in a unified way. Thus, two applications of majorant condition are provided: one is for functions whose derivative satisfies Hölder-like condition and the other is for functions that satisfies a Smale condition, which includes a substantial class of analytic functions. From the applicability viewpoint, we report some preliminary numerical experiments of the proposed method for medium and large problems and compare its performance with the constrained dogleg method [2].

This paper is organized as follows. Subsection 1.1 presents some notation and basic assumptions. Section 2 describes the inexact Newton-like conditional gradient method and presents its convergence theorem whose proof is postponed to Section 3. Two applications of the main convergence theorem are also present in Section 2. Section 4 presents some preliminary numerical experiments of the proposed method. We conclude the paper with some remarks.

1.1 Notation and basic assumptions

This subsection presents some notations and assumptions which will be used in the paper. We assume that $F : \Omega \rightarrow \mathbb{R}^n$ is a continuously differentiable nonlinear function, where $\Omega \subset \mathbb{R}^n$ is an open set containing a nonempty convex compact set C . The Jacobian matrix of F at $x \in \Omega$ is denoted by $F'(x)$. We also assume that there exists $x_* \in C$ such that $F(x_*) = 0$ and $F'(x_*)$ is nonsingular. Let the inner product and its associated norm in \mathbb{R}^n be denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. The open ball centered at $a \in \mathbb{R}^n$ and radius $\delta > 0$ is denoted by $B(a, \delta)$. For a given linear operator $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, we also use $\|\cdot\|$ to denote its norm, which is defined by $\|T\| := \sup\{\|Tx\|, \|x\| \leq 1\}$. The condition number of a continuous linear operator $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is denoted by $\text{cond}(A)$ and it is defined as $\text{cond}(A) := \|A^{-1}\| \|A\|$.

2 The method and its local convergence analysis

In this section, we present the inexact Newton-like conditional gradient (INL-CondG) method for solving (1) as well as its local convergence theorem whose proof is postponed to Section 3. Our analysis is done by using a majorant condition, which allows to unify the convergence results for two classes of nonlinear functions, namely, one satisfying a Hölder-like condition and another one satisfying a Smale condition. The convergence results for these special cases are established in this section.

The INL-CondG method is formally described as follows.

INL-CondG method

Step 0. Let $x_0 \in C$ and $\{\theta_j\} \subset [0, \infty)$ be given. Set $k = 0$ and go to step 1.

Step 1. If $F(x_k) = 0$, then **stop**; otherwise, choose an invertible approximation M_k of $F'(x_k)$ and compute a triple $(s_k, r_k, y_k) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ such that

$$M_k s_k = -F(x_k) + r_k, \quad y_k = x_k + s_k. \quad (2)$$

Step 2. Use CondG procedure to obtain $x_{k+1} \in C$ as

$$x_{k+1} = \text{CondG}(y_k, x_k, \theta_k \|s_k\|^2).$$

Step 3. Set $k \leftarrow k + 1$, and go to step 1.

end

We now describe the subroutine CondG procedure.

CondG procedure $z = \text{CondG}(y, x, \varepsilon)$

P0. Set $z_1 = x$ and $t = 1$.

P1. Use the LO oracle to compute an optimal solution u_t of

$$g_t^* = \min_{u \in C} \{\langle z_t - y, u - z_t \rangle\}.$$

P2. If $g_t^* \geq -\varepsilon$, set $z = z_t$ and **stop** the procedure; otherwise, compute $\alpha_t \in (0, 1]$ and z_{t+1} as

$$\alpha_t := \min \left\{ 1, \frac{-g_t^*}{\|u_t - z_t\|^2} \right\}, \quad z_{t+1} = z_t + \alpha_t(u_t - z_t).$$

P3. Set $t \leftarrow t + 1$, and go to **P1**.

end procedure

Remarks. 1) In our local analysis of the INL-CondG method, the invertible approximation M_k of $F'(x_k)$ and the residual r_k will satisfy classic conditions (see (5) and (6)). The inexact Newton-like method with these conditions on M_k and r_k was proposed in [33] and, subsequently, also studied in, for example, [6, 10]. 2) The INL-CondG method can be seen as a class of methods, depending on the choices of the invertible approximation M_k of $F'(x_k)$ and residual r_k . Indeed, by letting $M_k = F'(x_k)$ and $r_k = 0$, the INL-CondG method corresponds to Newton conditional gradient method which was studied in [15]. Another classical choice of M_k would be $M_k = F'(x_0)$. We also emphasize that there are some approach to built M_k that do not involve derivatives, see, for example, [24, 26, 30]. 3) Due to existence of constraint set C , the point y_k in Step 1 may be infeasible and hence the INL-CondG method use an inexact conditional gradient method in order to obtain the new iterate x_{k+1} in C . 4) The CondG procedure requires an oracle which is assumed to be able to minimize linear functions over the constraint set. 5) Finally, in the CondG procedure, if $g_t^* \geq -\varepsilon$, then $z_t \in C$ and stop. However, if $g_t^* < -\varepsilon \leq 0$, the procedure continues. In this case, the stepsize α_t is well defined and belong to $(0, 1]$.

In the following, we state our main local convergence result for the INL-CondG method whose proof is given in Section 3.

Theorem 1. *Let $x_* \in C$, $R > 0$ and $\kappa := \kappa(\Omega, R) = \sup \{t \in [0, R) : B(x_*, t) \subset \Omega\}$. Suppose that there exist a $f : [0, R) \rightarrow \mathbb{R}$ continuously differentiable function such that*

$$\|F'(x_*)^{-1} [F'(x) - F'(x_* + \tau(x - x_*))]\| \leq f'(\|x - x_*\|) - f'(\tau\|x - x_*\|), \quad (3)$$

for all $\tau \in [0, 1]$ and $x \in B(x_*, \kappa)$, where

h1. $f(0) = 0$ and $f'(0) = -1$;

h2. f' is strictly increasing.

Take the constants ϑ , ω_1 , ω_2 and λ such that

$$0 \leq \vartheta < 1, \quad 0 \leq \omega_2 < \omega_1, \quad \omega_1\vartheta + \omega_2 < 1, \quad \lambda \in [0, (1 - \omega_2 - \omega_1\vartheta)/(\omega_1(1 + \vartheta))].$$

Let the scalars ν , ρ and σ defined as

$$\begin{aligned} \nu &:= \sup\{t \in [0, R) : f'(t) < 0\}, \\ \rho &:= \sup\left\{\delta \in (0, \nu) : \omega_1(1 + \vartheta)(1 + \lambda) \left(\frac{f(t)}{tf'(t)} - 1\right) + \omega_1[(1 + \vartheta)\lambda + \vartheta] + \omega_2 < 1, t \in (0, \delta)\right\}, \\ \sigma &:= \min\{\kappa, \rho\}. \end{aligned} \quad (4)$$

Let $\{\theta_k\}$ and x_0 be given in step 0 of the INL-CondG method and let also $\{M_k\}$ and $\{(x_k, r_k)\}$ be generated by the INL-CondG method. Assume that the invertible approximation M_k of $F'(x_k)$ satisfies

$$\|M_k^{-1}F'(x_k)\| \leq \omega_1, \quad \|M_k^{-1}F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots, \quad (5)$$

and the residual r_k is such that

$$\|P_k r_k\| \leq \eta_k \|P_k F(x_k)\|, \quad 0 \leq \eta_k \text{ cond}(P_k F'(x_k)) \leq \vartheta, \quad k = 0, 1, \dots, \quad (6)$$

where $\{P_k\}$ is a sequence of invertible matrix (preconditioners for the linear system in (2)) and $\{\eta_k\}$ is a forcing sequence. If $x_0 \in C \cap B(x_*, \sigma) \setminus \{x_*\}$ and $\{\theta_k\} \subset [0, \lambda^2/2]$, then $\{x_k\}$ is contained in $B(x_*, \sigma) \cap C$, converges to x_* and there holds

$$\|x_{k+1} - x_*\| < \|x_k - x_*\|, \quad \limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leq \omega_1[(1 + \vartheta)\sqrt{2\tilde{\theta}} + \vartheta] + \omega_2, \quad (7)$$

where $\tilde{\theta} = \limsup_{k \rightarrow \infty} \theta_k$. Additionally, given $0 < p \leq 1$, assume that the following condition holds:

h3. the function $(0, \nu) \ni t \mapsto [f(t)/f'(t) - t]/t^{p+1}$ is strictly increasing.

Then, for all integer $k \geq 0$, we have

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \omega_1(1 + \vartheta)(1 + \lambda) \left(\frac{f(\|x_0 - x_*\|)}{f'(\|x_0 - x_*\|)} - \|x_0 - x_*\|\right) \left(\frac{\|x_k - x_*\|}{\|x_0 - x_*\|}\right)^{p+1} \\ &\quad + (\omega_1[(1 + \vartheta)\lambda + \vartheta] + \omega_2) \|x_k - x_*\|. \end{aligned} \quad (8)$$

Remark 1. As mentioned before, the INL-CondG method as be viewed as a class of methods. Hence, the above Theorem implies, in particular, the convergence of some new methods, which

are named below. We obtain, from Theorem 2, the convergence for the inexact modified Newton conditional gradient method if $M_k = F'(x_0)$, the inexact Newton conditional gradient method if $\omega_1 = 1$ and $\omega_2 = 0$ (i.e., $M_k = F'(x_k)$), and the Newton-like conditional gradient method if $\vartheta = 0$ (in this case $\eta_k \equiv 0$ and $r_k \equiv 0$). We also mention that when $\omega_1 = 1$, $\omega_2 = 0$ and $\vartheta = 0$ (i.e., $M_k = F'(x_k)$, $\eta_k \equiv 0$ and $r_k \equiv 0$), Theorem 2 is similar to Theorem 6 in [15].

Remark 2. It is worth pointing out that when $\tilde{\theta} = 0$, then (7) implies that the sequence $\{x_k\}$ converge linearly to x_* . Additionally, if $\omega_1 = 1$, $\omega_2 = 0$ and $\vartheta = 0$, then $\{x_k\}$ converge superlinear to x_* . On the other hand, if f' is convex, i.e., **h3** holds with $p = 1$, it follows from (8), the first inequality in (7), definition (4) and the fact that $x_0 \in C \cap B(x_*, \sigma) \setminus \{x_*\}$ that $\{x_k\}$ converge linearly to x_* . Additionally, if $\omega_1 = 1$, $\omega_2 = \vartheta = \lambda = 0$, it follows from (8), that $\{x_k\}$ converge quadratically to x_* .

We now specialize Theorem 1 for two important classes of functions. In the first one, F' satisfies a Hölder-like condition [13, 14, 17], and in the second one, F is an analytic function satisfying a Smale condition [36, 37].

Corollary 2. Let $\kappa = \kappa(\Omega, \infty)$ as defined in Theorem 1. Assume that there exist a constant $K > 0$ and $0 < p \leq 1$ such that

$$\|F'(x_*)^{-1}[F'(x) - F'(x_* + \tau(x - x_*))]\| \leq K(1 - \tau^p)\|x - x_*\|^p, \quad \tau \in [0, 1], \quad x \in B(x_*, \kappa). \quad (9)$$

Take $0 \leq \vartheta < 1$, $0 \leq \omega_2 < \omega_1$ such that $\omega_1\vartheta + \omega_2 < 1$ and $\lambda \in [0, (1 - \omega_2 - \omega_1\vartheta)/(\omega_1(1 + \vartheta))]$. Let

$$\bar{\sigma} := \min \left\{ \kappa, \left[\frac{(1 - \omega_1[(1 + \vartheta)\lambda + \vartheta] - \omega_2)(p + 1)}{K(p - \omega_1[(1 + \vartheta)\lambda + \vartheta - p] - \omega_2(p + 1) + 1)} \right]^{1/p} \right\}.$$

Let $\{\theta_k\}$ and x_0 be given in step 0 of the INL-CondG method and let also $\{M_k\}$ and $\{(x_k, r_k)\}$ be generated by the INL-CondG method. Assume that the invertible approximation M_k of $F'(x_k)$ satisfies

$$\|M_k^{-1}F'(x_k)\| \leq \omega_1, \quad \|M_k^{-1}F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

and the residual r_k is such that

$$\|P_k r_k\| \leq \eta_k \|P_k F(x_k)\|, \quad 0 \leq \eta_k \text{ cond}(P_k F'(x_k)) \leq \vartheta, \quad k = 0, 1, \dots$$

where $\{P_k\}$ is a sequence of invertible matrix (preconditioners for the linear system in (2)) and $\{\eta_k\}$ is a forcing sequence. If $x_0 \in C \cap B(x_*, \sigma) \setminus \{x_*\}$ and $\{\theta_k\} \subset [0, \lambda^2/2]$, then $\{x_k\}$ is contained in $B(x_*, \sigma) \cap C$, converges to x_* and there hold

$$\|x_{k+1} - x_*\| < \|x_k - x_*\|, \quad \limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leq \omega_1[(1 + \vartheta)\sqrt{2\tilde{\theta}} + \vartheta] + \omega_2,$$

$$\|x_{k+1} - x_*\| \leq \frac{\omega_1(1+\vartheta)(1+\lambda)pK}{(p+1)[1 - K\|x_0 - x_*\|^p]} \|x_k - x_*\|^{p+1} + (\omega_1[(1+\vartheta)\lambda + \vartheta] + \omega_2)\|x_k - x_*\|, \quad k \geq 0,$$

where $\tilde{\theta} = \limsup_{k \rightarrow \infty} \theta_k$.

Proof. It is immediate to prove that F , x_* and $f : [0, \infty) \rightarrow \mathbb{R}$ defined by $f(t) = Kt^{p+1}/(p+1) - t$ satisfy the inequality (3), conditions **h1**, **h2** and **h3** in Theorem 1. Moreover, in this case, it is easily seen that ν and ρ , as defined in Theorem 1, satisfy

$$\rho = \left[\frac{(1 - \omega_1[(1+\vartheta)\lambda + \vartheta] - \omega_2)(p+1)}{K(p - \omega_1[(1+\vartheta)\lambda + \vartheta - p] - \omega_2(p+1) + 1)} \right]^{1/p} < \nu = \left[\frac{1}{K} \right]^{1/p},$$

as a consequence, $\bar{\sigma} = \min\{\kappa, \rho\} = \sigma$ (see (4)). Therefore, the statements of the corollary follow directly from Theorem 1. \square

Remarks. 1) If a function F is such that its derivative is L-Lipschitz continuous, i.e., $\|F'(x) - F'(y)\| \leq L\|x - y\|$, for all $x, y \in B(x_*, \kappa)$ where $L > 0$, then it also satisfies condition (9) with $p = 1$ and $K = L\|F'(x_*)^{-1}\|$. Hence, we obtain the convergence of the INL-CondG method under a Lipschitz condition. In this case, $\{x_k\}$ converges linearly to x_* , and if additionally $\omega_1 = 1$ and $\omega_2 = \vartheta = \lambda = 0$, it converges to x_* quadratically. 2) It is worth mentioning that if $\omega_1 = 1$ and $\omega_2 = \vartheta = 0$ in the previous corollary, we obtain the convergence of the Newton-CondG method under a Hölder-like condition, as obtained in [15, Theorem 7].

We next specialize Theorem 1 for the class of analytic functions satisfying a Smale condition.

Corollary 3. *Let $\kappa = \kappa(\Omega, 1/\gamma)$ as defined in Theorem 1. Assume that $F : \Omega \rightarrow \mathbb{R}^n$ is an analytic function and*

$$\gamma := \sup_{n>1} \left\| \frac{F'(x_*)^{-1}F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} < +\infty.$$

Take $0 \leq \vartheta < 1$, $0 \leq \omega_2 < \omega_1$ such that $\omega_1\vartheta + \omega_2 < 1$ and $\lambda \in [0, (1 - \omega_2 - \omega_1\vartheta)/(\omega_1(1 + \vartheta))]$. Let $a := \omega_1(1 + \vartheta)(1 - 3\lambda) + 4(1 - \omega_1\vartheta - \omega_2)$, $b := 1 - \omega_1[(1 + \vartheta)\lambda + \vartheta] - \omega_2$ and

$$\bar{\sigma} := \min \left\{ \kappa, \frac{a - \sqrt{a^2 - 8b^2}}{4\gamma b} \right\}.$$

Let $\{\theta_k\}$ and x_0 be given in step 0 of the INL-CondG method and let also $\{M_k\}$ and $\{(x_k, r_k)\}$ be generated by the INL-CondG method. Assume that the invertible approximation M_k of $F'(x_k)$ satisfies

$$\|M_k^{-1}F'(x_k)\| \leq \omega_1, \quad \|M_k^{-1}F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

and the residual r_k is such that

$$\|P_k r_k\| \leq \eta_k \|P_k F(x_k)\|, \quad 0 \leq \eta_k \operatorname{cond}(P_k F'(x_k)) \leq \vartheta, \quad k = 0, 1, \dots$$

where $\{P_k\}$ is a sequence of invertible matrix (preconditioners for the linear system in (2)) and $\{\eta_k\}$ is a forcing sequence. If $x_0 \in C \cap B(x_*, \sigma) \setminus \{x_*\}$ and $\{\theta_k\} \subset [0, \lambda^2/2]$, then $\{x_k\}$ is contained in $B(x_*, \sigma) \cap C$, converges to x_* and there holds

$$\|x_{k+1} - x_*\| < \|x_k - x_*\|, \quad \limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leq \omega_1 [(1 + \vartheta)\sqrt{2\tilde{\theta}} + \vartheta] + \omega_2,$$

$$\|x_{k+1} - x_*\| \leq \frac{\omega_1(1 + \vartheta)(1 + \lambda)\gamma}{2(1 - \gamma\|x_0 - x_*\|)^2 - 1} \|x_k - x_*\|^2 + (\omega_1[(1 + \vartheta)\lambda + \vartheta] + \omega_2)\|x_k - x_*\|, \quad k \geq 0,$$

where $\tilde{\theta} = \limsup_{k \rightarrow \infty} \theta_k$.

Proof. Under the assumptions of the corollary, the real function $f : [0, 1/\gamma) \rightarrow \mathbb{R}$, defined by $f(t) = t/(1 - \gamma t) - 2t$, is a majorant function for F on $B(x_*, 1/\gamma)$; see for instance, [10, Theorem 15]. Since f' is convex, it satisfies **h3** in Theorem 1 with $p = 1$; see [10, Proposition 7]. Moreover, in this case, it is easily seen that ν and ρ , as defined in Theorem 1, satisfy

$$\rho = \frac{a - \sqrt{a^2 - 8b^2}}{4\gamma b}, \quad \nu = \frac{\sqrt{2} - 1}{\sqrt{2}\gamma}, \quad \rho < \nu < \frac{1}{\gamma},$$

and, as a consequence, $\bar{\sigma} = \min\{\kappa, \rho\} = \sigma$ (see (4)). Therefore, the statements of the corollary follow from Theorem 1. \square

Remark. The convergence of the Newton-CondG method under a Smale condition, as obtained in [15, Theorem 8], follows from the previous corollary with $\omega_1 = 1$ and $\omega_2 = \vartheta = 0$.

3 Proof of Theorem 1

The main goal of this section is to prove Theorem 1. First, we establish some properties involving the majorant function and its Newton iteration map. Then, some properties of the CondG procedure are discussed. Finally, the desired proof is presented.

From now on, we assume that all the assumptions of Theorem 1 hold, with the exception of **h3**, which will be considered to hold only when explicitly stated.

Proposition 4. *The constant ν is positive and $f'(t) < 0$ for all $t \in [0, \nu)$. As a consequence, the Newton iteration map $n_f : [0, \nu) \rightarrow \mathbb{R}$ defined by*

$$n_f(t) = t - f(t)/f'(t) \tag{10}$$

is well defined and satisfies

$$n_f(t) < 0 \text{ for all } t \in (0, \nu) \text{ and } \lim_{t \downarrow 0} \frac{|n_f(t)|}{t} = 0. \quad (11)$$

Moreover, the constants ρ and σ are positive and

$$0 < \omega_1(1 + \vartheta)(1 + \lambda)|n_f(t)| + (\omega_1[(1 + \vartheta)\lambda + \vartheta] + \omega_2)t < t, \quad t \in (0, \rho). \quad (12)$$

Proof. Firstly, since f' is continuous and $f'(0) = -1$, it follows that the constant ν is positive. Hence, **h2** implies that $f'(t) < 0$ for all $t \in [0, \nu)$, from which we conclude that n_f is well defined. On the other hand, in view of **h2**, we have f is strictly convex in $[0, R)$. Therefore, using $\nu \leq R$, we obtain $f(0) > f(t) + f'(t)(0 - t)$, for any $t \in (0, \nu)$ which, combined with $f(0) = 0$ and $f'(t) < 0$ for any $t \in (0, \nu)$, proves the inequality in (11). Now, using the fact that $f(0) = 0$ and $n_f(t) < 0$ for all $t \in (0, \nu)$, we obtain

$$\frac{|n_f(t)|}{t} = \frac{1}{t} \left(\frac{f(t)}{f'(t)} - t \right) = \frac{1}{f'(t)} \left(\frac{f(t) - f(0)}{t - 0} \right) - 1, \quad t \in (0, \nu). \quad (13)$$

Since $f'(0) \neq 0$, the second statement in (11) follows by taking limit in (13), as $t \downarrow 0$.

It remains to prove the last part of the proposition. First, as $\lambda < [1 - \omega_2 - \omega_1\vartheta]/\omega_1(1 + \vartheta)$, we have $[1 - \omega_1(1 + \vartheta)\lambda - \omega_1\vartheta - \omega_2]/\omega_1(1 + \vartheta)(1 + \lambda) > 0$. Hence, using (11), we conclude that there exists $\delta > 0$ such that

$$0 < \frac{|n_f(t)|}{t} < \frac{1 - \omega_1(1 + \vartheta)\lambda - \omega_1\vartheta - \omega_2}{\omega_1(1 + \vartheta)(1 + \lambda)}, \quad t \in (0, \delta),$$

or, equivalently,

$$0 < \omega_1(1 + \vartheta)(1 + \lambda) \frac{|n_f(t)|}{t} + \omega_1[(1 + \vartheta)\lambda + \vartheta] + \omega_2 < 1, \quad t \in (0, \delta).$$

Hence, ρ is positive which in turn implies that σ is positive and (12) holds. \square

The following lemma gives the some relationships between the majorant function f and the nonlinear operator F .

Lemma 5. *Let $x \in B(x_*, \min\{\kappa, \nu\})$. Then the function $F'(x)$ is invertible and the following estimates hold:*

- a) $\|F'(x)^{-1}F'(x_*)\| \leq 1/|f'(\|x - x_*\|)|$;
- b) $\|F'(x)^{-1}F(x)\| \leq f(\|x - x_*\|)/f'(\|x - x_*\|)$;
- c) $\|F'(x_*)^{-1}[F(x_*) - F(x) - F'(x)(x_* - x)]\| \leq f'(\|x - x_*\|)\|x - x_*\| - f(\|x - x_*\|)$.

Proof. The proof follows the same pattern as the proofs of Lemmas 10, 11 and 12 in [10]. \square

The next result presents a basic property of the CondG procedure whose the proof can be found in [15, lemma 4].

Lemma 6. *For any $y, \tilde{y} \in \mathbb{R}^n$, $x, \tilde{x} \in C$ and $\mu \geq 0$, we have*

$$\| \text{CondG}(y, x, \mu) - \text{CondG}(\tilde{y}, \tilde{x}, 0) \| \leq \|y - \tilde{y}\| + \sqrt{2\mu}.$$

Before presenting the proof of Theorem 1, we first establish a technical result which will be used to prove that the sequence $\{x_k\}$ is contained in $B(x_*, \sigma) \cap C$ and the sequence $\{\|x_k - x_*\|\}$ is strictly decreasing.

Lemma 7. *Assume that $x_k \in C \cap B(x_*, \sigma) \setminus \{x_*\}$. Then, for every $k \geq 0$,*

$$\|x_{k+1} - x_*\| \leq \omega_1(1 + \vartheta)(1 + \sqrt{2\theta_k})|n_f(\|x_k - x_*\|)| + (\omega_1[(1 + \vartheta)\sqrt{2\theta_k} + \vartheta] + \omega_2)\|x_k - x_*\|, \quad (14)$$

where n_f is as in (10). As a consequence,

$$\|x_{k+1} - x_*\| < \|x_k - x_*\|, \quad k \geq 0. \quad (15)$$

Proof. First of all, since $\text{CondG}(x, x, 0) = x$, for all $x \in C$, it follows from INL-CondG method that

$$\|x_{k+1} - x_*\| = \left\| \text{CondG} \left(x_k - M_k^{-1}(F(x_k) - r_k), x_k, \theta_k \|M_k^{-1}(F(x_k) - r_k)\|^2 \right) - \text{CondG}(x_*, x_*, 0) \right\|.$$

Hence, using the Lemma 6 with

$$y = x_k - M_k^{-1}(F(x_k) - r_k), \quad x = x_k, \quad \mu = \theta_k \|M_k^{-1}(F(x_k) - r_k)\|^2, \quad \tilde{y} = x_*, \quad \tilde{x} = x_*,$$

we obtain

$$\|x_{k+1} - x_*\| \leq \|x_k - M_k^{-1}(F(x_k) - r_k) - x_*\| + \sqrt{2\theta_k} \|M_k^{-1}(F(x_k) - r_k)\|.$$

Now, simple calculus yields

$$\begin{aligned} x_k - M_k^{-1}(F(x_k) - r_k) - x_* &= M_k^{-1}(F(x_*) - F(x_k) - F'(x_k)(x_* - x_k)) + (M_k^{-1}F'(x_k) - I)(x_* - x_k) + M_k^{-1}r_k. \end{aligned}$$

Since, $x_k \in C \cap B(x_*, \sigma) \setminus \{x_*\}$, it follows from Lemma 5 that $F'(x_k)$ is invertible. Thus, combining

the last two inequalities, we obtain

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \|M_k^{-1}F'(x_k)\| \|F'(x_k)^{-1}F'(x_*)\| \|F'(x_*)^{-1}[F(x_*) - F(x_k) - F'(x_k)(x_* - x_k)]\| \\ &\quad + \|M_k^{-1}F'(x_k) - I\| \|x_* - x_k\| + \|M_k^{-1}F'(x_k)\| \|F'(x_k)^{-1}P_k^{-1}\| \|P_k r_k\| \\ &\quad + \sqrt{2\theta_k} \|M_k^{-1}F'(x_k)\| \|F'(x_k)^{-1}F(x_k)\| + \sqrt{2\theta_k} \|M_k^{-1}F'(x_k)\| \|F'(x_k)^{-1}P_k^{-1}\| \|P_k r_k\|, \end{aligned}$$

which, combined with (5) and (6), yields

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \omega_1 \|F'(x_k)^{-1}F'(x_*)\| \|F'(x_*)^{-1}[F(x_*) - F(x_k) - F'(x_k)(x_* - x_k)]\| \\ &\quad + \omega_2 \|x_k - x_*\| + \omega_1 \eta_k \|F'(x_k)^{-1}P_k^{-1}\| \|P_k F(x_k)\| \\ &\quad + \omega_1 \sqrt{2\theta_k} \|F'(x_k)^{-1}F(x_k)\| + \omega_1 \eta_k \sqrt{2\theta_k} \|F'(x_k)^{-1}P_k^{-1}\| \|P_k F(x_k)\|. \end{aligned} \quad (16)$$

On the other hand, using the third inequality in (6), we find

$$\begin{aligned} \omega_1 \eta_k \|F'(x_k)^{-1}P_k^{-1}\| \|P_k F(x_k)\| &\leq \omega_1 \eta_k \|(P_k F(x_k))^{-1}\| \|P_k F'(x_k)\| \|F'(x_k)^{-1}F(x_k)\| \\ &\leq \omega_1 \vartheta \|F'(x_k)^{-1}F(x_k)\|. \end{aligned}$$

Hence, it follows from (16) that

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \omega_1 \|F'(x_k)^{-1}F'(x_*)\| \|F'(x_*)^{-1}[F(x_*) - F(x_k) - F'(x_k)(x_* - x_k)]\| + \omega_2 \|x_k - x_*\| \\ &\quad + \omega_1 \vartheta \|F'(x_k)^{-1}F(x_k)\| + \omega_1 \sqrt{2\theta_k} \|F'(x_k)^{-1}F(x_k)\| + \omega_1 \vartheta \sqrt{2\theta_k} \|F'(x_k)^{-1}F(x_k)\|. \end{aligned}$$

Combining the last inequality with items (a), (b) and (c) of Lemma 5, we conclude that

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \omega_1 \left(\frac{f'(\|x_k - x_*\|)\|x_k - x_*\| - f(\|x_k - x_*\|)}{|f'(\|x_k - x_*\|)|} \right) + \omega_2 \|x_k - x_*\| \\ &\quad + \left(\omega_1 \vartheta + \omega_1 \sqrt{2\theta_k} + \omega_1 \vartheta \sqrt{2\theta_k} \right) \frac{f(\|x_k - x_*\|)}{f'(\|x_k - x_*\|)}. \end{aligned}$$

The latter inequality, definition of n_f in (10) and the fact that $f'(\|x_k - x_*\|) < 0$ imply that

$$\begin{aligned} &\|x_{k+1} - x_*\| \\ &\leq \omega_1 \left(\frac{f(\|x_k - x_*\|)}{f'(\|x_k - x_*\|)} - \|x_k - x_*\| \right) + \omega_2 \|x_k - x_*\| + \omega_1 [(1 + \vartheta)\sqrt{2\theta_k} + \vartheta] \frac{f(\|x_k - x_*\|)}{f'(\|x_k - x_*\|)} \\ &= \omega_1 |n_f(\|x_k - x_*\|)| + \omega_2 \|x_k - x_*\| + \omega_1 [(1 + \vartheta)\sqrt{2\theta_k} + \vartheta] (|n_f(\|x_k - x_*\|)| + \|x_k - x_*\|). \end{aligned}$$

Hence, inequality (14) now follows by simple calculus. Since $\sqrt{2\theta_k} \leq \lambda$ and $0 < \|x_k - x_*\| < \sigma \leq \rho$,

it follows from (12) with $t = \|x_k - x_*\|$ that

$$\omega_1(1 + \vartheta)(1 + \sqrt{2\theta_k})|n_f(\|x_k - x_*\|)| + (\omega_1[(1 + \vartheta)\sqrt{2\theta_k} + \vartheta] + \omega_2)\|x_k - x_*\| < \|x_k - x_*\|,$$

which, combined with (14), yields (15). \square

We are now ready to prove Theorem 1.

Proof of Theorem 1: Since $x_0 \in C \cap B(x_*, \sigma) \setminus \{x_*\}$, combining the first statement of Lemma 5, inequality (15) and an induction argument, it is immediate to conclude that the sequence $\{x_k\}$ is contained in $B(x_*, \sigma) \cap C$.

Let us prove that $\{x_k\}$ converges to x_* . Since, $\|x_k - x_*\| < \sigma \leq \rho$ for all $k \geq 0$, the first inequality in (7) follows trivially from (15). Therefore, $\{\|x_k - x_*\|\}$ is a bounded strictly decreasing sequence and hence it converges to some $\ell_* \in [0, \rho)$. Moreover, taking into account that $n_f(\cdot)$ is continuous in $[0, \rho)$, in particular, from (14) we have

$$\ell_* \leq \omega_1(1 + \vartheta)(1 + \lambda)|n_f(\ell_*)| + (\omega_1[(1 + \vartheta)\lambda + \vartheta] + \omega_2)\ell_*,$$

which, combined with (12), implies that $\ell_* = 0$ and consequently $x_k \rightarrow x_*$.

Now, from (14) we obtain

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leq \omega_1(1 + \vartheta)(1 + \sqrt{2\theta_k}) \frac{|n_f(\|x_k - x_*\|)|}{\|x_k - x_*\|} + \omega_1[(1 + \vartheta)\sqrt{2\theta_k} + \vartheta] + \omega_2, \quad k \geq 0.$$

In order to prove the asymptotic rate in (7), just take the limit superior in the last inequality as $k \rightarrow \infty$ and use $\|x_k - x_*\| \rightarrow 0$, equality in (11) and $\limsup_{k \rightarrow \infty} \theta_k = \tilde{\theta}$.

It remains to show the last part of the theorem. For this purpose, let us assume that **h3** holds. It follows from (15) and (14) respectively that, for all $k \geq 0$, $\|x_k - x_*\| < \|x_0 - x_*\|$ and

$$\|x_{k+1} - x_*\| \leq \omega_1(1 + \vartheta)(1 + \sqrt{2\theta_k}) \frac{|n_f(\|x_k - x_*\|)|}{\|x_k - x_*\|^{p+1}} \|x_k - x_*\|^{p+1} + (\omega_1[(1 + \vartheta)\sqrt{2\theta_k} + \vartheta] + \omega_2)\|x_k - x_*\|.$$

Therefore, the inequality (8) follows due to assumption **h3**, $\sqrt{2\theta_k} \leq \lambda$ and the definition of n_f in (10). \square

Remark 3. *Similar to the analysis in [15], we could have defined a scalar sequence $\{t_k\}$, associated to the majorant function f , such that*

$$\|x_k - x_*\| \leq t_k, \quad k \geq 0. \tag{17}$$

Indeed, if $\{t_k\}$ is defined as

$$t_0 := \|x_0 - x_*\|, \quad t_{k+1} := \omega_1(1 + \vartheta)(1 + \sqrt{2\theta_k})|n_f(t_k)| + (\omega_1[(1 + \vartheta)\sqrt{2\theta_k} + \vartheta] + \omega_2)t_k, \quad k \geq 0,$$

it is possible to prove that (17) holds, and $\{t_k\}$ is well defined, is strictly decreasing and converges to 0. Moreover, $\limsup_{k \rightarrow \infty} t_{k+1}/t_k = \omega_1[(1 + \vartheta)\sqrt{2\tilde{\theta}} + \vartheta] + \omega_2$, where $\tilde{\theta} = \limsup_{k \rightarrow \infty} \theta_k$.

4 Numerical experiments

In this section, we present the results of some preliminaries numerical tests which show the computational feasibility of the INL-CondG method. The experiments were carried out on a set of 15 well-known box-constrained nonlinear systems (i.e., problem (1) with $C = \{x \in \mathbb{R}^n : l \leq x \leq u\}$, where $l \in \mathbb{R}^n$ and $u \in (\mathbb{R} \cup \{\infty\})^n$) with dimensions between $n = 400$ and $n = 10000$ (see Table 1). We made three implementations of our INL-CondG method which differ in the way that the approximation matrices M_k 's are built. In the first implementation, the matrices M_k 's were approximated by finite difference (FD), whereas in the second and third ones, we used the Broyden-Schubert Update (BSU) [5, 35] and Bogle-Perkins Update (BPU) [4], respectively. The three resulting methods described above are denoted by FD-INL-CondG, BSU-INL-CondG and BPU-INL-CondG, respectively. In the implementations of the BSU-INL-CondG and BPU-INL-CondG methods, the matrices M_k 's were approximated by finite difference when $k = 0$ and $\text{mod}(k - 1, 5) = 0$. This strategy to periodically compute the Jacobian matrix F' seems to be crucial for the robustness of these derivative-free methods. For a comparison purpose, we also run the constrained Dogleg solver (CoDoSol) which is a MATLAB package based on the constrained Dogleg method [2], and available on the web site <http://codosol.de.unifi.it>.

The computational results were obtained using MATLAB R2016a on a 2.5GHz Intel(R) i5 with 6GB of RAM and Windows 7 ultimate system. The stopping criterion $\|F(x_k)\|_\infty \leq 10^{-6}$ was used, and a failure was declared if the number of iterations was greater than 300 or no progress was detected. The starting points were defined as $x_0(\gamma) = l + 0.25\gamma(u - l)$ where $\gamma = 1, 2, 3$ for problems having finite lower and upper bounds and $x_0(\gamma) = 10^\gamma(1, \dots, 1)^T$ with $\gamma = 0, 1, 2$, for problems with infinite upper bound. However, since $x_0(0)$ is a solution of Pb13, we used $x_0(3)$ instead. In the implementations of the INL-CondG method, the error parameter θ_k was set equal to 10^{-5} for all k and the CondG Procedure stopped when either the required accuracy was obtained or the maximum of 300 iterations were performed. The parameters of CoDoSol were set as the default choice recommended by the authors (see [2, Subsection 4.1]). It worth pointing out that the Jacobian matrices in the latter solver are approximated by finite difference.

Table 2 reports the performance of the FD-INL-CondG, BSU-INL-CondG and BPU-INL-CondG methods, and CoDoSol for solving the 15 problems considered. In table 2, “ γ ” is the scalar used to compute the starting point $x_0(\gamma)$, “Iter” is the number of iterations of the methods and “ $\|F\|_\infty$ ” is the infinity norm of F at the final iterate x_k . Finally, the symbol “*” indicates a failure.

From Table 2, we see that the FD-INL-CondG, BSU-INL-CondG and BPU-INL-CondG methods and CoDoSol successfully ended 42, 40, 37 and 42 times, respectively, on a total of 45 runs. The FD-INL-CondG method is comparable to or even slightly better than CoDoSol, because it required less iterations in 35 cases in which both methods successfully ended. This behavior also has been

Table 1: Test problems

Problem	Name and source	n	Box
Pb 1	Chandrasekhar's H-equation $c = 0.99$ [21]	400	$[0, 5]$
Pb 2	Discrete boundary value function [32, Problem 28]	500	$[-100, 100]$
Pb 3	Troesch [27, Problem 4.21]	500	$[-1, 1]$
Pb 4	Discrete integral [32, Problem 29]	1000	$[-10, 10]$
Pb 5	Trigexp 1 [27, Problem 4.4]	1000	$[-100, 100]$
Pb 6	Problem 74 [28]	1000	$[0, 10]$
Pb 7	Problem 77 [28]	2000	$[0, 10]$
Pb 8	Function 15 [7]	2000	$[-10, 0]$
Pb 9	Tridiagonal exponential [27, Problem 4.18]	2000	$[e^{-1}, e]$
Pb 10	Trigonometric function [7, Problem 8]	2000	$[5, 15]$
Pb 11	Zero Jacobian function [7, Problem 19]	2000	$[0, 10]$
Pb 12	Trigonometric system [27, Problem 4.3]	5000	$[\pi, 2\pi]$
Pb 13	Five diagonal [27, Problem 4.8]	5000	$[1, \infty]$
Pb 14	Seven diagonal [27, Problem 4.9]	5000	$[0, \infty]$
Pb 15	Countercurrent reactors [27, Problem 4.1]	10000	$[-1, 10]$

observed in [15] for some small and medium scale problems. Regarding the methods whose the F' is not evaluated at each iteration, the BSU-INL-CondG method solved 3 problems more than BPU-INL-CondG method, while the BPU-INL-CondG method required less (resp. more) iterations than BSU-INL-CondG method in 11 (resp. 7) cases in which both methods successfully ended. Hence, we may say that latter two methods had similar numerical performance and, for some problems, they are comparable to the methods in which all M_k 's are approximated by finite difference. Finally, based on the previous discussion, we conclude that the INL-CondG method seems to be a promising tool for solving medium and large box-constrained systems of nonlinear equations.

Final remarks

We proposed a method for solving constrained systems of nonlinear equations which is a combination of inexact Newton-like and conditional gradient methods. Under appropriate hypotheses and using a majorant condition, it was showed that the sequence generated by new method converge locally. Additionally, we were able to provide convergence results for two important classes of nonlinear functions, namely, one is for functions whose the derivative satisfies Hölder-like condition and the other is for functions that satisfies a Smale condition. In order to show the practical behavior of the proposed method, we tested it on medium- and large-scale problems from the literature. The numerical experiments showed that it works quite well and compares favorably with the constrained dogleg method [2].

References

- [1] BELLAVIA, S., MACCONI, M., AND MORINI, B. An affine scaling trust-region approach to bound-constrained nonlinear systems. *Appl. Num. Math.* 44, 3 (2003), 257–280.
- [2] BELLAVIA, S., MACCONI, M., AND PIERACCINI, S. Constrained dogleg methods for nonlinear systems with simple bounds. *Comput. Optim. Appl.* 53, 3 (2012), 771–794.
- [3] BELLAVIA, S., AND MORINI, B. Subspace trust-region methods for large bound-constrained nonlinear equations. *SIAM J. Numer. Anal.* 44, 4 (2006), 1535–1555.
- [4] BOGLE, I. D. L., AND PERKINS, J. D. A new sparsity preserving quasi-Newton update for solving nonlinear equations. *SIAM J. Sci. Statist. Comput.* 11, 4 (1990), 621–630.
- [5] BROYDEN, C. G. The convergence of an algorithm for solving sparse nonlinear systems. *Math. Comp.* 25 (1971), 285–294.
- [6] CHEN, J., AND LI, W. Convergence behaviour of inexact Newton methods under weak Lipschitz condition. *J. Comput. Appl. Math.* 191, 1 (2006), 143–164.
- [7] CRUZ, W. L., AND RAYDAN, M. Nonmonotone spectral methods for large-scale nonlinear systems. *Optim. Methods Softw.* 18, 5 (2003), 583–599.
- [8] DUNN, J. C. Convergence rates for conditional gradient sequences generated by implicit step length rules. *SIAM J. Control Optim.* 18, 5 (1980), 473–487.
- [9] ECHEBEST, N., SCHUVERDT, M. L., AND VIGNAU, R. P. A derivative-free method for solving box-constrained underdetermined nonlinear systems of equations. *Appl. Math. Comput.* 219, 6 (2012), 3198–3208.
- [10] FERREIRA, O. P., AND GONÇALVES, M. L. N. Local convergence analysis of inexact Newton-like methods under majorant condition. *Comput. Optim. Appl.* 48, 1 (2011), 1–21.
- [11] FRANK, M., AND WOLFE, P. An algorithm for quadratic programming. *Naval Res. Logist.* 3, 1-2 (1956), 95–110.
- [12] FREUND, R., AND GRIGAS, P. New analysis and results for the Frank-Wolfe method. *Math. Program.* (2014), 1–32.
- [13] GONÇALVES, M. L. N. Local convergence of the Gauss-Newton method for injective-overdetermined systems of equations under a majorant condition. *Comput. Math. Appl.* 66, 4 (2013), 490–499.
- [14] GONÇALVES, M. L. N. Inexact Gauss-Newton like methods for injective-overdetermined systems of equations under a majorant condition. *Numer. Algorithms* 72, 2 (2016), 377–392.
- [15] GONÇALVES, M. L. N., AND MELO, J. G. A Newton conditional gradient method for constrained nonlinear systems. *J. Comput. Appl. Math.* 311 (2017), 473–483.
- [16] HARCHAOU, Z., JUDITSKY, A., AND NEMIROVSKI, A. Conditional gradient algorithms for norm-regularized smooth convex optimization. *Math. Program.* (2014), 1–38.
- [17] HUANG, Z. The convergence ball of Newton’s method and the uniqueness ball of equations under Hölder-type continuous derivatives. *Comput. Math. Appl.* 47, 2-3 (2004), 247–251.
- [18] JAGGI, M. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)* (2013), vol. 28, pp. 427–435.
- [19] KANZOW, C. An active set-type Newton method for constrained nonlinear systems. In *Complementarity: Applications, Algorithms and Extensions*, M. Ferris, O. Mangasarian, and J.-S. Pang, Eds., vol. 50 of *Appl. Optim.* Springer US, 2001, pp. 179–200.
- [20] KANZOW, C. Strictly feasible equation-based methods for mixed complementarity problems. *Numer. Math.* 89, 1 (2001), 135–160.
- [21] KELLEY, C. *Iterative Methods for Linear and Nonlinear Equations*. Society for Industrial and Applied Mathematics, 1995.
- [22] KOZAKEVICH, D. N., MARTINEZ, J. M., AND SANTOS, S. A. Solving nonlinear systems of equations with simple constraints. *Comput. Appl. Math.* 16 (1997), 215–235.
- [23] LA CRUZ, W. A projected derivative-free algorithm for nonlinear equations with convex constraints. *Optim. Methods Softw.* 29, 1 (2014), 24–41.
- [24] LA CRUZ, W., MARTÍNEZ, J. M., AND RAYDAN, M. Spectral residual method without gradient information for solving large-scale nonlinear systems of equations. *Math. Comp.* 75, 255 (2006), 1429–1448.
- [25] LAN, G., AND ZHOU, Y. Conditional gradient sliding for convex optimization. *SIAM J. Optim.* 26, 2 (2016), 1379–1409.
- [26] LUKŠAN, L., AND VLČEK, J. Computational experience with globally convergent descent methods for large sparse systems of nonlinear equations. *Optim. Methods and Softw.* 8 (1998), 201–223.
- [27] LUKŠAN, L., AND VLČEK, J. Sparse and partially separable test problems for unconstrained and equality constrained optimization. *Technical Report N. 767, Institute of Computer Science, Academy of Sciences of the Czech Republic* (1999).
- [28] LUKŠAN, L., AND VLČEK, J. Test problems for unconstrained optimization. *Technical Report N. 897, Institute of Computer Science, Academy of Sciences of the Czech Republic* (2003).
- [29] LUSS, R., AND TEBoulLE, M. Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint. *SIAM Review* 55, 1 (2013), 65–98.

- [30] MARINI, L., MORINI, B., AND PORCELLI, M. Quasi-Newton methods for convex constrained nonlinear systems and their application. Available on <http://www.optimization-online.org/>.
- [31] MARTINEZ, M. J. Quasi-inexact-Newton methods with global convergent for solving constrained nonlinear systems. *Non-linear Anal.* 30, 1 (1997), 1–7.
- [32] MORÉ, J. J., GARROW, B. S., AND HILLSTROM, K. E. Testing unconstrained optimization software. *ACM Trans. Math. Softw.* 7, 1 (1981), 17–41.
- [33] MORINI, B. Convergence behaviour of inexact Newton methods. *Math. Comp.* 68, 228 (1999), 1605–1613.
- [34] PORCELLI, M. On the convergence of an inexact Gauss-Newton trust-region method for nonlinear least-squares problems with simple bounds. *Optim. Letters* 7, 3 (2013), 447–465.
- [35] SCHUBERT, L. K. Modification of a quasi-Newton method for nonlinear equations with a sparse Jacobian. *Math. Comp.* 24 (1970), 27–30.
- [36] SHEN, W., AND LI, C. Smale’s α -theory for inexact Newton methods under the γ -condition. *J. Math. Anal. Appl.* 369, 1 (2010), 29 – 42.
- [37] SMALE, S. Newton’s method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*. Springer, New York, 1986, pp. 185–196.
- [38] WANG, P., AND ZHU, D. An inexact derivative-free levenberg–marquardt method for linear inequality constrained nonlinear systems under local error bound conditions. *Appl. Math. Comput.* 282 (2016), 32–52.
- [39] ZHANG, Y., AND ZHU, D.-T. Inexact Newton method via Lanczos decomposed technique for solving box-constrained nonlinear systems. *Appl. Math. Mech.* 31, 12 (2010), 1593–1602.
- [40] ZHU, D. An affine scaling trust-region algorithm with interior backtracking technique for solving bound-constrained nonlinear systems. *J. Comput. App. Math.* 184, 2 (2005), 343 – 361.

Table 2: Performance of the FD-INL-CondG, BSU-INL-CondG and BPU-INL-CondG methods and CoDoSol

Problem	γ	FD-INL-CondG		BSU-INL-CondG		BPU-INL-CondG		CoDoSol	
		Iter	$\ F\ _\infty$	Iter	$\ F\ _\infty$	Iter	$\ F\ _\infty$	Iter	$\ F\ _\infty$
Pb 1	1	5	1.39e-11	5	9.22e-7	5	6.99e-7	7	1.74e-11
	2	6	4.07e-9	7	3.58e-8	7	2.42e-8	7	8.25e-7
	3	5	1.71e-7	7	5.88e-11	7	1.13e-9	*	
Pb 2	1	9	3.30e-8	12	4.78e-07	12	3.10e-8	14	3.67e-10
	2	1	2.59e-7	1	2.59e-7	1	2.59e-7	2	2.19e-9
	3	9	2.22e-7	12	2.92e-7	12	2.06e-8	14	9.18e-9
Pb 3	1	6	2.18e-7	13	7.99e-7	8	3.60e-8	9	1.91e-9
	2	7	7.71e-8	11	2.35e-7	9	6.65e-8	6	5.58e-9
	3	6	2.18e-7	9	2.02e-7	8	4.03e-8	7	1.56e-7
Pb 4	1	5	2.66e-10	7	9.12e-12	7	2.62e-12	9	1.34e-10
	2	3	2.72e-11	3	2.00e-9	3	2.00e-9	3	1.04e-7
	3	6	4.64e-11	7	2.23e-8	*		9	3.49e-8
Pb 5	1	20	2.60e-8	169	5.46e-9	28	2.68e-8	21	9.01e-7
	2	9	3.40e-11	13	1.05e-7	12	8.00e-10	10	3.46e-9
	3	13	5.21e-9	19	3.64e-10	17	2.78e-8	23	2.41e-8
Pb 6	1	5	5.72e-7	7	1.67e-7	7	2.01e-8	7	3.39e-7
	2	13	2.72e-11	42	1.29e-8	69	2.91e-9	9	2.32e-8
	3	9	2.22e-8	25	1.73e-7	12	2.27e-8	12	3.07e-8
Pb 7	1	6	2.19e-10	7	9.34e-9	7	9.95e-9	9	4.22e-7
	2	8	6.44e-11	10	1.34e-9	10	3.26e-8	12	3.45e-9
	3	9	7.83e-11	11	8.40e-8	12	1.40e-11	13	2.73e-7
Pb 8	1	7	4.75e-10	11	6.48e-8	9	4.36e-8	13	6.70e-11
	2	6	3.00e-7	10	1.36e-7	8	2.61e-7	12	1.75e-7
	3	6	2.88e-13	8	9.36e-8	7	4.29e-7	11	1.93e-9
Pb 9	1	2	4.84e-14	2	4.84e-14	2	4.84e-14	8	6.03e-14
	2	2	1.39e-13	2	1.39e-13	2	1.39e-13	7	6.23e-14
	3	2	2.98e-14	2	2.98e-14	2	2.98e-14	7	3.96e-14
Pb 10	1	7	6.88e-10	9	4.46e-10	16	4.74e-9	10	2.12e-11
	2	3	1.45e-7	4	2.22e-10	4	1.62e-10	4	3.36e-8
	3	10	5.65e-7	14	7.83e-8	73	6.41e-9	12	1.28e-8
Pb 11	1	17	7.28e-7	22	9.58e-7	27	7.51e-7	22	2.56e-7
	2	18	7.28e-7	24	6.13e-7	28	9.21e-7	23	5.56e-7
	3	19	4.09e-7	25	5.39e-7	30	9.31e-7	24	4.60e-7
Pb 12	1	*		*		*		18	2.60e-8
	2	*		*		*		17	9.00e-8
	3	*		*		*		16	4.94e-9
Pb 13	1	8	0.00e+0	7	0.00e+0	5	0.00e+0	17	8.72e-10
	2	12	0.00e+0	22	0.00e+0	*		*	
	3	16	0.00e+0	*		7	0.00e+0	*	
Pb 14	0	4	3.64e-10	5	1.65e-7	5	2.47e-8	4	4.13e-10
	1	12	3.69e-7	17	1.32e-9	17	2.03e-11	17	1.29e-8
	2	20	1.09e-12	28	6.93e-10	*		25	1.29e-11
Pb 15	1	11	6.82e-9	*		21	3.79e-7	17	2.25e-9
	2	12	3.06e-8	19	6.05e-7	*		19	2.06e-9
	3	13	5.98e-9	20	3.24e-7	*		20	2.10e-9