

ON AFFINE SCALING INEXACT DOGLEG METHODS FOR
BOUND-CONSTRAINED NONLINEAR SYSTEMS

by S. Bellavia¹, M. Macconi¹, S. Pieraccini²

Dipartimento di Energetica “Sergio Stecco”,
Università di Firenze, Pubblicazione n. 5/2009.

¹ Dipartimento di Energetica “S. Stecco”,
Università di Firenze,
via C. Lombroso 6/17, 50134 Firenze, Italia,
Email: stefania.bellavia@unifi.it, benedetta.morini@unifi.it.

² Dipartimento di Matematica, Politecnico di Torino, corso Duca degli
Abruzzi, 24, 10129 Torino, Italia, e-mail: pieraccini@calvino.polito.it

On affine scaling inexact dogleg methods for bound-constrained nonlinear systems*

Stefania Bellavia,[†] Maria Macconi,^{*} Sandra Pieraccini[‡]

Abstract

A class of trust-region methods for large scale bound-constrained systems of nonlinear equations is presented. The methods in this class follow the so called affine-scaling approach and can efficiently handle large scale problems. At each iteration, a suitably scaled region around the current approximate solution is defined and, within such a region, the norm of the linear model of F is trusted to be an adequate representation of the merit function $\|F\|$. Both spherical and elliptical trust-regions are allowed. An *inexact dogleg* method is used to obtain an approximate minimizer of the linear model within the trust-region and the feasible set Ω . Thus, strictly feasible iterates are formed and a strictly feasible approximation of the solution is ensured. Global convergence results are established without any Lipschitz assumption on $F'(x)$ and locally fast convergence is showed under standard assumptions. Convergence analysis is performed without specifying the scaling matrix used to handle the bounds. In fact, a rather general class of scaling matrices is allowed in actual algorithms. An algorithm based on the standard spherical trust-region and the pioneer scaling matrix given by Coleman and Li is implemented in a Matlab code and its numerical performance is shown.

Key words: bound-constrained equations, affine scaling, trust-region methods, dogleg methods, inexact Newton methods, global convergence.

1 Introduction

In this paper we discuss affine scaling trust-region methods for large scale systems of nonlinear equations with simple bounds. The problem is to find a vector $x \in \mathbb{R}^n$ satisfying

$$F(x) = 0, \quad x \in \Omega, \quad (1)$$

where $F : X \mapsto \mathbb{R}^n$ is a continuously differentiable mapping, $X \subseteq \mathbb{R}^n$ is an open set containing the n -dimensional box $\Omega = \{x \in \mathbb{R}^n \mid l \leq x \leq u\}$. Here,

*Work supported by the MIUR, Rome, Italy, through grants PRIN07.

[†]Dipartimento di Energetica "S. Stecco", Università di Firenze, via C. Lombroso 6/17, 50134 Firenze, Italia, e-mail: stefania.bellavia@unifi.it, maria.macconi@unifi.it

[‡]Dipartimento di Matematica, Politecnico di Torino, corso Duca degli Abruzzi, 24, 10129 Torino, Italia, e-mail: pieraccini@calvino.polito.it

the inequalities are meant component-wise and the vectors $l \in (\mathbb{R} \cup -\infty)^n$, $u \in (\mathbb{R} \cup +\infty)^n$ are specified lower and upper bounds on the variables such that Ω has nonempty interior.

The numerical solution of large scale bound constrained nonlinear systems is required in many contexts from different areas of scientific, economic and engineering computations [15, 11, 14, 19, 5, 27] and their numerical solution has recently attracted several researchers [4, 24, 25, 12].

Inexact Newton methods augmented with trust-region procedures can efficiently solve constrained systems when the number of equations is large. In particular, the recent works [4, 2] focus on subspace affine scaling globally convergent methods where Newton-Krylov methods are augmented with a strategy for handling the constraints and with a trust-region procedure defined on a low dimensional subspace. These methods generate a sequence of strictly feasible points. Both spherical and elliptical trust-regions are allowed and the trust-region step is approximated by a procedure which basically consists in a double dogleg method. In fact, the trust-region problem is approximately solved by a dogleg strategy on a low (two) dimensional subspace. Then this approximate trust-region step is projected onto Ω and a convex combination of this projected step and the generalized Cauchy step gives the required approximated solution of the trust-region problem in \mathbb{R}^n .

The main motivation for the current work is the recent paper by Pawlowski et al. [23]. In such paper, which deals with the numerical solution of systems of nonlinear equations, the authors provide a very general dogleg method in an inexact Newton framework. However, in [23] only nonlinear systems without constraints on the solutions are addressed.

The present paper aims at generalizing the method proposed in [23] (extending it to systems with simple constraints) and looking for a simple procedure to approximate the trust-region step thus possibly obtaining more efficient algorithms than those previously proposed for constrained nonlinear systems. Nevertheless, some desirable features of the algorithms given in [4] should be maintained. In particular, we require that all the iterates produced lie in the strict interior of Ω . This way, the method is able to solve constrained problems even if F is not defined outside the feasible set Ω . Further, global and locally fast convergence must be ensured independently of the position of the asked solution. This way, we expect good numerical performance both on problems with solution within the interior of Ω and on problems with solution on the boundary of the feasible set.

To define a robust iterative process for the solution of large scale problems, an inexact Newton method is incorporated into a globally convergent affine scaling trust-region scheme. At each iteration, a suitably scaled region around the current approximate solution x_k is defined and, within such a region, the norm of the linear model of F is trusted to be an adequate representation of the merit function $\|F\|$. Both spherical and elliptical trust-regions are allowed. Therefore, an approximate minimizer of the linear model within the trust-region and the feasible set Ω is required at each iteration. This is accomplished by an *inexact dogleg* method that, taking into account the given constraints, ensures a

strictly feasible approximation of the solution. More precisely, the linear model is minimized along a path whose foundations are a scaled Cauchy step and a projected Inexact Newton step. Unlike the standard dogleg, in this approach the path is not a convex combination of these two steps, but a more general combination: this is due to the fact that, as the Newton step is not exact and it may be projected, many theoretical properties of the dogleg curve are lost and more flexibility is allowed in the choice of the step. This is the reason why we refer to this procedure as an *inexact dogleg*.

We stress the significant differences between our proposal and all the affine scaling trust-region methods previously given for the numerical solution of constrained nonlinear systems. The inexact dogleg strategy employed here is relatively simpler than those used in the methods discussed in the literature [3, 4, 16, 21, 26, 28]. Moreover, here, the foundation is the linear model of the merit function $\|F\|$, whereas in the previous works the merit function $\frac{1}{2}\|F\|^2$ is approximated by a quadratic model.

The aforementioned paper [23] has played an important role for the design of our algorithm and a few theoretical results given in [23] are useful in our context, too. In the convergence analysis carried out in the present paper, we point out where we follow the approach suggested in [23]. However, as we work in a constrained context within an affine scaling trust-region framework, different tools are required in order to prove the convergence properties of our method. On the other hand, proofs of theoretical properties of our method beat different paths with respect to those commonly used in the analysis of affine scaling methods for constrained nonlinear systems. This is mainly due to the adopted linear model of F . Moreover, global convergence is proven without any Lipschitz assumption on the Jacobian of $F(x)$ and locally fast convergence is ensured under standard assumptions.

Finally, we remark that the theoretical analysis performed here is carried out without specifying the scaling matrix used to handle the bounds. This distinguishes the present work from previous papers on affine scaling trust-region methods to solve (1) and gives rise to an algorithm with some flexibility in choosing the scaling matrix. In fact, a rather general class of scaling matrices is allowed in actual implementations of the method.

The paper is organized as follows. In Section 2 we derive our affine scaling inexact dogleg method for the solution of (1) and discuss some implementation issues. In Section 3 we show convergence properties of the method. In Section 4 we give details of our implementation of the proposed algorithm and report on our computational experience. Some conclusions and perspectives are presented in Section 5.

1.1 Notation

Throughout the paper we use the following notation. For any mapping $F : X \rightarrow \mathbb{R}^n$, differentiable at a point $x \in X \subset \mathbb{R}^n$, the Jacobian matrix of F at x is denoted by $F'(x)$. The subscript k is used as index for a sequence and when clear from the context the argument of a mapping is omitted. Then, for

any function F , the notation F_k is used to denote $F(x_k)$. To represent the i -th component of a vector $x \in \mathbb{R}^n$ the symbol $(x)_i$ is used but, when clear from the context, the brackets are omitted. For any vector $y \in \mathbb{R}^n$, the 2-norm is denoted by $\|y\|$ and the open ball with center y and radius ρ is indicated by $B_\rho(y)$, i.e. $B_\rho(y) = \{x : \|x - y\| < \rho\}$.

2 The method

We start by noting that every solution of (1) is a solution of the following constrained optimization problem:

$$\min_{x \in \Omega} f(x) = \min_{x \in \Omega} \frac{1}{2} \|F(x)\|^2. \quad (2)$$

Then, in a framework for (1), we expect that a numerical method generates a sequence of points converging to a solution x^* of (2).

As originally shown by Coleman and Li [7], a solution x^* to (2) satisfies

$$D(x^*)\nabla f(x^*) = 0, \quad (3)$$

where $\nabla f(x) = F'(x)^T F(x)$ and $D(x)$ is a proper diagonal scaling matrix of order n with diagonal elements given by

$$d_i^{CL}(x) = \begin{cases} u_i - x_i & \text{if } (\nabla f(x))_i < 0 \quad \text{and } u_i < \infty, \\ x_i - l_i & \text{if } (\nabla f(x))_i > 0 \quad \text{and } l_i > -\infty, \\ \min\{x_i - l_i, u_i - x_i\} & \text{if } (\nabla f(x))_i = 0 \quad \text{and } l_i > -\infty \text{ or } u_i < \infty, \\ 1 & \text{otherwise.} \end{cases} \quad (4)$$

We remark that this scaling matrix is possibly discontinuous in points x for which $\nabla f(x)_i = 0$ for some component i .

The original scaling matrix proposed in [7] was generalized in [13]. In this latter paper it is shown that first order optimality conditions for problem (2) are given by (3) for any diagonal matrix $D(x)$ with diagonal elements satisfying

$$d_i(x) \begin{cases} = 0 & \text{if } x_i = l_i \text{ and } \nabla f(x)_i > 0, \\ = 0 & \text{if } x_i = u_i \text{ and } \nabla f(x)_i < 0, \\ \geq 0 & \text{if } x_i \in \{l_i, u_i\} \text{ and } \nabla f(x)_i = 0, \\ > 0 & \text{otherwise.} \end{cases} \quad (5)$$

It is straightforward to note that elements given by (4) satisfy (5).

In this paper, we consider a quite general scaling matrix $D(x)$ satisfying (5). In fact, in the next section, we will state that, for global convergence purposes $D(x)$ must satisfy suitable assumptions that are very natural in a constrained context. We will show that the choice (4) satisfies such requirements and other existing scaling matrices from the recent literature are allowed in our context.

We remark that our aim is to define an iterative method that generates a sequence $\{x_k\}$ of strictly feasible points converging to a solution of (1). Then,

given $x_k \in \text{int}(\Omega)$ at hand, in order to ensure a stepsize large enough to produce a strictly feasible point and an acceptable progress towards a solution, we need to move towards the interior of Ω along search directions well-angled with respect to the bounds. The direction of the scaled gradient d_k , i.e.

$$d_k = -D_k \nabla f_k, \quad (6)$$

is an useful tool to achieve our aim, allowing implicit handling of the bounds by means of the diagonal matrix D_k and providing global convergence, as it will be shown in the next section. However, in order to go beyond a scaled gradient direction and obtain fast convergence of the method, we adopt a trust-region strategy. At iteration k , given an iterate $x_k \in \text{int}(\Omega)$ and the trust-region size $\Delta_k > 0$, we consider the scaled bound-constrained trust-region problem

$$\min_{p \in \mathbb{R}^n} \{m_k(p) : \|G_k p\| \leq \Delta_k, \quad x_k + p \in \text{int}(\Omega)\} \quad (7)$$

where m_k is the linear model for $\|F(x)\|$ at x_k , i.e.

$$m_k(p) = \|F_k + F'_k p\|$$

and $G_k = G(x_k) \in \mathbb{R}^{n \times n}$ with $G : \mathbb{R}^n \mapsto \mathbb{R}^{n \times n}$. Here the two choices $G_k = I$ or $G_k = D_k^{-1/2}$ are allowed. The first choice of G yields the standard spherical trust-region problem. The second one leads to an elliptical trust-region problem.

In the context of large scale problems, a suitable way of approximating the solution of (7) is a dogleg-type method where an inexact Newton method is employed. Let p_k^{IN} be an inexact Newton step satisfying

$$F'_k p_k^{IN} = -F_k + r_k, \quad \|r_k\| \leq \eta_k \|F_k\|, \quad (8)$$

where $\eta_k \in [0, 1)$ is called forcing term.

In order to guarantee that the new step is in the interior of Ω a projection, followed by a step back, is performed. That is, we consider the step \bar{p}_k^{IN} given by:

$$\bar{p}_k^{IN} = \alpha_k (P(x_k + p_k^{IN}) - x_k), \quad \alpha_k \in (0, 1), \quad (9)$$

where $P(x)_i = \max\{l_i, \min\{x_i, u_i\}\}$. We clearly have

$$\|\bar{p}_k^{IN}\| < \|p_k^{IN}\|. \quad (10)$$

To define the dogleg curve and find the next iterate, we look for the so-called generalized Cauchy point $p_c(\Delta_k)$ along the scaled gradient direction d_k , i.e. the minimizer of the model along d_k , constrained to be in the trust-region and to satisfy $x_k + p_c(\Delta_k) \in \text{int}(\Omega)$. The vector $p_c(\Delta_k)$ has the form

$$p_c(\Delta_k) = \tau_k d_k, \quad (11)$$

where d_k is given by (6) and the scalar τ_k is so that $x_k + p_c(\Delta_k) \in \text{int}(\Omega)$. More precisely, the value of τ_k is fixed in the following way. We consider

$$\tau'_k = \underset{\|\tau G_k d_k\| \leq \Delta_k}{\operatorname{argmin}} m_k(\tau d_k) = \min \left\{ -\frac{F_k^T F'_k d_k}{\|F'_k D_k \nabla f_k\|^2}, \frac{\Delta_k}{\|G_k D_k \nabla f_k\|} \right\} \quad (12)$$

and, if $x_k + \tau'_k d_k \in \text{int}(\Omega)$, we let $\tau_k = \tau'_k$ in (11). Otherwise we let λ_k be the stepsize along d_k to the boundary, i.e.

$$\lambda_k = \min_{1 \leq i \leq n} \Lambda_i, \quad \text{where} \quad \Lambda_i = \begin{cases} \max \left\{ \frac{l_i - (x_k)_i}{(d_k)_i}, \frac{u_i - (x_k)_i}{(d_k)_i} \right\} & \text{if } (d_k)_i \neq 0 \\ \infty & \text{if } (d_k)_i = 0 \end{cases}, \quad (13)$$

and set τ_k smaller than λ_k .

Summarizing, the parameter τ_k in (11) is given by

$$\tau_k = \begin{cases} \tau'_k & \text{if } x_k + \tau'_k d_k \in \text{int}(\Omega) \\ \theta \lambda_k, \theta \in (0, 1) & \text{otherwise.} \end{cases} \quad (14)$$

Now, we let

$$p_k^{gC} = - \frac{F_k^T F'_k d_k}{\|F'_k D_k \nabla f_k\|^2} d_k. \quad (15)$$

The vector p_k^{gC} is the *unconstrained* Cauchy step, i.e. the minimizer of the model along d_k without taking into account the trust-region constraint and the feasible set Ω . From the previous discussion it is straightforward to note that the generalized Cauchy step $p_c(\Delta_k)$ has always the form

$$p_c(\Delta_k) = \beta_k p_k^{gC}, \quad \beta_k \in (0, 1], \quad (16)$$

where $\beta_k = 1$ if $\tau_k = - \frac{F_k^T F'_k d_k}{\|F'_k D_k \nabla f_k\|^2}$ and $\beta_k < 1$ if $\tau_k = \frac{\Delta_k}{\|G_k D_k \nabla f_k\|}$ or $\tau_k = \theta \lambda_k$.

Then, $p_c(\Delta_k)$ is our reference direction that produces a reasonable progress towards satisfying

$$\lim_{k \rightarrow \infty} \|D_k \nabla f_k\| = 0.$$

Next, in order to produce an approximate trust-region step, we follow [23] and consider the linear path $p(\gamma)$ given by:

$$p(\gamma) = (1 - \gamma) p_c(\Delta_k) + \gamma \bar{p}_k^{IN}, \quad \gamma \in \mathbb{R}. \quad (17)$$

Note that γ is not necessarily confined to the interval $[0, 1]$, as in a standard dogleg approach: this is due to the fact that, as the Newton step is not exact and it may be projected, many theoretical properties of the dogleg curve are lost and we allow more flexibility in the choice of the step (see [23]). We point out a (slight) difference with the approach in [23]: here, the dogleg is based on $p_c(\Delta_k)$, whereas in [23] it is based on p_k^{gC} . This difference is due to the fact that we cannot simply consider p_k^{gC} as we have to take into account the presence of the bounds.

Then, we minimize our model along $p(\gamma)$ within the trust-region and the strictly feasible set. That is, we consider the function

$$\begin{aligned} \phi(\gamma) &= \|F_k + F'_k p(\gamma)\| \\ &= \|F_k + (1 - \gamma) F'_k p_c(\Delta_k) + \gamma F'_k \bar{p}_k^{IN}\| \\ &= \|F_k + F'_k p_c(\Delta_k) + \gamma F'_k (\bar{p}_k^{IN} - p_c(\Delta_k))\|. \end{aligned}$$

The function $\phi(\gamma)$ is convex, as it is the composition of an affine function and a convex function [6].

For the sake of simplicity let us set $a = F_k + F'_k p_c(\Delta_k)$ and $b = F'_k(\bar{p}_k^{IN} - p_c(\Delta_k))$. We have

$$\phi'(\gamma) = \frac{\sum_{i=1}^n (a_i + \gamma b_i) b_i}{\|a + \gamma b\|} = \frac{a^T b + \gamma b^T b}{\|a + \gamma b\|}.$$

Hence $\phi'(\gamma) = 0$ for $\gamma = \hat{\gamma}$ with

$$\hat{\gamma} = -\frac{a^T b}{b^T b} = -\frac{(F_k + F'_k p_c(\Delta_k))^T F'_k(\bar{p}_k^{IN} - p_c(\Delta_k))}{\|F'_k(\bar{p}_k^{IN} - p_c(\Delta_k))\|^2}. \quad (18)$$

Further, we have

$$\phi''(\gamma) = \frac{b^T b \|a + \gamma b\| - (a^T b + \gamma b^T b) \phi'(\gamma)}{\|a + \gamma b\|^2}.$$

Hence

$$\phi''(\hat{\gamma}) = \frac{b^T b}{\|a + \gamma b\|} > 0$$

and $\phi(\gamma)$ is therefore attaining a minimum at $\hat{\gamma}$.

Let us observe that $\hat{\gamma} > 0$ whenever $\|F_k + F'_k p_c(\Delta_k)\| > \|F_k + F'_k \bar{p}_k^{IN}\|$. In fact,

$$\begin{aligned} a^T b &= (F_k + F'_k p_c(\Delta_k))^T ((F_k + F'_k \bar{p}_k^{IN}) - F_k - F'_k p_c(\Delta_k)) \\ &= (F_k + F'_k p_c(\Delta_k))^T (F_k + F'_k \bar{p}_k^{IN}) - \|F_k + F'_k p_c(\Delta_k)\|^2 \\ &\leq \|F_k + F'_k p_c(\Delta_k)\| (\|F_k + F'_k \bar{p}_k^{IN}\| - \|F_k + F'_k p_c(\Delta_k)\|). \end{aligned}$$

Now, since we need $p(\gamma)$ in the trust-region, we compute the values of γ for which $p(\gamma)$ intersects the trust-region boundary.

Taking squares of $\|G_k p(\gamma)\| = \Delta_k$, we obtain

$$(1 - \gamma)^2 \|G_k p_c(\Delta_k)\|^2 + \gamma^2 \|G_k \bar{p}_k^{IN}\|^2 + 2\gamma(1 - \gamma) p_c(\Delta_k)^T G_k^2 \bar{p}_k^{IN} = \Delta_k^2$$

which rearranged gives

$$\gamma^2 \|G_k(p_c(\Delta_k) - \bar{p}_k^{IN})\|^2 - 2\gamma p_c(\Delta_k)^T G_k^2(p_c(\Delta_k) - \bar{p}_k^{IN}) + \|G_k p_c(\Delta_k)\|^2 - \Delta_k^2 = 0 \quad (19)$$

Then, two values γ_{\pm} , given by

$$\begin{aligned} \gamma_{\pm} &= \left(p_c(\Delta_k)^T G_k^2(p_c(\Delta_k) - \bar{p}_k^{IN}) \pm \left((p_c(\Delta_k)^T G_k^2(p_c(\Delta_k) - \bar{p}_k^{IN}))^2 - \right. \right. \\ &\quad \left. \left. \|G_k(p_c(\Delta_k) - \bar{p}_k^{IN})\|^2 (\|G_k p_c(\Delta_k)\|^2 - \Delta_k^2) \right)^{\frac{1}{2}} \right) / \|G_k(p_c(\Delta_k) - \bar{p}_k^{IN})\|^2 \end{aligned}$$

solve the above equation. Note that condition $p_c(\Delta_k) \neq \bar{p}_k^{IN}$ is always verified, as otherwise the path $p(\gamma)$ degenerates to the Cauchy point. Then, the argument

of the square root in γ_{\pm} is non negative if $\|G_k p_c(\Delta_k)\| \leq \Delta_k$. Clearly, we cannot have $\|G_k p_c(\Delta_k)\| > \Delta_k$, so existence of two real solutions of (19) is ensured.

Let us comment on the case $\|G_k p_c(\Delta_k)\| = \Delta_k$. We have

$$\gamma_{\pm} = \frac{p_c(\Delta_k)^T G_k^2 (p_c(\Delta_k) - \bar{p}_k^{IN}) \pm |p_c(\Delta_k)^T G_k^2 (p_c(\Delta_k) - \bar{p}_k^{IN})|}{\|G_k (p_c(\Delta_k) - \bar{p}_k^{IN})\|^2}$$

i.e., one of the two solutions is given by $\gamma = 0$, that is $p(\gamma) = p_c(\Delta_k)$. Indeed, if the Cauchy point lies on the boundary of the trust-region ($\|G_k p_c(\Delta_k)\| = \Delta_k$), one of the solutions of (19) is trivially the Cauchy point itself. Furthermore, if $G_k p_c(\Delta_k)$ and $G_k (p_c(\Delta_k) - \bar{p}_k^{IN})$ are orthogonal vectors, the path $p(\gamma)$ cuts the boundary of the trust-region in a unique point corresponding to $p_c(\Delta_k)$. Then, the path degenerates to the Cauchy point. We underline that in finite precision this event is unlikely to happen.

Finally, we have to take into account that $p(\gamma)$ is required to produce a strictly feasible point, as our aim is to generate a strictly feasible sequence. We note that the new point $x_k + p(\gamma)$ belongs to the interior of Ω if $\gamma \in [0, 1]$, because both $p_c(\Delta_k)$ and \bar{p}_k^{IN} are feasible steps. On the other hand, if we move along $p(\gamma)$ with γ negative or $\gamma > 1$ we need to check if strict feasibility of the new point is maintained and shorten the step if necessary. Recalling that

$$p(\gamma) = p_c(\Delta_k) + \gamma(\bar{p}_k^{IN} - p_c(\Delta_k)),$$

let us consider the stepsize to the boundary from $x_k + p_c(\Delta_k)$ along $\bar{p}_k^{IN} - p_c(\Delta_k)$. Then, if $\gamma > 1$, we set:

$$\Lambda_i = \begin{cases} \max \left\{ \frac{l_i - ((x_k)_i + (p_c(\Delta_k))_i)}{(\bar{p}_k^{IN} - p_c(\Delta_k))_i}, \frac{u_i - ((x_k)_i + (p_c(\Delta_k))_i)}{(\bar{p}_k^{IN} - p_c(\Delta_k))_i} \right\} & \text{if } (\bar{p}_k^{IN} - p_c(\Delta_k))_i \neq 0 \\ +\infty & \text{if } (\bar{p}_k^{IN} - p_c(\Delta_k))_i = 0 \end{cases}$$

and take

$$\bar{\gamma}_+ = \min_i \Lambda_i(p), \quad (20)$$

whereas if $\gamma < 0$ we set:

$$\Lambda_i = \begin{cases} \max \left\{ \frac{l_i - ((x_k)_i + (p_c(\Delta_k))_i)}{-(\bar{p}_k^{IN} - p_c(\Delta_k))_i}, \frac{u_i - ((x_k)_i + (p_c(\Delta_k))_i)}{-(\bar{p}_k^{IN} - p_c(\Delta_k))_i} \right\} & \text{if } (-\bar{p}_k^{IN} + p_c(\Delta_k))_i \neq 0 \\ +\infty & \text{if } (-\bar{p}_k^{IN} + p_c(\Delta_k))_i = 0 \end{cases}$$

and

$$\bar{\gamma}_- = -\min_i \Lambda_i(p). \quad (21)$$

To summarize, the choice of γ is made as follows. Since we want to minimize $\|F_k + F_k' p(\gamma)\|$, we seek $\gamma = \hat{\gamma}$. Moreover, since $p(\gamma)$ must belong to the trust-region and $x_k + p(\gamma)$ is required to be strictly feasible, we perform the following choice: if $\hat{\gamma} > 0$, we choose $\gamma = \min(\hat{\gamma}, \bar{\gamma}_+, \theta \bar{\gamma}_+)$, whereas if $\hat{\gamma} < 0$, we choose $\gamma = \max(\hat{\gamma}, \bar{\gamma}_-, \theta \bar{\gamma}_-)$, with $\theta \in (0, 1)$ and we set the trial step $p(\Delta_k) = p(\gamma)$.

Next, we sketch the process for finding $p(\Delta_k)$. It is based on Procedure 3.6 of [23], modified in order to take into account the presence of bounds. We

underline that this procedure generates a step that produces a decrease in the value of the model which is at least the decrease provided by the generalized Cauchy step. In other words, the step satisfies the condition

$$\rho_c(p(\Delta_k)) = \frac{\|F_k\| - \|F_k + F'_k p(\Delta_k)\|}{\|F_k\| - \|F_k + F'_k p_c(\Delta_k)\|} \geq 1. \quad (22)$$

STEP SELECTION. FINDING A STEP $p(\Delta_k)$ THAT SATISFIES THE MODEL DECREASE (22).

Input parameters: $x_k \in \text{int}(\Omega)$, $\Delta_k > 0$, d_k , \bar{p}_k^{IN} , $\theta \in (0, 1)$

Compute $p_c(\Delta_k)$ by (11) and (14).

Compute $\hat{\gamma}$ by (18).

If $\hat{\gamma} > 0$

compute γ_+ and $\bar{\gamma}_+$ and set $\gamma = \min\{\hat{\gamma}, \gamma_+, \theta\bar{\gamma}_+\}$

Else

compute γ_- and $\bar{\gamma}_-$ and set $\gamma = \min\{\hat{\gamma}, \gamma_-, \theta\bar{\gamma}_-\}$

Set $p(\Delta_k) = (1 - \gamma)p_c(\Delta_k) + \gamma\bar{p}_k^{IN}$

As for the unconstrained problems, the sufficient improvement condition

$$\rho_f(p(\Delta_k)) = \frac{\|F_k\| - \|F(x_k + p(\Delta_k))\|}{\|F_k\| - \|F_k + F'_k p(\Delta_k)\|} \geq \beta \quad (23)$$

is required to hold for a given constant $\beta \in (0, 1)$. Namely, if (23) is satisfied, then $p(\Delta_k)$ is accepted, the new iterate $x_{k+1} = x_k + p(\Delta_k)$ is formed and the trust-region radius may be increased. Otherwise, $p(\Delta_k)$ is rejected and Δ_k is shrunk.

Below we summarize the overall procedure.

AFFINE SCALING INEXACT DOGLEG (AS_ID) METHOD

Input parameters: the starting point $x_0 \in \text{int}(\Omega)$, the function G , $\Delta_{min} > 0$, the initial trust-region size $\bar{\Delta}_0 > \Delta_{min}$, β , δ , $\theta \in (0, 1)$.

For $k = 0, 1, \dots$

1. Set $\Delta_k = \bar{\Delta}_k$.
2. Choose $\alpha_k \in (0, 1)$, $\eta_k \in [0, 1)$.
3. Compute the solution p_k^{IN} to (8).
4. Form \bar{p}_k^{IN} by (9).
5. Set $d_k = -D_k \nabla f_k$.
6. Find $p(\Delta_k)$ by the Step Selection procedure.
7. While $\rho_f(p(\Delta_k)) < \beta$
 - 7.1 Set $\Delta_k = \delta \Delta_k$.
 - 7.2 Find $p(\Delta_k)$ by the Step Selection procedure.
8. Set $x_{k+1} = x_k + p(\Delta_k)$.
9. Choose $\bar{\Delta}_{k+1} > \Delta_{min}$.

Now, we make some comments that mainly focus on some algorithmic questions. First, we point out that the trust region size is updated according to standard rules, i.e. on the basis of accuracy of the adopted model to the merit function. So, at each iteration Δ_k is enlarged if (23) is satisfied. Otherwise, the trust region radius is reduced. The positive constant Δ_{min} is employed as a lower bound on the initial trust region size allowed at each iteration.

We underline that \bar{p}_k^{IN} does not depend on the trust-region radius and so it is computed only once at each iteration, even if reductions of the trust-region radius are needed.

Furthermore, it is worth mentioning that, as suggested in [23], before computing the inexact Newton step we could first check if the Cauchy step $p_c(\bar{\Delta}_k)$ satisfies the inexact Newton condition. If this is the case, the Cauchy step itself is an inexact Newton step. Furthermore, condition (22) is certainly satisfied and no linear system needs be solved in order to find \bar{p}_k^{IN} and consequently $p(\Delta_k)$. This may lead to computer time savings.

Another important point is that each iteration of the above method is well-defined because the while-loop at step 7 cannot continue indefinitely. In fact, we prove that there exists a sufficiently small Δ_k such that condition (23) is verified. So, after at most a finite numbers of repetitions, the while-loop terminates (see the following Proposition 3.1).

Finally, we underline that our method cannot be straightforwardly implemented in a matrix-free manner. In fact, the gradient of $\|F\|^2$ (i.e. the product $F'(x)^T F(x)$) must be computed and the sign of its components is of crucial importance for a good behaviour of the method.

3 Convergence analysis

The convergence analysis of AS_ID method is organized as follows. First, in Section 3.1 we will prove well-posedness of the scheme, i.e. finite termination of the while loop in step 7 of AS_ID method. Then, in Section 3.2 global convergence is proved without any Lipschitz assumption on F' . This distinguishes the results given here from those of [1], as global convergence is proved under weaker assumptions. Finally, in Section 3.4 the local convergence properties of the method are investigated, after introducing some technical results in Section 3.3.

3.1 Finite termination

Finite termination of the while loop in step 7 of AS_ID method is proved under the following assumptions.

Assumption 1: The sequence $\{x_k\}$ generated by the AS_ID method is bounded

Assumption 2: The scaling matrix $D(x)$:

- (i) satisfies (5);

- (ii) is bounded in $\Omega \cap B_\rho(x)$ for any $x \in \Omega$ and $\rho > 0$;
- (iii) there exists a $\bar{\lambda} > 0$ such that the stepsize λ_k to the boundary from x_k along d_k (see (13)) satisfies $\lambda_k > \bar{\lambda}$ whenever $\|\nabla f_k\|$ is uniformly bounded above.

Assumption 3: $F'(x)$ is uniformly continuous in Ω .

Note that Assumption 2-(iii) implies the constraint compatibility of d_k : this property avoids the problem of running directly into a bound by ensuring that the stepsize to the boundary remains bounded away from zero. Furthermore, it is straightforward to note that, as $D(x)$ satisfies (5), it is nonsingular for $x \in \text{int}(\Omega)$.

We remark that the scaling matrix (4) proposed by Coleman and Li, satisfies Assumption 2. In fact, we have already noted that it satisfies (5). Moreover, condition (ii) holds as for any $\bar{x} \in B_\rho(x)$, we have $u_i - x_i \leq u_i - \bar{x}_i + \rho$ and $x_i - l_i \leq \bar{x}_i + \rho - l_i$. Note that the term $u_i - x_i$ appears only if u_i is finite and the same is true for $x_i - l_i$. Finally, note that in the definition (13) of λ_k , for any $1 \leq i \leq n$ the maximum is $\frac{l_i - (x_k)_i}{(d_k)_i}$ whenever $(\nabla f_k)_i > 0$ and $\frac{u_i - (x_k)_i}{(d_k)_i}$ whenever $(\nabla f_k)_i < 0$. Then, from the definition of the scaling matrix (4) we have

$$\lambda_k = \min_i \left(\frac{1}{|(\nabla f_k)_i|} \right) \geq \frac{1}{\|\nabla f_k\|_\infty},$$

and condition (iii) is satisfied.

The following technical lemma paves the way for proving finite termination of the while loop in Step 7 of AS_ID method.

Lemma 3.1 *Let x_k be generated by Method AS_ID and assume that $\|D_k^{1/2}\nabla f_k\| \neq 0$. Then*

$$\|F_k\| - \|F_k + F'_k p(\Delta_k)\| \geq (1 - \sqrt{1 - \omega_k}) \|F_k\| \quad (24)$$

with

$$\omega_k = \min \left(\theta \lambda_k, \frac{\Delta_k}{\|G_k D_k \nabla f_k\|}, \frac{1}{\|F'_k\|^2 \|D_k\|} \right) \frac{\|D_k^{1/2} \nabla f_k\|^2}{\|F_k\|^2}. \quad (25)$$

Proof. First of all, we note that $\|F_k + F'_k p_c(\Delta_k)\| < \|F_k\|$ and from (22) it follows that

$$\|F_k\| - \|F_k + F'_k p(\Delta_k)\| \geq \|F_k\| - \|F_k + F'_k p_c(\Delta_k)\|. \quad (26)$$

Then, we proceed proving some inequalities for the Cauchy step $p_c(\Delta_k)$. Let us set

$$\eta_k^c := \frac{\|F_k + F'_k p_c(\Delta_k)\|}{\|F_k\|}.$$

First, assume that the step has the form $p_c(\Delta_k) = p_k^{gC}$, given in (15), i.e. $p_c(\Delta_k) = \sigma_k d_k$ with

$$\sigma_k := -\frac{F_k^T F'_k d_k}{\|F'_k D_k \nabla f_k\|^2} = -\frac{F_k^T F'_k d_k}{\|F'_k d_k\|^2}.$$

Then, we have

$$\begin{aligned}
(\eta_k^c)^2 &= \frac{\|F_k + F'_k D_c(\Delta_k)\|^2}{\|F_k\|^2} = \frac{\|F_k + \sigma_k F'_k d_k\|^2}{\|F_k\|^2} \\
&= \frac{F_k^T F_k + 2\sigma_k F_k^T F'_k d_k + \sigma_k^2 d_k^T (F'_k)^T F'_k d_k}{\|F_k\|^2} \\
&= 1 + 2\sigma_k \frac{F_k^T F'_k d_k}{\|F_k\|^2} + \sigma_k^2 \frac{\|F'_k d_k\|^2}{\|F_k\|^2} \\
&= 1 - 2 \frac{(F_k^T F'_k d_k)^2}{\|F'_k d_k\|^2 \|F_k\|^2} + \frac{(F_k^T F'_k d_k)^2 \|F'_k d_k\|^2}{\|F'_k d_k\|^4 \|F_k\|^2} \\
&= 1 - \left(\frac{F_k^T F'_k d_k}{\|F'_k d_k\| \|F_k\|} \right)^2 = 1 - \left(\frac{\|D_k^{1/2} \nabla f_k\|^2}{\|F'_k d_k\| \|F_k\|} \right)^2 \\
&\leq 1 - \left(\frac{\|D_k^{1/2} \nabla f_k\|}{\|F'_k\| \|F_k\| \|D_k^{1/2}\|} \right)^2. \tag{27}
\end{aligned}$$

Next, assume that $p_c(\Delta_k)$ has the form $p_c(\Delta_k) = \sigma_k d_k$ with

$$\sigma_k = \min \left(\theta \lambda_k, \frac{\Delta_k}{\|G_k D_k \nabla f_k\|} \right) < - \frac{F_k^T F'_k d_k}{\|F'_k d_k\|^2}.$$

Then, proceeding as to prove (27) we get:

$$\begin{aligned}
(\eta_k^c)^2 &= 1 + \sigma_k \left(2 \frac{F_k^T F'_k d_k}{\|F_k\|^2} + \sigma_k \frac{\|F'_k d_k\|^2}{\|F_k\|^2} \right) \\
&\leq 1 + \sigma_k \left(2 \frac{F_k^T F'_k d_k}{\|F_k\|^2} - \frac{(F_k^T F'_k d_k) \|F'_k d_k\|^2}{\|F'_k d_k\|^2 \|F_k\|^2} \right) \\
&= 1 - \min \left(\theta \lambda_k, \frac{\Delta_k}{\|G_k D_k \nabla f_k\|} \right) \frac{\|D_k^{1/2} \nabla f_k\|^2}{\|F_k\|^2}. \tag{28}
\end{aligned}$$

Relations (27) and (28) straightforwardly give

$$(\eta_k^c)^2 \leq 1 - \omega_k,$$

with ω_k given in (25). Note that, as $(\eta_k^c)^2$ is positive, it follows that $0 < \omega_k < 1$. The previous inequality immediately yields (24). \square

Now, we are ready to show that the while loop terminates.

Proposition 3.1 *Let x_k be generated by Method AS-ID and assume that $\|D_k^{1/2} \nabla f_k\| \neq 0$. Then the while loop in Step 7 terminates.*

Proof. In order to prove the thesis we need to consider the quantity

$$\rho_f(p(\Delta_k)) = 1 - \frac{\|F(x_k + p(\Delta_k))\| - \|F_k + F'_k p(\Delta_k)\|}{\|F_k\| - \|F_k + F'_k p(\Delta_k)\|}.$$

Taylor's theorem gives

$$F(x_k + p(\Delta_k)) = F(x_k) + F'_k p(\Delta_k) + \int_0^1 (F'(x_k + tp(\Delta_k)) - F'_k) p(\Delta_k) dt,$$

then the triangle inequality yields

$$\|F(x_k + p(\Delta_k))\| \leq \|F_k + F'_k p(\Delta_k)\| + \|p(\Delta_k)\| \int_0^1 \|F'(x_k + tp(\Delta_k)) - F'_k\| dt \quad (29)$$

and we obtain:

$$\|F(x_k + p(\Delta_k))\| - \|F_k + F'_k p(\Delta_k)\| \leq \|G_k^{-1}\| \Delta_k \int_0^1 \|F'(x_k + tp(\Delta_k)) - F'_k\| dt \quad (30)$$

Moreover, for

$$\Delta_k \leq \min \left(\theta \lambda_k, \frac{1}{\|F'_k\|^2 \|D_k\|} \right) \|G_k D_k \nabla f_k\|$$

from (25) we get:

$$\omega_k = \frac{\Delta_k \|D_k^{1/2} \nabla f_k\|^2}{\|G_k D_k \nabla f_k\| \|F_k\|^2}.$$

Note that, for $\omega_k \in (0, 1)$ it follows $1 - \sqrt{1 - \omega_k} > \frac{1}{2} \omega_k$. Then (24) yields

$$\|F_k\| - \|F_k + F'_k p(\Delta_k)\| \geq \frac{1}{2} \Delta_k \frac{\|D_k^{1/2} \nabla f_k\|^2}{\|G_k D_k \nabla f_k\| \|F_k\|}.$$

Then, from (30) we obtain

$$\rho_f(p(\Delta_k)) \geq 1 - 2 \frac{\|G_k^{-1}\| \int_0^1 \|F'(x_k + tp(\Delta_k)) - F'_k\| dt}{\frac{\|D_k^{1/2} \nabla f_k\|^2}{\|G_k D_k \nabla f_k\| \|F_k\|}}.$$

Since F' is continuous, the limit of the right-hand-side of the previous inequality goes to 1 as Δ_k tends to 0, therefore for Δ_k sufficiently small we have $\rho_f(p(\Delta_k)) > \beta$ and the while loop terminates. \square

3.2 Global convergence

Now we are ready to prove global convergence. Following [23], our proof is based on [9, Corollary 3.6], which is reported here for the reader's convenience.

Theorem 3.1 [9, Corollary 3.6] *Let $F : X \mapsto \mathbb{R}^n$ be continuously differentiable and assume that $\{x_k\} \in X$ is such that the conditions*

$$\|F_k\| - \|F(x_k + p_k)\| \geq \beta_1 (\|F_k\| - \|F_k + F'_k p_k\|) \quad (31)$$

$$\|F_k\| - \|F_k + F'_k p_k\| \geq 0 \quad (32)$$

are satisfied for each k , with $\beta_1 \in (0, 1)$ independent of k and $p_k = x_{k+1} - x_k$. If

$$\sum_{k \geq 0} \frac{\text{pred}_k}{\|F_k\|} = \sum_{k \geq 0} \frac{\|F_k\| - \|F_k + F'_k p_k\|}{\|F_k\|}$$

diverges, then $F_k \rightarrow 0$. If in addition x^* is a limit point of $\{x_k\}$ such that $F'(x^*)$ is invertible, then $F(x^*) = 0$ and $x_k \rightarrow x^*$.

Our convergence result is stated in the following Theorem.

Theorem 3.2 *Let Assumptions 1 and 2 be satisfied. Then all the limit points of $\{x_k\}$ are stationary points for problem (2). Further, if there exists a limit point $x^* \in \text{int}(\Omega)$ of $\{x_k\}$ such that $F'(x^*)$ is nonsingular, then $\|F_k\| \rightarrow 0$ and all the accumulation points of $\{x_k\}$ solve problem (1). If, in addition, there exists a limit point $x^* \in \Omega$ such that $F(x^*) = 0$ and $F'(x^*)$ is invertible, then $x_k \rightarrow x^*$.*

Proof. Let x^* be a limit point of $\{x_k\}$ and suppose that x^* is not a stationary point. Then $F(x^*) \neq 0$ and

$$D(x^*)\nabla f(x^*) \neq 0.$$

In particular, for at least an index $i \in \{1, \dots, n\}$ we have $d_i(x^*) \neq 0$ and $\nabla f(x^*)_i \neq 0$. We also have as an immediate consequence

$$D^{1/2}(x^*)\nabla f(x^*) \neq 0.$$

Looking at (5), this means we are in one of the following three situations: i) $l_i < x_i^* < u_i$, ii) $x_i^* = l_i$ and $\nabla f(x^*)_i < 0$, iii) $x_i^* = u_i$ and $\nabla f(x^*)_i > 0$. Due to the continuity of $\nabla f(x)$, there exists a neighborhood \tilde{N} of x^* such that $\nabla f(x^*)_i \nabla f(x)_i > 0 \forall x \in \tilde{N} \cap \Omega$. Let us consider case i) and let $\varepsilon = \min\{x_i^* - l_i, u_i - x_i^*\}$ and $N_* \subset \tilde{N}$ such that $N_* \cap \Omega \subset B_{\varepsilon/2}(x^*)$. Then $d_i(x)\nabla f(x)_i \neq 0$ for any x in $N_* \cap \Omega$. Then, we have $D(x)\nabla f(x) \neq 0$ and $D(x)^{1/2}\nabla f(x) \neq 0$ for any $x \in N_* \cap \Omega$. The same result is obtained in case ii) and case iii) using analogous arguments, letting $\varepsilon = u_i - x_i^*$ and $\varepsilon = x_i^* - l_i$, respectively.

Further by continuity we have

$$\sup_{x \in N_* \cap \Omega} \|F'(x)\| < +\infty.$$

The previous arguments imply that there exist constants $m > 0$ and $M > 0$ such that $\forall x \in N_* \cap \Omega$ we have

$$\|D(x)^{1/2}\nabla f(x)\| \geq m, \quad \|F'(x)\| \leq M.$$

Assumption 2 implies that there exist constants $\chi_D > 1$ and $\bar{\lambda} > 0$ such that

$$\|D_k^{1/2}\| \leq \chi_D, \quad \|G_k D_k^{1/2}\| \leq \chi_D,$$

and

$$\lambda_k > \bar{\lambda}$$

for any $x_k \in N_* \cap \Omega$. Hence for $x_k \in N_* \cap \Omega$, recalling (25) we have

$$\begin{aligned} \omega_k &= \frac{\|D_k^{1/2} \nabla f_k\|}{\|F_k\|^2} \min \left(\theta \lambda_k \|D_k^{1/2} \nabla f_k\|, \frac{\Delta_k \|D_k^{1/2} \nabla f_k\|}{\|G_k D_k \nabla f_k\|}, \frac{\|D_k^{1/2} \nabla f_k\|}{\|F_k'\|^2 \|D_k\|} \right) \\ &\geq \frac{m}{\|F_0\| \|F_k\|} \min \left(\theta \bar{\lambda} m, \frac{\Delta_k}{\chi_D}, \frac{m}{M^2 \chi_D^2} \right). \end{aligned} \quad (33)$$

Then, for

$$\Delta_k < m \chi_D \min \left(\theta \bar{\lambda}, \frac{1}{M^2 \chi_D^2} \right) =: \tilde{\Delta}$$

we have from (33)

$$\omega_k \geq \frac{\Delta_k m}{\chi_D \|F_0\| \|F_k\|}.$$

Therefore, as $1 - \sqrt{1 - \omega_k} > \frac{1}{2} \omega_k$, proceeding as in Proposition 3.1 we obtain:

$$\begin{aligned} \rho_f(p(\Delta_k)) &\geq 1 - \frac{2 \|G_k^{-1}\| \Delta_k \int_0^1 \|F'(x_k + tp(\Delta_k)) - F_k'\| dt}{\frac{\Delta_k m}{\chi_D \|F_0\| \|F_k\|} \|F_k\|} \\ &\geq 1 - 2 \frac{\chi_D^2}{m} \|F_0\| \int_0^1 \|F'(x_k + tp(\Delta_k)) - F_k'\| dt. \end{aligned}$$

The uniform continuity of F' implies that there exists $\hat{\Delta} < \tilde{\Delta}$ such that for $\Delta_k \leq \hat{\Delta}$ we have $\rho_f(p(\Delta_k)) > \beta$ for any $x_k \in N_* \cap \Omega$. This implies, for the updating rule of the trust-region radius, that $\Delta_k > \delta \hat{\Delta}$. Then, from (24) and (33) we have

$$\frac{\|F_k\| - \|F_k + F_k' p(\Delta_k)\|}{\|F_k\|} \geq 1 - \sqrt{1 - \bar{\omega}},$$

with

$$\bar{\omega} = \frac{m}{\|F_0\|^2} \frac{\delta \hat{\Delta}}{\chi_D} > 0.$$

Since $x_k \in N_*$ for infinitely many k , the series

$$\sum_{k=1}^{\infty} \frac{pred_k}{\|F_k\|}$$

diverges.

Note that the sequence $\{x_k\}$ satisfies (31) and (32). Then we are in a position to apply Theorem 3.1 and to conclude that $F(x^*) = 0$. This is a contradiction and therefore x^* is a stationary point.

Moreover, $\|F_k\|$ is a bounded and strictly decreasing sequence, hence it is convergent, then, if there exists a limit point $x^* \in \text{int}(\Omega)$ such that $F'(x^*)$ is nonsingular, it follows that $\lim_{k \rightarrow \infty} \|F_k\| = 0$. Finally, if there exists $x^* \in \Omega$ such that $F(x^*) = 0$ and $F'(x^*)$ is invertible, we can apply Theorem 3.3 of [9] with $\eta = 1$, that proves the convergence of x_k to x^* . \square

3.3 Superlinear convergence: technical results

In this section we develop a theoretical foundation for proving superlinear convergence of the AS_ID method. We start recalling from [4] some technical results tailored on the AS_ID method. In the sequel we assume the following:

Assumption 4: $\|F'\|$ is bounded above on

$$L = \cup_{k=0}^{\infty} \{x \in X : \|x - x_k\| \leq r\}, \quad r > 0,$$

and $\chi_J = \sup_{x \in L} \|F'(x)\|$.

Assumption 5: F' is Lipschitz continuous in an open, convex set containing L , with constant $2\gamma_L$.

Assumption 6: For any \bar{x} in $\text{int}(\Omega)$ there exists $\bar{\rho} > 0$ and $\chi_{\bar{x}}$ such that $B_{\bar{\rho}}(\bar{x}) \subset \text{int}(\Omega)$ and $\|D(x)^{-1}\| \leq \chi_{\bar{x}}$ for any x in $B_{\bar{\rho}/2}(\bar{x})$.

Note that Assumption 4 is always satisfied if Ω is compact as F' is continuous.

In the following lemma we report some useful inequalities whose proof is not reported here as it is a straightforward adaption of that of [4, Lemma 4.2].

Lemma 3.2 *Let $x^* \in \Omega$ be a limit point of the sequence of iterates $\{x_k\}$ generated by the AS_ID method such that $F(x^*) = 0$ and $F'(x^*)$ is nonsingular. Let $K_1 = 2\|F'(x^*)\|$, $K_2 = 2\|F'(x^*)^{-1}\|$, $\mu = \max\{K_1, K_2\}/2$, $\Gamma \in (0, 1/\mu)$ be given. Then, there exists $\rho > 0$ so that if $x \in B_{\rho}(x^*)$ then $x \in L$ and*

$$\|x - x^*\| \leq K_2 \|F(x)\|, \quad (34)$$

$$\|F(x)\| \leq K_1 \|x - x^*\|, \quad (35)$$

$$\|F'(x)^{-1}\| \leq K_2, \quad (36)$$

$$\|F(x) - F(z) - F'(z)(x - z)\| \leq \Gamma \|x - z\|^2 \text{ for all } z \in B_{\rho}(x^*). \quad (37)$$

In order to discuss the convergence rate issues, we make the additional hypothesis $\|G_k p_k^{IN}\| \rightarrow 0$ as $k \rightarrow \infty$. In practice, this condition may fail to hold only when $G_k = D_k^{-1/2}$ and x^* belongs to the boundary of Ω . On the other hand, it is guaranteed when $G_k = I$ or when $G_k = D_k^{-1/2}$ and x^* lies in the interior of Ω . To show this, note that by (8) and (36) we get

$$\|p_k^{IN}\| = \|F'_k{}^{-1}(-F_k + r_k)\| \leq (1 + \eta_k) \|F'_k{}^{-1}\| \|F_k\| \leq 2K_2 \|F_k\|, \quad (38)$$

whenever $x_k \in B_{\rho}(x^*)$. This implies that $\|p_k^{IN}\| \rightarrow 0$ as $k \rightarrow \infty$. Further, when $x^* \in \text{int}(\Omega)$, by Assumption 6 we have $\|D_k^{-1}\| \leq \chi_{x^*}$ for any $x_k \in B_{\bar{\rho}/2}(x^*)$. Letting $\tilde{\rho} = \min(\bar{\rho}, \rho)$, as x^* is a limit point of $\{x_k\}$, there exists \tilde{k} such that $x_k \in B_{\tilde{\rho}/2}(x^*)$ for any $k > \tilde{k}$. Then, eventually we have $\|D_k^{-1/2} p_k^{IN}\| \leq \chi_{x^*}^{1/2} \|p_k^{IN}\| \rightarrow 0$ as $k \rightarrow \infty$.

From now on, with γ_L and χ_J as in Assumptions 4-5, K_1 , K_2 and Γ as in Lemma 3.2, we let

$$K^* = \|F'(x^*)\| \|F'(x^*)^{-1}\|, \quad (39)$$

$$\nu = 8 K^* (K_2 \chi_J + 1), \quad (40)$$

$$\delta_k = K_2 \Gamma \nu^2 \|x_k - x^*\| + 4 K^* \eta_k, \quad (41)$$

$$\psi_k = \chi_J \delta_k + \gamma_L \nu^2 \|x_k - x^*\| + K_1 (1 - \alpha_k), \quad (42)$$

$$\sigma_k = \max\{\psi_k, K_2 (\Gamma \nu^2 \|x_k - x^*\| + \psi_k)\}. \quad (43)$$

In the next two lemmas we give some technical results which pave the way for proving fast convergence rate.

Lemma 3.3 *Assume that there exists a solution x^* of (1) such that $F'(x^*)$ is nonsingular and that the sequence $\{x_k\}$ generated by the AS_ID method converges to x^* . Suppose that*

- either $G_k = I$, $k \geq 0$, or
- $G_k = D_k^{-1/2}$, $k \geq 0$, and $\|G_k p_k^{IN}\| \rightarrow 0$ as $k \rightarrow \infty$.

Then there exists $\rho_1 \leq \rho$ such that for all $x_k \in B_{\rho_1}(x^*) \cap \text{int}(\Omega)$

$$\|G_k \bar{p}_k^{IN}\| \leq \bar{\Delta}_k, \quad (44)$$

where $\bar{\Delta}_k$ is the initial trust-region radius at k th iteration. Further, when $x_k \in B_{\rho_1}(x^*) \cap \text{int}(\Omega)$ we have

$$\|F_k + F'_k p_k^{IN}\| \leq K_1 \eta_k \|x_k - x^*\|, \quad (45)$$

$$\|\bar{p}_k^{IN}\| < \|p_k^{IN}\| \leq \nu \|x_k - x^*\|, \quad (46)$$

$$\|p(\bar{\Delta}_k)\| \leq \nu \|x_k - x^*\|. \quad (47)$$

Proof. Relation (44) is proven by using the fact that $\bar{\Delta}_k \geq \Delta_{min}$, i.e. $\bar{\Delta}_k$ is bounded below from zero for each $k \geq 0$. First let us consider the case $G_k = I$, $\forall k \geq 0$. Inequality (38) yields $\lim_{k \rightarrow \infty} \|p_k^{IN}\| = 0$. Then, with ρ as in Lemma 3.2, there exists $\rho_1 \leq \rho$ such that $\|\bar{p}_k^{IN}\| \leq \|p_k^{IN}\| \leq \bar{\Delta}_k$ when $x_k \in B_{\rho_1}(x^*) \cap \text{int}(\Omega)$ and the thesis (44) follows. Now consider the case $G_k = D_k^{-1/2}$, $\forall k \geq 0$. Noting that $|(\bar{p}_k^{IN})_i| \leq |(p_k^{IN})_i|$, we immediately have $\|G_k \bar{p}_k^{IN}\| \leq \|G_k p_k^{IN}\|$. Hence, from the assumption $\lim_{k \rightarrow \infty} \|G_k p_k^{IN}\| = 0$, we get (44).

The remaining results are proven independently of the form of G_k . Result (45) is easily proven as by (8) and (35) we obtain

$$\|F_k + F'_k p_k^{IN}\| \leq \eta_k \|F_k\| \leq K_1 \eta_k \|x_k - x^*\|.$$

Result (46) is derived noting that by (10), (38) and (35) we get

$$\|\bar{p}_k^{IN}\| < \|p_k^{IN}\| \leq 2 K_2 \|F_k\| \leq 8 K^* \|x_k - x^*\|. \quad (48)$$

Then, by (40) relation (46) follows.

Finally, we move on proving (47). We have

$$\|p(\bar{\Delta}_k)\| = \|p(\gamma)\| \leq \|p_c(\bar{\Delta}_k)\| + |\gamma| \|\bar{p}_k^{IN} - p_c(\bar{\Delta}_k)\|. \quad (49)$$

Let us consider the value of $|\gamma|$. From (18) we have

$$\begin{aligned} |\gamma| \leq |\hat{\gamma}| &\leq \frac{\|F_k + F'_k p_c(\bar{\Delta}_k)\|}{\|F'_k(\bar{p}_k^{IN} - p_c(\bar{\Delta}_k))\|} \\ &\leq \frac{\|F_k + F'_k p_c(\bar{\Delta}_k)\| \|(F'_k)^{-1}\|}{\|\bar{p}_k^{IN} - p_c(\bar{\Delta}_k)\|} \\ &\leq \frac{K_2 \|F_k + F'_k p_c(\bar{\Delta}_k)\|}{\|\bar{p}_k^{IN} - p_c(\bar{\Delta}_k)\|}. \end{aligned} \quad (50)$$

Moreover, recalling (11) we have

$$\begin{aligned} \|p_c(\bar{\Delta}_k)\| &= \tau_k \|D_k \nabla f_k\| \leq \frac{-F_k^T F'_k d_k}{\|F'_k D_k \nabla f_k\|^2} \|D_k \nabla f_k\| \\ &\leq \frac{\|F'_k\| \|F_k\| \|D_k \nabla f_k\|^2}{\|F'_k D_k \nabla f_k\|^2} \\ &\leq \|F'_k\| \|F_k\| \|(F'_k)^{-1}\|^2. \end{aligned}$$

From Assumption 4, (35) and (36), the last inequality yields

$$\|p_c(\bar{\Delta}_k)\| \leq \chi_J K_1 K_2^2 \|x_k - x^*\|. \quad (51)$$

Finally, letting

$$\eta_{gC} = \frac{\|F_k + F'_k p_k^{gC}\|}{\|F_k\|},$$

from (16) it follows:

$$\begin{aligned} \|F_k + F'_k p_c(\bar{\Delta}_k)\| &= \|F_k + F'_k \beta_k p_k^{gC}\| \\ &\leq (1 - \beta_k) \|F_k\| + \beta_k \eta_{gC} \|F_k\| \\ &= (1 - \beta_k (1 - \eta_{gC})) \|F_k\|. \end{aligned}$$

Then, from (35) and $\eta_{gC} \leq 1$, it follows:

$$\|F_k + F'_k p_c(\bar{\Delta}_k)\| \leq K_1 \|x_k - x^*\|.$$

Hence, by (49) and (50) we have

$$\begin{aligned} \|p(\bar{\Delta}_k)\| &\leq \|p_c(\bar{\Delta}_k)\| + K_2 \|F_k + F'_k p_c(\bar{\Delta}_k)\| \\ &\leq \chi_J K_1 K_2^2 \|x_k - x^*\| + K_2 K_1 \|x_k - x^*\| \\ &= K_1 K_2 (\chi_J K_2 + 1) \|x_k - x^*\|. \end{aligned}$$

This proves (47). \square

Lemma 3.4 *Assume that there exists a solution x^* of (1) such that $F'(x^*)$ is nonsingular and that the sequence $\{x_k\}$ generated by the AS_ID method converges to x^* . Suppose that*

- either $G_k = I$, $k \geq 0$, or
- $G_k = D_k^{-1/2}$, $k \geq 0$, and $\|G_k p_k\| \rightarrow 0$ as $k \rightarrow \infty$.

Then, there exists ρ_2 such that for all $x_k \in B_{\rho_2}(x^) \cap \text{int}(\Omega)$ we have*

$$\|F_k + F'_k \bar{p}_k^{IN}\| \leq \sigma_k \|x_k - x^*\|, \quad (52)$$

$$\|x_k + \bar{p}_k^{IN} - x^*\| \leq \sigma_k \|x_k - x^*\|. \quad (53)$$

Proof. The thesis straightforwardly follows from [4, Lemma 4.4] replacing $p_{tr}(\bar{\Delta}_k)$ with p_k^{IN} and $\bar{p}_{tr}(\bar{\Delta}_k)$ with \bar{p}_k^{IN} . \square

3.4 Superlinear convergence: main result

In this section we show that, if $\eta_k \rightarrow 0$ as $k \rightarrow \infty$, eventually the step $p(\bar{\Delta}_k)$ satisfies condition (23). Then it is not necessary to reduce the trust-region radius. Moreover, for k sufficiently large, $p(\bar{\Delta}_k)$ is an inexact Newton step that provides a linear residual bounded by $\sigma_k \|x_k - x^*\|$. This yields the superlinear/quadratic convergence of the procedure.

Theorem 3.3 *Assume that there exists a solution x^* of (1) such that $F'(x^*)$ is nonsingular and that the sequence $\{x_k\}$ generated by the AS_ID method converges to x^* . Suppose that $\eta_k \rightarrow 0$, $\alpha_k \rightarrow 1$, as $k \rightarrow \infty$, and*

- either $G_k = I$, $k \geq 0$, or
- $G_k = D_k^{-1/2}$, $k \geq 0$, and $\|G_k p_k^{IN}\| \rightarrow 0$ as $k \rightarrow \infty$.

Then, eventually, $p(\bar{\Delta}_k)$ satisfies (23) and the sequence $\{x_k\}$ converges to x^ superlinearly. Moreover, if*

$$\eta_k = O(\|F_k\|), \quad \alpha_k = 1 - O(\|F_k\|) \quad \text{as } k \rightarrow \infty,$$

the convergence rate is quadratic.

Proof. Recall that, in the step selection rule, $p(\bar{\Delta}_k)$ is obtained minimizing the convex function $\phi(\gamma)$ within the intersection of the trust-region and the interior of the feasible set Ω . Further, $x_k + \bar{p}_k^{IN}$ belongs to the trust-region for $x_k \in B_{\rho_1}(x^*)$, with ρ_1 as in Lemma 3.3. Moreover it belongs to the path described by $\phi(\gamma)$ and it is a feasible point. From these arguments and from (52), it follows

$$\|F_k + F'_k p(\bar{\Delta}_k)\| \leq \|F_k + F'_k \bar{p}_k^{IN}\| \leq \sigma_k \|x_k - x^*\| \quad (54)$$

for $x_k \in B_{\rho_2}(x^*)$ with ρ_2 as in Lemma 3.4.

Now we prove that eventually $p(\bar{\Delta}_k)$ satisfies (23). First note that, from (47), we get

$$\|x_k + p(\bar{\Delta}_k) - x^*\| \leq \|x_k - x^*\| + \|p(\bar{\Delta}_k)\| \leq (1 + \nu)\|x_k - x^*\| \leq (1 + \nu)\rho_2 \leq \rho_1.$$

Hence, $x_k + p(\bar{\Delta}_k)$ belongs to $B_{\rho_1}(x^*)$. Then,

$$\begin{aligned} \rho_f(p(\bar{\Delta}_k)) &= \frac{\|F_k\| - \|F(x_k + p(\bar{\Delta}_k))\|}{\|F_k\| - \|F_k + F'_k p(\bar{\Delta}_k)\|} \\ &= \frac{\|F_k\| - \|F(x_k + p(\bar{\Delta}_k)) - F_k - F'_k p(\bar{\Delta}_k) + F_k + F'_k p(\bar{\Delta}_k)\|}{\|F_k\| - \|F_k + F'_k p(\bar{\Delta}_k)\|} \\ &\geq \frac{\|F_k\| - \|F(x_k + p(\bar{\Delta}_k)) - F_k - F'_k p(\bar{\Delta}_k)\| - \|F_k + F'_k p(\bar{\Delta}_k)\|}{\|F_k\| - \|F_k + F'_k p(\bar{\Delta}_k)\|} \\ &= 1 - \frac{\|F(x_k + p(\bar{\Delta}_k)) - F_k - F'_k p(\bar{\Delta}_k)\|}{\|F_k\| - \|F_k + F'_k p(\bar{\Delta}_k)\|} \\ &\geq 1 - \frac{\Gamma \|p(\bar{\Delta}_k)\|^2}{\|F_k\| - \|F_k + F'_k p(\bar{\Delta}_k)\|} \\ &\geq 1 - \frac{\Gamma \|p(\bar{\Delta}_k)\|^2}{\frac{1}{K_2} \|x_k - x^*\| - \sigma_k \|x_k - x^*\|} \\ &\geq 1 - \frac{K_2 \Gamma \nu^2 \|x_k - x^*\|^2}{(1 - K_2 \sigma_k) \|x_k - x^*\|} \\ &= 1 - \frac{K_2 \Gamma \nu^2 \|x_k - x^*\|}{1 - K_2 \sigma_k} \end{aligned}$$

Note that $\sigma_k \rightarrow 0$ and $\|x_k - x^*\| \rightarrow 0$. Then, for k large enough we certainly have

$$1 - \frac{K_2 \Gamma \nu^2 \|x_k - x^*\|}{1 - K_2 \sigma_k} > \beta.$$

Hence, $x_{k+1} = x_k + p(\bar{\Delta}_k)$.

From (34) and (37) we have

$$\begin{aligned} \|x_k + p(\bar{\Delta}_k) - x^*\| &\leq K_2 \|F(x_k + p(\bar{\Delta}_k))\| \\ &\leq K_2 (\|F(x_k + p(\bar{\Delta}_k)) - F_k - F'_k p(\bar{\Delta}_k)\| + \|F_k + F'_k p(\bar{\Delta}_k)\|) \\ &\leq K_2 (\Gamma \|p(\bar{\Delta}_k)\|^2 + \|F_k + F'_k p(\bar{\Delta}_k)\|). \end{aligned} \quad (55)$$

Further, from (54) and (47) we get

$$\begin{aligned} \|x_k + p(\bar{\Delta}_k) - x^*\| &\leq K_2 (\Gamma \|p(\bar{\Delta}_k)\|^2 + \sigma_k \|x_k - x^*\|) \\ &\leq K_2 (\Gamma \nu^2 \|x_k - x^*\| + \sigma_k) \|x_k - x^*\|. \end{aligned} \quad (56)$$

Since we have

$$\sigma_k = O(\|x_k - x^*\| + \eta_k + (1 - \alpha_k)), \quad k \rightarrow \infty,$$

(56) ensures superlinear convergence rate if $\eta_k \rightarrow 0$ and $\alpha_k \rightarrow 1$ as $k \rightarrow \infty$. Moreover, if $\eta_k = O(\|F_k\|)$ and $1 - \alpha_k = O(\|F_k\|)$, by (34)-(35) we get $\sigma_k = O(\|x_k - x^*\|)$ and this yields quadratic convergence rate. \square

4 Numerical experiments

In this section, we report on numerical experiments with the AS_ID method. Our aim is to prove the computational feasibility of the method, give general information about its numerical performance and demonstrate the feasible efficiency of the proposed inexact dogleg method compared with the globalization technique given in [4].

Several choices have been made here to perform this preliminary experimental investigation and we are conscious of the fact that all the algorithmic possibilities of the AS_ID method are not fully investigated here. We briefly discuss our choices and give details of the major issues addressed in the implementation of AS_ID method. Furthermore, we describe the set of test problems used for the numerical experiments, discuss the results obtained and compare them with those obtained by the SIATR_2D algorithm given in [4].

4.1 Our implementation

We implemented the method AS_ID in a Matlab code. The numerical experiments were performed on Intel Xeon (TM) 3.4 Ghz, 1GB RAM using Matlab 7.6. and machine precision $\epsilon_m \approx 2.10^{-16}$.

Several choices made here, although very reasonable from our point of view, were due to the aim of investigating the numerical efficiency of the proposed inexact dogleg compared with the subspace dogleg strategy (i.e. the SIATR_2D method) given in [4]. Then, we decided to verify the basic effectiveness of AS_ID algorithm by using the standard spherical trust-region and the pioneer scaling matrix given by Coleman and Li. Furthermore, we have chosen to test a Newton-GMRES implementation. That is, we used GMRES as iterative linear solver for computing p_k^{IN} . Also for sake of comparison, we implemented the SIATR_2D method in a Matlab code and run both solvers with the same values of the common parameters and the same choices of algorithmic options. More specifically, we set $\Delta_0 = 1$, $\Delta_{min} = \sqrt{\epsilon_m}$, $\beta = 0.75$, $\theta = 0.99995$, $\delta = 0.25$ and updated the trust-region size as in [4], i.e. at step 7.1 of the AS_ID algorithm we reduced the trust region radius by setting $\Delta_k = \min\{0.25 \Delta_k, 0.5 \|p_k\|\}$ and, at step 9, we allowed the next iteration with an increased trust-region radius if condition (23) holds (in this case, we set $\bar{\Delta}_{k+1} = \max\{\Delta_k, 2\|p_k\|\}$), otherwise, we left unchanged the radius.

The forcing term were computed by the adaptive choice given in [10], i.e. $\eta_0 = 0.9$, and $\eta_k = 0.9 \|F_k\|^2 / \|F_{k-1}\|^2$, for $k \geq 1$, with the safeguards suggested in [10, p. 305].

The Jacobian matrices were evaluated analytically and the inexact Newton step p_k^{IN} was computed by GMRES with the null initial guess and a restart parameter equal to 50. A maximum of 150 GMRES iterations was allowed for each nonlinear iteration. If all these linear iterations were performed, the algorithm continued with p_k^{IN} given by the last computed GMRES iterate.

For the computation of the projected step \bar{p}_k^{IN} we used $\alpha_k = \max\{0.95, 1 - \|F_k\|\}$ for all k .

We stopped all the runs when the condition

$$\|F_k\| \leq 10^{-6}$$

was met. Such occurrence was indicated as a successful termination. On the other hand, the code declares a failure either if the number of iterations was greater than 150 or if the number of F -evaluations was greater than 1000. In addition, a failure was declared if the trust-region size was reduced below 10^{-8} or if

$$\|F_{k+1} - F_k\| \leq 10^{-6}\|F_k\|.$$

This condition is commonly called stagnation and may indicate that the method does not manage to escape from a local minimizer of the merit function which is not a solution of (1). We remark that these two last situations never happened in our tests.

4.2 The test problems

The experiments were carried out on a set of widely studied problems. These problems have proved to be reasonably hard in the past and many of them are regarded as challenging problems in the literature. All problems have more than one solution and we fixed the bounds to select specified solutions. We performed our experiments using four different initial guesses x_0 for each test problem, for a total of 32 tests. In the following, we give a short description of the test problems. For each of them, we specify the source, the dimension selected, the lower and upper bounds specified, the different starting points used. For any vector $z \in \mathbb{R}^n$ and $a \in \mathbb{R}$ the notation $z = a$ is used as a shorthand for $z_i = a$, for $i = 1, \dots, n$.

P1. The Chandrasekhar H-equation. The discrete analogue of the Chandrasekhar H-equation [17, p. 87] is a nonlinear system that depends on a parameter $c \in (0, 1]$. This system has two solutions but only one has a physical meaning. The bounds $l = 0$, $u = \infty$ are used to select the significant solution. We set $n = 400$ and $c = 0.9999$ and we used the bounds $l = 0$, $u = 10^{300}$, to approximate the required solution. We solved this problem starting from $x_0 = 10^{\gamma-2}$, $\gamma = 0, 1, 2, 3$.

P2. Integral equation. A discretization of the integral equation given in [18] leads to a nonlinear system of equations of variable dimension. We fixed $n = 10^3$ and approximated the negative solution by setting $l = -10^{300}$, $u = 0$. We started from $x_0 = -10^{\gamma-2}$, $\gamma = 0, 1, 2, 3$.

P3. Five diagonal system. This nonlinear system is Problem 4.8 in [20]. It has variable dimension and several solutions. We decided to solve this problem with dimension $n = 5 \cdot 10^3$, bounds $l = 1$, $u = 10^{300}$, and initial guesses $x_0 = 1 + 10^{\gamma-2}$, $\gamma = 0, 1, 2, 3$.

P4. Seven diagonal system. This nonlinear system is Problem 4.9 in [20]. Again the problem has variable dimension and several solutions. We set $n = 5 \cdot 10^3$ and used the bounds $l = 0$, $u = 10^{300}$, to avoid negative solutions. The vectors $x_0 = 10^{\gamma-2}$, $\gamma = 0, 1, 2, 3$ have been used to start.

P5. Poisson problem. This nonlinear system of variable dimension follows from the standard five-points finite differences analogue of the nonlinear partial differential equation given in [20, Problem 4.25]. We solved this problem with $n = 10^4$, $l = -5$, $u = 5$, and initial guesses given by $x_0 = l + (\gamma + 1)(u - l)/5$, $\gamma = 0, 1, 2, 3$.

P6. Bratu problem. This is the standard five-points finite differences analogue of the nonlinear partial differential equation given in [20, Problem 4.24]. We set $n = 10^4$ and we fixed $l = -10^{300}$, $u = 1.5$, $i = 1, \dots, n$. This problem was solved starting from $x_0 = -10^{\gamma-2}$, $\gamma = 0, 1, 2, 3$.

P7. A chemical equilibrium system. The System 1 given in Ref. [22] is a system of $n = 11$ nonlinear equations. A solution of this problem has a physical meaning if all its components are real and positive. We augmented this problem to the size $n = 11 \cdot 10^3$, fixed the bounds $l = 0$ and $u = 10^{300}$ and solved the augmented problem starting from $x_0 = 10^{\gamma-2}$, $\gamma = 0, 1, 2, 3$.

P8. Bratu NCP The Bratu nonlinear complementarity problem [11] was reformulated as a system of $n = 12500$ nonlinear equations with $l = 0$ and $u = 10^{300}$. It depends on a parameter λ and we set $\lambda = 6$. The solution of this problem lies on the boundary of the feasible set and it was included in the tests set for testing the considered methods on problems with solutions on the boundary of Ω . We solved this problem starting from $x_0 = 10^{\gamma-2}$, $\gamma = 0, 1, 2, 3$.

We underline that in the solution of Problems **P5**, **P6** and **P8** the preconditioner given by the incomplete LU factorization $ILLU(0)$ of the Jacobian was used to speed up the convergence of GMRES.

4.3 Numerical results

In the following Table 1 we summarize the results of the numerical experiments. For the sake of comparison, the table includes results for our Matlab implementations of `AS_ID` algorithm and `SIATR_2D` algorithm.

For each problem, we give: the value (γ) of the parameter used to compute the initial guess x_0 and the value ($\|F(x_0)\|$) of $\|F\|$ at x_0 . Furthermore, for both solvers `AS_ID` and `SIATR_2D`, we show the number (It) of iteration performed by the algorithm, the number (Fe) of function evaluations, the value ($\|F(x_c)\|$) of $\|F\|$ at the computed solution x_c . Finally, the symbol * indicates a failure. We use the rather standard notation $m(c)$ to denote $m \cdot 10^c$.

We remark that we have chosen to measure the algorithm efficiency by the

number It of iterations and the number Fe of F -evaluations performed to obtain convergence. Note that the measure in terms of iteration count is equivalent to considering the number of function evaluations if the initial trial step is accepted and no reduction of the trust-region size is required. In fact, in this case we have $Fe = It + 1$. On the other hand, the number of F -evaluations increases when reductions of the trust-region size are needed. In this case, $Fe > It + 1$ and the quantity $R = Fe - (It + 1)$ gives the number of the performed reductions of the trust-region size. We remark that the computational cost of the algorithm increases as the number of rejected trial steps. Then, severe reductions of the trust-region size damage its numerical efficiency.

We now see in more details our numerical results. In Table 1, the values $\|F(x_0)\|$ show that we started all runs far from the solution. Both algorithms seldom failed (on a total of 32 tests we count 3 failures for AS_ID and 2 failures for SIATR_2D) and reached good approximations of the solution.

We note that the AS_ID method seems to be quite efficient. All the successful runs required at most $It = 27$ iterations and $Fe = 28$ function evaluations. On a total of 32 runs, these maximum values are reached only once. Further, we count $It > 20$ and $Fe > 20$ five times whereas 17 runs have been ended with It and Fe ranging from 10 to 20. We note that most problems have been successfully solved without employing reductions of the trust-region radius, that is the first trial step is always accepted. In fact, only three runs required reduction of the trust-region size. In particular, we count one test where two reductions have been performed ($R = 2$ in the solution of Pb. 4 when $\gamma = 3$ is used to compute x_0) and two tests requiring one reduction of the trust-region (it results $R = 1$ in the solution of Pb. 7 and Pb. 3 starting from x_0 with $\gamma = 0$).

From the results obtained by the SIATR_2D solver we count $It \geq 27$ four times and $Fe \geq 28$ six times. Further, often convergence is reached with a reduction of the trust-region size. More specifically, it results $R = 1$ one time, $R = 2$ two times, $R = 3$ four times, $R = 22$ two times and $R = 30$ one time. Then, in our numerical experiments, the procedure given in [4] failed more frequently than the AS_ID method to produce good trial steps and severe reductions of the trust-region size occurred, too.

To make in more significant evidence our results we conclude this section comparing AS_ID and SIATR_2D algorithms by adopting the performance profile approach [8]. In Figure 1 the number of F -evaluations is used as a measure of the computational effort of the solvers under comparison, while in Figure 2 the computational effort is measured in terms of nonlinear iterations. Again, the conclusions follow the lines stated above and the profiles indicate that the method proposed here is an efficient way to solve constrained systems of nonlinear equations. More precisely, the performance profiles show that the AS_ID method outperforms SIATR_2D in our numerical tests, both in terms of F evaluations and nonlinear iterations. At this regard we underline that the number of linear systems approximately solved to compute the inexact Newton steps is equal to the number of nonlinear iterations. Then, from Figure 2, it can be seen that in about the 78% of tests the AS_ID method requires a smaller number of nonlinear iterations than SIATR_2D. Then, we have that in about the

78% of tests the AS_ID method solves a smaller number of linear systems than SIATR_2D. Moreover, the remaining 12% of tests successfully solved, requires a number of nonlinear iterations that is at most the double of those required by SIATR_2D.

5 Final remarks

We have introduced a new dogleg method for bound constrained nonlinear systems of equations. The method follows the so called affine-scaling approach and generates strictly feasible iterates only. It differs from other methods of the same type in the way of approximating the trust-region step and is suitable for numerical solution of large scale problems. In fact, an Inexact Newton method is incorporated into an affine scaling trust-region scheme. At each iteration, a suitably scaled region around the current approximate solution is defined and, within such a region, the norm of the linear model of F is trusted to be an adequate representation of the merit function $\|F\|$. The adopted linear model is minimized along a path founded on a scaled Cauchy step and a projected Inexact Newton step. Unlike the standard dogleg, the path is not a convex combination of these two steps, but a more general combination. Theoretical properties of the method are stated and global locally fast convergence is ensured both on problems with solution within the interior of the feasible region Ω and on problems with solution on the boundary of Ω . These good performances are guaranteed by using tools that are quite different from those commonly used in performing convergence analysis of affine scaling trust-region methods for constrained nonlinear systems. The method also allows a great deal of flexibility in specifying the scaling matrix used to handle the bounds. For global convergence purposes, quite general scaling matrices are allowed and several existing scaling matrices from the recent literature satisfy our requirements. To verify the basic effectiveness of the proposed algorithm and perform a comparison with the approach given in [4], the method has been implemented by using the standard spherical trust-region and the scaling matrix given by Coleman and Li. From the obtained numerical results the new method seems an efficient tool for solving large scale systems of nonlinear equations with bound constraints. We believe, however, that more pronounced differences in numerical performances of the method are possible by using different scaling matrices. We leave this topic as part of our future research since the effect of different scaling matrices on the practical behaviour of the method was not the main aim of this work.

References

- [1] S. BELLAVIA, M. MACCONI, AND B. MORINI, *An affine scaling trust-region approach to bound-constrained nonlinear systems*, Applied Numerical Mathematics, 44 (2003), pp. 257–280.

			AS_ID			SIATR-2D		
Pb#	γ	$\ F(x_0)\ $	It	Fe	$\ F(x_c)\ $	It	Fe	$\ F(x_c)\ $
1	0	3.1(1)	12	13	2.6(-8)	13	14	1.2(-8)
	1	2.9(1)	12	13	3.3(-10)	13	14	2.8(-9)
	2	1.2(1)	11	12	3.3(-10)	11	12	3.4(-7)
	3	5.5(2)	*	*	*	*	*	*
2	0	1.5(2)	8	9	3.2(-9)	10	11	1.3(-8)
	1	1.4(2)	8	9	9.9(-9)	9	10	2.5(-8)
	2	2.7(1)	4	5	2.9(-8)	6	7	1.1(-10)
	3	1.7(4)	13	14	6.9(-7)	16	17	4.2(-9)
3	0	4.3(0)	6	8	4.7(-10)	5	6	6.5(-7)
	1	5.1(1)	7	8	2.4(-7)	8	10	1.1(-11)
	2	1.8(3)	13	14	3.8(-11)	13	14	1.3(-10)
	3	6.5(5)	*	*	*	61	84	2.4(-10)
4	0	1.4(2)	17	18	6.8(-7)	20	24	2.9(-7)
	1	1.1(2)	16	17	8.9(-7)	16	17	9.6(-8)
	2	2.0(0)	8	9	2.0(-7)	7	8	3.1(-7)
	3	4.8(5)	23	26	4.5(-7)	26	30	1.1(-7)
5	0	3.6(1)	14	15	8.6(-8)	15	16	3.3(-11)
	1	2.1(1)	11	12	3.8(-12)	12	13	2.2(-12)
	2	5.6(0)	8	9	7.2(-13)	9	10	4.7(-9)
	3	1.3(1)	14	15	5.5(-10)	13	14	9.6(-7)
6	0	5.7(-1)	10	11	3.7(-10)	10	11	1.1(-8)
	1	1.0(0)	10	11	2.4(-10)	11	12	9.7(-10)
	2	7.0(0)	11	12	9.6(-8)	12	13	2.1(-7)
	3	6.9(1)	15	16	1.5(-7)	17	18	6.8(-11)
7	0	1.3(2)	24	26	1.2(-7)	46	77	1.8(-7)
	1	1.3(2)	23	24	3.7(-8)	42	65	9.3(-9)
	2	3.0(2)	24	25	5.2(-7)	23	26	3.9(-7)
	3	1.3(5)	27	28	1.2(-7)	29	30	7.4(-9)
8	0	2.0(2)	15	19	7.0(-8)	18	21	1.8(-7)
	1	1.9(2)	16	20	4.0(-8)	20	24	1.7(-9)
	2	1.2(2)	15	18	9.4(-7)	25	29	3.0(-8)
	3	7.5(3)	*	*	*	*	*	*

Table 1: Comparative numerical results

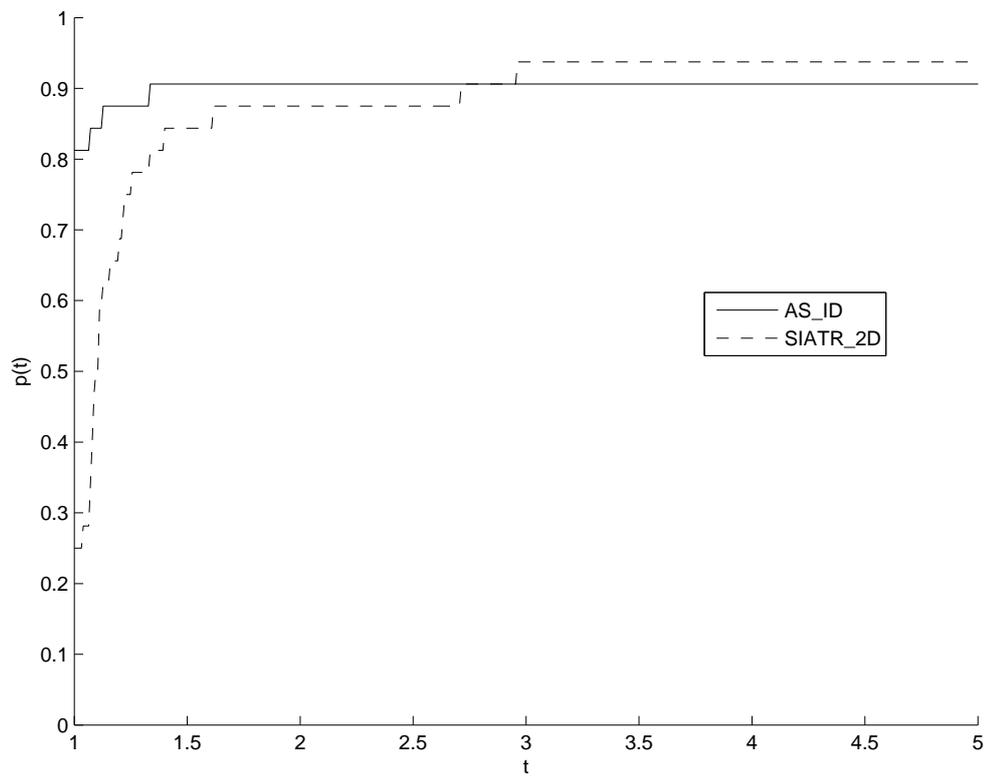


Figure 1: Performance profile in $[1,5]$ in terms of F-evaluations

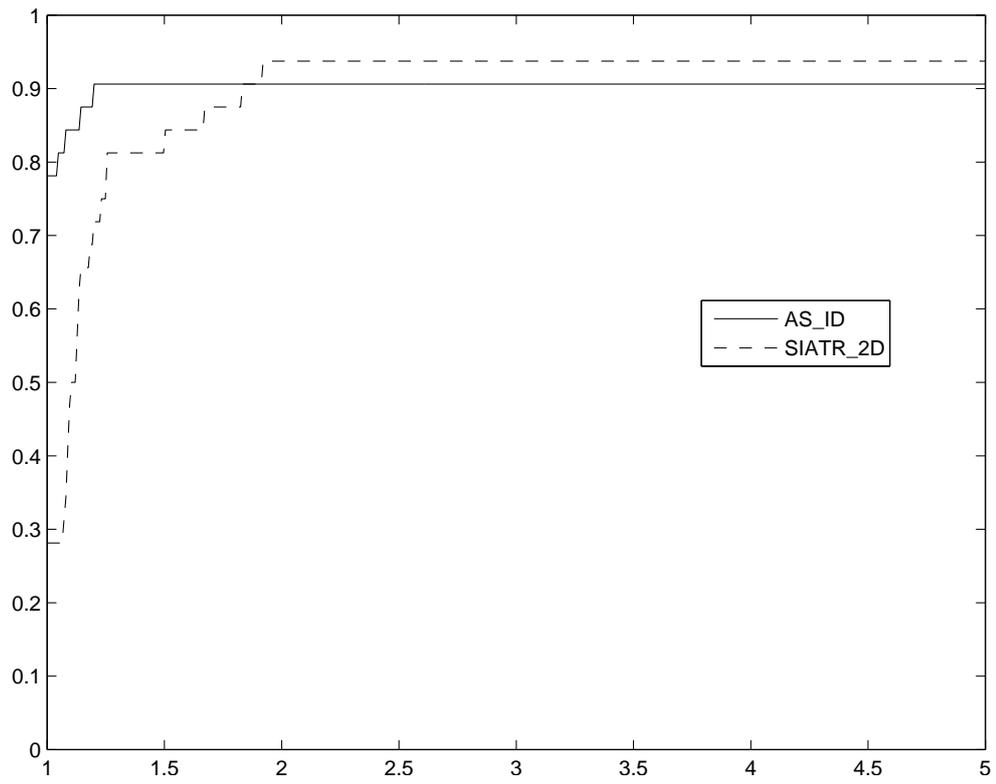


Figure 2: Performance profile in $[1,5]$ in terms of Nonlinear Iterations

- [2] ———, *A two-dimensional trust-region method for large scale bound-constrained nonlinear systems*, in Applied and Industrial Mathematics in Italy II, V. Cutello, G. Fotia, and L. Puccio, eds., Series on Advances in Mathematics for Applied Sciences, Vol.75, World Scientific, 2007, pp. 137–148.
- [3] S. BELLAVIA AND B. MORINI, *An interior global method for nonlinear systems with simple bounds*, Optimization Methods and Software, 20 (2005), pp. 1–22.
- [4] ———, *Subspace trust-region methods for large bound constrained nonlinear equations*, SIAM J. Numer. Anal., 44 (2006), pp. 1535–1555.
- [5] L. T. BIEGLER, O. GHATTAS, M. HEINKENSCHLOSS, AND B. VAN BLOEMEN WAANDERS, *Large-scale PDE-constrained optimization*, Lecture Notes in Computational Science and Engineering, 30, Springer, 2003.
- [6] S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, 2004.
- [7] T. F. COLEMAN AND Y. LI, *An interior trust-region approach for nonlinear minimization subject to bounds*, SIAM J. Optim., 6 (1996), pp. 418–445.
- [8] E. D. DOLAN AND J. J. MORÉ, *Benchmarking optimization software with performance profiles*, Mathematical Programming, 91 (2002), pp. 201–213.
- [9] S. C. EISENSTAT AND H. F. WALKER, *Globally convergent inexact Newton methods*, SIAM J. Optim., 4 (1994), pp. 393–422.
- [10] ———, *Choosing the forcing term in an inexact Newton method*, SIAM J. Scientific Computing, 17 (1996), pp. 16–32.
- [11] M. C. FERRIS AND J.-S. PANG, *Engineering and economic applications of complementarity problems*, SIAM Review, 39 (1997), pp. 669–713.
- [12] N. I. M. GOULD AND P. L. TOINT, *FILTRANE: a fortran 95 filter-trust-region package for solving nonlinear least squares and nonlinear feasibility problems*, ACM Trans. Math. Software, 33 (2007), pp. 3–25.
- [13] M. HEINKENSCHLOSS, M. ULBRICH, AND S. ULBRICH, *Superlinear and quadratic convergence of affine-scaling interior-point Newton methods for problems with simple bounds without strict complementarity assumptions*, Mathematical Programming, 86 (1999), pp. 615–635.
- [14] P. HOLLINGER, *Beyond Newton: Robust method for solving large nonlinear models in TROLL*. <http://EconPapers.repec.org/RePEc:sce:scecf0:308>, 2000. Presented at the Sixth International Conference on Computing in Economics and Finance.

- [15] V. A. JUVEKAR, C. V. ANOOP, S. K. PATTANAYEK, AND V. M. NAIK, *A continuum model for polymer absorption at the solid-liquid interface*, *Macromolecules*, 32 (1999), pp. 863–873.
- [16] C. KANZOW AND A. KLUG, *An interior-point affine-scaling trust-region method for semismooth equations with box constraints*, *Computational Optimization and Applications*, 37 (2007), pp. 329–353.
- [17] C. T. KELLEY, *Iterative Methods for Linear and Nonlinear Equations*, *Frontiers in Applied Mathematics*, SIAM, 1995.
- [18] C. T. KELLEY AND J. I. NORTHRUP, *A pointwise Quasi-Newton method for integral equations*, *SIAM J. Numer. Anal.*, 25 (1988), pp. 1138–1155.
- [19] D. LAXTON, P. ISARD, H. FARUQEE, E. PRASAD, AND B. TURTELBOOM, *MULTIMOD Mark III: The core dynamic and steady-state models*. Occasional Paper 164, International Monetary Fund, Washington DC, 1998.
- [20] L. LUKSAN AND J. VLCEK, *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.
- [21] M. MACCONI, B. MORINI, AND M. PORCELLI, *Trust-region quadratic methods for nonlinear systems of mixed equalities and inequalities*, *Applied Numerical Mathematics*, 59 (2009), pp. 859–876.
- [22] H. D. MITTELMANN AND H. MAURER, *Solving elliptic control problems with interior point and SQP methods: Control and state constraints*, *Journal of Computational and Applied Mathematics*, 120 (2000), pp. 175–195.
- [23] R. P. PAWLOWSKI, J. P. SIMONIS, H. F. WALKER, AND J. N. SHADID, *Inexact Newton dogleg methods*, *SIAM J. Numer. Anal.*, 46 (2008), pp. 2112–2132.
- [24] H. QI, L. QI, AND D. SUN, *Solving Karush-Kuhn-Tucker systems via the trust-region and the conjugate gradient methods*, *SIAM J. Optim.*, 14 (2003), pp. 439–463.
- [25] L. QI, X. J. TONG, AND D. H. LI, *An active-set projected trust-region algorithm for box-constrained nonsmooth equations*, *Journal of Optimization Theory and Applications*, 120 (2004), pp. 627–649.
- [26] M. ULBRICH, *Nonmonotone trust-region methods for bound-constrained semismooth equations with applications to nonlinear mixed complementarity problems*, *SIAM J. Optim.*, 11 (2000), pp. 889–917.
- [27] A. J. WOOD AND B. F. WOLLENBERG, *Power generation, Operation and Control*, John Wiley & Sons, 1996.

- [28] D. ZHU, *An affine scaling trust-region algorithm with interior backtracking technique for solving bound-constrained nonlinear systems*, Journal of Computational and Applied Mathematics, 184 (2005), pp. 343–361.