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

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

Stefania Bellavia†, Sandra Pieraccini‡

Abstract

Within the framework of affine scaling trust-region methods for bound constrained problems, we discuss the use of a inexact dogleg method as a tool for simultaneously handling the trust-region and the bound constraints while seeking for an approximate minimizer of the model. Focusing on bound-constrained systems of nonlinear equations, an inexact affine scaling method for large scale problems, employing the inexact dogleg procedure, is described. Global convergence results are established without any Lipschitz assumption on the Jacobian matrix, and locally fast convergence is shown under standard assumptions. Convergence analysis is performed without specifying the scaling matrix used to handle the bounds, and a rather general class of scaling matrices is allowed in actual algorithms. Numerical results showing the performance of the method are also given.

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

1 Introduction

Affine scaling methods have been originally proposed by Coleman and Li for the solution of bound-constrained optimization problems in [8, 9, 10] and further extended to the solution of different classes of problems and tailored for handling large dimension problems. Generalization of such methods to nonlinear minimization subject to linear inequality constraints [11] and to nonlinear programming problems [34] have been proposed. Moreover, these methods have been also modified in order to solve bound constrained nonlinear systems

*This work was supported by Italian funds from Miur (PRIN grant) and INdAM-GNCS funds. This is an Author's Original Manuscript of an article submitted for consideration in the Optimization Mathods and Software [copyright Taylor & Francis]; Optimization Methods and Software is available online at http://www.tandfonline.com/toc/goms20/current#.UqHMmte9wfE.

†Dipartimento di Ingegneria Industriale, Università di Firenze, viale Morgagni 40/44, 50134 Firenze, Italia, e-mail: stefania.bellavia@unifi.it

‡Dipartimento di Scienze Matematiche, Politecnico di Torino, corso Duca degli Abruzzi, 24, 10129 Torino, Italia, e-mail: sandra.pieraccini@polito.it

[33, 1, 4, 2, 3, 24, 38, 19] and systems of equalities and inequalities [27, 26, 29]. Affine scaling methods have been also used in conjunction with conic models [37] and Barzilai-Borwein gradient methods [20]. Methods suitable for large scale problems have been proposed, too [5, 31, 7]. Furthermore, we mention that affine scaling methods described in [8, 9, 10] are at the core of several functions implemented in the Matlab Optimization Toolbox.

An attractive aspect of affine-scaling interior point method is that they show strong local and global convergence properties: in most of the papers cited above the affine scaling scheme is combined with a trust-region approach in order to get a globally convergent scheme. Then, common ingredients of these approaches are the scaling diagonal matrix and the trust-region subproblem. The scaling matrix plays a key role in the definition of scaling-affine methods. The pioneering scaling matrix was proposed by Coleman and Li in [9]; Heinkenschloss et al. made a step further and proposed new scaling matrices that combined with a projection allowed to establish superlinear and quadratic converge without the strict complementarity assumption. Kanzow and Klug proposed in [23] a further scaling matrix with improved smoothness properties. Recently, a new scaling matrix is proposed and employed in [36]. Finally, we mention that a scaling matrix is implicitly defined also in [20].

Affine-scaling methods for bound constrained problems embedded in a trust-region framework [1, 2, 4, 5, 7, 9, 26, 27, 29, 31, 33, 36, 38] require, at a generic nonlinear iteration k, the solution of the following subproblem:

$$\begin{array}{ll} & \min_p m_k(p) \\ \text{s.}t. & x_k + p \in int(\Omega) \\ & \|D(x_k)^{-1/2} p\|_2 \leq \Delta_k \end{array} \tag{1}$$

where, given the current iterate x_k, $m_k(p)$ is a suitable local model for the function that has to be minimized, Ω is the n-dimensional box $\Omega = \{x \in \mathbb{R}^n \mid l \leq x \leq u\}$, $int(\Omega)$ is the interior of Ω, $D(x)$ is the diagonal positive definite scaling matrix and $\Delta_k > 0$ is the trust-region radius. Here, 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.

We mention that the subproblem (1) can be solved invoking an iterative optimization procedure, as adopted in [33, 36], but this may be computationally demanding. A different approach has been followed in [1, 2, 4, 5, 7, 9, 26, 27, 29, 31] where essentially a double-dogleg strategy is used: the classical trust-region is solved, then the trust-region step is projected/truncated onto Ω and a convex combination of this projected step and the generalized Cauchy step gives the required approximated solution of the sub-problem (1). In the solution of large scale problems the trust-region subproblem is approximately solved using two dimensional subspace approaches [5, 7]. Note that in the above approaches, in the first phase the bound constraints are not taken into account and only in the second phase they come into play.

Here, we first give a general framework describing a inexact dogleg procedure for computing an approximate solution to (1) where the trust-region constraints

and the bounds are simultaneously handled. This procedure can also be employed in conjunction with non-quadratic models as tensor models for nonlinear systems [32] or conic models [12, 37] for minimization problems. In the inexact dogleg procedure the model $m_k(p)$ is minimized along a path whose foundations are a scaled Cauchy step and a projected approximate unconstrained minimizer of $m_k(p)$. 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 the minimizer of $m_k(p)$ is computed in an approximate way, in the spirit of Inexact or Truncated Newton methods and moreover it may be projected. Then, 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*. The framework given here is a generalization to the bound constrained setting of the inexact dogleg procedure introduced in [30].

Then, we show that the Inexact Dogleg procedure, used within affine scaling methods, gives rise to methods with strong local and global convergence properties. In particular, we focus on the solution of bound constrained nonlinear systems and we analyze from a theoretical and computational point of view the behavior of Inexact affine scaling methods for large scale bound constrained nonlinear systems employing the inexact dogleg procedure to compute the trial step. The problem we are interested in is to find a vector $x \in \mathbb{R}^n$ satisfying

$$F(x) = 0, \qquad x \in \Omega, \tag{2}$$

where $F : X \mapsto \mathbb{R}^n$ is a continuosly differentiable mapping and $X \subseteq \mathbb{R}^n$ is an open set containing the n-dimensional box Ω.

We remark that the theoretical analysis performed here is carried out without specifying the scaling matrix used to handle the bounds. This 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. Moreover, global convergence is proven without any Lipschitz assumption on the Jacobian of $F(x)$ and locally fast convergence is ensured under standard assumptions.

The procedures employed in [3] and in [37] belong to the inexact dogleg framework outlined here. However, in both cases the inexact dogleg path is built around the exact minimizer of the model $m_k(p)$ instead of around an approximate one. In [3] the properties of scaling matrices proposed in literature have been analyzed and the computational behavior of an affine-scaling method employing such inexact dogleg path has been studied. A numerical comparison among the scaling matrices is also carrried out and the matlab code `Codosol` is introduced. The convergence theory carried out here covers also the method proposed in [3] where the convergence theory is not given. In [37] the dogleg path is used in conjunction with conic models.

We close this section mentioning that all the trust-region affine scaling methods we are aware of use the ℓ_2 norm trust-region subproblem. Since in (1) also the bounds constraints must be imposed, an appealing proposal is to employ

the ℓ_∞ norm trust-region. However, this latter subproblem is expensive to be solved, whereas both the inexact dogleg procedure and the double-dogleg procedures used in [1, 2, 4, 5, 7, 9, 26, 27, 29, 31] are not computational demanding, and their cost reduces to the cost of computing a, possibly approximate, unconstrained minimizer of the quadratic model.

The paper is organized as follows. In Section 2 we briefly recall the framework for trust-region affine scaling methods for bound constrained minimization; in Section 3 the inexact dogleg procedure is described; the affine scaling inexact dogleg method is then described in Section 4, and its convergence properties are analyzed in Section 5. Finally, some numerical experiments are proposed in Section 6.

1.1 Notation

Throughout the paper we use the following notation. For any mapping $F : X \to \mathbb{R}^n$, differentiable at a point $x \in X \subset \mathbb{R}^n$, the Jacobian matrix of F at x is denoted by $F'(x)$ and for any mapping $f : X \to \mathbb{R}$, the gradient of f at x is denoted by $\nabla f(x)$. 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 affine scaling trust-region framework

In this section we briefly describe the framework of trust-region affine scaling methods for bound constrained minimization problem, giving the basic concepts that will be useful for the development of the next sections.

Let us consider the minimization problem

$$\min_{x \in \Omega} f(x) \tag{3}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable function. As originally shown by Coleman and Li [9], a solution x^* to (3) satisfies

$$D(x^*)\nabla f(x^*) = 0, \tag{4}$$

where $D(x)$ is a proper diagonal scaling matrix of order n with diagonal elements given by

$$d_i^{CL}(x) = \begin{cases} u_i - x_i & \text{if } (\nabla f(x))_i < 0 \quad \text{and } u_i < \infty, \\ x_i - l_i & \text{if } (\nabla f(x))_i > 0 \quad \text{and } l_i > -\infty, \\ \min\{x_i - l_i,\, u_i - x_i\} & \text{if } (\nabla f(x))_i = 0 \quad \text{and } l_i > -\infty \text{ or } u_i < \infty, \\ 1 & \text{otherwise.} \end{cases} \tag{5}$$

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 [9] was generalized in [21]. In this latter paper it is shown that first order optimality conditions for problem (3) are given by (4) for any diagonal matrix $D(x)$ with diagonal elements satisfying

$$d_i(x) \begin{cases} = 0 & \text{if } x_i = l_i \text{ and } \nabla f(x)_i > 0, \\ = 0 & \text{if } x_i = u_i \text{ and } \nabla f(x)_i < 0, \\ \geq 0 & \text{if } x_i \in \{l_i, u_i\} \text{ and } \nabla f(x)_i = 0, \\ > 0 & \text{otherwise.} \end{cases} \tag{6}$$

It is straightforward to note that elements given by (5) satisfy (6).

Affine scaling methods aim at building a strictly feasible sequence $\{x_k\}$. Then, given $x_k \in int(\Omega)$ at hand, in order to ensure a stepsize large enough to produce a strictly feasible point and an acceptable progress towards a solution, affine scaling methods 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.

$$d_k = -D_k \nabla f_k,$$

is an useful tool as it allows to implicitly handle the bounds by means of the diagonal matrix D_k. However, in order to go beyond a scaled gradient direction and obtain fast convergence of the method, trust-region strategies are adopted. Then, at iteration k, given an iterate $x_k \in int(\Omega)$ and the trust-region radius Δ_k, an approximate solution $p(\Delta_k)$ of the following subproblem is required:

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

where m_k is a suitable model for f around x_k. Note that problem (7) differs from problem (1) in the definition of the trust-region shape. In fact, in (7) 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 the elliptical trust-region given in (1). Spherical trust-regions have been used in methods [5, 3, 26, 27, 29, 31, 37]. Further, let us define the so-called generalized Cauchy step $p_c(\Delta_k)$, that is the minimizer of the model along the scaled gradient d_k, constrained to be in the trust-region and to satisfy $x_k + p_c(\Delta_k) \in int(\Omega)$. Then,

$$p_c(\Delta_k) = \tau_k d_k \tag{8}$$

where τ_k is the solution of the one dimensional problem:

$$\min_{\tau} \{m_k(\tau d_k) : \quad \tau \|G_k\, d_k\| \leq \Delta_k, \quad x_k + \tau d_k \in int(\Omega)\}.$$

In order to get global convergence, the approximate minimizer $p(\Delta_k)$ is required to satisfy the following Cauchy decrease condition:

$$m_k(p(\Delta_k)) \leq m_k(p_c(\Delta_k)). \tag{9}$$

Finally, following the standard trust-region philosophy, the sufficient improvement condition

$$\rho_f(p(\Delta_k)) = \frac{f(x_k) - f(x_k + p(\Delta_k))}{m_k(0) - m_k(p(\Delta_k))} \geq \beta \tag{10}$$

is required to hold for a given constant $\beta \in (0,1)$. Namely, if (10) 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.

3 The inexact-Dogleg procedure

This section gives a detailed description of the inexact dogleg procedure for computing the trial step $p(\Delta_k)$. Let us assume to have at disposal an approximate minimizer p_k^{IN} of the model $m_k(p)$. 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:

$$\bar{p}_k^{IN} = \alpha_k(P(x_k + p_k^{IN}) - x_k), \qquad \alpha_k \in (0,1), \tag{11}$$

where $P(x)_i = \max\{l_i, \min\{x_i, u_i\}\}$ and the scalar $\alpha_k \in (0,1)$ is employed in order to move back from the boundary.

We clearly have

$$\|\bar{p}_k^{IN}\| < \|p_k^{IN}\|. \tag{12}$$

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

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

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 p_k^{IN} is not the exact minimizer of the model and it may also be projected, many theoretical properties of the dogleg curve are lost and we allow more flexibility in the choice of the step.

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

$$\phi(\gamma) = m_k(p(\gamma)). \tag{13}$$

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

$$\gamma^2\|G_k(p_c(\Delta_k) - \bar{p}_k^{IN})\|^2 - 2\gamma p_c(\Delta_k)^T G_k^2(p_c(\Delta_k) - \bar{p}_k^{IN}) + \|G_k p_c(\Delta_k)\|^2 - \Delta_k^2 = 0 \tag{14}$$

Then, two values $\gamma_\pm$, given by

$$\gamma_\pm = \Big(p_c(\Delta_k)^T G_k^2 (p_c(\Delta_k) - \bar{p}_k^{IN}) \pm \Big(\big(p_c(\Delta_k)^T G_k^2 (p_c(\Delta_k) - \bar{p}_k^{IN}) \big)^2 - \|G_k(p_c(\Delta_k) - \bar{p}_k^{IN})\|^2 (\|G_k p_c(\Delta_k)\|^2 - \Delta_k^2) \Big)^{\frac{1}{2}} \Big) / \|G_k(p_c(\Delta_k) - \bar{p}_k^{IN})\|^2$$

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 $\gamma_\pm$ 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 (14) 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 (14) 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 in unlikely to happen.

Finally, we have to take into account that $p(\gamma)$ is required to produce a strictly feasible point. 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 mantained and shorten the step if necessary. 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

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

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

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

To summarize, the choice of γ is made as follows. Since we want to minimize $m_k(p(\gamma))$, we first seek for the unconstrained minimizer

$$\hat{\gamma} = \operatorname*{argmin}_{\gamma \in \mathbb{R}} \phi(\gamma). \tag{15}$$

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 $\gamma = \min(\hat{\gamma}, \gamma_+, \theta\bar{\gamma}_+)$, whereas if $\hat{\gamma} < 0$, we choose $\gamma = \max(\hat{\gamma}, \gamma_-, \theta\bar{\gamma}_-)$, with $\theta \in (0,1)$ and we set the trial step $p(\Delta_k) = p(\gamma)$.

Next, we sketch the process for finding $p(\Delta_k)$. We underline that this procedure generates a step that satisfies the decrease condition (9) as $p_c(\Delta_k)$ belongs to the path $p(\gamma)$.

Inexact Dogleg framework

Input parameters: $x_k \in int(\Omega)$, $\Delta_k > 0$, d_k, $\bar{p}_k^{IN}$, $\theta \in (0,1)$
  Compute $p_c(\Delta_k)$ by (8).
  Compute $\hat{\gamma}$ by (15).
  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 $\gamma = \max\{\hat{\gamma}, \gamma_-, \theta\bar{\gamma}_-\}$
  Set $p(\Delta_k) = p_c(\Delta_k) + \gamma(\bar{p}_k^{IN} - p_c(\Delta_k))$.

4 The Affine Scaling Inexact-Dogleg method

This section is devoted to describe an affine scaling procedure for large scale nonlinear systems of the form (2) where the trial step $p(\Delta_k)$ is computed by an inexact Dogleg procedure belonging to the framework outlined in the previous section.

We start by noting that every solution of (2) is a solution of the following bound constrained optimization problem:

$$\min_{x\in\Omega} f(x) = \min_{x\in\Omega} \|F(x)\|, \tag{16}$$

and that $\nabla f(x) = F'(x)^T F(x)$.

Our method belongs to the trust-region affine scaling methods described in Section 2 applied to (16). Here we specify the main ingredients of our method: the choice of the model $m_k(p)$, the computation of the unconstrained approximate minimizer p_k^{IN} of the model and the computation of scalars τ_k in (8) and $\hat{\gamma}$ in (15).

As standard in the solution of nonlinear systems, we do not employ second order derivatives of f and at the k-th nonlinear iteration we take, as a model for $f(x) = \|F(x)\|$ around x_k, the norm of the linear model for F, i.e.

$$m_k(p) = \|F_k + F_k' p\|.$$

We consider a quite general scaling matrix $D(x)$ satisfying (6) and some additional assumptions which are given in the next section. In (7) both spherical and elliptical trust-regions are allowed.

The computation of the scalar τ_k needed in (8) to build the generalized Cauchy step $p_c(\Delta_k)$ is carried out as follows. First,

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

is computed and if $x_k + \tau_k' d_k \in int(\Omega)$, we let $\tau_k = \tau_k'$ in (8). Otherwise we let λ_k be the stepsize along d_k to the boundary, i.e.

$$\lambda_k = \min_{1 \leq i \leq n} \Lambda_i, \quad \text{where} \quad \Lambda_i = \begin{cases} \max\left\{ \frac{l_i - (x_k)_i}{(d_k)_i}, \frac{u_i - (x_k)_i}{(d_k)_i} \right\} & \text{if } (d_k)_i \neq 0 \\ \infty & \text{if } (d_k)_i = 0 \end{cases}, \tag{17}$$

and set τ_k smaller than λ_k. Summarizing, the parameter τ_k in (8) is given by

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

Focusing on the definition of the path $p(\gamma)$ and more precisely on the computation of an approximate minimizer of the model, we take into account that we are in a large scale setting. Then, p_k^{IN} is chosen as an Inexact Newton step satisfying

$$F_k' \, p_k^{IN} = -F_k + r_k, \qquad \|r_k\| \leq \eta_k \|F_k\|, \tag{18}$$

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

With this ingredient at hand we have that the function $\phi(\gamma)$ in (13) is given by

$$\phi(\gamma) = \|F_k + F_k' p_c(\Delta_k) + \gamma F_k'(\bar{p}_k^{IN} - p_c(\Delta_k))\|. \tag{19}$$

The function $\phi(\gamma)$ is convex, as it is the composition of an affine function and a convex function [6]. Then $\hat{\gamma}$ exists and can be esaily computed as follows. 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

$$\hat{\gamma} = -\frac{a^T b}{b^T b} = -\frac{(F_k + F_k' p_c(\Delta_k))^T F_k'(\bar{p}_k^{IN} - p_c(\Delta_k))}{\|F_k'(\bar{p}_k^{IN} - p_c(\Delta_k))\|^2}. \tag{20}$$

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 $\hat{\gamma}$.

Therefore, the trial step is $p(\Delta_k) = p_c(\Delta_k) + \gamma(\bar{p}_k^{IN} - p_c(\Delta_k))$, where $p_c(\Delta_k)$ is given in (8), $\bar{p}_k^{IN}$ is the projection onto Ω of the inexact Newton step p_k^{IN} satisfying (18), γ is computed by the Inexact-Dogleg procedure with $\phi(\gamma)$ and $\hat{\gamma}$ given in (19) and (20), respectively.

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,

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

We are now ready to sketch our procedure:

Affine Scaling Inexact Dogleg (AS_ID) Method

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

For $k = 0, 1, \ldots$

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 (18).
4. Form $\bar{p}_k^{IN}$ by (11).
5. Set $d_k = -D_k \nabla f_k$.
6. Find $p(\Delta_k)$ by the Inexact Dogleg 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 Inexact Dogleg 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 issues. First, we point out that the trust-region size is updated according to standard rules, i.e. on the basis of agreement between the adopted model and the merit function. So, at each iteration Δ_k is enlarged if (10) 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.

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 (10) is verified. So, after a finite number of repetitions, the while-loop terminates (see the following Proposition 5.1).

Finally, we underline that our method can be implemented in a matrix-free manner provided that an operator performing the products F' times a vector and ${F'}^T$ times a vector is available.

5 Convergence analysis

The convergence analysis of AS_ID method is organized as follows. First, in Section 5.1 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 5.2 global convergence to a stationary point of (16) is proved without any Lipschitz assumption on F'. Finally, in Section 5.3 the local convergence properties of the method are investigated.

5.1 Finite termination

Given an iterate x_k such that $\|D_k \nabla f_k\| \neq 0$, 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 (6);

(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 (17)) 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 (6), it is nonsingular for $x \in int(\Omega)$.

In [3], the authors showed that the scaling matrix (5) as well as those proposed in [23] and in [20] satisfy Assumption 2. Then, these three matrices can be used within the AS_ID method.

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

Lemma 5.1 *Let x_k be generated by Method AS_ID and assume that $\|D_k^{1/2} \nabla f_k\| \neq 0$. Then*

$$\|F_k\| - \|F_k + F_k' p(\Delta_k)\| \geq (1 - \sqrt{1 - \omega_k})\|F_k\| \tag{21}$$

with

$$\omega_k = \min\left(\theta\lambda_k, \frac{\Delta_k}{\|G_k D_k \nabla f_k\|}, \frac{1}{\|F_k'\|^2 \|D_k\|}\right) \frac{\|D_k^{1/2} \nabla f_k\|^2}{\|F_k\|^2}. \tag{22}$$

Proof. First of all, we note that $\|F_k + F_k' p_c(\Delta_k)\| < \|F_k\|$ and from (9) it follows that

$$\|F_k\| - \|F_k + F_k' p(\Delta_k)\| \geq \|F_k\| - \|F_k + F_k' p_c(\Delta_k)\|.$$

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 $p_c(\Delta_k) = \tau_k d_k$ with

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

Then, we have

$$\begin{aligned}
(\eta_k^c)^2 &= \frac{\|F_k + F_k' p_c(\Delta_k)\|^2}{\|F_k\|^2} = \frac{\|F_k + \tau_k F_k' d_k\|^2}{\|F_k\|^2} \\
&= \frac{F_k^T F_k + 2\tau_k F_k^T F_k' d_k + \tau_k^2 d_k^T (F_k')^T F_k' d_k}{\|F_k\|^2} \\
&= 1 + 2\tau_k \frac{F_k^T F_k' d_k}{\|F_k\|^2} + \tau_k^2 \frac{\|F_k' d_k\|^2}{\|F_k\|^2} \\
&= 1 - 2\frac{(F_k^T F_k' d_k)^2}{\|F_k' d_k\|^2 \|F_k\|^2} + \frac{(F_k^T F_k' d_k)^2 \|F_k' d_k\|^2}{\|F_k' d_k\|^4 \|F_k\|^2} \\
&= 1 - \left(\frac{F_k^T F_k' d_k}{\|F_k' d_k\| \|F_k\|}\right)^2 = 1 - \left(\frac{\|D_k^{1/2} \nabla f_k\|^2}{\|F_k' d_k\| \|F_k\|}\right)^2 \\
&\leq 1 - \left(\frac{\|D_k^{1/2} \nabla f_k\|}{\|F_k'\| \|F_k\| \|D_k^{1/2}\|}\right)^2. \qquad (23)
\end{aligned}$$

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

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

Then, proceeding as to prove (23) we get:

$$\begin{aligned}
(\eta_k^c)^2 &= 1 + \tau_k\left(2\frac{F_k^T F_k' d_k}{\|F_k\|^2} + \tau_k \frac{\|F_k' d_k\|^2}{\|F_k\|^2}\right) \leq 1 + \tau_k \left(2\frac{F_k^T F_k' d_k}{\|F_k\|^2} - \frac{(F_k^T F_k' d_k)\|F_k' d_k\|^2}{\|F_k' d_k\|^2 \|F_k\|^2}\right) \\
&= 1 - \min\left(\theta\lambda_k, \frac{\Delta_k}{\|G_k D_k \nabla f_k\|}\right) \frac{\|D_k^{1/2} \nabla f_k\|^2}{\|F_k\|^2}. \qquad (24)
\end{aligned}$$

Relations (23) and (24) straightforwardly give

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

with ω_k given in (22). Note that, as $(\eta_k^c)^2$ is positive, it follows that $0 < \omega_k < 1$. The previous inequality immediately yields (21). □

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

Proposition 5.1 *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

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

and we obtain:

$$\|F(x_k + p(\Delta_k))\| - \|F_k + F_k' p(\Delta_k)\| \leq \|G_k^{-1}\|\Delta_k \int_0^1 \|F'(x_k + tp(\Delta_k)) - F_k'\|\, dt \tag{25}$$

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 (22) 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 (21) 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 (25) 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. □

5.2 Global convergence

Now we are ready to prove global convergence. Following [30], our proof is based on [17, Corollary 3.6], which is reported here for the reader's convenience.

Theorem 5.1 *[17, 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*

$$\|F_k\| - \|F(x_k + p_k)\| \geq \beta_1 \left(\|F_k\| - \|F_k + F_k' p_k\|\right) \tag{26}$$

$$\|F_k\| - \|F_k + F_k' p_k\| \geq 0 \tag{27}$$

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

$$\sum_{k\geq 0} \frac{pred_k}{\|F_k\|} = \sum_{k\geq 0} \frac{\|F_k\| - \|F_k + F_k' p_k\|}{\|F_k\|}$$

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

Our convergence result is stated in the following Theorem.

Theorem 5.2 *Let Assumptions 1 and 2 be satisfied. Then all the limit points of $\{x_k\}$ are stationary points for problem (16). Further, if there exists a limit point $x^* \in int(\Omega)$ of $\{x_k\}$ such that $F'(x^*)$ is nonsingular, then $\|F_k\| \to 0$ and all the accumulation points of $\{x_k\}$ solve problem (2). If, in addition, there exists a limit point $x^* \in \Omega$ such that $F(x^*) = 0$ and $F'(x^*)$ is invertible, then $x_k \to x^*$.*

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 $i \in \{1, ..., n\}$ 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 (6), 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

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

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

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

and

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

for any $x_k \in N_* \cap \Omega$. Hence for $x_k \in N_* \cap \Omega$, recalling (22) we have

$$\begin{aligned} \omega_k &= \frac{\|D_k^{1/2}\nabla f_k\|}{\|F_k\|^2} \min\left(\theta\lambda_k\|D_k^{1/2}\nabla f_k\|, \frac{\Delta_k\|D_k^{1/2}\nabla f_k\|}{\|G_k D_k \nabla f_k\|}, \frac{\|D_k^{1/2}\nabla f_k\|}{\|F_k'\|^2\|D_k\|}\right) \\ &\geq \frac{m}{\|F_0\|\|F_k\|} \min\left(\theta\bar{\lambda}m, \frac{\Delta_k}{\chi_D}, \frac{m}{M^2\chi_D^2}\right). \end{aligned} \tag{28}$$

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 (28)

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

Therefore, as $1 - \sqrt{1-\omega_k} > \frac{1}{2}\omega_k$, proceeding as in Proposition 5.1 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 (21) and (28) we have

$$\frac{\|F_k\| - \|F_k + F_k' p(\Delta_k)\|}{\|F_k\|} \geq 1 - \sqrt{1-\bar{\omega}}, \qquad \text{with} \qquad \bar{\omega} = \frac{m}{\|F_0\|^2}\frac{\delta\hat{\Delta}}{\chi_D} > 0.$$

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 (26) and (27). Then we are in a position to apply Theorem 5.1 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 $x^* \in int(\Omega)$ such that $F'(x^*)$ is nonsingular, it follows that $\lim_{k\to\infty} \|F_k\| = 0$. Finally, if there exists $x^* \in \Omega$ such that $F(x^*) = 0$ and $F'(x^*)$ is invertible, we can apply Theorem 3.3 of [17] with $\eta = 1$, that proves the convergence of x_k to x^*. □

5.3 Superlinear convergence

In this subsection we will prove superlinear convergence of the AS_ID method. In the sequel we assume the following:

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

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

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 $int(\Omega)$ there exist $\bar{\rho} > 0$ and $\chi_{\bar{x}}$ such that $B_{\bar{\rho}}(\bar{x}) \subset int(\Omega)$ and $\|D(x)^{-1}\| \le \chi_{\bar{x}}$ 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 [5, Lemma 4.2].

Lemma 5.2 *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*

$$\|x - x^*\| \le K_2\,\|F(x)\|, \tag{29}$$

$$\|F(x)\| \le K_1\,\|x - x^*\|, \tag{30}$$

$$\|F'(x)^{-1}\| \le K_2, \tag{31}$$

$$\|F(x) - F(z) - F'(z)(x - z)\| \le \Gamma\|x - z\|^2 \text{ for all } z \in B_\rho(x^*). \tag{32}$$

In order to discuss the convergence rate issues, we make the additional hypothesis $\|G_k p_k^{IN}\| \to 0$ as $k \to \infty$. 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 (18) and (31) we get

$$\|p_k^{IN}\| = \|{F'}_k^{-1}(-F_k + r_k)\| \le (1 + \eta_k)\,\|{F'}_k^{-1}\|\,\|F_k\| \le 2\,K_2\,\|F_k\|, \tag{33}$$

whenever $x_k \in B_\rho(x^*)$. This implies that $\|p_k^{IN}\| \to 0$ as $k \to \infty$. Further, when $x^* \in int(\Omega)$, 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 $\|D_k^{-1/2} p_k^{IN}\| \leq \chi_{x^*}^{1/2} \|p_k^{IN}\| \to 0$ as $k \to \infty$.

From now on, with γ_L and χ_J as in Assumptions 4-5, K_1, K_2 and Γ as in Lemma 5.2, we let

$$\begin{aligned} &K^* = \|F'(x^*)\| \, \|F'(x^*)^{-1}\|, \\ &\nu = 8\, K^* \, (K_2 \chi_J + 1), \\ &\delta_k = K_2 \Gamma \nu^2 \|x_k - x^*\| + 4\, K^* \, \eta_k, \\ &\psi_k = \chi_J \delta_k + \gamma_L \nu^2 \|x_k - x^*\| + K_1 (1 - \alpha_k), \\ &\sigma_k = \max\{\psi_k,\, K_2 (\, \Gamma \nu^2 \|x_k - x^*\| + \psi_k)\}. \end{aligned} \tag{34}$$

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

Lemma 5.3 *Assume that there exists a solution x^* of (2) 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 $\|G_k p_k^{IN}\| \to 0$ as $k \to \infty$.*

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

$$\|G_k \bar{p}_k^{IN}\| \leq \bar{\Delta}_k, \tag{35}$$

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

$$\|F_k + F_k' p_k^{IN}\| \leq K_1 \, \eta_k \|x_k - x^*\|, \tag{36}$$

$$\|\bar{p}_k^{IN}\| < \|p_k^{IN}\| \leq \nu \|x_k - x^*\|, \tag{37}$$

$$\|p(\bar{\Delta}_k)\| \leq \nu \|x_k - x^*\|. \tag{38}$$

Proof. Relation (35) 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 (33) yields $\lim_{k\to\infty} \|p_k^{IN}\| = 0$. Then, with ρ as in Lemma 5.2, there exists $\rho_1 \leq \rho$ such that $\|\bar{p}_k^{IN}\| \leq \|p_k^{IN}\| \leq \bar{\Delta}_k$ when $x_k \in B_{\rho_1}(x^*) \cap int(\Omega)$ and the thesis (35) 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 $\lim_{k\to\infty} \|G_k p_k^{IN}\| = 0$, we get (35).

The remaining results are proven independently of the form of G_k. Result (36) is easily proven as by (18) and (30) we obtain

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

Result (37) is derived noting that by (12), (33) and (30) we get

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

Then, by (34) relation (37) follows.

Finally, we move on proving (38). We have

$$\|p(\bar{\Delta}_k)\| = \|p(\gamma)\| \leq \|p_c(\bar{\Delta}_k)\| + |\gamma|\|\bar{p}_k^{IN} - p_c(\bar{\Delta}_k)\|. \tag{39}$$

Let us consider the value of $|\gamma|$. From (20) we have

$$\begin{aligned} |\gamma| \leq |\hat{\gamma}| &\leq \frac{\|F_k + F_k' p_c(\bar{\Delta}_k)\|}{\|F_k'(\bar{p}_k^{IN} - p_c(\bar{\Delta}_k))\|} \leq \frac{\|F_k + F_k' p_c(\bar{\Delta}_k)\|\|(F_k')^{-1}\|}{\|\bar{p}_k^{IN} - p_c(\bar{\Delta}_k)\|} \\ &\leq \frac{K_2\|F_k + F_k' p_c(\bar{\Delta}_k)\|}{\|\bar{p}_k^{IN} - p_c(\bar{\Delta}_k)\|}. \end{aligned} \tag{40}$$

Moreover, recalling (8) 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, (30) and (31), the last inequality yields

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

Finally, let

$$p_k^{gC} = -\frac{F_k^T F_k' d_k}{\|F_k' D_k \nabla f_k\|^2} d_k, \qquad \eta_{gC} = \frac{\|F_k + F_k' p_k^{gC}\|}{\|F_k\|}.$$

Then, $\eta_{gC} \leq 1$ as p_k^{gC} is the unconstrained minimizer of the model along d_k. Moreover, note that $p_c(\Delta_k) = \beta_k p_k^{gC}$ 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 $\tau_k = \theta\lambda_k$. Therefore, it follows:

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

Then, (30) yields:

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

Hence, by (39) and (40) 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 (38). □

Lemma 5.4 *Assume that there exists a solution x^* of (2) 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 $\|G_k p_k\| \to 0$ as $k \to \infty$.*

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

$$\begin{aligned} &\|F_k + F_k' \, \bar{p}_k^{IN}\| \leq \sigma_k \|x_k - x^*\|, \\ &\|x_k + \bar{p}_k^{IN} - x^*\| \leq \sigma_k \|x_k - x^*\|. \end{aligned} \tag{41}$$

Proof. The thesis straightforwardly follows from [5, Lemma 4.4] replacing $p_{tr}(\bar{\Delta}_k)$ with p_k^{IN} and $\bar{p}_{tr}(\bar{\Delta}_k)$ with $\bar{p}_k^{IN}$. □

In the next Theorem we show that, if $\eta_k \to 0$ as $k \to \infty$, eventually the step $p(\bar{\Delta}_k)$ satisfies condition (10). 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 5.3 *Assume that there exists a solution x^* of (2) such that $F'(x^*)$ is nonsingular and that the sequence $\{x_k\}$ generated by the AS_ID method converges to x^*. Suppose that $\eta_k \to 0$, $\alpha_k \to 1$, as $k \to \infty$, and*

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

Then, eventually, $p(\bar{\Delta}_k)$ satisfies (10) and the sequence $\{x_k\}$ converges to x^* superlinearly. Moreover, if

$$\eta_k = O(\|F_k\|), \quad \alpha_k = 1 - O(\|F_k\|) \ \ as \quad k \to \infty,$$

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 5.3. Moreover it belongs to the path described by $\phi(\gamma)$ and it is a feasible point. From these arguments and from (41), it follows

$$\|F_k + F_k' p(\bar{\Delta}_k)\| \leq \|F_k + F_k' \bar{p}_k^{IN}\| \leq \sigma_k \|x_k - x^*\| \tag{42}$$

for $x_k \in B_{\rho_2}(x^*)$ with ρ_2 as in Lemma 5.4.

Now we prove that eventually $p(\bar{\Delta}_k)$ satisfies (10). First note that, from (38), 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 $\sigma_k \to 0$ and $\|x_k - x^*\| \to 0$. 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 (29) and (32) we have

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

Further, from (42) and (38) we get

$$\begin{aligned}
\|x_k + p(\bar{\Delta}_k) - x^*\| &\leq K_2(\Gamma\|p(\bar{\Delta}_k)\|^2 + \sigma_k\|x_k - x^*\|) \\
&\leq K_2(\Gamma\nu^2\|x_k - x^*\| + \sigma_k)\|x_k - x^*\|. \qquad (43)
\end{aligned}$$

Since we have

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

(43) ensures superlinear convergence rate if $\eta_k \to 0$ and $\alpha_k \to 1$ as $k \to \infty$. Moreover, if $\eta_k = O(\|F_k\|)$ and $1 - \alpha_k = O(\|F_k\|)$, by (29)-(30) we get $\sigma_k = O(\|x_k - x^*\|)$ and this yields quadratic convergence rate. □

6 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 compare its performance with that of the affine scaling method for large scale problems provided in the Matlab function `lsqnonlin`. The solver `lsqnonlin` is available in the Matlab Optimization Toolbox and it is designed to solve square or overdetermined bound constrained nonlinear least-squares problems. The solver is based on the method decribed in [7] where an inexact affine scaling trust region approach is used in conjunction with the Coleman and Li scaling matrix and elliptical trust regions. An approximate solution of the trust region subproblem (1) is computed using a 2D-subspace procedure involving the computation of an inexact Newton step. Then, in order to enforce the bounds the trial step is chosen as the best of three steps: the cauchy step (8), the truncated 2D-trust region solution and the reflection of the 2D trust region solution, truncated to remain strictly feasible. The Inexact Newton step is computed solving the arising linear system by the CG method with an adaptive choice of the tolerance in the stopping criterion. Note that, the effort to compute the Inexact Newton step may be completely lost in case the procedure for choosing the trial step selects the Cauchy step.

The experiments were carried out on a set of 14 widely studied problems, with dimension between $n = 500$ and $n = 12500$. Problems are listed in Table 1. Problems Pb1-Pb3 and Pb14 are already equipped with bounds on the variables. Problems Pb4-Pb13 have more than one solution and bounds have been added in order to select specific solutions. The bounds used in the numerical results are reported in Table 1 along with problem dimensions. Concerning Pb14, the lower bound reported in the table does not apply to the first 2020 components of x.

We performed our experiments starting from both good and poor initial guesses. As a general rule, we used the following starting points:

- $x_0 = l + \frac{\nu}{5}(u - l)$, $\nu = 1, 2, 3, 4$ for problems having finite lower and upper bounds;
- $x_0 = l + 10^{\nu-2}e^T$, with $\nu = 0, 1, 2, 3$ and $e = (1, \ldots, 1)$, for problems with infinite upper bound;
- $x_0 = -10^{\nu-2}e^T$, $\nu = 0, 1, 2, 3$ for the problem with infinite lower bound.

For problem Pb14 we used one single starting point used in [16]. As a whole, we performed 53 tests.

For a detailed description of the problems we refer the reader to the references reported in the table; here we only mention that Pb2 and Pb3 are nonlinear complementarity problems which have been reformulated as a system of $n = 12500$ smooth box-constrained nonlinear equations (see [35]). Pb2 depends on a parameter λ and we set $\lambda = 6$. Pb14 comes from the KKT condition of a convex Nash Equilibrium problem [16]. These latter problems were included in

Table 1: Test Problems

Pb #	Name and Source	n	Box
1	Chemical equilibrium system [22, system 1]	11000	$[0, \infty]$
2	Bratu NCP [13]	12500	$[0, \infty]$
3	Obstacle [13]	12500	$[0, \infty]$
4	Discrete boundary value function [28, Problem 28]	500	$[-100, 100]$
5	Trigexp1 [25, Problem 4.4]	1000	$[-100, 100]$
6	Troesch [25, Problem 4.21]	500	$[-1, 1]$
7	Trigonometric system [25, Problem 4.3]	5000	$[\pi, 2\pi]$
8	Tridiagonal exponential [25, Problem 4.18]	2000	$[e^{-1}, e]$
9	Countercurrent Reactors [25, Problem 4.1]	10000	$[-1, 10]$
10	Five Diagonal [25, Problem 4.8]	5000	$[1, \infty]$
11	Seven Diagonal [25, Problem 4.9]	5000	$[0, \infty]$
12	Bratu Problem [25, Problem 4.24]	10000	$[-\infty, 1.5]$
13	Poisson Problem [25, Problem 4.25]	10000	$[-5, 5]$
14	Spam Problem [16]	10100	$[0, +\infty]$

the tests set for testing the considered methods on problems with solutions on the boundary of the feasible set, as their solutions lie on $\partial\Omega$.

The Jacobian matrices of all the problems were evaluated analitically.

We implemented the AS_ID method in a Matlab code, using the standard spherical trust-region and the pioneer scaling matrix given by Coleman and Li. Furthermore, we have choosen to test a Newton-GMRES implementation. That is, we used Restarted-GMRES as iterative linear solver for computing p_k^{IN} satisfying (18). GMRES was restarted every 50 iterations, allowing a maximum of 20 restarts. We used the null initial guess. The forcing terms were computed by the adaptive choice given in [18], i.e. $\eta_0 = 0.9$, and $\eta_k = 0.9 \left\|F_k\right\|^2 / \left\|F_{k-1}\right\|^2$, for $k \geq 1$, with the safeguards suggested in [18, p. 305]. If GMRES fails in computing p_k^{IN} satisfying (18), the algorithm continued with p_k^{IN} given by the last computed GMRES iterate.

We avoided to use a preconditioner when GMRES managed to compute the inexact Newton step with the prescribed accuracy. Therefore, we used a preconditioning strategy only in the solution of Pb2, Pb7, Pb9, Pb12 and Pb13. We used the following preconditioning technique. At the first nonlinear iteration we compute a preconditioner for F_0' using the `ilu` Matlab function with drop tolerance set to 0.1. Then, the preconditioner is reused for preconditioning the linear systems arising in the subsequent nonlinear iterations. The preconditioner is recomputed whenever GMRES fails in satisfying the accuracy requirement (18).

We set $\Delta_0 = 1$, $\Delta_{min} = \sqrt{\epsilon_m}$, $\beta = 0.75$, $\theta = 0.99995$, $\delta = 0.25$. Moreover, at step 7.1 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 (10) holds (in this case, we set $\bar{\Delta}_{k+1} = \max\{\Delta_k,\, 2\|p_k\|\}$), otherwise, we left unchanged the radius. 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

$$\|F_k\| \leq 10^{-6} \tag{44}$$

was met. Such occurrence was indicated as a successful termination. On the other hand, the code declares a failure either if the number of nonlinear iterations was greater than 400 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 $\|F_{k+1} - F_k\| \leq 100\varepsilon_{\mathrm{m}}\|F_k\|$. This condition may indicate that the method does not manage to escape from a local minimizer of the merit function which is not a solution of (2). We remark that these two last situations never happened in our tests.

`Lsqnonlin` declares successful termination when the first-order optimality conditions for (3) are below a prescribed tolerance ϵ_1, i.e.

$$\|D_k \nabla f_k\|_\infty \leq \epsilon_1. \tag{45}$$

Moreover, the code terminates either if the norm of the current step is less than a prescribed tolerance ϵ_2, i.e.

$$\|x_{k+1} - x_k\| \leq \epsilon_2, \tag{46}$$

or if the relative change in the function value satisfies

$$\|F_{k+1} - F_k\| \leq \epsilon_1 \|F_k\|. \tag{47}$$

In this two latter situations `lsqnonlin` declares neither a failure nor a success of the procedure. Moreover, `lsqnonlin` is stopped when none of the above three stopping criteria is satisfied within 400 nonlinear iterations or the line search could not sufficiently decrease the residual along the current search direction. In our experiments we decided to consider a test successfully solved by `lsqnonlin` either if (45) is verified or one of the two conditions (46) and (47) is met and $\|F_k\| \leq 10^{-6}$.

In the first set of experiments we run `lsqnonlin` with the default tolerances for the stopping criteria. In the second set of experiments we followed [15] and, in order to compare the performance of the codes when the approximate solutions returned have the same level of accuracy, we re-run all the tests where (44) is not satisfied with the default tolerances, reducing the tolerances by a factor 10. This process is repeated until (44) is met or the tolerances provided to the solver reach 10^{-16}. In this latter case, a failure is declared. All the tests are run using the diagonal proconditioner provided by `lsqnonlin`.

We have chosen to measure the algorithms efficiency by the number It of iterations and the number Fe of F-evaluations. In fact, the measure in terms of iteration count takes into account the number of linear systems solved by each method and the number of Jacobian evaluations. On the other hand, the number of F-evaluations depends on the number of nonlinear iterations and the number of the performed reductions of the trust-region size when the step is

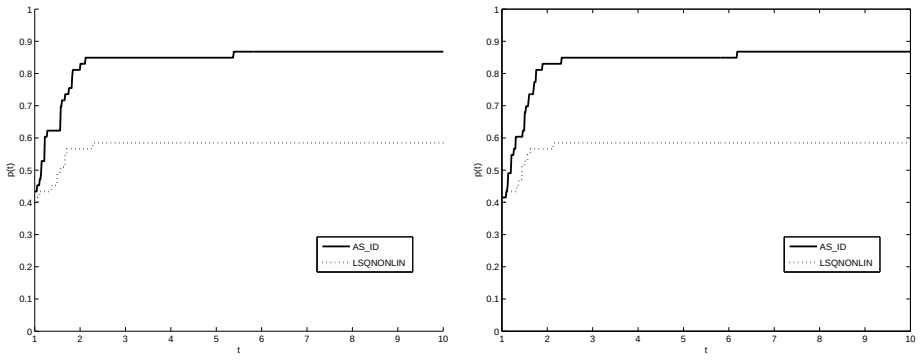

Figure 1: Performance profiles in terms of nonlinear iterations (left) and function evaluations (right); `lsqnonlin` run with default tolerances

rejected. Note that the computational cost of the algorithm increases as the number of rejected trial steps increases.

The results of the numerical experimentation are illustrated using the performace profile approach [14]. In this approach, when m solvers are compared on a test set, the performance of each solver in the solution of a test is measured by the ratio of its computational effort and the best computational effort by any solver on this test. In Figure 1 we report the performance profiles in terms of nonlinear iterations (left) and of F-evaluations (right); the figure refers to use of `lsqnonlin` with default tolerances. We recall that the right side of the plot gives the percentage of the test problems that are successfully solved by the solver. On the other hand, the left side of the plot gives the percentage of test problems for which the solver is the most efficient. As a first comment, we can observe that AS_ID successfully solves the 87% of tests. Namely, it fails in the solution of seven problems. Problems successfully solved required a reasonable number of nonlinear iterations and F evaluations: 16 nonlinear iterations and 18 F-evaluations as an average, except for Pb3 with $\nu = 0, 3$ and Pb7 with $\nu = 3$. On the other hand, `lsqnonlin` successfully solves the 59% of test. Then, the solver fails in the solution of 22 tests. Focusing on the tests successfully solved we underline that the stopping criterion (44) is not met in 13 out of 31 tests. The performance profiles also show that, using the number of linear iteration as performance measure, AS_ID is the best code in the solution of about the 43% of tests and it is within a factor two from the best code in the solution of 81% of tests. These percentages slightly decrease if we consider the number of F-evaluation as a performance measure.

Figure 2 displays performance profiles obtained running `lsqnonlin` with tolerances stricter than the standard ones, in order to enforce the convergence test (44). It is quite evident, as expected, that the number of successes of `lsqnonlin` increases. In fact, `lsqnonlin` manages to solve about the 75% of tests. At this regard, we mention that 13 failures of `lsqnonlin` with default

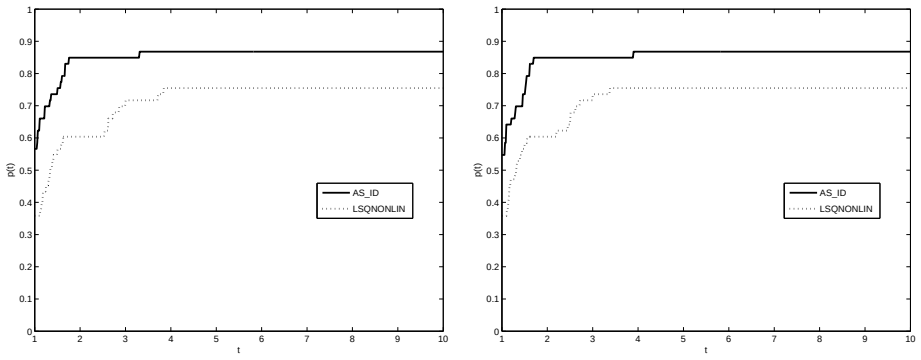

Figure 2: Performance profiles in terms of nonlinear iterations (left) and function evaluations (right); `lsqnonlin` run with tolerances chosen in order to satisfy (44)

tolerances are to be ascribed to the stopping criterion (47) that prematurely terminates the run. The use of stricter tolerance values allows to recover these failures. On the other hand, an increase in the number of nonlinear iterations and F evaluations was obviously observed.

In order to compare the two approaches in terms of efficiency, in Figure 3 we report the performance profiles obtained considering only tests successfully solved by both methods. In this figure, data concerning `lsqnonlin` refer to the stricter tolerances used in order to fulfill condition (44). This way, the two codes can be fairly compared, as the approximate solutions returned are obtained to the same level of accuracy. The major conclusion that can be drawn from Figure 3 is that method AS_ID is more efficient, both in terms of nonlinear iterations and F-evaluations than `lsqnonlin` in about the 58% of the tests successfully solved by both codes. Morever, AS_ID is within a factor two from the best code in the majority of these runs. All things considered, the approach taken in AS_ID seems to outperform `lsqnonlin` both in terms of efficiency and robustness.

Acknowledgments

The authors are grateful to Prof. Christian Kanzow and Dr. Axel Dreves for providing the Matlab files concerning Spam test problem.

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.

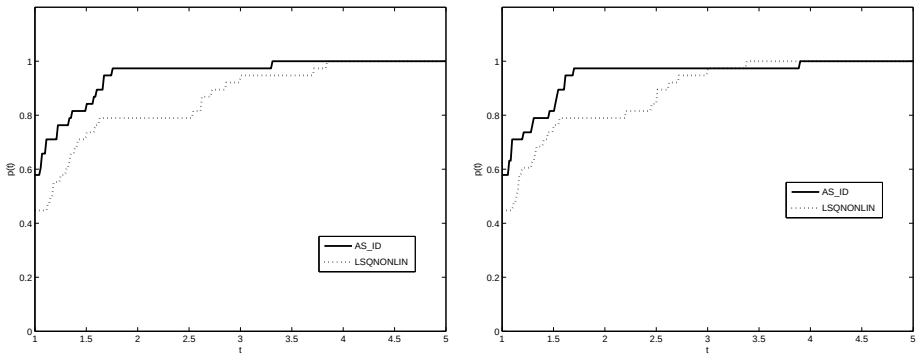

Figure 3: Performance profiles in terms of nonlinear iterations (left) and function evaluations (right) on tests successfully solved by both codes; `lsqnonlin` run with tolerances chosen in order to satisfy (44)

[2] ——, *STRSCNE: A scaled trust-region solver for constrained nonlinear equations*, Computational Optimization and Applications, 28 (2004), pp. 31–50.

[3] S. Bellavia, M. Macconi, and S. Pieraccini, *Constrained dogleg methods for nonlinear systems with simple bounds*, Computational Optimization and Applications, 53 (2012), pp. 771–794.

[4] S. Bellavia and B. Morini, *An interior global method for nonlinear systems with simple bounds*, Optimization Methods and Software, 20 (2005), pp. 1–22.

[5] ——, *Subspace trust-region methods for large bound constrained nonlinear equations*, SIAM J. Numer. Anal., 44 (2006), pp. 1535–1555.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[7] M. A. Branch, T. F. Coleman, and Y. Li, *A subspace, interior, and conjugate gradient method for large scale bound-constrained minimization problems*, SIAM J. Scientific Computing, 21 (1999), pp. 1–23.

[8] T. F. Coleman and Y. Li, *On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds*, Mathematical Programming, 67 (1994), pp. 189–224.

[9] ——, *An interior trust-region approach for nonlinear minimization subject to bounds*, SIAM J. Optim., 6 (1996), pp. 418–445.

[10] ——, *A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables*, SIAM J. Optim., 6 (1996), pp. 1040–1058.

[11] ——, *A trust region and affine scaling interior point method for nonconvex minimization with linear inequality constraints*, Mathematical Programming (Series A), 88 (1999), pp. 1–31.

[12] S. Di and W. Sun, *Trust region method for conic model to solve unconstrained optimization*, Optimization Methods and Software, 6 (1996), pp. 273–263.

[13] S. P. Dirkse and M. C. Ferris, *Mcplib: A collection of nonlinear mixed complementary problems*, Optimization Methods and Software, 5 (1995), pp. 319–345.

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

[15] E. D. Dolan, J. J. Moré, and T. Munson, *Optimality measures for performance profiles*, SIAM J. Optim., 16 (2006), pp. 891–909.

[16] A. Dreves, F. Facchinei, C. Kanzow, and S. Sagratella, *On the solution of the kkt conditions of generalized nash equilibrium problems*, SIAM J. Optim., 21 (2011), pp. 1082–1108.

[17] S. C. Eisenstat and H. F. Walker, *Globally convergent inexact Newton methods*, SIAM J. Optim., 4 (1994), pp. 393–422.

[18] ——, *Choosing the forcing term in an inexact Newton method*, SIAM J. Scientific Computing, 17 (1996), pp. 16–32.

[19] J. B. Francisco, N. Krejic, and J. M. Martinez, *An interior-point method for solving box-constrained underdetermined nonlinear systems*, Journal of Computational and Applied Mathematics, 177 (2005), pp. 67–88.

[20] W. W. Hager, B. A. Mair, and H. Zhang, *An affine-scaling interior-point CBB method for box-constrained optimization*, Mathematical Programming, (2007). Published online.

[21] 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.

[22] A. P. M. K. Meintjes, *Chemical equilibrium systems as numerical tests problems*, ACM Trans. Math. Soft., 16 (1990), pp. 143–151.

[23] C. Kanzow and A. Klug, *On affine-scaling interior-point Newton methods for nonlinear minimization with bound constraints*, Computational Optimization and Applications, 35 (2006), pp. 177–197.

[24] ——, *An interior-point affine-scaling trust-region method for semismooth equations with box constraints*, Computational Optimization and Applications, 37 (2007), pp. 329–353.

[25] 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.

[26] M. Macconi, B. Morini, and M. Porcelli, *A gauss-newton method for solving bound-constrained underdetermined nonlinear systems*, Optimization Methods and Software, 24 (2009), pp. 219–235.

[27] ——, *Trust-region quadratic methods for nonlinear systems of mixed equalities and inequalities*, Applied Numerical Mathematics, 59 (2009), pp. 859–876.

[28] J. J. Moré, B. Garbow, and K. Hillstrom, *Testing unconstrained optimization software*, ACM Trans. Math. Softw., 7 (1981), pp. 136–140.

[29] B. Morini and M. Porcelli, *Tresnei, a matlab trust-region solver for systems of nonlinear equalities and inequalities*, Computational Optimization and Applications, 51 (2012), pp. 27–49.

[30] R. P. Pawlowski, J. P. Simonis, H. F. Walker, and J. N. Shadid, *Inexact Newton dogleg methods*, SIAM J. Numer. Anal., (2008).

[31] M. Porcelli, *On the convergence of an inexact gauss-newton trust-region method for nonlinear least-squares problems with simple bounds*, Optimization Letters, 7 (2013), pp. 447–465.

[32] R. B. Schnabel and P. D. Frank, *Tensor methods for nonlinear equations*, SIAM J. Numer. Anal., 21 (1984), pp. 815–843.

[33] 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.

[34] L. N. Vicente, *Local convergence of the affine-scaling interior point algorithm for nonlinear programming*, Computational Optimization and Applications, 17 (2000), pp. 23–35.

[35] T. Wang, R. D. C. Monteiro, and J.-S. Pang, *An interior point potential reduction method for constrained equations*, Mathematical Programming, (1996), pp. 159–195.

[36] X. Wang and Y. X. Yuan, *A trust region method based on a new affine scaling technique for simple bounded optimization*, Optimization Methods and Software, 28 (2013), pp. 871–888.

[37] L. Zhao and W. Sun, *A conic affine scaling dogleg method for nonlinear optimization with bound constraints*, Asia-Pacific journal of Operation Research, 30 (2013).

[38] 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.