AN INTERIOR POINT NEWTON-LIKE METHOD FOR NONNEGATIVE LEAST SQUARES PROBLEMS WITH DEGENERATE SOLUTION *

STEFANIA BELLAVIA, MARIA MACCONI AND BENEDETTA MORINI †

Abstract. An interior point approach for medium and large nonnegative linear least-squares problems is proposed. Global and locally quadratic convergence is shown even if a degenerate solution is approached. Viable approaches for implementation are discussed and numerical results are provided.

Key words. convex quadratic programming, degeneracy, Interior Point methods, Inexact Newton methods, global convergence.

1. Introduction. Many applications, e.g. nonnegative image restoration, contact problems for mechanical systems, control problems, involve the numerical solution of Nonnegative Least Squares (NNLS) problems, i.e.

$$\min_{x\geq 0} q(x) = \frac{1}{2}\|Ax-b\|_2^2, \tag{1.1}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ are given and $m \geq n$, [1].

We assume that A has full column rank so that the NNLS problem (1.1) is a strictly convex optimization problem and there exists a unique solution $x^* \in \mathbb{R}^n_+$ for any vector b, [19]. We allow the solution x^* to be degenerate, that is strict complementarity does not hold at x^*.

Active set methods form a robust theoretical and practical class of methods containing a number of optimization algorithms. These methods applied to (1.1) generate a piecewise linear path following faces of the polytope defined by the constraints. These procedures are naturally effective for small problems while they can be slow to converge if the problem is large and the set of active constraints at x^* cannot be predicted well [1, 12, 17]. Active set methods and gradient projection algorithms have been combined to solve NNLS problems and bound-constrained minimization problems of medium and large dimension. These methods have been shown to be an efficient tool to solve these problems even in the presence of degenerate solutions, see e.g. [2, 18, 20, 21, 24]

In addition to active set approaches, Interior Point methods can be used to solve NNLS problems. They generate an infinite sequence of strictly feasible points converging to the solution and are known to be competitive with active set methods for medium and large problems. Several interior point methods for general bound-constrained problems, quadratic programming and NNLS problems have been proposed, see e.g [3, 4, 5, 6, 7, 13, 14, 15, 24]. To our knowledge, strict complementarity at the solution has been assumed to establish fast convergence properties of all interior-point methods for quadratic programming and NNLS problems. On the other hand, enhanced interior methods for general bound-constrained optimization problems have been given in [14, 15]. These methods are not ensured to converge for any arbitrary starting guess but show fast local convergence even in the presence of degeneracy.

*Work supported by MIUR, Rome, Italy, through MIUR, “Cofinanziamenti Programmi di Ricerca di Interesse Nazionale” and GNCS/INDAM.

†Dipartimento di Energetica “S. Stecco”, Università di Firenze, via C. Lombroso 6/17, 50134 Firenze, Italia, e-mail: stefania.bellavia@unifi.it, macconi@de.unifi.it, benedetta.morini@unifi.it

However, their application to (1.1) does not allow to take advantage of the structure of the NNLS problem.

The main motivation of this paper is to fill this gap providing an interior-point approach suited for NNLS problems. Global and locally fast convergence is guaranteed even if a degenerate solution is approached and the structure of the given problem is exploited both in the linear algebra phase and in the globalization strategy.

We consider the class of locally convergent Newton-like methods given by M. Heinkenschloss, M. Ulbrich and S. Ulbrich in [14] for solving the constrained nonlinear system that states the first-order necessary conditions for (1.1). We propose a modification of a locally convergent Newton-like method belonging to such class and and we embed it into a new globalization strategy which results to be very simple and cheap to implement. The method is formulated in a general way since the solution of the Newton equation, i.e. the linear system arising at each iteration, is allowed by either a direct method or inexactly with appropriately controlled accuracy.

The main computational work of the method in [14] is the solution of the Newton equation. A straightforward application of this method gives rise to a Newton equation which is related to the normal equation. But the symmetry and positive definiteness of the coefficient matrix is lost due to the presence of the constraints in (1.1). Our modification consists in a reformulation of the Newton equation that does not deteriorate the conditioning of the coefficient matrix and gives rise to a positive definite matrix. Hence, the solution of this system can be performed by Cholesky factorization or by short-recurrence iterative linear solvers. Further, in the case the system is solved by an iterative solver, a proper accuracy requirement is proposed.

The new globalization strategy adopts a simple rule for the transition from global to local convergence. Taking advantage from the fact that the function $q(x)$ is quadratic, a step producing a sufficient reduction of $q(x)$ is computed by a linear combination of the projected Newton step and a constrained scaled Cauchy step.

The computational experience highlights that the method is effective and has some distinguished features. Its convergence behaviour is insensitive to the dimension of the problem, to the number of active constraints at x^* and to the possible degeneracy of x^*. On the other hand, the conditioning of A may influence the cost of the globalization strategy.

Section 2 presents the local and global strategy for solving the NNLS problem and the resulting algorithm. Section 3 is devoted to the analysis of the local convergence rate while Section 4 provides several computational experiments.

2. A globally convergent interior point approach. In this section, we introduce a globally convergent Newton-like method for solving NNLS problems. It allows arbitrary positive starting vectors and shows local fast convergence to both degenerate and nondegenerate solutions. The iterates generated lie in the interior of the feasible region i.e. they are strictly positive. At each iteration the solution of a positive definite linear system is required and we investigate the possibility of solving such system by either a direct method or a sparse method. In the latter case, the resulting method belongs to the class of the inexact Newton procedures [8] and it can be effective in the solution of large NNLS problems where the matrix A is sparse.

This section is divided into three subsections. In Subsection 2.1 we present a locally convergent interior point method which is suitable for NNLS problems. In Subsection 2.2 we present a globalization strategy that is simple and cheap to implement. Finally, in Subsection 2.3 we sketch an algorithm which combines the ideas presented.

2.1. An affine scaling interior point method. In [14] a class of locally convergent affine scaling methods is introduced for bound-constrained minimization problems where the objective function is twice continuously differentiable on a open set containing the feasible region and the Hessian matrix is locally Lipschitz continuos on the feasible region. Such methods are locally fast convergent even if a degenerate solution is approached.

The core of our method is a modification of a method belonging to the framework given in [14]. We first show the application of such method to NNLS problems and then we sketch our modifications aimed to take advantage of the structure of NNLS problems.

The Karush-Kuhn-Tucker conditions state that the solution x^* to (1.1) satisfies

$$x^* \geq 0, \qquad g(x^*) \geq 0, \qquad g(x^*)^T x^* = 0,$$

where g is the gradient of the objective function, i.e. $g(x) = \nabla q(x) = A^T(Ax - b)$. Therefore, x^* can be found seeking for the positive solution of the following system of nonlinear equations

$$D(x)g(x) = 0, \tag{2.1}$$

where $D(x) = diag(d_1(x), \ldots, d_n(x))$, $x \geq 0$, has entries ([6])

$$d_i(x) = \begin{cases} x_i & \text{if} \quad g_i(x) \geq 0, \\ 1 & \text{otherwise} . \end{cases} \tag{2.2}$$

In (2.2) and in the sequel, x_i or $(x)_i$ denote the i-th component of a vector x.

In connection to the solution x^* to (1.1), we partition the index set $\{1, \ldots, n\}$ into the sets

$$\mathcal{I} = \{i \in \{1, \ldots, n\} : x_i^* > 0\}, \tag{2.3}$$

$$\mathcal{A} = \{i \in \{1, \ldots, n\} : x_i^* = 0,\ g_i^* > 0\}, \tag{2.4}$$

$$\mathcal{N} = \{i \in \{1, \ldots, n\} : x_i^* = 0,\ g_i^* = 0\}. \tag{2.5}$$

The set $\mathcal{I}$ represents the indices of the inactive components of x^*, the sets $\mathcal{A}$ and $\mathcal{N}$ contain the indices of the active components with and without strict complementarity, respectively. We say that strict complementarity at x^* holds if the set $\mathcal{N}$ is empty. Otherwise, the solution x^* is said to be degenerate.

In the sequel, we use the subscript k as index for any sequence and for any function f we denote $f(x_k)$ by f_k. Given $x_k > 0$, if the Newton method is used to solve (2.1), the Newton equation takes the form

$$(D_k A^T A + E_k)p = -D_k g_k, \tag{2.6}$$

where the coefficient matrix is obtained by formal application of the product rule, $E_k = E(x_k)$ and $E(x) = diag(e_1(x), \ldots, e_n(x))$ is a diagonal matrix.

In [5] the not everywhere existing derivatives $\frac{\partial}{\partial x_i} d_i(x)$, $i = 1, \ldots, n$, are substituted by the real valued functions $d_i^1(x) = \max(0, \text{sign}(g_i(x))$ and consequently, $E(x)$ is defined setting $e_i(x) = d_i^1(x) g_i(x) \geq 0$, $i = 1, \ldots, n$. The corresponding method was shown to be locally quadratic convergent under certain assumptions including strict complementarity at the solution x^*, [5].

In [14], to develop fast convergent methods without assuming strict complementarity at x^*, two modifications are introduced in (2.6). One modification concerns the scaling matrix $D(x)$ and/or the matrix $E(x)$. In the method we are interested in, the matrix $D(x)$ is left unchanged while the following entries in the matrix $E(x)$ are employed:

$$(2.7)\qquad e_i(x) = \begin{cases} g_i(x) & \text{if} \quad 0 \le g_i(x) < x_i^s \text{ or } g_i(x)^s > x_i \\ 0 & \text{otherwise} \end{cases}, \qquad 1 < s \le 2\,.$$

Note that $e_i(x) = d_i^1(x)g_i(x)$ for all $i \in \mathcal{I} \cup \mathcal{A}$ and x sufficiently close to x^*. Then, around a solution x^*, the modification introduced affects the matrix $E(x)$ only in the presence of degeneracy.

The second modification in (2.6) is due to the fact that approaching a degenerate solution x^*, the coefficient matrix tends to become singular. In particular, if the subset $\mathcal{N}$ is not empty and $\{x_k\}$ is a positive sequence such that $\lim_{k\to\infty} x_k = x^*$, then for $i \in \mathcal{N}$ we have $\lim_{k\to\infty} d_i(x_k) = 0$, $\lim_{k\to\infty} e_i(x_k) = 0$. Hence, the ith row of $(D_k\, A^T A + E_k)$ tends to zero and $\lim_{k\to\infty} \|(D_k\, A^T A + E_k)^{-1}\| = \infty$.

To overcome this pitfall, the Newton equation (2.6) is restated as

$$(2.8)\qquad W_k D_k M_k p = -W_k D_k g_k,$$

where the matrices $M(x)$ and $W(x)$ are definite for $x > 0$ by

$$(2.9)\qquad M(x) = A^T A + D(x)^{-1} E(x),$$

$$(2.10)\qquad W(x) = diag(w_1(x), \ldots, w_n(x)), \qquad w_i(x) = \frac{1}{d_i(x) + e_i(x)}.$$

Clearly, the matrices $D(x)$ and $W(x)$ are invertible for $x > 0$ and are positive definite. Further, the matrix $(W(x)D(x)M(x))^{-1}$ exists and is uniformly bounded for all positive x, see [14, Lemma 2].

Direct methods for (2.8) need to form $A^T A$. In the case A is large and sparse, $A^T A$ may be almost dense and this makes direct methods prohibitively costly in terms of storage and operations. On the other hand, iterative methods provide an effective mean to solve (2.8) as only the action of A and A^T on vectors is required and the matrix A need not be stored. This feature is also relevant in the case where A is regarded as an operator for obtaining the action of A and A^T on vectors cheaply, [13].

If the linear system (2.8) is not exactly solved, the resulting Newton method for (2.1) belongs to the class of inexact Newton methods [8] and the Newton equation takes the form

$$(2.11)\qquad W_k D_k M_k p_k = -W_k D_k g_k + r_k.$$

The convergence analysis developed in [14] is carried out assuming that the residual vector r_k satisfies

$$(2.12)\qquad \|r_k\|_2 \le \eta_k \|W_k D_k g_k\|_2, \qquad \eta_k \in [0, 1).$$

The scalar η_k in (2.12) affects the properties of the inexact method. Since (2.8) can be viewed as a special case of (2.11) where $\eta_k = 0$, for sake of generality, in the rest of the paper we will consider the inexact framework.

Let p_k be the inexact Newton step satisfying (2.11) and (2.12). To ensure strict feasibility of the iterates, the new iterate is given by

$$(2.13)\qquad x_{k+1} = x_k + \hat{p}_k,$$

where

$$(2.14) \qquad \hat{p}_k = \max\{\sigma,\, 1 - \|P(x_k + p) - x_k\|\,\}\,(P(x_k + p) - x_k), \qquad \sigma < 1,$$

and $P(x)$ is the projection of x onto the positive orthant, i.e. $P(x) = \max\{0, x\}$ with max meant componentwise. Here and in the sequel $\|\cdot\|$ is the 2-norm.

The theorem below establishes that the local convergence rate depends on the scalar s used in (2.7). We remark that the theorem holds even if x^* is degenerate.

THEOREM 2.1. *Let $s \in (1, 2]$ be the scalar s in the definition (2.7) of matrix E. There exists $\rho > 0$ such that for $x_0 \in B_\rho(x^*)$,*

i) if $\eta_k = 0$, $k \geq 0$, the sequence $\{x_k\}$ generated by the Newton-like method converges locally with q-order s toward x^*;

ii) if $\eta_k \leq \Lambda \|W_k D_k g_k\|^{s-1}$ for some $\Lambda > 0$ and $k \geq 0$, then the sequence $\{x_k\}$ generated by the inexact Newton-like method converges locally with q-order s toward x^*.

Proof. See [14, Th. 4, Th. 6]

Despite the good theoretical properties of the Newton methods aforementioned, their straightforward application to NNLS problems may be inappropriate. The Hessian $A^T A$ of the quadratic function q is symmetric and positive definite but the transformation $W_k D_k$ destroys this feature in the coefficient matrix of (2.8). On the other hand, it is apparent the potential benefit of handling the Newton equation with symmetric positive definite linear systems. Algorithms and software for factorizing nonsymmetric matrices are more complicated and computationally demanding than Cholesky algorithm both in case of medium and large dimensional matrix A. Also, the nonsymmetric iterative methods based on short recurrences, e.g. QMR [11], BI-CG [16, 10], BI-CGSTAB [26], are not optimal in the sense of error or residual norm minimization. GMRES method [25] is based on the residual norm minimization but requires full-term recurrence while a restarted version may stagnate.

The above remarks make clear the properties which are useful for the development of Newton methods for NNLS. Note that it is straightforward to reformulate (2.11) as

$$W_k D_k M_k W_k D_k \tilde{p}_k = -W_k D_k g_k + r_k,$$

where $\tilde{p}_k = D_k^{-1} W_k^{-1} p_k$. This way, the coefficient matrix is symmetric positive definite for $x_k > 0$ but the existence and uniform boundness of the inverse is not ensured in a neighborhood of x^*. For example, consider a sequence $\{x_k\}$ s.t. $\lim_{k\to\infty} x_k = x^*$ and let some i belong to the set $\mathcal{A}$ in (2.4). Then, we have $\lim_{k\to\infty}(d_k)_i = 0$, $\lim_{k\to\infty}(e_k)_i \neq 0$ and consequently the ith row of $W_k D_k M_k W_k D_k$ tends to zero.

Next we propose a reformulation of the Newton equation which is advantageous in a number of respects. In particular, the coefficient matrix of the reformulated Newton equation is symmetric and positive definite, its inverse is uniformly bounded and the condition number is at most equal to the condition number of the matrix in (2.11). This enables us to use a possibly sparse Cholesky direct method or an iterative method that converges in at most n iterations for all nonsingular coefficient matrices and is based on short-recurrence such as CGLS or LSQR [23].

Let us introduce, for any $x > 0$, two matrices $S(x)$ and $Z(x)$ as follows:

$$(2.15) \qquad S(x) = W(x)^{\frac{1}{2}}\, D(x)^{\frac{1}{2}},$$

$$(2.16) \qquad Z(x) = S(x) M(x) S(x) = S(x)^T A^T A S(x) + W(x) E(x).$$

Note that the matrix $S(x)$ is invertible for any $x > 0$ and $\|S(x)\| \le 1$. Moreover, since $W(x)$ and $E(x)$ are diagonal matrices with positive entries, it follows that $Z(x)$ is symmetric positive definite for $x > 0$.

The next lemma shows that for $x > 0$, the matrices $(S(x)M(x))^{-1}$ and $Z(x)^{-1}$ exist and are uniformly bounded and that the condition number of $Z(x)$ does not deteriorate with respect to $W(x)D(x)M(x)$ in (2.11).

LEMMA 2.1. *There exist two positive constants C_1, C_2 such that for all $x > 0$, the matrices $W(x)D(x)M(x)$, $S(x)M(x)$ and $Z(x)$ are nonsingular and*

$$\|(W(x)D(x)M(x))^{-1}\| \le C_1, \tag{2.17}$$

$$\|(S(x)M(x))^{-1}\| \le C_1, \tag{2.18}$$

$$\|Z(x)^{-1}\| \le C_2. \tag{2.19}$$

Further,

$$k_2(Z(x)) \le k_2(W(x)D(x)M(x)), \tag{2.20}$$

where k_2 is the condition number in the 2-norm.

Proof. Assume $x > 0$. From [14, Lemma 2] we know that $W(x)D(x)M(x)$ is invertible and inequality (2.17) holds for some $C_1 > 0$.

Since $S(x)M(x) = S(x)^{-1}W(x)D(x)M(x)$, it follows that $S(x)M(x)$ is invertible and

$$\|(S(x)M(x))^{-1}\| \le \|S(x)\|C_1 \le C_1.$$

Showing (2.19) is equivalent to prove that there exists $C_2 > 0$ such that

$$\|Z(x)u\| \ge \frac{1}{C_2}\,\|u\|, \qquad \forall u \in \mathbb{R}^n. \tag{2.21}$$

Given an arbitrary $u \in \mathbb{R}^n$, set $u_d = W(x)D(x)u$, $u_e = W(x)E(x)u$. Clearly, $W(x)D(x) + W(x)E(x) = I$ implies $u = u_d + u_e$. Let $\beta \in (0,1)$, $\mu = \frac{1-\beta}{1+\|A^TA\|}$. We distinguish the case $\|u_d\| \ge \mu\|u\|$ from the case $\|u_d\| < \mu\|u\|$.

If $\|u_d\| \ge \mu\|u\|$, we have

$$\|S(x)u\| = \|S(x)^{-1}u_d\| \ge \frac{\|u_d\|}{\|S(x)\|} \ge \mu\|u\|.$$

By $u^TW(x)E(x)u \ge 0$ and (2.16) we get

$$\|u\|\|Z(x)u\| \ge u^TZ(x)u \ge \|AS(x)u\|^2.$$

Furthermore, since A has full column rank, it follows that

$$\|AS(x)u\|^2 \ge \sigma_{min}^2(A)\|S(x)u\|^2 \ge \frac{\mu^2}{\|A^+\|^2}\,\|u\|^2,$$

where $\sigma_{min}(A)$ is the smallest nonzero singular value of A and A^+ is the Moore-Penrose pseudoinverse of A. Hence, we obtain

$$\|Z(x)u\| \ge \frac{\mu^2}{\|A^+\|^2}\,\|u\|. \tag{2.22}$$

If $\|u_d\| < \mu\|u\|$, by $u = u_d + u_e$ we get

$$\begin{aligned}\|u_e\|\|Z(x)u\| &\geq u_e^T Z(x)u \\ &= u_e^T S(x)A^T AS(x)u_d + \|AS(x)u_e\|^2 + \|u_e\|^2 \\ &\geq \|u_e\|^2 - \|A^T A\|\|u_e\|\|u_d\|,\end{aligned}$$

where the last inequality derives from $|u_e^T S(x)A^T AS(x)u_d| \leq \|A^T A\|\|u_d\|\|u_e\|$. Then, using again $u = u_d + u_e$, we obtain

$$\begin{aligned}\|Z(x)u\| &\geq \|u_e\| - \|A^T A\|\,\|u_d\| \geq \|u\| - \|u_d\| - \mu\|A^T A\|\|u\| \\ &> (1 - \mu\|A^T A\| - \mu)\,\|u\| = \beta\,\|u\|.\end{aligned} \tag{2.23}$$

Thus, from (2.22) and (2.23) inequality (2.21) holds with $C_2 = \max\{\frac{\|A^+\|^2}{\mu^2}, \frac{1}{\beta}\}$.

We now prove (2.20). Note that

$$S(x)Z(x)S(x)^{-1} = W(x)D(x)M(x).$$

Then, $Z(x)$ and $W(x)D(x)M(x)$ are similar and therefore they have the same spectrum. Denoting λ_{max} and λ_{min} the maximum and the minimum eigenvalue of the matrices, since Z is positive definite we get

$$k_2(W(x)D(x)M(x)) \geq \frac{\lambda_{max}}{\lambda_{min}} = k_2(Z(x)).$$

□

With this result at hand, we can reformulate the Newton equation to handle linear systems with symmetric positive definite matrices while the properties of the overall methods are unaltered. From (2.11) we have

$$S_k M_k p_k = -S_k g_k + \tilde{r}_k, \tag{2.24}$$

with $\tilde{r}_k = S_k^{-1} r_k$ and letting $\tilde{p}_k = S_k^{-1} p_k$, one obtain the following linear system

$$Z_k \tilde{p}_k = -S_k g_k + \tilde{r}_k. \tag{2.25}$$

In this linear system, Z_k is positive definite, its inverse is uniformly bounded and its conditioning is not worse than that of $W_k D_k M_k$.

We conclude this section with some computational remarks. At each iteration, we need an inexact step p_k satisfying (2.11) and (2.12). The computation of such step is performed solving (2.25) for $\tilde{p}_k$ and setting $p_k = S_k\tilde{p}_k$. By construction, $p_k = S_k\tilde{p}_k$ solves (2.11) with $r_k = S_k\tilde{r}_k$. Then, to enforce (2.12) one can apply CGLS to (2.25) and compute a vector $\tilde{p}_k$ such that

$$\|\tilde{r}_k\| \leq \eta_k\|W_k D_k g_k\|, \qquad \eta_k \in [0, 1). \tag{2.26}$$

This residual control is not the classical one associated to (2.25) but, as $\|S_k\| \leq 1$, it forces $r_k = S_k\tilde{r}_k$ to meet (2.12).

Alternatively, we can reformulate (2.25) as the least-square problem

$$\min_{\tilde{p} \in \mathbb{R}^n} \|B_k\tilde{p} + z_k\|, \tag{2.27}$$

with

$$B_k = \begin{pmatrix} AS_k \\ W_k^{\frac{1}{2}} E_k^{\frac{1}{2}} \end{pmatrix}, \qquad z_k = \begin{pmatrix} Ax_k - b \\ 0 \end{pmatrix},$$

and apply LSQR to (2.27) to compute an inexact solution $\tilde{p}_k$. Repeating the above argument, the residual control

$$\|\tilde{r}_k\| = \|B^T(B\tilde{p}_k - z_k)\| \leq \eta_k \|W_k D_k g_k\|, \tag{2.28}$$

ensures that the vector $p_k = S_k \tilde{p}_k$ satisfies (2.11) and (2.12).

2.2. The globalization technique. In this subsection we wrap the previously given Newton-like method into a globalization technique.

Chosen a positive starting point x_0 in some arbitrary way, we generate a sequence of strictly feasible iterates $\{x_k\}$ such that $\{q_k\}$ is monotonically decreasing. To this end, let $x_k > 0$ be given and, for $p \in \mathbb{R}^n$, define the quadratic function $\psi_k(p)$ as follows:

$$\begin{aligned} \psi_k(p) &= \frac{1}{2} p^T M_k p + p^T g_k \\ &= \frac{1}{2} p^T (A^T A + D_k^{-1} E_k) p + p^T A^T (A x_k - b). \end{aligned} \tag{2.29}$$

The functions $\psi_k(p)$ and $q(x)$ are related so that

$$q(x_k) - q(x_k + p) = -\psi_k(p) + \frac{1}{2} p^T D_k^{-1} E_k p.$$

Since $D_k^{-1} E_k$ is semipositive definite, it follows that

$$q(x_k) - q(x_k + p) \geq -\psi_k(p). \tag{2.30}$$

Further, it is clear that $\psi_k(0) = 0$ and any step $p \in \mathbb{R}^n$ such that $\psi_k(p) < 0$ yields a reduction in the value of q.

To design a global strategy theoretically well founded, we employ a constrained scaled Cauchy step p_k^C of the form ([5])

$$p_k^C = -\tau_k D_k g_k, \qquad \tau_k > 0. \tag{2.31}$$

We require the point $x_k + p_k^C > 0$ to be strictly positive and therefore we choose τ_k as either the unconstrained minimizer of the quadratic function $\psi_k(-\tau D_k g_k)$ or a fraction of the distance to the boundary along $-D_k g_k$, i.e

$$\tau_k = \begin{cases} \tau_1 = \operatorname{argmin}_\tau \psi_k(-\tau D_k g_k) = \frac{g_k^T D_k g_k}{g_k^T D_k^T M_k D_k g_k}, & \text{if } x_k - \tau_1 D_k g_k > 0 \\ \tau_2 = \theta \min_{1 \leq i \leq n} \{ \frac{(x_k)_i}{(D_k g_k)_i} : (D_k g_k)_i > 0 \} & \text{otherwise} \end{cases} \tag{2.32}$$

with $\theta \in (0, 1)$.

Trivially $\psi_k(p_k^C) < 0$, and global convergence depends on taking a point x_{k+1} satisfying

$$red(x_{k+1} - x_k) \geq \beta, \qquad \beta \in (0,\, 1), \tag{2.33}$$

where

$$red(p) = \frac{\psi_k(p)}{\psi_k(p_k^C)}, \qquad p \in \mathbb{R}^n. \tag{2.34}$$

Enforcing (2.33), implies $\psi_k(p) < 0$ and by (2.30) we get

$$q(x_k) - q(x_{k+1}) \geq -\beta \psi_k(p_k^C), \tag{2.35}$$

i.e. we obtain a reduction in the value of q that is at least a fixed multiple of $-\psi_k(p_k^C)$.

To form the new iterate, we set $x_{k+1} = x_k + s_k$ where

$$s_k = t\,(p_k^C - \hat{p}_k) + \hat{p}_k, \tag{2.36}$$

with $t \in [0, 1)$ and $\hat{p}_k$ given in (2.14). In particular, we set $x_{k+1} = x_k + \hat{p}_k$ if $\hat{p}_k$ satisfies (2.33). Otherwise, we compute the step satisfying $red(s_k) = \beta$. By (2.36) this problem is equivalent to solve the scalar problem

$$\pi(t) = \psi_k(t(p_k^C - \hat{p}_k) + \hat{p}_k) - \beta \psi_k(p_k^C) = 0. \tag{2.37}$$

Note that by (2.29), $\pi(t)$ has the form

$$\pi(t) = \frac{1}{2}(p_k^C - \hat{p}_k)^T M_k (p_k^C - \hat{p}_k)\, t^2 + (p_k^C - \hat{p}_k)^T (M_k \hat{p}_k + g_k)\, t + \psi_k(\hat{p}_k) - \beta \psi_k(p_k^C).$$

Since $\pi(0) > 0$, $\pi(1) < 0$, we conclude that $\pi(t)$ has two positive roots and one of them lies in $(0, 1)$. Consequently, the scalar t we need is the smaller root of π.

Next we prove global convergence of our procedure.

THEOREM 2.2. *Let $x_0 \in \mathbb{R}^n$ be an arbitrary positive initial point. The sequence $\{x_k\}$ converges to the solution of the NNLS problem.*

Proof. Since the sequence $\{q_k\}$ is decreasing and bounded below from zero, it follows $\lim_{k\to\infty}(q_k - q_{k+1}) = 0$ and by (2.35)

$$\lim_{k\to\infty} \psi(p_k^C) = 0. \tag{2.38}$$

Moreover, note that

$$q(x) \geq \frac{1}{2}\|Ax\|^2 - |b^T Ax| \geq \left(\frac{1}{2\|A^+\|^2} - \frac{\|b^T A\|}{\|x\|} \right) \|x\|^2.$$

Then, as $\{q_k\}$ is bounded above by construction, we can conclude that $\{x_k\}$ is a bounded sequence. Due to the boundness of $\{x_k\}$, there exists a constant $\chi > 0$ such that for $k \geq 0$

$$\|D_k\| \leq \chi, \qquad \|E_k\| \leq \|g_k\|_\infty \leq \chi. \tag{2.39}$$

Now we show that there exists a constant $\phi > 0$ such that

$$-\psi(p_k^C) \geq \phi \|D_k g_k\|^2. \tag{2.40}$$

To this end, we consider the possible values of τ_k in (2.32).

If $\tau_k = \tau_1$, with τ_1 given in (2.32), then from (2.31) we have

$$-\psi_k(p_k^C) = \frac{1}{2}\,\frac{\|D_k^{\frac{1}{2}} g_k\|^4}{g_k^T D_k^{\frac{1}{2}} D_k^{\frac{1}{2}} M_k D_k^{\frac{1}{2}} D_k^{\frac{1}{2}} g_k} \geq \frac{1}{2}\,\frac{\|D_k^{\frac{1}{2}} g_k\|^2}{\|D_k^{\frac{1}{2}} M_k D_k^{\frac{1}{2}}\|},$$

and from (2.39) we get

$$\begin{aligned} &\|D_k^{\frac{1}{2}} M_k D_k^{\frac{1}{2}}\| = \|D_k^{\frac{1}{2}} A^T A D_k^{\frac{1}{2}} + E_k\| \leq \chi\,(\|A^T A\| + 1), \\ &\|D_k g_k\|^2 \leq \chi\,\|D_k^{\frac{1}{2}} g_k\|^2. \end{aligned} \tag{2.41}$$

Hence,

$$-\psi(p_k^C) \geq \frac{1}{2\chi^2(\|A^T A\| + 1)}\,\|D_k g_k\|^2.$$

Suppose $\tau_k = \tau_2$ with τ_2 given in (2.32). Recalling that $\tau_2 < \tau_1$, we obtain

$$\begin{aligned} -\psi_k(p_k^C) &= \theta\tau_2\,(\|D_k^{\frac{1}{2}} g_k\|^2 - \frac{1}{2}\theta\,\tau_2\,\|M_k^{\frac{1}{2}} D_k g_k\|^2) \\ &\geq \theta\tau_2\,(\|D_k^{\frac{1}{2}} g_k\|^2 - \frac{1}{2}\tau_1\,\|M_k^{\frac{1}{2}} D_k g_k\|^2) = \frac{\theta\tau_2}{2}\,\|D_k^{\frac{1}{2}} g_k\|^2. \end{aligned} \tag{2.42}$$

Since $\tau_2 = \frac{(x_k)_{\hat{\imath}}}{(D_k g_k)_{\hat{\imath}}}$ for some $\hat{\imath} \in \{1, \ldots, n\}$ with $(D_k g_k)_{\hat{\imath}} > 0$, from (2.2) trivially follows $(g_k)_{\hat{\imath}} > 0$, $(d_k)_{\hat{\imath}} = (x_k)_{\hat{\imath}}$ and $\tau_2 = 1/(g_k)_{\hat{\imath}} \geq 1/\|g_k\|_\infty$. Thus, from (2.41) and (2.42) we obtain

$$-\psi_k(p_k^C) \geq \frac{\theta}{2\|g_k\|_\infty}\|D_k^{\frac{1}{2}} g_k\|^2 \geq \frac{\theta}{2\chi^2}\|D_k g_k\|^2.$$

Therefore, condition (2.40) follows with $\phi = \frac{1}{2\chi^2}\min\{\frac{1}{\|A^T A\|+1}, \theta\}$ and (2.38) yields $\lim_{k\to\infty}\|D_k g_k\| = 0$. Thus, any limit point of $\{x_k\}$ is a stationary point for (1.1).

Since q is a strictly convex function, each limit point is a global minimum (1.1). Uniqueness of the solution to problem (1.1) implies that $\lim_{k\to\infty} x_k = x^*$. □

2.3. The Algorithm. We now summarize all above ideas into an algorithm where a flag `IN` is used to distinguish the Inexact Newton method (`IN=true`) from the exact one (`IN=false`). In other words, `IN=false` means $\eta_k = 0$ for $k \geq 0$.

The resulting method is summarized below:

THE BASIC ALGORITHM

Choose $x_0 > 0$, `IN`, $\beta \in (0,1)$, $\theta \in (0,1)$, $\sigma \in (0,1)$. Set $k = 0$.
Repeat until tests for termination fail,

1. Compute p_k^C given in (2.31).
2. If `IN=true` then choose $\eta_k \in (0,1)$
   else set $\eta_k = 0$.
3. Solve (2.25) for $\tilde{p}_k$ with $\tilde{r}_k$ satisfying (2.26).
4. Set $p_k = S_k\,\tilde{p}_k$.
5. Form $\hat{p}_k$ by (2.14).
6. If $red(\hat{p}_k) \geq \beta$ then set $t = 0$

else compute the smaller root t of (2.37).

7. Form s_k from (2.36).
8. Set $x_{k+1} = x_k + s_k, \quad k = k+1$.

We conclude the section with some remarks on the computational work required at each iteration of the above algorithm. The most computationally intensive step is the solution of the linear system (2.25) at Step 3 and it is clear that the associated computational cost depends on the linear solver employed in the implementation of the algorithm. Regarding the work required to perform the other steps of the algorithm, we neglect scalar products and the products of diagonal matrices times a vector, and we give a brief account of the number of matrix-vector multiplications of the form Ax, $A^T y$ required.

Note that for $v \in \mathbb{R}^n$ and $M(x)$ given in (2.9), computing $v^T M_k v$ requires one product of A times v as $v^T M_k v = \|Av\|^2 + v^T D_k E_k v$. Then, each iteration of the algorithm requires three products of A times a vector and one product of A^T times a vector. Specifically, at Step 1, g_k and τ_1 given in (2.32) must be computed to evaluate p_k^C. These tasks require a total of two matrix-vector products employing A and one matrix-vector product employing A^T. The remaining matrix-vector product is needed at Step 6 to evaluate $\psi_k(\hat{p}_k)$ and then form $red(\hat{p}_k)$. At this regard, note that in (2.34) the scalar $\psi_k(p_k^C)$ is evaluated using a matrix-vector product already computed to form τ_1. Finally, we remark that solving (2.37) is inexpensive because this is a quadratic equation and it is easy to see that the computation of its exact roots requires matrix-vector products available from the previously performed steps of the algorithm.

3. Local convergence rate. In this section we discuss the local convergence rate of our method. We prove that for $\eta_k = 0$, $k \geq 0$, the procedure has q-order of convergence equal to s where s is the scalar in (2.7). Under a suitable choice of the forcing terms, this order of convergence is attained by the inexact method, too. Namely, the convergence rate of the local Newton-like methods given in Section 2 is preserved.

The convergence rate depends on the steps taken in the vicinity of the solution x^* i.e. on the step providing (2.33). Next, we call inexact step the vector p_k computed in Step 4 of our algorithm and we let p_k^N be the Newton step i.e. the solution to (2.8). Note that p_k^N is the global minimizer of the function $\psi_k(p)$.

To prove our thesis, we show that for k sufficiently large, the projected inexact step $\hat{p}_k$ computed in Step 5 of the Algorithm satisfies (2.33). Thus, Theorem 2.1 provides the final result.

We begin recalling the following result from [14], that characterizes the local behaviour of the inexact step. In the sequel, for any vector $y \in \mathbb{R}^n$, the open ball with center y and radius ρ is indicated by $B_\rho(y)$, i.e. $B_\rho(y) = \{x : \|x - y\| < \rho\}$.

Lemma 3.1. *Let $x_k > 0$, $s \in (1, 2]$ be the scalar in (2.7) and $\eta_k \leq \Lambda \|W_k D_k g_k\|^{s-1}$ for some $\Lambda > 0$. Then, there are $\zeta > 0$ and $\rho > 0$ such that if $x_k \in B_\rho(x^*)$,*

$$\|x_k + p_k - x^*\| \leq \zeta \|x_k - x^*\|^s. \tag{3.1}$$

Proof. See [14, Th. 4, Th. 6] □

To investigate on the satisfaction of (2.33), the features of ψ_k are analyzed. In the vicinity of x^*, the value of ψ_k at p_k is estimated by the following result.

LEMMA 3.2. *Let $\{x_k\}$ converge to the solution x^* of the NNLS problem, $s \in (1, 2]$ be the scalar in (2.7). Then, there are positive constants Γ_1 and τ such that for $x_k \in B_\tau(x^*)$, $\eta_k \le \Lambda \|W_k D_k g_k\|^{s-1}$ for some $\Lambda > 0$, we have*

$$\psi_k(p_k) \le \psi_k(p_k^N) + \Gamma_1 \eta_k^2 \|x_k - x^*\|^2. \tag{3.2}$$

Proof. Let $\tau \in (0, \rho]$ where ρ is given in Lemma 3.1. If the set $\mathcal{A}$ in (2.4) is not empty, we reduce τ so that when $x_k \in B_\tau(x^*)$, by continuity for any index $i \in \mathcal{A}$ the condition $(g^*)_i \neq 0$ implies $(g^*)_i (g_k)_i > 0$ and

$$(d_k)_i = (x_k)_i, \qquad (e_k)_i = (g_k)_i, \quad i \in \mathcal{A}. \tag{3.3}$$

By (2.8) and (2.24) we get

$$p_k^N - p_k = -M_k^{-1} S_k^{-1} \tilde{r}_k, \qquad M_k p_k + g_k = S_k^{-1} \tilde{r}_k.$$

Then, since M_k is positive definite we get

$$\begin{aligned} \psi_k(p_k^N) = \psi_k(p_k + p_k^N - p_k) &= \psi_k(p_k) + (p_k^N - p_k)^T (M_k p_k + g_k) + \frac{1}{2}(p_k^N - p_k)^T M_k (p_k^N - p_k) \\ &\ge \psi_k(p_k) + (p_k^N - p_k)^T (M_k p_k + g_k), \\ &\ge \psi_k(p_k) - \tilde{r}_k^T Z_k^{-1} \tilde{r}_k, \end{aligned}$$

with $Z(x)$ given in (2.16). Hence, by using (2.19) and (2.26) we obtain

$$\psi_k(p_k^N) \ge \psi_k(p_k) - C_2 \|\tilde{r}_k\|^2 \ge \psi_k(p_k) - C_2 \eta_k^2 \|W_k D_k g_k\|^2.$$

Now we investigate the term $\|W_k D_k g_k\|$. From (2.3)-(2.5) and the positivity of $(d_k)_i$, $(e_k)_i$, $i = 1, \ldots, n$ we have

$$\begin{aligned} \|W_k D_k g_k\|^2 &= \sum_{i \in \mathcal{I} \cup \mathcal{N}} \left(\frac{(d_k)_i (g_k)_i}{(d_k)_i + (e_k)_i} \right)^2 + \sum_{i \in \mathcal{A}} \left(\frac{(d_k)_i (g_k)_i}{(d_k)_i + (e_k)_i} \right)^2 \\ &\le \sum_{i \in \mathcal{I} \cup \mathcal{N}} \left(\frac{(d_k)_i (g_k)_i}{(d_k)_i} \right)^2 + \sum_{i \in \mathcal{A}} \left(\frac{(d_k)_i (g_k)_i}{(e_k)_i} \right)^2. \end{aligned}$$

Further, since $(g^*)_i = 0$ if $i \in I \cup N$ and (3.3) holds if $i \in A$, it follows

$$\begin{aligned} \|W_k D_k g_k\|^2 &\le \sum_{i \in I \cup N} (g_k - g_i^*)^2 + \sum_{i \in A} (x_k - x^*)_i^2 \\ &\le \|g_k - g^*\|^2 + \|x_k - x^*\|^2 \\ &= \|A^T A (x_k - x^*)\|^2 + \|x_k - x^*\|^2. \end{aligned} \tag{3.4}$$

Thus, letting $\Gamma_1 = C_2(1 + \|A^T A\|^2)$, the thesis follows. □

Now, we estimate the value of ψ_k when the inexact Newton step p_k is projected onto the feasible set.

LEMMA 3.3. *Let $\{x_k\}$ converge to the solution x^* of the NNLS problem, $s \in (1, 2]$ be the scalar in (2.7). Let $\hat{p}_k$ be given in (2.14) and $p_{k,P}$ be the vector such that*

$$p_{k,P} = P(x_k + p_k) - x_k. \tag{3.5}$$

Then, there are positive constants Γ_2 *and* $\nu > 0$ *such that for* $x_k \in B_\nu(x^*)$, $\eta_k \leq \Lambda \|W_k D_k g_k\|^{s-1}$ *for some* $\Lambda > 0$, *we have*

$$\psi_k(\hat{p}_k) \leq \max\{\sigma, 1 - \|p_{k,P}\|\}\, \psi_k(p_{k,P}), \tag{3.6}$$

$$\psi_k(p_{k,P}) \leq \psi_k(p_k) + \Gamma_2 \|x_k - x^*\|^s (\|x_k - x^*\|^{1/s} + \|p_{k,P}\|). \tag{3.7}$$

Proof. To prove (3.6), note that (2.14) and $\max\{\sigma, 1 - \|p_{k,P}\|\} < 1$ yield

$$\begin{aligned}
\psi_k(\hat{p}_k) &= \max\{\sigma, 1 - \|p_{k,P}\|\} \left(\frac{1}{2} \max\{\sigma, 1 - \|p_{k,P}\|\} p_{k,P}^T M_k p_{k,P} + p_{k,P}^T g_k \right) \\
&\leq \max\{\sigma, 1 - \|p_{k,P}\|\} \left(\frac{1}{2} p_{k,P}^T M_k p_{k,P} + p_{k,P}^T g_k \right).
\end{aligned}$$

Regarding (3.7), if $p_{k,P} = p_k$ for all k sufficiently large, the thesis trivially follows. This is the case when when $x^* > 0$. In this case, there exists an open ball $B_\mu(x^*)$ within the positive orthant. Let x_m such that $x_m \in B_{\mu/2}(x^*)$ and all the sequence $\{x_k\}_{k \geq m}$ belongs to $B_{\mu/2}(x^*)$. Assume $k \geq m$. Then $(x_k)_i > \mu/2$ for all i. By (3.4), $\lim_{k\to\infty} x_k = x^*$ implies $\lim_{k\to\infty} \|W_k D_k g_k\| = 0$. This along with the uniform boundness of $(W_k D_k M_k)^{-1}$ and (2.12) gives $\lim_{k\to\infty} \|p_k\| = 0$. Thus, for k sufficiently large, $\|p_k\| \leq \mu/2$ and $|(x_k + p_k)_i| \leq \mu$, i.e. no projection is needed.

Therefore, we restrict to the case where x^* belongs to the boundary of the feasible set i.e. $(x^*)_i = 0$ for some i. From the above argument, we claim that there exists $\nu > 0$ such that if $x_k \in B_\nu(x^*)$ then $(p_{k,P})_i = (p_k)_i$ for any i such that $(x^*)_i \neq 0$. Then, if $x_k \in B_\nu(x^*)$ and the set

$$Q_k = \{i \in \{1, \ldots, n\} : (p_k)_i \neq (p_{k,P})_i\},$$

is not empty, it follows

$$Q_k \subset \mathcal{A} \cup \mathcal{N}, \quad \text{i.e.} \quad (x^*)_i = 0, \ \text{ if } \ i \in Q_k. \tag{3.8}$$

Furthermore, let $\nu < \tau$ where τ is given in Lemma 3.2. Since M_k is positive definite we have

$$\begin{aligned}
\psi_k(p_k) &= ((p_k - p_{k,P}) + p_{k,P})^T g_k + \frac{1}{2}((p_k - p_{k,P}) + p_{k,P})^T M_k ((p_k - p_{k,P}) + p_{k,P}) \\
&\geq \psi_k(p_{k,P}) + (p_k - p_{k,P})^T (M_k p_{k,P} + g_k), \qquad (3.9)
\end{aligned}$$

and

$$\begin{aligned}
(p_k - p_{k,P})^T (M_k p_{k,P} + g_k) &= \sum_{i \in Q_k} (p_k - p_{k,P})_i \left(\frac{(e_k)_i\, (p_{k,P})_i}{(d_k)_i} + (g_k)_i \right) \\
&\quad + (p_k - p_{k,P})^T A^T A p_{k,P} \\
&\geq \sum_{i \in Q_k} (p_k - p_{k,P})_i \left(\frac{(e_k)_i\, (p_{k,P})_i}{(d_k)_i} + (g_k)_i \right) - |(p_k - p_{k,P})^T A^T A p_{k,P}|. \qquad (3.10)
\end{aligned}$$

Let

$$Q_{k,1} = \{i \in Q_k : (e_k)_i = (g_k)_i\}, \qquad Q_{k,2} = \{i \in Q_k : (e_k)_i = 0\}, \tag{3.11}$$

$$Q_{k,3} = \{i \in Q_{k,2} : (g_k)_i < 0\}. \tag{3.12}$$

Note that $Q_k = Q_{k,1} \cup Q_{k,2}$ and

$$(p_k - p_{k,P})_i < 0, \quad (x_k)_i = -(p_{k,P})_i, \quad \text{if} \quad i \in Q_k, \tag{3.13}$$
$$(g_k)_i \geq 0, \quad (d_k)_i = (x_k)_i, \quad \text{if} \quad i \in Q_{k,1}. \tag{3.14}$$

Therefore, we get

$$\sum_{i \in Q_{k,1}} (p_k - p_{k,P})_i (g_k)_i \left(1 + \frac{(p_{k,P})_i}{(d_k)_i}\right) = 0, \qquad \sum_{i \in Q_{k,3}} (p_k - p_{k,P})_i (g_k)_i \geq 0.$$

Then

$$\begin{aligned}
\sum_{i \in Q_k} (p_k - p_{k,P})_i \left((g_k)_i + \frac{(e_k)_i \, (p_{k,P})_i}{(d_k)_i}\right) &= \sum_{i \in Q_{k,1}} (p_k - p_{k,P})_i (g_k)_i \left(1 + \frac{(p_{k,P})_i}{(d_k)_i}\right) \\
&\quad + \sum_{i \in Q_{k,2}} (p_k - p_{k,P})_i (g_k)_i \\
&\geq \sum_{i \in Q_{k,2} \setminus Q_{k,3}} (p_k - p_{k,P})_i (g_k)_i. \qquad (3.15)
\end{aligned}$$

Hence, (3.9), (3.10) and (3.15) yield

$$\psi_k(p_{k,P}) \leq \psi_k(p_k) - \sum_{i \in Q_{k,2} \setminus Q_{k,3}} (p_k - p_{k,P})_i (g_k)_i + |(p_k - p_{k,P})^T A^T A p_{k,P}|. \tag{3.16}$$

From (2.7) it is easy to see that if $i \in Q_{k,2} \setminus Q_{k,3}$ we have $x_i^s \leq (g_k)_i \leq x_i^{1/s}$. Moreover from (3.13) and the definition of $Q_{k,3}$ we get $\sum_{i \in Q_{k,2} \setminus Q_{k,3}} (p_k - p_{k,P})_i (g_k)_i < 0$. Hence, by using (3.13), (3.8) and (3.1), we obtain

$$\begin{aligned}
-\sum_{i \in Q_{k,2} \setminus Q_{k,3}} (p_k - p_{k,P})_i (g_k)_i &\leq \sum_{i \in Q_{k,2} \setminus Q_{k,3}} |(p_k - p_{k,P})_i| \max_{i \in Q_{k,2} \setminus Q_{k,3}} (g_k)_i \\
&\leq \sum_{i \in Q_{k,2} \setminus Q_{k,3}} |(p_k + x_k)_i| \max_{i \in Q_{k,2} \setminus Q_{k,3}} (x_k - x^*)_i^{1/s} \\
&\leq \sum_{i \in Q_{k,2} \setminus Q_{k,3}} |(p_k + x_k - x^*)_i| \, \|x_k - x^*\|_\infty^{1/s} \\
&\leq \|p_k + x_k - x^*\|_1 \|x_k - x^*\|_\infty^{1/s} \\
&\leq \sqrt{n}\, \zeta \|x_k - x^*\|^{s+1/s}.
\end{aligned}$$

Analogously,

$$\begin{aligned}
|(p_k - p_{k,P})^T A^T A p_{k,P}| &\leq \|A^T A\| \, \|p_{k,P}\| \left(\sum_{i \in Q_k} (p_k + x_k - x^*)_i^2\right)^{\frac{1}{2}} \\
&\leq \zeta \, \|A^T A\| \, \|p_{k,P}\| \, \|x_k - x^*\|^s.
\end{aligned}$$

Hence, the thesis follows from (3.16) letting $\Gamma_2 = \zeta(\sqrt{n} + \|A^T A\|)$. $\square$

We conclude with the main result of the section.

THEOREM 3.1. *Let $s \in (1, 2]$ be the scalar in (2.7). Assume $\eta_k \le \Lambda \|W_k D_k g_k\|^{s-1}$ for some $\Lambda > 0$ and k sufficiently large. The sequence $\{x_k\}$ generated by the global inexact Newton-like method converges to x^* with q-order s.*

Proof. By Theorem 2.2, for any $x_0 \in \mathbb{R}^n$ we have $\lim_{k\to\infty} x_k = x^*$.

Let ζ and ν be given in Lemma 3.1 and Lemma 3.3, respectively. Let $\omega \le \nu$ such that $\zeta\, \omega^{s-1} < 1$. In the sequel we assume that $x_k \in B_\omega(x^*)$. The step $p_{k,P}$ in (3.5) is such that $\|p_{k,P}\| \le \|p_k\|$ and from (3.1) we get

$$\|p_{k,P}\| \le \|x_k + p_k - x^*\| + \|x_k - x^*\| \le \zeta \|x_k - x^*\|^s + \|x_k - x^*\|.$$

Consequently, (3.7) yields

$$\psi_k(p_{k,P}) \le \psi_k(p_k) + O(\|x_k - x^*\|^{s+1/s}). \tag{3.17}$$

Since p_k^N is the global minimum of ψ_k, it follows $\psi_k(p_k^N) \le \psi_k(p_k^C) < 0$ and $red(\hat{p}_k) \ge \psi_k(\hat{p}_k)/\psi_k(p_k^N)$. Then, by (3.6), (3.17) and (3.2) we get

$$\begin{aligned} red(\hat{p}_k) &\ge \max\{\sigma,\, 1 - \|p_{k,P}\|\} \frac{\psi_k(p_{k,P})}{\psi_k(p_k^N)} \\ &\ge \max\{\sigma,\, 1 - \|p_{k,P}\|\} \frac{\psi_k(p_k) + O(\|x_k - x^*\|^{s+1/s})}{\psi_k(p_k^N)} \\ &\ge \max\{\sigma,\, 1 - \|p_{k,P}\|\} \left(1 + \frac{\Gamma_1 \eta_k^2 \|x_k - x^*\|^2 + O(\|x_k - x^*\|^{s+1/s})}{\psi_k(p_k^N)}\right). \end{aligned}$$

In addition, $\psi_k(p_k^N) = -\frac{1}{2}(p_k^N)^T M_k p_k^N$ and

$$\frac{1}{2}(p_k^N)^T M_k p_k^N = \frac{1}{2}\|A p_k^N\|^2 + \frac{1}{2}(p_k^N)^T D_k^{-1} E_k p_k^N \ge \frac{\|p_k^N\|^2}{\|A^+\|^2}.$$

This relation and (3.1) imply

$$\begin{aligned} \frac{1}{2}(p_k^N)^T M_k p_k^N &\ge \frac{(\|x_k - x^*\| - \|x_k + p_k^N - x^*\|)^2}{\|A^+\|^2} \\ &\ge \frac{(1 - \zeta \|x_k - x^*\|^{s-1})^2 \|x_k - x^*\|^2}{\|A^+\|^2} \\ &\ge \frac{(1 - \zeta \omega^{s-1})^2 \|x_k - x^*\|^2}{\|A^+\|^2}. \end{aligned}$$

Then, we have shown that

$$red(\hat{p}_k) \ge \max\{\sigma,\, 1 - \|p_{k,P}\|\} \left(1 - \|A^+\|^2 \, \frac{\Gamma_1 \eta_k^2 + O(\|x_k - x^*\|^{s+1/s-2})}{(1 - \zeta \omega^{s-1})^2}\right).$$

Since $s + 1/s > 2$ for $s > 0$ and $\lim_{k\to\infty} \eta_k = 0$, it follows $\lim_{k\to\infty} red(\hat{p}_k) = 1$. Then (2.33) is met eventually and the step $\hat{p}_k$ is taken. Finally, Theorem 2.1 states that the local q-order of convergence is s. □

4. Computational experience. To show the computational feasibility of our method, we give the numerical results obtained on a large number of sparse NNLS problems.

Our algorithm was implemented in `Matlab` 7.0 and the numerical experiments were run on a `HP Workstation xw6200` with machine precision $\epsilon_m \simeq 2.\,10^{-16}$. For all tests we used the initial guess $x_0 = (1, \ldots, 1)^T$ and the algorithm's constants

$$\beta = 0.3, \qquad \theta = 0.9995, \qquad \sigma = 0.9995, \qquad s = 2.$$

A successful termination at the point x_k is declared whenever

$$\begin{cases} q_{k-1} - q_k < \tau(1 + q_{k-1}), \\ \|x_k - x_{k-1}\| \le \sqrt{\tau}(1 + \|x_k\|) \\ \|P(x_k + g_k) - x_k\| < \tau^{\frac{1}{3}}(1 + \|g_k\|) \end{cases} \qquad \text{or} \qquad \|D_k\, g_k\| \le \tau,$$

with $\tau = 10^{-9}$, see [12]. A failure is declared if the stopping criteria are not met within 300 iterations.

The numerical solution of the Newton equations by a direct method is carried out applying a minimum degree preordering algorithm and a sparse Cholesky decomposition with one step of fixed iterative refinement, (see [1] p. 264).

In the case we perform the inexact Newton method, we solve the linear systems (2.25) by the preconditioned CGLS method. The initial guess is the null vector and the preconditioners used are either the diagonal preconditioner or an incomplete Cholesky factorization with minimum degree preordering. Taking into account (2.26) and Theorem 3.1 we iterate CGLS until

$$\|\tilde{r}_k\| \le \max\{\, 500\epsilon_m,\ \min\{10^{-1}, \|W_k D_k g_k\|\}\, \|W_k D_k g_k\|\ \}. \tag{4.1}$$

The safeguards in (4.1) are employed so that the forcing term in (2.26) is less than 1 and not excessively small. Eventually, for $\|W_k D_k g_k\|^2 > 500\epsilon_m$, condition (2.26) is met with $\eta_k = \|W_k D_k g_k\|$. This way, theoretical quadratic convergence rate is ensured.

The experiments are carried out on five sets of problems.

The first set (`Set1`) is made up of the following problems from the Harwell Boeing collection [9]: `add20.rua, illc1033, illc1850, well1033, well1850`. These problems are well-conditioned or moderately ill-conditioned and may be degenerate.

Other three sets (`Set2, Set3, Set4`) were randomly generated by the test generator proposed by Portugal, Judice and Vicente in [24]. This generator constructs NNLS problems where the solution x^* to (1.1), the gradient $g(x^*)$ and the condition number of A are prescribed. We fixed x^* and $g(x^*)$ in the following way:

$$x^* = (\overbrace{\mathbf{c}}^{d_I}, \overbrace{\mathbf{0}}^{d_A}, \overbrace{\mathbf{0}}^{d_N})^T, \qquad g(x^*) = (\overbrace{\mathbf{0}}^{d_I}, \overbrace{\mathbf{1}}^{d_A}, \overbrace{\mathbf{0}}^{d_N})^T,$$

where d_I, d_A, d_N are the dimensions of the sets $\mathcal{I}$, $\mathcal{A}$, $\mathcal{N}$ given in (2.3)-(2.5), $\mathbf{c} = (1, 2, 3, \ldots, d_I)^T$, $\mathbf{0}$ and $\mathbf{1}$ are vectors with components equal to zero and one, respectively, and dimension dependent on the context. The matrix A is such that $m = 5000$, $n = 2000$ and has density $dens = \frac{nnz(A)}{m\,n} = 5.\,10^{-3}$ where $nnz(A)$ is the number of nonzero elements of A.

`Set2-Set4` have been formed to investigate the effect of degeneracy and conditioning of A on the behaviour of our approach. Therefore, we considered NNLS problems

Test name	$\|x_0 - x^*\|$	m	n	$dens$	$k_2(A)$	it
add20.rua	$1.7\,10^4$	2395	2395	$2.3\,10^{-3}$	$4.9,\,10^1$	13
illc1033	$5.8\,10^3$	1033	320	$1.4\,10^{-2}$	$1.9\,10^4$	35
illc1850	$6.1\,10^3$	1850	712	$6.7\,10^{-3}$	$1.4\,10^4$	16
well1033	$5.8\,10^3$	1033	320	$1.4\,10^{-2}$	$1.7\,10^2$	14
well1850	$5.2\,10^3$	1850	712	$6.7\,10^{-3}$	$1.1\,10^2$	16

TABLE 4.1
Newton-like method: Set1 *from Harwell Boeing collection*

with solution x^* belonging to the following three cases:

$$\begin{aligned}
\text{Highly degenerate case :}\quad & [d_I,\, d_A,\, d_N]=\, [\frac{1}{2}\,n,\, \frac{9}{20}\,n,\, \frac{1}{20}\,n];\\
\text{Mildly degenerate case :}\quad & [d_I,\, d_A,\, d_N]=\, [\frac{1}{4}\,n,\, \frac{3}{4}\,n - 10,\, 10];\\
\text{Non degenerate case :}\quad & [d_I,\, d_A,\, d_N]=\, [\frac{3}{4}n,\, \frac{1}{4}\,n,\, 0],
\end{aligned}$$

and let the matrices A be randomly generated with prescribed condition condition number of the order $O(10^\gamma)$, $\gamma = 1,\, 3,\, 5$.

We let Set2, Set3, Set4 be made up of tests with x^* highly degenerate, mildly degenerate and non degenerate, respectively. In particular, each set contains 10 tests with $k_2(A) = O(10^\gamma)$, for each value of $\gamma = 1, 2, 3$ and has a total amount of 30 tests.

Finally, the last set of problems (Set5) is made of 30 larger tests, where the features of x^* are not prescribed. The matrices A have dimension 5000×5000 and density $dens = 5.\,10^{-3}$. They are randomly generated for varying values of $k_2(A)$. As above, for each value of γ, $\gamma = 1,\, 3,\, 5$, we constructed 10 tests where the conditioning of A is of order $O(10^\gamma)$.

The Newton-like method was applied on tests from Set1- Set4 and the results are presented in Tables 4.1- 4.2. More precisely, in Table 4.1 we summarize the results obtained on Set1. For each test we report the distance of x_0 from x^*, the dimension of A, the density of A, the condition number of A and the number it of nonlinear iterations performed. The results obtained solving the problems from Set2-Set4 are displayed in Table 4.2. For each set and for each order of $k_2(A)$ considered, we report the average number Avg and the maximum number Max of nonlinear iterations performed over 10 tests.

The inexact Newton-like method was applied to solve the problems in Set2-Set5 and the corresponding results are summarized in Table 4.3. In this table, the meaning of Avg and Max is the same as in Table 4.2 while Itl indicates the average number of iterations performed by the iterative linear solver over 10 tests.

We remark that, even if we started from a very crude approximation to the solution, the considered NNLS problems have been solved in a reasonable number of iterations.

The overall experiments indicate that the number of nonlinear iterations is not sensitive to problem's dimension. Moreover, from Tables 4.2 and 4.3 it is clear that the

		$k_2(A) = O(10^1)$		$k_2(A) = O(10^3)$		$k_2(A) = O(10^5)$	
Set	$\|x_0 - x^*\|$	Avg	Max	Avg	Max	Avg	Max
Set2	$1.8\,10^4$	25	42	48	90	62	99
Set3	$6.4\,10^3$	21	39	49	85	58	125
Set4	$3.3\,10^4$	23	39	43	81	57	84

TABLE 4.2
Newton-like method: `Set2-Set4`, *highly degenerate, mildly degenerate and nondegenerate problems,* $m = 5000$, $n = 2000$

	$k_2(A) = O(10^1)$			$k_2(A) = O(10^3)$			$k_2(A) = O(10^5)$		
Set	Avg	Max	Itl	Avg	Max	Itl	Avg	Max	Itl
Set2	30	54	23	52	82	32	68	119	36
Set3	24	28	22	47	86	24	64	126	55
Set4	26	38	25	55	83	38	69	140	55
Set5	12	12	25	21	28	23	30	46	32

TABLE 4.3
Inexact Newton-like method: `Set2-Set5`

number of nonlinear iterations is insensitive to both the number of active constraints at the solution x^* and the degeneracy of x^*. On the contrary, the number of nonlinear iterations increases when the conditioning deteriorates.

The dependence of the nonlinear iteration counts on the conditioning of A can be ascribed to the globalization strategy. In fact, the examination of the results reveals fast local convergence for all runs while the value $\|D_k g_k\|$ oscillates for ill-conditioned problems and it can exhibit a large growth at some iterations. An analogous behaviour of the gradient norm in the steepest descent method for optimization problems was analized in [22]. In the globalization strategy we select a point on the segment connecting the projected inexact step $\hat{p}_k$ and the Cauchy step p_k^C. Since the direction of p_k^C is $D_k g_k$, the oscillating behaviour of $\|D_k g_k\|$ may slower the iterative process so that several iterations are needed to reach the vicinity of x^*. However, despite the growth of the computational work for increasing values of the condition number, the values of `Avg` and `Max` make evident that few runs are very expensive. The good results in the solution of `Set5` for all values of $k_2(A)$ are in accordance with these observations as the distance $\|x_0 - x^*\|$ is of the order of 0(10) i.e. considerably smaller than for the other sets.

Finally, it is evident that the inexact solution of the linear systems does not deteriorate the overall performance of the Newton-like method. In fact, on problems in `Set2-Set4` the Inexact approach requires an average number of nonlinear iterations comparable to that performed by the Newton-like method.

REFERENCES

[1] Björck A, *Numerical methods for least squares problems*, SIAM, 1996.

[2] M. Bierlaire, Ph. L. Toint, D. Tuyttens, *On iterative algorithms for linear least-squares problems with bound constraints*, Linear Algebra Appl., 143 (1991), pp. 111-143.

[3] S.S. Chen, D.L. Donoho, M.A. Saunders *Atomic decomposition by basis pursuit*, SIAM Review 43 (2001), pp. 129-159.

[4] T.F. Coleman, L.A. Hulbert, *A globally and superlinearly convergent algorithm for convex quadratic programs with simple bounds*, SIAM, J. Optim., 3 (1993), pp. 298-321.

[5] T.F. Coleman, Y. Li, *An interior trust-region approach for nonlinear minimization subject to bounds*, SIAM J. Optim., 6 (1996), pp. 418-445.

[6] T.F. Coleman, Y. Li, *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.

[7] T.F. Coleman, J. Liu, *An exterior Newton method for strictly convex quadratic programming*, Comput. Optim. Appl. 15 (2000), pp. 5-32.

[8] R.S. Dembo, S.C. Eisenstat, T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), pp. 400-408.

[9] I. Duff, R. Grimes, J. Lewis, *Sparse Matrix Test Problems*, ACM Trans. Math. Softw., 15 (1989), pp. 1-14.

[10] R. Fletcher, *Conjugate gradient methods for indefinite systems*, Numerical Analysis Dundee 1975, G. Watson ed.,Springer-Verlag, 1976, pp. 73-89.

[11] R. Freund, N. Nachtigal, *QMR: A Quasi-Minimal Residual Method for Non-Hermitian Linear Systems*, Numer. Math., 60 (1991), pp. 315–339.

[12] P.E. Gill, W. Murray, M.H. Wright, *Practical Optimization*, Academic Press, 198l.

[13] M. Hanke, J.G. Nagy, C. Vogel, *Quasi-Newton approach to nonnegative image restoration*, Linear Algebra Appl., 316 (2000), pp. 223-236.

[14] M.Heinkenschloss, M. Ulbrich, S. Ulbrich, *Superlinear and quadratic convergence of affine-scaling interior-point Newton methods for problems with simple bounds without strict complementarity assumptions*, Math. Program., 86 (1999), pp. 615-635.

[15] C. Kanzow, A. Klug, *On affine-scaling interior-point Newton methods for nonlinear minimization with bound constraints*, Computational Optimization and Applications, to appear.

[16] C. Lanczos, *Solution of linear equations by minimized iterations*, J. Res. Natl. Bur. Stand., 49 (1952), pp. 33-53.

[17] C.L. Lawson, R.J. Hanson, *Solving least-squares problems*, Prentice Hall, 1974.

[18] C.J. Lin, J.J. Moré, *Newton's method for large bound-constrained optimization problems*, SIAM J. Optim. 9 (1999), pp. 1100-1127.

[19] P. Lotstedt, *Perturbation bounds for the linear least-squares problem subject to linear inequality constraints*, BIT, 23 (1983), pp. 500-519.

[20] J.J. Moré, G. Toraldo, *On the solution of large quadratic programming problems with bound constraints*, SIAM J. Opim., 1 (1991), pp. 93-113.

[21] S. Morigi, L. Reichel, F. Sgallari, F. Zama *An iterative method for linear discrete ill-posed problems with box constraints*, J. Comput. Appl. Math., in press.

[22] J. Nocedal, A. Sartenear, C. Zhu, *On the behaviour of the gradient norm in the steepest descent method*, Comput. Optim. Appl. 22 (2002), pp. 5-35.

[23] C.C. Paige, M.A. Saunders *LSQR: an algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Softw., 8 (1982), pp. 43-71.

[24] L.F. Portugal, J.J. Judice, L.N. Vicente, *A comparison of block pivoting and interior-point algorithms for linear least squares problems with nonnegative variables*, Math. Comput., 63 (1994), pp. 625-643.

[25] Y. Saad, M.H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856-869.

[26] H.A. van der Vorst, *Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonlinear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631-644.