A Nonmonotone Inexact Newton Method *

Silvia Bonettini
Dipartimento di Matematica, Università di Modena e Reggio Emilia

Abstract

In this paper we describe a variant of the Inexact Newton method for solving nonlinear systems of equations. We define a nonmonotone Inexact Newton step and a nonmonotone backtracking strategy. For this nonmonotone Inexact Newton scheme we present the convergence theorems. Finally, we show how we can apply these strategies to Inexact Newton Interior–Point method and we present some numerical examples.
Keywords: Nonlinear Systems, Inexact Newton Methods, Nonmonotone Convergence, Newton Interior–Point Methods.

1 Introduction

A classical way to solve a system of nonlinear equations

$$F(x) = 0 \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable, is the Newton method: given a starting point x_0, at each step k the Newton equation

$$F'(x_k)s_k = -F(x_k) \tag{2}$$

has to be solved, in order to determine the Newton direction s_k. Then, the new iterate is computed by the rule

$$x_{k+1} = x_k + s_k.$$

Convergence theorems for Newton method can be found for example in [11]. The main computational task is the solution of (2), which can be very

*This research was supported by the Italian Ministry for Education, University and Research (MIUR), FIRB Project RBAU01JYPN.

expensive if n is large. The idea of Inexact Newton method introduced in [2] is to substitute (2) with a condition on its residual:

$$\|F(x_k) + F'(x_k)s_k\| \le \eta_k \|F(x_k)\|,$$

where $\eta_k \in [0,1)$ and $\|\cdot\|$ is an n–dimensional vector norm. In order to obtain global convergence properties, the Global Inexact Newton method presented in [5] also requires another condition that guarantees a "sufficient decrease" of the norm of F at each iterate. A general scheme for this method can be written as follows:

Let $x_0 \in \mathbb{R}^n$ and $\beta \in (0,1)$ be given.
For $k = 0, 1, 2, \ldots$
Find some $\eta_k \in [0,1)$ and a vector s_k that satisfy

$$\|F(x_k) + F'(x_k)s_k\| \le \eta_k \|F(x_k)\| \tag{3}$$

and

$$\|F(x_k + s_k)\| \le (1 - \beta(1 - \eta_k))\|F(x_k)\|. \tag{4}$$

Set $x_{k+1} = x_k + s_k$.
A vector that satisfies (3) is called *Inexact Newton step* at the level η_k and the parameter η_k is the *forcing term*. In [5], global convergence theorems have been established for some particular algorithms following this scheme, under the assumption that the sequence of the iterates $\{x_k\}$ has a limit point where the jacobian matrix F' is nonsingular.
We observe that condition (3) is a generalization of the Newton equation; hence, in order to satisfy (3), it may be sufficient to solve (2) inexactly, for example by means of an iterative solver. Furthermore, the accuracy of the solution is given by the norm of F at the current iterate, so, when we are far from the solution, unnecessary computations can be avoided. This is an advantage of Inexact Newton methods, especially for large scale problems.
Note that, when $\|\cdot\|$ is the euclidean norm $\|\cdot\|_2$, condition (3) guarantees that the Inexact Newton step is a descent direction for the scalar function

$$\Phi(x) = \frac{1}{2}\|F(x)\|_2^2. \tag{5}$$

Indeed we have the following inequality (we omit the iteration index):

$$\begin{aligned}
\nabla\Phi(x)^t s &= F(x)^t F'(x)s \\
&= F(x)^t[-F(x) + F'(x)s + F(x)] \\
&= -\|F(x)\|_2^2 + F(x)^t(F'(x)s + F(x)) \\
&\le -(1-\eta)\|F(x)\|_2^2 \le 0.
\end{aligned}$$

Condition (4) provides that at every iteration the norm of F is reduced, so the sequence $\{\|F(x_k)\|\}$ is monotone nonincreasing. Then, we conclude that Inexact Newton method with the euclidean norm can be considered as a descent method with line search (4) for the merit function $\Phi(x)$. However, it can be observed that if x_* is a root of $F(x)$, it is also a minimizer of the norm of F, but the converse is not true.

In this paper we present a nonmonotone version of Inexact Newton method, where both the conditions (3) and (4) have been relaxed. First of all, it is useful to introduce the following notations. Given $N \in \mathbb{N}$ and a sequence $\{x_k\}$, we denote by $x_{\ell(k)}$ the element with the following property

$$\|F(x_{\ell(k)})\| = \max_{0 \le j \le \min(N,k)} \|F(x_{k-j})\|. \tag{6}$$

Note that we have $k - min(N,k) \le \ell(k) \le k$. The modified scheme can be written as follows:

Let $x_0 \in \mathbb{R}^n$ and $\beta \in (0,1)$ be given.

For $k = 0, 1, 2, \ldots$

Find some $\eta_k \in [0,1)$ and a vector s_k that satisfy

$$\|F(x_k) + F'(x_k)s_k\| \le \eta_k \|F(x_{\ell(k)})\| \tag{7}$$

and

$$\|F(x_k + s_k)\| \le (1 - \beta(1-\eta_k))\|F(x_{\ell(k)})\|. \tag{8}$$

Set $x_{k+1} = x_k + s_k$.

According to (3), we define the vector s_k satisfying (7) *nonmonotone Inexact Newton step* at the level η_k. Note that the sequence $\{\|F(x_k)\|\}$ satisfying (7) and (8) is nonmonotone, but $\{\|F(x_{\ell(k)})\|\}$ is a monotone nonincreasing subsequence of it. Furthermore, the nonmonotone step is not a descent direction for the merit function defined in (5). This fact may be useful in some cases to avoid local minima of the merit function where F' is singular.

In the next section, we present a backtracking algorithm following the nonmonotone scheme, for which, in section 3, we state convergence theorems. In section 4, as a special case, we consider the Newton Inexact interior–point method and we show that, by applying nonmonotone strategies, we can choose the perturbation parameter of the interior methods in a larger range of values.

Finally, in section 5, we present some numerical experiments related to nonmonotone interior–point method.

For the remainder of the paper, we denote $N_\delta(x) = \{y \in \mathbb{R}^n : \|y - x\| < \delta\}$ for $\delta > 0$ and we use the following results:

Lemma 1.1 [11, 2.3.3] Assume that $F'(x)$ is invertible. Then, for any $\epsilon > 0$ there exists $\delta > 0$ such that $F'(x)$ is invertible and

$$\|F'(x)^{-1} - F'(y)^{-1}\| < \epsilon,$$

for all $y \in N_\delta(x)$.

Lemma 1.2 [11, 3.1.5] For any x and $\epsilon > 0$, there exists $\delta > 0$ such that

$$\|F(z) - F(y) - F'(y)(z - y)\| \leq \epsilon\|z - y\|,$$

for all $z, y \in N_\delta(x)$.

2 A nonmonotone Inexact Newton method

The nonmonotone Inexact Newton method can be implemented by using a backtracking strategy. At each step k, we determine a forcing term $\bar{\eta}_k$ and a vector $\bar{s}_k$ that satisfy the nonmonotone condition (7); then, we reduce $\bar{s}_k$ by means of a damping parameter α_k obtained by a nonomonotone backtracking rule; the nonmonotone Inexact Newton step $s_k = \alpha_k \bar{s}_k$ satisfies condition (7) and (8) with $\eta_k = (1 - \alpha_k(1 - \bar{\eta}_k))$. The algorithm can be stated as follows.

Algorithm 2.1

Step 1. Set $x_0 \in \mathbb{R}^n$, $\beta \in (0, 1)$, $0 < \theta_{min} < \theta_{max} < 1$, $\eta_{max} \in (0, 1)$, $k = 0$.

Step 2. Determine $\bar{\eta}_k \in [0, \eta_{max}]$, $\bar{s}_k$ that satisfy

$$\|F(x_k) + F'(x_k)\bar{s}_k\| \leq \bar{\eta}_k \|F(x_{\ell(k)})\|.$$

Set $\alpha_k = 1$.

Step 3. While $\|F(x_k + \alpha_k \bar{s}_k)\| > (1 - \alpha_k \beta(1 - \bar{\eta}_k))\|F(x_{\ell(k)})\|$

Step 3a. Choose $\theta \in [\theta_{min}, \theta_{max}]$;

Step 3b. Set $\alpha_k = \theta \alpha_k$.

Step 4. Set $x_{k+1} = x_k + \alpha_k \bar{s}_k$.
$k = k + 1$
Go to Step 2.

The following lemma shows that if $\bar{\eta}_k \in [0, \eta_{max}]$ and $\bar{s}_k$ satisfy the condition at the step 2, then the vector $\alpha\bar{s}_k$ is a nonmonotone Inexact Newton step at the level $\eta(\alpha) = (1 - \alpha(1 - \bar{\eta}_k))$ for any $\alpha \in (0, \alpha_{max}]$, where $\alpha_{max} < 1$. Furthermore, in $(0, \alpha_{max}]$ the condition

$$\|F(x_k + \alpha\bar{s}_k)\| < (1 - \alpha\beta(1 - \bar{\eta}_k))\|F(x_{\ell(k)})\|$$

is verified.

Lemma 2.1 Let $\beta \in (0, 1)$; suppose that there exist $\bar{\eta} \in [0, 1)$, $\bar{s}$ satisfying

$$\|F(x_k) + F'(x_k)\bar{s}\| \leq \bar{\eta}\|F(x_{\ell(k)})\|.$$

Then, there exist $\alpha_{max} \in (0, 1]$ and a vector s such that

$$\|F(x_k) + F'(x_k)s\| \leq \eta\|F(x_{\ell(k)})\| \tag{9}$$

$$\|F(x_k + s)\| \leq (1 - \beta\alpha(1 - \eta))\|F(x_{\ell(k)})\| \tag{10}$$

hold for any $\alpha \in (0, \alpha_{max}]$, where $\eta \in [\bar{\eta}, 1)$, $\eta = (1 - \alpha(1 - \bar{\eta}))$.

Proof. Let $s = \alpha\bar{s}$. Then we have

$$\begin{aligned} \|F(x_k) + F'(x_k)s\| &= \|F(x_k) - \alpha F(x_k) + \alpha F(x_k) + \alpha F'(x_k)\bar{s}\| \\ &\leq (1 - \alpha)\|F(x_k)\| + \alpha\|F(x_k) + F'(x_k)\bar{s}\| \\ &\leq (1 - \alpha)\|F(x_{\ell(k)})\| + \alpha\bar{\eta}\|F(x_{\ell(k)})\| \\ &= \eta\|F(x_{\ell(k)})\|, \end{aligned}$$

so (9) is proved. Now let

$$\varepsilon = \frac{(1 - \beta)(1 - \bar{\eta})}{\|\bar{s}\|}\|F(x_{\ell(k)})\|, \tag{11}$$

and $\delta > 0$ be sufficiently small (see Lemma 1.2) that

$$\|F(x_k + s) - F(x_k) - F'(x_k)s\| \leq \varepsilon\|s\| \tag{12}$$

whenever $\|s\| < \delta$. Choosing $\alpha_{max} = \min(1, \frac{\delta}{\|\bar{s}\|})$, for any $\alpha \in (0, \alpha_{max}]$ we have $\|s\| < \delta$ and then, using (11) and (12), we obtain the following inequality

$$\begin{aligned} \|F(x_k + s)\| &\leq \|F(x_k + s) - F(x_k) - F'(x_k)s\| + \|F(x_k) + F'(x_k)s\| \\ &\leq \varepsilon\alpha\|\bar{s}\| + \eta\|F(x_{\ell(k)})\| \\ &= ((1 - \beta)(1 - \bar{\eta})\alpha + (1 - \alpha(1 - \bar{\eta})))\|F(x_{\ell(k)}\| \\ &= (1 - \beta\alpha(1 - \bar{\eta}))\|F(x_{\ell(k)})\| \\ &\leq (1 - \beta\alpha(1 - \eta))\|F(x_{\ell(k)})\|, \end{aligned}$$

that completes the proof. $\square$

A consequence of the previous lemma is that the *while loop* at the step 3 terminates. Indeed, at each iterate k the backtracking condition

$$\|F(x_k + \alpha\bar{s}_k)\| \leq (1 - \alpha\beta(1 - \bar{\eta}))\|F(x_{\ell(k)})\| \tag{13}$$

is satisfied for $\alpha < \alpha_{max}$, where α_{max} depends on k. Since the value of α_k is reduced by a factor $\theta < \theta_{max} < 1$ at the step 3a, then there exists a positive integer p such that $(\theta_{max})^p < \alpha_{max}$ and so *the while loop* terminates at most after p steps. When it is impossible to determine x_{k+1} we say that the algorithm *breaks down*. Then, Lemma 2.1 yields that algorithm 2.1 breaks down if and only if is impossible to find a nonmonotone inexact Newton step at any level.

Theorem 2.1 Let $\{x_k\}$ a sequence such that $\lim_{k\to\infty} F(x_k) = 0$ and for each k the following conditions hold:

$$\|F(x_k) + F'(x_k)s_k\| \leq \eta\|F(x_{\ell(k)})\|, \tag{14}$$

$$\|F(x_{k+1})\| \leq \|F(x_{\ell(k)})\|, \tag{15}$$

where $s_k = x_{k+1} - x_k$ and $\eta < 1$. If x_* is a limit point of $\{x_k\}$, then $F(x_*) = 0$ and if $F'(x_*)$ is nonsingular, then the sequence $\{x_k\}$ converges to x_*.

Proof. If x_* is a limit point of the sequence $\{x_k\}$, there exists a subsequence $\{x_{k_j}\}$ of $\{x_k\}$ convergent to x_*. By the continuity of F, we obtain

$$F(x_*) = F\left(\lim_{j\to\infty} x_{k_j}\right) = \lim_{j\to\infty} F(x_{k_j}) = 0.$$

Furthermore, since $\{x_{\ell(k)}\}$ is a subsequence of $\{x_k\}$, also the sequence $\{F(x_{\ell(k)})\}$ converges to zero when k diverges. Denote $K = \|F'(x_*)^{-1}\|$ and $\delta > 0$ be sufficiently small that $F'(y)^{-1}$ exists whenever $y \in N_\delta(x_*)$; thus we can suppose

$$\|F'(y)^{-1}\| \leq 2K, \tag{16}$$

$$\|F(y) - F(x_*) - F'(x_*)(y - x_*)\| \leq \frac{1}{2K}\|y - x_*\|.$$

Then for any $y \in N_\delta(x_*)$ we have

$$\begin{array}{rcl} \|F(y)\| & = & \|F'(x_*)(y-x_*) + F(y) - F(x_*) - F'(x_*)(y-x_*)\| \\ & \geq & \|F'(x_*)(y-x_*)\| - \|F(y) - F(x_*) - F'(x_*)(y-x_*)\| \\ & \geq & \frac{1}{K}\|y-x_*\| - \frac{1}{2K}\|y-x_*\| \\ & = & \frac{1}{2K}\|y-x_*\|. \end{array}$$

Then

$$\|y-x_*\| \leq 2K\|F(y)\| \tag{17}$$

holds for any $y \in N_\delta(x_*)$. Now let $\epsilon \in (0, \frac{\delta}{4})$ and since x_* is a limit point of $\{x_k\}$, there exists a k sufficiently large that

$$x_k \in N_{\frac{\delta}{2}}(x_*)$$

and

$$x_{\ell(k)} \in S_\epsilon \equiv \left\{ y : \|F(y)\| < \frac{\epsilon}{K(1+\eta)} \right\}.$$

Note that since $x_{\ell(k)} \in S_\epsilon$ then also $x_{k+1} \in S_\epsilon$ because $\|F(x_{k+1})\| \leq \|F(x_{\ell(k)})\|$. For the direction s_k, by (14), (15) and (16), the following inequality holds:

$$\begin{array}{rcl} \|s_k\| & \leq & \|F'(x_k)^{-1}\|(\|F(x_k)\| + \|F(x_k) + F'(x_k)s_k\|) \\ & \leq & 2K(\|F(x_{\ell(k)})\| + \eta\|F(x_{\ell(k)})\|) \\ & = & 2K(1+\eta)\|F(x_{\ell(k)})\| < 2\epsilon < \frac{\delta}{2}. \end{array}$$

Since $s_k = x_{k+1} - x_k$, the previous inequality implies $\|x_{k+1} - x_*\| < \delta$ and from (17) we obtain

$$\|x_{k+1} - x_*\| \leq 2K\|F(x_{k+1})\| < 2K\frac{\epsilon}{K(1+\eta)} < \frac{\delta}{2}$$

that implies $x_{k+1} \in N_{\frac{\delta}{2}}(x_*)$. Therefore $x_{\ell(k+1)} \in S_\epsilon$, since $\|F(x_{\ell(k+1)})\| \leq \|F(x_{\ell(k)})\|$. It follows that, for any j sufficiently large, $x_j \in N_\delta(x_*)$, and from (17)

$$\|x_j - x_*\| \leq 2K\|F(x_j)\|.$$

Since $F(x_j)$ converges to 0 we can conclude that x_j converges to x_*. □

Lemma 2.2 Suppose that Algorithm 2.1 does not break down. If x_* is a limit point of $\{x_k\}$ such that $F'(x_*)$ is nonsingular then there exist infinitely many k such that $\alpha_k > \tau > 0$.

Proof. Denoting $\|F'(x_*)^{-1}\| = K$, we can find $\delta > 0$ such that

(i) $F'(x)^{-1}$ exists whenever $x \in N_\delta(x_*)$,

(ii) $\|F'(x)^{-1}\| \leq 2K \qquad \forall x \in N_\delta(x_*)$

(iii) $\|F(x) - F(y) - F'(y)(x-y)\| \leq \frac{(1-\beta)(1-\eta_{max})}{2K(1+\eta_{max})}\|y-x\| \qquad \forall x, y \in N_{2\delta}(x_*)$.

Since x_* is a limit point, there exist infinitely many k such that $x_k \in N_\delta(x_*)$ for which the following condition holds:

$$\begin{array}{rcl} \|\bar{s}_k\| & \leq & \|F'(x_k)^{-1}\|(\|F'(x_k)\bar{s}_k + F(x_k)\| + \|F(x_k)\|) \\ & \leq & 2K(1+\eta_{max})\|F(x_{\ell(k)})\|. \end{array} \tag{18}$$

Since $s_k = \alpha \bar{s}_k$, formula (18) can be written as

$$\|s_k\| \leq \Gamma\alpha\|F(x_{\ell(k)})\| \tag{19}$$

where $\Gamma = 2K(1+\eta_{max})$.
Now we show that if $\alpha \leq \frac{\delta}{\Gamma\|F(x_{\ell(k)})\|}$, then the *while loop* terminates. We can write by means of condition (ii), Lemma 2.1 and formula (19)

$$\begin{array}{rcl} \|F(x_k+s_k)\| & \leq & \|F(x_k) + F'(x_k)s_k\| + \|F(x_k+s_k) - F(x_k) - F'(x_k)s_k\| \\ & \leq & \eta\|F(x_{\ell(k)})\| + \frac{(1-\beta)(1-\eta_{max})}{\Gamma}\|s_k\| \\ & \leq & ((1-\alpha)(1-\bar{\eta}) + (1-\beta)\alpha(1-\bar{\eta}))\|F(x_{\ell(k)})\|. \end{array}$$

Thus

$$\|F(x_k + \alpha\bar{s}_k)\| \leq (1 - \alpha\beta(1-\bar{\eta}))\|F(x_{\ell(k)})\|$$

This inequality shows that the backtracking condition (13) is satisfied for $\alpha \leq \frac{\delta}{\Gamma\|F(x_{\ell(k)})\|}$ and since α is reduced at every step by a factor $\theta \leq \theta_{max} < 1$ the *while loop* terminates. Suppose now that the *while loop* has been executed at least once, let denote α_k the final value (i.e. the value of α for which (13) is satisfied) and $\bar{\alpha_k}$ the previous one. At the penultimate step the condition (13) is not satisfied, so necessarily we have

$$\bar{\alpha_k} > \frac{\delta}{\Gamma\|F(x_{\ell(k)})\|}$$

and so

$$\alpha_k = \theta\bar{\alpha_k} > \frac{\delta\theta_{min}}{\Gamma\|F(x_{\ell(k)})\|} \geq \frac{\delta\theta_{min}}{\Gamma\|F(x_0)\|}.$$

Hence Lemma (2.2) has been proved with $\tau = min(1, \frac{\delta\theta_{min}}{\Gamma\|F(x_0)\|})$. $\square$

From Lemma 2.2 we can derive the following corollary, which is used in the proof of the convergence theorem.

Corollary 2.1 Suppose that Algorithm 2.1 does not break down. If x_* is a limit point of $\{x_k\}$ such that $F'(x_*)$ is nonsingular and $\{x_{k_j}\}$ is a subsequence converging to x_* then the sequence $\{\alpha_{k_j}\}$ is bounded away from zero.

Now we can state the convergence theorem. The proof is similar to the one of theorem in section 3 of [7].

Theorem 2.2 Suppose that Algorithm 2.1 does not break down and that the norm of inexact Newton step is bounded for every k by a positive constant M

$$\|\bar{s}_k\| \leq M. \tag{20}$$

Assume also that one of the two following properties holds:

$$F \text{ is Lipschitz continuous;} \tag{21}$$

$$\text{the set } \Omega(0) = \{x \in \mathbb{R}^n : \|F(x)\| \leq \|F(x_0)\|\} \text{ is compact.} \tag{22}$$

If x_* is a limit point of x_k such that $F'(x_*)$ is invertible then $F(x_*) = 0$ and $\{x_k\}$ converges to x_* when k diverges.

Proof. Since $\|F(x_{\ell(k)})\|$ is a monotone nonincreasing, bounded sequence, then there exists $L \geq 0$ such that

$$L = \lim_{k\to\infty} \|F(x_{\ell(k)})\|.$$

Thus, writing the backtracking condition (13) for the iterate $\ell(k)$, we obtain

$$\|F(x_{\ell(k)})\| \leq (1 - \alpha_{\ell(k)-1}\beta(1 - \bar{\eta}_{\ell(k)-1}))\|F(x_{\ell(\ell(k)-1)})\|. \tag{23}$$

When k diverges, we can write

$$L \leq L - L \cdot \lim_{k\to\infty} \alpha_{\ell(k)-1}\beta(1 - \bar{\eta}_{\ell(k)-1}). \tag{24}$$

Since β is a constant and $1 - \bar{\eta}_j \geq 1 - \eta_{max} > 0$ for any j, (24) yields

$$L \cdot \lim_{k\to\infty} \alpha_{\ell(k)-1} \leq 0$$

that implies

$$L = 0$$

or

$$\lim_{k\to\infty} \alpha_{\ell(k)-1} = 0. \tag{25}$$

Suppose that $L \neq 0$, so that (25) holds. Let $\hat{\ell}(k) = \ell(k+N+1)$ so that $k+N+1\hat{\ell}(k) > k$ and we show by induction that for any $j \geq 0$ we have

$$\lim_{k\to\infty} \alpha_{\hat{\ell}(k)-j} = 0 \tag{26}$$

and

$$\lim_{k\to\infty} \|F(x_{\hat{\ell}(k)-j})\| = L. \tag{27}$$

For $j = 1$, since $\{\alpha_{\hat{\ell}(k)-1}\}$ is a subsequence of $\{\alpha_{\ell(k)-1}\}$, (25) implies (26). From (20) we also obtain

$$\lim_{k\to\infty} \|x_{\hat{\ell}(k)} - x_{\hat{\ell}(k)-1}\| = 0. \tag{28}$$

If (21) holds, from $|\|F(x)\| - \|F(y)\|| \leq \|F(x) - F(y)\|$ and (28) we obtain

$$\lim_{k\to\infty} \|F(x_{\hat{\ell}(k)-1})\| = L. \tag{29}$$

If, instead of (21), (22) holds, then, exploiting the uniform continuity of F in $\Omega(0)$, we can again derive (29). Assume now that (26) and (27) hold for a given j. We have

$$\|F(x_{\ell(k)-j})\| \leq (1 - \alpha_{\ell(k)-(j+1)}\beta(1 - \eta_{\ell(k)-(j+1)}))\|F(x_{\ell(\ell(k)-(j+1))})\|.$$

Using the same arguments employed above, since $L > 0$, we obtain

$$\lim_{k\to\infty} \alpha_{\hat{\ell}(k)-(j+1)} = 0$$

and so

$$\lim_{k\to\infty} \|x_{\hat{\ell}(k)-j} - x_{\hat{\ell}(k)-(j+1)}\| = 0,$$
$$\lim_{k\to\infty} \|F(x_{\hat{\ell}(k)-(j+1)})\| = L.$$

Thus, we conclude that (26) and (27) hold for any $j \geq 1$. Now, for any k, we can write

$$\|x_{k+1} - x_{\hat{\ell}(k)}\| \leq \sum_{j=1}^{\hat{\ell}(k)-k-1} \alpha_{\hat{\ell}(k)-j}\|\bar{s}_{\hat{\ell}(k)-j}\|$$

so that, since we have $\hat{\ell}(k) - k - 1 \leq N$, we have

$$\lim_{k\to\infty} \|x_{k+1} - x_{\hat{\ell}(k)}\| = 0. \tag{30}$$

Furthermore, we have

$$\|x_{\hat{\ell}(k)} - x_*\| \leq \|x_{\hat{\ell}(k)} - x_{k+1}\| + \|x_{k+1} - x_*\| \tag{31}$$

Since x_* is a limit point of $\{x_{k+1}\}$ and (30) holds, (31) implies that x_* is a limit point for the sequence $\{x_{\hat{\ell}(k)}\}$. From (28) we conclude that x_* is a limit point also for the sequence $\{x_{\hat{\ell}(k)-1}\}$, which contradicts the Corollary 2.1. Indeed, there exists a $\tau > 0$ such that $\alpha_{\hat{\ell}(k)-1} > \tau$ for infinitely many k. Hence, we necessarily have $L = 0$, that implies

$$\lim_{k\to\infty} \|F(x_k)\| = 0.$$

Now Theorem 2.1 completes the proof. □

Theorem 2.3 Under the hypothesis of Theorem 2.2 we have that the sequence $\{\|F(x_k)\|\}$ converges and

$$\lim_{k\to\infty} \|F(x_k)\| = \lim_{k\to\infty} \|F(x_{\ell(k)})\|.$$

Proof. If $\lim_{k\to\infty} \|F(x_{\ell(k)})\| = 0$, then $\lim_{k\to\infty} \|F(x_k)\| = 0$.
If $\lim_{k\to\infty} \|F(x_{\ell(k)})\| = L > 0$, using the same arguments in the first part of the proof of Theorem 2.2, we can conclude that (30) holds. If (21) or (22) holds, then $\lim_{k\to\infty} \|F(x_k)\| = L = \lim_{k\to\infty} \|F(x_{\ell(k)})\|$. □

3 An application: a nonmonotone Inexact Newton Interior–Point Method

First, we recall the basic concepts of Newton Inexact interior–point method, as a special case of Inexact Newton method. For the details we refer to [4]. Here and for the remainder, we assume $\|\cdot\| = \|\cdot\|_2$.
Consider now the nonlinear programming problem

$$\begin{array}{rl} \min & f(x) \\ g_1(x) = 0 & \\ g_2(x) \geq 0 & \end{array} \tag{32}$$

where $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $g_1 : \mathbb{R}^n \to \mathbb{R}^{neq}$, $g_2 : \mathbb{R}^n \to \mathbb{R}^p$; by introducing the slack variables s on the inequality constraints, the Karush–Kuhn–Tucker (KKT) optimality conditions for problem (32) are given by the following system of nonlinear equations:

$$H(v) \equiv \begin{pmatrix} \nabla f(x) - \nabla g_1(x)\lambda - \nabla g_2(x)w \\ -g_1(x) \\ -g_2(x) + s \\ WSe_p \end{pmatrix} = 0, \tag{33}$$

with

$$s, w \geq 0,$$

where $\lambda \in \mathbb{R}^{neq}$ $s, w \in \mathbb{R}^p$ and $W = \text{diag}(w); S = \text{diag}(s)$. Here λ and w are the Lagrange multipliers related to the equality and inequality constraint respectively; the vector e_j indicates the vector of j components whose values are equal to 1. Furthermore we set $v = (x^t, \lambda^t, w^t, s^t)^t$ and $\tilde{n} \equiv n + neq + 2p$ (the size of the system (33)). The first $n + neq + p$ components of the vector $H(v)$,

$$G(v) = \begin{pmatrix} \nabla f(x) - \nabla g_1(x)\lambda - \nabla g_2(x)w \\ -g_1(x) \\ -g_2(x) + s \end{pmatrix}$$

represent the gradient of the lagrangian function of the minimum problem, while the last p equations in (33),

$$SWe_p = 0,$$

are called *complementarity conditions*.
In the framework of Newton interior–point method, instead of (33), we consider the perturbed KKT conditions

$$\begin{array}{c} H(v) = \rho\tilde{e} \\ s, w > 0, \end{array} \tag{34}$$

with $\rho > 0$ and $\tilde{e} = (0^t_{n+neq+p}, e^t_p)^t$ and, given a starting point v_0 with $(s_0, w_0) > 0$, at the iteration k we have to solve the perturbed Newton equation

$$H'(v_k)\Delta v = -H(v_k) + \rho_k\tilde{e}, \tag{35}$$

so that the iterates satisfy the positivity condition on (s_k, w_k). The perturbation parameter ρ_k can be defined as

$$\rho_k = \sigma_k\mu_k, \tag{36}$$

with $\sigma_k \in (0,1)$ and $\mu_k > 0$.
Now we will briefly recall the conditions that enable us to view the Newton interior–point method as an Inexact Newton method applied to the the system (33). Consider the Newton equation for (33):

$$H'(v_k)\Delta v_k = -H(v_k). \tag{37}$$

The residual vector $r_k \in \mathbb{R}^{\tilde{n}}$ for (37) can be written as

$$r_k = H'(v_k)\Delta v_k + H(v_k). \tag{38}$$

If we suppose that r_k is given by the following expression

$$r_k = \begin{pmatrix} 0_{n+neq+p} \\ \rho_k \sigma_k e_p \end{pmatrix}, \tag{39}$$

then we obtain

$$\|r_k\| = \rho_k \|e_p\| = \rho_k \sqrt{p}.$$

Note that if we choose

$$\mu_k \le \frac{\|H(v_k)\|}{\sqrt{p}}, \tag{40}$$

as in [3], and if Δv_k satisfies (38) where r_k is given by (39), then Δv_k is an inexact Newton step at the level σ_k for the system (33). In interior–point methods a suitable choice of μ_k is $\mu_k = \frac{s_k^t w_k}{p}$; we have $\frac{s_k^t w_k}{p} \le \frac{\|H(v_k)\|}{\sqrt{p}}$, so (40) is satisfied. Furthermore, a sufficient condition for (39) is that Δv_k is an exact solution for the perturbed equation (35), so (40) guarantees that the vector computed at every step of the interior–point method by solving (35) exactly is an Inexact Newton step.
Suppose now that the residual of (37) at the iteration k has the following expression, instead of (39):

$$r_k = \begin{pmatrix} \bar{r}_k \\ \rho_k e_p \end{pmatrix}, \tag{41}$$

where $\bar{r}_k \in \mathbb{R}^{n+neq+p}$ satisfies the condition

$$\|\bar{r}_k\| \le \delta_k \|H(v_k)\|. \tag{42}$$

Now, if (40) and (42) hold and $\sigma_k + \delta_k < 1$, then Δv_k in (38) is an Inexact Newton step at the level $\sigma_k + \delta_k$ for the system (33). Indeed, we have

$$\|r_k\|^2 = p\rho_k^2 + \delta_k^2 \|\bar{r}_k\|^2 \le (\sigma_k^2 + \delta_k^2)\|H(v_k)\|^2 \le (\sigma_k + \delta_k)^2 \|H(v_k)\|^2,$$

which implies
$$\|r_k\| \leq (\delta_k + \sigma_k)\|H(v_k)\|.$$
In order to obtain a residual vector as in (41), one may solve the equation (35) inexactly on the first $n + neq + p$ equations, by means of an iterative solver, using condition (42) as inner stopping criterion. So, the conditions on $\|\bar{r}_k\|$, δ_k, σ_k and μ_k allow us to calculate only an inexact solution of (35), obtaining again an Inexact Newton step. This approach is useful when $\tilde{n}$ is large and the computation of an exact solution can be too expensive.
If we replace (40) and (42) with

$$\mu_k \in \left[\frac{{s_k}^t w_k}{p}, \frac{\|H(v_{\ell(k)})\|}{\sqrt{p}}\right] \tag{43}$$

and

$$\|\bar{r}_k\| \leq \delta_k \|H(v_{\ell(k)})\| \tag{44}$$

then it is easy to verify, using the same observations employed above, that a vector Δv for which the residual r_k has the form in (41) is a nonmonotone Inexact Newton step. After the computation of the direction Δv_k, the following iterate in an interior–point method is determined by the updating rule
$$v_{k+1} = v_k + \alpha_k \Delta v_k,$$
where $\alpha_k \in (0, 1]$ has to be chosen in order to guarantee the positivity of the components of s_{k+1} and w_{k+1}. Furthermore the parameter $\alpha_k \in (0, 1]$ must be selected so that the *centrality conditions* (see e.g. [6]) are satisfied. Finally, we include in the method the nonmonotone backtracking strategy seen in the previous section. Now we present the nonmonotone Newton Inexact interior–point method.

Algorithm 3.1

Step 1. Fix v_0 such that $(s_0, w_0) > 0$ and choose the positive parameters as follows:

- $\tau_1 < 1$ and $\tau_1 \leq p\left(\min_{i=1,\dots,p}(s_0)_i(w_0)_i\right)/s_0^t w_0$;
- $\tau_2 \leq (s_0^t w_0)/\|G(v_0)\|$;
- $\tilde{\delta}, \beta, \theta, tol \in (0, 1)$;
- $\delta_{max} + \sigma_{max} < 1$ and $\sigma_{max} \geq \sqrt{2}\tau_2\delta_{max}/\min(1, \tau_2) + \tilde{\delta}$
- $\tilde{\delta} \leq \sigma_{min} < \sigma_{max}$.

Set $k \leftarrow 0$.

Step 2. If $\|H(v_k)\| \leq tol$ then stop, else choose the positive parameters $\sigma_k, \delta_k, \mu_k$ such that:

- $0 \leq \delta_k$;
- $\sigma_{min} \leq \sigma_k \leq \sigma_{max}$ and $\sigma_k \geq \tilde{\delta} + \sqrt{2}\delta_k / \min(1, \tau_2)$;
- $\mu_k \in \left[\frac{{s_k}^t w_k}{p}, \frac{\|H(v_{\ell(k)})\|}{\sqrt{p}}\right]$as in (43).

Step 3. Find $\Delta v_k = (\Delta x_k^t, \Delta\lambda_k^t, \Delta w_k^t, \Delta s_k^t)^t$ such that (35) hold with r_k defined in (41) and

$$\|\bar{r}_k\| \leq \delta_k \|H(v_{\ell(k)})\|$$

as in (44).

Step 4. Compute $\tilde{\alpha_k} = \min\left(\alpha_k^{(1)}, \alpha_k^{(2)}\right)$, where $\alpha_k^{(1)}$ and $\alpha_k^{(2)}$ are the largest numbers in $(0, 1]$ such that the following *centrality conditions* hold for any $\alpha \in (0, \alpha_k^{(1)}]$ and $\alpha \in (0, \alpha_k^{(2)}]$ respectively:

$$\min_{i=1,\ldots,p} s_k(\alpha)_i w_k(\alpha)_i \geq (\tau_1/p) s_k(\alpha)^t w_k(\alpha), \tag{45}$$

$$s_k(\alpha)^t w_k(\alpha) \geq \tau_2 \|G(v(\alpha))\|, \tag{46}$$

where $v(\alpha) = v_k + \alpha\Delta v_k$.

Step 5. If

$$\|H(v_k + \tilde{\alpha_k}\Delta v_k)\| \leq (1 - \tilde{\alpha}\beta(1 - (\delta_k + \sigma_k))\|H(v_{\ell(k)})\|, \tag{47}$$

go to Step 6, else update $\tilde{\alpha} = \theta\tilde{\alpha}$. and go to Step 5.
Denote α_k the last value of $\tilde{\alpha_k}$.

Step 6. Update $v_{k+1} = v_k + \alpha_k \Delta v_k$.
Set $k \leftarrow k + 1$.
Go to Step 2.

Step 2, 3, 5 and 6 enable us to consider Algorithm 3.1 as a special case of Algorithm 2.1. At the first step all the parameters are set, while at the step 4 the centrality conditions are stated. One can observe that conditions (45) and (46) avoid the last p components of the vector $H(v)$ (related to complementarity equations) to become smaller than $\|G(v)\|$ at every iterate. For the analysis of the convergence, it is useful to introduce the set

$$\begin{aligned} \Omega(\epsilon) = \quad & \{v \in \mathbb{R}^{\tilde{n}} : \epsilon \leq \|H(v)\| \leq \|H(v_0)\|, \\ & \text{s.t. } v \text{ satisfies conditions (45) and (46)}\}. \end{aligned} \tag{48}$$

We observe that all the iterates v_k belong to $\Omega(0)$. For the convergence of Algorithm 3.1 we make the following assumptions:

A1. In $\Omega(0)$ $f(x), g_1(x), g_2(x)$ are twice continuously differentiable and the derivative of $G(v)$ is Lipschitz continuous. Moreover the columns of $\nabla g_1(x)$ are linearly independent.

A2. $\Omega(0)$ is a compact set.

A3. The matrix $H'(v_k)$ is nonsingular for any $k \geq 0$.

The assumption A2 implies that the iteration sequence $\{v_k\}$ is bounded. First we prove some lemmas used in the proof of the convergence theorem presented below.

Lemma 3.1 Let $\{v_k\}$ generated by Algorithm 3.1. Under the assumption A1–A3 there exists a positive constant M such that

$$\|\Delta v_k\| \leq M.$$

Proof. Recalling that the direction Δv_k computed at the step 3 is a nonmonotone Inexact Newton step at the level $\sigma_k + \delta_k$, we obtain the following inequality:

$$\begin{aligned} \|\Delta v_k\| & \leq \|H'(v_k)^{-1}\| \cdot (\|H'(v_k)\Delta v_k + H(v_k)\| + \|H(v_k)\|) \\ & \leq \|H'(v_k)^{-1}\| \cdot (1 + \sigma_k + \delta_k)\|H(v_{\ell(k)})\|. \end{aligned} \tag{49}$$

Denoting $M = \left(\max_{v\in\Omega(0)} \|H(v)^{-1}\|\right)(1 + \sigma_{max} + \delta_{max})\|H(v_0)\|$, (49) yields

$$\|\Delta v_k\| \leq M.$$

□

For the proof of the following lemma we refer to [4]: tacking into account Lemma 3.1 and (43), it is possible to use the same arguments.

Lemma 3.2 Let $\{v_k\}$ generated by Algorithm 3.1, so that the settings at the step 1 and 2 hold. Assume that $\{v_k\} \subset \Omega(\epsilon)$with $\epsilon > 0$. Then $\tilde{\alpha}_k$ computed at the step 4 is bounded away from zero.

Now we prove the following convergence result.

Theorem 3.1 Under the assumptions A1–A3, the Algorithm 3.1 with $tol = 0$ generates a sequence $\{v_k\}$ such that $\{\|H(v_k)\|\}$ converges to zero and each limit point of $\{v_k\}$ satisfies the KKT conditions for (32). Furthermore, if v_* is a limit point of $\{v_k\}$ such that $H'(v_*)$ is nonsingular, then the sequence $\{v_k\}$ converges to v_*.

Proof. Denote $L = \lim_{k\to\infty} \|H(v_{\ell(k)})\|$. From Lemma 3.1 and Theorem 2.3 we obtain that $\lim_{k\to\infty} \|H(v_k)\| = L$. Suppose now that $L > 0$. This implies that $\{v_k\} \subset \Omega(\epsilon)$, with $\epsilon > 0$ (at least for k large). Consequently, from Lemma 3.2, $\tilde{\alpha_k}$ is bounded away from zero. If v_* is a limit point of $\{v_k\}$, then $\{v_*\} \in \Omega(\epsilon)$ and $H'(v_*)$ is a nonsingular matrix. Then, from Theorem 2.2, we deduce that $L = 0$ and this is a contradiction. Notice that we can use Theorem 2.2 even if the starting value of the backtracking procedure is $\tilde{\alpha_k}$ instead of 1 because $\tilde{\alpha_k}$ is bounded away from zero. Then $\{\|H(v_k)\|\}$ has to converge to zero. So, if v_* is a limit point of $\{v_k\}$ such that $H'(v_*)$ is nonsingular, using Theorem 2.1, then the sequence $\{v_k\}$ converges to v_*. □

4 Numerical examples

In this section we report some numerical experiments, obtained by coding Algorithm 3.1 in FORTRAN90 using double precision on a Compaq XP1000 workstation. In particular we set $\beta = 10^{-4}, \theta = 0.5, tol = 10^{-8}$. We declare the failure of the algorithm when the tolerance tol is not satisfied after 500 iterations or when at some step the backtracking reductions are more than 10. Furthermore we set $\mu_k = \frac{s_k^t w_k}{p}$ as in [6]. The aim of the numerical experiments is to compare the behaviour of the monotone and nonmonotone algorithms; for the nonmonotone one, the parameter N has been chosen equal to 2, 4 and 9. Furthermore, the comparison has been performed in two different cases. In the first one, δ_k is set equal to 0: this means that the perturbed Newton equation (35) is solved exactly at each iteration. The solution of the linear system is computed by the MA27 subroutine of the Harwell library that performs a LU factorization. In the second case, Hestenes multipliers scheme has been adopted as iterative inner solver (see

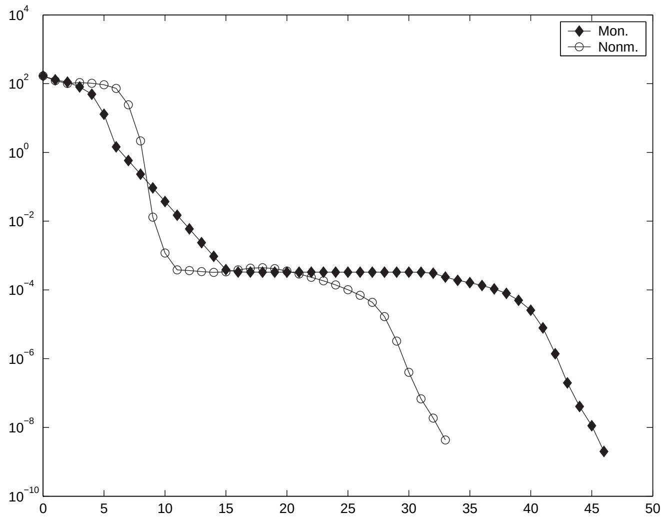
Table 1: References and starting points

P	Reference	$(x_0)_i$ direct	$(x_0)_i$ iterative	χ
P1	Example 5.5 in [8]	1	1	$10^7 - 10^8$
P2	Example 5.6 in [8]	0.01	2	$10^6 - 10^7$
P3	Example 5.7 in [8]	0.5	1.5	$10^6 - 10^7$
P4	Example 5.8 in [8]	3.995	2 ; 7	$10^6 - 10^7$
P5	Example 4 in [9]		0 ; 3	$10^6 - 10^7$
P6	Example 5 in [9]		0 ; 3	$10^6 - 10^7$
P7	Example 4.2, $M = 1, K = 0.8$ in [10]		1.75 ; 5	$10^7 - 10^8$
P8	Example 4.2, $M = 0, K = 1$ in [10]		3	$10^7 - 10^8$

[1]), so an inexact solution of (35) is calculated. The nonlinear programming problems considered here arise from the discretization by finite difference of elliptic control problems described in [8], [9] and [10]. The references are listed in Table 1. In the third and fourth column of Table 1, the starting points for the two choices of solver, direct and iterative, have been reported; when two different values are listed on the same row, the first one is the value of the components of x_0 related to the state variables, while the second one is related to the control variable. Only the value of the variable x_0 are reported, while the other components of the vector v_0 are always been set equal to 1. In the last column of Table 1 is specified the interval which the parameter χ in [1] belongs to. In Table 2 the results of the monotone and nonmonotone algorithms with the direct inner solver are compared in terms of number of iterations (it.) and total number of backtracking reductions (b.). Each test problem has been executed three times, by changing the meshsize: the values on the first column indicate the number of meshpoints on the x and y axis. In this case, since an exact solver has been adopted, the nonmonotone scheme differs to the monotone one only on the backtracking rule. From the results, in some cases the two algorithms seem to behave in a similar way. In more critical cases, in order to satisfy the monotone backtracking rule the damping parameter is reduced to a very small value; this fact yields the failure of the algorithm, while the nonmonotone rule allows to accept larger values of the damping parameter, avoiding in many cases the stagnation of the iterates. In general, a reduction of the number of the backtracking steps can be observed.

Figure 1 illustrates the decrease of $\|H(v_k)\|$ for P1 with $n = 10593$. The results Table 3 have been obtained employing the iterative solver, so the number of the inner iterations (inn.) is reported. Now the difference between the monotone and the nonmonotone schemes are not only in the backtracking rule, but in the stopping criterion of the inner solver too. In the last column are listed the final values of the objective function (obj.). A

Figure 1



general reduction of the number of inner iterations can be observed, and in many cases the number of external iterations (ext.) and of the backtracking reductions (b.) is also reduced.

5 Conclusions

We proposed a variant of Inexact Newton Method in which monotonicity requirements have been relaxed. For the modified scheme we devised conditions under which we proved the convergence theorems. Then we applied the nonmonotone techniques to the inexact interior–point method, as special case of Inexact Newton Method, and we proved the convergence of the whole scheme. As shown in the tables 2 and 3, the nonmonotone approach can reduce the number of the backtracking steps and of the inner iterations when an iterative solver is employed.

References

[1] S. Bonettini, E. Galligani and V. Ruggiero (2003). *A Newton Inexact Interior–Point method combined with Hestenes' multipliers scheme*, Technical Report n.334, Dipartimento di Matematica dell'Università di Ferrara.

Table 2: Numerical results: direct inner solver

			Monotone		Nonmonotone					
			$N=0$		$N=2$		$N=4$		$N=9$	
Grid	n	P	it	b.	it	b.	it	b.	it	b.
50	2793	P1	27	1	26	0	26	0	26	0
		P2	30	15	–	–	26	2	26	2
		P3	–	–	25	0	25	0	25	0
		P4	33	6	–	–	–	–	34	2
100	10593	P1	46	55	–	–	–	–	33	0
		P2	–	–	–	–	–	–	32	13
		P3	–	–	26	1	27	0	27	0
		P4	31	0	31	0	31	0	31	0
150	23393	P1	–	–	–	–	–	–	–	–
		P2	–	–	–	–	–	–	41	19
		P3	–	–	26	1	27	0	27	0
		P4	29	2	32	1	32	1	32	1

Table 3: Numerical results: iterative inner solver

			Monotone			Nonmonotone									
			$N=0$			$N=2$			$N=4$			$N=9$			
P	Grid	n	ext.	inn.	b.	ext.	inn.	b.	ext.	inn.	b.	ext.	inn.	b.	obj.
P1	50	2793	25	27	1	22	22	0	22	22	0	22	22	0	.5479649
	100	10593	35	36	1	29	29	1	29	29	1	29	29	1	.5522459
	200	41193	–	–	–	53	53	1	53	53	1	53	53	1	.5543686
P2	50	2793	23	24	0	23	23	0	23	23	0	22	23	0	.0140651
	100	10593	31	33	0	31	31	0	31	31	0	31	31	0	.0150786
	150	23393	–	–	–	39	39	0	39	39	0	39	39	0	.0154262
P3	50	2793	17	19	0	18	18	0	18	18	0	18	18	0	.2575581
	100	10593	24	26	0	23	23	0	23	23	0	23	23	0	.2638984
	200	41193	31	34	0	32	32	0	32	32	0	32	32	0	.2671221
P4	50	2793	18	19	0	18	18	0	18	18	0	18	18	0	.1539771
	100	10593	26	28	0	26	26	0	26	26	0	26	26	0	.1616639
	200	41193	39	41	0	37	37	0	37	37	0	37	37	0	.1657634
P5	50	4998	17	18	0	17	17	0	17	17	0	17	17	0	.0773888
	100	19998	17	18	0	17	17	0	17	17	0	17	17	0	.0780638
	200	79998	20	23	0	19	19	0	19	19	0	19	19	0	.0784259
P6	50	4998	29	30	0	29	29	0	29	29	0	29	29	0	.0521892
	100	19998	41	42	0	41	41	0	41	41	0	41	41	0	.0526638
	200	79998	69	72	0	59	59	0	59	59	0	59	59	0	.0529328
P7	50	4998	21	21	0	21	21	0	21	21	0	21	21	0	-6.4857811
	100	19998	27	28	0	27	27	0	27	27	0	27	27	0	-6.5764272
	200	79998	46	47	0	46	46	0	46	46	0	46	46	0	-6.6200922
P8	50	4998	28	28	0	28	28	0	28	28	0	28	28	0	-18.4825400
	100	19998	42	43	0	42	42	0	42	42	0	42	42	0	-18.7361482
	200	79998	51	96	0	51	66	0	51	51	0	51	51	0	-18.8633116

[2] R. S. Dembo, S. C. Eisenstat and T. Steihaug(1982). *Inexact Newton methods*, SIAM Journal on Numerical Analysis, **19**, 400–408.

[3] C. Durazzi (2000) *On the Newton interior–point method for nonlinear programming problems*, Journal of optimization Theory and Applications, **104**, 73-90.

[4] C. Durazzi and V. Ruggiero (2003). *A Newton Inexact Interior–Point method for large scale nonlinear optimization problems*, Annali dell'Università di Ferrara, Sezione VII Scienze Matematiche, **49**, 333-357.

[5] S. C. Eisenstat and H. F. Walker (1994). *Globally convergent Inexact Newton methods*, SIAM Journal on Optimization, **4**, 393–422.

[6] A. S. El Backry, R. A. Tapia, T. Tsuchiya and Y. Zhang (1996). *On the Formulation and Theory of Newton Interior–Point Method for Nonlinear Programming*, Journal of Optimization Theory and Applications, **89**, 507–541.

[7] L. Grippo, F. Lampariello and S. Lucidi (1986). *A Nonmonotone line search technique for Newton's method*, SIAM Journal on Numerical Analysis **23**, 707–716.

[8] H. D. Mittelmann and H. Maurer (1999). *Optimization techiques for solving elliptic control problems with control and state constraints: Part 1. Boundary control*, Computational Optimization and Applications, **16**, 29–55.

[9] H. D. Mittelmann and H. Maurer (2001). *Optimization techiques for solving elliptic control problems with control and state constraints: Part 2. Distributed control*, Computational Optimization and Applications, **18**, 141–160.

[10] H. D. Mittelmann and H. Maurer (2000). *Solving elliptic control problems with Interior Point and SQP Methods: control and state constraints*, Journal of Computational and Applied Mathematics, **120**, 175–195.

[11] J. M. Ortega and W. C. Rheimboldt (1970). *Iterative solution of nonlinear equations in several variables*, Academic Press, New York.