Solving complementarity problems by means of a new smooth constrained nonlinear solver

Roberto Andreani *   José Mario Martínez †

October 17, 1997 (revised)

Abstract

Given $F : \mathbb{R}^n \to \mathbb{R}^m$ and Ω a closed and convex set, the problem of finding $x \in \mathbb{R}^n$ such that $x \in \Omega$ and $F(x) = 0$ is considered. For solving this problem an algorithm of Inexact-Newton type is defined. Global and local convergence proofs are presented. As a practical application, the Horizontal Nonlinear Complementarity Problem is introduced. It is shown that the Inexact-Newton algorithm can be applied to this problem. Numerical experiments are performed and commented.

Keywords. Nonlinear systems, Inexact–Newton method, global convergence, convex constraints, box constraints, complementarity.

AMS: 65H10, 90C33, 90C30

---

*Department of Applied Mathematics, IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas SP, Brazil . This author was supported by FAPESP (Grant 90-3724-6, 93/02479-6), FINEP and FAEP-UNICAMP.

†Department of Mathematics, IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas SP, Brazil. This author was supported by FAPESP (Grant 90-3724-6), FINEP and FAEP-UNICAMP. E-Mail: martinez@ime.unicamp.br

1 Introduction

The problem considered in this paper is to find $x \in \Omega \subset I\!R^n$ such that

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

where Ω is closed and convex, $F : \Omega \to I\!R^m$, and first derivatives of F exist and are continuous on an open set that contains Ω, except, perhaps, at the solutions of (1).

When $m = n$ this is the constrained nonlinear system problem, considered recently in [17]. By means of the introduction of slack variables, any nonlinear feasibility problem given by a set of equations and inequalities can be reduced to the form (1).

We introduce an Inexact-Newton-like algorithm (see [6]) for solving (1). At each iteration of this algorithm, a search direction d is computed such that

$$\|F'(x_k)d + F(x_k)\| \le \theta_k \|F(x_k)\|, \quad x_k + d \in \Omega, \quad |d| \le \Delta,$$

where $\|\cdot\| = \|\cdot\|_2$, $|\cdot|$ is an arbitrary norm, Δ is large and $\theta_k \in [0, 1)$. If the Inexact-Newton step does not exist, the tolerance θ_k is increased. Otherwise, we try to find a new point in the direction of the computed step. We prove that global convergence holds in the sense that, under some conditions, a solution of (1) is met and, in general, a stationary point of the problem can be found. We establish conditions under which local convergence results can also be proved. A practical implementation of the algorithm is given for the case in which Ω is an $n-$dimensional box.

As an application of the Inexact-Newton algorithm we define the Horizontal Nonlinear Complementarity Problem (HNCP). This is a generalization of the Horizontal Linear Complementarity Problem (HLCP) (see [2]) for which several interesting applications exist. The HNCP also generalizes the well-known Nonlinear Complementarity Problem. See [5], [9], [14], [15], [20] and [21]. We prove that, under some conditions on the HNCP, stationary points of this problem coincide with their global solutions and we also analyze the conditions under which local convergence results hold. Finally, we show some numerical experiments.

The Inexact-Newton method for solving (1) is given in Section 2 of the paper. In Section 3 we prove the convergence results. In Section 4 we discuss the application to HNCP. In Section 5 we show the numerical experiments and in Section 6 we state some conclusions.

2 The Inexact-Newton Algorithm

Let $\Omega \subset I\!R^n$ be closed and convex, $\Omega \subset \mathcal{A}$, $\mathcal{A}$ an open set of $I\!R^n$. We assume that $F : \mathcal{A} \to I\!R^m$ is continuous and that the Jacobian matrix $F'(x) \in I\!R^{m\times n}$ exists and is continuous for all $x \in \mathcal{A}$ such that $F(x) \neq 0$.

We say that x_* is a *stationary point* of problem (1) if it is a minimizer of $\|F(x_*) + F'(x_*)(x - x_*)\|$ subject to $x \in \Omega$. Clearly, this is equivalent to say that x_* is a minimizer of $q(x) \equiv \|F(x_*) + F'(x_*)(x - x_*)\|^2$ subject to $x \in \Omega$. If the convex set Ω satisfies a constraint qualification (as it is the case of the application of Section 4, where Ω is the positive orthant) this is equivalent to say that the first order optimality conditions (Karush-Kuhn-Tucker) of the problem of minimizing the convex quadratic $q(x)$ on Ω are satisfied at x_*. But, since $\nabla q(x_*) = \nabla f(x_*)$ (where $f(x) \equiv \|F(x)\|^2$) it turns out that the stationarity of x_* corresponds exactly to the satisfaction of the Karush-Kuhn-Tucker conditions for minimizers of f on Ω.

Below we describe the main model algorithm of this paper. The algorithm computes a sequence of points $x_k \in \Omega$, starting from an arbitrary x_0. The following parameters are given independently of the iteration index k: $\sigma \in (0,1)$ (generally $\sigma \in [10^{-4}, 0.1]$) is associated to a sufficient decrease condition for the norm of F. The reduction of the stepsize α_k is related to two backtracking parameters $\kappa_1, \kappa_2 \in (0,1)$ (generally $\kappa_1 = 0.1, \kappa_2 = 0.9$). The scale-dependent large number $\Delta > 0$ is the maximum size admitted, initially, for the steplength. The parameter $\bar{\rho} \geq 1$ is an upper bound for the extrapolation factor related to the pure Inexact-Newton iteration. Finally $\theta_0 \in (0,1)$ is the initial value for the Inexact-Newton parameter θ_k. In Algorithm 2.1 it is explained how to obtain x_{k+1}, given the current iterate x_k, for all $k = 0, 1, 2, \ldots$ We initialize $\alpha_0 = 1$, which reflects our original intention of performing full Inexact-Newton steps.

Algorithm 2.1.

Given $x_k \in \Omega$, $\alpha_k \in (0,1]$, $\theta_k \in [0,1)$, the steps for obtaining x_{k+1}, α_{k+1}, and θ_{k+1} or for declaring finite convergence, are the following:

Step 0. *Stop if x_k is a solution.*
If $F(x_k) = 0$, stop. (A solution has been found.)

Step 1. *Find the Inexact-Newton direction.*
Try to find $d_k \in I\!R^n$ such that

$$x_k + d_k \in \Omega, \quad |d_k| \leq \Delta \tag{2}$$

and

$$\|F'(x_k)d_k + F(x_k)\| \leq \theta_k \|F(x_k)\|. \tag{3}$$

If a direction d_k satisfying (2)–(3) does not exist, define $x_{k+1} = x_k$, $\alpha_{k+1} = \alpha_k$, $\theta_{k+1} = (1+\theta_k)/2$, and finish the iteration. If a direction d_k satisfying (2)–(3) is found, define $\theta_{k+1} = \theta_k$.

Step 2. *Reject the trial point if the norm increased.*
If

$$\|F(x_k + \alpha_k d_k)\| \geq \|F(x_k)\|, \tag{4}$$

define $x_{k+1} = x_k$, $\alpha_{k+1} \in [\kappa_1 \alpha_k, \kappa_2 \alpha_k]$ and finish the iteration.

Step 3. *Test sufficient decrease and try to extrapolate.*
Define $\gamma_k = 1 - \theta_k^2$. If

$$\|F(x_k + \alpha_k d_k)\| \leq (1 - \frac{\sigma \gamma_k \alpha_k}{2})\|F(x_k)\|, \tag{5}$$

find $\rho_k \in [1, \bar{\rho}]$ such that $x_k + \rho_k \alpha_k d_k \in \Omega$ and

$$\|F(x_k + \alpha_k \rho_k d_k)\| \leq \|F(x_k + \alpha_k d_k)\|, \tag{6}$$

define $x_{k+1} = x_k + \alpha_k \rho_k d_k$, $\alpha_{k+1} = 1$ and finish iteration k.

If (5) does not hold, define $x_{k+1} = x_k + \alpha_k d_k$, choose

$$\alpha_{k+1} \in [\kappa_1 \alpha_k, \kappa_2 \alpha_k]. \tag{7}$$

and finish the iteration.

Some algorithms for solving $n \times n$ nonlinear systems accept trial points generated by Newton-like procedures only when some sufficient decrease condition holds. See, for example, [7]. Here we adopt the point of view that a trial point deserves to be accepted if the norm of F decreases, since this feature tends to produce larger steps far from the solution, as it is generally desired in minimization algorithms. Moreover, using the extrapolation step we try to find a point along the Inexact-Newton direction, the distance of which to x_k is larger than the distance of the pure Inexact-Newton trial point to the current point. In this way, we try to alleviate the tendency of backtracking algorithms to produce short steps.

Let us now describe a particular version of Algorithm 2.1, which corresponds to the case in which Ω is an $n-$dimensional box, perhaps with some infinite bounds. So,

$$\Omega = \{x \in I\!R^n \quad | \quad x_i \geq \ell_i \quad \text{for} \quad i \in I \quad \text{and} \quad x_i \leq u_i \quad \text{for} \quad i \in J\},$$

where I and J are subsets of $\{1, \ldots, n\}$. In this case we use $|\cdot| = \|\cdot\|_\infty$, so that the constraints (2) also define a box. The key point is that the step d_k is computed by means of the consideration of the following box-constrained quadratic minimization problem:

$$\text{Minimize } \Phi(d) \quad \text{subject to} \quad x_k + d \in \Omega \quad \text{and} \quad \|d\|_\infty \leq \Delta, \tag{8}$$

where $\Phi(d) = \|F'(x_k)d + F(x_k)\|^2$. A necessary and sufficient condition for a global minimizer of (8) is that $\|P(d - \nabla\Phi(d)) - d\| = 0$, where P is the orthogonal projection on the feasible region of (8). Moreover, if $0 \in I\!R^n$ is a solution of (8) it turns out that x_k is a stationary point of (1). The particular case of the model algorithm that we are going to describe also uses the scale-dependent parameter $\Delta > 0$, the sufficient descent constant $\sigma \in (0, 1)$, the maximum extrapolation parameter $\bar{\rho} \geq 1$ and the initial Inexact-Newton tolerance $\theta_0 \in [0, 1)$. We initialize, as before, $\alpha_0 = 1$. Moreover, we use an additional tolerance $\eta_k \in [0, 1)$ for deciding termination of the method used for solving the subproblem (8). Roughly speaking, η_k is the degree of accuracy used in the approximate solution of (8). Different strategies for the choice of η_k at each iteration will be discussed in Section 5.

Algorithm 2.2.

Given $x_k \in \Omega$, $\alpha_k \in (0, 1]$, $\theta_k \in [0, 1)$, $\eta_k \in [0, 1)$, the steps for obtaining x_{k+1}, α_{k+1}, and θ_{k+1} or for declaring finite convergence, are the following:

Step 0. *Stop the execution if x_k is a solution.*

If $F(x_k) = 0$, stop. (A solution has been found.)

Step 1. *Stop the execution if x_k is stationary.*

If $\|P(-\nabla\Phi(0))\| = 0$, stop. (A stationary point has been found.)

Step 2. *Compute the step d_k.*

If $k > 0$ and $x_k = x_{k-1}$ take $d_k = d_{k-1}$. Otherwise, choose $d_k \in I\!R^n$ such that

$$x_k + d_k \in \Omega, \quad \|F'(x_k)d_k + F(x_k)\| \leq \theta_k \|F(x_k)\| \quad \text{and} \quad \|d_k\|_\infty \leq \Delta. \tag{9}$$

If such a choice is not possible, replace $\theta_k \leftarrow (1 + \theta_k)/2$ and repeat Step 2.

Step 3. *Repeat the current iterate if the norm did not decrease.*

Define $\theta_{k+1} = \theta_k$. If (4) takes place, define $x_{k+1} = x_k$, $\alpha_{k+1} = \alpha_k/2$ and finish the iteration.

Step 4. *Test sufficient decrease.*

Define $\gamma_k = 1 - \theta_k^2$. If (5) does not hold, define $x_{k+1} = x_k + \alpha_k d_k$, $\alpha_{k+1} = \alpha_k/2$ and finish iteration k.

Step 5. *Try extrapolation step.*
Step 5.1. Set $\underline{\rho} \leftarrow 1$, $\rho \leftarrow 2$.
Step 5.2. If

$$\rho > \bar{\rho} \quad \text{or} \quad \|F(x_k + \alpha_k \rho d_k)\| \geq \|F(x_k + \alpha_k \underline{\rho} d_k)\|$$

set

$$\rho_k = \underline{\rho}, \;\; x_{k+1} = x_k + \alpha_k \rho_k d_k, \;\; \alpha_{k+1} = 1$$

and finish iteration k.
Step 5.3. Set $\underline{\rho} \leftarrow \rho$, $\rho \leftarrow 2\rho$ and go to Step 5.2.

As we mentioned above, for computing a direction d_k satisfying (9), we apply an iterative method to (8), stopping when an increment $\bar{d}$ is obtained such that $x_k + \bar{d} \in \Omega$, $\|\bar{d}\|_\infty \leq \Delta$ and

$$\|P(\bar{d} - \nabla\Phi(\bar{d})) - \bar{d}\| \leq \eta_k \|P(-\nabla\Phi(0))\|. \tag{10}$$

If the step $\bar{d}$ obtained in this way satisfies (3), we define $d_k = \bar{d}$. Otherwise, we continue the execution of the quadratic solver stopping only when an iterate $\bar{d}$ satisfies $\|F'(x_k)\bar{d} + F(x_k)\| \leq \theta_k \|F(x_k)\|$ or when a global minimizer of (8) is found.

If a minimizer $\bar{d}$ of (8) is found but $\|F'(x_k)\bar{d} + F(x_k)\| > \theta_k \|F(x_k)\|$, then the condition (9) cannot be fulfilled, so we increase θ_k. In Algorithm 2.1 we changed the iteration index in this case while in Algorithm 2.2 we preferred to redefine θ_k without increasing the iteration number. Clearly, this is only a formal modification that does not affect the convergence properties of the algorithm. However, from the description adopted in Algorithm 2.2 it is clear that no additional work is needed when one increases θ_k. In fact, if $\bar{d}$ is a minimizer of (8) and the null vector is not a solution of (8), the condition $\|F'(x_k)\bar{d} + F(x_k)\| \leq \theta \|F(x_k)\|$ is necessarily satisfied for large enough $\theta < 1$. Therefore, this rejection involves only a redefinition of θ_k.

3 Convergence results

The first theorem in this section proves that, if θ_k is increased a finite number of times then Algorithm 2.1 finds a solution of (1).

Theorem 3.1. *Let $\{x_k\}$ be a sequence generated by Algorithm 2.1, such that $\theta_k = \bar{\theta} < 1$ for all $k \geq k_0$. Then, every limit point of $\{x_k\}$ is solution of (1).*

Proof. Without loss of generality, assume that $\theta_k = \bar{\theta}$ for all $k = 0, 1, 2, \ldots$

A straightforward calculation shows that (3) implies that

$$\langle F'(x_k)d_k, F(x_k)\rangle \leq -\frac{\gamma_k}{2}\|F(x_k)\|^2.$$

Define $\varphi(x) = \frac{1}{2}\|F(x)\|^2$ and $\bar{\gamma} = 1 - \bar{\theta}^2$. Since $\nabla\varphi(x) = F'(x)^T F(x)$, we have that

$$\langle \nabla\varphi(x_k), d_k\rangle \leq -\frac{\bar{\gamma}}{2}\|F(x_k)\|^2, \tag{11}$$

for all $k = 0, 1, 2, \ldots$.

Let x_* be a limit point of $\{x_k\}$ and $K_0 \subset I\!\!N$, such that

$$\lim_{k\in K_0} x_k = x_*.$$

Let us call

$$K_1 = \{k \in K_0 \text{ such that (5) holds}\}.$$

Assume first that K_1 is infinite and $\lim\sup_{k\in K_1}\alpha_k > 0$. Let K_2 be an infinite subset of K_1 such that

$$\alpha_k \geq \overline{\alpha} > 0$$

for all $k \in K_2$. Therefore we obtain

$$1 - \frac{\sigma\bar{\gamma}\alpha_k}{2} \leq 1 - \frac{\sigma\bar{\gamma}\overline{\alpha}}{2} \equiv r < 1$$

for all $k \in K_2$. Therefore $\{\|F(x_k)\|\}$ is a nonincreasing sequence such that $\|F(x_{k+1})\| \leq r\|F(x_k)\|$ for all $k \in K_2$. This implies that $\|F(x_k)\| \to 0$. Therefore $F(x_*) = 0$.

The following possibilities remain to be considered:

(i)

$$K_1 \text{ is infinite but } \lim_{k\in K_1} \alpha_k = 0. \tag{12}$$

(ii)

$$K_1 \text{ is finite;} \tag{13}$$

Let us consider first (i). Assume, by contradiction, that $F(x_*) \neq 0$. Without loss of generality, assume that $\alpha_k < 1$ for all $k \in K_1$. So, by (5) and (7), we have that, for all $k \in K_1$,

$$\alpha_k \in [\kappa_1\alpha_{k-1}, \kappa_2\alpha_{k-1}] \tag{14}$$

and

$$\|F(x_{k-1} + \alpha_{k-1} d_{k-1})\|^2 > (1 - \frac{\sigma \bar{\gamma} \alpha_{k-1}}{2})^2 \|F(x_{k-1})\|^2. \tag{15}$$

By (12) and (14) we have that

$$\lim_{k \in K_1} \alpha_{k-1} = 0.$$

So, since $|d_k|$ is bounded, we see that

$$\lim_{k \in K_1} x_{k-1} = x_*.$$

Now, by (15),

$$\frac{\varphi(x_{k-1} + \alpha_{k-1} d_{k-1}) - \varphi(x_{k-1})}{\alpha_{k-1}} > -\frac{\sigma \bar{\gamma} \|F(x_{k-1})\|^2}{2} + \frac{\sigma^2 \bar{\gamma}^2 \alpha_{k-1} \|F(x_{k-1})\|^2}{8}.$$

for all $k \in K_1$. Since $\alpha_{k-1} \to 0$ and $|d_{k-1}|$ is bounded, we have that $x_{k-1} + \alpha_{k-1} d_{k-1}$ tends to x_* for $k \in K_1$. But $F(x_*) \neq 0$ so $F'(x)$ exists and is continuous in a neighborhood of x_*. For large enough $k \in K_1$ we have that both x_{k-1} and $x_{k-1} + \alpha_{k-1} d_{k-1}$ belong to that neighborhood. So, we can apply the Mean Value Theorem, which implies that for large enough $k \in K_1$, there exists $\xi_{k-1} \in [0, 1]$ such that

$$\langle \nabla \varphi(x_{k-1} + \xi_{k-1} \alpha_{k-1} d_{k-1}), d_{k-1} \rangle > -\frac{\sigma \bar{\gamma} \|F(x_{k-1})\|^2}{2} + \frac{\sigma^2 \bar{\gamma}^2 \alpha_{k-1} \|F(x_{k-1})\|^2}{8}. \tag{16}$$

Since $|d_k| \leq \Delta$ for all k, there exists K_3, an infinite subset of K_1, such that

$$\lim_{k \in K_3} d_{k-1} = d.$$

Taking limits for $k \in K_3$ on both sides of (16), we obtain

$$\langle \nabla \varphi(x_*), d \rangle \geq -\frac{\sigma \bar{\gamma} \|F(x_*)\|^2}{2}.$$

So, for large enough $k \in K_3$, defining $\sigma' = \frac{\sigma + 1}{2}$, we have that

$$\langle \nabla \varphi(x_{k-1}), d_{k-1} \rangle \geq -\frac{\sigma' \bar{\gamma} \|F(x_{k-1})\|^2}{2} > -\frac{\bar{\gamma} \|F(x_{k-1})\|^2}{2} \tag{17}$$

Therefore, (17) contradicts (11). This proves that $F(x_*) = 0$.

Let us now assume (ii). Since K_1 is finite, there exists $k_3 \in K_0$ such that (5) does not hold for all $k \in K_0$ and $k \geq k_3$. Therefore

$$\lim_{k \in K_0} \alpha_k = 0,$$

and we can repeat the former proof, with minor modifications, for proving that $F(x_*) = 0$. $\square$

In Theorem 3.1 we proved that if θ_k needs to be increased only a finite number of times, then any acumulation point of a sequence generated by Algorithm 2.1 is a solution of problem. In the next theorem we show what happens if θ_k needs to be increased infinitely many times. Essentially, we show that, in that case, we find a stationary point of the problem.

Theorem 3.2. *Suppose that, in Algorithm 2.1, θ_k is increased infinitely many times and define*

$$K_4 = \{k \in \{0, 1, 2, \ldots\} \mid \theta_{k+1} > \theta_k\}.$$

Then, every limit point of the sequence $\{x_k\}_{k \in K_4}$ is stationary.

Proof. Let $x_* \in \Omega$ be a limit point of $\{x_k\}_{k \in K_4}$. If $F(x_*) = 0$ we are done, so let us assume that $\|F(x_*)\| > 0$. This implies that $F'(x)$ exists and is continuous at x_*. Suppose, by contradiction, that x_* is not stationary. Therefore, there exists $d \in I\!\!R^n$ such that $|d| \leq \Delta/2$, $x_* + d \in \Omega$ and

$$\|F'(x_*)d + F(x_*)\| < \|F(x_*)\|.$$

Define

$$\frac{\|F'(x_*)d + F(x_*)\|}{\|F(x_*)\|} = r < 1.$$

Choose $r' \in (r, 1)$. By continuity of F and F', we have that

$$\frac{\|F'(x_k)(x_* + d - x_k) + F(x_k)\|}{\|F(x_k)\|} \leq r' \tag{18}$$

for large enough $k \in K_4$. But, since $|d| \leq \Delta/2$ for large enough $k \in K_4$, $|x_* + d - x_k| \leq \Delta$. So, (18) contradicts the fact that (3) cannot be achieved for $k \in K_4$ and $\theta_k \to 1$. $\square$

Remark. The hypotheses of Theorems 3.1 and 3.2 have an interesting interpretation from a non strictly algorithmic point of view. In fact, the operator $L(x) = F(\bar{x}) + F'(\bar{x})(x - \bar{x})$ is the first-order affine model of $F(x)$ that uses information available at $\bar{x}$. In Theorem 3.1 we assume that the inequality $\|L(x)\| \leq \bar{\theta}\|L(\bar{x})\|$, $\bar{\theta} < 1$, holds for some $x \in \Omega$, $\|x - \bar{x}\| \leq \Delta$ for all algorithmic iterates. Its conclusion implies that whenever Ω is bounded or $\|F(x)\|$ has bounded level sets, a solution of the system *exists*. In other words, Theorem 3.1 can be interpreted as an existence theorem that says that solutions of a bounded problem exist when it is possible a sufficient decrease of the model of the system on a neighborhood of fixed (not necessarily small) arbitrary size.

Now, many times one applies nonlinear system solvers to problems where a solution is not known to exist and sometimes the solution does not exist at all. In this case, according to Theorem 3.1, sufficient reduction of the model cannot be possible at all points. However, Theorem 3.2 ensures that cluster points of iterates where sufficient reduction is not possible are stationary.

Now we prove two auxiliary lemmas which, in turn, will be useful in the local convergence analysis.

Lemma 3.3. *If $F(x_k) \neq 0$ and (3) holds, then $\|F'(x_k)\| \neq 0$ and*

$$\|d_k\| \geq \frac{1-\theta_k}{\|F'(x_k)\|}\|F(x_k)\|. \tag{19}$$

Proof. Since $\theta_k < 1$ and $F(x_k) \neq 0$, we obtain that $F'(x_k) \neq 0$. Now, by (3),

$$\|F(x_k)\| - \|F'(x_k)\|\|d_k\| \leq \|F'(x_k)d_k + F(x_k)\| \leq \theta_k\|F(x_k)\|.$$

So,

$$\|F'(x_k)\|\|d_k\| \geq (1-\theta_k)\|F(x_k)\|$$

and the thesis follows from this inequality. □

Lemma 3.4. *Let $\{w_k, k = 0, 1, 2, \ldots\}$ be a sequence in $I\!R^n$ and assume that*

$$\lim_{k\to\infty} \frac{\|w_{k+1} - w_k\|}{\|w_k - w_{k-1}\|} = 0. \tag{20}$$

Then, $\{w_k\}$ converges R-superlinearly to some $w_* \in I\!R^n$.

Proof. By (20), there exists $k_0 \in I\!N$ such that

$$\frac{\|w_{k+1} - w_k\|}{\|w_k - w_{k-1}\|} \leq \frac{1}{2} \quad \text{for all} \quad k \geq k_0. \tag{21}$$

So,

$$\|w_{k+1} - w_k\| \leq \frac{1}{2^{k-k_0}}\|w_{k_0+1} - w_{k_0}\| \tag{22}$$

for all $k \geq k_0$.

Now, by (22), given $\varepsilon > 0$, there exists $k_1 \in I\!N$ such that

$$\|w_{k+1} - w_k\| \leq \frac{\varepsilon}{2} \text{ for all } \quad k \geq k_1$$

Define $k_2 = \max\{k_1, k_0\}$. For all $l, j \geq k_2$ we have that

$$\|w_j - w_l\| \leq \sum_{i=l}^{j-1} \|w_{i+1} - w_i\| \leq (\sum_{i=l}^{j-1} \frac{1}{2^{i-l}})\|w_{l+1} - w_l\| \leq (\sum_{i=l}^{j-1} \frac{1}{2^{i-l}})\frac{\varepsilon}{2} \leq \varepsilon$$

Then, $\{w_k\}$, is a Cauchy sequence and so, it converges to some $w_* \in I\!R^n$. Now, if $k \geq k_2$, we have that

$$\|w_k - w_*\| \leq \sum_{i=k}^{\infty} \|w_{i+1} - w_i\| \leq 2\|w_{k+1} - w_k\|. \tag{23}$$

Therefore, the sequence $\{\|w_k - w_*\|\}$ is bounded by the Q-superlinear sequence $\{2\|w_{k+1} - w_k\|\}$. This implies that the convergence of $\{w_k\}$ is R-superlinear. □

The following assumption states that x_* is a solution of the problem such that, in some neighborhood of this point, the norm of $F(x)$ is bounded below by a quantity of the order of the distance between x and x_*. As a consequence, x_* is the unique solution of the system on that neighborhood.

Assumption 3.1. *Strong Local Unicity (SLU)*

There exist $x_* \in \Omega$, $\varepsilon, c > 0$ such that $F(x_*) = 0$ and $\|x - x_*\| \leq c\|F(x)\|$ for all $x \in \Omega$ such that $\|x - x_*\| \leq \varepsilon$.

In the following theorem we prove that, under some conditions on the direction d_k, Algorithm 2.1 converges superlinearly to a solution of the problem. The assumptions include a Lipschitz condition on the Jacobian, the fact that the Inexact-Newton direction tends to be a Newton direction (the direction is an approximate solution of $F'(x_k)d + F(x_k) = 0$ with accuracy tending to zero) and, finally, that the norm of d_k has the same order as $\|F(x_k)\|$. In this way R-superlinear convergence to some solution x_* is obtained. If, in addition, the SLU assumption holds, the convergence is Q-superlinear.

Theorem 3.5. *Assume that*

$$\{x \in \Omega \text{ such that } \|F(x)\| \leq \|F(x_0)\|\} \text{ is bounded} \tag{24}$$

and the following Lipschitz condition on $F'(x)$ holds:

$$\|F'(x) - F'(y)\| \leq L\|x - y\| \tag{25}$$

for all $x, y \in \Omega$.

Assume that the sequence $\{x_k\}$ *is generated by Algorithm 2.1 and that there exists* $t_k \to 0$, $\beta > 0$ *such that*

$$\|F'(x_k)d_k + F(x_k)\| \le t_k\|F(x_k)\| \tag{26}$$

and

$$\|d_k\| \le \beta\|F(x_k)\| \tag{27}$$

for all $k = 0, 1, 2, \ldots$. *Then, there exists* $k_0 \in \{0, 1, 2, \ldots\}$ *such that* $\alpha_k = 1$ *for all* $k \ge k_0$. *Moreover,*

$$\lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|F(x_k)\|} = 0, \tag{28}$$

and the sequence is R-superlinearly convergent to some $x_* \in \Omega$ *such that* $F(x_*) = 0$. *Finally, if* x_* *satisfies the SLU Assumption, the convergence is Q-superlinear.*

Proof. By (25), (26) and (27), for all $k = 0, 1, 2, \ldots$, we have that

$$\begin{aligned} \|F(x_k + \alpha_k d_k)\| &\le \|F(x_k) + \alpha_k F'(x_k)d_k\| + \frac{L}{2}\alpha_k^2\|d_k\|^2 \\ &\le \|\alpha_k[F(x_k) + F'(x_k)d_k]\| + (1-\alpha_k)\|F(x_k)\| + \frac{L}{2}\alpha_k^2\|d_k\|^2 \\ &\le \alpha_k t_k\|F(x_k)\| + (1-\alpha_k)\|F(x_k)\| + \frac{L}{2}\alpha_k^2\|d_k\|^2 \\ &\le \alpha_k t_k\|F(x_k)\| + (1-\alpha_k)\|F(x_k)\| + \frac{L}{2}\alpha_k^2\beta^2\|F(x_k)\|^2 \\ &= (\alpha_k t_k + 1 - \alpha_k + \frac{L}{2}\alpha_k^2\beta^2\|F(x_k)\|)\|F(x_k)\| \\ &= [1 - \alpha_k(1 - t_k - \frac{L}{2}\alpha_k\beta^2\|F(x_k)\|)]\|F(x_k)\|. \end{aligned} \tag{29}$$

By (24), (26) and Theorem 3.1 we have that $\|F(x_k)\| \to 0$. Since $t_k \to 0$ and $\sigma\gamma_k/2 \in (0, 1)$, (29) implies that there exists $k_1 \in \{0, 1, 2, \ldots\}$ such that

$$\|F(x_k + \alpha_k d_k)\| \le (1 - \frac{\sigma\gamma_k\alpha_k}{2})\|F(x_k)\|$$

if $k \ge k_1$. By the definition of Algorithm 2.1, this implies that $\alpha_k = 1$ for all $k \ge k_0 \equiv k_1 + 1$.

Now, by (29) with $\alpha_k = 1$ for $k \ge k_0$, we obtain

$$\|F(x_{k+1})\| \le \|F(x_k + d_k)\| \le [1 - (1 - t_k - \frac{L}{2}\beta^2\|F(x_k)\|)]\|F(x_k)\|.$$

Therefore, (28) holds.

By the compactness assumption and Lipschitz condition there exists $c_1 \in \mathbb{R}$ such that $\|F'(x_k)\| \leq c_1$ for all $k = 0, 1, 2, ...$, then , by Lemma 3.3, there exists $\beta_1 > 0$ such that

$$\|d_k\| \geq \beta_1 \|F(x_k)\|. \tag{30}$$

Since $\rho_k \leq \bar{\rho}$ for all k, by (27), (28) and (30), we obtain

$$\|x_{k+1} - x_k\| = \rho_k \|d_k\| \leq \bar{\rho}\beta \|F(x_k)\| \leq \bar{\rho}\beta(\frac{\|F(x_k)\|}{\|F(x_{k-1})\|})\|F(x_{k-1})\|$$

$$\leq \frac{\bar{\rho}\beta}{\beta_1} u_k \|d_{k-1}\| = \frac{\bar{\rho}\beta}{\beta_1 \rho_{k-1}} u_k \|x_k - x_{k-1}\| \leq M u_k \|x_k - x_{k-1}\|$$

for all $k \geq k_1$, where $M = \frac{\bar{\rho}\beta}{\beta_1}$ and $u_k = \frac{\|F(x_k)\|}{\|F(x_{k-1})\|} \to 0$. Therefore

$$\frac{\|x_{k+1} - x_k\|}{\|x_k - x_{k-1}\|} \to 0$$

So, by Lemma 3.4, $\{x_k\}$ converges R-superlinearly to some solution $x_* \in \Omega$.

If Assumption 3.1 holds, there exists $k_2 \in \mathbb{N}$ such that, for all $k \geq k_2$,

$$\|x_{k+1} - x_*\| \leq c\|F(x_{k+1})\|. \tag{31}$$

But, by (25) there exists $L_1 > 0$ such that

$$\|x_k - x_*\| \geq \frac{1}{L_1}\|F(x_k)\|. \tag{32}$$

By (31) and (32) we have that

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leq c L_1 \frac{\|F(x_{k+1})\|}{\|F(x_k)\|}$$

Therefore, by (28), $\{x_k\}$ converges Q-superlinearly to x_*. $\square$

The following auxiliary lemma will be useful to prove a local convergence result in Section 4. Its proof follows from elementary analysis considerations and will be omitted here.

Lemma 3.6. *Assume that $F \in C^1$, Ω convex, $x_* \in \Omega$ and let $\mathcal{V}_1(\Omega, x_*)$ be the set of directions $d \in \mathbb{R}^n$ such that $d = \lim_{k\to\infty} \frac{x_k - x_*}{\|x_k - x_*\|}$ for some sequence $\{x_k\} \subset \Omega$ that converges to x_*. Then the following propositions are equivalent:*

(i)

$$\inf_{v \in \mathcal{V}_1(\Omega, x_*)} \|F'(x_*)v\| = \zeta > 0.$$

(ii) *There exists* $c, \varepsilon > 0$, *such that*

$$\|x - x_*\| \leq c\|F(x) - F(x_*)\|,$$

for all $x \in \Omega$ *such that* $\|x - x_*\| \leq \varepsilon$.

4 Application to Complementarity Problems

Assume that $G, H : I\!R^q \to I\!R^q$ represent two transformation processes that depend on nonnegative sequential inputs $y_1, \ldots, y_q$ (inputs of G) and $z_1, \ldots, z_q$ (inputs of H). However, at each time stage $i \in \{1, \ldots, q\}$ only one of the two controls $\{y_i, z_i\}$ can be positive. In other words, the very essence of the work does not admit performing two jobs simultaneously. Suppose that, as a result of the transformations, the following synchronization equation must hold:

$$G(y) + H(z) = 0. \tag{33}$$

Clearly, the conditions stated above on the controls impose that

$$y^T z = 0 \quad \text{and} \quad y, z \geq 0. \tag{34}$$

The system (33)-(34) defines the Horizontal Nonlinear Complementarity Problem (HNCP). If G and H are affine functions, we have the Horizontal Linear Complementarity Problem (HLCP) (see [2, 12]). It can be shown that the optimality conditions of quadratic programming problems form an HLCP and that the optimality conditions of linearly constrained minimization problems form an HNCP. If $H(z) = -z$ for all $z \in I\!R^q$, the HNCP reduces to the classical Nonlinear Complementarity Problem (NCP).

Let us define $n = 2q, m = q + 1$, $x = (y, z)$, $\Omega = \{x \in I\!R^n \mid x \geq 0\}$, and

$$F(x) = F(y, z) = \begin{pmatrix} G(y) + H(z) \\ y^T z \end{pmatrix}$$

Then, the HNCP consists on finding a solution of

$$F(y, z) = 0, \quad y, z \in \Omega. \tag{35}$$

We seek to solve problem (35) using the algorithms defined in Section 2. With this purpose, it will be useful to characterize the conditions under which stationary points of (35) are, in fact, global solutions. An answer to this question is given in Theorem 4.1. Let us recall first that, since Ω satisfies a constraint qualification, the points that satisfy Karush-Kuhn-Tucker conditions for $\|F(y,z)\|^2$ on Ω are exactly the stationary points defined at the beginning of Section 2.

Theorem 4.1. *Assume that (y_*, z_*) is stationary point of problem (35) and $H'(z_*), G'(y_*)$ are such that for all $u \in I\!R^q$, $u \neq 0$, there exists $D(u) \in I\!R^{q\times q}$, a diagonal matrix with strictly positive diagonal elements, such that*

$$u^T H'(z_*) D(u) G'(y_*)^T u < 0.$$

Then, (y_*, z_*) is solution of (33)–(34).

Proof. The solutions of (35) are minimizers of the problem

$$\text{Minimize } \|G(y) + H(z)\|^2 + (y^T z)^2 \quad \text{subject to} \quad y, z \geq 0. \tag{36}$$

Let us call

$$u_* = G(y_*) + H(z_*). \tag{37}$$

The first-order optimality conditions of (36) are:

$$2G'(y_*)^T u_* + 2(y_*^T z_*) z_* - \mu_1 = 0, \tag{38}$$

$$2H'(z_*)^T u_* + 2(y_*^T z_*) y_* - \mu_2 = 0, \tag{39}$$

$$y_*^T \mu_1 = 0, \tag{40}$$

$$z_*^T \mu_2 = 0, \tag{41}$$

$$y_* \geq 0, \quad z_* \geq 0, \quad \mu_1 \geq 0, \quad \mu_2 \geq 0. \tag{42}$$

By (38)- (42), we obtain

$$4[H'(z_*)^T u_*]_i [G'(y_*]^T u_*]_i = 4(y_*^T z_*)^2 [y_*]_i [z_*]_i + [\mu_1]_i [\mu_2]_i \geq 0, \tag{43}$$

for all $i = 1, \ldots, q$.

Suppose, for a moment, that $u_* \neq 0$. Then, by the hypothesis of the theorem, there exists a diagonal matrix $D(u_*) \in I\!R^{q\times q}$ with strictly positive diagonal entries, such that

$$\sum_{i=1}^{q} [H'(z_*)^T u_*]_i [D(u_*)]_{i,i} [G'(y_*)^T u_*]_i < 0. \tag{44}$$

Since $[D(u_*)]_{i,i} > 0$, the equality (43) cannot be true. This contradiction came from the assumption $u_* \neq 0$. Therefore, $G(y_*) + H(z_*) = 0$. Thus, by (38) and (39), we have that

$$2(y_*^T z_*)z_* - \mu_1 = 0 \tag{45}$$

and

$$2(y_*^T z_*)y_* - \mu_2 = 0. \tag{46}$$

Multiplying (45) and (46), we obtain, using (40) and (41):

$$4(y_*^T z_*)^3 + \mu_1^T \mu_2 = 0.$$

So, by (42),

$$y_*^T z_* = 0,$$

as we wanted to prove. □

Remark. In the Nonlinear Complementarity Problem $NCP(G)$ ($H(z) = -z$), the hypothesis of Theorem 4.1 reads: For all $u_* \in I\!R^q$, $u_* \neq 0$, there exists a diagonal matrix $D(u_*)$ with strictly positive diagonal entries such that $u_*^T D(u_*) G'(y_*)^T u_* > 0$. Friedler and Pták [13] proved that this hypothesis holds if $G'(y_*)$ is a $P-$matrix (all principal minors of $G'(y_*)$ are positive).

Let us recall (see [4]) that a matrix $A \in I\!R^{n\times n}$ is called an S-matrix if there exists $x \in I\!R^n, x \geq 0$ such that $Ax > 0$. It can be proved (see [13]) that A is an S-matrix if, and only if, $\{y \in I\!R^n \mid y \geq 0, y \neq 0,\ A^T y \leq 0\}$ is empty.

A matrix $A \in R^{n\times n}$ is said to be column-sufficient (see [4]) if for all $u \in R^n$ such that

$$[u]_i[Au]_i \geq 0 \text{ for all } i = 1, \ldots, n$$

one necessarily has that

$$[u]_i[Au]_i = 0 \text{ for all } i = 1, \ldots, n.$$

A matrix A is row-sufficient if A^T is column-sufficient and it is called sufficient if it is both column-sufficient and row-sufficient.

Although any P-matrix is necessarily a sufficient S-matrix, the reciprocal is not true. For example, the sufficient S-matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

is not a P-matrix.

Using the definitions above we can prove a new sufficient condition that guarantees that a stationary point of $NCP(G)$ is a solution of the problem.

Theorem 4.2. *Assume that (y_*, z_*) is a stationary point of system (35) associated to NCP(G). If $G'(y_*)$ is a row-sufficient S-matrix then y_* is solution of $NCP(G)$.*
Proof. In this case $H(z) = -z$ and, in Theorem 4.1, $u_* = G(y_*) - z_*$. Then, by (43), we obtain

$$[u_*]_i[G'(y_*)^T u_*]_i \leq 0 \tag{47}$$

for all $i = 1, \ldots, q$. Since $G'(y_*)$ is row-sufficient, we have that

$$[u_*]_i[G'(y_*)^T u_*]_i = 0 \tag{48}$$

for all $i = 1, \ldots, q$. By (43), we obtain $[y_*]_i[z_*]_i = 0$ for all $i = 1, \ldots, q$, so

$$y_*^T z_* = 0. \tag{49}$$

Thus, by (38) and (39),

$$2G'(y_*)^T(-u_*) = -\mu_1 \leq 0, \tag{50}$$

$$-2u_* = \mu_2 \geq 0, \tag{51}$$

Suppose that $u_* \neq 0$, then by (50) and (51) $G'(y_*)$ could not be an S-matrix. Therefore $u_* = 0$. So, by (49), y_* is solution of $NCP(G)$. $\square$

As we mentioned above, the Horizontal Linear Complementarity Problem (HLCP) is the particular case of HNCP in which G and H are affine functions. So, it can be stated in the following way:

Find $y, z \in I\!R^q$, $y, z \geq 0$ such that

$$Qy + Rz = b, \quad y^T z = 0 \tag{52}$$

We are going to prove that the stationary points of (52) are solutions of the problem under different hypotheses than those stated for the general HNCP. This result generalizes the Equivalence Theorem proved in [12]. As in [12], we say that the HLCP defined by (52) is feasible if there exist $y, z \geq 0$ such that

$$Qy + Rz = b.$$

An additional definition concerning pairs of square matrices is necessary. (See [19].) Given $M, N \in I\!R^{q\times q}$, $A = (M, N) \in I\!R^{q\times 2q}$, we say that the pair (M, N) is row-sufficient if

$$\left[\begin{pmatrix} r \\ s \end{pmatrix} \in \text{ range } A^T, \quad r \circ s \leq 0\right]$$

implies that

$$r \circ s = 0.$$

Here $\circ$ denotes the Hadamard product; i.e., $x \circ y$ is the vector whose components are the products of the corresponding components of x and y.

Theorem 4.3. *Assume that the HLCP (52) is feasible, (y_*, z_*) is a stationary point and the pair $(Q, -R)$ is row-sufficient. Then, (y_*, z_*) is solution of the Horizontal Linear Complementarity Problem (52).*

Proof. As in Theorem 4.1, the minimization problem associated to (52) is:

$$\text{Minimize } \|Qy + Rz - b\|^2 + (y^T z)^2 \quad \text{subject to} \quad y, z \geq 0. \tag{53}$$

Define $G(y) = Qy - b$ and $H(z) = Rz$, then $G'(y) = Q$ and $H'(z) = R$. Calling

$$u_* = Qy_* + Rz_* - b, \tag{54}$$

the inequality (43) of Theorem 4.1 remains

$$4[R^T u_*]_i[Q^T u_*]_i = 4(y_*^T z_*)^2[y_*]_i[z_*]_i + [\mu_1]_i[\mu_2]_i \geq 0 \quad \text{for all } i = 1, \dots, q. \tag{55}$$

Since the pair $(Q, -R)$ is row-sufficient, whenever

$$[R^T u_*]_i[Q^T u_*]_i \geq 0 \ \text{ for all } \quad i = 1, \dots, q$$

one necessarily has that

$$[R^T u_*]_i[Q^T u_*]_i = 0 \ \text{ for all } \quad i = 1, \dots, q.$$

Therefore, by (55) we obtain:

$$[y_*]_i[z_*]_i = 0 \ \text{ for all } i = 1, \dots, q. \tag{56}$$

Now, considering (38)-(42) in Theorem 4.1, with the obvious specifications relative to the linear case, the first-order conditions of (53) turn out to be

$$2Q^T u_* - \mu_1 = 0, \tag{57}$$

$$2R^T u_* - \mu_2 = 0, \tag{58}$$

$$y_*^T \mu_1 = 0, \tag{59}$$

$$z_*^T \mu_2 = 0, \tag{60}$$

$$y_* \geq 0, \quad z_* \geq 0, \quad \mu_1 \geq 0, \quad \mu_2 \geq 0. \tag{61}$$

Now, (57)-(61), are necessary and sufficient conditions for global minimizers of the following convex quadratic minimization problem:

$$\text{Minimize } \|Qy + Rz - b\|^2 \quad \text{subject to} \quad y \geq 0, \ z \geq 0. \tag{62}$$

So, by the feasibility of the HLCP, it turns out that (y_*, z_*) is global solution of (62) with minimun value zero, that is

$$Qy_* + Rz_* - b = 0. \tag{63}$$

Therefore, by (56) and (63), (y_*, z_*) is a solution of the Horizontal Linear Complementarity Problem. □

In Section 3 we showed that the key condition that allows us to prove superlinear convergence of some variations of Algorithm 2.1 is the Strong Local Unicity Assumption (SLCU). Therefore, it is interesting to study sufficient conditions under which this assumption holds in the case of the Horizontal Nonlinear Complementarity Problem. Probably, when these sufficient conditions are fulfilled, any reasonable method will converge quickly, if started from a good enough initial approximation. On the other hand, difficult problems will be probably characterized by situations where those sufficient conditions are not true. When applied to the Nonlinear Complementarity Problem, the condition introduced in Theorem 4.4 corresponds to conditions given by Mangasarian [18] for the uniqueness of the solution of NCP and Moré [20] for superlinear convergence of a specific NCP algorithm.

Given the set of indices $\mathcal{L} \subset \{1, \ldots, q\}$ and $\mathcal{K} \subset \{1, \ldots, q\}$ and the matrix $A \in I\!R^{q \times q}$, we denote $A_{\mathcal{L},\mathcal{K}}$ the submatrix of A such that $A_{\mathcal{L},\mathcal{K}} \equiv (a_{i,j})$ with $i \in \mathcal{L}$ and $j \in \mathcal{K}$.

Theorem 4.4. *Assume that (y_*, z_*) is a solution of the HNCP(G,H) such that*

$$[y_*]_i + [z_*]_i > 0 \quad \textit{for all} \quad i = 1, \ldots, q. \tag{64}$$

Moreover, suppose that the matrix formed by the columns of $G'(y_*)$ that correspond to $[y_*]_i > 0$ and the columns of $H'(z_*)$ that correspond to $[z_*]_i > 0$

is nonsingular. Then (y_*, z_*) satisfies Assumption 3.1.

Proof. Let us call

$$\mathcal{J} = \{i \in \{1, \ldots, q\} \mid [y_*]_i > 0\}, \quad \mathcal{M} = \{i \in \{1, \ldots, q\} \mid [z_*]_i > 0\}.$$

Since (y_*, z_*) is solution of $HNCP(G, H)$, by (64) we obtain

$$[y_*]_{\mathcal{J}} > 0, \;\; [y_*]_{\mathcal{M}} = 0, \;\; [z_*]_{\mathcal{J}} = 0, \;\; [z_*]_{\mathcal{M}} > 0,$$

$$\mathcal{I} = \{1, \ldots, q\} = \mathcal{J} \cup \mathcal{M} \;\text{ and }\; \mathcal{J} \cap \mathcal{M} = \emptyset \tag{65}$$

Assume that $F(y, z)$ does not verify the strong unicity condition at (y_*, z_*), then by Lemma 3.6 there exists $\tau = (\tilde{y}, \tilde{z}) \in I\!R^{2q}$, $\|\tau\| = 1$ and $\{(\tilde{y}_k, \tilde{z}_k)\} \subset \Omega = I\!R^{2q}_+$ such that

$$(\tilde{y}_k, \tilde{z}_k) \to (y_*, z_*),$$

$$\tau = \begin{pmatrix} \tilde{y} \\ \tilde{z} \end{pmatrix} = \lim_{k\to\infty} \frac{\begin{pmatrix} \tilde{y}_k \\ \tilde{z}_k \end{pmatrix} - \begin{pmatrix} y_* \\ z_* \end{pmatrix}}{\left\| \begin{pmatrix} \tilde{y}_k \\ \tilde{z}_k \end{pmatrix} - \begin{pmatrix} y_* \\ z_* \end{pmatrix} \right\|} \tag{66}$$

and

$$F'(y_*, z_*)\tau =$$

$$\begin{bmatrix} [G'(y_*)]_{\mathcal{I},\mathcal{J}} & [G'(y_*)]_{\mathcal{I},\mathcal{M}} & [H'(z_*)]_{\mathcal{I},\mathcal{J}} & [H'(z_*)]_{\mathcal{I},\mathcal{M}} \\ [z_*]_{\mathcal{J}}^T & [z_*]_{\mathcal{M}}^T & [y_*]_{\mathcal{J}}^T & [y_*]_{\mathcal{M}}^T \end{bmatrix} \begin{bmatrix} \tilde{y}_{\mathcal{J}} \\ \tilde{y}_{\mathcal{M}} \\ \tilde{z}_{\mathcal{J}} \\ \tilde{z}_{\mathcal{M}} \end{bmatrix} = 0. \tag{67}$$

By (65) and (67), we have that

$$[z_*]_{\mathcal{M}}^T \tilde{y}_{\mathcal{M}} + [y_*]_{\mathcal{J}}^T \tilde{z}_{\mathcal{J}} = 0. \tag{68}$$

Now, by (65) and (66), calling

$$\pi_k = \left\| \begin{pmatrix} \tilde{y}_k \\ \tilde{z}_k \end{pmatrix} - \begin{pmatrix} y_* \\ z_* \end{pmatrix} \right\|,$$

we obtain

$$\tilde{y}_{\mathcal{M}} = \lim_{k\to\infty} \frac{[\tilde{y}_k]_{\mathcal{M}}}{\pi_k} \geq 0 \;\text{ and }\; \tilde{z}_{\mathcal{J}} = \lim_{k\to\infty} \frac{[\tilde{z}_k]_{\mathcal{J}}}{\pi_k} \geq 0. \tag{69}$$

Then, by (65), (68) and (69), we have that

$$\tilde{y}_{\mathcal{M}} = 0 \;\text{ and }\; \tilde{z}_{\mathcal{J}} = 0. \tag{70}$$

Therefore, by (67),

$$\left[\begin{array}{cc} [G'(y_*)]_{\mathcal{I},\mathcal{J}} & [H'(z_*)]_{\mathcal{I},\mathcal{M}} \end{array}\right] \left[\begin{array}{c} \tilde{y}_{\mathcal{J}} \\ \tilde{z}_{\mathcal{M}} \end{array}\right] = 0. \tag{71}$$

By the hypothesis of nonsingularity, we obtain

$$\tilde{y}_{\mathcal{J}} = 0 \text{ and } \tilde{z}_{\mathcal{M}} = 0. \tag{72}$$

By (70) and (72), we have that $\tau = 0$, which is a contradiction. □

5 Numerical Implementation

In this section we describe a computer implementation of Algorithm 2.2. The computation of the direction d_k satisfying (9) takes into account, as mentioned in Section 2, the box-constrained quadratic subproblem (8), with a first stopping criterion given by (10). The second stopping criterion used in the computer implementation was

$$\begin{array}{c} \|F'(x_k)\bar{d} + F(x_k)\| \le \theta_k \|F(x_k)\| \\ \text{or} \\ \|P(\bar{d} - \nabla\Phi(\bar{d})) - \bar{d}\| \le 10^{-12} \|P(-\nabla\Phi(0))\| \end{array} \tag{73}$$

This means that at Step 2 of Algorithm 2.2 we run the quadratic solver with the stopping criterion (10), obtaining the increment $\bar{d}$. If $\|F'(x_k)\bar{d} + F(x_k)\| \le \theta_k \|F(x_k)\|$ we accept $d_k = \bar{d}$. Otherwise, we continue the execution of the quadratic solver with the stopping criterion (73).

We wish to deal with large-scale problems. So, we used, for solving (8), the algorithm introduced in [10] and improved in [1]. This algorithm does not use matrix factorizations at all. In fact, conjugate gradient iterations are used within the faces of the box and, when the current face needs to be abandoned, orthogonal chopped-gradient directions are employed (see [1, 10, 11]).

The algorithmic parameters used in the experiments were: $\Delta = 100$, $\bar{\rho} = 8$, $\sigma = 10^{-4}$ and $\theta_0 = 0.9999$. Practical convergence to a solution was declared when $\|F(x_k)\| \le 10^{-8}$. A current iterate was declared stationary if the choice (9) is not possible with $\theta_k = 0.999995$.

We wish to test the efficiency of different strategies for the choice of η_k. The tested strategies were:

(1) $\eta_k = 0.1$ for all $k = 0, 1, 2, \ldots$

(2) $\eta_k = 10^{-3}$ for all $k = 0, 1, 2, \ldots$

(3) $\eta_k = 10^{-8}$ for all $k = 0, 1, 2, \ldots$

(4) $\eta_k = 1/(k+2)$ for all $k = 0, 1, 2, \ldots$
(5) An adaptive strategy given by

$$\eta_0 = 0.1, \quad \eta_k = (\|F(x_k)\|/\|F(x_{k-1}\|)^2 \quad \text{if} \quad \eta_{k-1}^2 \leq 0.1,$$

$$\text{and} \quad \eta_k = \eta_{k-1}^2 \quad \text{if} \quad \eta_{k-1}^2 > 0.1.$$

The strategy (5) was inspired on a choice suggested in [8] for the forcing term of Inexact-Newton methods for solving nonlinear systems.

We generated a set of test problems with known solution defining

$$[y_*]_i = 1 \quad \text{if} \quad i \quad \text{is odd, and} \quad [y_*]_i = 0 \quad \text{otherwise,} \tag{74}$$

and

$$[z_*]_i = 1 \quad \text{if} \quad i \quad \text{is even, and} \quad [z_*]_i = 0 \quad \text{otherwise.} \tag{75}$$

The following nonlinear mappings $T_i : I\!R^q \to I\!R^q$, $i \in \{1, 2, 3, -1, -2, -3\}$, were considered:

$$[T_1(x)]_i = 10(x_{i+1}^2 - x_i) \quad \text{if} \quad i \quad \text{is odd,}$$

$$[T_1(x)]_i = x_{i-1} - 1 \quad \text{if} \quad i \quad \text{is even,}$$

$$[T_2(x)]_1 = 2x_1 + e^{x_1} + x_2,$$

$$[T_2(x)]_i = -x_{i-1} + 2x_i + e^{x_1} + x_{i+1}, i = 2, \ldots, q-1,$$

$$[T_2(x)]_q = -x_{q-1} + 2x_q + e^{x_q},$$

$$T_3(x) = x,$$

$$T_{-i}(x) = -T_i(x), i = 1, 2, 3.$$

We generated 12 problems $P_{ij}, (i, j) \in \mathcal{P}$, where

$$\mathcal{P} = \{(i, j) \ | \ i = 1, 2, 3, \ \ |j| \leq 3, |j| \leq |i|\}.$$

in the following way:

For each $(i, j) \in \mathcal{P}$, we computed $b = T_i(y_*) + T_j(z_*)$ and we defined $G(y) = T_i(y) - b/2$ and $H(y) = T_j(y) - b/2$. So, (y_*, z_*) is a solution of the Horizontal Nonlinear Complementarity Problem P_{ij}, defined by G and H.

We tried to solve the problems using three initial points:
(1) $x_0 = x_* + (\underline{2} - x_*)$;
(2) $x_0 = x_* + 10(\underline{2} - x_*)$;
(3) $x_0 = x_* + 100(\underline{2} - x_*)$;
where $\underline{2} = (2, \ldots, 2)^T$.

Problems P_{ij} with $|i| = 2$ or $|j| = 2$ cannot be run for the third initial point because an overflow is produced in the computation of the nonlinear system at x_0. Therefore, 30 different problems were generated. We tested the five strategies for η_k described above for each of these problems, using $q = 500$ (so $n = 1000$, $m = 501$).
The solution points of these problems are nondegenerate in the sense that $[y_*]_i + [z_*]_i > 0$ for all $i = 1, \ldots q$. With the aim of investigating the local behavior of the algorithm in degenerate problems we defined an additional set of problems $P'_{ij}, (i,j) \in \mathcal{P}$ using

$$[y_*]_i = 1 \quad \text{if} \quad i \leq q/2,\ i \text{ odd, and} \quad [y_*]_i = 0 \quad \text{otherwise,} \tag{76}$$

and

$$[z_*]_i = 1 \quad \text{if} \quad i \leq q/2,\ i \text{ even, and} \quad [z_*]_i = 0 \quad \text{otherwise.} \tag{77}$$

Since our theoretical results show that, in this case, problems may arise with the local convergence, we used only the initial approximation $x_0 = x_* + 0.1(\underline{2} - x_*)$.

Our codes were written in Fortran 77 with double precision, we used the DOS-Microsoft compiler, and the tests were run in a Pentium-166 MHz. The complete numerical results, in six different tables, and the source codes are available from the authors. Here we are only going to state the conclusions of the numerical study.

Surprisingly, the behavior of different accuracy strategies in the 30 nondegenerate problems is remarkably similar, with one exception: the performance of Strategy 3 in problem $(1, -2)$ with the second initial point is much poorer than the performance of the other strategies. In this problem, Strategy 5 was more efficient than the others. Roughly speaking, with all the strategies we obtained solutions of the problems, except in three cases, in which we obtained stationary points. In most cases, the computer time was very moderate, in spite of the modest computer environment used.

As expected, problems with degenerate solution were, in general, more difficult than the nondegenerate ones. Although no meaningful differences were detected regarding the performance of the three strategies tested, we verified that in two problems ($(1, -2)$ and $(1, -3)$) 100 iterations were not enough for achieving convergence. In four problems convergence was achieved but the distance between the approximate solution and the true one was greater than 10^{-6}. Finally, in general, the computer time used for solving these degenerate problems with initial point close to the solution is of the same order (if not greater) than the computer time used to solve nondegenerate problems starting far from the solution.

6 Conclusions

The method presented in this paper is able to deal with large-scale box-constrained, non-necessarily square, nonlinear systems of equations. Usually, these problems are treated as ordinary large-scale minimization problems in practical optimization. See, for example, [3, 11] and the CUTE collection of problems of Conn, Gould and Toint. We think that to consider the specific nonlinear-system structure can represent a practical advantage. On one hand, the sufficient decrease condition (3) on the model imposes a residual decrease proportional to the operator norm, and not only to some projected gradient measure, as a minimization algorithm would do. This probably implies that the nonlinear-system algorithm tends to be more active trying to find roots of the system, and less prone to converge to local nonglobal minimizers. On the other hand, high order convergence can be obtained using only first order information. In [17] a similar idea for square systems was applied successfully to turning point problems, with a different concept of near-stationarity.

Inexact-Newton methods using conditions similar to (3) are well developed and largely used for solving large-scale (unconstrained) nonlinear systems. The contribution of this paper is a natural extension of those methods to constrained problems.

The essential tool for computing the search direction at each iteration is a box-constrained quadratic solver that uses conjugate gradient directions inside the faces of the box and theoretically justified procedures for leaving the faces. Most of the computer time used by the algorithm is spent in the computation of these directions, at least in the numerical examples tested.

We used the new method for solving a set of Horizontal Nonlinear Complementarity Problems (HNCP), which are generalizations of Nonlinear Complementarity Problems (NCP) and Horizontal Linear Complementarity Problems (HLCP). We proved that, in many cases, stationary points of the HNCP are global solutions. These results generalize well-known existing results for the HLCP and the NCP.

We tried five different strategies related to the accuracy required in the solution of the quadratic subproblem at each iteration. Theoretical results concerning order of convergence seem to suggest that degenerate problems are more difficult than problems with nondegenerate solution, at least from the local convergence point of view. This theoretical prediction was confirmed by the experiments.

References

[1] R. H. Bielschowsky, A. Friedlander, F. M. Gomes, J. M. Martínez and M. Raydan [1995], An adaptive algorithm for bound constrained quadratic minimization, to appear in *Investigación Operativa.*

[2] J. F. Bonnans and C. C. Gonzaga [1996], Convergence of interior point algorithms for the monotone linear complementarity problem, *Mathematics of Operations Research* 21, pp. 1-25

[3] Conn, A. R.; Gould, N. I. M.; Toint, Ph. L. [1988], Global convergence of a class of trust region algorithms for optimization with simple bounds, *SIAM Journal on Numerical Analysis* 25, pp. 433 - 460. See also *SIAM Journal on Numerical Analysis* 26 [1989] pp. 764 - 767.

[4] R. W. Cottle, J-S. Pang and R. E. Stone [1992], The linear complementarity problem, Academic Press, Boston.

[5] T. De Luca, F. Facchinei and C. Kanzow [1994], A semismooth equation approach to the solution of nonlinear complementarity problem, *Mathematical Programming* 75, pp. 407-439.

[6] R. S. Dembo, S. C. Eisenstat and T. Steihaug [1982], Inexact Newton methods, *SIAM Journal on Numerical Analysis* 14, pp. 400-408.

[7] S. C. Eisenstat and H. F. Walker [1994], Globally convergent inexact Newton methods, *SIAM Journal on Optimization* 4, pp. 393-422.

[8] S. C. Eisenstat and H. F. Walker [1996], Choosing the forcing terms in an Inexact Newton method, *SIAM Journal on Scientific Computing* 17, pp. 16-32.

[9] A. Fischer [1994], An NCP-function and its use for the solution of complementarity problems, Technical Report MATH-NM-12-1994, Technische Universität Dresden, Dresden, Germany.

[10] A. Friedlander and J. M. Martínez [1994], On the maximization of a concave quadratic function with box constraints, *SIAM Journal on Optimization* 4, pp. 177-192.

[11] A. Friedlander, J. M. Martínez and S. A. Santos [1994], A new trust region algorithm for bound constrained minimization, *Applied Mathematics and Optimization* 30, pp. 235-266.

[12] A. Friedlander, J. M. Martínez and S. A. Santos [1995], Solution of linear complementarity problems using minimization with simple bounds, *Journal of Global Optimization* 6, pp. 1-15.

[13] M. Friedler and V. Pták [1966], Some generalizations of positive definiteness and monotonicity. *Numerische Mathematik* 9, pp. 163-172.

[14] S. A. Gabriel and J-S. Pang [1992], An inexact NE/SQP method for solving the nonlinear complementarity problem, *Computational Optimization and Applications* 1, pp. 67-91.

[15] H. Jiang and L. Qi [1994], A new nonsmooth equations approach to nonlinear complementarities, *Applied Mathematics Report, AMR94/31*, School of Mathematics, The University of South Wales, Sydney Australia.

[16] J. J. Júdice [1994], Algorithms for linear complementarity problems, in *Algorithms for Continuous Optimization*, edited by E. Spedicato, Kluwer, pp. 435-474.

[17] D. N. Kozakevich, J. M. Martínez and S. A. Santos [1996], Solving nonlinear systems of equations with simple constraints, to appear in *Computational and Applied Mathematics.*

[18] O. L. Mangasarian [1980], Locally unique solutions of quadratic programs, linear and nonlinear complementarity problems, *Mathematical Programming* 19, pp. 200-212.

[19] O. L. Mangasarian and J-S. Pang[1995], The extended linear complementarity problem, *SIAM Journal of Matrix Analysis and Applications*, 16, pp. 359-368.

[20] J. J. Moré [1996], Global methods for nonlinear complementarity problems, *Mathematics of Operations Research*, 21, pp. 589-614.

[21] N. Yamashita and M. Fukushima [1995], On stationary points of the implicit Lagrangian for the nonlinear complementarity problems, *Journal of Optimization Theory and Applications* 84, pp. 653-663.