Hamburger Beiträge
zur Angewandten Mathematik

An Active Set-type Newton Method for Constrained Nonlinear Systems

Christian Kanzow

Reihe A
Preprint 147
September 1999

Hamburger Beiträge zur Angewandten Mathematik

Reihe A	Preprints
Reihe B	Berichte
Reihe C	Mathematische Modelle und Simulation
Reihe D	Elektrische Netzwerke und Bauelemente
Reihe E	Scientific Computing
Reihe F	Computational Fluid Dynamics and Data Analysis

AN ACTIVE SET-TYPE NEWTON METHOD FOR CONSTRAINED NONLINEAR SYSTEMS

Christian Kanzow[1]

University of Hamburg
Institute of Applied Mathematics
Bundesstrasse 55
D-20146 Hamburg
Germany
e-mail: kanzow@math.uni-hamburg.de

September 23, 1999

Abstract. We consider the problem of finding a solution of a nonlinear system of equations subject to some box constraints. To this end, we introduce a new active set-type Newton method with global and local fast convergence properties. The method generates feasible iterates and has to solve only one linear system of equations at each iteration. Due to our active set strategy, this linear system is of reduced dimension.

Key Words. Nonlinear equations, box constraints, Newton's method, active set strategy, projected gradient, global convergence, quadratic convergence.

[1]The research of this author was partially supported by the DFG (Deutsche Forschungsgemeinschaft).

1 Introduction

The problem we address in this paper is to find a solution of the constrained nonlinear system

$$F(x) = 0, \quad x \in [l, u], \tag{1}$$

where $F : [l, u] \to \mathbb{R}^n$ is a given function which is assumed to be continuously differentiable in an open set containing the box $[l, u]$, and where

$$l = (l_1, \ldots, l_n)^T, \quad u = (u_1, \ldots, u_n)^T$$

with $l_i \in \mathbb{R} \cup \{-\infty\}$ and $u_i \in \mathbb{R} \cup \{+\infty\}$ for $i = 1, \ldots, n$ are given lower and upper bounds satisfying $l_i < u_i$ for all $i = 1, \ldots, n$.

Our motivation for studying this type of problem comes from different areas:

1. There are a couple of different mathematical programming problems like Karush-Kuhn-Tucker systems or complementarity problems which immediately result into a problem of the kind (1), see [20] for some more details.

2. The component functions F_i of the mapping F in (1) are not always defined on the whole space $\mathbb{R}^n$, so we are urged to restrict the domain of F, and this can usually be done by putting suitable bounds on some or all of the variables.

3. In a couple of applications, one already has an idea to which area a solution of an unconstrained nonlinear system $F(x) = 0$ belongs to; in this case, it might be very helpful to have the possibility to describe this area by some lower and upper bounds on the variables.

In view of these comments, we think that a constrained nonlinear system of the type (1) is a very important problem. Of course, it is possible to solve this problem by reformulating it either as a box constrained optimization problem or as a mixed complementarity problem and then applying some standard software to this reformulated problem. However, such an approach does not seem to take into account the particular structure of the constrained nonlinear system (1) and, therefore, does not seem to be an adequate way for solving it.

Hence we believe that it is very important to design algorithms for problem (1) which exploit the special properties of this system. Surprisingly, it seems that there is not a rich literature on this topic. In fact, we are only aware of the papers [13, 20, 19, 18]. In [13], Kozakevich, Martínez and Santos describe an inexact Gauss-Newton-type method for constrained nonlinear systems of equations, while Wang, Monteiro and Pang design a potential reduction-type interior-point method in [20]. Both references allow more general constraints than box constraints, but are completely different from our method. Somewhat closer to our approch is the trust-region framework used by Ulbrich [19] as well as the Newton-type method presented by Sun [18] on a recent conference.

However, the method we are going to present here is the only one using an active-set strategy. As a consequence, we only have to deal with linear systems of reduced dimension. This can be a significant advantage for large-scale problems since reduced problems can be solved much more efficiently than full dimensional systems.

Note that, in contrast to some of the above mentioned papers, we only deal with nonlinear systems having box constraints here. This is partially due to our motivation (the problems discussed there can usually be solved by introducing bounds on the variables) and partially due to the way we will try to solve the constrained nonlinear system (1) (e.g., we will compute projections onto the feasible set which is a trivial task for box constraints but becomes a much more difficult task for more general constraints). However, we stress that it is easily possible to deal also with, e.g., general inequality constraints. To this end, one just has to add a nonnegative slack variable in order to get a problem of the form (1) with one more equation and one more variable.

The algorithm which we will develop in this paper has a number of desirable properties: First, it generates feasible iterates only. This is highly important in view of the motivation we gave directly after the statement of problem (1). Second, it solves at each iteration only one linear system, and this linear system is actually of reduced dimension. Finally, the algorithm is globally and, under certain assumptions, locally superlinearly/quadratically convergent.

Basically, the idea of our method is to solve the unconstrained nonlinear system $F(x) = 0$ by using Newton's method. We then use some recent ideas from the identification of active constraints in order to take into account the box constraints. This, more or less, gives us a feasible and locally fast convergent method. In order to make it globally convergent, we will combine it in a simple way with a projected gradient method.

The organization of this paper is as follows: Section 2 gives a detailed statement of our algorithm and contains further motivation for the design of this method. In Section 3, we then investigate the global and local convergence properties of our algorithm. We conclude this paper with some final remarks in Section 4.

Notation: If $x \in \mathbb{R}^n$ is any given vector with components x_i and $\mathcal{K} \subseteq \{1, \dots, n\}$ is any given subset, we write $x_{\mathcal{K}}$ for the subvector having components $x_i, i \in \mathcal{K}$. Similarly, if $A \in \mathbb{R}^{n \times n}$ is a given matrix with entries a_{ij} and $\mathcal{K}, \mathcal{L} \subseteq \{1, \dots, n\}$ are two subsets, we denote by $A_{\mathcal{K}\mathcal{L}}$ the $|\mathcal{K}| \times |\mathcal{L}|$ submatrix with elements $a_{ij}, i \in \mathcal{K}, j \in \mathcal{L}$. The Jacobian of our mapping $F : [l, u] \to \mathbb{R}^n$ is denoted by $F'(x)$; F_i is the ith component function of F, and $\nabla F_i(x)$ is the gradient of F_i evaluated at the point x. Note that, throughout this manuscript, we view all gradient vectors as column vectors. Hence, the rows of the Jacobian $F'(x)$ are the transposed gradients of the component functions. Furthermore, we denote by $\| \cdot \|$ the Euclidean vector norm or any consistent matrix norm so that $\|Ax\| \leq \|A\| \, \|x\|$ for any matrix A and any vector x such that the matrix-vector product Ax is defined. Finally, the projection of a vector $y \in \mathbb{R}^n$ on the feasible set $[l, u]$ is denoted by y_+.

2 Algorithm

In this section, we give a detailed description of our algorithm for the solution of the constrained nonlinear system (1). Basically, our algorithm is a Newton-type method for the solution of the nonlinear system

$$F(x) = 0$$

which incorporates an active-set strategy in order to stay feasible with respect to the box $[l, u]$ and which measures progress by minimizing the related optimization problem

$$\min f(x) \quad \text{s.t.} \quad x \in [l, u], \tag{2}$$

where, throughout this paper, $f : [l, u] \to \mathbb{R}$ is defined by

$$f(x) := \frac{1}{2} F(x)^T F(x) = \frac{1}{2} \|F(x)\|^2.$$

The precise statement of our algorithm is as follows.

Algorithm 2.1 *(Active Set-type Newton Method)*

(S.0) *(Initialization)*
Choose $x^0 \in [l, u], \beta \in (0,1), \gamma \in (0,1), s > 0, \delta > 0, c > 0,$ *and set* $k := 0$.

(S.1) *(Termination Criterion)*
If x^k *is a stationary point of (2): STOP.*

(S.2) *(Active Set Strategy)*
Define

$$\delta_k := \min\{\delta, c\sqrt{\|F(x^k)\|}\}$$

and

$$\begin{aligned} \mathcal{A}_k &:= \{i \,|\, x_i^k - l_i \le \delta_k \text{ or } u_i - x_i^k \le \delta_k\}, \\ \mathcal{I}_k &:= \{1, \ldots, n\} \setminus \mathcal{A}_k. \end{aligned}$$

(S.3) *(Search Direction Calculation)*
Try to compute a vector $d^k \in \mathbb{R}^n$ *in the following way: For* $i \in \mathcal{A}_k$, *set*

$$d_i^k := \begin{cases} l_i - x_i^k & \text{if } x_i^k - l_i \le \delta_k, \\ u_i - x_i^k & \text{if } u_i - x_i^k \le \delta_k, \end{cases}$$

then solve the linear system

$$F'(x^k)_{\mathcal{I}_k \mathcal{I}_k} d_{\mathcal{I}_k} = -F(x^k)_{\mathcal{I}_k} - F'(x^k)_{\mathcal{I}_k \mathcal{A}_k} d_{\mathcal{A}_k}^k \tag{3}$$

in order to get the components d_i^k *for* $i \in \mathcal{I}_k$.

(S.4) *(Feasibility Preserving Safeguard)*
Compute

$$\begin{aligned} \bar{\tau}_k &:= \sup\{\tau > 0 \,|\, x_i^k + \tau d_i^k \in [l_i, u_i] \ \ \forall i \in \mathcal{I}_k\}, \\ \tau_k &:= \min\{1, \bar{\tau}_k\}. \end{aligned}$$

(S.5) (Computation of New Iterate)
If d^k could have been computed in Step (S.3) and satisfies the descent condition

$$f(x^k + \tau_k d^k) \le \gamma f(x^k), \tag{4}$$

set $x^{k+1} := x^k + \tau_k d^k$; otherwise use a projected gradient step, i.e., set $x^{k+1} := x^k(t_k)$, where $x^k(t) := [x^k - t\nabla f(x^k)]_+$ and $t_k = \max\{s\beta^\ell \mid \ell = 0, 1, 2, \dots\}$ such that

$$f(x^k(t_k)) \le f(x^k) - \sigma \nabla f(x^k)^T (x^k - x^k(t_k)).$$

(S.6) (Update)
Set $k \leftarrow k+1$, and go to (S.1).

Throughout this paper, we assume that the constant $\delta > 0$ is chosen sufficiently small such that

$$\delta < \frac{1}{2} \min_{i=1,\dots,n} |u_i - l_i|.$$

This, in particular, implies that we cannot have $x_i^k - l_i \le \delta_k$ and $u_i - x_i^k \le \delta_k$ for the same index $i \in \mathcal{A}_k$ since this would imply

$$|u_i - l_i| = u_i - l_i = (u_i - x_i^k) + (x_i^k - l_i) \le 2\delta_k \le 2\delta,$$

a contradiction to the choice of δ. Hence, there is no ambiguity in the definition of the components d_i^k for $i \in \mathcal{A}_k$ in Step (S.3) of Algorithm 2.1.

Let us now give some further comments on each step of Algorithm 2.1 in order to get a better understanding of the method.

Apart from some parameters which have an obvious meaning, we only assume in Step (S.0) that we have a feasible starting point x^0. Of course, finding such a starting point is a trivial task due to our simple constraints. Step (S.1) contains a standard termination criterion (which will be modified in our actual implementation).

In Step (S.2), we define a set $\mathcal{A}_k$ which is viewed as a kind of active set: It contains all those indices which, at a current iterate x^k, are close to the boundary (but not necessarily equal to one of the bounds). Later, we will see that the set $\mathcal{A}_k$ is eventually equal to the set of active constraints at a solution x^* of the constrained nonlinear system (1) if this solution satisfies an appropriate regularity condition. For some further motivation regarding the definition of $\mathcal{A}_k$, we refer the interested reader to the paper [6].

In Step (S.3) of Algorithm 2.1, we then try to compute a Newton-type search direction d^k. For the components d_i^k with $i \in \mathcal{A}_k$, we use a simple formula whose aim is, basically, to push the corresponding components of our iterates to come closer to the boundary. This makes sense since, due to our current information, we believe that these components will be active at a solution. In order to understand the formula for the computation of the components d_i^k for $i \in \mathcal{I}_k$, note that, after a possible permutation of the rows and columns, we can rewrite the standard (unconstrained) Newton equation $F'(x^k)d = -F(x^k)$ as

$$\begin{pmatrix} F'(x^k)_{\mathcal{I}_k\mathcal{I}_k} & F'(x^k)_{\mathcal{I}_k\mathcal{A}_k} \\ F'(x^k)_{\mathcal{A}_k\mathcal{I}_k} & F'(x^k)_{\mathcal{A}_k\mathcal{A}_k} \end{pmatrix} \begin{pmatrix} d_{\mathcal{I}_k} \\ d_{\mathcal{A}_k} \end{pmatrix} = - \begin{pmatrix} F(x^k)_{\mathcal{I}_k} \\ F(x^k)_{\mathcal{A}_k} \end{pmatrix}. \tag{5}$$

Now it is easy to see that the linear system (3) is nothing else but the first block row in (5) (recall that we already have an expression for $d^k_{\mathcal{A}_k}$). Note that our Newton-type search direction is somewhat different from those suggested in some projected Newton methods for constrained optimization problems in, e.g., [1, 14].

In Step (S.4), we compute the largest possible step $\tau \in [0,1]$ such that the point $x^k + \tau d^k$ would still be feasible. Note that there is a simple closed-form expression for the value of $\tau = \tau_k$, namely

$$\tau_k = \min\{1, \bar{\tau}_k\}$$

with

$$\bar{\tau}_k := \min \left\{ \min_{\{i \in \mathcal{I}_k \,|\, d_i^k < 0\}} \frac{l_i - x_i^k}{d_i^k}, \min_{\{i \in \mathcal{I}_k \,|\, d_i^k > 0\}} \frac{u_i - x_i^k}{d_i^k} \right\}.$$

Here, we assume that $\bar{\tau}_k = +\infty$ if the minimum is taken over the empty set.

Step (S.5) of Algorithm 2.1 then tries to accept the point $x^k + \tau_k d^k$ as the new iterate x^{k+1} if this gives a sufficient decrease in the objective function f in the sense that condition (4) is satisfied. From a practical point of view, this test is somewhat restrictive and will be replaced by some kind of arc search, i.e., in practice we will try to find a stepsize $t_k \in [0,1]$ which minimizes the value of f along the projected arc $[x^k + td^k]_+$ approximately (note that this is a piecewise linear arc). As long as the step t_k is chosen in such a way that

$$f([x^k + t_k d^k]_+) \le f(x^k + \tau_k d^k),$$

this modification does not destroy the global and local convergence properties of Algorithm 2.1. If, however, the test (4) (or its modification just described) is not satisfied either because the search direction d^k is a bad direction or because we were not able to compute d^k in Step (S.3), we switch to a projected gradient step, cf. [10, 3, 2] for further details. (Note that the projections onto $[l,u]$ can be carried out trivially in our situation.)

However, the projected gradient step is there not only as a safeguard. It is also used in order to change the active set $\mathcal{A}_k$. In fact, this set can change rapidly when using a projected gradient step, i.e., $\mathcal{A}_{k+1}$ can differ from $\mathcal{A}_k$ by more than just one element in contrast to many classical active set methods in constrained optimization [9].

On the other hand, it is also possible that the set $\mathcal{A}_k$ changes when using a Newton-type step in Algorithm 2.1. This is because $\mathcal{A}_k$ contains also those indices which are close to the boundary but not necessarily on the boundary. In fact, this is a major difference to some other active set Newton-type methods for constrained optimization, see, e.g., [1, 14] and references therein.

3 Convergence

In this section, we discuss the theoretical properties of Algorithm 2.1. Our first result shows that Algorithm 2.1 generates only feasible iterates.

Lemma 3.1 *Any sequence $\{x^k\}$ generated by Algorithm 2.1 belongs to the feasible set $[l,u]$.*

Proof. The proof is by induction on k: The starting point x^0 was chosen from $[l, u]$ in Step (S.0), so assume that x^k belongs to $[l, u]$ for some $k \geq 0$. If x^{k+1} is computed by using a projected gradient step at the kth iteration, then $x^{k+1} \in [l, u]$ is obvious. Otherwise we have $x^{k+1} = x^k + \tau_k d^k$ with $\tau_k \in (0, 1]$ being computed in Step (S.4) of Algorithm 2.1. Since $x_i^{k+1} \in [l_i, u_i]$ for all $i \in \mathcal{I}_k$ in view of the definition of τ_k, it remains to consider an index $i \in \mathcal{A}_k$. Then $x_i^k - l_i \leq \delta_k$ or $u_i - x_i^k \leq \delta_k$.

If $x_i^k - l_i \leq \delta_k$, we have $d_i^k = l_i - x_i^k$ and therefore

$$x_i^{k+1} = x_i^k + \tau_k d_i^k = \tau_k l_i + (1 - \tau_k) x_i^k.$$

Since this is a convex combination of $l_i, x_i^k \in [l_i, u_i]$, it follows that $x_i^{k+1} \in [l_i, u_i]$. Similarly, we can show that $x_i^{k+1} \in [l_i, u_i]$ if $u_i - x_i^k \leq \delta_k$. □

The global convergence properties of Algorithm 2.1 are discussed in our next result.

Theorem 3.2 *Any accumulation point of a sequence $\{x^k\}$ generated by Algorithm 2.1 is at least a stationary point of the equivalent minimization problem (2).*

Proof. Let x^* be an accumulation point of a sequence $\{x^k\}$ generated by Algorithm 2.1. If condition (4) is satisfied infinitely many times, it follows immediately that $f(x^*) = 0$, i.e., x^* is a global minimum of f and therefore, in particular, a stationary point of (2).

On the other hand, if condition (4) holds for only finitely many iterations, we eventually take projected gradient steps only, and in this case the assertion follows from, e.g., Proposition 2.3.3 (b) in Bertsekas [2]. □

Let us denote by

$$\mathcal{A}_* := \{i \,|\, x_i^* \in \{l_i, u_i\}\}$$

the set of active constraints at a given solution x^* of our constrained nonlinear system (1). In analogy to our notation used for the index set $\mathcal{A}_k$, we denote by $\mathcal{I}_*$ the complement of $\mathcal{A}_*$ in $\{1, \dots, n\}$:

$$\mathcal{I}_* := \{1, \dots, n\} \setminus \mathcal{A}_* = \{i \,|\, x_i^* \in (l_i, u_i)\}.$$

We next show that $\mathcal{A}_k$ is always a subset of $\mathcal{A}_*$ if x^k is sufficiently close to a solution x^* of the constrained nonlinear system (1). Later, we will show that these two sets actually coincide under a suitable regularity assumption, cf. Lemma 3.10.

Lemma 3.3 *Let x^* be a solution of the constrained nonlinear system (1). Then $\mathcal{A}_k \subseteq \mathcal{A}_*$ for all $x^k \in [l, u]$ sufficiently close to x^*.*

Proof. Define

$$\nu := \min\{\min\{x_i^* - l_i, u_i - x_i^*\} \,|\, i \in \mathcal{I}_*\} > 0,$$

i.e., ν is the smallest distance of the inactive components x_i^* to the boundary of the feasible set $[l, u]$. Let $x^k \in [l, u]$ be sufficiently close to x^* such that

$$\|x^k - x^*\| \leq \frac{\nu}{4}$$

and

$$c\sqrt{\|F(x^k)\|} \le \frac{\nu}{4}.$$

Choose $i \in \mathcal{A}_k$ arbitrarily. Then $x_i^k - l_i \le \delta_k$ or $u_i - x_i^k \le \delta_k$. If $x_i^k - l_i \le \delta_k$, we get from

$$|x_i^k - x_i^*| \le \|x^k - x^*\| \le \frac{\nu}{4}$$

and (using Lemma 3.1)

$$|x_i^k - l_i| = x_i^k - l_i \le \delta_k \le c\sqrt{\|F(x^k)\|} \le \frac{\nu}{4}$$

that

$$|x_i^* - l_i| \le |x_i^* - x_i^k| + |x_i^k - l_i| \le \frac{\nu}{2},$$

i.e., we have $i \in \mathcal{A}_*$ in view of the very definition of the constant ν.

If, on the other hand, we have $u_i - x_i^k \le \delta_k$, we can show in a similar way that

$$|x_i^* - u_i| \le \frac{\nu}{2},$$

so that $i \in \mathcal{A}_*$ follows also in this case. Hence we have $\mathcal{A}_k \subseteq \mathcal{A}_*$ for all $x^k \in [l, u]$ sufficiently close to x^*. □

As a simple consequence of Lemma 3.3, we obtain the following corollary.

Corollary 3.4 *Let x^* be a solution of the constrained nonlinear system (1). Then $d^k_{\mathcal{A}_k} = x^*_{\mathcal{A}_k} - x^k_{\mathcal{A}_k}$ for all $x^k \in [l, u]$ sufficiently close to x^*.*

Proof. Assume that $x^k \in [l, u]$ is close enough to x^* such that Lemma 3.3 holds, i.e., assume that $\mathcal{A}_k$ is a subset of $\mathcal{A}_*$.

Take any index $i \in \mathcal{A}_k$. Then $i \in \mathcal{A}_*$ so that $x_i^* = l_i$ or $x_i^* = u_i$. In fact, if x^k is sufficiently close to x^*, it follows from our choice of δ that $x_i^* = l_i$ if $x_i^k - l_i \le \delta_k$ and $x_i^* = u_i$ if $u_i - x_i^k \le \delta_k$. Hence we have

$$d_i^k = \begin{cases} l_i - x_i^k = x_i^* - x_i^k & \text{if } x_i^k - l_i \le \delta_k, \\ u_i - x_i^k = x_i^* - x_i^k & \text{if } u_i - x_i^k \le \delta_k. \end{cases}$$

This completes the proof. □

We next introduce our central assumption which will be used in order to show local fast convergence of Algorithm 2.1.

Assumption 3.5 *Let x^* be a solution of the constrained nonlinear system (1). We call x^** regular *if the submatrices $F'(x^*)_{II}$ are nonsingular for all index sets I satisfying $\mathcal{I}_* \subseteq I \subseteq \{1, \dots, n\}$.*

Note that, in particular, $\mathcal{A}_*$ is the empty set if all lower and upper bounds are infinite, so in this case the assumption that x^* is a regular solution reduces to the standard nonsingularity condition for the Jacobian $F'(x^*)$. Moreover, since Assumption 3.5 implies at least the nonsingularity of this Jacobian, it follows from well-known results for the solution of nonlinear systems of equations (see, e.g., [4]) that there is a constant $\alpha > 0$ such that

$$\|F(x)\| \geq \alpha \|x - x^*\| \tag{6}$$

for all x sufficiently close to x^*. In particular, this implies that x^* is a locally isolated solution of the standard nonlinear system $F(x) = 0$ and, therefore, also of our constrained nonlinear system (1). These facts will be used later.

Another consequence of Assumption 3.5 is stated next.

Lemma 3.6 *Let x^* be a regular solution of the constrained nonlinear system (1). Then there exist constants $\kappa_1 > 0$ and $\kappa_2 > 0$ such that*

(a) the submatrices $F'(x^k)_{\mathcal{I}_k\mathcal{I}_k}$ are nonsingular with

$$\|F'(x^k)^{-1}_{\mathcal{I}_k\mathcal{I}_k}\| \leq \kappa_1,$$

and

(b) the linear systems (3) have a unique solution satisfying the inequality

$$\|d^k_{\mathcal{I}_k}\| \leq \kappa_2 \|F(x^k)\|$$

for all $x^k \in [l, u]$ sufficiently close to x^*.

Proof. In view of Lemma 3.3, we have $\mathcal{A}_k \subseteq \mathcal{A}_*$ for all $x^k \in [l, u]$ sufficiently close to x^*. This implies $\mathcal{I}_* \subseteq \mathcal{I}_k \subseteq \{1, \dots, n\}$ for all $x^k \in [l, u]$ close enough to x^*. However, Assumption 3.5 guarantees that all matrices of the form $F'(x^*)_{II}$ with $\mathcal{I}_* \subseteq I \subseteq \{1, \dots, n\}$ are nonsingular. Hence, using a standard perturbation result (see [4]), it follows that for each index set I with $\mathcal{I}_* \subseteq I \subseteq \{1, \dots, n\}$, there exist constants $\kappa_I > 0$ and $\varepsilon_I > 0$ with $F'(x)_{II}$ being nonsingular and

$$\|F'(x)^{-1}_{II}\| \leq \kappa_I$$

for all $x \in [l, u]$ satisfying $\|x - x^*\| \leq \varepsilon_I$. Since there are only finitely many index sets I with $\mathcal{I}_* \subseteq I \subseteq \{1, \dots, n\}$, we have

$$\|F'(x^k)^{-1}_{\mathcal{I}_k\mathcal{I}_k}\| \leq \kappa_1$$

for all $x^k \in [l, u]$ with $\|x^k - x^*\| \leq \varepsilon$, where $\kappa_1 := \max_I \kappa_I$ and $\varepsilon := \min_I \varepsilon_I$. This proves the first statement.

In particular, the linear systems (3) have a unique solution $d^k_{\mathcal{I}_k}$. In order to verify the second statement, we first note that Corollary 3.4 and (6) imply that

$$\|d^k_{\mathcal{A}_k}\| = \|x^k_{\mathcal{A}_k} - x^*_{\mathcal{A}_k}\| \leq \|x^k - x^*\| \leq \|F(x^k)\|/\alpha$$

for some constant $\alpha > 0$ and all x^k close enough to x^*. Moreover, since F' is continuous, there is a constant $\kappa > 0$ such that

$$\|F'(x^k)_{\mathcal{I}_k\mathcal{A}_k}\| \leq \kappa$$

for all x^k close to x^*. Combining these inequalities with the first statement, we finally obtain from (3):

$$\begin{aligned} \|d^k_{\mathcal{I}_k}\| &\leq \|F'(x^k)^{-1}_{\mathcal{I}_k\mathcal{I}_k}\| \left(\|F(x^k)_{\mathcal{I}_k}\| + \|F'(x^k)_{\mathcal{I}_k\mathcal{A}_k}\| \, \|d^k_{\mathcal{A}_k}\|\right) \\ &\leq \kappa_1 \left(\|F(x^k)\| + \frac{\kappa}{\alpha}\|F(x^k)\|\right) \\ &= \kappa_2 \|F(x^k)\| \end{aligned}$$

for all $x^k \in [l, u]$ sufficiently close to x^*, where $\kappa_2 := \kappa_1 + \kappa/\alpha$. □

We next want to show that the entire sequence $\{x^k\}$ generated by Algorithm 2.1 converges to a solution x^* of the constrained nonlinear system if this solution is regular and is at least an accumulation point of the sequence $\{x^k\}$. To this end, we need the following result by Moré and Sorensen [16] (in fact, the result stated here is a refinement of the one given in [16], however, a closer look at the proof given in [16] shows that actually this refined result holds, cf. [12]).

Proposition 3.7 *Assume that $x^* \in \mathbb{R}^n$ is an isolated accumulation point of an arbitrary sequence $\{x^k\} \subseteq \mathbb{R}^n$ (not necessarily generated by Algorithm 2.1) such that, for every subsequence $\{x^k\}_K$ converging to x^*, there is an infinite subset $\tilde{K} \subseteq K$ with $\{\|x^{k+1} - x^k\|\}_{\tilde{K}} \to 0$. Then the whole sequence $\{x^k\}$ converges to x^*.*

Verifying the condition in Proposition 3.7, we are now able to state the following convergence result.

Proposition 3.8 *Let x^* be a regular solution of the constrained nonlinear system (1), and assume that x^* is an accumulation point of a sequence $\{x^k\}$ generated by Algorithm 2.1. Then the entire sequence $\{x^k\}$ converges to x^*.*

Proof. As noted after Assumption 3.5, a regular solution is an isolated solution of the constrained nonlinear system (1). Since the sequence $\{f(x^k)\}$ is monotonically decreasing, it can easily be seen that this implies that x^* is an isolated accumulation point of the sequence $\{x^k\}$.

Next, let $\{x^k\}_K$ be a subsequence converging to x^*. Then, by continuity, we have

$$\{\|F(x^k)\|\}_K \to \|F(x^*)\| = 0. \tag{7}$$

We now consider two cases:

Case 1: There is an infinite subset $\tilde{K} \subseteq K$ such that Algorithm 2.1 takes projected gradient steps for all $k \in \tilde{K}$.

In the same way as in the proof of Theorem 4.5 in [8], we then obtain from Lemma 3.1 and the fact that the projection operator is nonexpansive:

$$\begin{aligned}
\|x^{k+1} - x^k\| &= \|x^k(t_k) - x^k\| \\
&= \|[x^k - t_k \nabla f(x^k)]_+ - x^k\| \\
&= \|[x^k - t_k \nabla f(x^k)]_+ - [x^k]_+\| \\
&\leq \|t_k \nabla f(x^k)\| \\
&\leq s\|\nabla f(x^k)\| \\
&\to_{\tilde{K}} 0
\end{aligned}$$

since x^* solves the constrained nonlinear system (1) so that x^* is a global minimizer and hence an unconstrained stationary point of f (recall that F and therefore also f is defined and continuously differentiable in an open set containing the box $[l, u]$).

Case 2: There are only finitely many indices $k \in K$ for which Algorithm 2.1 takes a projected gradient step.

Then we can assume without loss of generality that the search direction d^k is always the one computed in Step (S.3) of Algorithm 2.1. Since

$$\|d^k_{\mathcal{I}_k}\| \leq \kappa_2 \|F(x^k)\|$$

for all k sufficiently large by Lemma 3.6, we have at least

$$\{\|d^k_{\mathcal{I}_k}\|\}_K \to 0$$

because of (7). If, on the other hand, we have an index $i \in \mathcal{A}_k$, then $i \in \mathcal{A}_*$ by Lemma 3.3, so that we either have

$$d_i^k = l_i - x_i^k \to l_i - x_i^* = 0 \qquad (\text{if } x_i^* = l_i)$$

or

$$d_i^k = u_i - x_i^k \to u_i - x_i^* = 0 \qquad (\text{if } x_i^* = u_i).$$

This shows that

$$\{\|d^k\|\}_K \to 0$$

which, in turn, implies that

$$\{\|x^{k+1} - x^k\|\}_K \to 0.$$

Summarizing both cases, we see that

$$\{\|x^{k+1} - x^k\|\}_{\tilde{K}} \to 0$$

for the infinite subset $\tilde{K} \subseteq K$ defined in Case 1. Hence our statement follows immediately from Proposition 3.7. □

The following result shows that the feasibility preserving safeguard in Step (S.4) of Algorithm 2.1 becomes inactive eventually, i.e., a full step stays within the feasible set $[l, u]$ if x^k is close enough to a regular solution.

Lemma 3.9 *Let x^* be a regular solution of the constrained nonlinear system (1). Then $\tau_k = 1$ for all $x^k \in [l, u]$ sufficiently close to x^*.*

Proof. In view of Lemma 3.6, there exists a constant $\kappa_2 > 0$ such that

$$\|d^k_{\mathcal{I}_k}\| \leq \kappa_2 \|F(x^k)\| \tag{8}$$

for all $x^k \in [l, u]$ close enough to x^*. Since we have $x^k_i - l_i > \delta_k$ and $u_i - x^k_i > \delta_k$ for $i \in \mathcal{I}_k$, we consider two cases.

Case 1: $x^k_i - l_i > \delta_k$.
Since $\|F(x^k)\| \to 0$ for $x^k \to x^*$, we obtain from (8) and the definition of δ_k:

$$\begin{aligned}
x^k_i + d^k_i &> l_i + \delta_k + d^k_i \\
&\geq l_i + \delta_k - \kappa_2 \|F(x^k)\| \\
&= l_i + \min\{\delta, c\sqrt{\|F(x^k)\|}\} - \kappa_2 \|F(x^k)\| \\
&= l_i + c\sqrt{\|F(x^k)\|} - \kappa_2 \|F(x^k)\| \\
&= l_i + \left(c/\sqrt{\|F(x^k)\|} - \kappa_2 \right) \|F(x^k)\| \\
&\geq l_i
\end{aligned}$$

for all $x^k \in [l, u]$ sufficiently close to x^*.

Case 2: $u_i - x^k_i > \delta_k$.
Similar to the previous case, we can show that

$$x^k_i + d^k_i < u_i$$

for all $x^k \in [l, u]$ in a small neighbourhood of x^*.

Combining these two cases together with the definition of $\bar{\tau}_k$, it follows that $\bar{\tau}_k \geq 1$ and therefore $\tau_k = 1$ for all $x^k \in [l, u]$ sufficiently close to x^*. □

We now prove that our approximation $\mathcal{A}_k$ of the active set $\mathcal{A}_*$ eventually coincides with $\mathcal{A}_*$.

Lemma 3.10 *Let x^* be a regular solution of the constrained nonlinear system (1). Then $\mathcal{A}_k = \mathcal{A}_*$ for all $x^k \in [l, u]$ sufficiently close to x^*.*

Proof. We first note that (6) implies

$$|x^k_i - x^*_i| \leq \|F(x^k)\|/\alpha \tag{9}$$

for some constant $\alpha > 0$, all $i \in \{1, \ldots, n\}$, and all $x^k \in [l, u]$ sufficiently close to x^*.

In order to prove our statement, it suffices to show that $\mathcal{A}_* \subseteq \mathcal{A}_k$ since the other inclusion has already been shown in Lemma 3.3. So let $i \in \mathcal{A}_*$ be any fixed index. Then $x^*_i = l_i$ or $x^*_i = u_i$.

Assume first that $x_i^* = l_i$. Then

$$x_i^k - l_i = |x_i^k - x_i^*| \leq \|F(x^k)\|/\alpha$$

because of (9) and Lemma 3.1. Since $\|F(x^k)\| \to 0$ for $x^k \to x^*$ and $\delta_k = O(\sqrt{\|F(x^k)\|})$ in view of the definition of δ_k, we have

$$\|F(x^k)\|/\alpha \leq \delta_k$$

for all $x^k \in [l, u]$ sufficiently close to x^*. Therefore, we obtain

$$x_i^k - l_i \leq \delta_k,$$

i.e., $i \in \mathcal{A}_k$.

If, on the other hand, we have $x_i^* = u_i$, we can show in an analogous way that

$$u_i - x_i^k \leq \delta_k.$$

This proves that $\mathcal{A}_* \subseteq \mathcal{A}_k$, so that both sets eventually coincide by Lemma 3.3. $\square$

We stress that the proofs of the previous two results clearly show why we took the square root of $\|F(x^k)\|$ in the definition of δ_k, cf. Step (S.2) of Algorithm 2.1. In fact, all results would still be true if we would take an arbitrary power $\|F(x^k)\|^p$ with $p \in (0, 1)$.

In order to prove local fast convergence, we will also employ the following result which is a specialization of a more general theorem in [12] (see also [7]).

Proposition 3.11 *Let $G : [l, u] \to \mathbb{R}^n$ be continuously differentiable in an open set containing $[l, u]$, $x^* \in [l, u]$ with $G(x^*) = 0$ and $G'(x^*)$ nonsingular be given, and assume that there are two sequences (not necessarily generated by Algorithm 2.1) $\{x^k\} \subseteq \mathbb{R}^n$ and $\{d^k\} \subseteq \mathbb{R}^n$ with $\{x^k\} \to x^*$ and $\|x^k + d^k - x^*\| = o(\|x^k - x^*\|)$. Then $\|G(x^k + d^k)\| = o(\|G(x^k)\|)$.*

In order to verify the assumptions of Proposition 3.11, we next show that the search directions computed in Step (S.3) of Algorithm 2.1 are "superlinearly convergent search directions".

Lemma 3.12 *Let x^* be a regular solution of the constrained nonlinear system (1), and let $\{x^k\}$ be any sequence (not necessarily generated by Algorithm 2.1) converging to x^*. Then*

$$\|x^k + d^k - x^*\| = o(\|x^k - x^*\|),$$

where d^k denotes the search direction calculated in Step (S.3) of Algorithm 2.1. Furthermore, if F' is locally Lipschitzian around x^*, then

$$\|x^k + d^k - x^*\| = O(\|x^k - x^*\|^2).$$

Proof. Since $x^k \to x^*$ by assumption, it follows from Lemma 3.6 that the search directions d^k as computed in Step (S.3) of Algorithm 2.1 actually exist for all k sufficiently large. Moreover, Lemma 3.10 guarantees that we have $\mathcal{A}_k = \mathcal{A}_*$ for all k large enough. In the rest of this proof, we therefore assume implicitly that k is sufficiently large so that the above two statements hold.

First consider an index $i \in \mathcal{A}_*$. Then $x_i^* = l_i$ or $x_i^* = u_i$. If $x_i^* = l_i$, then $d_i^k = l_i - x_i^k$, and it follows that

$$|x_i^k + d_i^k - x_i^*| = |x_i^k + l_i - x_i^k - l_i| = 0 = o(\|x^k - x^*\|). \tag{10}$$

Similarly, if $x_i^* = u_i$, then $d_i^k = u_i - x_i^k$, and we obtain

$$|x_i^k + d_i^k - x_i^*| = |x_i^k + u_i - x_i^k - u_i| = 0 = o(\|x^k - x^*\|). \tag{11}$$

Next consider indices $i \in \mathcal{I}_*$. Using the notation $\mathcal{N} := \{1, \dots, n\}$, we obtain from the linear system (3), Corollary 3.4 and the fact that $\mathcal{I}_k = \mathcal{I}_*$:

$$\begin{aligned} F'(x^k)_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} + d^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) &= F'(x^k)_{\mathcal{I}_*\mathcal{I}_*} d^k_{\mathcal{I}_*} + F'(x^k)_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) \\ &= -F(x^k)_{\mathcal{I}_*} - F'(x^k)_{\mathcal{I}_*\mathcal{A}_*} d^k_{\mathcal{A}_*} + F'(x^k)_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) \\ &= -F(x^k)_{\mathcal{I}_*} + F(x^*)_{\mathcal{I}_*} + F'(x^k)_{\mathcal{I}_*\mathcal{A}_*}(x^k_{\mathcal{A}_*} - x^*_{\mathcal{A}_*}) \\ &\quad + F'(x^k)_{\mathcal{I}_*\mathcal{I}_*}(x^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}) \\ &= -F(x^k)_{\mathcal{I}_*} + F(x^*)_{\mathcal{I}_*} + F'(x^k)_{\mathcal{I}_*\mathcal{N}}(x^k - x^*), \end{aligned} \tag{12}$$

since the solution x^* satisfies $F(x^*) = 0$. Since F is continuously differentiable, we have

$$|F_i(x^k) - F_i(x^*) - \nabla F_i(x^k)^T(x^k - x^*)| = o(\|x^k - x^*\|)$$

for all $i \in \{1, \dots, n\}$. Hence we obtain

$$\|F_{\mathcal{I}_*}(x^k) - F_{\mathcal{I}_*}(x^*) - F'(x^k)_{\mathcal{I}_*\mathcal{N}}(x^k - x^*)\| = o(\|x^k - x^*\|),$$

so it follows from (12) and Lemma 3.6 that

$$\|x^k_{\mathcal{I}_*} + d^k_{\mathcal{I}_*} - x^*_{\mathcal{I}_*}\| = o(\|x^k - x^*\|). \tag{13}$$

Using (10), (11) and (13), we have

$$\|x^k + d^k - x^*\| = o(\|x^k - x^*\|).$$

In a similar way, one can show that

$$\|x^k + d^k - x^*\| = O(\|x^k - x^*\|^2)$$

holds if F' is locally Lipschitzian around x^*. □

Summarizing the previous results, we are now in the position to prove our main local convergence theorem.

Theorem 3.13 *Let x^* be a regular solution of the constrained nonlinear system (1), and assume that x^* is an accumulation point of a sequence $\{x^k\}$ generated by Algorithm 2.1. Then the following statements hold:*

(a) The entire sequence $\{x^k\}$ converges to x^*.

(b) Eventually, the algorithm takes only Newton steps, i.e., no projected gradient steps are carried out for k sufficiently large.

(c) The rate of convergence is Q-superlinear.

(d) The rate of convergence is Q-quadratic if F' is locally Lipschitzian around x^*.

Proof. Statement (a) follows directly from Proposition 3.8.

Furthermore, Lemma 3.6 implies that the search directions d^k do at least exist in Step (S.3) of Algorithm 2.1 for all k sufficiently large, and Proposition 3.11 together with Lemma 3.12 show that

$$\|F(x^k + d^k)\| = o(\|F(x^k)\|).$$

Hence

$$f(x^k + d^k) = o(f(x^k))$$

so that the descent condition (4) is eventually satisfied with $\tau_k = 1$. This proves the second statement.

Statements (c) and (d) are now immediate consequences of Lemma 3.12 and the fact that $x^{k+1} = x^k + d^k$ for all sufficiently large k (recall that $\tau_k = 1$ for k large enough by Lemma 3.9). $\square$

4 Concluding Remarks

We presented an active set-type Newton method for the solution of constrained nonlinear systems. This method generates only feasible iterates and has some strong theoretical convergence properties. A modification of this method (using a somewhat different globalization strategy and generating strictly feasible iterates) has recently been investigated in [11], where it is used in order to solve reformulations of mixed complementarity problems as box constrained nonlinear systems of equations. The manuscript [11] also contains some numerical results which indicate that Newton-type methods combined with an active set strategy as described in this paper seem to work quite well in practice.

References

[1] D.P. Bertsekas: *Projected Newton methods for optimization problems with simple constraints.* SIAM Journal on Control and Optimization 20, 1982, pp. 221–246.

[2] D.P. Bertsekas: *Nonlinear Programming.* Athena Scientific, Belmont, MA, 1995.

[3] P.H. Calamai and J.J. Moré: *Projected gradient methods for linearly constrained problems.* Mathematical Programming 39, 1987, pp. 93–116.

[4] J.E. Dennis, Jr., and R.B. Schnabel: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Englewood Cliffs, NJ, 1983 (reprinted by SIAM, Philadelphia, PA, 1996).

[5] F. Facchinei, A. Fischer and C. Kanzow: *A semismooth Newton method for variational inequalities: The case of box constraints.* In: M.C. Ferris and J.-S. Pang (eds.): *Complementarity and Variational Problems: State of the Art.* SIAM, Philadelphia, PA, 1997, pp. 76–90.

[6] F. Facchinei, A. Fischer and C. Kanzow: *On the accurate identification of active constraints.* SIAM Journal on Optimization 9, 1999, pp. 14–32.

[7] F. Facchinei and J. Soares: *A new merit function for nonlinear complementarity problems and a related algorithm.* SIAM Journal on Optimization 7, 1997, pp. 225–247.

[8] M.C. Ferris, C. Kanzow and T.S. Munson: *Feasible descent algorithms for mixed complementarity problems.* Mathematical Programming, to appear.

[9] R. Fletcher: *Practical Methods of Optimization.* John Wiley & Sons, New York, NY, 1987.

[10] E.M. Gafni and D.P. Bertsekas: *Convergence of a gradient projection method.* Report LIDS-P-1201, Laboratory for Information and Decisision Systems, Massachusetts Institute of Technology, MA, 1982.

[11] C. Kanzow: *Strictly feasible equation-based methods for mixed complementarity problems.* Preprint 145, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, August 1999.

[12] C. Kanzow and H.-D. Qi: *A QP-free constrained Newton-type method for variational inequality problems.* Mathematical Programming 85, 1999, pp. 81–106.

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

[14] C.-J. Lin and J.J. Moré: *Newton's method for large bound-constrained optimization problems.* Preprint ANL/MCS-P724-0898, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, August 1998 (revised April 1999).

[15] J. Liu: *Strong stability in variational inequalities.* SIAM Journal on Control and Optimization 33, 1995, pp. 725–749.

[16] J.J. Moré and D.C. Sorensen: *Computing a trust region step.* SIAM Journal on Scientific and Statistical Computing 4, 1983, pp. 553–572.

[17] S.M. Robinson: *Strongly regular generalized equations.* Mathematics of Operations Research 5, 1980, pp. 43–62.

[18] D. Sun: *Projected Newton-type methods for constrained nonsmooth equations.* Talk presented at the "Workshop on Nonsmooth and Smoothing Methods" in Hongkong, China, December 1998.

[19] M. Ulbrich: *Non-monotone trust-region methods for bound-constrained semismooth equations with applications to nonlinear mixed complementarity problems.* Technical Report TUM-M9906, Technical University of Munich, Munich, Germany, April 1999.

[20] T. Wang, R.D.C. Monteiro and J.-S. Pang: *An interior point potential reduction method for constrained equations.* Mathematical Programming 74, 1996, pp. 159–195.