An Inexact Hybrid Algorithm for Nonlinear Systems of Equations

Mohammedi El Hallabi

CRPC-TR94372
January 1994

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Revised: May, 1995. Formerly entitled: "An Inexact Trust-Region Globalization of Newton's Method." Also available as CAAM-TR94-26 from the Department of Computational and Applied Mathematics, Rice University

AN INEXACT HYBRID ALGORITHM FOR NONLINEAR SYSTEMS OF EQUATIONS[1]

M. EL HALLABI[2]

Abstract. In this work we define a hybrid algorithm for approximating zeros of the nonlinear systems $F(x) = 0$, where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable. We are concerned with the possibility that n may be large and the Jacobian $F^{'}(x)$ sparse and singular. Trust region globalization methods are known to be robust and can be applied successfully to obtain global convergence results under rather weak hypotheses. However, these algorithms can be expensive, especially for large problems, if the trust region radius needs to be reduced quite often before an acceptable step is obtained. Exploiting the convex structure of the local model subproblem, we propose a hybrid algorithm that uses both trust region and linesearch globalization strategies. It solves, once and not accurately, a local model to obtain a search direction and then uses linesearch techniques to obtain an acceptable steplength . We demonstrate, under rather weak hypotheses, that the algorithm is globally convergent and that the sequence of residuals converges to zero. Moreover, under standard assumptions of Newton's method theory, we prove that the rate of convergence is q-superlinear. Furthermore, q-quadratic convergence can be obtained by requiring sufficient accuracy in the solution of the local model trust region subproblem.

Key Words: nonlinear systems, hybrid method, trust region, linesearch, inexact Newton's method, singular Newton's method, global convergence, superlinear convergence, quadratic convergence.

AMS subject classifications. 65K05, 49D37

Introduction. In this paper we consider the problem of solving the nonlinear system of equations

$$F(x) = 0 \tag{1.1}$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable. We will be concerned with the possibility that n may be large, and that the Jacobian of F at x, say $F'(x)$, may be sparse and singular.

It is well known that, locally, problem (??) is often solved by Newton's method. Globally, trust region algorithms can be used successfully to minimize a given norm of the residual, leading to quite satisfactory convergence results. We refer to Moré (1977) [?] for the ℓ_2-norm, to Duff, Nocedal and Reid (1987)[?] for the ℓ_∞-norm, and to El Hallabi and Tapia (1987) [?] and El Hallabi (1993)[?] for an arbitrary norm. However, trust region algorithms may be expensive, especially for large problems, if the local model needs to be solved more that once before an acceptable step is obtained.

Exploiting the convex structure of the local model, we propose a hybrid algorithm that uses both trust region and linesearch globalization strategies to solve problem in its equivalent form

$$minimize \quad f(x) = ||F(x)||_a \tag{1.2}$$

where $||\ ||_a$ is an arbitrary (but fixed) norm on $\mathbb{R}^n$ and F is given in (??). The algorithm solves, once and for an approximate solution only, the local model trust region subproblem

[1] This work was partially done while the author was visiting the Center for Research on Parallel Computation (CRPC) and the Computational and Applied Mathematics Department at Rice University, Houston,Tx 77251-1892

[2] Departement des Sciences de Base, Ecole Hassania des Travaux Publics, B.P. 8108, Route d'El Jadida, Km.7, Oasis, Casablanca, Morocco.

$$
\text{(1.3)} \qquad \begin{array}{ll} \textit{minimize} & m_k(s) = \|F(x_k) + F'(x_k)s\|_a \\ \textit{subject to} & \|s\|_b \leq \Delta_k, \end{array}
$$

where $\| \ \|_b$ is an arbitrary (but fixed) norm on $\mathbb{R}^n$; and then uses linesearch techniques to obtain an acceptable step.

In (??) and (1.3), we use arbitrary norms for the convenience of the presentation and for the sake of mathematical generalization. Motivated by the recent developments in the linear programming research area (primal-dual interior-point methods, simplex type methods), we mainly aim to use polyhedral norms, in which case the local model trust region subproblem can be formulated as a linear programming problem.

In Section 2 we define the optimality conditions for solving problem (1.2) and derive a necessary and sufficient condition for stationary points to solve problem (??). The inexact hybrid algorithm for nonlinear systems of equations (IHANSE) is described in Section 3. In Section 4 we demonstrate that the IHANSE Algorithm is globally convergent. In Section 5 we prove, under rather weak assumptions, that the sequence of residuals $\{F(x_k)\}$ converges to zero. Moreover, we prove that if the iteration sequence has an accumulation point, say x_*, such that $F'(x_*)$ is nonsingular, then actually converges to such a point. Furthermore, the q-superlinear convergence of the iteration sequence is demonstrated in Section 6; so is the fact that the rate of convergence is q-quadratic if more accuracy is required in the minimization of the local model subproblem. Finally, in Section 7 we present a summary and some concluding remarks.

2. Optimality Conditions. In this section, we define the optimality conditions for problem (1.3). We also give a necessary and sufficient condition for stationary points to be solutions of problem (??).

The locally Lipschitz composite function $f = \|F\|_a$ is regular, i.e. its *generalized directional derivative* denoted $f^0(x;s)$ and its *one-sided directional derivative* denoted $f'(x;s)$ exist and are equal (see Clarke (1983) [?]). They are respectively defined by

$$
\text{(2.1)} \qquad f^0(x;s) = \limsup_{y \to x,\ t \downarrow 0} \frac{f(x+ts) - f(x)}{t} \ .
$$

and

$$
\text{(2.2)} \qquad f'(x;s) = \lim_{t \downarrow 0} \frac{f(x+ts) - f(x)}{t} \ .
$$

Also its *generalized gradient* at x, denoted $\partial f(x)$, is the subset of $\mathbb{R}^n$ defined by

$$
\text{(2.3)} \qquad \partial f(x) = \{ g \in \mathbb{R}^n \mid f^0(x;s) \geq g^T s, \quad \forall s \in \mathbb{R}^n \} \ .
$$

In this research, we use both derivatives although they are equal. To study the optimality conditions, working with the one-sided directional derivative is sufficient. But to analyze the behavior of the algorithm at an iterate that is not a stationary point of f, the generalized directional derivative is a powerful tool because its definition uses a hall neighborhood of x rather just the point x.

The usual definition of a stationary point x in non differentiable optimization is that $0 \in \partial f(x)$ or equivalently $f^0(x; s \geq 0$ for all $s \in \mathbb{R}^n$. But since in our case the function f is regular, we will use the following definition of stationarity.

DEFINITION 1. *Let $f = ||F||_a$, where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a continuously differentiable function, and let $x \in \mathbb{R}^n$. Then x is a stationary point of f if*

$$f'(x; s) \geq 0 \quad \forall s \in \mathbb{R}^n \ . \tag{2.4}$$

In the following two lemmas we define the local model of f, say m_x, and we show that it has the same directional derivatives than f. Moreover we show that the notion of stationarity can be defined in terms of the set of minimizers of the local model. These properties are important from an algorithmic point of view.

LEMMA 2. *Let $f = ||F||_a$, where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a continuously differentiable function, and let $x \in \mathbb{R}^n$. Then*

$$f'(x; s) = m_x'(0; s) \ , \quad \forall \ s \in \mathbb{R}^n \ . \tag{2.5}$$

where

$$m_x(s) = ||F(x) + F'(x)s||_a \ . \tag{2.6}$$

Moreover, we have

$$f'(0; s) \leq m_x(s) - m_x(0). \tag{2.7}$$

Proof. For (??), we refer to El Hallabi and Tapia (1987)[?], and Inequality (??) is an obvious consequence of (??) and the convexity of $m_x()$. □

LEMMA 3. [El Hallabi and Tapia (1987)[?]]. *Let $f = ||F||_a$ where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable. Then $x_* \in \mathbb{R}^n$ is a stationary point of f if and only if for all $s \in \mathbb{R}^n$*

$$||F(x_*)||_a \leq ||F(x_*) + F^{'}(x_*)s||_a$$

or equivalently $m_{x_*}(0) \leq m_{x_*}(s)$ for all $s \in \mathbb{R}^n$ where m_x is given in (??).

In the following theorem, we establish a necessary and sufficient condition for a stationary point of f to be a solution of the nonlinear system (??).

LEMMA 4. *Let x_* be a stationary point of $f = ||F||$. Then x_* is a solution of the nonlinear system (??), i.e.*

$$F(x_*) = 0 \tag{2.8a}$$

is and only if the linearized system

$$F(x_*) + F'(x_*)s = 0 \tag{2.8b}$$

is consistent.

Proof. The proof is an obvious consequence of Lemma ??. □

3. The Inexact Hybrid Algorithm. In this section we define our hybrid algorithm for approximating a solution of the non differentiable optimization problem

$$\min_{x \in \mathbb{R}^n} f(x) = ||F(x)||_a$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable and where $|| \, ||_a$ is an arbitrary norm on $\mathbb{R}^n$.

At each iteration, the algorithm solves the local model for an approximate solution in the sense given in the following definition.

DEFINITION 5. *Assume that x is not a stationary point of f, $\varepsilon > 0$, $\Delta > 0$, and $|| \, ||_a$ and $|| \, ||_b$ are arbitrary norms on $\mathbb{R}^n$. We say that s_ε is an ε-solution of the local model trust region subproblem*

$$\begin{array}{ll} \textit{minimize} & m_x(s) = ||F(x) + F'(x)s||_a \\ \textit{subject to} & ||s||_b \leq \Delta \end{array}$$

if s_ε satisfies

$$m_x(s_\varepsilon) - m_x(0) < 0 \quad \textit{and} \quad m_x(s_\varepsilon) \leq m_x(s) + \varepsilon$$

for all s satisfying $||s||_b \leq \Delta$.

Inexact Hybrid Algorithm for Nonlinear Systems of Equations (IHANSE)

Let $c_i, i = 0, \ldots, 5, \Delta_{\min}, \Delta_0$ and β_0 be constants satisfying:

$$\begin{array}{ll} 0 < c_1 < c_2 < 1 \leq c_3 & 0 < c_4 < c_5 < 1 \\ 0 < \Delta_{\min} \ll 1 & 1 \ll \Delta_{\max} < \infty \\ < \beta_0 & 0 < \Delta_{\min} \leq \Delta_0 \end{array}$$

Let x_0 be any point in $\mathbb{R}^n$, and let $|| \, ||_a$ and $|| \, ||_b$ be arbitrary (but fixed) two norms on $\mathbb{R}^n$.

The algorithm will generate a sequence $\{(x_k, \Delta_k, \beta_k)\}$, where x_k is the iterate, Δ_k the trust region radius, and β_k is used to measure the required accuracy in the ε_k-solution.

Suppose that x_k, Δ_k, and β_k have been determined by the algorithm at the k^{th} iteration. The algorithm determines x_{k+1}, Δ_{k+1}, and β_{k+1} in the following manner:

STEP 1. Obtain an ε_k-solution with

$$\varepsilon_k = \beta_k \min\Big(||s_k||_b, ||F(x_k)||_a\Big),$$

of the model trust region subproblem

$$(LMTR) \equiv \left\{ \begin{array}{l} \textit{minimize} \quad m_k(s) = ||F(x_k) + F'(x_k)s||_a \\ \textit{subject to } ||s||_b \leq \Delta_k \end{array} \right.$$

STEP 2. Set $t_k = 1$

Until

$$f(x_k + t_k s_k) \leq f(x_k) + c_1[m_k(t_k s_k) - f(x_k)]$$

Choose $\bar{t_k}$ such that

$$c_4 t_k \leq \bar{t_k} \leq c_5 t_k;$$

set $t_k = \bar{t_k}$.

End (Until)

STEP 3. If $f(x_k + t_k s_k) \le f(x_k) + c_2[m_k(t_k s_k) - f(x_k)]$

choose Δ_{k+1} so that

$$\Delta_k \le \Delta_{k+1} \le \max(\Delta_k, c_3 \|t_k s_k\|)$$

Else

Choose Δ_{k+1} such that

$$c_4 \|t_k s_k\| \le \Delta_{k+1} \le \|t_k s_k\|$$

STEP 4. Set $\Delta_{k+1} = \min(\max(\Delta_{k+1}, \Delta_{\min}), \Delta_{\max})$

Choose $0 \le \beta_{k+1} \le \beta_0$.

REMARK 3.1. We could use $\varepsilon_k = \beta_k \|F(x_k)\|_a$ instead of $\varepsilon_k = \beta_k \min\Big(\|s_k\|_b, \|F(x_k)\|_a\Big)$. IN Appendix A, we discuss the use of the later choice versus the first one.

Throughout this paper, we use the following definition.

DEFINITION 6. . *The successful steplength t_k obtained in STEP 2 of the IHANSE Algorithm will be said to be acceptable with respect to (x_k, Δ_k, β_k). Moreover the iterate $x_{k+1} = x_k + t_k s_k$ will be referred to as a successor of x_k.*

4. Global Convergence for the IHANSE Algorithm. In this section we will establish the global convergence of the inexact hybrid algorithm for nonlinear systems of equations described in Section 3. Throughout this section, unless otherwise stated, $\varepsilon_k(s_k)$ is defined by

$$\varepsilon_k(s_k) = \beta_k \min\Big(\|s_k\|_b, \|F(x_k)\|_a\Big) . \tag{4.1}$$

We start by proving that any ε_k-solution, in the sense of Definition ??, of the local model trust region subproblem LMTR is a descent direction for $f = \|F\|_a$ at the current iterate, and consequently, we can obtain an acceptable step by using linesearch techniques.

PROPOSITION 7. *Assume that x_k is not a stationary point of f. Then*

$$f'(x_k; s_k) < 0 \tag{4.2}$$

holds for any ε_k-solution s_k of the local model subproblem LMTR. Moreover, there exists $t_k \in (0, 1]$ such that

$$f(x_k + t_k s_k) \le f(x_k) + c_1[m_k(t_k s_k) - f(x_k)] . \tag{4.3}$$

Proof. The proof follows from Definition ??, Lemma ?? and the inequality (**??**).□

PROPOSITION 8. . *Let $\{(x_k, \Delta_k, \beta_k)\}$ be a sequence converging to $(x_*, \Delta_*, 0)$, where x_* and x_k are not stationary points of f, and where $\Delta_k \ge \Delta_{min}$. Let s_k be an $\varepsilon_k(s_k)$-solution of the local subproblem*

$$\begin{array}{ll} \textit{minimize} & m_k(s) = \|F(x_k) + F'(x_k)s\|_a \\ \textit{subject to} & \|s\|_b \le \Delta_k . \end{array} \tag{4.4}$$

Then any accumulation point of $\{s_k\}$, *say* s_*, *is an exact solution of the local subproblem*

$$(4.5) \qquad \begin{array}{ll} \textit{minimize} & m_*(s) = ||F(x_*) + F'(x_*)s||_a \\ \textit{subject to} & ||s||_b \leq \Delta_k, \end{array}$$

Proof. First, observe that the condition $\Delta_k \geq \Delta_{\min}$ implies that $\Delta_* > 0$. Since $\{\Delta_k\}$ converges to Δ_* and $||s_k||_b \leq \Delta_k$ for all k, the sequence $\{s_k\}$ is bounded. Consider any accumulation point s_* of this sequence. We prove that

$$(4.6) \qquad ||F(x_*) + F'(x_*)s_*||_a \leq ||F(x_*) + F'(x_*)s||a$$

holds for all s such that $||s||_b \leq \Delta_*$, i.e., s_* is an exact solution of (4.5). Let s satisfy $||s||_b \leq \Delta_*$. We consider two cases:

i) First, we assume that $||s||_b < \Delta_*$. Since $\{\Delta_k\}$ converges to Δ_*, $||s|| \leq \Delta_k$ holds for sufficiently large k; and because s_k is an $\varepsilon_k(s_k)$-solution of the local subproblem (4.4), we obtain

$$(4.7) \qquad ||F(x_k) + F'(x_k)s_k||_a \leq ||F(x_k) + F'(x_k)s||_a + \varepsilon_k(s_k)$$

which implies that (??) is satisfied.

ii) Now, we assume that $||s||_b = \Delta_*$. Consider $y_k = \frac{\Delta_k}{||s||_b}s$, which satisfies $||y_k||_b = \Delta_k$. Because s_k is an $\varepsilon_k(s_k)$-solution of the local subproblem (4.4), we obtain

$$(4.8) \qquad ||F(x_k) + F'(x_k)s_k||_a \leq ||F(x_k) + \frac{\Delta_k}{||s||_b}F'(x_k)s||_a + \varepsilon_k(s_k)\ .$$

By passing to the limit when $k \to +\infty$, we obtain

$$(4.9) \qquad ||F(x_*) + F'(x_*)s_*||_a \leq ||F(x_*) + \frac{\Delta_*}{||s||_b}F'(x_*)s||_a\ ,$$

and since $||s||_b = \Delta_*$, this implies (??).□

The condition $\Delta_k \geq \Delta_{\min}$ implied by STEP 4 of the IHANSE algorithm was first introduced in El Hallabi and Tapia (1987)[?]. In a trust region framework, it forces the algorithm, before reducing the radius if needed, to start each iteration with a radius at least as large as some arbitrary small fixed constant $\Delta_{\min}$. This safeguard led to quite powerful global convergence in both unconstrained and constrained optimization. We refer to El Hallabi (1993) [?] and El Hallabi and Tapia (1987)[?] for the first case, and to Alexandrov (1993)[?], Dennis, El Alem and Maciel (1992)[?] and El Hallabi (1993)[?] for the second. The main implication of this safeguard is that the actual radius, i.e. that determines an acceptable step, which may be less than $\Delta_{\min}$ in case the trust region radius has been reduced, remains bounded away from zero at a non stationary point. In the following theorem we establish that this safeguard has the same implication with respect to the steplength t_k that determines an acceptable step in the hybrid algorithm under consideration.

THEOREM 9. *Let* $\{(x_k, \Delta_k, \beta_k)\}$ *be a sequence that converges to some* $(x_*, \Delta_*, 0)$. *Assume that* x_* *and* x_k *are not stationary points of* f *and that* $\Delta_k \geq \Delta_{\min}$ *for all* k. *Then there exists a positive integer* $t(x_*, \Delta_*) > 0$ *such that*

$$(4.10) \qquad t_* \geq t(x_*, \Delta_*)$$

holds for any accumulation point t_* of $\{t_k\}$ where t_k determines an acceptable step with respect to (x_k, Δ_k, β_k).

Proof. Assume that for any constant $\gamma > 0$, there exists an accumulation point of $\{t_k\}$, say $t_{*,\gamma}$, such that

$$0 \leq t_{*,\gamma} < \gamma \ .$$

Therefore there exists a subsequence $\{t_k, k \in N \subseteq \mathbb{N}\}$ converging to zero. Without loss of generality, we can assume that $\{t_k\}$ converges to zero. This implies that for sufficiently large k, we have $0 < t_k < 1$, i.e. a steplength of one is never accepted. Let $\bar{t_k}$ be the last non acceptable steplength in the direction s_k, an $\varepsilon_k(s_k)$-solution of the local model trust region subproblem LMTR. We have that

$$0 < c_4\bar{t_k} \leq t_k \leq c_5\bar{t_k} \ , \tag{4.11}$$

which implies that $\{\bar{t_k}\}$ converges to zero, and

$$f(x_k + \bar{t_k}s_k) - f(x_k) > c_1 \left[\|F(x_k) + \bar{t_k}F'(x_k)s_k\|_a - \|F(x_k)\|_a\right] \tag{4.12}$$

which can be written as

$$\begin{aligned} \frac{f(x_k + \bar{t_k}s_*) - f(x_k)}{\bar{t_k}} &> c_1 \frac{\|F(x_k) + \bar{t_k}F'(x_k)s_*\|_a - \|F(x_k)\|_a}{\bar{t_k}} \\ &\quad + \frac{f(x_k + \bar{t_k}s_*) - f(x_k + \bar{t_k}s_k)}{\bar{t_k}} \\ &+ c_1 \frac{\|F(x_k) + \bar{t_k}F'(x_k)s_*\|_a - \|F(x_k) + \bar{t_k}F'(x_k)s_*\|_a}{\bar{t_k}} \ . \end{aligned} \tag{4.13}$$

But because F is continuously differentiable, we have

$$\|(Fx_k) + \bar{t_k}F'(x_k)s_*\|_a - \|F(x_k)\|_a \geq f(x_k + \bar{t_k}s_*) - f(x_k) + o(\bar{t_k}) \ , \tag{4.14}$$

where $\lim_{k\to+\infty} \frac{o(\bar{t_k})}{\bar{t_k}} = 0$. ¿From (??) and (??) we obtain that

$$\begin{aligned} (1 - c_1) \frac{f(x_k + \bar{t_k}s_*) - f(x_k)}{\bar{t_k}} &> \frac{f(x_k + \bar{t_k}s_*) - f(x_k + \bar{t_k}s_k)}{\bar{t_k}} \\ + c_1 \frac{\|F(x_k) + \bar{t_k}F'(x_k)s_k\|_a - \|F(x_k) + \bar{t_k}F'(x_k)s_*\|_a}{\bar{t_k}} &+ \frac{o(\bar{t_k})}{\bar{t_k}} \ , \end{aligned} \tag{4.15}$$

which implies, because f and the norm $\| \ \|_a$ are locally Lipschitz and because $0 < c_1 < 1$, that

$$\limsup_{k\to+\infty} \frac{f(x_k + \bar{t_k}s_*) - f(x_k)}{\bar{t_k}} \geq 0 \ , \tag{4.16}$$

and hence

$$\limsup_{y\to x_*,\ t\downarrow 0} \frac{f(y + ts_*) - f(y)}{t} \geq 0. \tag{4.17}$$

¿From the regularity property of f, (??), (??), and (??), we obtain

(4.18) $$f'(x_*, s_*) \geq 0 \ ,$$

which, together with Lemma ?? and the convexity of $m_{x_*}(\ )$, implies that

(4.19) $$||F(x_*)||_a \leq ||F(x_*) + F'(x_*)s_*||_a \ .$$

On the other hand the sequence $\{s_k\}$ is bounded. Let s_* be any accumulation point of this sequence. Without loss of generality we can assume that $\{s_k\}$ converges to s_*. From Proposition ??, we obtain that s_* is an exact solution of the local model trust region subproblem (4.5), which, together with (??), implies that zero solves the local model trust region subproblem (4.5). Therefore, by Proposition ??, we conclude that x_* must be a stationary point of f, which contradicts our hypothesis.

Consequently, there exists a positive scalar $t(x_*, \Delta_*)$ such that (??) holds for any accumulation point t_* of $\{t_k\}$.□

Now we establish that the IHANSE Algorithm satisfies a property we call we *Local Uniform Decrease*. We believe that this property is a very powerful tool to obtain a global convergence result (see El Hallabi (1993)[?] and El Hallabi (1993)[?]). This property is the most important hypothesis of the global convergence theory of Polak (1970)[?] and Huard (1979)[?] concerning some conceptual algorithms.

THEOREM 10. *Consider $(x_*, \Delta_*, 0)$ where $\Delta_* > 0$ and x_* is not a stationary point of f. Then there exists a neighborhood of $(x_*, \Delta_*, 0)$, denoted $N_* = N(x_*, \Delta_*, 0)$, and a positive scalar $\rho_* = \rho(x_*, \Delta_*)$ such that for any $(x, \Delta, \beta) \in N_*$ with $\beta \geq 0$*

(4.20) $$f(x_+) < f(x_*) - \rho_*$$

holds for any successor (x_+, Δ_+, β_+) of (x, Δ, β).

Proof. Assume that the theorem does not hold. Then there exists a sequence $\{(x_k, \Delta_k, \beta_k)\}$, with $\beta_k > 0$, converging to $(x_*, \Delta_*, 0)$, a sequence of positive scalars $\{\rho_k\}$ converging to zero, and a sequence of successors $\{(x_{k+}, \delta_{k+}, \beta_{k+})\}$ such that

(4.21) $$f(x_{k+}) \geq f(x_*) - \rho_k$$

holds for all k. Therefore, for all k, there exists an $\varepsilon_k(s_k)$-solution s_k of the subproblem (4.4) and $0 < t_k \leq 1$ such that $x_{k+} = x_k + t_k s_k$ satisfies (??) and

(4.22) $$f(x_{k+}) \leq f(x_k) + c_1 \left[||F(x_k) + t_k F'(x_k)s_k||_a - ||F(x_k)||_a \right] \ .$$

¿From (??) and (??) we obtain

$$f(x_*) - \rho_k \leq f(x_k) + c_1 \left[||F(x_k) + t_k F'(x_k)s_k||_a - ||F(x_k)||_a \right] \ ,$$

and then, since the sequence $\{(t_k, s_k)\}$ is bounded,

(4.23) $$||F(x_*) + t_* F'(x_*)s_*||_a - ||F(x_*)||_a \geq 0 \ .$$

where (t_*, s_*) is an accumulation point $\{(t_k, s_k)\}$. Observe that, by Theorem ??, $t_* > 0$. Let us set

(4.24) $$\phi_*(t) = \|F(x_*) + tF'(x_*)s_*\|_a ,$$

and rewrite (??) as

(4.25) $$\phi_*(0) \leq \phi_*(t_*) .$$

Because ϕ_* is convex and $0 < t_* \leq 1$, we obtain necessarily from (??)

$$\phi_*(0) \leq \phi_*(1) ,$$

or equivalently

(4.26) $$\|F(x_*)\|_a \leq \|F(x_*) + F'(x_*)s_*\|_a .$$

On the other hand we obtain from Proposition ?? that s_* is an exact minimizer of the local model trust region subproblem (4.5). Consequently we obtain from (??) that zero is a solution of subproblem (4.5). This, together with Lemma ??, implies that x_* is a stationary point of f, which contradicts our hypothesis. □

In the following theorem, we demonstrate that the inexact hybrid algorithm for nonlinear systems of equations IHANSE described in Section 3 is globally convergent.

THEOREM 11. *Consider a continuously differentiable function $F : \mathbb{R}^n \to \mathbb{R}^n$. Let $\| \|_a$ and $\| \|_b$ be arbitrary (but fixed) norms on $\mathbb{R}^n$, x_0 be an arbitrary point in $\mathbb{R}^n$, and $f(x) = \|F(x)\|_a$. Assume that the the sequence $\{\beta_k\}$ converges to zero. Then any accumulation point of the sequence $\{x_k\}$ generated by the IHANSE algorithm of Section 3 using x_0 as initial iterate is a stationary point of f.*

Proof. Let x_* be an accumulation point of the sequence $\{x_k\}$ generated by the IHANSE algorithm. Without loss of generality (by considering a subsequence if necessary), we can assume that the sequence converges to x_*. The sequence $\{(x_k, \Delta_k, \beta_k)\}$ is bounded. Let $\{(x_j, \Delta_j, \beta_j)\}$ be a subsequence that converges to $(x_*, \Delta_*, 0)$. Because the sequence $\{f(x_k)\}$ is decreasing, we have

$$f(x_j) \leq f(x_k) \quad \forall j \geq k ,$$

which implies that

(4.27) $$f(x_*) \leq f(x_k) \quad \forall k \in \mathbb{N} .$$

Suppose that x_* is not a stationary point of f. Since the sequence $\{(x_j, \Delta_j, \beta_j)\}$ converges to $(x_*, \Delta_*, 0)$, there exists an integer j_* such that $(x_j, \Delta_j, \beta_j) \in N_*$ for all $j \geq j_*$, where N_* is defined in Theorem 4.3. Hence, we obtain

(4.28) $$f(x_{j+1}) < f(x_*) - \rho_* \quad \forall j \geq j_* ,$$

which contradicts (??). Consequently, any accumulation point of the sequence $\{x_k\}$ generated by the IHANSE algorithm in Section 3 is a stationary point of $f = \|F\|_a$. □

REMARK 4.1. Actually, Theorem ?? can be obtained as an application of Theorem ?? and the work of either Huard (1979)[?] or Polak (1970)[?] concerning the global convergence of conceptual algorithms. We choose to give a direct proof because that proof is not long and contributes to the completeness of the presentation.

5. Convergence to a Solution of $F(x) = 0$. In this section we demonstrate that, under rather weak hypotheses, the sequence of residuals $\{F(x_k)\}$ converges to zero. We also demonstrate that if the iteration sequence generated by the IHANSE Algorithm has an accumulation point, say x_*, such that $F^{'}(x_*)$ is nonsingular, then, the iteration sequence converges to x_*.

In the following two theorems, under rather weak assumptions that do not include the non-singularity of the Jacobian, we prove that the sequence of residuals converges to zero. This can be considered as a convergence result for the singular Newton's method.

THEOREM 12. *Assume the hypotheses of Theorem ??. Also assume that there exists a bounded subsequence $\{x_k, k \in N \subset \mathbb{N}\}$ and a constant $\eta \in [0, 1)$ such that*

$$||F(x_k) + F'(x_k)s_k|| \leq \eta ||F(x_k)|| \tag{5.1}$$

holds for all $k \in N \subset \mathbb{N}$. Then any accumulation point of the iteration sequence, generated by the IHANSE Algorithm, is a solution of the nonlinear system (??). Moreover, the sequence of residuals $\{F(x_k)\}$ converges to zero.

Proof. Let x_* be an arbitrary accumulation point of the subsequence $\{x_k, k \in N \subset \mathbb{N}\}$. ¿From Theorem ??, we obtain that x_* is a stationary point of f, which implies that

$$||F(x_*)|| \leq ||F(x_*) + F^{'}(x_*)s|| \tag{5.2}$$

for all $s \in \mathbb{R}^n$. On the other hand, since $\{s_k\}$ is bounded, we can assume without loss of generality that it converges to s_*. Therefore, inequality (??) implies that

$$||F(x_*) + F'(x_*)s_*|| \leq \eta ||F(x_*)||. \tag{5.3}$$

¿From (??), (??), and $0 \leq \eta < 1$, we obtain

$$F(x_*) = 0.$$

This implies, since the sequence $\{||F(x_k)||_a\}$ is decreasing, that $\{F(x_k)\}$ converges to zero, hence any accumulation point of the iteration sequence is a solution of the nonlinear system (??).□

REMARK 5.1. Condition (??) can be written as

$$m_k(s_k) \;\leq\; \eta \quad m_k(0). \tag{5.4}$$

Because, first, at each iteration we minimize, within some tolerance (see Definition 3.1), the local model trust-region subproblem LMTR, second, zero is a feasible point for such minimization problem, and third, we are considering a zero residual problem, the assumption that (??) holds for a subsequence does not seem to be restrictive.

Now, we prove that the iteration sequence actually converges.

THEOREM 13. *Assume the hypotheses of Theorem ??. Also assume that the iteration sequence has an accumulation point, say x_*, such that the linear system*

$$F(x_*) + F'(x_*)s = 0 \tag{5.5}$$

is consistent. Then any accumulation point of the iterate sequence is a solution of the nonlinear system (1.1). Moreover the sequence of residuals $\{F(x_k)\}$ converges to zero.

Proof. Let x_* be an arbitrary accumulation point of $\{x_k\}$. By Theorem ??, x_* is a stationary point of $f = \|F\|_a$. Assume that (??) holds. Then, by Lemma ??, we have

$$F(x_*) = 0, \tag{5.6}$$

which implies, since the sequence $\{\|F(x_k)\|_a\}$ is decreasing, that $\{F(x_k)\}$ converges to zero. Finally, we obtain that any accumulation point accumulation point of the iteration sequence is a solution of the nonlinear system (??). □.

THEOREM 14. *Assume the hypotheses of Theorem ??. If the sequence $\{x_k\}$ generated by the IHANSE Algorithm has an accumulation point, say x_*, such that $F'(x_*)$ is nonsingular, then $\{x_k\}$ converges to x_*, and $F(x_*) = 0$.*

Proof. Let x_* be an accumulation point of $\{x_k\}$ such that $F'(x_*)$ is nonsingular. Then the linear system (??) is consistent, and hence, by Theorem ??, $\{F(x_k)\}$ converges to zero. On the other hand, we have

$$\|F(x_k) + t_k F'(x_k)s_k\|_a = \|t_k(F(x_k) + F'(x_k)s_k) + (1 - t_k)F(x_k)\|_a$$

where s_k is an ε_k-solution of the local model trust region subproblem LMTR, and $t_k \in (0, 1]$ is an acceptable steplength with respect to (x_k, Δ_k, β_k). Therefore we have

$$\|F(x_k) + F'(x_k)(t_k s_k)\|_a \leq t_k\|(F(x_k) + F'(x_k)s_k\|_a + (1 - t_k)\|F(x_k)\|_a \ ,$$

and by using the Definition 3.1 of an approximate solution of LMTR we obtain

$$\|F(x_k) + F'(x_k)(t_k s_k)\|_a \leq \|F(x_k)\|_a \ .$$

Now, the convergence of the sequence $\{x_k\}$ to x_* follows from Theorem 3.3 of Eisenstat and Walker (1993) [?].

Observe that if we were solving the local model exactly, the next section would not be needed. Indeed, under the standard assumptions of Newton's method, the Newton step converges to zero, and because $\Delta_k \geq \Delta_{min}$, it becomes feasible for the local model subproblem LMTR for sufficiently large k. Therefore, we could conclude that the IHANSE Algorithm reduces to Newton's method for sufficiently large k, and hence it is q-quadratically convergent.

6. Convergence Rate of the IHANSE Algorithm. In this section, we prove that, under the standard assumptions of Newton's method, the IHANSE Algorithm is q-superlinearly convergent and that it is q-quadratically convergent if either $\beta_k = O(\|F(x_k)\|)$ or $\beta_k = O(\|s_k\|)$. We also prove that, for sufficiently large k, the trust region radius is not decreased, which implies that a very small safeguard for global convergence $\Delta_{\min}$ is of no importance for the convergence rate.

LEMMA 15. *Assume the hypothesis of Theorem ??. Also assume that the sequence $\{x_k\}$ generated by the IHANSE Algorithm has an accumulation point, say x_*, such that $F'(x_*)$ is nonsingular and F' is Lipschitz near x_*. Then there exists a positive integer k_* such that $\Delta_k \geq \Delta_{k_*}$ for $k \geq k_*$ and a steplength of one is acceptable with respect to (x_k, Δ_k, β_k).*

Proof. By Theorem ??, the iteration sequence converges to x_* and $\{\|F(x_k)\|\}$ converges to zero. Let us show that the inequality

$$f(x_{k+1}) < f(x_k) + c_2[m_k(s_k) - m_k(0)] \ , \tag{6.1}$$

where s_k is an $\varepsilon_k(s_k)$-solution of LMTR subproblem, is satisfied for sufficiently large k. Because $0 < c_1 < c_2 < 1$, this will answer both questions of the lemma. Since F is continuously differentiable and $\{x_k\}$ converges to x_*, we have

$$f(x_k + s_k) = ||F(x_k) + F'(x_k)s_k + o(||s_k||_b)||_a \ ,$$

and hence

$$f(x_k) - f(x_k + s_k) \geq f(x_k) - (m_k(s_k) - ||o(||s_k||_b)||_a) \ . \tag{6.2}$$

Because $f(x_k) - m_k(s_k) > 0$, this implies

$$\frac{f(x_k) - f(x_k + s_k)}{f(x_k) - m_k(s_k)} \geq 1 - \frac{||o(||s_k||_b)||_a}{||s_k||_b} \frac{||s_k||_b}{f(x_k) - m_k(s_k)} \ . \tag{6.3}$$

Let us show that

$$\frac{f(x_k) - m_k(s_k)}{||s_k||_b} \geq M_* \tag{6.4}$$

for some positive constant M_*. Since $\{x_k\}$ converges to x_*, $F'(x_*)$ is nonsingular, and F is continuously differentiable, there exists a positive integer k_* and a positive constant λ_* such that $F'(x_k)$ is nonsingular and

$$||F'(x_k)d||_a \geq \lambda_* ||d||_b \quad \forall d \in \mathbb{R}^n \ \text{ and } \ \forall k \geq k_* \ . \tag{6.5}$$

Consider (x_k, Δ_k, β_k) for $k \geq k_*$, and denote by s_k^N the Newton step, i.e. the solution of

$$F'(x_k)s_k^N + F(x_k) = 0 \ .$$

We consider two cases:
Case 1. Assume that $\Delta_k < ||s_k^N||$. Let us define

$$\alpha_k = \frac{||s_k||_b}{||s_k^N||_b} \ \text{ and } \ \hat{s}_k = \alpha_k s_k^N \ . \tag{6.6}$$

Since $0 < \alpha_k < 1$, we have

$$m_k(\hat{s}_k) \ = \ (1 - \alpha_k)||F(x_k)||_a. \tag{6.7}$$

On the other hand, because $||\hat{s}_k||_b = ||s_k||_b$ and s_k is an $\varepsilon_k(s_k)$-solution of the local model, we have

$$f(x_k) - m_k(\hat{s}_k) \leq f(x_k) - m_k(s_k) + \varepsilon_k(\beta_k) \ , \tag{6.8}$$

which, together with (??) and (??), implies

$$\frac{||F'(x_k)s_k^N||_a}{||s_k^N||_b} - \beta_k \leq \frac{f(x_k) - m_k(s_k)}{||s_k||_b} \ . \tag{6.9}$$

Therefore, since $\{\beta_k\}$ converges to zero, we obtain (??) from (??) and (??). Observe that we needed $\varepsilon_k \leq \beta_k ||s_k||_b$ (It is the only place in the paper where our choice of ε_k is needed).

Case 2. Now we assume that $||s_k^N||_b \le \Delta_k$. This implies that the Newton step is feasible for subproblem LMTR. Because s_k is an ε_k-solution of LMTR subproblem, we obtain

$$||F(x_k) + F'(x_k)s_k||_a \le \beta_k ||F(x_k)||_a$$

or equivalently, since the norms on $\mathbb{R}^n$ are equivalent,

$$||F'(x_k)(s_k - s_k^N)||_a \le \mu\beta_k ||F'(x_k)||_a ||s_k^N||_b. \tag{6.10}$$

for some constant μ. From (??) and (??) we obtain that

$$\left| \frac{||s_k||_b}{||s_k^N||_a} - 1 \right| \le M_1 \beta_k \ , \tag{6.11}$$

where M_1 is a constant depending on x_*, holds for sufficiently large k, which implies, by passing to the limit as β_k converges to zero,

$$\lim_{k\to+\infty} \frac{||s_k^N||_b}{||s_k||_b} = 1. \tag{6.12}$$

Therefore, for sufficiently large k, say for $k \ge k_*$ for convenience, we have

$$\frac{||s_k^N||_b}{||s_k||_b} \ge \frac{1}{2}. \tag{6.13}$$

Also, because s_k is an ε_k-solution of the local model trust region subproblem LMTR, we have

$$\frac{f(x_k) - m_k(s_k)}{||s_k||_b} \ge \frac{||F(x_k)||_b}{||s_k||_b} - \frac{\varepsilon_k}{||s_k||_b} \ . \tag{6.14}$$

Since $\varepsilon_k \le \beta_k ||F(x_k)||_a$, we obtain from (??)

$$\frac{f(x_k) - m_k(s_k)}{||s_k||_b} \ge (1-\beta_k) \frac{||F'(x_k)s_k^N||_a}{||s_k||_b}. \tag{6.15}$$

which, together with (??), implies

$$\frac{f(x_k) - m_k(s_k)}{||s_k||_b} \ge (1-\beta_k)\lambda_* \frac{||s_k^N||_a}{||s_k||_b}. \tag{6.16}$$

¿From (??) and (??), we obtain (??). Observe that we only needed $\varepsilon_k = \beta_k ||F(x_k)||_a$ (see the derivation of (??))
Inequalities (??) and (??) imply that for $k \ge k_*$ we have

$$\frac{f(x_k) - f(x_k + s_k)}{f(x_k) - m_k(s)} \ge 1 - \frac{1}{M_*} \frac{||o(||s_k||_b)||}{||s_k||_b} \ . \tag{6.17}$$

On the other hand, there exists a positive integer, say k_* for convenience, such that

$$1 - \frac{1}{M_*} \frac{||o(||s_k||_b)||}{||s_k||_b} > c_2 \tag{6.18}$$

for all $k > k_*$. Finally, we obtain from (??) and (??) that equality (??) holds for $k \ge k_*$. Therefore, for $k \ge k_*$, the trust region radius Δ_k satisfies

$$\Delta_{k_*} \le \Delta_k, \tag{6.19}$$

and a steplength of one is acceptable with respect to (x_k, Δ_k, β_k). Observe that we only needed $\varepsilon_k = \beta_k \|F(x_k)\|_a$ (see the derivation of (??)).□

Now, let us prove that the IHANSE Algorithm converges q-superlinearly.

THEOREM 16. *Assume the hypothesis of Lemma ??. Then the iteration sequence converges q-superlinearly to x_*.*

Proof. Let k_* be given by Lemma ??. Also let s_k^N denote the Newton's step, i.e. $s_k^N = -F'(x_k)^{-1}F(x_k)$. For for $k \geq k_*$, we have

$$F(x_k) + F'(x_k)s_k = F'(x_k)(s_k - s_k^N) \ . \tag{6.20}$$

On the other hand, since the sequence $\{F(x_k)\}$ converges to zero, the Newton step s_k^N is feasible for the local model subproblem LMTR, i.e.

$$\|s_k^N\|_b \leq \Delta_k \ , \tag{6.21}$$

for sufficiently large k. Since s_k is an ε_k-solution of LMTR subproblem, we obtain from (??) and (??)

$$\|F'(x_k)(s_k - s_k^N)\|_a \leq \beta_k \|F(x_k)\|_a \tag{6.22}$$

and consequently

$$\begin{aligned} \|s_k - s_k^N\|_a &\leq \lambda_*^{-1}\beta_k \|F(x_k) - F(x_*)\|_a \\ &\leq L_* \beta_k \|x_k - x_*\|_b \end{aligned} \tag{6.23}$$

for some positive constant L_*. On the other hand we have

$$\begin{aligned} x_{k+1} - x_* &= (x_k + s_k^N - x_*) + (s_k - s_k^N) \\ &= (x_{k+1}^N - x_*) + (s_k - s_k^N) \ , \end{aligned} \tag{6.24}$$

where x_{k+1}^N is the Newton iterate obtained from x_k. Therefore we have

$$\|x_{k+1} - x_*\|_b \leq \|x_{k+1}^N - x_*\|_b + \|s_k - s_k^N\|_b \ . \tag{6.25}$$

¿From (??) and (??), we obtain

$$\|x_{k+1} - x_*\|_b \leq \|x_{k+1}^N - x_*\|_b + L_*\beta_k \|x_k - x_*\|_b \ . \tag{6.26}$$

Consider D_* a convex neighborhood of x_* contained in the domain of the q-quadratic convergence of Newton's method (see Dennis and Schnabel [?]). Since $\{x_k\}$ converges to x_*, there exists an integer, say k_* for convenience, such that $x_k \in D_*$ for all $k \geq k_*$. Then we have

$$\|x_{k+1}^N - x_*\|_b \leq L_2 \|x_k - x_*\|_b^2 \ \ \forall \ \ k \geq k_* \ , \tag{6.27}$$

where L_2 is a positive constant. From (??) and (??) we obtain

$$\|x_{k+1} - x_*\|_b \leq L_2 \|x_k - x_*\|_b^2 + L_*\beta_k \|x_k - x_*\|_b \ . \tag{6.28}$$

Therefore, because $\{\beta_k\}$ converges to zero, (??) implies that

$$\lim_{k\to+\infty} \frac{\|x_{k+1}-x_*\|_b}{\|x_k-x_*\|_b} = 0 \ , \tag{6.29}$$

i.e. the iteration sequence $\{x_k\}$ generated by the algorithm converges q-superlinearly. □

THEOREM 17. . *Assume the hypothesis of Theorem ??.*
i) If $\beta_k = O(\|F(x_k)\|)$ or $\beta_k = O(\|s_k\|)$, the iteration sequence converges q-quadratically to x_*, and
ii) If $\beta_k = 0$ for sufficiently large k, x_k is the Newton iterate for the nonlinear equation $F(x) = 0$ and consequently the rate of convergence of $\{x_k\}$ to x_* is q-quadratic.

Proof. Assume that

$$\beta_k = O(\|F(x_k)\|_a) \tag{6.34a}$$

or

$$\beta_k = O(\|(s_k)\|_a) \ . \tag{6.34b}$$

¿From (??) we obtain

$$\lim_{k\to+\infty} \frac{\|s_k\|_b}{\|x_k-x_*\|_b} = 1 \ . \tag{6.31}$$

Since $F(x_*) = 0$ and F is continuously differentiable, we have

$$\begin{aligned} \|F(x_k)\|_a &= \|F(x_k) - F(x_*)\|_a \\ &\leq L_2\|x_k - x_*\|_b \ . \end{aligned} \tag{6.32}$$

¿From (6.34a), (6.34b) and (??) we obtain

$$\beta_k = O(\|x_k - x_*\|_b) \ . \tag{6.33}$$

Therefore (??) becomes

$$\|x_{k+1}-x_*\|_b \leq L^*\|x_k - x_*\|_b^2 \tag{6.34}$$

i.e. the iteration sequence $\{x_k\}$ generated by the IHANSE algorithm converges q-quadratically to x_*.

Now assume that $\beta_k = 0$ for $k \geq k_*$. This means that we are solving the local model trust region (LMTR) exactly. The proof is similar to the one given for Theorem 8.1 of El Hallabi and Tapia [?]. □

7. Summary and Concluding Remarks. To solve nonlinear systems of equations using an arbitrary starting point, trust region strategies are known to lead to quite robust globally convergent algorithms. However, these algorithms can be expensive if the trust region radius needs to be decreased quite often before an acceptable steplength is obtained, especially for large nonlinear systems.

To remedy to this possible situation, we proposed a hybrid algorithm for approximating a solution of $F(x) = 0$. To maintain the robustness of the trust region globalization strategy, the

IHANSE algorithm described in Section 3, determines a search direction s_k as an approximate solution of the local model trust region subproblem

$$(LMTR) \equiv \begin{cases} \textit{minimize} & m_k(s) = \|F(x_k) + F'(x_k)s\|_a \\ \textit{subject to} & \|s\|_b \le \Delta_k \end{cases}$$

where $\| \|_a$ and $\| \|_b$ are arbitrary (but fixed) norms on $\mathbb{R}^n$, then it uses linesearch techniques in the direction s_k to obtain an acceptable steplength t_k.

Observe that if $\| \|_a$ and $\| \|_b$ are polyhedral norms, in particular $\| \|_a = \ell_1$-norm and $\| \|_b = \ell_\infty$-norm, then the local model subproblem LMTR can be formulated as a linear programming problem.

We proved, under rather weak hypotheses, that the inexact hybrid algorithm for nonlinear systems of Equations is globally convergent.

Under the forcing condition that

$$\|F(x_k) + F'(x_k)s_k\|_a \le \eta \|F(x_k)\|_a, \tag{7.1}$$

where $0 \ll \eta < 1$, holds asymptotically and only for a subsequence, we proved that the sequence of residuals $\{F(x_k)\}$ converges to zero. Observe that since we are solving a nonlinear system of equations, condition (??) is more likely to hold. Moreover, we proved that if the optimal residual is not zero then, for any accumulation point of the iteration sequence, say x_*, not only the Jacobian matrix $F'(x_*)$ is singular, but the linear system $F(x_*) + F'(x_*)s = 0$ is inconsistent.

Also, under standard assumptions of the inexact Newton's method, we showed that the iteration sequence is q-superlinearly convergent and that it is q-quadratically convergent if more accurate, but not exact, minimization of the local model trust region subproblem is performed.

Finally, let us emphasize that the hybrid approach stems from the observation that as long as, first, the local model and the objective function have the same directional derivatives and, second, the local model is convex in the direction of the approximate solution , decreasing the trust region radius to obtain an acceptable step is irrelevant.

8. Appendix A. In the IHANSE algorithm, to obtain an approximate solution, we use the accuracy test

$$\varepsilon_k = \beta_k \min\Big(\|s_k\|_b, \|F(x_k)\|_a\Big) \tag{8.1}$$

that is *a posteriori* defined in the sense that the algorithm updates the required accuracy while solving the local model LMTR. We could just use

$$\varepsilon_k = \beta_k \|F(x_k)\|_a. \tag{8.2}$$

In fact the derivation of global convergence uses (??). But when deriving the superlinear convergence, with (??), we need not to consider **Case 1** in the proof of Lemma 6.1, i.e. we need to use the fact that the Newton step is inside the ball of radius Δ_k. This will occur for sufficiently large k since, by Theorem 5.3, s_k^N converges to zero. But this implies that $\Delta_{\min}$ should not be very small. On the other hand, a not very small $\Delta_{\min}$ would cost more linesearches for the global convergence, and make the algorithm $\Delta_{\min}$-dependent. Moreover, we believe that, although $\Delta_{\min}$ is important to obtain a global convergence result, its use should be for theoretical purposes only. Our approach, then, is to choose a very small $\Delta_{\min}$, and let the algorithm work with the trust region Δ_k that is automatically updated since it will be in general greater than $\Delta_{\min}$. This can be done using the a posteriori defined accuracy

$$\varepsilon_k = \beta_k \min\Big(\|s_k\|_b, \|F(x_k)\|_a\Big).$$

In the following lemma, we show that the approximate solution used in the algorithm is well defined.

LEMMA 18. *Assume that x is not a stationary point of f and $\Delta \geq \Delta_{\min}$. Then IHANSE algorithm always finds an ε_k-solution in the sense of of Definition ??.*

Proof. The local model subproblem LMTR is convex. Therefore its dual program is well defined (see Rockafellar (1970)[?],(1981)[?], Janh (1994)[?]). Let $p_j = m(s_j)$ be the primal objective value obtained at the j^{th} inner-iteration in the process of solving LMTR subproblem, and let d_j be the corresponding dual objective value. We call *primal-dual gap* the difference $m(s_j) - d_j$, and we denote it by pdg_j. Observe that the optimal primal-dual gap pdg_* is zero. On the other hand, since x is not a stationary point of f, there exist $\|s_0\|_b \leq \Delta$ such that

$$\|F(x) + F'(x)s_0\|_a < \|F(x)\|_a. \tag{8.3}$$

Let $\{s_j\}$ be a sequence generated when solving subproblem LMTR such that such that $\{\|F(x) + F'(x)s_j\|_a\}$ is decreasing and

$$\|F(x) + F'(x)s_j\|_a \leq \|F(x) + F'(x)s_0\|_a < \|F(x)\|_a, \tag{8.4}$$

Assume that there exists a subsequence of t $\{s_j\}$ that converges to zero. Then, from (??) we obtain

$$\|F(x)\|_a \leq \|F(x) + F'(x)s_0\|_a < \|F(x)\|_a, \tag{8.5}$$

which is impossible. Therefore there exists a positive constant depending on (x, Δ), say ω, such that

$$\|s_j\|_b \geq \omega \tag{8.6}$$

holds for all j. Now, because the primal-dual gap converges to zero, we have

$$pdg_j \leq \beta \min\Big(\omega, \|F(x)\|_a\Big) \tag{8.7}$$

for sufficiently large j. Let j_* be the smallest integer such that (??) holds. From (??) and (??), we obtain

$$pdg_{j_*} \leq \beta_k \min\Big(\|s_j\|_b, \|F(x)\|_a\Big). \tag{8.8}$$

Using the definition of the primal-dual gap, we rewrite (??) as

$$m_x(s_{j_*}) \leq p_{j_*} + \epsilon(s_{j_*}, \beta) \tag{8.9}$$

which implies that

$$m_x(s_{j_*}) \leq m_x(s) + \epsilon(s_{j_*}, \beta) \tag{8.10}$$

for all s such that $\|s\|_b \leq \Delta_k$, i.e. s_{j_*} is an $\epsilon(s_{j_*}, \beta)$-solution of LMTR subproblem.

REFERENCES

topsep

[1] N. Alexandrov, *Multilevel algorithms for nonlinear equations and equality constrained optimization*, Ph.D. Thesis, Technical Report TR93-20 (1993), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892.

[2] F.H. Clarke, *Optimization and Nonsmooth Analysis.* Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley and Sons Publications, 1983.

[3] J.E. Dennis, M. El-Alem and M.C. Maciel, *A global convergence theory for general trust region-based algorithms for equality constrained optimization.* Technical Report TR92-28 (1992), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892.

[4] J.E. Dennis, Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Englewood Cliffs, New Jersey, 1983.

[5] I.S. Duff, J. Nocedal and J.K. Reid, *The use of linear programming for the solution of sparse sets of nonlinear equations.* SIAM Journal on Scientific and Statistical Computing 2(1987), pp. 98-108.

[6] S.C. Eisenstat and H.F. Walker, *Globally convergent inexact Newton methods.* Research Report, February/91/51, Department of Mathematics and Statistics, Utah State University. Revised August 1992 and March 1993. Siam Journal on Optimization, to appear.

[7] M. El Hallabi, *An inexact minimization trust region algorithm: Globalization of Newton's method.* Technical Report 93-43 (1993), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892.

[8] M. El Hallabi, *A globally convergent theory for arbitrary norm trust region algorithms for equality constrained optimization*, Technical Report 93-60 (1993), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892. (revised May 1995)

[9] M. El Hallabi and R. Tapia, *A global convergence theory for arbitrary norm trust region methods for nonlinear equations.* Technical Report TR87-25, Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892 (revised as TR93-41).

[10] M.D. Hebden, *An algorithm for minimization using exact second derivatives.* Technical Report TP515, Atomic Energy Research Establishment, Harwell, England, 1973.

[11] P. Huard, *Point-set maps and mathematical programming.* Mathematical Programming Study 10, North Holland, Amsterdam, 1979.

[12] J. JAHN, *Introduction to the theory of nonlinear optimization*, Spinger-Verlag, Berlin, Heidelberg, 1994.

[13] J.J. Moré, *The Levenberg-Marquardt algorithm: Implementation and theory.* in G.A. Watson, ed., Lecture Notes in Mathematics 630, Springer-Verlag, Berlin, Heidelberg and New York, 1977, pp. 105-119.

[14] E. Polak, *On the implementation of conceptual algorithm.* In O.L. Mangasarian, K. Ritter and J.B. Rosen, eds., Nonlinear Programming, Academic Press, New York, 1970, pp. 275-291.

[15] R.T. ROCKAFELLAR, *Convex analysis.* Princeton University Press, Princeton, 1970.

[16] R.T. ROCKAFELLAR, *The Theory of subgradients and its applications to problems of optimization: convex and nonconvex functions.* Heldermann, Berlin, 1981.