A Global Convergence Theory for Sequential Linear Programming Inexact Hybrid Algorithms

Mohammedi El Hallabi

CRPC-TR94371
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: "A Global Convergence Theory for SLP and SQP Trust-Region Algorithms." Also available as CAAM-TR95-08 from the Department of Computational and Applied Mathematics, Rice University

A GLOBAL CONVERGENCE THEORY FOR SEQUENTIAL LINEAR PROGRAMMING INEXACT HYBRID ALGORITHMS[1]

MOHAMMEDI EL HALLABI[2]

Abstract. In this paper, we propose a sequential linear programming hybrid algorithm to minimize a nonlinear function $f : \mathbb{R}^n \to \mathbb{R}$ subject to nonlinear equality constraints $h_i(x) = 0, i = 1, \cdots, m$ where $h_i : \mathbb{R}^n \to \mathbb{R}$. We adopt the approach taken in Vardi (1985). We also replace the ℓ_2-norm in the trust-region constraint by the ℓ_∞-norm. At each iteration, a linear programming subproblem is solved within some tolerance. Instead of the regularity assumption of linear independent gradients, we assume that the system of linearized constraints is consistent at any point of the iteration sequence, and that, at any accumulation point of the iteration sequence, the largest singular value of the constraints gradient is bounded away from zero. Also, we assume that the functions f and $h_i, i = 1 \cdots m$, are continuously differentiable. We demonstrate that any accumulation point of the iteration sequence, obtained from an arbitrary starting point, is a Karush-Kuhn-Tucker point of he constrained minimization problem.

Key Words: Sequential Linear Programming, Global Convergence, Constrained Optimization, Consistency, Non Regularity, Equality Constrained, Linear dependency, Trust-Region, Linesearch, Hybrid Methods.

Subject Classifications: 65K05, 4D37

Introduction. In this paper we present a sequential linear programming algorithm for approximating a solution of the equality constrained optimization problem

$$(1.1) \qquad (EQCP) \quad \equiv \quad \begin{cases} \text{minimize} & f(x) \\ \text{subject to} & h_i(x) = 0, i = 1 \cdots m, \end{cases}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h_i : \mathbb{R}^n \to \mathbb{R}$, $\quad i = 1 \cdots$, are continuously differentiable and nonlinear. We are concerned with the possibility that the dimensions of (EQCP), i.e. n and m, might be large.

Problem EQCP can be solved by many trust region methods proposed in the literature; for example, the methods in Vardi [?], Byrd, Schnabel, Omokojun and Shultz [?], El-Alem [?], Powell and Yuan [?], Maciel [?], Dennis, El-Alem and Maciel [?], Alexandrov [?], Marucha, Nocedal, and Plantega [?], El Hallabi [?], and Dennis and Vicente [?]. However, These methods might be expensive if the trust region needs to be decreased quite often before an acceptable step is obtained.

In this research, we adopt the approach used in El Hallabi [?], and propose a sequential linear programming inexact hybrid algorithm (SLPIHA) to solve problem EQCP. This algorithm solves, once, a linear programming subproblem to obtain a descent direction of some merit function, and then uses linesearch techniques to obtain an acceptable steplength. We assume that the functions f and $h_i, i = 1 \cdots m$, are continuously differentiable, that the linear system $h(x) + \nabla h(x)^T s = 0$ is consistent, and that the largest singular value of the constraints gradient is bounded away from zero.

In Section 2, we recall from El Hallabi [?] a sufficient condition for the translation parameter

[1] This work was performed in part while the author was visiting the Computational and Applied Mathematics Department and the Center for Research on Parallel Computation, Rice University, Houston, Texas.

[2] Currently at Ecole Hassania des Travaux Publics, B.P. 8108 Route d'El Jadida, Km.7, Oasis, Casablanca, Morocco.

α_k to define a nonempty feasible region for the translated linear programming subproblem

$$(1.2)\qquad (TLPS) \quad \equiv \quad \begin{cases} \text{minimize} & \nabla f(x_k)^T s \\ \text{subject to} & \alpha_k h(x_k) + \nabla h(x_k)^T s = 0, \\ & \qquad\qquad ||s||_\infty \le \Delta_k. \end{cases}$$

In Section 3, we derive a characterization of stationarity in terms of minimizers of subproblem TPLS. We define the sequential linear programming inexact hybrid algorithm (SLPIHA) in Section 4. In Section 5, we prove that any accumulation point of the sequence generated by the SLPIHA algorithm, from an arbitrary starting point x_0, is a Karush-Kuhn-Tucker point of (EQCP). We end this paper by giving some concluding remarks in Section 6.

2. Linearized Constraint Translation. In this section we recall from El Hallabi [?] a sufficient condition for the translation parameter α to define a nonempty feasible region for the subproblem TLPS in (??). We also relate this parameter to the smallest nonzero singular value of $\nabla h(x_k)$. This relation is needed later.

PROPOSITION 2.1 [El Hallabi [?]]. *Assume that $x \in \mathbb{R}^n$ is not feasible for (EQCP), i.e. $h(x) \neq 0$, and that the linear system*

$$(2.1)\qquad h(x) + \nabla h(x)^T s = 0,$$

is consistent. Assume further that $\Delta > 0$. Let σ_x be the smallest positive singular value of $\nabla h(x)$. If

$$(2.2)\qquad 0 \le \alpha \le \min\Big(1, \frac{\sqrt{2}}{2}\,\Delta\,\frac{\sigma_x}{||h(x)||_2}\Big),$$

then the subset

$$(2.3)\qquad \mathcal{F}(x,\Delta) = \Big\{ s \in \mathbb{R}^n \;\Big|\; \alpha h(x) + \nabla h(x)^T s = 0, \quad ||s||_\infty \le \Delta \Big\} ,$$

is not empty. Moreover its is not a singleton set.

Actually in El Hallabi [?], it shown that

$$(2.4)\qquad \mathcal{M}(x,\Delta) = \left\{ s \in \mathbb{R}^n \;\middle|\; \alpha h(x) + \nabla h(x)^T s = 0, \quad ||s||_2 \le \frac{\sqrt{2}}{2}\,\Delta \right\} ,$$

which is contained in $\mathcal{F}(x,\Delta)$, is not empty.

REMARK 2.1. By chosing $\alpha = 0$ we can generalize Inequality (??) to the case where $h(x) \neq 0$ but $\nabla h(x) = 0$.

REMARK 2.1. The smallest positive singular value of $\nabla h(x_k)$ can be estimated using the QR decopmosition.

3. Characterization of stationary points of problem EQCP. In this section we derive a useful notion of stationarity in terms of minimizers of the linear programming subproblem LPS.

PROPOSITION 3.1. *Let $\Delta_k > 0$, and consider x_k satisfying $h(x_k) = 0$. Then $s_k = 0$ is a solution of the linear programming subproblem*

$$(3.1)\qquad (LPS) \quad \equiv \quad \begin{cases} \text{minimize} & \nabla f(x_k)^T s \\ \text{subject to} & \nabla h(x_k)^T s = 0 \\ & \qquad ||s||_\infty \le \Delta_k, \end{cases}$$

if and only if x_k is a Karush-Kuhn-Tucker point of (EQCP).

Proof. Because of the Slater condition, the proof follows obviously from the necessary and sufficient conditions for zero to be a minimizer of LPS subproblem. □

4. Sequential Linear Programming Inexact Hybrid Algorithm. In this section we propose a sequential linear programming inexact hybrid algorithm (SPLIHA) for solving (EQCP). We also show that the choice of the penalty parameter fits well with the objective function and the constraints.

Approximate solution of the linear programming subproblem.
At each iteration, we solve a translated linear programming subproblem TLPS

$$(TLPS) \quad \equiv \quad \begin{cases} \text{minimize} & \nabla f(x_k)^T s \\ \text{subject to} & \alpha_k h(x_k) + \nabla h(x_k)^T s = 0, \\ & \qquad\qquad ||s||_\infty \leq \Delta_k, \end{cases} \tag{4.1}$$

for some fixed $(x_k, \alpha_k, \Delta_k)$, and within some tolerance ϵ_k in the sense given in the following definition.

DEFINITION 4.1. *Let $x \in \mathbb{R}$, $0 < \alpha$, and $0 < \Delta$. Assume that x is not a Karush-Kuhn-Tucker point of (EQCP). Then we say that s_ϵ is an ϵ-solution of subproblem TLPS if s_ϵ is feasible,*

$$\nabla f(x)^T s_\epsilon \leq \nabla f(x)^T s + \epsilon \tag{4.2}$$

for any feasible s, and if in addition $h(x) = 0$, we also ask that

$$\nabla f(x)^T s_\epsilon < 0. \tag{4.3}$$

Our trial step s_k will be any ϵ_k-solution of the subproblem TLPS for fixed $(x_k, \alpha_k, \Delta_k)$, and with the tolerance

$$\epsilon_k = \beta_k \begin{cases} \alpha_k ||h(x_k)|| & \text{if } h(x_k) \neq 0 \\ ||s_k||_\infty & \text{otherwise} \end{cases} \tag{4.4}$$

for some $0 < \beta_k$ that will be set by the algorithm. Observe that in (??)

$$\alpha_k ||h(x_k)|| = ||h(x_k)|| - ||h(x_k) + \nabla h(x_k) s||,$$

i.e. the decrease in the norm of the constraints obtained for any feasible point of subproblem TLPS. Also, when $h(x_k) = 0$ and x_k is not stationary point, the test in (??) implies that a small approximate solution will point toward the feasible steepest descent, i.e the gradient projection (see El Hallabi and Tapia [?]). This property enables the algorithm to never fail at a nonstationary point.

In the following lemma and its corollary, we show that the ϵ_k-solution is well defined.

LEMMA 4.1. *Let $p_k(s_j)$ and d_k^j denote respectively the primal and the dual objective function values obtained at the j^{th} iteration for solving the subproblem TLPS. Let $\epsilon_k > 0$. If*

$$|p_k(s_j) - d_k^j| \leq \epsilon_k \tag{4.5}$$

holds, i.e. the duality gap is less that ϵ_k, then s_j is an ϵ_k-solution of (TLPS).

Proof. Assume that (??) holds. Then we have

$$p_k(s_j) \leq d_k^j + \epsilon_k \leq p_k(s_*) + \epsilon_k \tag{4.6}$$

where s_* is an exact solution of (TLPS). From (??) we obtain that

$$p_k(s_j) \leq p_k(s) + \epsilon_k$$

for all feasible s for (TLPS), i.e. s_j is an ϵ_k-solution of (TLPS). □

Penalty parameter and merit function.
To accept or reject a trial step s_k, we will use the actual reduction

$$Ared_k(s) = \Phi_k(s) - \Phi_k(0) \tag{4.7}$$

and the predicted reduction

$$Pred_k(s) = \Psi_k(s) - \Psi_k(0) \tag{4.8}$$

where

$$\Phi(x_k, \mu_k; s) = f(x_k + s) + \mu_k \|h(x_k + s)\| \tag{4.9}$$

is the merit function approximated by

$$\Psi(x_k, \mu_k; s) = f(x_k) + \nabla f(x_k)^T s + \mu_k \|h(x_k) + \nabla h(x_k)^T s\|. \tag{4.10}$$

For convenience, we will denote them respectively $\Phi_k(s)$ and $\Psi_k(s)$. In (??) and (??), μ_k denotes the penalty parameter, and $\| \quad \|$ denotes an arbitrary (but fixed) norm on $\mathbb{R}^m$.

The penalty parameter will be defined by

$$\mu_k = \begin{cases} \mu_{k-1} & \text{if } \mu_{k-1} \geq \bar{\mu_k} + \rho \\ \bar{\mu_k} + 2\rho & \text{otherwise,} \end{cases} \tag{4.11}$$

where ρ is a positive constant, and $\bar{\mu_k}$ is given by

$$\bar{\mu_k} = \begin{cases} 0 & \text{if } h(x_k) = 0 \\ 2 \max(0, \frac{\nabla f(x_k)^T s_k}{\alpha_k \|h(x_k)\|}) & \text{otherwise.} \end{cases} \tag{4.12}$$

The functions $s \rightarrow \Phi_k(s)$ and $s \rightarrow \Psi_k(s)$ have the same one-sided directional derivative at the origin. This is given in the following Lemma.

LEMMA 4.2. *Let $x_k \in \mathbb{R}^n$, and $\mu_k > 0$. Then for all $s \in \mathbb{R}^n$, we have*

$$\Phi_k'(0; s) = \Psi_k'(0; s)\ . \tag{4.13}$$

Proof. For all positive t and all $s \in \mathbb{R}^n$, we have

$$\Phi_k(ts) - \Phi_k(0) = f(x_k + ts) - f(x_k) + \mu_k \left[\|h(x_k + ts)\| - \|h(x_k)\| \right] \tag{4.14}$$

or, because f and h_i, $i = 1 \ldots m$ are continuously differentiable

$$(4.15)\quad \begin{aligned} \Phi_k(ts) - \Phi_k(0) &= t\nabla f(x_k)^T s + \mu_k \left[\left\| h(x_k) + t\nabla h(x_k)^T s \right\| - \|h(x_k)\| \right] + o(t) \\ &+ \mu_k \left[\left\| h(x_k) + t\nabla h(x_k)^T s + o(t) \right\| - \left\| h(x_k) + t\nabla h(x_k)^T s \right\| \right] \end{aligned}$$

¿From (??) and the Lipschitz continuity of the norm, we obtain

$$(4.16)\quad \frac{\Phi_k(ts) - \Phi_k(0)}{t} = \frac{\Psi_k(ts) - \Psi_k(0)}{t} + \frac{o(t)}{t}$$

which, by passing to the limit when t converges to zero, implies (??).□

In the following proposition, we show that any ϵ_k-solution of subproblem TPLS is a descent direction of Φ_k at the origin.

PROPOSITION 4.1. *Let x_k, and $\Delta_k > 0$. If x_k is not a non Karush-Kuhn-Tucker point of (EQCP) and s_k is an ϵ_k-solution of (TLPS), then*

$$(4.17)\quad Pred_k(ts_k) = -t \left[\left| \nabla f(x_k)^T s_k \right| - \rho\alpha_k \|h(x_k)\| \right]$$

holds for all $t \in (0, 1]$. Moreover we have

$$(4.18)\quad \Phi'_k(0; s_k) \leq - \left| \nabla f(x_k)^T s_k \right| - \rho\alpha_k \|h(x_k)\| ,$$

and consequently s_k is a descent direction of Φ_k at the origin.

Proof. We have

$$(4.19)\quad \mathrm{Pred}_k(ts_k) = t\nabla f(x_k)^T s_k + \mu_k \left[\left\| h(x_k) + t\nabla h(x_k)^T s_k \right\| - \|h(x_k)\| \right]$$

On the other hand, because $\alpha_k h(x_k) + \nabla h(x_k)^T s_k = 0$, we have for all $t \in (0, 1]$

$$h(x_k) + t\nabla h(x_k)^T s_k = (1 - t\alpha_k)h(x_k) ,$$

which together with (??), implies

$$(4.20)\quad \mathrm{Pred}_k(ts_k) = t \left[\nabla f(x_k)^T s_k - \mu_k\alpha_k \|h(x_k)\| \right] .$$

First we assume that

$$(4.21)\quad \nabla f(x_k)^T s_k > 0.$$

Therefore $h(x_k) \neq 0$ must hold. We have

$$(4.22)\quad \mu_k \geq 2\frac{\nabla f(x_k)^T s}{\alpha_k \|h(x_k)\|} + \rho$$

which, together with (??), implies that

$$(4.23)\quad \mathrm{Pred}_k(ts) \leq t \left[-\nabla f(x_k)^T s_k - \rho\alpha_k \|h(x_k)\| \right] .$$

¿From (??) and (??), we obtain (??).

Now we assume that (??) does not hold, i.e. we have

$$\nabla f(x_k)^T s_k \leq 0 \,. \tag{4.24}$$

We have $\mu_k \geq \rho$. Therefore (??) implies that

$$\mathrm{Pred}_k(ts) \leq t\left[\nabla f(x_k)^T s_k - \rho\alpha_k \|h(x_k)\|\,\right] \tag{4.25}$$

¿From (??) and (??) we obtain (??) .

By passing to the limit when t converges to zero in (??), we obtain (??). Moreover, if $h(x_k) \neq 0$ we obtain from (??) that $\Phi_k'(0; s_k) < 0$, and if $h(x_k) = 0$, we obtain from Definition 4.1 that

$$\nabla f(x_k)^T s_k < 0$$

and hence $\Phi_k'(0; s_k) < 0$ must hold. Consequently s_k is a descent direction of Φ_k at the origin. □

Definition of the algorithm SLPIHA.

Let $c_i, i = 1, \cdots, 5, \rho, \beta$, $\Delta_{\min}$, and $\Delta_{\max}$ be constants satisfying

$$\begin{array}{lll} 0 < c_1 < c_2 < 1 & , \quad 0 < c_3 < c_4 < 1 & , \quad 1 < c_5 \\ 0 < \gamma < 1 & , \quad 0 < \rho & , \quad 0 < \beta \\ 0 < \Delta_{\min} < \Delta_{\max}. & & \end{array}$$

Let $x_0 \in \mathbb{R}^n$ be an arbitrary point, $\Delta_{\min} \leq \Delta_0 \leq \Delta_{\max}$, $0 \leq \beta_0 < \beta$, and $\mu_0 = \rho$.

Let (x_k, Δ_k, β_k) be given by the k^{th} iteration. The algorithm generates $(x_{k+1}, \Delta_{k+1}, \beta_{k+1})$ by the following iterative scheme:

STEP 1. If $h(x_k) = 0$ set $\alpha_k = 1$ and go to STEP 3,

STEP 2. Obtain a lower bound of the positive singular values of $\nabla h(x_k)$, say ω_k, and set

$$\alpha_k = \min\Big(1, \frac{\sqrt{2}}{2}\, \Delta_k\, \frac{\omega_k}{\|h(x_k)\|_2}\Big),$$

STEP 3. Obtain an ϵ_k-solution of the subproblem TPLS with ϵ_k defined in (??),

STEP 4. Update the penalty parameter μ_k using (??) and (??)

STEP 5. Set $t_k = 1$

Until $Ared_k(t_k s_k) \leq c_1 Pred_k(t_k s_k)\}$

choose $\bar{t_k}$ such that $c_3 t_k \leq \bar{t_k} \leq c_4 t_k$,

set $t_k = \bar{t_k}$

End Until

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

STEP 6. If $Ared_k(t_k s_k) \leq c_2 Pred_k(t_k s_k)$

then choose δ_{k+1} such that $\Delta_k \leq \delta_{k+1} \leq \max(\Delta_k, c_5 t_k \|s_k\|_\infty)$

Else choose δ_{k+1} such that $c_4 t_k \|s_k\|_\infty \leq \delta_{k+1} \leq t_k \|s_k\|_\infty$.

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

STEP 7. Choose $0 \leq \beta_{k+1} < \beta$.

Observe that $\Delta_k \geq \Delta_{\min}$ holds for all k. Throughout this paper, we will use the following definition.

DEFINITION 4.1. *If for some t_k, the test in STEP 9 is satisfied, we say that t_k is an accepted steplength with respect to (x_k, Δ_k, β_k). Moreover, we say that (Δ_k, β_k) determines an acceptable step (or steplength). Furthermore, we will refer to x_{k+1} as a successor of x_k and to $(x_{k+1}, \Delta_{k+1}, \beta_{k+1})$ as a successor of (x_k, Δ_k, β_k).*

5. Global Convergence. In this section, we demonstrate that any accumulation point of the iteration sequence generated by the SLPIHA Algorithm is a Karush-Kuhn-Tucker point of (EQCP).

We make the following hypotheses:

H.1) The functions f and hi, $i = 1 \ldots, m$, are continuously differentiable.

H.2) The systems of linearized constraints $h(x_k) + \nabla h(x_k)^T s = 0$ are consistent for all k.

H.3) At any accumulation point of the iteration sequence $\{x_k\}$, say x_*, there exists $\nu_* > 0$ such that $||\nabla h(x_*)|| \geq \nu_*$, and

H.4) The sequence $\{\beta_k\}$ converges to zero.

To obtain our global convergence result, given by Theorem 5.4, we derive some important properties of the algorithm near non stationary points. These properties will play a crucial role in our global convergence theory analysis.

We start by analyzing, in the following lemma and its corollary, the behavior of the penalty parameter μ_k.

LEMMA 5.1 [El Hallabi [?]]. *Let $\{(x_k, \Delta_k, \beta_k)\}$ converge to $(x_*, , \Delta_*, 0)$. Then the there exists a positive constant μ_* such that $\bar{\mu_k} \leq \mu_*$.*

COROLLARY 5.1 [El Hallabi [?]]. *Assume that the hypothesis of Lemma 5.1 holds. Then there exists an integer k^* such that $\mu_k = \mu_{k^*}$ for all $k \geq k^*$.*

The following technical lemma will be used later.

LEMMA 5.2. *Assume that $\{(x_k, \Delta_k, \beta_k)\}$ converges to $(x_*, , \Delta_*, 0)$. Assume further that*

$$\lim_{k \to +\infty} \alpha_k h(x_k) = 0 \tag{5.1}$$

holds. Then

$$h(x_*) = 0 \tag{5.2}$$

holds. Moreover $\alpha_k = 1$ holds for sufficiently large k.

Proof. ¿From the equivalence of norms, the definition of α_k, and (??), we obtain

$$\lim_{k \to +\infty} \min\Big(||h(x_k)||_2, \frac{\sqrt{2}}{2}\ \Delta_k\ \omega_k\Big) = 0. \tag{5.3}$$

On the other hand, the singular values of $\nabla h(x)$ are continuous functions of x. Let σ_* denote the smallest nonzero singular value of $\nabla h(x_*)$. We obtain form hypothesis H.3 that, necessarily

$$\sigma_{k,r_k} \geq \frac{\sigma_*}{2} > 0 \tag{5.4}$$

holds for sufficiently large k. Therefore we can assume that the lower bound of the singular values of $\nabla h(x_k)$, i.e. ω_k, satisfies

$$\omega_k \geq \tau \sigma_* \tag{5.5}$$

where $\tau_* \in (0,1)$ is an arbitrary small positive constant, depending on x_*. Now, from (??) and (??), we obtain

$$h(x_*) = 0. \tag{5.6}$$

Finally, we obtain from the definition of α_k, $\Delta_k \geq \Delta_{\min}$, and (??) that

$$\alpha_k = 1 \tag{5.7}$$

holds for sufficiently large k. □

In the following proposition, we show that the algorithm cannot stop at a nonstationary point.

PROPOSITION 5.1. *Let x_k be a non Karush-Kuhn-Tucker point of (EQCP), and let s_k be a ϵ_k- solution of the linear programming subproblem TLPS. Then there exists $t_k \in (0,1]$ that is an acceptable steplength.*

Proof. ¿From Proposition 4.1, we obtain that s_k is a descent direction of the merit function Φ_k at the origin.

Assume that the algorithm reduces the steplength t indefinitely without obtaining an acceptable one. Then we have

$$\frac{\Phi_k(t_j s_k) - \Phi_k(0)}{t_j} > c_1 \frac{\Psi_k(t_j s_k) - \Psi_k(0)}{t_j}$$

where $\{0 < t_j\}$ converges to zero. By passing to the limit when $j \to +\infty$, we obtain

$$\Phi_k'(0; s_k) \geq c_1 \Psi_k'(0; s_k) \ . \tag{5.8}$$

¿From Lemma 4.2, (??) and $c_1 \in (0,1)$, we obtain

$$\Phi_k'(0; s_k) \geq 0 \ ,$$

which contradicts the fact that s_k is a descent direction of Φ_k at the origin.□

In the following theorem, we analyze the behavior of the steplength near a non stationary point.

THEOREM 5.2. *Let $\{(x_k, \Delta_k, \beta_k)\}$ converge to $(x_*, \Delta_*, 0)$, where x_k and $x*$ are not Karush-Kuhn-Tucker points of (EQCP) and $\Delta_k \geq \Delta_{\min}$. If t_k is an acceptable steplength with respect to (x_k, Δ_k, β_k), then there exists a positive scalar $t(x_*, B_*, \Delta_*)$ such that*

$$t_* \geq t(x_*, B_*, \Delta_*) \tag{5.9}$$

holds for any accumulation point t_* of $\{t_k\}$.

Proof. Assume that the theorem does not hold. Then, for all integer j, there exists an accumulation point $t_{*,j} \leq \frac{1}{j}$. This implies that there exists a subsequence $\{t_j, j \in N \subset \mathbb{N}\}$ that converges to zero. Without loss of generality we can assume that the sequence $\{t_j\}$ converges to zero. Consequently $0 < t_j < 1$ holds for sufficiently large j, which implies that $t_j = 1$ is never an acceptable steplength with respect to (x_j, δ_j, η_j). Let $\bar{t}_j$ be the last non acceptable steplength with respect to (x_j, δ_j, η_j). We have

$$c_3 \bar{t}_j \leq t_j \leq c_4 \bar{t}_j \ , \tag{5.10}$$

which implies that $\{\bar{t}_j\}$ converges to zero. Since $\bar{t}_j$ is not acceptable, we have

$$\Phi_j(\bar{t}_j s_j) - \Phi_j(0) > c_1 \left[\Psi_j(\bar{t}_j s_j) - \Psi_j(0)\right],$$

or equivalently

$$f(x_j + \bar{t}_j s_j) - f(s_j) + \mu_j \left[\, \|h(x_j + \bar{t}_j s_j)\| - \|h(x_j)\| \, \right] > \tag{5.11}$$
$$c_1 \left\{ \bar{t}_j \nabla f(x_j)^T s_j + \mu_j \left[\, \left\| h(x_j) + \bar{t}_j \nabla h(x_j)^T s_j \right\| - \|h(x_j)\| \, \right] \right\} \ .$$

Because $0 < c_1 < 1$, f and h_i, $i = 1 \ldots m$ are continuously differentiable, and the norm is locally Lipshitz, we obtain from (??)

$$\left\{ \nabla f(x_j)^T s_j + \mu_j \, \frac{\left\| h(x_j) + \bar{t}_j \nabla h(x_j)^T s_j \right\| - \|h(x_j)\|}{\bar{t}_j} \right\} + \frac{o(\bar{t}_j)}{\bar{t}_j} > 0 \ . \tag{5.12}$$

But since $\alpha_k h(x_k) + \nabla h)x_k)^T s_k = 0$, we have

$$\left\| h(x_j) + \bar{t}_j \nabla h(x_j)^T s_j \right\| - \|h(x_j)\| = -\mu_j \bar{t}_j \alpha_j \|h(x_j)\| \ . \tag{5.13}$$

Using (??), we rewrite (??) as

$$\nabla f(x_j)^T s_j - \mu_j \alpha_j \|h(x_j)\| > \frac{o(\bar{t}_j)}{\bar{t}_j} \ . \tag{5.14}$$

Therefore we obtain

$$\nabla f(x_j)^T s_j > \frac{o(\bar{t}_j)}{\bar{t}_j} \ . \tag{5.15}$$

Also, from the definition of μ_k (see (??)), and (??), we obtain

$$-\rho \alpha_j \|h(x_j)\| > \nabla f(x_j)^T s_j + \frac{o(\bar{t}_j)}{\bar{t}_j} \ , \tag{5.16}$$

which, together with (??), implies that

$$-\rho \alpha_j \|h(x_j)\| > \frac{o(\bar{t}_j)}{\bar{t}_j}. \tag{5.17}$$

Therefore, we obtain from (??) and Lemma 5.2 that

$$\|h(x_*)\| = 0 \tag{5.18}$$

and that $\alpha_j = 1$ holds for sufficiently large j. Let Δ_* and s_* be respectively accumulation points of $\{\Delta_j\}$ and $\{s_j\}$. Without loss of generality, we can assume that these sequences converge respectively to zero Δ_* and s_*. Therefore, we obtain from Huard [?]) that s_* is an exact solution of the linear programming suproblem

$$(LPS) \left\{ \begin{array}{ll} \text{minimize} & \nabla f(x_*)^T s \\ \text{subject to} & \nabla h(x_*)^T s = 0, \|s\|_\infty \le \Delta_*. \end{array} \right. \tag{5.19}$$

(The ϵ_j-solution is considered as a function of (x_j, Δ_j, β_j)). Also, from (??) we obtain

$$\nabla f(x_*)^T s_* \geq 0\,, \tag{5.20}$$

which implies that zero solves the subproblem LPS. Therefore we conclude from Proposition 3.1 that x_* is a Karush-Kuhn-Tucker point of (EQCP) which contradicts the hypothesis. Consequently (??) holds. □

Before we give our global convergence result,we establish that the SLPIHA Algorithm satisfies the very important *local uniform decrease* property. This property played a pivotal role to obtain the global convergence in El Hallabi [?].

Since, when $\{(x_k, \Delta_k, \beta_k)\}$ converges to some $(x_*, \Delta_*, 0)$, the penalty parameter becomes constant for all sufficiently large k, and since we assume that the iteration sequence is infinite, the merit function $\Phi(\mu_k, x_k; s)$ is constant with respect to this parameter; therefore, in the following theorem, we denote $\Phi(x_k + s)$ instead of $\Phi(\mu_k, x_k, s)$.

THEOREM 5.3. *(Local Uniform Decrease). Let $\{x_k, \Delta_k, \beta_k)\}$ converges to some $(x_*, \Delta_*, 0)$. If x_* is not a Karush-Kuhn-Tucker point of (EQCP) then there exists a positive integer k_*, depending on (x_*, Δ_*), such that for all $k \geq k_*$*

$$\Phi(x_{k+}) < \Phi(x_*) \tag{5.21}$$

holds for any successor $(x_{k+}, \Delta_+, \beta_{k+})$ of (x_k, Δ_k, β_k).

Proof. Assume that the theorem dose not hold. Then there exists a subsequence $\{(x_k, \Delta_k, \beta_k)\}$ converging to $(x_*, \Delta_*, 0)$ and a subsequence of successors $\{(x_{k+}, \Delta_+, \beta_{k+})\}$ such that

$$\Phi(x_{k+}) \geq \Phi(x_*) \tag{5.22}$$

holds for all k. Therefore there exists an ϵ_k-solution of the linear programming subproblem

$$(TLPS)\begin{cases} \text{minimize} & \nabla f(x_k)^T s \\ \text{subject to} & \alpha_k h(x_k) + \nabla h(x_k)^T s = o\,, ||s||_\infty \leq \Delta_k \end{cases}$$

such that $x_{k+} = x_k + t_k s_k$, $t_k \in (0, 1]$, and

$$\Phi(x_k + t_k s_k) \leq \Phi(x_k) + c_1 \mathrm{Pred}_k(t_k s_k)\,. \tag{5.23}$$

Inequalities (??) and (??) imply that

$$\Phi(x*) - \Phi(x_k) \leq c_1 \mathrm{Pred}_k(t_k s_k). \tag{5.24}$$

¿From Proposition 4.1 and (??) we obtain

$$\Phi(x_*) - \Phi(x_k) \leq -c_1 t_k \left[\left|\nabla f(x_k)^T s_k\right| + \rho\alpha_k ||h(x_k)\right]\,, \tag{5.25}$$

which, together with Theorem 5.2, implies that

$$\lim_{k\to+\infty} \alpha_k ||h(x_k)|| = 0\,. \tag{5.26}$$

Therefore, we obtain from Lemma 5.2 and (??) that

$$h(x_*) = 0 \tag{5.27}$$

and that $\alpha_k = 1$ holds for sufficiently large k . On the other hand, let us consider s_k, an ϵ_k-solution of subproblem TLPS, as a function of (x_k, Δ_k, β_k). Then we obtain from Huard [?], that any accumulation point, say s_*, of $\{s_k\}$ is a solution of the linear programming problem

$$(LPS) \begin{cases} \text{minimize} & \nabla f(x_*)^T s \\ \text{subject to} & \nabla h(x_*)^T s = 0, \ \|s\|_{+\infty} \leq \Delta_* \end{cases} \tag{5.28}$$

Also, we obtain from (??) that

$$\nabla f(x_*)^T s_* = 0 \ . \tag{5.29}$$

Consequently zero solves the linear programming subproblem LPS in (??), which, by Proposition 3.1, contradicts the hypothesis that x_* is not a Karush-Kuhn-Tucker point of (EQCP).□

Finally, we give our global convergence result which detracts from the matter at hand.

THEOREM 5.4. *Let* $\{x_k\}$ *be an iteration sequence generated by the SLPIHA Algorithm described in Section 3. Assume that*
H.1) the functions f *and* $h_i, i = 1 \ldots m$ *are continuously differentiable,*
H.2) for all k*, the linearized constraints are consistent,*
H.3) at any accumulation point of the iteration sequence $\{x_k\}$*, say* x_**, there exists* $\nu_* > 0$ *such that* $\|\nabla h(x_*)\| \geq \nu_*$*, and*
H.4) the sequence $\{\beta_k\}$ *converges to zero.*
Then any accumulation point of $\{x_k\}$ *is Karush-Kuhn-Tucker point of (EQCP).*

Proof. Let x_* be an arbitrary accumulation point of $\{x_k\}$. Recall that, because penalty parameter μ_k is constant for sufficiently large k, the merit function Φ is constant with respect to this parameter and therefore will be denoted $\Phi(x)$. Let $\bar{k} = \max(k_*, k^*)$, where k^* and k_* are respectively given if Corollary 5.1 and Theorem 5.3. Since, for all $k \geq \bar{k}$, the iterate x_k is not a Karush-Kuhn-Tucker point of (EQCP), we have

$$\Phi(x_{k+1}) < \Phi(x_k) \quad \forall k \geq \bar{k} \ .$$

Let $\{x_j, j \geq \bar{k}\}$ be a subsequence that converges to x_*. Consider $k \geq \bar{k}$.There exists $j(k) > k$ such that

$$\Phi(x_{j(k)}) < \Phi(x_k)$$

and consequently

$$\Phi(x_j) < \Phi(x_k)$$

holds for all $j \geq j(k)$. Therefore, we obtain

$$\Phi(x_*) < \Phi(x_k) \quad \forall k \geq \bar{k} \ . \tag{5.30}$$

Assume that x_* is not a Karush-Kuhn-Tucker point of (EQCP). Therefore, we obtain from Theorem 5.3

$$\Phi(x_{j+1}) < \Phi(x_*) \quad \forall j \geq \bar{k} \ .$$

This contradicts (??). Consequently x_* is a Karush-Kuhn-Tucker point of EQCP). □

6. Concluding Remarks. Motivated by the recent advances in linear programming research area (interior point methods and simplex type methods), we have presented a sequentially linear programming problem (SLPIHA) to solve the nonlinear equality constrained minimization problem EQCP. At each iteration the linear programming subproblem is solved within some tolerance. We proved, under rather weak hypotheses, that the SLPIHA Algorithm is globally convergent in the sense that any accumulation of the iteration sequence is a Karush-Kuhn-Tucker point of (EQCP).

Acknowledgment. This work was done partially while the author was visiting the Department of Computational and Applied Mathematics and the Center for Research on Parallel Computation at Rice University, Houston, Texas. The author would like to thank all the member of this department, especially Richard Tapia and Amr El-Bakry for fruitful discussions and a supportive and pleasant environment.

REFERENCES

topsep

[1] N. AlEXANDROV, *Multilevel Algorithms for Nonlinear Equations and Equality Constrained Optimization*, Ph.D. thesis, Rice University, Houston, Texas, Technical Report TR93-20, 1993.

[2] R.H. BYRD, R.B. SCHNABEL and G.A. SHULTZ, *A trust-region algorithm for nonlinearly constrained optimization,* SIAM J. Numer. Anal., 24 (1985), pp. 1152-170.

[3] R.H. BYRD, E.O. OMOJOJUN, R.B. SCHNABEL and G.A. SHULTZ, *Robust trust-region methods for nonlinearly constrained optimization*, presented at 1987 Conference on Optimization, Houston, TX 1987.

[4] M.R. CELIS, J.E. DENNIS and R.A. TAPIA, *A trust-region strategy for nonlinear equality constrained optimization*, in Numerical Optimization 1984, P. Boggs, R.H. Byrd and R. Schnabel, eds., Proceedings 20, Society for Industrial and Applied Mathematics, Philadelphia.

[5] J.E. DENNIS, M. EL-ALEM, and M.C. MACIEL, *A global convergence theory for general trust-region-based algorithm for equality constrained optimization*, Technical Report TR92-28 (1992), Department of Computational and Applied Mathematics, Rice university, Houston, Texas, 77251-1892.

[6] J.E. DENNIS, and M.C. VICENTE, *On the convergence theory of trust-region based algorithms for equality constrained optimization*, Technical Report TR92-94 (1994), Department of Computational and Applied Mathematics, Rice university, Houston, Texas, 77251-1892.

[7] M. EL-ALEM, *Robust trust-region algorithm with non parametric penalty parameter scheme for constrained optimization*, Technical Report CRPC-TR92245, Center for Research on Parallel Computation, Rice University, Houston, Texas, 77251-1892.

[8] M. EL HALLABI, *An inexact trust-region algorithm: Globalization of Newton's method*, Technical Report TR93-41 (1993), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892 (revised August 1994).

[9] M. EL HALLABI, *A global convergence theory for arbitrary norm trust-region algorithms for equality constrained optimization*, Technical Report TR93-60 (1993), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892 (revised may 1995).

[10] M. El Hallabi, *An arbitrary norm inexact hybrid algorithm for nonlinear equations* Technical Report TR94-26 (1994), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892 (revised may 1995).

[11] M. El Hallabi, *A necessary and sufficient condition for intersecting a translated hyperplane and a ball*, Technical Report TR95-14 (1995), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892 .

[12] M. El Hallabi,and R.A. TAPIA, *A global convergence theory for arbitrary norm trust-region methods for nonlinear equations*, Technical Report TR87-25 (1987), Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892 (revised as TR93-41).

[13] R. FLETCHER, *Practical Optimization*, John Wiley and Sons, New York, 1987.

[14] G. H. GOLUB, and C. F. VAN LOAN, *Matrix Computation*, John Hopkins University Press, Baltimore, Maryland, 1983.

[15] P. HUARD, *Optimization algorithms and Point-to-set maps*, Math. Programming, 8 (1975), pp. 308-331.

[16] M.C. MACIEL, *A global convergence theory for general class of trust-region algorithm for equality constrained optimization*, Ph.D. thesis, Rice University, Houston, Texas 77251-1892, 1992.

[17] L. MARUCHA, J. NOCEDAL and T. PLANTEGA, *On the implementation of an algorithm for large-scale equality optimization*, Technical Report NAM 09-93, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, Illinois.

[18] M.J.D. POWELL, *Convergence properties of a class of minimization algorithms*, in O.L. Mangasarian, R.R. Meyer, S.M. Robinson, eds., Nonlinear Programming 2, Academic Press, New York, 1975.

[19] M.J.D. POWELL, *General algorithms for discrete nonlinear approximation calculations*, Rep. DAMPT 1983/NA2, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.

[20] M.J.D. POWELL and Y. YUAN, *A trust-region algorithm for equality constrained optimization*, Math. Programming, 49 (1991), pp. 189-211.

[21] A. VARDI, *A trust-region algorithm for equality constrained minimization and implementation*, SIAM J. Numer. Anal. 22 (1985), pp. 575-591.