Zheng *Journal of Inequalities and Applications* 2012, **2012**:180
http://www.journalofinequalitiesandapplications.com/content/2012/1/180

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

A superlinearly convergent hybrid algorithm for systems of nonlinear equations

Lian Zheng*

*Correspondence:
zhenglian66@yahoo.com.cn
Department of Mathematics and Computer Science, Yangtze Normal University, Fuling, Chongqing, 408100, China

Abstract

We propose a new algorithm for solving systems of nonlinear equations with convex constraints which combines elements of Newton, the proximal point, and the projection method. The convergence of the whole sequence is established under weaker conditions than the ones used in existing projection-type methods. We study the superlinear convergence rate of the new method if in addition a certain error bound condition holds. Preliminary numerical experiments show that our method is efficient.

MSC: 90C25; 90C30

Keywords: nonlinear equations; projection method; global convergence; superlinear convergence

1 Introduction

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous mapping and $C \subset \mathbb{R}^n$ be a nonempty, closed, and convex set. The inner product and norm are denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. Consider the problem of finding

$$x^* \in C \quad \text{such that} \quad F(x^*) = 0. \tag{1.1}$$

Let S denote the solution set of (1.1). Throughout this paper, we assume that S is nonempty and F has the property that

$$\langle F(y), y - x^* \rangle \geq 0, \quad \text{for all } y \in C \text{ and all } x^* \in S. \tag{1.2}$$

The property (1.2) holds if F is monotone or more generally pseudomonotone on C in the sense of Karamardian [1].

Nonlinear equations have wide applications in reality. For example, many problems arising from chemical technology, economy, and communications can be transformed into nonlinear equations; see [2–5]. In recent years, many numerical methods for problem (1.1) with smooth mapping F have been proposed. These methods include the Newton method, quasi-Newton method, Levenberg-Marquardt method, trust region method, and their variants; see [6–14].

Recently, the literature [15] proposed a hybrid method for solving problem (1.1), which combines the Newton, proximal point, and projection methodologies. The method possesses a very nice globally convergent property if F is monotone and continuous. Under the assumptions of differentiability and nonsingularity, locally superlinear convergence of the method is proved. However, the condition of nonsingularity is too strong. Relaxing the nonsingularity assumption, the literature [16] proposed a modified version for the method by changing the projection way, and showed that under the local error bound condition which is weaker than nonsingularity, the proposed method converges superlinearly to the solution of problem (1.1). The numerical performances given in [16] show that the method is really efficient. However, the literatures [15, 16] need the mapping F to be monotone, which seems too stringent a requirement for the purpose of ensuring global convergence property and locally superlinear convergence of the hybrid method.

To further relax the assumption of monotonicity of F, in this paper, we propose a new hybrid algorithm for problem (1.1) which covers one in [16]. The global convergence of our method needs only to assume that F satisfies the property (1.2), which is much weaker than monotone or more generally pseudomonotone. We also discuss the superlinear convergence of our method under mild conditions. Preliminary numerical experiments show that our method is efficient.

2 Preliminaries and algorithms

For a nonempty, closed, and convex set $\Omega \subset \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, the projection of x onto Ω is defined as

$$\Pi_\Omega(x) = \arg\min\{\|y - x\| \,|\, y \in \Omega\}.$$

We have the following property on the projection operator; see [17].

Lemma 2.1 *Let $\Omega \subset \mathbb{R}^n$ be a closed convex set. Then it holds that*

$$\|x - \Pi_\Omega(y)\|^2 \le \|x - y\|^2 - \|y - \Pi_\Omega(y)\|^2, \quad \forall x \in \Omega, y \in \mathbb{R}^n.$$

Algorithm 2.1 Choose $x_0 \in C$, parameters $\kappa_0 \in [0,1)$, λ, $\beta \in (0,1)$, γ_1, $\gamma_2 > 0$, $a, b \ge 0$, $\max\{a,b\} > 0$, and set $k := 0$.

Step 1. Compute $F(x^k)$. If $F(x^k) = 0$, stop. Otherwise, let $\mu_k = \gamma_1\|F(x^k)\|^{1/2}$, $\sigma_k = \min\{\kappa_0, \gamma_2\|F(x^k)\|^{1/2}\}$. Choose a positive semidefinite matrix $G_k \in \mathbb{R}^{n\times n}$. Compute $d^k \in \mathbb{R}^n$ such that

$$F(x^k) + (G_k + \mu_k I)d^k = r^k, \tag{2.1}$$

where

$$\|r^k\| \le \sigma_k \mu_k \|d^k\|. \tag{2.2}$$

Stop if $d^k = 0$. Otherwise,

Step 2. Compute $y^k = x^k + t_k d^k$, where $t_k = \beta^{m_k}$ and m_k is the smallest nonnegative integer m satisfying

$$-\langle F(x^k + \beta^m d^k), d^k\rangle \ge \lambda(1 - \sigma_k)\mu_k \|d^k\|^2. \tag{2.3}$$

Step 3. Compute

$$x^{k+1} = \Pi_{C_k}\big(x^k - \alpha_k F(y^k)\big),$$

where $C_k := \{x \in C : h_k(x) \le 0\}$ and

$$h_k(x) := \big\langle aF(x^k) + bF(y^k), x - y^k\big\rangle + at_k\big\langle F(x^k), d^k\big\rangle, \qquad (2.4)$$
$$\alpha_k := \frac{\langle F(y^k), x^k - y^k\rangle}{\|F(y^k)\|^2}.$$

Let $k = k + 1$ and return to Step 1.

Remark 2.1 When we take parameters $a = 0$, $b = 1$, and the search direction $d^k = \bar{x}^k - x^k$, our algorithm degrades into one in [16]. At this step of getting the next iterate, our projection way and projection region are also different from the one in [15].

Now we analyze the feasibility of Algorithm 2.1. It is obvious that d^k satisfying conditions (2.1) and (2.2) exists. In fact, when we take $d^k = -(G_k + \mu_k I)^{-1}F(x^k)$, d^k satisfies (2.1) and (2.2). Next, we need only to show the feasibility of (2.3).

Lemma 2.2 *For all nonnegative integer k, there exists a nonnegative integer m_k satisfying (2.3).*

Proof If $d^k = 0$, then it follows from (2.1) and (2.2) that $F(x^k) = 0$, which means Algorithm 2.1 terminates with x^k being a solution of problem (1.1).

Now, we assume that $d^k \neq 0$, for all k. By the definition of r^k, the Cauchy-Schwarz inequality and the positive semidefiniteness of G_k, we have

$$\begin{aligned} -\big\langle F(x^k), d^k\big\rangle &= \big(d^k\big)^T(G_k + \mu_k I)\big(d^k\big) - \big(d^k\big)^T r^k \\ &\ge \mu_k\big\|d^k\big\|^2 - \big\|r^k\big\|\big\|d^k\big\| \\ &\ge (1-\sigma_k)\mu_k\big\|d^k\big\|^2. \end{aligned} \qquad (2.5)$$

Suppose that the conclusion of Lemma 2.2 does not hold. Then there exists a nonnegative integer $k_0 \ge 0$ such that (2.3) is not satisfied for any nonnegative integer m, *i.e.*,

$$-\big\langle F\big(x^{k_0} + \beta^m d^{k_0}\big), d^{k_0}\big\rangle < \lambda\mu_{k_0}(1-\sigma_{k_0})\big\|d^{k_0}\big\|^2, \quad \forall m. \qquad (2.6)$$

Letting $m \to \infty$ and by the continuity of F, we have

$$-\big\langle F\big(x^{k_0}\big), d^{k_0}\big\rangle \le \lambda\mu_{k_0}(1-\sigma_{k_0})\big\|d^{k_0}\big\|^2.$$

Which, together with (2.5), $d^{k_0} \neq 0$, and $\sigma_k \le \kappa_0 < 1$, we conclude that $\lambda \ge 1$, which contradicts $\lambda \in (0,1)$. This completes the proof. □

3 Convergence analysis

In this section, we first prove two lemmas, and then analyze the global convergence of Algorithm 2.1.

Lemma 3.1 *If the sequences* $\{x^k\}$ *and* $\{y^k\}$ *are generated by Algorithm 2.1,* $\{x^k\}$ *is bounded and F is continuous, then* $\{y^k\}$ *is also bounded.*

Proof Combining inequality (2.5) with the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\mu_k(1-\sigma_k)\|d^k\|^2 &\le -\langle F(x^k), d^k\rangle \\ &\le \|F(x^k)\|\,\|d^k\|.\end{aligned}$$

By the definition of μ_k and σ_k, it follows that

$$\|d^k\| \le \frac{\|F(x^k)\|}{\mu_k(1-\sigma_k)} \le \frac{\|F(x^k)\|^{1/2}}{\gamma_1(1-\kappa_0)}.$$

From the boundedness of $\{x_k\}$ and the continuity of F, we conclude that $\{d^k\}$ is bounded, and hence so is $\{y^k\}$. This completes the proof. □

Lemma 3.2 *Let* x^* *be a solution of problem (1.1) and the function* h_k *be defined by (2.4). If condition (1.2) holds, then*

$$h_k(x^k) \ge \lambda b t_k(1-\sigma_k)\mu_k\|d^k\|^2 \quad \text{and} \quad h_k(x^*) \le 0. \tag{3.1}$$

In particular, if $d^k \neq 0$, *then* $h_k(x^k) > 0$.

Proof

$$\begin{aligned}h_k(x^k) &= \langle aF(x^k) + bF(y^k), x^k - y^k\rangle + at_k\langle F(x^k), d^k\rangle \\ &= a\langle F(x^k), -t_k d^k\rangle + b\langle F(y^k), -t_k d^k\rangle + at_k\langle F(x^k), d^k\rangle \\ &= -bt_k\langle F(y^k), d^k\rangle \qquad (3.2)\\ &\ge \lambda b t_k(1-\sigma_k)\mu_k\|d^k\|^2, \qquad (3.3)\end{aligned}$$

where the inequality follows from (2.3).

$$\begin{aligned}h_k(x^*) &= \langle aF(x^k) + bF(y^k), x^* - y^k\rangle + at_k\langle F(x^k), d^k\rangle \\ &= a\langle F(x^k), x^* - x^k\rangle + a\langle F(x^k), x^k - y^k\rangle + b\langle F(y^k), x^* - y^k\rangle + at_k\langle F(x^k), d^k\rangle \\ &\le 0,\end{aligned}$$

where the inequality follows from condition (1.2) and the definition of y^k.

If $d^k \neq 0$, then $h_k(x^k) > 0$ because $\sigma_k \le \kappa_0 < 1$. The proof is completed. □

Remark 3.1 Lemma 3.2 means that the hyperplane

$$H_k := \{x \in \mathbb{R}^n \mid \langle aF(x^k) + bF(y^k), x - y^k\rangle + at_k\langle F(x^k), d^k\rangle = 0\}$$

strictly separates the current iterate from the solution set of problem (1.1).

Let $x^* \in S$ and $d^k \neq 0$. Since

$$\begin{aligned}\langle aF(x^k) + bF(y^k), x^k - x^*\rangle &= a\langle F(x^k), x^k - x^*\rangle + b\langle F(y^k), x^k - x^*\rangle \\ &= a\langle F(x^k), x^k - x^*\rangle + b\langle F(y^k), x^k - y^k\rangle + b\langle F(y^k), y^k - x^*\rangle \\ &\geq b\langle F(y^k), x^k - y^k\rangle \\ &\geq \lambda b t_k \mu_k (1-\sigma_k)\|d^k\|^2 \\ &> 0,\end{aligned}$$

where the first inequality follows from condition (1.2), the second one follows from (2.3), and the last one follows $d^k \neq 0$, which shows that $-(aF(x^k) + bF(y^k))$ is a descent direction of the function $\frac{1}{2}\|x - x^*\|^2$ at the point x^k.

We next prove our main result. Certainly, if Algorithm 2.1 terminates at Step k, then x^k is a solution of problem (1.1). Therefore, in the following analysis, we assume that Algorithm 2.1 always generates an infinite sequence.

Theorem 3.1 *If F is continuous on C, condition (1.2) holds and* $\sup_k \|G_k\| < \infty$, *then the sequence $\{x^k\} \subset \mathbb{R}^n$ generated by Algorithm 2.1 globally converges to a solution of (1.1).*

Proof Let x^* be a solution of problem (1.1). Since $x^{k+1} = \Pi_{C_k}(x^k - \alpha_k F(y^k))$, it follows from Lemma 2.1 that

$$\begin{aligned}\|x^{k+1} - x^*\|^2 &\leq \|x^k - \alpha_k F(y^k) - x^*\|^2 - \|x^{k+1} - x^k + \alpha_k F(y^k)\|^2 \\ &= \|x^k - x^*\|^2 - 2\alpha_k\langle F(y^k), x^k - x^*\rangle - \|x^{k+1} - x^k\|^2 - 2\alpha_k\langle F(y^k), x^{k+1} - x^k\rangle,\end{aligned}$$

i.e.,

$$\begin{aligned}\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 &\geq 2\alpha_k\langle F(y^k), x^k - x^*\rangle + \|x^{k+1} - x^k\|^2 + 2\alpha_k\langle F(y^k), x^{k+1} - x^k\rangle \\ &\geq 2\alpha_k\langle F(y^k), x^k - y^k\rangle + \|x^{k+1} - x^k + \alpha_k F(y^k)\|^2 - \alpha_k^2\|F(y^k)\|^2 \\ &\geq 2\alpha_k\langle F(y^k), x^k - y^k\rangle - \alpha_k^2\|F(y^k)\|^2 \\ &= \frac{\langle F(y^k), x^k - y^k\rangle^2}{\|F(y^k)\|^2},\end{aligned}$$

which shows that the sequence $\{\|x^{k+1} - x^*\|\}$ is nonincreasing, and hence is a convergent sequence. Therefore, $\{x^k\}$ is bounded and

$$\lim_{k\to\infty} \frac{\langle F(y^k), x^k - y^k\rangle^2}{\|F(y^k)\|^2} = 0. \tag{3.4}$$

From Lemma 3.1 and the continuity of F, we have that $\{F(y^k)\}$ is bounded; that is, there exists a positive constant M such that

$$\|F(y^k)\| \leq M, \quad \text{for all } k.$$

By (2.3) and the choices of σ_k and λ, we have

$$\begin{aligned}\frac{\langle F(y^k), x^k - y^k\rangle^2}{\|F(y^k)\|^2} &= \frac{t_k^2\langle F(y^k), d^k\rangle^2}{\|F(y^k)\|^2}\\ &\geq \frac{t_k^2\lambda^2(1-\sigma_k)^2\mu_k^2\|d^k\|^4}{M^2}\\ &\geq \frac{\lambda^2(1-\kappa_0)^2t_k^2\mu_k^2\|d^k\|^4}{M^2}.\end{aligned}$$

This, together with inequality (3.4), we deduce that

$$\lim_{k\to\infty} t_k\mu_k\|d^k\| = 0.$$

Now, we consider the following two possible cases:

Suppose first that $\limsup_{k\to\infty} t_k > 0$. Then we must have

$$\liminf_{k\to\infty}\mu_k = 0 \quad \text{or} \quad \liminf_{k\to\infty}\|d^k\| = 0.$$

From the definition of μ_k, the choice of d^k and $\sup_k \|G_k\| < \infty$, each case of them follows that

$$\liminf_{k\to\infty}\|F(x^k)\| = 0.$$

Since F is continuous and $\{x^k\}$ is bounded, which implies that the sequence $\{x^k\}$ has some accumulation point $\hat{x}$ such that

$$F(\hat{x}) = 0.$$

This shows that $\hat{x}$ is a solution of problem (1.1). Replacing x^* by $\hat{x}$ in the preceding argument, we obtain that the sequence $\{\|x^k - \hat{x}\|\}$ is nonincreasing, and hence converges. Since $\hat{x}$ is an accumulation point of $\{x_k\}$, some subsequence of $\{\|x^k - \hat{x}\|\}$ converges to zero, which implies that the whole sequence $\{\|x^k - \hat{x}\|\}$ converges to zero, and hence $\lim_{k\to\infty} x^k = \hat{x}$.

Suppose now that $\lim_{k\to\infty} t_k = 0$. Let $\bar{x}$ be any accumulation point of $\{x^k\}$ and $\{x^{k_j}\}$ be the corresponding subsequence converging to $\bar{x}$. By the choice of t_k, (2.3) implies that

$$-\langle F(x^{k_j} + t_{k_j}\beta^{-1}d^{k_j}), d^{k_j}\rangle < \lambda(1-\sigma_{k_j})\mu_{k_j}\|d^{k_j}\|^2, \quad \text{for all } j.$$

Since F is continuous, we obtain by letting $j \to \infty$ that

$$-\langle F(x^{k_j}), d^{k_j}\rangle \leq \lambda(1-\sigma_{k_j})\mu_{k_j}\|d^{k_j}\|^2. \tag{3.5}$$

From (2.5) and (3.5), we conclude that $\lambda \geq 1$, which contradicts $\lambda \in (0,1)$. Hence, the case of $\lim_{k\to\infty} t_k = 0$ is not possible. This completes the proof. □

Remark 3.2 Compared to the conditions of the global convergence used in literatures [15, 16], our conditions are weaker.

4 Convergence rate

In this section, we provide a result on the convergence rate of the iterative sequence generated by Algorithm 2.1. To establish this result, we need the following conditions (4.1) and (4.2).

For $x^* \in S$, there are positive constants δ, c_1, and c_2 such that

$$c_1 \operatorname{dist}(x, S) \le \|F(x)\|, \quad \forall x \in N(x^*, \delta), \tag{4.1}$$

and

$$\|F(x) - F(y) - G_k(x-y)\| \le c_2 \|x-y\|^2, \quad \forall x, y \in N(x^*, \delta), \tag{4.2}$$

where $\operatorname{dist}(x, S)$ denotes the distance from x to solution set S, and

$$N(x^*, \delta) = \{x \in R^n \mid \|x - x^*\| \le \delta\}.$$

If F is differentiable and $\nabla F(\cdot)$ is locally Lipschitz continuous with modulus $\theta > 0$, then there exists a constant $L_1 > 0$ such that

$$\|F(y) - F(x) - \nabla F(x)(y-x)\| \le L_1 \|y-x\|^2, \quad \forall x, y \in N(x^*, \delta). \tag{4.3}$$

In fact, by the mean value theorem of vector valued function, we have

$$\begin{aligned}
&\|F(y) - F(x) - \nabla F(x)(y-x)\| \\
&\quad = \left\| \int_0^1 \nabla F(\tau y + (1-\tau)x)(y-x)\,d\tau - \int_0^1 \nabla F(x)(y-x)\,d\tau \right\| \\
&\quad \le \int_0^1 \|\nabla F(\tau y + (1-\tau)x) - \nabla F(x)\| \|y-x\|\,d\tau \\
&\quad \le \theta \|y-x\|^2 \int_0^1 \tau\,d\tau \\
&\quad = L_1 \|y-x\|^2,
\end{aligned}$$

where $L_1 = \theta/2$. Under assumptions (4.2) or (4.3), it is readily shown that there exists a constant $L_2 > 0$ such that

$$\|F(y) - F(x)\| \le L_2 \|y-x\|, \quad \forall x, y \in N(x^*, \delta). \tag{4.4}$$

In 1998, the literature [15] showed that their proposed method converged superlinearly when the underlying function F is monotone, differentiable with $\nabla F(x^*)$ being nonsingular, and ∇F is locally Lipschitz continuous. It is known that the local error bound condition given in (4.1) is weaker than the nonsingular. Recently, under conditions (4.1), (4.2), and the underlying function F being monotone and continuous, the literature [16] obtained the locally superlinear rate of convergence of the proposed method.

Next, we analyze the superlinear convergence rate of the iterative sequence under a weaker condition. In the rest of section, we assume that $x^k \to x^*$, $k \to \infty$, where $x^* \in S$.

Lemma 4.1 *Let $G \in R^{n\times n}$ be a positive semidefinite matrix and $\mu > 0$. Then*
(1) $\|(G+\mu I)^{-1}\| \le \frac{1}{\mu}$;
(2) $\|(G+\mu I)^{-1}G\| \le 2$.

Proof See [18]. □

Lemma 4.2 *Suppose that F is continuous and satisfies conditions (1.2), (4.1), and (4.2). If there exists a positive constant N such that $\|G_k\| \le N$ for all k, then for all k sufficiently large,*
(1) $c_3\|d^k\| \le \|F(x^k)\| \le c_4\|d^k\|$;
(2) $\|F(x^k)+G_kd^k\| \le c_5\|d^k\|^{3/2}$, *where c_3, c_4 and c_5 are all positive constants.*

Proof For (1), let $x^k \in N(x^*, \frac{1}{2}\delta)$ and $\hat{x}^k \in S$ be the closest solution to x^k. We have

$$\|\hat{x}^k - x^*\| \le \|\hat{x}^k - x^k\| + \|x^k - x^*\| \le \delta,$$

i.e., $\hat{x}^k \in N(x^*, \delta)$. Thus, by (2.1), (2.2), (4.2), and Lemma 4.1, we have

$$\begin{aligned}\|d^k\| &\le \|(G_k+\mu_kI)^{-1}F(x^k)\| + \|(G_k+\mu_kI)^{-1}r^k\| \\ &\le \|(G_k+\mu_kI)^{-1}[F(\hat{x}^k)-F(x^k)-G_k(\hat{x}^k-x^k)]\| \\ &\quad + \|(G_k+\mu_kI)^{-1}G_k(\hat{x}^k-x^k)\| + \frac{1}{\mu_k}\|r^k\| \\ &\le \frac{c_2}{\mu_k}\|\hat{x}^k-x^k\|^2 + 2\|\hat{x}^k-x^k\| + \sigma_k\|d^k\|.\end{aligned}$$

By $\|x^k-\hat{x}^k\| = \operatorname{dist}(x^k,S)$ and $\sigma_k \le \kappa_0$, it follows that

$$(1-\kappa_0)\|d^k\| \le \left(\frac{c_2}{\mu_k}\operatorname{dist}(x^k,S)+2\right)\operatorname{dist}(x^k,S).$$

From (4.1) and the choice of μ_k, it holds that

$$\begin{aligned}\frac{c_2}{\mu_k}\operatorname{dist}(x^k,S) &\le \frac{c_1^{-1}c_2\|F(x^k)\|}{\gamma_1\|F(x^k)\|^{1/2}} \\ &= \frac{c_2}{\gamma_1c_1}\|F(x^k)\|^{1/2}.\end{aligned}$$

From the boundedness of $\{\|F(x^k)\|\}$, there exists a positive constant M_1 such that

$$\|F(x^k)\|^{1/2} \le M_1.$$

Therefore,

$$\begin{aligned}\|d^k\| &\le \frac{c_2M_1+2\gamma_1c_1}{c_1\gamma_1(1-\kappa_0)}\operatorname{dist}(x^k,S) \\ &\le \frac{c_2M_1+2\gamma_1c_1}{c_1^2\gamma_1(1-\kappa_0)}\|F(x^k)\|.\end{aligned} \tag{4.5}$$

We obtain that the left-hand side of (1) by setting $c_3 := \frac{c_1^2\gamma_1(1-\kappa_0)}{c_2M_1+2\gamma_1c_1}$.

For the right-hand side part, it follows from (2.1) and (2.2) that

$$\begin{aligned}\|F(x^k)\| &\le \|G_k + \mu_k I\|\|d^k\| + \|r^k\| \\ &\le (\|G_k + \mu_k I\| + \sigma_k\mu_k)\|d^k\| \\ &\le (N + \gamma_1 M_1 + \kappa_0\gamma_1 M_1)\|d^k\|.\end{aligned}$$

We obtain the right-hand side part by setting $c_4 := N + \gamma_1 M_1 + \kappa_0\gamma_1 M_1$.

For (2), using (2.1) and (2.2), we have

$$\begin{aligned}\|F(x^k) + G_k d^k\| &\le \mu_k\|d^k\| + \|r^k\| \\ &\le (1+\sigma_k)\mu_k\|d^k\| \\ &\le (1+\kappa_0)\gamma_1\|F(x^k)\|^{1/2}\|d^k\| \\ &\le (1+\kappa_0)\gamma_1 c_4^{1/2}\|d^k\|^{3/2}.\end{aligned}$$

By setting $c_5 := (1+\kappa_0)\gamma_1 c_4^{1/2}$, we obtain the desired result. □

Lemma 4.3 *Suppose that the assumptions in Lemma 4.2 hold. Then for all k sufficiently large, it holds that*

$$y^k = x^k + d^k.$$

Proof By $\lim_{k\to\infty} x^k = x^*$ and the continuity of F, we have

$$\lim_{k\to\infty} F(x^k) = F(x^*) = 0.$$

By Lemma 4.2(1), we obtain that

$$\lim_{k\to\infty} \|d^k\| = 0,$$

which means that $x^k + d^k \in N(x^*, \delta)$ for all k sufficiently large. Hence, it follows from (4.2) that

$$F(x^k + d^k) = F(x^k) + G_k d^k + R^k, \tag{4.6}$$

where $\|R^k\| \le c_2\|d^k\|^2$. Using (2.1) and (2.2), (4.6) can be written as

$$F(x^k + d^k) = -\mu_k d^k + r^k + R^k. \tag{4.7}$$

Hence,

$$\begin{aligned}-\langle F(x^k + d^k), d^k\rangle &= \langle \mu_k d^k, d^k\rangle - r^k d^k - R^k d^k \\ &\ge \mu_k\|d^k\|^2 - \sigma_k\mu_k\|d^k\|^2 - c_2\|d^k\|^3 \\ &= \left(1 - \frac{c_2\|d^k\|}{\mu_k(1-\sigma_k)}\right)\mu_k(1-\sigma_k)\|d^k\|^2.\end{aligned}$$

By Lemma 4.2(1) and the choices of μ_k and σ_k, for k sufficiently large, we obtain

$$\begin{aligned} 1 &\geq 1 - \frac{c_2\|d^k\|}{\mu_k(1-\sigma_k)} \\ &\geq 1 - \frac{c_2 c_3^{-1}\|F(x^k)\|}{(1-\kappa_0)\gamma_1\|F(x^k)\|^{1/2}} \\ &= 1 - \frac{c_2 c_3^{-1}\|F(x^k)\|^{1/2}}{(1-\kappa_0)\gamma_1} \geq \lambda, \end{aligned}$$

where the last inequality follows from $\lim_{k\to\infty} F(x^k) = 0$.

Therefore,

$$-\langle F(x^k + d^k), d^k\rangle \geq \lambda\mu_k(1-\sigma_k)\|d^k\|^2,$$

which implies that (2.3) holds with $t_k = 1$ for all k sufficiently large, *i.e.*, $y^k = x^k + d^k$. This completes the proof. □

From now on, we assume that k is large enough so that $y^k = x^k + d^k$.

Lemma 4.4 *Suppose that the assumptions in Lemma 4.2 hold. Set* $\tilde{x}^k := x^k - \alpha_k F(y^k)$. *Then for all k sufficiently large, there exists a positive constant* c_6 *such that*

$$\|\tilde{x}^k - y^k\| \leq c_6\|d^k\|^{3/2}.$$

Proof Set

$$H_k^1 = \{x \in R^n \mid \langle F(y^k), x - y^k\rangle = 0\}.$$

Then $\tilde{x}^k = \Pi_{H_k^1}(x^k)$ and $y^k \in H_k^1$. Hence, the vectors $x^k - \tilde{x}^k$ and $y^k - \tilde{x}^k$ are orthogonal. That is,

$$\|y^k - \tilde{x}^k\| = \|y^k - x^k\|\sin\theta_k = \|d^k\|\sin\theta_k, \tag{4.8}$$

where θ_k is the angle between $\tilde{x}^k - x^k$ and $y^k - x^k$. Because $\tilde{x}^k - x^k = -\alpha_k F(y^k)$ and $y^k - x^k = d^k$, the angle between $F(y^k)$ and $-\mu_k d^k$ is also θ_k. By (4.7), we obtain

$$F(y^k) - (-\mu_k d^k) = R^k + r^k,$$

which implies that the vectors $F(y^k)$, $-\mu_k d^k$ and $R^k + r^k$ constitute a triangle. Since $\lim_{k\to\infty}\mu_k = \lim_{k\to\infty}\gamma_1\|F(x^k)\|^{1/2} = 0$ and $\lim_{k\to\infty}\alpha_k = 0$. So for all k sufficiently large, we have

$$\begin{aligned} \sin\theta_k &\leq \frac{\|r^k + R^k\|}{\mu_k\|d^k\|} \\ &\leq \sigma_k + \frac{c_2\|d^k\|}{\mu_k} \end{aligned}$$

$$\begin{aligned} &\le \gamma_2 \| F(x^k) \|^{1/2} + \frac{c_2 \|F(x^k)\|}{c_3 \gamma_1 \|F(x^k)\|^{1/2}} \\ &= \left(\gamma_2 + \frac{c_2}{c_3 \gamma_1}\right) \| F(x^k) \|^{1/2}, \end{aligned}$$

which, together with (4.8) and Lemma 4.2(1), we obtain

$$\begin{aligned} \| y^k - \tilde{x}^k \| &\le \left(\gamma_2 + \frac{c_2}{c_3 \gamma_1}\right) \| F(x^k) \|^{1/2} \| d^k \| \\ &\le c_6 \| d^k \|^{3/2}, \end{aligned}$$

where $c_6 = c_4^{1/2}(\gamma_2 + \frac{c_2}{c_3 \gamma_1})$. This completes the proof. □

Now, we turn our attention to local rate of convergence analysis.

Theorem 4.1 *Suppose that the assumptions in Lemma 4.2 hold. Then the sequence* $\{\operatorname{dist}(x^k, S)\}$ *Q-superlinearly converges to* 0.

Proof By the definition of $\tilde{x}^k$, Lemma 4.2(1) and (4.4), for sufficiently large k, we have

$$\begin{aligned} \| \tilde{x}^k - x^* \| &\le \| x^k - \alpha_k F(y^k) - x^* \| \\ &\le \| x^k - x^* \| + \left\| \frac{\langle F(y^k), x^k - y^k \rangle}{\|F(y^k)\|^2} F(y^k) \right\| \\ &\le \| x^k - x^* \| + \| d^k \| \\ &\le \| x^k - x^* \| + c_3^{-1} \| F(x^k) \| \\ &= \| x^k - x^* \| + c_3^{-1} \| F(x^k) - F(x^*) \| \\ &\le (1 + L_2 c_3^{-1}) \| x^k - x^* \|, \end{aligned}$$

which implies that $\lim_{k\to\infty} \|\tilde{x}^k - x^*\| = 0$ because $\lim_{k\to\infty} \|x^k - x^*\| = 0$. Thus, $\tilde{x}^k \in N(x^*, \delta)$ for k sufficiently large, which, together with (4.2), Lemma 4.2, Lemma 4.4, and the definition of $\tilde{x}^k$, we obtain

$$\begin{aligned} \| F(\tilde{x}^k) \| &\le \| F(x^k) + G_k(\tilde{x}^k - x^k) \| + c_2 \| \tilde{x}^k - x^k \|^2 \\ &\le \| F(x^k) + G_k(y^k - x^k) \| + \|G_k\| \| \tilde{x}^k - y^k \| + c_2 \| \tilde{x}^k - x^k \|^2 \\ &\le c_5 \| d^k \|^{3/2} + N c_6 \| d^k \|^{3/2} + c_2 \| \alpha_k F(y^k) \|^2 \\ &\le (c_5 + N c_6) \| d^k \|^{3/2} + c_2 \| d^k \|^2 \\ &= (c_5 + N c_6 + c_2 \| d^k \|^{1/2}) \| d^k \|^{3/2} \\ &\le (c_5 + N c_6 + c_2 c_3^{-1/2} \| F(x^k) \|^{1/2}) \| d^k \|^{3/2}. \end{aligned}$$

Because $\{\|F(x^k)\|\}$ is bounded, there exists a positive constant c_7 such that

$$\| F(\tilde{x}^k) \| \le c_7 \| d^k \|^{3/2}. \tag{4.9}$$

On the other hand, from Lemma 3.2, we know that

$$S \subseteq C \cap H_k,$$

where S is the solution set of problem (1.1). Since $x^{k+1} = \Pi_{C\cap H_k}(\tilde{x}^k)$, it follows from Lemma 2.1 that

$$\|x^{k+1} - x^*\|^2 \le \|\tilde{x}^k - x^*\|^2 - \|x^{k+1} - \tilde{x}^k\|^2, \quad \forall x^* \in S,$$

which implies that

$$\|x^{k+1} - x^*\| \le \|\tilde{x}^k - x^*\|.$$

Therefore, together with inequalities (4.1), (4.5), and (4.9), we have

$$\begin{aligned}\operatorname{dist}(x^{k+1}, S) &\le \operatorname{dist}(\tilde{x}^k, S) \le \frac{1}{c_1}\|F(\tilde{x}^k)\| \\ &\le \frac{c_7}{c_1}\|d^k\|^{3/2} \le \frac{c_7}{c_1}\left(\frac{c_2 M_1 + 2\gamma_1 c_1}{c_1\gamma_1(1-\kappa_0)}\right)^{3/2} \operatorname{dist}^{3/2}(x^k, S),\end{aligned}$$

which implies that the order of superlinear convergence is at least 1.5. This completes the proof. □

Remark 4.1 Compared with the proof of the locally superlinear convergence in literatures [15, 16], our conditions are weaker.

5 Numerical experiments

In this section, we present some numerical experiments results to show the efficiency of our method. The MATLAB codes are run on a notebook computer with CPU2.10GHZ under MATLAB Version 7.0. Just as done in [16], we take $G_k = F'(x^k)$ and use the left division operation in MATLAB to solve the system of linear equations (2.1) at each iteration. We choose $b = 1$, $\lambda = 0.96$, $\kappa_0 = 0$, $\beta = 0.7$, and $\gamma_1 = 1$. 'Iter.' denotes the number of iteration and 'CPU' denotes the CPU time in seconds. We choose $\|F(x^k)\| \le 10^{-6}$ as the stop criterion. The example is tested in [16].

Example Let

$$F(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} x_1^3 \\ x_2^3 \\ 2x_3^3 \\ 2x_4^3 \end{pmatrix} + \begin{pmatrix} -10 \\ 1 \\ -3 \\ 0 \end{pmatrix}$$

and the constraint set C be taken as

$$C = \left\{ x \in \mathbb{R}^4 \,\middle|\, \sum_{i=1}^{4} x_i \le 3, x_i \ge 0, i = 1, 2, 3, 4 \right\}.$$

Table 1 Numerical results of Example with $a = 10^{-15}$

Initial point	Iter.	CPU	$\|F(x^*)\|$
(3,0,0,0)	11	0.10	1.07×10^{-8}
(1,1,0,0)	13	0.09	1.62×10^{-9}
(0,1,0,1)	15	0.04	2.46×10^{-9}
(0,0,0,1)	21	0.18	9.92×10^{-10}
(1,0,0,2)	16	0.54	5.66×10^{-10}

Table 2 Numerical results of Example with $a = 0$

Initial point	Iter.	CPU	$\|F(x^*)\|$
(3,0,0,0)	11	0.10	1.07×10^{-8}
(1,1,0,0)	13	0.12	1.62×10^{-9}
(0,1,0,1)	19	0.14	1.17×10^{-9}
(0,0,0,1)	18	0.18	1.44×10^{-9}
(1,0,0,2)	15	0.21	7.88×10^{-9}

From Tables 1-2, we can see that our algorithm is efficient if parameters are chosen properly. We can also observe that the algorithm's operation results change with the value of a. When we take $a = 0$, the operation results are not best, that is to say, the direction $F(y^k)$ is not an optimal one.

Competing interests
The author declares that they have no competing interests.

Acknowledgements
The author wish to thank the anonymous referees for their suggestions and comments. This work is also supported by the Educational Science Foundation of Chongqing, Chongqing of China (Grant No. KJ111309).

Received: 28 March 2012 Accepted: 6 August 2012 Published: 24 August 2012

References

1. Karamardian, S: Complementarity problems over cones with monotone and pseudomonotone maps. J. Optim. Theory Appl. **18**, 445-454 (1976)
2. Dirkse, SP, Ferris, MC: MCPLIB: a collection of nonlinear mixed complementarity problems. Optim. Methods Softw. **5**, 319-345 (1995)
3. El-Hawary, ME: Optimal Power Flow: Solution Techniques, Requirement and Challenges. IEEE Service Center, Piscataway (1996)
4. Meintjes, K, Morgan, AP: A methodology for solving chemical equilibrium system. Appl. Math. Comput. **22**, 333-361 (1987)
5. Wood, AJ, Wollenberg, BF: Power Generations, Operations, and Control. Wiley, New York (1996)
6. Bertsekas, DP: Nonlinear Programming. Athena Scientific, Belmont (1995)
7. Dennis, JE, Schnabel, RB: Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice Hall, Englewood Cliffs (1983)
8. Ortega, JM, Rheinboldt, WC: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, San Diego (1970)
9. Polyak, BT: Introduction to Optimization. Optimization Software, Inc. Publications Division, New York (1987)
10. Tong, XJ, Qi, L: On the convergence of a trust-region method for solving constrained nonlinear equations with degenerate solution. J. Optim. Theory Appl. **123**, 187-211 (2004)
11. Zhang, JL, Wang, Y: A new trust region method for nonlinear equations. Math. Methods Oper. Res. **58**, 283-298 (2003)
12. Fan, JY, Yuan, YX: On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption. Computing **74**, 23-39 (2005)
13. Fan, JY: Convergence rate of the trust region method for nonlinear equations under local error bound condition. Comput. Optim. Appl. **34**, 215-227 (2006)
14. Fan, JY, Pan, JY: An improved trust region algorithm for nonlinear equations. Comput. Optim. Appl. **48**, 59-70 (2011)
15. Solodov, MV, Svaiter, BF: A globally convergent inexact Newton method for systems of monotone equations. In: Fukushima, M, Qi, L (eds.) Reformulation: Piecewise Smooth, Semismooth and Smoothing Methods, pp. 355-369. Kluwer Academic, Dordrecht (1998)
16. Wang, CW, Wang, YJ: A superlinearly convergent projection method for constrained systems of nonlinear equations. J. Glob. Optim. **44**, 283-296 (2009)
17. Zarantonello, EH: Projections on Convex Sets in Hilbert Spaces and Spectral Theory. Academic Press, New York (1971)
18. Zhou, GL, Toh, KC: Superlinear convergence of a Newton-type algorithm for monotone equations. J. Optim. Theory Appl. **125**, 205-221 (2005)

doi:10.1186/1029-242X-2012-180
Cite this article as: Zheng: A superlinearly convergent hybrid algorithm for systems of nonlinear equations. *Journal of Inequalities and Applications* 2012 **2012**:180.