Globally Convergent Broyden-like Methods for Semismooth Equations and Applications to VIP, NCP and MCP [1]

Donghui Li
Department of Applied Mathematics
Hunan University
Changsha, China 410082
e-mail: dhli@mail.hunu.edu.cn

Masao Fukushima
Department of Applied Mathematics and Physics
Graduate School of Informatics
Kyoto University
Kyoto 606-8501, Japan
e-mail: fuku@kuamp.kyoto-u.ac.jp

September 26, 1999

Abstract In this paper, we propose a general smoothing Broyden-like quasi-Newton method for solving a class of nonsmooth equations. Under appropriate conditions, the proposed method converges to a solution of the equation globally and superlinearly. In particular, the proposed method provides the possibility of developing a quasi-Newton method that enjoys superlinear convergence even if strict complementarity fails to hld. We pay particular attention to semismooth equations arising from nonlinear complementarity problems, mixed complementarity problems and variational inequality problems. We show that under certain conditions, the related methods based on the perturbed Fischer-Burmeister function, Chen-Harker-Kanzow-Smale smoothing function and Gabriel-Moré class of smoothing functions converge globally and superlinearly.

Key words: Nonsmooth equations, Broyden-like method, global convergence, superlinear convergence.

[1] The work of the first author was done while he was visiting Kyoto University. This paper is an extended version of the unpublished technical report [25] by the authors.

1 Introduction

Let H be a function from R^n into itself that is not necessarily differentiable in the sense of Frechét but semismooth in the sense of Mifflin [27] and Qi [29]. We are concerned with the solution of the semismooth equation

$$H(x) = 0. \tag{1.1}$$

Semismooth equations have a rich practical background. Many problems such as nonlinear complementarity problems, mixed complementarity problems and variational inequality problems can be reformulated as semismooth equations.

Newton's method for solving (1.1) has been extensively studied in recent years. In particular, Chen et al. [5] proposed a smoothing Newton method for solving a semismooth equation reformulation arising from the mixed complementarity problem that possesses global and superlinear convergence properties. We refer to [14] and [31] for reviews about Newton-type methods for nonsmooth and semismooth equations. Local convergence of quasi-Newton methods for nonsmooth and semismooth equations has also been studied by some authors [3, 4, 17, 21, 28, 33]. There are a few papers that deal with globally convergent quasi-Newton methods for semismooth equations arising from nonlinear complementarity problems and mixed complementarity problems [24, 26, 35]. In [26], by using a splitting technique of [30], a well defined quasi-Newton method that possesses global and superlinear convergence properties was first proposed for nonlinear complementarity problems. In [24], a smoothing quasi-Newton method was developed for a semismooth equation reformulation of the mixed complementarity problem. The method proposed in [24] is an improvement of the method proposed in [26]. In particular, when applied to nonlinear complementarity problems, the method proposed in [24] converges globally whenever the problem involves a P_0 function with Lipschitz continuous derivatives. Moreover, if strict complementarity condition holds at the limit point, then the method converges superlinearly.

The purpose of this paper is to develop a general smoothing quasi-Newton approach to the semismooth equation (1.1). Under appropriate conditions, the proposed method converges globally and superlinearly. In particular, the results obtained in this paper provides the possibility of developing a globally convergent quasi-Newton method that enjoys superlinear convergence even if strict complementarity fails to hold at the limit point. We pay particular attention to semismooth equations arising from nonlinear complementarity problems, mixed complementarity problems and variational inequality problems. The smoothing functions considered include the perturbed Fischer-Burmeister function, Chen-Harker-Kanzow-Smale smoothing function and Gabriel-Moré class of smoothing functions.

The paper is organized as follows: In the next section, we propose a general quasi-Newton approach to the solution of semismooth equations. In Sections 3 and 4, we prove global and superlinear convergence of a general Broyden-like method, respectively. In Section 5, we specialize the method to the nonlinear complementarity problem, the variational inequality problem and the mixed complementarity problem. Section 6 concludes the paper.

2 General Quasi-Newton Approach

This section develops a general quasi-Newton approach to the solution of equation (1.1). Throughout, we assume that H is a semismooth function. The proposed method is a smoothing quasi-Newton method. The feature of smoothing methods is to construct a smooth function H^ϵ as an approximation of H satisfying $H^0(z) = H(z)$ for any $z \in R^n$ and

$$\|H^\epsilon(z) - H^{\epsilon'}(z)\| \leq \mu(\epsilon - \epsilon'), \quad \forall \epsilon > \epsilon' \geq 0, \qquad \forall z \in R^n, \tag{2.1}$$

and then to generate a sequence $\{x_k\}$ where x_k is a solution of the smooth equation

$$H^{\epsilon_k}(x) = 0$$

with $\epsilon_k \to 0$, so that the sequence $\{x_k\}$ may converge to a solution of (1.1). Smoothing Newton methods have been received much attention in the literature [1, 2, 5, 15, 19, 30, 31]. The quasi-Newton method developed here is based on the smoothing Newton method in [5]. For simplicity, we abbreviate H^{ϵ_k} as H^k. In the proposed method, the subproblem is the linear equation

$$B_k p + H^0(x_k) = 0, \tag{2.2}$$

where matrix $B_k \in R^{n\times n}$ is an approximation of $\nabla H^k(x_k)$. If $B_k = \nabla H^k(x_k)$, the linear equation (2.2) reduces to the subproblem of smoothing Newton methods [5, 30].

Unlike Newton's method, in a quasi-Newton method, derivative-free line search is indispensable because the computation of derivatives should be avoided. We now propose a derivative-free line search that is similar to those in [24]. Let $\{\eta_k\}$ be a positive sequence satisfying

$$\sum_{k=1}^{\infty} \eta_k \leq \eta < \infty, \tag{2.3}$$

where η is a positive constant. At iteration k, given the current iterate z_k and a direction p_k, we determine a positive steplength $\alpha_k > 0$ such that the following inequality holds for $\alpha = \alpha_k$:

$$\begin{aligned} &\|H^k(z_k + \alpha p_k)\| - \|H^k(z_k)\| \\ &\quad \leq \ -\sigma_1\|H^k(z_k + \alpha p_k) - H^k(z_k)\| - \sigma_2\|\alpha p_k\|^2 + \eta_k, \end{aligned} \tag{2.4}$$

where $\sigma_1, \sigma_2 > 0$ are suitable constants. Note that (2.4) is satisfied for all sufficiently small $\alpha > 0$ since η_k is positive and independent of α. Thus α_k can be obtained by a backtracking process.

We state a general smoothing quasi-Newton method for (1.1) as follows.

Algorithm 1

Step 0. Choose constants $\rho, \beta \in (0,1)$, $0 < \gamma < 1/\mu$, $\sigma_1 \in (0, \frac{1-\mu\gamma}{1+\mu\gamma})$, $\sigma_2 > 0$, $\bar{\theta} \in (0,1)$. Select a positive sequence $\{\eta_k\}$ satisfying (2.3). Choose an initial point $z_1 \in R^n$ and an initial nonsingular matrix $B_1 \in R^{n\times n}$. Choose an $\epsilon_1 > 0$ satisfying $\epsilon_1 \le \frac{\gamma}{2}\|H^0(z_1)\|$. Let $k := 1$.

Step 1. If $H^0(z_k) = 0$, stop. Otherwise, solve the linear equation (2.2) to get p_k.

Step 2. Let α_k be the maximum element in the set $\{1, \rho, \rho^2, \ldots\}$ such that $\alpha = \rho^i$ satisfies the line search condition (2.4).

Step 3. Let $z_{k+1} = z_k + \alpha_k p_k$.

Step 4. Update B_k to get B_{k+1}.

Step 5. Update ϵ_{k+1} by the following rule: If $\epsilon_k < \gamma\|H^0(z_{k+1})\|$, set $\epsilon_{k+1} = \epsilon_k$. Otherwise, choose ϵ_{k+1} such that

$$0 < \epsilon_{k+1} \le \min\{\frac{\gamma}{2}\|H^0(z_{k+1})\|, \frac{1}{2}\epsilon_k\}.$$

Step 6. Let $k := k+1$. Go to Step 1.

Remark Step 5 of Algorithm 1 ensures that $\{\epsilon_k\}$ is monotonically nonincreasing and $\epsilon_k \le \gamma\|H^0(z_k)\|$ holds for every k. Moreover, if the algorithm does not terminate finitely, i.e., $H^0(z_k) \neq 0$ for all k, then $\epsilon_k > 0$ for all k. We also notice that $\{z_k\}$ satisfies the inequality

$$\|H^k(z_{k+1})\| - \|H^k(z_k)\| \le -\sigma_1\|H^k(z_{k+1}) - H^k(z_k)\| - \sigma_2\|s_k\|^2 + \eta_k,$$

where $s_k = z_{k+1} - z_k$. By (2.1), it then follows that for every k

$$\begin{aligned} \|H^{k+1}(z_{k+1})\| &\le \|H^k(z_{k+1})\| + \|H^{k+1}(z_{k+1}) - H^k(z_{k+1})\| \\ &\le \|H^k(z_{k+1})\| + \mu(\epsilon_k - \epsilon_{k+1}) \\ &\le \|H^k(z_k)\| - \sigma_1\|H^k(z_{k+1}) - H^k(z_k)\| - \sigma_2\|s_k\|^2 \\ &\quad + \eta_k + \mu(\epsilon_k - \epsilon_{k+1}). \end{aligned} \tag{2.5}$$

Moreover, (2.5) implies

$$\|H^{k+1}(z_{k+1})\| \le \|H^k(z_k)\| + \eta_k + \mu(\epsilon_k - \epsilon_{k+1}). \tag{2.6}$$

Since η_k satisfies (2.3), we immediately get the following result.

Lemma 2.1 *Let $\{z_k\}$ be generated by Algorithm 1. Then $\{\|H^k(z_k)\|\}$ converges. Moreover, we have*

$$\sum_{k=1}^{\infty} \|H^k(z_{k+1}) - H^k(z_k)\| < \infty \tag{2.7}$$

and

$$\sum_{k=1}^{\infty} \|s_k\|^2 = \sum_{k=1}^{\infty} \|z_{k+1} - z_k\|^2 < \infty. \tag{2.8}$$

Proof Since (2.3) implies

$$\sum_{k=1}^{\infty}\{\eta_k + \mu(\epsilon_k - \epsilon_{k+1})\} < \infty,$$

it follows from (2.6) that $\{\|H^k(z_k)\|\}$ converges. The inequality (2.7) then follows from (2.5) directly. Moreover, by (2.5), it is easy to see that

$$\sigma_2\|s_k\|^2 \leq \|H^k(z_k)\| - \|H^{k+1}(z_{k+1})\| + \eta_k + \mu(\epsilon_k - \epsilon_{k+1}).$$

Summing the above inequalities yields (2.8). □

Define the level set

$$\Omega = \{z \in R^n \mid \|H^0(z)\| \leq \frac{1+\mu\gamma}{1-\mu\gamma}\|H^0(z_1)\| + \frac{1}{1-\mu\gamma}\eta\}, \tag{2.9}$$

where η is a positive constant such that (2.3) holds.

Notice that (2.1) particularly implies

$$\|H^\epsilon(z) - H(z)\| \leq \mu\epsilon, \qquad \forall z \in R^n. \tag{2.10}$$

The following lemma shows that the sequence $\{z_k\}$ generated by Algorithm 1 is contained in the level set Ω.

Lemma 2.2 *Let $\{z_k\}$ be generated by Algorithm 1. Then $\{z_k\} \subset \Omega$.*

Proof Summing (2.6), we get for every k

$$\begin{aligned}\|H^{k+1}(z_{k+1})\| &\leq \|H^1(z_1)\| + \eta + \mu\epsilon_1 \\ &\leq \|H^0(z_1)\| + \eta + 2\mu\epsilon_1 \\ &\leq (1+\mu\gamma)\|H^0(z_1)\| + \eta,\end{aligned} \tag{2.11}$$

where the second inequality follows from (2.10) and the last inequality follows from the choice of ϵ_1. By using (2.10) again, we have

$$\begin{aligned}\|H^0(z_k)\| &\leq \|H^k(z_k)\| + \|H^k(z_k) - H^0(z_k)\| \\ &\leq \|H^k(z_k)\| + \mu\epsilon_k \\ &\leq \|H^k(z_k)\| + \mu\gamma\|H^0(z_k)\|,\end{aligned}$$

where the last inequality holds since by Step 5, $\epsilon_k \leq \gamma\|H^0(z_k)\|$ for every k. So, we deduce from (2.11) and the last inequality

$$\|H^0(z_k)\| \leq \frac{1}{1-\mu\gamma}\|H^k(z_k)\| \leq \frac{1+\mu\gamma}{1-\mu\gamma}\|H^0(z_1)\| + \frac{1}{1-\mu\gamma}\eta.$$

This means that $\{z_k\} \subset \Omega$ as desired. □

3 Global Convergence of Broyden-Like Methods

In this section, we prove global convergence of Algorithm 1. Let us define the index set $\bar{K} \subset \{1, 2, \ldots\}$ by

$$\bar{K} = \{k \mid \epsilon_k \geq \gamma \|H^0(z_{k+1})\|\}. \tag{3.1}$$

Then we have the following result.

Theorem 3.1 *Let $\{z_k\}$ be generated by Algorithm 1. If $\bar{K}$ is infinite, then*

$$\liminf_{k \to \infty} \|H^0(z_k)\| = 0. \tag{3.2}$$

Proof Let $\bar{K} = \{k_0 < k_1 < k_2 < \cdots\}$. Then by Step 5, for every k such that $k_{i-1} < k \leq k_i$, we have $\epsilon_k = \epsilon_{k_i}$. So, it follows that

$$\begin{aligned}
\|H^0(z_{k_i+1})\| &\leq \gamma^{-1}\epsilon_{k_i} = \gamma^{-1}\epsilon_{k_{i-1}+1} \leq \frac{1}{2}\|H^0(z_{k_{i-1}+1})\| \\
&\leq \cdots \leq (\frac{1}{2})^i \|H^0(z_{k_0+1})\| \leq \gamma^{-1}(\frac{1}{2})^i \epsilon_{k_0} \leq (\frac{1}{2})^i \|H^0(z_{k_0})\|,
\end{aligned} \tag{3.3}$$

where the last inequality holds since by Step 5 of Algorithm 1, $\epsilon_k \leq \gamma \|H^0(z_k)\|$ holds for every k. This implies (3.2). $\square$

Theorem 3.1 reveals that to show that Algorithm 1 converges globally, it suffices to verify that the index set $\bar{K}$ defined by (3.1) is infinite. We focus our attention to Broyden-like methods. That is, in Step 4 of Algorithm 1, B_{k+1} is updated by the Broyden-like formula

$$B_{k+1} = B_k + \theta_k \frac{(y_k - B_k s_k)s_k^T}{\|s_k\|^2}, \tag{3.4}$$

where $s_k = z_{k+1} - z_k$ and $y_k = H^k(z_{k+1}) - H^k(z_k)$. The parameter θ_k is chosen to satisfy $|\theta_k - 1| \leq \bar{\theta}$ and B_{k+1} is nonsingular [9]. We call the algorithm with (3.4) Broyden-like method.

To prove global convergence of Broyden-like method, we need the following assumption.

Assumption A (i) The level set Ω defined by (2.9) is contained in a bounded convex set D.

(ii) For each $\epsilon > 0$, function H^ϵ is continuously differentiable and ∇H^ϵ is Lipschitz continuous on D. That is, there is a positive constant τ_ϵ such that

$$\|\nabla H^\epsilon(z) - \nabla H^\epsilon(z')\| \leq \tau_\epsilon \|z - z'\|, \qquad \forall z, z' \in D. \tag{3.5}$$

Let scalar δ_k be defined by

$$\delta_k = \frac{\|y_k - B_k s_k\|}{\|s_k\|} \tag{3.6}$$

and matrix A_k be defined by

$$A_k = \int_0^1 \nabla H^k(z_k + t s_k)dt. \tag{3.7}$$

Then by the mean value theorem, we have $y_k = A_k s_k$ and hence

$$\delta_k = \frac{\|(A_k - B_k)s_k\|}{\|s_k\|} = \frac{\|(A_k - B_k)p_k\|}{\|p_k\|}. \tag{3.8}$$

The next useful lemma follows from (2.8) and Lemma 2.6 in [23].

Lemma 3.1 *Let Assumption A hold. If there is an index $\bar{k}$ such that $H^k(z) \equiv H^{\bar{k}}(z)$ for all $k \geq \bar{k}$, then*

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} \delta_i^2 = 0.$$

In particular, there is an infinite index set $\bar{K}_1$ such that $\delta_k \to 0$ as $k \to \infty$ with $k \in \bar{K}_1$.

Now we prove global convergence of Broyden-like method. From Theorem 3.1, it suffices to show that the index set $\bar{K}$ defined by (3.1) is infinite.

Theorem 3.2 *Let Assumption A hold. Assume that, for any accumulation point $\bar{z}$ of the sequence $\{z_k\}$ generated by Broyden-like method, $\nabla H^k(\bar{z})$ is nonsingular for all sufficiently large k. Then the index set $\bar{K}$ defined by (3.1) must be infinite. Moreover, we have*

$$\lim_{k\to\infty} \|H^0(z_k)\| = 0 \tag{3.9}$$

and every accumulation point of $\{z_k\}$ is a solution of (1.1).

Proof We assume that $\bar{K}$ is finite to deduce a contradiction. By Step 5 of Algorithm 1, there is an index $\bar{k}$ such that $\epsilon_k = \bar{\epsilon}$ for all $k \geq \bar{k}$ with some constant $\bar{\epsilon} > 0$. Hence $H^k = H^{\bar{\epsilon}}$ for all $k \geq \bar{k}$. Without loss of generality, we assume that $\epsilon_k = \bar{\epsilon}$ and $H^k = \bar{H}$ for all k with some function $\bar{H} : R^n \to R^n$.

From Lemma 3.1, there is an infinite index set $\bar{K}_1$ such that $\lim_{k\in\bar{K}_1, k\to\infty} \delta_k = 0$. Since $\{z_k\}_{k\in\bar{K}_1} \subset \Omega$ is bounded, there exists a subsequence $\{z_k\}_{k\in K} \subset \{z_k\}_{k\in\bar{K}_1}$ converging to some point $\bar{z}$. Let A_k be defined by (3.7). Then $A_k \to \nabla\bar{H}(\bar{z})$ as $k \to \infty$ with $k \in K$. Moreover, we get from (2.2) and (3.8)

$$\begin{aligned}
\|H^0(z_k)\| &= \|B_k p_k\| \\
&\leq \|(B_k - A_k)p_k\| + \|A_k p_k\| \\
&\leq \delta_k \|p_k\| + \|A_k\| \, \|p_k\|.
\end{aligned}$$

Since $\delta_k \to 0$ as $k \to \infty$ with $k \in \bar{K}_1$, there is a constant $M_1 > 0$ such that for all $k \in K$ large enough

$$\|H^0(z_k)\| \leq M_1 \|p_k\|. \tag{3.10}$$

Let $\bar{\alpha} = \limsup_{k\in K, k\to\infty} \alpha_k$. If $\bar{\alpha} > 0$, then we have $\lim_{k\in K,k\to\infty} \|p_k\| = 0$ by (2.8). Hence (3.10) implies (3.9). Now consider the case $\bar{\alpha} = 0$, that is, $\lim_{k\in K,k\to\infty} \alpha_k = 0$. By the assumption that $\nabla \bar{H}(\bar{z})$ is nonsingular, it is clear that when $k \in K$ is sufficiently large, $\nabla \bar{H}(z_k)$ and A_k are uniformly nonsingular. In particular, there exists a positive constant $M_2 > 0$ such that $\|\nabla \bar{H}(z_k)^{-1}\| \le M_2$ and $\|A_k^{-1}\| \le M_2$ hold for all $k \in K$ large enough. This together with (2.2) yields

$$\begin{aligned}
\|p_k\| &= \|A_k^{-1}[(A_k - B_k)p_k - H^0(z_k)]\| \\
&\le \|A_k^{-1}\| \left(\|(A_k - B_k)p_k\| + \|H^0(z_k)\|\right) \\
&\le M_2\delta_k\|p_k\| + M_2\|H^0(z_k)\|.
\end{aligned}$$

Since $\delta_k \to 0$ as $k \to \infty$ with $k \in K$, the last inequality implies that there is a constant $M_3 > 0$ such that for all $k \in K$ sufficiently large

$$\|p_k\| \le M_3\|H^0(z_k)\|. \tag{3.11}$$

In particular, $\{p_k\}_{k\in K}$ is bounded. Without loss of generality, we assume that $p_k \to \bar{p}$ as $k \to \infty$ with $k \in K$. Again by (2.2) we have

$$\begin{aligned}
0 &= \|B_k p_k + H^0(z_k)\| \\
&\ge \|H^0(z_k) + A_k p_k\| - \|(B_k - A_k)p_k\| \\
&= \|H^0(z_k) + A_k p_k\| - \delta_k\|p_k\|.
\end{aligned}$$

Letting $k \in K$ go to infinity yields

$$\nabla \bar{H}(\bar{z})^T \bar{p} + H^0(\bar{z}) = 0. \tag{3.12}$$

On the other hand, the assumption that $\lim_{k\in K,k\to\infty} \alpha_k = 0$ means that $\alpha_k < 1$ for all $k \in K$ sufficiently large. By the line search rule, it means that when $k \in K$ is sufficiently large, $\alpha'_k \equiv \alpha_k/\rho$ does not satisfy (2.4). In other words, we have

$$\begin{aligned}
\|\bar{H}(z_k + \alpha'_k p_k)\| - \|\bar{H}(z_k)\| &> -\sigma_1\|\bar{H}(z_k + \alpha'_k p_k) - \bar{H}(z_k)\| - \sigma_2\|\alpha' p_k\|^2 + \epsilon_k \\
&\ge -\sigma_1\|\bar{H}(z_k + \alpha'_k p_k) - \bar{H}(z_k)\| - \sigma_2\|\alpha' p_k\|^2.
\end{aligned}$$

Multiplying the both sides by $\|\bar{H}(z_k + \alpha'_k p_k)\| + \|\bar{H}(z_k)\|$ yields

$$\begin{aligned}
&\|\bar{H}(z_k + \alpha'_k p_k)\|^2 - \|\bar{H}(z_k)\|^2 \\
&\quad \ge -\sigma_1\left(\|\bar{H}(z_k + \alpha'_k p_k)\| + \|\bar{H}(z_k)\|\right)\|\bar{H}(z_k + \alpha'_k p_k) - \bar{H}(z_k)\| \\
&\qquad - \sigma_2\|\alpha'_k p_k\|^2\left(\|\bar{H}(z_k + \alpha'_k p_k)\| + \|\bar{H}(z_k)\|\right).
\end{aligned}$$

Dividing the both sides by α'_k and then taking the limit as $k \to \infty$ with $k \in K$ yield

$$2\bar{H}(\bar{z})^T \nabla \bar{H}(\bar{z})^T \bar{p} \geq -2\sigma_1 \|\bar{H}(\bar{z})\| \, \|\nabla \bar{H}(\bar{z})\bar{p}\|.$$

This together with (3.12) implies

$$\bar{H}(\bar{z})^T H^0(\bar{z}) \leq \sigma_1 \|\bar{H}(\bar{z})\| \, \|H^0(\bar{z})\|. \tag{3.13}$$

Since $\bar{\epsilon} \leq \gamma \|H^0(z_k)\|$ for all k, we have $\bar{\epsilon} \leq \gamma \|H^0(\bar{z})\|$. It then follows from (2.10) that

$$\|\bar{H}(\bar{z}) - H^0(\bar{z})\| = \|H^{\bar{\epsilon}}(\bar{z}) - H^0(\bar{z})\| \leq \mu\bar{\epsilon} \leq \mu\gamma \|H^0(\bar{z})\|, \tag{3.14}$$

and hence

$$\|\bar{H}(\bar{z})\| \leq \|H^0(\bar{z})\| + \|\bar{H}(\bar{z}) - H^0(\bar{z})\| \leq (1 + \mu\gamma)\|H^0(\bar{z})\|.$$

Consequently, we get from the last inequality and (3.13)

$$\begin{aligned}
\sigma_1(1+\mu\gamma)\|H^0(\bar{z})\|^2 &\geq \sigma_1 \|\bar{H}(\bar{z})\| \, \|H^0(\bar{z})\| \\
&\geq \bar{H}(\bar{z})^T H^0(\bar{z}) \\
&= \|H^0(\bar{z})\|^2 + (\bar{H}(\bar{z}) - H^0(\bar{z}))^T H^0(\bar{z}) \\
&\geq \|H^0(\bar{z})\|^2 - \|\bar{H}(\bar{z}) - H^0(\bar{z})\| \, \|H^0(\bar{z})\| \\
&\geq (1-\mu\gamma)\|H^0(\bar{z})\|^2,
\end{aligned}$$

where the last inequality follows from (3.14). However, since $\sigma_1 < \frac{1-\mu\gamma}{1+\mu\gamma}$, we get $H^0(\bar{z}) = 0$, which contradicts the fact that $0 < \bar{\epsilon} \leq \|H(\bar{z})\|$. The contradiction shows that $\bar{K}$ must be infinite. Therefore by Theorem 3.1, we have $\liminf_{k\to\infty} \|H^0(z_k)\| = 0$.

Now we show that (3.9) holds. Indeed, from (2.10), we get

$$\|H^k(z_k)\| \leq \|H^0(z_k)\| + \mu\epsilon_k \leq (1+\mu\gamma)\|H^0(z_k)\|.$$

This means that there is a subsequence of $\{H^k(z_k)\}$ converging to zero. However, Lemma 2.1 shows that $\{\|H^k(z_k)\|\}$ itself converges and hence it converges to zero. Again by (2.10) we get

$$\|H^0(z_k)\| \leq \|H^k(z_k)\| + \mu\epsilon_k \leq \|H^k(z_k)\| + \mu\gamma\|H^0(z_k)\|,$$

which implies that

$$\|H^0(z_k)\| \leq \frac{1}{1-\mu\gamma}\|H^k(z_k)\|.$$

This shows that (3.9) holds. Consequently, every accumulation point of $\{z_k\}$ is a solution of (1.1). $\square$

Remark (i) In Theorem 3.2, the assumption that $\nabla H^k(\bar{z})$ is nonsingular for every accumulation point $\bar{z}$ can be replaced by the weaker assumption that there is an accumulation point $\bar{z}$ of the subsequence $\{z_k\}_{k\in\bar{K}_1}$ such that $\nabla H^k(\bar{z})$ is nonsingular for all k large enough, where $\bar{K}_1$ is the index set as specified in Lemma 3.1.

(ii) Since by Step 5 in Broyden-like method, $\epsilon_k \le \gamma\|H(z_k)\|$ holds for every k and $\{\epsilon_k\}$ is monotonically nonincreasing, Theorem 3.2 shows that

$$\lim_{k\to\infty} \epsilon_k = 0. \tag{3.15}$$

4 Superlinear Convergence

In this section, we prove superlinear convergence of Broyden-like method. To this end, we need further assumptions.

Assumption B (i) The sequence $\{z_k\}$ generated by Broyden-like method converges to a solution $\bar{z}$ of (1.1) at which every element in $\partial_C H(\bar{z})$ is nonsingular, where $\partial_C H(z) = \partial H_1(z)\times\partial H_2(z)\times\cdots\times\partial H_n(z)$, and $\partial h(z)$ denotes the generalized Jacobian of function $h: R^n \to R$ at z in the sense of Clarke [6].

(ii) Functions H^k and H satisfy

$$\lim_{k\to\infty} \text{dist }(\nabla H^k(z_k), \partial_C H(\bar{z})) = 0. \tag{4.1}$$

(iii) There exist a neighbourhood $U(\bar{z})$ of $\bar{z}$ and a constant $\tau > 0$ such that for all k sufficiently large

$$\|\nabla H^k(z) - \nabla H^{k-1}(z)\| \le \tau(\epsilon_{k-1} - \epsilon_k), \quad \forall z \in U(\bar{z}) \tag{4.2}$$

and

$$\|\nabla H^k(z_k + ts_k) - \nabla H^k(z_{k-1} + ts_{k-1})\| \le \tau(\|s_k\| + \|s_{k-1}\|), \quad \forall t \in (0,1). \tag{4.3}$$

(iv) As k goes to infinity, we have

$$\|H^k(\bar{z}) - H^0(\bar{z})\| = o(\|z_k - \bar{z}\|), \quad \|H^k(z_k) - H^0(z_k)\| = o(\|z_k - \bar{z}\|). \tag{4.4}$$

Remark (i) Since $\partial_C H(\bar{z})$ is compact and $\partial_C H(\cdot)$ is upper semicontinuous, it is not difficult to see that conditions (i) and (ii) of Assumptions B imply that there is a neighbourhood $U(\bar{z})$ of $\bar{z}$, a constant $\tau' > 0$ and an index $\bar{k}$ such that inequality $\|\nabla H^k(z)\| \le \tau'$ holds for all $k \ge \bar{k}$ and any $z \in U(\bar{z})$. Moreover, there are constants $0 < c \le C$ such that

$$c\|z - z'\| \le \|H^k(z) - H^k(z')\| \le C\|z - z'\|, \quad \forall k \ge \bar{k}, \ \ \forall z, z' \in U(\bar{z}). \tag{4.5}$$

(ii) Inequality (4.3) holds particularly when H^k satisfies

$$\|\nabla H^k(z) - \nabla H^k(z')\| \le \tau\|z - z'\|, \quad \forall z, z' \in U(\bar{z}).$$

Lemma 4.1 *Let Assumptions A and B hold and δ_k be defined by (3.6). Then we have*

$$\lim_{k\to\infty} \delta_k = 0. \tag{4.6}$$

Proof It follows from (2.7) in Lemma 2.1 and (4.5) that

$$\sum_{k=1}^{\infty} \|s_k\| = \sum_{k=1}^{\infty} \|z_{k+1} - z_k\| < \infty. \tag{4.7}$$

Let A_k be defined by (3.7). Then we have $y_k = A_k s_k$. Moreover, it follows from (3.4) that

$$\begin{aligned}
\|B_{k+1} - A_k\|_F^2 &= \Big\|B_k - A_k + \theta_k(A_k - B_k)\frac{s_k s_k^T}{\|s_k\|^2}\Big\|_F^2 \\
&= \|B_k - A_k\|_F^2 + (\theta_k^2 - 2\theta_k)\frac{\|(B_k - A_k)s_k\|^2}{\|s_k\|^2} \\
&= \|B_k - A_k\|_F^2 + (\theta_k^2 - 2\theta_k)\frac{\|y_k - B_k s_k\|^2}{\|s_k\|^2} \\
&= \|B_k - A_k\|_F^2 + (\theta_k^2 - 2\theta_k)\delta_k^2 \\
&\le \|B_k - A_k\|_F^2 - (1 - \bar{\theta}^2)\delta_k^2, \qquad (4.8)
\end{aligned}$$

where the last inequality holds since θ_k is chosen to satisfy $|1-\theta_k| \le \bar{\theta}$ in Step 4. Then we deduce

$$\begin{aligned}
\|A_k - A_{k-1}\| &\le \int_0^1 \|\nabla H^k(z_k + ts_k) - \nabla H^{k-1}(z_{k-1} + ts_{k-1})\|dt \\
&\le \int_0^1 \|\nabla H^k(z_k + ts_k) - \nabla H^k(z_{k-1} + ts_{k-1})\|dt \\
&\quad + \int_0^1 \|\nabla H^k(z_{k-1} + ts_{k-1}) - \nabla H^{k-1}(z_{k-1} + ts_{k-1})\|dt \\
&\le \tau\Big(\|s_k\| + \|s_{k-1}\|\Big) + \tau(\epsilon_{k-1} - \epsilon_k),
\end{aligned}$$

where the last inequality follows from (4.2) and (4.3). This together with (4.7) implies

$$\sum_{k=2}^{\infty} \|A_k - A_{k-1}\| < \infty. \tag{4.9}$$

Therefore, we get from (4.8)

$$\|B_{k+1} - A_k\|_F \le \|B_k - A_k\|_F \le \|B_k - A_{k-1}\|_F + \|A_k - A_{k-1}\|_F.$$

This along with (4.9) implies that $\{\|B_{k+1} - A_k\|_F\}$ converges. Thus we deduce from (4.8) again that (4.6) holds. □

Lemma 4.2 *Let $\bar{z}$ be a solution of (1.1). Then under Assumptions A and B, we have*

$$\lim_{k\to\infty} \frac{\|z_k + p_k - \bar{z}\|}{\|z_k - \bar{z}\|} = 0. \tag{4.10}$$

Proof Let A_k be defined by (3.7). It then follows from (4.3) that $A_k - \nabla H^k(z_k) \to 0$. Moreover, since (4.6) holds, using the arguments for showing (3.11) in the proof of Theorem 3.2, we can deduce that there is a positive constant M_3 such that

$$\|p_k\| \le M_3 \|H^0(z_k)\| \le LM_3 \|z_k - \bar{z}\|, \tag{4.11}$$

where $L > 0$ is the Lipschitz constant of H^0. We then get from (2.2)

$$\begin{aligned}
&\|z_k + p_k - \bar{z}\| \\
&= \|\nabla H^k(z_k)^{-1}(\nabla H^k(z_k)(z_k - \bar{z}) + \nabla H^k(z_k)p_k - B_k p_k - H^0(z_k))\| \\
&\le \|\nabla H^k(z_k)^{-1}\| \Big(\|(\nabla H^k(z_k) - B_k)p_k\| + \|\nabla H^k(z_k)(z_k - \bar{z}) + H^0(\bar{z}) - H^0(z_k)\| \Big) \\
&\le \tau' \Big(\|(A_k - B_k)p_k\| + \|(\nabla H^k(z_k) - A_k)p_k\| + \|\nabla H^k(z_k)(z_k - \bar{z}) + H^0(\bar{z}) - H^0(z_k)\| \Big) \\
&\le \tau' \Big(\delta_k \|p_k\| + \|\nabla H^k(z_k) - A_k\| \, \|p_k\| + \|\nabla H^k(z_k)(z_k - \bar{z}) + H^k(\bar{z}) - H^k(z_k)\| \\
&\qquad + \|[H^0(\bar{z}) - H^0(z_k)] - [H^k(\bar{z}) - H^k(z_k)]\| \Big) \\
&\le LM_3 \tau' \Big(\delta_k + \|\nabla H^k(z_k) - A_k\| \Big) \|z_k - \bar{z}\| + o(\|z_k - \bar{z}\|) \\
&= o(\|z_k - \bar{z}\|),
\end{aligned}$$

where the third inequality follows from (3.6) and $y_k = A_k s_k$, the last inequality follows from (4.4) and (4.11), and the last equality follows from (4.6). □

Now, we establish superlinear convergence of Broyden-like method. By Lemma 4.2, it suffices to show that when k is sufficiently large, the unit steplength is always accepted.

Theorem 4.1 *Under Assumptions A and B, the sequence $\{z_k\}$ generated by Broyden-like method converges superlinearly.*

Proof By Lemma 4.2, it suffices to show that when k is sufficiently large, $\alpha = 1$ satisfies (2.4). Indeed, it follows from (4.4), (4.5) and (4.10) that for all k large enough

$$\begin{aligned}
\|H^k(z_k + p_k)\| &= \|H^k(z_k + p_k) - H^0(\bar{z})\| \\
&\le \|H^k(z_k + p_k) - H^k(\bar{z})\| + \|H^k(\bar{z}) - H^0(\bar{z})\| \\
&\le C\|z_k + p_k - \bar{z}\| + o(\|z_k - \bar{z}\|) \\
&= o(\|z_k - \bar{z}\|).
\end{aligned} \tag{4.12}$$

On the other hand, we have

$$\begin{aligned}
&\|H^k(z_k)\| - \sigma_1 \|H^k(z_k + p_k) - H^k(z_k)\| - \sigma_2 \|p_k\|^2 \\
&= \|H^k(z_k) - H^0(\bar{z})\| - \sigma_1 \|H^k(z_k + p_k) - H^k(z_k)\| + o(\|z_k - \bar{z}\|)
\end{aligned}$$

$$
\begin{aligned}
&\geq \Big(\|H^k(z_k) - H^k(\bar{z})\| - \|H^k(\bar{z}) - H^0(\bar{z})\|\Big) \\
&\quad - \sigma_1\Big(\|H^k(z_k + p_k) - H^k(\bar{z})\| + \|H^k(\bar{z}) - H^k(z_k)\|\Big) + o(\|z_k - \bar{z}\|) \\
&= (1 - \sigma_1)\|H^k(z_k) - H^k(\bar{z})\| + o(\|z_k - \bar{z}\|) \\
&\geq (1 - \sigma_1)c\|z_k - \bar{z}\| + o(\|z_k - \bar{z}\|), \qquad (4.13)
\end{aligned}
$$

where the first equality follows from (4.11), the last equality follows from (4.4), (4.5) and (4.10), and the last inequality follows from (4.5). Comparing (4.12) with (4.13), it is clear that when k is sufficiently large

$$
\begin{aligned}
\|H^k(z_k + p_k)\| &\leq \|H^k(z_k)\| - \sigma_1\|H^k(z_k + p_k) - H^k(z_k)\| - \sigma_2\|p_k\|^2 \\
&\leq \|H^k(z_k)\| - \sigma_1\|H^k(z_k + p_k) - H^k(z_k)\| - \sigma_2\|p_k\|^2 + \eta_k.
\end{aligned}
$$

This means that when k is sufficiently large, $\alpha = 1$ is always accepted. □

Theorem 4.1 reveals that to ensure superlinear convergence of Broyden-like method, the strict complementarity at $\bar{z}$ is not necessary.

5 Applications

In this section, we apply the general smoothing quasi-Newton method proposed in Section 2 to nonlinear complementarity problems, mixed complementarity problems and variational inequality problems. We show that for some well known smoothing functions, the conditions in Assumptions A and B hold, and hence the related algorithms converge globally and superlinearly.

5.1 Nonlinear complementarity problems

Let $f : R^n \to R^n$ be continuously differentiable. Consider the nonlinear complementarity problem of finding a $z \in R^n$ such that

$$z \geq 0, \quad f(z) \geq 0, \quad z^T f(z) = 0. \qquad (5.1)$$

There have been developed various reformulations of (5.1) and related smoothing functions. Fischer-Burmeister function as well as its perturbed version and Chen-Harker-Kanzow-Smale smoothing function are widely used [7, 10, 11, 12, 16, 18, 19, 34]. Fischer-Burmeister function $\phi_{FB} : R^2 \to R$ is defined by

$$\phi_{FB}(a, b) = \sqrt{a^2 + b^2} - (a + b). \qquad (5.2)$$

It is easy to see that $\phi_{FB}(a, b) = 0$ holds if and only if $a \geq 0$, $b \geq 0$ and $ab = 0$. Fischer-Burmeister function is differentiable at every point except the origin and is semismooth at the origin. By introducing a perturbation parameter, we get a smoothing Fischer-Burmeister function

$$\phi_{FB}^{\epsilon}(a,b) = \sqrt{a^2+b^2+\epsilon^2} - (a+b), \tag{5.3}$$

where ϵ is a nonnegative parameter. It is clear that $\phi_{FB}^{0}(a,b) \equiv \phi_{FB}(a,b)$.

Let $H^{\epsilon}(z) = (H_1^{\epsilon}(z), H_2^{\epsilon}(z), \dots, H_n^{\epsilon}(z))^T$, where

$$H_i^{\epsilon}(z) = \phi_{FB}^{\epsilon}(z_i, f_i(z)), \quad i = 1,2,\dots,n. \tag{5.4}$$

Then H^{ϵ} is continuously differentiable everywhere if f is continuously differentiable and $\epsilon > 0$. Moreover, (5.1) is equivalent to the semismooth equation $H^0(z) = 0$. Let

$$a_i^{\epsilon}(z) = \frac{z_i}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2}} - 1, \quad b_i^{\epsilon}(z) = \frac{f_i(z)}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2}} - 1.$$

Then it is obvious that for every $\epsilon > 0$ and any x, we have

$$-2 < a_i^{\epsilon}(z) < 0, \quad -2 < b_i^{\epsilon}(z) < 0. \tag{5.5}$$

Moreover, it is easy to deduce

$$\nabla H^{\epsilon}(z) = \text{ diag } (a_i^{\epsilon}(z)) + \nabla f(z) \text{ diag } (b_i^{\epsilon}(z)). \tag{5.6}$$

The following lemma gives some properties of ϕ_{FB}^{ϵ} and H^{ϵ}.

Lemma 5.1 *Let ϕ_{FB}^{ϵ} and H^{ϵ} be defined by (5.3) and (5.4), respectively. Suppose that f is continuously differentiable. Then the following statements hold.*

(i) For any $\epsilon > 0$, functions ϕ_{FB}^{ϵ} and H^{ϵ} are continuously differentiable everywhere.

(ii) For any $\epsilon > \epsilon' > 0$, we have

$$|\phi_{FB}^{\epsilon}(a,b) - \phi_{FB}^{\epsilon'}(a,b)| \le \epsilon - \epsilon', \quad \forall (a,b) \in R^2 \tag{5.7}$$

and hence

$$\|H^{\epsilon}(z) - H^{\epsilon'}(z)\| \le \sqrt{n}(\epsilon - \epsilon'), \quad \forall z \in R^n.$$

(iii) For a given constant $\xi > 0$, the inequality

$$|\phi_{FB}^{\epsilon}(a,b) - \phi_{FB}^{0}(a,b)| \le \frac{1}{2}\xi^{-1}\epsilon^2$$

holds for all $(a,b) \in R^2$ satisfying $a^2 + b^2 \ge \xi^2$. Consequently, the inequality

$$\|H^{\epsilon}(z) - H^{\epsilon'}(z)\| \le \frac{1}{2}\sqrt{n}\xi^{-1}\epsilon^2$$

holds for all $z \in R^n$ satisfying $z_i^2 + f_i(z)^2 \ge \xi^2$, $i = 1,2,\dots,n$.

(iv) If f is a P_0 function, then $\nabla H^{\epsilon}(z)$ is nonsingular for every $\epsilon > 0$ and any $z \in R^n$.

Proof The statements (i), (ii) and (iv) have been proved in [34]. We prove (iii). Let $a^2+b^2 \geq \xi^2$. Then, we get

$$\begin{aligned}
|\phi_{FB}^{\epsilon}(a,b) - \phi_{FB}^{0}(a,b)| &= \sqrt{a^2+b^2+\epsilon^2} - \sqrt{a^2+b^2} \\
&= \frac{\epsilon^2}{\sqrt{a^2+b^2+\epsilon^2} + \sqrt{a^2+b^2}} \\
&\leq \frac{1}{2}\xi^{-1}\epsilon^2.
\end{aligned}$$

$\square$

Now we make the following assumption.

Assumption A$'$ (i) The level set Ω defined by (2.9) is contained in a bounded convex set D.

(ii) Function f is continuously differentiable and ∇f is Lipschitz continuous on D.

Lemma 5.2 *Let Assumption A$'$ hold. Then for any $\epsilon > 0$, there is a constant $\tau_1^{\epsilon} > 0$ such that*

$$\|\nabla H^{\epsilon}(z) - \nabla H^{\epsilon}(z')\| \leq \tau_1^{\epsilon}\|z - z'\|, \qquad \forall z, z' \in D. \tag{5.8}$$

Proof It follows from (5.5) and (5.6) that

$$\begin{aligned}
&\|\nabla H^{\epsilon}(z) - \nabla H^{\epsilon}(z')\| \\
&= \Big\| \text{ diag } (a_i^{\epsilon}(z) - a_i^{\epsilon}(z')) + (\nabla f(z) - \nabla f(z')) \text{ diag } (b_i^{\epsilon}(z)) + \nabla f(z') \text{ diag } (b_i^{\epsilon}(z) - b_i^{\epsilon}(z))\Big\| \\
&\leq \| \text{ diag } (a_i^{\epsilon}(z) - a_i^{\epsilon}(z'))\| + 2\|\nabla f(z) - \nabla f(z')\| + \|\nabla f(z')\| \, \| \text{ diag } (b_i^{\epsilon}(z) - b_i^{\epsilon}(z))\|.
\end{aligned}$$

It then suffices to show that there is a constant $\bar{\tau}_1^{\epsilon} > 0$ such that inequalities

$$|a_i^{\epsilon}(z) - a_i^{\epsilon}(z')| \leq \bar{\tau}_1^{\epsilon}\|z - z'\| \quad \text{ and } \quad |b_i^{\epsilon}(z) - b_i^{\epsilon}(z')| \leq \bar{\tau}_1^{\epsilon}\|z - z'\| \tag{5.9}$$

hold for any $z, z' \in D$. By the definition of a_i^{ϵ}, we have

$$\begin{aligned}
|a_i^{\epsilon}(z) - a_i^{\epsilon}(z')| &= \Big|\frac{z_i}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2}} - \frac{z'}{\sqrt{(z_i')^2 + f_i(z')^2 + \epsilon^2}}\Big| \\
&\leq \Big|\frac{z_i - z_i'}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2}}\Big| + |z_i'|\Big|\frac{1}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2}} - \frac{1}{\sqrt{(z_i')^2 + f_i(z')^2 + \epsilon^2}}\Big| \\
&\leq \epsilon^{-1}|z_i - z_i'| + \bar{M}\epsilon^{-2}\Big|\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2} - \sqrt{(z_i')^2 + f_i(z')^2 + \epsilon^2}\Big| \\
&= \epsilon^{-1}|z - z'| + \bar{M}\epsilon^{-2}\frac{|z_i^2 - (z_i')^2 + f_i(z)^2 - f_i(z')^2|}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2} + \sqrt{(z_i')^2 + f_i(z')^2 + \epsilon^2}} \\
&\leq \epsilon^{-1}|z_i - z_i'| + 2^{-1}\bar{M}\epsilon^{-3}(|z_i^2 - (z_i')^2| + |f_i(z)^2 - f_i(z')^2|),
\end{aligned} \tag{5.10}$$

where $\bar{M}$ is a constant such that $\|z\| \le \bar{M}$ for all $z \in D$. The last inequality implies the first inequality of (5.9) with a suitable constant $\bar{\tau}_1^\epsilon > 0$. The second inequality of (5.9) can be proved in a similar way. $\square$

Lemmas 5.1 and 5.2 imply the following global convergence theorem.

Theorem 5.1 *Let Assumption A′ hold and H^ϵ be defined by (5.4). Suppose that f is a P_0 function. Then the sequence $\{x_k\}$ generated by the related Broyden-like method satisfies (3.9).*

We say that $\bar{z}$ is a strictly complementary solution of (5.1) if it is a solution of (5.1) and satisfies

$$\bar{z}_i + f_i(\bar{z}) > 0, \quad i = 1, 2, \ldots, n. \tag{5.11}$$

We make another assumption as follows.

Assumption B′ (i) The sequence $\{z_k\}$ generated by Broyden-like method converges to a solution $\bar{z}$ of (1.1) that is a strictly complementary solution of (5.1).

(ii) $\nabla H^0(\bar{z})$ is nonsingular.

Remark The strict complementarity (5.11) implies that H is differentiable near $\bar{z}$. Hence, $\nabla H^0(\bar{z})$ exists. Moreover, $\partial_C H^0(\bar{z})$ reduces to the singleton $\{\nabla H^0(\bar{z})\}$. Therefore, Assumption B′ implies condition (i) in Assumption B.

The following lemma shows that Assumptions A′ and B′ imply Assumption B.

Lemma 5.3 *Let ϕ_{FB}^ϵ and H^ϵ be defined by (5.3) and (5.4), respectively. If Assumptions A′ and B′ hold, then Assumption B holds.*

Proof The inequality (5.11) implies that there is a neighbourhood $U(\bar{z})$ of $\bar{z}$ such that the inequality $z_i^2 + f_i^2(z) \ge \frac{1}{2}(\bar{z}_i^2 + f_i^2(\bar{z})) \stackrel{\triangle}{=} \zeta^2 > 0$ holds for any $z \in U(\bar{z})$. It is then clear that H^0 is differentiable in $U(\bar{z})$ and hence equality (4.1) holds. Inequality (4.3) can be obtained in a way similar to Lemma 5.2. Inequality (4.4) follows from Lemma 5.1 (iii) and the fact $\epsilon_k \le \gamma \|H^k(z_k)\|$. It remains to show (4.2). By (5.6), it suffices to show that there exists a constant $\bar{\tau} > 0$ such that the following inequalities hold for any $\epsilon > \epsilon'$ and all $z \in U(\bar{z})$:

$$|a_i^\epsilon(z) - a_i^{\epsilon'}(z)| \le \bar{\tau}(\epsilon - \epsilon'), \quad |b_i^\epsilon(z) - b_i^{\epsilon'}(z)| \le \bar{\tau}(\epsilon - \epsilon'). \tag{5.12}$$

Indeed, by the definition of a_i^ϵ, we have

$$\begin{aligned}
|a_i^\epsilon(z) - a_i^{\epsilon'}(z)| &= |z_i| \Big| \frac{1}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2}} - \frac{1}{\sqrt{z_i^2 + f_i(z)^2 + (\epsilon')^2}} \Big| \\
&\le \zeta^{-1} |z_i| \Big| \sqrt{z_i^2 + f_i(z)^2 + \epsilon^2} - \sqrt{z_i^2 + f_i(z)^2 + (\epsilon')^2} \Big| \\
&= \zeta^{-1} |z_i| \Big| \frac{\epsilon^2 - (\epsilon')^2}{\sqrt{z_i^2 + f_i(z)^2 + \epsilon^2} + \sqrt{z_i^2 + f_i(z)^2 + (\epsilon')^2}} \Big| \\
&\le \zeta^{-1} |z_i| (\epsilon - \epsilon').
\end{aligned}$$

This shows the first inequality of (5.12). The second inequality of (5.12) can be proved in a similar way. □

In view of Lemmas 5.1, 5.2 and 5.3, and Theorem 4.1, we get the following superlinear convergence theorem.

Theorem 5.2 *Let Assumptions A' and B' hold and H^ϵ be defined by (5.4). Then the sequence $\{z_k\}$ generated by the related Broyden-like method converges superlinearly.*

Chen-Harker-Kanzow-Smale smoothing function is defined by

$$\phi^\epsilon_{CHKS}(a,b) = \frac{a+b-\sqrt{(a-b)^2+\epsilon^2}}{2}. \tag{5.13}$$

It is a smoothing function of $\phi^0_{CHKS}(a,b) = \min\{a,b\}$. Functions ϕ^ϵ_{CHKS} and ϕ^0_{CHKS} enjoy some important properties that ϕ^ϵ_{FB} and ϕ^0_{FB} have. Let

$$H^\epsilon(z) = (\phi^\epsilon_{CHKS}(z_1, f_1(z)), \ldots, \phi^\epsilon_{CHKS}(z_n, f_n(z)))^T. \tag{5.14}$$

It is not difficult to prove that Lemmas 5.1, 5.2 and 5.3 remain true. It then follows that Theorems 5.1 and 5.2 hold even if ϕ^ϵ_{FB} is replaced by ϕ^ϵ_{CHKS}. We omit the detail.

5.2 KKT system of variational inequality problems

Let $F : R^n \to R^n$ be continuously differentiable and X be a closed convex subset of R^n. The variational inequality problem VI(X, F) is to find an $x \in R^n$ such that

$$F(x)^T(y-x) \geq 0, \qquad \forall y \in X. \tag{5.15}$$

In many cases, X is specified as

$$X = \{x \in R^n \,|\, g_i(x) \geq 0, i = 1,2,\ldots,m_1, h_j(x) = 0, j = 1,2,\ldots,m_2\},$$

where $g_i : R^n \to R$ and $h_j : R^n \to R$ are twice continuously differentiable functions. It has been shown that under a suitable constraint qualification, a solution of (5.15) together with some vectors $\lambda \in R^{m_1}$ and $\mu \in R^{m_2}$ is a solution of the following KKT system:

$$\begin{pmatrix} L(z) \\ \min(\lambda, g(x)) \\ h(x) \end{pmatrix} = 0, \tag{5.16}$$

where $z = (x, \lambda, \mu) \in R^{n+m_1+m_2}$,

$$L(z) = F(x) - \nabla g(x)\lambda - \nabla h(x)\mu, \tag{5.17}$$

and

$$\min(\lambda, g(x)) = (\min(\lambda_1, g_1(x)), \min(\lambda_2, g_2(x)), \ldots, \min(\lambda_{m_1}, g_{m_1}(x)))^T.$$

By using Fischer-Burmeister function, (5.16) can be reformulated as

$$H(z) = \begin{pmatrix} L(z) \\ \Phi_{FB}(\lambda, g(x)) \\ h(x) \end{pmatrix} = 0, \tag{5.18}$$

where $\Phi_{FB}(\lambda, g(x)) = (\phi_{FB}(\lambda_1, g_1(x)), \ldots, \phi_{FB}(\lambda_{m_1}, g_{m_1}(x)))^T$ and ϕ_{FB} is defined by (5.2). By using the perturbed Fischer-Burmeister function, we get a smoothing function to H

$$H^\epsilon(z) = \begin{pmatrix} L(z) \\ \Phi^\epsilon_{FB}(\lambda, g(x)) \\ h(x) \end{pmatrix} = 0, \tag{5.19}$$

where $\Phi^\epsilon_{FB}(\lambda, g(x)) = (\phi^\epsilon_{FB}(\lambda_1, g_1(x)), \ldots, \phi^\epsilon_{FB}(\lambda_{m_1}, g_{m_1}(x)))^T$ and ϕ^ϵ_{FB} is defined by (5.3). Put

$$\bar{a}^\epsilon_i(x, \lambda) = \frac{\lambda_i}{\sqrt{\lambda_i^2 + g_i(x)^2 + \epsilon^2}} - 1, \quad \bar{b}^\epsilon_i(x, \lambda) = \frac{g_i(x)}{\sqrt{\lambda_i^2 + g_i(x)^2 + \epsilon^2}} - 1.$$

It is easy to show the following lemma by using the chain rule of derivatives for composite functions.

Lemma 5.4 *Let F be continuously differentiable and g_i, $i = 1, 2, \ldots, m_1$, and h_j, $j = 1, 2, \ldots, m_2$, be twice continuously differentiable. Then Φ^ϵ and H^ϵ are continuously differentiable for every $\epsilon \neq 0$. Moreover, for every $z = (x, \lambda, \mu) \in R^{n+m_1+m_2}$ and any $\epsilon \neq 0$, we have the following expressions:*

$$\nabla_x \Phi^\epsilon(x, \lambda) = \nabla g(x) \text{ diag } (\bar{b}^\epsilon_i(x, \lambda)), \qquad \nabla_\lambda \Phi^\epsilon(x, \lambda) = \text{ diag } (\bar{a}^\epsilon_i(x, \lambda)),$$

and

$$\nabla H^\epsilon(z) = \begin{pmatrix} \nabla_x L(z) & -\nabla g(x) \text{ diag } (\bar{b}^\epsilon_i(x, \lambda)) & -\nabla h(x) \\ -\nabla g(x)^T & -\text{ diag } (\bar{a}^\epsilon_i(x, \lambda)) & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix}. \tag{5.20}$$

In the remainder of this section, we abbreviate $\bar{a}^\epsilon_i(x, \lambda)$ and $\bar{b}^\epsilon_i(x, \lambda)$ as $\bar{a}^\epsilon_i$ and $\bar{b}^\epsilon_i$, respectively. It is clear that $\text{diag}(\bar{a}^\epsilon_i)$ and $\text{diag}(\bar{b}^\epsilon_i)$ are nonsingular. Moreover, $\bar{a}^\epsilon_i$ and $\bar{b}^\epsilon_i$ are negative. This implies that for any $(x, \lambda) \in R^{n+m_1}$, the diagonal matrix $\text{diag}\,((\bar{a}^\epsilon_i)^{-1}\bar{b}^\epsilon_i)$ is positive definite, and hence, the matrix $\nabla g(x)\text{diag}((\bar{a}^\epsilon_i)^{-1}\bar{b}^\epsilon_i)\nabla g(x)^T$ is symmetric and positive semidefinite for any $(x, \lambda) \in R^{n+m_1}$. The following lemma shows some conditions that ensure regularity of H^ϵ.

Lemma 5.5 *Let $z = (x, \lambda, \mu) \in R^{n+m_1+m_2}$ and $\epsilon \neq 0$. If $\nabla h_j(x), j = 1, 2, \ldots, m_2$, are linearly independent and the matrix*

$$\nabla_x L(z) + \nabla g(x) \mathrm{diag}((\bar{a}_i^\epsilon)^{-1} \bar{b}_i^\epsilon) \nabla g(x)^T$$

is positive definite on the subspace $\{p \,|\, \nabla h(x)^T p = 0\}$, then $\nabla H^\epsilon(z)$ is nonsingular. In particular, if $\nabla_x L(z)$ is positive definite, then $\nabla H^\epsilon(z)$ is nonsingular.

Proof Let $w = (p, q, r) \in R^{n+m_1+m_2}$ be a solution of the system of linear equations $\nabla H^\epsilon(z) w = 0$, i.e.,

$$\begin{pmatrix} \nabla_x L(z) & -\nabla g(x)\ \mathrm{diag}\ (\bar{b}_i^\epsilon) & -\nabla h(x) \\ -\nabla g(x)^T & -\ \mathrm{diag}\ (\bar{a}_i^\epsilon) & 0 \\ -\nabla h(x)^T & 0 & 0. \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = 0.$$

We will show that $w = 0$. Writing the above equations blockwise, we get

$$\left\{ \begin{array}{lllllll} \nabla_x L(z) p & - & \nabla g(x)\ \mathrm{diag}\ (\bar{b}_i^\epsilon) q & - & \nabla h(x) r & = & 0, \\ -\nabla g(x)^T p & - & \mathrm{diag}\ (\bar{a}_i^\epsilon) q & & & = & 0, \\ -\nabla h(x)^T p & & & & & = & 0, \end{array} \right. \tag{5.21}$$

which implies

$$\left\{ \begin{array}{l} \left(\nabla_x L(z) + \nabla g(x) \mathrm{diag}((\bar{a}_i^\epsilon)^{-1} \bar{b}_i^\epsilon) \nabla g(x)^T\right) p - \nabla h(x) r = 0, \\ \nabla h(x)^T p = 0. \end{array} \right.$$

Multiplying the first equation by p and taking into account the second equation yield

$$p^T \left(\nabla_x L(z) + \nabla g(x) \mathrm{diag}((\bar{a}_i^\epsilon)^{-1} \bar{b}_i^\epsilon) \nabla g(x)^T\right) p = 0. \tag{5.22}$$

However, since p lies in the subspace $\{p \,|\, \nabla h(x)^T p = 0\}$, it follows from the positive definiteness of $\nabla_x L(z) + \nabla g(x) \mathrm{diag}((\bar{a}_i^\epsilon)^{-1} \bar{b}_i^\epsilon) \nabla g(x)^T$ on the subspace that $p = 0$. Then, the second equation in (5.21) implies that $q = 0$. Finally, since $\nabla h_j(x)$, $j = 1, 2, \ldots, m_2$, are linearly independent, it follows from the first equation of (5.21) that $r = 0$. Consequently, zero is the unique solution of (5.21), or equivalently, $\nabla H^\epsilon(z)$ is nonsingular. The last part of the lemma is a direct consequence of the first part by the positive semidefiniteness of $\nabla g(x)$ diag $((\bar{a}_i^\epsilon)^{-1} \bar{b}_i^\epsilon) \nabla g(x)^T$. $\square$

Similar to Assumption A$'$, we make the following assumption.

Assumption A$''$ (i) The level set Ω defined by (2.9) is contained in a bounded convex set D.

(ii) The function F is continuously differentiable on D_x, and g_i, $i = 1, 2, \ldots, m_1$, and h_j, $j = 1, 2, \ldots, m_2$, are twice continuously differentiable on D_x, where $D_x = \{x \in R^n \,|\, (x, \lambda, \mu) \in D\}$.

(iii) ∇F, $\nabla^2 g$ and $\nabla^2 h$ are Lipschitz continuous on D_x, i.e., there is a constant $\tau > 0$ such that

$$\begin{cases} \|\nabla F(x) - \nabla F(y)\| & \leq & \tau\|x-y\|, & \forall x, y \in D_x \\ \|\nabla^2 g(x) - \nabla^2 g(y)\| & \leq & \tau\|x-y\|, & \forall x, y \in D_x \\ \|\nabla^2 h(x) - \nabla^2 h(y)\| & \leq & \tau\|x-y\|, & \forall x, y \in D_x. \end{cases}$$

In a way similar to Lemma 5.2, we can prove that under Assumption A$''$, ∇H^ϵ is Lipschitz continuous on D. To be specific, we have the following lemma.

Lemma 5.6 *Let Assumption A$''$ hold. Then for any $\epsilon > 0$, there is a constant $\tau_2^\epsilon > 0$ such that*

$$\|\nabla H^\epsilon(z) - \nabla H^\epsilon(z')\| \leq \tau_2^\epsilon \|z - z'\|, \qquad \forall z, z' \in D. \tag{5.23}$$

Similar to Lemma 5.3, we can show that the following assumption implies Assumption B.

Assumption B$''$ (i) The sequence $\{z_k\}$ generated by Algorithm 1 converges to a solution of (5.16), say $\bar{z} = (\bar{x}, \bar{\lambda}, \bar{\mu})$, at which the strict complementarity condition holds, that is,

$$\bar{\lambda}_i + g_i(\bar{x}) > 0, \qquad i = 1, 2, \ldots, m_1.$$

(ii) $\nabla H^0(\bar{z})$ is nonsingular.

Therefore, we have the following global and superlinear convergence theorem.

Theorem 5.3 *Let Assumption A$''$ hold and the sequence $\{z_k\}$ be generated by Broyden-like method with H^ϵ defined by (5.19). Suppose that, for any accumulation point $\bar{z}$ of $\{z_k\}$, matrices $\nabla H^k(\bar{z})$ are nonsingular for all sufficiently large k. Then we have*

$$\lim_{k\to\infty} \|H^0(z_k)\| = 0$$

and every accumulation point of $\{z_k\}$ is a solution of (5.16). If in addition, Assumption B$''$ holds, then $\{z_k\}$ converges superlinearly.

Remark (i) A condition that ensures nonsingularity of $\nabla H^k(\bar{z})$ has been given in Lemma 5.5.

(ii) A convergence theorem similar to Theorem 5.3 can be established even if ϕ_{FB}^ϵ is replaced by Chen-Harker-Kanzow-Smale smoothing function.

5.3 Mixed complementarity problems

Let $X = \Pi_{i=1}^n [l_i, u_i]$, where $-\infty \leq l_i < u_i \leq \infty$ for each i. Then problem (5.15) is called the mixed complementarity problem. The mixed complementarity problem is equivalent to finding a $z \in R^n$ such that

$$\begin{cases} z_i = l_i, & \text{if} \quad f_i(z) > 0, \\ z_i = u_i, & \text{if} \quad f_i(z) < 0, \qquad i = 1, 2, \ldots, n, \\ l_i \leq z_i \leq u_i, & \text{if} \quad f_i(z) = 0, \end{cases}$$

which can be further reformulated as

$$\text{mid }\{z_i - l_i, z_i - u_i, f_i(z)\} = 0, \quad i = 1, 2, \ldots, n, \tag{5.24}$$

where mid $\{a, b, c\}$ denotes the median of a, b, c. Gabriel and Moré [15] introduced a class of smoothing functions for mixed complementarity problems, which is an extension of the class of smoothing functions introduced by Chen and Mangasarian [2]. We show that under suitable conditions, Gabriel-Moré smoothing function associated with the semismooth equation (5.24) satisfies Assumptions A and B.

Let $\rho : R \to R_+$ be a continuous density function with a bounded absolute mean, that is,

$$\kappa \triangleq \int_{-\infty}^{\infty} |s|\rho(s)ds < \infty.$$

Then the Gabriel-Moré smoothing function is defined by $H^\epsilon(z) = (H_1^\epsilon(z), \ldots, H_n^\epsilon(z))^T$, where

$$H_i^\epsilon(z) = \int_{-\infty}^{\infty} \text{mid }\{z_i - l_i, z_i - u_i, f_i(z) - \epsilon^2 s\}\rho(s)ds, \quad i = 1, 2, \ldots, n. \tag{5.25}$$

Note that the function H^ϵ given by (5.25) is slightly different from the original Gabriel-Moré smoothing function where ϵ^2 is replaced by ϵ. Gabriel-Moré smoothing function contains as a special case Chen-Harker-Kanzow-Smale smoothing function for (5.24), which corresponds to the choice

$$\rho(s) = \frac{1}{(s^2+1)^{3/2}}.$$

For any $\epsilon > 0$, H^ϵ is continuously differentiable, as long as f is continuously differentiable. Moreover, the mixed complementarity problem is equivalent to the semismooth equation $H^0(z) = 0$. Function H^ϵ possesses many good properties. Here we mention some properties that are particularly useful for the related smoothing quasi-Newton method. First, it has been proved in [5] that the density function ρ satisfies the inequality

$$\rho(s) < \left(\frac{1}{s}\right)^2 \tag{5.26}$$

for all s such that $|s|$ is sufficiently large.

The following lemma comes from [5] and [15].

Lemma 5.7 *Let H^ϵ be defined by (5.25). Suppose that f is continuously differentiable. Then the following statements are true.*

(i) For any $\epsilon > \epsilon' \geq 0$ and all $z \in R^n$, we have

$$|H_i^\epsilon(z) - H_i^{\epsilon'}(z)| \leq \kappa(\epsilon^2 - (\epsilon')^2). \tag{5.27}$$

(ii) For every $\epsilon > 0$, we have

$$\nabla H_i^\epsilon(z) = e_i - \Big(\int_{(z_i - f_i(z) - u_i)/\epsilon^2}^{(z_i - f_i(z) - l_i)/\epsilon^2} \rho(s)ds\Big)(e_i - \nabla f_i(z)), \quad i = 1, 2, \ldots, n, \tag{5.28}$$

where e_i is the i-th column of the $n \times n$ identity matrix.

The following lemma gives a sufficient condition for Lipschitz continuity of ∇H^ϵ.

Lemma 5.8 *Let H^ϵ be defined by (5.25). Suppose that f is continuously differentiable and ∇f is Lipschitz continuous. Then for each $\epsilon > 0$, there is a constant $\tau_3^\epsilon > 0$ such that for any $z, z' \in R^n$*

$$\|\nabla H^\epsilon(z) - \nabla H^\epsilon(z')\| \le \tau_3^\epsilon \|z - z'\|. \tag{5.29}$$

Proof By using (5.28), it suffices to verify that there is a constant $\bar{\tau}_3^\epsilon > 0$ such that for each $i = 1, 2, \ldots, n$

$$\Big|\int_{(z_i - f_i(z) - u_i)/\epsilon^2}^{(z_i - f_i(z) - l_i)/\epsilon^2} \rho(s)ds - \int_{(z_i' - f_i(z') - u_i)/\epsilon^2}^{(z_i' - f_i(z') - l_i)/\epsilon^2} \rho(s)ds\Big| \le \bar{\tau}_2^\epsilon \|z - z'\|. \tag{5.30}$$

It is clear that ρ is bounded. Let $\bar{\kappa} = \sup \rho(s)$. By direct deduction, we have

$$\begin{aligned}
&\Big|\int_{(z_i - f_i(z) - u_i)/\epsilon^2}^{(z_i - f_i(z) - l_i)/\epsilon^2} \rho(s)ds - \int_{(z_i' - f_i(z') - u_i)/\epsilon^2}^{(z_i' - f_i(z') - l_i)/\epsilon^2} \rho(s)ds\Big| \\
= \ & \Big|\int_{(z_i - f_i(z) - l_i)/\epsilon^2}^{(z_i' - f_i(z') - l_i)/\epsilon^2} \rho(s)ds - \int_{(z_i' - f_i(z') - u_i)/\epsilon^2}^{(z_i - f_i(z) - u_i)/\epsilon^2} \rho(s)ds\Big| \\
\le \ & \Big|\int_{(z_i - f_i(z) - l_i)/\epsilon^2}^{(z_i' - f_i(z') - l_i)/\epsilon^2} \rho(s)ds\Big| + \Big|\int_{(z_i' - f_i(z') - u_i)/\epsilon^2}^{(z_i - f_i(z) - u_i)/\epsilon^2} \rho(s)ds\Big| \\
\le \ & 2\bar{\kappa}\epsilon^{-2}(|z_i' - z_i| + |f_i(z') - f_i(z)|).
\end{aligned}$$

This together with Lipschitz continuity of f implies (5.30) with some constant $\bar{\tau}_3^\epsilon > 0$ as desired. □

Lemmas 5.7 and 5.8 show that Assumption A′ implies Assumption A. We turn to finding a sufficient condition for Assumption B to hold. We say $\bar{z}$ is a strictly complementary solution of MCP if it is a solution of MCP and satisfies

$$\bar{z}_i - l_i + f_i(\bar{z}) > 0, \quad \bar{z}_i - u_i + f_i(\bar{z}) < 0, \quad i = 1, 2, \ldots, n. \tag{5.31}$$

Let

$$\bar{\xi} = \frac{1}{2} \min_{1 \le i \le n} \{|\bar{z}_i - l_i + f_i(\bar{z})|, \ |\bar{z}_i - u_i + f_i(\bar{z})|\}.$$

If (5.31) holds, then $\bar{\xi} > 0$. Let $U(\bar{z})$ be a neighbourhood of $\bar{z}$ such that

$$U(\bar{z}) \subset \{z \mid |z_i - l_i + f_i(z)| \ge \bar{\xi}, \ |z_i - u_i + f_i(z)| \ge \bar{\xi}\}.$$

Then H^0 is continuously differentiable on $U(\bar{z})$. Moreover, it has been shown in [5] that for any $z \in U(\bar{z})$

$$\nabla H_i^0(z) = \begin{cases} \nabla f_i(z), & \text{if } z_i - f_i(z) \in (l_i, u_i), \\ e_i, & \text{if } z_i - f_i(z) \notin [l_i, u_i]. \end{cases} \tag{5.32}$$

The following lemma gives sufficient conditions for Assumption B to hold.

Lemma 5.9 *Let Assumption A′ and condition (i) in Assumption B hold. Suppose that $\bar{z}$ is a strictly complementary solution of MCP. Then Assumption B holds.*

Proof The semismoothness of H^0 follows from the continuous differentiability of f. In particular, $\nabla H^0(\bar{z})$ exists and is given by (5.32) with $z = \bar{z}$. If $\bar{z}_i - f_i(\bar{z}) \in (l_i, u_i)$, then we get from (5.28) and (5.32)

$$\begin{aligned}
\nabla H_i^k(z_k) - \nabla H_i^0(\bar{z}) &= \Big(1 - \int_{[(z_k - f(z_k))_i - u_i]/\epsilon_k^2}^{[(z_k - f(z_k))_i - l_i]/\epsilon_k^2} \rho(s)ds\Big)(e_i - \nabla f_i(z_k)) + \nabla f_i(z_k) - \nabla f_i(\bar{z}) \\
&= \Big(\int_{[(z_k - f(z_k))_i - l_i]/\epsilon_k^2}^{\infty} \rho(s)ds + \int_{-\infty}^{[(z_k - f(z_k))_i - u_i]/\epsilon_k^2} \rho(s)ds\Big)(e_i - \nabla f_i(z_k)) \\
&\quad + \nabla f_i(z_k) - \nabla f_i(\bar{z}).
\end{aligned}$$

This implies $\nabla H_i^k(z_k) - \nabla H_i^0(\bar{z}) \to 0$. If $\bar{z}_i - f_i(\bar{z}) \notin [l_i, u_i]$, we also get from (5.28) and (5.32)

$$\nabla H_i^k(z_k) - \nabla H_i^0(\bar{z}) = -\int_{[(z_k - f(z_k))_i - u_i]/\epsilon_k^2}^{[(z_k - f(z_k))_i - l_i]/\epsilon_k^2} \rho(s)ds\ (e_i - \nabla f_i(z_k)).$$

This also implies $\nabla H_i^k(z_k) - \nabla H_i^0(\bar{z}) \to 0$. Therefore, we have shown (4.1).

Since $\epsilon_k \to 0$, we have for all k sufficiently large and any $z \in U(\bar{z})$

$$\begin{aligned}
&\Big|\int_{(z_i - f_i(z) - u_i)/\epsilon_k^2}^{(z_i - f_i(z) - l_i)/\epsilon_k^2} \rho(s)ds - \int_{(z_i - f_i(z) - u_i)/\epsilon_{k-1}^2}^{(z_i - f_i(z) - l_i)/\epsilon_{k-1}^2} \rho(s)ds\Big| \\
\le\ & \Big|\int_{(z_i - f_i(z) - l_i)/\epsilon_{k-1}^2}^{(z_i - f_i(z) - l_i)/\epsilon_k^2} \rho(s)ds\Big| + \Big|\int_{(z_i - f_i(z) - u_i)/\epsilon_k^2}^{(z_i - f_i(z) - u_i)/\epsilon_{k-1}^2} \rho(s)ds\Big| \\
\le\ & \Big|\int_{(z_i - f_i(z) - l_i)/\epsilon_{k-1}^2}^{(z_i - f_i(z) - l_i)/\epsilon_k^2} \frac{1}{s^2}ds\Big| + \Big|\int_{(z_i - f_i(z) - u_i)/\epsilon_k^2}^{(z_i - f_i(z) - u_i)/\epsilon_{k-1}^2} \frac{1}{s^2}ds\Big| \\
=\ & \frac{\epsilon_{k-1}^2 - \epsilon_k^2}{|z_i - f_i(z) - l_i|} + \frac{\epsilon_{k-1}^2 - \epsilon_k^2}{|z_i - f_i(z) - u_i|} \\
\le\ & 2\bar{\xi}^{-1}(\epsilon_{k-1}^2 - \epsilon_k^2),
\end{aligned}$$

where the second inequality follows from (5.26). This together with (5.28) implies (4.2). Similarly, we have for all k sufficiently large and any $z, z' \in U(\bar{z})$

$$\Big|\int_{(z_i - f_i(z) - u_i)/\epsilon_k^2}^{(z_i - f_i(z) - l_i)/\epsilon_k^2} \rho(s)ds - \int_{(z'_i - f_i(z') - u_i)/\epsilon_k^2}^{(z'_i - f_i(z') - l_i)/\epsilon_k^2} \rho(s)ds\Big|$$

$$
\begin{aligned}
&\le \Big|\int_{(z_i-f_i(z)-l_i)/\epsilon_k^2}^{(z_i'-f_i(z')-l_i)/\epsilon_k^2}\rho(s)ds\Big| + \Big|\int_{(z_i'-f_i(z')-u_i)/\epsilon_k^2}^{(z_i-f_i(z)-u_i)/\epsilon_k^2}\rho(s)ds\Big| \\
&\le \Big|\int_{(z_i-f_i(z)-l_i)/\epsilon_k^2}^{(z_i'-f_i(z')-l_i)/\epsilon_k^2}\frac{1}{s^2}ds\Big| + \Big|\int_{(z_i'-f_i(z')-u_i)/\epsilon_k^2}^{(z_i-f_i(z)-u_i)/\epsilon_k^2}\frac{1}{s^2}ds\Big| \\
&= \epsilon_k^2\Big|\frac{1}{z_i-f_i(z)-l_i} - \frac{1}{z_i'-f_i(z')-l_i}\Big| + \epsilon_k^2\Big|\frac{1}{z_i'-f_i(z')-u_i} - \frac{1}{z_i-f_i(z)-u_i}\Big| \\
&\le 2\epsilon_k^2\bar{\xi}^{-2}(|z_i-z_i'| + |f_i(z)-f_i(z')|).
\end{aligned}
$$

This together with (5.28) yields (4.3).

Equations (4.4) follow from (5.27) and the fact $\epsilon_k \le \gamma\|H^0(z_k)\|$ immediately. The proof is then complete. □

From Lemma 5.9, we see that the following assumption together with Assumption A$'$ ensures the conditions of Assumption B.

Assumption B$'''$ (i) The sequence $\{z_k\}$ generated by Algorithm 1 converges to a strictly complementary solution $\bar{z}$ of the mixed complementarity problem.

(ii) $\nabla H^0(\bar{z})$ is nonsingular.

We now state a global and superlinear convergence theorem for Broyden-like method for MCP.

Theorem 5.4 *Let Assumption A$'$ hold and the sequence $\{z_k\}$ be generated by Broyden-like method with H^ϵ defined by (5.25). Suppose that, for any accumulation point $\bar{z}$ of $\{z_k\}$, matrices $\nabla H^k(\bar{z})$ are nonsingular for all sufficiently large k. Then we have*

$$\lim_{k\to\infty}\|H^0(z_k)\| = 0$$

and every accumulation point of $\{z_k\}$ is a solution of (5.16). If in addition, Assumption B$'''$ holds, then $\{z_k\}$ converges superlinearly.

6 Discussion

In this paper, we have presented a general Broyden-like method for solving semismooth equations including semismooth reformulations of nonlinear complementarity problems, variational inequality problems and mixed complementarity problems. We have established global and superlinear convergence theorems under suitable conditions. For nonlinear complementarity problems, global convergence only requires that F is a P_0 function with Lipschitz continuous Jacobian. On the other hand, for variational inequality problems, global convergence needs the assumption that $\nabla H^k(\bar{z})$ is nonsingular for all k sufficiently large, as in the global convergence theorems established in [26], [30] and [35]. Although this condition might appear somewhat artificial, it is implied, as shown in Lemma 5.5, by a condition that is reminiscent of the second order sufficient condition in nonlinear programming. Conditions that guarantee nonsingularity of ∇H^ϵ for some smoothing functions related to the mixed complementarity problem can be found in [32]. Theorem 4.1 is an important contribution of the paper. It shows that superlinear convergence of a smoothing Broyden-like method can be established without assuming the strict complementarity condition at the limit point. This provides the possibility of developing a globally convergent Broyden-like method that enjoys superlinear convergence even if strict complementarity fails to hold.

References

[1] B. Chen and P.T. Harker, Smooth approximations to nonlinear complementarity problems, *SIAM J. Optim.*, 7 (1997), 403-420.

[2] C. Chen and O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, *Comput. Optim. Appl.*, 5 (1996), 97-138.

[3] X. Chen, Superlinear convergence of smoothing quasi-Newton methods for nonsmooth equations, *J. Comp. Appl. Math.*, 80 (1997), 105-126.

[4] X. Chen and L. Qi, A parameterized Newton method and a quasi-Newton method for nonsmooth equations, *Comput. Optim. Appl.*, 3 (1994), 157–179.

[5] X. Chen, L. Qi and D. Sun, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, *Math. Comput.*, 67 (1998), 519-540.

[6] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley New York, 1983.

[7] T. De Luca, F. Facchinei and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, *Math. Programming*, 75 (1996), 407-440.

[8] J.E. Dennis, Jr. and J.J. Moré, A characterization of superlinear convergence and its application to quasi-Newton methods, *Math. Comput.*, 28 (1974), 549-560.

[9] J.E. Dennis, Jr. and J.J. Moré, Quasi-Newton methods, Motivation and the theory, *SIAM Rev.*, 19 (1977), 46–89.

[10] F. Facchinei and J. Soares, A new merit function for nonlinear complementarity problems and related algorithm, *SIAM J. Optim.*, 7 (1997), 225-247.

[11] A. Fischer, A special Newton-type optimization method, *Optimization*, 24 (1992), 269-284.

[12] A. Fischer, A Newton-type method for positive semidefinite linear complementarity problems, *J. Optim. Theory Appl.*, 86 (1995), 585-608.

[13] R. Fletcher, *Practical Methods of Optimization,* Second Edition, John Wiley & Sons, Chichester, 1987.

[14] M. Fukushima and L. Qi, eds., *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Kluwer Academic Publishers, Boston, 1998.

[15] S.A. Gabriel and J.J. Moré, Smoothing of mixed complementarity problems, M.C. Ferris and J.S. Pang, eds., *Complementarity and Variational Problems: State of the Art*, SIAM Philadelpha, Pennsylvania, 1997, pp. 105-116.

[16] C. Geiger and C. Kanzow, On the solution of monotone complementarity problems, *Comput. Optim. Appl.*, 5 (1996), 155-173.

[17] C. M. Ip and J. Kyparisis, Local convergence of quasi-Newton methods for B-differentiable equations, *Math. Programming*, 56 (1992), 71-89.

[18] H. Jiang and L. Qi, A new nonsmooth equations approach to nonlinear complementarity problems, *SIAM J. Control Optim.*, 35 (1997), 178-193.

[19] C. Kanzow and H. Jiang, A continuation method for (strongly) monotone variational inequalities, *Math. Programming*, 81 (1998), 103-125.

[20] C. Kanzow and M. Fukushima, Equivalence of the generalized complementarity problem to differentiable unconstrained minimization, *J. Optim. Theory Appl.*, 90 (1996), 581-603.

[21] M. Kojima and S. Shindo, Extension of Newton and quasi-Newton methods to systems of PC^1 equations, *J. Oper. Res. Soc. Japan*, 29 (1986), 352–374.

[22] D. Li, Global convergence of nonsingular Broyden's method for solving unconstrained optimization, *Chinese J. Numer. Math. Appl.*, 17 (1995), No.4 85-94.

[23] D. Li and M. Fukushima, A derivative-free line seach and global convergence of Broyden-like method for nonlinear equations, *Optim. Methods Softw.*, to appear.

[24] D. Li and M. Fukushima, Smoothing Newton and quasi-Newton methods for mixed complementarity problems, *Comput. Optim. Appl.*, to appear.

[25] D. Li and M. Fukushima, Broyden-like methods for variational inequality and nonlinear complementarity problems with global and superlinear convergence, Technical Report 98016, Department of Applied Mathematics and Physics, Kyoto University, September 1998.

[26] D. Li and Y. Yang, A globally convergent quasi-Newton method for nonlinear complementarity problems, *Math. Programming*, submitted.

[27] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, *SIAM J. Control Optim.*, 15 (1977), 957-972.

[28] F.A. Potra, L. Qi and D. Sun, Secant methods for semismooth equations, *Numerische Mathematik*, 80 (1998), 305–324.

[29] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.*, 18 (1993), 227-244.

[30] L. Qi and X. Chen, A globally convergent successive approximation method for severely nonsmooth equations, *SIAM J. Control Optim.*, 33 (1995), 402-418.

[31] L. Qi and J. Sun, Nonsmooth equations and smoothing Newton methods, Technical Report, School of Mathematics, University of New South Wales, Sydney, April 1998.

[32] L. Qi, D. Sun and G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, AMR 97/13, Applied Mathematics Report, University of New South Wales, Sydney, June 1997.

[33] D. Sun and J. Han, Newton and quasi-Newton methods for a class of nonsmooth equations and related problems, *SIAM J. Optim.*, 7 (1997), 463-480.

[34] N. Yamashita and M. Fukushima, Modified Newton methods for solving a semismooth reformulation of monotone complementarity problems, *Math. Programming*, 76 (1997), 469-491.

[35] S. Zhou, D. Li and J. Zeng, A successive approximation quasi-Newton process for nonlinear complementarity problems, D.Z. Du, L. Qi and R.S. Womersley, eds., *Recent Advances in Nonsmooth Optimization*, World Scientific Publishers, Singapore, 1995, pp. 459-472.