Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 14, No. 4, pp. 723–737, 2018

ON THE Q-LINEAR CONVERGENCE RATE OF A CLASS OF METHODS FOR MONOTONE NONLINEAR EQUATIONS*

Weijun Zhou and Dong-Hui Li

Abstract: Recently, Li and Li [IMA J. Numer. Anal., 31 (2011) 1625-1635] proposed a class of methods for solving large-scale systems of monotone equations which may be nonsmooth. Those methods have been proved to possess strong global convergence property in the sense that the whole iterative sequence converges to a solution of the equation. However, the convergence rate of the methods is not known. In this paper, we present a new line search and show that the class of methods with that line search still remain strong global convergence property. In addition, the sequence generated by any one of this class of methods converges Q-linearly to a solution of the monotone equation when the underlying function is Lipschitz continuous and all elements of its generalized Jacobian at the solution are nonsingular. Some numerical results are also reported.

Key words: *systems of monotone equations; Nonsmooth; Line search; Global convergence; Q-linear convergence*

Mathematics Subject Classification: *90C30, 65K05*

1 Introduction

In this paper, we consider iterative methods for solving large-scale systems of monotone equations. Such equations arise in various practical situations [10, 12, 13]. The study of numerical methods for this kind of equations has received much attention. By exploiting the structure of monotonicity, Solodov and Svaiter [10] presented a Newton type method, which has the truly global convergence property that the iterative sequence converges to a solution of the equation. Zhou and Toh [13] extended this method to monotone equations with singular solutions and established its superlinear convergence under the local error bound condition, which is weaker than the nonsingularity condition. Zhou and Li [14] introduced a globally convergent quasi-Newton method for monotone equations.

Recently, Li and Li [7] introduced a class of iterative methods for solving large-scale systems of monotone equations. These methods converge globally. However, no local convergence properties of the methods has been studied. In this paper, we will propose a new line search and show that the class of methods, with the proposed line search, are globally and Q-linearly convergent even if the equation is not differentiable. Here the Q stands for quotient.

*This work was supported by the NSF (11371073, 11371154 and 11461015) of China, the NSF (14JJ3084) of Hunan Province, and Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering

The paper is organized as follows. In Section 2, we present the new line search. In Section 3, we discuss the convergence properties of the class of methods with the proposed line search. In Section 4, we do some numerical experiments to show its efficiency. In Section 5, we make some conclusions.

2 Algorithm and Line Search

Consider the following nonlinear equation:

$$F(x) = 0, \tag{2.1}$$

where the function $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and monotone, but not necessarily differentiable. The so-called monotonicity here means that F satisfies

$$< F(x) - F(y), x - y > \geq 0, \quad \forall x, y \in \mathbb{R}^n, \tag{2.2}$$

where $< \cdot, \cdot >$ denotes the inner product in $\mathbb{R}^n$. Throughout the paper, we denote $F_k := F(x_k)$, $y_k := F_{k+1} - F_k$, $s_k := x_{k+1} - x_k$, and assume that the solution set of (2.1) is not empty.

Note that the methods in [10, 13, 14, 9] possess global convergence properties by using the hyperplane projection technique, but these methods are not suitable for large-scale problems since they need compute or store matrices. On the other hand, it is well-known that the nonlinear conjugate gradient methods are efficient algorithms for large-scale unconstrained optimization problems. In order to take advantage of these methods, Li and Li [7] proposed a class of iterative methods with global convergence for solving the large-scale problem (2.1), whose steps are given below.

Algorithm 2.1. Choose an initial point $x_0 \in \mathbb{R}^n$, and constants $\sigma > 0$, $\beta > 0$ and $\rho \in (0, 1)$. Let $k = 0$.

Step 1. Compute a direction d_k.

Step 2. Compute the stepsize $\alpha_k := \max\{\beta\rho^i : i = 0, 1, \cdots ,\}$ such that

$$- < F(x_k + \alpha_k d_k), d_k > \geq \sigma \|F(x_k + \alpha_k d_k)\| \alpha_k \|d_k\|^2. \tag{2.3}$$

Step 3. Set

$$z_k := x_k + \alpha_k d_k. \tag{2.4}$$

Step 4. Let next iterate x_{k+1} be given by

$$x_{k+1} := x_k - \frac{< F(z_k), x_k - z_k >}{\|F(z_k)\|^2} F(z_k). \tag{2.5}$$

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

Remark 2.2. It is clear that Algorithm 2.1 is well defined if d_k satisfies $< F_k, d_k >< 0$. The iterative point x_{k+1} given by (2.5) is the projection of x_k on the hyperplane

$$H_k := \{x \in \mathbb{R}^n : < F(z_k), x - z_k >= 0\}.$$

This hyperplane projection can ensure that the sequence of the distances between the iterates and the solution set of (2.1) is decreasing, which is independent of the line search used. This important result is specified as follows.

Lemma 2.3 ([10, Lemma 2.1]). *If* $F(x^*) = 0$ *and* $< F(z_k), x_k - z_k >> 0$, *then* x_{k+1}, *determined by* (2.5), *satisfies*

$$\|x_{k+1} - x^*\|^2 \le \|x_k - x^*\|^2 - \|x_{k+1} - x_k\|^2. \tag{2.6}$$

In [7], Li and Li studied the following three practical methods:

(i) the SG-like method: $d_k := -\delta_k F_k$ with $\delta_{\min} \le \delta_k \le \delta_{\max}$, and $\delta_{\min}, \delta_{\max}$ are two positive constants;

(ii) the MPRP-based method: $d_0 := -F_0$ and

$$d_k := -F_k + \beta_k^{PRP} d_{k-1} - \theta_k y_{k-1}, \quad k \ge 1, \tag{2.7}$$

where

$$y_{k-1} := F_k - F_{k-1}, \quad \beta_k^{PRP} := \frac{< F_k, y_{k-1} >}{\|F_{k-1}\|^2}, \quad \theta_k := \frac{< F_k, d_{k-1} >}{\|F_{k-1}\|^2}; \tag{2.8}$$

(iii) the TPRP-based method: $d_0 := -F_0$ and

$$d_k := -F_k + \beta_k^{PRP} \Big(I - \frac{F_k F_k^T}{\|F_k\|^2} \Big) d_{k-1}, \quad k \ge 1, \tag{2.9}$$

where I is the identity matrix and β_k^{PRP} is given by (2.8).

It is easy to verify that the search direction d_k in the MPRP-based and TPRP-based methods satisfy the important relation

$$< d_k, F_k >= -\|F_k\|^2. \tag{2.10}$$

Throughout the paper, we always suppose that the following assumption holds, which is the same as that in [7].

Assumption 2.4. The function $F : \mathbb{R}^n \to \mathbb{R}^n$ is monotone and Lipschitz continuous; that is, there exists a constant $L > 0$ such that

$$\|F(x) - F(y)\| \le L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n. \tag{2.11}$$

The following theorem comes from [7].

Theorem 2.5. *Let the iterative sequence* $\{x_k\}$ *be generated by the SG-like method, or the MPRP-based method, or the TPRP-based method with the line search* (2.3). *Then, the whole sequence* $\{x_k\}$ *converges to* x^* *with* $F(x^*) = 0$.

The above theorem shows that the sequence $\{x_k\}$ generated by Algorithm 2.1 converges globally. However, no local convergence properties of Algorithm 2.1 has been studied. In this paper, we further study the class of derivative-free methods. We will propose a new line search and show that the methods with this line search are globally and Q-linearly convergent even if F is nonsmooth.

We will focus on the MPRP-based method. The results can be extended to the other two methods easily in a similar way.

Let $\beta > 0, \rho \in (0,1)$ and $\sigma \in (0,1)$ be constants. We propose a line search technique to compute the stepsize $\alpha_k := \max\{\beta\rho^i : i = 0, 1, \cdots,\}$ such that

$$-< F(x_k + \alpha_k d_k), d_k > \geq \sigma \|F(x_k + \alpha_k d_k)\| \|F_k\|. \tag{2.12}$$

Clearly, this line search is well defined if d_k satisfies $< F_k, d_k > = -\|F_k\|^2$. In fact, as $\alpha \to 0^+$, the left and right sides of (2.12) tend to $- < F_k, d_k > = \|F_k\|^2$ and $\sigma\|F_k\|^2$, respectively. Hence (2.12) is satisfied for all $\alpha_k > 0$ sufficiently small. Moreover, it is not difficult to show that Lemma 2.3 still holds.

In the case d_k is determined by the SG-like method, we can further restrict $\sigma \in (0, \delta_{\min})$. The above properties obviously remain true.

3 Global and Q-linear Convergence

Throughout this section, we suppose that $\{x_k\}$ is generated by Algorithm 2.1 with α_k determined by (2.12). We also suppose that the conditions in Assumption 2.1 hold.

Clearly, (2.6) implies that $\{\|x_k - x^*\|\}$ is decreasing. Then

$$\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0. \tag{3.1}$$

From (2.4), (2.5) and the line search condition (2.12), we have

$$\|x_{k+1} - x_k\| = \Big| \frac{< F(z_k), x_k - z_k >}{\|F(z_k)\|^2} F(z_k) \Big| \leq \|z_k - x_k\| = \alpha_k \|d_k\|, \tag{3.2}$$

and

$$\|x_{k+1} - x_k\| = \Big| \frac{< F(z_k), x_k - z_k)}{\|F(z_k)\|^2} F(z_k) \Big| = \alpha_k \frac{- < F(z_k), d_k >}{\|F(z_k)\|} \geq \sigma\alpha_k \|F_k\|. \tag{3.3}$$

It follows from (3.1) and (3.3) that

$$\lim_{k\to\infty} \alpha_k \|F_k\| = 0. \tag{3.4}$$

The following lemma gives a lower bound to the stepsize α_k.

Lemma 3.1. *Let the sequence $\{x_k\}$ be generated by the MPRP-based method with line search* (2.12)*. Then we have*

$$\alpha_k \geq \min\Big\{\beta, \frac{\rho}{L} \frac{\|F_k\|^2 - \sigma\|F(z_k')\|\|F_k\|}{\|d_k\|^2}\Big\}, \tag{3.5}$$

where $z_k' := x_k + \alpha_k' d_k$ and $\alpha_k' := \alpha_k/\rho$.

Proof. If $\alpha_k \neq \beta$, then $\alpha_k' = \alpha_k/\rho$ does not satisfy the line search condition (2.12), that is,

$$- < F(z_k'), d_k > < \sigma\|F(z_k')\|\|F_k\|, \tag{3.6}$$

which together with (2.11) implies

$$- < F_k, d_k > -\sigma\|F(z_k')\|\|F_k\| \leq < F(z_k') - F_k, d_k > \leq L\alpha_k'\|d_k\|^2. \tag{3.7}$$

This and (2.10) yield (3.5). □

Theorem 3.2. *Let the sequence* $\{x_k\}$ *be generated by the MPRP-based method with line search* (2.12). *Then we have*

$$\liminf_{k\to\infty} \|F_k\| = 0. \tag{3.8}$$

Proof. We prove this theorem by contradiction. Suppose that (3.8) is not true. Then there exists a constant $\eta > 0$ such that

$$\|F_k\| \geq \eta, \quad \forall k \geq 0. \tag{3.9}$$

From (2.6), we know that $\{x_k\}$ is bounded. Then the sequence $\{\|F_k\|\}$ is also bounded with some upper bound $M > 0$. By (2.7)-(2.8) and (2.11), we get

$$\|d_k\| \leq \|F_k\| + \frac{2L\|F_k\|\|x_k - x_{k-1}\|\|d_{k-1}\|}{\|F_{k-1}\|^2} \leq M + \frac{2LM}{\eta^2}\|x_k - x_{k-1}\|\|d_{k-1}\|,$$

which together with (3.1) implies that these exists a positive constant M_1 such that

$$\|d_k\| \leq M_1. \tag{3.10}$$

Therefore, by (2.4), $\{z'_k\}$ and $\{\|F(z'_k)\|\}$ are bounded.
Case (i). If $\limsup_{k\to\infty} \alpha_k > 0$, we can easily get from (3.4) that $\liminf_{k\to\infty} \|F_k\| = 0$, which contradicts (3.9).
Case (ii). If $\limsup_{k\to\infty} \alpha_k = 0$, then

$$\lim_{k\to\infty} \alpha_k = 0. \tag{3.11}$$

Moreover, it follows from (2.11) and (3.10) that

$$\|F(z'_k)\| \leq \|F_k\| + L\alpha'_k\|d_k\| \leq \|F_k\| + L\rho^{-1}\alpha_k M_1.$$

This, together with (3.5) and (3.9), shows that

$$\begin{aligned}\alpha_k &\geq \min\Big\{\beta, \frac{\rho}{L}\frac{(1-\sigma)\|F_k\|^2 - \sigma L\rho^{-1}\alpha_k M_1\|F_k\|}{\|d_k\|^2}\Big\} \\ &\geq \min\Big\{\beta, \frac{\rho}{L}\frac{(1-\sigma)\eta^2 - \sigma L\rho^{-1}\alpha_k M_1 M}{M_1^2}\Big\},\end{aligned}$$

which implies

$$\lim_{k\to\infty} \alpha_k \geq \min\Big\{\beta, \frac{\rho}{L}\frac{(1-\sigma)\eta^2}{M_1^2}\Big\} > 0.$$

This contradicts (3.11). The proof is then complete. □

Lemma 2.3 and Theorem 3.2 imply the following strong global convergence theorem.

Theorem 3.3. *Let the sequence* $\{x_k\}$ *be generated by the MPRP-based method with line search* (2.12). *Then the whole sequence* $\{x_k\}$ *converges to a solution* x^* *of* (2.1).

Proof. Lemma 2.3 and Theorem 3.2 show that there exists a subsequence of $\{x_k\}$ which converges to a solution x^* of $F(x) = 0$. Therefore, the whole sequence $\{x_k\}$ converges to x^* since Lemma 2.3 implies that $\{\|x_k - x^*\|\}$ is monotone decreasing and converges. □

In what follows, we are going to investigate the local convergence property of the MPRP-based method with line search (2.12). We note that the Lipschitz continuouity of F implies that function F is differentiable almost everywhere. To obtain the linear convergence of the MPRP-based method, we further make the following assumption which will be assumed to hold in the rest of the paper.

Assumption 3.4. All $V \in \partial F(x^*)$ are nonsingular, where $F(x^*) = 0$ and $\partial F(x^*)$ is the generalized Jacobian of F at x^* in the sense of Clarke [1, 8].

It is well-known that, under the conditions of Assumption 3.1, x^* is an isolated solution of (2.1). Consequently, the sequence $\{x_k\}$ generated by Algorithm 2.1 with line search (2.12) converges to x^*.

To derive the linear convergence of $\{x_k\}$, we first show some useful lemmas.

Lemma 3.5 ([8, Proposition 3.1])**.** *If all $V \in \partial F(x)$ are nonsingular, then there is a neighborhood $N(x)$ of x and a constant C such that, for any $y \in N(x)$ and any $M \in \partial F(y)$, M is nonsingular and satisfies*

$$\|M^{-1}\| \le C. \tag{3.12}$$

Lemma 3.6 ([5, Corollary 3.4])**.** *Let $F : S \to \mathbb{R}^n$ be locally Lipschitz continuous. Then, F is monotone if and only if, for each $x \in S$, the matrices $M \in \partial F(x)$ are positive semidefinite, where $S \subseteq \mathbb{R}^n$ is an open and convex set.*

Lemma 3.7. *There exist a neighborhood $N(x^*)$ of x^* and positive constants m_1, m_2 and m_3 such that*

$$< x - y, F(x) - F(y) >\ \ge m_1\|x - y\|^2, \quad \forall x, y \in N(x^*). \tag{3.13}$$

$$m_2\|x_k - x^*\| \le \|F_k\| \le m_3\|x_k - x^*\|. \tag{3.14}$$

Proof. We prove (3.13) by contradiction. If (3.13) does not hold, then there exist sequences $\{x_l\}$, $\{y_l\}$, $\{u_{l,j}\}$, $\{V_{l,j}\}$ and $\lambda_{l,j} \ge 0, j = 1, 2, \cdots n + 1$ such that $x_l \to x^*$, $y_l \to x^*$, $u_{l,j} \to x^*$, $V_{l,j} \in \partial F(u_{l,j})$, $\sum_{j=1}^{n+1} \lambda_{l,j} = 1$ for any l and

$$< x_l - y_l, F(x_l) - F(y_l) >= \sum_{j=1}^{n+1} \lambda_{l,j} < x_l - y_l, V_{l,j}(x_l - y_l) >< \frac{1}{l}\|x_l - y_l\|^2.$$

Without any loss of generality, we assume that $\frac{x_l - y_l}{\|x_l - y_l\|} \to d$ and $V_{l,j} \to V^*$ for $j = 1, 2, \cdots n+1$. Let $l \to \infty$; then we have

$$< d, V^*d >\le 0,$$

which leads to a contradiction because Lemma 3.5 and Lemma 3.6 imply that $V^* \in \partial F(x^*)$ is positive definite. The inequality (3.14) holds clearly. The proof is then completed. □

Lemma 3.8. *Let the sequence $\{x_k\}$ be generated by the MPRP-based method with line search* (2.12)*. Then we have*

$$\|F_k\|^2 \ge m_1\alpha_k\|d_k\|^2. \tag{3.15}$$

Proof. Without any loss of generality, we assume that $\{x_k\} \subset N(x^*)$. By the line search (2.12) and (2.10), we have

$$\begin{aligned} \sigma\|F_k\|\|F(x_k + \alpha_k d_k)\| &\le - < F(x_k + \alpha_k d_k), d_k > \\ &= - < F_k, d_k > + < F_k - F(x_k + \alpha_k d_k), d_k > \end{aligned}$$

$$= \|F_k\|^2 + < F_k - F(x_k + \alpha_k d_k), d_k > .$$

This and (3.13) yield

$$\begin{aligned} \|F_k\|^2 &\geq \sigma \|F_k\| \|F(x_k + \alpha_k d_k)\| + < F(x_k + \alpha_k d_k) - F_k, d_k > \\ &\geq m_1 \alpha_k \|d_k\|^2. \end{aligned}$$

This shows that (3.15) holds. □

It follows from (2.7) and (2.11) that

$$\begin{aligned} \|d_k\| &\leq \|F_k\| + \|\beta_k^{PRP} d_{k-1}\| + \|\theta_k y_{k-1}\| \\ &\leq \|F_k\| + \frac{2L\|F_k\|\|x_k - x_{k-1}\|\|d_{k-1}\|}{\|F_{k-1}\|^2} \\ &\leq \big(1 + \frac{2L}{m_1}\big)\|F_k\| = C_1\|F_k\|, \end{aligned} \tag{3.16}$$

where $C_1 := 1 + \frac{2L}{m_1}$, the last inequality follows from (3.15) and (3.2). This and (3.3) show that

$$\lim_{k\to\infty} \alpha_k \|d_k\| = 0.$$

From (3.16) and (2.11), we know

$$\|F(z_k')\| \leq \|F_k\| + L\rho^{-1}\alpha_k\|d_k\| \leq \|F_k\| + L\beta C_1 \rho^{-1}\|F_k\|.$$

Therefore, we get from (3.5)

$$\begin{aligned} \alpha_k &\geq \min\Big\{\beta, \frac{\rho}{L}\frac{(1-\sigma)\|F_k\|^2 - \sigma L\beta C_1\rho^{-1}\alpha_k\|F_k\|^2}{\|d_k\|^2}\Big\} \\ &= \min\Big\{\beta, \frac{\rho\big((1-\sigma) - \sigma L\beta C_1\rho^{-1}\alpha_k\big)}{L}\frac{\|F_k\|^2}{\|d_k\|^2}\Big\}. \end{aligned}$$

Without any loss of generality, we assume $\alpha_k \leq \frac{\rho(1-\sigma)}{2\sigma L\beta C_1}$. Then from (3.16), we have

$$\alpha_k \geq C_2 := \min\Big\{\beta, \frac{\rho(1-\sigma)}{2LC_1^2}\Big\} > 0. \tag{3.17}$$

The following result shows that the MPRP-based method with line search (2.12) converges Q-linearly.

Theorem 3.9. *Let the sequence $\{x_k\}$ be generated by the MPRP-based method with line search* (2.12)*. Then there exists a constant $r \in (0,1)$ such that*

$$\|x_{k+1} - x^*\| \leq r\|x_k - x^*\|. \tag{3.18}$$

Proof. From (3.3) and (3.17), we have

$$\|x_{k+1} - x_k\| \geq \sigma\alpha_k\|F_k\| \geq \sigma C_2\|F_k\|.$$

This together with (2.6) and (3.14) implies

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - \sigma^2 C_2^2\|F_k\|^2 \leq (1 - \sigma^2 C_2^2 m_2^2)\|x_k - x^*\|^2. \tag{3.19}$$

This yields (3.18) with $r := \sqrt{1 - \sigma^2 C_2^2 m_2^2} < 1$. The proof is then complete. □

Remark 3.1 It is not difficult to prove the global and linear convergence of the SG-like method and the TPRP-based method with line search (2.12) by using the completely same argument as that of the MPRP-based method.

4 Numerical Experiments

In this section, we test the performance of the MPRP-based method with line search (2.12) which we call MPRP-2, and compare its performance with that of the MPRP-based method in [7] which we call MPRP-1. The parameters of the algorithms are specified as follows.

• In MPRP-1, we set the parameters as same as those in [7], namely, $\rho = 0.5$, $\sigma = 2$ and $\beta = \frac{|<F_k, d_k>|}{|<d_k, F(x_k+\epsilon d_k)-F_k>/\epsilon|}$ with $\epsilon = 10^{-8}$;

• In MPRP-2, we set $\rho = 0.1$, $\sigma = 0.5$. We use the same parameter β as that in the MPRP-1 method.

The codes were written in Matlab 7.4 and run on a personal computer with a 2.66 GHz CPU processor and 1 GB RAM memory. We stopped the iteration if the total number of iterations exceeds 10^4 or the inequality $\|F_k\| \le 10^{-4}$ is satisfied. We tested the two methods on the following 11 examples with different sizes and initial points.

Example 4.1. The discretized two-point boundary value problem [6]:

$$F(x) := \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix} x + \frac{1}{(n+1)^2}(\sin x_1 - 1, \cdots, \sin x_n - 1)^T.$$

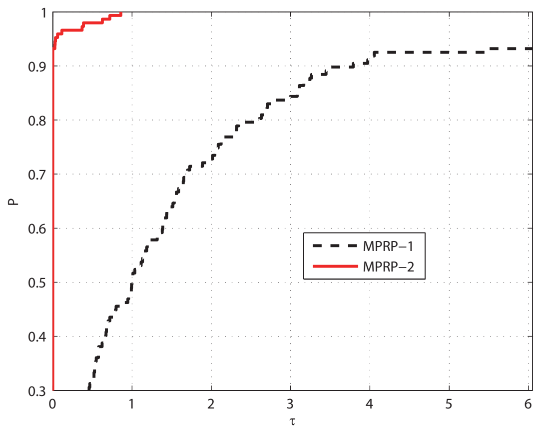

Fig. 1: Performance profiles with respect to the CPU time.

Example 4.2. The elements of $F(x)$ are given by [7]:

$$\begin{aligned} F_1(x) &:= 2x_1 + \sin x_1 - 1, \\ F_i(x) &:= -2x_{i-1} + 2x_i + \sin x_i - 1, \quad i = 2, 3, \cdots, n-1, \\ F_n(x) &:= 2x_n + \sin x_n - 1. \end{aligned}$$

Example 4.3. The gradient function of the Engval function [6]:

$$F_1(x) := x_1(x_1^2 + x_2^2) - 1,$$

Table 1: Test results for the MPRP-1 and MPRP-2 methods on the test examples.

			MPRP-1				MPRP-2			
Exa	$(x_0)_i$	n	Iter	Fcnt	Time	$\|F_k\|$	Iter	Fcnt	Time	$\|F_k\|$
4.1	0.1	50	1051	3151	2.215	9.99e-005	780	2338	1.570	1.00e-004
		100	1734	5200	3.716	9.99e-005	1538	4612	3.323	9.99e-005
		200	4529	13585	15.274	1.00e-004	3969	11905	14.528	1.00e-004
		500	*	*	*	1.32e-004	9857	29569	409.140	1.00e-004
4.1	1	20	1142	3424	2.113	9.94e-005	1107	3319	1.929	9.96e-005
		30	2311	6931	3.273	9.98e-005	2220	6658	2.483	1.00e-004
		50	5583	16747	5.982	9.99e-005	5314	15940	6.131	9.99e-005
4.1	-0.1	20	1171	3511	2.319	9.95e-005	1132	3394	1.753	9.97e-005
		30	2376	7126	3.186	9.98e-005	2283	6847	2.867	1.00e-004
		50	5754	17260	5.701	9.99e-005	5483	16447	5.636	9.99e-005
4.2	0.1	500	1032	3412	2.463	9.87e-005	992	2972	4.467	9.98e-005
		1000	1924	6257	4.160	9.97e-005	1803	5405	4.055	9.97e-005
		2000	3205	10416	8.193	9.99e-005	2851	8549	7.525	9.99e-005
		5000	6070	18522	28.661	9.98e-005	4264	12789	20.857	9.99e-005
		10000	7471	23711	66.095	9.99e-005	5384	16149	62.533	1.00e-004
4.2	1	500	988	2998	4.425	9.86e-005	978	2932	2.602	9.85e-005
		1000	1861	5678	4.837	9.98e-005	1788	5362	4.246	9.99e-005
		2000	3118	9535	8.189	9.99e-005	2835	8503	7.469	1.00e-004
		5000	5838	16915	27.034	9.99e-005	4251	12751	27.547	1.00e-004
		10000	7274	21357	65.066	1.00e-004	5374	16120	58.607	9.99e-005
4.2	10	50	578	4882	3.036	8.90e-005	340	1020	1.508	9.05e-005
		100	1126	9952	2.858	9.34e-005	662	1983	1.993	9.67e-005
		500	5490	48924	13.290	5.94e-005	3142	9425	4.592	9.44e-005
		1000	*	*	*	2.86e+000	6278	18835	11.502	9.89e-005
4.3	0.01	1000	282	1913	1.956	9.77e-005	125	377	1.846	9.78e-005
		5000	585	4888	4.978	8.94e-005	133	401	3.180	9.41e-005
		8000	743	6430	9.240	9.64e-005	135	407	3.519	9.64e-005
		10000	806	7115	12.356	9.60e-005	136	410	3.732	9.70e-005
		15000	938	8353	18.110	9.22e-005	138	416	4.656	9.62e-005
4.3	0.1	1000	105	499	1.822	9.89e-005	125	374	1.737	9.05e-005
		5000	157	918	3.360	9.37e-005	133	398	2.086	9.08e-005
		8000	180	1105	7.599	9.15e-005	135	404	2.780	9.42e-005
		10000	196	1211	3.850	9.79e-005	136	407	3.192	9.55e-005
		15000	208	1388	5.631	9.69e-005	138	413	3.419	9.58e-005
4.3	1	1000	111	398	1.569	9.52e-005	103	304	1.561	9.06e-005
		5000	159	728	2.559	9.53e-005	102	301	1.948	9.34e-005
		8000	188	929	2.872	9.19e-005	101	298	2.477	9.44e-005
		10000	193	1001	3.890	9.08e-005	101	298	2.661	9.02e-005
		15000	206	1141	4.526	9.85e-005	100	295	5.922	9.27e-005
4.3	10	1000	939	7989	3.008	9.54e-005	112	326	1.493	9.52e-005
		5000	1488	13606	10.782	9.63e-005	114	331	3.659	9.11e-005
		8000	1643	15145	17.507	9.29e-005	115	334	3.515	9.24e-005
		10000	1710	15838	22.408	9.17e-005	115	334	6.052	9.70e-005
		15000	1823	17044	34.172	8.91e-005	116	337	12.602	9.35e-005
4.4	1	1000	93	648	3.534	1.76e-005	4	7	1.515	2.45e-007
		5000	201	1625	4.057	4.94e-006	4	7	1.851	5.50e-007
		10000	296	2578	8.506	4.00e-006	4	7	2.831	7.78e-007
4.4	10	1000	585	5357	3.753	1.77e-005	6	11	1.692	2.60e-007
		5000	904	8621	15.089	4.77e-006	6	11	2.318	5.81e-007
		10000	1031	9928	30.902	4.00e-006	6	11	3.215	8.21e-007
4.4	100	1000	1166	11169	6.273	2.29e-005	13	31	2.005	2.81e-007
		5000	1485	14431	25.238	4.79e-006	13	31	2.810	6.31e-007
		10000	1612	15738	50.270	4.02e-006	13	31	3.113	8.86e-007
4.5	10	1000	311	2170	3.198	6.93e-005	174	510	2.143	9.40e-005
		2000	437	3145	8.661	6.01e-005	184	540	3.201	9.39e-005
		5000	591	4796	20.319	7.64e-005	197	578	9.212	9.28e-005
		10000	880	7140	57.712	9.49e-005	211	621	13.065	9.04e-005
4.5	100	5000	249	1454	8.541	8.28e-005	195	572	5.864	9.44e-005
		8000	287	1799	15.744	9.41e-005	202	593	9.044	9.06e-005
		10000	416	2750	25.079	9.60e-005	205	602	10.000	9.22e-005
		15000	398	2686	34.999	7.50e-005	210	617	15.315	9.96e-005
4.5	-10	3000	342	2514	8.596	9.53e-005	173	503	4.192	9.19e-005
		5000	434	3255	15.555	6.25e-005	180	524	5.440	9.47e-005
		8000	494	3990	28.301	9.61e-005	187	545	8.712	9.17e-005
		10000	542	4423	37.212	6.78e-005	190	554	11.451	9.35e-005
		15000	630	5271	69.307	6.06e-005	196	572	13.882	9.14e-005
4.5	-1	2000	228	1374	3.177	7.96e-005	180	531	1.994	9.33e-005
		5000	344	2276	6.961	7.55e-005	186	524	4.787	9.23e-005
		8000	420	2982	12.695	9.55e-005	197	581	8.493	9.64e-005
		10000	469	3383	17.407	9.16e-005	201	594	8.738	9.28e-005
		15000	539	4044	29.390	8.49e-005	207	611	9.980	9.19e-005

Table 1 continued.

			MPRP-1				MPRP-2			
Exa	$(x_0)_i$	n	Iter	Fcnt	Time	$\|F_k\|$	Iter	Fcnt	Time	$\|F_k\|$
4.6	-1	1000	78	336	1.574	9.23e-005	113	336	1.561	9.69e-005
		5000	136	731	2.974	9.45e-005	122	363	2.503	9.01e-005
		8000	160	948	3.166	8.93e-005	124	369	2.453	9.49e-005
		10000	170	1051	4.345	8.15e-005	126	375	3.140	9.25e-005
		15000	180	1221	4.691	9.36e-005	128	381	3.422	9.03e-005
		20000	212	1482	5.972	9.32e-005	127	377	4.612	9.25e-005
4.6	-0.1	1000	147	992	1.868	9.21e-005	116	346	1.713	9.85e-005
		5000	255	1979	3.486	9.62e-005	122	363	4.507	9.47e-005
		8000	291	2355	7.085	9.74e-005	124	369	3.117	9.23e-005
		10000	330	2589	6.004	9.62e-005	124	369	4.025	9.99e-005
		15000	348	2925	7.573	9.16e-005	127	377	3.605	8.62e-005
		20000	386	3221	10.791	9.84e-005	127	378	6.240	9.61e-005
4.6	0.1	1000	143	936	1.687	8.62e-005	121	361	1.723	9.72e-005
		5000	165	1098	2.550	9.04e-005	126	375	2.659	9.41e-005
		8000	185	1102	4.279	9.45e-005	128	381	2.694	9.20e-005
		10000	162	1021	4.604	8.69e-005	129	384	3.863	9.32e-005
4.7	10	1000	864	7656	5.068	8.86e-005	113	324	2.034	9.21e-005
		2000	1098	9996	10.375	7.47e-005	124	360	2.443	9.81e-005
		5000	1361	12759	27.060	8.93e-005	131	383	3.993	6.15e-005
		10000	1532	14549	60.354	9.94e-005	141	413	6.375	9.15e-005
4.7	100	1000	2808	26926	15.139	6.64e-005	204	531	3.744	9.21e-005
		5000	3287	31918	66.569	7.47e-005	205	532	4.239	9.88e-005
		10000	3415	33250	154.424	7.18e-005	202	521	25.386	9.28e-005
4.7	1000	500	4939	48568	17.460	8.32e-005	991	2104	3.556	9.42e-005
		1000	5178	51095	27.339	6.61e-005	994	2110	4.182	9.18e-005
		2000	5344	52850	48.390	8.89e-005	1000	2122	5.600	9.23e-005
		5000	5457	53980	130.397	8.88e-005	1015	2154	11.957	9.19e-005
4.8	10	100	316	2614	1.392	9.51e-005	145	432	1.107	9.38e-005
		200	404	3499	1.853	9.80e-005	139	414	1.401	9.59e-005
		500	513	4600	7.587	7.71e-005	150	447	2.914	9.16e-005
		1000	610	5545	36.432	7.47e-005	159	473	11.587	9.86e-005
		2000	679	6241	150.141	8.57e-005	170	506	24.257	9.51e-005
		5000	803	7447	1071.404	7.15e-005	164	489	255.306	9.40e-005
		10000	865	8032	4439.149	9.02e-005	153	456	699.303	9.67e-005
4.8	-10	100	342	2720	1.742	9.69e-005	151	449	1.214	9.13e-005
		200	416	3570	2.478	7.60e-005	159	473	1.276	9.32e-005
		500	532	4752	7.669	8.85e-005	163	485	3.883	9.65e-005
		1000	626	5654	38.388	7.24e-005	170	506	9.303	9.48e-005
		2000	721	6600	159.368	7.20e-005	169	503	31.378	9.27e-005
		5000	842	7795	1130.288	7.65e-005	177	527	868.606	9.98e-005
4.9	0	10	891	3117	1.481	9.91e-005	636	1906	1.228	9.99e-005
		20	5468	18149	3.383	9.94e-005	4081	12241	5.573	1.00e-004
		50	*	*	*	3.84e+000	8334	25000	12.904	1.00e-004
		80	*	*	*	1.42e+000	9090	27268	33.670	1.00e-004
		100	9174	25791	47.453	9.99e-005	7024	21070	43.767	9.98e-005
4.9	i	10	975	3276	1.457	9.96e-005	740	2218	1.011	9.98e-005
		20	5657	19406	3.848	9.91e-005	4126	12376	2.876	9.95e-005
		50	*	*	*	4.65e+000	8513	25537	14.698	1.00e-004
		80	*	*	*	1.81e+000	9695	29083	48.076	9.99e-005
		100	*	*	*	9.57e-004	7813	23437	51.648	9.96e-005
4.9	10	10	981	3338	2.093	9.93e-005	774	2320	1.140	9.94e-005
		20	5670	19081	3.438	1.00e-004	4093	12277	2.523	9.98e-005
		50	*	*	*	2.26e-004	8357	25069	27.211	9.99e-005
		80	*	*	*	1.62e+000	9286	27856	33.677	9.97e-005
		100	*	*	*	4.73e+000	7286	21856	46.005	9.99e-005
4.10	10^3	4	1184	10579	0.399	9.34e-005	193	577	0.024	9.89e-005
	10^2	4	592	4684	0.178	9.42e-005	171	511	0.021	9.98e-005
	10	4	177	728	0.040	9.27e-005	150	448	0.020	9.33e-005
	0	4	25	91	0.012	5.88e-005	109	326	0.013	9.78e-005
	-10^3	4	2488	23013	0.860	9.39e-005	157	444	0.019	9.57e-005
	-10^2	4	685	4954	0.200	9.38e-005	150	431	0.018	9.87e-005
	-10	4	177	676	0.039	9.39e-005	145	424	0.018	9.77e-005
4.11	100	500	1177	10736	4.916	8.97e-005	162	484	1.878	9.61e-005
		1000	2190	13407	10.866	8.60e-005	181	540	5.631	9.77e-005
		5000	1498	13723	41.868	8.84e-005	193	576	21.057	9.18e-005
		10000	1729	14902	90.250	8.93e-005	195	582	34.414	9.66e-005
4.11	$\frac{1}{i}$	500	121	495	2.124	9.52e-005	115	343	1.758	9.33e-005
		1000	130	629	2.776	9.00e-005	118	352	1.711	9.75e-005
		5000	220	1293	8.355	9.62e-005	126	376	5.227	9.50e-005
		10000	2161	5596	55.960	7.89e-005	129	385	86.318	9.82e-005
4.11	i	500	1413	13402	5.562	9.24e-005	166	456	1.845	9.46e-005
		1000	1683	16057	11.683	9.16e-005	224	633	2.680	9.67e-005
		5000	2286	22176	66.070	9.02e-005	224	639	8.282	9.85e-005
		10000	2546	24815	143.606	9.93e-005	251	737	16.651	9.56e-005
		15000	2701	26369	227.273	8.95e-005	243	669	147.564	9.68e-005
		20000	2811	27469	313.991	8.97e-005	233	672	22.694	9.21e-005

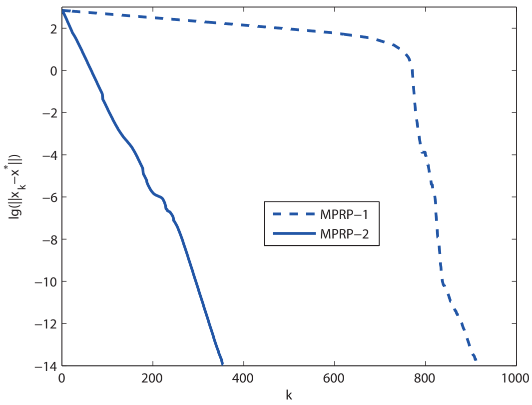

Fig. 2: Convergence rate of the two methods for Example 4.8 with $n = 5000$.

$$\begin{aligned} F_i(x) &:= x_i(x_{i-1}^2 + 2x_i^2 + x_{i+1}^2) - 1, \quad i = 2, 3, \cdots, n-1, \\ F_n(x) &:= x_n(x_{n-1}^2 + x_n^2). \end{aligned}$$

Example 4.4. A nonsmooth and monotone function [7]:

$$F_i(x) := 2x_i - \sin|x_i|, \quad i = 1, 2, \cdots, n.$$

Example 4.5. The trigonometric function [2]: for $i = 1, 2, \cdots, n$,

$$F_i(x) := 2\big(n + i(1 - \cos x_i) - \sin x_i - \sum_{j=1}^{n} \cos x_j\big)(2\sin x_i - \cos x_i).$$

Example 4.6. The Broyden tridiagonal function [2]:

$$\begin{aligned} F_1(x) &:= (3 - 0.5x_1)x_1 - 2x_2 + 1, \\ F_i(x) &:= (3 - 0.5x_i)x_i - x_{i-1} - 2x_{i+1} + 1, \quad i = 2, 3, \cdots, n-1, \\ F_n(x) &:= (3 - 0.5)x_n - x_{n-1} + 1. \end{aligned}$$

Example 4.7. The trigexp function [2]:

$$\begin{aligned} F_1(x) &:= 3x_1^3 + 2x_2 - 5 + \sin(x_1 - x_2)\sin(x_1 + x_2), \\ F_i(x) &:= -x_{i-1}e^{(x_{i-1}-x_i)} + x_i(4 + 3x_i^2) + 2x_{i+1} \\ &\quad + \sin(x_i - x_{i+1})\sin(x_i + x_{i+1}) - 8, \quad i = 2, \cdots, n-1, \\ F_n(x) &:= -x_{n-1}e^{(x_{n-1}-x_n)} + 4x_n - 3. \end{aligned}$$

The following four examples are nonsmooth equations, which come from the variational inequality problem (VIP). Let S be a nonempty and closed subset of $\mathbb{R}^n$ and H be a continuous monotone mapping from $\mathbb{R}^n$ into itself. The VIP is to find a vector $x^* \in S$ such that

$$< y - x^*, H(x^*) > \geq 0, \quad \forall y \in S. \tag{4.1}$$

Let $F(x) := x - P_S(x - H(x))$, where $P_S(u)$ denotes the projection of u onto S. It is well-known that the VIP (4.1) is equivalent to the system of equations $F(x) = 0$. In Examples

4.8-4.10, $S := \{x :\ x \geq 0\}$; in Example 4.11, $S := \{x :\ 0 \leq x \leq 1\}$. For these four examples, $H(x)$ is given as follows.

Example 4.8

$$H(x) := \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 4 \end{pmatrix} x + \big(-1, 1, -1, \cdots, (-1)^n\big)^T.$$

Example 4.9

$$H(x) := D(x) + Mx + q,$$

where $D(x) := (d_1 \arctan(x_1), \cdots, d_n \arctan(x_n))^T$ and $M := A^T A + B$. Here A, B, d, q are generated by the following Matlab code [4].

```
A=zeros(n,n); t=0; for i=1:n for j=1:n t=mod(t*31416+13846,46261);
A(i,j)=t*(10/46261)-5; end; end; B=zeros(n,n); t=0; for i=1:n for j=i+1:n
t=mod(t*42108+13846,46273); B(i,j)=t*10/46273-5; B(j,i)= - B(i,j); end; end;
M= A¨ *A+B; t=0; q=zeros(n,1); for j=1:n t=mod(t*45278+13846,46219); q(j)=t; end;
q=(q/46219 -0.5)*1000; d=zeros(n,1); for j=1:n t=mod(t*45278+13846,46219); d(j)=t;
end; d=d/46219;
```

Example 4.10 The problem [11]:

$$H(x) := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} x + \big(x_1^3 - 8, x_2^3 + 3, 2x_3^3 - 3, 2x_4^3\big)^T.$$

Example 4.11

$$\begin{aligned} H_1(x) &:= x_1 - x_2 + \frac{1}{3}(x_1 - x_2)^3 - 1, \\ H_i(x) &:= -x_{i-1} + 2x_i - x_{i+1} + \frac{i}{3}(x_i - x_{i+1})^3 \\ &\quad -\frac{i-1}{3}(x_{i-1} - x_i)^3 + (-1)^i i, \quad i = 2, \cdots, n-1, \\ H_n(x) &:= -x_{n-1} + x_n - \frac{n-1}{3}(x_{n-1} - x_n)^3 + (-1)^n n. \end{aligned}$$

Table 1 lists the numerical results for the methods on the test problems, where each column has the following meaning:

- Exa: the test example;
- Iter: the total number of iterations;
- Fcnt: the total number of function evaluations;
- Time: the CPU time in seconds;
- $\|F_k\|$: the norm of the residual at the stopping point;
- $*$: the method failed to find a solution within 10^4 iterations.

We can see from Table 1 that the MPRP-2 method performs much better than the MPRP-1 method, since the MPRP-2 method need much less iterations and function evaluations and CPU time. In order to show the performance of the two methods clearly, we plotted Fig. 1 according to the data about CPU time in Table 1 by using the performance

Table 2: Numerical results for the Newton-based method and the MPRP-2 method on Example 4.12.

			Newton-based				MPRP-2			
a_i	x_0	n	Iter	Fcnt	Time	$\|F_k\|$	Iter	Fcnt	Time	$\|F_k\|$
1	$\bar{x}$	10	54	275	1.133	9.69e-005	269	805	1.033	9.82e-005
		50	448	3355	1.453	9.79e-005	3222	9664	1.604	9.99e-005
		100	856	6583	2.866	9.15e-005	6708	20122	3.540	1.00e-004
		500	883	6712	15.828	9.72e-005	6740	20218	11.960	1.00e-004
		1000	818	6230	71.450	9.75e-005	6740	20218	85.116	1.00e-004
		2000	848	6454	405.167	1.00e-004	6740	20218	**38.463**	1.00e-004
1	$\hat{x}$	10	62	298	1.083	8.97e-005	331	985	0.862	9.99e-005
		50	456	3280	1.027	9.95e-005	3798	11386	1.634	1.00e-004
		100	1090	8089	2.387	9.65e-005	8110	24322	4.033	1.00e-004
		500	802	5952	14.077	9.26e-005	6461	19374	11.679	1.00e-004
		1000	764	5607	66.040	9.79e-005	6447	19332	16.802	1.00e-004
		2000	914	6880	441.646	9.53e-005	6444	19323	**30.286**	1.00e-004
i	$\bar{x}$	10	56	289	1.677	9.05e-005	269	805	0.950	9.86e-005
		50	468	3539	2.196	8.88e-005	3222	9664	1.929	1.00e-004
		100	891	6716	2.437	9.08e-005	6709	20125	3.990	1.00e-004
		500	881	6719	17.073	9.40e-005	6741	20221	13.572	1.00e-004
		1000	863	6641	74.573	9.47e-005	6741	20221	30.308	1.00e-004
		2000	954	7290	455.479	8.97e-005	6741	20221	**38.248**	1.00e-004
i	$\hat{x}$	10	67	297	1.299	9.15e-005	330	974	0.836	9.94e-005
		50	472	3432	1.270	9.85e-005	3805	11381	1.629	1.00e-004
		100	1082	7994	2.359	9.85e-005	8128	24345	3.747	1.00e-004
		500	823	5871	15.203	9.54e-005	6544	19563	10.911	1.00e-004
		1000	871	6267	74.682	9.75e-005	6601	19717	17.467	1.00e-004
		2000	834	5942	400.186	9.81e-005	6683	19941	**27.861**	1.00e-004

profiles of Dolan and Moré [3]. Fig. 1 indicates that the MPRP-2 method completely overcomes the MPRP-1 method since its corresponding curve is much higher than that of the MPRP-1 method. Moreover, to verify the linear convergence rate of the methods, we plotted the curve "$k - \lg \|x_k - x^*\|$" on Example 4.8 with $x_0 = 10 * \text{ones}(n, 1)$ and $n = 5000$, where $x^* = (\frac{1}{4}, 0, \frac{1}{4}, 0, \cdots, \frac{1}{4}, 0)^T$ is the solution. Fig. 2 shows clearly Q-linear convergence of the MPRP-2 method for Example 4.8.

In order to compare the performance of the Newton-based method by Zhou and Toh in [13] and the MPRP-2 method, we also did some experiments on the following example.
Example 4.12 $F(x) := \nabla f(x) = 0$, where $\nabla f(x)$ is the gradient of the function

$$f(x) := \frac{1}{2}\sum_{i=1}^{n-1}(x_i - x_{i+1})^2 + \frac{1}{12}\sum_{i=1}^{n-1} a_i(x_i - x_{i+1})^4, \quad a_i \geq 0.$$

We set the parameters $\gamma_1 = 1, \delta = 0.8, \sigma = 0.1, t = \frac{1}{2}$ and $\kappa = 0$ in the Newton-based method. It is easy to see that Example 4.12 has the solution set $X^* = \{x \in \mathbb{R}^n : x_1 = x_2 = \cdots = x_n\}$. Numerical results are listed in Table 2, where $\bar{x} = (1, \frac{1}{2}, \cdots, \frac{1}{n})^T, \hat{x} = (10, 0, 10, 0, \cdots, 10, 0)^T$. From Table 2, we can see that the Newton-based method need less iterations and function evaluations, but it requires more CPU time when the size is relatively large.

5 Conclusions

In this paper, we proposed a line search technique and established the global and Q-linear convergence of the MPRP-based method with this line search for solving large-scale monotone nonlinear equations. The new line search is well-defined and the new algorithms are Q-linearly convergent due to the important relation (2.10). In our numerical experiments, we noted that the initial stepsize choices have important impact on computational efficiency of the methods. How to choose a suitable initial stepsize, such as adopting self-adaptive technique, is our further study. Moreover, it is worth discussing the convergence rate of the methods under the weaker local error bound condition and extending the methods to general nonlinear equations without monotonicity.

References

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

[2] W. La Cruz, J.M. Martinez and M. Raydan, Spectral residual method without gradient information for solving large-scale nonlinear systems of equations, *Math. Comput.* 75 (2006) 1429–1448.

[3] E.D. Dolan and J.J. Moré, Benchmarking optimization software with performance profiles, *Math. Program.* 91 (2002) 201–213.

[4] B. He, X. Yuan and J. Zhang, Comparison of two kinds of prediction-correction methods for monotone variational inequalities, *Comput. Optim. Appl.* 27 (2004) 247–267.

[5] V. Jeyakumar, D.T. Luc and S. Schaible, Characterizations of generalized monotone nonsmooth continuous maps using approximate Jacobians, *J. Convex Anal.* 5 (1998) 119–132.

[6] D. Li and M. Fukushima, A globally and superlinearly convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations, *SIAM J. Numer. Anal.* 37 (1999) 152–172.

[7] Q. Li and D. Li, A class of derivative-free methods for large-scale nonlinear monotone equations, *IMA J. Numer. Anal.* 31 (2001) 1625–1635.

[8] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Math. Program.* 58 (1993) 353–367.

[9] M.V. Solodov and B.F. Svaiter, A new projection method for variational inequality problems, *SIAM J. Contr. Optim.* 37 (1999) 765–776.

[10] M.V. Solodov and B.F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations, in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth, and Smoothing Methods*, M. Fukushima and L. Qi (eds.), Kluwer Academic Publishers, Dordrecht, 1998, pp. 355–369.

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

[12] Y.B. Zhao and D. Li, Monotonicity of fixed point and normal mapping associated with variational inequality and its application, *SIAM J. Optim.* 4 (2001) 962–973.

[13] G. Zhou and K.C. Toh, Superlinear convergence of a Newton-type algorithm for monotone equations, *J. Optim. Theory Appl.* 125 (2005) 205–221.

[14] W. Zhou and D. Li, A globally convergent BFGS method for nonlinear monotone equations without any merit functions, *Math. Comput.* 77 (2008) 2231–2240.

Manuscript received 6 July 2014
revised 14 November 2014
accepted for publication 31 December 2014

Weijun Zhou
Department of Mathematics, Changsha University of Science and Technology
Changsha 410004, China
E-mail address: weijunzhou@126.com

Dong-Hui Li
School of Mathematical Sciences, South China Normal University
Guangzhou 510631, China
E-mail address: dhli@scnu.edu.cn