Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 9175496, 11 pages
https://doi.org/10.1155/2020/9175496

Research Article

Two Improved Conjugate Gradient Methods with Application in Compressive Sensing and Motion Control

Min Sun,[1] **Jing Liu**,[2] **and Yaru Wang**[3]

[1]*School of Mathematics and Statistics, Zaozhuang University, Zaozhuang, Shandong 277160, China*
[2]*School of Data Sciences, Zhejiang University of Finance and Economics, Hangzhou, Zhejiang 310018, China*
[3]*School of Opto-Electronic Engineering, Zaozhuang University, Zaozhuang, Shandong 277160, China*

Correspondence should be addressed to Min Sun; ziyouxiaodou@163.com

Received 27 December 2019; Accepted 27 February 2020; Published 5 May 2020

Academic Editor: Thomas Schuster

To solve the monotone equations with convex constraints, a novel multiparameterized conjugate gradient method (MPCGM) is designed and analyzed. This kind of conjugate gradient method is derivative-free and can be viewed as a modified version of the famous Fletcher–Reeves (FR) conjugate gradient method. Under approximate conditions, we show that the proposed method has global convergence property. Furthermore, we generalize the MPCGM to solve unconstrained optimization problem and offer another novel conjugate gradient method (NCGM), which satisfies the sufficient descent property without any line search. Global convergence of the NCGM is also proved. Finally, we report some numerical results to show the efficiency of two novel methods. Specifically, their practical applications in compressive sensing and motion control of robot manipulator are also investigated.

1. Introduction

Let $\mathscr{R}, \mathscr{R}^n$, and $\mathscr{R}^{m\times n}$ be the sets of real numbers, n dimensional real column vectors, and $m\times n$ dimensional real matrices, respectively. This paper is concerned with the following two active subjects in numerical analysis.

(i) Monotone equations with convex constraints: finding a vector $x^* \in \mathscr{X}$ such that

$$F(x^*) = 0, \quad x^* \in \mathscr{X}, \tag{1}$$

where F: $\mathscr{R}^n \longrightarrow \mathscr{R}^n$ is a continuous nonlinear mapping (not necessarily smooth) and $\mathscr{X} \subseteq \mathscr{R}^n$ is a nonempty convex set.

(ii) Unconstrained optimization problem: finding a vector $x^* \in \mathscr{R}^n$ such that

$$x^* \in \operatorname{argmin}_{x\in\mathscr{R}^n} f(x), \tag{2}$$

where f: $\mathscr{R}^n \longrightarrow \mathscr{R}$ is a continuously differentiable function whose gradient is denoted by $g(x)$.

Problems (1) and (2) are interchangeable in some sense. In fact, setting $f(x) = (1/2)\|F(x)\|^2$, problem (1) with $\mathscr{X} = \mathscr{R}^n$ can be transformed into problem (2). Similarly, the necessary condition of problem (2), i.e., $g(x)^* = 0$, is a special case of problem (1). Therefore, the design of numerical methods for the two problems often inspires each other and gives each other inspiration. For example, the first conjugate gradient method was developed by Hestenes and Stiefel to solve the system of linear equations [1], and then this method was generalized to solve the unconstrained optimization problem by Fletcher and Reeves [2].

Problems (1) and (2) appear frequently in many areas of applied mathematics and play important roles in many applications, such as compressive sensing, image processing, control theory, and motion control of robot manipulator [3–8]. For example, in the numerical solution theory of partial differential equations, the finite difference schemes of elliptic equations can be transformed into the following Sylvester equations:

$$AX + XB = C, \tag{3}$$

where $A \in \mathscr{R}^{p\times m}, B \in \mathscr{R}^{n\times q}$, and $C \in \mathscr{R}^{p\times q}$ are given matrices and $X \in \mathscr{R}^{m\times n}$ is the unknown matrix. Then, using the Kronecker product $\otimes$ and the vectorization operator $\mathrm{vec}(\cdot)$, we can transform the above Sylvester equations into a linear system of equations as follows [9]:

$$\left(I_n \otimes A + B^\top \otimes I_m\right)\mathrm{vec}(X) = \mathrm{vec}(C), \tag{4}$$

which is a special case of problem (1) with

$$\begin{aligned} x &:= \mathrm{vec}(X), \\ F(x) &:= \left(I_n \otimes A + B^\top \otimes I_m\right)\mathrm{vec}(X) - \mathrm{vec}(C), \\ \mathscr{X} &= \mathscr{R}^{mn}. \end{aligned} \tag{5}$$

Due to the numerous applications in diverse scientific areas, problems (1) and (2) have been extensively studied during the past few decades and many numerical methods have been proposed. The numerical methods for problem (1) can be roughly divided into two categories: the iterative methods for smooth case and the iterative methods for nonsmooth case. More specifically, the methods in the first category need to assume that the mapping $F(x)$ is smooth, which includes the Newton method, quasi-Newton method, Levenberg–Marquardt method, and their variants [10–13]. The methods in this category often need to solve a linear system of equations at each iteration, which indicates that they are not suitable to solve large-scale problem (1). The methods in the second category remove this restriction. For example, based on the spectral gradient method for unconstrained optimization problem, Cruz et al. [14, 15], Zhang and Zhou [16], and Liu and Duan [17] have successively proposed some spectral gradient projection methods or spectral residual methods for solving problem (1) with $\mathscr{X} = \mathscr{R}^n$. Motivated by the studies in [14–16], Cheng [18] extended the Polak–Ribiére–Polyak (PRP) method to solve problem (1) with $\mathscr{X} = \mathscr{R}^n$. Other similar methods include the two-term PRP-based method [19], the CG_DESCENT method [3], the Hestenes–Stiefel projection method [20], and the hybrid conjugate gradient projection method [21]. After careful analysis and comparison, we find that the above methods mainly consist of the following three steps at each iteration: (i) a sufficient descent direction is first generated, along which a step size is obtained by Armijo-like line search; (ii) a temporal iterate z_k is generated, and then a hyperplane

$$\mathscr{H}_k = \{x \in \mathscr{R}^n \mid \langle F(z_k), x - z_k\rangle = 0\}, \tag{6}$$

is defined, which strictly separates the current iterate x_k and the solution set $\mathscr{X}^*$ of problem (1); (iii) the next iterate x_{k+1} is defined by the projection of x_k onto the hyperplane $\mathscr{H}_k$.

On the other hand, the conjugate gradient method is one of the most efficient solvers for large-scale problem (2), whose iteration sequence $\{x_k\}$ is generated by

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, \ldots, \tag{7}$$

where $\alpha_k > 0$ is the step size and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{if } k \geq 1, \end{cases} \tag{8}$$

in which β_k is the so-called conjugate gradient formula which is the main difference in conjugate gradient methods. Since 1952, many conjugate gradient methods have been offered, such as the Hestenes–Stiefel (HS) method [1], the Fletcher–Reeves (FR) conjugate gradient method [2], the Polak–Ribiére–Polyak (PRP) conjugate gradient method [22], the Liu–Storey (LS) conjugate gradient method [23], and the Dai–Yuan (DY) conjugate gradient method [24]. During the last two decades, many conjugate gradient methods with sufficient descent property were proposed, and the first one is that proposed by Shi and Shen [25], which has not aroused continuing concern. Lately, the CG_DESCENT method designed by Hager and Zhang [26] is another one with sufficient descent property, which has inspired to benefit much research and design in this direction, and many efficient conjugate gradient methods have been developed, such as the modified FR in [27], the modified PRP in [28], and the descent memory gradient method in [29], in which the modified FR in [27] accomplished a theoretical breakthrough of great significance.

In this paper, based on the Fletcher–Reeves (FR) conjugate gradient method, we firstly propose a multiparameterized conjugate gradient method (MPCGM) for problem (1), which is derivative-free and thus only needs to compute the value of mapping $F(x)$ at each iteration. Then, the method is generalized to solve problem (2), and a novel conjugate gradient method (NCGM) is obtained. Both methods' convergence property is analyzed under traditional conditions and their practical application in compressive sensing and motion control of robot manipulator is investigated.

The remainder of this paper is organized as follows. In Section 2, we describe the MPCGM for nonsmooth problem (1). Moreover, the proof of its global convergence is also presented. In Section 3, we generalize MPCGM to solve nonconvex problem (2) and analyze the convergence property of the generalized method. In Section 4, some numerical results and comparisons are presented, and finally a brief conclusion is drawn in Section 5. Before ending this section, it is worth pointing out the main contributions of this paper as below.

(i) A multiparameterized conjugate gradient method is proposed for nonsmooth problem (1), which is used to solve compressive sensing.

(ii) A novel conjugate gradient method is proposed for nonconvex problem (2), which is used to solve motion control of robot manipulator.

(iii) Global convergence property of two novel methods is proved under mild conditions.

2. Multiparameterized Conjugate Gradient Method

Projection operator $P_\Omega[x]$ is defined as a mapping from the n dimensional Euclidean space $\mathscr{R}^n$ to a nonempty closed convex subset $\Omega \subseteq \mathscr{R}^n$:

$$P_\Omega[x] := \arg\min\{\|y-x\| \mid y \in \Omega\}, \quad \forall x \in \mathscr{R}^n, \tag{9}$$

which satisfies the following property [30].

Lemma 1. *Let Ω be a closed convex subset of $\mathscr{R}^n$. For any $x, y \in \mathscr{R}^n$, we have*

$$\left\|P_\Omega[x] - P_\Omega[y]\right\| \le \|x-y\|. \tag{10}$$

Assumption 1

(1) The solution set of problem (1), denoted by $\mathscr{X}^*$, is nonempty.

(2) The mapping $F(x)$ is monotone on $\mathscr{R}^n$, i.e.,

$$\langle x-y, F(x)-F(y)\rangle \ge 0, \quad \forall x, y \in \mathscr{R}^n. \tag{11}$$

(3) The mapping $F(x)$ is Lipschitz continuous on $\mathscr{X}$, i.e., there exists a constant $L>0$ such that

$$\|F(x)-F(y)\| \le L\|x-y\|, \quad \forall x, y \in \mathscr{X}. \tag{12}$$

Based on the research in [27, 29], we present a multiparameterized conjugate gradient method for nonsmooth problem (1) as follows.

Algorithm 1. *Multiparameterized conjugate gradient method (MPCGM).*

Step 0: choose constants $0<\rho<1$, $c>0$, $\sigma>0$, $\nu\ge 0$, $\beta>0$, $0<\gamma<2$, and tolerance error $\varepsilon>0$. Set an initial point $x_0 \in \mathscr{X}$, and let $k=0$.

Step 1: if $\|F(x_k)\|<\varepsilon$, then stop; otherwise, go to step 2.

Step 2: compute d_k by

$$d_k = \begin{cases} -F(x_k), & \text{if } k=0, \\ -\theta_k F(x_k) + \beta_k d_{k-1}, & \text{if } k \ge 1, \end{cases} \tag{13}$$

where θ_k and β_k are two parameters defined by

$$\begin{aligned} \theta_k &= c + \frac{F(x_k)^\top d_{k-1}}{\left\|d_{k-1}\right\|^2}, \quad \forall k \ge 1, \\ \beta_k &= \frac{\left\|F(x_k)\right\|^2}{\left\|d_{k-1}\right\|^2}, \quad \forall k \ge 1. \end{aligned} \tag{14}$$

Step 3: compute a temporal iterate $z_k = x_k + \alpha_k d_k$, where $\alpha_k = \beta\rho^{m_k}$ with m_k being the smallest nonnegative integer m such that

$$-\langle F(x_k+\beta\rho^m d_k), d_k\rangle \ge \sigma\beta\rho^m \left\|\nu F(x_k) + F(x_k+\beta\rho^m d_k)\right\| \left\|d_k\right\|^2. \tag{15}$$

Step 4: if $z_k \in \mathscr{X}$ and $\|F(z_k)\|<\varepsilon$, then stop; otherwise, compute the new iterate x_{k+1} by

$$x_{k+1} = P_{\mathscr{X}}\left[x_k - \gamma\xi_k(\nu F(x_k) + F(z_k))\right], \tag{16}$$

where

$$\xi_k = \frac{\langle F(z_k), x_k - z_k\rangle}{\left\|\nu F(x_k) + F(z_k)\right\|^2}. \tag{17}$$

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

Remark 1. Parameter β_k is obtained by replacing the denominator $\|F(x_{k-1})\|^2$ of β_k in the classical FR conjugate gradient method by $\|d_{k-1}\|^2$, and parameter θ_k makes the generated direction d_k satisfy sufficient descent property, which is proved in the next lemma.

Lemma 2. *For $c>0$ and any $k\ge 0$, the direction d_k defined by (13) satisfies*

$$F(x_k)^\top d_k \le -C\left\|F(x_k)\right\|^2, \tag{18}$$

where $C=\min\{1, c\}>0$.

Proof. If $k=0$, from (13), it holds that

$$F(x_0)^\top d_0 = -\left\|F(x_0)\right\|^2 \le -C\left\|F(x_k)\right\|^2. \tag{19}$$

If $k\ge 1$, from (13) again, we have

$$\begin{aligned} F(x_k)^\top d_k &= F(x_k)^\top(-\theta_k F(x_k) + \beta_k d_{k-1}) \\ &= -\left(c + \frac{F(x_k)^\top d_{k-1}}{\left\|d_{k-1}\right\|^2}\right)\left\|F(x_k)\right\|^2 + \frac{\left\|F(x_k)\right\|^2}{\left\|d_{k-1}\right\|^2} F(x_k)^\top d_{k-1} \\ &= -c\left\|F(x_k)\right\|^2 \le -C\left\|F(x_k)\right\|^2. \end{aligned} \tag{20}$$

Therefore, for all $k\ge 0$, inequality (18) always holds. This completes the proof. □

Remark 2. By Cauchy–Schwarz inequality, it holds that

$$\left\|d_k\right\| \ge C\left\|F(x_k)\right\|. \tag{21}$$

Remark 3. Parameter ξ_k in Step 4 of MPCGM is well defined, which is analyzed as follows.

(i) For $\nu=0$: if $\|F(z_k)\|=0$, from the line search (15), we have $\|d_k\|=0$, which together with (21) implies $\|F(x_k)\|=0$. This indicates that $\|\nu F(x_k)+F(z_k)=\|F(z_k)\|\| \ne 0$ if $\|F(x_k)\| \ne 0$.

(ii) For $\nu > 0$: if $\|\nu F(x_k) + F(z_k)\| = 0$, we have $F(z_k) = -\nu F(x_k)$. This together with the line search (15) gives $\langle \nu F(x_k), d_k\rangle \geq \sigma\alpha_k\|d_k\|^2$. From this inequality and (18), we have

$$-C\nu\|F(x_k)\|^2 \geq \sigma\alpha_k\|d_k\|^2, \tag{22}$$

i.e.,

$$\|F(x_k)\|^2 \leq -\frac{\sigma\alpha_k}{C\nu}\|d_k\|^2, \tag{23}$$

which indicates $\|F(x_k)\| = 0$. Therefore, we again get $\|\nu F(x_k) + F(z_k)\| \neq 0$ if $\|F(x_k)\| \neq 0$.

The following lemma indicates that the Armijo-type line search (15) is well defined.

Lemma 3. *For each $k \geq 0$, there exists a nonnegative integer m_k satisfying inequality (15).*

Proof. If the Armijo line search (15) is executed, then $\|F(x_k)\| \geq \varepsilon > 0$. Assume that there exists an integer $k_0 \geq 0$ such that inequality (15) does not hold for any nonnegative integer m, i.e.,

$$\begin{aligned}-\langle F(x_{k_0} + \beta_{k_0}\rho^m d_{k_0}), d_{k_0}\rangle &< \sigma\beta_{k_0}\rho^m\|\nu F(x_{k_0})\\ &\quad + F(x_{k_0} + \beta\rho^m d_{k_0})\|\|d_{k_0}\|^2, \quad \forall m \geq 1.\end{aligned} \tag{24}$$

Setting $m \longrightarrow +\infty$ and taking limits on both sides of the above inequality, we get

$$-\langle F(x_{k_0}), d_{k_0}\rangle \leq 0. \tag{25}$$

This together with inequality (18) gives $\|F(x_{k_0})\|^2 \leq 0$, i.e., $\|F(x_{k_0})\| = 0$ which contradicts $\|F(x_{k_0})\| \geq \varepsilon$. This completes the proof. □

Lemma 4. *Let $\{x_k\}$ be the sequence generated by MPCGM. Then, for any fixed $k \geq 0$, the step size α_k is bigger than a positive number, i.e., there exists $c_k > 0$, such that*

$$\alpha_k \geq c_k. \tag{26}$$

Furthermore, we can deduce that

$$c_k = \min\left\{\beta, \frac{\rho C\|F(x_k)\|^2}{\left(L + \sigma\|\nu F(x_k) + F(x_k + \alpha_k\rho^{-1}d_k)\|\right)\|d_k\|^2}\right\}. \tag{27}$$

Proof. If $\alpha_k \neq \beta$, then according to principle of the Armijo line search (15), the positive number $\alpha_k' = \alpha_k/\rho$ does not satisfy the following inequality:

$$-\langle F(x_k + \alpha_k' d_k), d_k\rangle < \sigma\alpha_k'\|\nu F(x_k) + F(x_k + \alpha_k' d_k)\|\|d_k\|^2. \tag{28}$$

So, by (15) and (18), we get

$$\begin{aligned}C\|F(x_k)\|^2 &\leq -F(x_k)^\top d_k\\ &= \langle F(x_k + \alpha_k' d_k) - F(x_k), d_k\rangle - \langle F(x_k + \alpha_k' d_k), d_k\rangle\\ &\leq L\alpha_k'\|d_k\|^2 + \sigma\alpha_k'\|\nu F(x_k) + F(x_k + \alpha_k' d_k)\|\|d_k\|^2\\ &= \left(L + \sigma\|\nu F(x_k) + F(x_k + \alpha_k\rho^{-1}d_k)\|\right)\alpha_k\rho^{-1}\|d_k\|^2,\end{aligned} \tag{29}$$

from which we get inequality (26). This completes the proof. □

Lemma 5. *Let $\{x_k\}$ and $\{z_k\}$ be two sequences generated by MPCGM. Then $\{x_k\}$ and $\{z_k\}$ are both bounded, and*

$$\lim_{k\longrightarrow\infty} \alpha_k\|d_k\| = 0. \tag{30}$$

Proof. From inequality (15), we have

$$\begin{aligned}\langle F(z_k), x_k - z_k\rangle &= -\alpha_k\langle F(z_k), d_k\rangle \geq \sigma\alpha_k^2\|\nu F(x_k)\\ &\quad + F(z_k)\|\|d_k\|^2 = \sigma\|\nu F(x_k) + F(z_k)\|\|x_k - z_k\|^2.\end{aligned} \tag{31}$$

Choose $x^* \in \mathcal{X}^*$; from the monotonicity of $F(x)$, we get

$$\langle F(z_k), z_k - x^*\rangle \geq \langle F(x^*), z_k - x^*\rangle = 0, \tag{32}$$

which together with (10) and (16) implies

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\leq \|x_k - \gamma\xi_k(\nu F(x_k) + F(z_k)) - x^*\|^2\\ &= \|x_k - x^*\|^2 - 2\gamma\xi_k\langle \nu F(x_k) + F(z_k), x_k - x^*\rangle + \gamma^2\xi_k^2\|\nu F(x_k) + F(z_k)\|^2\\ &\leq \|x_k - x^*\|^2 - 2\gamma\xi_k\langle F(z_k), x_k - x^*\rangle + \gamma^2\xi_k^2\|\nu F(x_k) + F(z_k)\|^2\\ &\leq \|x_k - x^*\|^2 - 2\gamma\xi_k\langle F(z_k), x_k - z_k\rangle + \gamma^2\xi_k^2\|\nu F(x_k) + F(z_k)\|^2\\ &= \|x_k - x^*\|^2 - \gamma(2-\gamma)\frac{\langle F(z_k), x_k - z_k\rangle^2}{\|\nu F(x_k) + F(z_k)\|^2}\\ &\leq \|x_k - x^*\|^2 - \sigma^2\gamma(2-\gamma)\|x_k - z_k\|^4,\end{aligned} \tag{33}$$

where the last inequality follows from (31). Therefore, the sequence $\{\|x_k - x^*\|\}$ is decreasing and convergent, and thus the sequence $\{x_k\}$ is bounded. From (33), we have

$$\sigma^2\gamma(2-\gamma)\sum_{k=0}^{\infty}\|x_k - z_k\|^4 \le \sum_{k=0}^{\infty}\left(\|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2\right) = \|x_0 - x^*\|^2 < \infty. \tag{34}$$

Then,

$$\lim_{k\longrightarrow\infty}\alpha_k\|d_k\| = \lim_{k\longrightarrow\infty}\|x_k - z_k\| = 0. \tag{35}$$

By the above inequality and the boundedness of the sequence $\{x_k\}$, it holds that the sequence $\{z_k\}$ is also bounded. The proof is completed.

Now, we are ready to establish the global convergence of MPCGM. □

Theorem 1. *Let $\{x_k\}$ be the sequence generated by MPCGM. Then, we have*

$$\lim_{k\longrightarrow\infty}\|F(x_k)\| = 0. \tag{36}$$

Proof. We prove (36) by using reduction to absurdity. Suppose that (36) is not true. Then, there is a constant $\varepsilon_0 > 0$ such that

$$\|F(x_k)\| \ge \varepsilon_0, \quad \forall k \ge 0. \tag{37}$$

By (21), we have

$$\|d_k\| \ge C\|F(x_k)\| \ge C\varepsilon_0, \quad \forall k \ge 0. \tag{38}$$

Combining this with (30), it holds that

$$\lim_{k\longrightarrow\infty}\alpha_k = 0. \tag{39}$$

On the other hand, by the boundedness of $\{x_k\}$, there exists a constant $M_1 > 0$ such that

$$\|F(x_k)\| \le M_1, \quad \forall k \ge 0. \tag{40}$$

Furthermore, by (39) and the continuity of $F(x)$, there exists $M_2 > 0$, such that

$$\left\|\nu F(x_k) + F\left(x_k + \alpha_k\rho^{-1}d_k\right)\right\| \le \nu\|F(x_k)\| + \left\|F\left(x_k + \alpha_k\rho^{-1}d_k\right)\right\| \le \nu M_1 + M_2, \quad \forall k \ge 0. \tag{41}$$

Then, by the definition of search direction d_k defined by (13), we have

$$\begin{aligned}\|d_k\| &\le c\|F(x_k)\| + \frac{|d_{k-1}^{\top}F(x_k)|}{\|d_{k-1}\|^2}\|F(x_k)\| + \frac{\|F(x_k)\|^2}{\|d_{k-1}\|^2}\|d_{k-1}\| \\ &\le \left(c + \frac{2M_1}{\|d_{k-1}\|}\right)\|F(x_k)\| \\ &\le \left(c + \frac{2M_1}{C\|F(x_{k-1})\|}\right)\|F(x_k)\| \\ &\le \left(c + \frac{2M_1}{C\varepsilon_0}\right)\|F(x_k)\|.\end{aligned} \tag{42}$$

This together with (26) implies that

$$\alpha_k \ge \min\left\{\beta, \frac{\rho C\varepsilon_0^2}{\left(L + \sigma(\nu M_1 + M_2)\right)(\varepsilon_0 + 2M_1)^2}\right\} > 0, \quad \forall k \ge 0, \tag{43}$$

which contradicts (39). Therefore, conclusion (36) holds and the proof is completed. □

3. Novel Conjugate Gradient Method

In this section, we will generalized MPCGM to solve problem (2) and prove its global convergence. Firstly, we make the following standard assumption.

Assumption 2

(1) The solution set of problem (2), denoted by $\mathcal{X}^*$, is nonempty.

(2) The level set $L_0 = \{x | f(x) \le f(x_0)\}$ is bounded, where $x_0 \in \mathcal{R}^n$ in an initial point.

(3) The gradient $g(x)$ is assumed to be Lipschitz continuous on $\mathcal{R}^n$, i.e., there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad \forall x, y \in \mathcal{R}^n. \tag{44}$$

Algorithm 1. *Novel conjugate gradient method (NCGM).*

Step 0: given an initial point $x_0 \in \mathcal{R}^n$, three constants $c > 0$, $0 < \rho, \gamma < 1$, and set $k = 0$.

Step 1: if $\|g_k\| = 0$, then stop; otherwise, go to step 2.

Step 2: compute d_k by

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -\theta_k g_k + \beta_k d_{k-1}, & \text{if } k \ge 1, \end{cases} \tag{45}$$

where θ_k and β_k are two parameters defined by

$$\theta_k = c + \frac{g_k^\top d_{k-1}}{\|d_{k-1}\|^2}, \forall k \geq 1,$$
$$\beta_k = \frac{\|g_k\|^2}{\|d_{k-1}\|^2}, \forall k \geq 1. \tag{46}$$

Determine the step size $\alpha_k = \rho^{m_k}$ with m_k being the smallest nonnegative integer m such that

$$f(x_k + \rho^m d_k) - f(x_k) \leq \gamma \rho^m g_k^\top d_k. \tag{47}$$

Step 3: set $x_{k+1} = x_k + \alpha_k d_k$ and $k = k + 1$; go to Step 1.

Similar to Lemma 2, it holds that

$$g_k^\top d_k \leq -C\|g_k\|^2, \tag{48}$$

where $C = \min\{1, c\}$. From inequality (48), it is easy to prove that the Armijo line search (47) is well defined. Moreover, from the Cauchy–Schwarz inequality and (48), it holds that

$$C\|g_k\| \leq \|d_k\|. \tag{49}$$

The next theorem indicates that NCGM is globally convergent.

Theorem 2. *If Assumption 2 holds and NCGM generates an infinite sequence $\{x_k\}$, we have*

$$\liminf_{k\longrightarrow\infty} \|g_k\| = 0. \tag{50}$$

Proof. First, we prove that there exists a constant $c_1 > 0$ such that the following inequality holds for all k:

$$\alpha_k \geq c_1 \frac{\|g_k\|^2}{\|d_k\|^2}. \tag{51}$$

The proof of (51) is divided into the following two cases.

Case (I): if $\alpha_k = 1$, then from (49), we have $\alpha_k = 1 \geq C^2 (\|g_k\|^2/\|d_k\|^2)$.

Case (II): if $\alpha_k < 1$, then by the Armijo line search condition, $\rho^{-1}\alpha_k$ does not satisfy inequality (47). That is,

$$f\left(x_k + \rho^{-1}\alpha_k d_k\right) - f(x_k) > \gamma\rho^{-1}\alpha_k g_k^\top d_k. \tag{52}$$

By the mean-value theorem of the continuous function, there exists a constant $t_k \in (0, 1)$ such that

$$\begin{aligned} & f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) \\ & = \rho^{-1}\alpha_k g(x_k + t_k\rho^{-1}\alpha_k d_k)^\top d_k \\ & = \rho^{-1}\alpha_k g_k^\top d_k + \rho^{-1}\alpha_k (g(x_k + t_k\rho^{-1}\alpha_k d_k) - g_k)^\top d_k \\ & \leq \rho^{-1}\alpha_k g_k^\top d_k + L\rho^{-2}\alpha_k^2 \|d_k\|^2. \end{aligned} \tag{53}$$

Substituting the last inequality into the left-hand side of (52), we get

$$\alpha_k \geq \frac{(1-\gamma)\rho C^2}{L} \frac{\|g_k\|^2}{\|d_k\|^2}. \tag{54}$$

Setting $c_1 = \min\{C^2, ((1-\delta)\rho C^2)/L\}$, we can get inequality (51). From (47), (48), and Assumption 2, it is easy to deduce that

$$\sum_{k=0}^{\infty} \alpha_k \|g_k\|^2 < \infty. \tag{55}$$

Substituting (51) into the left-hand side of (55), we can derive the famous Zoutendijk condition

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{56}$$

Suppose that conclusion (50) is not true, so there is a constant $\varepsilon_0 > 0$ such that

$$\|g_k\| \geq \varepsilon_0, \forall k \geq 0. \tag{57}$$

By the definition of d_k, we have

$$\begin{aligned} \|d_k\| & \leq c\|g_k\| + \frac{|g_k^\top d_{k-1}|}{\|d_{k-1}\|^2}\|g_k\| + \frac{\|g_k\|^2}{\|d_{k-1}\|^2}\|d_{k-1}\| \\ & \leq \left(c + \frac{2M_2}{\|d_{k-1}\|}\right)\|g_k\| \\ & \leq \left(c + \frac{2M_2}{C\|g_{k-1}\|}\right)\|g_k\| \\ & \leq \left(c + \frac{2M_2}{C\varepsilon_0}\right)\|g_k\|, \end{aligned} \tag{58}$$

where $M_2 > 0$ is the upper bound of $f(x)$ in the level set L_0. From this inequality and (56), we get

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \left(c + \frac{2M_2}{C\varepsilon_0}\right)^2 \frac{1}{\|g_k\|^2} \leq \left(c + \frac{2M_2}{C\varepsilon_0}\right)^2 \frac{1}{\varepsilon_0^2}. \tag{59}$$

Thus,

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \sum_{k=0}^{\infty} \varepsilon_0^2 \left(c + \frac{2M_2}{C\varepsilon_0}\right)^{-2} = \infty, \tag{60}$$

which contradicts (56). Therefore, conclusion (50) holds. The proof is completed. □

4. Numerical Results

In this section, to show the efficiency of MPCGM and NCGM, we apply them to solve problems (1) and (2). Furthermore, we compare the performance of MPCGM with the spectral gradient projection method in [31] (SGPM) and the conjugate gradient method in [3] (CGM). All codes were written in MATLAB R2014a, and run on a notebook computer with Intel Core 2 CPU 2.10 GHZ and RAM 2.00 GM.

4.1. Numerical Test of MPCGM. We consider two synthesized problems and one practical problem, which are drawn from [3, 32, 33].

Problem 1. Set $F(x) = (f_1(x), f_2(x), \ldots, f_n(x))^\top$, where

$$f_i(x) = e^{x_i} - 1, \quad \text{for } i = 1, 2, \ldots, n, \tag{61}$$

and $\mathcal{X} = R^n_+$. This problem has a unique solution $x^* = (0, 0, \ldots, 0)^\top$.

Problem 2. Set $F(x) = (f_1(x), f_2(x), \ldots, f_n(x))^\top$, where

$$f_i(x) = x_i - \sin\left(\left|x_i - 1\right|\right), \quad \text{for } i = 1, 2, \ldots, n, \tag{62}$$

and $\mathcal{X} = \{x \in \mathcal{R}^n \mid \sum_{i=1}^n x_i \leq n, x_i \geq 0, i = 1, 2, \ldots, n\}$. This problem is nonsmooth at the point $(1, 1, \ldots, 1)^\top$.

It is easy to prove that the above two mappings are monotone. The parameters in the three tested methods for Problem 1 and Problem 2 are set as follows:

SGPM: $r = 0.01$, $\sigma = 0.01$, $\beta = 0.5$.

CGM: $\rho = 0.39$, $\sigma = 10^{-4}$, $\beta = 1$.

MPCGM: $\rho = 0.2$, $c = 1$, $\sigma = 0.01$, $\beta = 1$, $\gamma = 1.7$.

In the experiment, we use the following termination condition:

$$\left\|F(x_k)\right\| \leq 10^{-6}. \tag{63}$$

In MPCGM, we have introduced two new parameters γ and ν. Now, we conduct some sensitivity tests on the two parameters to determine their optimal choices. Here, we use the tentative method and analyze the fluctuation of the number of iterations with respect to different values of γ and ν. Specifically, we set γ or ν as abscissa and we set the number of iterations as ordinate.

(i) We use Problem 1 with $x_0 = (1, 1, \ldots, 1)$ and $n = 10000$ to analyze the influence of γ on the number of iterations. Moreover, we set $\nu = 0$ and choose different values of $\gamma \in \{0.5, 0.6, \ldots, 1.9\}$.

(ii) We use Problem 2 with $x_0 = (1, 1, \ldots, 1)$ and $n = 10000$ to analyze the influence of ν on the number of iterations. Moreover, we set $\gamma = 1.7$ and choose different values of $\nu \in \{0, 0.01, \ldots, 0.1\}$.

The numerical results are graphically shown in Figure 1, from which we can see that for Problem 1, larger values of γ can accelerate the convergence of MPCGM, and for Problem 2, the positive values of ν can also accelerate the convergence of MPCGM. Therefore, the advantage of incorporating the parameters γ and ν into MPCGM is verified. In the following, we set $\gamma = 1.7$ and $\nu = 0.07$.

Now, we give more numerical results about Problem 1 and Problem 2 with the number of variables $n = 1000$, 2000, 5000, 10000, 20000, 50000, 100000, 1000000, and the initial point is set as $x_0 = (1, 1, \ldots, 1)$. The numerical results are reported in Tables 1 and 2, which contain the dimension of the problem (Dim), the number of iterations (Iter), the CPU time required in seconds (Time), and the final norm of equations (Fn) when the termination condition is satisfied. It is well known that when a set is a polyhedral, that is, all the constraint functions defining the set are linear, then computing the projection on it reduces to solving a quadratic problem. Here, we use the quadratic program solver quadprog.m from the MATLAB optimization toolbox to perform the projection operator.

The numerical results in Tables 1 and 2 verify that the gradient methods perform well on the large-scale constrained monotone equations. For Problem 1, the performance of CGM and MPCGM is obviously better than that of SGPM, and the performance of MPCGM is obviously better than that of CGM. That is, MPCGM performs the best among the three tested methods. As the dimension increased, the advantage on the required CPU time of MPCGM becomes prominent gradually. For Problem 2, there seems to be not much difference among the performance of the three tested methods, and MPCGM still performs a little better than the other two methods because it is the fastest for most scenarios. In a word, the numerical experiments show that the proposed method provides an efficient tool to solve nonlinear constrained equations.

Problem 3. Consider the compressive sensing (CS):

$$\min_{x \in \mathcal{R}^n} \frac{1}{2}\|Ax - b\|_2^2 + \mu\|x\|_1, \tag{64}$$

where $A \in \mathcal{R}^{m \times n}$ $(m \ll n)$ is a linear operator, $b \in \mathcal{R}^m$ is an observation, is the unknown vector, $\|x\|_1 = \sum_{i=1}^n |x_i|$ is the ℓ_1-norm of x, and parameter $\mu > 0$ is used to trade off both terms of the objective function of (64). Following the procedure of Figueiredo et al. [33], we can set $x = u - \nu$, $u \geq 0$, $\nu \geq 0$, where $u \in \mathcal{R}^n$, $\nu \in \mathcal{R}^n$, and $u_i = (x_i)_+$, $\nu_i = (-x_i)_+$ for all $i = 1, 2, \ldots, n$ with $(\cdot)_+ = \max\{0, \cdot\}$. Then, CS can be rewritten as

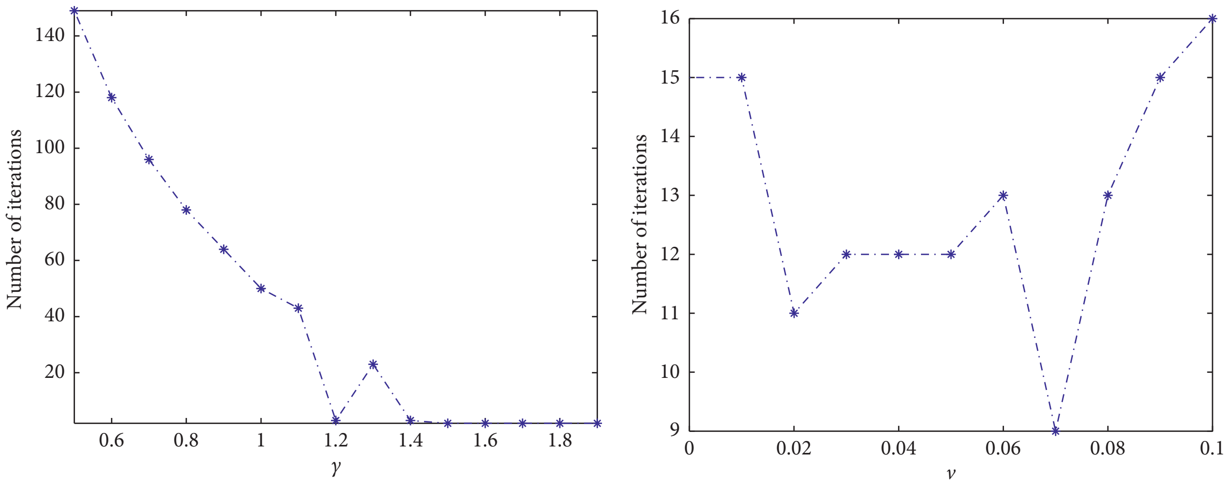
$$\begin{aligned} \min_{u,\nu} \quad & \frac{1}{2}\|b - A(u - \nu)\|_2^2 + \mu e_n^\top u + \mu e_n^\top \nu \\ \text{s.t.} \quad & u \geq 0, \nu \geq 0. \end{aligned} \tag{65}$$

That is,

$$\begin{aligned} \min_{u,\nu} \quad & \frac{1}{2} z^\top H z + c^\top z \\ \text{s.t.} \quad & z \geq 0, \end{aligned} \tag{66}$$

where

$$\begin{aligned} z &= \begin{bmatrix} u \\ \nu \end{bmatrix}, \\ y &= A^\top b, \\ c &= \mu e_{2n} + \begin{bmatrix} -y \\ y \end{bmatrix}, \\ H &= \begin{bmatrix} A^\top A & -A^\top A \\ -A^\top A & A^\top A \end{bmatrix}. \end{aligned} \tag{67}$$



(a) (b)

FIGURE 1: Sensitivity test on γ and ν. (a) Sensitivity test on the parameter γ. (b) Sensitivity test on the parameter ν.

TABLE 1: Numerical results of Problem 1.

Dim	SGPM			CGM			MPCGM		
	Iter	Time	Fn	Iter	Time	Fn	Iter	Time	Fn
1000	13	0.02	$5.62314e-07$	13	0.01	$4.76573e-07$	1	0.01	$0.00000e+00$
2000	13	0.03	$7.95231e-07$	13	0.01	$6.73977e-07$	1	0.01	$0.00000e+00$
5000	14	0.07	$8.85789e-08$	14	0.02	$4.36958e-07$	1	0.01	$0.00000e+00$
10000	14	0.13	$1.25269e-07$	14	0.04	$6.17952e-07$	2	0.02	$0.00000e+00$
20000	14	0.26	$1.77161e-07$	14	0.08	$8.73916e-07$	2	0.02	$0.00000e+00$
50000	14	0.68	$2.80126e-07$	15	0.23	$3.56664e-07$	2	0.03	$0.00000e+00$
100000	14	1.35	$3.96172e-07$	15	0.44	$5.04400e-07$	4	0.11	$0.00000e+00$
1000000	15	17.70	$2.30082e-07$	16	5.97	$8.61927e-07$	12	3.92	$0.00000e+00$

TABLE 2: Numerical results of Problem 2.

Dim	SGPM			CGM			MPCGM		
	Iter	Time	Fn	Iter	Time	Fn	Iter	Time	Fn
1000	11	0.01	$1.53710e-08$	13	0.01	$2.36758e-07$	8	0.01	$4.29514e-07$
2000	11	0.01	$2.17380e-08$	13	0.01	$3.34826e-07$	8	0.01	$6.07424e-07$
5000	11	0.02	$3.43707e-08$	13	0.01	$5.29406e-07$	8	0.01	$9.60422e-07$
10000	11	0.03	$4.86070e-08$	13	0.02	$7.48694e-07$	9	0.01	$7.05926e-07$
20000	11	0.05	$6.87439e-08$	14	0.04	$4.65956e-07$	9	0.02	$9.98330e-07$
50000	11	0.11	$1.08699e-07$	14	0.09	$7.36741e-07$	14	0.07	$8.57326e-08$
100000	11	0.24	$1.53719e-07$	15	0.19	$5.16018e-07$	14	0.12	$1.21244e-07$
1000000	11	3.23	$4.85879e-07$	16	2.36	$8.99942e-07$	21	2.35	$6.98408e-07$

Then, Xiao et al. [3] further transformed the above optimization problem as the constrained nonlinear equations:

$$F(z) = \min\{z, Hz + c\} = 0, \quad z \geq 0. \tag{68}$$

The following relative error (RelErr) to the original signal $\tilde{x}$ is used to measure the quality of restoration:

$$\text{RelErr} = \frac{\|\tilde{x} - x^*\|_2}{\|\tilde{x}\|_2}, \tag{69}$$

where x^* is the restored signal. In the experiment, our goal is to reconstruct a length-n sparse signal from m observation. Here, we set $n = 2048$ and $m = 512$, and the original signal contains 64 randomly placed spikes. The $m \times n$ matrix A is obtained by first filling it with independent samples of a standard Gaussian distribution and then orthonormalizing the rows. The observation b is generated by $b = A\tilde{x} + \omega$, where ω is the Gaussian noise distributed as $N(0, \delta^2 I)$ with $\delta = 10^{-3}$. We set $\mu = 0.01\|A^\top b\|_\infty$ and use $f(x) = \mu\|x\|_1 + \|Ax - b\|_2^2/2$ as the merit function and stop the tested methods if $\|f_k\| \leq 10^{-2}$. The parameters in the three tested methods for Problem 3 are listed as follows:

SGPM: $r = 10$, $\sigma = 0.01$, $\beta = 0.3$.

CGM: $\rho = 0.39$, $\sigma = 10^{-4}$, $\beta = 1$.

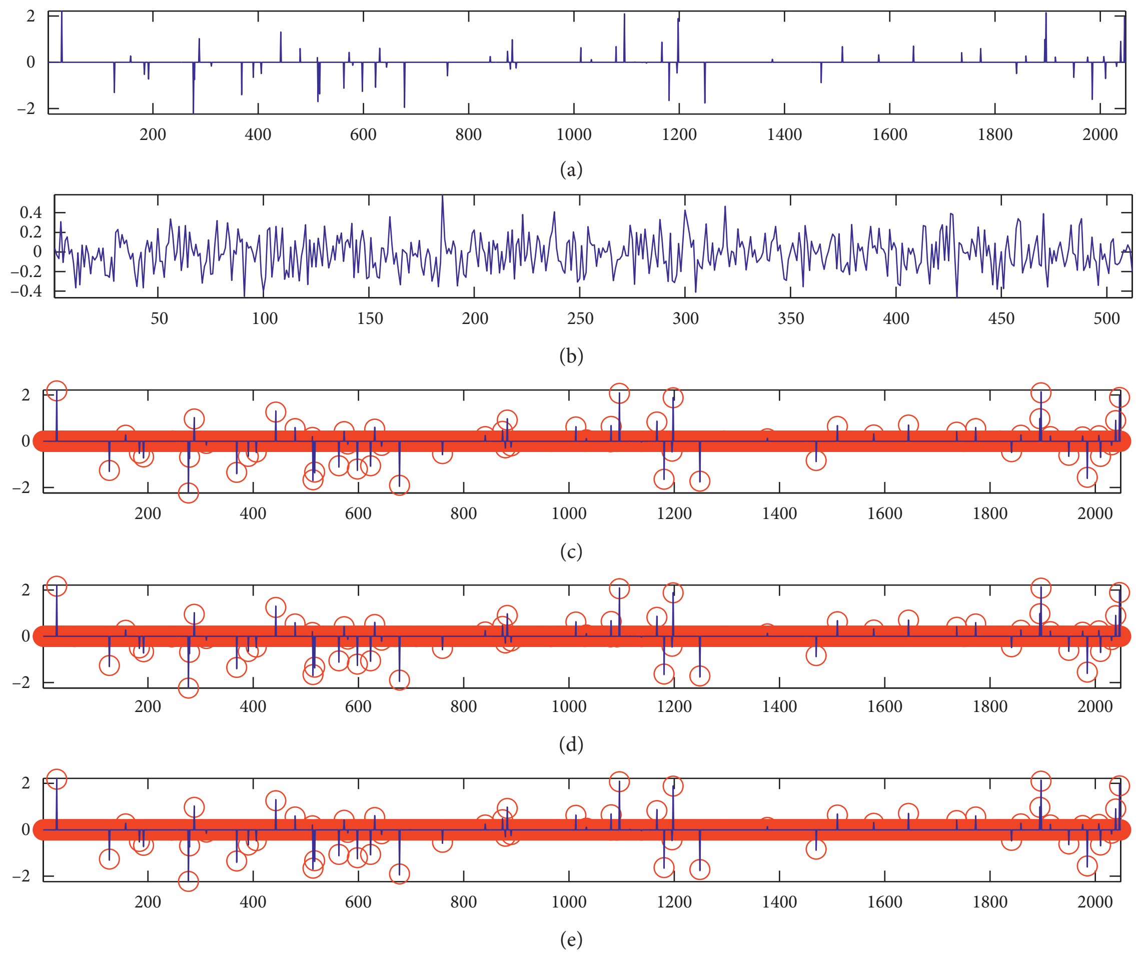


Figure 2: The original signal, noisy measurement, and recovered results. (a) Original signal. (b) Noisy measurement. (c) SGPM (RelErr = 3.91%, Iter = 347, Time = 3.04 s). (d) CGM (RelErr = 4.40%, Iter = 287, Time = 3.71 s). (e) MPCGM (RelErr = 3.81%, Iter = 156, Time = 1.48 s).

MPCGM: $\rho = 0.4$, $c = 1$, $\sigma = 0$, $\beta = 1$, $\gamma = 1.9$.

The numerical results generated by the three tested methods are given in Figure 2.

From Figures 2(c)–2(e), we can see that the three tested methods recover the original signal with high precision, and MPCGM still performs the best among the three methods because it takes the least number of iterations and CPU time.

4.2. Numerical Test of NCGM. In this section, the motion control of a two-joint planar robotic manipulator is solved by NCGM. As stated in [34], the discrete-time kinematics equation of two-joint planar robot manipulator at the position level is given as

$$f(\theta_k) = r_k, \tag{70}$$

where $f(\cdot)$ is the kinematics mapping function with known structure and defined as

$$f(\theta) = \begin{bmatrix} l_1 c_1 + l_2 c_2 \\ l_1 s_1 + l_2 s_2 \end{bmatrix}, \tag{71}$$

in which l_i is the length of the i-th rod, $c_1 = \cos(\theta_1)$, $s_1 = \sin(\theta_1)$, $c_2 = \cos(\theta_1 + \theta_2)$, and $s_2 = \sin(\theta_1 + \theta_2)$. Besides, $\theta_k \in \mathscr{R}^2$ is the joint angle vector and $r_k \in \mathscr{R}^2$ is the end effector position vector. Then, we need to solve a series of optimization problem defined at each time instant $t_k \in [0, t_f]$ as follows:

$$\min_{r_k \in \mathscr{R}^2} \frac{1}{2} \left\| r_k - r_{dk} \right\|^2. \tag{72}$$

In this experiment, we set $l_i = 1 (i = 1, 2)$ and the end effector is controlled to track a Lissajous curve, which is expressed as [34]

$$r_{dk} = \begin{bmatrix} 1.5 + 0.2 \sin \dfrac{\pi t_k}{5} \\ \sqrt{3}/2 + 0.2 \sin\left(\dfrac{2\pi t_k}{5} + \left(\dfrac{\pi}{3}\right)\right) \end{bmatrix}. \tag{73}$$

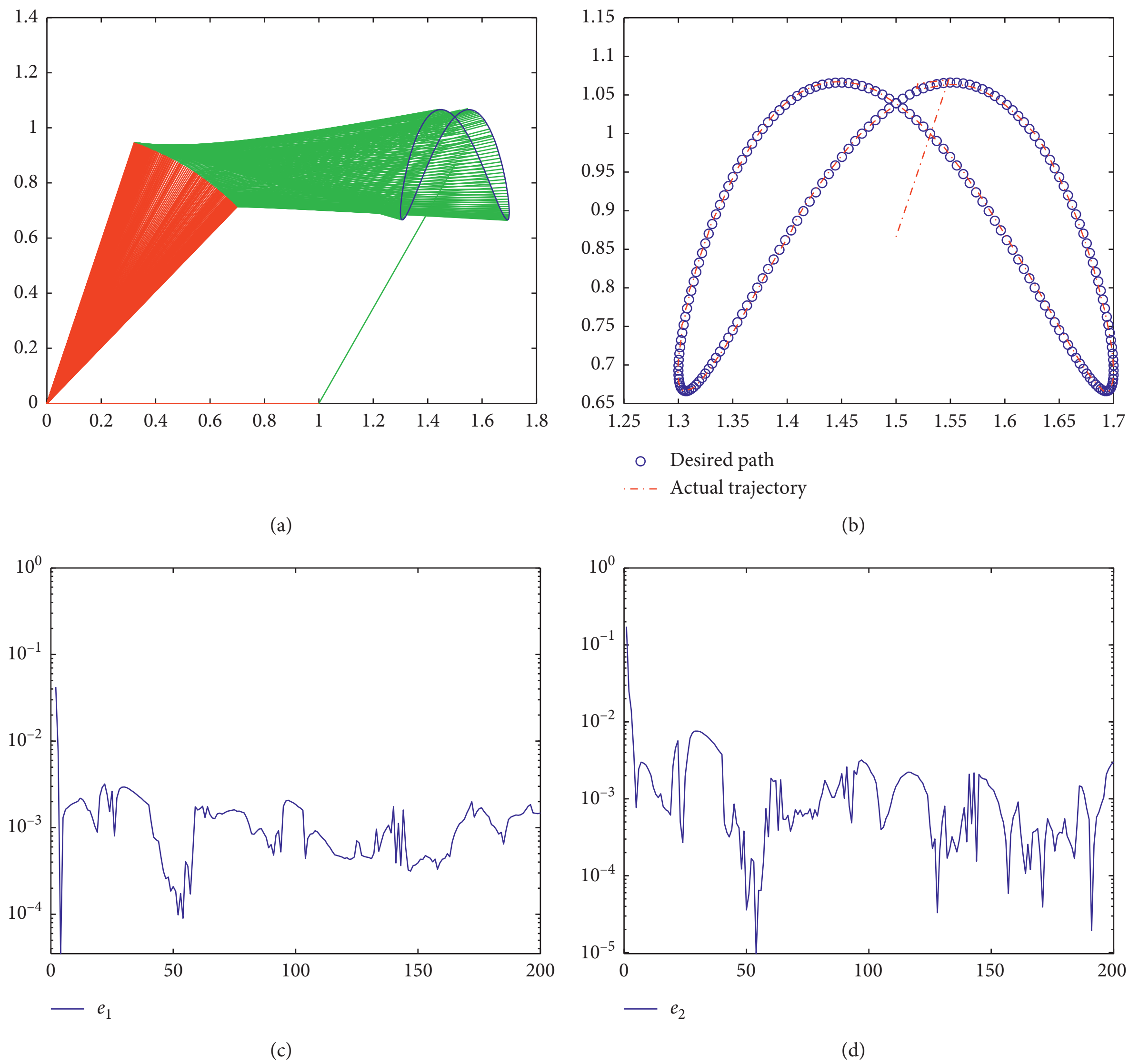


Figure 3: Numerical results generated by NCGM. (a) Manipulator trajectories. (b) End effector trajectory and desired path. (c) Tracking errors on the horizontal x-axis. (d) Tracking errors on the vertical y-axis.

For NCGM, we set $c = 0.01$, $\rho = 0.2$, and $\gamma = 0.08$. The initial point is set as $\theta_0 = [0, \pi/3]^\top$, the length of rod $l_i = 1 (i = 1, 2)$, the end of task duration $t_f = 10$s, and the task duration [0, 10] is divided into 200 equal parts. The numerical results generated by NCGM are plotted in Figure 3. Specifically, Figure 3(a) shows robot trajectories synthesized by NCGM. Figure 3(b) plots end effector trajectory and desired path. Figures 3(c) and 3(d) show the error of NCGM on x-axis and y-axis, respectively. From Figures 3(a) and 3(b), it is clear that NCGM successfully completes the given task. Furthermore, Figures 3(c) and 3(d) indicate that the generated error is about 10^{-3}.

5. Conclusion

In this paper, we have proposed a multiparameterized conjugate gradient method for nonlinear equations with convex constraints. Under the condition that the underlying mapping is monotone and Lipschitz continuous, we have established its global convergence. Furthermore, we have generalized this method to solve unconstrained optimization and get a new conjugate gradient method, whose global convergence is analyzed under mild conditions. Preliminary numerical results are reported which indicate that the proposed methods perform better than some well-developed methods.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China and Shandong Province (nos. 11671228, 11601475, and ZR2016AL05) and the PhD research startup foundation of Zaozhuang University.

References

[1] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.

[2] R. Fletcher and C. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.

[3] Y. Xiao and H. Zhu, "A conjugate gradient method to solve convex constrained monotone equations with applications in compressive sensing," *Journal of Mathematical Analysis and Applications*, vol. 405, no. 1, pp. 310–319, 2013.

[4] M. Sun and J. Liu, "The convergence rate of the proximal alternating direction method of multipliers with indefinite proximal regularization," *Journal of Inequalities and Applications*, vol. 2017, no. 1, p. 19, 2017.

[5] B. Zhou, G.-R. Duan, and Z. Lin, "A parametric periodic Lyapunov equation with application in semi-global stabilization of discrete-time periodic systems subject to actuator saturation," *Automatica*, vol. 47, no. 2, pp. 316–325, 2011.

[6] P. Bing, J. Sui, and M. Sun, "A prediction-correction primal-dual hybrid gradient method for convex programming with linear constraints," *ScienceAsia*, vol. 44, no. 1, pp. 34–39, 2018.

[7] M. Sun, M. Y. Tian, and Y. J. Wang, "Discrete-time Zhang neural networks for time-varying nonlinear optimization," *Discrete Dynamics in Nature and Society*, vol. 2019, Article ID 4745759, 14 pages, 2019.

[8] M. Sun and Y. Wang, "General five-step discrete-time Zhang neural network for time-varying nonlinear optimization," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 43, no. 2, pp. 1741–1760, 2020.

[9] M. Sun and J. Liu, "Iterative algorithms for symmetric positive semidefinite solutions of the Lyapunov matrix equations," *Mathematical Problems in Engineering*, vol. 2020, Article ID 6968402, 10 pages, 2020.

[10] J. E. Dennis and J. J. Moré, "A characterization of superlinear convergence and its application to quasi-Newton methods," *Mathematics of Computation*, vol. 28, no. 126, p. 549, 1974.

[11] D. Li and M. Fukushima, "A globally and superlinearly convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations," *SIAM Journal on Numerical Analysis*, vol. 37, no. 1, pp. 152–172, 1999.

[12] G. Zhou and K. C. Toh, "Superlinear convergence of a Newton-type algorithm for monotone equations," *Journal of Optimization Theory and Applications*, vol. 125, no. 1, pp. 205–221, 2005.

[13] W. J. Zhou and D. H. Li, "A globally convergent BFGS method for nonlinear monotone equations without any merit functions," *Mathematics of Computation*, vol. 77, no. 264, pp. 2231–2240, 2008.

[14] W. L. Cruz and M. Raydan, "Nonmonotone spectral methods for large-scale nonlinear systems," *Optimization Methods and Software*, vol. 18, no. 5, pp. 583–599, 2003.

[15] W. La Cruz, J. M. Martínez, and M. Raydan, "Spectral residual method without gradient information for solving large-scale nonlinear systems of equations," *Mathematics of Computation*, vol. 75, no. 255, pp. 1429–1449, 2006.

[16] L. Zhang and W. Zhou, "Spectral gradient projection method for solving nonlinear monotone equations," *Journal of Computational and Applied Mathematics*, vol. 196, no. 2, pp. 478–484, 2006.

[17] J. Liu and Y. R. Duan, "Two spectral gradient projection methods for constrained equations and their linear convergence rate," *Journal of Inequalities and Applications*, vol. 2015, no. 1, p. 8, 2015.

[18] W. Cheng, "A PRP type method for systems of monotone equations," *Mathematical and Computer Modelling*, vol. 50, no. 1-2, pp. 15–20, 2009.

[19] Q. Li and D.-H. Li, "A class of derivative-free methods for large-scale nonlinear monotone equations," *IMA Journal of Numerical Analysis*, vol. 31, no. 4, pp. 1625–1635, 2011.

[20] M. Sun and J. Liu, "A modified Hestenes-Stiefel projection method for constrained nonlinear equations and its linear convergence rate," *Journal of Applied Mathematics and Computating*, vol. 49, no. 1-2, pp. 145–156, 2015.

[21] M. Sun and J. Liu, "New hybrid conjugate gradient projection method for the convex constrained equations," *Calcolo*, vol. 53, no. 3, pp. 399–411, 2016.

[22] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées," *Revue française d'informatique et de recherche opérationnelle. Série rouge*, vol. 3, no. 16, pp. 35–43, 1969.

[23] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: theory," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991.

[24] Y. H. Dai and Y. X. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 2000.

[25] Z.-J. Shi and J. Shen, "A gradient-related algorithm with inexact line searches," *Journal of Computational and Applied Mathematics*, vol. 170, no. 2, pp. 349–370, 2004.

[26] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170–192, 2005.

[27] L. Zhang, W. Zhou, and D. Li, "Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search," *Numerische Mathematik*, vol. 104, no. 4, pp. 561–572, 2006.

[28] L. Zhang, W. Zhou, and D.-H. Li, "A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence," *IMA Journal of Numerical Analysis*, vol. 26, no. 4, pp. 629–640, 2006.

[29] M. Sun and Q. Bai, "A new descent memory gradient method and its global convergence," *Journal of Systems Science and Complexity*, vol. 24, no. 4, pp. 784–794, 2011.

[30] E. H. Zarantonello, *Projections on Convex Sets in Hilbert Space and Spectral Theory*, Academic Press, New York, NY, USA, 1971.

[31] Z. Yu, J. Lin, J. Sun, Y. Xiao, L. Liu, and Z. Li, "Spectral gradient projection method for monotone nonlinear equations with convex constraints," *Applied Numerical Mathematics*, vol. 59, no. 10, pp. 2416–2423, 2009.

[32] C. Wang, Y. Wang, and C. Xu, "A projection method for a system of nonlinear monotone equations with convex constraints," *Mathematical Methods of Operations Research*, vol. 66, no. 1, pp. 33–46, 2007.

[33] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[34] Y. Zhang, L. He, C. Hu, J. Guo, J. Li, and Y. Shi, "General four-step discrete-time zeroing and derivative dynamics applied to time-varying nonlinear optimization," *Journal of Computational and Applied Mathematics*, vol. 347, pp. 314–329, 2019.