

## Research Article

# The Smoothing FR Conjugate Gradient Method for Solving a Kind of Nonsmooth Optimization Problem with $l_1$ -Norm

Miao Chen and Shou-qiang Du 

School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China

Correspondence should be addressed to Shou-qiang Du; sqdu@qdu.edu.cn

Received 9 October 2017; Accepted 27 December 2017; Published 23 January 2018

Academic Editor: Elisa Francomano

Copyright © 2018 Miao Chen and Shou-qiang Du. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We study the method for solving a kind of nonsmooth optimization problems with  $l_1$ -norm, which is widely used in the problem of compressed sensing, image processing, and some related optimization problems with wide application background in engineering technology. Transformed by the absolute value equations, this kind of nonsmooth optimization problem is rewritten as a general unconstrained optimization problem, and the transformed problem is solved by a smoothing FR conjugate gradient method. Finally, the numerical experiments show the effectiveness of the given smoothing FR conjugate gradient method.

## 1. Introduction

In the last few years, the study work of finding sparsest solutions to undermined system of equations has been extensively done. Finding the sparsest solution of an undermined system of equations is equivalent to solving the  $l_0$ -norm regularized minimization problem as the following:

$$\begin{aligned} \min \quad & \|x\|_0, \\ \text{s.t.} \quad & x \in D, \end{aligned} \quad (1)$$

where  $D = \{x \mid Ax = b, x \in R^n\}$ ,  $A \in R^{m \times n}$  ( $m \ll n$ ),  $b \in R^m$ , and  $\|\cdot\|_0$  denotes 0-norm. From [1–4], we know that the above problem is difficult to solve in a straight way. In order to solve the  $l_0$ -norm problem effectively, an approximation model is to replace the  $l_0$ -norm by the  $l_1$ -norm to solve the Basis Pursuit problem, such as in [5, 6]:

$$\min_{x \in R^n} \{ \|x\|_1 : Ax = b \}. \quad (2)$$

Therefore the convex envelope of  $\|x\|_0$  is  $\|x\|_1$ , where  $\|x\|_1 = \sum_{i=1}^n |x_i|$  is the  $l_1$ -norm of  $x$ . And this problem is also NP-hard problem. When  $b$  contains some noise in practice application,

the above problem is rewritten as the following nonsmooth optimization problem with  $l_1$ -norm:

$$\min_{x \in R^n} \frac{1}{2} \|Ax - b\|_2^2 + \tau \|x\|_1, \quad (3)$$

where  $A \in R^{m \times n}$  ( $m \ll n$ ),  $b \in R^m$ ,  $\tau > 0$ ,  $\|\cdot\|_2$  denotes 2-norm, and  $\|\cdot\|_1$  denotes 1-norm. For the minimization of the  $l_1$ -norm having a good recovery, (3) is widely used in the sense compression, image processing, and other related fields in engineering technology, one can see [7, 8] and the references therein. For some  $\tau > 0$ ,  $(1/2)\|Ax - b\|_2^2 + \tau\|x\|_1$  is a convex function but not a differentiable function. Recently, many scholars have studied the method for solving (3). For instance, gradient projection for sparse reconstruction was proposed by Figueiredo et al. in [9], a two-step iterative shrinkage thresholding (IST) method was proposed by Bioucas-Dias and Figueiredo in [10], a fast IST algorithm was presented by Beck and Teboulle in [11], SPGL1 (A solver for large-scale sparse reconstruction) was proposed by Ewout and Friedlander in [12], who consider a least-squares problem with  $l_1$ -norm constraint and use a spectral gradient projection method, and ADM method was proposed by Yang and Zhang in [13]. Problem (3) was formulated to a convex quadratic problem in [14]. Among all the references mentioned above, there is no use of the relationship between

the linear complementarity problems and the absolute value equations to solve (3). They do not use the structure of the absolute value equation to propose new method to solve (3). They also do not translate the original problem into an absolute value equation problem and use the effective methods of absolute value equations. Just recently, only in [15], a smoothing gradient method is given for solving (3) based on the absolute value equations. Therefore, in this paper, we study how to use this new transformation to solve (3) by the useful FR conjugate method.

As we all know, the transformed linear complementarity problem is rewritten as the absolute value equation problem mainly based on the equivalence between the linear complementarity problem and the absolute value equation problem, such as [16–18]. The absolute value equation can be solved easily now. On the other hand, the conjugate gradient method is suitable for solving large-scale optimization problems and has simple structure and global convergence [19–25]. In addition, the smoothing methods are used to solve the related nonsmooth optimization problems, such as [26–28] and the references therein. Therefore, based on the above analysis, we present a new smoothing FR conjugate gradient method to solve (3); this is also our motivation to write this paper. The global convergence analysis of the given method is also presented. Finally, some computational results show that the smoothing FR conjugate gradient method is efficient in practice.

The remainder of this paper is organized as follows: In Section 2, the preliminaries are proposed, which include the description of how the linear complementarity problem is transformed into the absolute value equation problem. In Section 3, we present the smoothing FR conjugate gradient method and give its convergence analysis. Finally, in Section 4, we give some numerical results of the given method, which show the effectiveness of the given method.

## 2. Preliminaries

Firstly, we give the transformation form of (3). For any vector  $x \in R^n$ , it can be formulated to

$$x = u - v, \quad u \geq 0, \quad v \geq 0, \quad (4)$$

where  $u, v \in R^n$ ,  $u_i = (x_i)_+$ , and  $v_i = (-x_i)_+$  for all  $i = 1, \dots, n$ . By the definition of  $\|x\|_1 = \sum_{i=1}^n |x_i|$ , we get  $\|x\|_1 = e_n^T u + e_n^T v$ , where  $e_n = [1, 1, \dots, 1]^T$ . Therefore, as in [13–15], problem (3) can be rewritten as

$$\min_{z=(u,v)^T \geq 0} \frac{1}{2} \|b - Az\|_2^2 + \tau e_n^T u + \tau e_n^T v. \quad (5)$$

The above problem can be transformed to

$$\begin{aligned} \min \quad & \frac{1}{2} z^T H z + c^T z, \\ \text{s.t.} \quad & z \geq 0, \end{aligned} \quad (6)$$

where

$$\begin{aligned} z &= (u, v)^T, \\ c &= \tau e_{2n} + \begin{pmatrix} -c_t \\ c_t \end{pmatrix}, \\ c_t &= A^T b, \\ H &= \begin{pmatrix} A^T A & -A^T A \\ -A^T A & A^T A \end{pmatrix}. \end{aligned} \quad (7)$$

Since  $H$  is a positive semidefinite matrix, problem (3) can be transformed into convex optimization problem. Then problem (6) can be transformed into a linear variable inequality problem, which is to find  $z \in R^{2n}$ , such that

$$\langle Hz + c, \tilde{z} - z \rangle \geq 0, \quad \forall \tilde{z} \geq 0. \quad (8)$$

Given that the feasible region of  $z$  is a special structure (such as nonnegative orthant), (8) can be rewritten as the linear complementary problem, which is used to find  $z \in R^{2n}$ , such that

$$\begin{aligned} z &\geq 0, \\ Hz + c &\geq 0, \\ z^T (Hz + c) &= 0. \end{aligned} \quad (9)$$

Now, we give some results about the absolute value equations and the linear complementarity problems as follows; one can see sources such as [16, 29, 30]. The absolute value equations have the form  $Ax - |x| = b$ , where  $A \in R^{n \times n}$ ,  $b \in R^n$ . The linear complementarity problems have the form  $0 \leq z \perp Hz + c \geq 0$ , where  $H \in R^{n \times n}$ ,  $c \in R^n$ .

**Proposition 1.** (i) Conversely, if 1 is not an eigenvalue of  $H$ , then the linear complementary problem can be reduced to the following equation:

$$\begin{aligned} (H - I)^{-1} (H + I) \left( \frac{1}{2} ((H - I)z + c) \right) \\ - \left| \frac{1}{2} ((H - I)z + c) \right| = (H - I)^{-1} c. \end{aligned} \quad (10)$$

(ii) The absolute value equations is equivalent to the bimultilinear program:

$$\begin{aligned} \min \quad & ((A + I)x - b)^T ((A - I)x - b) = 0, \\ & ((A + I)x - b) \geq 0, \\ & ((A - I)x - b) \geq 0. \end{aligned} \quad (11)$$

(iii) And the absolute value equations is equivalent to the generalized linear complementary problem:

$$\begin{aligned} ((A + I)x - b) \perp ((A - I)x - b), \\ ((A + I)x - b) \geq 0, \\ ((A - I)x - b) \geq 0. \end{aligned} \quad (12)$$

**Proposition 2.** Equation (9) can be transformed into the following absolute value equation problem, which is defined as

$$(H + I)z + c = |(H - I)z + c|. \quad (13)$$

*Proof.* Based on Proposition 1 and (9), we know that

$$\begin{aligned} (A + I)x - b &= Hz + c, \\ (A - I)x - b &= z. \end{aligned} \quad (14)$$

Then by (14), we have

$$\begin{aligned} x &= (A - I)^{-1}(z + b), \\ Hz + c &= H(A - I)x - Hb + c = (A + I)x - b. \end{aligned} \quad (15)$$

To satisfy the last equation of (15) for all  $x \in R^n$ , denote

$$\begin{aligned} b &= (H - I)^{-1}c, \\ A &= (H - I)^{-1}(H + I). \end{aligned} \quad (16)$$

Substituting (16) in (15), we get

$$\begin{aligned} x &= (A - I)^{-1}(z + b) \\ &= ((H - I)^{-1}(H + I) - I)^{-1}(z + (H - I)^{-1}c) \\ &= ((H + I) - (H - I))^{-1}((H - I)z + c) \\ &= \frac{1}{2}((H - I)z + c). \end{aligned} \quad (17)$$

Due to the above (i) in Proposition 1, (9) can be reduced to the absolute value equations:

$$(H - I)^{-1}(H + I)x - |x| = (H - I)^{-1}c, \quad (18)$$

and substituting form (17) in (18), we get the following absolute value equation problem, which has the form

$$(H + I)z + c = |(H - I)z + c|. \quad (19)$$

Thus, we get (13). Then, problem (3) can be transformed into the following problem:

$$\min_{z \in R^{2n}} f(z) = \frac{1}{2} \|(H + I)z + c - |(H - I)z + c|\|_2^2. \quad (20) \quad \square$$

### 3. The Smoothing FR Conjugate Gradient Method

In this section, we give the smoothing FR conjugate gradient method for solving (20). Firstly, we give the definition of smoothing function and the smoothing approximation function of the absolute value function; one can see [15, 26, 27].

*Definition 3.* Let  $f : R^n \rightarrow R$  be a local Lipschitz continuous function. We call  $\tilde{f} : R^n \times R_+ \rightarrow R$  a smoothing function of

$f$ , if  $f(\cdot, \mu)$  is continuously differentiable in  $R^n$  for any fixed  $\mu > 0$ , and

$$\lim_{\mu \rightarrow 0} \tilde{f}(x, \mu) = f(x). \quad (21)$$

There are so many smoothing functions; for example, Chen and Mangasarian introduced a class of smooth approximations of the function  $(x)_+$ . Let  $\rho : R \rightarrow R^+$  be a density function satisfying

$$k := \int_{-\infty}^{\infty} |s| \rho(s) ds < \infty, \quad \rho(s) = \rho(-s). \quad (22)$$

Then

$$\phi(x, \mu) := \int_{-\infty}^{\infty} (x - \mu s) + \rho(s) ds. \quad (23)$$

Then  $\phi$  is a smoothing function of  $(x)_+$ .

In this paper, we use the smoothing approximation function of the absolute value function as

$$\begin{aligned} \Phi_i(z, \mu) &= \sqrt{((H - I)z + c)_i^2 + \mu^2}, \\ \mu &\in R_+, \quad i = 1, \dots, 2n. \end{aligned} \quad (24)$$

And (24) is the smoothing approximation function of (20). Then, problem (20) is solved by smoothing conjugate gradient method. And formula (24) also satisfies

$$\lim_{\mu \rightarrow 0} \Phi_i(z, \mu) = |((H - I)z + c)_i|, \quad i = 1, \dots, 2n. \quad (25)$$

Based on (24), we can get the following smoothing function of (20):

$$\tilde{f}(z, \mu) = \frac{1}{2} \sum_{i=1}^{2n} \tilde{f}_i^2(z, \mu), \quad (26)$$

where  $\tilde{f}_i(z, \mu) = ((H + I)z + c)_i - \Phi_i(z, \mu)$ ,  $i = 1, \dots, 2n$ .

Now, we give the smoothing FR conjugate gradient method for solving (20).

*Algorithm 4* (the smoothing FR conjugate gradient method).

*Step 1.* Choose  $0 < \sigma < 1$ ,  $0 < \rho < 1$ ,  $r > 0$ ,  $\mu_0 > 1$ ,  $\epsilon > 0$  and consider an initial point  $x_0$  in  $R^n$ . Let  $k = 0$ ; compute  $\tilde{g}_0 = \nabla \tilde{f}(z_0, \mu_0)$ . Let  $d_0 = -\tilde{g}_0$ .

*Step 2.* If  $\|\nabla \tilde{f}\| \leq \epsilon$ , then terminate the method; otherwise, let  $z_{k+1} = z_k + \alpha_k d_k$ , where

$$\begin{aligned} d_k &= -\tilde{g}_k + \beta_{k-1}^{\text{FR}} \cdot d_{k-1}, \\ \tilde{g}_k &= \nabla \tilde{f}(z_k, \mu_k), \\ \beta_k^{\text{FR}} &= \frac{\tilde{g}_{k+1}^T \tilde{g}_{k+1}}{\tilde{g}_k^T \tilde{g}_k}. \end{aligned} \quad (27)$$

*Step 3.* Compute  $\alpha_k$  by Armijo line search, where  $\alpha_k = \max\{\rho^0, \rho^1, \rho^2, \dots\}$  and  $\rho^i$  satisfies

$$\tilde{f}(z_k + \rho^i d_k, \mu_k) \leq \tilde{f}(z_k, \mu_k) + \sigma \rho^i \tilde{g}_k^T d_k. \quad (28)$$

*Step 4.* If  $\|\nabla_z \tilde{f}(z_{k+1}, \mu_k)\| \geq r\mu_k$ , then set  $\mu_{k+1} = \mu_k$ ; otherwise, let  $\mu_{k+1} = \sigma\mu_k$ .

*Step 5.* Set  $k := k + 1$ ; go to Step 2.

Now, we give the convergence analysis of Algorithm 4.

**Theorem 5.** Suppose  $\tilde{g}_k^T d_{k-1} < 0$ , then  $\tilde{g}_k^T d_k \leq -\|\tilde{g}_k\|^2$ .

*Proof.* According to definition  $\tilde{g}_k^T d_k = -\|\tilde{g}_k\|^2 + \beta_k^{\text{FR}} \tilde{g}_k^T d_{k-1}$  and  $\tilde{g}_k^T d_{k-1} < 0$ , we can obtain

$$\tilde{g}_k^T d_k = -\|\tilde{g}_k\|^2 + \beta_k^{\text{FR}} \tilde{g}_k^T d_{k-1} \leq -\|\tilde{g}_k\|^2. \quad (29)$$

It is obviously that  $\tilde{g}_k^T d_k \leq -\|\tilde{g}_k\|^2$ .  $\square$

**Theorem 6.** Suppose  $\tilde{f}(\cdot, \mu)$  is a smoothing function of  $f$ . And if for any constant  $\mu > 0$ ,  $\tilde{f}(\cdot, \mu)$  is bounded on the level set  $L(X_0) = \{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\}$ , then the  $\{z^k\}$  generated by Algorithm 4 satisfies

$$\begin{aligned} \lim_{k \rightarrow \infty} \mu_k &= 0, \\ \liminf_{k \rightarrow \infty} \|\nabla_z \tilde{f}(z_k, \mu_{k-1})\| &= 0. \end{aligned} \quad (30)$$

*Proof.* Define  $K = \{k \mid \mu_{k+1} = \sigma\mu_k\}$ . If  $K$  is a finite set, then there exists an integer  $\bar{k}$ , such that for all  $k > \bar{k}$

$$\|\nabla_z \tilde{f}(z_k, \mu_{k-1})\| \geq r\mu_{k-1}. \quad (31)$$

And  $\mu_k = \mu_{\bar{k}} = \bar{\mu}$  in Step 4 of Algorithm 4. By considering  $\tilde{f}(\cdot, \bar{\mu})$  a smooth function, Theorems 1 and 2 in [26], the conjugate gradient method for solving

$$\min \tilde{f}(z, \bar{\mu}) \quad (32)$$

satisfies

$$\liminf_{k \rightarrow \infty} \|\nabla_z \tilde{f}(z_k, \bar{\mu})\| = 0, \quad (33)$$

which contradicts (31). This shows that  $K$  must be infinite and

$$\lim_{k \rightarrow \infty} \mu_k = 0. \quad (34)$$

Because  $K$  is infinite, we can suppose  $K = \{k_0, k_1, \dots\}$ , where  $k_0 < k_1 < \dots$ , and then we can get

$$\lim_{i \rightarrow \infty} \|\nabla_z \tilde{f}(z_{k+1}, \mu_{k_i})\| \leq r \lim_{i \rightarrow \infty} \mu_{k_i} = 0. \quad (35) \quad \square$$

## 4. Numerical Tests

In this section, we give numerical experiment results of Algorithm 4. The numerical experiments are also considered in [9, 14, 15]. In computing Examples 1, 2, 3, and 4, we compare Algorithm 4 with the smoothing gradient method in [15]. In computing Example 5, we compare Algorithm 4 with GPSR, debiased and minimum norm methods proposed

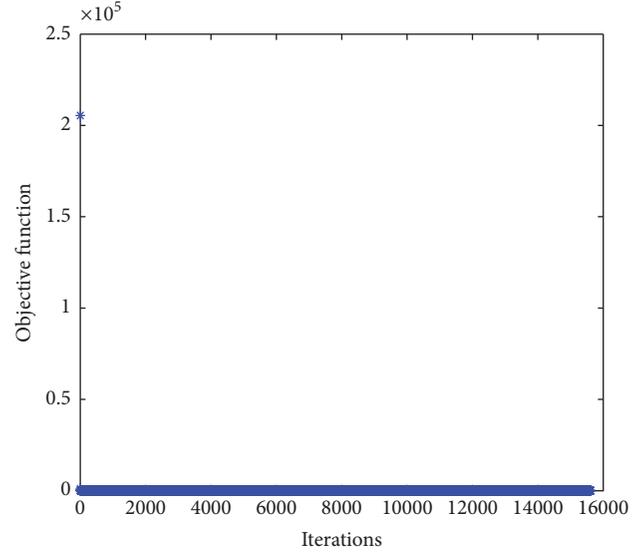


FIGURE 1: Numerical results for solving Example 1 with Algorithm 4.

in [4, 9, 13]. The numerical results of all the examples illustrate that Algorithm 4 is effective. All codes for the test problems are finished in the MATLAB8.0. For example, 1–4, the parameters used in Algorithm 4, are chosen as  $\sigma = 0.2$ ,  $\rho = 0.4$ ,  $r = 0.5$ ,  $\epsilon = 10^{-6}$ .

*Example 1.* Consider the following optimization problem:

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \tau \|x\|_1, \quad (36)$$

where

$$\tau = 5,$$

$$b = (2, 4, 1, 7)',$$

$$A = \begin{pmatrix} 3 & 5 & 8 & 4 & 1 & 5 \\ 2 & 9 & 6 & 5 & 7 & 4 \\ 3 & 4 & 7 & 2 & 1 & 6 \\ 8 & 9 & 6 & 5 & 7 & 4 \end{pmatrix}. \quad (37)$$

The problem has an optimal solution  $x^* = (0.3461, 0.0850, 0, 0, 0.3719, 0)$  in [14, 15]. The optimal solution of our method is  $x_1^* = (0.3461, 0.0850, 0, 0.0003, 0.3720, -0.0001)$ . The optimal solution of smoothing gradient method is  $x_2^* = (0.3461, 0.0850, 0, 0.0003, 0.3720, -0.0001)$ . In Figures 1 and 2, we, respectively, plot the evolution of the objective function versus iteration number when solving Example 1 with Algorithm 4 and smoothing gradient method in [15].

*Example 2.* Consider the following optimization problem:

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \tau \|x\|_1, \quad (38)$$

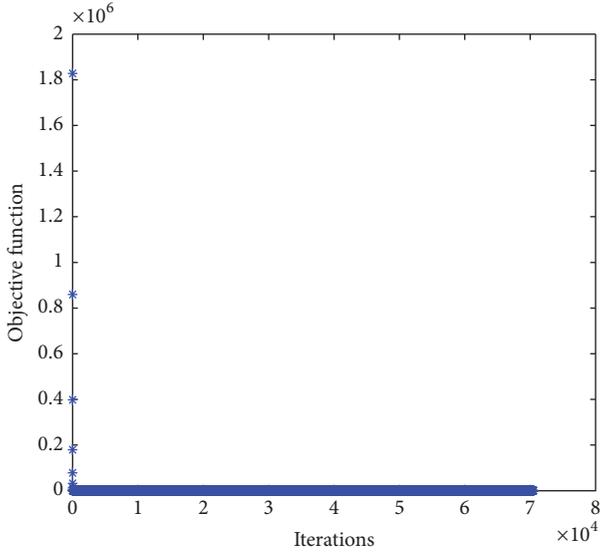


FIGURE 2: Numerical results for solving Example 1 with smoothing gradient method.

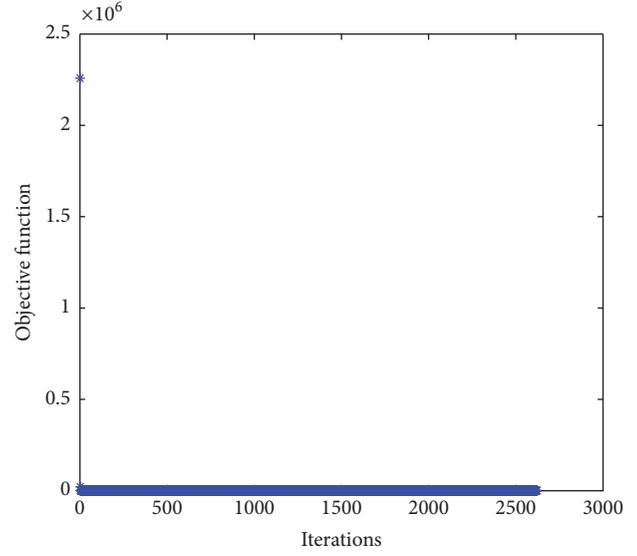


FIGURE 3: Numerical results for solving Example 2 with Algorithm 4 ( $m = 100, n = 110$ ).

where

$$\tau = 10,$$

$$b = (1, 1, \dots, 1),$$

$$A = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ -1 & 4 & -1 & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & -1 & 4 & \ddots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & -1 & 4 & 1 & \cdots & 1 \end{pmatrix}_{n \times (n+10)} \quad (39)$$

Figures 3 and 4 plotted the evolution of the objective function versus iteration number ( $m = 100, n = 110$ ) when solving Example 2 with Algorithm 4 and smoothing gradient method. Figures 5 and 6 plotted the evolution of the objective function versus iteration number ( $m = 200, n = 210$ ) when solving Example 2 with Algorithm 4 and smoothing gradient method. By comparison, we know that the number of iterations of the Algorithm 4 is less than that of the smoothing gradient method in [15].

*Example 3.* Consider the following optimization problem:

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \tau \|x\|_1, \quad (40)$$

where

$$\tau = 10,$$

$$b = (1, 1, \dots, 1),$$

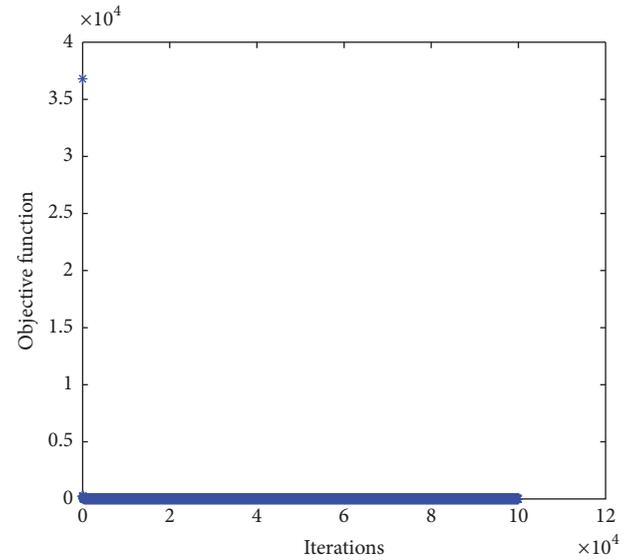


FIGURE 4: Numerical results for solving Example 2 with smoothing gradient method ( $m = 100, n = 110$ ).

$$A = \begin{pmatrix} 2 & 1 & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 1 & 2 & 0 & 0 & \cdots & 0 & \vdots & \cdots & \vdots \\ 1 & 0 & 2 & 0 & \cdots & 0 & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 2 & \cdots & 0 & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 2 & 1 & \cdots & 1 \end{pmatrix}_{n \times (n+10)} \quad (41)$$

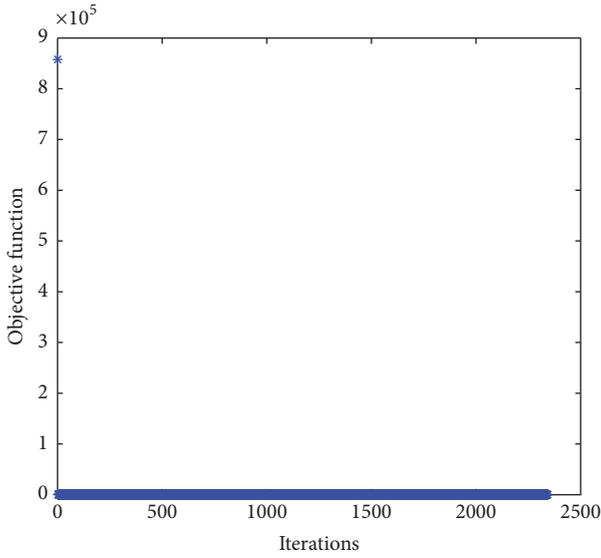


FIGURE 5: Numerical results for solving Example 2 with Algorithm 4 ( $m = 200, n = 210$ ).

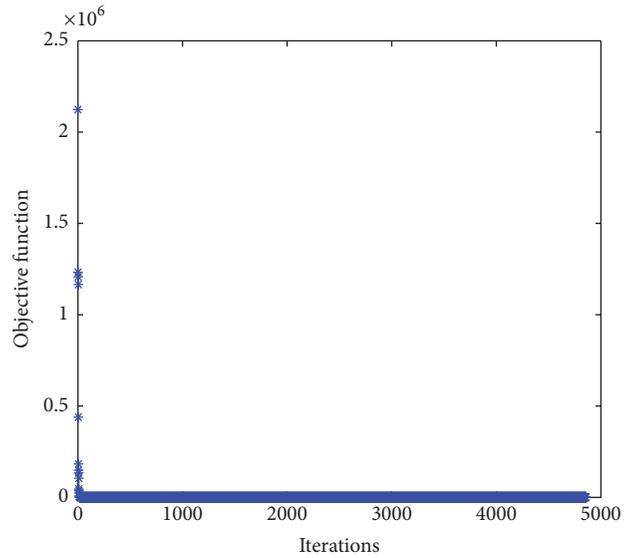


FIGURE 7: Numerical results for solving Example 3 with Algorithm 4 ( $m = 100, n = 110$ ).

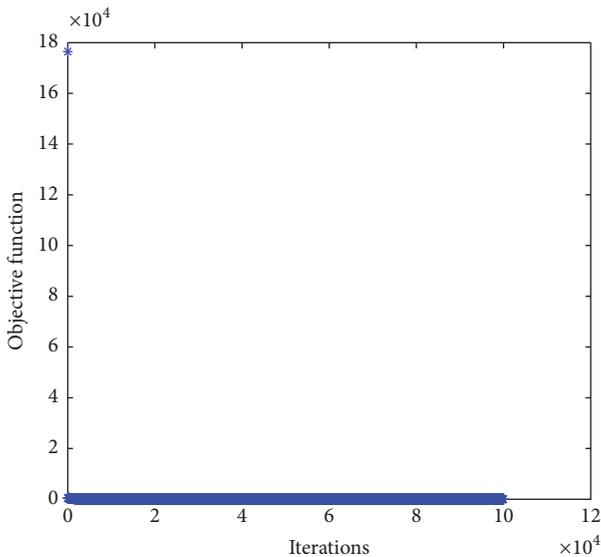


FIGURE 6: Numerical results for solving Example 2 with smoothing gradient method ( $m = 200, n = 210$ ).

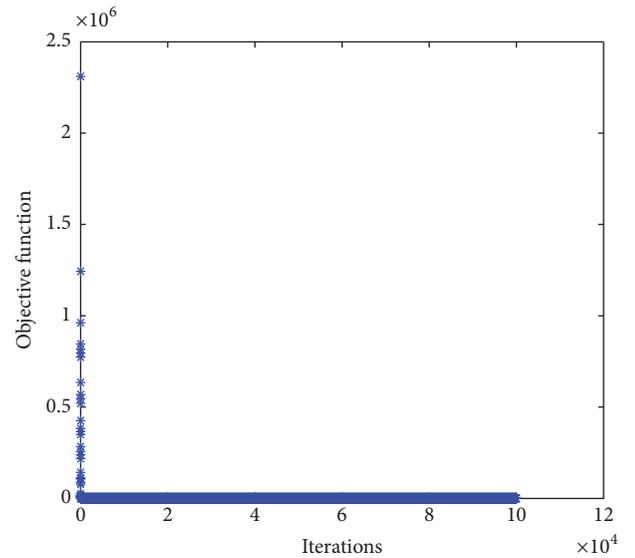


FIGURE 8: Numerical results for solving Example 3 with smoothing gradient method ( $m = 100, n = 110$ ).

Figures 7 and 8, respectively, show the results while  $m = 100, n = 110$ , the change in the number of iterations of the objective function when solving Example 3 with Algorithm 4 and smoothing gradient method. Figures 9 and 10, respectively, show the results while  $m = 150, n = 160$ , the change in the number of iterations of the objective function when solving Example 3 with Algorithm 4 and smoothing gradient method. By comparison, the objective function variants are faster in Algorithm 4 than that of the smoothing gradient method in [15].

*Example 4.* Consider the following optimization problem:

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \tau \|x\|_1, \quad (42)$$

where

$$\tau = 10,$$

$$b = (1, 1, \dots, 1),$$

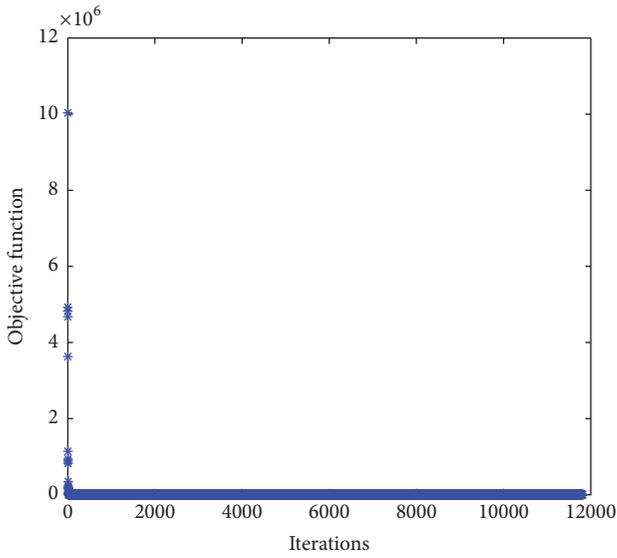


FIGURE 9: Numerical results for solving Example 3 with Algorithm 4 ( $m = 150, n = 160$ ).

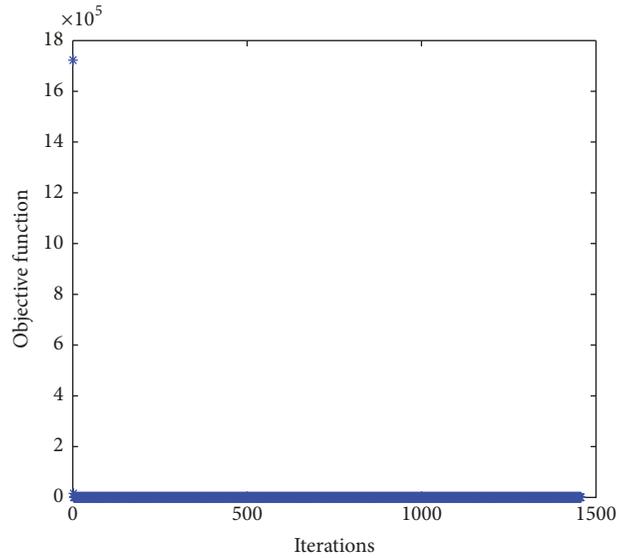


FIGURE 11: Numerical results for solving Example 4 with Algorithm 4 ( $m = 100, n = 110$ ).

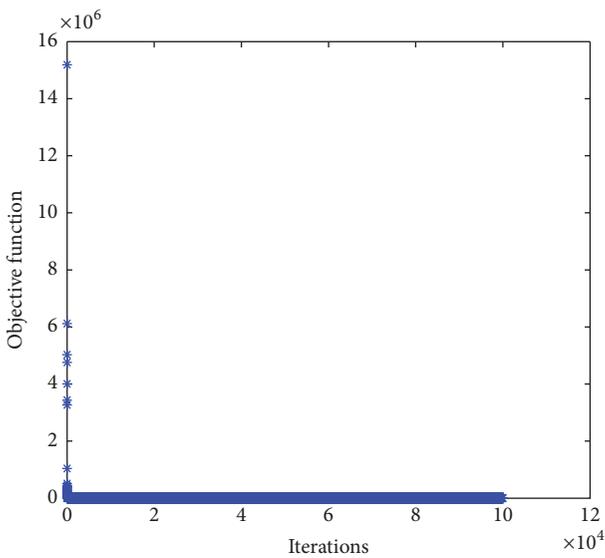


FIGURE 10: Numerical results for solving Example 3 with smoothing gradient method ( $m = 150, n = 160$ ).

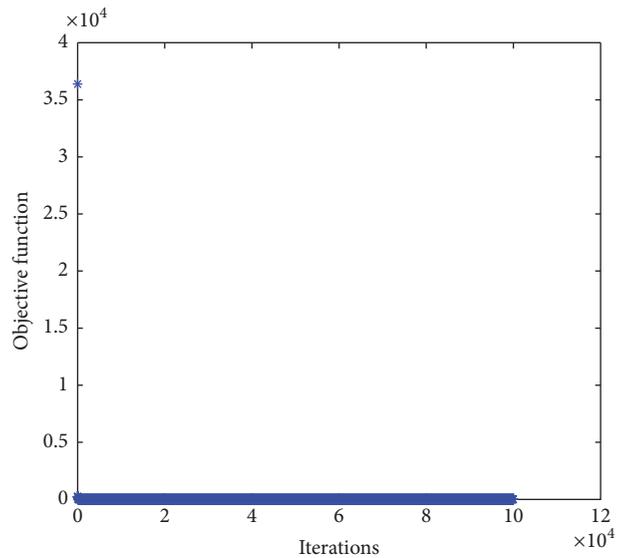


FIGURE 12: Numerical results for solving Example 4 with smoothing gradient method ( $m = 100, n = 110$ ).

$$A = \begin{pmatrix} \frac{1}{n} & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ 0 & \frac{2}{n} & 0 & 0 & \cdots & 0 & \vdots & \cdots & \vdots \\ 0 & 0 & \frac{3}{n} & 0 & \cdots & 0 & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \frac{4}{n} & \cdots & 0 & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 & \cdots & 1 \end{pmatrix}_{n \times (n+10)} \quad (43)$$

Figures 11 and 12, respectively, show the objective function plotted against iteration number while  $m = 100, n = 110$ , when solving Example 4 with Algorithm 4 and smoothing gradient method. Figures 13 and 14, respectively, show the objective function plotted against iteration number while  $m = 150, n = 160$ , when solving Example 4 with Algorithm 4 and smoothing gradient method in [15]. By comparison, Algorithm 4 is more effective than the smoothing gradient method.

*Example 5.* We consider a typical CS scenario, which is also considered in [1–4, 13–15]. The goal is to reconstruct a length- $n$  sparse signal from  $m$  observations, where  $m < n$ . In this

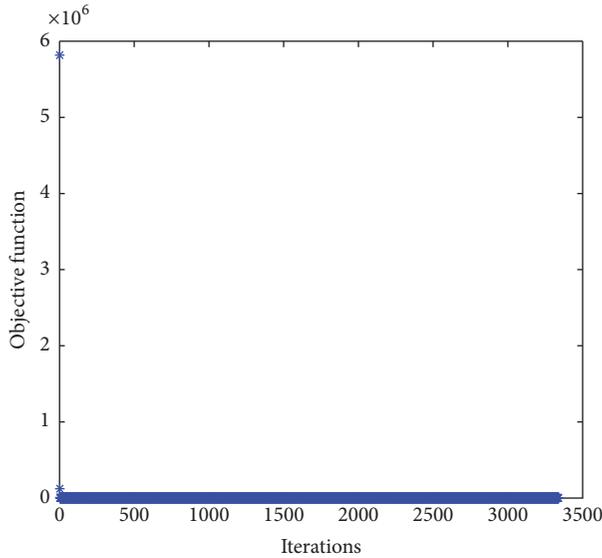


FIGURE 13: Numerical results for solving Example 4 with Algorithm 4 ( $m = 150, n = 160$ ).

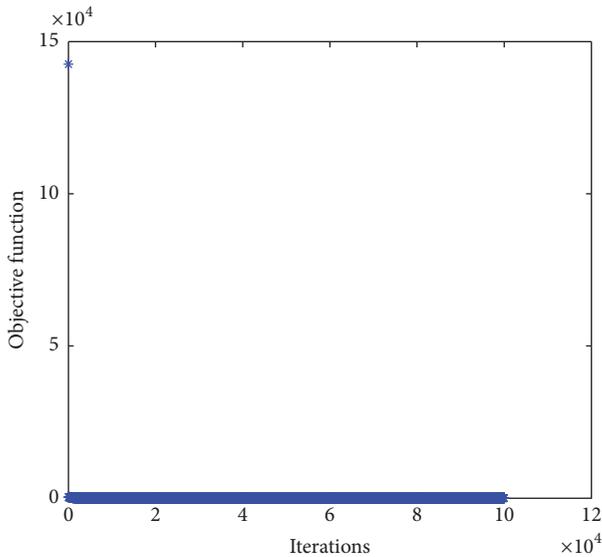


FIGURE 14: Numerical results for solving Example 4 with smoothing gradient method ( $m = 150, n = 160$ ).

example, we choose  $m = 2^3, n = 2^5, \sigma = 0.1, \rho = 0.5, r = 0.5, \epsilon = 10^{-6}$ , and the original signal contains 15 randomly generated  $\pm 1$  spikes. The observation  $b$  is generated according to  $b = Ax + n_e$ , where  $b$  is contaminated by noise. Furthermore, the  $m \times n$  matrix  $A$  is obtained by first filling it with independent samples of a standard Gaussian distribution and then orthonormalizing its rows. In this example, we choose  $\sigma^2 = 10^{-4}$  and  $\tau = 0.1 \|A^T y\|_\infty$  as suggested in [14]. Figure 15 shows the results of the signal reconstruction.

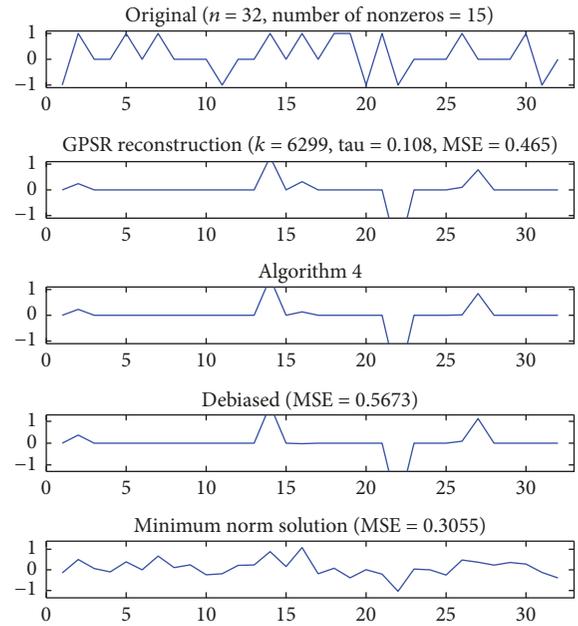


FIGURE 15: Numerical results for solving Example 5 with Algorithm 4.

## 5. Conclusion

Compared with the GPSR method and other methods in [4, 9, 13–15], the smoothing FR conjugate gradient method is simple and needs small storage. The establishment and continuous improvement of the smoothing method for (3) provide a very useful tool to meet the challenges of many practical problems. For example, Figure 15 shows that the smoothing FR conjugate gradient method works well, and it provides an efficient approach to denoise sparse signals. Compare with the smoothing gradient method in [15], the smoothing FR conjugate gradient method is significantly faster than the smoothing gradient method, especially in large-scale iterations. We have also shown that, under weak conditions, the smoothing FR conjugate gradient method converges globally.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by National Natural Science Foundation of China (no. 11671220) and Natural Science Foundation of Shandong Province (no. ZR2016AM29).

## References

- [1] M. Hosein, B. Z. Massoud, and J. Christian, "A fast approach for overcomplete sparse decomposition based on smoothed  $l_0$  norm," *IEEE Transaction on Signal Processing*, vol. 57, no. 1, pp. 289–301, 2009.

- [2] J. Jin, Y. Gu, and S. Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 409–420, 2010.
- [3] K. Shi and P. Shi, "Adaptive sparse Volterra system identification with  $l_0$ -norm penalty," *Signal Processing*, vol. 91, no. 10, pp. 2432–2436, 2011.
- [4] Y. Liu and J. Hu, "A neural network for  $l_1 - l_2$  minimization based on scaled gradient projection: application to compressed sensing," *Neurocomputing*, vol. 173, pp. 988–993, 2016.
- [5] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [6] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.
- [7] A. Y. Yang, S. S. Sastry, A. Ganesh, and Y. Ma, "Fast  $l_1$ -minimization algorithms and an application in robust face recognition: a review," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 1849–1852, Hong Kong, September 2010.
- [8] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.
- [9] 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.
- [10] J. M. Bioucas-Dias and M. A. Figueiredo, "A new TwIst: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.
- [11] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.
- [12] V. D. B. Ewout and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008/09.
- [13] J. Yang and Y. Zhang, "Alternating direction algorithms for  $l_1$ -problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.
- [14] Y. Xiao, Q. Wang, and Q. Hu, "Non-smooth equations based method for  $l_1$ -norm problems with applications to compressed sensing," *Nonlinear Analysis. Theory, Methods and Applications. An International Multidisciplinary Journal*, vol. 74, no. 11, pp. 3570–3577, 2011.
- [15] Y. Chen, Y. Gao, Z. Liu, and S. Du, "The smoothing gradient method for a kind of special optimization problem," *OperationsResearchTransactions*, pp. 119–125, 2017.
- [16] L. Caccetta, B. Qu, and G. Zhou, "A globally and quadratically convergent method for absolute value equations," *Computational optimization and applications*, vol. 48, no. 1, pp. 45–58, 2011.
- [17] O. L. Mangasarian, "Absolute value programming," *Computational optimization and applications*, vol. 36, no. 1, pp. 43–53, 2007.
- [18] O. L. Mangasarian, "A generalized Newton method for absolute value equations," *Optimization Letters*, vol. 3, no. 1, pp. 101–108, 2009.
- [19] G. Yu, L. Qi, and Y. Dai, "On nonmonotone Chambolle gradient projection algorithms for total variation image restoration," *Journal of Mathematical Imaging and Vision*, vol. 35, no. 2, pp. 143–154, 2009.
- [20] 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.
- [21] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, pp. 149–154, 1964.
- [22] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.
- [23] S. Du and Y. Chen, "Global convergence of a modified spectral FR conjugate gradient method," *Applied Mathematics and Computation*, vol. 202, no. 2, pp. 766–770, 2008.
- [24] Y. Dai and Y. Yuan, "A class of globally convergent conjugate gradient methods," *Science China Mathematics*, vol. 46, no. 2, pp. 251–261, 2003.
- [25] Y. Dai, J. Han, G. Liu, D. Sun, H. Yin, and Y.-X. Yuan, "Convergence properties of nonlinear conjugate gradient methods," *SIAM Journal on Optimization*, vol. 10, no. 2, pp. 345–358, 1999.
- [26] D. Pang, S. Du, and J. Ju, "The smoothing Fletcher-Reeves conjugate gradient method for solving finite minimax problems," *ScienceAsia*, vol. 42, no. 1, pp. 40–45, 2016.
- [27] X. Chen, "Smoothing methods for nonsmooth, nonconvex minimization," *Mathematical Programming*, vol. 134, no. 1, Ser. B, pp. 71–99, 2012.
- [28] L. Zhang, S.-Y. Wu, and T. Gao, "Improved smoothing Newton methods for  $P_0$  nonlinear complementarity problems," *Applied Mathematics and Computation*, vol. 215, no. 1, pp. 324–332, 2009.
- [29] O. L. Mangasarian, "Absolute value equation solution via concave minimization," *Optimization Letters*, vol. 1, no. 1, pp. 3–8, 2007.
- [30] O. L. Mangasarian and R. R. Meyer, "Absolute value equations," *Linear Algebra and its Applications*, vol. 419, no. 2–3, pp. 359–367, 2006.

