

ERA, 32(11): 6153–6174. [DOI: 10.3934](https://dx.doi.org/10.3934/era.2024286)/era.2024286 Received: 06 September 2024 Revised: 20 October 2024 Accepted: 06 November 2024 Published: 15 November 2024

https://[www.aimspress.com](https://www.aimspress.com/journal/era)/journal/era

## *Research article*

# Global convergence in a modified RMIL-type conjugate gradient algorithm for nonlinear systems of equations and signal recovery

## Yan Xia<sup>1</sup> and Songhua Wang<sup>2,\*</sup>

<sup>1</sup> School of Artificial Intelligence, Guangzhou Huashang College, Guangzhou 511300, China

<sup>2</sup> School of Mathematics, Physics and Statistics, Baise University, Baise 533099, China

\* Correspondence: Email: wang6600789@bsuc.edu.cn.

Abstract: This paper proposes a modified Rivaie-Mohd-Ismail-Leong (RMIL)-type conjugate gradient algorithm for solving nonlinear systems of equations with convex constraints. The proposed algorithm offers several key characteristics: (1) The modified conjugate parameter is non-negative, thereby enhancing the proposed algorithm's stability. (2) The search direction satisfies sufficient descent and trust region properties without relying on any line search technique. (3) The global convergence of the proposed algorithm is established under general assumptions without requiring the Lipschitz continuity condition for nonlinear systems of equations. (4) Numerical experiments indicated that the proposed algorithm surpasses existing similar algorithms in both efficiency and stability, particularly when applied to large scale nonlinear systems of equations and signal recovery problems in compressed sensing.

Keywords: nonlinear systems of equations; conjugate gradient method; large scale; global convergence; signal recovery

## 1. Introduction

### *1.1. Background*

This paper focuses on solving the problem of finding  $x \in C$  such that:

<span id="page-0-0"></span>
$$
F(x) = 0,\t\t(1.1)
$$

where  $C \subseteq \mathbb{R}^n$  is a bounded, closed, and convex set, and the function  $F: C \to \mathbb{R}^n$  is both continuous and monotone. Specifically, the monotonicity condition of the function is defined by the following inequality:

$$
(F(x) - F(y))^{\mathrm{T}}(x - y) \ge 0, \quad \text{for any } x, y \in C. \tag{1.2}
$$

The study of such a function plays a crucial role in various scientific fields, such as economic equilibrium systems, chemical equilibrium, and signal recovery [\[1](#page-19-0)[–4\]](#page-19-1). Extensive researches have been devoted to developing numerical methods for solving this class of nonlinear systems of equations, leading to a variety of methods such as Newton's methods, quasi-Newton methods, Gauss-Newton methods, Levenberg-Marquardt methods, and so on [\[5](#page-19-2)[–9\]](#page-19-3).

Newton's methods and quasi-Newton methods have been extensively utilized to solve the problem [\(1.1\)](#page-0-0) due to their straightforward implementation and rapid convergence properties. However, these methods necessitate the computation of the Jacobian matrix or its approximation value at every iteration. This can become inefficient when addressing large scale problems. As a result, conjugate gradient methods [\[10–](#page-20-0)[13\]](#page-20-1) have emerged as an effective numerical solution for large scale problems, primarily due to their lower memory requirements. In these conjugate gradient methods, the iterative sequence of points is defined as follows:

$$
x_{k+1} = x_k + t_k d_k, \quad k = 0, 1, 2, \dots
$$

where  $x_{k+1}$  and  $x_k$  are the next and current iteration points, respectively. The step size  $t_k$  is determined through a line search technique, and  $d_k$  denotes the search direction. Generally, the search direction  $d_k$ is updated by using the following formula:

<span id="page-1-0"></span>
$$
d_k = -F_k + \beta_k d_{k-1}, \quad k \ge 1, \quad d_0 = -F_0,\tag{1.3}
$$

where  $F(x_k)$  is abbreviated as  $F_k$ . Here,  $\beta_k$  is a crucial parameter in the iterative process, reflecting the conjugate fundamental characteristics of conjugate gradient methods, and is commonly known as the conjugate fundamental characteristics of conjugate gradient methods, and is commonly known as the conjugate parameter.

A key aspect of conjugate gradient methods ensures that the search direction exhibits sufficient descent and trust region properties, which are critical for guaranteeing global convergence. For example, Narushima et al. [\[14\]](#page-20-2) developed a method capable of generating a descent direction and rigorously proves its global convergence under certain conditions. Moreover, Narushima et al. [\[15\]](#page-20-3) proposed a novel conjugate gradient method that combines smoothing techniques with the Polak-Ribière-Polyak (PRP) method to address unconstrained non-smooth equations. Huang et al. [\[16\]](#page-20-4) proposed a biased stochastic conjugate gradient algorithm for nonconvex problems, which integrates the stochastic recursive gradient method and the modified Barzilai-Borwein technique into the typical stochastic gradient algorithm. Jiang et al. [\[17\]](#page-20-5) proposed two restart conjugate gradient methods for unconstrained optimization, employing different restart procedures. Their restart conditions are based on the Fletcher-Reeves and Dai-Yuan conjugate gradient parameters, with the objective of ensuring that the search directions maintain sufficient descent. In recent years, conjugate gradient methods for unconstrained optimization problems have been a widespread application in solving nonlinear systems of equations. For example, Cheng et al. [\[18\]](#page-20-6) presented a PRP-type method designed for monotone equations, which effectively combines the PRP method from unconstrained optimization with hyperplane techniques. Subsequently, Yu et al. [\[19\]](#page-20-7) introduced a spectral gradient method for large scale nonlinear systems of equations, together with the classical PRP method used in unconstrained minimization problems. Lastly, Waziri et al. [\[20\]](#page-20-8) developed a Hager-Zhang-type conjugate gradient method aimed at solving monotone nonlinear systems of equations. Recently, Liu et al. [\[21\]](#page-20-9) extended a classical three-term framework to a family of three-term conjugate gradient projection methods with a restart procedure and their inertial versions for constrained nonlinear pseudo-monotone equations. Ibrahim et al. [\[22\]](#page-20-10) extended a spectral three-term conjugate gradient method framework for unconstrained optimization to two derivative-free methods.

#### *1.2. Conjugate parameter design*

We will discuss the design of the conjugate parameter in detail. For conjugate gradient methods in unconstrained optimization problems, Rivaie et al. [\[23\]](#page-21-0) proposed a significant modification to the PRP-type conjugate parameter, with the modified expression as follows:

$$
\beta_k^{\text{RMIL}} = \frac{F_k^{\text{T}}(F_k - F_{k-1})}{||d_{k-1}||^2}
$$

It is important to note that if the denominator  $||d_{k-1}||^2$  in the above expression is replaced by  $||F_k||^2$ , then the expression degenerates into the classical PRP-type conjugate parameter. To address the nonnegativity issue of the conjugate parameter  $\beta_k^{RMIL}$ <br>parameter, specifically formulated as follows: *k* , Dai et al. [\[24\]](#page-21-1) introduced a modified RMIL conjugate parameter, specifically formulated as follows:

<span id="page-2-0"></span>
$$
\beta_k^{\text{RMIL+}} = \begin{cases} \frac{F_k^{\text{T}}(F_k - F_{k-1})}{||d_{k-1}||^2}, & \text{if } 0 \le F_k^{\text{T}} F_{k-1} \le ||F_k||^2, \\ 0, & \text{otherwise.} \end{cases}
$$
(1.4)

It is particularly noteworthy that if  $F_k^T F_{k-1}$  is negative or greater than or equal to  $||F_k||^2$ , then the conjugate parameter  $\beta_k^{\text{RML+}}$  becomes redundant, causing the search direction to degenerate into the classical<br>steenest descent direction. Therefore, to ensure that the search direction exhibits both sufficient descent steepest descent direction. Therefore, to ensure that the search direction exhibits both sufficient descent and trust region properties, we further refine [\(1.4\)](#page-2-0) and design a revised RMIL conjugate parameter, i.e.,

<span id="page-2-3"></span>
$$
\beta_k^{\text{MRMIL}} = \frac{\min\left\{ |F_k^{\text{T}}(F_k - F_{k-1})|, ||F_k||^2 \right\}}{\mu \left( ||F_k||^2 + ||d_{k-1}||^2 \right) + ||d_{k-1}||^2},\tag{1.5}
$$

which implies that the non-negativity holds, i.e.,

<span id="page-2-1"></span>
$$
0 \le \beta_k^{\text{MRMIL}} \le \frac{\|F_k\|^2}{\mu \left( \|F_k\|^2 + \|d_{k-1}\|^2 \right) + \|d_{k-1}\|^2}.
$$
\n(1.6)

#### 2. Search direction properties and algorithm framework

In this section, we delve into the intricate aspects of the search direction properties, the line search technique, and the fundamental concept of the projection operator, all of which are critical components in the formulation and analysis of the proposed algorithm.

First, we begin by discussing the sufficient descent and trust-region properties, which are essential for ensuring the convergence and stability of the proposed algorithm. The sufficient descent property ensures that each iteration of the algorithm decreases the objective function by a significant amount. To this end, we multiply both sides of  $(1.3)$  by  $F_k^T$  on the left, yielding:

$$
F_k^{\rm T} d_k = F_k^{\rm T}(-F_k + \beta_k^{\rm MRMIL} d_{k-1}) = -||F_k||^2 + \beta_k^{\rm MRMIL} F_k^{\rm T} d_{k-1}.
$$

Incorporating the non-negatively condition from [\(1.6\)](#page-2-1) and  $\mu (||F_k||^2 + ||d_{k-1}||^2) + ||d_{k-1}||^2 \ge 2\mu ||F_k|| ||d_{k-1}||$ , we can further derive: we can further derive:

<span id="page-2-2"></span>
$$
F_k^{\mathrm{T}} d_k \le -\|F_k\|^2 + \frac{\|F_k\|^2}{2\mu \|F_k\| \|d_{k-1}\|} \|F_k\| \|d_{k-1}\| = -P_1 \|F_k\|^2, \quad P_1 = 1 - \frac{1}{2\mu},\tag{2.1}
$$

which clearly demonstrates that the sufficient descent property is satisfied. The trust-region property is another critical aspect that ensures the stability of the algorithm by bounding the search direction. By applying the Cauchy-Schwarz inequality to [\(2.1\)](#page-2-2), we can derive the following inequality:

$$
-||F_k||||d_k|| \le F_k^{\mathrm{T}} d_k \le -P_1||F_k||^2,
$$

which implies that  $P_1||F_k|| \le ||d_k||$  holds. In addition, from [\(1.3\)](#page-1-0) and [\(1.6\)](#page-2-1), we obtain:

$$
||d_k|| \le ||F_k|| + \frac{||F_k||^2}{2\mu||F_k||||d_{k-1}||} ||d_{k-1}|| = ||F_k|| + \frac{1}{2\mu}||F_k|| = P_2||F_k||, \quad P_2 = 1 + \frac{1}{2\mu}.
$$
  
Hence, the trust-region property is satisfied, i.e.,

<span id="page-3-1"></span>
$$
P_1||F_k|| \le ||d_k|| \le P_2||F_k||. \tag{2.2}
$$

Next, we turn our attention to the line search technique, a method used to determine the appropriate step size  $t_k$  along the search direction  $d_k$ . The step size is calculated by using the formula  $t_k = \xi \rho^{m_k}$ , where  $m_k$  is the smallest non-negative integer  $m$  that satisfies the following condition:

<span id="page-3-0"></span>
$$
-F(x_k + \xi \rho^m d_k)^T d_k \ge \sigma \xi \rho^m \|F(x_k + \xi \rho^m d_k)\| \|d_k\|^2. \tag{2.3}
$$

### Algorithm 1

1: Initialize  $x_0 \in \mathbb{R}^n$ ,  $d_0 = -F_0$ ,  $\mu > 1$ ,  $\sigma > 0$ ,  $\rho \in (0, 1)$ ,  $\xi > 0$ ,  $0 < \epsilon < 1$ , and set the counter  $k := 0$ .<br>2. while  $||F|| > \epsilon$  do 2: **while**  $||F_k|| > \epsilon$  do<br>3: Calculate the s Calculate the step size  $t_k$  by [\(2.3\)](#page-3-0). 4: Set the point  $w_k = x_k + t_k d_k$ . 5: **if**  $x_k \in C$  and  $||F(w_k)|| \le \epsilon$  **then** 6: **Break**. Break. 7: else 8: Calculate the next iteration by  $x_{k+1} = P_C[x_k - \gamma \theta_k F(w_k)], \quad \theta_k =$  $F(w_k)^{\mathrm{T}}(x_k - w_k)$ ∥*F*(*wk*)∥ 2 .  $(2.4)$ 9: end if 10: Calculate  $\theta_k^{\text{MRMIL}}$  by [\(1.5\)](#page-2-3) and  $d_k$  by [\(1.3\)](#page-1-0). 11: Set the counter  $k := k + 1$ .

12: end while

Lastly, we define the projection operator, an essential tool in optimization, particularly in constrained problems. The projection operator ensures that the iteration point of the algorithm remain within the feasible region defined by the problem's constraints. The projection operator  $P_C$  onto a convex set  $C$  is defined as follows:

$$
P_C[x] = \arg \min \{ ||x - y|| \mid y \in C \}, \quad x \in \mathbb{R}^n,
$$

where ∥*x* − *y*∥ denotes the Euclidean distance between *x* and *y*. To be specific, the projection operator possesses a well-known non-expansive property, which can be expressed as:

$$
||P_C[x] - P_C[y]|| \le ||x - y||, \quad \forall x, y \in \mathbb{R}^n.
$$

To sum up, we propose a modified RMIL conjugate gradient algorithm (abbreviated as the MRMIL algorithm). The procedure of this algorithm can be outlined and discussed as shown in Algorithm 1.

#### 3. Convergence analysis

In the following discussion, we consistently assume that the sequences  $\{t_k\}$ ,  $\{d_k\}$ ,  $\{x_k\}$ ,  $\{w_k\}$ , and  $\{F_k\}$ are generated by the MRMIL algorithm. Additionally, we introduce the following general assumptions: Assumption S:

- (S1) The solution set  $C_*$  of the problem [\(1.1\)](#page-0-0) is non-empty.
- (S2) The function *F* is monotone and continuous on  $\mathbb{R}^n$ .

It is important to note that the continuity condition in Assumption S2 is less restrictive than the Lipschitz continuity, which is the more traditional assumption in algorithms designed to solve nonlinear systems of equations.

**Lemma 1.** The line search technique is well-defined. That is, for all  $k \geq 0$ , there exists a non-negative *integer*  $m_k$  *such that* [\(2.3\)](#page-3-0) *holds.* 

*Proof.* Similar to Lemma 3.2 in [\[4\]](#page-19-1), the proof can be completed straightforwardly. □

<span id="page-4-0"></span>Lemma 2. *The following conclusions hold:*

- *(1) The sequence*  $\{||x_k x_*||\}$  *converges for any*  $x_* \in C_*$ *.*
- *(2) The sequence* {*xk*} *is bounded.*
- *(3)*  $\lim_{k \to \infty} t_k ||d_k|| = 0.$

*Proof.* Similar to Lemma 3.3 in [\[4\]](#page-19-1), the proof can be completed easily. □

Theorem 1. *The sequence* {*xk*} *globally converges to a solution of the problem* [\(1.1\)](#page-0-0)*, i.e.,*

$$
\lim_{k\to\infty}\inf||F_k||=0.
$$

*Proof.* To establish the global convergence of the sequence  $\{x_k\}$ , we consider two possible cases based on the behavior of  $||F_k||$  as  $k \to \infty$ .

Case 1 ( $\lim_{k \to \infty} \inf ||F_k|| = 0$ ): In this case, the continuity of *F* implies that the sequence {*x<sub>k</sub>*} has an accumulation point  $x_*$  such that  $F(x_*) = 0$ . Additionally, the convergence of  $\{\|x_k - x_*\|\}$  further indicates that the sequence  $\{x_k\}$  converges to  $x_*$ .

Case 2 ( $\lim_{k \to \infty} \inf ||F_k|| > 0$ ): Assume, for the sake of contradiction, that there exists a constant  $r > 0$ such that  $||F_k|| \ge r$  for any  $k \ge 0$ . This assumption, together with [\(2.2\)](#page-3-1), implies that  $||d_k|| \ge P_1||F_k|| \ge$ *P*<sub>1</sub>*r*. Combining this and Lemma [2\(](#page-4-0)3), it follows that  $\lim_{k \to \infty} t_k = 0$ . Due to the boundary of the sequences  ${x<sub>k</sub>}$  and  ${d<sub>k</sub>}$ , there exists an infinite index set  $\mathcal L$  such that

$$
\lim_{l\to\infty,l\in\mathcal{L}}x_{k_l}=\ddot{a}\quad\text{and}\quad\lim_{l\to\infty,l\in\mathcal{L}}d_{k_l}=\ddot{d}.
$$

From the sufficient descent condition defined in [\(2.1\)](#page-2-2), we have  $-F(x_k)^T d_{k_l} \ge P_1 ||F(x_{k_l})||^2$  for any  $l \in \mathcal{L}$ . Taking the limit as  $l \rightarrow \infty$  in the above inequality, we obtain

$$
-F(\ddot{a})^{\mathrm{T}}\ddot{d} \ge P_1 ||F(\ddot{a})||^2 > P_1 r^2 > 0.
$$

Additionally, applying the line search technique defined in [\(2.3\)](#page-3-0), we obtain  $-F(x_{k_l} + \rho^{-1}t_{k_l}d_{k_l})^T d_{k_l} < \tau^{-1}t$ ,  $||F(x_{k_l} + \rho^{-1}t_{k_l}d_{k_l})||d_{k_l}||^2$  for any  $l \in \mathcal{L}$ . Taking the limit as  $l \to \infty$  in this inequality,  $\sigma \rho^{-1} t_{k_l} || F(x_{k_1} + \rho^{-1} t_{k_l} d_{k_1}) || ||d_{k_1}||^2$  for any  $l \in \mathcal{L}$ . Taking the limit as  $l \to \infty$  in this inequality, we obtain

$$
-F(\ddot{a})^{\mathrm{T}}\ddot{d} \leq 0.
$$

The two inequalities are clearly contradictory. Therefore, the assumption that  $\lim_{k \to \infty} \inf ||F_k|| > 0$  must  $\lim_{k \to \infty} \inf ||F_k|| > 0$ be false, which implies that  $\lim_{k \to \infty} \inf ||F_k|| = 0.$ 

#### 4. Numerical experiments

In this section, we present the general experimental setups and conduct two distinct experiments: one focused on solving nonlinear systems of equations and the other on sparse signal restoration.

#### *4.1. General setups*

To validate the effectiveness of the MRMIL algorithm, we conduct a comparative analysis against other similar algorithms, including the modified fletcher-reeves method (MFRM) algorithm [\[25\]](#page-21-2) and the hybrid three-term conjugate gradient projection (HTTCGP) algorithm [\[26\]](#page-21-3). This comparative study provides a robust framework for evaluating the performance of the MRMIL algorithm, particularly in large scale nonlinear systems of equations and sparse signal restoration problems. All experimental codes are executed on a system running the Ubuntu 20.04.2 LTS 64-bit operating system, equipped with an Intel(R) Xeon(R) Gold 5115 processor with a clock speed of 2.40 GHz. In the comparative analysis, the MRMIL algorithm is directly compared to the MFRM and HTTCGP algorithms. A key aspect of the analysis is the modification of step 10 in the MRMIL algorithm, where the search direction is calculated by using both the MFRM and HTTCGP algorithms. This modification allows us to isolate and assess the impact of different search direction strategies on the overall performance of the algorithm. Importantly, all other steps of the MRMIL algorithm remain unchanged during this comparison, ensuring that any differences in performance could be attributed to the search direction computation.

The parameters for the MRMIL algorithm are carefully selected to optimize its performance. Specifically, the parameters are set as follows:  $\mu = 2$ ,  $\sigma = 10^{-4}$ ,  $\rho = 0.74$ ,  $\xi = 1$ , and  $\epsilon = 10^{-5}$ .<br>For the MERM and HTTCGP algorithms, the parameters are maintained according to their default For the MFRM and HTTCGP algorithms, the parameters are maintained according to their default setting as specified in the original references.

#### *4.2. Experiments on nonlinear systems of equations*

To thoroughly evaluate the performance of the proposed algorithm, we selecte a set of benchmark problems with varying dimensions, where the number of variables ranges from [1000 5000 10,000 50,000 100,000]. The initial points for these benchmark problems are set as follows to provide diverse starting conditions, ensuring a comprehensive assessment of the algorithm:  $x_1 = [0, 1]^n$ ,<br>  $(1, 1, 2, \ldots, n)$ <sup>T</sup>  $x_2 = \left(1 - \frac{1}{n}\right)$  $\frac{1}{n}$ , 1 –  $\frac{2}{n}$  $\frac{2}{n}, \ldots, 1 - \frac{n}{n}$  $\left(\frac{n}{3}\right)^{\text{T}}, \ x_3 = \left(\frac{1}{3}\right)^{\text{T}}$ 3 , 1  $\frac{3^2}{1}$ 1  $\left(\frac{1}{3^n}\right)^T$ ,  $x_4 = \left(\frac{1}{n}\right)$ *n* , 2  $n, \ldots,$ *n*  $\left(1, \frac{1}{2}\right)$ <sup>T</sup>,  $x_5 = \left(1, \frac{1}{2}\right)$  $_2,\ldots,$ 1  $\left(\frac{1}{n}\right)^{\mathrm{T}},$  $x_6 = (1, 1, ..., 1)^T, x_7 = \left(\frac{1}{2}\right)$ 2 mulated as  $F(x) = (F_1(x), F_2(x), ..., F_n(x))^T$  with  $x = (x_1, x_2, ..., x_n)^T$ , as defined in Problems 1–8. 1  $\frac{1}{2^2}, \ldots, \frac{1}{2^n}$  $\left(\frac{1}{2^n}\right)^T$ ,  $x_8 = \left(0, \frac{1}{n}\right)^T$ *n n*−1  $\left(\frac{-1}{n}\right)^{T}$ . The benchmark problems are for-Each algorithm is executed on these benchmark problems with a termination setting:  $||F_k|| \leq \epsilon$  or when the number of iterations exceeded 2000.

<span id="page-6-0"></span>

Figure 1. Performance profiles on NIter.

<span id="page-6-1"></span>

Figure 2. Performance profiles on NF.

Problem 1. Let the function  $F(x)$  be defined as:

$$
F_1(x) = e^{x_1} - 1,
$$
  
\n
$$
F_i(x) = e^{x_i} + x_i - 1, \text{ for } i = 2, 3, \dots, n,
$$

where the feasible region is  $C = \mathbb{R}^n_+$ .

Problem 2. Let the function  $F(x)$  be defined as:

$$
F_i(x) = e^{x_i} - 1
$$
, for  $i = 1, 2, \dots, n$ ,

where the feasible region is  $C = \mathbb{R}^n_+$ .

Problem 3. Let the function  $F(x)$  be defined as:

$$
F_i(x) = 2x_i - \sin(x_i)
$$
, for  $i = 1, 2, \dots, n$ ,

where the feasible region is  $C = [-2, +\infty)$ .

Problem 4. Let the function  $F(x)$  be defined as:

$$
F_i(x) = \log(x_i + 1) - \frac{x_i}{n}
$$
, for  $i = 1, 2, \dots, n$ ,

where the feasible region is  $C = [-1, +\infty)$ .

Problem 5. Let the function  $F(x)$  be defined as:

$$
F_i(x) = (e^{x_i})^2 + 3\sin(x_i)\cos(x_i) - 1
$$
, for  $i = 1, 2, \dots, n$ ,

where the feasible region is  $C = \mathbb{R}^n_+$ .

Problem 6. Let the function  $F(x)$  be defined as:

$$
F_1(x) = 2x_1 + \sin(x_1) - 1,
$$
  
\n
$$
F_i(x) = 2x_{i-1} + 2x_i + \sin(x_i) - 1, \text{ for } i = 2, 3, \dots, n - 1,
$$
  
\n
$$
F_n(x) = 2x_n + \sin(x_n) - 1,
$$

where the feasible region is  $C = \mathbb{R}^n_+$ .

Problem 7. Let the function  $F(x)$  be defined as:

$$
F_i(x) = \frac{1}{n}e^{x_i} - 1
$$
, for  $i = 1, 2, \dots, n$ ,

where the feasible region is  $C = \mathbb{R}^n_+$ .

Problem 8. Let the function  $F(x)$  be defined as:

$$
F_i(x) = x_i - 2\sin(|x_i - 1|)
$$
, for  $i = 1, 2, \dots, n$ ,

<span id="page-7-0"></span>where the feasible region is  $C = \mathbb{R}^n_+$ .



Figure 3. Performance profiles on CPUT.

The performance of the MRMIL, MFRM, and HTTCGP algorithms is evaluated through a series of benchmark problems, with the results shown in Tables [1–](#page-9-0)[8.](#page-16-0) In these tables, the term "Init(n)" denotes the initial points and the corresponding dimension multiplied by 1000. The results are presented in the format CPUT/NF/NIter/Norm, where "CPUT" denotes the CPU time in seconds, "NF" denotes the number of function evaluations, "NIter" denotes the number of iterations, and "Norm" denotes the norm of the function at the approximate optimal point. The tables demonstrate that all three algorithms effectively solved the benchmark problems across a range of initial points and dimensions. It is particularly noteworthy that the MRMIL algorithm consistently outperforms the others in most cases. To thoroughly analyze the performance of these three algorithms, we can employ the performance profiles methodology developed by Dolan and Moré  $[27]$  $[27]$ , which is defined by the following fraction:

$$
\rho(\tau) \triangleq \frac{1}{|T_p|} \left| \left\{ t_p \in T_p \; : \; \log_2 \left( \frac{t_{p,q}}{\min\{ t_{p,q} : q \in Q \}} \right) \leq \tau \right\} \right|,
$$

where  $T_p$  is the test set,  $|T_p|$  is the number of problems in the test set, Q is the set of solvers, and  $t_{p,q}$ is the performance indicator (e.g., NIter, NF, and CPUT) for  $t_p \in T_p$  and  $q \in Q$ . These profiles offer a comprehensive way for comparing algorithms, where a higher curve indicates better performance of the associated method. By plotting these performance profiles for the three algorithms in terms of CPUT, NF, and NIter, we can visually assess and compare their efficiency and effectiveness (see Figures [1](#page-6-0)[–3\)](#page-7-0). According to Figure [1,](#page-6-0) the MRMIL algorithm demonstrates remarkable efficiency, solving 75% of the benchmark problems with the least NIter. Furthermore, Figures [2](#page-6-1) and [3](#page-7-0) provide additional insights into the the algorithm's effectiveness: the MRMIL algorithm successfully solves approximately 72% (73%) of the benchmark problems with less NF (CPUT) compared to the MFRM and HTTCGP algorithms, which solve around 20% (19%) and 27% (21%) of the problems, respectively. These results underscore the superior efficiency of the proposed algorithm, highlighting its ability to optimize computational resources while maintaining high performance across a broad range of test cases.

#### *4.3. Experiments on sparse signal restoration*

To further illustrate the effectiveness of the MRMIL algorithm in addressing real-world challenges, we apply it to a sparse signal restoration problem [\[28\]](#page-21-5) as a case study. The objective of this problem is to accurately restore a sparse signal, denoted as  $\hat{a}$ , from an observed signal  $\tilde{a}$ , where both signals have a dimension of *n*. The quality of the signal restoration is assessed by the mean squared error (MSE), a widely accepted evaluation metric in signal processing. The MSE is defined as MSE =  $||\hat{a} - \tilde{a}||^2/n$ ,<br>where  $\hat{a}$  represents the original signal, and  $\tilde{a}$  represents the restored signal. In this experiment, the where  $\hat{a}$  represents the original signal, and  $\tilde{a}$  represents the restored signal. In this experiment, the original signal is configured with a length of  $n = 5120$  and  $k = 1280$ . Notably, these non-zero elements are randomly distributed across the signal, with a total of 160 such elements.

Figure [4](#page-17-0) provides the results of the three algorithms for the sparse signal restoration problem, where the MRMIL algorithm clearly outperforms the others in terms of both NIter and CPUT. To account for the inherent randomness in the experiment, we conducted 10 independent trials for each algorithm, utilizing different random seeds in each trial. The average results from these experiments are shown in Table [9.](#page-17-1) The data clearly indicate that the MRMIL algorithm excels in restoring the original signal, surpassing both the MFRM and HTTCGP algorithms by requiring fewer iterations and less computational time. These findings not only validate the superior efficiency of the MRMIL algorithm but also highlight its robustness and consistency across multiple trials.

<span id="page-9-0"></span>

Table 1. Numerical results for Problem 1.



Table 2. Numerical results for Problem 2.



Table 3. Numerical results for Problem 3.

Inti $(n)$	<b>MRMIL</b>	<b>MFRM</b>	<b>HTTCGP</b>
	CPUT/NF/NIter/Norm	CPUT/NF/NIter/Norm	CPUT/NF/NIter/Norm
x1(1)	9.47e-03/23/10/7.78e-08	1.47e-03/11/5/3.60e-08	6.54e-03/53/25/8.18e-07
x2(1)	2.36e-03/19/8/5.31e-07	8.59e-04/7/3/5.17e-07	5.73e-03/49/23/9.64e-07
x3(1)	3.79e-03/40/15/4.65e-07	1.96e-02/208/75/9.12e-07	6.22e-03/63/24/8.79e-07
x4(1)	2.96e-03/26/11/3.98e-07	1.58e-03/13/6/1.74e-08	5.97e-03/51/24/8.99e-07
x5(1)	4.14e-03/43/17/4.58e-07	1.84e-02/217/78/8.85e-07	5.64e-03/64/23/3.29e-07
x6(1)	2.93e-03/46/17/5.90e-07	1.86e-02/291/109/8.46e-07	5.77e-03/73/26/9.67e-07
x7(1)	3.95e-03/44/17/7.42e-07	1.36e-02/205/74/9.36e-07	6.73e-03/95/34/9.34e-07
x8(1)	3.55e-03/52/19/2.11e-07	1.53e-02/237/84/8.63e-07	4.28e-03/65/25/7.98e-07
x1(5)	4.96e-03/23/10/2.08e-07	2.40e-03/11/5/6.26e-09	1.26e-02/52/24/6.55e-07
x2(5)	3.93e-03/21/9/5.73e-07	1.62e-03/7/3/1.75e-07	1.25e-02/53/25/8.20e-07
x3(5)	$3.00e-03/16/7/1.83e-07$	1.98e-03/9/4/2.00e-09	8.10e-03/36/16/8.24e-07
x4(5)	4.79e-03/26/11/8.90e-07	2.60e-03/13/6/2.36e-09	1.70e-02/55/26/7.71e-07
x5(5)	9.46e-03/45/17/2.06e-07	4.25e-02/177/62/8.65e-07	1.81e-02/66/26/9.53e-07
x6(5)	1.18e-02/50/19/3.77e-07	9.35e-02/271/97/9.78e-07	3.62e-02/85/31/6.06e-07
x7(5)	1.48e-02/44/17/1.40e-07	7.01e-02/211/76/2.42e-07	3.40e-02/116/37/3.72e-07
x8(5)	1.46e-02/51/19/3.17e-07	6.78e-02/263/93/9.99e-07	3.11e-02/105/36/5.28e-07
x1(10)	$1.10e-02/23/10/3.00e-07$	6.53e-03/11/5/3.62e-09	3.28e-02/52/24/9.35e-07
x2(10)	1.05e-02/21/9/8.62e-07	3.14e-03/7/3/1.21e-07	2.72e-02/50/23/6.50e-07
x3(10)	6.52e-03/16/7/2.03e-07	3.37e-03/9/4/9.73e-10	1.83e-02/36/16/8.28e-07
x4(10)	1.53e-02/28/12/1.39e-07	5.89e-03/13/6/1.24e-09	3.80e-02/52/24/6.11e-07
x5(10)	1.82e-02/43/16/7.79e-07	7.53e-02/178/63/7.91e-07	3.82e-02/78/28/8.98e-07
x6(10)	2.26e-02/49/18/8.21e-07	7.14e-02/163/60/5.61e-07	4.75e-02/91/31/1.42e-07
x7(10)	2.15e-02/49/18/9.55e-07	1.08e-01/231/83/5.07e-07	4.71e-02/91/32/8.91e-07
x8(10)	1.62e-02/46/18/5.27e-07	5.20e-02/163/59/6.17e-07	5.75e-02/126/42/8.53e-07
x1(50)	3.24e-02/23/10/6.82e-07	2.04e-02/11/5/1.34e-09	9.80e-02/51/23/7.02e-07
x2(50)	2.97e-02/25/10/2.55e-07	8.31e-03/7/3/6.32e-08	8.49e-02/52/24/8.77e-07
x3(50)	2.88e-02/16/7/2.21e-07	2.05e-02/9/4/2.18e-10	4.89e-02/31/13/7.71e-07
x4(50)	3.35e-02/28/12/3.12e-07	2.54e-02/13/6/4.01e-10	1.05e-01/54/25/8.26e-07
x5(50)	4.97e-02/42/16/7.40e-07	1.95e-01/156/54/8.16e-07	8.89e-02/72/27/6.90e-07
x6(50)	6.08e-02/50/19/7.08e-07	2.92e-01/210/74/9.78e-07	1.82e-01/119/37/8.99e-07
x7(50)	6.98e-02/50/19/9.13e-07	3.01e-01/260/95/4.00e-07	1.70e-01/102/38/9.01e-07
x8(50)	6.69e-02/56/21/2.67e-07	2.42e-01/212/75/6.80e-07	1.42e-01/82/31/1.80e-07
x1(100)	3.66e-02/23/10/9.66e-07	$1.70e-02/11/5/9.27e-10$	1.00e-01/51/23/9.93e-07
x2(100)	3.99e-02/25/10/3.61e-07	1.08e-02/7/3/5.40e-08	8.99e-02/49/22/6.89e-07
x3(100)	2.96e-02/16/7/2.23e-07	1.41e-02/9/4/1.28e-10	5.88e-02/31/13/7.71e-07
x4(100)	6.17e-02/28/12/4.42e-07	2.77e-02/13/6/2.71e-10	1.08e-01/51/23/6.49e-07
x5(100)	7.02e-02/50/18/2.25e-07	3.47e-01/216/74/2.32e-07	1.31e-01/73/28/8.05e-07
x6(100)	9.04e-02/53/20/1.91e-07	3.31e-01/197/68/6.68e-07	1.81e-01/94/35/6.66e-07
x7(100)	9.46e-02/53/20/2.18e-07	4.07e-01/221/79/4.32e-07	1.95e-01/100/37/7.41e-07
x8(100)	1.25e-01/54/20/2.42e-07	4.69e-01/267/94/8.29e-07	2.64e-01/120/42/8.12e-07

Table 4. Numerical results for Problem 4.





Table 6. Numerical results for Problem 6.



Table 7. Numerical results for Problem 7.

<span id="page-16-0"></span>

Table 8. Numerical results for Problem 8.

<span id="page-17-0"></span>

Figure 4. From top to bottom: the original signal, the observed signal, and the recovered signal of three algorithms.

<span id="page-17-1"></span>

No.	<b>MRMIL</b>	<b>MFRM</b>	<b>HTTCGP</b>
	NIter/CPUT/MSE	NIter/CPUT/MSE	NIter/CPUT/MSE
	68/0.54/1.136e-05	130/0.95/2.379e-04	94/0.70/1.323e-05
2	67/0.49/2.418e-05	141/1.03/3.226e-04	153/1.07/2.637e-05
3	68/0.51/8.482e-06	142/1.00/1.182e-04	96/0.68/9.684e-06
4	70/0.53/7.291e-06	140/1.01/1.741e-04	127/0.90/8.306e-06
5	77/0.59/1.827e-05	130/0.92/3.736e-04	153/1.08/2.019e-05
6	70/0.53/5.458e-06	132/0.94/7.814e-05	135/0.96/6.194e-06
7	73/0.54/1.109e-05	130/0.97/1.831e-04	124/0.88/1.250e-05
8	62/0.45/8.374e-06	124/0.88/1.468e-04	107/0.75/9.071e-06
9	72/0.56/1.150e-05	139/1.00/2.949e-04	106/0.75/1.312e-05
10	69/0.52/1.229e-05	139/0.98/2.162e-04	140/0.98/1.370e-05
Average	69.60/0.53/1.183e-05	134.70/0.97/2.146e-04	123.50/0.87/1.324e-05

Table 9. The comparison results of three algorithms under different randomness.



Figure 5. The comparison results of three algorithms.

## 5. Conclusions

In this paper, we presented a modified RMIL-type conjugate gradient algorithm specifically tailored to solve nonlinear systems of equations with convex constraints. The proposed algorithm incorporates several key innovations that enhance both its theoretical foundation and practical performance. A central feature of the algorithm is the modification of the conjugate parameter, ensuring its non-negativity. This adjustment plays a critical role in improving the overall stability of the algorithm by preventing negative conjugate parameters that could destabilize the iterative process. The proposed algorithm also guarantees sufficient descent and trust region properties in the search direction, eliminating the need for traditional line search techniques, which often add complexity and computational cost. We established the global convergence of the algorithm under general conditions, without relying on the commonly imposed assumption of Lipschitz continuity. Extensive numerical experiments were conducted to evaluate the performance of the proposed algorithm. The results demonstrate that it consistently outperforms existing algorithms, particularly in terms of computational efficiency. This is especially notable in large scale nonlinear systems of equations. Furthermore, the algorithm shows superior performance in signal recovery problems.

## Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

## Acknowledgments

This work is supported by the Natural Science Foundation in Guangxi Province, PR China (grant number 2024GXNSFAA010478) and the Guangzhou Huashang College Daoshi Project (grant number 2023HSDS38).

## Conflict of interest

The authors declare there are no conflicts of interest.

## References

- <span id="page-19-0"></span>1. M. Sun, Y. Wang, General five-step discrete-time Zhang neural network for timevarying nonlinear optimization, *Bull. Malays. Math. Sci. Soc.*, 43 (2020), 1741–1760. https://doi.org/10.1007/[s40840-019-00770-4](https://dx.doi.org/https://doi.org/10.1007/s40840-019-00770-4)
- 2. K. Meintjes, A. P. Morgan, A methodology for solving chemical equilibrium systems, *Appl. Math. Comput.*, 22 (1987), 333–361. https://doi.org/10.1016/[0096-3003\(87\)90076-2](https://dx.doi.org/https://doi.org/10.1016/0096-3003(87)90076-2)
- 3. D. Li, S. Wang, Y. Li, J. Wu, A projection-based hybrid PRP-DY type conjugate gradient algorithm for constrained nonlinear equations with applications, *Appl. Numer. Math.*, 195 (2024), 105–125. https://doi.org/10.1016/[j.apnum.2023.09.009](https://dx.doi.org/https://doi.org/10.1016/j.apnum.2023.09.009)
- <span id="page-19-1"></span>4. D. Li, J. Wu, Y. Li, S. Wang, A modified spectral gradient projection-based algorithm for largescale constrained nonlinear equations with applications in compressive sensing, *J. Comput. Appl. Math.*, 424 (2023), 115006. https://doi.org/10.1016/[j.cam.2022.115006](https://dx.doi.org/https://doi.org/10.1016/j.cam.2022.115006)
- <span id="page-19-2"></span>5. M. W. Yusuf, L. W. June, M. A. Hassan, Jacobian-free diagonal Newton's method for solving nonlinear systems with singular Jacobian, *Malays. J. Math. Sci.*, 5 (2011), 241–255.
- 6. Q. Yan, X. Peng, D. Li, A globally convergent derivative-free method for solving large-scale nonlinear monotone equations, *J. Comput. Appl. Math.*, 234 (2010), 649–657. https://doi.org/10.1016/[j.cam.2010.01.001](https://dx.doi.org/https://doi.org/10.1016/j.cam.2010.01.001)
- 7. H. Abdullahi, A. S. Halilu, M. Y. Waziri, A modified conjugate gradient method via a double direction approach for solving large-scale symmetric nonlinear equations, *J. Numer. Math. Stoch.*, 10 (2018), 32–44.
- 8. I. Yusuf, A. S. Halilu, M. Y. Waziri, Efficient matrix-free direction method with line search for solving large scale systems of nonlinear equations, *Yugosl. J. Oper. Res.*, 30 (2020), 399–412. https://doi.org/10.2298/[YJOR160515005H](https://dx.doi.org/https://doi.org/10.2298/YJOR160515005H)
- <span id="page-19-3"></span>9. D. Q. Huynh, F. N. Hwang, An accelerated structured quasi-Newton method with a diagonal second-order Hessian approximation for nonlinear least squares problems, *J. Comput. Appl. Math.*, 442 (2024), 115718. https://doi.org/10.1016/[j.cam.2023.115718](https://dx.doi.org/https://doi.org/10.1016/j.cam.2023.115718)
- <span id="page-20-0"></span>10. X. Wu, H. Shao, P. Liu, An efficient conjugate gradient-based algorithm for unconstrained optimization and its projection extension to large-scale constrained nonlinear equations with applications in signal recovery and image denoising problems, *J. Comput. Appl. Math.*, 422 (2023), 114879. https://doi.org/10.1016/[j.cam.2022.114879](https://dx.doi.org/https://doi.org/10.1016/j.cam.2022.114879)
- 11. G. Ma, J. Jiang, J. Jian, A modified inertial three-term conjugate gradient projection method for constrained nonlinear equations with applications in compressed sensing, *Numer. Algor.*, 92 (2023), 1621–1653. https://doi.org/10.1007/[s11075-022-01356-1](https://dx.doi.org/https://doi.org/10.1007/s11075-022-01356-1)
- 12. W. Liu, J. Jian, J. Yin, An inertial spectral conjugate gradient projection method for constrained nonlinear pseudo-monotone equations, *Numer. Algor.*, 97 (2024), 985–1015. https://doi.org/10.1007/[s11075-023-01736-1](https://dx.doi.org/https://doi.org/10.1007/s11075-023-01736-1)
- <span id="page-20-1"></span>13. S. B. Salihu, A. S. Halilu, M. Abdullahi, An improved spectral conjugate gradient projection method for monotone nonlinear equations with application, *J. Appl. Math. Comput.*, 70 (2024), 3879–3915. https://doi.org/10.1007/[s12190-024-02121-4](https://dx.doi.org/https://doi.org/10.1007/s12190-024-02121-4)
- <span id="page-20-2"></span>14. Y. Narushima, H. Yabe, J. A. Ford, A three-term conjugate gradient method with sufficient descent property for unconstrained optimization, *SIAM J. Optim.*, 21 (2011), 212–230. https://doi.org/10.1137/[080743573](https://dx.doi.org/https://doi.org/10.1137/080743573)
- <span id="page-20-3"></span>15. Y. Narushima, A smoothing conjugate gradient method for solving systems of nonsmooth equations, *Appl. Math. Comput.*, 219 (2013), 8646–8655. https://doi.org/10.1016/[j.amc.2013.02.060](https://dx.doi.org/https://doi.org/10.1016/j.amc.2013.02.060)
- <span id="page-20-4"></span>16. R. Huang, Y. Qin, K. Liu, G. Yuan, Biased stochastic conjugate gradient algorithm with adaptive step size for nonconvex problems, *Expert Syst. Appl.*, 238 (2024), 121556. https://doi.org/10.1016/[j.eswa.2023.121556](https://dx.doi.org/https://doi.org/10.1016/j.eswa.2023.121556)
- <span id="page-20-5"></span>17. X. Jiang, Y. Zhu, J. Jian, Two efficient nonlinear conjugate gradient methods with restart procedures and their applications in image restoration, *Nonlinear Dyn.*, 111 (2023), 5469–5498. https://doi.org/10.1007/[s11071-022-08013-1](https://dx.doi.org/https://doi.org/10.1007/s11071-022-08013-1)
- <span id="page-20-6"></span>18. W. Cheng, A PRP type method for systems of monotone equations, *Math. Comput. Model.*, 50 (2009), 15–20. https://doi.org/10.1016/[j.mcm.2009.04.007](https://dx.doi.org/https://doi.org/10.1016/j.mcm.2009.04.007)
- <span id="page-20-7"></span>19. G. Yu, A derivative-free method for solving large-scale nonlinear systems of equations, *J. Ind. Manag. Optim.*, 6 (2009), 149–160. https://doi.org/10.3934/[jimo.2010.6.149](https://dx.doi.org/https://doi.org/10.3934/jimo.2010.6.149)
- <span id="page-20-8"></span>20. M. Y. Waziri, K. Ahmed, J. Sabi'u, A family of Hager-Zhang conjugate gradient methods for system of monotone nonlinear equations, *Appl. Math. Comput.*, 361 (2019), 645–660. https://doi.org/10.1016/[j.amc.2019.06.012](https://dx.doi.org/https://doi.org/10.1016/j.amc.2019.06.012)
- <span id="page-20-9"></span>21. P. Liu, H. Shao, Z. Yuan, T. Zheng, A family of three-term conjugate gradient projection methods with a restart procedure and their relaxed-inertial extensions for the constrained nonlinear pseudo-monotone equations with applications, *Numer. Algor.*, 94 (2023), 1055–1083. https://doi.org/10.1007/[s11075-023-01527-8](https://dx.doi.org/https://doi.org/10.1007/s11075-023-01527-8)
- <span id="page-20-10"></span>22. A. Ibrahim, M. Alshahrani, S. Al-Homidan, Two classes of spectral three-term derivative-free method for solving nonlinear equations with application, *Numer. Algor.*, 96 (2024), 1625–1645. https://doi.org/10.1007/[s11075-023-01679-7](https://dx.doi.org/https://doi.org/10.1007/s11075-023-01679-7)
- <span id="page-21-0"></span>23. M. Rivaie, M. Mamat, L. W. June, I. Mohd, A new class of nonlinear conjugate gradient coefficients with global convergence properties, *Appl. Math. Comput.*, 218 (2012), 11323–11332. https://doi.org/10.1016/[j.amc.2012.05.030](https://dx.doi.org/https://doi.org/10.1016/j.amc.2012.05.030)
- <span id="page-21-1"></span>24. Z. Dai, Comments on a new class of nonlinear conjugate gradient coefficients with global convergence properties, *Appl. Math. Comput.*, 276 (2016), 297–300. https://doi.org/10.1016/[j.amc.2015.11.085](https://dx.doi.org/https://doi.org/10.1016/j.amc.2015.11.085)
- <span id="page-21-2"></span>25. A. B. Abubakar, P. Kumam, H. Mohammad, A modified Fletcher-Reeves conjugate gradient method for monotone nonlinear equations with some applications, *Mathematics*, 7 (2019), 745. https://doi.org/10.3390/[math7080745](https://dx.doi.org/https://doi.org/10.3390/math7080745)
- <span id="page-21-3"></span>26. J. Yin, J. Jian, X. Jiang, M. Liu, L. Wang, A hybrid three-term conjugate gradient projection method for constrained nonlinear monotone equations with applications, *Numer. Algor.*, 88 (2021), 389–418. https://doi.org/10.1007/[s11075-020-01043-z](https://dx.doi.org/https://doi.org/10.1007/s11075-020-01043-z)
- <span id="page-21-4"></span>27. E. D. Dolan, J. Jorge, Benchmarking optimization software with performance profiles, *Math. Program.*, 91 (2001), 201–213. https://doi.org/10.1007/[s101070100263](https://dx.doi.org/https://doi.org/10.1007/s101070100263)
- <span id="page-21-5"></span>28. D. Li, S. Wang, Y. Li, J. Wu, A convergence analysis of hybrid gradient projection algorithm for constrained nonlinear equations with applications in compressed sensing, *Numer. Algor.*, 95 (2024), 1325–1345. https://doi.org/10.1007/[s11075-023-01610-0](https://dx.doi.org/https://doi.org/10.1007/s11075-023-01610-0)



© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://[creativecommons.org](https://creativecommons.org/licenses/by/4.0)/licenses/by/4.0)