Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://journal.nsps.org.ng/index.php/jnsps

J. Nig. Soc. Phys. Sci. 8 (2026) 3051

Journal of the Nigerian Society of Physical Sciences

# A modified Dai–Yuan method with restart mechanism for nonlinear system of equations with a signal recovery

A. S. Halilu[a,b,c,d,*], M.A. Mohamed[a], I. A. R. Moghrabi[e,f], K. Ahmed[c,g,i], S. M. Ibrahim[h], M. Y. Waziri[c,g], S. Murtala[c,e,i], H. Abdullahi[b,d,g], M. A. Jada[d,g]

[a]*Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Campus Besut, 22200 Terengganu, Malaysia*
[b]*Department of Mathematics, Sule Lamido University Kafin Hausa, Nigeria*
[c]*Numerical Optimization Research Group, Bayero University, Kano, Nigeria*
[d]*Mathematical Innovation and Applications Research Group, Sule Lamido University Kafin Hausa, Nigeria*
[e]*Information Systems and Technology Department, Kuwait Technical College, Kuwait*
[f]*Department of Computer Science, College of Arts and Sciences, University of Central Asia, Naryn, Kyrgyz Republic*
[g]*Department of Mathematical Sciences, Bayero University, Kano, Nigeria*
[h]*College of Applied and Health Sciences, A'Sharqiyah University, 400 Ibra, Sultanate of Oman*
[i]*Department of Mathematics, Federal University, Dutse, Nigeria*

## Abstract

The Dai–Yuan method is an important iterative scheme for solving unconstrained optimization problems, but it often requires exact or Wolfe-type line searches to satisfy descent or sufficient descent conditions. It may also perform poorly due to the jamming phenomenon. This paper presents a modified three-term Dai–Yuan method for solving constrained systems of nonlinear monotone equations. The search direction of the proposed method includes a nonnegative parameter whose value is determined through singular-value analysis of the iteration matrix. The method also incorporates a restart mechanism that ensures global convergence regardless of the line-search procedure used. Theoretical analysis establishes the global convergence and convergence rate of the proposed scheme under suitable assumptions. Numerical experiments on constrained nonlinear equations and sparse signal reconstruction demonstrate the efficiency of the method compared with related algorithms.

DOI: 10.46481/jnsps.2026.3051

*Keywords:* Conjugate gradient methods, Lipschitz condition, Signal reconstruction, Singular values

*Article History:*
Received: 14 July 2025
Received in revised form: 06 March 2026
Accepted for publication: 16 March 2026
Available online: 09 June 2026



Communicated by: P. Thakur

## 1. Introduction

This paper is dedicated to constrained system of nonlinear equations and its application in sparse signal reconstruction. Generally, the constrained systems of nonlinear equations is formulated as

$$F(\bar{x}) = 0, \quad \bar{x} \in C \subseteq \mathbb{R}^n, \tag{1}$$

in which $C$ is a nonempty, closed convex set and $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous and monotone mapping. This implies that $F$ satisfies the inequality

$$(F(x) - F(y))^T (x - y) \geq 0, \quad \forall x, y \in C. \tag{2}$$

Some applications of equation (1) are found in the general equilibrium problems [1, 2] as well as in compressed sensing [1–3], where sparse signals are reconstructed and images are deblurred. Since the focus of this work is on large-scale prob-

*Corresponding author Tel. No.: +234-803-614-3352
*Email address:* abubakars.halilu@slu.edu.ng (A. S. Halilu)

lems in the form of equation (1), we require an iterative method that avoids computing and storing matrices at each iteration. Recent developments have shown that the conjugate gradient (CG) method, which requires less memory to implement, is the method that fits the bill. Primarily, CG method is used to solve the large-scale optimization problem

$$\min_{x\in\mathbb{R}^n} f(x), \tag{3}$$

in which $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function with gradient at $x_k$ given as $\nabla f(x_k) = g(x_k)$. Starting with a guess $x_0 \in \mathbb{R}^n$, the formula for generating the method's iterates is given by

$$x_{k+1} = x_k + s_k, \quad s_k = \vartheta_k d_k, \quad k = 0, 1, ..., \tag{4}$$

where $x_k$ represents the $k^{th}$ iterate, $\vartheta_k > 0$ is a step-size often obtained using a line search procedure along the method's direction $d_k$ which is defined by

$$d_0 = -g_0, \quad d_{k+1} = -g_{k+1} + \beta_k d_k, \quad k \geq 0. \tag{5}$$

In equation (5), $g_{k+1} = g(x_{k+1})$ and $\beta_k$ is the CG (update) parameter that is crucial in the method's implementation. For different CG methods with different choices of $\beta_k$, see the comprehensive survey in [4, 5]. The earlier or classical $\beta_k$ parameters are defined as follows:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}[6], \ \beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k}[7], \ \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T(g_{k+1} - g_k)}[8], \tag{6}$$

$$\beta_k^{HS} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{d_k^T(g_{k+1} - g_k)}\ [9],$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{\|g_k\|^2}\ [10, 11], \tag{7}$$

$$\beta_k^{LS} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{-d_k^T g_k}\ [12].$$

where $\|.\|$ is the $\ell_2$–norm and $g_k = g(x_k)$. For global convergence of a CG method that is implemented via equation (4) and equation (5), it is expected to satisfy the inequality

$$d_{k+1}^T g_{k+1} \leq -\vartheta\|g_{k+1}\|^2, \quad \vartheta > 0. \tag{8}$$

As with the other methods in equation (6), the DY scheme with $\beta_k^{DY}$ suffers from the jamming phenomenon, where small steps are taken without much progress to the minimum of $f$ in equation (8). Also, as with other methods in equation (6) and the other group of the classical CG methods in [9–12], the DY scheme does not automatically satisfy equation (8). To address these shortcomings, modified variants of the DY method for solving equation (3) have emerged in recent years. Andrei [13] proposed a scaled DY-type method for the problem in equation (3) with sufficient descent and conjugacy conditions. Andrei [14] also proposed a hybrid method for equation (3), where the update parameter is obtained as a convex combination of the HS and DY parameters. Based on ideas by Li and Fukushima [15], Wei et al. [16], and Zhang et al. [17], two DY-type algorithms were presented by Zhang [18] for solving equation (3). One of the methods converges globally for nonconvex functions, while the second method exhibits the good performance properties it inherits from the HS method. For other DY-type algorithms, see [19–26] and the references therein.

Owing to the useful properties of CG methods, researchers dealing with high-dimensional problems have proposed CG-type methods for solving the constrained problem in equation (1), where the direction is defined as

$$d_0 = -F_0, \quad d_{k+1} = -F_{k+1} + \beta_k d_k, \quad F_{k+1} = F(x_{k+1}), \quad k = 0, 1..., $$

where $\beta_k$ corresponds to a modified version of any of the classical ones in equation (6) and equation (7).

Recently, some DY-type methods for solving problem (1) have been developed. These include the method by Liu and Li [27], where the algorithm was obtained by combining the classical DY parameter [8], the spectral gradient method [28] and the method proposed in [29]. The authors also proved convergence of the method without the differentiability assumption. Also, in [30], Liu and Li combined the method in [31] with the DY method [32] to present a spectral DY-type method for the problem in equation (1). Motivated by the work in Refs. [27, 32], Liu and Feng [33] presented a DY-type method for solving problem (1). The method's derivative-free structure makes it ideal for nonsmooth problems. By employing the Lipschitz continuity assumption, the authors proved global convergence of the new method. Inspired by the work in Ref. [33], Sani et al. [34] proposed a DY-type algorithm for solving equation (1), where the method's search direction was obtained as a convex combination of the unmodified DY and CD parameters. Only recently, Alhobaiti et al. [35] proposed two scaled DY-type algorithms for solving equation (1), where two different approaches were applied to compute the scaling parameter. The authors also showed that both methods satisfy the inequality in equation (8). The reader can explore Refs. [1–3, 36–45] for other works in the literature, and also view the recent works in [46–48] to have a broader comprehension of the underlined concept outside the literature discussed above.

We outline the objectives of the work as follows:

- To construct a DY-type algorithm for solving the constrained problem in equation (1) and add to the few that already exist in the literature.
- To present a method with the vital property for analyzing the convergence of CG-type methods for the constrained problem in equation (1).
- To analyze global convergence and convergence rate of the new method.
- To analyze numerical performance of the method in solving problem (1).
- To describe the method's application in sparse signal recovery.

The paper is structured as follows: Motivation and details of the method are given in the following section. Analysis of its convergence and rate of convergence are given in Section 3. Results of some numerical experiments as well as application of the method are presented in Section 4. Conclusions are given in Section 5.

## 2. Motivation and algorithm of the method

As known theoretically, the vital property expressed in equation (8) holds for all the classical CG methods when $f$ in equation (3) is convex quadratic and $\vartheta_k$ in equation (4) is computed exactly. In actual computations, however, where inexact line searches are used, equation (8) does not hold in general for these methods. For example, by substituting the HS parameter in equation (5) and multiplying through by $g_{k+1}^T$, we obtain

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \beta_k^{HS} g_{k+1}^T d_k. \tag{9}$$

Clearly, equation (9) satisfies equation (8) if $\beta_k^{HS} \geq 0$ and $g_{k+1}^T d_k \leq 0$. On the other hand, if $\beta_k^{HS} \geq 0$ and $g_{k+1}^T d_k > 0$ the condition may not hold as the quantity $\beta_k^{HS} g_{k+1}^T d_k$ may become larger than $-\|g_{k+1}\|^2$. To remedy this defect of the HS method, Dong et al. [49] made some modifications to the method that not only satisfies equation (8), but also inherits the nice attributes of the method. The authors proposed the following modified HS search direction:

$$d_{k+1}^D = \begin{cases} -\lambda_{k+1} g_{k+1} + \beta_k^D d_k, & g_{k+1}^T d_k > 0, \quad k \geq 0; \\ -g_{k+1} + \beta_k^{HS} d_k, & g_{k+1}^T d_k \leq 0, \quad k \geq 0, \text{ otherwise.} \end{cases} \tag{10}$$

In equation (10), $\lambda_{k+1}$ is given by

$$\lambda_{k+1} = 1 + \frac{g_{k+1}^T d_k}{d_k^T y_k} \cdot \frac{g_{k+1}^T y_k}{\|g_{k+1}\|^2}, \quad \beta_k^D = \max\{\beta_k^{DHS}, \eta_k\},$$

where

$$\eta_k = \frac{-1}{\|d_k\| \min\{\eta, \|g_k\|\}},$$

and

$$\beta_k^{DHS} = \left(1 - \frac{g_k^T d_k}{d_k^T y_k}\right)\beta_k^{HS} - t\frac{\|y_k\|^2 g_{k+1}^T d_k}{d_k^T y_k}, \quad t > 0.$$

Motivated by this approach, Aminifard and Babaie-Kafaki [50] made an almost similar modification to the classical PRP method, which is also a method with the same built-in mechanism as the HS method but also fails to satisfy the condition in equation (8) when inexact line searches are employed in general as demonstrated in the case of the HS method. The search direction put forward by the authors in Ref. [50] for which equation (8) holds, while retaining nice attributes of the unmodified PRP method is defined as follows:

$$d_0 = -g_0,$$
$$d_{k+1}^M = \begin{cases} -\lambda_{k+1} g_{k+1} + \beta_k^{MPRP} d_k, & \text{if } g_{k+1}^T d_k > 0, \quad k \geq 0, \\ -g_{k+1} + \beta_k^{PRP} d_k, & \text{if } g_{k+1}^T d_k \leq 0, \quad k \geq 0. \end{cases} \tag{11}$$

where

$$\lambda_{k+1} = 1 + \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}\beta_k^{PRP},$$

and

$$\beta_k^{MPRP} = \left(1 - \frac{g_{k+1}^T s_k}{\|g_k\|^2}\right)\beta_k^{PRP} - t\frac{\|y_k\|^2 g_{k+1}^T s_k}{\|g_k\|^4}, \quad t \geq 0.$$

Inspired by equation (10), equation (11), and the classical DY method, we propose the following DY-type search direction:

$$d_0 = -F_0, \quad d_{k+1} = \begin{cases} -F_{k+1} + \beta_k^{DY1} d_k, & F_{k+1}^T d_k > 0, \quad k \geq 0; \\ -F_{k+1} + \beta_k^{DY2} d_k, & F_{k+1}^T d_k \leq 0, \quad k \geq 0, \\ \text{otherwise,} \end{cases} \tag{12}$$

in which

$$\beta_k^{DY1} = \Gamma_k \beta_k^{DY2} - t_k \frac{\|F_{k+1}\|^2 F_{k+1}^T d_k}{(d_k^T w_k)^2}, \quad k = 0, 1, ..., \tag{13}$$

$$\beta_k^{DY2} = \frac{\|F_{k+1}\|^2}{d_k^T w_k}, \tag{14}$$

with $t_k$ being a nonnegative parameter which is to be determined and

$$w_k = y_k + r\frac{\|F_{k+1}\| s_k}{\|s_k\|}, \quad y_k = F(\psi_k) - F(x_k), \quad r > 1, \tag{15}$$

and

$$\psi_k = x_k + \vartheta_k d_k.$$

From equation (15) and monotonicity of $F$, we obtain

$$d_k^T w_k = \frac{s_k^T y_k}{\vartheta_k} + \frac{r}{\vartheta_k}\frac{\|F_{k+1}\|}{\|s_k\|}\|s_k\|^2 \geq \frac{r}{\vartheta_k}\|F_{k+1}\|\|s_k\| > 0.$$

Consequently, we get

$$s_k^T w_k = s_k^T y_k + r\|F_{k+1}\|\|s_k\| \geq r\|F_{k+1}\|\|s_k\| > 0. \tag{16}$$

To find suitable approximations for $t_k$, we conduct singular value analysis of the matrix associated with the first case of the directions in equation (12), where $F_{k+1}^T d_k > 0$. Observe that the direction in the first case of equation (12) can be written as

$$d_{k+1} = -M_{k+1} F_{k+1}, \quad k = 0, 1, ...,$$

in which $M_{k+1}$ is known as the matrix associated with the first case of the search direction in equation (12) and is given as

$$M_{k+1} = I - \Gamma_k \frac{d_k F_{k+1}^T}{d_k^T w_k} + t_k \frac{\|F_{k+1}\|^2 d_k d_k^T}{(d_k^T w_k)^2}.$$

Now, from the first case in equation (12), $F_{k+1}^T d_k > 0$. So, we have that $F_{k+1}$ and $d_k$ are nonzero vectors. Hence, there exists mutually orthonormal vectors $\{\tau_k^i\}_{i=1}^{n-2} \subset \mathcal{V}^\perp$ satisfying

$$F_{k+1}^T \tau_k^i = d_k^T \tau_k^i = 0, \quad i = 1, ..., n-2,$$

which ultimately yields

$$M_{k+1}\tau_k^i = M_{k+1}^T\tau_k^i = \tau_k^i, \quad i = 1, ..., n-2.$$

Hence, $M_{k+1}$ contains $n-2$ singular values with multiplicity 1 each. Let $\sigma_k^+$ and $\sigma_k^-$ represent the two singular values left. Considering that the matrix $M_{k+1}$ represents a rank-two update, its determinant can be computed by employing the algebra formula in [51], i.e.,

$$\det(I + b_1b_2^T + b_3b_4^T) = (1 + b_1^Tb_2)(1 + b_3^Tb_4) - (b_1^Tb_4)(b_2^Tb_3),$$

and setting $b_1 = \frac{-\Gamma_k d_k}{d_k^T w_k}$, $b_2 = F_{k+1}$, $b_3 = t_k\frac{\|F_{k+1}\|^2 d_k}{(d_k^T w_k)^2}$, $b_4 = d_k$, we get

$$\det(M_{k+1}) = 1 + t_k\frac{\|F_{k+1}\|^2\|d_k\|^2}{(d_k^T w_k)^2} - \frac{\Gamma_k F_{k+1}^T d_k}{d_k^T w_k}, \tag{17}$$

which is clearly greater than zero for $t_k \geq \Gamma_k$. Now by employing the property of Frobenius norm, we have

$$\begin{aligned}\|M_{k+1}\|_F^2 &= \operatorname{tr}(M_{k+1}^T M_{k+1})\\ &= n - \frac{2\Gamma_k d_k^T F_{k+1}}{d_k^T w_k} + \frac{2t_k\|F_{k+1}\|^2\|d_k\|^2}{(d_k^T w_k)^2}\\ &\quad - \frac{2\Gamma_k t_k\|F_{k+1}\|^2\|d_k\|^2 F_{k+1}^T d_k}{(d_k^T w_k)^3} + \frac{\Gamma_k^2\|F_{k+1}\|^2\|d_k\|^2}{(d_k^T w_k)^2}\\ &\quad + t_k^2\frac{\|F_{k+1}\|^4\|d_k\|^4}{(d_k^T w_k)^4}\\ &= \underbrace{1 + \cdots + 1}_{(n\text{-}2)\text{ times}} + (\sigma_k^+)^2 + (\sigma_k^-)^2.\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}(\sigma_k^+)^2 + (\sigma_k^-)^2 &= 2 - \frac{2\Gamma_k d_k^T F_{k+1}}{d_k^T w_k} + \frac{2t_k\|F_{k+1}\|^2\|d_k\|^2}{(d_k^T w_k)^2}\\ &\quad - \frac{2\Gamma_k t_k\|F_{k+1}\|^2\|d_k\|^2 F_{k+1}^T d_k}{(d_k^T w_k)^3} + \frac{\Gamma_k^2\|F_{k+1}\|^2\|d_k\|^2}{(d_k^T w_k)^2}\\ &\quad + t_k^2\frac{\|F_{k+1}\|^4\|d_k\|^4}{(d_k^T w_k)^4}.\end{aligned} \tag{18}$$

By employing equation (17) and equation (18), and setting

$$\chi_k = \frac{\|F_{k+1}\|^2\|d_k\|^2}{(d_k^T w_k)^2}, \quad \varpi_k = \frac{F_{k+1}^T d_k}{d_k^T w_k}$$

with some algebraic simplifications, $\sigma_k^+$ and $\sigma_k^-$ are obtained as

$$\begin{aligned}\sigma_k^{\pm} &= \frac{1}{2}\sqrt{(t_k\chi_k - \Gamma_k\varpi_k + 2)^2 + \Gamma_k^2\chi_k - \Gamma_k^2\varpi_k^2}\\ &\quad \pm \sqrt{(t_k\chi_k - \Gamma_k\varpi_k)^2 + \Gamma_k^2\chi_k - \Gamma_k^2\varpi_k^2},\end{aligned} \tag{19}$$

which after some algebraic simplifies satisfies the inequality $0 < \sigma_k^- \leq 1 \leq \sigma_k^+$.

It has been stated in Ref. [52] that a matrix condition number is a vital factor in sensitivity analysis of numerical computations of the matrix. Let $Q$ be an arbitrary nonsingular matrix. The condition number associated with $Q$ is given as

$$\kappa(Q) = \|Q\|\|Q^{-1}\|. \tag{20}$$

It was equally noted in Ref. [52] that the larger $\kappa(Q)$ gets, the closer or more likely is $Q$ to being ill-conditioned. The matrix $Q$ is said to be well-conditioned if $\kappa(Q)$ is not too large. The most preferred condition number is 1. Also, note that the value of $\kappa(Q)$ is dependent on the type of norm used to compute it. Apart from equation (20), the condition number can also be obtained by employing the smallest $\sigma_n$ and largest $\sigma_1$ singular values of the matrix $Q$, namely

$$\kappa_2(Q) = \frac{\sigma_1}{\sigma_n}. \tag{21}$$

The condition number defined in equation (21) is known as the spectral condition number. Now, considering equation (19) and equation (21), we see that as the distance between $\sigma_k^+$ and $\sigma_k^-$ increases, the larger the value of $\kappa_2(Q)$ gets. So, to obtain a suitable value for the parameter $t_k$ which will improve the numerical stability of the proposed method, we obtain it as the solution of the minimization problem

$$\min_t(\sigma_k^+ - \sigma_k^-), \tag{22}$$

so that $\sigma_k^-$ is made as close as possible to $\sigma_k^+$, and ensuring that $\kappa_2(M_{k+1})$ gets as close to 1 as possible. By conducting some algebraic simplifications and setting $\Gamma_k = \frac{F_{k+1}^T d_k}{d_k^T w_k}$ we obtain

$$t_k = \frac{(F_{k+1}^T d_k)^2}{\|F_{k+1}\|^2\|d_k\|^2}, \tag{23}$$

as the minimizer of equation (22), which we take as the optimal value of $t_k$. In order to ensure that $t_k \geq \Gamma_k$ as required by equation (17), we present the revised form inside a scale-adjusted constraint container:

$$t_k^* = \begin{cases} t_k, & \text{if } t_k \geq \frac{F_{k+1}^T d_k}{d_k^T w_k},\\ \frac{\|F_{k+1}\|\|d_k\|}{d_k^T w_k}, & \text{if } t_k < \frac{F_{k+1}^T d_k}{d_k^T w_k}.\end{cases} \tag{24}$$

**Lemma 1.** *Given that $C$ is as defined in equation (1), then the projection operator $P_C[x]$ is defined as*

$$P_C(x) = \arg\min\{\|x - y\| : y \in C\}, \quad \forall x \in \mathbb{R}^n.$$

*Also, $P_C[\cdot]$ is nonexpansive, namely, it satisfies the inequality*

$$\|P_C(x) - P_C(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$

*Also,*

$$\|P_C(x) - y\| \leq \|x - y\|, \quad \forall y \in C. \tag{25}$$

We describe the algorithm as follows:

**Algorithm 1**

**Data**: Set $\epsilon > 0$, $x_0 \in C$, $\beta \in (0, 1)$, $\bar{\zeta} \in (0, 1)$, $\delta \in (0, 1)$, $0 < \phi < 2$, $r > 1$.

**Initialization**: Set $k = 0$ and $d_0 = -F_0$.

**Step 1**: Compute $F(x_k)$ and test if $\|F(x_k)\| \leq \epsilon$. If yes, stop, otherwise move to Step 2.

**Step 2**: Determine $\psi_k = x_k + \vartheta_k d_k$, where $\vartheta_k = \overline{\zeta}\beta^{m_k}$, and $m_k$ is the smallest nonnegative integer satisfying

$$-F(x_k + \vartheta_k d_k)^T d_k \geq \delta\vartheta_k \|F(\psi_k)\|\|d_k\|^2. \tag{26}$$

**Step 3**: If $\psi_k \in C$ and $\|F(\psi_k)\| \leq \epsilon$, end the process, else obtain

$$x_{k+1} = P_C\left[x_k - \phi\zeta_k F(\psi_k)\right], \quad \text{where} \tag{27}$$

$$\zeta_k = \frac{F(\psi_k)^T(x_k - \psi_k)}{\|F(\psi_k)\|^2}. \tag{28}$$

**Step 4**: Determine $d_{k+1}$ by equation (12) with equations (13), (14), (15), and (24).

**Step 5**: Set $k = k + 1$ and move back to **Step 1**.

## 3. Convergence results

**Assumption 1**. The solution set, say $\overline{C}$ of equation (1) is not empty, namely there exists $\bar{x} \in C$ such that $F(\bar{x}) = 0$.

**Assumption 2**. Given $x, y \in C$, there exists a positive constant $L$ such that

$$\|F(x) - F(y)\| \leq L\|x - y\|. \tag{29}$$

**Lemma 2.** *The search directions determined in Step 4 by Algorithm 1 satisfy the inequality*

$$d_{k+1}^T F_{k+1} \leq -\gamma\|F_{k+1}\|^2, \quad \gamma = \left(1 - \frac{1}{r^2}\right), \tag{30}$$

*and the trust region property*

$$\begin{aligned}\left(1 - \frac{1}{r^2}\right)\|F_{k+1}\| &\leq \|d_{k+1}\| \\ &\leq \max\left\{1, \left(1 + \frac{2}{r^2}\right), \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right), \right. \\ &\quad \left.\left(1 + \frac{1}{r}\right)\right\}\|F_{k+1}\|.\end{aligned} \tag{31}$$

*Proof.* Notice from equation (12) that for $k = 0$, $d_0^T F_0 = -F_0^T(F_0) = -\|F_0\|^2$. Next, we prove the result for $k = 1, 2, \ldots$, in the following manner:

**Case (1)**. $F_{k+1}^T d_k > 0$, $t_k^* = t_k$ and $\Gamma_k = \frac{F_{k+1}^T d_k}{d_k^T w_k}$. Multiplying through equation (12) by $F_{k+1}$ and applying the Cauchy-Schwarz inequality, we have

$$\begin{aligned} d_{k+1}^T F_{k+1} &= -\|F_{k+1}\|^2 \\ &\quad + \Gamma_k \frac{\|F_{k+1}\|^2 F_{k+1}^T d_k}{d_k^T w_k} \\ &\quad - t_k^* \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)^2}{(d_k^T w_k)^2} \\ &= -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)^2}{(d_k^T w_k)^2} \\ &\quad - \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)^4}{\|F_{k+1}\|^2\|d_k\|^2 (d_k^T w_k)^2} \\ &= -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)^2}{(d_k^T w_k)^2} \\ &\quad - \frac{(F_{k+1}^T d_k)^4}{\|d_k\|^2 (d_k^T w_k)^2} \\ &\leq -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)^2}{(d_k^T w_k)^2} \\ &\leq -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^4 \|d_k\|^2}{r^2\|F_{k+1}\|^2\|d_k\|^2} \\ &= -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^2}{r^2} \\ &= -\left(1 - \frac{1}{r^2}\right)\|F_{k+1}\|^2. \end{aligned}$$

For $F_{k+1}^T d_k > 0$, with $t_k^* = \frac{\|F_{k+1}\|\|d_k\|}{d_k^T w_k}$ and $\Gamma_k = \frac{F_{k+1}^T d_k}{d_k^T w_k}$, we have:

$$\begin{aligned} d_{k+1}^T F_{k+1} &= -\|F_{k+1}\|^2 \\ &\quad + \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)^2}{(d_k^T w_k)^2} \\ &\quad - \frac{\|F_{k+1}\|^3 \|d_k\| (F_{k+1}^T d_k)^2}{(d_k^T w_k)^3} \\ &\leq -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)^2}{(d_k^T w_k)^2} \\ &\leq -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^4 \|d_k\|^2}{r^2\|F_{k+1}\|^2\|d_k\|^2} \\ &= -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^2}{r^2} \\ &= -\left(1 - \frac{1}{r^2}\right)\|F_{k+1}\|^2. \end{aligned}$$

**Case (2)**. $F_{k+1}^T d_k \leq 0$. Since $\|F_{k+1}\|^2$ and $d_k^T w_k$ are both greater than zero, taking the inner product of the direction in the second

case with $F_{k+1}$ yields

$$d_{k+1}^T F_{k+1} = -\|F_{k+1}\|^2 + \frac{\|F_{k+1}\|^2 F_{k+1}^T d_k}{d_k^T w_k} \leq -\|F_{k+1}\|^2.$$

Setting $\gamma = 1 - \frac{1}{r^2}$, which is $\min\left\{1, 1 - \frac{1}{r^2}\right\}$, we established the inequality in equation (30).

Next, from equation (30) and applying the Cauchy-Schwarz inequality, the first part of equation (31) holds. Now, from equation (12), $d_0 = -F_0$, which implies that $\|d_0\| = \|F_0\|$ for $k = 0$. For $k = 1, 2, \ldots$, we proceed as follows:

Suppose $t_k^* = t_k$, then from equation (23) and equation (24), we have

$$|t_k^*| = \frac{(F_{k+1}^T d_k)^2}{\|F_{k+1}\|^2\|d_k\|^2} \leq \frac{\|F_{k+1}\|^2\|d_k\|^2}{\|F_{k+1}\|^2\|d_k\|^2} = 1. \tag{32}$$

If $t_k^* = \frac{\|F_{k+1}\|\|d_k\|}{d_k^T w_k}$, then by equation (24), we obtain

$$|t_k^*| = \frac{\|F_{k+1}\|\|d_k\|}{d_k^T w_k} \leq \frac{\|F_{k+1}\|\|d_k\|}{r\|F_{k+1}\|\|d_k\|} = \frac{1}{r}. \tag{33}$$

So, if $F_{k+1}^T d_k > 0$, employing equation (12), equation (16), and equation (32), yields

$$\begin{aligned}
\|d_{k+1}\| &= \| - F_{k+1} + \Gamma_k \frac{\|F_{k+1}\|^2}{s_k^T w_k} s_k - t_k^* \frac{\|F_{k+1}\|^2 F_{k+1}^T s_k}{(s_k^T w_k)^2} s_k\| \\
&\leq \|F_{k+1}\| + \frac{\|F_{k+1}\|^3\|s_k\|^2}{(s_k^T w_k)^2} + |t_k^*|\frac{\|F_{k+1}\|^3\|s_k\|^2}{(s_k^T w_k)^2} \\
&\leq \|F_{k+1}\| + \frac{\|F_{k+1}\|^3\|s_k\|^2}{r^2\|F_{k+1}\|^2\|s_k\|^2} + \frac{\|F_{k+1}\|^3\|s_k\|^2}{r^2\|F_{k+1}\|^2\|s_k\|^2} \\
&\leq \|F_{k+1}\| + \frac{\|F_{k+1}\|}{r^2} + \frac{\|F_{k+1}\|}{r^2} \\
&\leq \left(1 + \frac{1}{r^2} + \frac{1}{r^2}\right)\|F_{k+1}\| \\
&\leq \left(1 + \frac{2}{r^2}\right)\|F_{k+1}\|.
\end{aligned} \tag{34}$$

Also, if $F_{k+1}^T d_k > 0$, utilizing equation (12), equation (16), equation (33), and the Cauchy-Schwarz inequality, we get

$$\begin{aligned}
\|d_{k+1}\| &= \| - F_{k+1} + \Gamma_k \frac{\|F_{k+1}\|^2}{s_k^T w_k} s_k - t_k^* \frac{\|F_{k+1}\|^2 F_{k+1}^T s_k}{(s_k^T w_k)^2} s_k\| \\
&\leq \|F_{k+1}\| + \frac{\|F_{k+1}\|^3\|s_k\|^2}{(s_k^T w_k)^2} + |t_k^*|\frac{\|F_{k+1}\|^3\|s_k\|^2}{(s_k^T w_k)^2} \\
&\leq \|F_{k+1}\| + \frac{\|F_{k+1}\|^3\|s_k\|^2}{r^2\|F_{k+1}\|^2\|s_k\|^2} + \frac{\|F_{k+1}\|^3\|s_k\|^2}{r^3\|F_{k+1}\|^2\|s_k\|^2} \\
&= \|F_{k+1}\| + \frac{\|F_{k+1}\|}{r^2} + \frac{\|F_{k+1}\|}{r^3} \\
&= \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right)\|F_{k+1}\|.
\end{aligned} \tag{35}$$

Now, if $F_{k+1}^T d_k \leq 0$. From equation (12) and equation (16), we have

$$\begin{aligned}
\|d_{k+1}\| &= \| - F_{k+1} + \frac{\|F_{k+1}\|^2}{s_k^T w_k} s_k\| \\
&\leq \|F_{k+1}\| + \frac{\|F_{k+1}\|^2\|s_k\|}{r\|F_{k+1}\|\|s_k\|} \\
&= \|F_{k+1}\| + \frac{\|F_{k+1}\|}{r} \\
&= \left(1 + \frac{1}{r}\right)\|F_{k+1}\|.
\end{aligned} \tag{36}$$

Considering all the cases analyzed in equations (34), (35), and (36), and noting that $r > 1$, we obtain that

$$\|d_{k+1}\| \leq \max\left\{1, \left(1 + \frac{2}{r^2}\right), \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right), \left(1 + \frac{1}{r}\right)\right\}\|F_{k+1}\| \tag{37}$$

which establishes the second inequality in equation (31). □

**Lemma 3.** *Given that equation (29) holds with $\bar{x}$ being a solution of equation (1) in $\overline{C}$. Then, for $\phi \in (0, 2)$, the sequence $\{\|x_k - \bar{x}\|\}$ is convergent and*

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_k - \bar{x}\|^2 - \phi(2 - \phi)\delta^2\|x_k - \psi_k\|^4,$$

*holds with $\{x_k\}$ and $\{\psi_k\}$ bounded.*

*Proof.* From equation (26) and the definition of $\psi_k$, we have

$$(x_k - \psi_k)^T F(\psi_k) \geq \delta\|F(\psi_k)\|\|x_k - \psi_k\|^2. \tag{38}$$

By equation (2) and for all $\bar{x} \in \overline{C}$, we have

$$\begin{aligned}
(x_k - \bar{x})^T F(\psi_k) &= (x_k - \psi_k)^T F(\psi_k) + (\psi_k - \bar{x})^T F(\psi_k) \\
&\geq (x_k - \psi_k)^T F(\psi_k) + (\psi_k - \bar{x})^T F(\bar{x}) \\
&= (x_k - \psi_k)^T F(\psi_k).
\end{aligned} \tag{39}$$

From equation (25), equation (27), equation (28), equation (38), and equation (39), we have

$$\begin{aligned}
\|x_{k+1} - \bar{x}\|^2 &= \|P_C[x_k - \phi\zeta_k F(\psi_k)] - \bar{x}\|^2 \\
&\leq \|x_k - \phi\zeta_k F(\psi_k) - \bar{x}\|^2 \\
&= \|(x_k - \bar{x}) - \phi\zeta_k F(\psi_k)\|^2 \\
&= \|x_k - \bar{x}\|^2 - 2\phi\zeta_k F(\psi_k)^T(x_k - \bar{x}) + \phi^2\zeta_k^2\|F(\psi_k)\|^2 \\
&\leq \|x_k - \bar{x}\|^2 - 2\phi\zeta_k F(\psi_k)^T(x_k - \psi_k) + \phi^2\zeta_k^2\|F(\psi_k)\|^2 \\
&= \|x_k - \bar{x}\|^2 - \phi(2 - \phi)\frac{(F(\psi_k)^T(x_k - \psi_k))^2}{\|F(\psi_k)\|^2} \\
&\leq \|x_k - \bar{x}\|^2 - \phi(2 - \phi)\frac{\delta^2\|F(\psi_k)\|^2\|x_k - \psi_k\|^4}{\|F(\psi_k)\|^2} \\
&= \|x_k - \bar{x}\|^2 - \phi(2 - \phi)\delta^2\|x_k - \psi_k\|^4.
\end{aligned} \tag{40}$$

Hence, the sequence $\{\|x_k - \bar{x}\|\}$ is decreasing, bounded and shows that $\{x_k\}$ is also bounded. Now, by applying equation (29) and the fact that $\|x_{k+1} - \bar{x}\| \leq \|x_k - \bar{x}\|$, $\forall k \geq 0$, we have

$$\|F(x_k)\| = \|F(x_k) - F(\bar{x})\| \leq L\|x_k - \bar{x}\| \leq L\|x_0 - \bar{x}\|. \tag{41}$$

Setting $L\|x_0 - \bar{x}\| = \bar{\kappa}$, shows that $\{F_k\}$ is bounded. Combining this result with equation (37) shows that there exists $M > 0$ for which

$$\|d_k\| \leq \max\left\{1, \left(1 + \frac{2}{r^2}\right), \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right), \left(1 + \frac{1}{r}\right)\right\} \|F_k\|$$
$$\leq \max\left\{1, \left(1 + \frac{2}{r^2}\right), \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right), \left(1 + \frac{1}{r}\right)\right\} M.$$

Setting $\max\left\{1, \left(1 + \frac{2}{r^2}\right), \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right), \left(1 + \frac{1}{r}\right)\right\} M = \overline{M}$ shows that the sequence of directions $\{d_k\}$ is bounded. Also, from the inequality (38) and the Cauchy-Schwarz inequality, we have

$$\delta\|F(\psi_k)\|\|x_k - \psi_k\|^2 \leq (x_k - \psi_k)^T F(\psi_k)$$
$$\leq \|F(\psi_k)\|\|x_k - \psi_k\|,$$

which further yields

$$\delta\|x_k - \psi_k\| \leq 1,$$

implying that $\{\psi_k\}$ is bounded. Therefore, by continuity of $F$, there exists a constant $\overline{m}$ such that

$$\|F(\psi_k)\| \leq \overline{m}. \tag{42}$$

□

**Lemma 4.** *Let $\{x_k\}$ be generated by Algorithm 1. Then*

$$\lim_{k\to\infty} \vartheta_k \|d_k\| = 0. \tag{43}$$

*Proof.* From equation (40) and the fact that both $\{x_k\}$ and $\{\psi_k\}$ are bounded, we obtain

$$\phi(2-\phi)\delta^2 \sum_{k=0}^{\infty} \|x_k - \psi_k\|^4 \leq \sum_{k=0}^{\infty} (\|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2) < \infty,$$

which indicates that

$$\lim_{k\to\infty} \|x_k - \psi_k\| = \lim_{k\to\infty} \vartheta_k \|d_k\| = 0.$$

□

In the following Lemma, we prove that the line search procedure defined in equation (26) is well-defined with a uniform lower bound for the sequence of step-sizes $\{\vartheta_k\}$.

**Lemma 5.** *(1) Let $\{d_k\}$ and $\{x_k\}$ be determined by Algorithm 1. If $F$ is continuous on $\mathbb{R}^n$, then, there exists $m_k \geq 0$ satisfying equation (26).*

*(2) Given that Assumption 2 holds with $\{x_k\}$ and $\{\psi_k\}$ being generated sequences by Algorithm 1. Then*

$$\vartheta_k \geq \min\left\{\overline{\zeta}, \frac{\beta\left(1-\frac{1}{r^2}\right)}{(L+\delta\overline{m})\left(\max\left\{1, \left(1+\frac{2}{r^2}\right), \left(1+\frac{1}{r^2}+\frac{1}{r^3}\right), \left(1+\frac{1}{r}\right)\right\}\right)^2}\right\}. \tag{44}$$

*Proof.* The first part has been proven in Ref. [41] and is left out here.

For (2), considering equation (26), if $\vartheta_k = \overline{\zeta}$, the inequality (26) holds, else $\overline{\vartheta}_k = \beta^{-1}\vartheta_k$ does not satisfy equation (26), i.e.,

$$-F(\overline{\psi}_k)^T d_k < \delta\overline{\vartheta}_k \|F(\overline{\psi}_k)\|\|d_k\|^2.$$

Now, by equation (29), we have

$$\left(1 - \frac{1}{r^2}\right)\|F_k\|^2 \leq -F_k^T d_k = (F(\overline{\psi}_k) - F_k)^T d_k - F(\overline{\psi}_k)^T d_k$$
$$\leq \overline{\vartheta}_k (L + \delta\|F(\overline{\psi}_k)\|)\|d_k\|^2$$
$$\leq \beta^{-1}\vartheta_k (L + \delta\overline{m})\|d_k\|^2.$$

Consequently, we get

$$\vartheta_k \geq \frac{\beta\left(1 - \frac{1}{r^2}\right)\|F_k\|^2}{(L + \delta\overline{m})\|d_k\|^2}$$
$$\geq \frac{\beta\left(1 - \frac{1}{r^2}\right)\|F_k\|^2}{(L + \delta\overline{m})\left(\max\left\{1, \left(1 + \frac{2}{r^2}\right), \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right), \left(1 + \frac{1}{r}\right)\right\}\right)^2 \|F_k\|^2}$$
$$= \frac{\beta\left(1 - \frac{1}{r^2}\right)}{(L + \delta\overline{m})\left(\max\left\{1, \left(1 + \frac{2}{r^2}\right), \left(1 + \frac{1}{r^2} + \frac{1}{r^3}\right), \left(1 + \frac{1}{r}\right)\right\}\right)^2}$$

to establish the proof of equation (44). □

**Theorem 1.** *The sequence $\{x_k\}$ generated by Algorithm 1 converges globally to the solution of equation (1).*

*Proof.* The proof involves two cases.

**1st case**: Suppose that

$$\liminf_{k\to\infty} \|F_k\| = 0. \tag{45}$$

Then, from boundedness of $\{x_k\}$ and continuity of $F$, some cluster points $\hat{x}$ exists for which $F(\hat{x}) = 0$. In addition, the convergence of $\{\|x_k - \hat{x}\|\}$ implies that $\{x_k\}$ converges to $\hat{x}$.

**2nd case**: $\liminf_{k\to\infty} \|F_k\| > 0$. This indicates the existence of a constant $\overline{\mu} > 0$ such that

$$\|F_k\| \geq \overline{\mu}. \tag{46}$$

Also, from the first inequality in equation (31), we have

$$\|d_k\| \geq \left(1 - \frac{1}{r^2}\right)\|F_k\| \geq \left(1 - \frac{1}{r^2}\right)\overline{\mu}.$$

Combining this with equation (43) yields

$$\lim_{k\to\infty} \vartheta_k = 0. \tag{47}$$

Now, the boundedness of $\{x_k\}$ and $\{d_k\}$ shows that there exists subsequences $\{x_{k_i}\}$ and $\{d_{k_i}\}$ such that

$$\lim_{i\to\infty, i\in\mathcal{K}} x_{k_i} = \hat{x}, \quad \lim_{i\to\infty, i\in\mathcal{K}} d_{k_i} = \hat{d},$$

where $\mathcal{K}$ is an infinite indexing set. Furthermore, from equation (30), we get

$$-F(x_{k_i})^T d_{k_i} \geq \left(1 - \frac{1}{r^2}\right)\|F(x_{k_i})\|^2, \quad \forall i \in \mathcal{K}. \tag{48}$$

Letting $i \to \infty$ in equation (48), with the continuity of $F$ and equation (46), yields

$$-F(\hat{x})^T \hat{d} \geq \left(1 - \frac{1}{r^2}\right)\|F(\hat{x})\|^2 > \left(1 - \frac{1}{r^2}\right)\overline{\mu}^2 > 0. \tag{49}$$

Also, we have seen that if $\vartheta_k \neq \overline{\zeta}$ in equation (26), then $\overline{\vartheta}_k = \beta^{-1}\vartheta_k$ will not satisfy equation (26), i.e.,

$$\begin{aligned} -F(x_{k_i} + \beta^{-1}\vartheta_{k_i} d_{k_i})^T d_{k_i} &< \overline{\vartheta}\beta^{-1}\vartheta_{k_i}[\|F(x_{k_i} + \beta^{-1}\vartheta_{k_i} d_{k_i})\|]\|d_{k_i}\|^2 \\ &\leq \overline{\vartheta}\beta^{-1}\vartheta_{k_i}\overline{m}\|d_{k_i}\|^2. \end{aligned} \tag{50}$$

Taking limit as $i \to \infty$, with $i \in \mathcal{K}$ in equation (50), equation (47) and continuity of $F$, we obtain

$$-F(\hat{x})^T \hat{d} \leq 0,$$

which contradicts equation (49). So, it is established that $\liminf_{k\to\infty} \|F_k\| = 0$.

We now prove that the sequences $\{x_k\}$ and $\{\psi_k\}$ converge to the same solution point of equation (1). Based on the boundedness of $\{x_k\}$, continuity of $F$, closedness of $C$, and equation (45), it can be shown that

$$\lim_{i\to\infty} x_{k_i} = \hat{x} \in C \text{ with } \lim_{k\to\infty} F(x_k) = F(\hat{x}) = 0,$$

which indicates that $\hat{x} \in \overline{C}$. If we set $\bar{x} = \hat{x}$ in the sequence $\{\|x_k - \bar{x}\|\}$, then we have that

$$\lim_{k\to\infty} \|x_k - \hat{x}\| = \lim_{i\to\infty} \|x_{k_i} - \hat{x}\| = 0.$$

This proves that $\{x_k\}$ converges to $\hat{x} \in \overline{C}$. Combining this result with equation (43) shows that $\{\psi_k\}$ also converges to $\hat{x} \in \overline{C}$, which concludes the proof. □

### 3.1. Rate of convergence

In this subsection, we prove the R-linear rate of convergence of Algorithm 1. Since $\{x_k\}$ converges to solution of (1), it can be stated that $x_k \to \bar{x}$ as $k \to \infty$ where $\bar{x} \in \overline{C}$. To proceed, the following assumption is required:

**Assumption** 3 Let $\bar{x} \in \overline{C}$, $\overline{\sigma} \in (0, 1)$, and $\alpha > 0$ exists such that

$$\overline{\sigma} dist(x, \overline{C}) \leq \|F(x)\|^2, \quad \forall x \in \mathcal{N}_\alpha(\bar{x}), \tag{51}$$

where $\mathcal{N}_\alpha(\bar{x})$ represents the neighbourhood of $\bar{x}$ defined by

$$\mathcal{N}_\alpha(\bar{x}) := \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \alpha\},$$

with $dist(x, \overline{C})$ being the distance from $x$ to $\overline{C}$.

**Theorem 2.** *Suppose that Assumptions* 1 − 3 *hold and the sequence* $\{x_k\}$ *is generated by Algorithm* 1. *Then the sequence* $\{dist(x_k, \overline{C})\}$ *converges Q-linearly to* 0, *implying* $\{x_k\}$ *is R-linearly convergent to* $\bar{x}$.

*Proof.* : We first define $\bar{x}_k := \arg\min\{\|x_k - x\| : x \in \overline{C}\}$ as the nearest solution to $x_k$, i.e,

$$\|x_k - \bar{x}_k\| = dist(x_k, \overline{C}). \tag{52}$$

Setting $\overline{\chi} = \phi(2 - \phi)\delta^2$ and employing (30), (40), (51) with $\bar{x}_k \in \overline{C}$, we have

$$\begin{aligned} dist(x_{k+1}, \overline{C})^2 &= \|x_{k+1} - \bar{x}_k\|^2 \\ &\leq dist(x_k, \overline{C})^2 - \overline{\chi}\|\vartheta_k d_k\|^4 \\ &\leq dist(x_k, \overline{C})^2 - \overline{\chi}\gamma^4\vartheta^4\|F(x_k)\|^4 \\ &\leq dist(u_k, \overline{C})^2 - \overline{\chi}\overline{\sigma}^2\gamma^4\vartheta^4(x_k, \overline{C})^2 \\ &= (1 - \overline{\chi}\overline{\sigma}^2\gamma^4\vartheta^4) dist(x_k, \overline{C})^2, \end{aligned} \tag{53}$$

which indicates that $\{dist(x_k, \overline{C})\}$ converges $Q$-linearly to 0, implying $\{x_k\}$ is $R$-linearly convergent to $\bar{x}$, since $\overline{\chi}$, $\overline{\sigma}$, $\gamma$, and $\vartheta$ are in the interval $(0, 1)$. □

## 4. Results of computational experiments and discussions

This section is used to analyze and verify the numerical effectiveness of Algorithm 1 (labelled MDYA). To that end, the section is divided into two subsections, where experiments are conducted for equation (1) and signal recovery in compressed sensing. In both experiments, the programs were written in MATLAB and executed on a PC with a (2.30 GHz CPU, 4 GB RAM) configuration.

### 4.1. Experiment 1: Constrained nonlinear equations

In the first experiment, the MDYA method is tested with the methods in [53] (labelled HHSLS), [54] (labelled T2DFP), and [55] (labelled PBDFP), respectively. The parameter values for the HHSLS, T2DFP, and PBDFP methods were set as applied by the respective authors. For the MDYA method, the parameters are $\beta = 0.5$, $\delta = 0.001$, $\phi = 1.97$, and $r = 5.5$. The iterations terminate when $\|F(x_k)\| \leq 10^{-10}$ or $\|F(z_k)\| \leq 10^{-10}$ or when 1000 iterations are exceeded.

Results of the first experiments are displayed in Tables 1, 2, and 3, where the labels "N", "Pdim", "Isp", "Nit", "Fval", and "Ptime(s)" represent test example number, dimension of test example, initial starting point, number of iterations, function evaluations, and processing time, respectively. Also, "$\|F(x_k)\|$" denotes the residual recorded at the approximate solution.

The following test examples with dimensions 1000, 10000, and 50000 were employed:

**Example 1.** [56]
$F_i(x) = 2x_i - \sin|x_i|, \quad i = 1, 2, \ldots, n,$
where $C = \mathbb{R}^n_+$.

**Example 2.** [3].
$F_1(x) = x_1 - e^{\left(\cos\frac{x_1 + x_2}{n+1}\right)},$
$F_i(x) = x_i - e^{\left(\cos\frac{x_{i-1} + x_i + x_{i+1}}{n+1}\right)}, \quad i = 2, 3, \ldots, n-1,$
$F_n(x) = x_n - e^{\left(\cos\frac{x_{n-1} + x_n}{n+1}\right)},$
where $C = \mathbb{R}^n_+$.

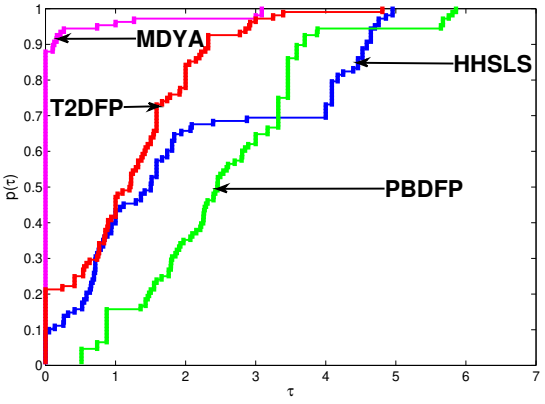


Figure 1: Iterations performance profile for MDYA, HHSLS, T2DFP, and PBDFP methods.

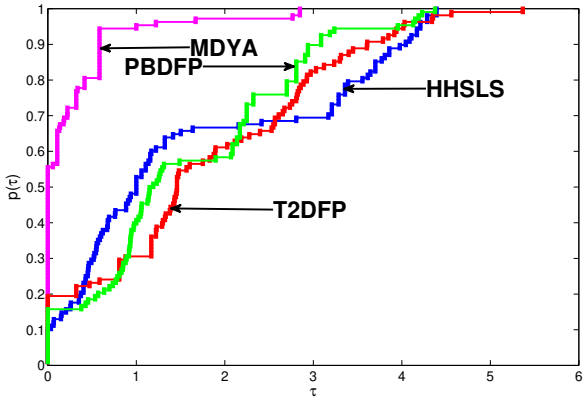


Figure 2: Function evaluations performance profile for MDYA, HHSLS, T2DFP, and PBDFP methods.

**Example 3.** Modification of Problem 1 in [57]
$F_1(x) = e^{\sin x_1} - 1,$
$F_i(x) = e^{\sin x_i} + x_i - 1, \quad i = 2, \ldots, n,$
where $C = \mathbb{R}^n_+$.

**Example 4.** Modification of Problem 2 in [56]
$F_i(x) = 3x_i - \sin(x_i), \quad i = 1, 2, \ldots, n,$
where $C = \mathbb{R}^n_+$.

**Example 5.** [58]
$F_1(x) = 2x_1 + \sin x_1 - 1,$
$F_i(x) = 2x_{i-1} + 2x_i + 2\sin x_i - 1,$
$F_n(x) = 2x_n + \sin x_n - 1, \quad i = 2, \ldots, n-1,$
where $C = \mathbb{R}^n_+$.

**Example 6.** Modification of test example 2 in [58]
$F_1(x) = x_1 - e^{\left(\cos \frac{x_1 + x_2}{2}\right)},$
$F_i(x) = x_i - e^{\left(\cos \frac{x_{i-1} + x_i + x_{i+1}}{i}\right)}, \quad i = 2, 3, \ldots, n-1,$
$F_n(x) = x_n - e^{\left(\cos \frac{x_{n-1} + x_n}{n}\right)},$
where $C = \mathbb{R}^n_+$.

Initial guesses used are listed as follows:
$\mathbf{x_0^1} = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}\right)^T$, $\mathbf{x_0^2} = \left(\frac{1}{2}, \frac{3}{2}, \ldots, -\frac{[(-1)^n - 2]}{2}\right)^T$, $\mathbf{x_0^3} = \left(1, 3, \ldots, -\frac{-2[(-1)^n - 2]}{2}\right)^T$,
$\mathbf{x_0^4} = \left(\frac{n-1}{n}, \frac{n-2}{n}, \ldots, 0\right)^T$, $\mathbf{x_0^5} = \left(\frac{1}{4}, \frac{3}{4}, \ldots, \frac{-[(-1)^n - 2]}{4}\right)^T$, $\mathbf{x_0^6} =$ rand$(n, 1)$, i.e., random numbers between 0 and 1.

In order to give a pictorial view of the results, Figures 1, 2, and 3 were plotted using the idea by Dolan and Moré's [59]. It can be seen from all the three figures presented that the MDYA method is the most successful of the four methods. This is backed by the fact that the top curve in all the three figures represents that of the MDYA method, indicating the former solved more of the test examples with lowest values of the three performance metrics than the latter. Going by the above analysis, we conclude that the MDYA method is effective and suitable for solving the constrained system of equations in (1).

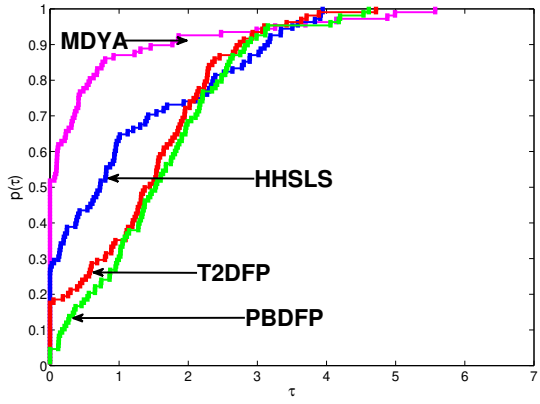


Figure 3: Processing time performance profile for MDYA, HHSLS, T2DFP, and PBDFP methods.

### *4.2. Experiment 2: Sparse signal reconstruction*

A concept associated with compressive sensing (CS) that has been trending over the past decade is sparse signal recovery or reconstruction (see [3, 60, 61] and other references). It involves the reconstruction of certain signals from fewer samplings or measurements. This is mainly achieved by finding sparse solutions to the under-determined linear system of equations $\mathcal{M}x = h$ via the $\ell_1 - \ell_2$ norm minimization problem:

$$\min_x \frac{1}{2}\|\mathcal{M}x - h\|_2^2 + \varpi\|x\|_1, \tag{54}$$

where $\varpi \geq 0$ is a parameter, $x \in \mathbb{R}^n$, $h \in \mathbb{R}^k$ is an observed value, $\mathcal{M} \in \mathbb{R}^{k \times n}$ $(k \ll n)$ is a linear mapping, and $\|x\|_1$ and $\|x\|_2$ are the $\ell_1$ and $\ell_2$ norms, respectively. It is quite clear that equation (54) is an unconstrained optimization problem.

Several of iterative methods exist for solving equation (54) (see [62–64]). According to Figueiredo et al. [65], equation (54) can be expressed as a convex quadratic problem, where $x \in \mathbb{R}^n$ is split into two parts as:

$$x = u - w, \quad u \geq 0, \quad w \geq 0, \quad u, w \in \mathbb{R}^n. \tag{55}$$

Table 1: Results of test examples 1-2 for MDYA, HHSLS, T2DFP and PBDFP methods.

| N | Pdim | | MDYA | | | | HHSLS | | | | T2DFP | | | | PBDFP | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Isp | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ |
| 1 | 1000 | $x_1$ | 1 | 3 | 0.0100 | 0 | 3 | 5 | 0.0167 | 0 | 1 | 2 | 0.1246 | 0 | 10 | 13 | 0.1839 | 9.23E-11 |
| | 1000 | $x_2$ | 1 | 3 | 0.0864 | 0 | 22 | 34 | 0.0180 | 5.81E-11 | 1 | 2 | 0.0109 | 0 | 13 | 17 | 0.0153 | 6.09E-12 |
| | 1000 | $x_3$ | 17 | 36 | 0.0198 | 3.09E-11 | 2 | 5 | 0.0055 | 0 | 21 | 72 | 0.0278 | 0 | 13 | 21 | 0.0091 | 8.70E-11 |
| | 1000 | $x_4$ | 1 | 3 | 0.0033 | 0 | 23 | 36 | 0.0191 | 9.15E-11 | 1 | 2 | 0.0049 | 0 | 11 | 14 | 0.007 | 2.05E-11 |
| | 1000 | $x_5$ | 1 | 3 | 0.0068 | 0 | 2 | 4 | 0.0040 | 0 | 1 | 2 | 0.0055 | 0 | 11 | 13 | 0.0121 | 5.82E-11 |
| | 1000 | $x_6$ | 1 | 3 | 0.1843 | 0 | 23 | 36 | 0.0196 | 8.49E-11 | 1 | 2 | 0.0063 | 0 | 11 | 14 | 0.0104 | 2.15E-11 |
| | 10000 | $x_1$ | 1 | 3 | 0.4807 | 0 | 3 | 5 | 0.0201 | 0 | 1 | 2 | 0.0101 | 0 | 10 | 13 | 0.0443 | 9.23E-11 |
| | 10000 | $x_2$ | 1 | 3 | 0.0110 | 0 | 24 | 37 | 0.0830 | 3.21E-11 | 1 | 2 | 0.0103 | 0 | 13 | 17 | 0.0691 | 1.93E-11 |
| | 10000 | $x_3$ | 17 | 36 | 0.0971 | 9.78E-11 | 2 | 5 | 0.0175 | 0 | 19 | 81 | 0.1470 | 0 | 14 | 22 | 0.0753 | 1.56E-11 |
| | 10000 | $x_4$ | 1 | 3 | 0.0117 | 0 | 25 | 39 | 0.0989 | 5.11E-11 | 1 | 2 | 0.0109 | 0 | 11 | 14 | 0.0555 | 6.51E-11 |
| | 10000 | $x_5$ | 1 | 3 | 0.0125 | 0 | 2 | 4 | 0.0144 | 0 | 1 | 2 | 0.0094 | 0 | 12 | 14 | 0.0524 | 1.04E-11 |
| | 10000 | $x_6$ | 1 | 3 | 0.0101 | 0 | 25 | 39 | 0.1019 | 5.22E-11 | 1 | 2 | 0.0121 | 0 | 11 | 14 | 0.0563 | 6.48E-11 |
| | 50000 | $x_1$ | 1 | 3 | 0.0672 | 0 | 3 | 5 | 0.0534 | 0 | 1 | 2 | 0.0245 | 0 | 10 | 13 | 0.1778 | 9.23E-11 |
| | 50000 | $x_2$ | 1 | 3 | 0.0329 | 0 | 24 | 37 | 0.4200 | 7.19E-11 | 1 | 2 | 0.0273 | 0 | 13 | 17 | 0.2364 | 4.3E-11 |
| | 50000 | $x_3$ | 16 | 34 | 0.4610 | 4.51E-11 | 2 | 5 | 0.0482 | 0 | 15 | 61 | 0.5167 | 0 | 14 | 22 | 0.2848 | 3.48E-11 |
| | 50000 | $x_4$ | 1 | 3 | 0.0339 | 0 | 27 | 42 | 0.4133 | 2.03E-11 | 1 | 2 | 0.0274 | 0 | 12 | 15 | 0.2433 | 8.24E-12 |
| | 50000 | $x_5$ | 1 | 3 | 0.0360 | 0 | 2 | 4 | 0.0407 | 0 | 1 | 2 | 0.0307 | 0 | 12 | 14 | 0.2298 | 2.33E-11 |
| | 50000 | $x_6$ | 1 | 3 | 0.0331 | 0 | 27 | 42 | 0.4575 | 2.02E-11 | 1 | 2 | 0.0304 | 0 | 12 | 15 | 0.2636 | 8.19E-12 |
| 2 | 1000 | $x_1$ | 7 | 15 | 0.1861 | 4.74E-12 | 18 | 25 | 0.0278 | 9.32E-11 | 31 | 181 | 0.0623 | 5.52E-11 | 10 | 12 | 0.0144 | 5.11E-11 |
| | 1000 | $x_2$ | 7 | 15 | 0.0118 | 3.13E-12 | 20 | 27 | 0.0238 | 8.21E-11 | 15 | 85 | 0.0425 | 4.61E-11 | 10 | 12 | 0.0139 | 3.37E-11 |
| | 1000 | $x_3$ | 6 | 14 | 0.0133 | 9.93E-15 | 23 | 32 | 0.0206 | 6.89E-11 | 18 | 98 | 0.0375 | 5.65E-11 | 10 | 12 | 0.0145 | 2.32E-11 |
| | 1000 | $x_4$ | 7 | 15 | 0.0111 | 3.91E-12 | 18 | 25 | 0.0193 | 4.72E-11 | 34 | 196 | 0.0712 | 6.40E-11 | 10 | 12 | 0.0152 | 4.22E-11 |
| | 1000 | $x_5$ | 7 | 15 | 0.0115 | 3.89E-12 | 25 | 34 | 0.0270 | 7.85E-11 | 24 | 146 | 0.0559 | 8.02E-11 | 10 | 12 | 0.0102 | 4.21E-11 |
| | 1000 | $x_6$ | 7 | 15 | 0.0191 | 3.92E-12 | 20 | 28 | 0.0225 | 8.71E-11 | 29 | 184 | 0.0641 | 7.18E-11 | 10 | 12 | 0.0137 | 4.23E-11 |
| | 10000 | $x_1$ | 6 | 14 | 0.0587 | 0 | 18 | 24 | 0.1031 | 2.06E-11 | 12 | 63 | 0.2003 | 1.94E-11 | 11 | 13 | 0.0739 | 9.18E-12 |
| | 10000 | $x_2$ | 6 | 14 | 0.0755 | 0 | 18 | 24 | 0.1201 | 5.10E-11 | 17 | 99 | 0.2855 | 9.01E-11 | 11 | 13 | 0.0826 | 6.05E-12 |
| | 10000 | $x_3$ | 6 | 14 | 0.0640 | 0 | 17 | 23 | 0.1048 | 5.11E-11 | 19 | 113 | 0.3734 | 2.63E-11 | 10 | 12 | 0.0596 | 7.35E-11 |
| | 10000 | $x_4$ | 6 | 14 | 0.0642 | 0 | 19 | 26 | 0.1223 | 4.44E-11 | 14 | 75 | 0.2244 | 9.79E-11 | 11 | 13 | 0.0819 | 7.56E-12 |
| | 10000 | $x_5$ | 6 | 14 | 0.0840 | 0 | 21 | 29 | 0.1267 | 2.75E-11 | 15 | 91 | 0.2845 | 7.41E-11 | 11 | 13 | 0.0787 | 7.53E-12 |
| | 10000 | $x_6$ | 6 | 14 | 0.0615 | 0 | 20 | 28 | 0.1202 | 4.71E-11 | 16 | 93 | 0.2733 | 7.02E-11 | 11 | 13 | 0.0668 | 7.56E-12 |
| | 50000 | $x_1$ | 6 | 14 | 0.3334 | 0 | 17 | 22 | 0.5003 | 4.77E-11 | 15 | 76 | 1.0440 | 1.16E-11 | 11 | 13 | 0.3766 | 2.05E-11 |
| | 50000 | $x_2$ | 6 | 14 | 0.3109 | 0 | 21 | 29 | 0.5845 | 6.42E-11 | 10 | 53 | 0.7367 | 5.20E-11 | 11 | 13 | 0.3403 | 1.35E-11 |
| | 50000 | $x_3$ | 4 | 10 | 0.1904 | 7.02E-14 | 17 | 22 | 0.4704 | 3.01E-11 | 12 | 63 | 0.9211 | 2.75E-11 | 11 | 13 | 0.3463 | 9.35E-12 |
| | 50000 | $x_4$ | 6 | 14 | 0.2789 | 0 | 21 | 28 | 0.5399 | 3.92E-11 | 14 | 77 | 1.0823 | 2.11E-11 | 11 | 13 | 0.338 | 1.69E-11 |
| | 50000 | $x_5$ | 6 | 14 | 0.2973 | 0 | 20 | 26 | 0.5701 | 9.60E-11 | 20 | 113 | 1.5497 | 9.40E-11 | 11 | 13 | 0.3442 | 1.69E-11 |
| | 50000 | $x_6$ | 6 | 14 | 0.3010 | 0 | 16 | 20 | 0.4383 | 2.85E-11 | 14 | 77 | 1.1200 | 3.63E-11 | 11 | 13 | 0.3332 | 1.69E-11 |

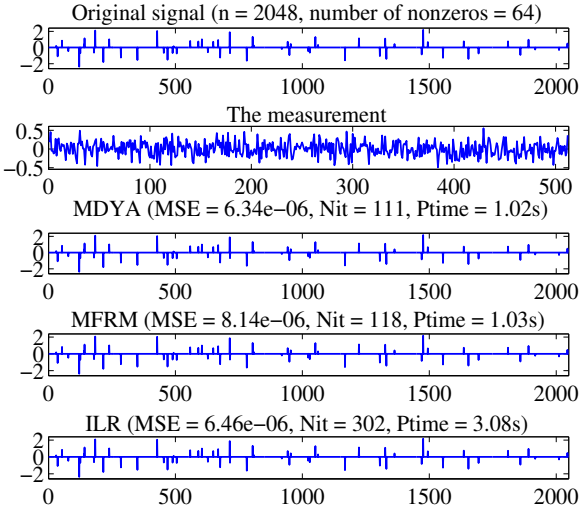


Figure 4: In descending order: the original signal, measurement, and recovered signals by MDYA, MFRM, and ILR methods.

Let $u_i = (x_i)_+$ and $w_i = (-x_i)_+$ for all $i = 1, 2, \ldots, n$, where $(\cdot)_+$ is a positive operator given as $(\cdot)_+ = \max\{0, \cdot\}$. Using the standard definition of the $\ell_1$-norm, we have $\|x\|_1 = A_n^T u + A_n^T w$, where $A_n = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$. Based on this, equation (54) is reformulated as:

$$\min_{u,w} \frac{1}{2}\|\mathcal{M}(u - w) - h\|_2^2 + \varpi A_n^T u + \varpi A_n^T w, \quad u, w \geq 0.$$

Following the framework in Ref. [65], this problem can be rewritten as:

$$\min_{\psi} \frac{1}{2}\psi^T G\psi + C^T\psi, \quad \psi \geq 0, \tag{56}$$

in which

$$\psi = \begin{pmatrix} u \\ w \end{pmatrix}, \qquad C = \varpi A_{2n} + \begin{pmatrix} -h_0 \\ h_0 \end{pmatrix},$$

$$h_0 = \mathcal{M}^T h, \qquad G = \begin{pmatrix} \mathcal{M}^T\mathcal{M} & -\mathcal{M}^T\mathcal{M} \\ -\mathcal{M}^T\mathcal{M} & \mathcal{M}^T\mathcal{M} \end{pmatrix}.$$

Here, $G$ is positive semidefinite. Therefore, equation (56) is a convex quadratic programming problem, which is equivalent to finding the zero of the map:

$$F(\psi) = \min\{\psi, G\psi + C\} = 0,$$

Table 2: Results of test examples 3-4 for MDYA, HHSLS, T2DFP and PBDFP methods.

| N | Pdim | | MDYA | | | | HHSLS | | | | T2DFP | | | | PBDFP | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Isp | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ |
| 3 | 1000 | $x_1$ | 11 | 31 | 0.7439 | 7.89E-11 | 18 | 41 | 0.0234 | 3.21E-11 | 22 | 115 | 0.0411 | 0 | 58 | 76 | 0.0341 | 9.03E-11 |
| | 1000 | $x_2$ | 12 | 32 | 0.0144 | 3.34E-11 | 19 | 44 | 0.0156 | 1.69E-11 | 11 | 48 | 0.0189 | 0 | 47 | 71 | 0.0311 | 7.91E-11 |
| | 1000 | $x_3$ | 15 | 39 | 0.0218 | 3.54E-11 | 20 | 45 | 0.0201 | 9.56E-11 | 22 | 107 | 0.0306 | 0 | 70 | 93 | 0.0489 | 9.07E-11 |
| | 1000 | $x_4$ | 1 | 4 | 0.0049 | 0 | 23 | 52 | 0.0225 | 1.26E-11 | 3 | 7 | 0.0085 | 0 | 50 | 62 | 0.0305 | 5.32E-11 |
| | 1000 | $x_5$ | 1 | 4 | 0.0050 | 0 | 17 | 36 | 0.0190 | 1.63E-11 | 3 | 7 | 0.0074 | 0 | 51 | 70 | 0.0348 | 7.67E-11 |
| | 1000 | $x_6$ | 5 | 14 | 0.0076 | 8.56E-12 | 22 | 49 | 0.0247 | 6.43E-11 | 3 | 7 | 0.0051 | 0 | 44 | 66 | 0.0312 | 4.71E-11 |
| | 10000 | $x_1$ | 11 | 31 | 0.0756 | 7.90E-11 | 18 | 41 | 0.0866 | 3.10E-11 | 13 | 57 | 0.1213 | 0 | 66 | 87 | 0.2725 | 6.37E-11 |
| | 10000 | $x_2$ | 12 | 30 | 0.0805 | 9.96E-11 | 18 | 44 | 0.0883 | 8.35E-11 | 16 | 70 | 0.1723 | 0 | 44 | 62 | 0.2059 | 8.34E-12 |
| | 10000 | $x_3$ | 15 | 40 | 0.1461 | 7.58E-11 | 18 | 41 | 0.0840 | 9.92E-11 | 20 | 104 | 0.2288 | 0 | 52 | 77 | 0.2410 | 9.97E-11 |
| | 10000 | $x_4$ | 1 | 4 | 0.0148 | 0 | 22 | 50 | 0.0921 | 3.94E-11 | 4 | 9 | 0.0319 | 0 | 54 | 71 | 0.2674 | 6.15E-11 |
| | 10000 | $x_5$ | 1 | 4 | 0.0150 | 0 | 22 | 49 | 0.1115 | 4.34E-11 | 4 | 9 | 0.0321 | 0 | 56 | 73 | 0.2501 | 7.57E-11 |
| | 10000 | $x_6$ | 8 | 21 | 0.0609 | 5.44E-11 | 21 | 48 | 0.1041 | 9.12E-11 | 4 | 9 | 0.0322 | 0 | 59 | 82 | 0.2553 | 6.67E-11 |
| | 50000 | $x_1$ | 11 | 31 | 0.3657 | 7.90E-11 | 18 | 41 | 0.3405 | 3.09E-11 | 11 | 43 | 0.4606 | 0 | 59 | 77 | 1.1486 | 7.54E-11 |
| | 50000 | $x_2$ | 12 | 35 | 0.4108 | 8.98E-11 | 21 | 49 | 0.4707 | 8.64E-11 | 5 | 11 | 0.1759 | 0 | 36 | 51 | 0.7250 | 9.85E-11 |
| | 50000 | $x_3$ | 15 | 41 | 0.5432 | 7.11E-11 | 18 | 43 | 0.4079 | 8.41E-11 | 20 | 101 | 0.9320 | 0 | 57 | 82 | 1.1145 | 7.19E-11 |
| | 50000 | $x_4$ | 1 | 4 | 0.0437 | 0 | 25 | 56 | 0.4960 | 2.01E-11 | 8 | 29 | 0.3049 | 0 | 51 | 75 | 1.0124 | 8.15E-11 |
| | 50000 | $x_5$ | 1 | 4 | 0.0505 | 0 | 19 | 42 | 0.4343 | 3.36E-11 | 28 | 165 | 1.3269 | 7.66E-11 | 58 | 83 | 1.2350 | 7.17E-11 |
| | 50000 | $x_6$ | 7 | 18 | 0.2521 | 6.85E-11 | 25 | 56 | 0.4802 | 2.82E-11 | 13 | 66 | 0.6295 | 0 | 52 | 67 | 1.1678 | 9.96E-11 |
| 4 | 1000 | $x_1$ | 1 | 4 | 0.1040 | 0 | 16 | 37 | 0.0136 | 4.45E-11 | 2 | 5 | 0.0059 | 0 | 8 | 17 | 0.0116 | 9.65E-11 |
| | 1000 | $x_2$ | 1 | 4 | 0.0043 | 0 | 17 | 39 | 0.0111 | 2.79E-11 | 3 | 7 | 0.0052 | 0 | 11 | 20 | 0.0130 | 6.8E-12 |
| | 1000 | $x_3$ | 1 | 4 | 0.0048 | 0 | 16 | 39 | 0.0192 | 4.63E-11 | 1 | 3 | 0.0037 | 0 | 11 | 23 | 0.0128 | 1.22E-11 |
| | 1000 | $x_4$ | 1 | 4 | 0.0054 | 0 | 25 | 58 | 0.0256 | 7.60E-11 | 3 | 7 | 0.0052 | 0 | 9 | 18 | 0.0103 | 3.72E-11 |
| | 1000 | $x_5$ | 1 | 4 | 0.0039 | 0 | 17 | 39 | 0.0200 | 3.79E-11 | 3 | 7 | 0.0073 | 0 | 10 | 18 | 0.0079 | 1.45E-11 |
| | 1000 | $x_6$ | 1 | 4 | 0.0057 | 0 | 21 | 47 | 0.0164 | 9.79E-11 | 3 | 7 | 0.0059 | 0 | 9 | 18 | 0.0089 | 3.79E-11 |
| | 10000 | $x_1$ | 1 | 4 | 0.0128 | 0 | 16 | 37 | 0.0641 | 4.42E-11 | 2 | 5 | 0.0145 | 0 | 8 | 17 | 0.0439 | 9.65E-11 |
| | 10000 | $x_2$ | 1 | 4 | 0.0162 | 0 | 17 | 39 | 0.0699 | 8.82E-11 | 4 | 9 | 0.0300 | 0 | 11 | 20 | 0.0519 | 2.15E-11 |
| | 10000 | $x_3$ | 1 | 4 | 0.0145 | 0 | 17 | 39 | 0.0857 | 4.97E-11 | 1 | 3 | 0.0125 | 0 | 11 | 23 | 0.0576 | 3.86E-11 |
| | 10000 | $x_4$ | 1 | 4 | 0.0125 | 0 | 31 | 66 | 0.1271 | 9.26E-11 | 4 | 9 | 0.0319 | 0 | 10 | 19 | 0.0562 | 3.44E-12 |
| | 10000 | $x_5$ | 1 | 4 | 0.0156 | 0 | 18 | 41 | 0.0796 | 1.39E-11 | 4 | 9 | 0.0301 | 0 | 10 | 18 | 0.0554 | 4.58E-11 |
| | 10000 | $x_6$ | 1 | 4 | 0.0131 | 0 | 29 | 63 | 0.1193 | 2.73E-11 | 4 | 9 | 0.0326 | 0 | 10 | 19 | 0.0688 | 3.43E-12 |
| | 50000 | $x_1$ | 1 | 4 | 0.0408 | 0 | 16 | 37 | 0.2849 | 4.41E-11 | 2 | 5 | 0.0616 | 0 | 8 | 17 | 0.1887 | 9.65E-11 |
| | 50000 | $x_2$ | 1 | 4 | 0.0414 | 0 | 17 | 41 | 0.3419 | 7.61E-11 | 5 | 11 | 0.1190 | 0 | 11 | 20 | 0.2420 | 4.81E-11 |
| | 50000 | $x_3$ | 1 | 4 | 0.0496 | 0 | 17 | 41 | 0.3504 | 3.49E-11 | 1 | 3 | 0.0395 | 0 | 11 | 23 | 0.2780 | 8.62E-11 |
| | 50000 | $x_4$ | 1 | 4 | 0.0397 | 0 | 31 | 69 | 0.5657 | 8.87E-11 | 5 | 11 | 0.1533 | 0 | 10 | 19 | 0.2301 | 7.69E-12 |
| | 50000 | $x_5$ | 1 | 4 | 0.0445 | 0 | 18 | 41 | 0.3301 | 3.12E-11 | 5 | 11 | 0.1110 | 0 | 11 | 19 | 0.2379 | 2.98E-12 |
| | 50000 | $x_6$ | 1 | 4 | 0.0402 | 0 | 26 | 58 | 0.4831 | 3.99E-11 | 5 | 11 | 0.1258 | 0 | 10 | 19 | 0.2413 | 7.62E-12 |

where the function $F$ is vector-valued, monotone, and Lipschitz continuous. Thus, according to Ref. [64, 66], the MDYA scheme can be applied effectively to solve it.

To further verify the practical efficacy of the MDYA method, we compare it against the methods presented in Ref. [67] (labelled MFRM) and Ref. [68] (labelled ILR) to reconstruct a sparse signal of length $n$ from $k$ observations. In this experiment, the mean squared error (MSE) relative to the original signal $\tilde{x}$, defined as:

$$MSE = \frac{1}{n}\|\tilde{x} - \bar{x}\|^2,$$

is employed to measure the quality of the reconstruction, where $\bar{x}$ stands for the restored signal.

The signal size is initially set to $n = 2^{11}$ with a sparsity level of $2^6$ (meaning it contains 64 randomly distributed nonzero elements). The number of measurements is taken as $k = 2^9$ contaminated with Gaussian noise $\xi \sim N(0_k, \sigma^2 I_k)$, where $0_k$ is the zero vector, $I_k$ is the $k \times k$ identity matrix, and the noise variance is set to $\sigma^2 = 10^{-4}$. The measurement vector $h$ is given by:

$$h = \mathcal{M}\tilde{x} + \xi.$$

We utilize $f(x) = \frac{1}{2}\|\mathcal{M}x - h\|_2^2 + \varpi\|x\|_1$ as our merit function. The experiment is initialized with $x_0 = \mathcal{M}^T h$ and is terminated when the relative change satisfies:

$$\frac{\|f_{k+1} - f_k\|}{\|f_k\|} < 10^{-5},$$

where $f_k$ represents the objective function value at iteration $k$.

The experiment was repeated ten times, and the comparative numerical results are reported in Table 4. From these results, it is clear that on average, the MDYA method outperformed both the MFRM and ILR algorithms. Furthermore, Figure 4 displays the original signal, measurement, and recovered signals for each algorithm. The visual plots demonstrate that all three methods can reconstruct the signal almost exactly, though MDYA consistently requires the shortest computational time, fewer iterations, and yields lower MSE values.

Additionally, Figure 5 presents four distinct graphs illustrating the convergence behavior of the algorithms across MSE, objective function values (ObjFun), number of iterations (Nit), and processing time (Ptime (s)). The figures reveal that the rate of descent for both MSE and the objective function values achieved by MDYA is faster than that of MFRM and ILR. This confirms that MDYA requires less computational overhead to reconstruct the original signal under identical conditions.

As noted in Ref. [69], reducing the required number of measurements without degrading performance remains a vital ob-

Table 3: Results of test examples 5-6 for MDYA, HHSLS, T2DFP and PBDFP methods.

| N | Pdim | | MDYA | | | | HHSLS | | | | T2DFP | | | | PBDFP | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Isp** | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ | Nit | Fval | Ptime(s) | $\|F(x_k)\|$ |
| 5 | 1000 | $x_1$ | 28 | 116 | 0.0756 | 9.88E-11 | 35 | 139 | 0.0472 | 7.95E-11 | 42 | 284 | 0.0676 | 9.97E-11 | 154 | 236 | 0.1220 | 4.77E-11 |
| | 1000 | $x_2$ | 29 | 120 | 0.0422 | 5.89E-11 | 25 | 103 | 0.0318 | 8.17E-11 | 58 | 366 | 0.1065 | 7.61E-11 | 174 | 247 | 0.1436 | 8.84E-11 |
| | 1000 | $x_3$ | 29 | 120 | 0.0512 | 3.24E-11 | 45 | 172 | 0.0400 | 9.66E-11 | 52 | 333 | 0.1035 | 9.25E-11 | 138 | 229 | 0.1010 | 7.59E-11 |
| | 1000 | $x_4$ | 30 | 125 | 0.0643 | 5.06E-11 | 25 | 99 | 0.0259 | 9.15E-11 | 51 | 368 | 0.1227 | 7.82E-11 | 120 | 206 | 0.1031 | 9.30E-11 |
| | 1000 | $x_5$ | 28 | 116 | 0.0422 | 5.28E-11 | 25 | 102 | 0.0316 | 8.10E-11 | 50 | 310 | 0.0933 | 6.45E-11 | 141 | 225 | 0.0982 | 8.54E-11 |
| | 1000 | $x_6$ | 30 | 125 | 0.1393 | 6.92E-11 | 44 | 171 | 0.0386 | 8.33E-11 | 51 | 327 | 0.0931 | 8.22E-11 | 145 | 230 | 0.1233 | 5.84E-11 |
| | 10000 | $x_1$ | 29 | 120 | 0.3051 | 4.52E-11 | 26 | 106 | 0.1970 | 9.86E-11 | 47 | 283 | 1.3221 | 7.64E-11 | 137 | 221 | 0.7728 | 4.16E-11 |
| | 10000 | $x_2$ | 28 | 116 | 0.3545 | 6.09E-11 | 26 | 106 | 0.2206 | 8.07E-11 | 47 | 296 | 0.6521 | 7.03E-11 | 151 | 242 | 0.8268 | 8.33E-11 |
| | 10000 | $x_3$ | 30 | 125 | 0.3359 | 7.74E-11 | 47 | 180 | 0.3123 | 6.89E-11 | 51 | 337 | 0.7298 | 7.42E-11 | 149 | 239 | 0.7909 | 9.84E-11 |
| | 10000 | $x_4$ | 31 | 128 | 0.3267 | 3.70E-11 | 34 | 134 | 0.2540 | 8.12E-11 | 45 | 300 | 0.6122 | 9.72E-11 | 130 | 213 | 0.7354 | 7.95E-11 |
| | 10000 | $x_5$ | 28 | 116 | 0.3069 | 6.54E-11 | 29 | 116 | 0.2122 | 5.55E-11 | 52 | 344 | 0.6848 | 8.59E-11 | 137 | 228 | 0.8259 | 9.92E-11 |
| | 10000 | $x_6$ | 32 | 133 | 0.3913 | 4.75E-11 | 46 | 179 | 0.3146 | 7.11E-11 | 54 | 335 | 0.7226 | 5.22E-11 | 148 | 237 | 0.7832 | 9.73E-11 |
| | 50000 | $x_1$ | 31 | 129 | 1.4230 | 3.84E-11 | 37 | 143 | 1.0537 | 9.53E-11 | 48 | 352 | 3.1313 | 8.30E-11 | 170 | 246 | 4.0945 | 5.89E-11 |
| | 50000 | $x_2$ | 29 | 120 | 1.3433 | 9.36E-11 | 27 | 109 | 0.8017 | 9.95E-11 | 62 | 367 | 3.5172 | 7.57E-11 | 148 | 238 | 3.5738 | 9.51E-11 |
| | 50000 | $x_3$ | 30 | 123 | 1.3700 | 8.97E-11 | 49 | 185 | 1.3946 | 9.97E-11 | 51 | 308 | 2.9978 | 7.55E-11 | 165 | 273 | 4.0138 | 9.98E-11 |
| | 50000 | $x_4$ | 31 | 128 | 1.3755 | 7.23E-11 | 31 | 123 | 0.9340 | 9.98E-11 | 45 | 287 | 2.7270 | 9.77E-11 | 148 | 228 | 3.5087 | 6.44E-11 |
| | 50000 | $x_5$ | 24 | 96 | 1.0521 | 6.09E-11 | 43 | 163 | 1.2482 | 9.61E-11 | 66 | 437 | 4.0610 | 9.50E-11 | 146 | 232 | 3.5339 | 7.31E-11 |
| | 50000 | $x_6$ | 32 | 133 | 1.4735 | 9.33E-11 | 51 | 194 | 1.3847 | 7.13E-11 | 59 | 359 | 3.3977 | 6.41E-11 | 153 | 240 | 3.5864 | 8.43E-11 |
| 6 | 1000 | $x_1$ | 12 | 25 | 0.1138 | 4.82E-11 | 23 | 40 | 0.0333 | 2.90E-11 | 56 | 273 | 0.1224 | 9.10E-11 | 47 | 55 | 0.0401 | 9.98E-11 |
| | 1000 | $x_2$ | 13 | 27 | 0.0306 | 6.60E-11 | 21 | 38 | 0.0319 | 5.08E-11 | 49 | 266 | 0.0889 | 8.92E-11 | 41 | 46 | 0.0452 | 8.57E-11 |
| | 1000 | $x_3$ | 11 | 25 | 0.0217 | 4.71E-11 | 23 | 38 | 0.0241 | 4.92E-11 | 32 | 190 | 0.0655 | 6.42E-11 | 38 | 45 | 0.0342 | 7.93E-11 |
| | 1000 | $x_4$ | 12 | 26 | 0.0195 | 9.79E-11 | 24 | 42 | 0.0262 | 9.99E-11 | 23 | 137 | 0.0564 | 7.65E-11 | 45 | 54 | 0.0395 | 4.19E-11 |
| | 1000 | $x_5$ | 13 | 29 | 0.0263 | 3.55E-11 | 22 | 37 | 0.0265 | 2.59E-11 | 94 | 592 | 0.2007 | 6.33E-11 | 35 | 39 | 0.0340 | 5.19E-11 |
| | 1000 | $x_6$ | 14 | 31 | 0.0267 | 5.47E-11 | 21 | 37 | 0.0254 | 2.62E-11 | 25 | 134 | 0.0469 | 6.12E-11 | 41 | 46 | 0.0527 | 8.82E-11 |
| | 10000 | $x_1$ | 12 | 27 | 0.1174 | 6.65E-11 | 20 | 35 | 0.1175 | 7.01E-11 | 34 | 183 | 0.4952 | 8.71E-11 | 36 | 42 | 0.2148 | 5.45E-11 |
| | 10000 | $x_2$ | 12 | 24 | 0.0996 | 6.21E-11 | 26 | 45 | 0.1746 | 8.06E-11 | 31 | 188 | 0.4852 | 3.66E-11 | 42 | 47 | 0.2762 | 9.22E-11 |
| | 10000 | $x_3$ | 9 | 20 | 0.0820 | 9.87E-11 | 22 | 38 | 0.1498 | 3.21E-11 | 22 | 122 | 0.3327 | 8.90E-11 | 40 | 47 | 0.3001 | 8.37E-11 |
| | 10000 | $x_4$ | 13 | 29 | 0.1198 | 6.97E-11 | 23 | 39 | 0.1377 | 4.80E-11 | 34 | 185 | 0.4871 | 8.85E-11 | 44 | 50 | 0.2477 | 8.27E-11 |
| | 10000 | $x_5$ | 14 | 29 | 0.1258 | 7.54E-12 | 22 | 37 | 0.1402 | 6.57E-11 | 53 | 306 | 0.8175 | 8.19E-11 | 41 | 47 | 0.2441 | 9.02E-11 |
| | 10000 | $x_6$ | 13 | 27 | 0.1197 | 3.32E-11 | 21 | 37 | 0.1319 | 1.71E-11 | 35 | 200 | 0.5895 | 1.76E-11 | 44 | 51 | 0.3206 | 8.40E-11 |
| | 50000 | $x_1$ | 13 | 29 | 0.5559 | 3.80E-11 | 23 | 40 | 0.6586 | 7.73E-11 | 38 | 204 | 2.7017 | 7.03E-11 | 47 | 55 | 1.3483 | 9.37E-11 |
| | 50000 | $x_2$ | 12 | 25 | 0.4927 | 9.59E-11 | 23 | 40 | 0.6420 | 8.02E-11 | 55 | 345 | 4.2239 | 4.03E-11 | 34 | 48 | 1.0881 | 1.72E-11 |
| | 50000 | $x_3$ | 16 | 35 | 0.6961 | 6.21E-11 | 23 | 39 | 0.5945 | 5.39E-11 | 38 | 233 | 2.8637 | 3.59E-11 | 41 | 48 | 1.1472 | 7.88E-11 |
| | 50000 | $x_4$ | 14 | 28 | 0.5600 | 2.15E-11 | 23 | 41 | 0.6515 | 9.94E-11 | 106 | 660 | 8.2994 | 9.21E-11 | 39 | 49 | 1.1350 | 7.74E-11 |
| | 50000 | $x_5$ | 13 | 27 | 0.5158 | 1.36E-11 | 24 | 43 | 0.6761 | 4.62E-11 | 89 | 547 | 6.7536 | 5.04E-11 | 45 | 51 | 1.2665 | 7.07E-11 |
| | 50000 | $x_6$ | 13 | 29 | 0.5451 | 6.23E-11 | 25 | 43 | 0.7032 | 9.21E-11 | 52 | 288 | 4.1197 | 7.11E-11 | 35 | 42 | 0.9937 | 9.18E-11 |

Table 4: Signal recovery results with sparsity level 64 and ten noise samples for MDYA, MFRM, and ILR methods.

| | MDYA | | | | MFRM | | | | ILR | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Exp. No | MSE | Nit | ObjFE | Ptime(s) | MSE | Nit | ObjFE | Ptime(s) | MSE | Nit | ObjFE | Ptime(s) |
| 1 | 5.177e-06 | 109 | 1.430e-01 | 1.14 | 5.286e-06 | 287 | 1.430e-01 | 3.36 | 6.517e-06 | 134 | 1.430e-01 | 1.30 |
| 2 | 4.240e-06 | 100 | 1.395e-01 | 0.97 | 1.591e-05 | 246 | 1.401e-01 | 2.16 | 5.424e-06 | 142 | 1.395e-01 | 1.42 |
| 3 | 1.153e-05 | 136 | 1.871e-01 | 1.83 | 1.228e-05 | 345 | 1.868e-01 | 3.20 | 1.541e-05 | 146 | 1.868e-01 | 1.50 |
| 4 | 6.039e-06 | 124 | 1.638e-01 | 1.22 | 6.039e-06 | 312 | 1.639e-01 | 3.31 | 7.495e-06 | 132 | 1.638e-01 | 1.25 |
| 5 | 1.007e-05 | 77 | 1.389e-01 | 0.81 | 1.060e-05 | 298 | 1.387e-01 | 2.86 | 1.288e-05 | 131 | 1.387e-01 | 1.75 |
| 6 | 1.148e-05 | 131 | 2.035e-01 | 1.48 | 1.168e-05 | 321 | 2.035e-01 | 3.59 | 1.427e-05 | 137 | 2.034e-01 | 1.44 |
| 7 | 7.623e-06 | 108 | 1.502e-01 | 1.05 | 7.674e-06 | 331 | 1.502e-01 | 3.30 | 9.651e-06 | 138 | 1.502e-01 | 1.28 |
| 8 | 1.230e-05 | 102 | 1.341e-01 | 1.11 | 5.192e-06 | 283 | 1.337e-01 | 2.55 | 6.450e-06 | 114 | 1.337e-01 | 1.13 |
| 9 | 1.222e-05 | 126 | 1.988e-01 | 1.47 | 1.125e-05 | 327 | 1.989e-01 | 3.36 | 1.373e-05 | 139 | 1.989e-01 | 1.59 |
| 10 | 1.099e-05 | 134 | 2.000e-01 | 1.38 | 1.138e-05 | 309 | 1.998e-01 | 2.75 | 1.431e-05 | 126 | 1.998e-01 | 1.58 |
| Average | 9.066e-06 | 114.7 | 1.658e-01 | 1.246 | 9.729e-06 | 305.9 | 1.658e-01 | 3.044 | 1.061e-05 | 133.9 | 1.658e-01 | 1.424 |

jective in CS. To evaluate this aspect, we test the three methods under different sparsity levels and CS ratios defined by $\overline{\lambda} = \frac{k}{n}$. Table 5 is divided into sub-tables (a) and (b) to track these distinct scenarios. Table 5(a) evaluates CS ratios $\overline{\lambda} = 0.5, 0.25$, and 0.125 with corresponding sparsity levels of 32 and 64 (shown in parentheses). Table 5(b) reports the performance across varying noise variances and sparsity levels. The optimal numerical results (underlined) in both sub-tables confirm that the MDYA algorithm recovers sparse signals more efficiently while reliably preserving reconstruction quality.

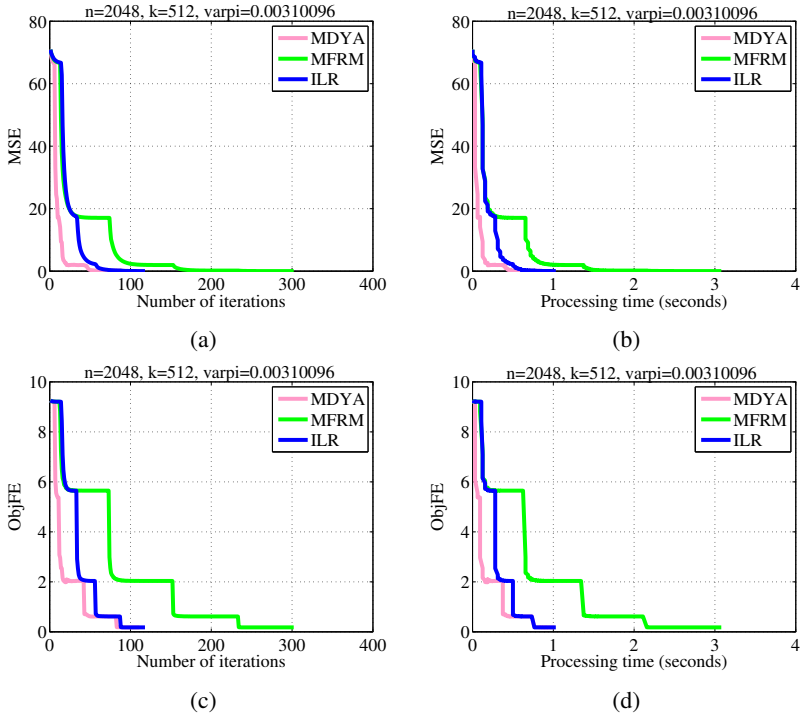


Figure 5: Comparison of the numerical results for MDYA, MFRM and ILR methods.

Table 5: Signal recovery results for MDYA, MFRM, and ILR under different compressed-sensing ratios, noise levels, and sparsity levels.

(a) Signal recovery results of MDYA, MFRM, and ILR with sparsity level 32(64) under different CS ratios

| | MDYA | | | | MFRM | | | | ILR | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CS ratio | MSE | ObjFun | Nit | Ptime | MSE | ObjFun | Nit | Ptime | MSE | ObjFun | Nit | Ptime |
| 0.50 | 1.301e-06 | 1.560e-01 | 27 | 1.66 | 1.427e-06 | 1.556e-01 | 57 | 3.70 | 1.455e-06 | 1.556e-01 | 58 | 3.42 |
| | (2.086e-06) | (2.611e-01) | (49) | (3.69) | (2.144e-06) | (2.610e-01) | (95) | (6.55) | (2.142e-06) | (2.610e-01) | (63) | (4.44) |
| 0.25 | 1.360e-06 | 6.600e-02 | 81 | 2.63 | 1.368e-06 | 6.600e-02 | 169 | 5.67 | 1.660e-06 | 6.599e-02 | 89 | 2.95 |
| | (3.152e-06) | (1.685e-01) | (101) | (4.11) | (9.591e-06) | (1.692e-01) | (187) | (6.27) | (3.931e-06) | (1.684e-01) | (111) | (3.64) |
| 0.125 | 1.194e-06 | 3.268e-02 | 110 | 2.59 | 1.227e-06 | 3.268e-02 | 463 | 9.02 | 1.487e-06 | 3.268e-02 | 221 | 4.61 |
| | (6.610e-06) | (9.832e-02) | (146) | (3.52) | (5.162e-06) | (9.844e-02) | (671) | (13.48) | (6.843e-06) | (9.845e-02) | (310) | (6.45) |

(b) Signal recovery results for MDYA, MFRM, and ILR under different noises and sparsity levels 32(64)

| | MDYA | | | | MFRM | | | | ILR | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma^2$ | MSE | ObjFun | Nit | Ptime | MSE | ObjFun | Nit | Ptime | MSE | ObjFun | Nit | Ptime |
| $10^{-4}$ | 7.504e-05 | 5.445e-02 | 161 | 3.97 | 7.666e-05 | 5.457e-02 | 850 | 18.72 | 7.609e-05 | 5.460e-02 | 254 | 7.08 |
| | (1.765e-04) | (1.298e-01) | (114) | (2.63) | (1.756e-04) | (1.298e-01) | (1250) | (23.73) | (1.722e-04) | (1.299e-01) | (326) | (6.42) |
| $10^{-6}$ | 2.179e-06 | 4.183e-02 | 136 | 3.08 | 2.201e-06 | 4.182e-02 | 409 | 8.16 | 2.649e-06 | 4.182e-02 | 237 | 4.89 |
| | (1.053e-05) | (7.473e-02) | (154) | (3.64) | (3.705e-06) | (3.705e-06) | (770) | (14.75) | (5.528e-06) | (7.473e-02) | (360) | (7.39) |
| $10^{-8}$ | 1.716e-06 | 4.836e-02 | 154 | 3.63 | 1.938e-06 | 4.834e-02 | 488 | 9.52 | 2.327e-06 | 4.834e-02 | 241 | 5.00 |
| | (7.392e-06) | (7.067e-02) | (190) | (3.89) | (2.998e-06) | (7.070e-02) | (918) | (18.11) | (4.446e-06) | (7.071e-02) | (433) | (10.06) |

## 5. Conclusion

In this paper, another DY-type method is proposed for solving constrained systems of nonlinear monotone equations as well as signal reconstruction. By conducting singular-value-analysis of the iteration matrix associated with a proposed DY-type direction, as well as performing a restart when certain conditions are not satisfied, the proposed method addresses the major shortcomings of the classical DY method. Applying some mild assumptions, global convergence of the new method was proven. Some Numerical tests conducted on the new scheme and three others proved that the proposed scheme is effective. Lastly, the scheme was used to reconstruct noisy signals.

## Data availability

The data supporting the results are provided in the relevant section of the article.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

## Funding

This research was supported by the Institutional-Based Research (IBR) grant of the Tertiary Education Trust Fund (TETFund), Sule Lamido University, Kafin Hausa, Nigeria.

## References

[1] A. S. Halilu, A. Majumder, M.Y. Waziri & K. Ahmed, "Signal recovery with convex constrained nonlinear monotone equations through conjugate gradient hybrid approach", Mathematics with computers in simulation **187** (2021) 520. https://doi.org/10.1016/j.matcom.2021.03.020.

[2] A. S. Halilu, A. Majumder, M. Y. Waziri, A. M. Awwal & K. Ahmed, "On solving double direction methods for convex constrained monotone nonlinear equations with image restoration", Computational and applied mathematics **40** (2021) 1. https://doi.org/10.1007/s40314-021-01624-1.

[3] J. K. Liu & S. J. Li, "A projection method for convex constrained monotone nonlinear equations with applications", Computers and mathematics with applications **70** (2015) 2442. https://doi.org/10.1016/j.camwa.2015.09.014.

[4] W. W. Hager & H. Zhang, "A survey of nonlinear conjugate gradient methods", Pacific journal of optimization **2** (2006) 35. https://www.researchgate.net/publication/228389054_A_survey_of_nonlinear_conjugate_gradient_method.

[5] Y. Narushima & H. Yabe, "A survey of sufficient descent conjugate gradient methods for unconstrained optimization", SUT Journal of mathematics **50** (2014) 167. https://doi.org/10.55937/sut/1424782608.

[6] R. Fletcher & C. Reeves, "Function minimization by conjugate gradients", The computer journal **7** (1964) 149. https://doi.org/10.1093/comjnl/7.2.149.

[7] R. Fletcher, *Practical methods of optimization*, 2nd ed., John Wiley & Sons, New York, NY, 1982. https://doi.org/10.1137/1024028.

[8] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property", SIAM journal on optimization **10** (1999) 177. https://doi.org/10.1137/S105262349731899,.

[9] M. R. Hestenes & E. L. Stiefel, "Methods of conjugate gradients for solving linear systems", Journal of research of the national bureau of standards **49** (1952) 409. https://doi.org/10.6028/JRES.049.044.

[10] E. Polak & G. Ribi´ere, "Note Sur la convergence de directions conjugees", ESAIM: Mathematical Modelling and Numerical Analysis ' **3** (1969) 35. https://doi.org/10.1051/m2an/196903R100351.

[11] B. T. Polyak, "The conjugate gradient method in extreme problems", USSR Computational Mathematics and Mathematical Physics 9 (1969) 94. https://doi.org/10.1016/0041-5553(69)90035-4.

[12] Y. Liu, & C. Storey, "Efficient generalized conjugate gradient algorithms", Journal of optimization theory and applications **69** (1991) 129. https://doi.org/10.1007/BF00940464.

[13] N. Andrei, "A Dai–yuan conjugate gradient algorithm with sufficient descent and conjugacy conditions for unconstrained optimization", Applied mathematics letters **21** (2008) 165. https://doi.org/10.1016/j.aml.2007.05.002.

[14] N. Andrei, "Another hybrid conjugate gradient algorithm for unconstrained optimization", Numerical algorithms **47** (2008) 143. https://doi.org/10.1007/s11075-007-9152-9.

[15] D. H. Li & M. Fukushima, "A modified BFGS method and its global convergence in nonconvex minimization", Journal of computational and applied mathematics **129** (2001) 15. https://doi.org/10.1016/S0377-0427(00)00540-9.

[16] Z. Wei, S. Yao, L. & Liu, "The convergence properties of some new conjugate gradient methods", Applied mathematics and computation **183** (2006) 1341. https://doi.org/10.1016/j.amc.2006.05.150.

[17] L. Zhang, W. Zhou & D. Li, "Global convergence of a modified Fletcher-reeves conjugate gradient method with Armijo-type line search", Numerische mathematik **104** (2006) 561. https://doi.org/10.1007/s00211-006-0028-z.

[18] L. Zhang, "Two modified Dai-yuan nonlinear conjugate gradient methods", Numerical algorithms **50** (2009) 1. https://doi.org/10.1007/s11075-008-9213-8.

[19] S. Babaie-Kafaki, "A hybrid conjugate gradient method based on a quadratic relaxation of the Dai–yuan hybrid conjugate gradient parameter", Optimization **62** (2013) 929. https://doi.org/10.1080/02331934.2011.611512.

[20] X. Jiang & J. Jian, "A sufficient descent Dai–yuan type nonlinear conjugate gradient method for unconstrained optimization problems", Nonlinear dynamics **72** (2013) 101. https://doi.org/10.1007/s11071-012-0694-6.

[21] G, Zhou & N. Qin, "A spectral Dai-yuan-type conjugate gradient method for unconstrained optimization", Mathematical problems in engineering **2015** (2015) 1. https://dx.doi.org/10.1155/2015/839659.

[22] X. Jiang, X. & J. Jian, "Improved fletcher-reeves and Dai-yuan conjugate gradient methods with the strong wolfe line search" Journal of computational and applied mathematics **348** (2018) 525. https://doi.org/10.1016/j.cam.2018.09.012.

[23] Z. Zhu, D. Zhang & S. Wang, "Two modified DY conjugate gradient methods for unconstrained optimization problems" Applied mathematics and computation **373** (2020) 125004. https://doi.org/10.1016/j.amc.2019.125004.

[24] S. K. Mishra, M. E. Samei, S. K. Chakraborty & B. Ram, "On q-variant of Dai–yuan conjugate gradient algorithm for unconstrained optimization problems", Nonlinear dynamics **104** (2021) 2471. https://doi.org/10.1007/s11071-021-06378-3.

[25] S. Narayanan & P. Kaelo, "A linear hybridization of Dai-Yuan and Hestenes-stiefel conjugate gradient method for unconstrained optimization", Numerical mathematics: theory, methods and application **14** (2021) 539. https://doi.org/10.4208/nmtma.OA-2020-0056.

[26] A. L. Yoksal, M. K. Hisham, M. N. Edrees, & K. A. Khalil, "A hybridization of the Hestenes-stiefel and Dai-yuan conjugate gradient methods", European journal of pure and applied mathematics **16** (2023) 1059. https://doi.org/10.29020/nybg.ejpam.v16i2.4746.

[27] J. K. Liu, & S. J. Li, "Spectral DY-type projection methods for nonlinear monotone system of equations", Journal of computational mathematics **33** (2015) 341. https://doi.org/10.4208/jcm.1412-m4494.

[28] J. M. Barzilai & M. Borwein, "Two point step size gradient methods", IMA journal of numerical analysis **8** (1988) 141. https://doi.org/10.1093/imanum/8.1.141.

[29] M. V. Solodov & B. F. Svaiter, "A globally convergent inexact newton method for systems of monotone equations", in: M. Fukushima, L. Qi (Eds.), reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods, Kluwer academic publishers **22** (1998) 355. https://doi.org/10.1007/978-1-4757-6388-1-18.

[30] J. Liu & S. Li, "Multivariate spectral projection method for convex constrained nonlinear monotone equations", Journal of industrial and management optimization **13** (2017) 283. https://doi.org/10.3934/jimo.2016017.

[31] G. H. Yu, S. Z. Niu & J.H. Ma, "Multivariate spectral gradient projection method for non-linear monotone equations with convex constraints", Journal of industrial management and optimization **9** (2013) 117. https://doi.org/10.3934/jimo.2013.9.117.

[32] Q. N. Li & D. H. Li, "A class of derivative-free methods for large-scale nonlinear monotone equations", IMA journal of numerical analysis **31** (2011) 1625. https://doi.org/10.1093/imanum/drq015.

[33] J. Liu & Y. Feng, "A derivative-free iterative method for nonlinear monotone equations with convex constraints", Numerical algorithms **82** (2019) 245. https://doi.org/10.1007/s11075-018-0603-2.

[34] S. Aji, P. Kumam, A. M. Awwal, M. M. Yahaya & K. Sitthithakerngkiet, "An efficient DY-type spectral conjugate gradient method for system of nonlinear monotone equations with application in signal recovery", AIMS Mathematics **6** (2021) 8078. https://doi.org/10.3934/math.2021469.

[35] A. Althobaiti, J. Sabi'u, H. Emadifar, P. Junsawang & S. K. Sahoo, "A scaled Dai–Yuan projection-based conjugate gradient method for solving monotone equations with applications", Symmetry **14** (2022) 1401. https://doi.org/10.3390/sym14071401.

[36] M. Y. Waziri, K. Ahmed & A. S. Halilu, "Adaptive three-term family of conjugate residual methods for system of monotone nonlinear equations", São Paulo Journal of mathematical sciences **16** (2023) 957. https://doi.org/10.1007/s40863-022-00293-0.

[37] K. Ahmed, M. Y. Waziri, A. S. Halilu, & S. Murtala, "Sparse signal reconstruction via Hager-zhang-type schemes for constrained system of nonlinear equations", Optimization **73** (2024) 1949. https://doi.org/10.1080/02331934.2023.2187255.

[38] A. S. Halilu, A. Majumder, M. Y. Waziri, K. Ahmed & A. M. Awwal, "Motion control of the two joint planar robotic manipulators through accelerated Dai-liao method for solving system of nonlinear equations", Engineering computations **39** (2022) 1802. https://doi.org/10.1108/EC-06-2021-0317.

[39] A. I. Kiri, M. Y. Waziri & K. Ahmed, "A modified Liu-storey scheme for nonlinear systems with an application to image recovery", Iranian journal of numerical analysis and optimization (2022) 1. https://doi.org/10.22067/ijnao.2022.75413.1107.

[40] M. Y. Waziri, K. Ahmed & J. Sabi'u, "A family of Hager–Zhang conjugate gradient methods for system of monotone nonlinear equations", Applied mathematics and computation **361** (2022) 645. https://doi.org/10.1016/j.amc.2019.06.012.

[41] M. Y. Waziri, K. Ahmed, A.S. Halilu & J. Sabi'u, "Two new Hager-Zhang iterative schemes with improved parameter choices for monotone nonlinear systems and their applications in compressed sensing", Rairo operations research **56** (2021) 239. https://doi.org/10.1051/ro/2021190.

[42] M. Y. Waziri, K. Ahmed & J. Sabi'u, "A Dai-Liao conjugate gradient method via modified secant equation for system of nonlinear equations", Arabian journal of mathematics 9 (2021) 443. https://doi.org/10.1007/s40065-019-0264-6.

[43] M. Y. Waziri, K. Ahmed, J. Sabi'u & A.S. Halilu, "Enhanced Dai-Liao conjugate gradient methods for systems of monotone nonlinear equations", SeMA Journal **78** (2020) 15. https://doi.org/10.1007/s40324-020-00228-9.

[44] K. Ahmed, M. Y. Waziri, A. S. Halilu & S. Murtala, "Image recovery via modified Dai–Kou method for constrained system of nonlinear equations", Journal of computational and applied mathematics **484** (2026) 117537. https://doi.org/10.1016/j.cam.2026.117537.

[45] 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", Numerical algorithms **88** (2021) 389. https://doi.org/10.1007/s11075-020-01043-z.

[46] P. Thakur & P. Gulial, "Nonlinear thermo-mechanics of smart functionally graded rotating disks for adaptive hydromechatronics", International journal of hydromechatronics. https://doi.org/10.1504/IJHM.2025.10073599.

[47] F. Arenas, R. Pérez, M. Gonzalez-Lima & C. A. Arias, "A centered newton method for nonlinear complementarity problem", Journal of computational and applied mathematics **484** (2026) 117473. https://doi.org/10.1016/j.cam.2026.117473.

[48] J. Puhan & A. Burmen "On the optimal parameter values of the Nelder–mead simplex algorithm", Journal of computational and applied mathematics **484** (2026) 117521. https://doi.org/10.1016/j.cam.2026.117521.

[49] X. L. Dong, H. W. Liu, Y. B. He & X. M. Yang, "A modified Hestenes-stiefel conjugate gradient method with sufficient descent and conjugacy condition", Journal of computational and applied mathematics **281** (2015) 239. https://doi.org/10.1016/j.cam.2014.11.058.

[50] Z. Aminifard & S. Babaie-Kafaki, "A modified descent Polak-ribiere-polyak conjugate gradient method with global convergence property for nonconvex functions", Calcolo **56** (2019) 1. https://doi.org/10.1007/s10092-019-0312-9.

[51] W. Sun & Y. H. Yuan, "Optimization theory and methods: Nonlinear programming", New York: Springer, (2006). https://doi.org/10.1007/b106451.

[52] D. S. Watkins, "Fundamentals of matrix computations", Wiley, New York, 2000. https://doi.org/10.1002/0471249718.

[53] A. B. Abubakar, P. Kumam, H. Mohammad, A H. Ibrahim, T. Seangwattana & B. A. Hassan, "A hybrid BFGS-Like method for monotone operator equations with applications", Journal of computational and applied mathematics **446** (2024) 115857. https://doi.org/10.1016/j.cam.2024.115857.

[54] L. Pengjie, "A three-term derivative-free projection method with BFGS-like update and its accelerated variant", Optimization, **75** (2026) 2055. https://doi.org/10.1080/02331934.2025.2526727.

[55] M. Rehman, J. Sabi'u, M. Sohaib & A. Shah, "A projection-based derivative free DFP approach for solving system of nonlinear convex constrained monotone equations with image restoration applications", Journal of applied mathematics and computation **69** (2023) 3645. https://doi.org/10.1007/s12190-023-01897-1.

[56] W. L. Cruz, "A Spectral algorithm for large-scale systems of nonlinear monotone equations", Numerical algorithms **76** (2017) 1109. https://doi.org/10.1007/s1107-017-0299-8.

[57] W. L. Cruz, J. M. Martinez & M. Raydan, "Spectral residual method without gradient information for solving large-scale nonlinear systems of equations", theory and experiments, Technical Report RT-04-08, **75** (2004) 1429. https://www.ams.org/journals/mcom/2006-75-255/S0025-5718-06-01840-0/S0025-5718-06-01840-0.pdf

[58] P. T. Gao & C. J. He, "An efficient three-term conjugate gradient method for nonlinear monotone equations with convex constraints" Calcolo **55** (2018) 1. https://doi.org/10.1007/s10092-018-0291-2.

[59] E. D. Dolan & J. J. More, "Benchmarking optimization software with performance profiles", Mathematical programming **91** (2002) 201. https://doi.org/10.1007/s101070100263.

[60] E. J. Candes & T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies", Institute of electrical and electronics engineers transactions on information theory **52** (2006) 5406. https//doi.org/10.1109/TIT.2006.885507.

[61] D. L. Donoho, "Compressed sensing", Institute of electrical and electronics engineers transactions on information theory **52** (2006) 1289. https://dx.doi.org/10.1109/TIT.2006.871582.

[62] T. Elaine, Y. Wotao & Z. Yin, "A fixed-point continuation method for $\ell_1$–regularized minimization with applications to compressed sensing", CAAM TR07-07, Rice University, (2007) 43. https://www.researchgate.net/publication/238445979_A_fixed-point_continuation_method_for_'1-Regularized_minimization_with_applications_to_compressed_sensing.

[63] A. T. Mario, R. Figueiredo & D. Nowak, "An EM algorithm for wavelet-based image restoration", IEEE transactions on image processing. **12**(2003) 906. https://doi.org/10.1109/TIP.2003.814255.

[64] J. S. Pang, "Inexact Newton methods for the nonlinear complementarity problem", Mathematical programming **36** (1986) 54. https://doi.org/10.1007/BF02591989.

[65] M.A.T. Figueiredo, R. Nowak & S. J. Wright, "Gradient projection for sparse reconstruction, application to compressed sensing and other inverse problems", IEEE J-STSP IEEE press, Piscataway, NJ. (2007) 586. https://doi.org/10.1137/070698920.

[66] Y. Xiao, Q. Wang & Q. Hu, "Non-smooth equations based method for $\ell_1 - norm$ problems with applications to compressed sensing", Nonlinear analysis theory methods and applications **74** (2011) 3570. https://doi.org/10.1016/j.na.2011.02.040.

[67] A. B. Abubakar, P. Kumam, H. Mohammad, A. M. Awwal & K. Sitthithakerngkiet, "A modified fletcher–reeves conjugate gradient method for monotone nonlinear equations with some applications", Mathematics **7** (2022) 745. https://doi.org/10.3390/math7080745.

[68] Y. Xia, X. Ma & D. Li, "An improved LS-RMIL-type conjugate gradient projection algorithm for systems of nonlinear equations and impulse noise image restoration", AIMS mathematics **10** (2025) 13640. https://doi.org/10.3934/math.2025614.

[69] Z. Wan, J. Guo, J. Liu & W. Liu, "A modified spectral conjugate gradient projection method for signal recovery", Signal, image and video processing **12** (2018) 1455. https://doi.org/10.1007/s11760-018-1300-2.