Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(8): 19675–19692.
DOI: 10.3934/math.2025877
Received: 18 May 2025
Revised: 01 August 2025
Accepted: 13 August 2025
Published: 27 August 2025

Research article

Sparse signal recovery through a modified Dai-Yuan algorithm

Kabiru Ahmed[1,2], **Mohammed Yusuf Waziri**[1,2], **Mohammed A. Saleh**[3,*], **Abdulgader Z. Almaymuni**[3], **Abubakar Sani Halilu**[2,4,5,6], **Mohamad Afendee Mohamed**[5], **Sulaiman M. Ibrahim**[7,8], **Salisu Murtala**[2,9] **and Habibu Abdullahi**[2,4,6]

[1] Department of Mathematics, Bayero University, Kano, Nigeria

[2] Numerical Optimization Research Group, Bayero University, Kano, Nigeria

[3] Department of Cybersecurity, College of Computer, Qassim University, Saudi Arabia

[4] Department of Mathematics, Sule Lamido University, Kafin Hausa, Nigeria

[5] Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Campus Besut, 22200 Terengganu, Malaysia

[6] Mathematical Innovation and Applications Research Group, Sule Lamido University, Kafin Hausa, Nigeria

[7] School of Quantitative Science, Universiti Utara Malaysia, Sintok, Malaysia

[8] Faculty of Education and Arts, Sohar University, Sohar 311, Oman

[9] Department of Mathematics, Federal University, Dutse, Nigeria

* **Correspondence:** Email: m.saleh@qu.edu.sa.

Abstract: Sparse signal recovery is a concept that is not only central to compressed sensing problems, but also apparent in magnetic resonance imaging (MRI) problems, machine learning, as well as statistical inference. In each of these fields, the target is finding sparse solutions to linear systems of equations that are underdetermined or ill-conditioned. In this paper, an efficient modified Dai-Yuan conjugate gradient method that is globally convergent irrespective of the line search procedure employed was developed to reconstruct sparse signals in compressed sensing. Results of the experiments conducted show that the method is promising.

Keywords: compressed sensing; conjugate gradient method; sparse signals; Lipschitz condition; global convergence; signaling; algorithms
Mathematics Subject Classification: 65K10, 68U10

1. Introduction

Often times, scientists and engineers encounter scenarios that require the solutions to underdetermined linear system

$$\mathcal{A}x = h, \tag{1.1}$$

where $\mathcal{A} \in \mathbb{R}^{k\times n}$ is a linear operator, $h \in \mathbb{R}^k$ is an estimation, and $x \in \mathbb{R}^n$. In (1.1), $k \ll n$ and the solutions are infinitely many. As a regular feature, (1.1) appears in signal processing via compressed sensing (CS) techniques [1–3], where the sparest of the solutions to (1.1) are of paramount interest. To solve (1.1), we first consider the general optimization problem (P_H)

$$(P_H) : \min_x H(x) \ \text{ s.t. } \ \mathcal{A}x = h. \tag{1.2}$$

Taking $H(x)$ to be strictly convex like the squared $\ell_2 - norm$ namely, "$\|x\|_2^2$" ensures a unique but not sparse solution. On the other hand, selecting $H(x)$ to be the non-convex $\ell_0 - norm$ guarantees a sparse solution and the new model is formulated as

$$\min \|x\|_0 \ \text{ s.t. } \ \mathcal{A}x = h, \tag{1.3}$$

with $\|.\|_0$ denoting the components of the vector x that are non-zeros, $\mathcal{A}$ and h remain as defined above, and k is the number of measurements. The minimization problem (1.3) is mostly referred to as the ℓ_0 minimization problem [4]. Even though the $\ell_0 - norm$ is often referred to in the measurement of a vector's sparsity, its implementation in the model (1.3) is combinatorial and computationally intractable [5]. To address this issue, researchers considered $H(x)$ to be the $\ell_1 - norm$, which is not only convex but also ensures sparse solutions to the problem in (1.1). Thus, to obtain the sparest solution to problem (1.1), researchers mostly consider the regularized convex optimization problem

$$\min_{x\in\mathbb{R}^n} H(x) := \frac{1}{2}\|\mathcal{A}x - h\|_2^2 + \tau\|x\|_1, \tag{1.4}$$

where the term $\|x\|_1$, which denotes the ℓ_1-norm, is incorporated to penalize large values in the objective function, while $\tau > 0$ serves as a parameter that regulates the balance between achieving sparsity and minimizing the residual error. Clearly, the objective function $H(x)$ in (1.4) is convex and nonsmooth. Nonetheless, suitable iterative schemes [6, 7] have been suggested in the past decades to obtain its solution. The scheme by Figueiredo et al. [6] is by far the most prominent method applied to solve (1.4). The main approach involves the reformulation of (1.4) by splitting $x \in \mathbb{R}^n$ as

$$x = a - b, \ \ a, b \in \mathbb{R}^n_+,$$

with $b_i = (-x_i)_+$, $a_i = (x_i)_+$, $\forall i = 1, 2, ..., n$, where $(x)_+ = \max\{0, x\}$. By employing this expression, the $\ell_1 - norm$ is further represented as $\|x\|_1 = e_n^T a + e_n^T b$, with $e_n = (1, 1, ..., 1)^T \in \mathbb{R}^n$, leading to the reformulation of (1.4) as

$$\min_{a,b} \left\{ \frac{1}{2}\|\mathcal{A}(a - b) - h\|_2^2 + \tau(e_n^T a + e_n^T b)| \ a, b \in \mathbb{R}^n_+ \right\}. \tag{1.5}$$

In addition, the authors in [6] showed that (1.5) can further be expressed as

$$\min_{\psi} \left\{ \frac{1}{2}\psi^T B\psi + \Omega^T\psi| \ \psi \in \mathbb{R}^{2n}_+ \right\}, \tag{1.6}$$

where

$$\psi = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \Omega = \tau e_{2n} + \begin{pmatrix} -q \\ q \end{pmatrix}, \quad q = \mathcal{A}^T h, \quad B = \begin{pmatrix} \mathcal{A}^T\mathcal{A} & -\mathcal{A}^T\mathcal{A} \\ -\mathcal{A}^T\mathcal{A} & \mathcal{A}^T\mathcal{A} \end{pmatrix}.$$

Considering that B is positive semi-definite, (1.6) is a convex quadratic programming problem [8]. The authors in [8] also stated that ψ is the solution to (1.6) provided it solves the following system of nonlinear equations:

$$F(\psi) = \min\{\psi, B\psi + \Omega\} = 0, \;\; \psi \in \mathbb{R}^{2n}_{+}. \tag{1.7}$$

Remark 1.1. *It was proven in [8] that F in* (1.7) *is monotone, namely the function F satisfies*

$$(F(x) - F(y))^T (x - y) \geq 0, \;\; \forall x, y \in \mathbb{R}^{2n}. \tag{1.8}$$

Among the various algorithms for solving (1.7), the Newton- and quasi-Newton-type algorithms [9, 10] are arguably the most prominent. This stems from the fact that they exhibit rapid convergence from initial points that are within a neighborhood of the solution point [11, 12]. However, the huge requirement for storage space and other tasking computations makes them not ideal for large-dimension problems. In the past decades, modified versions of famous conjugate gradient (CG) formulas for the unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} f(x), \tag{1.9}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is smooth with gradient $\nabla f(x) = g(x)$, have been employed to solve (1.7) [13]. This is mainly because the CG schemes do not require huge memory space to be implemented. Starting with a guess of $x_0 \in \mathbb{R}^n$, iterates of the CG method are generated by the formulas

$$x_{k+1} = x_k + s_k, \;\; s_k = \vartheta_k d_k, \;\; k \geq 0, \tag{1.10}$$

where x_k represents the k^{th} iterative point, $\vartheta_k > 0$ is a steplength generated by a line search process along the scheme's direction d_k which is computed using the following formulas:

$$d_0 = -g_0, \;\; d_{k+1} = -g_{k+1} + \beta_k d_k, \;\; k = 0, 1, ..., \tag{1.11}$$

with $g_{k+1} = g(x_{k+1})$ and β_k denoting the CG updating parameter which is crucial in the scheme's implementation and is different for each method. The reader should see the comprehensive survey in [14, 15] for different CG methods with different parameters β_k. Below is a list of the classical β_k parameters:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \text{ [16]}, \;\; \beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k} \text{ [17]}, \;\; \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T (g_{k+1} - g_k)} \text{ [18]}, \tag{1.12}$$

$$\beta_k^{HS} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T (g_{k+1} - g_k)} \text{ [19]}, \;\; \beta_k^{PRP} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{\|g_k\|^2} \text{ [20,21]}, \;\; \beta_k^{LS} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{-d_k^T g_k} \text{ [22]}, \tag{1.13}$$

where $\|.\|$ is the $\ell_2 - norm$ and $g_k = g(x_k)$. In recent decades, CG methods for solving (1.7) have also been employed in solving compressed sensing problems (see [23, 24] for instances). Apart from CG methods, some other solvers have also been developed for solving compressed sensing problems. These includes the iterative shrinkage/thresholing algorithm (ISTA) [25], which is a scheme that depends only on the gradient information, making it computationally cheap, the fast iterative shrinkage/thresholing algorithm (FISTA) [26], which is an improvement of ISTA by way of incorporating a momentum term for faster convergence, the alternating direction method of multipliers (ADMM) [27], which is a scheme with the attribute of handling complex functions, and alternating projection gradient linearization (APGL) [28], which recovers sparse signals by combining the principle of alternating projections and ℓ_1 minimization.

An essential version of CG methods that has not gained much attention for solving (1.7) is the modified form of the classical Dai-Yuan (DY) method in (1.12). Recently, some DY-type schemes for (1.7) have been proposed. One of these is the scheme by Liu and Li [29], where their formula was constructed by integrating the famous DY coefficient [18], the projection method by Solodov and Svaiter [30], and the spectral gradient scheme [31]. The authors also proved the convergence of the method independent of the well-known differentiability assumption. In [32], Liu and Li proposed a spectral DY-type formula for (1.7) which was constructed by combining the classical DY formula [18] with the technique in [33]. Inspired by the work in [34] and [29], Liu and Feng [35] defined a variant of DY method for constrained nonlinear systems. The method's structure makes it ideal for solving nonsmooth nonlinear problems. The convergence analysis of the scheme was discussed under the Lipschitz condition. Motivated by the idea from [35], Sani et al. [36] proposed a variant DY algorithm for constrained nonlinear systems, with the scheme's direction obtained as a combination of the classical CD and DY coefficients. Recently, Alhobaiti et al. [37] introduced two scaled DY-based schemes for the solution of (1.7) by employing distinct strategies to determine the scaling coefficient. Even though the DY-type methods are promising, numerical experiments and analysis have shown that they are quite sensitive to inexact line searches and often lack global convergence guarantees. The reader can explore the works in [23, 24, 38] for more on methods for constrained nonlinear systems. These new developments show that there is increasing interest in creating reliable and effective DY-based techniques that are suited to the complexity of constrained nonlinear situations. To improve the theoretical comprehension and practical usability of optimization algorithms in various scientific and technical fields, more study in this field is essential.

We outline the objectives of the work as follows:

- To construct a DY-type algorithm for constrained nonlinear systems.
- To present a method with the vital property for analyzing the convergence of CG-type methods.
- To analyze global convergence of the proposed scheme.
- To describe the method's application in sparse signals recovery.

The article is organized as follows: Motivation and details of the proposed method are given in Section 2. Analysis of its convergence are given in Section 3. The results of some numerical experiments to test effectiveness of the method in signal recovery are presented in Section 4. Conclusions are made in Section 5.

2. Motivation and algorithm

To begin this section, we first introduce an important concept that is required in implementing our algorithm and proving its global convergence. Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. The projection operator $P_C : \mathbb{R}^{2n} \to C$ is defined as

$$P_C(x) = \arg\min \|x - y\| : y \in C, \ \ \forall x \in \mathbb{R}^{2n},$$

with the properties

$$\|P_C(x) - P_C(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^{2n},$$

and

$$\|P_C(x) - y\| \leq \|x - y\|, \quad \forall y \in C. \tag{2.1}$$

We also give the following lemma that is equally required in proving global convergence of the proposed method:

Lemma 2.1. *(**Cauchy-Schwarz inequality**) Let V be an inner product space over the field of complex numbers $\mathbb{C}$ with inner product $\langle ., . \rangle$. For every pair of vectors $x, y \in V$, the following inequality holds:*

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle. \tag{2.2}$$

Theoretically, classical CG methods produce conjugate directions under the assumption of an exact line search and a convex quadratic objective function. However, in practical implementations where inexact line searches are employed, this condition generally fails to hold. For instance, if the coefficient HS is substituted into (1.11) and then both sides of the resulting equation are multiplied by g_{k+1}^T, we obtain

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \beta_k^{HS} g_{k+1}^T d_k. \tag{2.3}$$

Clearly, if $\beta_k^{HS} \geq 0$ and $g_{k+1}^T d_k \leq 0$ in (2.3), the descent condition holds. In contrast, when $g_{k+1}^T d_k > 0$ and $\beta_k^{HS} \geq 0$, the condition may not hold, since the term $\beta_k^{HS} g_{k+1}^T d_k$ can exceed $-\|g_{k+1}\|^2$. To address the shortcoming of the HS scheme, Dong et al. [39] introduced variants that ensure the condition is satisfied while retaining the desirable properties of the original method. They presented a new variant of HS search direction whose formula is described as follows:

$$d_{k+1}^D = \begin{cases} -\lambda_{k+1} g_{k+1} + \beta_k^D d_k, & g_{k+1}^T d_k > 0, \quad k \geq 0; \\ -g_{k+1} + \beta_k^{HS} d_k, & g_{k+1}^T d_k \leq 0, \quad k \geq 0, \ \text{otherwise}. \end{cases} \tag{2.4}$$

In (2.4), λ_{k+1} is given by

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

where

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

and

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

Inspired by this strategy, Aminifard and Babaie-Kafaki [40] introduced a closely related modification to the famous PRP scheme [20, 21]. Like the HS method, the PRP algorithm possesses a similar inherent structure but also fails in satisfying the descent property under inexact line searches, as previously observed for the HS algorithm. The direction proposed in [40] ensures the sufficient descent condition is met while preserving the favorable properties of the original PRP method, and is defined as follows:

$$d_0 = -g_0, \ \ d_{k+1}^M = \begin{cases} -\lambda_{k+1} g_{k+1} + \beta_k^{MPRP} d_k, & g_{k+1}^T d_k > 0, \quad k \geq 0; \\ -g_{k+1} + \beta_k^{PRP} d_k, & g_{k+1}^T d_k \leq 0, \quad k \geq 0, \end{cases} \tag{2.5}$$

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 (2.4), (2.5), and the original DY formula, we present a new DY-type search direction with the structure as follows:

$$d_0 = -F_0, \quad d_{k+1} = \begin{cases} -\lambda_{k+1} F_{k+1} + \beta_k^{NDY} d_k, & F_{k+1}^T d_k > 0, \quad k \geq 0; \\ -F_{k+1} + \beta_k^{MDY} d_k, & F_{k+1}^T d_k \leq 0, \quad k \geq 0, \text{ otherwise}, \end{cases} \tag{2.6}$$

where

$$\lambda_{k+1} = 1 + \frac{F_{k+1}^T d_k}{\|F_{k+1}\|^2}\beta_k^{MDY}, \quad \beta_k^{NDY} = \left(1 - \frac{F_{k+1}^T s_k}{d_k^T \bar{y}_k}\right)\beta_k^{MDY} - t\frac{\|F_{k+1}\|^2 F_{k+1}^T s_k}{(d_k^T \bar{y}_k)^2}, \tag{2.7}$$

with $t > 0$ being a positive parameter. Also, β_k^{MDY} in (2.7) is defined as

$$\beta_k^{MDY} = \frac{\|F_{k+1}\|^2}{d_k^T \bar{y}_k},$$

and

$$\bar{y}_k = y_k + \gamma\frac{\|F_{k+1}\| s_k}{\|s_k\|}, \quad \gamma > 0, \tag{2.8}$$

$$y_k = F(\psi_k) - F(x_k),$$

where

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

From (2.8) and the monotonicity of F, we obtain

$$d_k^T \bar{y}_k = \frac{s_k^T y_k}{\vartheta_k} + \frac{\gamma}{\vartheta_k}\frac{\|F_{k+1}\|}{\|s_k\|}\|s_k\|^2 \geq \frac{\gamma}{\vartheta_k}\|F_{k+1}\|\|s_k\| > 0, \tag{2.9}$$

which consequently implies that

$$s_k^T \bar{y}_k \geq \gamma\|F_{k+1}\|\|s_k\| > 0. \tag{2.10}$$

Lemma 2.2. *The search directions generated from* (2.6)–(2.8) *satisfy the inequality*

$$d_{k+1}^T F_{k+1} \leq -c\|F_{k+1}\|^2, \quad c = 1. \tag{2.11}$$

Proof. It is clear from (2.6) that for $k = 0$, $d_0^T F_0 = -\|F_0\|^2$. Then for $k = 1, 2, ...$, we consider the two cases presented in (2.6).

Case 1. $F_{k+1}^T d_k > 0$. Pre-multiplying both sides of (2.6) by F_{k+1} yields

$$\begin{aligned} d_{k+1}^T F_{k+1} &= -\lambda_{k+1}\|F_{k+1}\|^2 + \beta_k^{NDY} F_{k+1}^T d_k \\ &= -\left(1 + \frac{\|F_{k+1}\|^2}{d_k^T \bar{y}_k}\frac{F_{k+1}^T d_k}{\|F_{k+1}\|^2}\right)\|F_{k+1}\|^2 + \left(1 - \frac{F_{k+1}^T s_k}{d_k^T \bar{y}_k}\right)\frac{\|F_{k+1}\|^2}{d_k^T \bar{y}_k}F_{k+1}^T d_k - t\frac{\|F_{k+1}\|^2 (F_{k+1}^T s_k)^2}{(d_k^T \bar{y}_k)^2} \\ &\leq -\|F_{k+1}\|^2 - \frac{\|F_{k+1}\|^2 (F_{k+1}^T d_k)}{d_k^T \bar{y}_k} + \left(1 - \frac{F_{k+1}^T s_k}{d_k^T \bar{y}_k}\right)\frac{\|F_{k+1}\|^2}{d_k^T \bar{y}_k}F_{k+1}^T d_k \\ &\leq -\|F_{k+1}\|^2. \end{aligned}$$

Case 2. $F_{k+1}^T d_k \le 0$. Considering that $||F_{k+1}||^2$ and $d_k^T \bar{y}_k$ are both greater than zero, pre-multiplying the direction in the second case by F_{k+1} yields

$$d_{k+1}^T F_{k+1} = -||F_{k+1}||^2 + \frac{||F_{k+1}||^2}{d_k^T \bar{y}_k} F_{k+1}^T d_k \le -||F_{k+1}||^2.$$

We now present the algorithm of the proposed method (see Figure 1).

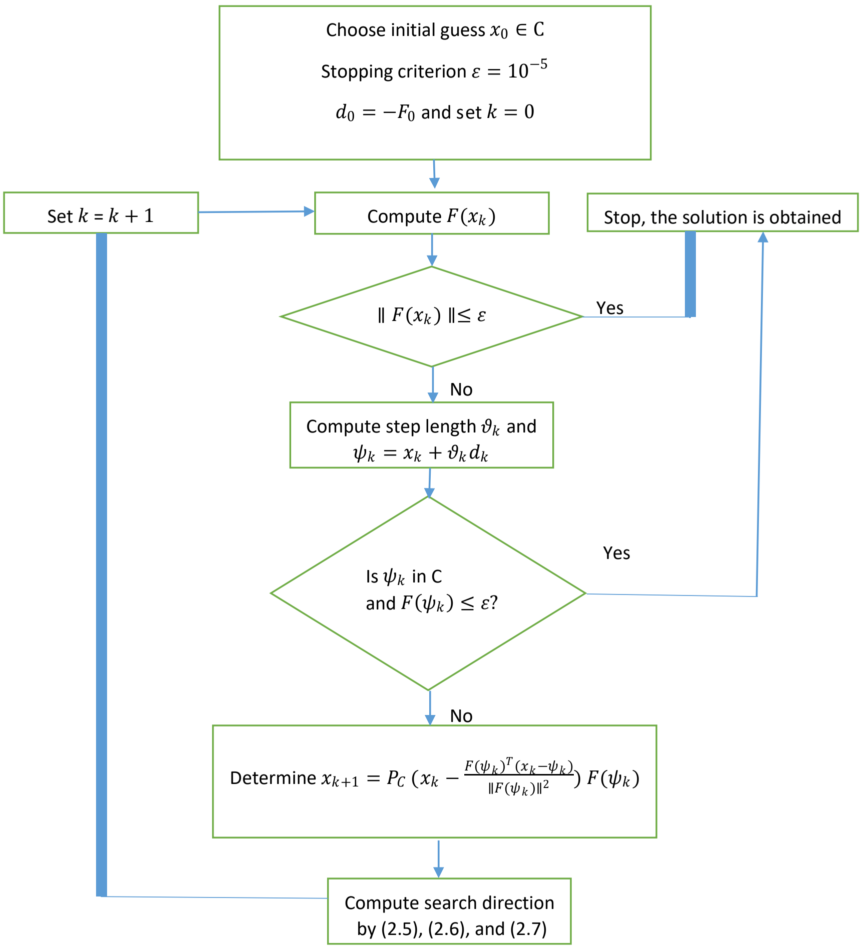

Figure 1. Flowchart representation of Algorithm 1.

Algorithm 1. Modified Dai-Yuan method

Input:

- Tolerance value $\epsilon > 0$.
- Initial starting point $x_0 \in C$.
- Line search and projection parameters $\beta \in (0, 1)$, $\delta \in (0, 1)$, $0 < \phi < 2$, $\gamma \in (0, +\infty)$.

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

1: Determine $F(x_k)$ and test if $||F(x_k)|| \le \epsilon$. If yes, stop. Otherwise go to step 2.

2: Calculate $\psi_k = x_k + \vartheta_k d_k$, with $\vartheta_k = \beta^{m_k}$, where m_k is the smallest nonnegative integer m satisfying

$$-F(x_k + \vartheta_k d_k)^T d_k \geq \delta\vartheta_k \|F(\psi_k)\| \|d_k\|^2. \tag{2.12}$$

3: If $\psi_k \in C$ and $\|F(\psi_k)\| \leq \epsilon$, then $x_{k+1} = \psi_k$.
End.
Otherwise, compute:

$$x_{k+1} = P_C\left[x_k - \phi\Phi_k F(\psi_k)\right], \quad \text{where} \tag{2.13}$$

$$\Phi_k = \frac{F(\psi_k)^T(x_k - \psi_k)}{\|F(\psi_k)\|^2}. \tag{2.14}$$

4: Determine d_{k+1} by (2.6) with (2.7) and (2.8).

5: Set $k = k + 1$ and go to 1.

3. Convergence results

To analyze the convergence of Algorithm 1, the following useful assumptions are required:

Assumption A. A solution $\bar{x} \in C$ exists such that $F(\bar{x}) = 0$.

Assumption B. The mapping F is Lipschitz continuous; namely, there exists a positive constant L such that

$$\|F(x) - F(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathbb{R}^{2n}. \tag{3.1}$$

The following lemma is used to prove that the linesearch procedure computed in Step 2 of Algorithm 1 is well-defined:

Lemma 3.1. *(1) Let $\{d_k\}$ and $\{x_k\}$ be generated by Algorithm* 1*. Also, let F be continuous on $\mathbb{R}^n$. Then, a non-negative integer m_k exists satisfying* (2.12)*.*

(2) Suppose Assumption B holds, where $\{x_k\}$ and $\{\psi_k\}$ follow from Algorithm 1*. Then the step-size $\vartheta_k > 0$ satisfies*

$$\vartheta_k \geq \min\left\{1, \frac{\beta\|F_k\|^2}{\|F(\bar{\psi}_k)\|(L + \delta)\|d_k\|^2}\right\}, \tag{3.2}$$

with $\bar{\psi}_k = x_k + \bar{\vartheta}_k d_k$ and $\bar{\vartheta}_k = \beta^{-1}\vartheta_k$.

Proof. Part one of the proof is omitted because it is similar to the proof in [41].

As for (2), if $\vartheta_k \neq 1$ in (2.12), $\bar{\vartheta}_k = \beta^{-1}\vartheta_k$ will not satisfy (2.12), i.e.,

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

Using (3.1) and (2.11), we get

$$\begin{aligned}\|F_k\|^2 &\leq -F_k^T d_k = (F(\bar{\psi}_k) - F_k)^T d_k - F(\bar{\psi}_k)^T d_k \\ &\leq \bar{\vartheta}_k(L + \delta\|F(\bar{\psi}_k)\|)\|d_k\|^2 \\ &= \beta^{-1}\vartheta_k(L + \delta\|F(\bar{\psi}_k)\|)\|d_k\|^2,\end{aligned}$$

which further yields

$$\vartheta_k \geq \frac{\beta\|F_k\|^2}{(L + \delta\|F(\bar{\psi}_k)\|)\|d_k\|^2}.$$

Lemma 3.2. *Suppose that F satisfies* (1.8) *and* (3.1) *holds with $\bar{x} \in \bar{C}$. 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 and $\{x_k\}$ and $\{\psi_k\}$ are bounded.

Proof. The proof begins by showing that $\{x_k\}$ and $\{\psi_k\}$ are bounded sequences. Considering (2.12) and the definition of ψ_k, we have

$$(x_k - \psi_k)^T F(\psi_k) \geq \delta\|F(\psi_k)\|\|x_k - \psi_k\|^2. \tag{3.3}$$

From (1.8) and for all $\bar{x} \in \bar{C}$, we get

$$\begin{aligned}(x_k - \bar{x})^T F(\psi_k) &= (x_k - \psi_k)^T F(\psi_k) + (\psi_k - \bar{x})^T F(\psi_k)\\ &\geq (x_k - \psi_k)^T F(\psi_k) + (\psi_k - \bar{x})^T F(\bar{x})\\ &= (x_k - \psi_k)^T F(\psi_k).\end{aligned} \tag{3.4}$$

From (2.1), (2.13), (2.14), (3.3), and (3.4), we have

$$\begin{aligned}\|x_{k+1} - \bar{x}\|^2 &= \|P_C[x_k - \phi\Phi_k F(\psi_k)] - \bar{x}\|^2\\ &\leq \|x_k - \phi\Phi_k F(\psi_k) - \bar{x}\|^2\\ &= \|(x_k - \bar{x}) - \phi\Phi_k F(\psi_k)\|^2\\ &= \|x_k - \bar{x}\|^2 - 2\phi\Phi_k F(\psi_k)^T(x_k - \bar{x}) + \phi^2\Phi_k^2\|F(\psi_k)\|^2\\ &\leq \|x_k - \bar{x}\|^2 - 2\phi\Phi_k F(\psi_k)^T(x_k - \psi_k) + \phi^2\Phi_k^2\|F(\psi_k)\|^2\\ &= \|x_k - \bar{x}\|^2 - \phi(2-\phi)\frac{(F(\psi_k)^T(x_k - \psi_k))^2}{\|F(\psi_k)\|^2}\\ &\leq \|x_k - \bar{x}\|^2 - \phi(2-\phi)\frac{\delta^2\|F(\psi_k)\|^2\|x_k - \psi_k\|^4}{\|F(\psi_k)\|^2}\\ &= \|x_k - \bar{x}\|^2 - \phi(2-\phi)\delta^2\|x_k - \psi_k\|^4,\end{aligned} \tag{3.5}$$

which further implies that

$$0 \leq \|x_{k+1} - \bar{x}\| \leq \|x_k - \bar{x}\| \leq \|x_{k-1} - \bar{x}\| \leq ... \leq \|x_0 - \bar{x}\|.$$

So, $\{\|x_k - \bar{x}\|\}$ is non-increasing, bounded below and therefore convergent. It also implies that $\{x_k\}$ is bounded. Now, by (3.1) and since $\|x_{k+1} - \bar{x}\| \leq \|x_k - \bar{x}\|$, we have

$$\|F(x_k)\| = \|F(x_k) - F(\bar{x}\| \leq L\|x_k - \bar{x}\| \leq L\|x_0 - \bar{x}\|. \tag{3.6}$$

Letting $L\|x_0 - \bar{x}\| = m_1$ indicates that $\{F_k\}$ is bounded. Also, from (3.3) and (2.2), we obtain

$$\begin{aligned}\delta\|F(\psi_k)\|\|x_k - \psi_k\|^2 &\leq (x_k - \psi_k)^T F(\psi_k)\\ &\leq \|F(\psi_k)\|\|x_k - \psi_k\|,\end{aligned}$$

which eventually yields

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

and indicates that $\{\psi_k\}$ is bounded. Hence, from (3.1) we conclude there exists $\bar{m}$ such that

$$\|F(\psi_k)\| \leq \bar{m}. \tag{3.7}$$

We now prove that $\{d_k\}$ is bounded, namely, there exists M such that

$$\|d_{k+1}\| \leq M. \tag{3.8}$$

From (2.6) and (2.7), we analyze two cases.

Case 1. $F_{k+1}^T d_k > 0$. By employing (2.6), (2.7), (2.9), and (2.2), we obtain

$$\begin{aligned}
\|d_{k+1}\| &= \| - \lambda_{k+1} F_{k+1} + \beta_k^{NDY} d_k\| \\
&\leq \left(1 + \frac{\|F_{k+1}\|\|d_k\|}{d_k^T \bar{y}_k}\right)\|F_{k+1}\| + \left(1 + \frac{\|F_{k+1}\|\|d_k\|}{d_k^T \bar{y}_k}\right)\frac{\|F_{k+1}\|^2}{d_k^T \bar{y}_k}\|d_k\| + t\frac{\|F_{k+1}\|^3\|d_k\|^2}{(d_k^T \bar{y}_k)^2} \\
&\leq \left(1 + \frac{\|F_{k+1}\|\|d_k\|}{\gamma\|F_{k+1}\|\|d_k\|}\right)\|F_{k+1}\| + \left(1 + \frac{\|F_{k+1}\|\|d_k\|}{\gamma\|F_{k+1}\|\|d_k\|}\right)\frac{\|F_{k+1}\|^2\|d_k\|}{\gamma\|F_{k+1}\|\|d_k\|} + t\frac{\|F_{k+1}\|^3\|d_k\|^2}{\gamma^2\|F_{k+1}\|^2\|d_k\|^2} \\
&= \left(1 + \frac{2}{\gamma} + \frac{2}{\gamma^2}\right)\|F_{k+1}\| \\
&\leq \left(1 + \frac{2}{\gamma} + \frac{1}{\gamma^2} + \frac{t}{\gamma^2}\right) m_1 \overset{\text{def}}{=} M_1.
\end{aligned} \tag{3.9}$$

Case 2. $F_{k+1}^T d_k \leq 0$. From (2.6) and (2.10), we have

$$\begin{aligned}
\|d_{k+1}\| &= \| - F_{k+1} + \frac{\|F_{k+1}\|^2}{s_k^T \bar{y}_k} s_k\| \\
&\leq \|F_{k+1}\| + \frac{\|F_{k+1}\|^2\|s_k\|}{\gamma\|F_{k+1}\|\|s_k\|} \\
&= \|F_{k+1}\| + \frac{\|F_{k+1}\|}{\gamma} = \left(1 + \frac{1}{\gamma}\right)\|F_{k+1}\| \\
&\leq \left(1 + \frac{1}{\gamma}\right) m_1 \overset{\text{def}}{=} M_2.
\end{aligned} \tag{3.10}$$

Setting $M = \max\{M_1, M_2\}$, we establish the proof. □

Lemma 3.3. *Let Assumptions A and B hold. Then*

$$\lim_{k\to\infty} \vartheta_k \|d_k\| = 0. \tag{3.11}$$

Proof. Combining the result in (3.5) with (3.7), we obtain

$$\delta^2 \|x_k - \psi_k\|^4 \leq \frac{\bar{m}^2}{\phi(2-\phi)}(\|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2). \tag{3.12}$$

Considering that $\{\|x_k - \bar{x}\|\}$ is convergent and $\{F(\psi_k)\}$ is bounded, taking the limit of the sum of both sides of (3.12) as k approaches infinity, we get

$$\delta^2 \lim_{k\to\infty} \vartheta_k^4 \|d_k\|^4 \leq 0,$$

which indicates that

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

The following theorem is used to prove global convergence of Algorithm 1:

Theorem 3.4. *Suppose Assumptions A and B hold, where the sequence $\{x_k\}$ is generated by Algorithm 1. Then,*

$$\liminf_{k\to\infty} \|F_k\| = 0. \tag{3.13}$$

Proof. Assume by contradiction that (3.13) is not true. Then, $m_2 > 0$ exists such that

$$\|F_k\| \geq m_2, \quad \forall k \geq 1. \tag{3.14}$$

By Cauchy-Schwarz inequality in (2.2), (2.11), and (3.14), we see that

$$\|d_k\| \geq \|F_k\| \geq m_2, \quad \forall k \geq 1. \tag{3.15}$$

Also, from (3.2), (3.7), and (3.8), we have

$$\begin{aligned}\vartheta_k \|d_k\| &\geq \min\left\{1, \frac{\beta\|F_k\|^2}{(L + \delta\|F(\bar{\psi}_k)\|)\|d_k\|^2}\right\}\|d_k\| \\ &\geq \min\left\{m_2, \frac{\beta m_2^2}{(L+\delta\bar{m})M}\right\} > 0,\end{aligned}$$

which is a clear contradiction with (3.11). Thus, (3.13) holds. □

Remark 3.5. *As stated in the introduction section, DY-type methods are sensitive to inexact line searches and do not always guarantee global convergence. The design of our method, which includes a restart mechanism, ensures global convergence irrespective of the line search procedure. Also, our method is derivative-free, which implies it does not require the computation of the gradient vector, unlike some of the methods for solving* (1.7) *discussed in the introduction section.*

4. Numerical experiments

To show the effectiveness of Algorithm 1, it was used with the methods in [42] (labelled PCG), [43] (labelled CGD), and [6] (labelled IST) to reconstruct a sparse signal with n lengths from some k observations. In the experiment, the mean of square error (MSE) to the original signal $\tilde{x}$, i.e.,

$$MSE = \frac{1}{n}\|\tilde{x} - \bar{x}\|_2^2,$$

is used in measuring the quality of recovery, where $\bar{x}$ stands for the signal restored. In the first experiment conducted, we take the size of the signal as $n = 2^{12}$ and the sparsity level as 2^7, implying it contains 128 randomly nonzero elements. We also take the number of measurement h as $k = 2^{10}$ with Gaussian noise (i.e., a random vector having normal distribution $N(0_k, \sigma^2 I_k)$), with 0_k representing the zero vector, I_k a k size identity matrix, and σ^2 is set as 10^{-4}. The measurement or estimation h is disturbed by noise, i.e.,

$$h = \mathcal{A}\bar{x} + \xi,$$

where ξ stands for the Gaussian noise. We also use $f(x) = \frac{1}{2}\|\mathcal{A}x - h\|_2^2 + \tau\|x\|_1$ for the merit function. The experiment was initiated with $x_0 = \mathcal{A}^T h$, which terminates whenever

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

where f_{k+1} represents the function value at x_{k+1}. Also, we set parameters of Algorithm 1 for the experiments as $\beta = 0.9$, $\delta = 0.01$, $\phi = 1.8$, $\gamma = 5.5$, and $t = 0.1$. Parameters for the PCG, CGD, and IST methods are set as used in each of the papers.

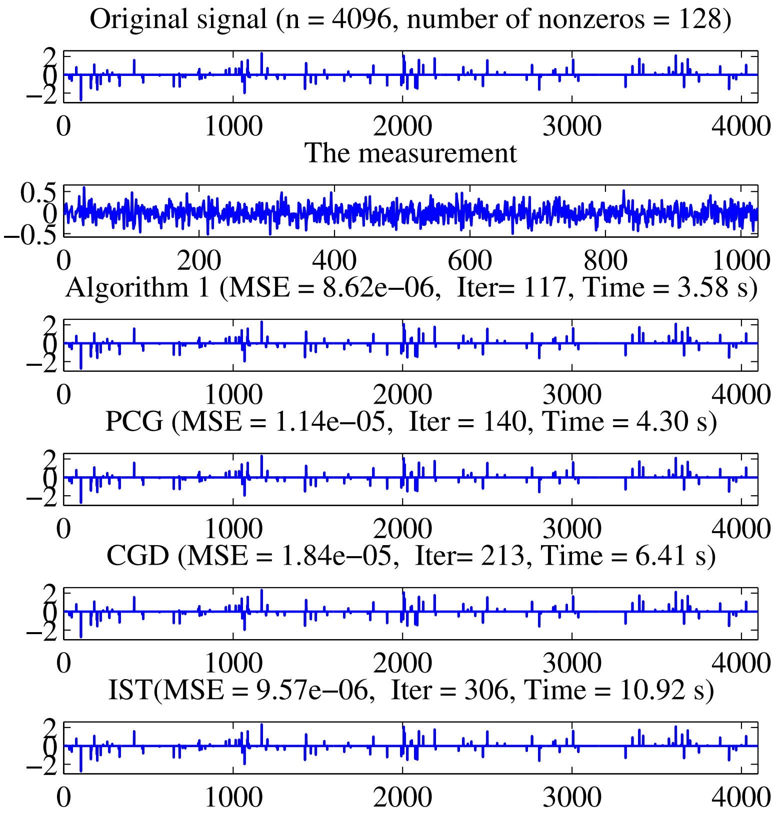

Figure 2. The original, measured, and recovered signals.

Figure 2 shows the original, estimated, and reconstructed signals by Algorithm 1, PCG, CGD, and IST. It also shows that the methods almost exactly reconstruct the signals from the estimation, where Algorithm 1 has the fastest time, less iterations, and MSE. Figure 3, on the other hand, displays four graphs that exhibit convergence behavior of the four algorithms through MSE and the objective function (ObjFun), as well as iterations and CPU time. Figure 3 also shows that the rate of descent for MSE and the objective function obtained for Algorithm 1 is much faster compared to the PCG, CGD, and IST methods. It also shows that Algorithm 1 requires less iterations and CPU time to reconstruct the original signal than PCG, CGD, and IST for the same process. It has been described in [44] that reducing the number of measurements without degrading reconstruction performance is an essential goal of the CS technique. With that in mind, we test the four algorithms with different sparsity levels and CS ratios given by $\varrho = \frac{k}{n}$, where k and n are as defined earlier.

In Table 1, which is made up of three tables labelled (a), (b), and (c), three signal reconstruction results are reported. In Table 1(a), the reconstruction results with CS ratios $\bar{\lambda}$ = 0.5, 0.25, and 0.125 and corresponding sparsity levels 128 and 32 (in brackets) are given. In Table 1(b), the reconstruction results with sparsity level 20 and a signal with length 500, as conducted in [45], are reported. Lastly, Table 1(c) displays the reconstruction results for different noise samples, sparsity levels, and standard deviations. In each of the Table 1(a), (b), and (c), the scheme with the highest number of underlined values is the most effective. Clearly, from Table 1(a), (b), and (c), Algorithm 1 records the highest number of underlined results. Thus, we conclude that Algorithm 1 is the most effective and was able to recover the sparse signals more efficiently while maintaining the quality of recovery.

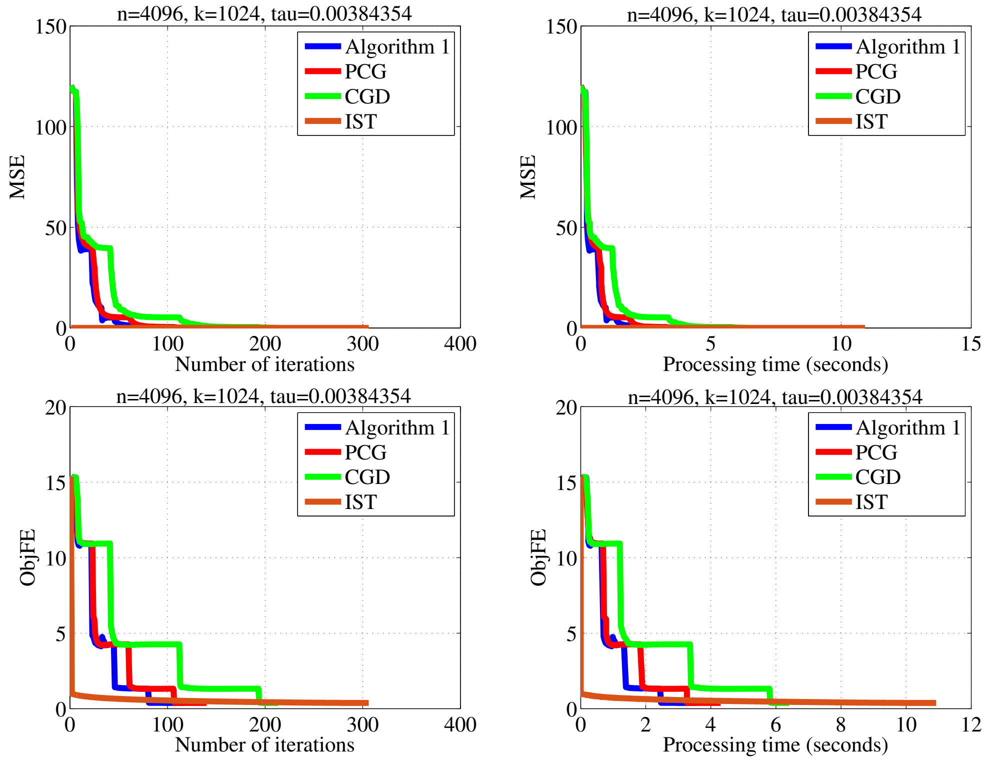

Figure 3. Comparison of the results for Algorithm 1, PCG, CGD, and IST solvers.

5. Conclusions

In this paper, an iterative method, which is employed to reconstruct sparse signals in compressed sensing, is presented. As a modified variant of the Dai-Yuan method, the new scheme addresses some of the shortcomings of the classical DY method. It also ensures the essential condition for the proof of global convergence holds for any line search strategy employed. It also converges globally under mild conditions. Furthermore, its effectiveness is shown by conducting some compressed sensing experiments, where different noise samples are analyzed.

Table 1. Three signal reconstruction results.

(a) Signal recovery results with sparsity level 128(32) under different CS ratios

CS ratio	Algorithm 1 MSE	ObjFun	Iter	Time	PCG MSE	ObjFun	Iter	Time	CGD MSE	ObjFun	Iter	Time	IST MSE	ObjFun	Iter	Time
0.5	4.67E-06	5.56E-01	47	3.05	4.79E-06	5.56E-01	57	3.61	4.72E-06	5.56E-01	69	4.41	4.79E-06	5.56E-01	63	3.64
	9.35E-07	1.36E-01	29	1.88	9.53E-07	1.36E-01	43	2.69	9.55E-07	1.36E-01	38	2.31	9.57E-07	1.36E-01	32	1.77
0.25	3.82E-06	1.87E-01	71	2.47	3.85E-06	1.87E-01	109	3.69	3.85E-06	1.87E-01	119	3.95	4.11E-06	1.87E-01	164	5.98
	1.26E-06	7.83E-02	69	2.22	1.29E-06	7.83E-02	112	3.72	1.30E-06	7.83E-02	123	4.02	1.37E-06	7.83E-02	150	5.63
0.125	3.80E-03	1.84E-01	85	1.66	3.01E-03	1.84E-01	284	5.56	4.23E-03	1.84E-01	273	5.11	6.33E-03	1.87E-01	1001	2.13
	7.62E-07	2.73E-02	168	2.98	8.08E-07	2.73E-02	212	3.92	1.10E-06	2.73E-02	243	4.42	8.73E-07	2.73E-02	666	15.64

(b) Signal recovery results for Algorithm 1, PCG, CGD, and IST under sparsity level 20 and signal with length 500

CS ratio	Algorithm 1 MSE	ObjFun	Iter	Time	PCG MSE	ObjFun	Iter	Time	CGD MSE	ObjFun	Iter	Time	IST MSE	ObjFun	Iter	Time
0.17	1.95E-03	2.60E-02	125	0.16	3.46E-03	2.68E-02	126	0.11	2.42E-03	2.62E-02	264	0.25	2.79E-03	2.62E-02	1001	1.48
0.20	1.27E-03	3.44E-02	141	0.16	2.55E-03	3.52E-02	183	0.31	2.11E-03	3.47E-02	221	0.28	1.26E-03	3.44E-02	1001	1.41
0.23	5.61E-04	2.85E-02	116	0.11	1.51E-03	2.92E-02	140	0.2	6.22E-04	2.85E-02	233	0.34	9.27E-04	2.87E-02	1001	2.27
0.26	7.56E-05	7.08E-02	68	0.03	2.60E-04	7.13E-02	139	0.13	1.70E-04	7.06E-02	193	0.25	1.00E-05	6.95E-02	460	0.61
0.29	4.04E-06	4.15E-02	91	0.41	5.99E-05	4.21E-02	123	0.16	1.91E-05	4.17E-02	177	0.16	5.01E-06	4.16E-02	407	0.28
0.32	7.63E-06	7.11E-02	77	0.06	7.78E-06	7.11E-02	124	0.13	1.01E-05	7.10E-02	138	0.16	8.26E-06	7.11E-02	236	0.33

(c) Signal recovery results for Algorithm 1, PCG, CGD and IST under different noise samples

σ^2	Algorithm 1 MSE	ObjFun	Iter	Time	PCG MSE	ObjFun	Iter	Time	CGD MSE	ObjFun	Iter	Time	IST MSE	ObjFun	Iter	Time
10^{-4}	3.23E-06	7.94E-02	74	0.81	3.32E-06	7.94E-02	122	1.36	4.78E-06	7.93E-02	104	1.72	3.55E-06	7.94E-02	188	2.81
10^{-6}	1.33E-06	5.42E-02	84	0.88	1.34E-06	5.42E-02	105	1.03	1.36E-06	5.42E-02	140	1.41	1.45E-06	5.42E-02	272	3.28
10^{-8}	1.30E-06	6.22E-02	73	0.67	1.99E-06	6.25E-02	120	1.20	2.01E-06	6.25E-02	153	1.44	2.11E-06	6.25E-02	210	2.72
10^{-10}	6.24E-06	1.18E-01	100	0.97	6.22E-06	1.18E-01	140	1.34	6.26E-06	1.18E-01	182	1.77	6.64E-06	1.18E-01	209	2.84

Author contributions

K. Ahmed: Conceptualization, Methodology, Investigation, Writing–original draft preparation; M. Y. Waziri: Methodology, Data curation, Writing–review and editing, Visualization, Supervision; M. A. Saleh: Software, Resources, Data curation, Writing–review and editing, Funding acquisition; A. Z. Almaymuni: Validation, Resources, Data curation, Writing–review and editing, Funding acquisition; A. S. Halilu: Formal analysis, Investigation, Writing–review and editing; M. A. Mohamed: Validation, Writing–review and editing, Visualization, Supervision; S. M. Ibrahim: Formal analysis, Investigation, Writing–review and editing; S. Murtala: Validation, Visualization, Supervision; H. Abdullahi: Software, Validation, Visualization, Supervision. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

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

Acknowledgments

The researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for the financial support (QU-APC-2025).

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. D. L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory*, **52** (2006), 1289–1306. https://doi.org/10.1109/TIT.2006.871582
2. E. J. Candes, T. Tao, Near-optimal signal recovery from random projections: universal encoding stragies, *IEEE Trans. Inform. Theory*, **52** (2006), 5406–5425. https://doi.org/10.1109/TIT.2006.885507
3. R. Arablouei, Fast reconstruction algorithm for perturbed compressive sensing based on total least-squares and proximal splitting, *Signal Process.*, **130** (2017), 57–63. https://doi.org/10.1016/j.sigpro.2016.06.009
4. H. F. Liu, J. G. Peng, Sparse signal recovery via alternating projection method, *Signal Process.*, **143** (2018), 161–170. https://doi.org/10.1016/j.sigpro.2017.09.003
5. S. Foucart, H. Rauhut, *A mathematical introduction to compressive sensing*, New York: Birkhäuser, 2013. https://doi.org/10.1007/978-0-8176-4948-7
6. M. A. T. Figueiredo, R. D. Nowak, S. J. Wright, Gradient projection for sparse reconstruction, application to compressed sensing and other inverse problems, *IEEE J. Sel. Top. Signal Process.*, **1** (2007), 586–597. https://doi.org/10.1109/JSTSP.2007.910281

7. Y. Nesterov, Gradient methods for minimizing composite functions, *Math. Program.*, **140** (2013), 125–161. https://doi.org/10.1007/s10107-012-0629-5

8. Y. H. Xiao, Q. Y. Wang, Q. J. Hu, Non-smooth equations based method for ℓ_1-norm problems with applications to compressed sensing, *Nonlinear Anal.*, **74** (2011), 3570–3577. https://doi.org/10.1016/j.na.2011.02.040

9. J. E. Dennis, Jr, J. J. Moré, Quasi-Newton methods, motivation and theory, *SIAM Rev.*, **19** (1977), 46–89. https://doi.org/10.1137/1019005

10. C. T. Kelly, *Iterative methods for optimization*, SIAM, 1999. https://doi.org/10.1137/1.9781611970920

11. J. M. Ortega, W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, New York: Academic Press, 1970. https://doi.org/10.1016/C2013-0-11263-9

12. J. E. Dennis, Jr, R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Englewood Cliffs: Prentice-Hall, 1983. https://doi.org/10.1137/1.9781611971200

13. M. Malik, M. Mamat, S. S. Abas, I. M. Sulaiman, A new spectral conjugate gradient method with descent condition and global convergence property for unconstrained optimization, *J. Math. Comput. Sci.*, **10** (2020), 2053–2069.

14. Y. Narushima, H. Yabe, A survey of sufficient descent conjugate gradient methods for unconstrained optimization, *SUT J. Math.*, **50** (2014), 167–203. https://doi.org/10.55937/sut/1424782608

15. W. W. Hager, H. Zhang, A survey of nonlinear conjugate gradient methods, *Pac. J. Optim.*, **2** (2006), 35–58.

16. R. Fletcher, C. M. Reeves, Function minimization by conjugate gradients, *Comput. J.*, **7** (1964), 149–154. https://doi.org/10.1093/comjnl/7.2.149

17. R. Fletcher, *Practical method of optimization*, John Wiley & Sons, 2000. https://doi.org/10.1002/9781118723203

18. Y. H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM J. Optim.*, **10** (1999), 177–182. https://doi.org/10.1137/S1052623497318992

19. M. R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bureau Standards*, **49** (1952), 409–436. http://dx.doi.org/10.6028/jres.049.044

20. E. Polak, G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, *Rev. Française D'informatique Rech. Opérationnelle Série Rouge*, **3** (1969), 35–43. https://doi.org/10.1051/m2an/196903R100351

21. B. T. Polyak, The conjugate gradient method in extreme problems, *USSR Comput. Math. Math. Phys.*, **9** (1969), 94–112. https://doi.org/10.1016/0041-5553(69)90035-4

22. Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms, part 1: Theory, *J. Optim. Theory Appl.*, **69** (1991), 129–137. https://doi.org/10.1007/BF00940464

23. K. Ahmed, M. Y. Waziri, A. S. Halilu, I. A. R. Moghrabi, H. Abdullahi, S. M. Ibrahim, et al., A Dai-Kou-type method with image de-blurring application, *J. Niger. Soc. Phys. Sci.*, **7** (2025), 2569. https://doi.org/10.46481/jnsps.2025.2569

24. K. Ahmed, M. Y. Waziri, A. S. Halilu, S. Murtala, H. Abdullahi, Signal and image reconstruction with a double parameter Hager-Zhang-type conjugate gradient method for system of nonlinear equations, *Numer. Linear Algebra Appl.*, **32** (2025), e2583. https://doi.org/10.1002/nla.2583

25. O. V. Michailovich, An iterative shrinkage approach to total-variation image restoration, *IEEE Trans. Image Process.*, **20** (2011), 1281–1299. https://doi.org/10.1109/TIP.2010.2090532

26. A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imaging Sci.*, **2** (2009), 183–202. https://doi.org/10.1137/080716542

27. S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends Mach. Learn.*, **3** (2011), 1–122. https://doi.org/10.1561/2200000016

28. H. X. Li, C. H. Shen, Q. F. Shi, Real-time visual tracking using compressive sensing, *CVPR 2011*, 2011, 1305–1312. https://doi.org/10.1109/CVPR.2011.5995483

29. J. K. Liu, S. J. Li, Spectral DY-type projection method for nonlinear monotone system of equations, *J. Comput. Math.*, **33** (2015), 341–355. https://doi.org/10.4208/jcm.1412-m4494

30. M. V. Solodov, B. F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations, In: *Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods*, New York: Springer, 1999, 355–369. https://doi.org/10.1007/978-1-4757-6388-1_18

31. J. Barizilai, J. M. Borwein, Two-point step size gradient methods, *IMA J. Numer. Anal.*, **8** (1988), 141–148. https://doi.org/10.1093/imanum/8.1.141

32. J. K. Liu, S. J. Li, Multivariate spectral DY-type projection method for convex constrained nonlinear monotone equations, *J. Indust. Manag. Optim.*, **13** (2017), 283–295. https://doi.org/10.3934/jimo.2016017

33. G. H. Yu, S. Z. Niu, J. H. Ma, Multivariate spectral gradient projection method for nonlinear monotone equations with convex constraints, *J. Indust. Manag. Optim.*, **9** (2013), 117–129. https://doi.org/10.3934/jimo.2013.9.117

34. Q. N. Li, D. H. Li, A class of derivative-free methods for large-scale nonlinear monotone equations, *IMA J. Numer. Anal.*, **31** (2011), 1625–1635. https://doi.org/10.1093/imanum/drq015

35. J. K. Liu, Y. M. Feng, A derivative-free itertive method for nonlinear monotone equations with convex constraints, *Numer. Algorithms*, **82** (2019), 245–262. https://doi.org/10.1007/s11075-018-0603-2

36. 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 Math.*, **6** (2021), 8078–8106. https://doi.org/10.3934/math.2021469

37. 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), 1–27. https://doi.org/10.3390/sym14071401

38. 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

39. 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, *J. Comput. Appl. Math.*, **281** (2015), 239–249. https://doi.org/10.1016/j.cam.2014.11.058

40. Z. Aminifard, S. Babaie-Kafaki, A modified descent Polak-Ribiere-Polyak conjugate gradient method with global convergence property for nonconvex functions, *Calcolo*, **56** (2019), 16. https://doi.org/10.1007/s10092-019-0312-9

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 Oper. Res.*, **56** (2022), 239–273. https://doi.org/10.1051/ro/2021190

42. J. K. Liu, S. J. Li, A projection method for convex constrained monotone nonlinear equations with applications, *Comput. Math. Appl.*, **70** (2015), 2442–2453. https://doi.org/10.1016/j.camwa.2015.09.014

43. Y. H. Xiao, H. Zhu, A conjugate gradient method to solve convex constrained monotone equations with applications in compressive sensing, *J. Math. Anal. Appl.*, **405** (2013), 310–319. https://doi.org/10.1016/j.jmaa.2013.04.017

44. Z. Wan, J. Guo, J. J. Liu, W. Y. Liu, A modified spectral conjugate gradient projection method for signal recovery, *Signal Image Video Process.*, **12** (2018), 1455–1462. https://doi.org/10.1007/s11760-018-1300-2

45. S. K. Ambat, K. V. S. Hari, An iterative framework for sparse signal reconstruction algorithms, *Signal Process.*, **108** (2015), 351–364. https://doi.org/10.1016/j.sigpro.2014.09.023

AIMS Press