Received: 17 November 2023 | Revised: 10 July 2024 | Accepted: 28 July 2024

DOI: 10.1002/mma.10396

RESEARCH ARTICLE

WILEY

An improved Dai-Liao-style hybrid conjugate gradient-based method for solving unconstrained nonconvex optimization and extension to constrained nonlinear monotone equations

Zihang Yuan[1] | **Hu Shao**[1] | **Xiaping Zeng**[2] | **Pengjie Liu**[1] | **Xianglin Rong**[3] | **Jianhao Zhou**[4]

[1]School of Mathematics, Jiangsu Center for Applied Mathematics, China University of Mining and Technology, Xuzhou, China

[2]Center for Applied Mathematics of Guangxi, Yulin Normal University, Yulin, China

[3]School of Computational Science and Electronics, Hunan Institute of Engineering, Xiangtan, China

[4]LSA mathematics, University of Michigan at Ann Arbor, Ann Arbor, Michigan, USA

Correspondence
Hu Shao, School of Mathematics, Jiangsu Center for Applied Mathematics, China University of Mining and Technology, Xuzhou 221116, China.
Email: shaohu@cumt.edu.cn

Xiaping Zeng, Center for Applied Mathematics of Guangxi, Yulin Normal University, Yulin 537000, China.
Email: xiaping062@163.com

Communicated by: P. Agarwal

Funding information
This work is supported by the National Natural Science Foundation of China (No. 72071202), the Natural Science Foundation of Guangxi Province (No. 2024GXNSFBA010395), the Enhancing Young and Middle-aged Teacher's Research Basic Ability in Colleges of Guangxi (No. 2021KY0592), and the Key Laboratory of Mathematics and Engineering Applications, Ministry of Education.

In this work, for unconstrained optimization, we introduce an improved Dai-Liao-style hybrid conjugate gradient method based on the hybridization-based self-adaptive technique, and the search direction generated fulfills the sufficient descent and trust region properties regardless of any line search. The global convergence is established under standard Wolfe line search and common assumptions. Then, combining the hyperplane projection technique and a new self-adaptive line search, we extend the proposed conjugate gradient method and obtain an improved Dai-Liao-style hybrid conjugate gradient projection method to solve constrained nonlinear monotone equations. Under mild conditions, we obtain its global convergence without Lipschitz continuity. In addition, the convergence rates for the two proposed methods are analyzed, respectively. Finally, numerical experiments are conducted to demonstrate the effectiveness of the proposed methods.

KEYWORDS

conjugate gradient method, global convergence, nonlinear monotone equations, unconstrained optimization

MSC CLASSIFICATION

65K0, 90C30, 90C90

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2024 The Author(s). *Mathematical Methods in the Applied Sciences* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

1.1 | Conjugate gradient method

In this subsection, the following unconstrained optimization problem is considered,

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Unconstrained optimization plays an important role in practical applications. For example, it is essential in machine learning, which is popular in many fields. Recently, Jandaghi et al. [1] proposed a novel machine learning approach to train a soft trunk robot model, giving a more accurate model estimation. When the scale of (1) is not large, the Newton method and quasi-Newton method are promising in solving it, and Javad Ebadi et al. [2] proposed a new BFGS method based on modified secant relations and verified its efficiency by applying it in solving (1). However, when solving (1) in large scale, the conjugate gradient (CG) method is more welcome since it requires lower memory in practical computation. Featured in its efficiency, CG method has been applied by many researchers in a substantial number of fields such as image restoration problem [3–8], compressive signals problem [9–12], and signal reconstruction problem [13]. In other practical fields, like disease model, optimization methods are used to solve fractional-order tuberculosis disease model [14] and fractional fascioliasis disease model [15], where the CG method can be considered to apply in the optimization problems appearing in the models. Regularly, the iterative process of a CG method starts from x_0 and generates the next iteration point by the recursive formula as follows:

$$x_{k+1} = x_k + l_k d_k,\ k = 0, 1, 2, \ \dots\ , \tag{2}$$

where $l_k > 0$ is the step length obtained by a chosen line search. The sequence of search direction $\{d_k\}$ is generated by

$$d_k = \begin{cases} -\omega_k, & k = 0, \\ -\omega_k + \beta_k d_{k-1}, & k \geq 1, \end{cases} \tag{3}$$

where $\omega_k = \omega(x_k) = \nabla f(x_k)$ is the gradient of f at point x_k and the scalar β_k is known as the conjugate coefficient.

Up to now, extensive studies have been established with the purpose to derive more effective CG methods with promising properties, and various CG methods have been proposed. It is a well-established fact that distinct CG methods vary mainly in the conjugate coefficients, which tend to bring different features in convergence and computation, respectively. Accordingly, classical CG methods are considered to be divided into two categories with respect to numerical performance and convergence property. Let $y_{k-1} = \omega_k - \omega_{k-1}$ and $\|\cdot\|$ be the Euclidean norm. The first category consists of the Fletcher–Reeves (FR) method [16], Dai–Yuan (DY) method [17], and conjugate descent (CD) method [18], where the conjugate coefficients are defined by

$$\beta_k^{\mathrm{FR}} = \frac{\|\omega_k\|^2}{\|\omega_{k-1}\|^2},\ \beta_k^{\mathrm{DY}} = \frac{\|\omega_k\|^2}{d_{k-1}^{\top} y_{k-1}},\ \beta_k^{\mathrm{CD}} = -\frac{\|\omega_k\|^2}{d_{k-1}^{\top}\omega_{k-1}}.$$

In general, it is certain for their global convergence. Unfortunately, these methods may exhibit limited numerical performances in practice due to the jamming phenomenon. Conversely, the convergence of the second category, including the Polak–Ribiére–Polyak (PRP) method [19, 20], Hestenes–Stiefel (HS) method [21], and Liu–Storey (LS) method [22] with

$$\beta_k^{\mathrm{PRP}} = \frac{\omega_k^{\top} y_{k-1}}{\|\omega_{k-1}\|^2},\ \beta_k^{\mathrm{HS}} = \frac{\omega_k^{\top} y_{k-1}}{d_{k-1}^{\top} y_{k-1}},\ \beta_k^{\mathrm{LS}} = -\frac{\omega_k^{\top} y_{k-1}}{d_{k-1}^{\top}\omega_{k-1}}$$

is more complicated to establish, but they usually behave more effectively in computation instead. With desire for new methods with both promising convergence and numerical performance, based on the aforementioned six classical methods, researchers have made substantial efforts and lots of work have been done. Recently, using the spectral technique, Ibrahim et al. [23] proposed two new spectral CG methods based on the PRP and Wei-Yao-Liu (WYL) methods. The efficiency was verified by applying them in real-life problems in regression analysis. Different from the commonly used first-order hybridization technique, Kaelo et al. [24] proposed a new hybrid CG method based on the modified quadratic

hybridization of the PRP and FR methods and demonstrated the competitiveness of the new hybrid method. We refer the readers to [25–35] for more information.

This work initially focuses on the hybridization-based self-adaptive technique. Appealed by the excellent numerical efficiency of the PRP method, Wei et al. [36] modified it and gave a variant method (WYL), where β_k is obtained by

$$\beta_k^{\mathrm{WYL}} = \frac{\|\omega_k\|^2 - \frac{\|\omega_k\|}{\|\omega_{k-1}\|}\omega_k^\top \omega_{k-1}}{\|\omega_{k-1}\|^2}.$$

They established the globally convergent results of the WYL method assuming that the sufficient descent property holds, which is significant in the global convergence of the CG method. Later, Huang et al. [37] proved that under the condition of strong Wolfe line search, the WYL method fulfills the sufficient descent property by restricting the parameter $\sigma < \frac{1}{4}$. Motivated by the WYL method and to take advantage of the modification scheme adopted in it, Yao et al. [38] extended the scheme to the HS method and suggested the YWH method, that is,

$$\beta_k^{\mathrm{YWH}} = \frac{\|\omega_k\|^2 - \frac{\|\omega_k\|}{\|\omega_{k-1}\|}\omega_k^\top \omega_{k-1}}{d_{k-1}^\top y_{k-1}}.$$

Unfortunately, in the YWH method, the sufficient descent property may not hold in general cases. It holds only when the strong Wolfe line search satisfies $\sigma < \frac{1}{3}$, in which case the method converges globally and performs well in numerical experiments. Recently, based on Ref. [39], a new hybrid CG method (Jian-Han-Jiang, JHJ) was suggested by Jian et al. [40], where

$$\beta_k^{\mathrm{JHJ}} = \frac{\|\omega_k\|^2 - \max\left\{\frac{\|\omega_k\|}{\|\omega_{k-1}\|}\omega_k^\top \omega_{k-1}, 0\right\}}{\max\left\{\|\omega_{k-1}\|^2, d_{k-1}^\top y_{k-1}\right\}},$$

and the globally convergent results were proven under standard Wolfe line search.

Next, the Dai-Liao (DL)-style CG method is taken into our consideration. In CG method, the traditional conjugacy condition is $d_k^\top y_{k-1} = 0$. However, not only is the condition hard to fulfill, which requires exact line search, but it can also occasionally bring up some disadvantages in computation [41, 42]. In order to avoid such phenomenon, Dai and Liao [43] explored a weaker conjugacy condition, that is,

$$d_k^\top y_{k-1} = -t\omega_k^\top s_{k-1},\ t \geq 0,$$

where $s_{k-1} = x_k - x_{k-1}$, and it yields a new CG method (DL) with

$$\beta_k^{\mathrm{DL}} = \frac{\omega_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}} - t\frac{\omega_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}}.$$

The DL method guarantees global convergence when the objective functions are uniformly convex. To get global convergence for general functions, they generalized it and gave a variant of β_k^{DL}, that is,

$$\beta_k^{\mathrm{DL+}} = \max\left\{\frac{\omega_k^\top y_{k-1}}{d_{k-1}^\top y_{k-1}}, 0\right\} - t\frac{\omega_k^\top s_{k-1}}{d_{k-1}^\top y_{k-1}}.$$

Numerical experiments demonstrate that the two methods above possess excellent properties. Recently, based on the JHJ and DL methods, Aminifard and Babaie-Kafaki [44] suggested a new CG method with the following β_k^{EJHJ},

$$\beta_k^{\mathrm{EJHJ}} = \frac{\|\omega_k\|^2 - \max\left\{\frac{\|\omega_k\|}{\|\omega_{k-1}\|}\omega_k^\top \omega_{k-1}, 0\right\}}{\max\{\|\omega_{k-1}\|^2, d_{k-1}^\top y_{k-1}\}} - t\frac{\omega_k^\top s_{k-1}}{\max\{\|\omega_{k-1}\|^2, d_{k-1}^\top y_{k-1}\}},$$

and numerical experiments conducted on the EJHJ method verified its promising performance.

In this work, we propose a new conjugate coefficient based on the DL-style CG methods [43, 44] and the hybridization-based self-adaptive technique in [36, 38, 40]. Furthermore, different from the aforementioned methods, we choose to utilize a new type of denominator for promising results. The improved DL-style hybrid CG (iDLHCG) method takes the following β_k^{idlhcg},

$$\beta_k^{\text{idlhcg}} = \frac{\|\omega_k\|^2 - \max\left\{\frac{\|\omega_k\|}{\|d_{k-1}\|}\omega_k^\top d_{k-1}, \frac{\|\omega_k\|}{\|\omega_{k-1}\|}\omega_k^\top \omega_{k-1}, 0\right\}}{d_{k-1}^\top p_{k-1}} - t\frac{\max\{0, \omega_k^\top s_{k-1}\}}{d_{k-1}^\top p_{k-1}}, \tag{4}$$

where

$$p_{k-1} = \omega_k + \lambda_k \frac{\|\omega_k\|}{\|d_{k-1}\|} d_{k-1},\ \lambda_k = \nu + \max\left\{0, -\frac{\omega_k^\top d_{k-1}}{\|\omega_k\|\|d_{k-1}\|}\right\}, \tag{5}$$

$t \geq 0$ and $\nu > 1$. It is worth mentioning that the denominator in β_k^{idlhcg} has a new construction $d_{k-1}^\top p_{k-1}$. This adjustment ensures that all the search directions will possess the sufficient descent and trust region properties, automatically.

Now, we lean our focus on another famous hybridization technique. Dai and Yuan [45] proposed a classical hybrid CG method (HD) with $\beta_k^{\text{HD}} = \max\left\{0, \min\left\{\beta_k^{\text{HS}}, \beta_k^{\text{DY}}\right\}\right\}$, which converges globally and performs effectively in computation. The technique is also adopted by Yang et al. [46]. Extended from the LS and CD methods, the so-called LSCD method was suggested, where $\beta_k^{\text{LSCD}} = \max\left\{0, \min\left\{\beta_k^{\text{LS}}, \beta_k^{\text{CD}}\right\}\right\}$. Effectiveness of this technique has been verified, which inspires us to exploit it in the iDLHCG method, too. Therefore, β_k^{iDLHCG} is calculated by

$$\beta_k^{\text{iDLHCG}} = \max\left\{0, \min\left\{\beta_k^{\text{idlhcg}}, \beta_k\right\}\right\}, \tag{6}$$

where β_k is an arbitrary conjugate coefficient. The search direction of the iDLHCG method is computed by (3)–(6).

1.2 | Conjugate gradient projection method

In this subsection, we make attempts to extend the iDLHCG method for constrained nonlinear monotone equations,

$$Q(x) = 0, x \in C, \tag{7}$$

where C is a nonempty closed convex set. The mapping $Q : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and monotone, where the monotonicity means the following inequality is satisfied,

$$(Q(x) - Q(y))^\top (x - y) \geq 0, \forall x, y \in \mathbb{R}^n. \tag{8}$$

The nonlinear monotone equations system (7) reveals its importance in many modern applications, including chemical equilibrium system problem [47], compressed sensing problem [48–53], and image restoration problem [54–59]. In addition, many constrained optimization problems can be converted into the form (7), and we refer the readers to [60–69] for more information about constrained optimization.

The numerous applications have appealed many researchers to derive new methods to solve (7). Up to now, various methods have been suggested, which can be categorized into two groups depending on whether the derivatives are required. The first group, featured in rapid convergence, includes the Newton method with projection technique [70] and the trust region method [71]. However, in practical computation, it is expensive to apply them in solving (7) in large scale, since in this case these methods require the derivative information, which means large-scale matrices need to be computed and stored. The second group consists of the derivative-free projection (DFP) methods, including the spectral gradient method [72, 73] and the conjugate gradient projection (CGP) method [74–76]. These methods behave better for solving (7) in large scale because they demand minimal storage.

Now, we only concentrate on the CGP method. First, we briefly recall the hyperplane projection technique suggested by Solodov and Svaiter [77]. Let $w_k = x_k + l_k d_k$. By the monotonicity of Q, the hyperplane $H_k = \{x \in \mathbb{R}^n \mid Q(w_k)^\top (x - w_k) = 0\}$ strictly separates the current iteration point x_k from the solutions of problem (7). Therefore, they suggested that the

projection of x_k onto the hyperplane could be the next iteration point; that is,

$$x_{k+1} = x_k - \frac{\mathcal{Q}(w_k)^\top (x_k - w_k)}{\|\mathcal{Q}(w_k)\|^2}\mathcal{Q}(w_k).$$

With the purpose to address convex constraints, we have to introduce an important projection operator P from $\mathbb{R}^n$ onto the feasible region $\mathcal{C}$, which is specified as $P_{\mathcal{C}}[x] = \arg\min\{\|y - x\| \mid y \in \mathcal{C}\}$. In the convergence analysis, the following non-expansive property is of great importance; that is,

$$\|P_{\mathcal{C}}[x] - P_{\mathcal{C}}[y]\| \leq \|x - y\|, \forall x, y \in \mathbb{R}^n. \tag{9}$$

Based on the projection technique, the CG method can be modified and applied in solving (7), and then the CGP method is obtained.

In the past decades, considering that the CGP method possesses many excellent properties, researchers have paid great attention to it and many well-known methods have been proposed so far. For instance, Cheng [78] combined the famous PRP method and the projection strategy, proposing a new PRP-type method. Later, Yan et al. [72] also proposed two new CGP methods based on the well-known HS method. Sun and Liu [75] introduced a self-adaptive hybrid method, extended from the JHJ method, and its numerical performance is promising.

Now, we give the improved DL-style hybrid CGP method extended from the iDLHCG method and denote it as the iDLHCGP method. The conjugate coefficient takes the form

$$\beta_k^{\text{idlhcgp}} = \frac{\|\mathcal{Q}_k\|^2 - \max\left\{0, \frac{\|\mathcal{Q}_k\|}{\|d_{k-1}\|}\mathcal{Q}_k^\top d_{k-1}, \frac{\|\mathcal{Q}_k\|}{\|\mathcal{Q}_{k-1}\|}\mathcal{Q}_k^\top \mathcal{Q}_{k-1}\right\}}{d_{k-1}^\top p_{k-1}} - t\frac{\max\{0, \mathcal{Q}_k^\top s_{k-1}\}}{d_{k-1}^\top p_{k-1}}, \tag{10}$$

where $\mathcal{Q}_k = \mathcal{Q}(x_k)$,

$$p_{k-1} = \mathcal{Q}_k + \lambda_k \frac{\|\mathcal{Q}_k\|}{\|d_{k-1}\|}d_{k-1},\ \lambda_k = \nu + \max\left\{0, -\frac{\mathcal{Q}_k^\top d_{k-1}}{\|\mathcal{Q}_k\|\|d_{k-1}\|}\right\}, \tag{11}$$

$t \geq 0$ and $\nu > 1$. Based on the similar hybridization technique in the iDLHCG method, and to stay consistent with it, the search direction of the iDLHCGP method is defined by

$$d_k = \begin{cases} -\mathcal{Q}_k, & k = 0, \\ -\mathcal{Q}_k + \beta_k^{\text{iDLHCGP}} d_{k-1}, & k \geq 1, \end{cases} \tag{12}$$

$$\beta_k^{\text{iDLHCGP}} = \max\left\{0, \min\left\{\beta_k^{\text{idlhcgp}}, \beta_k\right\}\right\}, \tag{13}$$

where β_k is an arbitrary conjugate coefficient.

1.3 | Motivation and main contributions

To sum up, the limitations of existing works and the motivation for our methods are listed in three aspects. (i) The DL-style CG method possesses many nice properties in solving unconstrained optimization, and how to propose a new DL-style CG method that performs better both in theory and computation? (ii) Based on the new CG method, a natural question is that how to extend it to constrained nonlinear monotone equations and obtain a new CGP method? (iii) Convergence property and convergence rate analysis are important in optimization fields, but besides the global convergence property, can we establish the corresponding convergence rates results for the proposed methods?

With these questions, we introduce the iDLHCG method for unconstrained optimization, whose search direction is defined by (3)–(6). Additionally, taking the hyperplane projection technique into consideration, we extend the iDLHCG method to constrained nonlinear monotone equations and obtain the iDLHCGP method, of which the search direction is calculated by (10)–(13). Moreover, we establish the convergence rates results for the proposed methods. The key contributions of this work have at least three aspects as follows:

- Methodically, we first introduce a new conjugate coefficient based on the DL-style CG methods [43, 44] and the hybridization-based self-adaptive technique. The new conjugate coefficient β_k^{iDLHCG} contains an uncertain parameter,

which allows more flexibility in computation. Then, combining the standard Wolfe line search, we obtain the iDLHCG method for unconstrained optimization. Further, combining the hyperplane projection technique and a new well-defined self-adaptive line search, the iDLHCGP method, an extension of the iDLHCG method, for constrained nonlinear monotone equations is obtained.

- Theoretically, the search direction generated by the iDLHCG method or the iDLHCGP method fulfills the sufficient descent and trust region properties independent of any line search. Then we establish the globally convergent results for the iDLHCG method under standard Wolfe line search. For the iDLHCGP method, under the new self-adaptive line search, the global convergence is obtained without Lipschitz continuity. Moreover, convergence rate analysis is an important aspect in optimization, but there have been relatively few works analyzing the convergence rates of the CG method and the CGP method. Therefore, we complete the convergence rates analysis for the iDLHCG method and the iDLHCGP method.
- Numerically, we conduct different numerical experiments for the iDLHCG method and the iDLHCGP method, respectively. For the iDLHCG method, we apply it in solving unconstrained optimization problems. And the iDLHCGP method is applied in constrained nonlinear monotone equations and signal recovery problem. The numerical results demonstrate the computational effectiveness of our new methods.

The remainder of this work is organized as follows. In Section 2, we introduce the iDLHCG method and develop the corresponding global convergence under usual assumptions and standard Wolfe line search. And its convergence rate is given. In Section 3, we suggest the iDLHCGP method and analyze its global convergence under a new self-adaptive line search. Then the convergence rate results are established. Numerical experiments are conducted in Section 4 to verify their computational effectiveness. Finally, we make the conclusions in Section 5.

2 | A NEW HYBRID CG METHOD

2.1 | Algorithm framework and properties

In this part, the iDLHCG method will be described under standard Wolfe line search. We give the framework of our algorithm as follows:

Algorithm 1 iDLHCG method

Initiate $x_1 \in \mathbb{R}$, $\epsilon > 0$, $0 < \sigma_1 < \sigma_2 < 1$, $\nu > 1$ and the maximum iteration number 'Itermax'. Set $d_1 = -\omega_1$.
for k=1:Itermax **do**
  Determine the step length l_k using the standard Wolfe line search, i.e.,

$$\begin{cases} f(x_k + l_k d_k) \le f(x_k) + \sigma_1 l_k \omega_k^\top d_k, \\ \omega(x_k + l_k d_k)^\top d_k \ge \sigma_2 \omega_k^\top d_k. \end{cases} \tag{14}$$

  Generate the next iteration point by $x_{k+1} = x_k + l_k d_k$.
  Calculate $\omega_{k+1} = \omega(x_{k+1})$.
  if $\|\omega_{k+1}\| \le \epsilon$ **then**
    break
  else
    Calculate $\beta_{k+1}^{\text{idlhcg}}$ by (4)-(5) and an appropriate conjugate coefficient β_{k+1}.
    Calculate $\beta_{k+1}^{\text{iDLHCG}} = \max\left\{0, \min\left\{\beta_{k+1}^{\text{idlhcg}}, \beta_{k+1}\right\}\right\}$.
    Compute the search direction $d_{k+1} = -\omega_{k+1} + \beta_{k+1}^{\text{iDLHCG}} d_k$.
  end if
end for

Next, we will prove the sufficient descent property of the iDLHCG method.

Lemma 1. *The following relation*

$$0 \le \beta_k^{iDLHCG} \le \frac{\|\omega_k\|}{\nu \|d_{k-1}\|}, \forall k \ge 0 \tag{15}$$

always holds.

Proof. From the definition of β_k^{iDLHCG}, we know that it is positive. Before proving the right-hand side of (15), we first give a lower bound for the denominator $d_{k-1}^\top p_{k-1}$, that is,

$$\begin{aligned} d_{k-1}^\top p_{k-1} &= d_{k-1}^\top \left(\omega_k + \lambda_k \frac{\|\omega_k\|}{\|d_{k-1}\|} d_{k-1}\right) \\ &= d_{k-1}^\top \omega_k + \lambda_k \|\omega_k\| \|d_{k-1}\| \\ &\geq d_{k-1}^\top \omega_k + \left(\nu - \frac{d_{k-1}^\top \omega_k}{\|\omega_k\| \|d_{k-1}\|}\right) \|\omega_k\| \|d_{k-1}\| \\ &= \nu \|\omega_k\| \|d_{k-1}\|. \end{aligned}$$

To obtain the upper bound for β_k^{iDLHCG}, noticing that $\beta_k^{\text{iDLHCG}} = \max\left\{0, \min\left\{\beta_k^{\text{idlhcg}}, \beta_k\right\}\right\}$, we examine the following three scenarios:

(i) $\beta_k^{\text{iDLHCG}} = 0$. Clearly, the relation holds.

(ii) $\beta_k^{\text{iDLHCG}} = \beta_k^{\text{idlhcg}}$. From (4) and the lower bound of $d_{k-1}^\top p_{k-1}$, we have

$$\begin{aligned} 0 \leq \beta_k^{\text{iDLHCG}} = \beta_k^{\text{idlhcg}} &= \frac{\|\omega_k\|^2 - \max\left\{0, \frac{\|\omega_k\|}{\|d_{k-1}\|}\omega_k^\top d_{k-1}, \frac{\|\omega_k\|}{\|\omega_{k-1}\|}\omega_k^\top \omega_{k-1}\right\}}{d_{k-1}^\top p_{k-1}} - t\frac{\max\{0, \omega_k^\top s_{k-1}\}}{d_{k-1}^\top p_{k-1}} \\ &\leq \frac{\|\omega_k\|^2 - \max\left\{0, \frac{\|\omega_k\|}{\|d_{k-1}\|}\omega_k^\top d_{k-1}, \frac{\|\omega_k\|}{\|\omega_{k-1}\|}\omega_k^\top \omega_{k-1}\right\}}{d_{k-1}^\top p_{k-1}} \\ &\leq \frac{\|\omega_k\|^2}{d_{k-1}^\top p_{k-1}} \leq \frac{\|\omega_k\|}{\nu \|d_{k-1}\|}. \end{aligned}$$

(iii) $\beta_k^{\text{iDLHCG}} = \beta_k$. It means that β_k should satisfy $\beta_k \leq \beta_k^{\text{idlhcg}}$, and then, together with the relation in case (ii), it follows that

$$0 \leq \beta_k^{\text{iDLHCG}} = \beta_k \leq \beta_k^{\text{idlhcg}} \leq \frac{\|\omega_k\|}{\nu \|d_{k-1}\|}.$$

The whole proof is completed. □

Lemma 2. *Suppose that the sequence $\{d_k\}$ is generated by the iDLHCG method, then the sufficient descent property holds; that is,*

$$\omega_k^\top d_k \leq -\left(1 - \frac{1}{\nu}\right) \|\omega_k\|^2, \forall k \geq 0. \tag{16}$$

Proof. For $k = 0$, $\omega_0^\top d_0 = -\|\omega_0\|^2$. And for all $k \geq 1$, from (3) and (15), we have

$$\begin{aligned} \omega_k^\top d_k &\leq -\|\omega_k\|^2 + \beta_k^{\text{iDLHCG}} \|\omega_k\| \|d_{k-1}\| \\ &\leq -\|\omega_k\|^2 + \frac{\|\omega_k\|}{\nu \|d_{k-1}\|} \|\omega_k\| \|d_{k-1}\| \\ &= -\left(1 - \frac{1}{\nu}\right) \|\omega_k\|^2. \end{aligned}$$

The proof is complete. □

2.2 | Convergence analysis

In this subsection, we concentrate on the global convergence of the iDLHCG method. To this end, we assume the following.

Assumption 1. The level set $\mathcal{L} = \{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\}$ is bounded.

Assumption 2. Suppose $\mathcal{N}$ is some neighborhood of $\mathcal{L}$, f is continuously differentiable, and its gradient is Lipschitz continuous; that is, a constant $L > 0$ can be found such that

$$\|\omega(x) - \omega(y)\| \leq L\|x - y\|, \forall x, y \in \mathcal{N}.$$

From Assumptions 1 and 2, we conclude that a positive constant Γ can be found such that $\|\omega(x)\| \leq \Gamma$ for all $x \in \mathcal{L}$.

Lemma 3 ([79]). *Assume that Assumptions 1 and 2 hold. For any iterative method taking the form (2), where d_k satisfies the descent condition and the step length l_k satisfies (14), we have*

$$\sum_{k=0}^{\infty} \frac{(\omega_k^{\top} d_k)^2}{\|d_k\|^2} < \infty.$$

Additionally, if the sufficient descent property holds, it yields that

$$\sum_{k=0}^{\infty} \frac{\|\omega_k\|^4}{\|d_k\|^2} < \infty. \tag{17}$$

Now, by Lemmas 1–3, we will give the globally convergent results completely.

Theorem 1. *Suppose that Assumptions 1 and 2 hold, the iDLHCG method generating the sequence $\{x_k\}$, where the step length l_k is calculated by standard Wolfe line search, we have*

$$\liminf_{k\to\infty} \|\omega_k\| = 0.$$

Proof. First, we prove that there exists a positive constant C, such that

$$\|d_k\| \leq C\|\omega_k\|. \tag{18}$$

By Lemma 1, we have

$$\begin{aligned} \|d_k\| &\leq \|\omega_k\| + |\beta_k^{\mathrm{iDLHCG}}|\|d_{k-1}\| \\ &\leq \|\omega_k\| + \frac{\|\omega_k\|}{\nu\|d_{k-1}\|}\|d_{k-1}\| \\ &\leq \left(1 + \frac{1}{\nu}\right)\|\omega_k\|. \end{aligned}$$

Letting $C = 1 + \frac{1}{\nu}$, then the relation (18) is proved. Next, combining it with (17), we have

$$\sum_{k=0}^{\infty} \|\omega_k\|^2 \leq \sum_{k=0}^{\infty} \frac{C^2\|\omega_k\|^4}{\|d_k\|^2} < \infty,$$

which implies that $\liminf_{k\to\infty}\|\omega_k\| = 0$ together with Lemma 3. Hence, the proof is completed. □

2.3 | Convergence rate analysis

In this subsection, we lean our focus on the r_1-linear convergence rate of the iDLHCG method when $f(x)$ is twice continuously differentiable and uniformly convex, that is,

$$N_1\|y\|^2 \leq y^{\top}\nabla^2 f(x)y \leq N_2\|y\|^2, \forall x, y \in \mathbb{R}^n,$$

where N_1 and N_2 ($\geq N_1$) are two positive constants and $\nabla^2 f(x)$ is the Hessian of f at x. Before the complete analysis, we provide two preparatory lemmas.

Lemma 4 ([80–82]). *Suppose that Assumptions 1 and 2 hold, and $f(x)$ is twice continuously differentiable and uniformly convex. Then there exist two positive constants m and $M(\geq m)$ such that*

$$\frac{1}{2}m\|x-x^*\|^2 \leq f(x)-f(x^*) \leq \frac{1}{2}M\|x-x^*\|^2, \forall x \in \mathbb{R}^n, \tag{19}$$

$$m\|x-x^*\| \leq \|\omega(x)\| \leq M\|x-x^*\|, \forall x \in \mathbb{R}^n, \tag{20}$$

and

$$\|\omega(x)-\omega(y)\| \leq M\|x-y\|, \forall x, y \in \mathbb{R}^n, \tag{21}$$

where x^* is the unique minimal point.

Then, Lemma 5 provides a lower bound for the step length l_k.

Lemma 5. *Suppose that Assumptions 1 and 2 hold, $f(x)$ is twice continuously differentiable and uniformly convex, and the sequence $\{x_k\}$ is generated by the iDLHCG method under the line search (14). Then for all $k > 0$, we can find a constant $\varphi > 0$ such that*

$$l_k \geq \varphi. \tag{22}$$

Proof. From inequality (18) in the proof of Theorem 1, we have

$$\frac{\|\omega_k\|^2}{\|d_k\|^2} \geq \frac{1}{C^2}. \tag{23}$$

Let $\tilde{\nu} = \left(1-\frac{1}{\nu}\right)$. From (14), (16), and (21), it yields that

$$(1-\sigma_2)\tilde{\nu}\|\omega_k\|^2 \leq (\sigma_2-1)\omega_k^\top d_k \leq (\omega_{k+1}-\omega_k)^\top d_k \leq M l_k \|d_k\|^2.$$

Together with (23), we have

$$l_k \geq \frac{(1-\sigma_2)\tilde{\nu}\|\omega_k\|^2}{M\|d_k\|^2} \geq \frac{(1-\sigma_2)\tilde{\nu}}{MC^2}.$$

Set $\varphi = \frac{(1-\sigma_2)\tilde{\nu}}{MC^2}$, which completes the proof. □

Theorem 2. *Suppose that Assumptions 1 and 2 hold, $f(x)$ is twice continuously differentiable and uniformly convex, and the sequence $\{x_k\}$ is generated by the iDLHCG method under the line search (14). Then we can find two constants $a > 0$ and $r_1 \in (0,1)$ such that*

$$\|x_k - x^*\| \leq a r_1^k.$$

Proof. From the line search (14), we have

$$\begin{aligned} f(x_{k+1})-f(x^*) &\leq f(x_k)-f(x^*)+\sigma_1 l_k \omega_k^\top d_k \\ &\leq f(x_k)-f(x^*)-\sigma_1 l_k \tilde{\nu}\|\omega_k\|^2 \\ &\leq f(x_k)-f(x^*)-\sigma_1 \varphi \tilde{\nu} m^2 \|x_k-x^*\|^2 \\ &\leq f(x_k)-f(x^*)-\frac{2\sigma_1\varphi\tilde{\nu}m^2}{M}[f(x_k)-f(x^*)] \\ &= \left[1-\frac{2\sigma_1\varphi\tilde{\nu}m^2}{M}\right][f(x_k)-f(x^*)], \end{aligned}$$

where the third inequality follows from (20) and (22) and the fourth one utilizes (19). Letting $r_1 = \left[1-\frac{2\sigma_1\varphi\tilde{\nu}m^2}{M}\right]^{\frac{1}{2}}$, it yields that

$$f(x_k)-f(x^*) \leq r_1^2[f(x_{k-1})-f(x^*)] \leq \cdots \leq r_1^{2k}[f(x_0)-f(x^*)],$$

Now we show that r_1 is bounded in $(0, 1)$. At first, (16) and (18) yield that

$$\|\omega_k\|\|d_k\| \leq C\|\omega_k\|^2 \leq -\frac{C}{\tilde{\nu}}\omega_k^\top d_k \leq \frac{C}{\tilde{\nu}}\|\omega_k\|\|d_k\|,$$

which implies that $\frac{C}{\tilde{\nu}} \geq 1$. Then we have

$$\frac{2\sigma_1\varphi\tilde{\nu}m^2}{M} \leq 2\sigma_1\varphi\tilde{\nu}m = \frac{2\sigma_1(1-\sigma_2)m\tilde{\nu}^2}{MC^2} \leq 2\sigma_1(1-\sigma_2) \leq 2\sigma_2(1-\sigma_2) < 1.$$

Hence, r_1 is bounded in $(0, 1)$. Combined with (19), it shows that

$$\|x_k - x^*\|^2 \leq \frac{2}{m}[f(x_k) - f(x^*)] \leq \frac{2}{m}r_1^{2k}[f(x_0) - f(x^*)].$$

Let $a = \sqrt{2[f(x_0) - f(x^*)]/m}$, which completes the proof. □

3 | A NEW HYBRID CGP METHOD

3.1 | Algorithm framework and properties

In this subsection, we describe the framework of the iDLHCGP method as follows.

Algorithm 2 iDLHCGP method

1: **Initiate** $x_1 \in C$, the constants $\varepsilon > 0,\ 0 < \tau < 2,\ 0 < \rho < 1,\ 0 < \zeta \leq 1,\ \delta > 0,\ c \geq 1,\ \eta \geq 0,\ \nu > 1$ and the maximum iteration number 'Itermax'. Set $d_1 = -Q_1$.
2: **for** k=1:Itermax **do**
3: Calculate $w_k = x_k + l_k d_k$ where the step length $l_k = \zeta\rho^{i_k}$ with i_k being the smallest nonnegative integer i such that

$$-Q(w_k)^\top d_k \geq \delta l_k \psi_k \|d_k\|^2, \tag{24}$$

4: where $\psi_k = \frac{\|Q(w_k)\|^{1/c}}{\eta + \|Q(w_k)\|^{1/c}}$ and ζ is the initial guess of the step length.
5: Compute

$$x_{k+1} = P_C[x_k - \tau\xi_k Q(w_k)], \tag{25}$$

6: where

$$\xi_k = \frac{Q(w_k)^\top(x_k - w_k)}{\|Q(w_k)\|^2}. \tag{26}$$

7: Calculate $Q_{k+1} = Q(x_{k+1})$.
8: **if** $\|Q_{k+1}\| \leq \varepsilon$ **then**
9: break
10: **else**
11: Calculate $\beta_{k+1}^{\text{idlhcgp}}$ by (10)-(11) and an appropriate conjugate coefficient β_{k+1}.
12: Calculate $\beta_{k+1}^{\text{iDLHCGP}} = \max\left\{0, \min\left\{\beta_{k+1}^{\text{idlhcgp}}, \beta_{k+1}\right\}\right\}$.
13: Calculate the search direction $d_{k+1} = -Q_{k+1} + \beta_{k+1}^{\text{iDLHCGP}} d_k$.
14: **end if**
15: **end for**

Remark 1. In CGP method, two famous line searches are often adopted. The first one is proposed by Zhang and Zhou [83], that is,

$$-Q(w_k)^\top d_k \geq \delta l_k \|d_k\|^2. \tag{27}$$

The second one, proposed by Cheng [78], takes the following form:

$$-Q(w_k)^\top d_k \geq \delta l_k \|Q(w_k)\|\|d_k\|^2. \tag{28}$$

It is worth noticing that when x_k is distant from the solution of (7), $\|\mathcal{Q}(w_k)\|$ will be very big, which can result in the right-hand side of (28) becoming large, therefore increasing the computational cost. To avoid such situation and reduce the cost, Muhammad et al. [84] proposed a new type of line search:

$$-\mathcal{Q}(w_k)^\top d_k \geq \delta l_k \|\mathcal{Q}(w_k)\|^{1/c} \|d_k\|^2,$$

where $c \geq 1$ and the power $\frac{1}{c}$ is introduced to reduce $\|\mathcal{Q}(w_k)\|$. However, in the case when $\|\mathcal{Q}(w_k)\|$ is exceptionally large, the reduction may not be as substantial as desired. To this end, based on (27) and (28), Amini and Kamandi [85] gave a new self-adaptive line search, that is,

$$-\mathcal{Q}(w_k)^\top d_k \geq \delta l_k \hat{\psi}_k \|d_k\|^2, \tag{29}$$

where $\hat{\psi}_k = \frac{\|\mathcal{Q}(w_k)\|}{1+\|\mathcal{Q}(w_k)\|}$. Inspired by (29), we choose to compute l_k by the self-adaptive line search strategy (24), where $\eta \geq 0$. The introduction of the parameters c and η allows us more flexibility than (29) in computation. Considering that $\psi_k \in (0, 1]$, this operation keeps the influence of $\|\mathcal{Q}(w_k)\|$ in a manageable range when it becomes excessively large, thereby empirically reducing the computational cost of the line search process. In fact, the line search (24) has the following two characteristics: (i) The line search (24) will be equivalent to (29) when $\eta = 1$ and $c = 1$. (ii) The line search (24) becomes (27) when $\eta = 0$.

Next, we give Lemma 6 stating that all the search directions generated by the iDLHCGP method, regardless of any line search, possess the sufficient descent and trust region properties.

Lemma 6. *The sequence $\{d_k\}$ generated by the iDLHCGP method fulfills the sufficient descent property*

$$\mathcal{Q}_k^\top d_k \leq -\left(1 - \frac{1}{\nu}\right) \|\mathcal{Q}_k\|^2, \forall k \geq 0, \tag{30}$$

and the trust region property

$$\left(1 - \frac{1}{\nu}\right) \|\mathcal{Q}_k\| \leq \|d_k\| \leq \left(1 + \frac{1}{\nu}\right) \|\mathcal{Q}_k\|. \tag{31}$$

Proof. Utilizing the same manipulation in Lemma 1, we obtain the following relation

$$0 \leq \beta_k^{\text{iDLHCGP}} \leq \frac{\|\mathcal{Q}_k\|}{\nu \|d_{k-1}\|}. \tag{32}$$

Therefore, after the same manipulation, (30), is at hand. Next, we will prove (31). From $|\mathcal{Q}_k^\top d_k| \leq \|\mathcal{Q}_k\| \|d_k\|$ and (30), we have

$$-\|\mathcal{Q}_k\| \|d_k\| \leq -\left(1 - \frac{1}{\nu}\right) \|\mathcal{Q}_k\|^2,$$

which implies that the left-hand side of (31) holds. Then, from (12) and (32), one has

$$\|d_k\| \leq \|\mathcal{Q}_k\| + |\beta_k^{\text{iDLHCGP}}| \|d_{k-1}\| \leq \|\mathcal{Q}_k\| + \frac{\|\mathcal{Q}_k\|}{\nu \|d_{k-1}\|} \|d_{k-1}\| = \left(1 + \frac{1}{\nu}\right) \|\mathcal{Q}_k\|.$$

Therefore, the proof is complete. □

3.2 | Convergence analysis

In this subsection, we will establish the globally convergent results of the iDLHCGP method, and the following common assumptions are necessary.

Assumption 3. The solution set of problem (7) is nonempty, denoting it as $\mathcal{S}$ and $\mathcal{S} \subseteq C$.

Assumption 4. The mapping $\mathcal{Q}$ is monotone and continuous on $\mathbb{R}^n$.

Notice that in Assumption 4, the condition is milder since the Lipschitz continuity is not required.

Lemma 7. *Suppose that Assumptions 3 and 4 hold and $\{x_k\}$ and $\{w_k\}$ being sequences generated by the iDLHCGP method. Then we have that the sequence $\{\|x_k - x^*\|\}$ is convergent for any $x^* \in S$. It concludes that $\{x_k\}$ is bounded, and in addition, we have*

$$\lim_{k\to\infty} l_k\|d_k\| = 0. \tag{33}$$

Proof. We can see from the definition of w_k that

$$Q(w_k)^\top(x_k - w_k) \geq \delta\psi_k l_k^2\|d_k\|^2 = \delta\psi_k\|x_k - w_k\|^2 > 0.$$

For any $x^* \in S$, from (8), we have

$$\begin{aligned} Q(w_k)^\top(x_k - x^*) &= Q(w_k)^\top(x_k - w_k) + Q(w_k)^\top(w_k - x^*) \\ &\geq Q(w_k)^\top(x_k - w_k) + Q(x^*)^\top(w_k - x^*) \\ &= Q(w_k)^\top(x_k - w_k) \\ &\geq \delta\psi_k\|x_k - w_k\|^2 > 0. \end{aligned}$$

The above relations, together with Assumption 3, (9), (25), and (26), imply that

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &= \|P_C[x_k - \tau\xi_k Q(w_k)] - P_C[x^*]\|^2 \\ &\leq \|x_k - \tau\xi_k Q(w_k) - x^*\|^2 \\ &= \|x_k - x^*\|^2 - 2\tau\xi_k Q(w_k)^\top(x_k - x^*) + \tau^2\xi_k^2\|Q(w_k)\|^2 \\ &\leq \|x_k - x^*\|^2 - 2\tau\xi_k Q(w_k)^\top(x_k - w_k) + \tau^2\xi_k^2\|Q(w_k)\|^2 \\ &= \|x_k - x^*\|^2 - \tau(2-\tau)\frac{\left[Q(w_k)^\top(x_k - w_k)\right]^2}{\|Q(w_k)\|^2} \\ &\leq \|x_k - x^*\|^2 - \tau(2-\tau)\frac{\delta^2\psi_k^2\|x_k - w_k\|^4}{\|Q(w_k)\|^2}. \end{aligned}$$

From the inequality, we know that $0 \leq \|x_{k+1} - x^*\| \leq \|x_k - x^*\|$ since $\tau \in (0, 2)$, which means the sequence $\{\|x_k - x^*\|\}$ is nonincreasing and bounded below. Therefore, the sequence $\{\|x_k - x^*\|\}$ is convergent, and then we conclude that $\{x_k\}$ is bounded. Deducing from the boundedness of $\{x_k\}$ and the continuity of Q, we can find a positive constant c_1 such that $\|Q_k\| \leq c_1$. Moreover, noticing the boundedness of $\{Q_k\}$ and the right-hand side of (31), we have that $\{d_k\}$ is bounded, which further implies that $\{w_k\}$ is bounded too. Then again from the continuity of Q, we obtain that there exists a positive constant c_2 such that $\|Q(w_k)\| \leq c_2$. As a result, the inequality above demonstrates that

$$\tau(2-\tau)\frac{\delta^2\psi_k^2}{c_2}\sum_{k=0}^{\infty}\|x_k - w_k\|^4 \leq \sum_{k=0}^{\infty}(\|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2) = \|x_0 - x^*\|^2 - \lim_{k\to\infty}\|x_{k+1} - x^*\|^2 < \infty,$$

which implies $\lim_{k\to\infty} l_k\|d_k\| = \lim_{k\to\infty}\|x_k - w_k\| = 0$, and the proof is complete. □

Next, we establish the globally convergent results of the iDLHCGP method.

Theorem 3. *Suppose that Assumptions 3 and 4 hold, $\{x_k\}$ being generated by the iDLHCGP method, and then we have*

$$\liminf_{k\to\infty}\|Q_k\| = 0.$$

Additionally, $\{x_k\}$ converges to a solution of (7) globally.

Proof. We proceed by contradiction. Assume that there exists a positive constant c_3 such that

$$\|Q_k\| \geq c_3, \forall k \geq 0. \tag{34}$$

Combing (31) and (34), we obtain for all $k \geq 0$

$$\|d_k\| \geq \left(1 - \frac{1}{\nu}\right) \|Q_k\| \geq \left(1 - \frac{1}{\nu}\right) c_3,$$

which together with (33) implies that

$$\lim_{k\to\infty} l_k = 0. \tag{35}$$

By the boundedness of sequences $\{x_k\}$ and $\{d_k\}$, we assume, without loss of generality, that there exist two convergent subsequences $\{x_{k_j}\}$ and $\{d_{k_j}\}$ such that

$$\lim_{j\to\infty} x_{k_j} = \tilde{x}, \lim_{j\to\infty} d_{k_j} = \tilde{d}. \tag{36}$$

Using (30), (34), (36), and the continuity of Q, we get

$$-Q(\tilde{x})^\top \tilde{d} \geq \left(1 - \frac{1}{\nu}\right) \|Q(\tilde{x})\|^2 \geq \left(1 - \frac{1}{\nu}\right) c_3^2 > 0. \tag{37}$$

Additionally, (24) implies that

$$-Q(x_{k_j} + \rho^{-1} l_{k_j} d_{k_j})^\top d_{k_j} < \delta \rho^{-1} l_{k_j} \psi_k \|d_{k_j}\|^2 \leq \delta \rho^{-1} l_{k_j} \|d_{k_j}\|^2.$$

Letting $j \to \infty$ in the inequality above, from the continuity of Q and (35), we have $-Q(\tilde{x})^\top \tilde{d} \leq 0$, which contradicts with (37) and we obtain $\liminf_{k\to\infty} \|Q_k\| = 0$. Since Q is continuous, there exists an accumulation point $\hat{x}$ of $\{x_k\}$ such that $Q(\hat{x}) = 0$. Combining the convergence of $\{\|x_k - x^*\|\}$, we conclude that $\{x_k\}$ converges to a solution of (7) globally.

3.3 | Convergence rate analysis

In this subsection, we are going to study the convergence rate of the iDLHCGP method. To this aim, we give the three following assumptions.

Assumption 5. The mapping Q is Lipschitz continuous on $\mathbb{R}^n$; that is, we find a positive constant $\hat{L}$ such that

$$\|Q(x) - Q(y)\| \leq \hat{L}\|x - y\|, \forall x, y \in \mathbb{R}^n. \tag{38}$$

Assumption 6. There exist two positive constants c_4 and δ^* such that for any $x^* \in \mathcal{S}$,

$$\|Q(x)\| \geq c_4 dist(x, \mathcal{S}), \forall x \in \mathcal{N}(x^*, \delta^*) \cap \mathcal{C}. \tag{39}$$

Assumption 7. The mapping Q is strongly monotone with modulus μ; that is,

$$(Q(x) - Q(y))^\top (x - y) \geq \mu \|x - y\|^2, \forall x, y \in \mathcal{C}.$$

Remark 2. It is noticeable that Assumption 5 is stronger than Assumption 4; that is, we require the Lipschitz continuity of Q, which is significant in the convergence rate analysis. Assumptions 6 and 7 are common assumptions in analyzing the convergence rate of the CGP method.

Now, we give three lemmas as follows to delve deeper into the convergence rate of the iDLHCGP method.

Lemma 8 ([86]). *Let $u_k > 0$ and*

$$u_{k+1} \leq u_k - a_k u_k^{1+p}, a_k \geq 0, p > 0.$$

Then we have

$$u_k \leq u_0 \left(1 + p u_0^p \sum_{i=0}^{k-1} a_i\right)^{-\frac{1}{p}},$$

and especially, if $a_k \equiv a, p = 1$,

$$u_k \leq \frac{u_0}{1 + au_0 k}. \tag{40}$$

Lemma 9. *Suppose that Assumptions 3–5 hold,* $x^* \in \mathcal{S}$ *and* $\{x_k\}$ *being generated by the iDLHCGP method. Then we find a positive constant* c_5 *such that*

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - c_5\|x_k - w_k\|^4. \tag{41}$$

Proof. From the proof of Lemma 7 and $\|\mathcal{Q}(w_k)\| \leq c_2$, we have

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\leq \|x_k - x^*\|^2 - \tau(2-\tau)\frac{\delta^2\psi_k^2\|x_k - w_k\|^4}{\|\mathcal{Q}(w_k)\|^2}\\ &= \|x_k - x^*\|^2 - \tau(2-\tau)\frac{\delta^2\|\mathcal{Q}(w_k)\|^{2/c-2}}{(\eta + \|\mathcal{Q}(w_k)\|^{1/c})^2}\|x_k - w_k\|^4\\ &\leq \|x_k - x^*\|^2 - \tau(2-\tau)\frac{\delta^2 c_2^{2/c-2}}{(\eta + c_2^{1/c})^2}\|x_k - w_k\|^4.\end{aligned}$$

By letting $c_5 = \frac{\tau(2-\tau)\delta^2 c_2^{2/c-2}}{(\eta+c_2^{1/c})^2}$, the proof is completed. □

Lemma 10. *Suppose that Assumptions 3–5 hold. Then there exists a positive constant* c_6 *such that*

$$l_k \geq \min\left\{\zeta, c_6\frac{\|\mathcal{Q}_k\|^2}{\|d_k\|^2}\right\}. \tag{42}$$

Proof. If $l_k = \zeta$, then (42) holds. Otherwise, the line search (24) with $\tilde{l}_k = \rho^{-1}l_k$ does not hold; that is,

$$-\mathcal{Q}(\tilde{w}_k)^\top d_k < \delta\tilde{l}_k\psi_k\|d_k\|^2 \leq \delta\tilde{l}_k\|d_k\|^2,$$

where $\tilde{w}_k = x_k + \tilde{l}_k d_k$. Combining (30) and (38), we obtain

$$\begin{aligned}\left(1 - \frac{1}{\nu}\right)\|\mathcal{Q}_k\|^2 &\leq (\mathcal{Q}(\tilde{w}_k) - \mathcal{Q}_k)^\top d_k - \mathcal{Q}(\tilde{w}_k)^\top d_k\\ &< \tilde{l}_k(\hat{L} + \delta)\|d_k\|^2\\ &= \rho^{-1}l_k(\hat{L} + \delta)\|d_k\|^2.\end{aligned}$$

Letting $\tilde{\nu} = \left(1 - \frac{1}{\nu}\right)$, it yields that

$$l_k \geq \min\left\{\zeta, \frac{\tilde{\nu}\rho\|\mathcal{Q}_k\|^2}{(\hat{L} + \delta)\|d_k\|^2}\right\}.$$

Letting $c_6 = \frac{\tilde{\nu}\rho}{(\hat{L}+\delta)}$, the whole proof is completed. □

Theorem 4. *Let* $x^* \in \mathcal{S}$ *and* $\{x_k\}$ *be generated by the iDLHCGP method. Then the following two conclusions are obtained as follows:*

(i) With Assumptions 3–6, the sequence $dist(x_k, \mathcal{S})$ *converges to 0 at a sublinear rate; that is, a positive constant* r_2 *can be found such that for sufficiently large k,*

$$dist(x_k, \mathcal{S}) \leq \frac{r_2}{\sqrt{k}}.$$

(ii) With Assumptions 3–5 and 7, the sequence $\{x_k\}$ *converges to* $x^* \in \mathcal{S}$ *at a sublinear rate.*

Proof. (i) Let $\tilde{x} \in \mathcal{S}$ and satisfy

$$\|x_k - \tilde{x}\| = dist(x_k, \mathcal{S}).$$

Combined with (41), it yields that

$$dist^2(x_{k+1}, \mathcal{S}) \leq \|x_{k+1} - \tilde{x}\|^2 \leq dist^2(x_k, \mathcal{Q}) - c_5\|x_k - w_k\|^4. \tag{43}$$

Letting $C = \left(1 + \frac{1}{\nu}\right)$, from (31), we obtain that $\|d_k\| \leq C\|\mathcal{Q}_k\|$. Together with (42), it yields that there exists a positive constant c_7 such that

$$l_k \geq \min\left\{\zeta, c_6\frac{\|\mathcal{Q}_k\|^2}{\|d_k\|^2}\right\} \geq c_7, \tag{44}$$

where $c_7 = \min\left\{\zeta, \frac{c_6}{C^2}\right\}$. Combining the left-hand side of (31), (39), and (44), it follows that for sufficiently large k,

$$\|x_k - w_k\|^2 = l_k^2\|d_k\|^2 \geq c_7^2\tilde{\nu}^2\|\mathcal{Q}_k\|^2 \geq (c_7\tilde{\nu}c_4)^2 dist^2(x_k, \mathcal{S}).$$

Together with (43), it further implies that

$$dist^2(x_{k+1}, \mathcal{S}) \leq dist^2(x_k, \mathcal{S}) - c_5(c_7\tilde{\nu}c_4)^4 dist^4(x_k, \mathcal{S}).$$

Denoting $c_8 = c_5(c_7\tilde{\nu}c_4)^4$, we have from Lemma 8 with $u_k = dist^2(x_k, \mathcal{S})$ and $a_k \equiv c_8$ that

$$dist^2(x_k, \mathcal{S}) \leq \frac{dist^2(x_0, \mathcal{S})}{1 + c_8 dist^2(x_0, \mathcal{S})k},$$

which further implies that for sufficiently large k,

$$dist(x_k, \mathcal{Q}) \leq \frac{1}{\sqrt{dist^{(-2)}(x_0, \mathcal{Q}) + c_8 k}} \leq \frac{1}{\sqrt{c_8 k}}.$$

It demonstrates that the sequence $dist(x_k, \mathcal{S})$ converges to 0 at a sublinear rate.

(ii) From the strong monotonicity of $\mathcal{Q}$ and Cauchy–Schwarz inequality, we have

$$\|\mathcal{Q}_k\| = \|\mathcal{Q}_k - \mathcal{Q}(x^*)\| \geq \mu\|x_k - x^*\|.$$

Together with the left-hand side of (31) and (44), we obtain

$$\|x_k - w_k\| = l_k\|d_k\| \geq c_7\tilde{\nu}\|\mathcal{Q}_k\| \geq c_7\tilde{\nu}\mu\|x_k - x^*\|,$$

which yields with (41) that

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - c_5\|x_k - w_k\|^4 \leq \|x_k - x^*\|^2 - c_5(c_7\tilde{\nu}\mu)^4\|x_k - x^*\|^4.$$

Denoting $c_9 = c_5(c_7\tilde{\nu}\mu)^4$, we have from Lemma 8 with $u_k = \|x_k - x^*\|^2$ and $a_k \equiv c_9$ that

$$\|x_k - x^*\|^2 \leq \frac{\|x_0 - x^*\|^2}{1 + c_9\|x_0 - x^*\|^2 k},$$

further implying that

$$\|x_k - x^*\| \leq \frac{1}{\sqrt{\|x_0 - x^*\|^{-2} + c_9 k}} \leq \frac{1}{\sqrt{c_9 k}},$$

which demonstrates that the sequence $\{x_k\}$ converges to a solution $x^* \in \mathcal{S}$ at a sublinear rate. Hence, the proof is completed.

4 | NUMERICAL EXPERIMENTS

In this section, we conduct two groups of numerical experiments to verify the effectiveness of our proposed methods. The iDLHCG method is applied for solving unconstrained optimization problems in Section 4.1. In Section 4.2, we apply the iDLHCGP method to constrained nonlinear monotone equations. Moreover, the application of the iDLHCGP method to signal recovery problem is also explored. All the methods were written in MATLAB R2021b and ran on a 64-bit Lenovo laptop (CPU 3.80 GHz, RAM 32.0 GB) with the Windows 11 operating system.

4.1 | Numerical experiments on the iDLHCG method

In this subsection, we are going to testify the efficiency of the iDLHCG method in comparison with the other well-known CG methods under standard Wolfe line search with the parameters set as $\sigma_1 = 0.1$ and $\sigma_2 = 0.01$. The process of iteration will be terminated when $\|\omega_k\| \leq 10^{-6}$ or the number of iterations exceeds 2000. A total of 111 problems from the CUTE library [87] and other commonly used collections [88, 89], whose dimensions range from 2 to 800,000, are selected out as tested examples. In the iDLHCG method, we set $\nu = 1.2$, $t = \frac{1}{l_k}$, and choose $\beta_k = \beta_k^{\mathrm{LS}}$ in (6). The other methods we choose to compare are as follows:

- The DPRP method proposed by Dai and Wen [90] with $\mu = 2$.
- The MMDL method proposed by Stanimirović et al. [91] with $\mu = 2$ and $\bar{t} = l_k$.
- The EJHJ method proposed by Aminifard and Babaie-Kafaki [44] with $t = \frac{\|y_k\|^2}{s_k^\top y_k}$.
- The KD method proposed by Kou and Dai [92] with $\xi = 0.5$ and $\zeta = 0.1$.
- The HYB method proposed by Ibrahim et al. [93] with $\tau = 0.001$.

Among these compared methods, the DPRP, MMDL, and EJHJ methods also utilize the hybridization-based self-adaptive technique, the KD method is a well-known and effective CG method, and the HYB method is a hybrid CG method using the convex combination technique. This selection is helpful for us to verify the effectiveness of our proposed iDLHCG method. Details of the numerical experiments are shown in https://www.cnblogs.com/poin07/p/18272306. "Name," "n," "Itr," "Time," and "$\|\omega_*\|$" represent the name of the tested problem, the corresponding dimension, the number of iterations, the running time, and the value of final gradient, respectively. "NaN" denotes the cases where the iteration cannot be terminated within 2000 times. Throughout the results, we conclude that the iDLHCG method, DPRP method, MMDL method, EJHJ method, KD method, and HYB method solve about 99.10% (110 out of 111), 93.69% (104 out of 111), 94.59% (105 out of 111), 92.79% (103 out of 111), 99.10% (110 out of 111), and 93.69% (104 out of 111) of the problems, respectively. Moreover, to help us visually compare the performances of different methods, the performance profiles proposed by Dolan and Moré [94] are introduced as the way to present the results. As interpreted in [94], the higher a curve is located, the more efficient the corresponding method is; see Figures 1 and 2. It can be seen from the performance profiles that the iDLHCG method behaves better both in Time and Itr. Hence, the efficiency of the iDLHCG method is verified.

4.2 | Numerical experiments on the iDLHCGP method

In this subsection, some other CGP methods are chosen for comparison to solve constrained nonlinear monotone equations (7). The selected methods are the WYL method [95], EPM method [96], PDY method [97], and IHCGP method [98]. As to the parameters and line searches in these methods, they are set in consistency with the original literature and so are the notations. And they are listed below as follows:

- The iDLHCGP method in this work: $\zeta = 1$, $\rho = 0.65$, $\delta = 0.3$, $\tau = 1.65$, $c = 2$, $\eta = 0.5$, $\nu = 1.2$, $t = \frac{1}{l_k}$, and $\beta_k = \beta_k^{\mathrm{LS}}$.
- The WYL method: $\rho = 0.6$, $\eta = 10^{-3}$, $\xi = 10^{-4}$, and initial trial step length $\sigma_k = \|s_k^2\|/s_k^\top y_k$.
- The EPM method: $\sigma = 10^{-4}$, $\rho = 0.55$, $c = 0.6$, and $\xi = 1.55$.
- The PDY method: $\xi = 1$, $\rho = 0.5$, $c = 1$, and $\sigma = 0.01$.
- The IHCGP method: $\rho = 0.5$, $\varsigma = 1$, $\sigma = 1$, $\lambda = 0.001$, $\nu = 0.8$, $\mu = 3$, $\gamma = 1.5$, and $\alpha = 0.2$.

We set the dimensions of the tested problems as $n = 10^4$, $n = 5 \times 10^4$, and $n = 10^5$, respectively. Initial guesses for all the problems are $x_{(1)} = (1, 1, \ldots, 1)^\top$, $x_{(2)} = (\frac{1}{3}, \frac{1}{3^2}, \ldots, \frac{1}{3^n})^\top$, $x_{(3)} = (\frac{1}{2}, \frac{1}{2^2}, \ldots, \frac{1}{2^n})^\top$, $x_{(4)} = (0, \frac{1}{n}, \ldots, \frac{n-1}{n})^\top$, $x_{(5)} = (1, \frac{1}{2}, \ldots, \frac{1}{n})^\top$, $x_{(6)} = (\frac{1}{n}, \frac{2}{n}, \ldots, 1)^\top$, $x_{(7)} = (1 - \frac{1}{n}, 1 - \frac{2}{n}, \ldots, 0)^\top$. The iteration procedure for all methods will be

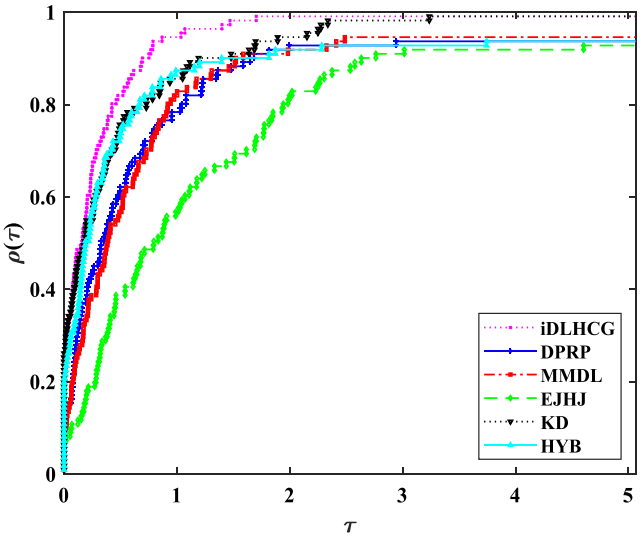


FIGURE 1 Performance profiles on Time. [Colour figure can be viewed at wileyonlinelibrary.com]

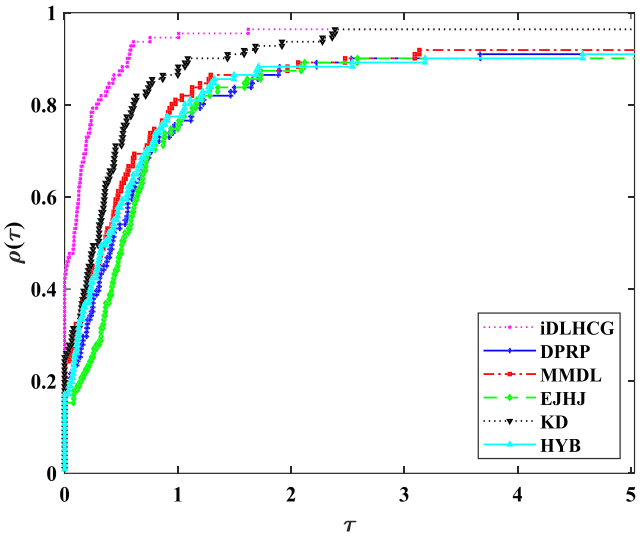


FIGURE 2 Performance profiles on Itr. [Colour figure can be viewed at wileyonlinelibrary.com]

terminated when one of the following conditions is satisfied:

$$\text{(i)}\,\|Q_k\| \le 10^{-6},\ \text{(ii)}\,\|d_k\| \le 10^{-7}\ \text{or}\ \ \text{(iii)}\,\text{Itr} > 2000.$$

Take $Q(x) = (q_1(x), q_2(x), \ldots, q_n(x))^\top$, and we choose 10 tested problems, which are listed as follows.

Problem 1 ([99]). Set $q_i(x) = e^{x_i} - 1$, for $i = 1, 2, \ldots, n$ and $C = \mathbb{R}^n_+$.

Problem 2 ([100]). Set

$$\begin{aligned} q_1(x) &= x_1 - e^{\cos\left(\frac{x_1+x_2}{n+1}\right)}, \\ q_i(x) &= x_i - e^{\cos\left(\frac{x_{i-1}+x_i+x_{i+1}}{n+1}\right)}, \text{ for } i = 2, 3, \ldots, n-1, \\ q_n(x) &= x_n - e^{\cos\left(\frac{x_{n-1}+x_n}{n+1}\right)}, \end{aligned}$$

and $C = \mathbb{R}^n_+$.

Problem 3 ([101]). Set

$$Q(x) = \begin{pmatrix} 5/2 & 1 & 0 & \dots & 0 \\ 1 & 5/2 & 1 & \dots & 0 \\ 0 & 1 & 5/2 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & 5/2 \end{pmatrix} x + (-1, -1, -1, \dots, -1)^\top$$

and $C = \mathbb{R}^n_+$.

Problem 4 ([102]). Set

$$\begin{aligned} q_1(x) &= 2x_1 + \sin(x_1) - 1, \\ q_i(x) &= 2x_{i-1} + 2x_i + \sin(x_i) - 1, \text{ for } i = 2, 3, \dots, n-1, \\ q_n(x) &= 2x_n + \sin(x_n) - 1, \end{aligned}$$

and $C = \mathbb{R}^n_+$.

Problem 5 ([102]). Set

$$\begin{aligned} q_1(x) &= x_1 - e^{\cos(\frac{x_1+x_2}{2})}, \\ q_i(x) &= x_i - e^{\cos(\frac{x_{i-1}+x_i+x_{i+1}}{i})}, \text{ for } i = 2, 3, \dots, n-1, \\ q_n(x) &= x_n - e^{\cos(\frac{x_{n-1}+x_n}{n})}, \end{aligned}$$

and $C = \mathbb{R}^n_+$.

Problem 6 ([102]). Set $q_i(x) = (e^{x_i})^2 + 3\sin(x_i)\cos(x_i) - 1$, for $i = 1, 2, \dots, n$ and $C = \mathbb{R}^n_+$.

Problem 7 ([103]). Set $q_i(x) = 2x_i - \sin|x_i|$, for $i = 1, 2, \dots, n$ and $C = \mathbb{R}^n_+$.

Problem 8 ([104]). Set $q_i(x) = \ln(|x_i| + 1) - \frac{x_i}{n}$, for $i = 1, 2, \dots, n$ and $C = \mathbb{R}^n_+$.

Problem 9 ([83]). Set $q_i(x) = x_i - 2\sin|x_i - 1|$, for $i = 1, 2, \dots, n$ and $C = \mathbb{R}^n_+$.

Problem 10 ([103]). Set

$$\begin{aligned} q_1(x) &= x_1(x_1^2 + x_2^2) - 1, \\ q_i(x) &= x_i(x_{i-1}^2 + 2x_i^2 + x_{i+1}^2) - 1, \text{ for } i = 2, 3, \dots, n-1, \\ q_n(x) &= x_n(x_{n-1}^2 + x_n^2), \end{aligned}$$

and $C = \mathbb{R}^n_+$.

To spare the space of the work, detailed results of the experiments are shown in https://www.cnblogs.com/poin07/p/18272303, where "Init" refers to the initial point, "n" is the dimension of the problem, "Itr" is the iteration numbers, "NF" means the number of function evaluations, "Time" represents the running time, and "$\|Q_*\|$" denotes the value of $\|Q_k\|$ when the program is terminated. Similar to Section 4.1, we apply the performance profiles suggested by Dolan and Moré [94] in terms of Time, NF, and Itr, respectively. From the numerical results in the tables, we see that all the five methods solve the constrained nonlinear monotone equations successfully, and the iDLHCGP method is superior to the other four methods by lower Time, NF, and Itr for all tested problems except for a few cases. Furthermore, from Figures 3–5, it is concluded that the iDLHCGP method overall outperforms the other four methods in Time, NF, and Itr. Therefore, The numerical effectiveness of the iDLHCGP method is verified.

4.3 | Application in compressive sensing

In this subsection, we will apply the iDLHCGP method to solve the sparse signal recovery problem, which requires finding sparse solutions of an under-determined linear system $Ax = b$. $A \in \mathbb{R}^{m\times n}(m \ll n)$ and the vector $b \in \mathbb{R}^m$ is the observation data. The problem is equivalent to solving the ℓ_1-norm regularized optimization problem

$$\min_{x\in\mathbb{R}^n} q(x) = \frac{1}{2}\|Ax - b\|_2^2 + \theta\|x\|_1, \tag{45}$$

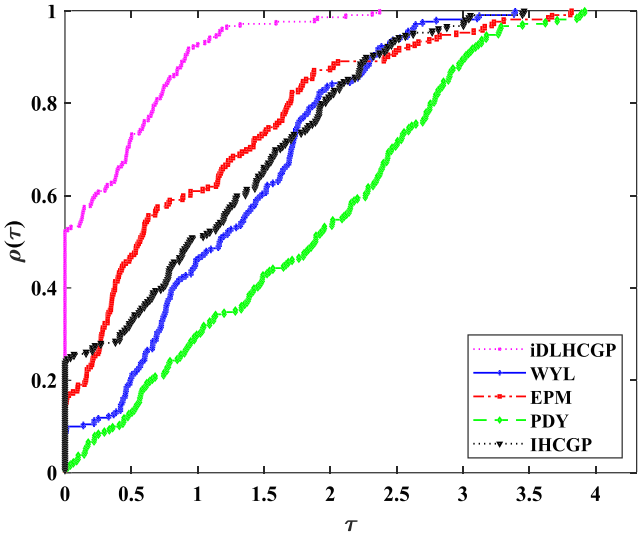


FIGURE 3 Performance profiles on Time. [Colour figure can be viewed at wileyonlinelibrary.com]

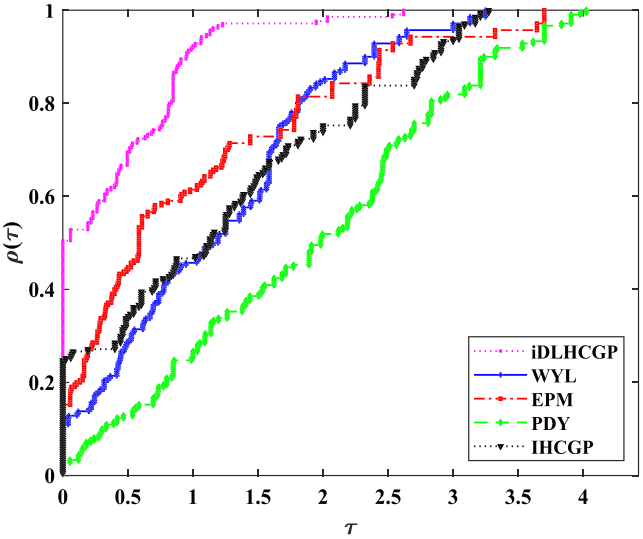


FIGURE 4 Performance profiles on NF. [Colour figure can be viewed at wileyonlinelibrary.com]

where $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the ℓ_1 and ℓ_2 norms, respectively, and $\theta > 0$ is a positive constant. Noticing the nonsmoothness of (45), Figureiredo et al. [105] proposed the earliest gradient projection method, that is, the GPSR method, for sparse signal reconstruction. The first step of the GPSR method is to reformulate problem (45) as a convex quadratic programming problem. For completeness, we briefly recall it here. For any vector $x \in \mathbb{R}^n$, we reformulate it as $x = u - v$, where $u_i = \max\{x_i, 0\}$ and $v_i = \max\{-x_i, 0\}$ for all $i = 1, 2, \ldots, n$ are two auxiliary vectors. From the definition of ℓ_1 norm, $\|x\|_1$ can be expressed as $\|x\|_1 = e_n^\top u + e_n^\top v$, where $e_n^\top = (1, 1, \ldots, 1) \in \mathbb{R}^n$. Based on [48, 105], (45) can be equivalently rewritten as the following equations:

$$\mathcal{Q}(z) = \min\{z, Hz + c\} = 0,$$

where

$$z = \begin{bmatrix} u \\ v \end{bmatrix}, \; c = \begin{bmatrix} \theta e_n - A^\top b \\ \theta e_n + A^\top b \end{bmatrix}, \; H = \begin{bmatrix} A^\top A & -A^\top A \\ -A^\top A & A^\top A \end{bmatrix}.$$

Furthermore, from [48, Lemma 2.2] and [106, Lemma 3], we obtain that $\mathcal{Q}(z)$ is Lipschitz continuous and monotone, which means problem (45) can be translated into problem (7). Therefore, the CGP method can be used to solve it.

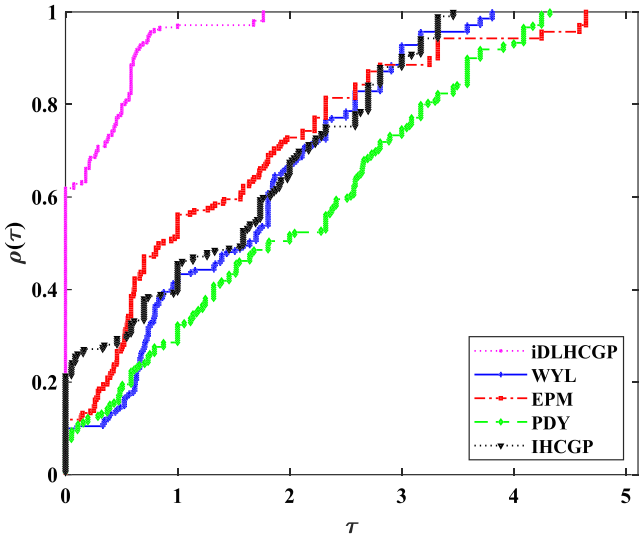


FIGURE 5 Performance profiles on Itr. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Numerical results on signal restoration.

Problem No.	Problem size (m,n,r)	iDLHCGP $\overline{Itr}/\overline{Time}/\overline{MSE}$	WYL $\overline{Itr}/\overline{Time}/\overline{MSE}$	EPM $\overline{Itr}/\overline{Time}/\overline{MSE}$	PDY $\overline{Itr}/\overline{Time}/\overline{MSE}$	IHCGP $\overline{Itr}/\overline{Time}/\overline{MSE}$
1	(256,1024,32)	80/0.02/7.470e-06	91/0.03/7.427e-06	166/0.04/8.020e-06	137/0.05/1.686e-05	158/0.04/4.794e-06
2	(384,1536,48)	81/0.03/7.641e-06	91/0.03/7.241e-06	167/0.06/8.340e-06	143/0.05/1.877e-05	156/0.05/5.460e-06
3	(512,2048,64)	75/0.05/8.151e-06	89/0.06/7.812e-06	156/0.09/8.938e-06	139/0.08/1.802e-05	154/0.10/6.780e-06
4	(640,2560,80)	77/0.14/8.833e-06	91/0.17/8.504e-06	161/0.26/9.782e-06	143/0.23/1.988e-05	156/0.27/7.681e-06
5	(768,3072,96)	79/0.25/1.135e-05	88/0.26/1.106e-05	165/0.53/1.255e-05	151/0.46/1.901e-05	163/0.52/1.088e-05
6	(896,3584,112)	77/0.36/9.946e-06	87/0.43/1.009e-05	161/0.78/1.097e-05	137/0.63/1.938e-05	161/0.76/9.467e-06
7	(1024,4096,128)	76/0.43/1.045e-05	90/0.52/1.024e-05	160/0.91/1.150e-05	136/0.78/2.401e-05	160/0.93/1.029e-05
8	(1152,4608,144)	76/0.56/1.014e-05	87/0.62/1.008e-05	158/1.09/1.117e-05	137/0.95/1.732e-05	154/1.07/1.010e-05
9	(1280,5120,160)	78/0.71/1.057e-05	85/0.78/1.017e-05	162/1.49/1.165e-05	139/1.27/2.826e-05	159/1.50/1.078e-05
10	(1408,5632,176)	78/0.93/1.026e-05	86/1.05/1.075e-05	162/1.92/1.126e-05	141/1.69/2.164e-05	156/1.90/1.038e-05
11	(1536,6144,192)	78/1.10/1.017e-05	88/1.22/1.074e-05	162/2.21/1.117e-05	141/1.87/1.935e-05	155/2.09/1.031e-05
12	(1664,6656,208)	78/1.24/1.006e-05	87/1.39/1.033e-05	160/2.55/1.104e-05	141/2.21/2.290e-05	151/2.38/1.019e-05
13	(1792,7168,224)	78/1.41/1.189e-05	87/1.54/1.167e-05	163/2.87/1.297e-05	144/2.50/1.992e-05	160/2.85/1.261e-05
14	(1920,7680,240)	76/1.61/1.030e-05	86/1.80/1.140e-05	160/3.30/1.123e-05	140/2.92/2.481e-05	158/3.31/1.077e-05
15	(2048,8192,256)	76/1.80/1.155e-05	87/2.08/1.106e-05	158/3.78/1.250e-05	140/3.37/2.701e-05	155/3.62/1.213e-05

Note: $\overline{\text{Itr}}$ is the average number of iterations, $\overline{\text{Time}}$ is the average running time in second, and $\overline{\text{MSE}}$ is the average value of MSE.Abbreviations: iDLHCGP, improved DL-style hybrid CGP; MSE, mean of squared error.

To measure the quality of restoration, we employ the mean of squared error (MSE), that is,

$$\text{MSE} = \frac{1}{n}\|x^* - x\|^2,$$

where x denotes the original signal and x^* is the restored one. It is noted that the lower value of MSE represents better quality of the restored signal. In experiment, denote the number of nonzero elements in the original signal x by r. The matrix A is given randomly and generated by the following codes in MATLAB and so is $b = Ax_0 + e$, the observation data, where e is Gaussian noise. Given a triple (m, n, r), the tested data are generated as follows:

```
x = zeros(n,1); % initialize the original sparse signal
q = randperm(n);
x(q(1:r)) = randn(r,1);
A = randn(m,n);
A = orth(A'); % orthonormalize the rows of A
b = A*x+0.001*randn(m,1);
```

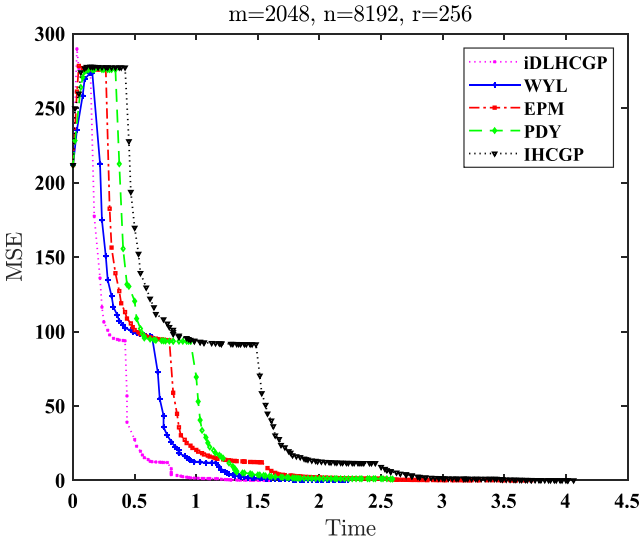


FIGURE 6 Performance profiles on $\overline{Time}$. [Colour figure can be viewed at wileyonlinelibrary.com]

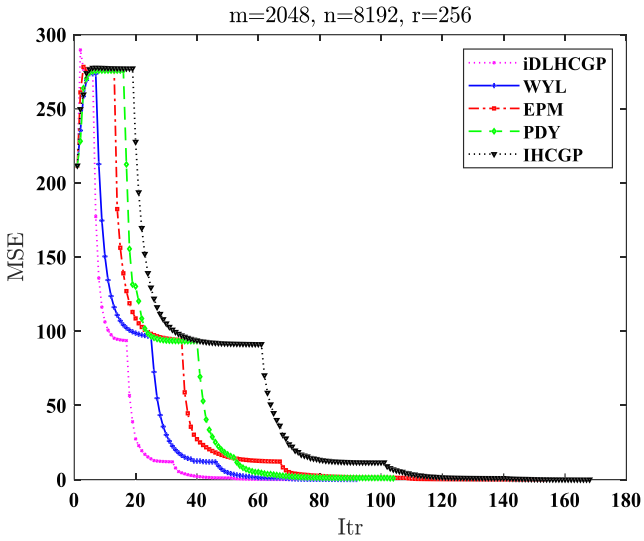


FIGURE 7 Performance profiles on $\overline{Itr}$. [Colour figure can be viewed at wileyonlinelibrary.com]

$\theta = 0.1 * \max(\text{abs}(\text{A}' * \text{b}))$;
Iterations start with $x_0 = A^\top b$, and they will be terminated when the following condition is satisfied, that is,

$$\frac{|q(x_k) - q(x_{k-1})|}{|q(x_{k-1})|} < 10^{-5}.$$

Set $(m, n, r) = (256i, 1024i, 32i)$, with $i = 1 : 0.5 : 8$, thus generating sparse signals of different lengths. For each i, we generate 20 random instances to compare the numerical performance of the iDLHCGP method with the WYL, EPM and PDY methods. Parameters in these methods are set as those in Section 4.2. Detailed numerical results are shown in Table 1. $\overline{\text{Itr}}$ is the average number of iterations, $\overline{\text{Time}}$ is the average running time in second, and $\overline{\text{MSE}}$ is the average value of MSE. In addition, Figures 6 and 7 visually display the trends of MSE concerning Time and Itr in the case when $(m, n, r) = (2048,8192, 256)$. From Table 1 and Figures 6–8, we see that all the five methods solve the signal recovery problem successfully, and the iDLHCGP method outperforms the other four methods both in $\overline{\text{Itr}}$ and $\overline{\text{Time}}$. Moreover, the iDLHCGP method is competitive in terms of $\overline{\text{MSE}}$, and from Figures 6 and 7, we conclude that the iDLHCGP method produces a lower MSE within a shorter Time and Itr. Thus, the effectiveness of the iDLHCGP method in signal recovery is indicated.

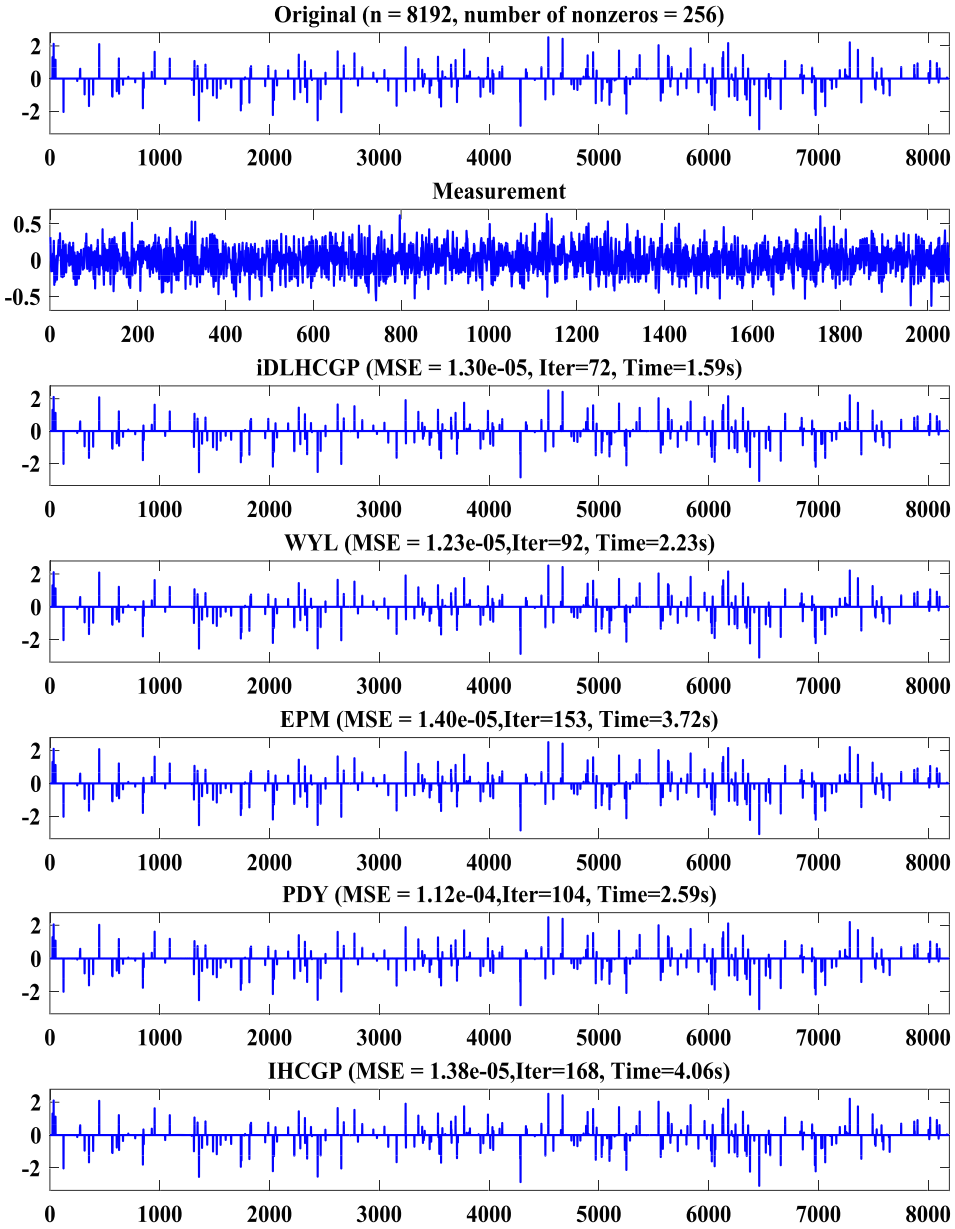


FIGURE 8 From top to bottom: the original signal, the measurement, and the recovered signals by five methods. [Colour figure can be viewed at wileyonlinelibrary.com]

5 | CONCLUSIONS

In this work, at first, a new hybrid CG method for efficiently solving unconstrained optimization is introduced based on the DL-style CG methods [43, 44] and the hybridization-based self-adaptive technique. The search direction generated by the proposed method possesses the sufficient descent and trust region properties regardless of any line search. The global convergence is established under common assumptions and standard Wolfe line search. Then we extend the iDLHCG method for constrained nonlinear monotone equations. With the hyperplane projection technique and a new well-defined self-adaptive line search, the iDLHCGP method is proposed. We establish the globally convergent results for the iDL-HCGP method under milder conditions, that is, the Lipschitz continuity is not required. Besides, the convergence rates analysis for the two methods are given in Sections 2.3 and 3.3, respectively. Numerical experiments conducted on unconstrained optimization and constrained nonlinear monotone equations verify the effectiveness of the proposed methods. In addition, the iDLHCGP method is applied in signal recovery problem, and the results of the practical application show the iDLHCGP method is promising.

In subsequent studies, there are at least five aspects intriguing and worthy of investigation. (i) For constrained convex optimization, it can be converted to unconstrained convex optimization based on the penalization technique, and then the iDLHCG method can be applied thereby. (ii) Since the iDLHCG method uses the gradient information, the stochastic gradient estimator can be considered. Hence, the iDLHCG method is adapted to stochastic optimization, lowering the computational cost when the scale of the problem is very large. (iii) In parallelization and distributed computing, the original problem, whose scale may be very large, is split into several subproblems. Each subproblem can be solved independently, and therefore, the iDLHCG method or the iDLHCGP method can be applied to solve them, which improves the computational efficiency and lowers the computational cost. (iv) Many problems appearing in machine learning and deep learning can be formulated as unconstrained optimization problems, using the iDLHCG method to solve them can accelerate the optimization process and improve the convergence speed. (v) In practical applications, the data may be noisy or incomplete, in which case the regularization techniques can be used to improve the robustness of the method. Then the iDLHCG method or the iDLHCGP method can be applied to solve the regularized problems.

AUTHOR CONTRIBUTIONS

Zihang Yuan: Software; writing—original draft; methodology. **Hu Shao:** Methodology; conceptualization; funding acquisition; supervision; project administration. **Xiaping Zeng:** Methodology; funding acquisition. **Pengjie Liu:** Methodology; software; conceptualization; writing—original draft; formal analysis. **Xianglin Rong:** Software; Formal analysis. **Jianhao Zhou:** Methodology.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China (No. 72071202), the Natural Science Foundation of Guangxi Province (No. 2024GXNSFBA010395), the Enhancing Young and Middle-aged Teacher's Research Basic Ability in Colleges of Guangxi (No. 2021KY0592), and the Key Laboratory of Mathematics and Engineering Applications, Ministry of Education.

CONFLICT OF INTEREST STATEMENT

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 paper.

ORCID

Hu Shao https://orcid.org/0000-0001-7816-0707
Pengjie Liu https://orcid.org/0000-0002-9176-6321

REFERENCES

1. E. Jandaghi, X. Chen, and C. Yuan, *Motion dynamics modeling and fault detection of a soft trunk robot*, 2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), 2023, pp. 1324–1329.
2. M. Javad Ebadi, A. Fahs, H. Fahs, and R. Dehghani, *Competitive secant (BFGS) methods based on modified secant relations for unconstrained optimization*, Optimization **72** (2023), no. 7, 1691–1706.
3. G. Yuan, Y. Zhou, and M. Zhang, *A hybrid conjugate gradient algorithm for nonconvex functions and its applications in image restoration problems*, J. Oper. Res. Soc. China **11** (2023), 759–781.
4. X. Wang, G. Yuan, and L. Pang, *A class of new three-term descent conjugate gradient algorithms for large-scale unconstrained optimization and applications to image restoration problems*, Numer. Algorithms **93** (2023), no. 3, 949–970.
5. J. Lu, G. Yuan, and Z. Wang, *A modified Dai–Liao conjugate gradient method for solving unconstrained optimization and image restoration problems*, J. Appl. Math. Comput. **68** (2022), 681–703.
6. G. Yuan, J. Lu, and Z. Wang, *The PRP conjugate gradient algorithm with a modified WWP line search and its application in the image restoration problems*, Appl. Numer. Math. **152** (2020), 1–11.
7. G. Yuan, J. Lu, and Z. Wang, *The modified PRP conjugate gradient algorithm under a non-descent line search and its application in the Muskingum model and image restoration problems*, Soft. Comput. **25** (2021), no. 8, 5867–5879.
8. X. Jiang, H. Yang, J. Jian, and X. Wu, *Two families of hybrid conjugate gradient methods with restart procedures and their applications*, Optim. Meth. Softw. **38** (2023), no. 5, 947–974.
9. J. Liu and X. L. Du, *A gradient projection method for the sparse signal reconstruction in compressive sensing*, Appl. Anal. **97** (2018), no. 12, 2122–2131.

10. Y. Liu, Z. Zhu, and B. Zhang, *Two sufficient descent three-term conjugate gradient methods for unconstrained optimization problems with applications in compressive sensing*, J. Appl. Math. Comput. **68** (2022), 1787–1816.
11. A. M. Awwal, L. Wang, P. Kumam, M. I. Sulaiman, S. Salisu, N. Salihu, and P. Yodjai, *Generalized RMIL conjugate gradient method under the strong Wolfe line search with application in image processing*, Math. Meth. Appl. Sci. **46** (2023), 17544–17556.
12. J. Yin, J. Jian, X. Jiang, and X. Wu, *A family of inertial-relaxed DFPM-based algorithms for solving large-scale monotone nonlinear equations with application to sparse signal restoration*, J. Comput. Appl. Math. **419** (2023), 114674.
13. Z. Zhu, J. Ma, and B. Zhang, *A new conjugate gradient hard thresholding pursuit algorithm for sparse signal recovery*, Comput. Appl. Math. **39** (2020), 1–20.
14. Z. Avazzadeh, H. Hassani, P. Agarwal, S. Mehrabi, M. Javad Ebadi, and M. S. Dahaghin, *An optimization method for studying fractional-order tuberculosis disease model via generalized Laguerre polynomials*, Soft Comput. **27** (2023), no. 14, 9519–9531.
15. Z. Avazzadeh, H. Hassani, P. Agarwal, S. Mehrabi, M. Javad Ebadi, and M. K. Hosseini Asl, *Optimal study on fractional fascioliasis disease model based on generalized fibonacci polynomials*, Math. Meth. Appl. Sci. **46** (2023), no. 8, 9332–9350.
16. R. Fletcher and C. M. Reeves, *Function minimization by conjugate gradients*, The Comput. J. **7** (1964), no. 2, 149–154.
17. Y.H. Dai and Y. Yuan, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optim. **10** (1999), no. 1, 177–182.
18. R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, 2000.
19. E. Polak and G. Ribiere, *Note sur la convergence de méthodes de directions conjuguées*, Rev. Fr. D'inform. Rech. Opér. Série Rouge **3** (1969), no. 16, 35–43.
20. B. T. Polyak, *The conjugate gradient method in extremal problems*, USSR Comput. Math. Math. Phys. **9** (1969), no. 4, 94–112.
21. M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Natl. Bureau Stand. **49** (1952), no. 6, 409–435.
22. Y. Liu and C. Storey, *Efficient generalized conjugate gradient algorithms, part 1: theory*, J. Optim. Theory Appl. **69** (1991), 129–137.
23. S. M. Ibrahim, U. A. Yakubu, and M. Mamat, *Application of spectral conjugate gradient methods for solving unconstrained optimization problems*, An Int. J. Optim. Control: Theor. Appl. (IJOCTA) **10** (2020), no. 2, 198–205.
24. P. Kaelo, S. Narayanan, and M. V. Thuto, *A modified quadratic hybridization of Polak-Ribiere-Polyak and Fletcher–Reeves conjugate gradient method for unconstrained optimization problems*, An Int. J. Optim. Control: Theor. Appl. (IJOCTA) **7** (2017), no. 2, 177–185.
25. Q. Hu, H. Zhang, Z. Zhou, and Y. Chen, *A class of improved conjugate gradient methods for nonconvex unconstrained optimization*, Numer. Linear Algebra Appl. **30** (2023), no. 4, e2482.
26. X. Jiang, Y. Zhu, and J. Jian, *Two efficient nonlinear conjugate gradient methods with restart procedures and their applications in image restoration*, Nonlinear Dyn. **111** (2023), no. 6, 5469–5498.
27. X. Jiang, W. Liao, J. Yin, and J. Jian, *A new family of hybrid three-term conjugate gradient methods with applications in image restoration*, Numer. Algorithms **91** (2022), no. 1, 161–191.
28. J. Jian, P. Liu, X. Jiang, and B. He, *Two improved nonlinear conjugate gradient methods with the strong Wolfe line search*, Bull. Iranian Math. Soc. **48** (2022), 2297–2319.
29. J. Jian, P. Liu, X. Jiang, and C. Zhang, *Two classes of spectral conjugate gradient methods for unconstrained optimizations*, J. Appl. Math. Comput. **68** (2022), no. 6, 4435–4456.
30. J. Jian, W. Chen, X. Jiang, and P. Liu, *A three-term conjugate gradient method with accelerated subspace quadratic optimization*, J. Appl. Math. Comput. **68** (2022), 2407–2433.
31. W. Cheng, Q. Hu, and D. Li, *A fast conjugate gradient algorithm with active set prediction for ℓ_1 optimization*, Optim. Meth. Softw. **34** (2019), no. 6, 1277–1305.
32. Q. Hu, H. Zhang, and Y. Chen, *Global convergence of a descent PRP type conjugate gradient method for nonconvex optimization*, Appl. Numer. Math. **173** (2022), 38–50.
33. J. K. Liu, Y. X. Zhao, and X. L. Wu, *Some three-term conjugate gradient methods with the new direction structure*, Appl. Numer. Math. **150** (2020), 433–443.
34. M. K. Riahi and I. A. Qattan, *On the convergence rate of fletcher-reeves nonlinear conjugate gradient methods satisfying strong Wolfe conditions: application to parameter identification in problems governed by general dynamics*, Math. Meth. Appl. Sci. **45** (2022), no. 7, 3644–3664.
35. P. Liu, Z. Yuan, Y. Zhuo, and H. Shao, *Two efficient spectral hybrid CG methods based on memoryless BFGS direction and Dai–Liao conjugacy condition*, Optim. Meth. Softw. **2024** (2024), 1–19.
36. Z. Wei, S. Yao, and L. Liu, *The convergence properties of some new conjugate gradient methods*, Appl. Math. Comput. **183** (2006), no. 2, 1341–1350.
37. H. Huang, Z. Wei, and S. Yao, *The proof of the sufficient descent condition of the Wei–Yao–Liu conjugate gradient method under the strong Wolfe–Powell line search*, Appl. Math. Comput. **189** (2007), no. 2, 1241–1245.
38. S. Yao, Z. Wei, and H. Huang, *A note about Wyl's conjugate gradient method and its applications*, Appl. Math. Comput. **191** (2007), no. 2, 381–388.
39. X. Jiang, L. Han, and J. Jian, *A globally convergent mixed conjugate gradient method with Wolfe line search*, Math. Numer. Sin. **34** (2012), no. 1, 103.
40. J. Jian, L. Han, and X. Jiang, *A hybrid conjugate gradient method with descent property for unconstrained optimization*, Appl. Math. Model. **39** (2015), no. 3-4, 1281–1290.
41. Y. X. Yuan and J. Stoer, *A subspace study on conjugate gradient algorithms*, Z. Angew. Math. Mech. **75** (1995), no. 1, 69–77.
42. J. K. Liu and S. J. Li, *New hybrid conjugate gradient method for unconstrained optimization*, Appl. Math. Comput. **245** (2014), 36–43.
43. Y. H. Dai and L. Liao, *New conjugacy conditions and related nonlinear conjugate gradient methods*, Appl. Math. Optim. **43** (2001), 87–101.

44. Z. Aminifard and S. Babaie-Kafaki, *Dai–Liao extensions of a descent hybrid nonlinear conjugate gradient method with application in signal processing*, Numer. Algorithms **89** (2022), no. 3, 1369–1387.
45. Y. H. Dai and Y. Yuan, *An efficient hybrid conjugate gradient method for unconstrained optimization*, Ann. Oper. Res. **103** (2001), 33–47.
46. X. Yang, Z. Luo, and X. Dai, *A global convergence of LS-CD hybrid conjugate gradient method*, Adv. Numer. Anal. **2013** (2013), no. 1, 517452.
47. K. Meintjes and A. P. Morgan, *A methodology for solving chemical equilibrium systems*, Appl. Math. Comput. **22** (1987), no. 4, 333–361.
48. Y. Xiao, Q. Wang, and Q. Hu, *Non-smooth equations based method for ℓ_1-norm problems with applications to compressed sensing*, Nonlinear Anal.: Theory, Meth. Appl. **74** (2011), no. 11, 3570–3577.
49. X. Wu, H. Shao, P. Liu, Y. Zhang, and Y. Zhuo, *An efficient conjugate gradient-based algorithm for unconstrained optimization and its projection extension to large-scale constrained nonlinear equations with applications in signal recovery and image denoising problems*, J. Comput. Appl. Math. **422** (2023), 114879.
50. J. Yin, J. Jian, X. Jiang, M. Liu, and L. Wang, *A hybrid three-term conjugate gradient projection method for constrained nonlinear monotone equations with applications*, Numer. Algorithms **88** (2021), 389–418.
51. P. Liu, H. Shao, Y. Wang, and X. Wu, *A three-term CGPM-based algorithm without Lipschitz continuity for constrained nonlinear monotone equations with applications*, Appl. Numer. Math. **175** (2022), 98–107.
52. Y. Xiao and H. Zhu, *A conjugate gradient method to solve convex constrained monotone equations with applications in compressive sensing*, J. Math. Anal. Appl. **405** (2013), no. 1, 310–319.
53. P. Kumam, A. B. Abubakar, A. H. Ibrahim, H. U. Kura, B. Panyanak, and N. Pakkaranang, *Another hybrid approach for solving monotone operator equations and application to signal processing*, Math. Meth. Appl. Sci. **45** (2022), no. 12, 7897–7922.
54. J. Yin, J. Jian, and X. Jiang, *A generalized hybrid CGPM-based algorithm for solving large-scale convex constrained equations with applications to image restoration*, J. Comput. Appl. Math. **391** (2021), 113423.
55. B. Zhang, Z. Zhu, and S. Li, *A modified spectral conjugate gradient projection algorithm for total variation image restoration*, Appl. Math. Lett. **27** (2014), 26–35.
56. P. Liu, X. Wu, H. Shao, Y. Zhang, and S. Cao, *Three adaptive hybrid derivative-free projection methods for constrained monotone nonlinear equations and their applications*, Numer. Linear Algebra Appl. **30** (2023), no. 2, e2471.
57. A. H. Ibrahim, P. Kumam, and W. Kumam, *A family of derivative-free conjugate gradient methods for constrained nonlinear equations and image restoration*, IEEE Access **8** (2020), 162714–162729.
58. W. Hu, J. Wu, and G. Yuan, *Some modified Hestenes-Stiefel conjugate gradient algorithms with application in image restoration*, Appl. Numer. Math. **158** (2020), 360–376.
59. A. B. Abubakar, H. Mohammad, P. Kumam, S. A. Rano, A. H. Ibrahim, and A. I. Kiri, *On the new spectral conjugate gradient-type method for monotone nonlinear equations and signal recovery*, Math. Meth. Appl. Sci. **1–23** (2022).
60. J. Sun, L. Kong, and S. Zhou, *Gradient projection newton algorithm for sparse collaborative learning using synthetic and real datasets of applications*, J. Comput. Appl. Math. **422** (2023), 114872.
61. H. Wang, J. Du, H. Su, and H. Sun, *A linearly convergent self-adaptive gradient projection algorithm for sparse signal reconstruction in compressive sensing*, AIMS Math. **8** (2023), no. 6, 14726–14746.
62. B. Xue, J. Du, H. Sun, and Y. Wang, *A linearly convergent proximal ADMM with new iterative format for BPDN in compressed sensing problem*, AIMS Math. **7** (2022), no. 6, 10513–10533.
63. H. Liu, Q. Chen, and P. Liu, *An optimization method of large-scale video stream concurrent transmission for edge computing*, Mathematics **11** (2023), no. 12, 2622.
64. H. C. Sun, M. Sun, and H. C. Zhou, *A proximal splitting method for separable convex programming and its application to compressive sensing*, J. Nonlinear Sci. Appl. **9** (2016), no. 2, 392–403.
65. H. Sun, M. Tian, and M. Sun, *The symmetric ADMM with indefinite proximal regularization and its application*, J. Inequal. Appl. **2017** (2017), 1–22.
66. M. Sun, H. Sun, and Y. Wang, *Two proximal splitting methods for multi-block separable programming with applications to stable principal component pursuit*, J. Appl. Math. Comput. **56** (2018), 411–438.
67. M. Sun and H. Sun, *Improved proximal ADMM with partially parallel splitting for multi-block separable convex programming*, J. Appl. Math. Comput. **58** (2018), 151–181.
68. J. Sun and W. Qu, *DCA for sparse quadratic kernel-free least squares semi-supervised support vector machine*, Mathematics **10** (2022), no. 15, 2714.
69. S. Ji, G. Li, D. Zhang, and X. Zhang, *Approximation algorithm for the capacitated correlation clustering problem with penalties*, J. Combinator. Optim. **45** (2023), 12.
70. X. Tong and S. Zhou, *A smoothing projected newton-type method for semismooth equations with bound constraints*, J. Ind. Manag. Optim. **1** (2005), no. 2, 235–250.
71. L. Qi, X. J. Tong, and D. H. Li, *Active-set projected trust-region algorithm for box-constrained nonsmooth equations*, J. Optim. Theory Appl. **120** (2004), 601–625.
72. Q. Yan, X. Peng, and D. Li, *A globally convergent derivative-free method for solving large-scale nonlinear monotone equations*, J. Comput. Appl. Math. **234** (2010), no. 3, 649–657.
73. J. Yin, J. Jian, and X. Jiang, *A spectral gradient projection algorithm for convex constrained nonsmooth equations based on an adaptive line search*, Math. Numer. Sin. **42** (2020), no. 4, 457.
74. P. Liu, J. Jian, and X. Jiang, *A new conjugate gradient projection method for convex constrained nonlinear equations*, Complexity **2020** (2020), no. 1, 8323865.
75. M. Sun and J. Liu, *New hybrid conjugate gradient projection method for the convex constrained equations*, Calcolo **53** (2016), 399–411.

76. A. H. Ibrahim, P. Kumam, A. B. Abubakar, and J. Abubakar, *A derivative-free projection method for nonlinear equations with non-Lipschitz operator: application to lasso problem*, Math. Meth. Appl. Sci. **46** (2023), no. 8, 9006–9027.
77. M. V. Solodov and B. F. Svaiter, *A globally convergent inexact newton method for systems of monotone equations*, Reform.: Nonsmooth, Piecewise Smooth, Semismooth Smoothing Methods **22** (1999), 355–369.
78. W. Cheng, *A PRP type method for systems of monotone equations*, Math. Comput. Model. **50** (2009), no. 1-2, 15–20.
79. G. Zoutendijk, *Nonlinear programming, computational methods*, Integer Nonlinear Programming, 1970, pp. 37–86.
80. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, 2000.
81. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
82. W. Sun and Y. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer Science & Business Media, 2006.
83. L. Zhang and W. Zhou, *Spectral gradient projection method for solving nonlinear monotone equations*, J. Comput. Appl. Math. **196** (2006), no. 2, 478–484.
84. A. B. Muhammad, C. Tammer, A. M. Awwal, R. Elster, and Z. Ma, *Inertial-type projection methods for solving convex constrained monotone nonlinear equations with applications to robotic motion control*, J. Nonlinear Var. Anal. **5** (2021), no. 5, 831–849.
85. K. Amini and A. Kamandi, *A new line search strategy for finding separating hyperplane in projection-based methods*, Numer. Algorithms **70** (2015), 559–570.
86. B. T. Polyak, *Introduction to optimization. Optimization software*, Vol. **1**, Inc., Publications Division, New York, 1987, pp. 1.
87. I. Bongartz, A. R. Conn, N. Gould, and P. L. Toint, *Cute: constrained and unconstrained testing environment*, ACM Trans. Math. Softw. (TOMS) **21** (1995), no. 1, 123–160.
88. N. Andrei, *An unconstrained optimization test functions collection*, Adv. Model. Optim. **10** (2008), no. 1, 147–161.
89. J. J. Moré, B. S. Garbow, and K. E. Hillstrom, *Testing unconstrained optimization software*, ACM Trans. Math. Softw. (TOMS) **7** (1981), no. 1, 17–41.
90. Z. Dai and F. Wen, *Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property*, Appl. Math. Comput. **218** (2012), no. 14, 7421–7430.
91. P. S. Stanimirović, B. Ivanov, S. Djordjević, and I. Brajević, *New hybrid conjugate gradient and Broyden–Fletcher–Goldfarb–Shanno conjugate gradient methods*, J. Optim. Theory Appl. **178** (2018), 860–884.
92. C. Kou and Y.-H. Dai, *A modified self-scaling memoryless Broyden–Fletcher–Goldfarb–Shanno method for unconstrained optimization*, J. Optim. Theory Appl. **165** (2015), 209–224.
93. A. H. Ibrahim, P. Kumam, A. Kamandi, and A. B. Abubakar, *An efficient hybrid conjugate gradient method for unconstrained optimization*, Optim. Meth. Softw. **37** (2022), no. 4, 1370–1383.
94. E. D. Dolan and J. J. Moré, *Benchmarking optimization software with performance profiles*, Math. Progr. **91** (2002), 201–213.
95. Y. Hu and Z. Wei, *Wei–Yao–Liu conjugate gradient projection algorithm for nonlinear monotone equations with convex constraints*, Int. J. Comput. Math. **92** (2015), no. 11, 2261–2272.
96. J. K. Liu, Z. L. Lu, J. L. Xu, S. Wu, and Z. W. Tu, *An efficient projection-based algorithm without Lipschitz continuity for large-scale nonlinear pseudo-monotone equations*, J. Comput. Appl. Math. **403** (2022), 113822.
97. J. Liu and Y. Feng, *A derivative-free iterative method for nonlinear monotone equations with convex constraints*, Numer. Algorithms **82** (2019), 245–262.
98. J. Jian, J. Yin, C. Tang, and D. Han, *A family of inertial derivative-free projection methods for constrained nonlinear pseudo-monotone equations with applications*, Comput. Appl. Math. **41** (2022), no. 7, 309.
99. C. Wang, Y. Wang, and C. Xu, *A projection method for a system of nonlinear monotone equations with convex constraints*, Math. Meth. Oper. Res. **66** (2007), 33–46.
100. G. Yu, S. Niu, and J. Ma, *Multivariate spectral gradient projection method for nonlinear monotone equations with convex constraints*, J. Ind. Manag. Optim. **9** (2013), no. 1, 117–129.
101. A. B. Abubakar and P. Kumam, *A descent Dai-Liao conjugate gradient method for nonlinear equations*, Numer. Algorithms **81** (2019), 197–210.
102. P. Gao, C. He, and Y. Liu, *An adaptive family of projection methods for constrained monotone nonlinear equations with applications*, Appl. Math. Comput. **359** (2019), 1–16.
103. W. Zhou and D. Li, *Limited memory BFGS method for nonlinear monotone equations*, J. Comput. Math. **25** (2007), 89–96.
104. W. L. Cruz and M. Raydan, *Nonmonotone spectral methods for large-scale nonlinear systems*, Optim. Meth. Softw. **18** (2003), no. 5, 583–599.
105. M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, *Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems*, IEEE J. Sel. Top. Sig. Process. **1** (2007), no. 4, 586–597.
106. J. S. Pang, *Inexact newton methods for the nonlinear complementarity problem*, Math. Progr. **36** (1986), no. 1, 54–71.

How to cite this article: Z. Yuan, H. Shao, X. Zeng, P. Liu, X. Rong, and J. Zhou, *An improved Dai-Liao-style hybrid conjugate gradient-based method for solving unconstrained nonconvex optimization and extension to constrained nonlinear monotone equations*, Math. Meth. Appl. Sci. **48** (2025), 1563–1588, DOI 10.1002/mma.10396.