Applied Mathematics and Computational Intelligence
Volume 13, No. 1, Feb 2024 [52 – 68]

# A New Hybrid Three-Term HS-DY Conjugate Gradient In Solving Unconstrained Optimization Problems

Ishak M. A. I.[1] and Marjugi S. M.[1*]

[1]Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, Jalan Universiti 1, 43400 Serdang, Selangor, Malaysia.

* Corresponding author: smahani@upm.edu.my

Received: 31 July 2023; Revised: 2 November 2023; Accepted: 26 January 2024; Published: 29 February 2024

**ABSTRACT**

*Conjugate Gradient (CG) method is an interesting tool to solve optimization problems in many fields, such design, economics, physics and engineering. Until now, many CG methods have been developed to improve computational performance and have applied in the real-world problems. Combining two CG parameters with distinct denominators may result in non-optimal outcomes and congestion.In this paper, a new hybrid three-term CG method is proposed for solving unconstrained optimization problems. The hybrid three-term search direction combines Hestenes-Stiefel (HS) and Dai-Yuan (DY) CG parameters which standardized by using a spectral to determine the suitable conjugate parameter choice and it satisfies the sufficient descent condition. Additionally, the global convergence was proved under standard Wolfe conditions and some suitable assumptions. Furthermore, the numerical experiments showed the proposed method is most robust and superior efficiency compared to some existing methods.*

**Keywords:** Unconstrained Optimization, Three-Term Conjugate Gradient, Memoryless Quasi-Newton Method, Line Search, Global Convergence.

## 1 INTRODUCTION

Consider the subsequent large-scale unconstrained optimization problem:

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

The function $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and its gradient $r_k := \nabla f(x_k)$ is a Lipschitz continuous. The Newton method, the quasi-Newton method, and some of their variants [1–3] are some approaches for addressing unconstrained optimization problems (1). However, the approaches are not preferable for large-scale problems since they involve computing and storing the Hessian matrix at each iteration. Specifically, the Hessian matrix turns singular when the approaches fail. As a result, Conjugate Gradient (CG) were developed to overcome those problems due to its ease of implementation, Hessian free and low storage requirements [4].
The sequence $\{x_k\}$ is generated by the iterative formula to solve equation (1) which as follows

$$x_{k+1} = x_k + s_k, \quad s_k = \alpha_k d_k, \quad k = 0, 1, 2, \ldots \tag{2}$$

where $\alpha_k$ is a step length which $\alpha_k > 0$ and $d_k$ s the search direction, defined by

$$d_k = \begin{cases} -r_k, & \text{if } k = 0, \\ -r_k + \beta_k d_{k-1}, & \text{if } k > 0. \end{cases} \tag{3}$$

The step length $\alpha_k$ is generated by computing the suitable the line search conditions. The step length satisfies the standard Wolfe line search, in which

$$f(x_k + \alpha_k d_k) - f(x_k) \le \eta_1 \alpha_k r_k^T d_k, \tag{4}$$

$$r(x_k + \alpha_k d_k)^T d_k \ge \eta_2 r_k^T d_k, \tag{5}$$

where $0 < \eta_1 < \eta_2 < 1$. The search direction is required to satisfy the sufficient descent condition

$$r_k^T d_k \le -t\|r_k\|^2, \quad t > 0. \tag{6}$$

The $\beta_k$ is the conjugate gradient parameter that defines the global convergence properties and numerical performance of various conjugate gradient methods. The most well-known conjugate gradient methods include Hestenes-Stiefel (HS) [5], Polak-Ribiere-Polyak (PRP) [6, 7], Liu-Storey (LS) [8], Dai-Yuan (DY) [9], Fletcher-Reeves (FR) [10], and Conjugate Descent (CD) [11]. These methods are described as follows:

$$\beta_k^{\text{FR}} = \frac{\|r_k\|^2}{\|r_{k-1}\|^2}, \quad \beta_k^{\text{CD}} = \frac{\|r_k\|^2}{-r_{k-1}^T d_{k-1}}, \quad \beta_k^{\text{DY}} = \frac{\|r_k\|^2}{d_{k-1}^T y_{k-1}}. \tag{7}$$

$$\beta_k^{\text{PRP}} = \frac{r_k^T y_{k-1}}{\|r_{k-1}\|^2}, \quad \beta_k^{\text{LS}} = \frac{r_k^T y_{k-1}}{-r_{k-1}^T d_{k-1}}, \quad \beta_k^{\text{HS}} = \frac{r_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \tag{8}$$

where $\|.\|$ is the Euclidean norm in $\mathbb{R}^n$ and $y_{k-1} = r_k - r_{k-1}$.

Babaie-Kafaki and Ghanbari [12] propose that methods incorporating a common term like $r_k^T y_{k-1}$ tend to perform well in practical situations. According to [13, 14], these methods may not consistently show improvement due to jamming and demonstrating differences when compared to methods utilizing a common term $\|r_k\|^2$. Babaie-Kafaki and Mahdavi-Amiri [15] emphasized the quest to enhance the effectiveness of these strategies and prevent potential issues, researchers are exploring the combination of methods from both groups. From a theoretical perspective, Hager and Zhang [4] contend that global convergence theorems for methods using a common term $\|r_k\|^2$ only require the Lipchitz assumption, diverging from other choices of update parameters that necessitate boundedness assumptions. Powell [16] underscores that jamming is the primary factor contributing to the nonoptimal practical performance of the FR method. Babaie-Kafaki [17] observes that when a bad direction and a small step occur between $x_k$ and $x_{k-1}$, the ensuing direction $d_k$ and step length $\alpha_k$ are likely to be less than optimal unless a gradient restart is employed. Nevertheless, Babaie-Kafaki [18] also highlighted that methods employing a common term inherently exhibit an inherent feature resembling an approximate restart to address the jamming issue. According to Andrei [19, 20], the newly computed search direction $d_k$ closely aligns with the steepest descent direction $-r_k$ when a small value of $\beta_k$ is generated due to the poor step $s_{k-1}$, where the gradient difference $y_{k-1}$ in the numerator approaches zero.

Wang [21] proposed a spectral that provides the optimal step size strategy in the gradient method as a new scheme for determining the conjugate parameters and the new search direction meets both the sufficient descent and spectral conditions. The global convergence is then proven under some suitable assumptions. The spectral parameter $\theta_k$ is described below

$$\theta_k = \max\{\min\{\alpha_k^*, \bar{\rho}_k\}, \rho_k\} \tag{9}$$

where $\alpha_k^* = -\dfrac{s_{k-1}^T r_{k-1}}{\upsilon \|y_{k-1}\|^2 \rho_k}$, $\bar{\rho}_k = \dfrac{\|s_{k-1}\|^2}{s_{k-1}^T y_{k-1}}$, $\rho_k = \dfrac{s_{k-1}^T y_{k-1}}{\|y_{k-1}\|^2}$ and $\upsilon$ is positive value.

Motivated by suitable conjugate parameters choice spectral proposed by Wang [21] and the issues discussed by Andrei [13, 14] and Babaie-Kafaki [12] in addressing convergence and jamming [12], this paper aims to tackle these challenges. The primary objective is to prevent jamming by considering a combined analysis of the norms $\|r_k\|^2$, $\|s_{k-1}\|^2$, and $\|y_{k-1}\|^2$. This modification involves computing the maximum of these norms, serving as a new adjustable parameter dynamically influencing the CG update. Equation (16) introduces a critical decision point in the CG update process, which depends on the value of $z_k$ calculated in Equation (15). If $z_k$ equals $\|y_{k-1}\|^2$, the update direction becomes $y_{k-1}$; otherwise, it remains $r_k$. This decision is pivotal in preventing jamming and ensuring convergence throughout the iterative process. The approach of these equations is to amalgamate the strengths of various CG schemes and dynamically adapt update parameters to address jamming issues. By assessing the norms and switching between update directions based on the value of $z_k$, these equations enhance the performance of CG methods and aligning with discussions by various authors in the existing literature.

To enhance the standard two-term direction, researchers have explored the development of hybrid and three-term CG methods. As the method presented by Andrei [22], it involves modifying the inverse Hessian approximation within the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula while ensuring the adherence of the search direction to descent and conjugacy principles. Liu and Li [23] purposed a hybrid CG approach combining features of LS and DY methods through a convex combination. This results in a search direction satisfying both the Dai-Liao (DL) conjugacy condition and the Newton direction which the added advantage of achieving global convergence through a strong Wolfe line search. Xu and Kong [24] proposed two hybrid methods that combine the PRP method with FR and the HS method with DY, respectively. Both hybrids yield descent directions and achieve global convergence through Wolfe line search. Dong [25] has devised a modified HS method that not only adheres to the descent condition but also closely approximates the Newton method. Min Li [26] suggests a three-term PRP CG method closely resembling the memoryless BFGS quasi-Newton method. This method reverts to the classical PRP approach under exact line search conditions and when the descent condition is satisfied regardless of the line search considerations. The satisfactory line search strategies contribute to its global convergence and numerical results indicate its effectiveness in solving unconstrained optimization problems. Additionally, Min Li [27] introduces a nonlinear CG algorithm generating a search direction akin to the memoryless BFGS quasi-Newton method. Notably, this search direction meets the descent condition and the global convergence has been established for both strongly convex and nonconvex functions under strong Wolfe line search. Abubakar [28] proposes a hybrid three-term CG algorithm where the search direction is determined using the limited memory BFGS method. This approach successfully meets the requirements of both sufficient descent and trust region, establishing global convergence under specific conditions and showcasing efficiency when compared to certain previously suggested methods. In addition, Kumam [29] and Deepho [30] introduce modifications to hybrid three-term CG approaches incorporating combinations of HS and LS, as well as CD and DY, and offering a scaled preconditioner for the hybrid parameters. These adjustments utilize the existing conjugate gradient parameters, yielding favorable outcomes in resolving a range of test problems for both methodologies. A comparable concept was implemented in [31] and [32] with diverse combinations of conjugate parameters.

Inspired by ideas presented in [28–30], we introduce a fresh hybrid three-term CG method designed for solving equation (1). Termed the hybrid three-term HS-DY (TTHD) direction, it amalgamates the three-term HS and DY directions. Additionally, this direction exhibits similarities to the memoryless BFGS quasi-Newton method and incorporates trust region properties. We establish global convergence under both Wolfe line search conditions. The reported numerical outcomes suggest the superiority of our hybrid method over those proposed in [27–30]. Which sets our approach apart is its unique advantage of possessing favorable properties from both HS and DY directions. For more comprehensive information on CG methods, interested readers can explore [33–35].

## 2 ALGORITHM AND THEORETICAL RESULTS

In [30], Deepho has introduced a hybrid three-term TTCDDY CG method with the following search direction

$$d_k^{\text{TTCDDY}} = -r_k + \left(\frac{r_k^T r_k}{v_k} - \frac{\|r_k\|^2 r_k^T d_{k-1}}{v_k^2}\right) d_{k-1} - c_k \frac{r_k^T d_{k-1}}{v_k} r_k, \quad k \geq 1, \tag{10}$$

where

$$v_k = \max\left(\zeta \|d_{k-1}\| \|r_k\|, -d_{k-1}^T r_{k-1}, d_{k-1}^T y_{k-1}\right), \quad \zeta > 1, \quad 0 \leq c_k \leq \bar{c}_k < 1.$$

Similarly, Kumam [29] proposed a hybrid three-term TTHSLS CG algorithm wherein the search direction possesses the form

$$d_k^{\text{HTTHSLS}} = -r_k + \left(\frac{r_k^T y_{k-1}}{u_k} - \frac{\|y_{k-1}\|^2 r_k^T d_{k-1}}{u_k^2}\right) d_{k-1} + c_k \frac{r_k^T d_{k-1}}{v_k} y_{k-1}, \quad k \geq 1, \tag{11}$$

where

$$u_k = \max\left(\zeta \|d_{k-1}\| \|y_{k-1}\|, -d_{k-1}^T r_{k-1}, d_{k-1}^T y_{k-1}\right), \quad \zeta > 1, \quad 0 \leq c_k \leq \bar{c}_k < 1.$$

Both TTCDDY and HTTHSLS methods satisfy the sufficient descent conditions and global convergence is proven under some assumptions. The numerical results indicated that both hybrid methods outperform the previous methods. Motivated by the TTCDDY and HTTHSLS, we propose a new hybrid three-term CG algorithm based on the LBFGS Quasi-Newton algorithm with standardization by the spectral proposed by Wang [21]. Next, we will recall the search direction of the memoryless BFGS method by Shanno [36] and Nocedal [37], which can be written as

$$d_k^{\text{BFGS}} = -\left(I - \frac{s_{k-1}^T y_{k-1}}{s_{k-1}^T y_{k-1}} - \frac{y_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} + \frac{s_{k-1} y_{k-1}^T y_{k-1} s_{k-1}}{s_{k-1}^T y_{k-1}} + \frac{s_{k-1} s_{k-1}^T}{s_{k-1}^T y_{k-1}}\right) r_k,$$

where $s_{k-1} = x_k - x_{k-1} = \alpha_{k-1} d_{k-1}$ and $I$ is the identity matrix. By simplifying the $d_k^{\text{BFGS}}$, it can be described as

$$d_k^{\text{BFGS}} = -r_k + \left( \frac{r_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} - \frac{\|y_{k-1}\|^2 r_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2} \right) d_{k-1} + \frac{r_k^T d_{k-1}}{d_{k-1}^T y_{k-1}} (y_{k-1} - s_{k-1}), \quad k \geq 1. \tag{12}$$

By recalling the three-term HS CG method proposed by Li [27] which defined as

$$d_k^{\text{THS}} = -r_k + \left( \frac{r_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} - \frac{\|y_{k-1}\|^2 r_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2} \right) d_{k-1} + c_k \frac{r_k^T d_{k-1}}{d_{k-1}^T y_{k-1}} y_{k-1}, \tag{13}$$

By recalling the three-term DY CG method proposed by Deepho [30] which defined as

$$d_k^{\text{TDY}} = -r_k + \left( \frac{r_k^T r_k}{d_{k-1}^T y_{k-1}} - \frac{\|r_k\|^2 r_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2} \right) d_{k-1} - c_k \frac{r_k^T d_{k-1}}{d_{k-1}^T y_{k-1}} r_k. \tag{14}$$

To achieve the standardization for both parameters, we propose a modification of spectral proposed by Wang [21]. This adjustment involves replacing terms associated with $\|r_k\|^2$, $\|s_{k-1}\|^2$, and $\|y_{k-1}\|^2$, thereby allowing the selection of suitable values for the conjugate parameter and search direction.

$$z_k = \max\left\{ \min\left\{ \|r_k\|^2, \|s_{k-1}\|^2 \right\}, \|y_{k-1}\|^2 \right\}. \tag{15}$$

where

$$\omega_k = \begin{cases} y_{k-1} & \text{if } z_k = \|y_{k-1}\|^2, \\ r_k & \text{otherwise.} \end{cases} \tag{16}$$

When $\omega_k = y_{k-1}$ then $d_k^{\text{TTHD}} = d_k^{\text{THS}}$; otherwise, $d_k^{\text{TTHD}} = d_k^{\text{TDY}}$. The standardization process applied to both search directions in equations (13) and (14) using equations (15) and (16) results in a similarity to the TTHD search direction. Consequently, the standardized search direction can be explicitly defined as follows

$$d_k^{\text{TTHD}} = -r_k + \left( \frac{r_k^T \omega_k}{d_{k-1}^T y_{k-1}} - \frac{\|\omega_k\|^2 r_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2} \right) d_{k-1} + c_k \frac{r_k^T d_{k-1}}{d_{k-1}^T y_{k-1}} \omega_k, \tag{17}$$

To solve the problem of finding the univariate minimum, it becomes necessary to determine the parameter $c_k$,

$$\min_{t \in \mathbb{R}} \|(y_{k-1} - s_{k-1}) - t\omega_k\|^2. \tag{18}$$

Let $A_k = (y_{k-1} - s_{k-1}) - c\omega_k$, then

$$\begin{aligned} A_k A_k^T &= [(y_{k-1} - s_{k-1}) - c\omega_k] \left[(y_{k-1} - s_{k-1}) - c\omega_k\right]^T, \\ &= c^2 \omega_k \omega_k^T - c\left[\omega_k^T (y_{k-1} - s_{k-1}) + (y_{k-1} - s_{k-1})^T \omega_k\right] + (y_{k-1} - s_{k-1})(y_{k-1} - s_{k-1})^T, \end{aligned}$$

Let $B_k = y_{k-1} - s_{k-1}$, then

$$\begin{aligned} A_k A_k^T &= c^2 \omega_k \omega_k^T - c(\omega_k^T B_k + B_k^T \omega_k) + B_k B_k^T \\ \mathrm{tr}(A_k A_k^T) &= c^2 \|\omega_k\|^2 - c(\mathrm{tr}(\omega_k^T B_k) + \mathrm{tr}(B_k^T \omega_k)) + \|B_k\|^2 \\ &= c^2 \|\omega_k\|^2 - 2c\omega_k^T B_k + \|B_k\|^2. \end{aligned}$$

By taking the derivative of the previous expression with respect to $c_k$ and equating it to zero, we derive the following result,

$$2c\|\omega_k\|^2 - 2\omega_k^T B_k = 0.$$

This yields

$$c = \frac{\omega_k^T (y_{k-1} - s_{k-1})}{\|\omega_k\|^2}. \tag{19}$$

Therefore, we choose $c_k$ to be

$$c_k = \min\{\bar{c}, \max\{0, c\}\}, \tag{20}$$

where $0 \leq c_k \leq \bar{c} < 1$.

In accordance with the search direction stated in equations (17), we introduce a new search direction for the hybrid three-term CG method which as follows,

$$d_0 = -g_0, \quad d_k^{\mathrm{TTHD}} = -r_k + \beta_k^{\mathrm{TTHD}} d_{k-1} + \gamma_k^{\mathrm{TTHD}} \omega_k, \quad k \geq 1, \tag{21}$$

where

$$\beta_k^{\mathrm{TTHD}} = \frac{r_k^T \omega_k}{d_{k-1}^T y_{k-1}} - \frac{\|\omega_k\|^2 r_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2}, \quad \gamma_k^{\mathrm{TTHD}} = c_k \frac{r_k^T d_{k-1}}{d_{k-1}^T y_{k-1}}. \tag{22}$$

## 3 CONVERGENCE ANALYSIS

Next, we discuss the global convergence results of the TTHD method based on the following set of assumptions

**Assumption 1** *The level set $\mathcal{B} = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ is bounded, where $\mathbf{x}_0$ is starting point.*

**Assumption 2** *Suppose some neighbourhood $\mathcal{B}$ of $B$ is gradient of $f$ that is Lipschitz continuous on $\mathcal{B}$ and continuously differentiable. In which, $B > 0$ such that for all $x$,*

$$\|r(x) - r(b)\|^2 \leq L\|x - b\|, \quad b \in B.$$

From Assumptions 1 and 2 indicates that there exists a constant $T_1$, $T_2 > 0$ for all $x \in B$, in which

$$\|x\| \leq T_1, \quad \|r(x)\| \leq T_2.$$

Moreover, $\{f(x_k)\}$ is decreasing when the sequence $\{x_k\} \in B$ is decreasing. Thus, suppose that the objective function is bounded below and the Assumption 1 and 2 hold.

Next, we present the sufficient descent condition for the TTHD method.

**Lemma 3.1** *Search direction* $d_k$ *determined in* (21) *satisfies* (6) *with* $t = \left(1 - \frac{1}{4}(1+\bar{c})^2\right)$.

*Proof:* Multiplying each sides of (21) with $r_k^T$, we get

$$\begin{aligned} r_k^T d_k &= -\|r_k\|^2 + \frac{r_k^T\omega_k}{d_{k-1}^T y_{k-1}} r_k^T d_{k-1} \quad - \frac{\|\omega_k\|^2}{(d_{k-1}^T y_{k-1})^2}(r_k^T d_{k-1})^2 + c_k \frac{r_k^T\omega_k}{d_{k-1}^T y_{k-1}} r_k^T d_{k-1} \\ &= -\|r_k\|^2 + (1+c_k)\frac{r_k^T\omega_k}{d_{k-1}^T y_{k-1}} r_k^T d_{k-1} - \frac{\|\omega_k\|^2}{(d_{k-1}^T y_{k-1})^2}(r_k^T d_{k-1})^2. \end{aligned} \tag{23}$$

We obtain $a_k$ and $b_k$ by using the inequality $a_k^T b_k \leq \frac{1}{2}\left(\|a_k\|^2 + \|b_k\|^2\right)$,

$$\alpha_k = \frac{1}{\sqrt{2}}(1+c_k)r_k, \quad \beta_k = \frac{\sqrt{2}(r_k^T d_{k-1})\omega_k}{d_{k-1}^T y_{k-1}}$$

$$(1+c_k)\frac{r_k^T\omega_k}{d_{k-1}^T y_{k-1}} r_k^T d_{k-1} \leq \frac{1}{4}(1+c_k)^2\|r_k\|^2 + \frac{\|\omega_k\|^2}{(d_{k-1}^T y_{k-1})^2}(r_k^T d_{k-1})^2. \tag{24}$$

Substitute (24) into (23), we obtain

$$\begin{aligned} r_k^T d_k &\leq -\|r_k\|^2 + \frac{1}{4}(1+c_k)^2\|r_k\|^2 + \frac{\|\omega_k\|^2}{(d_{k-1}^T y_{k-1})^2}(r_k^T d_{k-1})^2 - \frac{\|\omega_k\|^2}{(d_{k-1}^T y_{k-1})^2}(r_k^T d_{k-1})^2 \\ &= -\|r_k\|^2 + \frac{1}{4}(1+c_k)^2\|r_k\|^2 \\ &\leq -\left(1 - \frac{1}{4}(1+\bar{c})^2\right)\|r_k\|^2. \end{aligned}$$

The proof is completed. □

**Remark 3.1** *The lemma 3.1 demonstrates that the TTHD always obeys the sufficient descent condition without necessitating a line search.*

Dai and Yuan [9] demonstrated that the Wolfe line search condition is satisfied by all conjugate gradient methods

**Theorem 3.1** *[9] Given the fulfillment of both Assumptions 1 and 2, and satisfied conditions* (4) *and* (5)*, if*

$$\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} = +\infty.$$

*Then*

$$\lim_{k\to\infty} \inf \|r_k\| = 0. \tag{25}$$

*Proof:* By employing a proof by contradiction, assume that equation (25) is not fulfilled. In this instance, there exists a positive scalar $\varsigma$ such that

$$\|r_k\| \geq \varsigma, \text{ for all positive value of k.} \tag{26}$$

**Lemma 3.2** *If the sequence* $\{d_k\}$ *is described in* (21)*, there exists a positive scalar* $\lambda_2$ *such that* $\|d_k\| \leq \|r_k\|\lambda_2$.

Recalling the expression for $d_k^{\text{TTHD}}$ as given in (22) for $\omega_k = y_{k-1}$ when $z_k = \|y_{k-1}\|^2$,

$$\begin{aligned}
\left\|d_k^{\text{TTHD}}\right\| &\leq \left\|-r_k + \beta_k^{\text{TTHD}} d_{k-1} + \gamma_k^{\text{TTHD}} y_{k-1}\right\| \\
&\leq \| - r_k\| + |\beta_k^{\text{TTHD}}|\|d_{k-1}\| + |\gamma_k^{\text{TTHD}}|\|y_{k-1}\| \\
&= \|r_k\| + \left|\frac{r_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} - \frac{\|y_{k-1}\|^2 r_k^T d_{k-1}}{(d_{k-1}^T y_{k-1})^2}\right| \|d_{k-1}\| + c_k \left|\frac{r_k^T d_{k-1}}{d_{k-1}^T y_{k-1}}\right| \|y_{k-1}\| \\
&\leq \|r_k\| + \left(\frac{\|r_k\|\|y_{k-1}\|}{\|r_{k-1}\|\|d_{k-1}\|} + \frac{\|y_{k-1}\|^2\|r_k\|\|d_{k-1}\|}{(\|r_{k-1}\|\|d_{k-1}\|)^2}\right)\|d_{k-1}\| + c_k \left(\frac{\|r_k\|\|d_{k-1}\|}{\|r_{k-1}\|\|d_{k-1}\|}\right)\|y_{k-1}\| \\
&\leq \|r_k\| + \left(\frac{\alpha_{k-1}\|r_k\|\|d_{k-1}\|}{f\alpha_{k-1}\|d_{k-1}\|^2} + \frac{\alpha_{k-1}^2\|r_k\|\|d_{k-1}\|^3}{\zeta^2\alpha_{k-1}^2\|d_{k-1}\|^4}\right)\|d_{k-1}\| \\
&\quad + c_k\left(\frac{\|r_k\|\|d_{k-1}\|}{\zeta\alpha_{k-1}\|d_{k-1}\|^2}\right)\alpha_{k-1}\|d_{k-1}\| \\
&= \|r_k\| + \left(\|r_k\|\frac{1}{\zeta} + \|r_k\|\frac{1}{\zeta^2}\right) + \|r_k\|c_k\left(\frac{1}{\zeta}\right) \\
&\leq \|r_k\|\left(1 + \frac{1}{\zeta} + \frac{1}{\zeta^2} + \frac{\bar{c}}{\zeta}\right).
\end{aligned}$$

In which $\lambda_2 = \|r_k\|\left(1 + \frac{1}{\zeta} + \frac{1}{\zeta^2} + \frac{\bar{c}}{\zeta}\right)$, where $\|d_k\| \leq \|r_k\|\lambda_2$.

The same proof technique is applied in an alternative scenario, where $z_k \neq |y_{k-1}|^2$ and $\omega_k = r_k$ holds. Consequently, the sequence $|d_k|$ has the upper bound produced by the TTHD method. □

Moreover, we present the well-known Zoutendijk condition [38], a pivotal factor in the global convergence analysis of the TTHD method.

**Lemma 3.3** *[38] Assuming the fulfillment of Assumptions 1 and 2, and considering the sequence* $\{x_k\}$ *generated by* (2)*, where* $d_k$ *adheres to the sufficient descent condition and* $\alpha_k$ *is determined by Wolfe line search, then*

$$\sum_{k=0}^{\infty} \frac{\left(r_k^T d_k\right)^2}{\|d_k\|^2} < +\infty. \tag{27}$$

Applied Mathematics and Computational Intelligence
Volume 13, No. 1, Feb 2024 [52 − 68]

Satisfying condition (4), under the conditions of $\alpha_k > 0$, $\eta_1 > 0$, and $0 \leq \bar{c} \leq 1$ and Lemma 3.1, we derive

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq f(x_k) + \eta_1 \alpha_k r_k^T d_k \\ &\leq f(x_k) - \eta_1 \alpha_k \left(1 - \frac{1}{4}(1+\bar{c})^2\right) \|r_k\|^2 \\ &\leq f(x_k). \end{aligned}$$

Elaborating on the above finding and considering Assumption 1, we deduce the following

$$f(x_{k+1}) \leq f(x_k) + \eta_1 \alpha_k r_k^T d_k \leq f(x_k) \leq f(x_{k-1}) \leq \ldots \leq f(x_0) < +\infty.$$

Incorporating condition (5) by adding $-r_k^T d_k$ yields

$$g(x_k + \alpha_k d_k)^T d_k - r_k^T d_k \geq \eta_2 r_k^T d_k - r_k^T d_k = -\left(1 - \eta_2\right) r_k^T d_k.$$

Using Lemma 3.1, along with condition (5) and Assumption 2, it deduces as follows

$$-(1-\eta_2) r_k^T d_k \leq (g_{k+1} - r_k)^T d_k \leq \|g_{k+1} - r_k\| \|d_k\| \leq \alpha_k L \|d_k\|^2. \tag{28}$$

Multiplying $-\eta_1 r_k^T d_k$ to the above inequality and combine it with (4), we deduce

$$\frac{\eta_1 \left(1 - \eta_2\right)}{L} \frac{\left(r_k^T d_k\right)^2}{\|d_k\|^2} \leq -\eta_1 \alpha_k r_k^T d_k \leq f(x_k) - f(x_{k+1})$$

and

$$\frac{\eta_1 \left(1 - \eta_2\right)}{L} \sum_{k=0}^{\infty} \frac{\left(r_k^T d_k\right)^2}{\|d_k\|^2} \leq \left(f(x_0) - f(x_1)\right) + \left(f(x_1) - f(x_2)\right) + \ldots \leq f(x_0) < +\infty.$$

As mentioned earlier, the sequence $f(x_k)$ is confined within specific bounds. This suggests that

$$\sum_{k=0}^{\infty} \frac{\left(r_k^T d_k\right)^2}{\|d_k\|^2} < +\infty.$$

The conjunction of inequality (26) and (6) implies that

$$r_k^T d_k \leq -\left(1 - \frac{1}{4}(1+\bar{c})^2\right) \|r_k\|^2 \leq -\left(1 - \frac{1}{4}(1+\bar{c})^2\right) \|\varsigma\|^2. \tag{29}$$

By squaring both sides and dividing equation (29) by $\|d_k\|^2$, where $\|d_k\| \neq 0$, we derive

$$\sum_{k=0}^{\infty} \frac{\left(r_k^T d_k\right)^2}{\|d_k\|^2} \geq \left(1 - \frac{1}{4}(1+\bar{c})^2\right)^2 \sum_{k=0}^{\infty} \frac{\|\varsigma\|^4}{\|d_k\|^2} = +\infty. \tag{30}$$

Since it contradicts the Zoutendijk condition (27), then it validates the theorem.

## 4 NUMERICAL EXPERIMENTS

The performance of the new TTHD CG algorithm is analysed in this section on 150 test functions taken into consideration from Moré [39], and Jamil [40], and Andrei [41]. The newly proposed method, denoted as TTHD, will undergo comparison with various other methods, such as NHS+ [27], HTT [28], TTCDDY [30], HTTHSLS [29]. The comparisons are made based on reductions in terms of the Number of Iterations (NOI) and Central Processing Unit (CPU) times with dimensions ranging from 2 to 1,000,000 as stateed in Table 1. All the comparative methods were implemented and executed using Matlab R2021B which equipped with an Intel® Core™ i5-9300H processor, 16 GB RAM, and 64-bit Windows 11 on a personal laptop.

Table 1 : List of Test Functions and their Dimensions.

| No. | Functions | Dimensions | No. | Functions | Dimensions |
|---|---|---|---|---|---|
| 1 | Extended White & Holst | 50,000 | 76 | Cube | 2 |
| 2 | Extended White & Holst | 100,000 | 77 | Cube | 50 |
| 3 | Extended White & Holst | 1,000,000 | 78 | Cube | 100 |
| 4 | Extended Rosenbrock | 50,000 | 79 | Extended Maratos | 10 |
| 5 | Extended Rosenbrock | 100,000 | 80 | Extended Maratos | 50 |
| 6 | Extended Rosenbrock | 1,000,000 | 81 | Extended Maratos | 100 |
| 7 | Extended Freudenstein and Roth | 1,000 | 82 | Generalized Tridiagonal 1 | 5 |
| 8 | Extended Freudenstein and Roth | 50,000 | 83 | Generalized Tridiagonal 1 | 10 |
| 9 | Extended Freudenstein and Roth | 100,000 | 84 | Generalized Tridiagonal 1 | 100 |
| 10 | Extended Beale | 1,000 | 85 | Trecanni | 2 |
| 11 | Extended Beale | 50,000 | 86 | Trecanni | 2 |
| 12 | Extended Beale | 100,000 | 87 | Zettl | 2 |
| 13 | Raydan 1 | 10 | 88 | Zettl | 2 |
| 14 | Raydan 1 | 50 | 89 | Shallow | 1,000 |
| 15 | Raydan 1 | 100 | 90 | Shallow | 50,000 |
| 16 | Extended Tridiagonal 1 | 10 | 91 | Shallow | 100,000 |
| 17 | Extended Tridiagonal 1 | 50 | 92 | Generalized Quartic | 100 |
| 18 | Extended Tridiagonal 1 | 10 | 93 | Generalized Quartic | 5,000 |
| 19 | Diagonal 4 | 1,000 | 94 | Generalized Quartic | 10,000 |
| 20 | Diagonal 4 | 5,000 | 95 | Quadratic QF2 | 10 |
| 21 | Diagonal 4 | 50,000 | 96 | Quadratic QF2 | 100 |
| 22 | Extended Himmelblau | 1,000 | 97 | Quadratic QF2 | 1,000 |
| 23 | Extended Himmelblau | 50,000 | 98 | Six Hump Camel | 2 |
| 24 | Extended Himmelblau | 100,000 | 99 | Six Hump Camel | 2 |
| 25 | FLETCHCR | 100 | 100 | Three Hump Camel | 2 |
| 26 | FLETCHCR | 5,000 | 101 | Three Hump Camel | 2 |
| 27 | FLETCHCR | 50,000 | 102 | Dixon and Price | 1,000 |
| 28 | Extended Powell | 100 | 103 | Dixon and Price | 10,000 |
| 29 | Extended Powell | 1,000 | 104 | Dixon and Price | 100,000 |
| 30 | NONSCOMP | 2 | 105 | POWER | 10 |
| 31 | NONSCOMP | 4 | 106 | POWER | 50 |
| 32 | NONSCOMP | 10 | 107 | POWER | 500 |

Continued on next page

Continued from previous page

| No. | Functions | Dimensions | No. | Functions | Dimensions |
|---|---|---|---|---|---|
| 33 | Extended DENSCHNB | 1,000 | 108 | Quadratic QF1 | 100 |
| 34 | Extended DENSCHNB | 50,000 | 109 | Quadratic QF1 | 1,000 |
| 35 | Extended DENSCHNB | 100,000 | 110 | Quadratic QF1 | 10,000 |
| 36 | Extended Penalty Function U52 | 5 | 111 | Generalized Tridiagonal 2 | 10 |
| 37 | Extended Penalty Function U52 | 10 | 112 | Generalized Tridiagonal 2 | 50 |
| 38 | Extended Penalty Function U52 | 50 | 113 | Generalized Tridiagonal 2 | 500 |
| 39 | Hager | 5 | 114 | Leon | 2 |
| 40 | Hager | 10 | 115 | Leon | 2 |
| 41 | Hager | 50 | 116 | Sphere | 1,000 |
| 42 | Booth | 2 | 117 | Sphere | 10,000 |
| 43 | Booth | 2 | 118 | Sphere | 100,000 |
| 44 | Sum Squares | 1,000 | 119 | Quartic | 4 |
| 45 | Sum Squares | 10,000 | 120 | Quartic | 4 |
| 46 | Sum Squares | 100,000 | 121 | Strait | 1,000 |
| 47 | Matyas | 2 | 122 | Strait | 100,000 |
| 48 | Matyas | 2 | 123 | Strait | 1,000,000 |
| 49 | Extended Quadratic Penalty QP3 | 5 | 124 | Zirilli or Aluffie-Petini's | 2 |
| 50 | Extended Quadratic Penalty QP3 | 10 | 125 | Zirilli or Aluffie-Petini's | 2 |
| 51 | Extended Quadratic Penalty QP3 | 100 | 126 | Extended Block-Diagonal BD1 | 100 |
| 52 | Extended Quadratic Penalty QP2 | 5 | 127 | Extended Block-Diagonal BD1 | 5,000 |
| 53 | Extended Quadratic Penalty QP2 | 50 | 128 | Extended Block-Diagonal BD1 | 50,000 |
| 54 | Extended Quadratic Penalty QP2 | 500 | 129 | Perturbed Quadratic | 2 |
| 55 | Extended Quadratic Penalty QP1 | 5 | 130 | Perturbed Quadratic | 2 |
| 56 | Extended Quadratic Penalty QP1 | 10 | 131 | Perturbed Quadratic | 2 |
| 57 | Extended Quadratic Penalty QP1 | 100 | 132 | Extended Hiebert | 1,000 |
| 58 | DENSCHNA | 1,000 | 133 | Extended Hiebert | 10,000 |
| 59 | DENSCHNA | 10,000 | 134 | Extended Hiebert | 100,000 |
| 60 | DENSCHNA | 100,000 | 135 | Linear Perturbed | 100 |
| 61 | DENSCHNB | 100 | 136 | Linear Perturbed | 5,000 |
| 62 | DENSCHNB | 5,000 | 137 | Linear Perturbed | 50,000 |
| 63 | DENSCHNB | 50,000 | 138 | QUARTICM | 1,000 |
| 64 | DENSCHNC | 100 | 139 | QUARTICM | 50,000 |
| 65 | DENSCHNC | 5,000 | 140 | QUARTICM | 100,000 |
| 66 | DENSCHNC | 50,000 | 141 | Diagonal 2 | 2 |
| 67 | DENSCHNF | 100 | 142 | Diagonal 2 | 5 |
| 68 | DENSCHNF | 5,000 | 143 | Diagonal 2 | 10 |
| 69 | DENSCHNF | 50,000 | 144 | Colville | 4 |
| 70 | HIMMELBG | 10 | 145 | Colville | 4 |
| 71 | HIMMELBG | 50 | 146 | Price Function 4 | 2 |
| 72 | HIMMELBG | 100 | 147 | Price Function 4 | 2 |
| 73 | HIMMELBH | 10 | 148 | DIAG-AUP1 | 10 |
| 74 | HIMMELBH | 50 | 149 | DIAG-AUP1 | 1,000 |

Continued on next page

Continued from previous page

| No. | Functions | Dimensions | No. | Functions | Dimensions |
|---|---|---|---|---|---|
| 75 | HIMMELBH | 100 | 150 | DIAG-AUP1 | 10,000 |

The numerical comparisons were conducted objectively using the standard Wolfe line search, where the parameter values for our proposed method are $\eta_1 = 0.0001$, $\eta_2 = 0.09$, and $\bar{c} = 0.3$, while the parameter values were maintained for NHS+, HTT, TTCDDY and HTTHSLS. When $\|r_k\| \leq 10^{-6}$, all methods were terminated and will fail if the optimal value is never reached or the number of iterations exceeds 10,000. For the step length, $\alpha_k$ will be selected when the search iterations of the standard Wolfe line search exceed 6. The overall numerical results for the all methods including the NOI and CPU times are provided at https://shorturl.at/dnxPS. Further evaluation and visual illustration of the results were conducted using the performance profile tool introduced by Dolan and Moré [42], as shown in Figure 1 and Figure 2, respectively. Generally, the highest curve in the performance profile indicates superiority and better efficiency of the algorithm. As demonstrated in [42], the performance profiles measure $\tau_{p,s}$, representing the time needed to solve each problem $p \in P$ by solver $s \in S$. The performance profile formula is defined as follows

$$\psi_s(\tau) = \frac{1}{n_p} size\{p \in P : \log_2 r_{p,s} \leq \tau\}.$$

where $\tau > 0$, *size* $\{p \in P : \log_2 r_{p,s} \leq \tau\}$ is the number of elements in the set $\{p \in P : \log_2 r_{p,s} \leq \tau\}$, and $r_{p,s}$ is the performance ratio formulated as $r_{p,s} = \tau_{p,s}/\min\{p \in P : \log_2 r_{p,s} \leq \tau\}$.

According to numerical reports and our plots in Figure 1 and Figure 2, the proposed TTHD method establishes a number of advantages which TTHD is effective for the 62% of the tested problems and more efficient compared to other methods in the comparison. In addition, the numerical performance of the TTHD method is comparatively stable due to the parameter choices (15), (16), and (22). According to the numerical results of the two comparisons and their respective performance profiles, all five methods have proved to be practically effective, at least for these particular sets of numerical experiments. The effectiveness of each method can be observed by referring to Figure 1 and Figure 2, where the NHS+ mmethod solves 91% of the problems, the HTTHSLS 97%, the TTCDDY 92%, the HTT 89% and the TTHD 100%. In this perspective, the TTHD method is most effective compared with other methods. Additionally, it is necessary to note that the TTHD performs robustly, particularly when confronting difficult problems.

## 5 CONCLUSION

In this article, the HS and DY CG parameters are combined to create a hybrid CG algorithm. Independent of the line search, the search direction of the algorithm is sufficiently descent and bounded. In addition, the step length was derived through standard Wolfe line searches. Under appropriate assumptions, global convergence of the algorithm was proven. On the basis of the numerical results, it is clear that the new hybrid method is more effective and robust than other methods, providing quicker and more stable convergence for the majority of the problems considered.

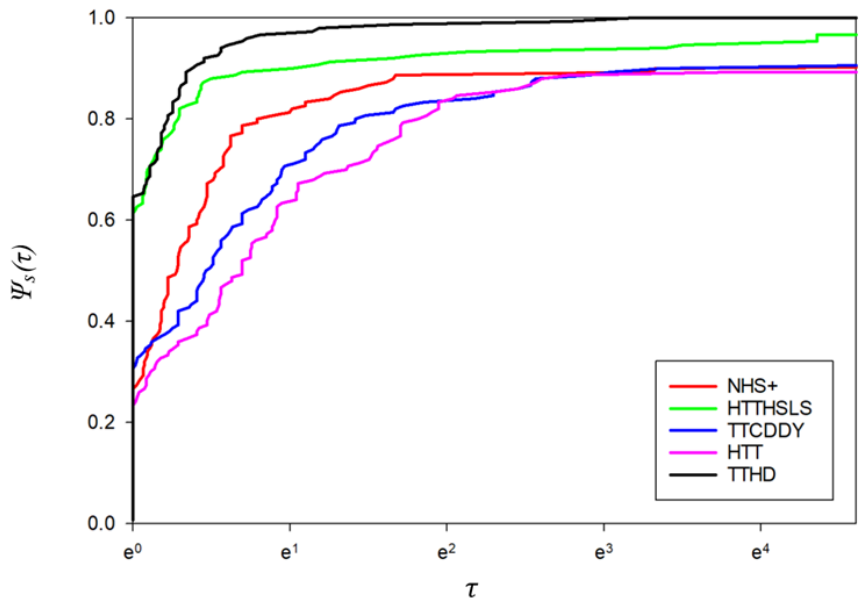


Figure 1 : Performance Profiles on NOI.

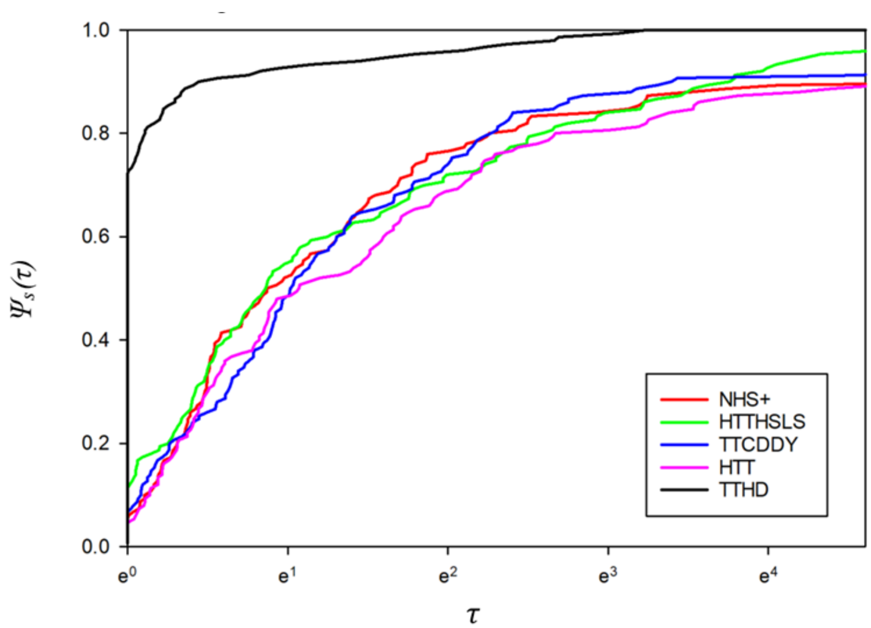


Figure 2 : Performance Profiles on CPU time.

## ACKNOWLEDGEMENT

The authors are grateful to the editor and the anonymous reviewers for their comments and suggestions which improved this paper substantially and acknowledges the generous support from

Dr. Mohamat Aidil Mohamat Johari for funding the Simposium Kebangsaan Sains Matematik ke-30 (SKSM30) conference fee under GERAN PUTRA 2021 (Grant No. 9300477). Special thanks to the Department of Mathematics and Statistics, Faculty of Science, UPM.

## REFERENCES

[1] J. E. Dennis and J. J. Moré, "A characterization of superlinear convergence and its application to quasi-newton methods," *Mathematics of Computation*, vol. 28, no. 126, pp. 549–560, 1974, https://doi.org/10.2307/2005926.

[2] ——, "Quasi–newton methods, motivation and theory," *SIAM review*, vol. 19, no. 1, pp. 46–89, 1977, https://www.jstor.org/stable/2029325.

[3] P. E. Gill and W. Murray, "Newton–type methods for unconstrained and linearly constrained optimization," *Mathematical Programming*, vol. 7, no. 1, pp. 311–350, 1974, https://doi.org/10.1007/BF01585529.

[4] W. Hager and W. Murray, "A survey of nonlinear conjugate gradient methods," *Pacific Journal of Optimization*, vol. 2, no. 1, pp. 35–58, 2006.

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952, https://api.semanticscholar.org/CorpusID:2207234.

[6] E. Polak and G. Ribiére, "Note sur la convergence de méthodes de directions conjugées," *Revue Francaise D'informatique et de Recherche Opérationnelle. Série rouge.*, vol. 3, no. 16, pp. 35–43, 1969, http://www.numdam.org/item/M2AN_1969__3_1_35_0/.

[7] B. T. Polyak, "The conjugate gradient method in extremal problems," *USSR Computational Mathematics and Mathematical Physics*, vol. 9, no. 4, pp. 94–112, 1969, https://doi.org/10.1016/0041-5553(69)90035-4.

[8] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: Theory," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991, https://doi.org/10.1007/BF00940464.

[9] Y. H. Dai and Y. X. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal of Optimization*, vol. 10, no. 1, pp. 177–182, 1999, https://doi.org/10.1137/S1052623497318992.

[10] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964, https://doi.org/10.1093/comjnl/7.2.1492.

[11] R. Fletcher, *Practical methods of optimization*, 2nd ed. John Wiley and Sons, 2013.

[12] S. Babaie-Kafaki, R. Ghanbari, and N. Mahdavi-Amiri, "Two new conjugate gradient methods based on modified secant equations," *Journal of Computational and Applied Mathematics*, vol. 234, no. 5, pp. 1374–1386, 2010, https://doi.org/10.1016/j.cam.2010.01.052.

[13] N. Andrei, “Another hybrid conjugate gradient algorithm for unconstrained optimization,” *Numerical Algorithm*, vol. 47, pp. 143–156, 2008, https://doi.org/10.1007/s11075-007-9152-9.

[14] ——, “A hybrid conjugate gradient algorithm for unconstrained optimization as convex combination of hestenes–steifel and dai–yuan,” *Studies in Informatics and Control*, vol. 47, pp. 55–70, 2008.

[15] S. Babaie-Kafaki and N. Mahdavi-Amiri, “Two modified hybrid conjugate gradient methods based on a hybrid secant equation,” *Mathematical Modelling and Analysis*, vol. 18, no. 1, pp. 32–52, 2013, https://doi.org/10.3846/13926292.2013.756832.

[16] M. J. D. Powell, “Nonconvex minimization calculations and the conjugate gradient method,” in *Numerical Analysis*, D. F. Griffiths, Ed. Berlin, Heidelberg: Springer, 1984, pp. 122–141.

[17] S. Babaie-Kafaki, M. Fatemi, and N. Mahdavi-Amiri, “Two effective hybrid conjugate gradient algorithms based on modified bfgs updates,” *Numerical Algorithm*, vol. 58, no. 3, pp. 315–331, 2011, https://doi.org/10.1007/s11075-011-9457-6.

[18] S. Babaie-Kafaki, “A hybrid conjugate gradient method based on quadratic relaxation of dai-yuan hybrid conjugate gradient parameter,” *Optimization*, vol. 62, no. 7, pp. 929–941, 2013, https://doi.org/10.1080/02331934.2011.611512.

[19] N. Andrei, “An unconstrained optimization test functions collection,” 2008. [Online]. Available: https://api.semanticscholar.org/CorpusID:63504217

[20] ——, “Hybrid conjugate gradient algorithm for unconstrained optimization,” *Journal of Optimization Theory and Applications*, vol. 141, no. 2, pp. 249–264, 2009, https://doi.org/10.1007/s10957-008-9505-0.

[21] L. Wang, M. Cao, F. Xing, and Y. Yang, “The new spectral conjugate gradient method for large-scale unconstrained optimisation,” *Journal of Inequalities and Applications*, vol. 2020, no. 1, p. 111, 2020, https://doi.org/10.1186/s13660-020-02375-z.

[22] N. Andrei, “A simple three-term conjugate gradient algorithm for unconstrained optimization,” *Journal of Computational and Applied Mathematics*, vol. 241, pp. 19–29, 2013, https://doi.org/10.1016/j.cam.2012.10.002.

[23] J. Liu and S. Li, “New hybrid conjugate gradient method for unconstrained optimization,” *Applied Mathematics and Computation*, vol. 245, pp. 36–43, 2014, https://doi.org/10.1016/j.amc.2014.07.096.

[24] X. Xu and F.-Y. Kong, “New hybrid conjugate gradient methods with the generalized wolfe line search,” *SpringerPlus*, vol. 5, no. 1, p. 881, 2016, https://doi.org/10.1186/s40064-016-2522-9.

[25] X.-L. Dong, D.-R. Han, R. Ghanbari, X.-L. Li, and Z.-F. Dai, “Some new three–term hestenes–stiefel conjugate gradient methods with affine combination,” *Optimization*, vol. 66, no. 5, pp. 759–776, 2017, https://doi.org/10.1080/02331934.2017.1295242.

[26] M. Li, "A three term polak-ribiére-polyak conjugate gradient method close to the memoryless bfgs quasi-newton method," *Journal of Industrial and Management Optimization*, vol. 16, no. 1, pp. 245–260, 2020, https://doi.org/10.3934/jimo.2018149.

[27] ——, "A modified hestense-stiefel conjugate gradient method close to the memoryless bfgs quasi-newton method," *Optimization Methods and Software*, vol. 33, no. 2, pp. 336–353, 2018, https://doi.org/10.1080/10556788.2017.1325885.

[28] A. B. Abubakar, P. Kumam, M. Malik, P. Chaipunya, and A. Ibrahim, "A hybrid fr-dy conjugate gradient algorithm for unconstrained optimization with application in portfolio selection," *AIMS Mathematics*, vol. 66, no. 6, pp. 6506–6527, 2021, https://doi.org/10.3934/math.2021383.

[29] P. Kumam, A. B. Abubakar, M. Malik, A. H. Ibrahim, N. Pakkaranang, and B. Panyanak, "A hybrid hs-ls conjugate gradient algorithm for unconstrained optimization with applications in motion control and image recovery," *Journal of Computational and Applied Mathematics*, vol. 433, p. 115304, 2023, https://doi.org/10.1016/j.cam.2023.115304.

[30] J. Deepho, A. B. Abubakar, M. Malik, and I. K. Argyros, "Solving unconstrained optimization problems via hybrid cd-dy conjugate gradient methods with applications," *Journal of Computational and Applied Mathematics*, vol. 405, p. 113823, 2022, https://doi.org/10.1016/j.cam.2021.113823.

[31] M. Malik, I. M. Sulaiman, A. B. Abubakar, G. Ardaneswari, and Sukono, "A new family of hybrid three-term conjugate gradient method for unconstrained optimization with application to image restoration and portfolio selection," *AIMS Mathematics*, vol. 8, no. 1, pp. 1–28, 2023, https://doi.org/10.3934/math.2023001.

[32] A. B. Abubakar, P. Kumam, M. Malik, and A. H. Ibrahim, "A hybrid conjugate gradient based approach for solving unconstrained optimization and motion control problems," *Mathematics and Computers in Simulation*, vol. 201, pp. 640–657, 2022, https://doi.org/10.1016/j.matcom.2021.05.038.

[33] D.-H. Li and M. Fukushima, "A modified bfgs method and its global convergence in nonconvex minimization," *Journal of Computational and Applied Mathematics*, vol. 129, pp. 15–35, 2001, https://doi.org/10.1016/S0377-0427(00)00540-9.

[34] Z. Khoshgam and A. Ashrafi, "A new hybrid conjugate gradient method for large-scale unconstrained optimization problem with non-convex objective function," *Computational and Applied Mathematics*, vol. 38, no. 4, pp. 186–35, 2019, https://doi.org/10.1007/s40314-019-0973-7.

[35] X. Li, J. Shi, X. Dong, and J. Yu, "A new conjugate gradient method based on quasi-newton equation for unconstrained optimization," *Journal of Computational and Applied Mathematics*, vol. 350, pp. 372–379, 2019, https://doi.org/10.1016/j.cam.2018.10.035.

[36] D. F. Shannoi, "Conjugate gradient methods with inexact searches," *Mathematics of Opera-*

*tions Research*, vol. 3, no. 3, pp. 244–256, 1978.

[37] J. Nocedal, "Updating quasi-newton matrices with limited storage," *Mathematics of Computation*, vol. 35, no. 151, pp. 773–782, 1980, https://doi.org/10.1090/S0025-5718-1980-0572855-7.

[38] G. Zoutendijk, "Nonlinear programming computational methods," in *Integer and Nonlinear Programming*, J. Abadie, Ed. Amsterdam: North–Holland, 1970, pp. 37–86.

[39] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "Testing unconstrained optimization software," *ACM Transactions on Mathematical Software*, vol. 7, no. 1, pp. 17–41, 1981, https://doi.org/10.1145/355934.355936.

[40] M. Jamil and X.-S. Yang, "A literature survey of benchmark functions for global optimization problems," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 4, pp. 150–194, 2013, https://doi.org/10.1504/IJMMNO.2013.055204.

[41] N. Andrei, *Nonlinear Conjugate Gradient Methods for Unconstrained Optimization*, 1st ed. Springer Cham, 2020.

[42] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002, https://doi.org/10.1007/s101070100263.

[43] W. Sun and Y.-X. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, 1st ed., ser. Springer Optimization and Its Applications. Springer New York, NY, 2006.

[44] J. Liu, Y. Feng, and L. Zou, "Some three-term conjugate gradient methods with the inexact line search condition," *Calcolo*, vol. 55, no. 2, p. 16, 2018, https://doi.org/10.1007/s10092-018-0258-3.

[45] S. Djordjevic, "New hybrid conjugate gradient method as a convex combination of ls and fr methods," *Acta Mathematica Scientia*, vol. 39, pp. 214–228, 2017, https://doi.org/10.1007/s10473-019-0117-6.

[46] Y. H. Dai, "New properties of nonlinear conjugate gradient method," *Numerische Mathematik*, vol. 89, no. 1, pp. 83–98, 2001, https://doi.org/10.1007/PL00005464.

[47] X. Yang, Z. Luo, and X. Dai, "A global convergence of ls-cd hybrid conjugate gradient method," *Advances in Numerical Analysis*, vol. 2013, no. 1, 1980, https://doi.org/10.1155/2013/517452.

[48] N. Salihu, M. R. Odekunle, M. Y. Waziri, and A. S. Halilu, "A new hybrid conjugate gradient method based on secant equation for solving large scale unconstrained optimization problems," *Optimization*, vol. 12, no. 1, pp. 33–44, 2020.