

A Dai-Liao Hybrid PRP and DY Schemes for Unconstrained Optimization

Huzaifa B. Aliyu^a, Modu Barma^b, Salihu Nasiru^a and Salihu Suraj^c

^aDepartment of Mathematics, Modibbo Adama University, Yola, Adamawa State, Nigeria.

^bDepartment of Operations Research, Modibbo Adama University, Yola, Adamawa State, Nigeria.

^cDepartment of Computer Science, Gombe State University, Gombe State, Nigeria

ARTICLE INFO

Article history:

Received 17 December 2023

Received in revised form 01 March 2024

Accepted 06 March 2024

Keywords:

Conjugate gradient method, Descent property, Dai-Liao conjugacy condition, Global convergence, Numerical methods

MSC 2020 Subject classification:

65K05, 90C30

ABSTRACT

This article presents a new conjugate gradient (CG) method that requires first-order derivatives but overcomes the slow convergence issue associated with the steepest descent method and does not require the computation of second-order derivatives, as needed in the Newton method. The CG update parameter is suggested from the extended conjugacy condition as a convex combination of Polak, Ribière, and Polyak (PRP) and Dai and Yuan (DY) algorithms by employing the optimal choice of the modulating parameter 't'. Numerical computations show that the algorithm is robust and efficient based on the number of iterations and CPU time. The scheme converges globally under Wolfe line search and adopts an inexact line search to obtain the step-size that generates a descent property, without requiring exact computation of the step size.

1. Introduction

Problems arise practically in almost every branch of science and in other disciplines. Some of these problems have no analytical solution or are too difficult to solve, so the development of numerical methods to obtain approximate solutions became necessary (Salihu *et al.*, 2020). Several numerical methods arise for solving such problems, including the steepest descent method, Newton method, quasi-Newton method, and conjugate gradient method, among others. Gradient methods are more efficient when the function to be minimized is continuous in its first derivative (Salihu *et al.* 2021). The Conjugate Gradient (CG) method was originally proposed for solving a linear system. Since solving a linear system is equivalent to the minimization of a positive definite quadratic function, the method was later modified by Hestenes-Stiefel (1952) to solve an unconstrained optimization problem. The CG algorithm is characterized by low memory requirements and global convergence properties. A typical unconstrained optimization problem has the following form:

$$\min f(x), x \in \mathbb{R}^n, \quad (1)$$

where $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth nonlinear function and the iterative procedure is calculated by

$$x_{k+1} = x_k + \alpha_k d_k, \quad (2)$$

in which $\alpha_k > 0$ is a step length obtained by a suitable exact or inexact line search. The CG scheme designed by exact line search is computationally expensive. Consequently, the step length is typically calculated by an inexact line search, such as

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (3)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k \quad (4)$$

or using the strong Wolfe condition, which contain of (3) and

$$|g(x_k + \alpha_k d_k)^T d_k| \leq -\sigma g_k^T d_k, \quad (5)$$

where $0 < \delta < \sigma < 1$ and d_k is the search direction given by

$$d_{k+1} = -g_{k+1} + \beta_k s_k, \quad (6)$$

* Corresponding author. Tel.: +2348069779085

E-mail address: hababando@mau.edu.ng (H. B. Aliyu).

<https://doi.org/10.62054/ijdm/0101.05>

and β_k is a scalar called CG (update) parameter, and $s_k = x_{k+1} - x_k$.

To obtain new β_k formulas, numerous CG parameters were suggested, corresponding to some excellent CG algorithms in (Dai and Yua (1999); Fletcher and Reeves (1964); Fletcher (1987); Lotfiand Hosseini (2019); Polak (1967); Salihu *et al.*, (2021); Salihu *et al.*, (2023a); Salihu *et al.*, (2023b), Salihu *et al.*, (2023c)). However, these schemes were characterized by certain strengths and weaknesses from both theoretical and numerical analysis perspectives. In theory, numerical experiments show that CG schemes proposed by Hestenes and Stiefel (HS) (1952), Polak, Ribière, and Polyak (PRP) (1969), Polak (1967) and Liu and Storey (LS) (1991) have better practical performances but may not always be convergent.

On the other hand, the CG schemes proposed by Fletcher and Reeves (FR) (1964), Dai and Yuan (DY) (1999), and Conjugate Descent (CD) (1987) are characterized by strong global convergence properties, but their numerical performance is marked by jamming Babaie-Kafaki and Ghanbari (2014c), i.e., taking short steps without meaningful progress toward the minimizer. The first global convergent property of the FR method with an exact line search was proved by Salihu *et al.* (2023c), where in Al-Baali (1985), this result was extended to an inexact line search and showed that the method generates a sufficient descent direction under the strong Wolfe conditions. However, the HS method and PRP schemes possesses an automatic restart feature which addresses jamming problem, making them numerically efficient (Babaie-Kafaki, 2013).

The idea of combining different algorithms to avoid jamming and improve theoretical analysis has attracted more attention. To create new CG hybrid methods, (see: Djordjevic (2016); Djordjevic (2017); Djordjevic (2018)), new β_k parameters were proposed as a convex combination of the hybrid parameter θ_k , satisfying certain conjugacy conditions. This was achieved using the strong Wolfe line search conditions. The methods have the following β_k formulas, respectively:

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{PRP} + \theta_k \beta_k^{FR}. \quad (7)$$

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{LS} + \theta_k \beta_k^{CD}. \quad (8)$$

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{HS} + \theta_k \beta_k^{FR}. \quad (9)$$

On the other hand, the schemes derived in (Al-Namat and Al-Naemi (2020); Djordjevic (2019); Salihu *et al.* (2021)) are sufficiently descent and satisfy certain secant equations. The algorithms are characterized by the following CG formulas:

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{LS} + \theta_k \beta_k^{FR}. \quad (10)$$

$$\beta_k^{CLCS} = (1 - \theta_k) \beta_k^{LS} + \theta_k \beta_k^{CD}. \quad (11)$$

$$\beta_k^{FG} = (1 - \theta_k) \beta_k^{MMWU} + \theta_k \beta_k^{RMAR}. \quad (12)$$

The numerical evaluations show that these algorithms perform better than some known approaches and converge globally under Wolfe conditions. Therefore, this article aims to modify classical PRP and DY CG methods by employing the optimal choice of the parameter 't' to solve equation (1) using the following CG parameters.

Let $\|\cdot\|$ denotes Euclidean norm and since $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$, then

$$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}, \quad (13)$$

and

$$\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{y_k^T s_k}. \quad (14)$$

This paper is organized as follows: Section 2 presents the literature review, Section 3 introduces the proposed method and convergence results, Section 4 reports some numerical results, and finally, Section 5 concludes the paper.

2. Literature Review

Esmaili *et al.* (2018) employed a new CG method to solve compressive sensing problems, which play an important role in medical and astronomical imaging, file restoration, image and video coding, and other applications. Xue *et al.* (2018) suggested the DY CG method for solving large-scale unconstrained optimization problems,

possessing a spectral CG parameter where the search direction generated at each iteration is descent independent of any line search. The global convergence of the method is also established using strong Wolfe conditions. Additionally, Liu and Du (2019) proposed a CG method by transforming the M-tensor system into a general unconstrained minimization problem and solving a type of non-smooth optimization problem. The numerical experiment shows the efficiency of the suggested method, with low storage requirements and nice theoretical properties. Similarly, Guo and Wan (2019) developed a CG algorithm to solve an engineering problem originating from compressed sensing of sparse signals. Numerical results and preliminary applications in recovering sparse signals indicate that the established algorithm outperforms comparable algorithms available in the literature, especially for solving large-scale problems (Perry, 1978).

Recently, considerable efforts have been made to extend CG methods to solve monotone nonlinear equations. Abubakar *et al.* (2019) presented a modification of the FR CG method for constrained monotone nonlinear equations, which possesses a sufficient descent property, and its global convergence was proven. Numerical experiments demonstrate the efficiency of the proposed method in solving benchmark test problems, as well as its application in signal and image recovery problems arising from compressive sensing. Ibrahim *et al.* (2020) utilized the HLSFR Algorithm of Djordjevic (2019) in the restoration of one-dimensional sparse signals and image restoration using mean squared error (MSE). The performance of HLSFR outperformed and proved to be more efficient in decoding sparse signals in compressive sensing, requiring fewer iterations, less computing time, and achieving a lower MSE through repeated experiments on 10 different noise samples. The HLSFR algorithm is similar to HHSFR in Djordjevic (2018).

Due to the simpler structure and low memory requirements of Dai-Liao conjugate gradient methods, Salihu *et al.* (2023b) combined the Dai-Liao conjugacy condition with a modified symmetric Perry matrix to propose a class of three-term Dai-Liao conjugate gradient algorithms. The method possesses both the Dai-Liao conjugacy condition and a sufficient descent condition. Meanwhile, global convergence is established under Wolfe line search for general objective functions. Numerical experiments show that the proposed method is promising. Recently, Esmaeili *et al.* (2018) suggested a kind of important tensor optimization problems with higher-order nonlinear equations, which are widely used in engineering and economics.

The standard conjugacy condition is important in designing certain CG algorithms. However, methods relying on exact computation are computationally expensive, particularly for large-scale problems. Therefore, formulating CG methods with a generalized form for the selection of the step size is often preferred for designing efficient CG algorithms Salihu *et al.* (2023c). The standard conjugacy condition is given by:

$$d_{k+1}^T y_k = 0. \quad (15)$$

Perry (1978) extended the result in equation (15) by manipulating the secant condition of a quasi-Newton scheme and quasi-Newton search direction:

$$d_{k+1}^T y_k = -g_{k+1}^T s_k. \quad (16)$$

This implies that (16) holds for exact line search, but practical numerical computations usually adopt inexact line search. To address this, Dai and Liao (2001) replaces (16) with a condition called Dai-Liao extended conjugacy condition:

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k, \quad t \geq 0. \quad (17)$$

Based on the modified BFGS method proposed in Li and Fukushima (2001) and Lotfi and Hosseini (2019), a new value of the parameter t is introduced in the Dai-Liao CG scheme. The global convergence property of the proposed method is established, and numerical results illustrate the computational efficiency of the new method. The structures of CG procedures require computation and storage of the Hessian matrix. However, practical numerical computations adopt inexact line searches instead of exact line searches to obtain the step-size. To address these shortcomings, this article presents a hybrid method based on the Dai-Liao conjugacy condition. If the modulating parameter t equals 0, it reduces to a method that uses the secant equation. The numerical performance of the Dai-Liao CG method depends on the parameter t , and the optimal choice of this parameter remains a subject of consideration (Babaie-Kafaki and Ghanbari, 2017). Several optimal choices for the parameter t were proposed in (Andrei (2017); Babaie-Kafaki and Ghanbari (2014a); Babaie-Kafaki and Ghanbari (2014b); Babaie-Kafaki and Ghanbari (2015); Babaie-Kafaki (2015); Babaie-Kafaki and Ghanbari (2017); Salihu *et al.* (2021)). Motivated by

the above, we propose a hybrid parameter by employing the choice of the parameter 't' using (17) to access and combine the strength of the CG update parameters.

3. Dai-Liao Convex Combination

In this section, we combine the CG parameters proposed in Polak (1967) with Dai and Yuan conjugate descent based on the Dai-Liao conjugacy condition as a convex combination as follows:

$$\beta_k^{ECCPD} = (1 - \theta_k) \beta_k^{PRP} + \theta_k \beta_k^{DY}. \quad (18)$$

Using equations (13) and (14), we get

$$\beta_k^{ECCPD} = (1 - \theta_k) \left(\frac{g_{k+1}^T y_k}{\|g_k\|^2} \right) + \theta_k \left(\frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \quad (19)$$

Taking the inner product with the vector y_k^T and using relation (6), we obtain

$$d_{k+1} = -g_{k+1} + \left((1 - \theta_k) \left(\frac{g_{k+1}^T y_k}{\|g_k\|^2} \right) + \theta_k \left(\frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \right) s_k, \quad (20)$$

$$d_{k+1}^T y_k = -g_{k+1}^T y_k + \left((1 - \theta_k) \left(\frac{g_{k+1}^T y_k}{\|g_k\|^2} \right) + \theta_k \left(\frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \right) s_k^T y_k. \quad (21)$$

Equating equations (15) and (21) lead to the hybridization parameter:

$$\theta_k = \frac{(y_k^T g_{k+1})(y_k^T s_k - \|g_k\|^2)}{(y_k^T g_{k+1})(y_k^T s_k) - \|g_{k+1}\|^2 \|g_k\|^2}, \quad (22)$$

which satisfies the pure conjugacy condition. However, for large-scale problems, choices for the update parameter that do not require the evaluation of the Hessian matrix are often required. Therefore, after some algebra, we propose another hybridization parameter using (17) on (21):

$$\theta_k = \frac{t (g_{k+1}^T s_k) \|g_k\|^2 + (y_k^T g_{k+1})(y_k^T s_k - \|g_k\|^2)}{(y_k^T g_{k+1})(y_k^T s_k) - \|g_{k+1}\|^2 \|g_k\|^2}, \quad (23)$$

where θ_k is the hybridization scalar parameter satisfying $\theta_k \in [0, 1]$. It is clear that if $\theta_k \leq 0$, set $\theta_k = 0$, then $\beta_k^{ECCPD} = \beta_k^{PRP}$ and if $\theta_k \geq 0$, set $\theta_k = 1$, then $\beta_k^{ECCPD} = \beta_k^{DY}$. On the other hand, if $0 < \theta_k < 1$, then β_k^{ECCPD} is a proper convex combination of β_k^{PRP} and β_k^{DY} . Therefore, the optimal choice of the proposed method depends on the first term of the numerator of (23):

$$t (g_{k+1}^T s_k) \|g_k\|^2, \quad (24)$$

where t is the modulating parameter, and we wish to employ simple choice of the parameter.

3.1 Dai-Liao Algorithm

Step 1. Select $x_0 \in \mathbb{R}^n$, $\varepsilon > 0$ and parameter $0 < \delta < \sigma < 1$.

Compute $f(x_0)$ and g_0 .

Step 2. If $\|g_k\| \leq \varepsilon$, then stop.

Step 3. Compute $\alpha_k > 0$ satisfying Wolfe conditions (3) and (5).

Step 4. If $(y_k^T g_{k+1})(y_k^T s_k) - \|g_{k+1}\|^2 \|g_k\|^2 = 0$, then set $\theta_k = 0$ else find θ_k from (23).

Step 5. If $0 < \theta_k < 1$, then compute β_k^{ECCPD} by (19) and set $t = 0.5$ from (24).

Step 6. If restart criterion of Powell

$$|g_{k+1}^T g_k| > c \|g_{k+1}\|^2, \quad (25)$$

is satisfied, then $d_{k+1} = -g_{k+1}$; else, $d_{k+1} = -g_{k+1} + \beta_k d_k$. Set $k = k + 1$ and go to step 1.

3.2 Convergence Analysis

In this section, for an algorithm to be considered convergent, it must possess both a sufficient descent condition and global convergence properties.

3.2.1. Sufficient Descent Condition

Definition: Search direction satisfies descent directions (or equivalently, satisfy the decent condition) if an only if

$$d_k^T g_k < 0, \quad (26)$$

and also satisfies sufficient descent condition if and only if

$$d_k^T g_k \leq -c \|g_k\|^2, \forall k \geq 0, \quad (27)$$

where c is positive constant.

Furthermore, we need the following basic assumptions in convergence analysis of CG algorithms.

Assumption (i). The level set $S = \{x \in \mathbb{R} : f(x) \leq f(x_0)\}$, is bounded from below. That is, there exist a positive constant B such that

$$\|x\| \leq B, \forall x \in S. \quad (28)$$

Assumption (ii). In a neighborhood N of S , the objective function f is continuously differentiable and its gradient $g(x)$ is Lipschitz continuous on N that is, there exist a constant $L > 0$ such that

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\|, \text{ for all } x, y \in N. \quad (29)$$

Under Assumptions (i) and (ii) on f , there exist a constant Γ such that

$$\|g(x)\| \leq \Gamma, \text{ for all } x \in S. \quad (30)$$

In further consideration of relation (20), we get

$$d_{k+1} = -(\theta_k g_{k+1} + (1 - \theta_k) g_{k+1}) + \beta_k^{ECCPD} s_k,$$

$$d_{k+1} = -(\theta_k g_{k+1} + (1 - \theta_k) g_{k+1}) + (1 - \theta_k) \beta_k^{PRP} s_k + \theta_k \beta_k^{DY} s_k.$$

The last relation yields

$$d_{k+1} = \theta_k (-g_{k+1} + \beta_k^{DY} s_k) + (1 - \theta_k) (-g_{k+1} + \beta_k^{PRP} s_k).$$

$$d_{k+1} = \theta_k d_{k+1}^{DY} + (1 - \theta_k) d_{k+1}^{PRP}. \quad (31)$$

Theorem 3.1. Assume that (28) and (29) hold. Consider a CG scheme generated by (19) and (20), where α_k is obtained from (3)-(5) with $\sigma < \frac{1}{2}$. If $\|s_k\|$ tends to zero, and there exist some nonnegative constants τ_1, τ_2 such that

$$\|g_k\|^2 \geq \tau_1 \|s_k\|^2, \quad (32)$$

$$\|g_{k+1}\|^2 \leq \tau_2 \|s_k\|. \quad (33)$$

Then d_{k+1} satisfies sufficient descent condition for all k .

Proof: If $k = 0$, it holds that $d_0 = -g_0$, so $g_0^T d_0 = -\|g_0\|^2$, then it can be concluded that (27) holds for $k = 0$. Next is to show that it holds for $k > 0$.

Multiply (31) by vector g_{k+1}^T , we get

$$g_{k+1}^T d_{k+1} = \theta_k g_{k+1}^T d_{k+1}^{DY} + (1 - \theta_k) g_{k+1}^T d_{k+1}^{PRP}. \quad (34)$$

Firstly, suppose that $\theta_k = 0$, then $\beta_k = \beta_k^{PRP}$, then

$$g_{k+1}^T d_{k+1} = g_{k+1}^T d_{k+1}^{PRP}. \quad (35)$$

It follows from (13) and (35) with triangular inequality that

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2 \|y_k\| \|s_k\|}{\|g_k\|^2}. \quad (36)$$

From Lipschitz condition $\|y_k\| \leq L\|s_k\|$, we have

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2 L \|s_k\|^2}{\|g_k\|^2}. \quad (37)$$

So that using (32) and (33), we get

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \frac{L \tau_2 \|s_k\|^3}{\tau_1 \|s_k\|^2},$$

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \frac{L \tau_2}{\tau_1} \|s_k\|. \quad (38)$$

Since $\|s_k\| \rightarrow 0$, the last summand in (38) tends to zero, if not tends to zero, let there exist a number $0 < \delta \ll 1$, such that

$$\frac{L \tau_2}{\tau_1} \|s_k\| \ll \delta \|g_{k+1}\|^2. \quad (39)$$

Now (38) becomes

$$\begin{aligned} d_{k+1}^T g_{k+1} &\leq -\|g_{k+1}\|^2 + \delta \|g_{k+1}\|^2, \\ d_{k+1}^T g_{k+1} &\leq -(1 - \delta) \|g_{k+1}\|^2. \end{aligned} \quad (40)$$

Secondly, suppose that $\theta_k = 1$, then $\beta_k = \beta_k^{DY}$, it follows from (14) and (34) with triangular inequality that

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{|y_k^T s_k|} |g_{k+1}^T s_k|.$$

But (27) holds for DY method in the presence of (3) and (5) is initially mentioned in Dai and Yuan (1999). This suggests that (14) is well defined because of line search condition (4) and implies that $y_k^T s_k > 0$. If line searches are exact, then the DY formula behave the same as the FR formula of Fletcher and Reeves (1964). Therefore, the sufficient descent condition for FR method was earlier mentioned in Hager and Zhang (2006) and later in (Djordjevic (2018) and Djordjevic (2019)).

So, let there exists a constant $c_2 > 0$, such that

$$d_{k+1}^T g_{k+1} \leq -c_2 \|g_{k+1}\|^2 \quad (41)$$

Now, let $0 < \theta_k < 1$, then $0 < a_1 < \theta_k < a_2 < 1$. Therefore, from (14) and (34) we can write

$$g_{k+1}^T d_{k+1}^{DHPD} \leq a_1 g_{k+1}^T d_{k+1}^{DY} + (1 - a_1) g_{k+1}^T d_{k+1}^{HS}.$$

Denote $c = a_1 c_2 + (1 - a_2) c_1$. Then finally, we get

$$d_{k+1}^T g_{k+1} \leq -c \|g_{k+1}\|^2. \quad (42)$$

Andrei (2008) showed that the PRP method is globally convergent for general functions, if the $\|S_k\| \ll \|S_{k-1}\|$ tends to zero, a condition sufficient for theoretical result. Therefore, the norm of the gradient can satisfy (32) and (33). The condition that was initially used in Andrei (2009) and later Djordjevic (2016).

3.3. Convergence Analysis

In this section, we are going to apply the following theorems to illustrate the global convergence of ECCPD method.

Lemma 3.1. Let Assumptions (i) and (ii) hold. Consider CG methods in (2) and (6), where d_k is a descent direction and α_k satisfies strong Wolfe condition, then either

$$\lim_{k \rightarrow \infty} \inf \|g_k\| = 0. \quad (43)$$

or

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \quad (44)$$

For any conjugate gradient method with strong Wolfe line search, the convergence holds. But, for general function, only weak form of the Zoutendijk condition is needed (Dai and Liao, 2001).

Theorem 3.2. Consider the iterative method, defined by method in (2) and (6). Let d_{k+1} be a descent direction, then either $g_k = 0$, for some k , or

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (45)$$

Let the function f be a uniformly convex function, i.e. there exists a constant $\mu \geq 0$ such that for all $x, y \in R^n$

$$(\nabla f(x) - \nabla f(y))^T(x - y) \geq \mu \|x - y\|^2,$$

then $\lim_{k \rightarrow \infty} g_k = 0$.

The remaining proof is using contradiction, that theorem (3.1) is not true.

Proof: Let $g_k \neq 0$, for all k . Then we have to prove (45). Suppose on the contrary, there exist a constant $r > 0$, such that

$$\|g_k\| \geq r. \quad (46)$$

It follows from the convexity assumption that

$$y_k^T s_k \geq \mu \|s_k\|^2. \quad (47)$$

Because the descent condition holds. Since $d_{k+1} \neq 0$, it is sufficient to prove that d_{k+1} is bounded above so from relation (31), we have

$$\|d_{k+1}\| \leq \|d_{k+1}^{DY}\| + \|d_{k+1}^{PRP}\|. \quad (48)$$

Next, it follows from (33) and (47) that

$$\|d_{k+1}^{DY}\| \leq \|g_k\| + |\beta_k^{DY}| \|s_k\|.$$

$$|\beta_k^{DY}| \leq \frac{\|g_{k+1}\|^2 \|s_k\|}{\mu \|s_k\|^2} \leq \frac{\tau_2}{\mu}.$$

Now, from (30) and the above inequalities, we have

$$\|d_{k+1}^{DY}\| \leq \Gamma + \frac{\tau_2}{\mu}. \quad (49)$$

Also, from (30) and (46), we get

$$|\beta_k^{PRP}| = \frac{g_{k+1}^T y_k}{\|g_k\|^2} \leq \frac{\|g_{k+1}\| \|y_k\|}{r^2} \leq \frac{\Gamma L \|s_k\|}{r^2} \leq \frac{\Gamma L D}{r^2}$$

$$\|d_{k+1}^{PRP}\| \leq \|g_k\| + |\beta_k^{PRP}| \|s_k\|.$$

$$\|d_{k+1}^{PRP}\| \leq \Gamma + \frac{\Gamma L D}{r^2}. \quad (50)$$

Finally, using (49) and (50) implies

$$\|d_{k+1}\| \leq \Gamma + \frac{\Gamma^2}{L} + \Gamma + \frac{\Gamma L D}{r^2} \leq 2\Gamma + \frac{\Gamma^2}{L} + \frac{\Gamma L D}{r^2}. \quad (51)$$

Therefore, from

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geq 2\Gamma + \frac{\Gamma^2}{L} + \frac{\Gamma LD}{r^2}.$$

Applying Lemma 3.1, we conclude that

$$\lim_{k \rightarrow \infty} \inf \|g_k\| = 0.$$

This is a contradiction; hence the proof is complete.

4. Numerical Results and Discussion

In this section, we present computational performance of ECCPD method and compare with hybrid methods of (Andrei (2008); Andrei (2009); Djordjevic (2016); Djordjevic (2018)). To implement CG parameters, the codes were written in MATLAB (R2018a) version. The test problems were obtained from (Andrei (2008); Gould *et al.* (2003)) unconstrained optimization problems as shown in Table 1. The stopping criterion is set to $\|g_k\|_{\infty} \leq 10^{-5}$. Numerical results were compared based on performance profile of Dolan and More (2002).

Benchmark results are generated by running a solver on a set of problems and recording information of interest such as the number of iterations and the computing time. A solver has higher efficiency when its value of $P_s(t)$ is higher. The $P_s(t)$ from the performance profile is the fraction of problem with a high ratio performance t . In a set of problem P and a set of optimization solver S , a performance comparison of problem $p \in P$ by a particular algorithm $s \in S$ is measured. Let, $t_{p,s}$ be the number of iterations or CPU time required when solving a problem $p \in P$ with solver $s \in S$. The performance ratio is defined by $r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s}:s \in S\}}$. From this expression, it is assumed that $r_{p,s} \in [1, r_M]$, where $r_M \geq r_{p,s}$ and $r_{p,s} = r_M$ only when problem P is not solved by solver. Then, graphically, a graph of $P_s(t)$ versus $t \in [1, r_M]$ is plotted. In a graph of performance profile, the smallest performance ratio is 1 and it will be located at the most left of t -axis hence, the top curve represents the most efficient method. In particular, if the set of problems P is suitably large and representative of problems that are likely to occur in applications, then solver with large probability $P_s(t)$ are to be preferred. So, if we allow $P_s(t) = P(\tau)$ and $S = \tau$, then the numerical results were compared graphically from Figures 1-2.

Figures 1-2 show the performance of the hybrid coefficients based on number of iteration and central processing time per unit respectively. The top left curve indicated fraction or percentage of how fast the coefficient converge while the top right determines the fraction or percentage how many test functions can be tested on given coefficient. The both figures, clearly indicated that the proposed hybrid is efficient and outperformed the other CG coefficients.

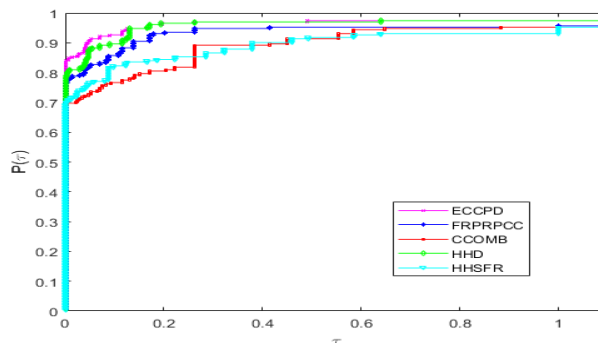


Figure 1. Number Iterations of ECCPD, FRPRPCC, CCOMB, HHD and HHSFR schemes

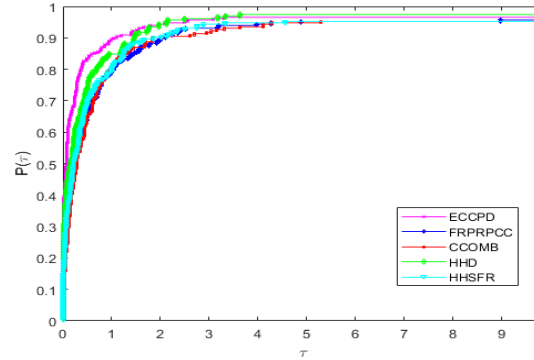


Figure 2. CPU Time of ECCPD, FRPRPCC, CCOMB, HHD and HHSFR schemes

5. Conclusion

The standard conjugacy condition is important in designing certain CG algorithms. However, these methods rely on exact computation of the step size which are computationally expensive, particularly for large-scale problems. Therefore, formulating CG methods with an extended conjugacy condition is often preferred for designing efficient CG algorithms. In this paper, we have presented a hybrid CG algorithm in which the CG parameter is computed as a convex combination of β_k^{RRP} and β_k^{DY} from extended conjugacy condition by using optimal choice of the modulating parameter t . Numerical computation adopts inexact line search which is compared with some known CG coefficients proposed in literature. Numerical results show that new coefficient outperform the other schemes and the algorithm converge globally using strong Wolfe condition.

Table 1. List of Test Functions

S/No	Functions
1	Extended White and Holst
2	Extended Rosenbrock
3	Extended Freudenstein and Roth
4	Extended Beale
5	Raydan 1
6	Extended Tridiagonal 1
7	Diagonal 4
8	Extended Himmelblau
9	Extended Powel 1
10	Fletcher Function (Cute)
11	Extended Powel
12	Nonscomp Function (Cute)
13	Extended Denschnb Function (Cute)
14	Extended Quadratic Penalty Qp1
15	Hager
16	Extended Maratos
17	Shallo
18	Generalized Quartic
19	Quadratic Qf2
20	Generalized Tridiagonal 1
21	Generalized Tridiagonal 2
22	Power
23	Quadratic Qf1
24	Extended Quadratic Penelty
25	Extended Penalty

References

- Abubakar, A. B., Kumam, P., Mohammad, H., Awwal., A. M. and Sitthithakerngkiet, K.(2019). A modified Fletcher-Reeves conjugate gradient method for monotone nonlinear equations with some applications, *MPDI, mathematics*, 7, 745 doi:10.3390/math7080745.
- Al-Baali, M. (1985). Descent property and global convergence of the Fletcher-Reeves method with inexact line search. *IMA Journal of Numerical Analysis*, 5, 121-124.
- Al-Namat, F.N. and Al-Naemi, G.M. (2020). Global convergence property with inexact line search for a new hybrid conjugate gradient method. *Open Access Library Journal*, 7: e6048. <https://doi.org/10.4236/oalib.1106048>
- Andrei, N. (2008). A hybrid conjugate gradient algorithm for unconstrained optimization as convex combination of Hestenes- Stiefel and Dai-Yuan. *Study Information Control*, 17, 55-70.
- Andrei, N. (2008). An unconstrained optimization test function. *Advanced Modeling and Optimization. An Electronic International Journal*,10, 172-182.
- Andrei, N. (2009). Hybrid conjugate gradient algorithm for unconstrained optimization. *Journal of Optimization Theory Application*, 141, 249-264.
- Andrei, N. (2017). A Dai-Liao nonlinear conjugate gradient algorithm. *Numer. Algor.*, DOI 10.1007/s11075-017-0362-5
- Babaie-Kafaki, S. (2011). A modified BFGS algorithm based on a hybrid secant equation. *Science China Mathematics*,58, 315-331.
- Babaie-Kafaki, S. (2013). A hybrid conjugate gradient method based on quadratic relaxation of Dai-Yuan hybrid conjugate gradient parameter. *Journal of Mathematical Programming and Operation Research*, 62(7), 929-941.
- Babaie-Kafaki, S. and Ghanbari, R. (2014c). Two hybrid nonlinear conjugate gradient methods based on a modified secant equation. *Journal of Mathematical Programming and Operation Research*,63(7), 1027-1042.
- Babaie-Kafaki, S., and Ghanbari, R. (2014a). The Dai-Liao nonlinear conjugate gradient method with optimal parameter choices. *Eur. J. Oper. Res.* 234(3), 625-630.
- Babaie-Kafaki, S., and Ghanbari, R. (2014b). A descent family of Dai-Liao conjugate gradient methods. *Optim. Methods Softw.*, 29(3): 583-591.
- Babaie-Kafaki, S., and Ghanbari, R. (2015). Two optimal Dai-Liao conjugate gradient methods. *Optimization*, 64(1): 2277-2287.
- Babaie-Kafaki, S. (2015). On optimality of two adaptive choices for the parameter of Dai-Liao method. *Optim. Lett.*, DOI 10.1007/s11590-015-0965-5.
- Babaie-Kafaki, S., and Ghanbari, R. (2017). Two adaptive Dai-Liao nonlinear conjugate gradient methods. *Iran J. Sci. Technol. Trans. Sci.*, DOI 10.1007/s40995-017- 027 1-4.
- Dai, Y.H. and Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.*,10, 177-182.
- Dai, Y.H and Liao, L.Z. (2001). New conjugacy conditions and related nonlinear conjugate gradient methods, *Appl. Math. Optim.*, 43, 87-101.

- Dai, Y.H. and Yuan, Y. (2001). An efficient hybrid conjugates gradient method for unconstrained optimization. *Annals of Operations Research*, 103, 33-47.
- Djordjevic, S.S. (2016). New hybrid conjugate gradient method as a convex combination of FR and PRP methods. *Published by faculty of science and mathematics. University of Nis, Serbia, Filomat*, 31, 3083-3100. <https://doi.org/10.2298/FIL16083D>
- Djordjevic, S. S. (2017). New hybrid conjugate gradient methods as a convex combination of LS and CD methods. *Published by faculty of science and mathematics. University of Nis, Serbia,31(6)*, 1813-1825.
- Djordjevic, S. S. (2018). New hybrid conjugate gradient method as a convex combination of HS and FR conjugate gradient methods. *Journal of Applied Mathematics and Computation*, 2, 366-378. <https://doi.org/10.26855/jamc.2018.09.002>
- Djordjevic, S.S. (2019). New hybrid conjugate gradient method as a convex combination of LS and FR conjugate gradient methods. *Acta Mathematica Scientia*, 39, 214-228.
- Dolan, E. D. and Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Journal of Math. Program*, 91(2), 201-213.
- Esmaeili H., Rostami M. and Kimiaei M. (2018). Extended Dai–Yuan conjugate gradient strategy for large-scale unconstrained optimization with applications to compressive sensing. *Published by faculty of sciences and mathematics, University of Nis, Serbia, Filomat*, 32(6), 2173–2191. Available at: <http://www.pmf.ni.ac.rs/filomat>
- Fletcher, R. and Reeves, C. (1964). Function minimization by conjugate gradients. *Computational Journal*,7, 149-154.
- Fletcher, R. (1987). *Practical methods of optimization*, (vol. 1), *Unconstrained optimization*. New York: John Wiley and Sons.
- Gould, N. I. M., Orban, D. and Toint, P.L. (2003). CUTER: a constrained and unconstrained testing environment, revisited. *ACM Transactions on Mathematical Software*,29(4), 373-394.
- Guo, J. and Wan, Z. (2019). A modified spectral PRP conjugate gradient projection method for solving large-scale monotone equations and its application in compressed sensing. *Hindawi Mathematical Problems in Engineering*. Volume 2019, Article ID 5261830, 17 pages. <https://doi.org/10.1155/2019/5261830>
- Hager, W.W. and Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. *Pacific J. Optim.*, 2, 35-58.
- Hestenes, M. R. and Stiefel, E. L. (1952). Methods of conjugate gradients for solving linear systems. *Journal Res. Nat. Bur. Stand.*, 49, 409-436
- Ibrahim, A.H., Kumam, P., Abubakar, A. B., Jirakitpuwapat, W., and Abubakar, J. (2020). A hybrid conjugate gradient algorithm for constrained monotone equations with application in compressive sensing. *Heliyon*.e03466. <https://doi.org/10.1016/j.heliyon.2020.e03466>
- Li, D. H. and Fukushima, A. M. (2001). modified BFGS method and its global convergence in nonconvex minimization, *J. Comput. Appl. Math.* 129(1) 15-35.
- Liu, K., and Du, S. (2019). Modified three-term conjugate gradient method and its applications. *Hindawi Mathematical Problems in Engineering*. Volume 2019, Article ID 5976595, 9 pages. <https://doi.org/10.1155/2019/5976595>

- Liu, J., Du, S. and Chen, Y. (2020). A sufficient descent nonlinear conjugate gradient method for solving M-tensor equations. *Journal of Computational and Applied Mathematics*, 371 112709
- Liu, Y. and Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: Theory. *Journal of optimization theory and applications*. 69, 129-137.
- Lotfi, M. and Hosseini, S. M. (2019). An efficient Dai–Liao type conjugate gradient method by reformulating the CG parameter in the search direction equation. *Journal of Computational and Applied Mathematics*, 371. 112708. <https://doi.org/10.1016/j.cam.2019.112708>
- Mohammed, N. S., Mustafa M., Mohd R. and Shazlyn M. S. (2020). A new hybrid coefficient of conjugate gradient method. *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 18, No. 3, June 2020, pp. 1454-1463
- Perry, A. (1978). A modified conjugate gradient algorithm. *Oper. Res. Tech. Notes*, 26(6): 1073-1078.
- Polak, B. T. (1969). The conjugate gradient method in extreme problems. *USSR Comput. Math. Math. Phys.* 4, 94-112.
- Polak, B.T. (1967). A general method of solving extremal problems. *Soviet Math. Doklady*, 8, 14-29.
- Powell, M.J.D. (1984). Nonconvex minimization calculations and the conjugate gradient method, *Numerical Analysis (Dundee, 1983)*, D.F. Griffiths, ed., *Lecture Notes in Mathematics*, Vol. 1066, Springer, Berlin. 122-141.
- Salihu, N., Odekunle, M., Waziri, M. and Halilu, A. (2020). A new hybrid conjugate gradient method based on secant equation for solving large scale unconstrained optimization problems. *Iranian Journal of Optimization*, 12(1): 33-44.
- Salihu, N., Odekunle, M., Waziri, M., Halilu, A., and Salihu, S (2021). A Dai- Liao hybrid conjugate gradient method for unconstrained optimization. *International Journal of Industrial Optimization* 2, 69–84.
- Salihu, N., Odekunle, M. R., Saleh, A. M., and Salihu, S (2021). A Dai-Liao hybrid Hestenes-Stiefel and Fletcher-Reeves methods for unconstrained optimization. *Int. J. Indust. Opt.*, 2(1):33–50.
- Salihu, N., Kumam, P., Awwal, A. M., Arzuka, I., and Seangwattana, T., A (2023a). structured Fletcher -Reeves spectral conjugate gradient method for unconstrained optimization with application in robotic model. *Oper. Res. Forum* 4, Paper No. 81, 25.
- Salihu, N., Kumam, P., Awwal, A. M., Sulaiman, I. M., and Seangwattana, T. (2023b). The global convergence of spectral RMIL conjugate gradient method for unconstrained optimization with applications to robotic model and image recovery. *Plos one* 18, 3, e0281250.
- Salihu, N., Kumam, P., Sulaiman, I. M., and Seangwattana, T (2023c). An efficient spectral minimization of the dai-yuan method with application to image reconstruction. *AIMS Mathematics* 8(12):30940–30962.
- Wang, X. A class of spectral three-term descent Hestenes-Stiefel conjugate gradient algorithms for large-scale unconstrained optimization and image restoration problems. *Appl. Numer. Math.* 192 (2023), 41–56
- Waziri, M. Y., Ahmed, K., and Sabi’u, J. (2019). A Dai-Liao conjugate gradient method via modified secant equation for system of nonlinear equations. *Arab. J. Math.*, <https://doi.org/10.1007/s40065-019-0264-6>.
- Xue, W., Ren, J., Zheng, X., Liu, Z., and Liang, Y. (2018). A New DY conjugate gradient method

and applications to image denoising. *IEICE TRANS. INF. and SYST.*, VOL.E101-D, NO.1 2,
December 2018. DOI: 10.1587/transinf.2018EDP7210

Yao, S., Feng, Q., Li, L., and Xu, J. (2019). A class of globally convergent three-term Dai-Liao conjugate gradient methods. *Applied Numerical Mathematics*, 151 354-366.
<https://doi.org/10.1016/j.apnum.2019.12.026>.

Zoutendijk, G. (1970). Nonlinear programming computational methods, in integer and nonlinear programming. *J. Abadie ed.*, North-Holland, Amsterdam 37-86.

