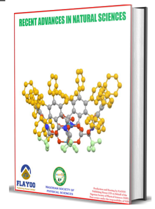
Published by Nigerian Society of Physical Sciences. **Hosted by** FLAYOO Publishing House LTD

Recent Advances in Natural Sciences

Journal Homepage: https://flayoophl.com/journals/index.php/rans



Global convergence properties of a Dai-Liao-type CGM for unconstrained optimization

Oluwaseun Biodun Onuoha*

Department of Mathematical Sciences, Adekunle Ajasin University, P.M.B. 001, Akungba Akoko, Ondo State, Nigeria

ARTICLE INFO

Article history:
Received: 05 September 2023
Received in revised form: 05 February 2024
Accepted: 12 February 2024
Available online: 04 March 2024

Keywords: Unconstrained optimization, Convergence properties, Descent directions, Objective function

DOI:10.61298/rans.2024.2.1.30

ABSTRACT

A popular optimization technique called the conjugate gradient method (CGM) is renowned for its effectiveness in addressing problems involving unconstrained optimization. Several conjugate gradient (CG) techniques have been proven to possess global convergence properties both theoretically and numerically. The Dai-Liao-type CGM is a variant that incorporates certain modifications to enhance its convergence properties. This paper examines the global convergence properties of a Dai-Liao-type CGM for unconstrained optimization problems. Theoretically, this study investigates the conditions under which the method ensures convergence to the global minimum of the objective function, focusing on the algorithm's descent directions, the necessary reduction in objective function values, and termination criteria. A numerical experiment is conducted on a set of unconstrained optimization problems to validate the theoretical results obtained in this work. The numerical findings of this study demonstrate the robustness and reliability of the Dai-Liao-type CGM, showing its ability to find the global optimal solution in a wide range of unconstrained optimization problems.

1. INTRODUCTION

Optimization entails choosing the best out of numerous possibilities [1]. Optimization theory is widely applied to optimal control problems [2, 3]. In this paper, we consider a nonlinear, unconstrained minimization problem:

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

where the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is smooth and its gradient is represented by $\nabla f(x) = g(x)$. To solve Eq. (1), the conjugate gradient method (CGM) is usually employed. The conjugate gradient (CG) method is developed iteratively by:

$$x_{n+1} = x_n + \varphi_n d_n, \qquad n \geq 0 \tag{2}$$

where φ_n is the steplength, and d_n is the direction of search. At the first iteration, the direction of search is the steepest descent one, that is, $d_0 = -g_0$. Consequently, d_n is defined as:

$$d_n = -g_n + \beta_n d_{n-1}, \tag{3}$$

where β_n is a scalar known as the CG update parameter. Some of the most frequently used update parameters are:

$$\beta_n^{FR} = \frac{\|g_n\|^2}{\|g_{n-1}\|^2}, \tag{4}$$

$$\beta_n^{HS} = \frac{g_n^T y_{n-1}}{d_{n-1}^T y_{n-1}}, \tag{5}$$

*Corresponding Author Tel.: +234-703-0826-908
e-mail: oluwaseun.akinduko@aaua.edu.ng (Oluwaseun Biodun Onuoha)

$$\beta_n^{DY} = \frac{||g_n||^2}{d_{n-1}^T y_{n-1}}, \tag{6}$$

$$\beta_n^{PRP} = \frac{g_n^T y_{n-1}}{||g_{n-1}||^2}, \tag{7}$$

namely, Fletcher-Reeves (FR) method [4], Hestenes-Stiefel (HS) method [5], Dai-Yuan (DY) method [6], and Polak-Ribiere-Polyak (PRP) method [7, 8], respectively.

In the above classical CG formulae, the difference between two gradients is denoted by $y_{n-1} = g_n - g_{n-1}$, $||.||$ denotes norm, T denotes transpose, $||g_n||^2 = g_n^T g_n$, and $||g_{n-1}||^2 = g_{n-1}^T g_{n-1}$. The convergence characteristics of the aforementioned CG methods have been investigated in Refs. [9–12], where the FR and DY methods have been shown to possess strong convergence properties with modest performances, while the PRP and HS methods were proven to have better computational performances, which may not be generally convergent.

Despite the DY method's strong convergence property, its efficacy can be affected by the line search technique and parameters used. The convergence may be affected by the precision of the line search in identifying the step size, and for non-convex functions, its convergence may be slower, and it could become trapped in local minima or saddle points [13, 14]. To overcome the drawbacks of the DY technique, Jiang & Jian in Ref. [15] presented a new hybrid CGM, called an improved Dai-Yuan method, where the update parameter is given by

$$\beta_n^{IDY} = \frac{\left|g_n^T d_{n-1}\right|}{-g_{n-1}^T} \cdot \frac{||g_n||^2}{d_{n-1}^T (g_n - g_{n-1})},$$

which they proved to be convergent under the strong Wolfe line search.

Many researchers have developed new methods in response to the need to produce some CG methods with good convergence properties that also perform well computationally. The Dai-Liao type CGM has drawn attention within the CGM family due to improvements targeted at improving its convergence features. Dai & Liao (DL) in Ref. [16] proposed a novel CG approach that can be deemed an improvement over the HS method by incorporating the conjugacy condition:

$$d_n^T y_{n-1} = -t g_n^T s_{n-1}, \tag{8}$$

with the resulting CG parameter given by

$$\beta_n^{DL} = \frac{g_n^T (y_{n-1} - t s_{n-1})}{d_{n-1}^T y_{n-1}}, \tag{9}$$

where $t \geq 0$.

Nevertheless, as the DL-CG method's performance is dependent on the parameter t, the optimal value of t in Eq. (9) is still being taken into consideration [17, 18]. Researchers have dedicated significant efforts to enhance the effectiveness of the DL method. For example, the authors in Ref. [17] suggested the following selections for the parameter t :

$$t_1 = \frac{||y_{n-1}||}{||s_{n-1}||}, \qquad t_2 = \frac{y_{n-1} s_{n-1}}{||s_{n-1}||^2} + \frac{||y_{n-1}||}{||s_{n-1}||}.$$

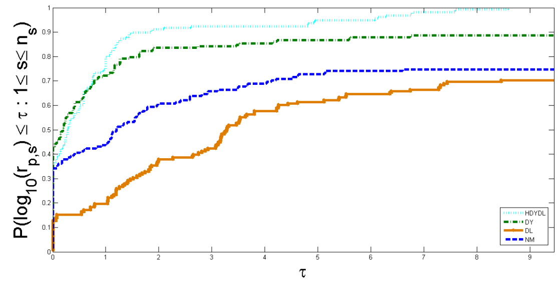


Figure 1. Iteration Profile of the HDYDL Method in Comparison to the DY, DL, and NM Methods

Salihu *et al.* [19] employed the optimal choice of the parameter t to modify CG methods using the classical HS and FR methods, while Lofti & Hosseini [20] presented a new value t based on a modified BFGS method.

The fulfillment of the sufficient descent criterion:

$$g_n^T d_n \leq -k\, ||g_n||^2, \qquad 0 < k \leq 1 \tag{10}$$

is very crucial to the global convergence of CG methods. Akinduko [21] proposed a new hybrid CGM of a Dai-Liao type, which was proved to fulfill the condition (10) and given by:

$$\beta_k^{NM} = \frac{g_k^T s_{k-1}}{g_{k-1}^T y_{k-1}}. \tag{11}$$

Furthermore, Onuoha [22] proposed a new hybrid CG method with sufficient descent property by combining the DY and DL methods; the resulting update parameter is given by:

$$\beta_n^{HDYDL} = \frac{g_n^T (g_n - t s_{n-1})}{d_{n-1}^T y_{n-1}}. \tag{12}$$

Recently, global convergence results for CG methods under various inexact line searches have been established in Refs. [23–25]. To demonstrate the convergence of CG techniques, the steplength φ_n typically needs to meet the strong Wolfe (SW) conditions put forth by Wolfe [26] and provided by:

$$f_n - f_{n+1} \geq -\eta \varphi_n g_n^T d_n \tag{13}$$

and

$$|g_{n+1}^T d_n| \leq \upsilon |g_n^T d_n|, \tag{14}$$

where $0 \leq \eta \leq \upsilon < 1, f_n = f(x_n), f_{n+1} = f(x_n + \varphi_n d_n)$, and $g_{n+1} = g(x_n + \varphi_n d_n)$.

Yousif *et al.* [27] introduced a criterion that ensures the establishment of the descent search direction property and the global convergence of CG techniques under SW line search. Many researchers (see e.g., Refs [28–32]) have demonstrated that several numerical techniques for unconstrained optimization converge under the SW condition.

This work aims to examine the global convergence features of the CGM, following the proposal of Onuoha [22]. Specifically, it focuses on how much the Dai-Liao-type CGM attains the global minimum of the objective function. The continuous quest for optimization techniques exhibiting robust theoretical convergence characteristics and practical efficacy in practical applications serves as the driving force behind this research.

Table 1. Test Problems and their Initial Points

S/N	Problem Names	Initial Points
1	Extended Block Diagonal BD1	$(0.1, 0.1, \cdots, 0.1), (0.3, 0.3, \cdots, 0.3), (0.5, 0.5, \cdots, 0.5), (0.7, 0.7, \cdots, 0.7)$
2	MODF COSINE	$(1./n, 1./n, \cdots, 1./n), (2./n, 2./n, \cdots, 2./n), (5./n, 5./n, \cdots, 5./n), (7./n, 7./n, \cdots, 7./n)$
3	Diagonal 4	$(1, 1, \cdots, 1), (3, 3, \cdots, 3), (5, 5, \cdots, 5), (7, 7, \cdots, 7)$
4	Diagonal 5	$(1.1, 1.1, \cdots, 1.1), (1.3, 1.3, \cdots, 1.3), (1.5, 1.5, \cdots, 1.5), (1.8, 1.8, \cdots, 1.8)$
5	MODF SINE	$(1./n, 1./n, \cdots, 1./n), (3./n, 3./n, \cdots, 3./n), (5./n, 5./n, \cdots, 5./n), (7./n, 7./n, \cdots, 7./n)$
6	Extended Beale	$(1, 0.8, \cdots, 1, 0.8), (1, 2, \cdots, 1, 2), (2, 2, \cdots, 2, 2), (1, 0.6, \cdots, 1, 0.6)$
7	RMODF COSINE	$(1, 1, \cdots, 1), (5, 5, \cdots, 5), (10, 10, \cdots, 10), (20, 20, \cdots, 20)$
8	MDF EXPLIN 1	$(1, 1, \cdots, 1), (3, 3, \cdots, 3), (5, 5, \cdots, 5), (7, 7, \cdots, 7)$
9	RMDF GENHUMPS	$(1, 1, \cdots, 1), (3, 3, \cdots, 3), (3.5, 3.5, \cdots, 3.5), (4.5, 4.5, \cdots, 4.5)$
10	RMODF SINE	$(1, 1, \cdots, 1), (2, 2, \cdots, 2), (3, 3, \cdots, 3), (7, 7, \cdots, 7)$
11	QUARTC	$(2, 2, \cdots, 2), (4, 4, \cdots, 4), (6, 6, \cdots, 6), (8, 8, \cdots, 8)$
12	Partial Perturbed Quadratic	$(0.5, 0.5, \cdots, 0.5), (0.7, 0.7, \cdots, 0.7), (0.9, 0.9, \cdots, 0.9), (1.0, 1.0, \cdots, 1.0)$
13	Generalized Quartic	$(1, 1, \cdots, 1), (0.1, 0.1, \cdots, 0.1), (0.3, 0.3, \cdots, 0.3), (0.5, 0.5, \cdots, 0.5)$
14	Extended DENSCHNB	$(1, 1, \cdots, 1), (5, 5, \cdots, 5), (10, 10, \cdots, 10), (15, 15, \cdots, 15)$
15	Diagonal 8	$(1, 1, \cdots, 1), (3, 3, \cdots, 3), (5, 5, \cdots, 5), (0.5, 0.5, \cdots, 0.5)$
16	Diagonal 7	$(1, 1, \cdots, 1), (0.1, 0.1, \cdots, 0.1), (0.3, 0.3, \cdots, 0.3), (0.5, 0.5, \cdots, 0.5)$
17	SINCOS	$(3, 0.1, \cdots, 3, 0.1), (3, 0.5, \cdots, 3, 0.5), (3, 3, \cdots, 3, 3), (1, 2, \cdots, 1, 2)$
18	Full Hessian FH3	$(1, 1, \cdots, 1), (3, 3, \cdots, 3), (5, 5, \cdots, 5), (7, 7, \cdots, 7)$
19	Extended Tridiagonal-1	$(2, 2, \cdots, 2), (4, 4, \cdots, 4), (5, 5, \cdots, 5), (6, 6, \cdots, 6)$
20	HIMMELBG	$(1.5, 1.5, \cdots, 1.5), (0.1, 0.1, \cdots, 0.1), (0.5, 0.5, \cdots, 0.5), (0.7, 0.7, \cdots, 0.7)$

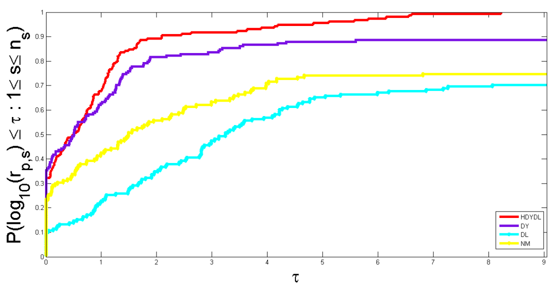


Figure 2. CPU Profile of the HDYDL Method in Comparison to the DY, DL, and NM Methods

2. GLOBAL CONVERGENCE ANALYSIS OF HDYDL-CGM

In this section, the global convergence results for the hybrid method proposed in Ref. [22] are provided. The method was implemented based on the following algorithm:

ALGORITHM 2.1: HDYDL METHOD

Step 1: Input x_0, set $n = 0$, $d_0 = -g_0$.
Step 2: Terminate the algorithm if $\|g_n\| \leq 10^{-6}$.
Step 3: Let d_n be calculated by Eq. (3)
Step 4: Let the step length φ_n be determined by Eqs. (13) and (14).
Step 5: Calculate x_n by Eq. (2).
Step 6: Calculate β_n by Eq. (12).
Step 7: Make $n := n + 1$, and go back to Step 2.

The following lemma is useful for the establishment of the global convergence of the HDYDL method:

Lemma 2.1 ([22]). *The HDYDL method satisfies the sufficient descent condition (10) where,*

$$k = \frac{1}{1-\sigma}, \quad \sigma \in [0, 1). \tag{15}$$

Definition 2.1. *A CGM is said to be globally convergent if, starting from any given initial iterate x_0, it meets the condition:*

$$\lim_{n\to\infty} \|g_n\| = 0,$$

where g_n is the gradient of the objective function f at the point x_n.

To determine the global convergence of the CG method, the following assumptions are made:

Assumption 2.1. *The level set $\omega = \{x : f(x) \leq f(x_0)\}$ is bounded.*

Assumption 2.1 implies that there is a constant B such that:

$$|g(x)| \leq B \quad \forall \quad x \in \omega. \tag{16}$$

Assumption 2.2. *In some neighborhood Z of ω, f is Lipschitz continuously differentiable, that is, there exists a positive constant M such that:*

$$\|g(a) - g(b)\| \leq M \|a - b\| \quad \forall \quad a, b \in Z. \tag{17}$$

Lemma 2.2. *Suppose that Assumption 2.1 is satisfied and consider any method of the forms (2) and (3), where d_n is a descent search direction and φ_n satisfies the SW conditions (13) and (14), then*

$$\sum_{n=0}^{\infty} \frac{(g_n^T d_n)^2}{\|d_n\|^2} < \infty. \tag{18}$$

Table 2. Numerical Results for HDYDL, DY, DL, and NM Methods

		Number of Iteration				Computational time			
TP	Dim	HDYDL	DY	DL	NM	HDYDL	DY	DL	NM
1 - 1	5000	114	116	71	31	2.428	4.211	3.303	1.771
1 - 2	5000	34	30	109	75	0.973	1.349	3.771	4.652
1 - 3	10000	118	F	F	42	4.748	F	F	3.695
1 - 4	10000	21	21	101	43	0.844	1.212	4.491	2.481
1 - 4	5000	21	21	75	53	0.445	1.144	5.294	5.374
1 - 4	7000	21	21	70	49	0.589	0.868	5.76	9.311
1 - 4	8000	21	21	50	37	0.681	0.914	3.092	1.405
2 - 1	5000	6	4	F	F	0.341	0.278	F	F
2 - 1	6000	6	4	F	F	0.126	0.36	F	F
2 - 2	7000	8	4	F	F	0.281	0.285	F	F
2 - 2	8000	6	4	F	F	0.298	0.135	F	F
2 - 3	9000	8	4	F	F	0.261	0.282	F	F
2 - 3	10000	8	4	F	F	0.512	0.12	F	F
2 - 4	5000	10	6	F	F	0.551	0.205	F	F
2 - 4	10000	8	4	F	F	0.308	0.122	F	F
3 - 1	5000	15	15	F	44	0.691	0.922	F	1.943
3 - 1	10000	15	15	F	45	0.697	0.789	F	2.054
3 - 2	5000	22	22	36	166	1.013	1.065	1.601	13.302
3 - 2	6000	22	22	36	167	1.097	1.032	1.904	10.875
3 - 3	7000	12	12	F	26	1.005	1.432	F	1.676
3 - 3	8000	12	12	F	26	0.753	2.769	F	1.855
3 - 4	9000	20	19	41	16	1.259	0.8	3.398	1.157
3 - 4	10000	20	19	33	16	1.165	2.252	3.424	1.036
4 - 1	5000	21	23	164	19	0.644	0.779	5.11	0.992
4 - 1	6000	21	23	165	19	1.132	0.706	5.144	0.755
4 - 2	7000	34	47	170	15	1.434	2.153	5.536	0.583
4 - 2	8000	34	47	172	15	1.501	1.503	5.467	0.625
4 - 3	9000	9	9	171	31	0.336	0.36	6.251	1.437
4 - 3	10000	9	9	172	31	0.555	0.264	6.18	1.481
4 - 4	5000	27	33	169	34	2.241	0.964	5.613	1.388
4 - 4	10000	28	33	F	61	1.285	1.18	F	2.911
5 - 1	5000	1	1	1	1	0.037	0.044	0.015	0.041
5 - 1	10000	1	1	1	1	0.035	0.029	0.041	0.038
5 - 2	6000	3	3	3	2	0.083	0.113	0.154	0.085
5 - 2	7000	1	1	1	1	0.046	0.028	0.032	0.039
5 - 3	8000	3	3	9	3	0.156	0.123	0.463	0.142
5 - 3	9000	3	3	7	2	0.156	0.079	0.207	0.072
5 - 4	5000	3	3	22	10	0.165	0.107	0.752	0.388
5 - 4	10000	3	3	12	4	0.085	0.081	0.466	0.161

Proof. The Lipschitz condition (17) implies that:

$$(g_{n+1} - g_n)^T d_n \le \|g_{n+1} - g_n\| \, \|d_n\| \le \omega \, \|x_{n+1} - x_n\| \, \|d_n\| .$$

By Eq. (2)

$$\varphi_n d_n = x_{n+1} - x_n.$$

Therefore,

$$\begin{aligned}(g_{n+1} - g_n)^T d_n &\le \omega \, \|x_{n+1} - x_n\| \, \|d_n\| , \\ &\le \omega \, |\varphi_n| \, \|d_n\| \, \|d_n\| , \\ &= \omega \, |\varphi_n| \, \|d_n\|^2 .\end{aligned}$$

Thus,

$$\varphi_n \ge \frac{(g_{n+1} - g_n)^T d_n}{\omega \, \|d_n\|^2}, \tag{19}$$

$$\ge \frac{d_n^T y_n}{\omega \, \|d_n\|^2}. \tag{20}$$

Eq. (14) implies that:

$$-\upsilon g_n^T d_n \le \left| g_{n+1}^T d_n \right| \le \upsilon g_n^T d_n.$$

Subtracting $g_n^T d_n$ from both sides of the inequalities above we have:

$$-\upsilon g_n^T d_n - g_n^T d_n \le \left| g_{n+1}^T d_n \right| - g_n^T d_n \le \upsilon g_n^T d_n - g_n^T d_n,$$

Table 3. Numerical Results for HDYDL, DY, DL, and NM Methods (cont'd)

		Number of Iteration				Computational time			
TP	Dim	HDYDL	DY	DL	NM	HDYDL	DY	DL	NM
6 - 1	5000	109	242	F	1223	8.21	22.916	F	82.014
6 - 1	10000	110	242	F	1243	11.758	42.67	F	158.244
6 - 2	5000	955	F	34	1206	66.223	F	4.485	80.894
6 - 2	10000	994	F	35	1222	178.082	F	10.181	161.16
6 - 3	5000	53	45	F	1112	3.673	3.319	F	77.315
6 - 3	10000	57	46	F	1139	7.542	6.037	F	152.966
6 - 4	5000	149	57	F	996	13.045	4.163	F	67.648
6 - 4	10000	149	60	55	1019	25.189	7.835	14.157	140.906
7 - 1	5000	15	F	40	F	0.832	F	1.593	F
7 - 1	10000	16	F	40	F	0.568	F	1.682	F
7 - 2	5000	24	24	126	14	1.021	1.167	3.285	0.463
7 - 2	7000	24	24	127	14	1.235	1.091	3.879	0.567
7 - 3	8000	20	9	221	F	0.825	0.462	8.133	F
7 - 3	9000	20	9	195	27	0.995	0.354	7.285	0.904
7 - 4	5000	27	F	F	F	1.189	F	F	F
8 - 1	5000	19	16	156	18	0.509	0.597	13.857	1.249
8 - 1	6000	19	16	157	18	0.822	0.765	11.71	1.444
8 - 2	7000	29	29	167	12	1.111	1.564	7.583	0.572
8 - 2	8000	29	29	167	12	1.259	1.49	7.121	0.627
8 - 3	5000	21	17	161	18	0.991	2.371	6.961	1.449
8 - 3	9000	21	17	164	18	2.026	2.185	7.053	2.182
8 - 4	5000	27	25	198	16	1.859	3.244	10.209	0.599
8 - 4	10000	27	26	202	16	1.041	3.323	6.902	0.982
9 - 1	5000	173	F	F	F	62.419	F	F	F
9 - 2	5000	37	37	145	F	1.512	1.287	3.608	F
9 - 2	6000	37	38	F	F	1.533	0.876	F	F
9 - 2	7000	37	38	147	F	1.596	1.048	4.396	F
9 - 3	5000	18	27	F	12	0.842	0.593	F	0.417
9 - 3	7000	18	27	132	12	0.623	0.782	4.398	0.411
9 - 3	8000	18	27	F	12	0.611	1.036	F	0.429
9 - 4	5000	34	F	F	F	1.322	F	F	F
10 - 1	5000	20	20	40	18	0.676	0.64	1.279	0.956
10 - 1	10000	20	20	29	40	0.504	0.569	1.072	5.242
10 - 2	6000	10	10	F	23	0.491	0.271	F	1.509
10 - 2	8000	10	10	62	23	0.519	0.251	4.69	1.991
10 - 3	9000	14	16	32	21	0.661	0.596	0.947	1.654
10 - 3	10000	14	16	32	30	0.867	0.41	0.723	3.439
10 - 4	5000	20	27	50	14	0.754	0.667	1.887	0.711
10 - 4	10000	20	38	45	14	0.783	1.815	2.484	0.844

$$(-\upsilon - 1)g_n^T d_n \leq \left|g_{n+1}^T d_n\right| - g_n^T d_n,$$

Thus

$$-(\upsilon + 1)g_n^T d_n \leq \left|g_{n+1}^T d_n\right| - g_n^T d_n,$$

and

$$\begin{aligned}-(\upsilon + 1)g_n^T d_n &\leq \left|g_{n+1}^T d_n\right| - g_n^T d_n,\\ &\leq \left|g_{n+1}^T d_n\right| - \left|g_n^T d_n\right|,\\ &\leq \left|(g_{n+1} - g_n)^T d_n\right|,\\ &\leq \left|d_n^T y_n\right|.\end{aligned}$$

Therefore

$$\left|d_n^T y_n\right| \geq -(\upsilon + 1)g_n^T d_n,$$

which also implies that:

$$\left|d_n^T y_n\right| \geq d_n^T y_n \geq -(\upsilon + 1)g_n^T d_n.$$

The above inequality, combined with Eq. (20), implies that

$$\varphi_n \geq \frac{-(1+\upsilon)}{\omega}\frac{g_n^T d_n}{||d_n||^2}.$$

Table 4. Numerical Results for HDYDL, DY, DL, and NM Methods (cont'd)

		Number of Iteration				Computational time			
TP	Dim	HDYDL	DY	DL	NM	HDYDL	DY	DL	NM
11 - 1	5000	1	1	1	1	0.054	0.118	0.033	0.041
11 - 1	6000	1	1	1	1	0.034	0.031	0.032	0.032
11 - 2	8000	386	1	173	5	11.382	0.038	4.446	0.2
11 - 2	9000	391	536	173	5	11.004	11.188	4.297	0.231
11 - 3	5000	337	88	231	5	7.299	2.395	5.753	0.211
11 - 3	7000	350	94	232	5	7.787	2.5	5.3	0.274
11 - 4	9000	398	625	242	13	8.493	15.737	6.012	0.777
11 - 4	10000	402	636	243	13	15.282	15.332	5.937	0.84
12 - 1	5000	4	3	321	49	0.266	0.17	12.489	2.212
12 - 1	10000	4	4	1401	51	0.289	0.642	77.692	3.22
12 - 2	6000	4	2	311	14	0.361	0.102	14.825	0.575
12 - 2	7000	4	2	332	46	0.241	0.133	20.411	2.083
12 - 3	8000	4	2	316	12	0.29	0.227	7.536	0.605
12 - 3	9000	4	4	333	24	0.513	0.393	9.278	1.346
12 - 4	8000	4	2	317	12	0.294	0.12	7.637	0.579
12 - 4	5000	4	4	328	50	0.23	0.231	9.193	1.829
13 - 1	5000	33	38	91	47	1.217	2.459	4.58	1.667
13 - 1	10000	33	36	F	27	1.389	2.301	F	0.921
13 - 2	6000	15	14	40	17	1.022	0.885	1.225	0.621
13 - 2	7000	15	14	37	15	0.961	0.971	1.261	0.527
13 - 3	8000	20	20	56	18	1.227	1.531	1.996	0.602
13 - 3	9000	20	20	35	17	1.486	1.267	1.339	0.638
13 - 4	7000	22	22	41	19	1.486	1.583	1.474	0.723
13 - 4	10000	23	22	74	19	1.587	1.321	2.564	0.632
14 - 1	5000	27	29	20	19	0.942	1.192	1.263	1.567
14 - 1	10000	29	29	20	19	1.055	1.455	1.217	1.587
14 - 2	6000	45	F	38	F	2.086	F	2.099	F
14 - 2	7000	46	F	38	F	1.907	F	2.088	F
14 - 3	8000	41	33	51	25	1.886	1.628	2.888	0.901
14 - 3	9000	41	33	51	25	1.584	1.75	4.05	0.903
14 - 4	5000	102	48	F	127	4.327	2.42	F	4.611
14 - 4	10000	92	49	34	132	4.292	3.183	2.38	5.011
15 - 1	5000	21	20	F	15	0.505	1.066	F	0.726
15 - 1	6000	21	20	F	15	0.765	0.669	F	0.753
15 - 2	7000	74	F	F	F	2.883	F	F	F
15 - 2	8000	74	F	F	F	4.422	F	F	F
15 - 3	9000	88	F	F	F	3.626	F	F	F
15 - 3	10000	88	F	F	F	3.245	F	F	F
15 - 4	5000	16	16	160	18	0.623	1.002	5.429	0.853
15 - 4	10000	18	18	163	1740	0.653	0.977	5.419	73.228

Eq. (13) implies that

$$\begin{aligned} f(x_n) - f(x_{n+1}) &\geq -\eta\varphi_n g_n^T d_n, \\ &\geq \frac{(-\eta)(-(1+\upsilon))}{\omega}\frac{(g_n^T d_n)^2}{\|d_n\|^2}, \\ &= \frac{\eta(1+\upsilon)}{\omega}\frac{(g_n^T d_n)^2}{\|d_n\|^2}, \\ &= \gamma\frac{(g_n^T d_n)^2}{\|d_n\|^2}, \end{aligned}$$

where

$$\gamma = \frac{\eta(1+\upsilon)}{\omega}.$$

Summing the above from $n = 0$ to $n = m$, we have

$$\begin{aligned} \sum_{n=0}^{m}\frac{\gamma(g_n^T d_n)^2}{\|d_n\|^2} &\leq \sum_{n=0}^{m} f(x_n) - f(x_{n+1}), \\ &= f(x_0) - f(x_{m+1}), \\ &\leq f(x_0) + |f(x_{m+1})|, \\ &\leq f(x_0) + M. \end{aligned}$$

Table 5. Numerical Results for HDYDL, DY, DL, and NM Methods (cont'd)

		Number of Iteration				Computational time			
TP	Dim	HDYDL	DY	DL	NM	HDYDL	DY	DL	NM
16 - 1	5000	91	88	152	F	3.587	2.975	4.979	F
16 - 1	10000	93	90	155	F	4.391	3.002	5.437	F
16 - 2	5000	19	18	151	F	1.749	0.583	5.843	F
16 - 2	6000	19	18	152	F	2.717	0.637	5.044	F
16 - 3	7000	14	23	161	F	0.508	0.84	6.114	F
16 - 3	8000	14	23	162	F	0.637	0.657	6.037	F
16 - 4	9000	19	20	170	F	0.644	0.677	6.774	F
16 - 4	10000	19	20	171	F	0.623	0.548	6.36	F
17 - 1	5000	374	391	162	251	17.358	62.973	5.337	9.819
17 - 1	10000	350	386	58	500	17.382	7.38	2.817	18.486
17 - 2	6000	58	44	64	175	2.092	1.6	2.356	6.177
17 - 2	7000	96	57	131	175	3.842	1.823	3.861	6.493
17 - 3	6000	101	123	83	268	4.107	4.097	2.742	9.159
17 - 3	8000	130	127	98	521	5.369	4.47	3.163	18.273
17 - 4	5000	297	F	69	151	11.634	F	2.127	5.215
17 - 4	10000	281	F	165	164	14.019	F	6.1	6.265
18 - 1	5000	30	26	F	20	1.781	1.811	F	1.066
18 - 1	10000	26	44	F	34	2.891	4.169	F	3.105
18 - 2	6000	64	60	276	41	3.392	3.584	8.463	2.168
18 - 2	7000	130	142	F	F	8.355	8.793	F	F
18 - 3	5000	43	38	F	21	1.827	2.066	F	1.095
18 - 3	8000	1002	1017	277	523	99.845	71.096	8.07	37.918
18 - 4	9000	29	36	F	15	1.272	3.322	F	1.389
18 - 4	10000	42	44	235	25	3.5	4.102	7.07	2.515
19 - 1	5000	87	269	F	F	3.085	13.935	F	F
19 - 1	7000	87	269	F	F	3.951	13.476	F	F
19 - 2	6000	247	130	F	F	6.965	5.525	F	F
19 - 2	8000	245	105	F	F	11.264	3.617	F	F
19 - 3	9000	378	290	F	F	19.614	10.109	F	F
19 - 3	10000	378	324	F	F	12.946	10.138	F	F
19 - 4	5000	242	268	F	F	8.629	8.153	F	F
19 - 4	10000	242	268	F	F	7.609	9.501	F	F
20 - 1	5000	431	44	168	4	11.512	1.418	4.706	0.163
20 - 1	6000	440	46	169	4	10.216	1.377	4.517	0.181
20 - 2	7000	28	23	24	67	1.406	0.79	0.936	2.217
20 - 2	8000	29	23	24	67	1.723	0.803	0.752	2.35
20 - 3	9000	79	77	37	77	3.154	3.692	1.278	2.891
20 - 3	10000	79	77	37	77	3.385	3.921	1.553	3.102
20 - 4	7000	955	F	35	5	20.726	F	1.153	0.212
20 - 4	8000	962	F	38	5	19.185	F	1.412	0.204

The last inequality follows from Assumption 2.1, indicating that for each $x \in \omega$ and a positive constant M, $|f(x)| \leq M$. Therefore, for all $m \geq 0$,

$$0 < \sum_{n=0}^{\infty} \frac{(g_n^T d_n)^2}{||d_n||^2} \leq \frac{1}{\gamma}(f(x_0) + M) < \infty.$$

Thus, proving Lemma 2.2 □

Theorem 2.1 ([33]). *Suppose that Assumption 2.1 and Assumption 2.2 are satisfied, and consider any method satisfying Eqs. (2) and (3), where φ_n is obtained with the SW conditions (13) and (14). If*

$$\sum_{n=0}^{\infty} ||g_n||^2 < \infty, \tag{21}$$

then

$$\lim_{n\to\infty} \inf ||g_n|| = 0. \tag{22}$$

Proof. Suppose Eq. (22) is not true, then there exists a constant $u > 0$ such that:

$$||g_n|| \geq u, \forall n. \tag{23}$$

Using Eq. (10) with k given by Eq. (15), Eqs. (21) and (23), we have

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{u^4}{||d_n||^2} &\leq \sum_{n=0}^{\infty}\frac{||g_n||^4}{||d_n||^2},\\ &\leq \sum_{n=0}^{\infty}\frac{(1-\sigma)^2(g_n^T d_n)^2}{||d_n||^2},\\ &= \sum_{n=0}^{\infty}\frac{(1-\sigma)^2(g_n^T d_n)^2}{||d_n||^2},\\ &\leq \sum_{n=0}^{\infty}\frac{(1-\sigma)^2\,||g_n||^2\,||d_n||^2}{||d_n||^2},\\ &= (1-\sigma)^2\sum_{n=0}^{\infty}||g_n||^2\,,\\ &= \infty,\end{aligned}$$

which contradicts Lemma 2.2. Thus, $\lim_{n\to\infty}\inf||g_n|| = 0$. □

3. NUMERICAL RESULTS AND DISCUSSION

In this section, a report on the performance of the HDYDL method in comparison with the DY and DL methods is presented. To further validate the method's efficiency, it is compared to the NM method, a recently proposed Dai-Liao-type CGM given by Akinduko [21].

A total of 20 unconstrained optimization problems taken from Bongartz *et al.* [34] and Andrei [35] were solved. For each solved test problem (TP), four initial points and dimensions (Dim) ranging from 5000 to 10000 were considered, totaling 158 computations. The computations were carried out on a computer with the following specifications: 4 GB of RAM, 2.2 GHz processor speed, and the Windows 10 operating system. The basis of comparison includes the computational time (CPU) and the number of iterations. The iterations were terminated when $||g_n|| \leq 10^{-6}$ or the number of iterations went beyond 2000. The notation "F" is used to denote a failed iteration. For the step size computation, the SW search technique was used.

Table 1 presents the solved test problems and their initial points, while Tables 2–5 give the details of the numerical results. In the presentation of results in Tables 2–5, the form (**a** − **b**) is used in the column labeled TP, where TP represents a particular solved test problem, **a** denotes the serial number of the TP as it appears in Table 1, and **b** denotes the initial point accompanying **a** as it appears in Table 1. Figures 1 and 2 depict the CPU and the iteration profiles, respectively. This is based on the method of Dolan & More [36]. The y-axes of the Figures show the percentage of successfully solved problems, while the top curve represents the fastest method. Based on this fact, the percentages for the compared methods are recorded: 100% for HDYDL, 88.6% for DY, 70.3% for DL, and 74.7% for NM. These results demonstrate that the HDYDL method is the most successful of the four under consideration, solving every test problem regardless of the starting point. The DY and NM methods follow this, while the DL method lags in performance.

4. CONCLUSION

This work investigates the global convergence properties of the Dai-Liao-type CGM for unconstrained optimization problems. It looked into the convergence requirements, such as termination criteria, descent directions, and reduction of the objective function. This work adds to the increasing body of knowledge on optimization techniques by concentrating on the Dai-Liao-type CGM. It provides practitioners with useful information to help them select efficient strategies for solving unconstrained optimization problems. The numerical experiments serve as practical validation, bridging the gap between theory and real-world applicability. Comparing the HDYDL approach to the DY, DL, and NM methods, the results show how resilient and reliable the HDYDL method is in locating global optimal solutions.

ACKNOWLEDGMENT

The author appreciate the editors and the anonymous reviewers for their valuable comments and suggestions for the improvement of the paper.

References

[1] E. K. Chong & S. H. Zak, *An Introduction to Optimization*, Wiley Inter-Science Publication Inc, New York, 2001. https://epdf.tips/an-introduction-to-optimization.html.
[2] O. Olotu, C. Aladesaye & K. A. Dawodu, "Modified gradient flow method for solving one-dimensional optimal control problem governed by linear equality constraint", Journal of the Nigerian society of Physical Sciences **4** (2022) 146. https://doi.org/10.46481/jnsps.2022.589.
[3] T. T. Yusuf, A. Abidemi, A. S. Afolabi & E. J. Dansu, "Optimal control of the coronavirus pandemic with impacts on implemented control measures", Journal of the Nigerian Society of Physical Sciences **4** (2022) 88. https://doi.org/10.46481/jnsps.2022.414.
[4] R. Fletcher & C. M. Reeves, "Function minimization by conjugate gradients", Comput. J. **7** (1964) 149. https://doi.org/10.1093/COMJNL/7.2.149.
[5] M. R. Hestenes & E. Stiefel, "Methods of conjugate gradients for solving linear systems", J. Res. Natl. Bur. Stand. **49** (1952) 409. https://nvlpubs.nist.gov/nistpubs/jres/049/jresv49n6p409_a1b.pdf.
[6] Y. H. Dai & Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property", SIAM J. Optim. **10** (1999) 177. https://doi.org/10.1137/S1052623497318992.
[7] E. Polak & G. Ribiere, "Note Sur la Convergence de directions conjugees", ESAIM: Math. Model. Numer. Anal. **3** (1969) 35. https://doi.org/10.1051/M2AN/196903R100351.
[8] B. T. Polyak, "The conjugate gradient method in extreme problems", Comput. Math. Math. Phys. **9** (1969) 94. https://doi.org/10.1016/0041-5553(69)90035-4.
[9] J. C. Gilbert & J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization", SIAM Journal on Optimization **2** (1992) 21. https://doi.org/10.1137/0802003.
[10] Z. Wei, S. Yao & L. Liu, "The convergence properties of some new conjugate gradient methods", Applied Mathematics and Computation **183** (2006) 1341. https://doi.org/10.1016/j.amc.2006.05.150.
[11] X. Jiang & J. Jian, "A sufficient descent Dai-Yuan type nonlinear conjugate gradient method for unconstrained optimization problems", Nonlinear Dyn. **72** (2013) 101. https://doi.org/10.1007/s11071-012-0694-6.
[12] Y. Yueting & C. Mingyuan, "The global convergence of a new mixed conjugate gradient method for unconstrained optimization", Journal of Applied Mathematics **2012** (2012) 932980. https://doi.org/10.1155/2012/932980.
[13] J. Zhang, Y. Xiao & Z. Wei, "Nonlinear conjugate gradient methods with sufficient descent condition for large-scale unconstrained optimization", Mathematical Problems in Engineering **2009** (2009) 243290. https://doi.org/10.1155/2009/243290.
[14] S. S. Djordjevic, "New hybrid conjugate gradient method as a convex combination of HS and FR methods", Journal of Applied Mathematics and Computation **2** (2018) 366. https://doi.org/10.26855/jamc.2018.09.002.
[15] X. Jiang & J. Jian, "Improved Fletcher-Reeves and Dai-Yuan conjugate gradient methods with strong Wolfe line search", Journal of Computational and Applied Mathematics **348** (2019) 525. https://doi.org/10.1016/j.cam.2018.09.012.
[16] Y. H. Dai & L. Z. Liao, "New conjugacy conditions and related nonlinear conjugate gradient method", Appl. Math. Optim. **43** (2001) 87. https://doi.org/10.1007/s002450010019.

[17] S. Babaie-Kafaki & R. Ghanbari, "The Dai-Liao nonlinear conjugate gradient method with optimal parameter choices", Eur. J. Opt. Res. **234** (2014) 625. https://doi.org/10.1016/j.ejor.2013.11.012.

[18] S. Babaie-Kafaki & R. Ghanbari, "Two adaptive Dai-Liao nonlinear conjugate gradient methods", Iran J. Sci. Technol. Trans. Sci. **42** (2018) 1505. https://doi.org/10.1007/s40995-017-0271-4.

[19] N. Salihu, M. R. Odekunle, A. M. Saleh & S. Salihu, "A Dai-Liao hybrid Hestenes-Stiefel and Fletcher-Reeves methods for unconstrained optimization", Int. J. Ind. Optim. **2** (2021) 33. https://doi.org/10.12928/IJIO.V2I1.3054.

[20] M. Lofti & S. M. Hosseini, "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** (2019) 112708. https://doi.org/10.1016/j.cam.2019.112708.

[21] O. B. Akinduko, "A new conjugate gradient method with sufficient descent property", Earthline J. Math. Sci. **6** (2021) 163. https://doi.org/10.34198/EJMS.6121.163174.

[22] O. B. Onuoha, "A sufficient descent Dai-Liao type conjugate gradient update parameter", Earthline Journal of Mathematical Sciences **13** (2023) 353. https://doi.org/10.34198/ejms.13223.353368.

[23] A. H. Sheeko & G. M. Al-Naemi, "Global convergence condition for a new spectra conjugate gradient method for large-scale optimization", J. Phys. Conf. Ser. **1879** (2021) 032001. https://doi.org/10.1088/1742-6596/1879/3/032001.

[24] C. Ahmed & B. Taher, "Global convergence of new conjugate gradient method with inexact line search", International Journal of Electrical and Computer Engineering **11** (2021) 1469. https://doi.org/10.11591/ijece.v11i2.pp1469-1475.

[25] A. Alhawarat, G. Alhamzi, I. Masmali & Z. Salleh, "A descent four-term conjugate gradient method with global convergence properties for large-scale unconstrained optimization problems", Mathematical Problems in Engineering **2021** (2021) 6219062. https://doi.org/10.1155/2021/6219062.

[26] P. Wolfe, "Convergence conditions for ascent methods", SIAM J. Optim. (1969) 226. https://doi.org/10.1137/1011036.

[27] O. O. O. Yousif, M. A. Y. Mohammed, M. A. Saleh & M. K. Elbashir, "A criterion for the global convergence of conjugate gradient methods under strong Wolfe line search", Journal of King Saud University-Science **34** (2022) 102281. https://doi.org/10.1016/j.jksus.2022.102281.

[28] P. Kaelo, P. Mtagulwa & M. V. Thuto, "A globally convergent hybrid conjugate gradient method with strong Wolfe conditions for unconstrained optimization", Mathematical Sciences **14** (2020) 1. https://doi.org/10.1007/s40096-019-00310-y.

[29] S. liu & Y. Huang, "Several guaranteed descent conjugate gradient methods for unconstrained optimization", Journal of Applied Mathematics **2014** (2014) 825958. https://doi.org/10.1155/2014/825958.

[30] O. O. O. Yousif, "The convergence properties of RMIL+ conjugate gradient method under strong Wolfe line search", Appl. Math. Comp. **367** (2020) 124777. https://doi.org/10.1016/j.amc.2019.124777.

[31] A. Alhawarat, N. T. Trung & Z. Salleh, "Conjugate gradient method: a developed version to resolve unconstrained optimization problems", Journal of Computer Science **16** (2020) 1220. https://doi.org/10.3844/jcssp.2020.1220.1228.

[32] P. Mtagulwa & P. Kaelo, "An efficient modified PRP-FR hybrid conjugate gradient method for solving unconstrained optimization problems", Applied Numerical Mathematics **145** (2019) 111. https://doi.org/10.1016/j.apnum.2019.06.003.

[33] J. Nocedal & S. J. Wright, *Numerical Optimization*, Springer-Verlag Inc., New York, 1999. https://doi.org/10.1007/b98874 .

[34] I. Bongartz, A. R. Conn, N. I. M. Gould & P. L. Toint, "CUTE: Constrained and unconstrained testing environments", ACM Trans. Math. Softw. **21** (1995) 123. https://doi.org/10.1145/200979.201043.

[35] N. Andrei, "An unconstrained optimization test functions collection", Adv. Model. Optim. **10** (2008a) 147. https://camo.ici.ro/journal/vol10/v10a10.pdf.

[36] E. D. Dolan & J. J. More, "Benchmarking optimization software with performance profiles", Math. Program. **91** (2002) 201. https://doi.org/10.1007/s101070100623.