RAIRO-Oper. Res. 55 (2021) 3281–3291
https://doi.org/10.1051/ro/2021159

RAIRO Operations Research
www.rairo-ro.org

A NEW FAMILY OF DAI-LIAO CONJUGATE GRADIENT METHODS WITH MODIFIED SECANT EQUATION FOR UNCONSTRAINED OPTIMIZATION

Yutao Zheng*

Abstract. In this paper, a new family of Dai-Liao–type conjugate gradient methods are proposed for unconstrained optimization problem. In the new methods, the modified secant equation used in [H. Yabe and M. Takano, *Comput. Optim. Appl.* **28** (2004) 203–225] is considered in Dai and Liao's conjugacy condition. Under some certain assumptions, we show that our methods are globally convergent for general functions with strong Wolfe line search. Numerical results illustrate that our proposed methods can outperform some existing ones.

Mathematics Subject Classification. 65K05, 90C26, 90C30.

Received May 19, 2021. Accepted October 16, 2021.

1. Introduction

Consider the following unconstrained optimization problem

$$\min f(x),\ \ x \in \mathbb{R}^n, \tag{1.1}$$

where the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and its gradient $g(x)$ is available. The problem (1.1) has a wide range of applications in areas of scientific computing and engineering. Therefore, its efficient and effective numerical solution methods have been intensively studied in the literature, including the spectral gradient methods [5, 15], conjugate gradient methods [4, 13] and memoryless BFGS methods [16]. Among them, conjugate gradient methods are popular and efficient for solving (1.1), especially for large scale problems.

Let x_k be the kth iterate point, g_k the gradient of $f(x)$ at x_k, *i.e.* $g_k = g(x_k)$. The (nonlinear) conjugate gradient method is given by

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

where α_k is the step length computed by carrying out an one-dimension line search and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k, & \text{if } k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{if } k \geq 1, \end{cases} \tag{1.3}$$

Keywords. Conjugate gradient method, Dai-Liao–type method, modified secant equation.

School of Mathematics and Information Science, Henan Normal University, Xinxiang 453007, PR China.
* Corresponding author: zhengyutao@htu.edu.cn

where β_k is a scalar.

Since exact line search for searching α_k is usually expensive and impractical, the strong Wolfe inexact line search is often considered in the convergence analysis and implementation of nonlinear conjugate gradient methods. It aims to find a step size α_k satisfying the following two strong Wolfe conditions

$$f\big(x_k + \alpha_k d_k\big) \le f\big(x_k\big) + \rho g_k^T d_k, \tag{1.4}$$
$$|g_{k+1}^T d_k| \le \sigma |g_k^T d_k|, \tag{1.5}$$

where $0 < \rho < \sigma < 1$.

Nonlinear conjugate gradient method for unconstrained optimization problem is generated from the linear conjugate gradient method for a special quadratic minimization problem

$$\min \frac{1}{2} x^T Q x + b^T x + c$$

or its equivalent line system $Qx = b$, where Q is a real symmetric positive definite matrix. Linear conjugate gradient methods generate a search direction such that the conjugacy condition holds, namely,

$$d_i^T Q d_j = 0, \forall i \ne j. \tag{1.6}$$

For general nonlinear functions, it follows from the mean value theorem that there exists some $\tau \in (0,1)$ such that

$$d_k^T y_{k-1} = \alpha_{k-1} d_k^T \nabla^2 f\big(x_{k-1} + \tau \alpha_{k-1} d_{k-1}\big) d_{k-1},$$

where $y_{k-1} = g_k - g_{k-1}$ denotes the gradient change. Therefore, it is reasonable to replace (1.6) by the following conjugacy condition:

$$d_k^T y_{k-1} = 0. \tag{1.7}$$

Let $d_k = -g_k + \beta_k d_{k-1}$ satisfy the above condition, we have the famous Hestenes-Stiefel formula [14]

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{y_{k-1}^T d_{k-1}}.$$

In 2001, Dai and Liao [6] suggested an extended one

$$d_k^T y_{k-1} = t g_k^T s_{k-1}, \tag{1.8}$$

which leads to the following conjugate gradient parameter

$$\beta_k^{\mathrm{DL+}} = \max\left\{ \frac{g_k^T y_{k-1}}{y_{k-1}^T d_{k-1}}, 0 \right\} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}, \tag{1.9}$$

where $t > 0$ is a scalar, $s_{k-1} = x_k - x_{k-1}$. Note that the first item has been restricted to be nonnegative like [10]. The DL+ method (1.2)–(1.3) with β_k in (1.9) is globally convergent for general functions under the sufficient descent condition

$$g_k^T d_k \le -c\|g_k\|^2, \ c > 0. \tag{1.10}$$

and some other suitable conditions, where and hereafter $\|\cdot\|$ denotes the Euclidean norm of vectors.

As a special case of Dai-Liao–type conjugate gradient method, the efficient CG_descent method [12] utilizes a particular t. The conjugacy parameter of CG_descent method is

$$\beta_k^N = \frac{g_k^T y_{k-1}}{y_{k-1}^T d_{k-1}} - 2\frac{\|y_{k-1}\|^2}{s_{k-1}^T y_{k-1}} \cdot \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}.$$

Two further developments of the Dai and Liao's method were made by Yabe and Takano [20] and Li *et al.* [17] based on different modified secant equations. Some more efficient Dai-Liao–type methods were designed and studied in [2, 9, 22, 23] by using different techniques. In this paper, we further give a new family of Dai-Liao–type conjugate gradient methods for unconstrained optimization problems, including their convergence analysis. Numerical experiments show that our methods can outperform the existing ones.

The rest of this paper is organized as follows. In Section 2, we introduce a new Dai-Liao–type method by modifying the conjugate gradient parameter. Based on the strong Wolfe line search rules, the global convergence for uniformly convex and general objective functions is studied in Section 3 and numerical experiments are performed in Section 4. Finally, in Section 5, we give some conclusions to end this paper.

2. New Dai-Liao–Type methods

We start with the original Dai and Liao's method in which the quasi-Newton techniques are used. In the quasi-Newton method, an approximation Hessian B_k is updated such that

$$B_k s_{k-1} = y_{k-1}$$

and the search direction d_k is calculated by

$$d_k = -B_k^{-1} g_k. \tag{2.1}$$

Combining the above two equations, we have

$$d_k^T y_{k-1} = -g_k^T s_{k-1}.$$

The above relation implies that (1.7) holds in case of $g_k^T d_{k-1} = 0$, *i.e.* the line search is exact. However, in practical numerical algorithms, the inexact line search is adopted instead of exact line search. Dai and Liao suggested the following conjugacy condition:

$$d_k^T y_{k-1} = -t g_k^T s_{k-1},\ \ (t \geq 0).$$

In 2004, Yabe and Takano [20] used the modified secant equation

$$B_{k+1} s_k = y_k + \rho_k \theta_k \frac{u_k}{u_k^T s_k}, \tag{2.2}$$

where $\rho_k \in [0,3]$ and $\theta_k = 2(f_k - f_{k+1}) + (g_k + g_{k+1})^T s_k$, u_k is chosen s.t. $u_k^T s_k \neq 0$, to derive a new conjugacy condition through replacing y_k by $z_k = y_k + \rho_k \theta_k \frac{u_k}{u_k^T s_k}$, the modified conjugacy parameter is

$$\beta_{k+1}^{YT+} = \max\left\{ \frac{g_{k+1}^T z_k}{d_k^T z_k}, 0 \right\} - t \frac{g_{k+1}^T s_k}{d_k^T z_k}.$$

In this paper, we will derive a new conjugacy condition from another view of point. Combining (2.2) with (2.1), we have

$$\begin{aligned} d_{k+1}^T y_k &= d_{k+1}^T \Big(B_{k+1} s_k - \rho_k \theta_k \frac{u_k}{u_k^T s_k} \Big) \\ &= d_{k+1}^T B_{k+1} s_k - \rho_k \theta_k \frac{u_k^T d_{k+1}}{u_k^T s_k} \\ &= -g_{k+1}^T s_k - \rho_k \theta_k \frac{u_k^T d_{k+1}}{u_k^T s_k}. \end{aligned}$$

Using the Dai-Liao's conjugacy condition

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k$$

and $d_{k+1} = -g_{k+1} + \beta_{k+1} d_k$, where $t \in [0,1]$, we have

$$\rho_k \theta_k \frac{u_k^T(-g_{k+1} + \beta_{k+1} d_k)}{u_k^T s_k} = (t-1) g_{k+1}^T s_k,$$

which yields a new conjugate gradient parameter

$$\begin{aligned}\beta_{k+1}^{\text{new}} &= \left[\frac{(t-1) g_{k+1}^T s_k \cdot u_k^T s_k}{\rho_k \theta_k} + g_{k+1}^T u_k\right] \Big/ d_k^T u_k \\ &= \frac{g_{k+1}^T u_k}{d_k^T u_k} - (1-t)\frac{u_k^T s_k}{\rho_k \theta_k} \cdot \frac{g_{k+1}^T s_k}{d_k^T u_k}\end{aligned}$$

if $\rho_k \theta_k \neq 0$, otherwise, Dai and Liao's conjugate gradient parameter $\beta_k^{\text{DL+}}$ will be used. According to the experience of the quasi-Newton methods with modified secant equations [21], we choose $u_k = y_k$.

In the case of $u_k = y_k$, the conjugacy parameter β_{k+1}^{new} can be written as

$$\begin{aligned}\beta_{k+1}^{\text{new}} &= \left[\frac{(t-1) g_{k+1}^T s_k \cdot y_k^T s_k}{\rho_k \theta_k} + g_{k+1}^T y_k\right] \Big/ d_k^T y_k \\ &= \frac{g_{k+1}^T y_k}{d_k^T y_k} + \frac{(t-1) y_k^T s_k}{\rho_k \theta_k} \cdot \frac{g_{k+1}^T s_k}{d_k^T y_k}. \end{aligned} \tag{2.3}$$

and we correct it as

$$\beta_{k+1}^{\text{new+}} = \max\left\{\frac{g_{k+1}^T y_k}{d_k^T y_k}, 0\right\} + \frac{(t-1) y_k^T s_k}{\rho_k |\theta_k|} \cdot \frac{g_{k+1}^T s_k}{d_k^T y_k}. \tag{2.4}$$

We call the method (1.2) and (1.3) with β_k given in (2.4) NEW+ method. The corresponding algorithm is given as below:

Algorithm 2.1. Improved Dai-Liao conjugate gradient method

Step 1: Given $x_0 \in R^n$, $\varepsilon, \eta > 0$, set $d_0 = -g_0$, $k := 0$; if $\|g_0\| \leq \varepsilon$, then stop;

Step 2: Compute α_k such that strong Wolfe line search (1.4) and (1.5) hold;

Step 3: Let $x_{k+1} = x_k + \alpha_k d_k$, if $\|g_{k+1}\| \leq \varepsilon$, then stop;

Step 4: Compute β_{k+1} by (2.4) if $|\theta_k| > \eta$, otherwise, compute β_{k+1} by (1.9); generate d_{k+1} by (1.3);

Step 5: Set $k := k+1$ and go to Step 2.

In the rest of the paper, we first analyze the convergence properties of the new algorithm, then give some numerical results which show the modified algorithms are robust and efficient.

3. Convergence Analysis

Throughout this section, we assume that $g_k \neq 0$ for all $k \geq 0$, otherwise a stationary point is found. We first give some standard assumptions.

Assumption 3.1. *The level set* $\mathcal{L} = \{x \in \mathbb{R}^n : f(x) \le f(x_0)\}$ *is bounded, where* $x_0 \in \mathbb{R}^n$ *is an initial point.*

Assumption 3.2. *In some neighborhood* $\mathcal{N}$ *of* $\mathcal{L}$*, the function* f *is continuously differentiable and its gradient* $g(x)$ *is Lipschitz continuous, i.e. there exists a positive constant* $L > 0$ *such that* $\|g(x) - g(y)\| \le L\|x - y\|$ *for all* $x, y \in \mathcal{N}$.

Assumption 3.1 guarantees that there exists some constant c such that $\|s_k\| \le 2\bar{c}, \forall k > 0$. Assumption 3.2 implies that $\|g\| \le \bar{\gamma}$ for any $x \in \mathcal{L}$, where $\bar{\gamma} = 2\bar{c}L + \|g_0\|$.

Firstly, we give some estimation on θ_k. We know by mean value theorem that

$$\begin{aligned}\theta_k &= 2(f_k - f_{k+1}) + (g_k + g_{k+1})^T s_k \\ &= -2\nabla f(\eta_k)^T s_k + (g_k + g_{k+1})^T s_k \\ &= -[\nabla f(x_k) - \nabla f(\eta_k) + \nabla f(x_{k+1}) - \nabla f(\eta_k)]^T s_k,\end{aligned}$$

where $\eta_k = x_k + \tau(x_{k+1} - x_k)$ and $\tau \in (0,1)$. Hence

$$\begin{aligned}|\theta_k| &\le (\|\nabla f(x_k) - \nabla f(\eta_k)\| + \|\nabla f(x_{k+1}) - \nabla f(\eta_k)\|)\|s_k\| \\ &\le L(\|x_k - \eta_k\| + \|x_{k+1} - \eta_k\|)\|s_k\| \\ &= L\|s_k\|^2 \le 4L\bar{c}^2.\end{aligned}$$

On the other hand, since θ_k is appeared in the denominator, too small value must be avoided for the numerical stability, we ask $|\theta_k|$ to satisfy $0 < \eta \le |\theta_k|$ as shown in Algorithm 2.1. Otherwise, $\beta_k^{\mathrm{DL+}}$ will be used.

Let f be a uniformly convex function, then there exists some constant $\mu > 0$ such that

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

which implies

$$\mu\|s_k\|^2 \le s_k^T y_k \le L\|s_k\|^2. \tag{3.1}$$

Then we have that

$$\begin{aligned}\theta_k &= 2\Big(f_k - f_{k+1}\Big) + \Big(g_k + g_{k+1}\Big)^T s_k \\ &\ge \Big(-g_{k+1}^T s_k + \frac{\mu}{2}\|s_k\|^2\Big) + \Big(g_k + g_{k+1}\Big)^T s_k \\ &= -s_k^T y_k + \mu\|s_k\|^2 \\ &\ge -\Big(1 - \frac{\mu}{L}\Big) s_k^T y_k\end{aligned}$$

and

$$|\theta_k| \le L\|s_k\|^2 \le \frac{L}{\mu} s_k^T y_k.$$

Thus θ_k locates in the interval $\Big(-(1-\mu/L)s_k^T y_k, -\eta\Big) \cup \Big(\eta, L/\mu s_k^T y_k\Big)$. Therefore, we assume that the following relationship always holds.

$$\frac{\mu}{L} \le \left|\frac{s_k^T y_k}{\theta_k}\right| \le \frac{1}{\epsilon}.$$

The following theorem states the global convergence property of new method for uniformly convex functions.

Theorem 3.3. *Let* f *be a uniformly convex function and Assumptions 3.1 and 3.2 hold. Suppose that* $\{x_k\}$ *is the sequence generated by Algorithm 2.1 with* β_k *in* (2.4)*, then*

$$\lim_{k\to\infty} \|g_k\| = 0. \tag{3.2}$$

Proof. It follows from f is uniformly convex function that

$$d_k^T y_k \geq \mu\alpha_{k-1}\|d_{k-1}\|^2.$$

By using Triangular and Cauchy-Schwartz inequalities, we have

$$\begin{aligned}\|d_k\| &\leq \|g_k\| + |\beta_k^{\text{new}}|\|d_{k-1}\| \\ &\leq \|g_k\| + \frac{(L+(1-t)/\epsilon)\,\|g_k\|\,\|s_{k-1}\|}{\mu\alpha_{k-1}\,\|d_{k-1}\|^2}\,\|d_{k-1}\| \\ &\leq \mu^{-1}(L+(1-t)/\epsilon+\mu)\|g_k\|,\end{aligned}$$

which means

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

Therefore, from Lemma 3.2 [6] and the fact f is a uniformly convex function, we have

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

□

For the general function, we only need to show the modified Dai-Liao method with β_k in (2.4) satisfies the Property(*) depicted by Gilbert and Nocedal [10]. The rest analysis is similar to the original Dai-Liao's method.

Definition 3.4. Consider a method of the form (1.2)–(1.3), and suppose that

$$0 < \gamma \leq \|g_k\| \leq \bar{\gamma} \tag{3.3}$$

for all $k \geq 1$. We say that the conjugate gradient method has the Property($*$), if for all k, there exist constants $b > 1, k > 0$ such that for all k,

$$|\beta_k| \leq b \text{ and } \|s_{k-1}\| \leq \lambda \text{ imply } |\beta_k| \leq \frac{1}{2b}. \tag{3.4}$$

By the strong Wolfe condition (1.5), (1.10) and (3.3), we have

$$d_{k-1}^T y_{k-1} \geq (\sigma-1)g_{k-1}^T d_{k-1} \geq (1-\sigma)c\gamma^2.$$

Using this and boundedness of $\|s_k\|$, we obtain

$$|\beta_k| \leq \frac{(L+(1-t)/\epsilon)\,\|g_k\|\,\|s_{k-1}\|}{(1-\sigma)c\gamma^2} \leq \frac{2(L+(1-t)/\epsilon)\bar{\gamma}\bar{c}}{(1-\sigma)c\gamma^2} =: b.$$

Note that b can be defined such that $b > 1$. If we set

$$\lambda := \frac{(1-\sigma)c\gamma^2}{b(L+(1-t)/\epsilon)\bar{\gamma}}$$

and $s_{k-1} \leq \lambda$, then

$$|\beta_k| \leq \frac{(L+(1-t)/\epsilon)\bar{\gamma}\lambda}{(1-\sigma)c\gamma^2} = \frac{1}{b}.$$

Therefore, the NEW+ method has Property (*). Thus, we have the following convergence theorem.

Theorem 3.5. *Let Assumptions 3.1 and 3.2 hold. If the sequence $\{x_k\}$ is generated by the NEW+ method with the strong Wolfe line search for $\sigma \in (0,1)$, where d_k satisfies condition (1.10) with $c > 0$. Then we have*

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

TABLE 1. Tested conjugate gradient algorithms.

β_k	Name of method	Abbreviation
β_k^{FR}	The Fletcher-Reeves method [8]	FR
β_k^{PRP+}	The Polak-Ribiere-Polyak [18, 19]	PRP+
β_k^{HS}	The Hestenes-Stiefel method [14]	HS
β_k^{DY}	The Dai-Yuan method [3]	DY
β_k^{N}	The Hager and Zhang's method [12]	HZ
β_k^{DL+}	The Dai-Liao method [6]	DL+
β_k^{YT+}	Yabe-Takano's method [20]	YT+
β_k^{new+}	Our method	NEW+

TABLE 2. Numerical results for Trigonometric with $n = 5000$.

FR	PRP+	HS	DY	HZ	CG_descent
167/292	133/340	39/109	213/291	96/195	73/278

	ρ_k	t = 0.1	0.3	0.5	0.7	0.9
DL+	0	21/100	49/136	20/92	64/156	69/158
YT+	0.1	21/100	82/177	20/92	61/152	63/151
NEW+	0.1	78/169	61/143	52/136	**19/91**	66/156
YT+	0.3	20/99	67/160	20/93	**21/95**	63/153
NEW	0.3	**20/94**	20/93	20/91	65/157	67/161
YT+	0.5	**19/95**	21/91	42/134	41/127	38/127
NEW+	0.5	20/92	**19/88**	**19/89**	64/155	71/162
YT+	0.7	22/107	21/93	67/154	69/162	23/100
NEW+	0.7	22/108	45/135	20/91	64/152	64/152
YT+	0.9	20/95	68/166	21/99	38/127	33/1
NEW+	0.9	20/97	47/140	20/98	67/158	**19/94**
YT+	1.0	21/97	**20/90**	21/95	40/132	22/95
NEW+	1.0	20/97	44/128	**19/89**	65/153	**19/95**

4. Numerical Experiments

In this section, some numerical results are reported on a a set of 76 unconstrained optimization problems selected from [1] and CUTEst library [11]. We tested the conjugate gradient algorithms with the conjugacy parameters given in Table 1.

For the algorithms DL+, YT+ and our new method, different scaled parameters ρ and t are used. In the case where an ascent direction is generated, we restart the algorithm by setting $d_k = -g_k$.

All codes were written in Fortran and in double precision arithmetic. (Note that, for the sake of fairness, at the beginning of experiments we do not directly run Hager and Zhang's CG_descent codes for the test problems, we just use their conjugate parameter under the same linear search in our test framework). The stopping rule is set as $\|g_k\|_\infty \leq 10^{-6}$. The iteration is also terminated if the total number of iterations exceeds 10,000. Partial numerical results are summarized in Tables 2–6 and given in the form of (number of iterations/number of function-gradient evaluations), the detailed complete numerical results can be downloaded from the website https://github.com/piratetwo/mdl.

TABLE 3. Numerical results for ENGVAL1 with $n = 5000$.

FR	PRP+	HS	DY	HZ	CG_descent
285/8484	195/5333	695/20246	1711/53963	229/6770	**24/78**

	ρ_k	t = 0.1	0.3	0.5	0.7	0.9
DL+	0	117/3068	**33/364**	**24/72**	105/2642	161/4520
YT+	0.1	101/2570	224/6557	145/3867	105/2774	**43/704**
NEW+	0.1	189/5398	162/4586	93/2162	197/5752	269/7994
YT+	0.3	166/4696	185/5189	184/5306	112/2886	181/5142
NEW+	0.3	228/6821	200/6005	307/9366	137/3683	255/7464
YT+	0.5	179/5123	171/4828	157/4448	140/3857	136/3666
NEW+	0.5	162/4693	164/4715	170/4931	**78/1778**	144/3926
YT+	0.7	58/1225	79/1805	81/1896	217/6311	164/4696
NEW+	0.7	**46/805**	253/7499	126/3339	157/4211	128/3529
YT+	0.9	200/6064	127/3524	60/1233	149/4241	259/7689
NEW+	0.9	**37/431**	177/4856	**72/1552**	234/6807	**116/2866**
YT+	1.0	120/3187	159/4454	109/2979	282/8316	145/4029
NEW+	1.0	94/2333	**62/1125**	120/3091	**102/2489**	119/3154

TABLE 4. Numerical results for Raydan 1 with $n = 5000$.

FR	PRP+	HS	DY	HZ	CG_descent
–/–	–/–	756/1019	739/819	856/1334	490/1472

	ρ_k	t = 0.1	0.3	0.5	0.7	0.9
DL+	0	816/1077	782/1082	724/972	766/1046	812/1101
YT+	0.1	834/1274	770/1041	766/1077	**540/711**	767/1001
NEW+	0.1	850/1166	689/933	808/1092	**649/853**	785/1037
YT+	0.3	809/1109	749/1038	782/1063	810/1122	863/1140
NEW+	0.3	839/1142	**628/858**	748/1010	791/1063	864/1154
YT+	0.5	752/1032	744/1015	749/1028	724/964	829/1135
NEW+	0.5	764/1015	732/1011	779/1058	710/964	801/1090
YT+	0.7	**622/853**	**679/915**	826/1144	842/1114	789/1065
NEW+	0.7	719/970	736/992	**704/953**	696/940	820/1081
YT+	0.9	712/974	824/1094	882/1224	777/1045	890/1187
NEW+	0.9	**671/891**	870/1151	786/1062	851/1179	**746/991**
YT+	1.0	708/957	859/1183	**669/897**	876/1176	815/1103
NEW+	1.0	722/969	715/972	759/1023	706/948	**680/905**

In the Tables 2–6, the boldface font is used to mark the first and second efficient method which performs better than the other two algorithms for each ρ_k and t. The number of the best performance for Algorithm NEW+, YT+, DL+ are 32, 15 and 3, respectively.

In most cases, our new method improves Yabe-Takano's method. For a special $\rho_k = 3$, which was used in the modified quasi-Newton method, we compare the numerical performance of YT+ and NEW+. We run the codes with different $t = 0.1, 0.2, \ldots, 1$ and compute the medians for each problem. The performance profiles introduced by Dolan and Moré [7] are used to display the behaviours of these two methods. Figure 1 shows that the NEW+ method performs the best result regarding the number of iterations and function-gradient evaluations, which is located at the top curve in Figure 1.

TABLE 5. Numerical results for SINQUAD with $n = 5000$.

FR	PRP+	HS	DY	HZ	CG_descent
1455/2665	–/–	–/–	1891/2579	2021/4044	1071/3813

	ρ_k	t = 0.1	0.3	0.5	0.7	0.9
DL+	0	481/1014	681/1424	728/1532	**475/1018**	588/1238
YT+	0.1	602/1270	**499/1043**	792/1676	703/1515	**390/874**
NEW+	0.1	550/1172	530/1145	**465/1021**	**451/965**	606/1282
YT+	0.3	533/1121	568/1225	1431/2951	477/1013	526/1110
NEW+	0.3	**482/1009**	721/1507	638/1325	599/1277	597/1249
YT+	0.5	528/1095	579/1272	544/1210	687/1474	553/1185
NEW+	0.5	532/1132	**491/1057**	804/1682	503/1059	500/1094
YT+	0.7	588/1295	842/1773	**496/1048**	555/1197	499/1102
NEW+	0.7	533/1105	543/1195	565/1189	452/1158	**476/1012**
YT+	0.9	546/1212	522/1088	610/1275	598/1287	568/1235
NEW+	0.9	483/1029	594/1280	502/1068	489/1046	644/1358
YT+	1.0	**423/873**	590/1272	538/1162	498/1048	483/1011
NEW+	1.0	593/1275	526/1119	567/1182	569/1204	506/1064

TABLE 6. Numerical results for Woods with $n = 5000$.

FR	PRP+	HS	DY	HZ	CG_descent
112/218	65/133	87/214	95/178	161/309	182/683

	ρ_k	t = 0.1	0.3	0.5	0.7	0.9
DL+	0	39/80	52/99	43/85	45/90	48/94
YT+	0.1	**32/68**	37/76	63/122	59/113	46/90
NEW+	0.1	38/77	**37/75**	70/123	43/90	44/89
YT+	0.3	48/93	40/101	53/103	58/116	61/114
NEW+	0.3	39/83	44/86	48/97	46/96	**36/74**
YT+	0.5	32/69	42/83	41/80	81/161	58/116
NEW+	0.5	62/113	**34/72**	50/100	**36/78**	53/103
YT+	0.7	35/74	49/99	40/81	58/114	45/91
NEW+	0.7	35/76	51/103	**36/70**	69/132	58/109
YT+	0.9	**31/66**	44/86	47/93	38/79	**39/79**
NEW+	0.9	40/83	49/97	57/108	**36/74**	59/99
YT+	1.0	41/83	41/83	36/74	39/79	59/118
NEW+	1.0	48/95	60/113	**32/67**	46/91	41/84

Finally, we run Hager and Zhang's method with the approximate Wolfe line search conditions (Hager and Zhang's CG_descent Fortran code Version 1.4[2]). From Figure 2, for about 62% of all problems, CG_descent needs the least iterations, it has the best performance. However, CG_descent has the poorer performance than the YT+ and NEW+ regarding the number of function-gradient evaluations which mainly affects the efficiency of the methods.

[2]https://people.clas.ufl.edu/hager/software/

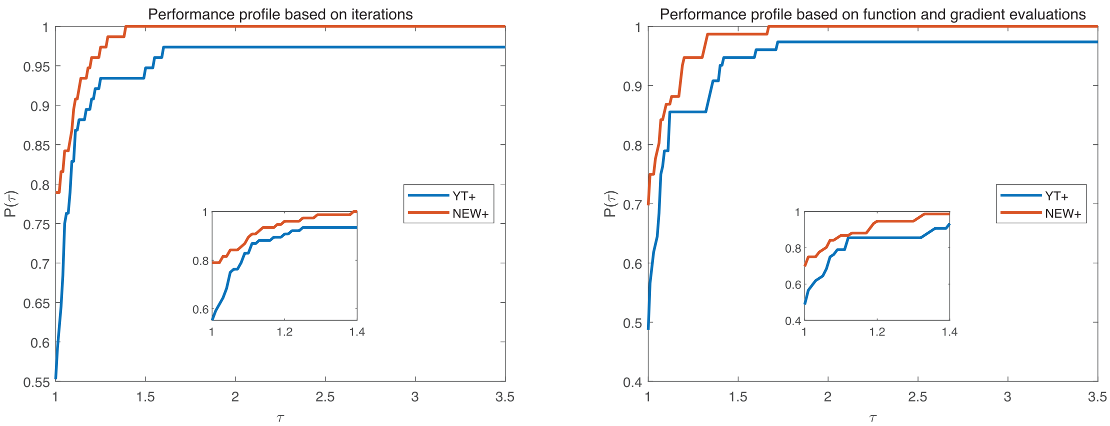

FIGURE 1. Performance profiles based on iterations and function-gradient evaluations for YT+ and NEW+.

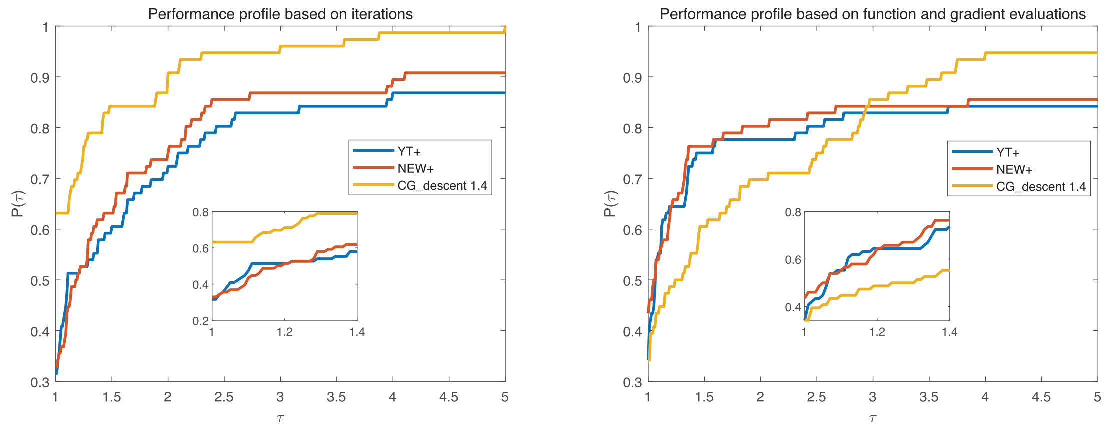

FIGURE 2. Performance profiles based on iterations and function-gradient evaluations for YT+, NEW+ and CG_descent.

5. Conclusions

In this paper, based on the Dai and Liao's conjugacy condition and the modified secant condition proposed by Zhang and Xu [21], we derived a new family of Dai-Liao–type conjugate gradient methods. Under some certain assumptions, we show that our methods are globally convergent for general functions. Numerical results show that our new methods can outperform some existing ones.

Acknowledgements. This work was supported by the Natural Science Foundation of Henan Province (Grant No. 202300410236) and the Training Program Foundation for Youth of Henan Normal University (Grant No. 2020PL28).

References

[1] N. Andrei, An unconstrained optimization test functions collection. *Adv. Model. Optim.* **10** (2008) 147–161.

[2] Y. Cheng, Q. Mou, X. Pan and S. Yao, A sufficient descent conjugate gradient method and its global convergence. *Optim. Methods Softw.* **31** (2016) 577–590.

[3] Y. Dai and Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.* **10** (1999) 177–182.

[4] Y. Dai, Nonlinear conjugate gradient methods, in Wiley Encyclopedia of Operations Research and Management Science (2011). DOI: 10.1002/9780470400531.eorms0183

[5] Y. Dai, Y. Huang and X. Liu, A family of spectral gradient methods for optimization. *Comput. Optim. Appl.* **74** (2019) 43–65.

[6] Y. Dai and L. Liao, New conjugacy conditions and related nonlinear conjugate gradient methods. *Appl. Math. Optim.* **43** (2001) 87–101.

[7] E. Dolan and J. Moré, Benchmarking optimization software with performance profiles. *Math. Program.* **91** (2002) 201–213.

[8] R. Fletcher, Function minimization by conjugate gradients. *Comput. J.* **7** (1964) 149–154.

[9] J. Ford, Y. Narushima and H. Yabe, Multi-step nonlinear conjugate gradient methods for unconstrained minimization. *Comput. Optim. Appl.* **40** (2008) 191–216.

[10] J. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim.* **2** (1992) 21–42.

[11] N. Gould, D. Orban and Ph. Toint, CUTEst: a Constrained and Unconstrained Testing Environment with safe threads for mathematical optimization. *Comput. Optim. Appl.* **60** (2015) 545–557.

[12] W. Hager and H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM J. Optim.* **16** (2005) 170–192.

[13] W. Hager and H. Zhang. A survey of nonlinear conjugate gradient methods. *Pacific J. Optim.* **2** (2006) 35–58.

[14] M. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.* **49** (1952) 409–436.

[15] Y. Huang, Y. Dai, X. Liu and H. Zhang, Gradient methods exploiting spectral properties. *Optim. Methods Softw.* **35** (2020) 681–705.

[16] C. Kou and Y. Dai, A modified self-scaling memoryless Broyden–Fletcher–Goldfarb–Shanno method for unconstrained optimization. *J. Optim. Theory Appl.* **165** (2015) 209–224.

[17] G. Li, C. Tang and Z. Wei, New conjugacy condition and related new conjugate gradient methods for unconstrained optimization. *J. Comput. Appl. Math.* **202** (2007) 523–539.

[18] E. Polak and G. Ribiere, Note sur la convergence de méthodes de directions conjuguées. *ESAIM: M2AN* **3** (1969) 35–43.

[19] B. Polyak, The conjugate gradient method in extremal problems. *USSR Comput. Math. Math. Phys.* **9** (1969) 94–112.

[20] H. Yabe and M. Takano, Global convergence properties of nonlinear conjugate gradient methods with modified secant condition. *Comput. Optim. Appl.* 28 (2004) 203–225.

[21] J. Zhang and C. Xu, Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations. *J. Comput. Appl. Math.* **137** (2001) 269–278.

[22] K. Zhang, H. Liu and Z. Liu, A new Dai-Liao conjugate gradient method with optimal parameter choice. *Numer. Funct. Anal. Optim.* **40** (2019) 194–215.

[23] Y. Zheng and B. Zheng, Two new Dai-Liao–type conjugate gradient methods for unconstrained optimization problems. *J. Optim. Theory Appl.* **175** (2017) 502–509.