Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2020-0056

Published online 4 November 2020

# A Linear Hybridization of Dai-Yuan and Hestenes-Stiefel Conjugate Gradient Method for Unconstrained Optimization

Sindhu Narayanan and P. Kaelo*

*Department of Mathematics, University of Botswana, Private Bag UB00704, Gaborone, Botswana*

Received 4 April 2020; Accepted (in revised version) 19 October 2020

**Abstract.** Conjugate gradient methods are interesting iterative methods that solve large scale unconstrained optimization problems. A lot of recent research has thus focussed on developing a number of conjugate gradient methods that are more effective. In this paper, we propose another hybrid conjugate gradient method as a linear combination of Dai-Yuan (DY) method and the Hestenes-Stiefel (HS) method. The sufficient descent condition and the global convergence of this method are established using the generalized Wolfe line search conditions. Compared to the other conjugate gradient methods, the proposed method gives good numerical results and is effective.

**AMS subject classifications**: 90C06, 90C30, 65K05

**Key words**: Unconstrained optimization, conjugate gradient method, global convergence.

## 1. Introduction

Consider the unconstrained optimization problem

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

where $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is a continuously differentiable function that is bounded below. There are several numerical methods for solving the unconstrained optimization problem (1.1). These include conjugate gradient methods [1, 6, 9, 11, 19], Newton methods [18, 39], quasi-Newton methods [8, 12, 33, 37, 42, 43] and steepest descent methods [7, 22, 26, 49]. These methods are iterative, that is, given an initial guess $x_0 \in \mathbb{R}^n$, they generate a sequence $\{x_k\}$ using

*Corresponding author. *Email address:* `kaelop@ub.ac.bw` (P. Kaelo)

 

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

where $\alpha_k > 0$ is a step length and $d_k$ is a descent direction. They also differ according to how the search direction $d_k$ is obtained or updated. Newton and quasi-Newton methods require the second derivative information for updating the direction $d_k$ and hence have good convergence rate. Conjugate gradient and steepest descent methods only require the first derivative information, which makes them more applicable to solving large-scale optimization problems.

The step length $\alpha_k > 0$ is chosen to satisfy certain line search conditions. Two of the usually used line searches are the strong Wolfe conditions

$$\begin{cases} f(x_k + \alpha_k d_k) \le f(x_k) + \sigma \alpha_k g_k^T d_k, \\ \left| g(x_k + \alpha_k d_k)^T d_k \right| \le \sigma_1 \left| g_k^T d_k \right|, \end{cases} \tag{1.3}$$

and the weak Wolfe conditions

$$\begin{cases} f(x_k + \alpha_k d_k) \le f(x_k) + \sigma \alpha_k g_k^T d_k, \\ g(x_k + \alpha_k d_k)^T d_k \ge \sigma_1 g_k^T d_k, \end{cases} \tag{1.4}$$

where $0 < \sigma < \sigma_1 < 1$.

In this paper, we consider solving problem (1.1) using a conjugate gradient method. Conjugate gradient methods generate the next iterate $x_{k+1}$ by updating the direction $d_k$ as

$$d_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k d_{k-1}, & k \ge 1, \end{cases} \tag{1.5}$$

where $g_k = \nabla f(x_k)$ is the gradient of the function $f$ at $x_k$, and $\beta_k \in \mathbb{R}$ is a parameter known as the conjugate gradient coefficient. Different choices of $\beta_k$ lead to different conjugate gradient methods, with the most well known methods being the Fletcher-Reeves (FR) [17], Polak-Ribière-Polyak (PRP) [36, 38], conjugate descent (CD) [16], Dai-Yuan (DY) [10], Liu-Storey (LS) [29] and Hestenes-Stiefel (HS) [21]. The PRP, LS and HS conjugate gradient methods have been shown to be numerically efficient while the others are theoretically effective. Other conjugate gradient methods have also been suggested in the literature [5,13,24,28,30,34,35,47] and a number of them are either modifications or hybridizations of the above methods. For instance, Wei *et al.* [44] proposed a conjugate gradient method

$$\beta_k^{WYL} = \frac{g_k^T \left( g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1} \right)}{\|g_{k-1}\|^2}, \tag{1.6}$$

which is a modification of the PRP method. It is globally convergent under weak Wolfe line search and numerically better than the PRP method. Similarly, Yao *et al.* [46] extended the above modification to HS method, that is,

$$\beta_k^{VHS} = \frac{g_k^T \left( g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1} \right)}{d_{k-1}^T y_{k-1}}, \tag{1.7}$$

where $y_{k-1} = g_k - g_{k-1}$, and showed it generates sufficient descent directions.

Another method was proposed by Li and Zhao [25] where

$$\beta_k^{P-W} = \max\left\{\beta_k^{PRP}, \beta_k^{WYL}\right\}, \tag{1.8}$$

which is a hybridization of the PRP and WYL methods. This method is also globally convergent under weak Wolfe conditions and is numerically effective. Liu [27] modified $\beta^{CD}$ as

$$\beta_k = \beta^{CD} + \min\left\{0, \varphi_k.\beta^{CD}\right\}, \tag{1.9}$$

where

$$\varphi_k = -\frac{g_k^T d_{k-1}}{d_{k-1}(g_k - g_{k-1})}. \tag{1.10}$$

The global convergence and numerical experiments of this method are done under generalized Wolfe conditions

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \sigma\alpha_k g_k^T d_k, \\ \sigma_1 g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k \leq -\sigma_2 g_k^T d_k, \end{cases} \tag{1.11}$$

where $0 < \sigma < \sigma_1 < 1$ and $\sigma_2 \geq 0$.

In 2008 and 2009, Andrei proposed two hybrids in [2] and [4], respectively,

$$\beta_k^C = (1 - \theta_k)\beta_k^{HS} + \theta_k\beta_k^{DY}, \tag{1.12}$$

which is a convex combination of HS and DY, and

$$\beta_k^N = (1 - \theta_k)\beta_k^{PRP} + \theta_k\beta_k^{FR}, \tag{1.13}$$

which combines PRP and FR methods. These methods are globally convergent under strong Wolfe conditions and standard Wolfe conditions, respectively. Other hybrids include those of [1, 6, 9, 15, 20, 23, 35, 40, 41, 48].

In this paper we present another hybrid conjugate gradient method that combines the good features of DY, VHS and $\beta_k^{**}$ by Mo *et al.* [31]. This method is presented in the next section. The rest of the paper is organised as follows. The sufficient decent condition and the global convergence of the proposed method is presented in Section 3. We then compare the numerical results of the new method with some other existing methods in Section 4. Finally, conclusion is presented in Section 5.

## 2. Proposed hybrid method

In this section we develop a new hybrid conjugate gradient method guided by the works of Mo *et al.* [31], and Xu and Kong [45]. In 2007, Mo *et al.* [31] proposed a hybrid method with

$$\beta_k = \max\left\{0, \min\left\{\beta_k^{HS}, \beta_k^{DY}, \beta_k^{**}\right\}\right\}, \tag{2.1}$$

where

$$\beta_k^{**} = \beta_k^{HS} + \frac{2g_k^T g_{k-1}}{d_{k-1}^T y_{k-1}}. \tag{2.2}$$

With the generalized Wolfe conditions (1.11), this method was shown to be globally convergent. Xu and Kong [45] proposed a linear combination of HS and DY methods as

$$\beta_k = \begin{cases} a_1\beta_k^{DY} + a_2\beta_k^{HS}, & \text{if} \quad \|g_k\|^2 > \left|g_k^T g_{k-1}\right|, \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

where $a_1, a_2$ are positive numbers which satisfy $0 < a_1 + 2a_2 < \frac{1}{1+\sigma_2} < 1$. To establish the method's global convergence, the authors suggested using the generalized Wolfe line search (1.11), when

$$d_{k-1}^T g_k \geq 0 \quad \text{and} \quad -\sigma_2 d_{k-1}^T g_{k-1} < -\sigma_2 d_{k-1}^T (g_{k-1} - g_k),$$

and the modified Wolfe line search

$$\sigma_1 d_{k-1}^T g_{k-1} \leq d_{k-1}^T g_k \leq -\sigma_2 d_{k-1}^T (g_{k-1} - g_k), \tag{2.4}$$

when

$$d_{k-1}^T g_k < 0, \quad \text{and} \quad -\sigma_2 d_{k-1}^T (g_{k-1} - g_k) < -\sigma_2 d_{k-1}^T g_{k-1}.$$

Motivated by the above two methods, here we propose a new hybrid as a linear combination of VHS, $\beta_k^{**}$ and DY methods as

$$\beta_k^{S_1} = \begin{cases} a_1\beta_k^{DY} + a_2\beta_k, & \text{if} \quad \|g_k\|^2 < \left|g_k^T g_{k-1}\right|, \\ \beta_k^{VHS}, & \text{otherwise}, \end{cases} \tag{2.5}$$

where

$$\beta_k = \max\left\{0, \min\left\{\beta_k^{VHS}, \beta_k^{**}\right\}\right\}$$

and $a_1$ and $a_2$ are positive constants such that $0 < a_1 + a_2 < \frac{1}{1+\sigma_2} < 1$ and $\sigma_2 \geq 0$ is as defined in (1.11). The sufficient descent condition and the global convergence of the method are established using the generalized Wolfe conditions (1.11). We describe the developed algorithm below.

**Algorithm 2.1** A new linear combination of HS and DY Conjugate Gradient (NLCHSDY) method.

1: Give the initial point $x_0 \in \mathbb{R}^n$, tolerance $\epsilon > 0$ and set $k = 0$.
2: Set $g_0 = \nabla f(x_0)$. If $\|g_0\| \leq \epsilon$ stop.
3: **for** $k = 0, 1, \dots$ **do**
4: Compute $d_k$ using (1.5).
5: Find $\alpha_k$ using (1.11).
6: Set $x_{k+1} = x_k + \alpha_k d_k$, and set $k = k + 1$.
7: If $\|g_k\| \leq \epsilon$ stop.
8: Compute $\beta_k$ using (2.5).
9: **end for**

## 3. Convergence analysis

In this section we provide the sufficient descent condition of our new hybrid conjugate gradient method and establish its global convergence under the generalized Wolfe conditions (1.11).

**Lemma 3.1.** *Let* $\{g_k\}$ *and* $\{d_k\}$ *be generated by Algorithm* 2.1. *Then*

$$d_k^T g_k < 0, \quad \forall k \geq 0. \tag{3.1}$$

*Proof.* When $k = 0$, we have $d_k = -g_k$ which implies $d_0^T g_0 = -\|g_0\|^2 < 0$ satisfies (3.1). Now, we show it also holds for $k \geq 1$. Suppose the result holds for $n = k - 1$, that is, $d_{k-1}^T g_{k-1} < 0$. From $\beta_k = \max\{0, \min\{\beta_k^{VHS}, \beta_k^{**}\}\}$ we have $\beta_k \geq 0$. Also

$$\min\left\{\beta_k^{VHS}, \beta_k^{**}\right\} \leq \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} = \beta_k^{DY}.$$

Notice that from the line search (1.11) we have that

$$d_{k-1}^T y_{k-1} \geq (\sigma_1 - 1) d_{k-1}^T g_{k-1} > 0,$$

and since $0 < a_1 + a_2 < \frac{1}{1+\sigma_2} < 1$, we get that $0 \leq \beta_k^{S_1} \leq \beta_k^{DY}$.

Now, for $n = k$, pre-multiplying both sides of $d_k = -g_k + \beta_k^{S_1} d_{k-1}$ by $g_k^T$ it follows that

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2 + \beta_k^{S_1} g_k^T d_{k-1} \leq -\|g_k\|^2 + \beta_k^{DY} g_k^T d_{k-1} \\ &= \|g_k\|^2 + \left(\frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}\right) g_k^T d_{k-1} = \left(-1 + \frac{g_k^T d_{k-1}}{d_{k-1}^T y_{k-1}}\right) \|g_k\|^2 \\ &= \left(\frac{g_{k-1}^T d_{k-1}}{d_{k-1}^T y_{k-1}}\right) \|g_k\|^2 = \beta_k^{DY} g_{k-1}^T d_{k-1} < 0. \end{aligned} \tag{3.2}$$

Thus, $g_k^T d_k < 0$ holds for all $k \geq 0$. This completes the proof. □

Now to prove the convergence of our new hybrid method we need the following assumptions.

**Assumption 3.1.** Let the level set

$$\Omega = \left\{x \in \mathbb{R}^n \mid f(x) \leq f(x_0)\right\},$$

where $x_0$ is the initial guess, be bounded. That is, there exists a positive constant $B$ such that

$$\|x\| \leq B, \quad \forall x \in \Omega. \tag{3.3}$$

**Assumption 3.2.** In some neighbourhood $N$ of $\Omega$ the function $f$ is continuously differentiable and its gradient $g(x)$ is Lipschitz continuous, i.e. there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall x, y \in N.$$

**Lemma 3.2.** *Consider any iteration of the form* $x_{k+1} = x_k + \alpha_k d_k$, *where* $d_k$ *is a descent direction and* $\alpha_k$ *satisfies the generalized Wolfe conditions* (1.11). *Suppose Assumptions* 3.1 *and* 3.2 *hold, then either*

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

*or*

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

**Theorem 3.1.** *Suppose Assumptions* 3.1 *and* 3.2 *hold, and let* $\{g_k\}$ *and* $\{d_k\}$ *be generated by Algorithm* 2.1. *Then*

$$\lim_{k \longrightarrow \infty} \inf \ \|g_k\| = 0. \tag{3.4}$$

*Proof.* We will prove this theorem by contradiction. Assume that (3.4) does not hold. Then there exists a constant $\gamma > 0$ such that

$$\|g_k\| \geq \gamma, \quad \forall k \geq 0. \tag{3.5}$$

Now, by squaring both sides of $d_k + g_k = \beta_k^{S_1} d_{k-1}$ and dividing throughout by $(g_k^T d_k)^2$, we get

$$\begin{aligned}
\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} &= \frac{\left(\beta_k^{S_1}\right)^2 \|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} - \frac{2}{g_k^T d_k} - \frac{\|g_k\|^2}{\left(g_k^T d_k\right)^2} \\
&= \frac{\left(\beta_k^{S_1}\right)^2 \|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} - \left(\frac{1}{\|g_k\|} + \frac{\|g_k\|}{g_k^T d_k}\right)^2 + \frac{1}{\|g_k\|^2} \\
&\leq \frac{\left(\beta_k^{DY}\right)^2 \|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} + \frac{1}{\|g_k\|^2}.
\end{aligned} \tag{3.6}$$

From (3.2) we have that

$$\beta_k^{DY} \leq \frac{d_k^T g_k}{d_{k-1}^T g_{k-1}},$$

which gives that

$$\begin{aligned}
\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} &\leq \left(\frac{d_k^T g_k}{d_{k-1}^T g_{k-1}}\right)^2 \frac{\|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} + \frac{1}{\|g_k\|^2} \\
&= \frac{\|d_{k-1}\|^2}{\left(g_{k-1}^T d_{k-1}\right)^2} + \frac{1}{\|g_k\|^2}.
\end{aligned} \tag{3.7}$$

Now, by noting that

$$\frac{\|d_0\|^2}{\left(d_0^T g_0\right)^2} = \frac{1}{\|g_0\|^2},$$

and using (3.7) recursively, together with (3.5), we get that

$$\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} \leq \sum_{i=0}^{k} \frac{1}{\|g_i\|^2} \leq \frac{k+1}{\gamma^2}.$$

Thus,

$$\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} \geq \frac{\gamma^2}{k+1},$$

and therefore

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

This is a contradiction to Lemma 3.2. Hence the proof is complete. □

## 4. Numerical results

In this section we present some numerical results of our proposed new hybrid conjugate gradient method, herein denoted NLCHSDY, and we compare the results with those of the hybrid methods by Mo *et al.* [31] and Xu and Kong [45], herein denoted BMHSDY and LCHSDY, respectively. All the algorithms are coded in MATLAB R2019b. We test the methods using a total of 72 test problems selected from Andrei [3] and Moré *et al.* [32], with dimensions ranging from 2 to 20000. We use the generalized Wolfe conditions (1.11) for the step size $\alpha_k$, and set the parameters as $\sigma = 0.01$, $\sigma_1 = 0.1$ and $\sigma_2 = 0.1$ for our proposed method. The other parameters $a_1$ and $a_2$ are set to be $0.1$ and $0.6$, respectively. We terminated all the algorithms when $\|g_k\| \leq \epsilon = 10^{-4}$, or the maximum number of iterations exceeds $5000$. The parameters for the other two methods, that is, LCHSDY and BMHSDY, are set as in respective papers.

We present the numerical results in Table 1 and also compare the methods using the performance profiles tool introduced by Dolan and Moré [14], in Figs. 1-3. In Table 1, 'FN' denotes the name of the test problem, 'DIM' denotes dimension of the test problem, 'NI' is the number of iterations, 'FE' is the number of function evaluations and 'NGE' gives the number of gradient evaluations. The dash '-' indicates that the algorithm (method) failed to reach the optimal solution within the maximum number of iterations. It can be seen from Table 1 that the LCHSDY method struggled on a number of problems as it failed to solve them within the maximum number of iterations. On the other hand, the other two methods were very competitive, with the proposed method managing to solve all the problems and the BMHSDY failing on only one problem.

The comparison of the methods is further presented in Figs. 1-3 using Dolan and Moré performance profiles tool. Fig. 1 presents the performance profiles of number of

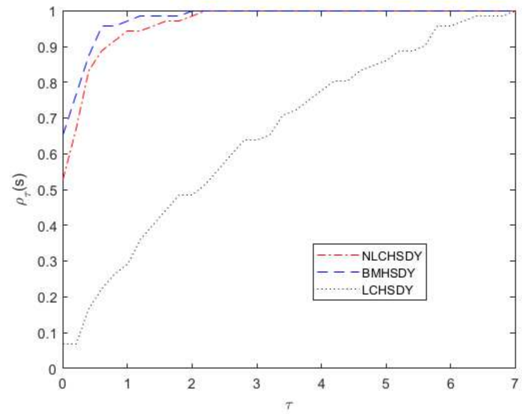


Figure 1: Iterations performance profiles.

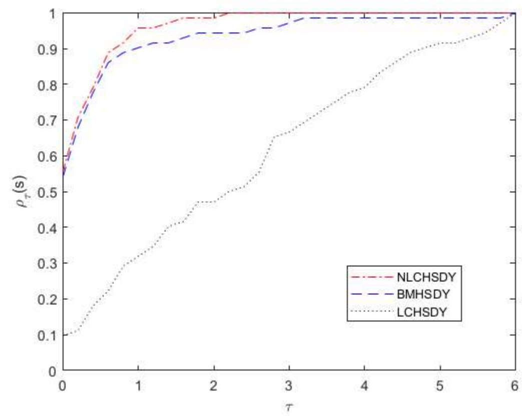


Figure 2: Function evaluations performance profiles.

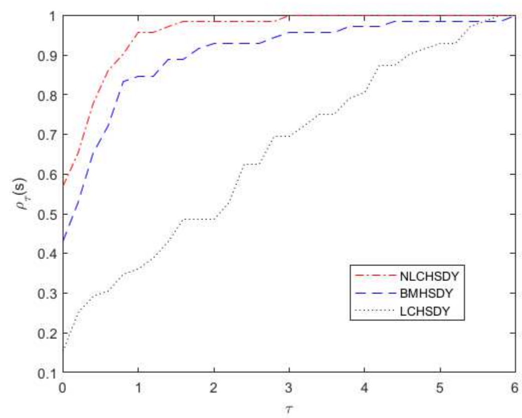


Figure 3: Number of gradient evaluations performance profiles.

Table 1: Table of numerical results.

| FN | DIM | NLCHSDY | | | BMHSDY | | | LCHSDY | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | NI | FE | lNGE | NI | FE | NGE | NI | FE | NGE |
| Brown & Denis | 4 | 41 | 332 | 110 | 30 | 275 | 93 | 105 | 1594 | 107 |
| Ext. Freudenstein & Roth | 5000 | 11 | 53 | 21 | 10 | 38 | 14 | 2488 | 10962 | 2650 |
| Ext. Beale | 20000 | 14 | 65 | 41 | 16 | 728 | 710 | 333 | 998 | 337 |
| Perturbed Quadratic | 2000 | 229 | 1147 | 231 | 346 | 1592 | 527 | 3280 | 16401 | 3281 |
| Helical | 3 | 34 | 167 | 65 | 37 | 172 | 80 | 940 | 3762 | 941 |
| Ext. Rosenbrock | 20000 | 41 | 208 | 93 | 30 | 377 | 312 | - | - | - |
| Diagonal 1 | 50 | 37 | 121 | 72 | 39 | 131 | 80 | 70 | 211 | 71 |
| Box 3 dimensional | 3 | 9 | 37 | 30 | 16 | 58327 | 58318 | 4600 | 14452 | 9843 |
| Freudenstein & Roth | 2 | 10 | 48 | 20 | 8 | 32 | 12 | 2216 | 8869 | 2217 |
| Biggs EXP6 | 6 | 98 | 296 | 244 | 94 | 286 | 244 | - | - | - |
| Powell | 4 | 43 | 162 | 87 | 12 | 45 | 22 | - | - | - |
| Diagonal 2 | 2000 | 156 | 673 | 672 | 205 | 827 | 826 | 822 | 2327 | 2324 |
| Diagonal 3 | 500 | 133 | 728 | 451 | 137 | 526 | 220 | 755 | 3031 | 757 |
| Raydan 1 | 2000 | 227 | 693 | 236 | 266 | 831 | 316 | 2340 | 7041 | 2341 |
| Raydan 2 | 20000 | 4 | 6 | 6 | 3 | 5 | 5 | 3 | 5 | 4 |
| Ext. White & Holst | 20000 | 46 | 246 | 107 | 30 | 672 | 600 | - | - | - |
| Ext. Himmelblau | 20000 | 8 | 32 | 15 | 9 | 35 | 17 | 19 | 60 | 22 |
| Wood | 4 | 122 | 630 | 176 | 157 | 779 | 365 | - | - | - |
| Ext. Tridiagonal 1 | 20000 | 14 | 61 | 51 | 2 | 40 | 38 | 2113 | 4225 | 2114 |
| Ext. BD1 | 20000 | 10 | 35 | 24 | 10 | 54 | 43 | 16 | 48 | 29 |
| BDEXP | 20000 | 5 | 49 | 49 | 5 | 49 | 49 | 16 | 152 | 152 |
| Gen. Rosenbrock | 200 | 1089 | 5467 | 1190 | 2662 | 12012 | 4237 | - | - | - |
| APQ | 5000 | 365 | 2191 | 727 | 493 | 2450 | 526 | - | - | - |
| Diagonal 7 | 20000 | 4 | 10 | 6 | 4 | 10 | 6 | 4 | 9 | 5 |
| Diagonal 8 | 20000 | 3 | 8 | 5 | 3 | 8 | 5 | 3 | 8 | 4 |
| Beale | 2 | 13 | 56 | 34 | 12 | 239 | 224 | 71 | 191 | 74 |
| Rosenbrock | 2 | 41 | 208 | 93 | 29 | 373 | 311 | 3393 | 13664 | 3484 |
| QUARTC | 20000 | 3 | 22 | 20 | 4 | 24 | 23 | 14 | 84 | 81 |
| Ext. Penalty | 20000 | 16 | 134 | 30 | 11 | 120 | 39 | 33 | 287 | 34 |
| BIGGSB1 | 100 | 100 | 206 | 107 | 63 | 132 | 70 | 4050 | 8101 | 4051 |
| Gaussian | 3 | 1 | 3 | 2 | 1 | 3 | 2 | 1 | 3 | 2 |
| Penalty I | 20000 | 21 | 149 | 57 | 27 | 363 | 290 | 85 | 374 | 146 |
| Penalty II | 10 | 14 | 58 | 32 | 13 | 89 | 70 | 101 | 308 | 103 |
| Ext. quad. pen. QP1 | 20000 | 4 | 25 | 16 | 7 | 29 | 12 | 12 | 1053 | 1009 |
| Ext. Tridiagonal 2 | 20000 | 19 | 45 | 26 | 23 | 61 | 40 | 26 | 1055 | 1017 |
| Ext. Wood | 20000 | 174 | 890 | 238 | 123 | 791 | 465 | - | - | - |
| Gen. Tridiagonal 1 | 20000 | 17 | 51 | 21 | 17 | 51 | 21 | 23 | 69 | 24 |
| Gen. Tridiagonal 2 | 20000 | 40 | 139 | 53 | 42 | 149 | 65 | 70 | 216 | 71 |
| Nondquar | 20000 | 500 | 1354 | 1170 | 664 | 1850 | 1568 | - | - | - |
| Hager | 10000 | 63 | 254 | 75 | 79 | 1287 | 1134 | 165 | 1525 | 1178 |

Table 1: Table of numerical results (cont'd).

| FN | DIM | NLCHSDY | | | BMHSDY | | | LCHSDY | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | NI | FE | lNGE | NI | FE | NGE | NI | FE | NGE |
| Himmelbg | 20000 | 4 | 47 | 47 | 4 | 47 | 47 | 18 | 209 | 209 |
| Sinquad | 20000 | 495 | 3560 | 1130 | 55 | 1480 | 1256 | - | - | - |
| Liarwhd | 20000 | 37 | 261 | 57 | 19 | 168 | 96 | 494 | 2682 | 495 |
| Cosine | 20000 | 7 | 24 | 18 | 7 | 24 | 17 | 7 | 24 | 11 |
| Gen. White & Holst | 100 | 1822 | 9929 | 2508 | - | - | - | - | - | - |
| Diagonal 4 | 20000 | 6 | 14 | 8 | 6 | 14 | 7 | 66 | 176 | 67 |
| Ext. Maratos | 5000 | 65 | 398 | 188 | 51 | 534 | 409 | 1836 | 7355 | 1849 |
| Bard | 3 | 18 | 61 | 39 | 13 | 52 | 35 | 4191 | 9786 | 4192 |
| Arwhead | 2000 | 18 | 61 | 15 | 5 | 25 | 8 | 61 | 1289 | 1052 |
| Ext. TET | 5000 | 5 | 14 | 9 | 5 | 14 | 9 | 19 | 42 | 20 |
| Ext. Denschnb | 20000 | 4 | 13 | 8 | 6 | 23 | 17 | 12 | 25 | 13 |
| Ext. Denschnf | 20000 | 10 | 61 | 33 | 8 | 34 | 13 | 15 | 61 | 16 |
| Ext. Powell singular | 10000 | 67 | 255 | 144 | 16 | 61 | 32 | - | - | - |
| Ext. PSC1 | 20000 | 10 | 26 | 14 | 10 | 25 | 14 | 13 | 31 | 14 |
| Gen. PSC1 | 20000 | 26 | 183 | 169 | 10 | 69 | 63 | 15 | 145 | 71 |
| BDQRTIC | 500 | 86 | 433 | 123 | 77 | 323 | 91 | - | - | - |
| Gen. Quartic | 20000 | 7 | 24 | 17 | 7 | 17 | 10 | 9 | 22 | 10 |
| Ext. quad. exp. EP1 | 10000 | 2 | 112 | 66 | 3 | 13 | 4 | 3 | 13 | 4 |
| Engval1 | 20000 | 19 | 47 | 25 | 19 | 1051 | 1024 | 26 | 100 | 27 |
| EG2 | 200 | 24 | 172 | 71 | 17 | 105 | 58 | - | - | - |
| Dqdrtic | 20000 | 12 | 50 | 19 | 5 | 19 | 7 | 408 | 1632 | 409 |
| Brown almost-linear | 200 | 7 | 71 | 35 | 50 | 1002 | 999 | 13 | 53 | 14 |
| Broyden tridiagonal | 20000 | 30 | 103 | 40 | 27 | 88 | 35 | 35 | 109 | 36 |
| Gulf | 3 | 60 | 258 | 181 | 34 | 258 | 202 | - | - | - |
| Dixon3dq | 100 | 100 | 206 | 107 | 177 | 370 | 235 | 4529 | 9059 | 4530 |
| Ext. quad. pen. QP2 | 20000 | 44 | 251 | 112 | 37 | 292 | 192 | 983 | 4035 | 1076 |
| Nondia | 2000 | 13 | 88 | 19 | 7 | 44 | 14 | 201 | 1104 | 202 |
| Nonscomp | 20000 | 41 | 150 | 67 | 35 | 139 | 75 | 50 | 156 | 51 |
| Perturbed Quad. Diagonal | 20000 | 309 | 1784 | 497 | 254 | 1097 | 338 | 3148 | 16897 | 3149 |
| Quadratic QF1 | 5000 | 365 | 1829 | 369 | 508 | 2409 | 796 | - | - | - |
| Quadratic QF2 | 5000 | 645 | 3880 | 652 | 660 | 3330 | 723 | - | - | - |
| Tridia | 500 | 476 | 2785 | 880 | 387 | 1903 | 457 | - | - | - |

iterations, Fig. 2 is function evaluations performance profiles and lastly, Fig. 3 gives the performance profiles of gradient evaluations. It is clear from these figures that our proposed NLCHSDY method outperforms the other two methods. Thus, the NLCHSDY method, compared to the other two methods, is very effective.

## 5. Conclusion

In this paper we proposed a new hybrid conjugate gradient method as a linear combination of the Dai-Yuan (DY) method and either Hestenes-Stiefel (HS) method or its modifications. The method was shown to satisfy the sufficient descent condition and its global convergence was established under the generalized Wolfe conditions. The method was tested on a number of benchmark problems from the literature and compared with other methods. Numerical results show that our new proposed hybrid conjugate gradient method is more effective compared to those other conjugate gradient methods.

## References


[1] A. ALHAWARAT, Z. SALLEH, M. MAMAT AND M. RIVAIE, *An efficient modified Polak-Ribière-Polak conjugate gradient method with global convergence properties*, Optim. Methods Softw., 32(6) (2017), 1299–1312.

[2] N. ANDREI, *A hybrid conjugate gradient algorithm for unconstrained optimization as a convex combination of Hestenes-Stiefel and Dai-Yuan*, Stud. Inform. Control, 17(1) (2008), 55–70.

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

[4] N. ANDREI, *Hybrid conjugate gradient algorithm for unconstrained optimization*, J. Optim. Theory Appl., 141(2) (2009), 249–264.

[5] S. BABAIE-KAFAKI AND R. GHANBARI, *A hybridization of Polak-Ribière-Polak and Fletcher-Reeves conjugate gradient methods*, Numer. Algor., 68(3) (2015), 481–495.

[6] S. BABAIE-KAFAKI AND N. MAHDAVI-AMIRI, *Two modified hybrid conjugate gradient methods based on a hybrid secant equation*, Math. Model. Anal., 18(1) (2013), 32–52.

[7] C. CARTIS, N. I. M. GOULD AND PH. L. TOINT, *On the complexity of steepest descent, Newton's and regularized Newton's methods for nonconvex unconstrained optimization problems*, SIAM J. Optim., 20(6) (2010), 2833–2852.

[8] S. K. CHAKRABORTY AND G. PANDA, *Two phase quasi-Newton method for unconstrained optimization problem*, Afr. Mat., 30 (2019), 737–746.

[9] W. CHENG, *A two-term PRP-based descent method*, Numer. Funct. Anal. Optim., 28 (2007), 1217–1230.

[10] Y. H. DAI AND Y. YUAN, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optim., 10 (1999), 177–182.

[11] R. DEHGHANI, N. BIDABADI, H. FAHS AND M. M. HOSSEINI, *A conjugaate gradient method based on a modified secant relation for unconstrained optimization*, Numer. Funct. Anal. Optim., 41(5) (2020), 621–634.

[12] R. DEHGHANI, M. M. HOSSEINI AND N. BIDABADI, *The modified quasi-Newton methods for solving unconstrained optimization problems*, Int. J. Numer. Model. El., 32 (2019), e2459.

[13] S. S. DJORDJEVIĆ, *New hybrid conjugate gardient method as a convex combination of Ls and Fr methods*, Acta Math. Sci., 39 (2019), 214–228.

[14] E. D. DOLAN AND J. J. MORÉ, *Benchmarking optimization software with performance profiles*, Math. Program., 91 (2001), 201–213.

[15] X. L. DONG, H. W. LIU, Y. B. HE, S. BABAIE-KAFAKI AND R. GHANBARI, *A new three-term conjugate gradient method with descent direction for unconstrained optimization*, Math. Model. Anal., 21 (2016), 399–411.

[16] R. FLETCHER, *Practical Methods of Optimization. Second Edition*, John Wiley and Sons, 1987.

[17] R. FLETCHER AND C. REEVES, *Function minimization by conjugate gradients*, Computer J., 7 (1964), 149–154.

[18] A. GALÁNTAI, *The theory of Newton's method*, J. Comput. Appl. Math., 24 (2000), 25–44.

[19] W. 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.

[20] M. HAMODA, M. MAMAT, M. RIVAIE AND Z. SALLEH, *A conjugate gradient method with strong Wolfe-Powell line search for unconstrained optimization*, Appl. Math. Sci. (Hikari), 10 (2016), 721–734.

[21] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. Natl. Bur. Stand., 49(6) (1952), 409–436.

[22] R. JOHARI, M. RIVAIE AND M. MAMAT, *A new scaled steepest descent method for unconstrained optimization with global convergence properties*, J. Eng. Appl. Sci., 13(6) (2018), 5442–5445.

[23] P. KAELO, *A hybrid nonlinear conjugate gradient method for unconstrained optimization problems*, Pacific J. Optim., 12(4) (2016), 847–859.

[24] P. KAELO, P. MTAGULWA AND M. V. THUTO, *A globally convergent hybrid conjugate gradient method with strong Wolfe conditions for unconstrained optimization*, Math. Sci. (Springer), 14 (2020), 1–9.

[25] X. LI AND X. ZHAO, *A hybrid conjugate gradient method for optimization problems*, Natural Science, 3(1) (2011), 85–90.

[26] C. S. LIU, J. R. CHANG AND Y. W. CHEN, *A modified algorithm of steepest descent method for solving unconstrained nonlinear optimization problems*, J. Mar. Sci. Tech. Japan, 23(1) (2015), 88–97.

[27] J. LIU, *A hybrid nonlinear conjugate gradient method*, Lobachevskii J. Math., 33(3) (2012), 195–199.

[28] J. LIU AND S. DU, *Modified three-term conjugate gradient method and its applications*, Math. Probl. Eng., 2019 (2019), Article ID 5976595.

[29] Y. LIU AND C. STOREY, *Efficient generalized conjugate gradient algorithms*, J. Optim. Theory Appl., 69(1) (1991), 129–137.

[30] I. E. LIVIERIS, V. TAMPAKAS AND P. PINTELAS, *A descent hybrid conjugate gradient method based on the memoryless BFGS update*, Numer. Algor., 79 (2018), 1169–1185.

[31] J. MO, N. GU AND Z. WEI, *Hybrid conjugate gradient methods for unconstrained optimization*, Optim. Methods Softw., 22(2) (2007), 297–307.

[32] J. J. MORÉ, B. S. GARBOW AND K. E. HILLSTROM, *Testing unconstrained optimization software*, ACM Trans. Math. Software, 7(1) (1981), 136–140.

[33] V. MOROVATI, H. BASIRZADEH AND L. POURKARIMI, *Quasi-Newton methods for multiobjective optimization problems*, 4OR, 16 (2018), 261–294.

[34] P. MTAGULWA AND P. KAELO, *An efficient modified PRP-FR hybrid conjugate gradient method for solving unconstrained optimization problems*, Appl. Numer. Math., 145 (2019), 111–120.

[35] S. NARAYANAN, P. KAELO AND M. V. THUTO, *A modified quadratic hybridization of Polak-Ribière-Polak and Fletcher-Reeves conjugate gradient method for unconstrained optimization problems*, Int. J. Optim. Control. Theor. Appl., 7(2) (2017), 177–185.

[36] E. POLAK AND G. RIBIÈRE, *Note sur la convergence de methodes de directions conjuguees*, Rev. Fr. Inf. Rech. Oper. 3e Année, 16 (1969), 35–43.

[37] Z. POLAREJ, *Quasi-Newton's method for multiobjective optimization*, J. Comput. Appl. Math., 255(1) (2014), 765–777.

[38] B. T. POLYAK, *The conjugate gradient method in extreme problems*, USSR Comput. Math. Math. Phys., 9 (1969), 94–112.

[39] B. T. POLYAK, *Newton's method and its use in optimization*, Eur. J. Oper. Res., 181(3) (2007), 1086–1096.

[40] M. RIVAIE, M. MAMAT, L. W. JUNE AND I. MOHD, *A new class of nonlinear conjugate gradient coefficients with global convergence properties*, Appl. Math. Comput., 218 (2012), 11323–11332.

[41] Z. SALLEH AND A. ALHAWARAT, *An efficient modification of the Hestenes-Stiefel nonlinear conjugate gradient method with restart property*, J. Inequal. Appl., 2016 (2016): 110.

[42] A. WAKILI, *Quasi-Newton method for solving non-linear optimization problems of convergence functional*, IOSR J. Math., 15(4) (2019), 42–47.

[43] Z. WEI, G. LI AND L. QI, *New quasi-Newton methods for unconstrained optimization problems*, Appl. Math. Comput., 175(2) (2006), 1156–1188.

[44] Z. WEI, S. YAO AND L. LIU, *The convergence properties of some new conjugate gradient methods*, Appl. Math. Comput., 183 (2006), 1341–1350.

[45] X. XU AND F. KONG, *New hybrid conjugate gradient methods with the generalized Wolfe line search*, Springer Plus, 5(1) (2016): 881.

[46] S. YAO, Z. WEI AND H. HUANG, *A note about WYL's conjugate gradient method and its application*, Appl. Math. Comput., 191(2) (2007), 381–388.

[47] G. YUAN AND W. HU, *A conjugate gradient algorithm for large-scale unconstrained optimization problems and nonlinear equations*, J. Inequal. Appl., 2018 (2018): 113.

[48] G. YUAN, X. DUAN, W. LIU, X. WANG, Z. CUI AND Z. SHENG, *Two new PRP conjugate gradient algorithms for minimization optimization models*, Plos One, 10(10) (2015), e0140071.

[49] Y.-X. YUAN, *A new stepsize for the steepest descent method*, J. Comput. Math., 24(2) (2006), 149–156.