

# A New Family of Hybrid Conjugate Gradient Methods for Unconstrained Optimization

Olawale J. Adeleke<sup>1,2,\*</sup>, Absalom E. Ezugwu<sup>3</sup>, Idowu A. Osinuga<sup>4</sup>

<sup>1</sup>*Department of Mathematical Sciences, redeemer's university, ede, Osun state, Nigeria*

<sup>2</sup>*DST-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS)*

<sup>3</sup>*School of Mathematics, Statistics, and Computer Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa*

<sup>4</sup>*Department of Mathematics, College of Physical Sciences, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria*

**Abstract** The conjugate gradient method is a very efficient iterative technique for solving large-scale unconstrained optimization problems. Motivated by recent modifications of some variants of the method and construction of hybrid methods, this study proposed four hybrid methods that are globally convergent as well as computationally efficient. The approach adopted for constructing the hybrid methods entails projecting ten recently modified conjugate gradient methods. Each of the hybrid methods is shown to satisfy the descent property independent of any line search technique and globally convergent under the influence of strong Wolfe line search. Results obtained from numerical implementation of these methods and performance profiling show that the methods are very competitive with well-known traditional methods.

**Keywords** Hybrid methods, nonlinear conjugate gradient method, unconstrained optimization problems, descent property, global convergence, strong Wolfe line search conditions

**AMS 2010 subject classifications** 90C26

**DOI:** 10.19139/soic.v7i3.480

## 1. Introduction

From the industrial perspective and the associated processes such as facility location, production, distribution, logistics, etc, the applications of optimization techniques can not be overemphasized. For instance, their applications take central place in the well studied vehicle routing problem, a problem frequently encountered by the distribution units of manufacturing industries (see [2]). Their importance has necessitated the design of many techniques that can be deployed for solving both constrained and unconstrained optimization problems (see [11, 12]). In this paper, however, we consider an optimization technique known as the conjugate gradient (CG) method for solving unconstrained optimization problems of the form

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

where  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is a continuously differentiable objective function for which the gradient can be evaluated. The nonlinear CG methods constitute a class of iterative scheme for solving (1) effectively especially for large values of  $n$ . The CG scheme for (1) can be written as follows:

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

\*Correspondence to: Olawale J. Adeleke (Email: wale.adeleke@covenantuniversity.edu.ng). Department of Mathematics, College of Science and Technology, Covenant University, Canaanland, PMB. 1023, Ota, Ogun State, Nigeria.

where  $\alpha_k$  is the step-length evaluated by an appropriate line search process, and  $d_k$  is the search direction given by the following rules:

$$d_0 = -g_0, \text{ for } k = 0; \quad d_k = -g_k + \beta_k d_{k-1}, \text{ for } k \geq 1, \quad (3)$$

where  $g_k = g(x_k)$  is the gradient of  $f$  at  $x_k$ ,  $\beta_k$  is the CG update parameter. Usually  $\beta_k$  is chosen such that the iterative scheme (2) and (3) reduces to the linear CG method when  $f$  is a strictly convex quadratic function and  $\alpha_k$  is determined by a one-dimensional exact line search technique.

The past few decades have seen many variants of the nonlinear CGM developed, some of which includes Hestenes and Stiefel (HS) [16], Fletcher and Reeves (FR) [13], Polak, Ribier and Polyak (PRP) [27, 28], Liu and Storey (LS) [24], Dai and Yuan (DY) [7] and Hager and Zhang (CG\_DESCENT (N)) [14]. The basic difference among these methods exists in the value of their  $\beta_k$  as may be observed in their respective value as follows:

$$\begin{aligned} \beta_k^{HS} &= \frac{g_k^T y_k}{d_{k-1}^T y_k}, & \beta_k^{FR} &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, & \beta_k^{PRP} &= \frac{g_k^T y_k}{\|g_{k-1}\|^2}, \\ \beta_k^{LS} &= \frac{g_k^T y_k}{-d_{k-1}^T g_{k-1}}, & \beta_k^{DY} &= \frac{\|g_k\|^2}{d_{k-1}^T y_k}, & \beta_k^N &= \left( y_k - 2d_{k-1} \frac{\|y_k\|^2}{d_{k-1}^T y_k} \right)^T \frac{g_{k+1}}{d_{k-1}^T y_k}, \end{aligned}$$

where  $\|\cdot\|$  denotes the Euclidean norm and  $y_k = g_k - g_{k-1}$ .

Experiments have shown that for strictly convex function  $f$  all the above methods are equivalent with exact line search. However, for non-quadratic objective functions, the  $\beta_k$  of these methods perform differently [1, 15]. Establishing the global convergence results of these methods usually requires that the step-length  $\alpha_k$  satisfies some inexact line search criteria because most exact line search approaches are computationally expensive. The most commonly used among the approximate line search criteria is the strong Wolfe line search (SWLS) criterion which is given as follows:

$$f(x_{k-1} + \alpha_k d_k) - f(x_k) \geq -\delta \alpha_k g_{k-1}^T d_{k-1} \quad (4)$$

and

$$|g_k^T d_{k-1}| \leq -\sigma g_{k-1}^T d_{k-1}, \quad (5)$$

with  $0 < \delta < \sigma < 1$ . The weaker version of this condition combines (4) with

$$|g_k^T d_{k-1}| \geq \sigma g_{k-1}^T d_{k-1} \quad (6)$$

Dai and Yuan [6] gave a generalization (GWLS) of the above conditions as a combination of (4) and

$$\sigma_1 g_{k-1}^T d_{k-1} \leq g_k^T d_{k-1} \leq -\sigma_2 g_{k-1}^T d_{k-1}, \quad (7)$$

where  $0 < \delta < \sigma_1 < 1$  and  $\sigma_2 \geq 0$ . Note that when  $\sigma_1 = \sigma_2 = \sigma$ , (4) and (7) reduces to (4)-(5). Similarly, when  $\sigma_1 = \sigma$  and  $\sigma_2 = \inf$ , the combination of (4) and (7) reduces to (4) and (6). In [14], Hager and Zhang introduced an approximation of the Wolfe conditions by replacing (7) with

$$\sigma g_{k-1}^T d_{k-1} \leq g_k^T d_{k-1} \leq (2\delta - 1) g_{k-1}^T d_{k-1}, \quad (8)$$

such that  $0 < \delta < \frac{1}{2}$  and  $\delta < \sigma < 1$ . This line search procedure implemented on the CG\_DESCENT method (See [14]) is among the fastest known CG algorithm.

However, an important requirement for deploying the above CG variants and others not mentioned here is the existence of descent property during the implementation of any of the line search techniques described above. A CG algorithm is said to satisfy the descent property if

$$g_k^T d_{k-1} < 0, \quad \forall k \geq 1. \quad (9)$$

A more natural way of guaranteeing descent for CG algorithms apart from (9) is by using the sufficient descent property. This is given by

$$g_k^T d_{k-1} \leq -c \|g_{k-1}\|^2, \quad \forall k \geq 1. \quad (10)$$

where  $c$  is a positive constant. For more information on the theories and applications of CG methods, we refer the reader to [20, 21, 22, 23, 25, 26, 32].

Each CG method has very striking features that makes it adaptable to some sets of unconstrained problems. For instance, the FR and DY methods have been identified as having the best convergence results (See [3, 8]). However, for general objective functions, the two methods have poor computational power. Conversely, the HS and PRP methods have good computational strength even though they exhibit less desirable convergence results. These contrasting features have led to the development of hybrid methods which are constructed with the aim of overcoming existing deficiencies in two or more methods. For instance, a well-constructed hybrid method of FR and PRP should perform well computationally as well as yield good convergence properties. These advantages associated with hybrid CG methods are the motivations behind this study. In this paper, new hybrid CG methods are proposed through the combination of recently proposed variants of the PRP and HS methods. Even though the two methods are specifically known to perform well computation-wise, it is important to note that their hybrids exhibit strong global convergence properties.

The rest of this paper is organized as follows. In Section 2, a review of some recent methods leading to the proposed methods is presented. In Section 3, the new methods are highlighted along with a suitable algorithm for testing their computational strength. The descent and global convergence properties of the proposed methods in are described in Section 4. Some preliminary numerical results are presented in Section 5, while some concluding remarks are given in Section 6.

## 2. Related Studies

In this section, we trace the development of methods that directly relates to the proposed hybrid methods. The PRP and HS methods form the basic building blocks for all that is discussed hereafter. In [30], a new variant of PRP method, named VPRP method, was proposed with the CG update parameter  $\beta_k$  given by

$$\beta_k^{VPRP} = \frac{g_k^T \left( g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1} \right)}{\|g_{k-1}\|^2} \quad (11)$$

The study assumed that the method satisfies (10) and established its global convergence under different line search techniques, ranging from exact to approximate methods. Specifically, Huang et al. [17] demonstrated the global convergence of the method under the influence of strong Wolfe line search (SWLS) (4-5) with  $\sigma < \frac{1}{4}$ . Motivated by the results in [30], Shengwei et al. [29] extended the approach to suggest CG variants of the HS and LS methods. Particularly, their proposed variant of HS method has the value of  $\beta_k$  given by

$$\beta_k^{VHS} = \frac{g_k^T \left( g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1} \right)}{d_{k-1}^T y_k} \quad (12)$$

This method was shown to possess the sufficient descent property and globally convergent with  $\sigma < \frac{1}{3}$ . Zhang [34] modified (11) and (12) to give new variants, which are referred to as MVPRP and MVHS here, respectively, with the update parameters given as

$$\beta_k^{MVPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\|g_{k-1}\|^2} \quad (13)$$

and

$$\beta_k^{MVHS} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{d_{k-1}^T y_k} \quad (14)$$

Both methods of (13) and (14) satisfy (10) and were shown to be globally convergent for SWLS (4-5) with  $\sigma < \frac{1}{2}$ . The numerical results reported for these two methods show considerable improvement compared to (11) and (12). In fact, Du et al. [10] noted that  $0 \leq \beta_k^{MVPRP} \leq \beta_k^{FR}$ , and thus according to [33], the MVPRP method possesses

the descent property (9) and is globally convergent under GWLS (4) and (7) with  $\sigma_1 + \sigma_2 \leq 1$ . The equality part of this inequality suggests that the MVPRP method is guaranteed to converge globally under line search (4-5) with  $\sigma = \frac{1}{2}$ .

A new variant of PRP method was introduced in [18] possessing property (10) and globally convergent for the SWLS with  $\sigma \in (0, \frac{1}{2})$ . Their variant, which we shall denote as HPRP, has its  $\beta_k$  given as

$$\beta_k^{HPRP} = \frac{g_k^T \left( g_k - \frac{g_k^T g_{k-1}}{\|g_{k-1}\|^2} g_{k-1} \right)}{\|g_{k-1}\|^2}. \quad (15)$$

What was said of MVPRP above is also applicable to HPRP since  $0 \leq \beta_k^{HPRP} \leq \beta_k^{FRR}$ . A similar study in [31] produced a new variant of HS method referred to as WHS method in this study. In this case  $\beta_k$  is given as

$$\beta_k^{WHS} = \frac{g_k^T \left( g_k - \frac{g_k^T g_{k-1}}{\|g_{k-1}\|^2} g_{k-1} \right)}{d_{k-1}^T y_k}. \quad (16)$$

This method satisfies (10) and is globally convergent for the SWLS with  $\sigma \in (0, \frac{1}{2})$ . Du et al. [10] later showed that WHS method is very closely related to the DY method. Particularly, they showed that  $\beta_k^{WHS} \leq \beta_k^{DY}$ , for all  $k \geq 1$  and thus claimed that the method has a similar global convergence property like the DY method. In other words, the method is a descent method and globally convergent under the Wolfe line search (4) and (6) with  $\sigma < 1$ .

Very recently, Du et al. [10] proposed four modified CG methods. Here, we only mention two, which are of interest in this paper. The first of these, a variant of the PRP method which we shall refer to as DPRP method, has its  $\beta_k$  value given as follows:

$$\beta_k^{DPRP} = \frac{g_k^T \left( g_k - \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_{k-1} \right)}{\|g_{k-1}\|^2} \quad (17)$$

and the second, referred to as DHS method, is a variant of the HS method and has its  $\beta_k$  given as follows:

$$\beta_k^{DHS} = \frac{g_k^T \left( g_k - \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_{k-1} \right)}{d_{k-1}^T y_k}. \quad (18)$$

It was shown that the DPRP and DHS methods possess the sufficient descent property and are globally convergent under SWLS with  $0 < \sigma < \frac{1}{4}$  and  $0 < \sigma < \frac{1}{3}$ , respectively. It was further shown that the two methods are always nonnegative under the same assumptions of SWLS and the respective intervals of convergence above.

Next, the newly proposed hybrid methods are presented by projecting the methods described above. The development of these methods was motivated by the observation that numerical results reported for these methods revealed that the methods are not so efficient compared to some classical methods, especially the CG\_DESCENT method.

### 3. Hybrid Conjugate Gradient Methods and Implementation Algorithm

In this section, the hybrid methods of PRP, HS, VPRP, VHS, MVPRP, MVHS, HPRP, WHS, DPRP and DHS are presented. The following are the four proposed hybrid methods:

i. The hybrid of DPRP, DHS, PRP and HS, where  $\beta_k$  is given as

$$\beta_k^{DPH} = \frac{\|g_k\|^2 - \max\left\{\frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}, g_k^T g_{k-1}\right\}}{\max\{\|g_{k-1}\|^2, d_{k-1}^T (g_k - g_{k-1})\}} \quad (19)$$

ii. The hybrid of DPRP, DHS, HPRP and WHS, where  $\beta_k$  is given as

$$\beta_k^{DHW} = \frac{\|g_k\|^2 - \max\left\{\frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}, \left(\frac{g_k^T g_{k-1}}{\|g_{k-1}\|}\right)^2\right\}}{\max\{\|g_{k-1}\|^2, d_{k-1}^T (g_k - g_{k-1})\}} \quad (20)$$

iii. The hybrid of DPRP, DHS, VPRP and VHS, where  $\beta_k$  is given as

$$\beta_k^{DV} = \frac{\|g_k\|^2 - \max\left\{\frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}, \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}\right\}}{\max\{\|g_{k-1}\|^2, d_{k-1}^T (g_k - g_{k-1})\}} \quad (21)$$

iv. The hybrid of DPRP, DHS, MVPRP and MVHS, where  $\beta_k$  is given as

$$\beta_k^{DM} = \frac{\|g_k\|^2 - \max\left\{\frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}, \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|\right\}}{\max\{\|g_{k-1}\|^2, d_{k-1}^T (g_k - g_{k-1})\}} \quad (22)$$

These methods exhibit the properties of both PRP and HS methods, and the analyzes that follow in the next sections specifically show that the proposed methods are always descent independent of the step-length line search procedure adopted. Based on (19)-(22), the following algorithm is proposed for the analyzes and numerical implementation of the methods.

---

**Algorithm 1:** Hybrid Conjugate Gradient Algorithm

---

- 1 Given  $x_0 \in \mathbb{R}^n$ ,  $\epsilon > 0$ ; set  $g(x_0) = g_0$ ,  $d_0 = -g_0$ ,  $k = 0$ ;
  - 2 Obtain  $\alpha_k$  such that conditions (4-5) are satisfied and evaluate  $x_{k+1} = x_k + \alpha_k d_k$ ;
  - 3 **while**  $\|g_k\| \leq \epsilon$  **do**
  - 4      $g_k = g(x_k) = \nabla f(x_k)$ ;
  - 5      $\beta_k = \beta_k = \{\beta_k^{DPH}, \beta_k^{DHW}, \beta_k^{DV}, \beta_k^{DM}\}$ ;
  - 6      $d_k = -g_k + \beta_k d_{k-1}$ ;
  - 7      $k = k + 1$ ;
  - 8 **end**
- 

#### 4. Descent and Global Convergence Properties of Algorithm 1

In this section, the properties of Algorithm 1 are presented. The analyzes begin by establishing the descent properties of each method followed by their global convergence results. The results that follow establish the descent properties of the sequence of search directions generated by Algorithm 1. Interestingly, these results were obtained independent of any line search procedure.

*Theorem 1*

Suppose that  $f(x)$  in (1) is a smooth function and  $d_k$  is generated by Algorithm 1. Then  $g_k^T d_k < 0$  for each  $k \geq 0$ .

*Proof*

Clearly,  $g_0^T d_0 = \|g_0\|^2$  for  $k = 0$ . By assumption, let  $g_k^T d_k < 0$  holds for  $k \geq 1$ . To show that  $g_k^T d_k < 0$  for all  $k$ , the proof is divided into four different cases for each of the hybrid method. Note that if  $\beta_k = 0$ , then, from (3),  $d_k = -g_k + \beta_k d_{k-1} = -g_k \rightarrow g_k^T d_k = -\|g_0\|^2 < 0$ . Thus, it is assumed that  $\beta_k \neq 0$  in all the cases.

#### DPH Method

**Case I:** If  $g_k^T g_{k-1} < (|g_k^T g_{k-1}|/\|g_k\|^2) g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then from (19),  $\beta_k^{DPH} = \beta_k^{DHS}$ . Notably, since  $\|g_k\|^2 > 0$ , the denominator of  $\beta_k^{DHS}$ , that is,  $d_{k-1}^T (g_k - g_{k-1}) > 0$ . Hence, starting with (3), we have

$$\begin{aligned} g_k^T d_k &= g_k (-g_k + \beta_k^{DHS} d_{k-1}) = -\|g_k\|^2 + \left( \frac{\|g_k\|^2 - \frac{|g_k^T g_{k-1}|}{\|g_k\|^2} g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} \right) g_k^T d_{k-1} = -\|g_k\|^2 + \frac{\|g_k\|^2 g_k^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} \\ &- \frac{\frac{|g_k^T g_{k-1}|}{\|g_k\|^2} g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} g_k^T d_{k-1} = \frac{-\|g_k\|^2 d_{k-1}^T (g_k - g_{k-1}) + \|g_k\|^2 g_k^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} - \frac{\frac{|g_k^T g_{k-1}|}{\|g_k\|^2} g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} g_k^T d_{k-1} \\ &= \frac{\|g_k\|^2 g_k^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} - \frac{\frac{|g_k^T g_{k-1}|}{\|g_k\|^2} g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} g_k^T d_{k-1} < \frac{\|g_k\|^2 g_k^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} < 0 \end{aligned} \tag{23}$$

In (23), the first equality was obtained by finding the inner product of (3) and substituting  $\beta_k^{DHS}$  for  $\beta_k$ . The first inequality is obvious because  $(|g_k^T g_{k-1}|/\|g_k\|^2) > 1$  from the first assumption above, while the second was obtained from the fact that  $g_{k-1}^T d_{k-1} < 0$ .

**Case II:** If  $g_k^T g_{k-1} < (|g_k^T g_{k-1}|/\|g_k\|^2) g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then (19) yields  $\beta_k^{DPH} = \beta_k^{DPRP}$ . Notice that the second assumption implies that  $g_k^T d_{k-1} - \|g_{k-1}\|^2 < g_{k-1}^T d_{k-1}$ . Thus, starting from (3), we obtain the following result:

$$\begin{aligned} g_k^T d_k &= g_k (-g_k + \beta_k^{DPRP} d_{k-1}) = -\|g_k\|^2 + \left\{ \|g_k\|^2 - \left( \frac{|g_k^T g_{k-1}|}{\|g_k\|^2} \right) g_k^T g_{k-1} \right\} g_k^T d_{k-1} \\ &= -\|g_k\|^2 + \frac{\|g_k\|^2 g_k^T d_{k-1}}{\|g_{k-1}\|^2} - \frac{\left( \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} \right) g_k^T g_{k-1}}{\|g_{k-1}\|^2} g_k^T d_{k-1} = \frac{(g_k^T d_{k-1} - \|g_{k-1}\|^2) \|g_k\|^2}{\|g_{k-1}\|^2} \\ &- \frac{\left( \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} \right) g_k^T g_{k-1}}{\|g_{k-1}\|^2} g_k^T d_{k-1} < \frac{(g_k^T d_{k-1} - \|g_{k-1}\|^2) \|g_k\|^2}{\|g_{k-1}\|^2} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_{k-1}^T d_{k-1} < 0 \end{aligned} \tag{24}$$

**Case III:** If  $g_k^T g_{k-1} \geq (|g_k^T g_{k-1}|/\|g_k\|^2) g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then from (19),  $\beta_k^{DPH} = \beta_k^{HS}$ . Proceeding from the inner product of (3) with  $g_k$  gives

$$\begin{aligned} g_k^T d_k &= g_k (-g_k + \beta_k^{HS} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} g_k^T d_{k-1} = \frac{\|g_k\|^2 g_{k-1}^T d_{k-1} - g_k^T g_{k-1} \cdot g_k^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} \\ &= \frac{\|g_k\|^2 g_{k-1}^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} - \frac{g_k^T g_{k-1} \cdot g_k^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} < \frac{\|g_k\|^2 g_{k-1}^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} < 0 \end{aligned} \tag{25}$$

**Case IV:** If  $g_k^T g_{k-1} \geq (|g_k^T g_{k-1}|/\|g_k\|^2) g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) < \|g_k\|^2$ , then from (19),  $\beta_k^{DPH} = \beta_k^{PRP}$ . The second assumption in this case allows us to set  $g_k^T d_{k-1} - \|g_k\|^2 < g_{k-1}^T d_{k-1}$  and beginning from (3) as in other cases, we obtain

$$\begin{aligned} g_k^T d_k &= g_k (-g_k + \beta_k^{PRP} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|g_{k-1}\|^2} g_k^T d_{k-1} = \frac{(g_k^T d_{k-1} - \|g_k\|^2) \|g_k\|^2}{\|g_{k-1}\|^2} \\ &- \frac{g_k^T g_{k-1} \cdot g_k^T d_{k-1}}{\|g_{k-1}\|^2} < \frac{\|g_k\|^2 (g_k^T d_{k-1} - \|g_k\|^2)}{\|g_{k-1}\|^2} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_{k-1}^T d_{k-1} < 0 \end{aligned} \tag{26}$$

The descent property is satisfied in (23)-(26). Hence, the sequence of search directions generated by the DPH method satisfies the descent property.

**DHW Method**

**Case I:** If  $(g_k^T g_{k-1} / \|g_{k-1}\|)^2 < (|g_k^T g_{k-1}| / \|g_{k-1}\|)^2 g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then (20) yields  $\beta_k^{DHW} = \beta_k^{DHS}$ . Clearly, this result may be established as in Case I of the DPH method.

**Case II:** If  $(g_k^T g_{k-1} / \|g_{k-1}\|)^2 < (|g_k^T g_{k-1}| / \|g_{k-1}\|)^2 g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then (20) yields  $\beta_k^{DHW} = \beta_k^{DPRP}$ . The proof of this result is similar to Case II in the DPH method.

**Case III:** If  $(g_k^T g_{k-1} / \|g_{k-1}\|)^2 \geq (|g_k^T g_{k-1}| / \|g_{k-1}\|)^2 g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then (20) yields  $\beta_k^{DHW} = \beta_k^{WHS}$ . Starting from (3) yields the following result:

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{WHS} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - (g_k^T g_{k-1} / \|g_{k-1}\|)^2}{d_{k-1}^T (g_k - g_{k-1})} \cdot g_k^T d_{k-1} \\ &= \frac{-\|g_k\|^2 (g_k^T d_{k-1} - g_{k-1}^T d_{k-1}) + \|g_k\|^2 g_k^T d_{k-1} - (g_k^T g_{k-1} / \|g_{k-1}\|)^2 g_k^T d_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} \\ &= \frac{\|g_k\|^2}{d_{k-1}^T (g_k - g_{k-1})} g_{k-1}^T d_{k-1} - \frac{(g_k^T g_{k-1} / \|g_{k-1}\|)^2}{d_{k-1}^T (g_k - g_{k-1})} g_k^T d_{k-1} < \frac{\|g_k\|^2}{d_{k-1}^T (g_k - g_{k-1})} g_{k-1}^T d_{k-1} < 0 \end{aligned} \tag{27}$$

The first inequality was obtained from the fact that  $g_k - g_{k-1} \geq |g_k - g_{k-1}|$  (from the first assumption) and  $d_{k-1}^T (g_k - g_{k-1}) > 0$  (the second assumption). These two inequalities enforced the positivity of the second term in the expression of the last equality. The last inequality was obtained from the fact that  $g_{k-1}^T d_{k-1} < 0$ .

**Case IV:** Let  $(g_k^T g_{k-1} / \|g_{k-1}\|)^2 \geq (|g_k^T g_{k-1}| / \|g_{k-1}\|)^2 g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) < \|g_{k-1}\|^2$ . In this case,  $\beta_k$  is computed as  $\beta_k^{DHW} = \beta_k^{HPRP}$ . The inner product of (3) and after some algebra give

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{HPRP} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - (g_k^T g_{k-1} / \|g_{k-1}\|)^2}{\|g_{k-1}\|^2} \cdot g_k^T d_{k-1} \\ &= \frac{-\|g_k\|^2 \|g_{k-1}\|^2 + \|g_k\|^2 g_k^T d_{k-1} - (g_k^T g_{k-1} / \|g_{k-1}\|)^2 g_k^T d_{k-1}}{\|g_{k-1}\|^2} = \frac{\|g_k\|^2 (g_k^T d_{k-1} - \|g_{k-1}\|^2)}{\|g_{k-1}\|^2} \\ &\quad - \frac{(g_k^T g_{k-1} / \|g_{k-1}\|)^2}{\|g_{k-1}\|^2} g_k^T d_{k-1} < \frac{\|g_k\|^2}{\|g_{k-1}\|} g_{k-1}^T d_{k-1} - \frac{(g_k^T g_{k-1} / \|g_{k-1}\|)^2}{\|g_{k-1}\|^2} g_k^T d_{k-1} \\ &< \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_{k-1}^T d_{k-1} < 0 \end{aligned} \tag{28}$$

Therefore, since  $g_k^T d_k < 0$  for all  $k \geq 1$  in all the cases, the DHW method generates descent directions.

**DV Method**

**Case I:** Suppose  $(\|g_k\| / \|g_{k-1}\|) g_k^T g_{k-1} < (|g_k^T g_{k-1}| / \|g_{k-1}\|) g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ ,  $\beta_k^{DV}$  in (21) reduces to  $\beta_k^{DHS}$ . The proof of the descent property of  $d_k$  in this case is similar to Case I of the DPH and DHW methods.

**Case II:** Suppose  $(\|g_k\| / \|g_{k-1}\|) g_k^T g_{k-1} < (|g_k^T g_{k-1}| / \|g_{k-1}\|) g_k^T g_{k-1}$  and  $d_{k-1}^T (g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then  $\beta_k$  is yielded by (17). The proof of the descent property of in this case is also similar to Case II of the DPH and DHW methods.

**Case III:** Let  $(\|g_k\|/\|g_{k-1}\|)g_k^T g_{k-1} \geq (|g_k^T g_{k-1}|/\|g_k\|^2)g_k^T g_{k-1}$  and  $d_{k-1}^T(g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ . These conditions yield  $\beta_k^{DV} = \beta_k^{VHS}$  and beginning with (3), the following result is obtained.

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{VHS} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - (\|g_k\|/\|g_{k-1}\|)g_k^T g_{k-1}}{d_{k-1}^T(g_k - g_{k-1})} \cdot g_k^T d_{k-1} \\ &= \frac{\|g_k\|^2 g_{k-1}^T d_{k-1}}{d_{k-1}^T(g_k - g_{k-1})} - \frac{(\|g_k\|/\|g_{k-1}\|)g_k^T g_{k-1}}{d_{k-1}^T(g_k - g_{k-1})} \cdot g_k^T d_{k-1} = \frac{\|g_k\|^2 g_{k-1}^T d_{k-1}}{d_{k-1}^T(g_k - g_{k-1})} \\ &\quad - \frac{\|g_k\|^2 \cos \theta_k}{d_{k-1}^T(g_k - g_{k-1})} \cdot g_k^T d_{k-1} < \frac{\|g_k\|^2 g_{k-1}^T d_{k-1}}{d_{k-1}^T(g_k - g_{k-1})} - \frac{\|g_k\|^2 \cos \theta_k}{d_{k-1}^T(g_k - g_{k-1})} \cdot g_{k-1}^T d_{k-1} \\ &= \frac{\|g_k\|^2 (1 - \cos \theta_k)}{d_{k-1}^T(g_k - g_{k-1})} g_{k-1}^T d_{k-1} < 0 \end{aligned} \quad (29)$$

This result is explained as follows: the first supposition holds if  $g_k^T g_{k-1} > 0$ , and so we have  $0 < \cos \theta_k < 1$ , where  $\theta_k$  is the angle between  $g_k$  and  $g_{k-1}$ . Thus, it is easy to see that  $(\|g_k\|/\|g_{k-1}\|)g_k^T g_{k-1} = \|g_k\|^2 \cos \theta_k$ , thus establishing the last equality.

**Case IV:** Suppose  $(\|g_k\|/\|g_{k-1}\|)g_k^T g_{k-1} \geq (|g_k^T g_{k-1}|/\|g_k\|^2)g_k^T g_{k-1}$  and  $d_{k-1}^T(g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then  $\beta_k^{DV} = \beta_k^{PRP}$  and beginning with (3), the following result is obtained.

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{PRP} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - (\|g_k\|/\|g_{k-1}\|)g_k^T g_{k-1}}{\|g_{k-1}\|^2} g_k^T d_{k-1} \\ &= -\|g_k\|^2 + \frac{\|g_k\|^2 - \|g_k\|^2 \cos \theta_k}{\|g_{k-1}\|^2} g_k^T d_{k-1} = -\|g_k\|^2 + \frac{\|g_k\|^2 (1 - \cos \theta_k)}{\|g_{k-1}\|^2} g_k^T d_{k-1} \\ &< -\|g_k\|^2 + \frac{\|g_k\|^2 (1 - \cos \theta_k)}{\|g_{k-1}\|^2} (\|g_{k-1}\|^2 + g_{k-1}^T d_{k-1}) < -\|g_k\|^2 \cos \theta_k \\ &\quad + \frac{\|g_k\|^2 (1 - \cos \theta_k)}{\|g_{k-1}\|^2} g_{k-1}^T d_{k-1} < \frac{\|g_k\|^2 (1 - \cos \theta_k)}{\|g_{k-1}\|^2} g_{k-1}^T d_{k-1} < 0 \end{aligned} \quad (30)$$

Results in (29) and (30) together with cases I and II show that the DV method generates descent directions.

### DM Method

**Case I:** If  $(\|g_k\|/\|g_{k-1}\|)|g_k^T g_{k-1}| < (|g_k^T g_{k-1}|/\|g_k\|^2)g_k^T g_{k-1}$  and  $d_{k-1}^T(g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then  $\beta_k^{DM} = \beta_k^{DHS}$ . In this case,  $g_k^T g_{k-1} < 0$ , similar to results in Cases I of the DPR, DHW and DV methods above.

**Case II:** If  $(\|g_k\|/\|g_{k-1}\|)|g_k^T g_{k-1}| < (|g_k^T g_{k-1}|/\|g_k\|^2)g_k^T g_{k-1}$  and  $d_{k-1}^T(g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then  $\beta_k^{DM} = \beta_k^{PRP}$ . In this case also  $g_k^T g_{k-1} < 0$  as in Cases II of the DPR, DHW and DV methods as earlier proved.

**Case III:** If  $(\|g_k\|/\|g_{k-1}\|)|g_k^T g_{k-1}| \geq (|g_k^T g_{k-1}|/\|g_k\|^2)g_k^T g_{k-1}$  and  $d_{k-1}^T(g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then  $\beta_k$  is yielded by  $\beta_k^{MHS}$  as in (14). It is clear that  $|g_k^T g_{k-1}| > 0$ , and so  $\beta_k^{MHS} < \beta_k^{DY}$ . This and the inner product of (3) with  $g_k$  give

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{MHS} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - (\|g_k\|/\|g_{k-1}\|)|g_k^T g_{k-1}|}{d_{k-1}^T(g_k - g_{k-1})} g_k^T d_{k-1} \\ &\leq -\|g_k\|^2 + \frac{\|g_k\|^2}{d_{k-1}^T(g_k - g_{k-1})} g_k^T d_{k-1} = \frac{-\|g_k\|^2 (g_k^T d_{k-1} - g_{k-1}^T d_{k-1}) + \|g_k\|^2 g_k^T d_{k-1}}{d_{k-1}^T(g_k - g_{k-1})} < 0 \end{aligned} \quad (31)$$



The last inequality follows as  $d_{k-1}^T(g_k - g_{k-1}) > 0$  since  $\|g_{k-1}\|^2 > 0$ .

**Case IV:** If  $(\|g_k\|/\|g_{k-1}\|)|g_k^T g_{k-1}| \geq (|g_k^T g_{k-1}|/\|g_k\|^2)g_k^T g_{k-1}$  and  $d_{k-1}^T(g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then  $\beta_k^{DM} = \beta_k^{MVPRP}$ . Starting from (3) yields

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{MVPRP} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - (\|g_k\|/\|g_{k-1}\|)|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T d_{k-1} \\ &\leq -\|g_k\|^2 + \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_k^T d_{k-1} < -\|g_k\|^2 + \frac{\|g_k\|^2}{\|g_{k-1}\|^2} (\|g_{k-1}\|^2 + g_{k-1}^T d_{k-1}) \\ &= \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_{k-1}^T d_{k-1} < 0 \end{aligned} \tag{32}$$

The first inequality follows the fact that  $\beta_k^{MVPRP} \leq \beta_k^{FR}$  as  $\|g_{k-1}\|^2 > 0$ , and the second inequality holds according to the second assumption. Hence, the DM method is a descent method. Therefore, the descent property is satisfied for all the methods.  $\square$

The following important result can easily be inferred from Theorem 1. This result is important in establishing the global convergence of the Hybrid Algorithm. The result is shown for only the DPH method as the same approach is valid for the other three methods.

*Lemma 1*

The inequality  $0 \leq \beta_k \leq \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}$  always holds for every  $k \geq 1$ , where  $\beta_k = \{\beta_k^{DPH}, \beta_k^{DHW}, \beta_k^{DV}, \beta_k^{DM}\}$ .

*Proof*

From (19) it is known that  $\beta_k^{DPH} \geq 0$ , which establishes the first part of the inequality. Now, suppose  $\beta_k^{DPH} = 0$  and  $g_k \neq 0$ , then  $\frac{g_k^T d_k}{g_{k-1}^T d_{k-1}} > 0$  since  $g_k^T d_k$  and  $g_{k-1}^T d_{k-1}$  (both from Theorem 1). By assuming  $\beta_k^{DPH} > 0$ , the following four cases are established.

**Case I:** If  $g_k^T g_{k-1} < \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}$  and  $d_{k-1}(g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then (19) yields  $\beta_k^{DPH} = \beta_k^{DHS}$ . Hence, Theorem 1 together with (23) gives

$$\beta_k^{DPH} = \frac{\|g_k\|^2 - \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}}{d_{k-1}(g_k - g_{k-1})} < \frac{\|g_k\|^2}{d_{k-1}(g_k - g_{k-1})} < \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}} \tag{33}$$

**Case II:** If  $g_k^T g_{k-1} < \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}$  and  $d_{k-1}(g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then (19) yields  $\beta_k^{DPH} = \beta_k^{DPRP}$ . Hence, Theorem 1 together with (24) gives

$$\beta_k^{DPH} = \frac{\|g_k\|^2 - \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}}{\|g_{k-1}\|^2} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2} < \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}} \tag{34}$$

**Case III:** If  $g_k^T g_{k-1} \geq \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}$  and  $d_{k-1}(g_k - g_{k-1}) \geq \|g_{k-1}\|^2$ , then (19) yields  $\beta_k^{DPH} = \beta_k^{HS}$ . Hence, Theorem 1 together with (25) gives

$$\beta_k^{DPH} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}(g_k - g_{k-1})} < \frac{\|g_k\|^2}{d_{k-1}(g_k - g_{k-1})} < \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}} \tag{35}$$

**Case IV:** If  $g_k^T g_{k-1} \geq \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} g_k^T g_{k-1}$  and  $d_{k-1}(g_k - g_{k-1}) < \|g_{k-1}\|^2$ , then (19) yields  $\beta_k^{DPH} = \beta_k^{PRP}$ . Hence, Theorem 1 together with (26) gives

$$\beta_k^{DPH} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2} < \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}} \tag{36}$$

Therefore, the inequality  $0 \leq \beta_k \leq \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}$  is true for  $\beta_k = \beta_k^{DPH}$ . □

Next, the global convergence result of Algorithm 1 is established. The result is established under the following assumptions.

**Assumption 1:** 1. Bound on the Objective Function:  $f(x)$  is bounded from below on the level set  $\Omega = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ , where  $x_0$  is the initial point.

**Assumption 2:** Lipschitz Condition: within some neighbourhood  $N$  of  $\Omega$ ,  $f(x)$  is continuously differentiable, and its gradient is Lipschitz continuous, that is, for all  $x, y \in N$ , there exists a constant  $L \geq 0$  such that

$$\|g(x) - g(y)\| \leq L\|x - y\| \tag{37}$$

Together with Assumptions 1 and 2, the following important results is usually employed to establish the global convergence results of CG algorithms.

*Lemma 2*

Let an iterative scheme of the form (2) where  $d_k$  is a descent direction and  $\alpha_k$  satisfies the Wolfe line search conditions (4) and (6). If assumptions 1 and 2 hold, then,

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \tag{38}$$

Lemma 2 is the famous Zoutendijk condition, the proof of which may be found in [7]. Using Lemma 2 and the above assumptions, the following global convergence result of the Algorithm 1 is presented.

*Theorem 2*

Let  $d_k$  be generated by the iterative rules (2)-(3) with  $\beta_k$  computed from Line 5 of Algorithm 1. If Assumptions 1 and 2 hold, then by Lemmas 1 and 2,  $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$ .

*Proof*

Suppose by contradiction that the conclusion does not hold, that is,  $\lim_{k \rightarrow \infty} \inf \|g_k\| \neq 0$ . However, since  $\|g_k\| > 0$ , then there exists a constant  $\zeta > 0$  such that  $\|g_k\| \geq \zeta, \forall k$ . Starting with (3) gives

$$\|d_k\|^2 = \beta_k^2 \|d_{k-1}\|^2 - 2d_k^T g_k - \|g_k\|^2 \tag{39}$$

Dividing both sides of (35) by  $(d_k^T g_k)^2$  and using Lemma 1, the following was obtained

$$\begin{aligned} \frac{\|d_k\|^2}{(d_k^T g_k)^2} &= \frac{\|d_{k-1}\|^2}{(d_{k-1}^T g_{k-1})^2} - \frac{2}{d_k^T g_k} - \frac{\|g_k\|^2}{(d_k^T g_k)^2} = \frac{\|d_{k-1}\|^2}{(d_{k-1}^T g_{k-1})^2} - \left( \frac{1}{\|g_k\|} + \frac{\|g_k\|}{d_k^T g_k} \right)^2 + \frac{1}{\|g_k\|^2} \\ &\leq \frac{\|d_{k-1}\|^2}{(d_{k-1}^T g_{k-1})^2} + \frac{1}{\|g_k\|^2} \end{aligned} \tag{40}$$

Since  $\|d_1\|^2 / (d_1^T g_1)^2 = 1 / (d_1^T g_1)^2$ , (35) gives  $\left( \|d_{k-1}\|^2 / (d_{k-1}^T g_{k-1})^2 \right) \leq \sum_{i=1}^k \left( 1 / \|g_i\|^2 \right) \leq k / \zeta, \forall k$ . This implies  $\sum_{k \geq 1} (g_k^T d_k)^2 / \|d_k\|^2 = \infty$ , which contradicts (34). □

### 5. Numerical Experiments

In this section, we describe the numerical implementations of the new hybrid methods and compare their performances with other existing methods based on four indicators. The indicators for profiling the performance of

our methods with other methods are: number of iterations (Table 1), the final value of the objective function (Table 2), the infinite norm of the final value of the gradient (Table 3) and the CPU time of executing Algorithm 1 (Table 4). All the test functions were drawn from [4], many of which are contained in the CUTer library described in [5].

Hybrid CG Algorithm was coded and implemented on Matlab R2016a installed on a PC with Windows 10 OS and 2G RAM. Varying dimensions of 500, 1000, 5000 and 10000 were used except for very few cases where, due to problem peculiarity, we used much-reduced dimensions to obtain complete result output by all the methods. The new methods were compared with other existing methods based on the indicators mentioned earlier adopting the Dolan and Mor [9] performance profile theory. The theory outlines the comparison of the performance of a set of methods  $S$  on a set of problem  $P$ . In this case, for instance,  $S$  denotes the set of different CG methods implemented on our algorithm while  $P$  is the set of unconstrained test functions. The theory is described in the following paragraphs.

Suppose there exist  $n_s$  methods and  $n_p$  problems, for each problem  $p$  and method  $s$ , then,  $r_{p,s}$  is the computing cost (number of iterations, CPU time, function or gradient evaluations) required to solve problem  $p$  by method  $s$ . If  $I_{p,s}$  is an attribute (say, number of iterations or function evaluation, etc) for problem  $p \in P$  by method  $s \in S$ , then the comparison between the different methods is based on the performance ratio given by

$$r_{p,s} = \frac{I_{p,s}}{\min \{I_{p,s} : s \in S\}} \quad (41)$$

Equation (37) gives the required number of iterations for solving problem  $p \in P$  with method  $s \in S$ . Consequently, if there exists a parameter  $r_K$ , which is large enough so that  $r_K \geq r_{p,s}$  for all  $p, s$ . Equality holds, that is  $r_K = r_{p,s}$ , only when the chosen method  $s$  does not solve the problem  $p$ . Based on (37), the cumulative distribution function for the performance ratio  $r_{p,s}$ , is defined by

$$\rho_s(\tau) = \frac{1}{n_p} |p \in P : \log r_{p,s} \leq \tau| \quad (42)$$

where  $|\cdot|$  represents cardinality,  $\rho_s(\tau)$  is the probability, in relation to method  $s$ , that  $r_{p,s}$  is within a factor  $\tau \in \mathbb{R}^n$ . At  $\tau = 1$ , the value of  $\rho_s(\tau)$  is the probability that the method will out-performs the other methods.

The comparison of methods using the theory described above is outlined for each of the performance indicators as follows: Set 1: DPRP, DHS, DPH, DHW, DV, DM and CG\_DESCENT and Set 2: VPRP, VHS, DPH, DHW, DV, DM and CG\_DESCENT. The CG\_DESCENT method, being one of the most efficient CG methods, was selected in both sets to profile the performance of the proposed methods.

At the start of the simulation, the value of  $\alpha_k$  at  $k = 0$  is set to zero (0) while other values of  $\alpha_k$  are computed using SWLS (4)-(5). The stopping criterion is given by  $\|g_k\| \leq \epsilon$ , where  $\epsilon = 10^{-6}$ . Other parameters of the algorithm are set as follows:  $\delta = 0.0001$  and  $\sigma = 0.9$ . The numerical results are presented in Tables 1-4. Note that in Tables (3) and (4), the entries of a few cells are NaN meaning undefined or unrepresentative floating numbers. In instances where this occurs, the numbers of iterations and the CPU times were still recorded in Table 1 and Table 2, respectively.

The following Figures 1-8 produced from the numerical values in Tables 1-4 indicate the performance profiles of our hybrid methods against some non-hybrid methods. The performance profile with respect to the numbers of iterations is illustrated in Figures 1. Clearly, in Figure 1, DPH and DV gave better performance than DHS, DPRP, CG\_DESCENT (Figure 1 (left) at  $\tau = 1$  and  $\tau = 2$ ) and VPRP, VHS, CG\_DESCENT (Figure 1 (right) at  $\tau = 1$  and  $\tau = 2$ ), respectively. The comparison based on CPU time of execution (see Figures 2) shows that DPH and DV methods remain very efficient and highly competitive (see Figure 2 (left) at  $\tau = 2$ ). Similarly, according to Figure 2 (right), DPH and DV methods perform better than the CG\_DESCENT method. In term of the function value, DM produced the best performance (Figure 3). From Figure 4, the average performance of DV and DM are approximately the better methods.

In summary, in terms of computation time and number of iterations, DPH and DV were observed to be very efficient and competitive. DM method gave a better performance based on objective function evaluation. On the average, all the new methods either out-perform or very competitive with CG\_DESCENT method. The following abbreviations were used in the tables: FN-Function Name; EBD1-Extended Block Diagonal-1, EH-Extended

Himmelblau, GR-Generalized Rosenbrock, EWH-Extended White & Holst, GPSC1-Generalized PSC1, ET1-Extended Tridiagonal-1, VD-Variable Dimension, PQD-Perturbed Quadratic Diagonal, GWH-Generalized White & Holst, CGD-CG\_DESCENT.

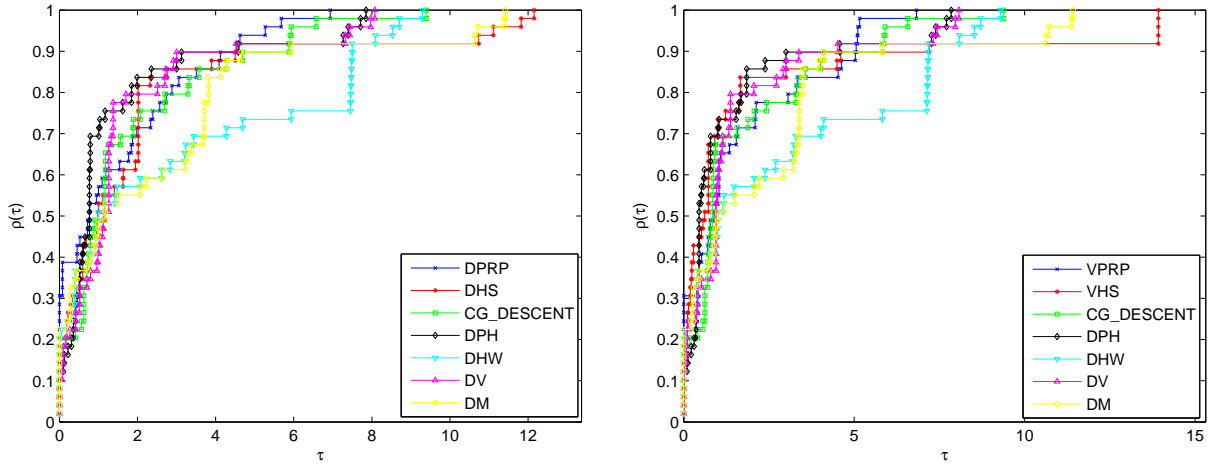


Figure 1. The performance profile of DPRP, DHS, DPH, DHW, DV, DM & CG\_DESCENT (left) and VPRP, VHS, DPH, DHW, DV, DM & CG\_DESCENT (right) with respect to the numbers of iterations.

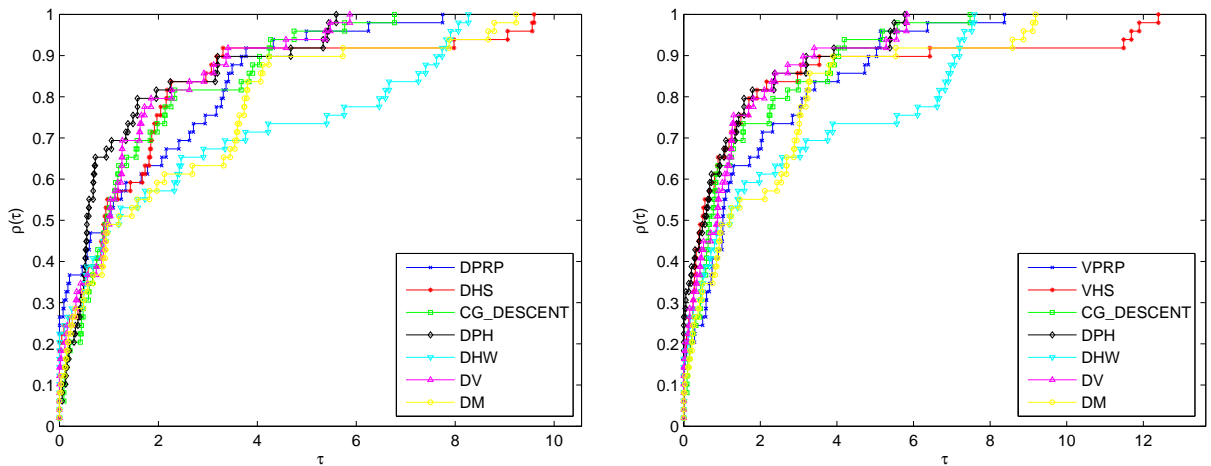


Figure 2. The performance profile of DPRP, DHS, DPH, DHW, DV, DM & CG\_DESCENT (left) and VPRP, VHS, DPH, DHW, DV, DM & CG\_DESCENT (right) with respect to CPU time.

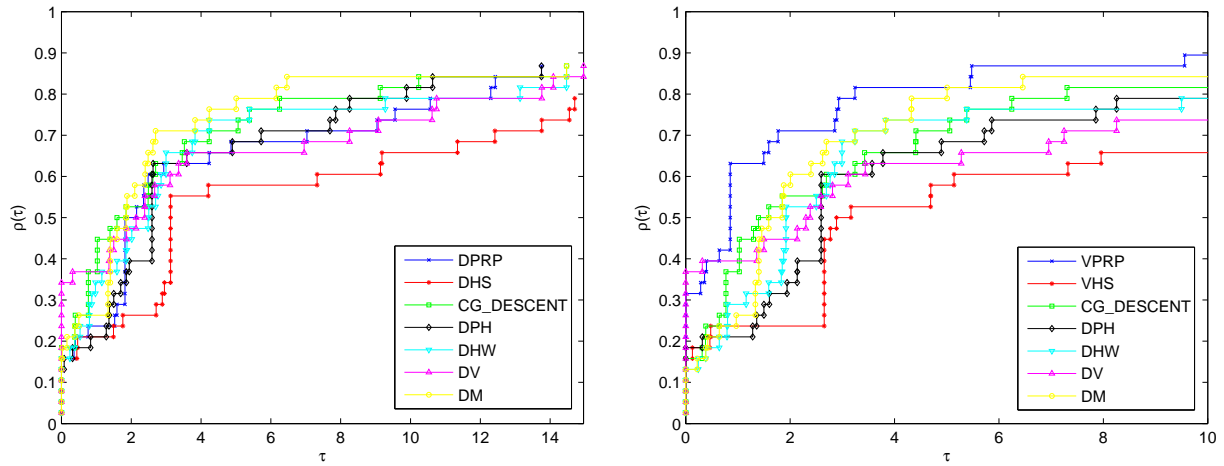


Figure 3. The performance profile of DPRP, DHS, DPH, DHW, DV, DM & CG\_DESCENT (left) and VPRP, VHS, DPH, DHW, DV, DM & CG\_DESCENT (right) with respect to the values of  $f$ .

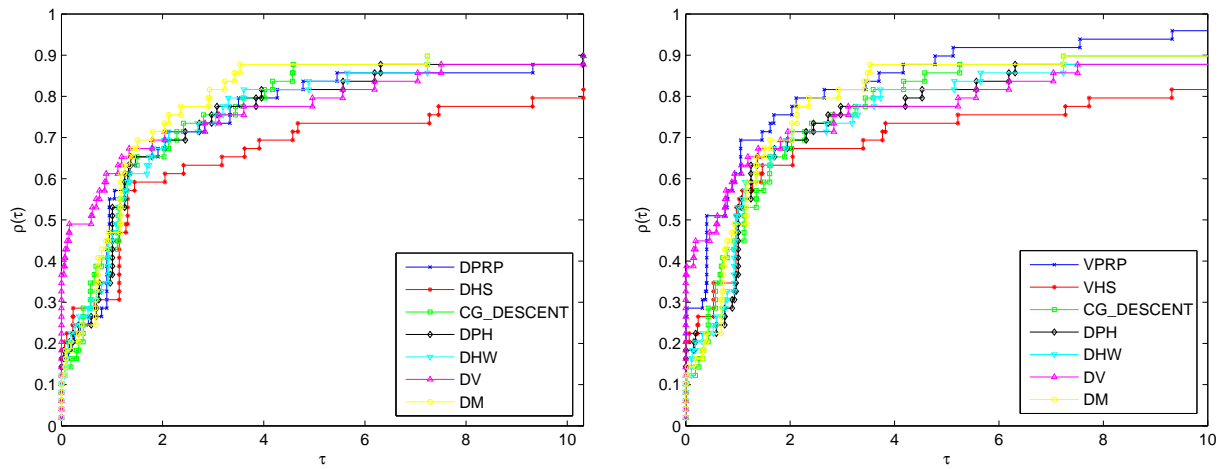


Figure 4. The performance profile of DPRP, DHS, DPH, DHW, DV, DM & CG\_DESCENT (left) and VPRP, VHS, DPH, DHW, DV, DM & CG\_DESCENT (right) with respect to the values of  $\|g\|$ .

Table 1. Numerical test results for the numbers of iterations after running Algorithm 1 for different values of  $n$ . Functions ARGLINB, HIMMELH and QUADRATIC behave differently from other functions, forcing the use of smaller dimensions. ET1 under VHS produces the largest number of iterations compared to other functions which use the same values for  $n$ . DV has the least average number of iterations (175.73) followed by DPH (191.10).

FN	$n$	CG METHOD								
		VPRP	VHS	DPRP	DHS	DPH	DHW	DV	DM	CGD
EBD1	500	66	61	95	69	80	52	57	26	96
	1000	78	79	69	82	52	55	73	42	62
	5000	89	72	60	61	72	9	68	42	107
	10000	110	77	71	101	67	54	52	71	59
Diagonal-4	500	167	5	85	5	18	46	40	47	49
	1000	171	5	87	5	18	47	33	48	50
	5000	179	5	91	8	18	49	52	50	52
	10000	181	22	91	29	18	50	45	51	53
EH	500	25	60	75	77	26	25	37	19	31
	1000	26	60	68	59	27	25	39	19	31
	5000	26	60	140	95	29	25	42	19	33
	10000	26	61	105	77	30	25	43	20	33
GR	500	76	40	39	56	64	37	96	45	57
	1000	76	40	39	56	64	37	96	45	57
	5000	76	40	39	56	64	37	96	45	57
	10000	76	40	39	56	64	37	96	45	57
EWH	500	46	76	37	150	63	6570	89	484	83
	1000	46	76	37	150	63	6473	89	484	83
	5000	46	76	37	150	63	6497	89	484	83
	10000	46	76	37	150	63	6497	89	484	83
GPSC1	500	2515	589	2095	353	3087	3824	1150	6146	1314
	1000	2021	463	1158	339	1344	2421	1932	3542	761
ET1	500	3	46520	3	5124	462	817	479	4749	2013
	1000	3	46520	3	6632	504	1107	520	5020	286
	5000	3	46520	3	10862	628	1253	809	8228	180
	10000	3	46520	3	13652	692	1883	752	8017	176
VD	500	41	3536	41	358	32	24	32	24	24
	1000	36	34	36	34	25	25	25	25	25
QUADRATIC	50	207	84	172	59	93	157	71	132	382
PQD	500	125	15	125	15	20	63	20	63	63
	1000	16	23	16	23	23	30	23	30	30
	5000	3179	28	3179	26	37	1590	37	1590	1590
	10000	965	28	965	25	36	483	36	483	483
ARGLINB	10	65	23	32	19	15	34	15	34	15
	20	1411	43	1386	27	138	698	44	679	694
HIMMELH	2	33	13	25	34	12	21	11	21	20
	4	31	21	26	41	19	22	20	22	20
LIARWHD	500	24	16	24	16	14	14	15	14	14
	1000	27	16	27	16	15	14	15	14	14
	5000	18	21	18	21	21	30	19	30	30
	10000	18	22	18	22	23	31	20	31	31
QUARTC	500	390	386	390	386	392	389	386	389	195
	1000	257	313	257	313	320	311	313	311	156
	5000	185	181	185	181	191	8	184	8	8
	10000	148	146	148	146	6	8	6	8	8
GWH	500	46	76	37	150	63	6677	89	522	84
	1000	46	76	37	150	63	6594	89	522	84
	5000	46	76	37	150	63	6691	89	522	84
	10000	46	76	37	150	63	6613	89	472	88

Table 2. Numerical test results for the CPU time of Algorithm 1. The average CPU time for the methods are: VPRP (2.3302), VHS (24.8496), DPRP (1.6083), DHS (5.7417), CGD (1.0447), DPH (0.9834), DHW (14.7620), DV (0.9938), DM (5.5049), with DPH giving the least value, followed by DV.

FN	n	CG METHOD								
		VPRP	VHS	DPRP	DHS	DPH	DHW	DV	DM	CGD
EBD1	500	0.46	0.268	0.481	0.338	0.573	0.327	0.304	0.226	1.078
	1000	0.649	0.168	0.358	0.366	0.204	0.187	0.313	0.204	0.3
	5000	1.417	0.582	1	0.452	0.459	0.05	0.531	0.322	0.824
	10000	3.18	0.521	1.619	0.693	0.638	0.617	0.728	0.742	0.386
Diagonal-4	500	1.061	0.099	0.398	0.031	1.498	0.162	0.32	0.536	0.465
	1000	0.973	0.028	0.506	0.051	0.143	0.255	0.184	0.18	0.223
	5000	2.635	0.075	1.53	0.112	0.161	0.616	0.369	0.439	0.492
	10000	5.478	0.273	2.212	0.348	0.207	0.622	0.491	0.571	0.606
EH	500	0.208	0.527	0.672	0.761	0.398	0.556	0.282	0.352	0.611
	1000	0.207	0.265	0.396	0.238	0.145	0.126	0.151	0.139	0.136
	5000	0.329	0.466	2.215	0.89	0.242	0.197	0.355	0.386	0.264
	10000	0.513	0.826	2.468	1.137	0.375	0.303	0.513	0.255	0.405
GR	500	0.381	0.213	0.205	0.33	0.369	0.178	0.428	0.192	0.242
	1000	0.395	0.19	0.217	0.267	0.302	0.206	0.449	0.247	0.287
	5000	0.899	0.401	0.429	0.54	0.611	0.381	0.919	0.447	0.651
	10000	1.539	0.627	0.686	0.865	0.972	0.644	1.547	0.73	1.086
EWH	500	0.338	0.52	0.208	0.815	0.3	44.343	1.574	2.872	0.741
	1000	0.524	0.438	0.255	0.965	0.371	68.216	0.782	3.094	0.653
	5000	0.795	0.999	0.563	2.005	0.909	127.357	1.745	7.491	1.237
	10000	1.419	1.674	0.939	3.379	1.379	184.787	2.253	12.561	2.154
GPSC1	500	9.479	2.328	8.886	1.36	12.066	13.826	4.292	22.807	5.355
	1000	6.886	1.657	3.968	1.154	4.486	8.682	7.126	13.447	2.615
ET1	500	0.046	174.715	0.075	18.874	1.917	3.172	1.793	17.626	8.196
	1000	0.036	192.869	0.049	26.142	1.971	4.328	2.029	19.921	1.312
	5000	0.097	319.511	0.094	72.565	4.089	9.518	5.501	56.464	1.774
	10000	0.172	493.666	0.182	136.981	7.749	17.523	8.103	80.542	2.324
VD	500	0.3	17.024	0.292	1.957	0.264	0.213	0.221	0.198	0.212
	1000	0.249	0.219	0.242	0.228	0.234	0.167	0.171	0.171	0.194
QUADRATIC	50	0.875	0.379	0.671	0.266	0.401	0.624	0.265	0.489	1.333
PQD	500	0.574	0.132	0.216	0.028	0.035	0.093	0.035	0.083	0.083
	1000	0.066	0.093	0.023	0.033	0.033	0.044	0.034	0.044	0.045
	5000	42.777	0.205	24.268	0.113	0.129	6.104	0.144	6.002	6.145
	10000	22.383	0.33	16.431	0.216	0.28	4.025	0.272	4.061	4.162
ARGLINB	10	0.225	0.202	0.176	0.033	0.156	0.177	0.159	0.053	0.045
	20	1.768	0.176	1.714	0.223	0.161	0.732	0.054	0.752	0.769
HIMMELH	2	0.179	0.026	0.038	0.046	0.026	0.035	0.025	0.035	0.035
	4	0.036	0.026	0.033	0.046	0.025	0.027	0.024	0.027	0.027
LIARWHD	500	0.062	0.039	0.165	0.178	0.037	0.037	0.04	0.161	0.041
	1000	0.07	0.034	0.192	0.033	0.034	0.043	0.034	0.031	0.033
	5000	0.12	0.141	0.125	0.131	0.34	0.206	0.114	0.207	0.212
	10000	0.259	0.322	0.475	0.318	0.343	0.453	0.308	0.458	0.463
QUARTC	500	0.423	0.415	0.422	0.404	0.436	0.418	0.418	0.429	0.226
	1000	0.395	0.439	0.398	0.436	0.504	0.444	0.439	0.443	0.193
	5000	0.432	0.392	0.423	0.395	0.43	0.047	0.408	0.048	0.05
	10000	0.545	0.502	0.661	0.498	0.064	0.091	0.065	0.09	0.097
GWH	500	0.271	0.317	0.104	0.379	0.17	24.934	0.213	1.248	0.212
	1000	0.395	0.339	0.121	0.425	0.194	37.292	0.247	1.504	0.23
	5000	0.59	0.685	0.359	1.145	0.484	60.736	0.67	3.93	0.66
	10000	1.071	1.288	0.649	2.153	0.875	99.219	1.253	6.483	1.305

Table 3. Numerical test results indicating the values of  $f$ . Notice that the proposed methods solved all the selected problems, with the exemption of DHW which failed on one instance.

FN	$n$	CG METHOD								
		VPRP	VHS	DPRP	DHS	DPH	DHW	DV	DM	CGD
EBD1	500	8.41E-17	4.35E-14	3.90E-15	7.76E-14	6.91E-15	2.69E-13	4.68E-14	2.62E-15	2.98E-17
	1000	3.80E-17	9.45E-15	2.33E-14	2.50E-14	4.19E-14	2.75E-14	7.62E-14	4.43E-17	5.34E-14
	5000	5.54E-14	4.30E-14	3.85E-15	6.79E-14	4.03E-15	NaN	8.07E-14	2.54E-18	1.93E-16
	10000	5.86E-14	8.09E-15	7.58E-15	2.57E-14	2.46E-14	3.07E-14	3.99E-14	3.25E-14	8.71E-14
Diagonal-4	500	1.08E-15	NaN	7.01E-16	NaN	5.89E-17	1.01E-15	2.42E-17	7.79E-16	8.07E-16
	1000	8.78E-16	NaN	6.03E-16	NaN	1.18E-16	9.41E-16	1.02E-15	7.27E-16	7.26E-16
	5000	7.19E-16	NaN	5.58E-16	NaN	5.89E-16	1.02E-15	3.55E-15	7.90E-16	7.34E-16
	10000	9.15E-16	2.68E-16	1.12E-15	5.10E-15	1.18E-15	9.55E-16	2.91E-15	7.36E-16	6.60E-16
EH	500	5.13E-15	7.99E-16	5.21E-16	2.16E-15	6.16E-15	1.17E-16	1.45E-14	6.18E-16	4.89E-15
	1000	6.18E-17	1.60E-15	1.74E-17	1.01E-14	3.60E-15	2.34E-16	9.40E-15	1.24E-15	9.77E-15
	5000	3.09E-16	7.99E-15	3.37E-15	3.29E-15	4.24E-15	1.17E-15	1.20E-14	6.18E-15	6.57E-15
	10000	6.18E-16	5.55E-15	6.11E-15	1.19E-15	7.34E-15	2.34E-15	4.34E-15	1.21E-15	1.31E-14
GR	500	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00
	1000	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00
	5000	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00
	10000	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00	3.99E+00
EWH	500	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	1.99E-16	3.51E-17	9.27E-17	5.99E-17
	1000	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	2.80E-16	3.51E-17	9.27E-17	5.99E-17
	5000	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	6.12E-17	3.51E-17	9.27E-17	5.99E-17
	10000	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	6.12E-17	3.51E-17	9.27E-17	5.99E-17
GPSC1	500	4.99E+02	4.99E+02	4.99E+02	4.99E+02	4.99E+02	4.99E+02	4.99E+02	4.99E+02	4.99E+02
	1000	9.87E+01	9.87E+01	9.99E+02	9.87E+01	9.87E+01	9.87E+01	9.87E+01	9.87E+01	9.87E+01
ET1	500	6.08E-61	2.64E-10	0	4.49E-09	4.10E-09	6.17E-09	4.11E-09	4.01E-09	6.06E-09
	1000	1.21E-60	5.29E-10	0	5.53E-09	5.23E-09	6.10E-09	5.16E-09	5.72E-09	2.07E-09
	5000	6.08E-60	2.64E-09	0	9.91E-09	8.82E-09	1.17E-08	8.83E-09	8.61E-09	6.58E-09
	10000	1.21E-59	5.29E-09	0	1.25E-08	1.11E-08	1.32E-08	1.11E-08	1.24E-08	4.34E-10
VD	500	7.03E-19	1.88E-17	7.03E-19	3.50E-18	1.37E-18	1.96E-18	1.37E-18	1.96E-18	5.33E-19
	1000	8.02E-21	9.99E-19	8.02E-21	9.99E-19	6.24E-21	8.87E-20	6.24E-21	8.87E-20	8.12E-21
QUADRATIC	50	9.97E-01	-1.00E+00	9.97E-01	-1.00E+00	9.97E-01	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00
PQD	500	1.47E-14	4.88E-15	1.47E-14	4.88E-15	1.38E-14	1.47E-14	1.38E-14	1.47E-14	1.47E-14
	1000	8.90E-16	9.74E-16	8.90E-16	1.21E-15	1.11E-15	3.21E-15	1.11E-15	3.21E-15	3.21E-15
	5000	3.60E-16	3.80E-17	3.60E-16	6.43E-17	1.67E-16	3.60E-16	1.67E-16	3.60E-16	3.60E-16
	10000	9.37E-17	8.37E-17	9.37E-17	4.98E-18	6.00E-17	9.37E-17	6.00E-17	9.37E-17	9.37E-17
ARGLNB	10	2.14E+00	2.14E+00	2.14E+00	2.14E+00	2.14E+00	2.14E+00	2.14E+00	2.14E+00	2.14E+00
	20	4.63E+00	4.63E+00	4.63E+00	NaN	4.63E+00	4.63E+00	4.63E+00	4.63E+00	4.63E+00
HIMMELH	2	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00	-1.00E+00
	4	-2.00E+00	-2.00E+00	-2.00E+00	-2.00E+00	-2.00E+00	-2.00E+00	-2.00E+00	-2.00E+00	-2.00E+00
LIARWHD	500	1.89E-20	3.42E-21	1.89E-20	3.42E-21	9.53E-21	3.24E-21	2.48E-21	3.24E-21	3.24E-21
	1000	2.15E-22	6.84E-21	2.15E-22	6.84E-21	2.85E-25	6.47E-21	4.97E-21	6.47E-21	6.47E-21
	5000	1.62E-23	1.20E-22	1.62E-23	1.20E-22	4.83E-22	1.05E-22	8.46E-23	1.05E-22	1.05E-22
	10000	3.23E-23	0	3.23E-23	0	2.32E-23	2.93E-23	6.67E-23	2.93E-23	2.93E-23
QUARTC	500	1.98E-10	1.98E-10	9.99E-07	1.98E-10	1.98E-10	1.98E-10	1.98E-10	1.98E-10	1.98E-10
	1000	1.57E-10	1.57E-10	1.57E-10	1.57E-10	1.57E-10	1.57E-10	1.57E-10	1.57E-10	1.57E-10
	5000	9.13E-11	9.17E-11	9.13E-11	9.17E-11	9.12E-11	6.61E-15	9.19E-11	6.61E-15	6.61E-15
	10000	7.28E-11	7.23E-11	7.28E-11	7.23E-11	4.03E-12	1.32E-14	4.03E-12	1.32E-14	1.32E-14
GWH	500	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	7.83E-17	3.51E-17	8.85E-17	7.16E-17
	1000	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	2.39E-16	3.51E-17	8.85E-17	7.16E-17
	5000	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	4.15E-17	3.51E-17	8.85E-17	7.16E-17
	10000	6.35E-17	2.21E-16	1.24E-16	3.07E-16	2.12E-16	2.54E-16	3.51E-17	1.27E-16	9.22E-17



Table 4. Numerical test results indicating the values of  $\|g\|$ . These values represent the values of the norm of the gradient of the objective functions,  $f$  at  $x^*$ , the optimal value of  $x$ .

FN	n	CG METHOD								
		VPRP	VHS	DPRP	DHS	DPH	DHW	DV	DM	CGD
EBD1	500	2.59E-08	9.62E-07	1.16E-07	7.89E-07	4.79E-07	9.17E-07	9.64E-07	2.90E-07	3.09E-08
	1000	3.49E-08	2.67E-07	3.62E-07	4.48E-07	8.92E-07	9.45E-07	8.87E-07	1.88E-08	4.50E-07
	5000	8.47E-07	9.55E-07	1.97E-07	7.92E-07	3.58E-07	NaN	8.21E-07	4.51E-09	8.15E-08
	10000	9.62E-07	2.22E-07	4.93E-07	3.41E-07	8.56E-07	8.42E-07	8.63E-07	9.09E-07	8.26E-07
Diagonal-4	500	9.99E-07	NaN	7.71E-07	NaN	1.98E-07	9.33E-07	7.66E-08	8.19E-07	8.31E-07
	1000	8.99E-07	NaN	7.15E-07	NaN	2.79E-07	9.01E-07	5.15E-07	7.91E-07	7.88E-07
	5000	8.14E-07	NaN	6.88E-07	NaN	6.25E-07	9.39E-07	9.53E-07	8.25E-07	7.93E-07
	10000	9.18E-07	4.76E-07	9.73E-07	8.82E-07	8.83E-07	9.07E-07	8.13E-07	7.96E-07	7.52E-07
EH	500	8.68E-07	2.93E-07	2.86E-07	5.70E-07	9.14E-07	1.38E-07	9.88E-07	3.11E-07	6.71E-07
	1000	1.00E-07	4.14E-07	5.87E-08	7.23E-07	4.95E-07	1.95E-07	7.15E-07	4.39E-07	9.48E-07
	5000	2.24E-07	9.27E-07	8.43E-07	6.82E-07	7.28E-07	4.36E-07	7.89E-07	9.83E-07	6.84E-07
	10000	3.17E-07	6.22E-07	9.50E-07	8.33E-07	6.19E-07	6.17E-07	8.30E-07	4.03E-07	8.22E-07
GR	500	7.85E-07	5.47E-07	7.29E-07	9.37E-07	9.00E-07	7.19E-07	3.78E-07	8.54E-07	8.22E-07
	1000	7.85E-07	5.47E-07	7.29E-07	9.37E-07	9.00E-07	7.19E-07	3.78E-07	8.54E-07	8.22E-07
	5000	7.85E-07	5.47E-07	7.29E-07	9.37E-07	9.00E-07	7.19E-07	3.78E-07	8.54E-07	8.22E-07
	10000	7.85E-07	5.47E-07	7.29E-07	9.37E-07	9.00E-07	7.19E-07	3.78E-07	8.54E-07	8.22E-07
EWH	500	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	8.95E-07	4.46E-07	7.14E-07	6.01E-07
	1000	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	9.84E-07	4.46E-07	7.14E-07	6.01E-07
	5000	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	4.81E-07	4.46E-07	7.14E-07	6.01E-07
	10000	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	4.81E-07	4.46E-07	7.14E-07	6.01E-07
GPSC1	500	9.91E-07	9.74E-07	9.43E-07	9.85E-07	9.76E-07	9.83E-07	9.86E-07	1.00E-06	9.93E-07
	1000	9.81E-07	5.81E-07	9.24E-07	7.72E-07	9.76E-07	9.84E-07	7.99E-07	9.78E-07	9.51E-07
ET1	500	9.79E-46	9.33E-08	0.00E+00	9.98E-07	9.86E-07	9.92E-07	9.88E-07	7.82E-07	9.96E-07
	1000	1.38E-45	1.32E-07	0.00E+00	9.66E-07	9.98E-07	9.74E-07	9.88E-07	9.33E-07	3.68E-07
	5000	3.10E-45	2.95E-07	0.00E+00	9.96E-07	9.92E-07	9.68E-07	9.93E-07	9.16E-07	9.21E-07
	10000	4.38E-45	4.17E-07	0.00E+00	9.98E-07	9.94E-07	8.73E-07	9.96E-07	9.06E-07	4.10E-07
VD	500	4.20E-07	3.18E-07	4.21E-07	8.78E-07	3.63E-07	8.13E-07	3.63E-07	8.13E-07	8.12E-07
	1000	2.42E-07	5.88E-07	2.42E-07	2.60E-07	3.64E-07	9.92E-07	3.64E-07	9.92E-07	9.87E-07
QUADRATIC		8.86E-07	8.54E-07	8.16E-07	7.30E-07	8.09E-08	7.40E-07	8.98E-08	4.15E-07	9.76E-07
PQD	500	8.47E-07	4.88E-07	8.47E-07	4.88E-07	8.22E-07	8.47E-07	8.22E-07	8.47E-07	8.47E-07
	1000	3.57E-07	3.74E-07	3.57E-07	4.18E-07	3.98E-07	6.79E-07	3.98E-07	6.79E-07	6.79E-07
	5000	9.84E-07	3.20E-07	9.84E-07	4.16E-07	6.71E-07	9.84E-07	6.71E-07	9.84E-07	9.84E-07
	10000	9.85E-07	9.31E-07	9.85E-07	2.27E-07	7.88E-07	9.85E-07	7.88E-07	9.85E-07	9.85E-07
ARGLINB	10	7.92E-07	9.21E-07	8.65E-07	2.10E-07	8.65E-07	7.35E-07	8.65E-07	7.35E-07	4.52E-07
	20	6.93E-07	1.99E-08	3.19E-08	NaN	4.59E-07	2.64E-07	3.36E-08	5.11E-08	7.56E-07
HIMMELH	2	9.63E-07	7.31E-07	6.07E-07	8.06E-07	4.20E-07	5.27E-07	5.36E-08	6.20E-07	2.85E-07
	4	4.71E-07	3.18E-07	8.16E-07	6.13E-07	1.15E-07	7.45E-07	9.96E-07	8.77E-07	8.05E-07
LIARWHD	500	5.54E-07	2.35E-07	5.54E-07	2.35E-07	3.93E-07	2.29E-07	2.01E-07	2.29E-07	2.29E-07
	1000	1.18E-07	6.66E-07	1.18E-07	6.66E-07	4.30E-09	6.48E-07	5.67E-07	6.48E-07	6.48E-07
	5000	1.62E-07	4.41E-07	1.62E-07	4.41E-07	8.85E-07	4.13E-07	3.70E-07	4.13E-07	4.13E-07
	10000	4.58E-07	0.00E+00	4.58E-07	0.00E+00	3.88E-07	4.36E-07	6.57E-07	4.36E-07	4.36E-07
QUARTC	500	9.99E-07	9.99E-07	9.99E-07	9.99E-07	1.00E-06	9.98E-07	9.99E-07	9.98E-07	9.98E-07
	1000	9.99E-07	9.98E-07	9.99E-07	9.98E-07	9.99E-07	1.00E-06	9.98E-07	1.00E-06	9.98E-07
	5000	9.94E-07	9.97E-07	9.94E-07	9.97E-07	9.92E-07	7.80E-10	9.98E-07	7.80E-10	7.80E-10
	10000	9.97E-07	9.92E-07	9.97E-07	9.92E-07	1.14E-07	1.56E-09	1.14E-07	1.56E-09	1.56E-09
GWH	500	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	6.43E-07	4.46E-07	7.36E-07	6.63E-07
	1000	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	9.83E-07	4.46E-07	7.36E-07	6.63E-07
	5000	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	4.87E-07	4.46E-07	7.36E-07	6.63E-07
	10000	5.90E-07	8.70E-07	8.30E-07	9.85E-07	8.97E-07	9.45E-07	4.46E-07	8.21E-07	7.16E-07

### 6. Conclusion

This study, based on previous ideas of hybrid CG method construction and recently proposed CG methods of the PRP and HS family, has described four hybrid CG methods which satisfy the descent property as well as globally convergent. Numerical experiments revealed that two of the proposed methods, DPH and DV, are computationally efficient compared to non-hybrid methods. As a major contribution to knowledge, the proposed methods (especially, DPH and DV) gave better performances than the CG\_DESCENT method, acclaimed to be one of the best known CG methods. The DHW and DM methods which are less efficient computation-wise also compete reasonably with the non-hybrid methods. In fact, it was observed that the weak performance of these methods could be addressed if hybridized with classical methods such as the FR and DY methods which are known to possess good convergence properties. Therefore, as part of future work, new hybrid methods comprising of the methods introduced in this paper and those exhibiting good global convergence properties would be developed, tested numerically and compared with highly efficient methods.

## Acknowledgement

The support of the DST-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the CoE. The authors also wish to appreciate Covenant University Centre for Research, Innovation and Discovery (CUCRID) for providing enabling facilities to conduct this research.

## REFERENCES

1. O. J. Adeleke, and I. A. Osinuga, *A five-term hybrid conjugate gradient method with global convergence and descent properties for unconstrained optimization problems*, Asian Journal of Scientific Research, vol. 11, pp. 185–194, 2018.
2. A. O Adewumi, and O. J. Adeleke, *A survey of recent advances in vehicle routing problems*, International Journal of System Assurance Engineering and Management, vol. 9, no. 1, pp. 155–172, 2018.
3. M. Al-Baali, *Descent property and global convergence of the Fletcher-Reeves method with inexact line search*, IMA Journal of Numerical Analysis, vol. 5, no. 1, pp. 121–124, 1985.
4. N. Andrei, *An unconstrained optimization test functions collection*, Adv. Mod. and Optim., vol. 10, no. 1, pp. 147–161, 2008.
5. I. Bongartz, A. R. Conn, N. I. M. Gould, and P. L. Toint, *CUTE: constrained and unconstrained testing environments*, ACM Trans. Math. Software, vol. 21, pp. 123–160, 1995.
6. Y. H. Dai, and Y. Yuan, *Convergence properties of the Fletcher-Reeves method*, IMA Journal of Numerical Analysis, vol. 16, no. 2, pp. 155–164, 1996.
7. Y. H. Dai, and Y. Yuan, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM J. Optim., vol. 10, pp. 177–182, 1999.
8. Y. H. Dai, and Y. Yuan, *An efficient hybrid conjugate gradient method for unconstrained optimization*, Annals of Operations Research, vol. 103, pp. 33–47, 2001.
9. E. D. Dolan, and J. J. Moré, *Benchmarking optimization software with performance profiles*, Math. Program., vol. 91, pp. 201–213, 2002.
10. X. Du, P. Zhang, and W. Ma, *Some modified conjugate gradient methods for unconstrained optimization*, Journal of Computational and Applied Mathematics, vol. 305, pp. 92–114, 2016.
11. A. E. Ezugwu, O. J. Adeleke, and S. Viriri, *Symbiotic organisms search algorithm for the unrelated parallel machines scheduling with sequence-dependent setup times*, PloS one, vol. 13, no. 7, 2018.
12. A. E. Ezugwu, O. J. Adeleke, A. A. Akinyelu, and S. Viriri, *A conceptual comparison of several metaheuristic algorithms on continuous optimisation problems*, Neural Computing and Applications, vol. 2019, pp. 1–45, 2019.
13. R. Fletcher, and C. Reeves, *Function minimization by conjugate gradients*, Comput. J., vol. 7, pp. 149–154, 1964.
14. W. W. Hager, and H. Zhang, *A new conjugate gradient method with guaranteed descent and an efficient line search*, SIAM J. Optim., vol. 16, no. 1, pp. 170–192, 2005.
15. W. W. Hager, and H. Zhang, *A survey of nonlinear conjugate gradient methods*, Pacific Journal of Optimization, vol. 2, pp. 35–58, 2006.
16. M. R. Hestenes, and E. L. Stiefel, *Method of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, vol. 49, pp. 409–436, 1952.
17. H. Huang, Z. Wei, and Y. Shengwei, *The proof of the sufficient descent condition of the WeiYaoLiu conjugate gradient method under the strong WolfePowell line search*, Applied Mathematics and Computation, vol. 189, no. 2, pp. 1241–1245, 2007.
18. H. D. Huang, Y. J. Li, and Z. W. Wei, *Global convergence of a modified PRP conjugate gradient method*, J. Math. Res. Exposition, vol. 30, no. 1, pp. 141–148, 2010.
19. J. Jian, L. Han, and X. Jiang, *A hybrid conjugate gradient method with descent property for unconstrained optimization*, Applied Mathematical Modelling, vol. 39, pp. 1281–1290, 2015.
20. X. Jiang, and J. Jian, *Improved FletcherReeves and DaiYuan conjugate gradient methods with the strong Wolfe line search*, Journal of Computational and Applied Mathematics, vol. 348, pp. 525–534, 2019.
21. Q. Li, *A modified Fletcher-Reeves-type method for nonsmooth convex minimization*, Statistics, Optimization & Information Computing, vol. 2, no. 3, pp. 200–210, 2014.
22. X. Li, W. Zhang, and X. Dong, *A class of modified FR conjugate gradient method and applications to non-negative matrix factorization*, Computers & Mathematics with Applications, vol. 73, no. 2, pp. 270–276, 2017.
23. X. Li, J. Shi, J. Dong, and J. Yu, *A new conjugate gradient method based on Quasi-Newton equation for unconstrained optimization*, Journal of Computational and Applied Mathematics, vol. 350, pp. 372–379, 2019.
24. Y. Liu, and C. Storey, *Efficient generalized conjugate gradient algorithms*, Part 1: Theory, J. Optim. Theory Appl., vol. 69, pp. 129–137, 1991.
25. P. Mtagulwa, and P. Kaelo, *An efficient modified PRP-FR hybrid conjugate gradient method for solving unconstrained optimization problems*, Applied Numerical Mathematics, vol. 145, pp. 111–120, 2019.
26. D. A. Oladepo, O. J. Adeleke, and C. T. Ako, *A mixed hybrid conjugate gradient method for unconstrained engineering optimization problems*, In Computer Science On-line Conference 2018, pp. 423–431. Springer, Cham.
27. E. Polak, and G. Ribière, *Note sur la convergence de directions conjuges*, Rev. Francaise Informat Recherche Operationelle, 3e Anne, vol. 16, pp. 35–43, 1969.
28. B. T. Polyak, *The conjugate gradient method in extreme problems*, USSR Comp. Math. Math. Phys., vol. 9, pp. 94–112, 1969.

29. Y. Shengwei, Z. Wei, and H. Huang, *A note about WYs conjugate gradient method and its applications*, Appl. Math. Comput., vol. 191, no. 2, pp. 381–388, 2007.
30. Z. Wei, S. Yao, and L. Liu, *The convergence properties of some new conjugate gradient methods*, Appl. Math. Comput., vol. 183, pp. 1341–1350, 2006.
31. Z. X. Wei, H. D. Huang and Y. R. Tao, *A modified hestenes stiefel conjugate gradient method and its convergence*, J. Math. Res. Exposition, vol. 30, no. 2, pp. 297–308, 2010.
32. W. Xue, and W. Zhang, *Learning unknown structure in CRFs via adaptive gradient projection method*, Statistics, Optimization & Information Computing, vol. 4, no. 3, pp. 243–251, 2016.
33. D. Xuewu, and X. Chengxian, *Global convergence properties of two classes of optimal algorithms constrained by the FR conjugate gradient method*, Numer. Math. J. of Chinese Univ., vol. 22, no. 4, pp. 311–318, 2000, (in Chinese).
34. L. Zhang, *An improved WeiYaoLiu nonlinear conjugate gradient method for optimization computation* Applied Mathematics and computation, vol. 215, no. 6, pp. 2269–2274, 2009.