Available online at http://scik.org

J. Math. Comput. Sci. 2 (2012), No. 4, 937-966

ISSN: 1927-5307

# A NEW HYBRID WC-FR CONJUGATE GRADIENT ALGORITHM WITH MODIFIED SECANT CONDITION FOR UNCONSTRAINED OPTIMIZATION

ABBAS Y. Al-BAYATI[1,*], RANA Z. Al-KAWAZ[2]

[1]College of Basic Education Telafer, Mosul University, Mosul, Iraq

[2]College of Computers Sciences and Mathematics, Mosul University, Mosul, Iraq

**Abstract**: An accelerated hybrid Conjugate Gradient (CG) algorithm represents the subject of this paper. The parameter $\beta_k^{\mathrm{HYWCFR}}$ is computed as a convex combination of Fletcher and Reeves, $\beta_k^{FR}$ [22] and Wu-Chen, $\beta_k^{WCh}$ [3], i.e. $\beta_k^{HYWCFR} = (1-\theta_k)\beta_k^{FR} + \theta_k \beta_k^{WCh}$. The parameter $\theta_k$ in the convex combination is computed in such a way that the direction corresponding to the CG algorithm is the best direction we know, i.e. the Newton direction, while the pair ( $s_k$ , $y_k$ ) satisfies the classical secant condition $B_{k+1}s_k = y_k$, where, $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$. It is shown that both for uniformly convex functions and for general nonlinear functions the new proposed algorithm with strong Wolfe line search is globally convergent. This algorithm uses an acceleration scheme modifying the step-length $\alpha_k$ for improving the reduction of the function values along the iterations. The technique was given by (Andrei [15]). Numerical comparisons with some similar CG algorithms show that the new proposed hybrid computational scheme outperforms the CG algorithms given by Wu-Chen and FR. A set of 35 unconstrained optimization problems with several different dimensions are used in

*Corresponding author

E-mail addresses: profabbasalbayati@yahoo.com (A.Y. Al-Bayati), rana_alkawaz@yahoo.com (R.Z. Al-Kawaz).

Received February 20, 2012

this paper.

**Keywords:** Unconstrained Optimization, Hybrid Conjugate Gradient, Newton Direction, Secant Condition , Numerical Comparisons.

**2000 AMS Subject Classification:** 47H17; 47H05; 47H09

## 1. Introduction.

Our problem is to minimize a function of n variables:

$$Min\ f(x),\ \text{where}\ f : R^n \to R \tag{1}$$

is a smooth nonlinear function and its gradient $\nabla f(x)$ is available. At the current iterative point $x_k$, the CG method has the following form:

$$x_{k+1} = x_k + \alpha_k d_k \tag{2}$$

$$d_{k+1} = \begin{cases} -g_{k+1}, & for \quad k = 0 \\ -g_{k+1} + \beta_k d_k & for \quad k \geq 1 \end{cases} \tag{3}$$

where $\alpha_k$ is a step-length; $d_k$ is a search direction; $g_k = \nabla f(x_k)$ and $\beta_k$ is a parameter. Consider $\|.\|$ the Euclidean norm and define $y_k = g_{k+1} - g_k$ . The line search in the CG-algorithm is often based on the Wolfe-Powell conditions [21]:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \tag{4a}$$

$$g_{k+1}^T d_k \geq \sigma\ g_k^T d_k \tag{4b}$$

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \tag{5a}$$

$$\left| g_{k+1}^T d_k \right| \leq -\sigma\ g_k^T d_k \tag{5b}$$

where $0 < \delta < 0.5 \leq \sigma < 1$

Equations (4a)-(4b) and (5a)-(5b) are called the “Standard Wolfe” and “Strong Wolfe” conditions, respectively. Different CG-algorithms correspond to different choices for the scalar parameter $\beta_k$ [26]. The methods of Fletcher and Reeves (FR) [22]; Dai and Yuan (DY) [29] and the Conjugate Descent (CD) proposed by Fletcher [23] are defined by:

$$\beta_k^{FR} = \frac{g_{k+1}{}^T g_{k+1}}{g_k{}^T g_k}, \quad \beta_k^{DY} = \frac{g_{k+1}{}^T g_{k+1}}{y_k{}^T s_k}, \quad \beta_k^{CD} = -\frac{g_{k+1}{}^T g_{k+1}}{g_k{}^T s_k},$$

and have strong convergence properties, but they may have modest practical performance due to jamming. On the other hand, the methods of Polak-Ribiere-Polyak (PRP) [2,6]; Hestenes and Stiefel (HS) [14] and Liu and Storey (LS) [27] are defined by:

$$\beta_k^{PR} = \frac{g_{k+1}{}^T y_k}{g_k{}^T g_k}, \quad \beta_k^{HS} = \frac{g_{k+1}{}^T y_k}{d_k{}^T y_k}, \quad \beta_k^{LS} = -\frac{g_{k+1}{}^T y_k}{g_k{}^T s_k},$$

and may not always be convergent, but they often have better computational performances. In this paper we focus on hybrid CG-methods. These methods have been devised to use the attractive features of the above CG-methods. They are defined by (2) and (3) where the parameter $\beta_k$ is computed as projections or as convex combinations of different CG-methods, as in Table (1.1). The hybrid computational schemes perform better than the classical CG-methods [15,20]. In [17] a hybrid CG-method is presented as a convex combination of the Hestenes-Stiefel and the Dai-Yuan methods, where the parameter in convex combination is computed so that the direction corresponding to the CG-method can be the best known Newton direction, while the pair $(s_k, y_k)$ satisfied the secant condition. Numerical experiments with this computational scheme proved to outperform the Hestenes-Stiefel and the Dai-Yuan CG methods, as well as some other hybrid CG-methods [17]. Here, we also present another variant of the hybrid CG-method for unconstrained optimization, which performs much better than the modified PRCG-method by considering the Wu-Chen CG-method and using the classical secant condition. The structure of this paper is as follows. Section 2 introduces a new hybrid CG-algorithm as a convex combination of FR and WC algorithms with classical secant condition, which is an hybrid modified version of our WC-algorithm presented in [3]. Section 3 presents the convergence analysis of the new proposed hybrid CG-algorithm, while in Section 4 some numerical experiments and performance corresponding to this new hybrid CG-algorithm are given. The performance percentages correspond to a set of 35 unconstrained optimization test problems in the CUTE test problem library [8] . Each test problem was tested several times for a gradually increasing number of variables n=500, 1000,...,4000. It is shown that this new hybrid CG-algorithm outperforms the classical FR and WC algorithms.

**Table (1.1): Hybrid CG–Parameters**

| No. | Formula | Author(s) |
|---|---|---|
| **1.** | $\beta_k^{hDY} = \max\left\{c\beta_k^{DY}, \min\left\{\beta_k^{HS}, \beta_k^{DY}\right\}\right\}$<br>$c = (1-\sigma)/(1+\sigma)$ | Hybrid Dai-Yuan [30] (hDY) |
| **2.** | $\beta_k^{hDYz} = \max\left\{0, \min\left\{\beta_k^{HS}, \beta_k^{DY}\right\}\right\}$ | Hybrid Dai-Yuan zero [30] (hDY) |
| **3.** | $\beta_k^{GN} = \max\left\{-\beta_k^{FR}, \min\left\{\beta_k^{PRP}, \beta_k^{FR}\right\}\right\}$ | Gilbert and Nocedal [10](GN) |
| **4.** | $\beta_k^{HuS} = \max\left\{0, \min\left\{\beta_k^{PRP}, \beta_k^{FR}\right\}\right\}$ | Hu and Storey [28] (HuS) |
| **5.** | $\beta_k^{TaS} = \begin{cases} \beta_k^{PRP} & 0 \le \beta_k^{PRP} \le \beta_k^{FR} \\ \beta_k^{FR} & otherwise \end{cases}$ | Touati-Ahmed and Storey [4] (TaS) |
| **6.** | $\beta_k^{LS-CD} = \max\left\{0, \min\left\{\beta_k^{LS}, \beta_k^{CD}\right\}\right\}$ | Hybrid Liu-Storey, Conjugate-Descent (LS-CD) |
| **7.** | $\beta_k^{C} = (1-\theta_k)\beta_k^{HS} + \theta_k\beta_k^{DY} \quad ,0<\theta_k<1$<br>$\theta_k = -\frac{s_k^T g_{k+1}}{g_k^T g_{k+1}}$ | Andrei [17] Convex combination of HS and DY with Newton direction. Secant condition. |
| **8.** | $\beta_k^{AN} = (1-\theta_k)\beta_k^{PR} + \theta_k\beta_k^{DY} \quad ,0<\theta_k<1$<br>$\theta_k = \frac{(y_k^T g_{k+1})(y_k^T s_k) - (y_k^T g_{k+1})(g_k^T g_k)}{(y_k^T g_{k+1})(y_k^T s_k) - (g_{k+1}^T g_{k+1})(g_k^T g_k)}$ | Andrei [18] Convex combination of PRP and DY with conjugacy Condition |

| | | |
|---|---|---|
| **9.** | $\beta_k^{AN} = (1-\theta_k)\beta_k^{PR} + \theta_k \beta_k^{DY} \quad ,0 < \theta_k < 1$ <br> $\theta_k = \dfrac{\left(\dfrac{\delta_1 \eta_k}{s_k^T s_k} - 1\right) s_k^T g_{k+1} - \dfrac{y_k^T g_{k+1}}{y_k^T s_k}\delta_1 \eta_k}{g_k^T g_{k+1} + \dfrac{g_k^T g_{k+1}}{y_k^T s_k}\delta_1 \eta_k}$ | Andrei [16] Accelerated Convex combination of HS and DY with Newton direction. Modified Secant condition |
| **10.** | $\beta_k = \phi_k \beta_k^{PR+} + (1-\phi_k)\alpha_k \beta_k^{AN}, \quad 0 \le \phi_k \le 1$ <br> $\beta_k^{AN} = \dfrac{1}{y_k^T s_k}\left( \|g_{k+1}\|^2 - \dfrac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1} \right)$ <br> $\beta_k^{PR+} = \max\left\{0, \dfrac{g_{k+1}^T y_k}{g_k^T g_k}\right\}$ | Al-Bayati and AL-Baro, [1] |
| **11.** | $\beta_k^{NEW2} = \phi_k \beta_k^{DL+} + (1-\phi_k)\alpha_k \beta_k^{A}, \quad 0 \le \phi_k \le 1$ <br> $\beta_k^{DL+} = \left\{ \max\left\{ \dfrac{g_{k+1}^T y_k}{d_k^T y_k}, 0 \right\} - t \dfrac{g_{k+1}^T s_k}{d_k^T y_k} \right\}$ <br> $\beta_k^{A} = \left\{ \dfrac{1}{y_k^T s_k}\left( \|g_{k+1}\|^2 - \dfrac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1} \right) \right\}$ | Al-Bayati and AL-Baro, [1] |
| **12.** | $\beta_k^{NEW3} = \phi_k \beta_k^{YT+} + (1-\phi_k)\alpha_k \beta_k^{A}, \quad 0 \le \phi_k \le 1$ <br> $\beta_k^{YT+} = \phi_k \left\{ \max\left\{ \dfrac{g_{k+1}^T z_k}{d_k^T z_k}, 0 \right\} - t \dfrac{g_{k+1}^T s_k}{d_k^T z_k} \right\}$ <br> $\beta_k^{A} = \left\{ \dfrac{1}{y_k^T s_k}\left( \|g_{k+1}\|^2 - \dfrac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1} \right) \right\}$ <br> $z_k = y_k + \rho\left( \dfrac{\theta_k}{s_k^T u_k} u_k \right) \rho \ge 0, \quad u_k \in R^n, \quad \ni \quad s_k^T u_k \ne 0$ <br> $\theta_k = 6(f_k - f_{k+1}) + 3(g_k + g_{k+1})^T s_k$ | Al-Bayati and AL-Baro, [1] |
| **13.** | $\beta_k = (1-\theta_k)\beta_k^{FR} + \theta_k \beta_k^{*}$ | Li, and Sun, [24] |

$$\beta_k^* = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_k\|} g_{k+1}^T g_k}{\|g_k\|^2}$$

$$\theta_k = \frac{-\|g_k\|^3 (g_{k+1}^T s_k - s_k^T \nabla^2 f(x_{k+1})^{-1} g_{k+1}) + \|s_k\|^2 \|g_{k+1}\|^2 \|g_k\|}{\|g_{k+1}\| \|s_k\|^2 g_{k+1}^T g_k}$$

## 2. A New Hybrid CG-Algorithm.

Our new proposed algorithm generates the iterates $x_0, x_1, x_2, \ldots$ computed by means of the recurrence (2), where the step-size $\alpha_k > 0$ is determined according to the Wolfe line search conditions (4) and (5), and the directions $d_k$ are generated by the rule:

$$d_{k+1} = \begin{cases} -g_{k+1}, & k = 0 \\ -g_{k+1} + \beta_k^{HYWCFR} d_k, & k \geq 1 \end{cases} \tag{6}$$

Where

$$\beta_k^{HYWCFR} = (1-\theta_k)\beta_k^{FR} + \theta_k \beta_k^{WCh}$$
$$= (1-\theta_k) \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} + \theta_k \left( \beta_k^{PR} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right) \tag{7}$$

and $\theta_k$ is a parameter satisfying $0 \leq \theta_k \leq 1$, which is to be determined. Observe that if $\theta_k = 0$, then $\beta_k^{HYWCFR} = \beta_k^{FR}$, and if $\theta_k = 1$, then $\beta_k^{HYWCFR} = \beta_k^{WCh}$. On the other hand, if $0 < \theta_k < 1$, then $\beta_k^{HYWCFR}$ is a convex combination of $\beta_k^{FR}$ and $\beta_k^{WCh}$. Generally, the global convergence results for the FR method using the strong Wolfe line search will be achieved with $\sigma < 1/2$. The PRP method performs much better than the FR method from the computation point of view. In addition PRP method [6] proved that, when the function $f$ is strongly convex and the line search is exact, then the PRP method is global convergent. As early as 1952, HS method had been proposed by Hestenes and Stiefel. In practical computation, the HS and PR methods, which share the common numerator $g_{k+1}^T y_k = 0$, are generally believed to be the most efficient CG methods, and have got meticulous in recent years. However, Powell [13] constructed an example showed that both methods can cycle infinitely without approaching any stationary point even if an exact line search is used. This counter-example also indicates that both methods have a drawback that they may not be globally convergent when the objective function is non-convex. Therefore, during the past few years, much effort has been investigated to create new formulae. For general function the PR and HS methods perform similarly in terms of theoretical property. Both methods are preferred to the FR method in its numerical performance, because the methods essentially perform a restart after it encounters a bad direction. On the other hand, the FR method always generates descent directions, this CG algorithm, relating the descent directions to the sufficient descent condition. It is shown that if there exist constants $\gamma_1$ and $\gamma_2$ such that $\gamma_1 \le \|g_k\| \le \gamma_2$ for all k, then for any $\sigma \in (0,1)$, there exists a constant $c > 0$ such that the sufficient descent condition:

$$g_{k+1}^T d_{k+1} \le -c \, \|g_{k+1}\|^2 \qquad (8)$$

Therefore, we combine these two methods in a convex combination in order to have a good performance CG-algorithm for unconstrained optimization.

From (6) and (7). It is obvious that:

$$d_{k+1} = -g_{k+1} + (1-\theta_k)\frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} s_k + \theta_k \left( \beta_k^{PR} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right) s_k \qquad (9)$$

As we know, when the initial point $x_0$ is close enough to a local minimum point $x^*$, then the best direction to be followed in the current point $x_{k+1}$ is the Newton direction $-\nabla^2 f(x_{k+1})^{-1} g_{k+1}$. Therefore, our motivation is to choose the parameter $\theta_k$ in (9) so that the direction $d_{k+1}$ can be the best direction we know, i.e. the Newton direction. Hence, using the Newton direction from the equality:

$$-\nabla^2 f(x_{k+1})^{-1} g_{k+1} = -g_{k+1} + (1-\theta_k)\frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} s_k + \theta_k \left( \beta_k^{PR} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right) s_k \quad (10)$$

Multiply both sides of equation (10) by $s_k^T \nabla^2 f(x_{k+1})$ have get:

$$-s_k^T g_{k+1} = -s_k^T \nabla^2 f(x_{k+1}) g_{k+1} + s_k^T \nabla^2 f(x_{k+1}) s_k \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} - \theta_k s_k^T \nabla^2 f(x_{k+1}) s_k \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2}$$
$$+ \theta_k s_k^T \nabla^2 f(x_{k+1}) s_k \left( \frac{g_{k+1}{}^T y_k}{\|g_k\|^2} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right)$$

$$-s_k^T g_{k+1} + s_k^T \nabla^2 f(x_{k+1}) g_{k+1} - s_k^T \nabla^2 f(x_{k+1}) s_k \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} = -\theta_k s_k^T \nabla^2 f(x_{k+1}) s_k \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2}$$
$$+ \theta_k s_k^T \nabla^2 f(x_{k+1}) s_k \left( \frac{g_{k+1}{}^T y_k}{\|g_k\|^2} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right)$$

We get:

$$\theta_k = \frac{s_k^T \nabla^2 f(x_{k+1}) g_{k+1} - s_k^T g_{k+1} - s_k^T \nabla^2 f(x_{k+1}) s_k \left( \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} \right)}{-s_k^T \nabla^2 f(x_{k+1}) s_k \left( \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} \right) + s_k^T \nabla^2 f(x_{k+1}) s_k \left( \frac{2(f_k - f_{k+1}) + g_k^T s_k + g_{k+1}{}^T y_k}{\|g_k\|^2} \right)}$$

$$\theta_k = \frac{-s_k^T g_{k+1}\|g_k\|^2 + s_k^T \nabla^2 f(x_{k+1}) g_{k+1}\|g_k\|^2 - s_k^T \nabla^2 f(x_{k+1}) s_k g_{k+1}{}^T g_{k+1}}{\left(2(f_k - f_{k+1}) + g_k^T s_k + g_{k+1}{}^T y_k - g_{k+1}{}^T g_{k+1}\right) s_k^T \nabla^2 f(x_{k+1}) s_k} \quad (11)$$

Observe that the Newton direction is being used here only as a motivation for formula (11). However, in formula (11) for $\theta_k$ the main drawback is the presence of the Hessian. One of the first CG-algorithm using the Hessian matrix was given by Daniel [11], where $\beta_k = (g_{k+1}^T \nabla^2 f(x_k) d_k)/(d_k^T \nabla^2 f(x_k) d_k)$. For large-scale problems, choices for the update parameter that do not require the evaluation of the Hessian matrix are often preferred in practice to the methods that require the Hessian. As we know, for Quasi-Newton (QN) methods an approximation matrix $B_k$ to the Hessian $\nabla^2 f_k$ is used and updated so that the new matrix $B_{k+1}$ satisfies the secant condition $B_{k+1}s_k = y_k$ . Therefore, in order to have an algorithm for solving large-scale problems in [17] it is assumed that the pair $(s_k,\ y_k)$ satisfies the secant condition. This leads us to a hybrid CG-algorithm, by using $s_k = -g_k$, where:

$$\theta_k = \frac{-s_k^T y_k g_{k+1}^T g_k - s_k^T g_{k+1}\|g_k\|^2 - y_k^T s_k \|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) - \|g_{k+1}\|^2 + g_k^T s_k + g_{k+1}{}^T y_k\right) y_k^T s_k}$$

$$\theta_k = \frac{-s_k^T (g_{k+1} - g_k) g_{k+1}^T g_k - s_k^T g_{k+1}\|g_k\|^2 - (g_{k+1} - g_k)^T s_k \|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) - \|g_{k+1}\|^2 + g_k^T s_k + g_{k+1}^T (g_{k+1} - g_k)\right)(g_{k+1} - g_k)^T s_k}$$

$$= \frac{-s_k^T g_k \|g_{k+1}\|^2 + s_k^T g_{k+1}\|g_k\|^2 - s_k^T g_{k+1}\|g_k\|^2 - g_{k+1}^T s_k \|g_{k+1}\|^2 + g_k^T s_k \|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) - \|g_{k+1}\|^2 + g_k^T s_k + \|g_{k+1}\|^2 - g_{k+1}^T g_k\right)(g_{k+1} - g_k)^T s_k}$$

$$\Rightarrow \theta_k = \frac{-s_k^T g_{k+1}\|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k\right) y_k^T s_k} \qquad (12)$$

**Theorem 2.1.** In the CG-algorithm (2), (9), (12), assume that $\alpha_k$ is determined by the Wolfe line search (4). If $0 < \theta_k < 1$, then the direction $d_{k+1}$ given by (9) is a **descent direction.**

**Proof**.

Since, $0 < \theta_k < 1$, from (9) we get:

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + (1-\theta_k)\frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k + \theta_k \left( \beta_k^{PR} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right) g_{k+1}^T s_k$$

$$\leq -\|g_{k+1}\|^2 + (1 - \frac{-s_k^T g_{k+1}\|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k\right) y_k^T s_k}) \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k$$

$$+ \frac{-s_k^T g_{k+1}\|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k\right) y_k^T s_k} (\beta_k^{PR} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2}) g_{k+1}^T s_k$$

$$= -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k + \frac{s_k^T g_{k+1}\|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k\right) y_k^T s_k} \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k$$

$$- \frac{s_k^T g_{k+1}\|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k\right) y_k^T s_k} (\frac{2(f_k - f_{k+1}) + g_k^T s_k + g_{k+1}^T y_k}{\|g_k\|^2}) g_{k+1}^T s_k$$

$$= (-1 + \frac{g_{k+1}^T s_k}{\|g_k\|^2})\|g_{k+1}\|^2 + \left( \frac{\|g_{k+1}\|^2}{\left(2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k\right) y_k^T s_k} - \frac{\left(2(f_k - f_{k+1}) + g_k^T s_k + g_{k+1}^T y_k\right)}{\left(2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k\right) y_k^T s_k} \right) \frac{(s_k^T g_{k+1})^2 \|g_{k+1}\|^2}{\|g_k\|^2} \quad (13)$$

But, $y_k^T s_k > 0$ by (4b) and, since $g_k^T s_k \leq 0$, it follows that $(-1 + \frac{g_{k+1}^T s_k}{\|g_k\|^2}) \leq 0$. When the iterations are in progress or when they jam, $y_k$ becomes tiny while $\|g_k\|$ is bounded away from zero and $2(f_k - f_{k+1}) + s_k^T g_k \leq 0$. Therefore, from (13), it follows that $g_{k+1}^T d_{k+1} \leq 0$, i.e. the direction $d_{k+1}$ is a descent one.

**Theorem 2.2.** If $0 < \theta_k < 1$, then the direction $d_{k+1}$ given by (9) satisfies the **sufficient descent condition,** for exact and inexact line searches, $d_{k+1}^T g_{k+1} \leq -c\|g_{k+1}\|^2; \quad c > 0$.

**Proof**.

From (9), we have

$$d_{k+1} = -g_{k+1} + (1 - \theta_k) \frac{g_{k+1}{}^T g_{k+1}}{\|g_k\|^2} s_k + \theta_k \left( \beta_k^{PR} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right) s_k$$

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + (1 - \theta_k) \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k + \theta_k \left( \beta_k^{PR} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{\|g_k\|^2} \right) g_{k+1}^T s_k$$

$$g_{k+1}^T d_{k+1} \leq -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k - \theta_k \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k + \theta_k \left( \frac{2(f_k - f_{k+1}) + g_k^T s_k + g_{k+1}^T y_k}{\|g_k\|^2} \right) g_{k+1}^T s_k$$

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k + \left( \frac{2(f_k - f_{k+1}) + g_k^T s_k + g_{k+1}^T y_k - \|g_{k+1}\|^2}{\|g_k\|^2} \right) \theta_k g_{k+1}^T s_k$$

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} g_{k+1}^T s_k - \left( \frac{(s_k^T g_{k+1})^2 \|g_{k+1}\|^2}{\|g_k\|^2 y_k^T s_k} \right) \tag{14}$$

Using exact line searches, in (14) yields: $g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2$, $c > 0$. Using, inexact line searches, in (14) yields:

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

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

$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2$, $c > 0$. This completes the proof.

Observe that $y_k$ becomes tiny while $\|g_k\|$ is bounded away from zero. Consequently, the last term in the above inequality becomes negligible, since $y_k^T s_k > 0$ by (4b) and since $s_k^T g_{k+1} = y_k^T s_k + s_k^T g_k < y_k^T s_k$. Observe that the parameter $\theta_k$ given by (12) can be outside the interval [0,1]. However, in order to have a real convex combination in (7) the following rule is used: if $\theta_k \le 0$, then set ($\theta_k = 0$) in (7), i.e. $\beta_k^{\text{HYWCFR}} = \beta_k^{FR}$; if $\theta_k \ge 1$, then take ($\theta_k = 1$) in (7), i.e. $\beta_k^{\text{HYWCFR}} = \beta_k^{FR}$. Therefore, under this rule for $\theta_k$ selection, the direction $d_{k+1}$ in (9) combines the properties of the FR and the WC algorithms in a convex way.

### 2.3 An Acceleration Scheme of the Line Search Parameter.

In [9] Nocedal pointed out that in CG-methods the step lengths may differ from 1 in a very unpredictable manner. They can be larger or smaller than 1 depending on how the problem is scaled. This is in very sharp contrast to the Newton and QN-methods, including the limited memory QN-methods, which accept the unit step-length most of the time along the iterations, and therefore usually they require only few function evaluations per search direction. Numerical comparisons between CG-methods and the limited memory QN-method by Liu and Nocedal [5], show that the latter is more successful [20]. One explanation of efficiency of this limited memory QN-method is given by its ability to accept unity step-lengths along the iterations. In this section we take advantage of this behavior of CG-algorithms and consider an acceleration scheme which was presented in [15]. In accelerated algorithm instead of (2) the new estimation of the minimum point is computed as:

$$x_{k+1} = x_k + \lambda_k \alpha_k d_k \tag{15}$$

where

$$\lambda_k = -\frac{a_k}{b_k} \tag{16}$$

$a_k = \alpha_k g_k^T d_k$ , $b_k = -\alpha_k (g_k - g_z)^T d_k$ , $g_z = \nabla f(z)$ *and* $z = x_k + \alpha_k d_k$. Hence, if $b_k = 0$, then the new estimation of the solution is computed as $x_{k+1} = x_k + \lambda_k \alpha_k d_k$, otherwise $x_{k+1} = x_k + \alpha_k d_k$. Therefore, using the definitions of $g_k$, $s_k$, $y_k$ and the above acceleration scheme (15) and (16) we can present the following hybrid CG-algorithm.

**2.4 Outline of the New Hybrid CG-Algorithm.**

**Step1:** **Initialization:** $x_1 \in R^n$; ($\varepsilon > 0$); set k=1, compute $f(x_1)$ and $g_1$.

Consider $d_1 = -g_1$.

**Step2:** **Test for Continuation of Iterations:** If $\|g_{k+1}\| < \varepsilon$, then stop.

**Step3:** **Line Search:** Compute $\alpha_k$ > 0 satisfying the Wolfe line search condition and compute $z = x_k + \alpha_k d_k$, $y_k = g_k - g_z$, $g_z = \nabla f(z)$. **Acceleration scheme:** compute, $a_k = \alpha_k g_k^T d_k$, $b_k = -\alpha_k y_k^T d_k$, If $b_k \neq 0$, then compute, $\lambda_k = -\frac{a_k}{b_k}$ and update the variables as $x_{k+1} = x_k + \lambda_k \alpha_k d_k$,

otherwise update the variables as $x_{k+1} = x_k + \alpha_k d_k$.

**Step4:** **Computation of $\theta_k$:** If $[2(f_k - f_{k+1}) + g_k^T s_k - g_{k+1}^T g_k] y_k^T s_k = 0$, then set $\theta_k = 0$; otherwise, compute $\theta_k$ as in (12).

**Step5:** **Computation of $\beta_k$:** If 0 < $\theta_k$ < 1, then compute $\beta_k^{\text{HYWCFR}}$ as in (7).

If $\theta_k \geq 1$, then set $\beta_k^{\text{HYWCFR}} = \beta_k^{WCh}$ else if $\theta_k \leq 0$, then set $\beta_k^{\text{HYWCFR}} = \beta_k^{FR}$

**Step6:** **Computation of $d_k$:** Compute $d_{k+1} = -g_{k+1} + \beta_k^{\text{HYWCFR}} d_k$. If the restart criterion of Powell:

$$\left| g_{k+1}^T g_k \right| \geq 0.2 \left\| g_{k+1} \right\|^2 \tag{17}$$

is satisfied, then set $d_{k+1} = -g_{k+1}$; otherwise, define $d_{k+1} = d$.

**Step7:** **Computation of iteration:** set $k = k + 1$ and continue with **Step 2**.

It is well known that, if f is bounded along the direction $d_k$, then there exists a step-size $\alpha_k$ satisfying the Wolfe line search conditions (4) and (5). In our algorithm, when the Powell restarting condition (15) is satisfied, then we restart the algorithm with the negative gradient. More sophisticated reasons for restarting the algorithms have been proposed in the literature [13,21], but we are interested in the performance of a CG-algorithm that uses this restart criterion associated to a direction satisfying the conjugacy condition when $0 < \theta_k < 1$. Under reasonable assumptions, conditions (4), (5) and (17) are sufficient to prove the global convergence of the algorithm.

## 3. Convergence Analysis.

Throughout this section, we assume that:

**(i)** The level set $L_{x_0} = \{x \in R^n \,|\, f(x) \leq f(x_0)\}$ is bounded.

**(ii)** In a neighborhood $U$ of $L_{x_0}$, the function $f$ is continuously differentiable and its gradient $\nabla f$ is Lipschitz continuous, i.e. there exists a constant $L>0$ such that $\|\nabla f(x)-\nabla f(y)\| \le L\|x-y\|$ for all x, y $\in U$. Under these assumptions on f , there exists a constant $\Gamma \ge 0$ such that $\|\nabla f(x)\| \le \Gamma$ , for all $x \in L_{x_0}$. The convergence of the Steepest Descent (SD) method with Armijo-type search is proved under very general conditions in [12]. On the other hand, in [31] it is proved that, for any CG method with strong Wolfe line search, the following general result holds.

**Lemma 3.1.** Let assumptions **(i)** and **(ii)** hold and consider any CG method defined by (2) and (3), where $d_k$ is a **descent direction** and $\alpha_k$ is obtained by the strong Wolfe line search. If

$$\sum_{k\ge 1}\frac{1}{\|d_k\|^2}=\infty \tag{18}$$

then

$$\liminf_{k\to\infty} \|g_k\| = 0 \tag{19}$$

For uniformly convex functions which satisfy the above assumptions, we can prove that the norm of $d_{k+1}$ given by (9) is bounded above. Assume that the function $f$ is a uniformly convex function, i.e. there exists a constant $\mu \ge 0$ such that, for all $x, y \in L_{x_0}$,

$$(\nabla f(x)-\nabla f(y))^T(x-y) \ge \mu\,\|x-y\|^2 \tag{20a}$$

and the step-length $\alpha_k$ is obtained by the strong Wolfe line search defined in (5). Using **Lemma-3.1** the following result can be proved.

**Theorem 3.2.** Suppose that the assumptions **(i)** and **(ii)** hold. Consider the algorithm (2), (9) and (12), where $0\le\theta_k\le 1$ and $\alpha_k$ is obtained by the strong Wolfe line search.

If $\|s_k\|$ tends to zero and there exists nonnegative constants $\eta_1$ and $\eta_2$ such that:

$$\|g_k\|^2 \geq \eta_1 \|s_k\|^2 \ , \ \|g_{k+1}\|^2 \leq \eta_2 \|s_k\| \tag{20b}$$

and $f$ is a uniformly convex function, then

$$\lim_{k\to\infty} \|g_k\|^2 = 0 \tag{21}$$

**Proof.**

From (20a) it follows that $y_k^T s_k \geq \mu \|s_k\|^2$. Now, since $0 \leq \theta_k \leq 1$, from uniform convexity and (20a) we have:

$$\left|\beta_k^{HYWCFR}\right| \leq \left|\frac{g_{k+1}{}^T g_{k+1}}{g_k{}^T g_k}\right| + \left|\frac{g_{k+1}{}^T y_k}{g_k{}^T g_k}\right| + \left|\frac{2(f_k - f_{k+1}) + g_k{}^T s_k}{g_k{}^T g_k}\right|$$

$$\leq \left|\frac{g_{k+1}{}^T g_{k+1}}{g_k{}^T g_k}\right| + \left|\frac{g_{k+1}{}^T y_k}{g_k{}^T g_k}\right| + \left|\frac{-2\delta\ (g_k{}^T s_k) + g_k{}^T s_k}{g_k{}^T g_k}\right|$$

$$\leq \left|\frac{g_{k+1}{}^T g_{k+1}}{g_k{}^T g_k}\right| + \left|\frac{g_{k+1}{}^T y_k}{g_k{}^T g_k}\right| + \left|\frac{(1-2\delta) g_k{}^T s_k}{g_k{}^T g_k}\right|$$

$$\leq \frac{\eta_2 \|s_k\|}{\eta_1 \|s_k\|^2} + \frac{\|g_{k+1}\| \|y_k\|}{\eta_1 \|s_k\|^2} + \alpha_k (2\delta - 1)$$

But $\|y_k\| \leq L \|s_k\|$,

$$\left|\beta_k^{HYWCFR}\right| \leq \frac{\eta_2}{\eta_1 \|s_k\|} + \frac{\Gamma L}{\eta_1 \|s_k\|} + \alpha_k (2\delta - 1)$$

Hence, $$\|d_{k+1}\| \leq \|g_{k+1}\| + \left|\beta_k^{HYWCFR}\right| \|s_k\| \leq \Gamma + \frac{\eta_2}{\eta_1} + \frac{\Gamma L}{\eta_1} + \alpha_k (2\delta - 1)$$

which implies that (18) is true. Therefore, by **Lemma-3.1** we have (19), which for uniformly convex functions is equivalent to (21). Powell [13] showed that, for general functions, the PRP method is globally convergent if the step-length $\|s_k\| = \|x_{k+1} - x_k\|$ tends to zero, i.e. $\|s_k\| \leq \|s_{k-1}\|$ is a condition of convergence. For the convergence of our algorithm from (5b), we see that, for k≥1, the gradient must be bounded as: $\eta_1 \|s_k\|^2 \leq \|g_k\|^2 \leq \eta_2 \|s_{k-1}\|$. If the Powell condition is satisfied, i.e. $\|s_k\|$ tends to zero, then $\|s_k\|^2 << \|s_{k-1}\|$ and therefore the norm of the gradient can satisfy (20). In the numerical experiments, we observed that (20) is constantly satisfied in the last part of the iterations. For general nonlinear functions, the convergence analysis of our algorithm exploits insights developed by Gilbert and Nocedal [10], by Dai and Liao [33] and by Hager and Zhang [25]. The global convergence proof of the new algorithm is based on the Zoutendijk condition combined with the analysis showing that the sufficient descent condition holds and $\|d_k\|$ is bounded. Suppose that the level set L is bounded and the function $f$ is bounded from below.

**Lemma 3.3.** Assume that $d_k$ is a descent direction and $\nabla f$ satisfied the Lipschitz condition $\|\nabla f(x) - \nabla f(x_k)\| \le L\|x - x_k\|$ for all $x$ on the line segment connecting $x_k$ and $x_{k+1}$, where L is a constant. If the line search satisfies the second Wolfe condition (5), then

$$\alpha_k \ge \frac{1-\sigma}{L}\frac{|g_k^T d_k|}{\|d_k\|^2} \qquad (22)$$

**Proof.**

Subtracting $g_k^T d_k$ from both sides of (4b) and using the Lipschitz condition we have

$$(\sigma - 1)g_k^T d_k \le (g_{k+1} - g_k)^T d_k \le L\alpha_k \|d_k\|^2 \qquad (23)$$

Since $d_k$ is a descent direction and σ< 1, then (22) follows immediately from (23).

**Theorem 3.4.** Let assumptions **(i)** and **(ii)** hold. Assume that $0 < \theta_k < 1$ and that, for every, k≥0, there exists a positive constant $\omega$ such that $1 \ge \omega > 0$ as well as the constants $\gamma$ and $\Gamma$ such that $\gamma \le \|g_k\| \le \Gamma$. Then, for the computational scheme (2), (9), (12), where $\alpha_k$ is determined by the Wolfe line search (4), either $g_k = 0$ for some k or $\lim_{k\to\infty}\inf \|g_k\| = 0$.

**Proof.**

From the Wolfe condition (4a) we have:

$$2(f_k - f_{k+1}) + g_k{}^T s_k \le (1-2\delta) g_k{}^T s_k \tag{24}$$

By **Theorem-2.2** and the assumption 1≥ ω, it follows that

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 \le -\omega \|g_{k+1}\|^2$$

Therefore, $-g_{k+1}^T d_{k+1} \ge \omega \|g_{k+1}\|^2$

On the other hand, $\|y_k\| = \|g_{k+1} - g_k\| \le L\|s_k\|$ . Hence, $|g_{k+1}^T y_k| \le \|g_{k+1}\|\|y_k\| \le \Gamma L \|s_k\|$

With these, from (7) we get:

$$|\beta_k^{HYWCFR}| \le \left|\frac{g_{k+1}{}^T g_{k+1}}{g_k{}^T g_k}\right| + \left|\frac{g_{k+1}{}^T y_k}{g_k{}^T g_k}\right| + \left|\frac{2(f_k - f_{k+1}) + g_k{}^T s_k}{g_k{}^T g_k}\right|$$

But ,

$$\left|\frac{g_{k+1}{}^T y_k}{g_k{}^T g_k}\right| \le \frac{\|g_{k+1}\|\|y_k\|}{\gamma^2} \le \frac{\Gamma L \|s_k\|}{\gamma^2} \le \frac{\Gamma L D}{\gamma^2}$$

where $D = \max\{\|y - z\| : y, z \in L_{x_0}\}$ is the diameter of the level set $L_{x_0}$ . On the other hand, $\left|\frac{g_{k+1}{}^T g_{k+1}}{g_k{}^T g_k}\right| \le \frac{\Gamma^2}{\gamma^2}$ and $\left|\frac{2(f_k - f_{k+1}) + g_k{}^T s_k}{g_k{}^T g_k}\right| \le \alpha_k (2\delta - 1)$ .

$$\text{Therefore, } |\beta_k^{HYWCFR}| \le \frac{\Gamma^2}{\gamma^2} + \frac{\Gamma L D}{\gamma^2} + \alpha_k (2\delta - 1) \equiv E \tag{25}$$

Now, we can write:

$$\|d_{k+1}\| \le \|g_{k+1}\| + |\beta_k^{HYWCFR}| \|s_k\| \le \Gamma + ED \tag{26}$$

Since the level set L is bounded and the function f is bounded from below, using **Lemma-3.3**, from (4) it follows that:

$$0 < \sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{27}$$

i.e. the Zoutendijk condition holds. Therefore, from **Theorem-2.2** using (27) the descent property yields:

$$\sum_{k=0}^{\infty}\frac{\gamma^4}{\|d_k\|^2}\leq\sum_{k=0}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2}\leq\sum_{k=0}^{\infty}\frac{1}{\omega^2}\frac{(g_k^T d_k)^2}{\|d_k\|^2}\leq\sum_{k=0}^{\infty}\frac{(g_k^T d_k)^2}{\|d_k\|^2}\leq\infty$$

which contradicts (26). Hence, $\gamma=\liminf_{k\to\infty}\|g_k\|=0$.

Therefore, when $0<\theta_k<1$, our CG-algorithm is globally convergent, meaning that either $g_k=0$ for some k or (19) holds. Observe that, in the conditions of **Theorem-2.2**, the direction $d_{k+1}$ satisfied the sufficient descent condition independently of the line search.

## 4. Numerical Experiments.

In this section, we present the computational performance of a Fortran implementation of the new algorithm on a set of 35 nonlinear unconstrained optimization test problems. All the tests are performed on a PC. These test problems are contributed in CUTE and their details are given in the Appendix. Each of them is tested several times for a gradually increasing number of variables n=500, 1000,...,4000. At the same time, we present comparisons with other CG algorithms, including the performance percentage of Dolan and Moré [7]. All algorithms implement the same stopping criterion $\|g_k\|_\infty\leq 10^{-6}$, where $\|.\|_\infty$ is the maximum absolute component of a vector. The comparisons of algorithms are given in the following context. All codes are written in double precision Fortran and compiled with visual Fortran (default compiler settings) on an Intel Pentium 4, 1.86 GHz workstation.

**Table (4.1)**

**COMPARISON BETWEEN :NEW ,(WU & CHEN) and (FR) METHODS**

**n= 500, 1000, … ,4000**

| Prob | FR/1964 | | | Wu & Chen/2010 | | | New Hybrid CG | | |
|---|---|---|---|---|---|---|---|---|---|
| . | **NOI/NOFG/TIME** | | | **NOI/NOFG/TIME** | | | **NOI/NOFG/TIME** | | |
| **1** | 362 | 692 | 0.8 | 1210 | 1611 | 2.34 | 300 | 579 | 1.04 |
| **2** | 265 | 639 | 0.15 | 207 | 428 | 0.12 | 203 | 424 | 0.14 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **3** | 122 | 339 | 0.25 | 63 | 134 | 0.16 | 63 | 134 | 0.18 |
| **4** | 1065 | 1976 | 2.24 | 2116 | 2239 | 5.26 | 477 | 586 | 1.21 |
| **5** | 289 | 566 | 0.18 | 834 | 921 | 0.49 | 182 | 267 | 0.12 |
| **6** | 245 | 475 | 0.79 | 283 | 305 | 0.99 | 405 | 424 | 1.52 |
| **7** | 296 | 777 | 0.18 | 6674 | 6724 | 3.19 | 71 | 96 | 0.02 |
| **8** | 93 | 294 | 0.53 | 141 | 166 | 0.72 | 141 | 166 | 0.73 |
| **9** | 215 | 498 | 0.09 | 278 | 399 | 0.15 | 150 | 248 | 0.09 |
| **10** | 182 | 425 | 0.37 | 152 | 223 | 0.35 | 144 | 206 | 0.31 |
| **11** | 352 | 705 | 0.39 | 392 | 511 | 0.5 | 254 | 390 | 0.34 |
| **12** | 165 | 430 | 0.1 | 149 | 267 | 0.09 | 152 | 245 | 0.07 |
| **13** | 94 | 323 | 0.13 | 33 | 57 | 0.04 | 33 | 57 | 0.03 |
| **14** | 614 | 1157 | 0.35 | 2962 | 3037 | 1.8 | 260 | 332 | 0.16 |
| **15** | 2033 | 3020 | 1.35 | 13542 | 13614 | 8.36 | 128 | 190 | 0.08 |
| **16** | 123 | 341 | 0.19 | 88 | 113 | 0.11 | 69 | 95 | 0.09 |
| **17** | 121 | 337 | 0.19 | 89 | 192 | 0.17 | 89 | 192 | 0.19 |
| **18** | 136 | 356 | 0.21 | 97 | 164 | 0.16 | 91 | 187 | 0.17 |
| **19** | 315 | 650 | 0.2 | 833 | 918 | 0.52 | 212 | 308 | 0.13 |
| **20** | 123 | 342 | 0.22 | 63 | 134 | 0.17 | 66 | 154 | 0.18 |
| **21** | 305 | 653 | 0.18 | 809 | 896 | 0.46 | 277 | 369 | 0.16 |
| **22** | 318 | 568 | 0.44 | 569 | 613 | 0.82 | 180 | 238 | 0.3 |
| **23** | 646 | 1262 | 0.98 | 817 | 890 | 1.44 | 373 | 407 | 0.71 |
| **24** | 133 | 369 | 0.07 | 109 | 219 | 0.07 | 73 | 178 | 0.05 |
| **25** | 215 | 505 | 0.16 | 192 | 280 | 0.14 | 204 | 313 | 0.14 |
| **26** | 599 | 1193 | 0.8 | 693 | 725 | 0.93 | 296 | 320 | 0.46 |
| **27** | 117 | 334 | 0.08 | 121 | 188 | 0.07 | 109 | 176 | 0.06 |
| **28** | 107 | 330 | 0.31 | 64 | 129 | 0.19 | 64 | 129 | 0.19 |
| **29** | 101 | 302 | 0.37 | 58 | 74 | 0.17 | 58 | 74 | 0.16 |
| **30** | 121 | 420 | 0.36 | 32 | 64 | 0.12 | 32 | 64 | 0.13 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **31** | 182 | 425 | 0.35 | 154 | 219 | 0.32 | 149 | 215 | 0.29 |
| **32** | 566 | 1096 | 0.3 | 1019 | 1159 | 0.59 | 299 | 430 | 0.17 |
| **33** | 42 | 369 | 0.09 | 36 | 95 | 0.02 | 38 | 117 | 0.05 |
| **34** | 78 | 94 | 0.08 | 66 | 90 | 0.1 | 66 | 90 | 0.09 |
| **35** | 176 | 416 | 0.1 | 174 | 251 | 0.09 | 106 | 196 | 0.06 |
| **Total** | **10916** | **22678** | **13.58** | **35119** | **38049** | **31.22** | **5814** | **8596** | **9.82** |

**TABLE (4.2)**

**PERCENTAGE PERFORMANCE OF TABLE (4.1)**

| **TOOLS** | **WU & Ch (2010)** | **New Hybrid** |
|---|---|---|
| NOI | 100 % | 16.6 % |
| NOFG | 100 % | 22.6 % |
| TIME | 100 % | 31.5 % |

| **TOOLS** | **FR (1964)** | **New Hybrid** |
|---|---|---|
| NOI | 100 % | 53.3 % |
| NOFG | 100 % | 37.9 % |
| TIME | 100 % | 72.3 % |

Clearly, from the above table, we have found that the new proposed algorithm beats FR algorithm in about (46%) NOI; (62%) NOFG and (27% )Time. From this table we have also concluded that the new algorithm beats the standard WC-algorithm in about (83%) NOI; (77%) NOFG and (68%)Time.

**Table (4.3)**

**COMPARISON BETWEEN: NEW ,(WU & CHEN) and (FR) METHODS**

**n= 100, 300, … ,900**

| Prob. | FR/1964 NOI/NOFG/TIME | | | Wu & Chen/2010 NOI/NOFG/TIME | | | New Hybrid CG NOI/NOFG/TIME | | |
|---|---|---|---|---|---|---|---|---|---|
| **1** | 288 | 545 | 0.16 | 936 | 1099 | 0.7 | 254 | 296 | 0.19 |
| **2** | 149 | 364 | 0.02 | 113 | 208 | 0.02 | 110 | 205 | 0.01 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **3** | 64 | 192 | 0.03 | 37 | 47 | 0 | 37 | 47 | 0 |
| **4** | 393 | 745 | 0.22 | 775 | 848 | 0.46 | 225 | 299 | 0.13 |
| **5** | 172 | 333 | 0.03 | 513 | 559 | 0.05 | 126 | 165 | 0.03 |
| **6** | 148 | 284 | 0.13 | 151 | 163 | 0.09 | 173 | 185 | 0.15 |
| **7** | 161 | 465 | 0.03 | 5474 | 5488 | 0.54 | 38 | 49 | 0 |
| **8** | 50 | 173 | 0.06 | 84 | 99 | 0.09 | 81 | 98 | 0.09 |
| **9** | 125 | 298 | 0.01 | 168 | 228 | 0.03 | 81 | 141 | 0 |
| **10** | 107 | 255 | 0.05 | 93 | 134 | 0.05 | 93 | 120 | 0.04 |
| **11** | 205 | 404 | 0.07 | 219 | 289 | 0.05 | 163 | 217 | 0.04 |
| **12** | 103 | 263 | 0.01 | 89 | 140 | 0 | 87 | 138 | 0.02 |
| **13** | 38 | 179 | 0.02 | 18 | 35 | 0.02 | 18 | 35 | 0 |
| **14** | 378 | 709 | 0.05 | 660 | 698 | 0.09 | 222 | 255 | 0.03 |
| **15** | 373 | 779 | 0.06 | 3311 | 3349 | 0.52 | 117 | 154 | 0.02 |
| **16** | 79 | 211 | 0.1 | 65 | 75 | 0.03 | 42 | 52 | 0.01 |
| **17** | 77 | 201 | 0.03 | 45 | 60 | 0.01 | 45 | 59 | 0.02 |
| **18** | 71 | 200 | 0.03 | 54 | 65 | 0.02 | 54 | 68 | 0.03 |
| **19** | 199 | 411 | 0.05 | 639 | 683 | 0.09 | 141 | 201 | 0.03 |
| **20** | 64 | 192 | 0.02 | 37 | 47 | 0.01 | 37 | 47 | 0 |
| **21** | 180 | 393 | 0.01 | 455 | 494 | 0.04 | 137 | 180 | 0.01 |
| **22** | 185 | 325 | 0.05 | 374 | 404 | 0.12 | 107 | 137 | 0.05 |
| **23** | 397 | 780 | 0.14 | 429 | 492 | 0.16 | 168 | 204 | 0.08 |
| **24** | 82 | 224 | 0 | 68 | 87 | 0 | 61 | 80 | 0 |
| **25** | 130 | 307 | 0.02 | 119 | 174 | 0.04 | 122 | 141 | 0.01 |
| **26** | 395 | 761 | 0.11 | 510 | 549 | 0.14 | 192 | 215 | 0.06 |
| **27** | 76 | 207 | 0.02 | 74 | 99 | 0.01 | 55 | 102 | 0.01 |
| **28** | 62 | 196 | 0.04 | 40 | 80 | 0.03 | 40 | 81 | 0.03 |
| **29** | 70 | 201 | 0.05 | 43 | 53 | 0.04 | 43 | 53 | 0.03 |
| **30** | 66 | 235 | 0.05 | 22 | 38 | 0.02 | 22 | 38 | 0.02 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **31** | 107 | 255 | 0.05 | 95 | 122 | 0.04 | 93 | 120 | 0.03 |
| **32** | 339 | 666 | 0.04 | 572 | 632 | 0.06 | 196 | 248 | 0.02 |
| **33** | 5 | 15 | 0 | 12 | 35 | 0 | 13 | 42 | 0 |
| **34** | 45 | 55 | 0.02 | 39 | 55 | 0 | 40 | 55 | 0.02 |
| **35** | 109 | 257 | 0.01 | 98 | 144 | 0.01 | 61 | 131 | 0.01 |
| **Total** | **5492** | **12080** | **1.79** | **16431** | **17772** | **3.58** | **3494** | **4658** | **1.22** |

**TABLE (4.4)**

**PERCENTAGE PERFORMANCE OF TABLE (4.3)**

| TOOLS | WU & CHEN (2010) | New Hybrid |
|---|---|---|
| NOI | 100 % | 21.3% |
| NOFG | 100 % | 26.2 % |
| TIME | 100 % | 34.1% |

| TOOLS | FR (1964) | New Hybrid |
|---|---|---|
| NOI | 100 % | 63.6% |
| NOFG | 100 % | 38.5% |
| TIME | 100 % | 68.2% |

Clearly, from the above table, we have found that the new proposed algorithm beats FR algorithm in about (37%) NOI; (61%) NOFG and (31% )Time. From this table we have also concluded that the new algorithm beats the standard WC-algorithm in about (78%) NOI; (73%) NOFG and (65% )Time.

**Table (4.5)**

**COMPARISON BETWEEN: NEW ,(WU & CHEN) and (FR) METHODS**

**n= 100, 400,700 ,1000**

| Prob. | FR/1964 NOI/NOFG/TIME | | | Wu & Chen/2010 NOI/NOFG/TIME | | | New Hybrid CG NOI/NOFG/TIME | | |
|---|---|---|---|---|---|---|---|---|---|
| **1** | 107 | 206 | 0.06 | 892 | 1008 | 0.67 | 230 | 346 | 0.18 |
| **2** | 123 | 296 | 0 | 87 | 164 | 0.01 | 84 | 161 | 0.01 |
| **3** | 57 | 160 | 0.03 | 29 | 38 | 0.02 | 35 | 78 | 0.03 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **4** | 330 | 626 | 0.2 | 627 | 681 | 0.41 | 173 | 214 | 0.09 |
| **5** | 138 | 270 | 0.03 | 406 | 443 | 0.06 | 92 | 125 | 0.03 |
| **6** | 111 | 219 | 0.11 | 137 | 147 | 0.1 | 115 | 124 | 0.1 |
| **7** | 144 | 391 | 0.03 | 3426 | 3457 | 0.46 | 29 | 40 | 0 |
| **8** | 38 | 140 | 0.08 | 66 | 79 | 0.1 | 66 | 79 | 0.11 |
| **9** | 102 | 242 | 0.01 | 141 | 183 | 0.01 | 65 | 113 | 0.02 |
| **10** | 84 | 202 | 0.05 | 84 | 119 | 0.03 | 75 | 94 | 0.04 |
| **11** | 149 | 303 | 0.03 | 233 | 285 | 0.07 | 115 | 168 | 0.03 |
| **12** | 82 | 208 | 0 | 98 | 138 | 0.02 | 87 | 140 | 0.01 |
| **13** | 31 | 142 | 0.02 | 14 | 28 | 0 | 14 | 28 | 0 |
| **14** | 295 | 559 | 0.05 | 510 | 538 | 0.08 | 134 | 164 | 0.01 |
| **15** | 318 | 615 | 0.04 | 2156 | 2188 | 0.42 | 59 | 91 | 0 |
| **16** | 66 | 171 | 0.02 | 53 | 61 | 0.06 | 33 | 41 | 0.04 |
| **17** | 63 | 162 | 0.03 | 36 | 48 | 0.01 | 36 | 47 | 0.01 |
| **18** | 69 | 178 | 0.03 | 43 | 52 | 0.02 | 43 | 54 | 0.02 |
| **19** | 158 | 324 | 0.03 | 510 | 545 | 0.07 | 102 | 141 | 0.02 |
| **20** | 57 | 160 | 0.02 | 29 | 38 | 0.01 | 30 | 47 | 0 |
| **21** | 151 | 324 | 0.02 | 379 | 412 | 0.07 | 121 | 150 | 0.02 |
| **22** | 159 | 272 | 0.03 | 334 | 355 | 0.14 | 85 | 106 | 0.04 |
| **23** | 308 | 618 | 0.12 | 386 | 428 | 0.19 | 99 | 147 | 0.07 |
| **24** | 68 | 184 | 0 | 56 | 84 | 0 | 41 | 79 | 0.01 |
| **25** | 103 | 246 | 0.03 | 98 | 124 | 0.01 | 102 | 119 | 0.03 |
| **26** | 289 | 576 | 0.11 | 333 | 357 | 0.14 | 138 | 170 | 0.07 |
| **27** | 58 | 161 | 0.01 | 61 | 84 | 0.02 | 49 | 80 | 0 |
| **28** | 49 | 156 | 0.05 | 32 | 64 | 0.03 | 32 | 66 | 0.01 |
| **29** | 55 | 158 | 0.05 | 35 | 43 | 0.03 | 35 | 43 | 0.03 |
| **30** | 54 | 191 | 0.04 | 18 | 30 | 0.01 | 18 | 30 | 0.02 |
| **31** | 84 | 202 | 0.05 | 76 | 95 | 0.05 | 75 | 94 | 0.03 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **32** | 266 | 531 | 0.03 | 393 | 437 | 0.05 | 180 | 224 | 0.03 |
| **33** | 4 | 12 | 0 | 9 | 27 | 0 | 7 | 21 | 0 |
| **34** | 36 | 44 | 0 | 32 | 44 | 0 | 32 | 44 | 0.02 |
| **35** | 87 | 205 | 0.02 | 78 | 115 | 0.02 | 49 | 108 | 0.01 |
| **Total** | **4293** | **9454** | **1.43** | **11897** | **12939** | **3.39** | **2680** | **3776** | **1.14** |

**TABLE (4.6)**

**PERCENTAGE PERFORMANCE OF TABLE( 4.5)**

| **TOOLS** | **WU & CHEN (2010)** | **New Hybrid** |
|---|---|---|
| NOI | 100 % | 22.5 % |
| NOFG | 100 % | 29.2 % |
| TIME | 100 % | 33.6 % |

| **TOOLS** | **FR (1964)** | **New Hybrid** |
|---|---|---|
| NOI | 100 % | 62.4 % |
| NOFG | 100 % | 39.9 % |
| TIME | 100 % | 79.7 % |

Clearly, from the above table, we have found that the new proposed algorithm beats FR algorithm in about (37%) NOI; (60%) NOFG and (20% )Time. From this table we have also concluded that the new algorithm beats the standard WC-algorithm in about (77%) NOI; (70%) NOFG and (66% )Time.

Figures (4.1 and 4.2) show the Dolan and Moré CPU performance percentages of HYWCFR versus FR and WC, respectively. When comparing the new method against FR (Fig.4.1) subject to the number of iterations, we see the new method was better in 33 problems (i.e. it achieved the minimum number of iterations in 33 problems), FR was better in 2 problems, etc.

**Figure (4.1)**

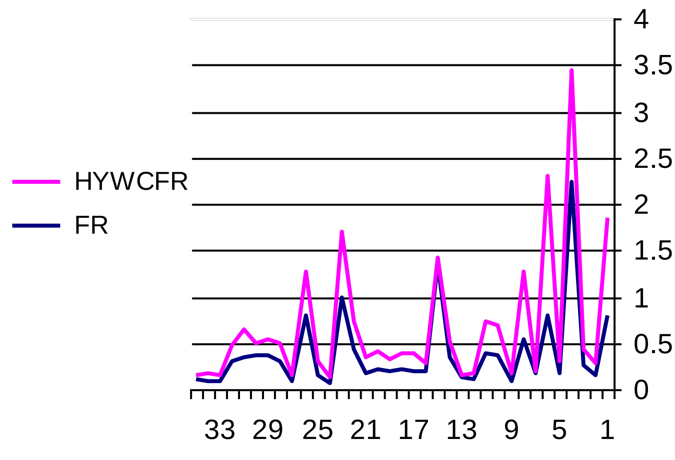
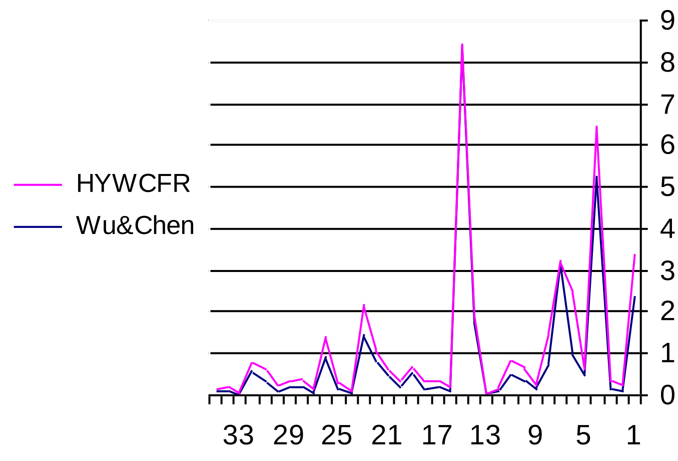



**Figure (4.2)**




Similarly, in (Fig.4.2) we see the number of problems for which HYWCFR is better than WC. Observe that the convex combination of FR and WC expressed as in (12) is far more successful than FR and WC algorithms. From the tables and figures, we can see that the new method was the top performer. Since these codes use the same Wolfe line search and the same stopping criterion, they differ in their choice of the search direction. Hence, among these CG-algorithms, the new algorithm appears to generate the best search direction.

## 5. Conclusions.

There exists a large variety of CG-algorithms. In this paper, we have presented a new hybrid CG-algorithm in which the parameter $\beta_k$ is computed as a convex combination of $\beta_k^{FR}$ and $\beta_k^{WCh}$. For uniformly convex functions, if the step-size $s_k$ approaches zero, the gradient is bounded in the sense that $\eta_1\|s_k\|^2 \le \|g_k\|^2 \le \eta_2\|s_{k-1}\|^2$ and the line search satisfies the strong Wolfe conditions, then our hybrid CG-algorithm is globally convergent. For general nonlinear functions, if the parameter $\theta_k$ from $\beta_k^{HYWCFR}$ definition is bounded, then our new hybrid CG-algorithm is globally convergent. The performance percentage of our new proposed algorithm is higher than those of the well established CG-algorithms for a test set consisting of (35) unconstrained optimization problems, some of them from the CUTE library. Additionally, the proposed hybrid CG-algorithm is more robust than the FR and the WC-algorithms.

## Appendix.

1- Extended Trigonometric Function.

2- Extended Penalty Function.

3- Raydan 2 Function.

4- Diagonal2 Function.

5- Generalized Tridiagonal-1 Function.

6- Extended Tridiagonal-1 Function.

7- Extended 3-Exponential Terms Function.

8- Diagonal4 Function.

9- Diagonal5 Function.

10- Extended Himmelblau Function.

11- Extended PSC1 Function.

12- Extended Block Diagonal BD1 Function.

13- Extended EP1 Function.

14- DIXMAANA CUTE- Function.

15- DIXMAANB CUTE- Function.

16- DIXMAANC CUTE- Function.

17- Broyden Tri-diagonal Function.

18- EDENSCH CUTE- Function.

19- VARDIM CUTE- Function.

20- LIARWHD CUTE- Function.

21- DIAGONAL 6 Function.

22- ENGVAL1 CUTE- Function.

23- DENSCHNA CUTE- Function.

24- DENSCHNB CUTE- Function.

25- DENSCHNF CUTE- Function.

26- Generalized Quartic GQ1 function.

27- Diagonal 7 Function.

28- Diagonal 8 Function.

29- Full Hessian Function.

30- SINCOS Function.

31- Generalized quartic GQ2 function.

32- ARGLINB CUTE-Function.

33- FLETCHCR CUTE-Function.

34- HIMMELBG CUTE-Function.

35- HIMMELBH CUTE-Function.

## REFERENCES

[1] A.Y. Al-Bayati and B.M. AL-Baro, New Hybrid Globally Convergent CG- Algorithms for Nonlinear Unconstrained Optimization, Australian Journal of Basic and Applied Sciences, 4(9), (2010), 4223-4239.

[2] B.T. Polyak, The conjugate gradient method in extreme problems). USSR Computational Mathematic Math. Phys., 9, (1969), 94–112.

[3] C.Y. Wu and G.Q. Chen, New type of conjugate gradient algorithm for unconstrained optimization problems, Journal of Systems Engineering and Electronics, 21(6), (2010), 1000-1007.

[4] D. Touati-Ahmed and C. Storey, Efficient hybrid conjugate gradient techniques, Journal Optimization Theory Application, 64, (1990), 379–397.

[5] D.C. Liu and J. Nocedal, On the limited memory BFGS method for large scale optimization, Mathematical Program, 45, (1989), 503–528.

[6] E. Polak and G. Ribière, Note Sur la Convergence de Directions Conjuguée, Rev. Francaise Informat Recherche Operationelle, 3e Année, 16, (1969), 35–43.

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

[8] I. Bongartz; A.R., Conn, N.I.M. Gould and P.L. Toint, CUTE: constrained and unconstrained testing environments, ACM Trans. Math. Software, 21, (1995), 123–160.

[9] J. Nocedal, Conjugate gradient methods and nonlinear optimization, In: L. Adams, J.L. Nazareth, (eds.) (Linear and Nonlinear Conjugate Gradient—Related Methods), (1996), 9–23. SIAM, Philadelphia.

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

[11] J.W Daniel, The conjugate gradient method for linear and nonlinear operator equations, SIAM Journal Numerical Analysis, 4, (1967) , 10–26.

[12] K.C. Kiwiel and K. Murty, Convergence of the steepest descent method for minimizing quasi-convex functions, Journal Optimization Theory Application, 89(1), (1996), 221–226.

[13] M.J.D. Powell, Restart procedures for the conjugate gradient method, Mathematical Program, 12, (1977), 241–254.

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

[15] N. Andrei, Acceleration of conjugate gradient algorithms for unconstrained optimization, Applied Mathematical Computational, 213, (2009), 361–369.

[16] N. Andrei, Accelerated hybrid conjugate gradient algorithm with modified secant condition for unconstrained optimization, Numerical Algorithms, 54, (2010) , 23–46.

[17] N. Andrei, Another hybrid conjugate gradient algorithm for unconstrained optimization, Numerical Algorithms, 47, (2008), 143–156.

[18] N. Andrei, Hybrid conjugate gradient algorithm for unconstrained optimization, Journal

Optimization Theory Application, 141, (2009), 249–264.

[19] N. Andrei, Numerical comparison of conjugate gradient algorithms for unconstrained optimization, Stud. Inform. Control, 16, (2007), 333–352.

[20] N. Andrei, Performance profiles of conjugate gradient algorithms for unconstrained optimization, In: Floudas, C.A., Pardalos, P.M. (eds.) Encyclopedia of Optimization, 2nd edition, (2009), 2938–2953. Springer, New York.

[21] P. Wolfe, Convergence conditions for ascent methods, SIAM Rev. , 11, (1969), 226–235.

[22] R. Fletcher and C. Reeves, Function minimization by conjugate Gradients, Computer Journal, 7, (1964), 149–154.

[23] R. Fletcher, Practical Methods of Optimization, Unconstrained Optimization. Wiley, New York, 1, (1987).

[24] Sh. Li, and Z. Sun, A New Hybrid Conjugate Gradient Method And Its Global Convergence For Unconstrained Optimization, International Journal of Pure and Applied Mathematics, 63(3), (2010), 285-296.

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

[26] W.W. Hager and H. Zhang, A survey of nonlinear conjugate gradient methods), Pacific Journal Optimization, 2, (2006), 35–58.

[27] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, part 1: theory, JOTA, 69, (1991), 129–137.

[28] Y.F. Hu and C. Storey, Global convergence result for conjugate gradient methods, Journal Optimization Theory Application , 71, (1991), 399–405.

[29] Y.H. Dai and Y. Yuan, (A nonlinear conjugate gradient method with a strong global convergence property). SIAM Journal of Optimization, 10, (1999), 177–182.

[30] Y.H. Dai and Y. Yuan, An efficient hybrid conjugate gradient method for unconstrained optimization, Ann. Oper. Res., 103, (2001), 33–47.

[31] Y.H. Dai; J.Y. Han, G.H. Liu, D.F. Sun, X. Yin and Y. Yuan, Convergence properties of nonlinear conjugate gradient methods, SIAM Journal Optimization, 10, (1999), 348–358.

[32] Y.H. Dai; L.Z. Liao and D. Li, On restart procedures for the conjugate gradient method). Numerical Algorithms, 35, (2004), 249–260.

[33] Y.H. Dai and L.Z. Liao, New conjugacy conditions and related nonlinear conjugate gradient methods, Application Mathematical Optimization, 43, (2001), 87–101.