

APPLICATION OF NEW HYBRID CONJUGATE GRADIENT ALGORITHMS TO NONLINEAR UNCONSTRAINED OPTIMIZATION PROBLEMS

SABREEN M. ABBAS¹, ABBAS Y. AL-BAYATI², MAYSOON M. AZIZ³

Abstract: In this research, we introduce a novel class of hybrid nonlinear conjugate gradient algorithms that generate an adequate search direction for descending gradients at each iteration. This property persists whether or not a line search is accurate and regardless of whether or not the objective function is convex. For functions that are not convex in general, we show that the suggested approaches converge globally under certain conditions. Numerical results demonstrate the effectiveness of these new hybrid methods in their application to selected testing problems compared to the standard Mevlana CG algorithm.

INTRODUCTION

To start by giving any issue of unconstrained optimization we need to know this issue as follows:

$$\min \{f(x): x \in R^n\} \quad (1)$$

where $f: R \rightarrow R^n$ is a nonlinear, smooth function, and its gradient is represented by $g(x) = \nabla f(x)$. The iterative formula of the conjugate gradient methods is given by:

$$x_{k+1} = x_k + \alpha_k d_k, k = 0, 1, 2 \quad (2)$$

where x_k is the current iterate point and α_k is the step length, which is computed by carrying out a line search, and d_k is the search direction defined by:

$$d_{k+1} = \begin{cases} -g_{k+1} & \text{if } k = 0 \\ -g_{k+1} + \beta_k d_k & \text{if } k \geq 1 \end{cases} \quad (3)$$

where β_k is scalar. The differences highlight CG methods in their speed and performance by specifying the numerical parameter β_k . This is a set of the effective versions of β_k as contained in their sources [17].

$$\beta_k^{HS} = g_{k+1}^T (g_{k+1} - g_k) / (g_{k+1} - g_k)^T d_k \quad (4)$$

$$\beta_k^{FR} = g_{k+1}^T g_{k+1} / g_k^T g_k$$

$$\beta_k^{CD} = -g_{k+1}^T g_{k+1} / d_k^T g_k$$

$$\beta_k^{LS} = -g_{k+1}^T (g_{k+1} - g_k) / d_k^T g_k$$

$$\beta_k^{PR} = g_{k+1}^T (g_{k+1} - g_k) / g_k^T g_k$$

$$\beta_k^{DY} = g_{k+1}^T g_{k+1} / (g_{k+1} - g_k)^T d_k$$

$$\beta_k^{RMIL} = g_{k+1}^T (g_{k+1} - g_k) / d_k^T d_k$$

where $y_k = g_{k+1} - g_k$ and $\| \cdot \|$ referees for the Euclidean norm of any vectors.

However, it is important to note that the behavior of the following approaches might vary significantly for general objective functions, even if they may appear equivalent [2] in cases where f is a strictly convex quadratic function and k is determined using exact line searches (ELS). Wolfe conditions, strong Wolfe conditions, and strong

¹ Department of Mathematics, College of Computer Sciences and Mathematics, University of Mosul, Iraq. e-mail: mssabreen6@gmail.com

² University of Telafer, Iraq. Email: profabbasalbavati@yahoo.com

³ University of Mosul, Iraq. Email: aziz_maysoon@uomosul.edu.iq

Keywords: Conjugate Gradient Method, Hybrid Conjugate Gradient, Exact Line Search, Sufficient Descent, Global Convergence

Wolfe conditions are examples of inexact line searches (ILS) that are hoped for in the convergence analysis of conjugate gradient methods.

(i) The standard Wolfe line search:

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \\ d_k^T g(x_k + \alpha_k d_k) \geq \sigma d_k^T g_k \end{cases} \quad (5a)$$

(ii) The strong Wolfe line search:

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \\ |d_k^T g(x_k + \alpha_k d_k)| \leq -\sigma d_k^T g_k \end{cases} \quad (5b)$$

(iii) Equations (ii) can be represented by:

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \\ \sigma d_k^T g_k \leq d_k^T g(x_k + \alpha_k d_k) \leq 0 \end{cases} \quad (6)$$

Where $0 < \delta < \frac{1}{2}$ and $\delta < \sigma < 1$

Global convergence of the FR technique was proven by Zoutendijk [18] and M. Al-Baali [8] for generic functions. Using several font-specific searches. Mr. Powell recited that the [10] PRP approach may cycle for non-convex functions, and even with precise Line searches, they still don't get close to the solution. Although achieving Universal convergence would be satisfactory, the FR approach is far less effective than the PRP approach (HS, LS) when dealing with actual data. To rephrase, PR encompasses both the HS and LS approaches to practical arithmetic. As far as conjugate gradient (CG) methods go, this one seems to be the most effective. Because a restart is executed by these techniques in the event of a negative trend. The DY and CD approaches experience the same problem. As a result, numerical results are poor for the CD, DY, and FR methods of convergence [1, 5, 14]. Okay, decent outcomes for the nonlinear conjugate gradient approach are provided. Recently, [3] and [4] proposed some hybrid techniques, which we will refer to as the H1 method and the H2 method, respectively, and which combine the good numerical performance of the PR and HS methods with the nice global convergence qualities of the FR and DY methods.

$$\beta_k^{H_1} = \max\{0, \min\{\beta_k^{PR}, \beta_k^{FR}\}\} \quad (7) \quad \beta_k^{H_2} = \max\{0, \min\{\beta_k^{HS}, \beta_k^{DY}\}\}$$

In the last year, Wei, et al. [15] introduced an odd plan of its predecessors from the formulas of the PR methodology referred to as the WYL methodology. Zhang recommended through his study a lot of and has improved WYL normal methodology victimization the new CG, additionally, it was well-tried that the strategy of NPR met the traditional demand ratios beneath conditions of line search SWP [9]. Finally, Zhang et al. proposed a modified version of the CG approach called the MPR methodology, whereas Dai and Wen [6] proposed a modified version of the NPR methodology known as the DPR methodology.

Some Important Updates for Recent Different CG-Algorithms [6, 11].

Researchers	Formulas
WYL	$\beta_k^{WYL} = \frac{g_{k+1}^T \left(g_{k+1} - \frac{\ g_{k+1}\ }{\ g_k\ } g_k \right)}{\ g_k\ ^2}$
NPR	$\beta_k^{NPR} = \frac{\ g_{k+1}\ ^2 - \frac{\ g_{k+1}\ }{\ g_k\ } g_{k+1}^T g_k }{\ g_k\ ^2}$
MPR	$\beta_k^{MPR} = \frac{\delta \ g_{k+1}\ ^2}{\ g_k\ ^2 + g_k^T d_k }$
DPR	$\beta_k^{DPR} = \frac{\ g_{k+1}\ ^2 - \frac{\ g_{k+1}\ }{\ g_k\ } g_{k+1}^T g_k }{\mu g_{k+1}^T d_k + \ g_k\ ^2}$

The researchers [12] have proposed a new hybrid CG algorithm. The β_k which is known

as $\beta_k^{Maulana}$, we designed a combination of RMIL-CG and a new formula for the numerator by adding another negative gradient of WYL-CG and retaining the original denominator of RMIL-CG. Hence, Maulana1 and Maulana2 become, in general,

$$\beta_k^{Maulana} = \max\{0, \min\{\beta_k^{RMIL}, \beta_k^{MMSSS}\}\} \quad (8)$$

where:

$$\beta_k^{MMSSS1} = \frac{g_k^T \left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1} - g_{k-1} \right)}{g_k^T (g_k - d_{k-1})} \quad (9)$$

and

$$\beta_k^{MMSSS2} = \frac{\frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} \|g_{k-1}\|}{(1-\mu)\|d_{k-1}\|^2 + \mu\|g_{k-1}\|^2} |g_k^T g_{k-1}| - |g_k^T g_{k-1}|}{\mu = 0.6.} \quad (10)$$

Similarly, four new hybrids (CG) algorithms, which we call the (SB1, SB2, NSB1, and NSB2) CG methods, respectively:

$$\beta_k^{SB1} = \max\{0, \min\{\beta_k^{FR}, \beta_k^{MMSSS1}\}\} \quad (12)$$

$$\beta_k^{SB2} = \max\{0, \min\{\beta_k^{FR}, \beta_k^{MMSSS2}\}\} \quad (13)$$

Now for these two new search directions algorithms, we have:

$$d_k^{SB1} = -g_k + \beta_k^{MMSSS1} d_{k-1} - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} g_k, \quad (14)$$

$$d_k^{SB2} = -g_k + \beta_k^{MMSSS2} d_{k-1} - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} g_k, \quad (15)$$

proceeding in the same manner, we can construct another two new CG algorithms:

$$d_k^{NSB1} = - \left(1 + \beta_k^{SB1} \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} \right) g_k + \beta_k^{SB1} d_{k-1}, \text{ call (NEW1)} \quad (16)$$

$$d_k^{NSB2} = - \left(1 + \beta_k^{SB2} \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} \right) g_k + \beta_k^{SB2} d_{k-1}, \text{ Call (NEW2)} \quad (17)$$

These new hybrid CG algorithms still satisfy

$$g_k^T d_k = -\|g_k\|^2 \quad (18)$$

OUTLINES OF SOME NEW AND PUBLISHED CG-ALGORITHMS.

Maulana et al. CG algorithms

1. For $X_0 \in R^n, 0 < \epsilon < 1, 0 < \delta < \frac{1}{2}$, and $\delta < \sigma < 1$,
set $d_0 = -g_0, k = 0$.
2. If $\|g_k\| < \epsilon$, then stop, otherwise go to the next step.
3. Compute step size α_k by strong Wolfe line search (6).
4. Let $x_{k+1} = x_k + \alpha_k d_k$, if $\|g_{k+1}\| < \epsilon$, then stop.
5. Calculate the search directions:
 $d_{k+1} = -g_{k+1} + \beta_k^{Maulana2} d_k$, where $\beta_k^{Maulana1}$ is defined by (10).
 $d_{k+1} = -g_{k+1} + \beta_k^{Maulana2} d_k$, where $\beta_k^{Maulana2}$ is defined by (11).
6. Set $k = k + 1$, and go to step2.

First Two New β_k^{SB1} and β_k^{SB2} for CG algorithms

1. For $X_0 \in R^n, 0 < \epsilon < 1, 0 < \delta < \frac{1}{2}$, and $\delta < \sigma < 1$,
set $d_0 = -g_0, k = 0$.
2. If $\|g_k\| < \epsilon$, then stop, otherwise go to the next step.
3. Compute step size α_k by strong Wolfe line search (6).
4. Let $x_{k+1} = x_k + \alpha_k d_k$, if $\|g_{k+1}\| < \epsilon$, then stop.
5. Calculate the first and second new search directions:
 $d_k^{SB1} = -g_k + \beta_k^{MMSSS1} d_{k-1} - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} g_k$, which is defined in (15)
 $d_k^{SB2} = -g_k + \beta_k^{MMSSS2} d_{k-1} - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} g_k$, which is defined by (16).
6. Set $k = k + 1$, and go to step2.

Second Two New Hybrid CG Algorithms, namely (NEW1 and NEW2).

1. For $X_0 \in R^n, 0 < \epsilon < 1, 0 < \delta < \frac{1}{2}$, and $\delta < \sigma < 1$,
set $d_0 = -g_0, k = 0$.
2. If $\|g_k\| < \epsilon$, then stop, otherwise go to the next step.
3. Compute step size α_k by Wolfe line search (6).
4. Let $x_{k+1} = x_k + \alpha_k d_k$, if $\|g_{k+1}\| < \epsilon$, then stop.
5. Calculate the new search directions NEW1 and NEW2 by:

$$d_k^{NSB1} = - \left(1 + \beta_k^{SB1} \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} \right) g_k + \beta_k^{SB1} d_{k-1}, \text{ defined in (17)}$$

$$d_k^{NSB2} = - \left(1 + \beta_k^{SB2} \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}} \right) g_k + \beta_k^{SB2} d_{k-1}, \text{ defined in (18)}$$
6. Set $k = k + 1$, and go to step2.

THE GLOBAL CONVERGENCE PROPERTY FOR THE NEW CG ALGORITHMS

Assumption A

Let $f(x)$ have its lower bound defined on the set

$$S = \{x \in R^n, f(x) \leq f(x_0)\}, \quad (19)$$

where x_0 is the initial point.

Assumption B

The objective function is continuously differentiated and the gradient of the objective function is Lipchitz continuous in some neighborhood N of S , such that:

$$\|g(x) - g(y)\| \leq L\|x - y\| \quad \forall x, y \in N. \quad (20)$$

The sequence $\{x_k\}$ produced by NEW1 and NEW2 is included in S since $\{f(x_k)\}$ is decreasing. Furthermore, from Assumption A, we can deduce that for any constant B and any B and $\gamma_1 > 0$, we have:

$$\|x\| \leq B, \|g(x)\| \leq \gamma_1, \forall x \in S \quad (21)$$

Assumption A's conditions are taken as given throughout the paper's subsequent sections. Then, there is a handy lemma that was first presented in [16, 18].

Lemma 3.3: The sequence $\{x_k\}$ is created by equation (2), where d_k represents a descending direction. If α_k is determined by the Wolfe line search (5), then we have:

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \quad (22)$$

From Lemma 3.3 and from (15)-(18) for the new SB1, SB2, NEW1, and NEW2 algorithms with the Wolfe line search, we can easily obtain the following condition:

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty \quad (23)$$

We now establish the global convergence theorem for NEW1 and NEW2 algorithms.

Theorem 3.4

Given that d_k obtained by NEW1 satisfies Equation (17), we have

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

Proof: Let's pretend that the desired result can't possibly hold, that is, that there is no constant $\epsilon > 0$ such that:

$$\|g_k\| \geq \epsilon \quad \forall k \geq 0 \quad (24)$$

$$\|g_k\|^2 \geq \epsilon^2 \Leftrightarrow \frac{1}{\|g_k\|^2} \leq \frac{1}{\epsilon^2}$$

Set $h_k = 1 + \beta_k^* \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}$, then we have:

$$d_k = -h_k g_k + \beta_k^* d_{k-1} \quad (25)$$

Multiply the (26) by g_k from the left

$$g_k^T d_k = -h_k \|g_k\|^2 + \beta_k^* g_k^T d_{k-1}$$

Where

$$g_k^T d_k = -\|g_k\|^2 \quad (26)$$

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

$$\beta_k^* g_k^T d_{k-1} = h_k \|g_k\|^2 - \|g_k\|^2$$

From (26) and (27) it follows that

$$\|d_k\|^2 = \|\beta_k^* d_{k-1} - h_k g_k\|^2$$

$$\begin{aligned} \|d_k\|^2 &= \beta_k^{*2} \|d_{k-1}\|^2 - 2h_k \beta_k^* d_{k-1}^T g_k + h_k^2 \|g_k\|^2 = (\beta_k^*)^2 \|d_{k-1}\|^2 - 2h_k (h_k \|g_k\|^2 - \|g_k\|^2) + \\ h_k^2 \|g_k\|^2 &= (\beta_k^*)^2 \|d_{k-1}\|^2 - 2h_k^2 \|g_k\|^2 + 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 = (\beta_k^*)^2 \|d_{k-1}\|^2 - 2h_k^2 \|g_k\|^2 + \\ 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 &= (\beta_k^*)^2 \|d_{k-1}\|^2 + 2h_k \|g_k\|^2 - h_k^2 \|g_k\|^2 \quad (27) \end{aligned}$$

Dividing both sides of (28) by $(g_k^T d_k)^2$ we get from (27) and the definition of β_k^* , that

$$\begin{aligned} \frac{\|d_k\|^2}{(g_k^T d_k)^2} &= (\beta_k^*)^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{2h_k}{g_k^T d_k} - \frac{h_k^2 \|g_k\|^2}{(g_k^T d_k)^2} \frac{\|d_k\|^2}{(g_k^T d_k)^2} = (\beta_k^*)^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{1}{\|g_k\|^2} (h_k^2 + 2h_k + 1 - 1) \\ &= (\beta_k^*)^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{1}{\|g_k\|^2} ((h_k + 1)^2 - 1) = (\beta_k^*)^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{(h_k + 1)^2}{\|g_k\|^2} + \frac{1}{\|g_k\|^2} \\ &\leq \frac{\|d_{k-1}\|^2}{\|g_k\|^4} \leq \sum_{i=0}^k \frac{1}{\|g_i\|^4} \leq \frac{k}{\epsilon^2} \end{aligned}$$

Theorem 3.5

Let $\{x_k\}$ be generated by the NEW1 method, where d_k satisfies (17), then we have then we have:

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

Proof: Suppose that by contradiction the desired conclusion is not that is to say there exists a constant $\epsilon > 0$ such that:

$$\|g_k\| \geq \epsilon \quad \forall k \geq 0 \quad (28)$$

$$\|g_k\|^2 \geq \epsilon^2 \Leftrightarrow \frac{1}{\|g_k\|^2} \leq \frac{1}{\epsilon^2}$$

Set $h_k = 1 + \beta_k^{SB1} \frac{g_k^T d_{k-1}}{\|g_k\|^2}$, then we have

$$d_k = -h_k g_k + \beta_k^{SB1} d_{k-1} \quad (29)$$

Multiply the (30) by g_k from the left

$$g_k^T d_k = -h_k \|g_k\|^2 + \beta_k^{SB1} g_k^T d_{k-1}$$

Where

$$g_k^T d_k = -\|g_k\|^2 \quad (30)$$

$$-\|g_k\|^2 = -h_k \|g_k\|^2 + \beta_k^{SB1} g_k^T d_{k-1}$$

$$\beta_k^{SB1} g_k^T d_{k-1} = h_k \|g_k\|^2 - \|g_k\|^2$$

From (30) and (31) it follows that

$$\|d_k\|^2 = \|\beta_k^{SB1} d_{k-1} - h_k g_k\|^2$$

$$\begin{aligned} \|d_k\|^2 &= (\beta_k^{SB1})^2 \|d_{k-1}\|^2 - 2h_k \beta_k^{SB1} d_{k-1}^T g_k + h_k^2 \|g_k\|^2 = (\beta_k^{SB1})^2 \|d_{k-1}\|^2 - 2h_k (h_k \|g_k\|^2 - \|g_k\|^2) + \\ h_k^2 \|g_k\|^2 &= (\beta_k^{SB1})^2 \|d_{k-1}\|^2 - 2h_k^2 \|g_k\|^2 + 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 = (\beta_k^{SB1})^2 \|d_{k-1}\|^2 - 2h_k^2 \|g_k\|^2 + \\ 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 &= (\beta_k^{SB1})^2 \|d_{k-1}\|^2 + 2h_k \|g_k\|^2 - h_k^2 \|g_k\|^2 \quad (31) \end{aligned}$$

Dividing both sides of (32) by $(g_k^T d_k)^2$ we get from (31) and the definition of β_k^{SB1} ,

$$\leq \sum_{i=0}^k \frac{1}{\|g_i\|^4} \leq \frac{k}{\epsilon^2}.$$

Theorem 3.6

Given that d_k obtained by NEW2 satisfies Equation (18), we have

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

Proof: Let's pretend that the desired result can't possibly hold, that is, that there is no constant $\epsilon > 0$ such that:

$$\|g_k\| \geq \epsilon \quad \forall k \geq 0 \quad (32)$$

$$\|g_k\|^2 \geq \epsilon^2 \Leftrightarrow \frac{1}{\|g_k\|^2} \leq \frac{1}{\epsilon^2}$$

Let $h_k = 1 + \beta_k^{MMSSS2} \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}$, then we have:

$$d_k = -h_k g_k + \beta_k^{MMSSS2} d_{k-1} \quad (33)$$

Multiply the (34) by g_k from the left

$$g_k^T d_k = -h_k \|g_k\|^2 + \beta_k^{MMSSS2} g_k^T d_{k-1}$$

where

$$- \|g_k\|^2 = -h_k \|g_k\|^2 + \beta_k^{MMSSS2} g_k^T d_{k-1} \beta_k^{MMSSS2} g_k^T d_{k-1} = h_k \|g_k\|^2 - \|g_k\|^2$$

From (34) and (35) it follows that

$$\begin{aligned} \|d_k\|^2 &= \|\beta_k^{MMSSS2} d_{k-1} - h_k g_k\|^2 \\ \|d_k\|^2 &= (\beta_k^{MMSSS2})^2 \|d_{k-1}\|^2 - 2h_k \beta_k^{MMSSS2} d_{k-1}^T g_k + h_k^2 \|g_k\|^2 = (\beta_k^{MMSSS2})^2 \|d_{k-1}\|^2 - 2h_k (h_k \|g_k\|^2 - \\ \|g_k\|^2) + h_k^2 \|g_k\|^2 &= (\beta_k^{MMSSS2})^2 \|d_{k-1}\|^2 - 2h_k^2 \|g_k\|^2 + 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 = (\beta_k^{MMSSS2})^2 \|d_{k-1}\|^2 - \\ 2h_k^2 \|g_k\|^2 + 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 &= (\beta_k^{MMSSS2})^2 \|d_{k-1}\|^2 + 2h_k \|g_k\|^2 - h_k^2 \|g_k\|^2 \quad (35) \end{aligned}$$

Dividing both sides of (36) by $(g_k^T d_k)^2$ we get from (35) and the definition of β_k^{MMSSS2} , that

$$\begin{aligned} \frac{\|d_k\|^2}{(g_k^T d_k)^2} &= (\beta_k^{MMSSS2})^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{2h_k}{g_k^T d_k} - \frac{h_k^2 \|g_k\|^2}{(g_k^T d_k)^2} \frac{\|d_k\|^2}{(g_k^T d_k)^2} = (\beta_k^{MMSSS2})^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{1}{\|g_k\|^2} (h_k^2 + 2h_k + 1 - 1) = \\ (\beta_k^{MMSSS2})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{1}{\|g_k\|^2} ((h_k + 1)^2 - 1) &= (\beta_k^{MMSSS2})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{(h_k + 1)^2}{\|g_k\|^2} + \frac{1}{\|g_k\|^2} \frac{\|d_k\|^2}{\|g_k\|^4} \leq \frac{\|d_{k-1}\|^2}{\|g_k\|^4} \leq \sum_{i=0}^k \frac{1}{\|g_i\|^4} \leq \frac{k}{\epsilon^2}. \end{aligned}$$

Theorem 3.7

let $\{x_k\}$ be generated by the NEW2 method, where d_k satisfies (18), then we have:

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

Proof: Suppose that by contradiction the desired conclusion is not that is to say there exists a constant $\epsilon > 0$ such that:

$$\|g_k\| \geq \epsilon \quad \forall k \geq 0 \quad (36)$$

$$\|g_k\|^2 \geq \epsilon^2 \Leftrightarrow \frac{1}{\|g_k\|^2} \leq \frac{1}{\epsilon^2}$$

Set $h_k = 1 + \beta_k^{SB2} \frac{g_k^T d_{k-1}}{\|g_k\|^2}$, then we have

$$d_k = -h_k g_k + \beta_k^{SB2} d_{k-1} \quad (37)$$

Multiply the (38) by g_k from the left

$$g_k^T d_k = -h_k \|g_k\|^2 + \beta_k^{SB2} g_k^T d_{k-1}$$

Where

$$- \|g_k\|^2 = -h_k \|g_k\|^2 + \beta_k^{SB2} g_k^T d_{k-1} \beta_k^{SB2} g_k^T d_{k-1} = h_k \|g_k\|^2 - \|g_k\|^2$$

From (38) and (39) it follows that

$$\begin{aligned} \|d_k\|^2 &= \|\beta_k^{SB2} d_{k-1} - h_k g_k\|^2 \\ \|d_k\|^2 &= (\beta_k^{SB2})^2 \|d_{k-1}\|^2 - 2h_k \beta_k^{SB2} d_{k-1}^T g_k + h_k^2 \|g_k\|^2 = (\beta_k^{SB2})^2 \|d_{k-1}\|^2 - 2h_k (h_k \|g_k\|^2 - \|g_k\|^2) + \\ h_k^2 \|g_k\|^2 &= (\beta_k^{SB2})^2 \|d_{k-1}\|^2 - 2h_k^2 \|g_k\|^2 + 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 = (\beta_k^{SB2})^2 \|d_{k-1}\|^2 - 2h_k^2 \|g_k\|^2 + \\ 2h_k \|g_k\|^2 + h_k^2 \|g_k\|^2 &= (\beta_k^{SB2})^2 \|d_{k-1}\|^2 + 2h_k \|g_k\|^2 - h_k^2 \|g_k\|^2 \quad (39) \end{aligned}$$

Dividing both sides of (40) by $(g_k^T d_k)^2$ we get from (39) and the definition of β_k^{SB2} , that

$$\begin{aligned} \frac{\|d_k\|^2}{(g_k^T d_k)^2} &= (\beta_k^{SB2})^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{2h_k}{g_k^T d_k} - \frac{h_k^2 \|g_k\|^2}{(g_k^T d_k)^2} \frac{\|d_k\|^2}{(g_k^T d_k)^2} = (\beta_k^{SB2})^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{1}{\|g_k\|^2} (h_k^2 + 2h_k + 1 - 1) = \\ (\beta_k^{SB2})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{1}{\|g_k\|^2} ((h_k + 1)^2 - 1) &= (\beta_k^{SB2})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{(h_k + 1)^2}{\|g_k\|^2} + \frac{1}{\|g_k\|^2} \frac{\|d_k\|^2}{\|g_k\|^4} \leq \frac{\|d_{k-1}\|^2}{\|g_k\|^4} \leq \sum_{i=0}^k \frac{1}{\|g_i\|^4} \leq \frac{k}{\epsilon^2}. \end{aligned}$$

NUMERICAL EXPERIMENTS

Here, we take a quick look at some computation tables to evaluate the relative efficacy of the mathematical implementations of forty test problems involving packaging, human resource management, and planned new standards (Maulana1 and 2) and (NEW1) and (NEW2) algorithms. You can learn more about it in [13, 17]. For each of our free and large-scale optimization problems, we have selected unique test functions. These modified Visual FORTRAN CG algorithms are enforced using the Wolfe-Powell line search technique. The Intel(R) Core (TM) i7-6700HQ processor in a single laptop was used for all the computational experiments, Core TM, i7-3612QM (2.10GHZ)

central processing unit, 6GB of random access memory, and Windows 10. To determine the efficacy of the NEW1 and NEW2 CG methods, we compared them to the Maulana1 and Maulana2 CG methods by subjecting them to constant problems, presuming that the best algorithmic rule would require fewer iterations, NOI; fewer function and gradient evaluations, NOFG; fewer seconds spent inducing the answer purpose processor; and less overall work, as defined by:

TOTAL WORK = NOI + NOFG + TIME.

All of those CG-Algorithms terminated once [7]:

$\|g_k\| < 10^{-6}$.

Table 4. 1. Comparison between (NEW1) against (Maulana1) CG-Algorithms.

Test Problems	Maulana1	NEW1
	NOI/NOFG/TIME	NOI/NOFG/TIME
1- Trigonometric	95/192/0.21	176/303/0.09
2- Extended Rosenbrock (CUTE)	1760/2674/0.36	1514/1573/0.07
3- Extended White & Holst	4004/4809/0.35	2293/2358/0.10
4- Extended Beale	357/728/0.01	301/323/0.02
5- Raydan 2	48/133/0.01	29/38/0.01
6- Extended Tridiagonal 1	606/960/0.03	801/854/0.05
7- Extended Three Expo Terms	78/172/0.02	67/79/0.03
8- Generalized Tridiagonal 2	2121/2774/0.17	214/263/0.01
9- Diagonal 4	247/521/0.02	229/241/0.01
10- Diagonal 5	32/112/0.02	54/68/0.03
11- Extended Himmelblau	82/206/0.00	97/129/0.00
12- Extended PSC1	78/183/0.02	76/84/0.01
13- Extended BD1	112/248/0.01	98/138/0.01
14- Extended Hiebert	4004/4340/0.16	4004/4076/0.20
15- Extended EP1	26/125/0.00	10/22/0.00
16- Extended Tridiagonal 2	203/386/0.00	165/190/0.04
17- ARROWHEAD (CUTE)	382/758/0.08	223/255/0.02
18- NONDIA (CUTE)	3533/4053/0.19	3181/3201/0.20
19- DQDRTIC (CUTE)	453/854/0.02	396/408/0.01
20- DIXMAANA (CUTE)	57/144/0.00	36/44/0.01
21- DIXMAANB (CUTE)	56/141/0.00	40/51/0.01
22- DIXMAANC (CUTE)	61/152/0.01	47/58/0.00
23- Broyden Tridiagonal	370/683/0.03	211/222/0.02
24- Tridiagonal Perturbed Quadratic	1332/2345/0.10	699/708/0.05
25- LIARWHD (CUTE)	1726/2728/0.10	1684/1699/0.11
26- DIAGONAL 6	48/133/0.01	33/59/0.00
27- DENSCHNA (CUTE)	108/202/0.02	99/107/0.02
28- DENSCHNC (CUTE)	207/434/0.03	121/154/0.02
29- DENSCHNB (CUTE)	57/157/0.00	43/59/0.00
30- DENSCHNF (CUTE)	91/212/0.01	83/112/0.01
31- Extended Block-Diagonal BD2	230/453/0.03	147/188/0.03
32- Generalized quartic GQ1	57/158/0.01	42/80/0.00
33- DIAGONAL 7	40/127/0.01	32/65/0.00
34- Full Hessian	45/165/0.02	18/30/0.00
35- SIN COS	78/183/0.02	62/100/0.02
36- Generalized quartic GQ2	229/442/0.01	144/183/0.01
37- ARGLINB (CUTE)	4/12/0.00	7/21/0.00
38- FLETCHCR (CUTE)	179/355/0.01	208/250/0.02
39- HIMMELBG (CUTE)	32/40/0.01	31/44/0.00
40- HIMMELBH (CUTE)	64/164/0.00	75/87/0.01
Total	23.292/33.658/2.11	17.790/18.924/1.25
Total Work= NOI+NOFG+TIME	59.06	37.964

Table (4.2). Percentage performance of (NEW1) against (Maulana1) algorithms.

NEW1	Maulana1	TOOLS
76.4%	100%	NOI
56.2%	100%	NOFG
59.2%	100%	TIME

Table (4.3). No. of best NOFG test problems.

No. of equal NOFG In both	No. of best NOFG (Maulana1)	No. of best NOFG (NEW1)	Tools
3	7	30	NOI
0	3	37	NOFG
11	12	17	TIME

Tables (4.2-4.3) show that compared to the baseline Maulana1 CG method, the (NEW1) algorithm improves upon it by a factor of (33.6%, 43.8%, and 40.8%) in terms of NOI, NOFG, and TIME. Comparing the (NEW2) algorithm with the standard Maulana1 algorithm in terms of NOI, NOFG, and CPU, as shown in Table (4.3), reveals that the (NEW1) algorithm achieves the strongest results in NOI, NOFG, and TIME under the accelerated Wolfe-Powell line search, demonstrating that the (NEW1) algorithm is significantly more effective than the Maulana1 CG-algorithm. The (NEW1) algorithm achieves its best performance by making full use of all available resources (including NOI, NOFG, and TIME).

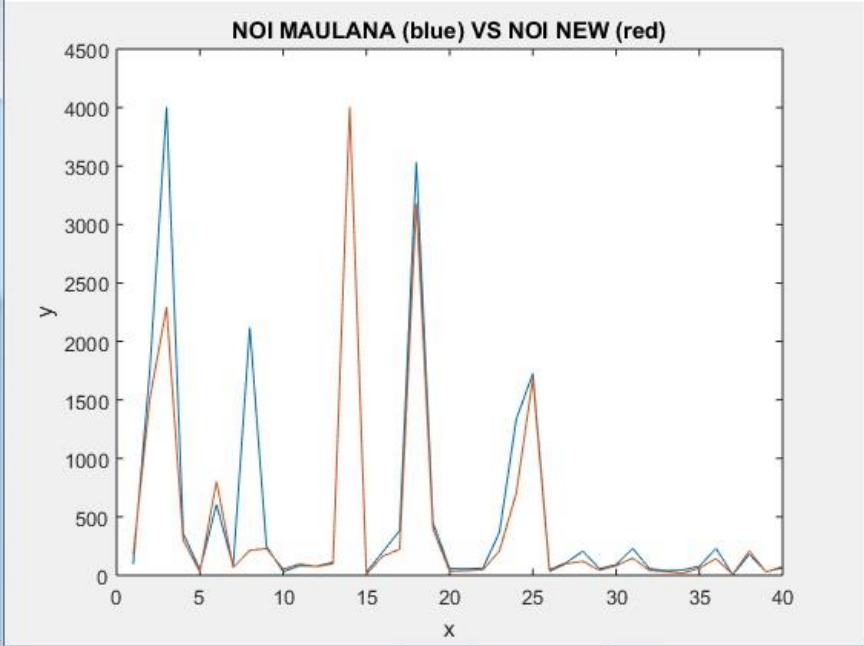


Figure1: Performance profile of (NEW1 against Maulana1) relative to the NOI.

Tables (4.4-4.6) show that compared to the baseline Maulana2 CG algorithm, the (NEW2) algorithm improves upon it by a factor of (59.4%, 67.2%, and 46.2%) in terms of NOI, NOFG, and TIME. Hence, the (NEW2) algorithm is significantly more effective than the Maulana2 CG algorithm. The (NEW2) algorithm achieves its best performance by making full use of all available resources (including NOI, NOFG, and TIME).

Table (4.4). Comparison between (NEW2) against (Maulana2) CG-Algorithms.

Test Problems	Maulana2 NOI/NOFG/TIME	NEW2 NOI/NOFG/TIME
1- Trigonometric	104/221/0.05	235/358/0.09
2- Extended Rosenbrock (CUTE)	932/1621/0.05	368/488/0.04
3- Extended White & Holst	2693/3512/0.09	411/546/0.01
4- Extended Beale	187/471/0.02	176/197/0.02
5- Raydan 2	48/133/0.01	29/38/0.00
6- Extended Tridiagonal 1	195/500/0.02	88/137/0.00
7- Extended Three Expo Terms	88/172/0.03	58/67/0.01
8- Generalized Tridiagonal 2	2101/2488/0.14	283/360/0.04
9- Diagonal 4	145/397/0.01	47/59/0.00
10- Diagonal 5	32/112/0.01	54/68/0.02
11- Extended Himmelblau	85/207/0.01	159/183/0.00
12- Extended PSC1	79/184/0.03	103/117/0.01
13- Extended BD1	104/235/0.01	105/148/0.02
14- Extended Hiebert	3942/5180/0.15	599/1066/0.03
15- Extended EP1	26/125/0.00	14/28/0.00
16- Extended Tridiagonal 2	189/375/0.00	176/209/0.01
17- ARROWHEAD (CUTE)	194/538/0.01	73/102/0.00
18- NONDIA (CUTE)	2602/2848/0.13	887/921/0.02
19- DQDRTIC (CUTE)	256/580/0.01	201/216/0.02
20- DIXMAANA (CUTE)	57/144/0.01	41/49/0.00
21- DIXMAANB (CUTE)	56/141/0.01	51/60/0.00
22- DIXMAANC (CUTE)	61/152/0.02	63/73/0.00
23- Broyden Tridiagonal	408/817/0.03	280/330/0.01
24- Tridiagonal Perturbed Quadratic	1223/2524/0.06	1757/1769/0.11
25- LIARWHD (CUTE)	1300/1931/0.08	187/205/0.01
26- DIAGONAL 6	48/133/0.01	35/78/0.02
27- DENSCHNA (CUTE)	90/189/0.01	102/110/0.01
28- DENSCHNC (CUTE)	164/387/0.05	114/147/0.02
29- DENSCHNB (CUTE)	57/157/0.01	73/86/0.00
30- DENSCHNF (CUTE)	90/206/0.00	110/128/0.02
31- Extended Block-Diagonal BD2	164/359/0.02	112/134/0.02
32- Generalized quartic GQ1	56/156/0.00	72/112/0.01
33- DIAGONAL 7	40/127/0.02	32/65/0.02
34- Full Hessian	45/165/0.02	18/30/0.00
35- SIN COS	79/184/0.02	112/132/0.02
36- Generalized quartic GQ2	186/379/0.01	164/212/0.01
37- ARGLINB (CUTE)	4/12/0.00	9/26/0.00
38- FLETCHCR (CUTE)	157/312/0.02	140/180/0.00
39- HIMMELBG (CUTE)	32/40/0.00	34/47/0.00
40- HIMMELBH (CUTE)	72/168/0.01	71/83/0.02
Total	18.391/28.572/1.19	7.643/9.364/0.64
Total Work= NOI+NOFG+TIME	48.153	17.647

Table (4.5). Percentage performance of (NEW2) against (Maulana2) algorithms.

NEW2	Maulana2	TOOLS
41.6%	100%	NOI
32.8%	100%	NOFG
53.8%	100%	TIME

Table (4.6). No. of best NOFG test problems.

No. of equal NOFG In both	No. of best NOFG (Maulana2)	No. of best NOFG (NEW2)	Tools
0	14	26	NOI
0	3	37	NOFG
9	10	21	TIME

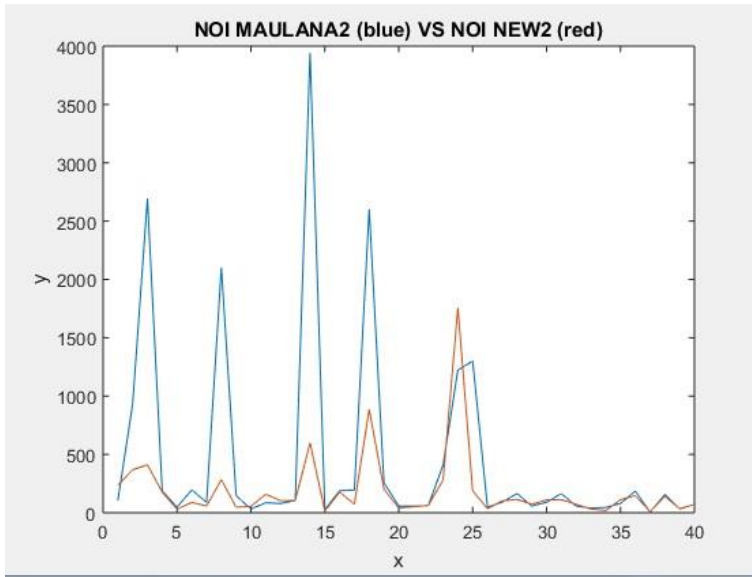


Figure 2: Performance profile of (NEWII against Maulana2) relative to the NOI.

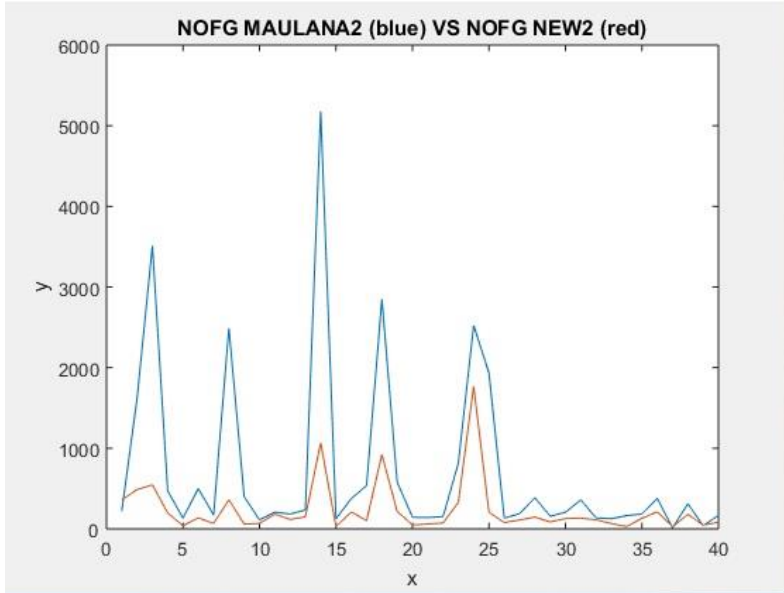


Figure 3: Performance profile of (NEWII against Maulana2) relative to the NOFG.

CONCLUSIONS

From this paper, we derive four new hybrid nonlinear conjugate gradient algorithms, namely (SB1, SB2, NEW1, and NEW2), which produce sufficient descending search directions in each iteration, and these proposed algorithms converge globally for non-convex general functions. The numerical results in Tables (4.1-4.6) show that all these new hybrid CG algorithms are effective for several large-scale test problems, according to the use of the calculation of NOI, NOFG, and TIME.

REFERENCES

- [1] C. M. R. R. Fletcher, *Function minimization by conjugate gradients*, Comput. J. **7** (1964), 149–154.
- [2] C. Cuiling, L. Liling, H. Caihong, C. Yu, *Nonlinear Conjugate Gradient Methods*, Shanghai Science and Technology Press of China, Shanghai, 2000.
- [3] D. Touati-Ahmed, C. Storey, *Efficient hybrid conjugate gradient techniques*, J. Optim. Theory Appl **64** (1990), 379–397.
- [4] Y. H. Dai, Y. Yuan, *An Efficient Hybrid Conjugate Gradient Method for Unconstrained Optimization*, Annals of Operations Research **103** (2001), 33–47.
- [5] Y. H. Dai, Y. Yuan, *A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property*, SIAM Journal on Optimization **10** (1999), 177–182.
- [6] Z. Dai, F. Wen, *Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property*, Applied Mathematics and Computation **218** (2012), 7421–7430.
- [7] E. D. Dolan, J. J. Moré, *Benchmarking optimization software with performance profiles*, Math. Program. Ser. B **91** (2002).
- [8] M. Al-Baali, *Descent property and global convergence of the fletcher-reeves method with inexact line search*, IMA J. Numer. Anal. **5** (1985).
- [9] M. Hamoda, M. Mamat, M. Rivaie, Z. Salleh, *A conjugate gradient method with Strong Wolfe-Powell line search for unconstrained optimization*, Appl. Math. Sci. **10** (2016), 13–16.
- [10] M. J. D. Powell, *Nonconvex minimization calculations and the conjugate gradient method*, Lect. Notes Math. **1066** (2006), 122–141.
- [11] M. Malik, S. S. Abas, M. Mamat, Sukono, I. S. Mohammed, *A new hybrid conjugate gradient method with global convergence properties*, Int. J. Adv. Sci. Technol. **29** (2020).
- [12] M. Maulana, M. Mustafa, S. A. Siti, M. S. Ibrahim, Sukono, *A new modification of nprp conjugate gradient method for unconstrained optimization*, Scientific Journal **9**, 4955–4970.
- [13] N. Andrei, *An unconstrained optimization test functions collection*, Adv. Model. Optim. **10** (2008), 147–161.
- [14] R. Fletcher, *Practical methods of optimization*, New York: Wiley-Interscience, 1987.
- [15] Z. Wei, S. Yao, L. Liu, *The convergence properties of some new conjugate gradient methods*, Applied Mathematics and Computation **183** (2006), 1341–1350.
- [16] Wolfe P, *Convergence Conditions for Ascent Methods*, SIAM Rev. **11** (1969).
- [17] Y. Yuan, W. Sun, *Theory and methods of optimization*, Science Press of China, Beijing, 1999.
- [18] G. Zoutendijk, *Nonlinear programming, computational methods*, Integer and Nonlinear Programming (1970), 37–86.