FUNCTION MINIMIZATION BY VARIANTS OF BFGS-CG METHOD

I. A. Osinuga and **I. O. Olofin**

Department of Mathematics,
Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.
**E-mail: osinuga08@gmail.com*

ABSTRACT

Some variants of Broyden-Fletcher-Goldfarb-Shanno conjugate gradient (BFGS-CG) method are developed in this work. This is achieved by combining BFGS method with coefficients of CG like Fletcher-Reeves, Hestenes-Stiefel, Dai and Yuan, etc. We prove the global convergence of one of these methods using Armijo-type line search. The purpose of this paper is to present these algorithms as well as their Dolan and More's performances to solve variety of large-scale unconstrained optimization problems. Some comparisons with conventional BFGS-CG algorithm are also presented.
Preliminary results show that among the variants of BFGS – CG method, BFGS – BAN competes well with the conventional BFGS – CG method in terms of number of iterations and CPU time.

Keywords: BFGS – CG method, Armijo - type line Search, global convergence, unconstrained optimization.

1. INTRODUCTION

The Conjugate Gradient (CG) methods are one of the important techniques used for seeking solutions to unconstrained optimization problems. They are popular because of their attractive features, such as, computer low memory requirements, relatively simple program, good convergence properties, among others. Applications of CG methods span across many field of endeavours, such as engineering, management sciences, operations research, social sciences, physical and behavioural sciences e.t.c.

CG algorithms can be applied to find the ideal feasible solution to a problem in a company where the problem has been modeled into an unconstrained optimization problem (Ibrahim and Rohanin, 2016). The first CG method proposed by (Hestenes and Stiefel, 1952) was established to solve positive definite symmetric matrices, linear equations. However, the first nonlinear CG method was proposed by (Fletcher and Reeves, 1964) to solve nonlinear equations. Despite the fact that CG methods are not the fastest or most robust optimization algorithms for nonlinear problems available today, they still remain very popular for Engineers, Mathematicians, and Scientists who engaged in solving large problems (Ibrahim and Rohanin, 2016).

A nonlinear optimization problem of the form below will be considered

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

where f is smooth (continuously differentiable) and g(x) is the gradient of the objective function $f(x)$. CG methods generate a sequence of points $\{x_k\}$ starting from an initial guess $x_0 \in \mathbb{R}^n$ by using the iterative scheme

$$x_{k+1} = x_k + s_k, \quad k = 0, 1, 2, ...$$

where $s_k = \alpha_k d_k$, x_k is the current iterate, and α_k is a step length which is determined by some line searches. In this paper α_k was computed using an Armijo - type line search given as follows

$$Given\ s > 0,\ \beta_k \in (0,1),\ \sigma \in (0,1) \qquad (1.2)$$

$$\alpha_k = \max\{s,\ s\beta, s\beta^2, s\beta^3, s\beta^4, \dots\} \qquad (1.3)$$

such that

*Corresponding Author

How to cite this paper: Osinuga, I. A., & Olofin, I. O. (2018). Function minimization by variants of BFGS-CG method. *Confluence Journal of Pure and Applied Sciences (CJPAS),* 1(2), 41-49.

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \sigma \alpha_k {g_k}^T d_k \qquad (1.4)$$

k = 0, 1, 2, ... Then, the sequence $\{x_k\}_{k=0}^{\infty}$ converges to the optimal point x^* which minimizes $f(x)$ .The search direction d_k is defined by

$$d_k = \begin{cases} -g_k, & k=0 \\ -g_k + \beta_k d_{k-1} & k \geq 1 \end{cases} \qquad (1.5)$$

where β_k is a scalar known as CG parameter. The formula for β_k should be chosen such that the method reduces to the linear CG method in case when f is strictly convex quadratic and the line search is exact. Well known formulae for β_k are:

Fletcher and Reeves Method, 1964

$$\beta_k^{FR} = \frac{{g_k}^T g_k}{\|g_{k-1}\|^2} \qquad (1.6)$$

Polak - Ribiere - Polyak (PRP) method, 1969

$$\beta_k^{PRP} = \frac{{g_k}^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} \qquad (1.7)$$

Hestenes and Stiefel (HS) method, 1952

$$\beta_k^{HS} = \frac{{g_k}^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \qquad (1.8)$$

Dai and Yan (DY) method, 1999

$$\beta_k^{DY} = \frac{-{d_{k-1}}^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} \qquad (1.9)$$

Liu and Storey (LS) method, 1991

$$\beta_k^{LS} = \frac{-{g_k}^T y_{k-1}}{{d_{k-1}}^T g_{k-1}} \qquad (1.10)$$

Conjugate Descent Method, (Fletcher, 1987)

$$\beta_k^{CD} = \frac{-{d_{k-1}}^T g_{k-1}}{\|g_{k-1}\|^2} \qquad (1.11)$$

Bamigbola, Ali and Nwaeze, 2010

$$\beta_k^{BAN} = \frac{-{g_k}^T y_{k-1}}{{g_{k-1}}^T y_{k-1}} \qquad (1.12)$$

where g_k and g_{k-1} are gradients of $f(x)$ at the points x_k and x_{k-1}, respectively, while $\|\cdot\|$ is a norm and d_{k-1} is a direction for the previous iteration.

Conjugate gradient methods differ in their way of defining the scalars β_k. Over the years, several choices of β_k, which give rise to different conjugate gradient methods, have been proposed. Despite the famous formulas (1.6 – 1.12), other parameters β_k for nonlinear CG methods have been proposed in literature (see for example Dai and Yuan (2001, 2003), Zhang et al. (2006 a, 2006b), Adeleke and Osinuga (2018), Osinuga and Olofin (2017), Kaelo (2016), Narayanan et al. (2017), Dai (2003) Ibrahim and Rohanin (2016), Liu and Wu (2014), Mehdi and Masoud (2001), Polyak (1969) and Hager and Zhang (2006)).

More often, the search direction satisfies the conjugacy property ${d_k}^T H d_j = 0, k \neq j$, where H is the positive definite matrix for linear CG. For nonlinear CG methods, the conjugacy condition is not satisfied since the Hessian $\nabla f(x)$ vary at different points.

In this paper, we suggest variants of hybrid BFGS-CG algorithm for nonlinear CG methods. In section 2, we present the description of the variants of the BFGS - CG method and show that the variants of BFGS - CG are descent. We present the convergence analysis for the proposed methods in section 3. Section 4 comprises numerical results of these variants against some other existing CG methods, the benchmark problems and lastly the conclusion in section 5.

2. Overview of Recent Hybrid CG Methods

CG methods such as β_k^{FR}, β_k^{CD} and β_k^{DY} are known for their strong global convergence properties but are poor in computational performance. However, methods such as β_k^{PRP}, β_k^{HS} and β_k^{LS} perform better than β_k^{FR}, β_k^{CD} and β_k^{DY} numerically. They are also known to be among the most efficient methods because of their restart capabilities if it encounters bad direction. Researchers have made spirited efforts for decades and still working to improve on the existing methods. One of such efforts is the establishment of hybrid CG methods. These hybrid CG methods combined the strengths of one or more CG methods, thereby resulting in strong convergence properties as well as good numerical performance (Babaie-Kafaki (2012, 2013), Ibrahim et al. 2014a, 2014b, 2014c, Mamat et al. (2009), Ibrahim (2014)). Although, the first work on hybrid CG methods can be traced to Touati-Ahmed-Storey (1990), this was later followed by the works of Hu and Storey (1991), Gilbert and Nocedal (1992) and

Dai and Yuan (2001) respectively. A brief overview of only recent hybrid methods is given in this section.

Kaelo (2014) proposed a hybrid method based on hybrid methods of Gilbert and Nocedal (1992) and those of Dai and Yuan (2001). The method is given as

$$\beta^k = \max\{\min\{-c\beta_k^{PRP}, \beta_k^{FR}\}, \min\{\beta_k^{PRP}, \beta_k^{FR}\}\}$$
(2.1)

where $c = \dfrac{1-\gamma}{1+\gamma} > 0$ and $\gamma \in [\frac{1}{2}, 1]$. Another recently proposed can be found in Xu and Kong (2016) as well as Djorjevic (2016, 2017) that employed the ideas of linear and convex combinations of classical CG algorithms respectively. Many other hybrids have been proposed by combining the classical algorithms and quasi-Newton methods to obtain the parameter β_k (see Ibrahim et al. (2014a, 2014b), Ibrahim (2014) and Wan Osman et al. (2017) and references therein). In Wan Osman et al. (2017), for example, a new method was proposed by combining the search direction between conjugate gradient method and quasi-Newton method using the Davidon-Fletcher- Powell (DFP) update formula as an approximation of Hessian. One other example of a hybrid that uses the attractive features of $\beta_k{}^{FR}$ and $\beta_k{}^{PRP}$ is that of Narayanan et al. (2017). The ideas used by these authors were based on the approaches of Mo, Gu and Wei (2007) and Babaie-Kafaki (2012, 2013). More recently, Adeleke and Osinuga (2018) suggested a five-term hybrid method based on the ideas of Wei et al.(2006) and Jiang et al. (2015) with β_k in this case computed as

$$\beta_k = \frac{\|g_k\|^2 - \max\left\{0, \dfrac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}\right\}}{\max\{\|g_k\|^2, d_{k-1}^T(g_k - g_{k-1}), -d_{k-1}^T g_{k-1}\}}$$

(2.2)

3. Description of Method

In this section, we introduce our methods; let us simply recall the well-known BFGS quasi - Newton method. The direction d_k in BFGS method is given by

$$d_k = -H_k g_k \qquad (3.1)$$

where H_k is obtained by the BFGS formula

$$H_{k+1} = H_k - \frac{H_k s_k s_k{}^T H_k}{s_k{}^T H_k s_k} + \frac{y_k y_k{}^T}{s_k{}^T y_k} \qquad (3.2)$$

with $s_k = x_k - x_{k-1}$ and $y_k = g_k - g_{k-1}$. The approximation that the Hessian must fulfill is

$$H_{k+1} s_k = y_k \qquad (3.3)$$

This condition is required to hold for the updated matrix H_{k+1}, which is the secant equation. This is only possible to fulfill the secant equation, if

$$s_k{}^T y_k > 0 \qquad (3.4)$$

which is known as the curvature condition.

By extension, Ibrahim et al. (2014a) proposed a new hybrid CG direction as

$$d_k = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k d_{k-1}) & k \geq 1 \end{cases}$$
(3.5)

where $\eta > 0$, $\beta_k = \frac{g_k{}^T g_{k-1}}{d_k{}^T g_{k-1}}$ and H_k is the approximate Hessian. The update formula for the BFGS is with $s_k = x_k - x_{k-1}$ and $y_k = g_k - g_{k-1}$.

In line with this progress, our aim is to establish variants of the hybrid BFGS - CG method by replacing the CG update parameter in (3.5) successively with the formulas as shown in (1.6 - 1.12). The direction generated by the proposed variants is always a descent direction of the objective function. These variants are described as follows:

$$d_k{}^{FR} = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k{}^{FR} d_{k-1}) & k \geq 1 \end{cases}$$
-(3.6)

$$d_k{}^{PRP} = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k{}^{PRP} d_{k-1}) & k \geq 1 \end{cases}$$
-(3.7)

$$d_k{}^{HS} = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k{}^{HS} d_{k-1}) & k \geq 1 \end{cases}$$
-(3.8)

$$d_k{}^{BAN} = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k{}^{BAN} d_{k-1}) & k \geq 1 \end{cases}$$
-(3.9)

Vol. 1, No. 2, Nov. 2018
ISSN: 2616-1303 | **Web:** *www.cjpas.fulokoja.edu.ng*

$$d_k^{DY} = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k^{DY} d_{k-1}) & k \geq 1 \end{cases}$$
-(3.10)

$$d_k^{LS} = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k^{LS} d_{k-1}) & k \geq 1 \end{cases}$$
-(3.11)

$$d_k^{CD} = \begin{cases} -H_k g_k, & k = 0 \\ -H_k g_k + \eta(-g_k + \beta_k^{CD} d_{k-1}) & k \geq 1 \end{cases}$$
-(3.12)

We prove that the variants of BFGS - CG proposed with Armijo - type line search is globally convergent.

Algorithm 3.1. Based on the BFGS - CG method, we propose our algorithm as follows:

Require: A starting point x_0, parameters $0 < \epsilon < 1, 0 < \delta < \frac{1}{2}, \delta < \sigma < 1, s > 0, \beta \in (0,1), \rho \in (0,1)$

Step 1: Set k = 0 and compute $d_0 = -g_0$.
Step 2: If $\| g_k \| < \epsilon$, STOP; else go to Step 3.
Step 3: Compute step size α_k, such that

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \quad (3.13)$$

and

$$g(x_k + \alpha_k d_k)^T d_k \geq \rho g_k^T d_k \quad (3.14)$$

Step 4: Compute $x_{k+1} = x_k + s_k$
Step 5: $g_{k+1}, y_k = g_{k+1} - g_k$ and go to Step 6.
Step 6: Compute the search direction by (3.5) using $\beta_k^{FR}, \beta_k^{PRP}, \beta_k^{HS}, \beta_k^{BAN}, \beta_k^{DY}, \beta_k^{LS}, \beta_k^{CD}$ respectively.
Step 7: Let k = k + 1; and go to Step 2.

Theorem 3.2: Let the sequences $\{x_k\}$ and $\{d_k\}$ be generated by the Algorithm 3.1 for β_k^{DY}. Then, $g_k^T d_k < 0$ holds true. Therefore, for k > 1, $g_k^T d_k = -(\emptyset + \eta(1 + \tau_k)) \| g_k \|^2$.

Proof:

$$d_k = -H_k g_k + \eta(-g_k + \beta_k d_{k-1}) \quad (3.15)$$

$$g_k^T d_k = -H_k g_k^T g_k + \eta g_k^T(-g_k + \beta_k d_{k-1}) \quad (3.16)$$

$$g_k^T d_k = -\emptyset \| d_k \|^2 - \eta \| d_k \|^2 + \eta \beta_k g_k^T d_{k-1} \quad (3.17)$$

$$g_k^T d_k = -(\emptyset + \eta) \| g_k \|^2 + \eta \beta_k g_k^T d_{k-1} \quad (3.18)$$

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

$$g_k^T d_k = -(\emptyset + \eta) \| g_k \|^2 + \eta\left(\frac{g_k^T g_k}{d_k^T(g_k - g_{k-1})}\right) g_k^T d_{k-1}$$
-(3.19)

$$g_k^T d_k = -(\emptyset + \eta) \| g_k \|^2 + \eta\left(\frac{\|g_k\|^2 g_k^T d_{k-1}}{d_k^T(g_k - g_{k-1})}\right)$$
-(3.20)

$$g_k^T d_k = -(\emptyset + \eta) \| g_k \|^2 + \eta \tau_k \| g_k \|^2 \quad (3.21)$$

$$g_k^T d_k = -(\emptyset + \eta(1 + \tau_k)) \| g_k \|^2 \quad (3.22)$$

$$g_k^T d_k = -c_1 \| g_k \|^2 \quad (3.23)$$

where $c_1 = -(\emptyset + \eta(1 + \tau_k))$ which is bound away from zero. Hence, $g_k^T d_k = -c_1 \| g_k \|^2$ holds. The proof is completed.

The descent conditions for other variants can be proved analogously.

4. Global Convergence of Proposed Variants

We will like to propose some basic assumptions based on the objective function in order to discuss the global convergence of these variants.

Assumption 4.1

A: f is bounded below on the level set $S = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ where x_0 is the starting point.

B: In some neighborhood N of S, the function f is continuously differentiable and its gradient, $g(x) = \nabla f(x)$, is Lipschitz continous, i.e. there exist a constant L > 0 such that

$$\| g(x) - g(y) \| \leq L \| x - y \| \quad (4.1)$$

for all $x, y \in$ N .

Lemma 4.2 [5]. From Assumption 4.1, positive constants ϖ_1 and ϖ_2 exist, such that for any x_k and any d_k with $g_k^T d_k < 0$ step size α_k produced by algorithm 3.1 will satisfy either

$$f(x_k + \alpha_k d_k) - f_k \leq -\varpi_1 \frac{g_k^T d_k}{\|d_k\|^2} \quad (4.2)$$

or

$$f(x_k + \alpha_k d_k) - f_k \leq -\varpi_1 g_k^T d_k \quad (4.3)$$

Theorem 4.3 [5]. Let H_k be generated by the BFGS formula (3.2) where H_k is symmetric and positive definite, and $y_k s_k > 0$ for all k. Furthermore, assume that $\{s_k\}$ and $\{y_k\}$ are such that

$$\frac{\|(y_k - G_*)s_k\|}{\|s_k\|} \leq \epsilon_k \quad (4.4)$$

Vol. 1, No. 2, Nov. 2018
ISSN: 2616-1303 | **Web:***www.cjpas.fulokoja.edu.ng*

for some symmetric and positive definite matrix $G(x^*)$ and for some sequence $\{\epsilon_k\}$ with the property $\sum_{k=1}^{\infty} \epsilon_k < \infty$. Then,

$$\lim_{k\to\infty} \frac{\|(H_k - G_*)d_k\|}{\|d_k\|} = 0 \quad (4.5)$$

and the sequence $\{\| H_k \|\}$, $\{\| H_k^{-1} \|\}$ are bound. Furthermore based on algorithms 3.1, the search directions d_k given by equation (3.5) satisfy the descent and sufficient descent condition (4.6) and (4.7) respectively.

$$g_k^T d_k < 0 \quad (4.6)$$

for all $k \geq 0$. If there exists a constant $c_1 > 0$ such that

$$g_k^T d_k \leq -c_1 \| g_k \|^2 \quad (4.7)$$

for all $k \geq 0$.

Theorem 4.4 (Global Convergence). Suppose that Assumption 4.1 and Theorem 4.3 hold. Then,

$$\lim_{k\to\infty} \| g_k \|^2 = 0 \quad (4.8)$$

Proof. By combining descent property (4.6) and Lemma 4.2, we have

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty \quad (4.9)$$

Hence, from Theorem 3.2, we can define that $\| d_k \| \leq -c \| g_k \|$. Then, (4.9) will be simplified as

$$\sum_{k=0}^{\infty} \| g_k \|^2 < \infty \quad (4.10)$$

Therefore, the proof is completed.

4. Numerical Experiments

4.1 Benchmark Problems

This section gives the presentation of the simulation results on the test problems for our proposed variants of the BFGS - CG method against the conventional BFGS – CG method. We consider some test problems from CUTEr and (Andrei, 2008) using inexact line search Conditions (1.3) and (1.4) for all methods in this paper for easy comparison where $\delta = 0.0001$ and $\rho = 0.01$. These problems are listed below and shown on table 1.

4.2 List of Problems

1. Extended Matyas Function: $f(x) = 0.26\,(x_1^2 + x_2^2) - 0.48\,x_1x_2$.
2. Extended Booth Function: $f(x) = (x_1 + 2x_2)^2 + (2x_1 + x_2 - 5)^5$.
3. The Six Hump: $f(x) = \left(4 - 2.1x_1^2 + \frac{x_1^4}{3}\right)x_1^2 + x_1x_2 + (4x_2^2 - 4)x_2^2$.
4. Extended Wood Function: $f(x) = \sum_{i=1}^{\frac{n}{4}} 100\,(x_{4i-2}^2 - x_{4i-2})^2 + (x_{4i-2} - 1)^2 + 90(x_{4i-1}^2 - x_{4i})^2 + (1 - x_{4i-1})^2 + 10.1\{\,(x_{4i-2} - 1)^2 - (x_{4i} - 1)^2\} + 19.8\,(x_{4i-2} - 1)(x_{4i} - 1)$.
5. Extended Freudenstein & Roth function: $f(x) = \sum_{i=1}^{\frac{n}{2}}(-13 + x_{2i} + ((5 - x_{2i})x_{2i} - 2)x_{2i})^2 + (-29 + x_{2i-1} + ((x_{2i-1} + 1)x_{2i} - 14)x_{2i})^2$.
6. Quadratic Function: $f(x) = -3803.84 - 138.08x_1 - 232.92x_2 + 128.08x_1^2 + 203.64x_2^2 + 182.25x_1x_2$.
7. Extended Maratos function: $f(x) = \sum_{i=1}^{\frac{n}{2}} x_{2i-1} + c(x_{2i-1} + x_{2i}^2 - 1)^2, \quad c = 100$
8. Raydan 1 function: $f(x) = \sum_{i=1}^{n} \frac{1}{10}(\exp(x_i) - x_i)$.
9. Quadratic QF1 function: $f(x) = \frac{1}{2}\sum_{i=1}^{n} i\, x_2^{\,i} - x_n$.
10. Extended White and Holst: $f(x) = \sum_{i=1}^{\frac{n}{2}} c(x_{2i} - x_{2i-1}^3)^2 + (1 - x_{2i-1})^2$.
11. Diagonal 4 function: $f(x) = \sum_{i=1}^{\frac{n}{2}} \frac{1}{2}(x_{2i-1}^2 + c x_{2i}^2)$.
12. Extended Rosenbrook function: $f(x) = \sum_{i=1}^{\frac{n}{2}} c(x_{2i} - x_{2i-1}^2) + (1 - x_{2i-1})^2, \quad c = 100$.

TABLE I: LIST OF PROBLEM FUNCTIONS

No	Test Problems	Dim	Initial points
1	Extended Matyas	2	[1, 1], [5, 5], [10, 10], [50, 50]
2	Extended Booth	2	[10,10], [20,20], [50,50], [100, 100]
3	The six-hump	4	[1,1], [2,2], [5,5], [10,10], [-10,-10], [8,8], [-8,-8]
4	Extended Wood	4	[-1,-1,-1,-1], [-3, -1, -3, -1], [-2, -1, -2, -1], [-4, -1, -4, -1]
5	Ext. Freud. & Roth	2,4	[2,2], [-2,-2], [5,5], [-5,-5], [8,8], [-8,-8], [10,10], [-10,-10]

Vol. 1, No. 2, Nov. 2018
ISSN: 2616-1303 | **Web:** *www.cjpas.fulokoja.edu.ng*

6	Quadratic	2,4,10	[5,5] , [20,20], [23,23], [50,50]
7	Extended Maratos	2,4,10	[0,0], [0.5,5], [10, 0.5], [70,70]
8	Raydan 1	2, 4, 10, 100	[1,1,1], [3,3,3], [5,5,5], [-10, -10, -10]
9	Quadratic QF1	2, 4, 10, 100	[5,5], [7,7], [10,10], [100,100]
10	White & Holst	2, 4, 10, 100, 500, 1000	[-3,-3], [6,6], [10,10], [3,3]
11	Diagonal 4	2, 4, 10, 100, 500, 1000	[2,2], [5,5], [10,10], [15,15]
12	Ext. Rosenbrock	2, 4, 10, 100, 500, 1000	[13,13], [16,16], [20,20], [30,30]

4.2 Parameter Settings

The parameters such as number of iterations (it) and CPU time (t) were considered to evaluate the computational capability of the proposed variants of BFGS - CG as compared with the conventional BFGS – CG method. For each test problem, the stopping criteria used are $\| g_k \| \leq 10^{-6}$ and the number of iterations exceeds a limit of 10,000. We implemented the methods using MATLAB R2013 with CPU 1.30 GHz and 3.00GB RAM, on SAMSUNG PC notebook.

4.3 DISCUSSION OF RESULTS

The performance profile of Dolan and More (2002) was used to compare the numerical strength of the proposed variants against the conventional BFGS -CG method based on number of iterations and CPU time. We plot fraction $P(\tau)$ of the test problems for which the method is within a factor τ of the best time for each method. The left hand side of the figures gives the % of how fast is a particular method in solving the test problems. The right hand side of the figures gives the % of test problems that are successfully solved by each method. The solver with large probability is regarded as the best solver for the test problems. Figures 1 – II, show that BFGS – BAN is the fastest solver on approximately 47% of the test problems for iterations and 19% of CPU time. However, it competes well with the conventional BFGS – CG method by solving approximately 98% of the test problems compared with BFGS-PRP (97%), BFGS-FR (96%), BFGS-HS (86%), BFGS-LS (85%), BFGS-DY (84%) and BFGS-CD (79%) methods.

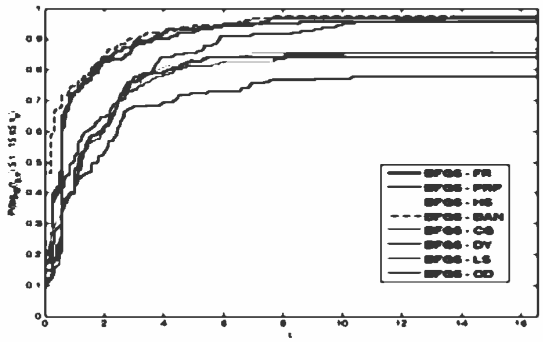

Figure I: Performance Profile based on numbers of iteration for BFGS – CG versus variants.

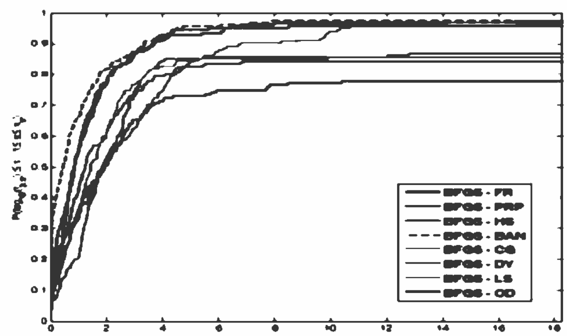

Figure II: Performance Profile based on CPU time for BFGS – CG versus variants.

5. CONCLUSION

In this paper, variants of the BFGS - CG method was proposed for solving unconstrained optimization problems. The proposed methods generate descent directions using Armijo line search condition. Under the line search conditions (1.3) and (1.4), we established the global convergence of the proposed method. The simulation results of the proposed variants are shown to be efficient for handling unconstrained optimization problems. We employed one of the best methods of comparison (Performance Profiles by (Dolan and More, 2002) to show the effectiveness of our proposed variants. Among the variants of BFGS – CG considered in this research, BFGS – BAN competes favourably well with the conventional BFGS – CG method in terms of function evaluation for number of iterations and CPU time. Moreover, BFGS - BAN is the fastest solver among the variants and the conventional BFGS – CG method.

ACKNOWLEDGEMENTS

The first author gratefully acknowledges support from Federal University of Agriculture, Abeokuta for the TETFUND Conference Intervention Fund to present this work during the 2017 Southern Africa Mathematical **Sciences** Association (SAMSA) conference organized by the Department of Mathematics, University of Dar-es-Salaam, Dar-es-Salaam, Tanzania. The authors are grateful to anonymous referees especially for their useful suggestions that greatly improved this work.

REFERENCES

Adeleke, O. J. and Osinuga, I. A. (2018), A five-term hybrid conjugate gradient method with global convergence and descent properties for unconstrained optimization problems, *Asian Journal of Scientific Research*, *11* (2), 185-194, DOI: 10.3923/ajsr.2018.

Andrei N. (2008), Unconstrained optimization test functions collection, *Advanced Modelling and Optimization*,*10*(1), 147 – 161.

Babaie – Kafaki, S., (2012). A quadratic hybridization of Polak-Ribiere-Polyak and Fletcher-Reeves conjugate gradient methods, *Journal of Optimization Theory and Application*, *154* (3), 916 - 932.

Babaie – Kafaki, S., (2013). A note on global convergence of the quadratic hybridization of Polak-Ribiere-Polyak and Fletcher-Reeves conjugate gradient methods, *Journal of Optimization Theory and Application*, *157* (1), 297 - 298.

Bamigbola, O. M., Ali, M. M. and Nwaeze, E. (2010), An efficient and convergent method for unconstrained optimization, In *Proceedings of International Congress of Mathematicians*, Hyderabad, Indian.

Dai Y. H. and Yuan Y. (1999), A nonlinear conjugate gradient with a strong convergence property, *SIAM Journal of Optimization, 10* (1), 177 – 182.

Dai Y. H. and Yuan Y. (2001), An efficient hybrid conjugate gradient method for unconstrained optimization, *Annals of Operations Research*, *103*, 33- 47.

Dai, Y. – H. (2003), A family of hybrid conjugate gradient methods for unconstrained optimization, *Mathematics of Computation*, *72* (243), 1317 – 1328.

Dolan E., More J. J. (2002), Benchmarking optimization software with performance profile, *Mathematical Programming*", *91*, 201 – 213.

Djordjevic S. S. (2016), New hybrid conjugate gradient method as a convex combination of FR and PRP Methods, *Filomat*, *30* (11), 3083 – 3100.

Djordjevic S. S. (2017), New hybrid conjugate gradient method as a convex combination of LS and CD Methods, *Filomat*, *31* (6), 1813 – 1825.

Fletcher, R. and Reeves, C. (1964), *Function minimization by conjugate gradients*, *Computer Journal*, *7*, 149-154.

Fletcher, R. (1987, *Practical Methods of Optimization vol. 1: Unconstrained Optimization*, New York :John Wiley & Sons.

Gilbert E. G. and Nocedal J. (1992), Global convergence properties of conjugate gradient methods for Optimization, *SIAM Journal of Optimization*, *2,* 21 -42.

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

Hestenes, M.R.andStiefel, E.L. (1952),*Methodofconjugategradientsforsolvinglinearsystems*, *Journal of Research National Bureau Standards*, *49*, 409-436.

Hu, Y. F. and Storey, C. (1991), Global convergence result of conjugate gradient methods, Journal of Optimization Theory and Applications, *71,*399-405.

Ibrahim, A. and Rohanin, A. (2016), Global Convergence Analysis of a Nonlinear Conjugate Gradient Method for Unconstrained Optimization Problems. *Indian Journal of Science,9* (41). Doi.10.17485/ijst/2016/v9i41/90175.

Ibrahim, M. A. H., Mustafa M., Leong W. J. (2014a), The hybrid BFGS - CG method in solving unconstrained optimization problem,. *Abstract and Applied Analysis*, *Vol. 2014* Article ID 507102, 6 pp.

Ibrahim, M. A. H., Mustafa M., Leong W. J., Azfi Z. S. (2014b), The Algorithms of Broyden – CG for unconstrained optimization problems, *International Journal of MathematicalAnalysis*, *8* (52), 2591 – 2600, http://dx.doi.org/10.12988/ijma.2014.49272

Ibrahim, M. A. H. (2014), The CG – BFGS method for unconstrained optimization problems. *AIP Conference Proceedings*, *1605*, 167 – 172.

Ibrahim S. M., Mustafa M., Kamil U. K., Abdelrhaman A. (2014c), A new hybrid WYL – AMRI conjugate gradient method with sufficient descent condition for unconstrained optimization, *International Journal of Technical Research and Applications*, e-ISSN: 2320 - 9163, 14 – 17.

Jiang, X., Han, L. and Jian, J. (2015), A globally convergent mixed conjugate gradient method with Wolfe line search, *Mathematics Numerica Sinica*, *34*, 103 – 112.

Kaelo, P. (2014), Another hybrid conjugate gradient method for unconstrained optimization problems, *Journal of Nonlinear Analysis and Optimization*, *5* (2), 127 – 137.

Kaelo, P. (2016), A hybrid conjugate gradient method for unconstrained optimization problems, *Pacific Journal of Optimization*, *12* (4), 847 – 859.

Liu, Y. and Storey, C. (1991), *Efficient generalized conjugate gradient algorithms,* Part 1: Theory, *Journal of Optimization Theory Application*, *69*, 129-137.

Liu, J. and Wu, X. (2014), New Three - Term Conjugate Gradient Method for Solving Unconstrained Optimization Problems. *Science Asia, 40*, 295 - 300, DOI: 10.2306/science asia1513-1874.2014.40.295

Mamat M., Ismail M., Leong W. J., Yosza D. (2009), Hybrid Broyden method for unconstrained optimization, *International Journal of Numerical Methods and Applications*, 121 – 130.

Mehdi, D. and Masoud, H. (2001), Convergence of the descent Dai - Yuan conjugate gradient method for unconstrained optimization. *Journal of Vibration and Control,18* (9), 1249 – 1253.

Mo, J., Gu, N. and Wei, Z. (2007), Hybrid conjugate gradient methods for unconstrained optimization, *Optimization Methods and Software*, *22* (2), 297 – 307.

Natarayann, S., Kaelo, P. and Thuto, M. V. (2017), A modified quadratic hybridization of Polak-Ribiere- Polyak and Fletcher- Reeves conjugate gradient method for unconstrained optimization problems. *An International Journal of Optimization and Control: Theory and Applications,7* (2), 177 – 185.

Osinuga, I. A. and Olofin, I. O. (2017), Extended hybrid conjugate gradient method for unconstrained optimization. *Journal of Computer Science and its Application, To appear*.Polak, E. and Ribieré, G. (1969), Note sur la convergence de directions conjugeés, *Rev. Francaise Informat Recherche Operationelle*, 3e Année, *16*, 35-43.

Polyak, B. T. (1969), *The conjugate gradient method in extreme problems*, *USSR Comp. Mathematics and Mathematical Physics,9*, 94-112.

Touati - Ahmed D., Storey C. (1990), Efficient hybrid conjugate gradient techniques, *Journal of Optimization Theory Appl*ication, *64* (2), 379 – 397.

Wan Osman W. F. H., Ibrahim, M. A. H., Mamat, M. (2017), Hybrid DFP-CG method for solving Unconstrained optimization problem, *Journal of Physics: Conf. Series,890*, DOI:10.1088/1742-6596/890/1/012033.

Wei, Z., Yao, S. and Liu, L. (2006), The convergence properties of some new conjugate gradient methods.

Vol. 1, No. 2, Nov. 2018
ISSN: 2616-1303 | **Web:***www.cjpas.fulokoja.edu.ng*

Applied Mathematics and Computation, 183, 1341-1350.

Xu X. and Kong F. (2016), New hybrid conjugate gradient methods with the generalized Wolfe line search. *Springerplus*, *5*, 881.

Zhang, L., Zhou, W. J. and Li, D. H. (2006a), Global convergence of a modified Fletcher Reeves conjugate gradient method with Armijo type line search. *Numerical Mathematics, 104*, 561 - 572.

Zhang L., Zhou W. J. and Li D. H. (2006b), A descent modified Polak – Ribiere – Polyak conjugate method and its global convergence", *IMA Journal of Numerical Anal*ysis, *26*, 629 – 649.