Indonesian Journal of Electrical Engineering and Computer Science
Vol. 19, No. 3, September 2020, pp. 1564~1573
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v19.i3.pp1564-1573

New multi-step three-term conjugate gradient algorithms with inexact line searches

Abbas Y. Al-Bayati[1], Muna M. M. Ali[2]
[1]Department of Mathematics, College of Basic Education, Telafer University, Iraq
[2]Department of Mathematics, College of Computers Sciences and Mathematics, Mosul University, Iraq

Article Info

Article history:

Received Nov 19, 2019
Revised Feb 8, 2020
Accepted Mar 29, 2020

Keywords:

Global convergence property
Inexact line searches
Multi-step conjugate gradient
Sufficient descent property
Three-term conjugate gradient
Scaled conjugate gradient

ABSTRACT

This work suggests several multi-step three-term Conjugate Gradient (CG)-algorithms that satisfies their sufficient descent property and conjugacy conditions. First, we considered (39) well-known three-term CG-method, and we have, therefore, suggested two new classes of this type of algorithms based on Hestenes and Stiefel (HS) and Polak-Ribière (PR) formulas with four different versions. Both descent and conjugacy conditions for all the proposed algorithms are satisfied, at each iteration by using the strong Wolfe line search condition and it's accelerated version. These new suggested algorithms are some sort of modifications to the original HS and PR methods. These CG-algorithms are considered as a sort of the memoryless BFGS update. All of our new suggested methods are proved to be globally convergent and numerically, more efficient than similar methods in the same area based on our selected set of used numerical problems.

Corresponding Author:

Abbas Y. Al-Bayati,
Department of Mathematics,
College of Basic Education,
Telafer University,
Tall ‘Afar, Mosul, Nineveh, Iraq.
Email: profabbasalbayati@yahoo.com

1. INTRODUCTION

This paper considers the calculation of a local minimizer x*, say, for the problem:

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

Is a nonlinear function and its gradient vectoris g_k is available. The Hessian matrix is not available. At the current iterative point x_k the CG-method has the following form; based on the quadratic form:

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

$$d_{k+1} = \begin{cases} -g_{k+1}, & k = 0 \\ -g_{k+1} + \beta_k d_k, & k \geq 1 \end{cases} \tag{2b}$$

Here, α_k; step-length, d_k; search direction, β_k; parameter. Standard algorithms for solving this problem include CG-algorithms, with very low memory requirements, which are iterative algorithms and generate a sequence of approximations of the f(x). For more details, see Hassan [1].

The first Three-Term CG-method was proposed by Beale [2] as:

$$d_{k+1}^{Beale} = -g_{k+1} + \beta_k d_k + \gamma_k d_t \ , \ \beta_k = \beta_k^{HS}, \beta_k^{FR}, \beta_k^{PR}\ etc... \tag{3a}$$

$$\gamma_k = \begin{cases} 0, & k = t+1 \\ g_{k+1}^T y_t / d_t^T y_t, & k > t+1 \end{cases} \tag{3b}$$

And d_t, is a restart direction and $y_k = -g_{k+1} - g_k$. Nazareth [3] proposed another three-term recurrence formula:

$$d_{k+1} = -y_k + (y_k^T y_k / y_k^T d_k) d_k + (y_{k-1}^T y_k / y_{k-1}^T d_{k-1}) d_{k-1}\ ;\ d_{-1} = 0\ and\ d_0 = 0 \tag{4}$$

Also, two different three-term CG-algorithms was considered by Zhang [4, 5], that is,

$$d_{k+1}^{ZPR} = -g_{k+1} + \beta_k^{PR} d_k - \theta_k^{(1)} y_k\ ;\ d_{k+1}^{ZHS} = -g_{k+1} + \beta_k^{HS} s_k - \theta_k^{(2)} y_k \tag{5}$$

$$\theta_k^{(1)} = g_{k+1}^T d_k / g_k^T g_k\ ;\ \theta_k^{(2)} = g_{k+1}^T s_k / s_k^T y_k \tag{6}$$

Zhang [6] introduced a three-term recurrence formula based on Dai and Liao [7] CG-method as follows:

$$d_{k+1}^{ZDL} = -g_{k+1} + [g_{k+1}^T (y_k - ts_k) / d_k^T y_k] d_k - g_{k+1}^T d_k / d_k^T y_k (y_k - ts_k) \tag{7a}$$

where $d_0 = -g_0,\ t \geq 0$. The sufficient descent condition also holds independent of the used line search procedure, i.e. for this method:

$$g_k^T d_k = -\|g_k\|^2 \text{ for all k.} \tag{7b}$$

A specialization of (7) was developed by Nazareth [8] where the search direction is computed as:

$$d_{k+1}^{BS} = -g_{k+1} + \beta_k^{DL+} s_k - g_{k+1}^T s_k / g_k^T g_k (y_k - ts_k) k \tag{8a}$$

$$\beta_k^{DL+} = \max[y_k^T g_{k+1} / y_k^T s_k, 0] - ts_k^T g_{k+1} / y_k^T s_k\ ; t = 2\|y_k\|^2 / y_k^T s_k \tag{8b}$$

Furthermore, it is easy to see that (8) satisfies the sufficient descent condition independent of the line search used. Moreover, a different three-term CG-method was improved by Al-Bayati and Hassan [9], and their search direction, with inexact line search (ILS), is as follows:

$$d_{k+1}^{BH} = -g_{k+1} + \beta_k^{PR}\ d_k + \gamma_k\ ; \gamma_k = -(d_k^T g_{k+1}) / g_k^T g_k \tag{9}$$

Recently, a three-term CG-method was introduced by Al-Bayati and Al-Khayat [10] and their search direction is as follows:

$$d_{k+1}^{BK} = -g_{k+1} + \beta_k^{HS}\ d_k + \phi_k (y_k - t_1 s_k); \phi_k = s_k^T g_{k+1} / y_k^T s_k, \quad t_1 = 1 + y_k^T y_k / y_k^T s_k \tag{10}$$

Conjugate Gradient algorithms could be regarded as a sort of the Memoryless QN-updates, especially, for the BFGS update. This type of method was suggested for the first time by Perry [11] noted that the scalar β_k has been chosen so that the search directions d_k and d_{k+1} are conjugate using ELS. Perry relaxed this requirement where β_k is defined by HS formula in an equivalent form, but assuming ILS, thus he obtained,

$$d_{k+1}^{P1} = -(I - d_k y_k^T / d_k^T y_k) g_{k+1} \tag{11a}$$

but this matrix is not of full rank; Perry modified it further as:

$$d_{k+1}^{P2} = -(I - s_k y_k^T / s_k^T y_k + s_k s_k^T / s_k^T y_k) g_{k+1} \tag{11b}$$

Then Shanno [12] addressed that (11) does not satisfy the actual QN-condition, so he modified it to obtain:

$$Q_{k+1}^{SH} = I - (s_k s_k^T + y_k s_k^T / y_k^T s_k) + (1 + y_k^T y_k / y_k^T s_k)(s_k s_k^T / y_k^T s_k) \tag{12}$$

This new form of the projection matrix Q_{k+1} has a special relationship with the BFGS update formula; defined by Dennis and More [13].

$$H_{k+1}^{BFGS} = H_k - \left[\frac{s_k y_k^T H_k + H_k y_k s_k^T}{y_k^T s_k}\right] + \left[1 + \frac{y_k^T H_k y_k}{y_k^T s_k}\right]\frac{s_k s_k^T}{y_k^T s_k} \tag{13}$$

It is easily seen that (12) is equivalent to (13) when H_k replaced by I, i.e. if $H_k = I$, the identity matrix. For more details, see Hassan [14]. The Memoryless BFGS method is defined by:

$$d_{k+1} = -Q_{k+1} g_{k+1} \tag{14}$$

$$d_{k+1} = -g_{k+1} + [(g_{k+1}^T y_k / s_k^T y_k) - (y_k^T y_k / s_k^T y_k)(s_k^T g_{k+1} / s_k^T y_k)] s_k + (s_k^T g_{k+1} / s_k^T y_k)(y_k - s_k)$$

Finally, Al-Bayati, et al. [15] introduced a new CG-algorithm with different parameters, namely for:

$$\lambda_k = (1+t) s_k^T g_{k+1} / y_k^T g_{k+1};\ \rho_k = (\lambda_k\ y_k^T g_{k+1} - t s_k^T g_{k+1} / s_k^T g_{k+1} \tag{15a}$$

$$d_{k+1}^{BHA} = -g_{k+1} + (\psi_k \rho_k y_k^T g_{k+1} + s_k^T y_k g_{k+1}^T s_k - \lambda_k s_k^T y_k g_{k+1}^T s_k / \|s_k\|^2\ s_k^T y_k)\ s_k \tag{15b}$$

2. ON MULTI-STEP THREE-TERM CG-ALGORITHMS

Nazareth and Nocedal [16] developed a multi-step CG-method which does not need ELS; by defining the following matrices:
$D = d_1, d_2, ..., d_n$; $G = g_1, g_2, ..., g_n$; $-G = DB$, where B is an (NxN) upper triangular matrix with, $\beta_{ii} = 1,\ for\ \ i = 1,2,3,,k$. Assuming the Hussain matrix G=I (identity matrix), to get a set of mutually orthogonal vectors $g_1, g_2, ..., g_n$. Let us define:

$$g_1^* = g_1 \tag{16a}$$

iterate for k = 2,3,4, ... with

$$g_k^* = g_k - \left[\left(g_k^T g_{k-1}^*\right) / \left(g_{k-1}^{*\ T} / g_{k-1}^*\right)\right] g_{k-1}^* \tag{16b}$$

$$g_k^* = g_k^* + c_{k-1} \tag{16c}$$

$$c_{k-1} = \begin{cases} \left(c_{k-2} + \left[\left(g_k^T g_{k-2}^*\right) / \left(g_{k-2}^{*\,T} g_{k-2}^*\right)\right] g_{k-2}^*; & k = 3,4,... \\ 0 & k = 1,2 \end{cases} \tag{16d}$$

Note that, (16), for the first iterate, is equivalent to the normal set of gradients, while it gives better approximations for the next iterates. In this paper; we have the following (four) new CG-algorithms:

$$d_{k+1}^{Ni} = -g_{k+1}^* + \beta_k^{HS} s_k + (s_k^T g_{k+1}^* / y_k^{*T} s_k)(y_k^* - t_i s_k)\ ,\ i = 1,\ 2 \tag{17a}$$

$$d_{k+1}^{N3} = -g_{k+1}^* + \beta_k^{HS} s_k + (s_k^T g_{k+1}^* / y_k^{*T} s_k)(y_k^* - t_1 s_k) + t_3 s_k \tag{17b}$$

$$d_{k+1}^{N4} = -\theta_k g_{k+1}^* + \beta_k^{SPR} s_k - (s_k^T g_{k+1}^* / y_k^{*T} s_k).P_k\ ::\quad y^* = g_{k+1}^* - g_k^*. \tag{17c}$$

Wolfe line search procedure is fully described by many researchers, see for example Nocedal [17] and Liu and Nocedal [18]. This line search scheme has been modified by Andrei [19]. The standard Wolfe line search conditions can be defined as:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \mu \alpha_k g_k^T d_k \tag{18a}$$

$$f(x_k + \alpha_k d_k) \ge f(x_k) + (1-\mu)\alpha_k g_k^T d_k \tag{18b}$$

The strong Wolfe line search conditions can be defined as:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \mu \alpha_k g_k^T d_k \tag{18c}$$

$$\left| g_{k+1}^T d_k \right| \le -\sigma\ g_k^T d_k \tag{18d}$$

The accelerating Scheme for Wolfe line search technique is as follows:

$$x_{k+1} = x_k + \lambda_k \alpha_k d_k \quad ;\ \lambda_k = -a_k / b_k \tag{18e}$$

$$a_k = \alpha_k g_k^T d_k\ ,\ \ b_k = -\alpha_k (g_k - g_z)^T d_k\ ,\ \ g_z = \nabla f(z)\ \ and\ \ z = x_k + \alpha_k d_k. \tag{18f}$$

Hence, if $b_k \neq 0$. The new estimation of the solution is computed as, $x_{k+1} = x_k + \lambda_k \alpha_k d_k$, else $x_{k+1} = x_k + \alpha_k d_k$. For this reason, using the definitions of g_k, s_k, y_k and the above acceleration scheme can present the accelerated Wolfe line search procedure.

<u>NOTE</u>: (For the rest of this paper, and for simplicity, set $g_k^* = g_k$, $y_k^* = y_k$).

First Multi-Step Three-Term CG-Method (N1)

To compute the new search direction, d_{k+1}^{N1} let us consider the QN-(BFGS update with H=I).

$$d_{k+1}^{N1} = -\theta_{k+1}^{/} g_{k+1} \tag{19}$$

$$\theta_{k+1}^{/} = I - (s_k^T y_k + y_k^T s_k / y_k^T s_k) + [1 + \|y_k\|^2 / y_k^T s_k](s_k s_k^T / y_k^T s_k) \tag{20}$$

$$d_{k+1} = -g_{k+1} + \beta_k^{HS} d_k + (s_k^T g_{k+1} / y_k^T s_k) y_k - [1 + \|y_k\|^2 / y_k^T s_k](s_k^T g_{k+1} / y_k^T s_k) s_k \tag{21}$$

Or, equivalently:

$$\phi_k = (s_k^T g_{k+1} / y_k^T s_k); \quad t_1 = (1 + \|y_k\|^2 / y_k^T s_k) \tag{22}$$

$$d_{k+1}^{N1} = -g_{k+1} + \beta_k^{HS} d_k + \varphi_k (y_k - t_1 s_k) \tag{23a}$$

$$d_{k+1}^{N1} = -g_{k+1} + \max[(g_{k+1}^T y_k / d_k^T y_k), 0] d_k + \varphi_k (y_k - t_1 s_k) \tag{23b}$$

Outline of N1-Algorithm.

St1. Given $x_0 \in R^n$, let $0 < \delta < \sigma < 1$, $t \geq 0$ and $d_0 = -g_0$. Set $k = 0$.

St2. If stopping criteria ($\|g_k\|_\infty \leq 10^{-6}$) satisfied, then stop.

St3. Compute α_k by accelerated Wolfe line search condition (18).

St4. The parameters φ_k , t_1 are computed from (22).

St5. The new search direction d_{k+1}^{N1} is computed from (23).

St6. Check whether ($\left|g_{k+1}^T g_k\right| > 0.2\|g_{k+1}\|^2$) is satisfied then set $d_{k+1}^{N1} = -g_{k+1}$

St7. Set $k = k + 1$, go to St2.

Second Multi-Step Three-Term CG-Method (N2)

To compute the new search direction d_{k+1}^{N2} let us consider Al-Bayati [20] QN-update with H=I

$$H_{k+1}^{Bayati} = H_k + (2 y_k^T H_k y_k / (s_k^T y_k)^2) s_k s_k^T - (H_k y_k s_k^T + s_k y_k^T H_k) / y_k^T s_k \tag{24}$$

Let $d_{k+1} = -Q_{k+1}^{(2)} g_{k+1}$; and H=I (25)

$$Q_{k+1}^{(2)} = I + (2 y_k^T y_k / (s_k^T y_k)^2) s_k s_k^T - (y_k s_k^T + s_k y_k^T) / y_k^T s_k \tag{26}$$

$$\phi_k = s_k^T g_{k+1} / y_k^T s_k; \quad t_2 = 2(\|y_k\|^2 / y_k^T s_k) \tag{27}$$

$$d_{k+1}^{N2} = -g_{k+1} + \beta_k^{HS} d_k + \varphi_k (y_k - t_2 s_k) \tag{28a}$$

$$d_{k+1}^{N2} = -g_{k+1} + \max[g_{k+1}^T y_k / d_k^T y_k, 0] d_k + \varphi_k (y_k - t_2 s_k) \tag{28b}$$

Outline of N2-Algorithm.
All the steps as in Algorithm (N1) except:

St4. The parameters φ_k , t_2 are computed from (27).

St5. The new search direction d_{k+1}^{N2} is computed from (28).

Third Multi-Step Three-Term CG-Method (N3)

To compute the new search direction, d_{k+1}^{N3}, let us consider Oren [21] QN-update with H=I.

$$H_{k+1}^{Oren} = H_k - (s_k y_k^T H_k + H_k y_k s_k^T) / y_k^T s_k + (\eta_k + y_k^T H_k y_k / y_k^T s_k) s_k s_k^T / y_k^T s_k \tag{29}$$

$$\eta_k = s_k^T y_k / y_k^T y_k; \quad \text{H=I} \tag{30}$$

$$Q_{k+1}^{(3)} = I + (s_k^T y_k / y_k^T y_k + y_k^T y_k / v) s_k s_k^T / s_k^T y_k - (y_k s_k^T + s_k y_k^T) / y_k^T s_k \quad (31)$$

$$d_{k+1} = -g_{k+1} + \frac{s_k^T g_{k+1}}{y_k^T y_k} y_k + \frac{y_k^T g_{k+1}}{y_k^T y_k} s_k - \left[\frac{s_k^T y_k}{y_k^T y_k} + \frac{y_k^T y_k}{y_k^T s_k} \right] \frac{s_k^T g_{k+1}}{y_k^T s_k} s_k \quad (32)$$

$$\phi_k = s_k^T g_{k+1} / v; \qquad t_3 = (s_k^T y_k / y_k^T y_k + y_k^T y_k / y_k^T s_k) \quad (33)$$

$$d_{k+1}^{N3} = -g_{k+1} + \beta_k^{HS} d_k + \varphi_k (y_k - t_3 s_k) \quad (34a)$$

$$d_{k+1}^{N3} = -g_{k+1} + \max[g_{k+1}^T y_k / d_k^T y_k, 0] d_k + \varphi_k (y_k - t_3 s_k) \quad (34b)$$

Outline of N3-Algorithm.
All the steps as in Algorithm (N1) except:

St4. The parameters φ_k , t_3 are computed from (33).

St5. The new search direction d_{k+1}^{N3} is computed from (34).

Fourth Scaled Multi-Step Three-Term CG-Method (N4)

Here, we describe our new CG-method (N4) as a scaled CG-algorithm; (for details of scaled CG-algorithms; see Al-Bayati, at el. [22] and Hassan et al. [23]). This algorithm is independent of the line search, at every step. The search direction d_{k+1}^{N4} is computed as:

$$d_{k+1}^{N4} = -\theta_{k+1} g_{k+1} + \beta_k^{SPR} s_k - \eta_k p_k \quad (35)$$

$$\theta_k = (s_k^T s_k / y_k^T s_k), \; p_k = \theta_k y_k - s_k, \; \eta_k = s_k^T g_{k+1} / \|g_k\|^2 \quad (36)$$

$$\beta_k^{SPR} = \theta_{k+1} g_{k+1}^T (g_{k+1} - g_k) / \alpha_k \theta_k g_k^T g_k, \quad (37)$$

$$d_{k+1}^{N4} = -\theta_{k+1} g_{k+1} + (\theta_{k+1} g_{k+1}^T (g_{k+1} - g_k) / \alpha_k \theta_k g_k^T g_k) s_k - s_k^T g_{k+1} / \|g_k\|^2 (\theta_k y_k - s_k) \quad (38)$$

Outline of N4-Algorithm.
All the steps as in Algorithm (N1) except:

St4. The parameters θ_k , p_k , η_k are computed from (36).

St5. β_k^{SPR} and the new search direction d_{k+1}^{N4} are computed from (37) and (38) respectively.

3.1 CONVERGENCE ANALYSIS

Descent and global convergence conditions properties of HS and PR methods can be found directly in Hestenes and Stiefel [24] and Polak and Ribière [25]. To show that our new multi-step TTCG-algorithms (N1, N2, N3, and N4) have a descent and global convergence properties using Wolfe conditions (18). Consider

Theorem-I.

Suppose that Wolfe conditions (18) are satisfied, then new search directions, N1, N2, N3, and N4 defined by (23), (28), (34), and (38) satisfy the descent property, i.e.

$$g_{k+1}^T d_{k+1} \le 0 \quad (39)$$

Proof.

I: For d_{k+1}^{N1} and from Wolfe conditions (18), we get $y_k^T s_k > 0$ then:

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 - \lfloor 1 + \|y_k\|^2 / y_k^T s_k \rfloor (s_k^T g_{k+1})^2 / y_k^T s_k \leq 0 \tag{40}$$

II: For d_{k+1}^{N2} and from Wolfe condition (18), we get $y_k^T s_k > 0$ then:

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 - \lfloor 2 \|y_k\|^2 / y_k^T s_k \rfloor (s_k^T g_{k+1})^2 / y_k^T s_k \leq 0 \tag{41}$$

III: For d_{k+1}^{N3} and from Wolfe condition (18), we get $y_k^T s_k > 0$ then:

$$y_k^T d_{k+1} = -\|g_{k+1}\|^2 - (y_k^T s_k / \|y_k\|^2 + \|y_k\|^2 / y_k^T s_k)(s_k^T g_{k+1}) \leq 0 \tag{42}$$

IV: For d_{k+1}^{N4} and (18) is held, then $y_k^T s_k \neq 0$. From (36):

$$g_{k+1}^T d_{k+1} = -s_k^T s_k / y_k^T s_k \|g_{k+1}\|^2 \tag{43}$$

$$s_k^T s_k > 0, \text{ then } s_k^T s_k / y_k^T s_k > 0, \text{ for all } k \geq 0 \tag{44}$$

Hence, d_{k+1}^{N4}, defined in (38), satisfies and

$$g_{k+1}^T d_{k+1} = -c\|g_{k+1}\|^2, \ c > 0 \tag{45}$$

Therefore, our new algorithms N1, N2, N3, and N4 are descent.

Assumption (H):
(a)- "The level set $S = \{x : x \in R^n, f(x) \leq f(x_1)\}$ is bounded, and x_1 is the starting point".
(b)- "In a neighborhood Ω of S, f is continuously differentiable and its gradient is Lipchitz continuously, namely, there exists a constant $L \geq 0$ such that "$\|g(x) - g(x_k)\| \leq L\|x - x_k\|, \forall x, \ x_k \in \Omega$".
under these assumptions on f there exists a constant $\Gamma \geq 0$ such that $\|g(x)\| \leq \Gamma, \forall x \in S$.
We know that the new search directions generated by (23), (28), (34) and (38) are always descent directions.
To ensure the global convergence property of these algorithms let us consider:
Theorem-II
Assume that (H, a) and (H, b) hold, and consider the algorithms (2), (23), (28), (34), and (38) where N1, N2, N3and N4 are descent directions and α_k computed by (18). Suppose that f is a uniformly convex function on S, i.e. there exists a constant δ such that:

$$\text{"} (\nabla f(x) - \nabla f(y))^T (x - y) \geq \delta \|x - y\|^2 \text{"} \tag{46}$$

$$\text{For all } x, \ y \in N \text{ then } \lim_{k \to \infty} \|g_k\| = 0 \tag{47}$$

Proof.
Consider CG-algorithms (2), (23), (28), (34) and (38). From "Lipschitz continuity", we get $\|y_k\| \leq L\|s_k\|$. Furthermore, from uniform convexity $y_k^T s_k \geq \mu \|s_k\|^2$. Using the Cauchy inequality, assumption (H, a) and (H, b) and the above inequalities, we get:

$$\left|\beta_k^{HS}\right| \le \frac{\left|g_{k+1}^T y_k\right|}{\left|s_k^T y_k / \alpha_k\right|} \le \frac{\alpha L \Gamma}{\Gamma \|s_k\|} \; ; \left|\varphi_k\right| \le \frac{\left|s_k^T g_{k+1}\right|}{\left|y_k^T s_k\right|} \le \frac{\Gamma}{\mu \|s_k\|} ; \left|t_1\right| \le 1 + \frac{\|y_k\|^2}{\left|y_k^T s_k\right|} \le 1 + \frac{L^2}{\mu} \tag{48}$$

$$\left|t_2\right| \le \frac{2\|y_k\|^2}{\left|y_k^T s_k\right|} \le 2\frac{L^2}{\mu} \; ; \left|t_3\right| \le \frac{\left|y_k^T s_k\right|}{\|y_k\|^2} + \frac{\|y_k\|^2}{\left|y_k^T s_k\right|} \le \frac{\mu}{L^2} + \frac{L^2}{\mu} \tag{49}$$

Therefore, putting (48) and (49) in (17a) and (17b) yields:

$$\left\|d_{k+1}^{N1}\right\| = \|g_{k+1}\| + \left|\beta_k^{HS} / \alpha_k\right| \|s_k\| + |\varphi_k| \|y_k\| + |\varphi_k| |t_1| \|s_k\| \le \Gamma + \frac{\Gamma}{M}(L + 1 + \frac{L^2}{M}) \tag{50}$$

$$\left\|d_{k+1}^{N2}\right\| = \|g_{k+1}\| + \left|\beta_k^{HS} / \alpha_k\right| \|s_k\| + |\varphi_k| \|y_k\| + |\varphi_k| |t_2| \|s_k\| \le \Gamma + \frac{\Gamma}{M}(L + \frac{2L^2}{M}) \tag{51}$$

$$\left\|d_{k+1}^{N3}\right\| = \|g_{k+1}\| + \left|\beta_k^{HS} / \alpha_k\right| \|s_k\| + |\varphi_k| \|y_k\| + |\varphi_k| |t_3| \|s_k\| \le \Gamma + \frac{\Gamma}{M}(L + \frac{M}{L^2} + \frac{2L^3}{M}) \tag{52}$$

Similarly, for d_{k+1}^{N4}, using assumption (H) and the above inequalities, we get:

$$\left|\theta_k\right| \le \frac{\left|s_k^T s_k\right|}{\left|y_k^T s_k\right|} \le \frac{\|s_k\|^2}{\eta \|s_k\|^2} \le \frac{1}{\eta} ; \left|\beta_k^{SPR}\right| \le \left|\frac{\theta_{k+1} g_{k+1}^T y_k}{\alpha_k \theta_k g_k^T g_k}\right| \le \frac{\Gamma L}{\|s_k\|} \; ; \left|\eta_k\right| \le \frac{\left|s_k^T g_{k+1}\right|}{\|g_k\|^2} \le \frac{\Gamma}{\|s_k\|} \tag{53}$$

Hence, putting (53) in (17c) yields:

$$\left\|d_{k+1}^{N4}\right\| = |\theta_k| \|g_{k+1}\| + \left|\beta_k^{SPR}\right| \|s_k\| + |y_k| \|\theta_k\| \|y_k\| + |y_k| \|s_k\| \le \frac{\Gamma}{M} + \Gamma L + \Gamma\left[\frac{1}{M} L + 1\right] \tag{54}$$

Hence, d_{k+1}^{N1}, d_{k+1}^{N2}, d_{k+1}^{N3}, and d_{k+1}^{N4} ensure the global convergence property and $\lim_{k\to\infty} \|g_k\| = 0$.

3. RESULTS AND DISCUSSION

Here, we report the performance of the new proposed CG-algorithms, namely (N1, N2, N3, N4) on a set of (39) large-scale nonlinear test problems; CUTE library (see Bongartz [26] for details of these test problems) using codes written in Fortran. We have taken (10) numerical experiments with N=1000, 4500, 10000 for each test problem. To assess the reliability of our new proposed method, we have tested it against other similar CG-algorithms with the stopping criterion $\|g_k\|_\infty \le 10^{-6}$ or when the iterations exceed 10000 or the number of function gradient evaluations (NOFG) reach 15000 without satisfying the stopping criteria. In all these tables: N = "Dimension of the problem"; NOI = "Number of iterations"; TIME = " Total time required to complete the evaluation process for each test Problem". All our numerical results are represented in Figures 1, 2, 3, 4, 5, and 6.

Figures 1 and 2 compares N1, N2, N3, and N4 against TT(BS; PR and HS) due to NOI. Figures 3 and 4 compares N1, N2, N3, and N4 against TT(BS; PR and HS) due to NOFG. Figiures 5 and 6 compares N1, N2, N3, and N4 against TT(BS; PR and HS) due to TIME.

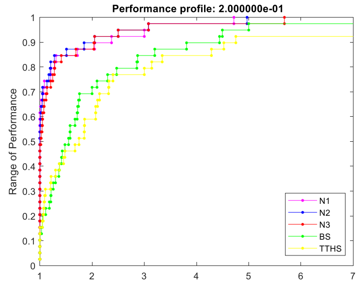

Figure 1. Performance due to NOI

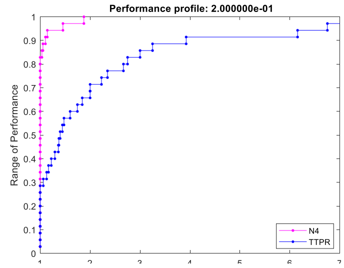

Figure 2. Performance due to NOI

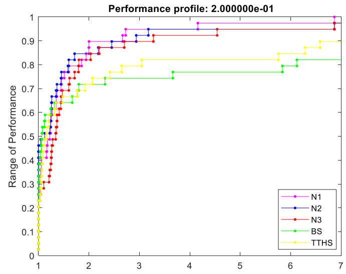

Figure 3. Performance due to NOFG

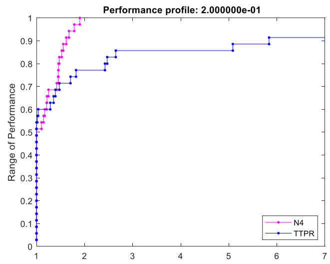

Figure 4. Performance due to NOFG

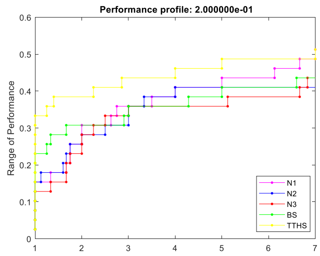

Figure 5. performance due to Time

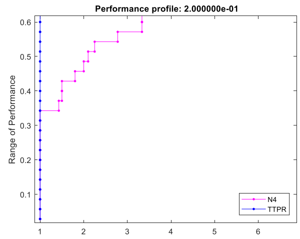

Figure 6. performance due to Time

4. CONCLUSION

In this work, we have investigated new three-term multi-step search directions defined in (17). The goal of these new algorithms is a multi-step property that combines three-term CG-techniques with memoryless QN-updates. Our theoretical implementation related to the requirement of sufficient descent property and ensures the property of global convergence. Also, we have presented four new three-term multi-step CG-algorithms, which they use Wolfe's and it's acceleration as a line search subprogram. Moreover, Our numerical results show that our new algorithms have robust numerical results as compared to other similar algorithms in the same field.

ACKNOWLEDGEMENTS

The research is supported by the College of Computer Sciences and Mathematics, University of Mosul, Republic of Iraq. The authors declare that there are no conflicts of interest regarding this work.

REFERENCES

[1] Hassan, B.A. “A new formula for conjugate parameter computation based on the quadratic model,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 13, no. 3, pp. 954-961, 2019.
[2] Beale, E.M.L., *A derivation of conjugate gradients, in numerical methods for nonlinear optimization*, New York: Lootsma, F.A., pp. 39-43, 1972.
[3] Nazareth, L., “Conjugate direction algorithm without line search,” *Journal of Optimization Theory and Applications*, vol. 23, no. 3, pp. 373–387, 1977.
[4] Zhang, L. Zhou, W. Li. D., “A descent modified Polak- Ribière-Polyak conjugate gradient method and its global convergence,” *IMA Journal of Numerical Analysis*, vol. 26, no. 4, pp. 629–640, 2006.
[5] Zhang, L. Zhou, W. Li, D., “Some descent three-term conjugate gradient methods and their global convergence,” *Optimization Methods and Software*, vol. 22, no. 4, pp. 697–711, 2007.
[6] Zhang, L. Xiao, Y. Wei. Z., “Nonlinear conjugate gradient methods with sufficient descent condition for large-scale unconstrained optimization,” *Math. Prob. Eng.*, vol. 2009, pp.1-16, 2009.
[7] Dai, Y. and Liao, L., “New conjugacy conditions and related nonlinear conjugate gradient methods, *Appl. Math Optim.*, vol. 43, no. 1, pp. 87-101, 2001.
[8] Al-Bayati, A.Y. and, Sharif, W.H., “A new three-term conjugate gradient method for unconstrained optimization,” *Canadian. J. Sci. Eng. Math.*, vol.1, no. 5, pp. 108–124, 2010.

[9] Al-Bayati, A.Y. and Hassan, B.A. "A modified three-term conjugate gradient method," *Advances in Modeling and Analysis A*, vol. 48, no. 2, pp. 1-14, 2011.
[10] Al-Bayati, A.Y. and Al-Khayat, H.N., "Different types of Three-Term CG-Methods with sufficient descent and conjugacy conditions," *Journal of Zankoy Suleimani Part-A for Pure and Applied Sciences*, vol. 16, no. 2, pp. 27-45, 2014.
[11] Perry, A., "A modified conjugate gradient algorithm," *Operations Research*, vol. 26, no. 6, pp. 1073–1078, 1978.
[12] Shanno, D., "Conjugate gradient methods with inexact searches," *Math. of Operation Research*, vol. 3, no. 3, pp. 244-256, 1978.
[13] Dennis, J. E., Jr., and Schnabel, R. E. *Numerical methods for unconstrained optimization and nonlinear equations*, New Jersey: Prentice-Hall, 1983.
[14] Hassan, B. A., "A new class of BFGS updating formula based on the new Quasi-Newton equation," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 13, no. 3, pp. 945-953, 2019.
[15] Al-Bayati A.Y; Ahmed, H.I. and Hamed, E.T., "A new hybrid algorithm for convex nonlinear unconstrained optimization," *Hindawi Journal of Applied Mathematics*, vol. 2019, pp. 1-6, 2019.
[16] Nazareth L. and Nocedal J., "Properties of CG methods with Inexact line searches," Report 780-1 System Optimization Laboratory, Department of Operation Research, Stanford University, USA, 1978.
[17] Nocedal, J., "Conjugate gradient methods and nonlinear optimization," *Linear and Nonlinear Conjugate Gradient Related Methods*, pp. 9–23, 1996.
[18] Liu, D.C. and Nocedal, J., "On the limited memory BFGS method for large scale optimization," *Mathematical Program*, vol. 45, no. 1-3, pp. 503–528, 1989.
[19] Andrei, N. "Acceleration of conjugate gradient algorithms for unconstrained optimization," *Applied Mathematical Computational*, vol. pp. 213, no. 2, pp. 361–369, 2009.
[20] Al-Bayati, A.Y., "A new family of self-scaling variable metric algorithms for unconstrained optimization," *Journal of Educ. and Sci*., *University of Mosul, Iraq*, vol. 12, pp. 25-54, 1991.
[21] Oren, S.S., "On the selection of parameters in self-scaling variable metric algorithms," *Mathematical Programming*, vol. 7, no. 1, pp. 351-367, 1974.
[22] Al-Bayati A.Y.; Hamed, E.T. and Al-Kawaz, R.Z., "New investigation for The Liu-Story Scaled conjugate gradient methods for nonlinear optimizations," *Hindawi, Journal of Mathematics*, vol. 2020, pp. 1-12, 2020.
[23] Hassan, B. A., *et al*., "A new kind of parameter conjugate gradient for unconstraint optimization," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 1, pp. 404-411, 2020.
[24] Hestenes, M.R. and Stiefel, E. "Methods of the conjugate gradient for solving a linear system," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.
[25] Polak, E. and Ribière, G. "Note Sur la Convergence de Directions Conjuguée," *Rev. Francaise Informat Recherche Operationally, 3e Année*, vol. 16, no. 3, pp. 35–43, 1969.
[26] Bongartz, K.E., Conn A.R., Gould N.I. and Toint, P.L., "CUTE, constrained and unconstrained testing environments," *ACM Trans. Math. Software*, vol. 21, no. 1, pp. 123-160, 1995.