Hindawi
Journal of Mathematics
Volume 2020, Article ID 3615208, 12 pages
https://doi.org/10.1155/2020/3615208

Research Article

New Investigation for the Liu-Story Scaled Conjugate Gradient Method for Nonlinear Optimization

Eman T. Hamed,[1] **Rana Z. Al-Kawaz**,[2] **and Abbas Y. Al-Bayati**[3]

[1]*Department of Operation Research and Intelligent Techniques, College of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq*
[2]*Department of Mathematics, College of Basic Education, University of Telafer, Tall'Afar, Iraq*
[3]*University of Telafer, Tall'Afar, Iraq*

Correspondence should be addressed to Eman T. Hamed; emantarik84@gmail.com

Received 30 September 2019; Revised 18 November 2019; Accepted 13 December 2019; Published 25 January 2020

Academic Editor: Hang Xu

This article considers modified formulas for the standard conjugate gradient (CG) technique that is planned by Li and Fukushima. A new scalar parameter θ_k^{New} for this CG technique of unconstrained optimization is planned. The descent condition and global convergent property are established below using strong Wolfe conditions. Our numerical experiments show that the new proposed algorithms are more stable and economic as compared to some well-known standard CG methods.

1. Introduction

Conjugate gradient (CG) strategies consists of a category of nonlinear optimization algorithms, which needs low memory and powerful local and global convergence properties [1,2]. Typically, a CG method is meant to resolve massive scaled nonlinear optimization problem:

$$\underset{x\in R^n}{\text{minimize}}\, f(x). \tag{1}$$

On the understanding that the function is defined in the form $f\colon R^n \longrightarrow R$ is smooth nonlinear function. The repetitive formula is in the form

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

The most important component of this formula is α_k step-size, and the search direction d_k consists of

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

whereas $g_k = g(x_k)$ denotes $\nabla f(x_k)$ and β_k denotes a positive scalar. The step-size α_k is sometimes chosen to satisfy bound line search condition [3]. Among these search direction conditions, the strong Wolfe line search condition is sometimes outlined as follows:

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

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k, \tag{5}$$

and $0 < \delta < \sigma < 1$. There are many different formulas for conjugate coefficients as in the following sources, e.g., Hestenes and Stiefel, HS [4]; Fletcher and Reeves, FR [5]; Polak and Ribière, PR [6]; Conjugate Descent, CD [7]; Li and Fukushima, LF [1]; and Liu and Story, LS [8], correspond to different choice for the scalar parameter β_k.

2. A New Scalar Formula for the Parameter θ_k^{New}

Here in this part of this article, we proposed a new version for the parameter θ_k by relying on the modified BFGS method proposed by Li and Fukushima [1]. In the BFGS method, the matrix B_{k+1} is updated to the following formula [9]:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \tag{6}$$

where $y_k = g_{k+1} - g_k$ and $s_k = x_{k+1} - x_k$. In addition, the normal secant relation is outlined consistent with the subsequent formula:

$$B_{k+1} s_k = y_k. \tag{7}$$

The researchers Li and Fukushima presented an appropriate modified BFGS technique which is globally and super-linearly convergent, even though while not requiring convex objective functions. The subsequent modified secant equation is outlined consistent with the subsequent formula as follows:

$$B_{k+1} s_k = \tilde{y}_k, \tag{8}$$

where

$$\tilde{y}_k = y_k + h_k \|g_k\|^r s_k, \tag{9}$$

and $r > 0$; $h_k > 0$, h_k is a parameter defined as

$$h_k = c + \max\left\{ -\frac{s_k^T y_k}{\|s_k\|^2}, 0 \right\} \|g_k\|^r. \tag{10}$$

Specifically, take value c is constant, and it is greater than zero.

There are three different cases for the term $s_k^T y_k$:

Case 1: if $s_k^T y_k \leq 0$, in this case we have the problem of the nonpositive definite matrix, so Li and Fukushima proposed $\tilde{y}_k$ formula as in (9) and developed the corresponding BFGS formula as follows:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{\tilde{y}_k \tilde{y}_k^T}{s_k^T \tilde{y}_k}. \tag{11}$$

Moreover, the form of h_k in (10), when max is used so that the value (0; zero), is not selected in this case. Through this formula, the researchers proved that the modified symmetric matrix is positive definite [10].

Case 2: if $s_k^T y_k > 0$, in this case, we can say surely that the BFGS update matrix is symmetric and positive definite when applied within this formula (in other words, when applying the inequality $s_k^T y_k > 0$ in the formula h_k, max = 0) [11].

If we use β_k^{LS} of Liu and Story (LS), we use any scalar θ_k; then, (3) becomes

$$d_{k+1} = -\theta_k g_{k+1} + \beta_k^{LS} s_k, \tag{12}$$

where

$$\beta_k^{LS} = -\frac{g_{k+1}^T y_k}{s_k^T g_k}. \tag{13}$$

When any positive value to k is greater than one, the new parallel search direction d_{k+1} provided in equation (12) is the Newton direction. Hence, Newton's direction is

$$\begin{aligned} -G_{k+1}^{-1} g_{k+1} &= -\theta_k g_{k+1} + \beta_k^{LS} s_k, \\ -G_{k+1}^{-1} g_{k+1} &= -\theta_k g_{k+1} - \frac{g_{k+1}^T y_k}{s_k^T g_k} s_k. \end{aligned} \tag{14}$$

Hence,

$$\theta_k = \frac{s_k^T g_{k+1} s_k^T g_k - g_{k+1}^T y_k s_k^T G_{k+1} s_k}{s_k^T g_k s_k^T G_{k+1} g_{k+1}}. \tag{15}$$

Using equation (8), the new scalar θ_k^{New} becomes

$$\theta_k^{\text{New}} = \frac{s_k^T g_{k+1} s_k^T g_k - g_{k+1}^T y_k \tilde{y}_k^T s_k}{s_k^T g_k \tilde{y}_k^T g_{k+1}}. \tag{16}$$

By substituting equations (9) and (10) and by taking max = 0 (because we use the strong Wolfe line search in equation (10), yields $h_k = c$). Therefore, the new scalar within the new search direction is

$$\theta_k^{\text{New}} = \frac{s_k^T g_{k+1} + \beta_k^{LS}\left(y_k^T s_k - c\|g_k\|^r \|s_k\|^2 \right)}{y_k^T g_{k+1} + c\|g_k\|^r s_k^T g_{k+1}}. \tag{17}$$

Hence, we conclude from equation (17) that the new parameter θ_k^{New} is best because it is up to date to find the value of y, and also we find different forms when changing the value of c as we will notice in the section of numerical results.

3. New Theorem (Sufficient Descent Direction)

If we presume that the line search satisfies conditions (4) and (5), then the new search direction which is generated from equations (12) and (17) could be a sufficient descent direction.

Proof. From equations (12) and (17) we obtained

$$
\begin{aligned}
d_{k+1} &= -\theta_k^{\mathrm{New}} g_{k+1} - \frac{g_{k+1}^T y_k}{s_k^T g_k} s_k \\
&= -\left[\frac{s_k^T g_{k+1} + \beta_k^{LS}\left(y_k^T s_k - c\|g_k\|^r \|s_k\|^2\right)}{y_k^T g_{k+1} + c\|g_k\|^r s_k^T g_{k+1}}\right] g_{k+1} - \frac{g_{k+1}^T y_k}{s_k^T g_k} s_k, \\
d_{k+1}^T g_{k+1} &= -\left[\frac{s_k^T g_{k+1} s_k^T g_k - g_{k+1}^T y_k\left(y_k^T s_k - c\|g_k\|^r \|s_k\|^2\right)}{s_k^T g_k\left(y_k^T g_{k+1} + c\|g_k\|^r s_k^T g_{k+1}\right)}\right] \cdot \|g_{k+1}\|^2 - \frac{y_k^T g_{k+1}}{s_k^T g_k} s_k^T g_{k+1} \\
&\leq -\left[\frac{y_k^T s_k s_k^T g_k - \|g_{k+1}\|^2\left(y_k^T s_k - c\|g_k\|^r \|s_k\|^2\right)}{s_k^T g_k\left(0.2\, g_{k+1}^2 + c\|g_k\|^r y_k^T s_k\right)}\right] \cdot \|g_{k+1}\|^2 - \frac{y_k^T s_k}{s_k^T g_k} s_k^T g_{k+1}.
\end{aligned} \tag{18}
$$

By using Powell restart equation (i.e., $|g_k^T g_{k+1}| \geq 0.2 g_{k+1}^2$),

$$
\begin{aligned}
d_{k+1}^T g_{k+1} &\leq -\left[\frac{y_k^T s_k s_k^T g_k - \|g_{k+1}\|^2\left(y_k^T s_k - c\|g_k\|^r \|s_k\|^2\right)}{s_k^T g_k\left(0.2\, \|g_{k+1}\|^2 + c\|g_k\|^r y_k^T s_k\right)}\right] \cdot \|g_{k+1}\|^2 - 0.2\, \frac{y_k^T s_k}{s_k^T g_k} \|g_{k+1}\|^2, \\
d_{k+1}^T g_{k+1} &\leq -\left[y_k^T s_k\, s_k^T g_k - \|g_{k+1}\|^2\left(y_k^T s_k - c\|g_k\|^r \|s_k\|^2\right) + 0.2 y_k^T s_k\left(0.2\|g_{k+1}\|^2 + y_k^T s_k c\|g_k\|^r\right)\right] \\
&\quad \cdot \frac{1}{\left[s_k^T g_k\left(0.2\, \|g_{k+1}\|^2 + c\|g_k\|^r y_k^T s_k\right)\right]} \|g_{k+1}\|^2.
\end{aligned} \tag{19}
$$

If $s_k^T g_k\, (0.2\, \|g_{k+1}\|^2 + c\|g_k\|^r s_k^T y_k) > 0$, the next inequality is true:

$$
\begin{aligned}
d_{k+1}^T g_{k+1} &\leq -\left[y_k^T s_k s_k^T g_k - \|g_{k+1}\|^2 y_k^T s_k + c\|g_k\|^r \|s_k\|^2 \|g_{k+1}\|^2 + 0.04\, y_k^T s_k \|g_{k+1}\|^2 + 0.2\|y_k\|^2 \|s_k\|^2 c\|g_k\|^r\right] \\
&\quad \cdot \frac{1}{\left[s_k^T g_k\left(0.2\, \|g_{k+1}\|^2 + c\|g_k\|^r y_k^T s_k\right)\right]} \cdot \|g_{k+1}\|^2 \\
&\leq -\left[\frac{y_k^T s_k s_k^T g_k - 0.96\|g_{k+1}\|^2 y_k^T s_k + c\|g_k\|^r \|s_k\|^2 \cdot \left(0.2\|y_k\|^2 + \|g_{k+1}\|^2\right)}{s_k^T g_k\left(0.2\, \|g_{k+1}\|^2 + c\|g_k\|^r y_k^T s_k\right)}\right] \cdot \|g_{k+1}\|^2.
\end{aligned} \tag{20}
$$

Using strong Wolfe line search condition (5a) yields

$$s_k^T y_k = s_k^T g_{k+1} - s_k^T g_k \geq \sigma s_k^T g_k - s_k^T g_k \geq -(1-\sigma) s_k^T g_k),$$

$$d_{k+1}^T g_{k+1} \leq -\frac{s_k^T g_k}{s_k^T g_k \left(0.2\left\|g_{k+1}\right\|^2 + c\left\|g_k\right\|^r y_k^T s_k\right)} \left\|g_{k+1}\right\|^2 \tag{21}$$

$$\cdot\left[-(1-\sigma) s_k^T g_k + 0.96\,(1-\sigma)\left\|g_{k+1}\right\|^2 + c\left\|g_k\right\|^{r-1} s_k\left(0.2\left\|y_k\right\|^2 + \left\|g_{k+1}\right\|^2\right)\right].$$

This latter equation implies that

$$d_{k+1}^T g_{k+1} \leq -v\left\|g_{k+1}\right\|^2. \tag{22}$$

Thus, our requirement is complete. □

3.1. Outlines of the New CG-Algorithms

Step 1: select the initial point $x_0 \in R^n$, $\epsilon > 0$, and select some positive values for δ and σ. Then, set $d_0 = -\nabla f(x_0)$ and set $k = 0$.

Step 2: test for stopping criterion. If satisfied, then stop; otherwise, continue.

Step 3: determine α_k by Wolfe conditions, which are defined in equations (4) and (5).

Step 4: compute the second iterative point x_{k+1} from equation (2).

Step 5: calculate the scalar parameter θ_k^{New} from equation (17).

Step 6: calculate the new search directions, namely, $d_{k+1} = -\theta_k^{\mathrm{New}} g_{k+1} + \beta_k^{LS} s_k$

Step 7: test Powell restarting criterion, namely, if $|g_{k+1}^T g_k| \geq 0.2 g_{k+1}^2$, then restart the new search direction with $-g_{k+1}$.

Step 8: set the next iteration $k = k+1$, and go to Step 2.

4. Convergence Analysis for the New Proposed Algorithm

In the following parts, we have a tendency to discuss the convergence analysis property for the new algorithm thoroughly. First, we offer an assumption for the convergence analysis property for the new algorithms. Then, we offer another well-known lemma needed within the study of convergence analysis property. Finally, we have a tendency to set new theorems aboard their proofs that area unit associated with the convergence analysis for the new algorithm.

Assumption

(i) The level set

$$S = \{x:\ f(x) \leq f(x_0)\}, \tag{23}$$

is bounded, that is, there exists a constant $z > 0$, such as [12]

$$\|x\| \leq z, \quad \forall x \in S. \tag{24}$$

(ii) In neighbourhood N of S, f is continuously differentiable, and its gradient is Lipschitz continuous, that is, there exists a constant $L > 0$, such as

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall\, x,\ y \in N. \tag{25}$$

From the assumptions (i) and (ii) on f, we are able to deduce that there exists $\gamma > 0$ such as

$$\gamma \leq \|\nabla f(x)\| \leq \overline{\gamma}. \tag{26}$$

Lemma. *If we suppose that [3,13]*

(1) Assumption holds.

(2) Search direction d_{k+1} in the standard CG method is a descent direction.

(3) Optimal step α_k is calculated by equations (4) and (5).

(4) The convergence condition is satisfied, i.e., if

$$\sum_{k\geq 1} \frac{1}{\left\|d_{k+1}\right\|^2} = \infty. \tag{27}$$

Then,

$$\lim_{k\longrightarrow\infty}\left(\inf\left\|g_k\right\|\right) = 0''. \tag{28}$$

5. New Theorem (Uniformly Convex Function)

If we suppose that

(1) Assumption holds.

(2) The new search direction d_{k+1} defined by equations (12) and (17) is a descent direction.

(3) The optimal step α_k is calculated by equations (4) and (5).

(4) The objective function f is uniformly convex; then,

$$\lim_{k\longrightarrow\infty}\left(\left\|g_k\right\|\right)=0. \tag{29}$$

Proof. Consider the new direction in equation (12) and the parameter of equation (17) satisfy the next absolute value condition:

$$\left|\theta_k^{\text{New}}\right|=\left|\frac{s_k^T g_{k+1}s_k^T g_k-g_{k+1}^T y_k\tilde{y}_k^T s_k}{s_k^T g_k\tilde{y}_k^T g_{k+1}}\right|$$

$$\leq\frac{\left|s_k^T g_{k+1}\right|}{\left|\tilde{y}_k^T g_{k+1}\right|}+\frac{\left|g_{k+1}^T y_k\right|.\left|\tilde{y}_k^T s_k\right|}{\left|s_k^T g_k\right|.\left|\tilde{y}_k^T g_{k+1}\right|},$$

$$\text{Since}\begin{cases}\tilde{y}_k^T s_k=y_k^T s_k+c\left\|g_k\right\|^r\left\|s_k\right\|^2\,(\text{s.t. } h_k=c)\\ \leq L\left\|s_k\right\|^2+c\overline{\gamma}^r\left\|s_k\right\|^2\\ \leq\left(c\overline{\gamma}^r+L\right)\left\|s_k\right\|^2,\end{cases}$$

$$\text{also}\begin{cases}\tilde{y}_k^T g_{k+1}=y_k^T g_{k+1}+c\left\|g_k\right\|^r s_k^T g_{k+1}\\ \geq c\left\|g_k\right\|^r y_k^T s_k\\ \geq c\gamma^r\mu\left\|s_k\right\|^2,\end{cases}$$

$$\left|\theta_k^{\text{New}}\right|\leq\frac{L\left\|s_k\right\|^2}{c\gamma^r\mu\left\|s_k\right\|^2}+\frac{\overline{\gamma}L\left\|s_k\right\|\left(c\overline{\gamma}^r+L\right)\left\|s_k\right\|^2}{\left\|s_k\right\|\gamma c\gamma^r\mu\left\|s_k\right\|^2}\leq\frac{L}{c\gamma^r\mu}+\frac{\overline{\gamma}L\left(c\overline{\gamma}^r+L\right)}{c\gamma^{r+1}\mu}. \tag{30}$$

Well parameter β_k^{LS}

$$\left|\beta_k^{LS}\right|\leq\frac{\overline{\gamma}L\left\|s_k\right\|}{\mu\left\|s_k\right\|^2}. \tag{31}$$

Moreover, by combining the results, we obtained

$$\left\|d_{k+1}\right\|\leq\left|\theta_k\right|\left\|g_{k+1}\right\|+\left|\beta_k^{LS}\right|.\left\|s_k\right\|$$

$$\leq\left(\frac{L}{c\gamma^r}+\frac{\overline{\gamma}L\left(c\overline{\gamma}^r+L\right)}{c\gamma^{r+1}\mu}\right)\overline{\gamma}+\frac{\overline{\gamma}L}{\mu}=D_1,$$

$$\sum_{k\geq1}\frac{1}{\left\|d_{k+1}\right\|^2}\geq\frac{1}{D_1}\sum_{k\geq1}1=\infty. \tag{32}$$

We got the required proof. We put similar points to the previous hypotheses, but there are some variations in the formulas. □

6. New Theorem (General Function)

If we suppose that

(1) Assumption holds.

(2) The new search direction d_{k+1} defined by equations (12) and (17)is a descent direction.

(3) The optimal step α_k is calculated by equations (4) and (5).

(4) The objective function f is general function; then,

$$\lim_{k\longrightarrow\infty}\left(\inf\left\|g_k\right\|\right)=0. \tag{33}$$

Proof. Using the same proof style of the previous theorem with the difference in the fact that the functions of the algorithm are general functions,

$$\text{Since}\begin{cases}\tilde{y}_k^T s_k=y_k^T s_k+c\left\|g_k\right\|^r\left\|s_k\right\|^2\\ \geq\gamma\left\|s_k\right\|+c\gamma^r\left\|s_k\right\|^2,\end{cases}$$

$$\left\|d_{k+1}\right\|\leq\left|\theta_k^{New}\right|\left\|g_{k+1}\right\|+\left|\beta_k^{LS}\right|.\left\|s_k\right\|$$
$$\leq\left[\frac{\left\|s_k\right\|\overline{\gamma}}{c\gamma^r L\left\|s_k\right\|}+\frac{\overline{\gamma}L^2\left\|s_k\right\|^2}{\left\|s_k\right\|^2 c\gamma^{r+1}L}\right]\overline{\gamma}+\frac{\overline{\gamma}L\left\|s_k\right\|^2}{L\left\|s_k\right\|^2} \tag{34}$$
$$\leq\frac{\overline{\gamma}^2}{c\gamma^r L}+\frac{\overline{\gamma}^2L^2}{c\gamma^{r+1}L}+\overline{\gamma}=D_2.$$

Then, we obtain

$$\sum_{k\geq1}\frac{1}{\left\|d_{k+1}\right\|^2}\geq\frac{1}{D_2}\sum_{k\geq1}1=\infty. \tag{35}$$

Therefore, the proof of the new theorems in regards to the convergence analyses of the proposed algorithms is complete. □

7. Numerical Experiments

In this section, we have reported some numerical experiments that are performed on a set of (60) unconstrained optimization test problems to analyse the efficiency of θ_k^{New}. Detail of these test problems, with their given initial points, can be found in [14,15]. We handled each of these (60) test functions by adding 1000 for each n to arrive at maximum number of n which is equal to 10000. The termination criterion used in our experiments is $\left\|g_k\right\|\leq10^{-6}$, where $\delta=0.01$ and $\sigma=0.1$.

In our comparisons below, we employ the following algorithms:

(i) LS: with the Wolfe line search

(ii) CD: with the Wolfe line search

(iii) HS: with the Wolfe line search

(iv) PR: with the Wolfe line search

(v) New Algorithm, using equation (17) and the scalar $c=0.1$

(vi) New Algorithm, using equation (17) and the scalar $c=0.001$

In Tables 1 and 2, we numerically compare the proposed new CG algorithms against other well-known CG algorithms to verify their performance using the known comparison tools for such algorithms which are as follows:

Table 1: Comparisons of new algorithm against (LS) & (CD) algorithms for the total of (60) test problems with ($1000 \leq n \leq 10000$, and the increasing size in n is equal to 1000) with ($c = 0.1$ and 0.001).

Prob.	New algorithm ($C = 0.1$) NOI/NOFG/TIME	New algorithm ($C = 0.001$) NOI/NOFG/TIME	LS NOI/NOFG/TIME	CD NOI/NOFG/TIME
1	99/252/0.14	101/263/0.14	323/7045/2.31	414/9608/2.64
2	408/884/2.02	405/880/1.97	412/888/1.98	395/878/2.08
3	853/2183/1.06	830/2171/0.98	823/2159/1.02	824/2154/0.97
4	123/308/0.14	123/308/0.15	114/300/0.17	121/314/0.15
5	100/388/0.15	100/388/0.20	274/5586/1.14	167/1918/0.77
6	585/995/1.56	585/995/1.54	6576/19689/4.54	6848/20496/5.70
7	30/80/0.15	30/80/0.13	40/100/0.21	40/100/0.16
8	1032/2705/1.58	1020/2668/1.42	1033/2768/1.49	1026/2972/1.52
9	2388/4881/2.40	2503/5048/3.06	3520/7182/7.49	3676/7485/8.83
10	477/871/2.66	478/858/2.62	18795/503441/7.60	19346/518146/9.14
11	182/423/1.03	174/418/1.09	8669/270951/5.83	8252/257171/6.83
12	113/302/0.16	113/302/0.20	318/3881/1.69	279/2792/5.80
13	80/226/0.11	78/222/0.08	73/241/0.10	73/241/0.09
14	61/131/0.50	61/131/0.56	124/2140/7.12	160/2833/3.54
15	460/991/0.62	452/969/0.59	456/953/0.61	467/997/0.64
16	66/132/0.03	66/132/0.04	67/134/0.14	68/136/0.06
17	70/160/0.11	70/160/0.06	69/158/0.10	71/162/0.08
18	753/1577/0.79	791/1668/0.85	691/1477/0.74	782/1687/0.82
19	74/158/0.33	74/158/0.42	123/1905/2.31	70/150/0.30
20	110/349/0.41	109/349/0.37	114/334/0.43	111/329/0.41
21	806/3224/1.60	600/1665/1.26	429/1895/4.93	561/5284/5.69
22	72/275/0.51	72/285/0.49	84/366/0.70	2084/2322/9.54
23	4470/9572/1.33	5033/10748/2.42	20010/168275/4.95	19334/128670/5.02
24	62/201/0.42	62/201/0.40	546/15413/3.73	495/13595/3.48
25	459/1091/0.60	521/1192/0.78	494/1101/0.65	531/1154/0.76
26	56/153/0.08	65/373/0.13	56/155/0.06	66/328/0.12
27	85/203/0.11	85/203/0.12	80/190/0.07	80/190/0.09
28	534/1139/0.65	514/1073/0.68	493/1055/0.63	492/1075/0.61
29	540/1274/0.62	537/1267/0.58	941/2237/1.03	943/2239/1.11
30	576/1440/0.96	591/1504/1.05	19026/581537/2.57	18144/543276/2.03
31	113/236/0.16	113/236/0.11	118/245/0.11	123/253/0.11
32	813/2181/3.28	735/1997/3.03	18140/560188/8.67	18320/564064/8.04
33	98/268/0.18	98/268/0.17	359/5355/2.21	387/6200/3.06
34	346/766/0.48	348/766/0.50	356/786/0.44	332/732/0.44
35	7635/12820/8.59	7554/12715/8.73	7463/12576/8.29	7183/12051/8.37
36	280/978/0.43	280/978/0.35	264/935/0.45	273/978/0.39
37	217/534/0.31	217/534/0.27	221/551/0.31	219/546/0.31
38	121/287/0.14	120/285/0.13	814/20969/8.83	786/20117/7.46
39	150/329/0.64	153/328/0.65	141/305/0.59	137/297/0.58
40	107/217/0.65	107/217/0.68	144/1165/3.67	124/476/1.35
41	120/330/0.23	120/330/0.25	100/290/0.17	100/290/0.18
42	3832/9998/2.51	3373/8928/7.18	3404/9046/8.47	3182/8618/7.82
43	40/80/0.11	40/80/0.08	40/80/0.12	40/80/0.12
44	50/110/0.10	50/110/0.05	50/110/0.09	50/110/0.05
45	43/184/0.08	43/184/0.11	3903/129109/7.10	7265/242193/4.78
46	427/1323/2.56	427/1320/2.60	413/1307/2.53	428/1252/2.44
47	64/249/0.09	64/249/0.06	118/447/0.22	114/425/0.19
48	308/798/1.39	293/772/1.28	773/7349/0.72	1731/23050/5.88
49	22/89/0.11	22/89/0.13	316/8561/9.47	298/8022/8.33
50	20/50/0.03	20/50/0.02	20/50/0.00	20/50/0.04
51	92/1755/2.25	42/136/0.17	361/9646/9.02	315/8202/8.61
52	107/418/0.16	107/418/0.15	107/418/0.16	107/418/0.16
53	6199/52426/6.75	6199/52426/6.78	8071/53291/6.20	8071/53291/9.30
54	51/151/0.20	51/151/0.19	51/151/0.21	51/151/0.18
55	60/140/0.19	60/140/0.19	60/140/0.18	60/140/0.20
56	70/140/0.20	70/140/0.17	70/140/0.19	70/140/0.16
57	79/158/0.22	79/158/0.22	79/158/0.24	74/148/0.22
58	143/570/0.25	143/570/0.25	143/570/0.23	143/570/0.21

Table 1: Continued.

Prob.	New algorithm ($C = 0.1$) NOI/NOFG/TIME	New algorithm ($C = 0.001$) NOI/NOFG/TIME	LS NOI/NOFG/TIME	CD NOI/NOFG/TIME
59	188/498/0.29	172/453/0.27	176/449/0.27	177/461/0.22
60	83/236/0.13	83/236/0.11	985/27409/9.27	917/25010/8.38
Total	37602/124886/55.54	37426/121643/59.2	131933/2455353/154.98	136949/1242169/166.56

Table 2: Comparisons of new algorithm against (HS) & (PR) algorithms for the total of (60) test problems with ($1000 \leq n \leq 10000$, and the increasing size in n is equal to 1000) with ($c = 0.1$ and 0.001).

Prob.	New algorithm ($C = 0.1$) NOI/NOFG/TIME	New algorithm ($C = 0.001$) NOI/NOFG/TIME	HS NOI/NOFG/TIME	PR NOI/NOFG/TIME
1	99/252/0.14	101/263/0.14	5902/173484/0.72	11798/270815/6.07
2	408/884/2.02	405/880/1.97	362/637/1.83	416/720/2.07
3	853/2183/1.06	830/2171/0.98	789/1817/0.86	989/1979/1.11
4	123/308/0.14	123/308/0.15	141/281/0.19	254/430/0.26
5	100/388/0.15	100/388/0.20	653/17073/3.79	410/7948/2.73
6	585/995/1.56	585/995/1.54	9881/16243/3.07	20010/22091/4.04
7	30/80/0.15	30/80/0.13	40/90/0.18	40/90/0.18
8	1032/2705/1.58	1020/2668/1.42	997/2279/1.28	8780/10109/1.57
9	2388/4881/2.40	2503/5048/3.06	4658/7644/2.48	14945/16013/9.92
10	477/871/2.66	478/858/2.62	20010/98744/4.70	20010/292528/9.39
11	182/423/1.03	174/418/1.09	14053/39805/1.28	15593/489487/8.58
12	113/302/0.16	113/302/0.20	428/6427/2.19	609/11752/4.81
13	80/226/0.11	78/222/0.08	113/234/0.08	291/509/0.26
14	61/131/0.50	61/131/0.56	906/23792/1.66	318/6032/3.76
15	460/991/0.62	452/969/0.59	636/1006/0.73	964/1479/1.25
16	66/132/0.03	66/132/0.04	60/120/0.03	1043/1116/0.38
17	70/160/0.11	70/160/0.06	207/339/0.19	110/230/0.10
18	753/1577/0.79	791/1668/0.85	821/1545/0.78	3732/4630/3.99
19	74/158/0.33	74/158/0.42	108/1352/2.44	303/7194/2.66
20	110/349/0.41	109/349/0.37	135/321/0.42	154/339/0.45
21	806/3224/1.60	600/1665/1.26	875/14122/5.89	929/10442/9.02
22	72/275/0.51	72/285/0.49	2104/2442/3.61	161/440/0.92
23	4470/9572/1.33	5033/10748/2.42	18912/38658/8.87	20010/25808/6.95
24	62/201/0.42	62/201/0.40	1853/6983/3.39	2527/76854/9.01
25	459/1091/0.60	521/1192/0.78	304/606/0.40	1199/1793/1.55
26	56/153/0.08	65/373/0.13	128/1103/0.37	2697/14700/7.09
27	85/203/0.11	85/203/0.12	91/193/0.13	132/264/0.16
28	534/1139/0.65	514/1073/0.68	288/558/0.35	556/925/0.60
29	540/1274/0.62	537/1267/0.58	852/1783/0.98	1014/2180/1.14
30	576/1440/0.96	591/1504/1.05	20010/98171/6.44	20010/317766/8.02
31	113/236/0.16	113/236/0.11	79/168/0.10	147/287/0.13
32	813/2181/3.28	735/1997/3.03	20010/91480/9.47	20010/179051/5.16
33	98/268/0.18	98/268/0.17	631/11069/5.15	837/17999/8.99
34	346/766/0.48	348/766/0.50	716/1148/0.90	744/1213/0.93
35	7635/12820/8.59	7554/12715/8.73	8375/13146/9.98	8539/12513/15.46
36	280/978/0.43	280/978/0.35	330/695/0.37	401/868/0.47
37	217/534/0.31	217/534/0.27	610/6778/2.40	820/11112/5.50
38	121/287/0.14	120/285/0.13	1565/42467/4.89	1624/45918/6.43
39	150/329/0.64	153/328/0.65	174/290/0.54	193/323/0.69
40	107/217/0.65	107/217/0.68	253/426/1.03	4281/4461/11.17
41	120/330/0.23	120/330/0.25	118/286/0.19	124/298/0.20
42	3832/9998/2.51	3373/8928/7.18	3685/8619/8.30	17419/22467/5.83
43	40/80/0.11	40/80/0.08	99/119/0.19	99/119/0.19
44	50/110/0.10	50/110/0.05	50/110/0.02	70/282/0.13
45	43/184/0.08	43/184/0.11	12047/91618/9.93	15601/521624/4.14
46	427/1323/2.56	427/1320/2.60	409/1040/2.07	597/1232/2.51
47	64/249/0.09	64/249/0.06	133/380/0.15	143/393/0.40

Table 2: Continued.

Prob.	New algorithm ($C=0.1$) NOI/NOFG/TIME	New algorithm ($C=0.001$) NOI/NOFG/TIME	HS NOI/NOFG/TIME	PR NOI/NOFG/TIME
48	308/798/1.39	293/772/1.28	2549/27198/2.61	15662/156182/6.10
49	22/89/0.11	22/89/0.13	1377/37674/8.58	1571/43940/8.14
50	20/50/0.03	20/50/0.02	20/50/0.02	20/50/0.03
51	92/1755/2.25	42/136/0.17	1620/5439/3.43	1680/46332/4.04
52	107/418/0.16	107/418/0.15	125/360/0.14	125/360/0.16
53	6199/52426/6.75	6199/52426/6.78	4160/9534/3.23	2233/8070/9.66
54	51/151/0.20	51/151/0.19	75/173/0.25	77/144/0.20
55	60/140/0.19	60/140/0.19	60/120/0.14	60/120/0.17
56	70/140/0.20	70/140/0.17	70/140/0.24	80/160/0.22
57	79/158/0.22	79/158/0.22	79/158/0.24	86/172/0.23
58	143/570/0.25	143/570/0.25	177/506/0.25	177/506/0.25
59	188/498/0.29	172/453/0.27	209/460/0.30	641/917/1.01
60	83/236/0.13	83/236/0.11	1615/43549/5.17	2005/57225/7.75
Total	37602/124896/60.1	37426/121943/70.14	167737/10741186/139.61	271227/7136001/214.33

NOI = the total number of calculated iterative iterations

NOFG = the total number of function and gradient calculations

TIME = the total CPU time required for the processor to execute the CG algorithm and reach the minimum value of the required function minimization

Therefore, among these CG algorithms, the new algorithm appears to generate the best search direction. In Table 3, there is a clear evidence that the new algorithm outperforms the standard LS and CD algorithms detailed as follows (when $c=0.1$):

(a) For 100% LS algorithm: the new algorithm is improved by (71.5%) NOI, improved by (94.92%) NOFG, and improved by (64.2%) time

(b) For 100% CD algorithm: the new algorithm is improved by (72.6%) NOI, improved by (89.95%) NOFG, and improved by (66.7%) time

And (when $c=0.001$):

(c) For 100% LS algorithm: the new algorithm is improved by (64.34%) NOI, improved by (16.99%) NOFG, and improved by (16.08%) time

(d) For 100% CD algorithm: the new algorithm is improved by (52.60%) NOI, improved by (14.63%) NOFG, and improved by (12.75%) time

In Table 4, there is a clear evident that the new algorithm outperforms the standard HS and PR algorithms as detailed below (when $c=0.1$):

(e) For 100% HS algorithm: the new algorithm is improved by (77.6%) NOI, improved by (98.9%) NOFG, and improved by (57%) time

(f) For 100% PR algorithm: the new algorithm is improved by (86.2%) NOI, improved by (98.3%) NOFG, and improved by (72%) time

and (when $c=0.001$)

Table 3: Standardizing perceptual of the new algorithm vs. LS and CD algorithms.

Algorithm	Tools	LS (1991) (%)	CD (1987) (%)
When $c=0.1$	NOI	28.5	27.4
	NOFG	5.08	10.05
	TIME	35.8	33.3
When $c=0.001$	NOI	28.3	27.3
	NOFG	4.9	9.7
	TIME	38.1	35.5

Table 4: Standardizing perceptual of the new algorithm vs. HS and PR algorithms.

Algorithm	Tools	HS (1952) (%)	PR (1969) (%)
When $c=0.1$	NOI	22.4	13.8
	NOFG	1.1	1.7
	TIME	43	28
When $c=0.001$	NOI	22.3	13.7
	NOFG	1.1	1.7
	TIME	50.2	32.7

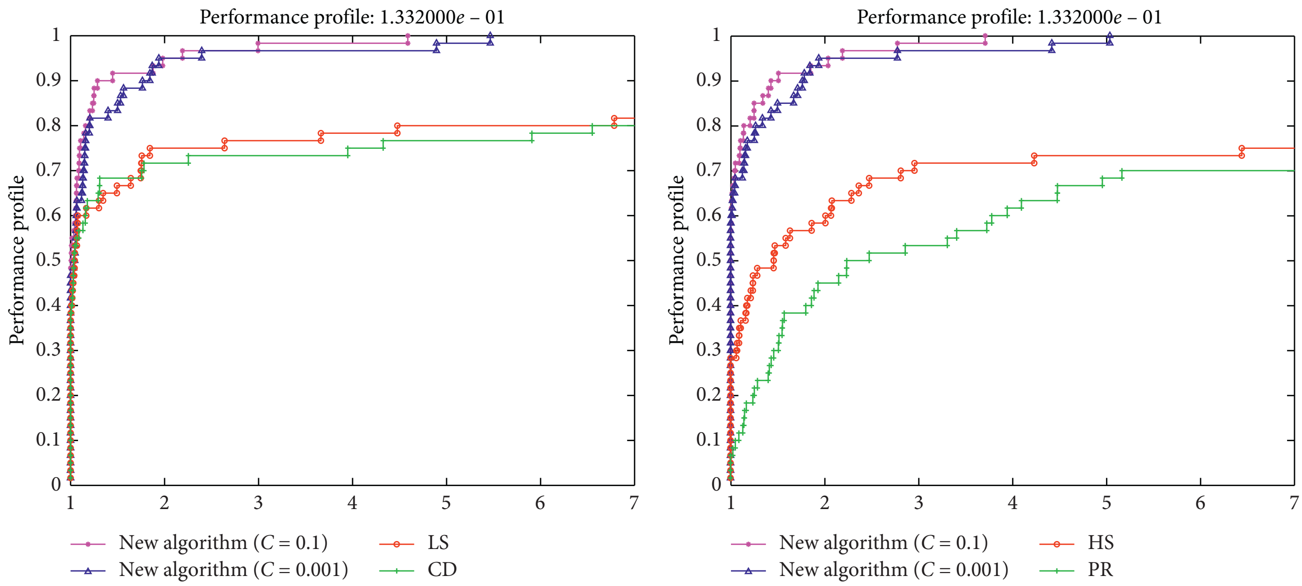
(g) For 100% HS algorithm: the new algorithm is improved by (77.7%) NOI, improved by (98.9%) NOFG, and improved by (49.8%) time

(h) For 100% PR algorithm: the new algorithm is improved by (86.3%) NOI, improved by (98.3%) NOFG, and improved by (67.3%) time

What can be deduced from the above tables and experiments are summarized in the following:

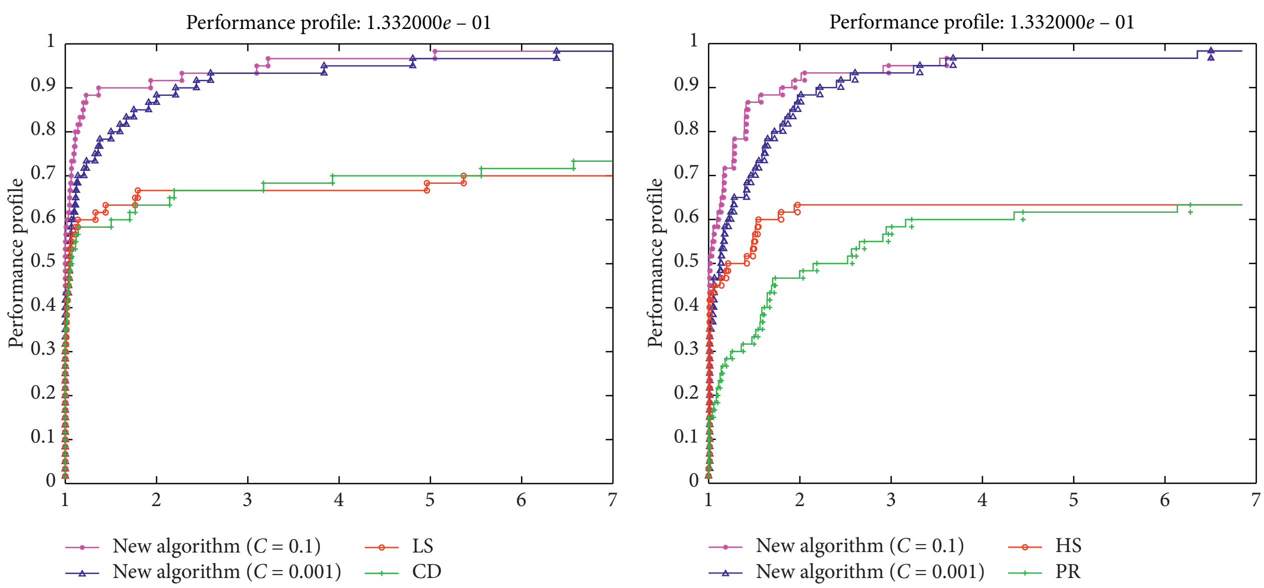
(i) Points (a to d) above are that our new proposed algorithms in the field of CG-type methods are economic and robust as compared to the standard LS and CD algorithms

(ii) The abovementioned points (e to h) are that our new proposed algorithms in the field of CG-type methods



(a) (b)

Figure 1: Performance profiles based on number of iterations (NOI). (a) Comparison of the new algorithm vs. LS and CD. (b) Comparison of the new algorithm vs. HS and PR.



(a) (b)

Figure 2: Performance profiles based on function and gradient evaluations (NOFG). (a) Comparison of the new algorithm vs. LS and CD. (b) Comparison of the new algorithm vs. HS and PR.

are economic and robust as compared to the standard HS and PR algorithms

All these comparisons were made using the performance profile of Dolan and Moré [16], and we can conclude that

(1) Figure 1 illustrates the new algorithm versus (LS, CD, HS, and PR) the activity of the new algorithms in calculating the number of iterations

(2) Figure 2 explains the new algorithm versus (LS, CD, HS, and PR) the activity of the new algorithms in

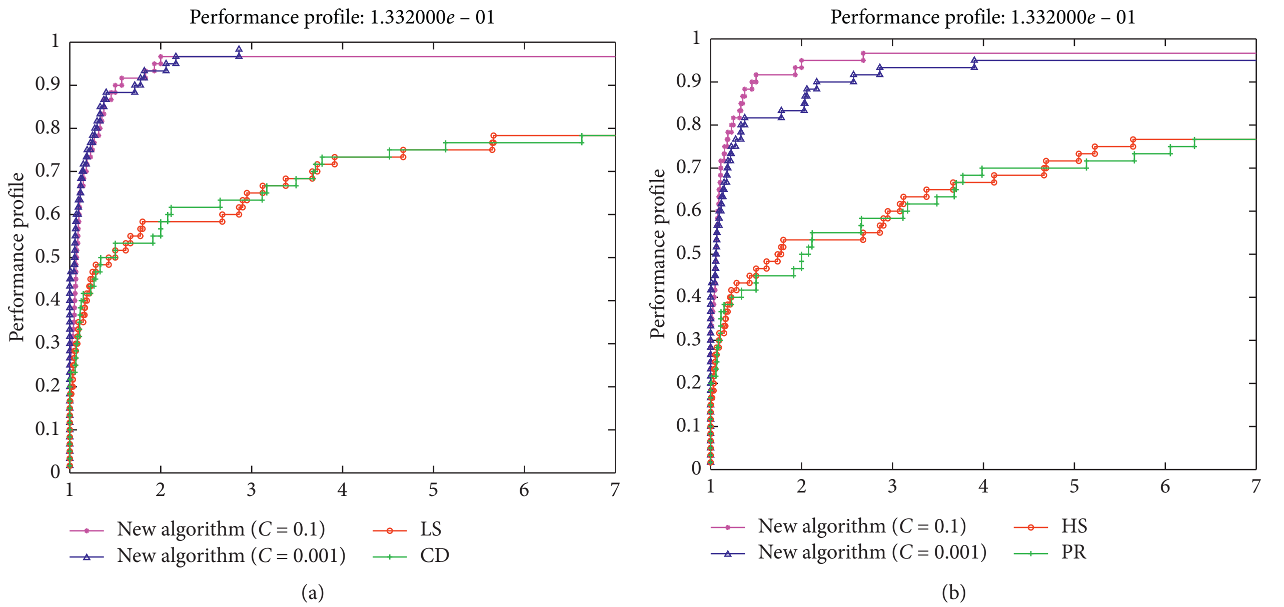


FIGURE 3: Performance profiles based on CPU Time. (a) Comparison of the new algorithm vs. LS and CD. (b) Comparison of the new algorithm vs. HS and PR.

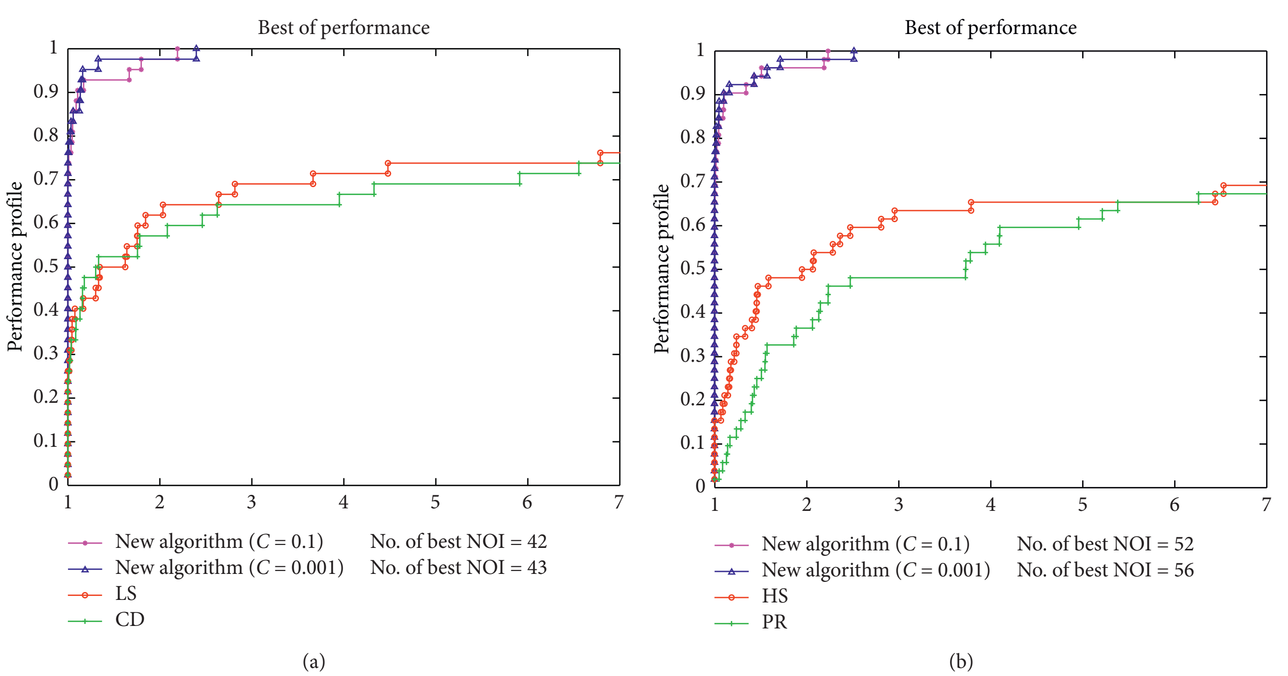


FIGURE 4: Performance profiles based on the best NOI. (a) Comparison of the new algorithm vs. LS and CD. (b) Comparison of the new algorithm vs. HS and PR.

calculating the number of function and gradient evaluations

(3) Figure 3 displays how long the algorithms take to reach the solution (i.e., the required CPU time)

(4) Figure 4 shows the functions that perform well in the new algorithm with two different constants compared to the basic algorithms (LS, CD, HS, and PR) based on the number of iterations

(5) Figure 5 demonstrates the outstanding performance of a number of functions in the new algorithm with two different constants compared to basic algorithms (LS, CD, HS, and PR) based on the number of function and gradient evaluations

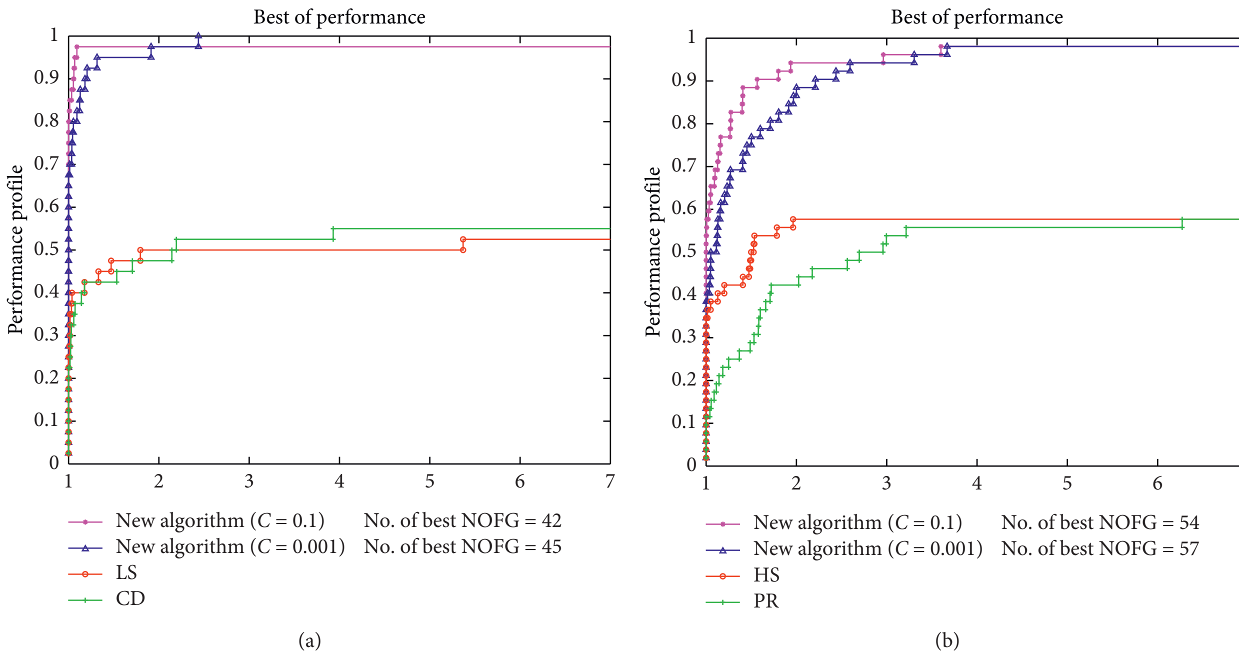


FIGURE 5: Performance profiles based on the best NOFG. (a) Comparison of the new algorithm vs. LS and CD. (b) Comparison of the new algorithm vs. HS and PR.

8. Conclusions

In this study, we have submitted two proposed new CG methods (by changing the value of c). A crucial property of proposed CG methods is that it secures sufficient descent directions. Under mild conditions, we have demonstrated that the new algorithms are globally convergent for each uniformly convex and general functions using the strong Wolfe line search conditions. The preliminary numerical results show that if we decide a good value of parameter c, the new algorithms perform very well. However, an optimal value of the parameter c can be handled theoretically (in future research studies) to achieve more best numerical results.

Data Availability

The data used the support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The research was supported by College of Computer Sciences and Mathematics, University of Mosul, Republic of Iraq, under Project no. 3615208.

References

[1] D.-H. Li and M. Fukushima, “A modified BFGS method and its global convergence in nonconvex minimization,” *Journal of Computational and Applied Mathematics*, vol. 129, no. 1-2, pp. 15–35, 2001.

[2] Y. H. Dai and Y. X. Yuan, “Convergent properties of nonlinear conjugate gradient methods,” *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 348–358, 1999.

[3] Y.-H. Dai, “New conjugacy conditions and related nonlinear conjugate gradient methods,” *Applied Mathematics and Optimization*, vol. 43, no. 1, pp. 87–101, 2001.

[4] M. R. Hestenes and E. Stiefel, “Methods of conjugate gradients for solving linear systems,” *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.

[5] R. Fletcher, “Function Minimization by conjugate gradients,” *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.

[6] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, Cham, Switzerland, 2nd edition, 2006.

[7] R. Fletcher, *Practical Methods of Optimization, Unconstrained Optimization*, Wiley, New York, NY, USA, 1987.

[8] Y. Liu and C. Storey, “Efficient generalized conjugate gradients algorithms, part 1: theory,” *Journal of Optimization Theory Applications*, vol. 69, pp. 129–137, 1999.

[9] Q. Guo, J.-G. Liu, D. H. Wang, and D.-H. Wang, “A modified BFGS method and its superlinear convergence in nonconvex minimization with general line search rule,” *Journal of Applied Mathematics and Computing*, vol. 28, no. 1-2, pp. 435–446, 2008.

[10] Z. Liu and H. Liu, “An efficient gradient method with approximately optimal stepsize based on tensor model for unconstrained optimization,” *Journal of Optimization Theory and Applications*, vol. 181, no. 2, pp. 608–633, 2019.

[11] X. Li, B. Wang, and W. Hu, “A modified nonmonotone BFGS algorithm for unconstrained optimization,” *Journal of Inequalities and Applications*, vol. 183, pp. 1–18, 2017.

[12] J. C. Gilbert and J. Nocedal, “Global convergence properties of conjugate gradient methods for optimization,” *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 21–42, 1992.

[13] X. Chen and J. Sun, “Global convergence of a two-parameter family of conjugate gradient methods without line search,”

Journal of Computational and Applied Mathematics, vol. 146, no. 1, pp. 37–45, 2002.

[14] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Model Optimization*, vol. 10, pp. 147–161, 2011.

[15] N. Andrei, "40 Conjugate gradients algorithms for unconstrained optimization," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 34, pp. 319–330, 2011.

[16] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.