Hindawi
Journal of Mathematics
Volume 2022, Article ID 1383129, 9 pages
https://doi.org/10.1155/2022/1383129

Research Article

Nonlinear Conjugate Gradient Coefficients with Exact and Strong Wolfe Line Searches Techniques

Awad Abdelrahman,[1] **Mogtaba Mohammed**,[2,3] **Osman O. O. Yousif**,[1] **and Murtada K. Elbashir**[4]

[1]*Department of Mathematics, College of Science, Sudan University of Science and Technology, Khartoum, Sudan*
[2]*Department of Mathematics, College of Science, Majmaah University, Zulfi 11932, Saudi Arabia*
[3]*Department of Mathematics, Faculty of Mathematical and Computer Sciences, University of Gezira, Wad Madani, Sudan*
[4]*Department of Information Systems, College of Computer and Information Sciences, Jouf University, Sakaka 72441, Saudi Arabia*

Correspondence should be addressed to Mogtaba Mohammed; mogtaba.mohammed@gmail.com

Received 15 April 2022; Revised 29 August 2022; Accepted 5 September 2022; Published 22 November 2022

Academic Editor: Sheng Du

Nonlinear conjugate gradient (CG) methods are very important for solving unconstrained optimization problems. These methods have been subjected to extensive researches in terms of enhancing them. Exact and strong Wolfe line search techniques are usually used in practice for the analysis and implementation of conjugate gradient methods. For better results, several studies have been carried out to modify classical CG methods. The method of Fletcher and Reeves (FR) is one of the most well-known CG methods. It has strong convergence properties, but it gives poor numerical results in practice. The main goal of this paper is to enhance this method in terms of numerical performance via a convexity type of modification on its coefficient β_k. We ensure that with this modification, the method is still achieving the sufficient descent condition and global convergence via both exact and strong Wolfe line searches. The numerical results show that this modified FR is more robust and effective.

1. Introduction

An unconstrained optimization problem is solved using the nonlinear conjugate gradient method to obtain the minimal value of a given function. The CG technique is usually expressed as follows:

$$\min_{x\in R^n} f(x), \quad (1)$$

such that $f : R^n \longrightarrow R, f \in C^1(R)$, and $g(x)$ are given function denoting its gradient. To apply CG methods in solving (1), we start with an initial point $x_1 \in R^n$ and use the iterative form

$$x_{k+1} = x_k + \alpha_k d_k, k = 1, 2, 3, \cdots, \quad (2)$$

such that x_k is the current iterative point, $\alpha_k > 0$ is a step-size determined by some line searches, and d_k is the search direction defined by

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

where $g_k = \nabla f(x_k)$ and β_k are scalar.

Different choices of the scalar β_k relate to different conjugate gradient algorithms. Over the years, several versions of this approach have been presented, some of which are now extensively utilized. β_k has at least six well-known formulas, which are listed below (Dai and Yaxiang [1], Conjugate descent [2], Fletcher and Colin [3], Liu and Storey [4], Hestenes and Stiefel [5], and Polak and Ribiere-Polyak [6]):

$$\beta_k^{\mathrm{HS}} = \frac{g_k^T (g_k - g_{k-1})}{(\mathrm{g_k} - \mathrm{g_{k-1}})^{\mathrm{T}} \mathrm{d_{k-1}}}, \quad (4)$$

$$\beta_k^{\mathrm{FR}} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \quad (5)$$

$$\beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}}, \tag{6}$$

$$\beta_k^{CD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}}, \tag{7}$$

$$\beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}}, \tag{8}$$

$$\beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}}. \tag{9}$$

Many authors have examined the behavior of global convergence for the β_k's formulas with several line search for several years (see, for example, [1–4, 7–18]). When the objective function is a strongly convex quadratic and the line search is exact, these methods are identical, since the gradients are mutually orthogonal, and the scalars β_k in these methods are equal. When applied for general nonlinear functions with inexact line searches, the behavior of these methods is clearly distinct (see [11, 17, 19–21]). One of the most important properties of the CG methods is global convergence. Zoutendijk [22] proved the global convergence of the FR method via the exact line search. Although FR, DY, and CD methods have strong convergence properties, they may not perform well in practice [14]. The CD and DY methods were proved to have a global convergence under strong Wolfe line search [3, 15]. Moreover, under the exact and strong Wolfe line searches, up to our knowledge, the global convergence and sufficient descent property of some CG methods such as the PRP and the HS have not been established [2, 14]. Andrei [9] classified the CG into groups, these are scaled CG method, classical CG method, and hybrid and parameterized CG methods.

Formulas (4) to (9) are in the classical group. One of the most important among them is the FR method, in which when a bad direction and a little step from x_{k-1} to x_k are generated, the next search direction and the next step $s_k = x_k - x_{k-1} = \alpha_{k-1} d_{k-1}$ will be delayed unless a rest along the gradient direction is performed. Despite this defect, it is well known that the FR method is globally convergent for general nonlinear functions with exact or inexact line search [18].

Al-Baali [7] proved the global convergence of the FR method if the strong Wolfe line search is used and the parameter σ is restricted in $(0, 1/2)$. Also, Liu and Li [17] extended Al-Baali's result to the case that $\sigma = 1/2$. Moreover, Gilbert and Jorge [14] investigated global convergence properties of the dependent FR conjugate gradient method with β_k satisfying $|\beta_k| \le \beta_k^{FR}$, provided that the line search satisfies the strong Wolfe line search. If β_k satisfies $|\beta_k| \le c\beta_k^{FR}$ (where $c > 1$), they have given an example to indicate that even exact line search cannot guarantee the global convergence property of the dependent FR conjugate gradient method. Nosratipour and Keyvan [23] proved that the dependent FR conjugate gradient method is globally convergent with the strong Wolfe line search if β_k satisfies

$$\sigma \left| \frac{\beta_k}{\beta_k^{FR}} \right| \le \bar{\sigma},$$
$$\|g_k\|^2 \sum_{j=1}^{k} \prod_{i=j}^{k-1} l_i \le \text{ck}, \tag{10}$$

where $0 < \sigma < 1$, $0 < \bar{\sigma} < 1/2$, $l_i = |\beta_i / \beta_i^{FR}|$, and $c > 0$.

Zhang and Donghui [24] proposed a modified FR method (called MFR) in which the direction d_k is defined by

$$d_k = \begin{cases} -g_k, & k = 0, \\ \dfrac{-d_{k-1} y_{k-1}}{\|g_{k-1}\|^2} g_k + \beta_k^{FR} d_{k-1}, & k > 0, \end{cases} \tag{11}$$

which is a descent direction independent of the line search.

Competitive numerical and global convergence results are obtained by newly introduced or modified techniques; see for instance [10] and the references therein.

This paper is organized as follows: Section 2 is devoted obtain the modification and introduce the algorithm for the modified FR method. In Section 3, we have determined sufficient descent and global convergence properties under exact and strong Wolfe line searches. Section 4 provides preliminary numerical results and considerations. Section 5 shows the conclusions.

2. Motivation and Properties

Several authors have attempted to modify classical CG methods such as FR, PRP, HS, and LS in order to produce new modifications with sufficient descent and global convergence properties. In addition to that, the new modifications are expected to have an efficient numerical performance. It is well known that the FR method has poor numerical performance but has strong convergence properties. The main aim from this paper is to overcome this flaw, using the following modification to the FR formula:

$$\beta_k^{NMFR} = \frac{\|g_k\|^2}{(1-\theta)\|d_{k-1}\|^2 + \theta \|g_{k-1}\|^2}, \tag{12}$$

where $0 < \theta < 1, \theta$ is a scalar parameter, which is to be determined later. Note that if $\theta = 1$, then $\beta_k^{NMFR} = \beta_k^{FR}$ and if $\theta = 0$, then $\beta_k^{NMFR} = \|g_k\|^2 / \|d_{k-1}\|^2$. On the other hand, if $0 < \theta < 1$, then β_k^{NMFR} is the modification of β_k^{FR}, where $\|.\|$ means the norm in R. (12) satisfies the following inequalities:

$$\beta_k^{NMFR} \le \frac{\|g_k\|^2}{\theta \|g_{k-1}\|^2}. \tag{13}$$

Algorithm 1. Step 1. Initialization. Given $x_0 \in R^n$, $d_0 = -g_0$, $\varepsilon > 0$ is tolerance, $k = 0$, when $g_0 = 0$ stop.

Step 2. Evaluate β_k according to (12).

Step 3. Evaluate d_k according to (3); if $\|g_k\| \leq \varepsilon$, then stop; otherwise, go to the next step.

Step 4. Evaluate an $\alpha_k > 0$ using exact line search, i.e., α_k that satisfies

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k), \tag{14}$$

and inexact line searches, i.e., α_k that satisfies

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \tag{15}$$

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k, \tag{16}$$

where $0 < \delta < \sigma < 1$.

Step 5. Renew the point according to (2), if $\|g_k\| \leq \varepsilon$, then stop

Step 6. Set $k = k + 1$, and go to step 2.

The ideal method to choose the step length is via exact line search. But since it is too expensive to choose it via exact line search in practice, some approximation methods called inexact line search such as strong Wolfe are used to define the step length that give suitable reductions in the objective function with minimal cost. However, the convergence properties of some CG methods such as FR, RMIL, and RMIL+ have been established under exact line search (see [3, 6, 25]).

3. Convergent Analysis

In this section, we will examine the convergence properties of β_k^{NMFR}. The main feature of Algorithm 1 is achieving sufficient descent conditions and global convergence properties according exact and strong Wolfe line searches.

3.1. Convergent Analysis via the Exact Line Search. In this subsection, we show that our modification (12) will possess sufficient descent condition and global convergence properties according to the exact line search.

Theorem 1. *Assume that (2) and (3) be generated by Algorithm 1 and α_k be determined by the exact line search (14), where $\beta_k = \beta_k^{NMFR}$ is given as (12); there exists $c > 0$, such that*

$$g_k^T d_k \leq -c\|g_k\|^2 \quad \forall k \geq 0. \tag{17}$$

The proof of this Theorem 1 is obvious, from (3), and multiply by g_{k+1}; then,

$$g_{k+1}^T d_{k+1} = g_{k+1}^T(-g_{k+1} + \beta_{k+1} d_k) = -\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k. \tag{18}$$

When $g_{k+1}^T d_k = 0$, then (17) holds true for all $k \geq 1$ and becomes

$$g_k^T d_k = -\|g_k\|^2. \tag{19}$$

3.1.1. Global Convergence Properties. In this subsection, we show that our modified (12) coefficient satisfies global convergence according to the exact line search.

Assumption 2.

(i) There exists some positive constant c such that $c \geq \|f(x)\|$ for all $x \in R^n$ and $f \in C^1(N)$ for some neighborhood N of $\Gamma = \{x \in R^n | f(x) \leq f(x_0)\}$. Also, assume that there exists a positive constant C_0 such that

$$\|\mathrm{x} - \mathrm{y}\| \leq \mathrm{C}_0, \text{for any } x, y \in N \tag{20}$$

(ii) There exists positive constant C_1

$$\|g(x) - g(y)\| \leq C_{1\|x-y\|}, \text{for all } x, y \in N \tag{21}$$

From the above assumption, we can easily see that

$$\|g(x)\| \leq \mathrm{C}_2 \quad \text{for any } x \in \Gamma \text{ and some positive constant } C_2. \tag{22}$$

The following lemma will be used in our analysis (see Zoutendijk [22]).

Lemma 3. *Assume that Assumption 2 holds, for any conjugate gradient method as in (2), such that d_k is a descent search direction and α_k satisfies the exact line search. Then,*

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{23}$$

Theorem 4. *Assume that Assumption 2 holds, for conjugate gradient method as in (2) and (3) such that α_k is achieved via exact line search. Furthermore, assume that the sufficient descent condition is satisfied. Then,*

$$\lim_{k \longrightarrow \infty} \|g_k\| = 0, \quad \text{or} \sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{24}$$

Proof. The proof will be conducted by contradiction argument. So, assume that the statement of Theorem 4 is false. Thus, there exists some positive constant c, where

$$\|\mathrm{g_k}\| \geq c. \tag{25}$$

We rewrite (3) as

$$d_{k+1} + g_{k+1} = \beta_{k+1} d_k. \tag{26}$$

Taking the square on both sides of above equation, we get

$$\|d_{k+1}\|^2 = -\|g_{k+1}\|^2 - 2g_{k+1}^T d_{k+1} + (\beta_{k+1})^2 \|d_k\|^2. \quad (27)$$

Dividing both sides of (27) by $\|g_{k+1}\|^4$, we obtain

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} = \frac{-1}{\|g_{k+1}\|^2} - \frac{2\,g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^4} + \frac{(\beta_{k+1})^2 \|d_k\|^2}{\|g_{k+1}\|^4}. \quad (28)$$

From (19) and (13), we obtain

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} = \frac{-1}{\|g_{k+1}\|^2} + \frac{(\beta_{k+1})^2 \|d_k\|^2}{\|g_{k+1}\|^4} \le \frac{\|d_k\|^2}{\theta^2 \|g_k\|^4}, \quad (29)$$

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \le \frac{\|d_k\|^2}{\theta^2 \|g_k\|^4}. \quad (30)$$

Recursively using (29) and noticing that $\|d_1\|^2 = -g_1^T d_1 = \|g_1\|^2$, we get

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \le \left(\frac{1}{\theta^2}\right) \sum_{i=0}^{k+1} \frac{1}{\|g_i\|^2}. \quad (31)$$

Hence,

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \le \frac{k}{\theta^2 c^2}, \quad (32)$$

$$\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \ge \frac{\theta^2 c^2}{k}. \quad (33)$$

As a result of (32) and (25), it is clear that

$$\sum_{k=0}^{\infty} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} = \infty, \quad (34)$$

which contradicts Lemma 3's Zoutendijk condition; hence, the proof is completed. □

3.2. Convergent Analysis according to Strong Wolf Line Searches. In this subsection, we show that our modified (12) coefficient satisfies sufficient descent condition and global convergence according to strong Wolfe line search (15) and (16).

3.2.1. Sufficient Descent Condition

Theorem 5. *Let Assumption 1 hold true, for $\sigma < 1/2$ and $\theta = 0.3$. Then, Algorithm 1 assures the sufficient descent condition (17) such that*

$$c = 2 - \frac{\theta}{\theta - \sigma}. \quad (35)$$

Proof. Multiplying (3) by g_{k+1} and (13) and taking the absolute value of the second term in (36), we obtain

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k \\ &\le -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\theta \|g_k\|^2} \left| g_{k+1}^T d_k \right|. \end{aligned} \quad (36)$$

From (16) and the Cauchy-Schwartz inequality, we get

$$\begin{aligned} g_{k+1}^T d_{k+1} &\le -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\theta \|g_k\|^2} \left| g_{k+1}^T d_k \right| \\ &\le -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\theta \|g_k\|^2} \left(-\sigma g_k^T d_k \right) \\ &\le -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\theta \|g_k\|^2} \left(\left| -\sigma g_k^T d_k \right| \right), \end{aligned} \quad (37)$$

$$\begin{aligned} g_{k+1}^T d_{k+1} &\le -\|g_{k+1}\|^2 + \sigma \frac{\|g_{k+1}\|^2}{\theta \|g_k\|^2} \|g_k\| \|d_k\| \\ &= -\|g_{k+1}\|^2 + \frac{\sigma}{\theta} \|g_{k+1}\|^2 \frac{\|d_k\|}{\|g_k\|}. \end{aligned}$$

Dividing both sides in the above inequality by $\|g_{k+1}\|^2$, we obtain

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -1 + \frac{\sigma}{\theta} \frac{\|d_k\|}{\|g_k\|}. \quad (38)$$

By repeating this process and the fact that $\|d_1\| = \|g_1\|$, we get

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -1 + \frac{\sigma}{\theta} \frac{\|d_k\|}{\|g_k\|}, \quad (39)$$

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2 + \sum_{i=0}^{\infty} \left(\frac{\sigma}{\theta}\right)^i = -2 + \frac{1}{1 - \sigma/\theta} = -2 + \frac{\theta}{\theta - \sigma}, \quad (40)$$

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -c. \quad (41)$$

Therefore, from (39), we can deduce that (17) holds for $k \ge 0$. The proof is completed. □

3.2.2. Global Convergence. This subsection is devoted to the global convergence in the case of inexact line search technique. The following lemma is collected from [22].

Lemma 6. *Let Assumption 1 and let any conjugate gradient method be in the form (2), where the descent direction is d_{k+1} and α_k satisfies the strong Wolfe line search (15) and (16).*

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \quad (42)$$

Theorem 7. *Let Assumption 2 hold. Then, Algorithm 1 is*

Table 1: A list of test problem functions.

No.	Function	DIM	FR IN/CPU	PRP IN/CPU	CD IN/CPU	NMFR IN/CPU
1	Extended White & Holst	10	631/4.449	11/0.106	341/2.473	28/0.246
		100	177/2.312	20/0.292	308/4.012	17/0.246
2	Extended Rosenbrock	4	205/1.368	30/0.226	341/2.300	26/0.202
		100	204/1.629	30/0.278	316/2.554	26/0.242
3	Extended Freudenstein & Roth	10	F/F	F/F	57/0.445	34/0.269
		100	117/1.114	F/F	42/0.446	59/0.573
4	Extended Beale	4	66/0.566	8/0.083	66/0.571	27/0.243
		1000	19/1.637	9/0.776	19/1.641	10/0.859
5	Raydan1	10	52/0.469	24/0.220	52/0.470	26/0.247
		10	190/2.244	104/1.257	186/2.210	131/1.563
6	Extended Tridiagonal1	100	25/0.399	14/0.227	25/0.406	42/0.656
		1000	48/3.760	14/1.104	48/3.763	76/5.948
7	Diagonal4	1000	5/0.124	4/0.099	5/0.122	5/0.118
		10000	5/0.646	4/0.525	5/0.658	5/0.644
8	Extended Himmelblau	10000	8/1.509	6/1.124	8/1.501	8/1.488
		50000	8/7.165	6/5.371	8/7.157	8/7.157
9	FLETCHCR	10	19/0.179	19/0.179	19/0.186	19/0.173
		100	24/0.265	24/0.264	25/0.274	24/0.260
10	Diagonal2	4	10/0.118	12/0.132	10/0.117	10/0.114
		10	148/1.414	30/0.303	147/1.405	30/0.300
11	NONSCOMP	2	34/0.309	8/0.086	34/0.309	10/0.104
		4	1725/14.289	344/2.972	1726/14.316	546/4.730
12	Extended DENSCHNB	4	11/0.112	9/0.102	11/0.117	9/0.091
		100	11/0.135	9/0.109	11/0.135	10/0.122
13	Extended Penalty	10	10/0.118	7/0.093	10/0.113	8/0.093
		100	29/0.357	F/F	49/0.569	17/0.209
14	Hager	10	14/0.155	13/0.142	14/0.153	13/0.142
		100	99/1.386	36/0.521	94/1.335	36/0.521
15	ARWHEAD	4	6/0.070	7/0.086	6/0.084	6/0.070
		10	7/0.086	7/0.082	7/0.095	7/0.082
16	Extended Maratos	4	28/0.267	18/0.182	28/0.267	26/0.246
		100	159/1.543	F/F	37/0.415	52/0.579
17	Six hump	2	4/0.065	4/0.066	4/0.061	4/0.056
18	Three hump	2	6/0.085	7/0.123	6/0.084	7/0.083
19	Booth	2	3/0.049	3/0.042	3/0.051	3/0.039
20	Trecanni	2	5/0.068	5/0.067	5/0.076	5/0.066
21	Zettl	2	16/0.169	9/0.109	16/0.173	9/0.109
22	Shallow	10	26/0.289	7/0.092	26/0.304	9/0.118
		1000	1216/30.119	14/0.369	1171/28.969	18/0.479
23	Generalized Quartic	100	6/0.092	6/0.095	6/0.091	6/0.088
		10000	6/1.211	6/1.217	6/1.1903	6/1.177
24	Quadratic QF2	10	875/8.867	31/0.341	746/7.476	35/0.382140/
		100	401/4.489	87/1.042	359/3.971	1.683
25	Leon	2	1681/14.026	43/0.399	557/4.697	36/0.343
26	Generalized Tridiagonal 1	10	27/0.326	23/0.274	27/0.321	23/0.274
		100	151/2.339	25/0.403	26/0.422	24/0.377
27	Generalized Tridiagonal 2	10	42/0.473	36/0.411	42/0.485	33/0.378
28	POWER	10	20/0.207	21/0.218	20/0.217	131/1.296

Table 1: Continued.

No.	Function	DIM	FR IN/CPU	PRP IN/CPU	CD IN/CPU	NMFR IN/CPU
29	Quadratic QF1	10	10/0.127	10/0.125	10/0.119	25/0.285
30	Extended quadratic penalty QP2	1000	F/F	41/2.460	122/9.486	73/3.679
31	Extended quadratic penalty QP1	100	314/3.803	F/F	14/0.209	14/0.207
32	Quartic	4	223/2.513	615/6.911	223/2.510	260/2.960
33	Colville	4	34/0.369	74/0.775	34/0.3661	59/0.626
34	Dixon and Price	10	74/0.906	42/0.533	74/0.897	42/0.531
		100	F/F	432/5.639	F/F	998/12.978
35	Extended Wood	4	F/F	267/2.979	F/F	208/2.350

convergent, that is,

$$\lim_{k\longrightarrow\infty}\|g_k\| = 0,\quad \text{or} \sum_{k=1}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{43}$$

Proof. The proof will be conducted by contradiction argument. So, assume that the statement of Theorem 7 is not true; then, there exists a constant $\epsilon > 0$, such that

$$\|g_k\| \geq \epsilon. \tag{44}$$

Equation (3) can be written as

$$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k, \tag{45}$$

and multiplying (45) on both sides by d_{k+1}, we obtain

$$\|d_{k+1}\|^2 = -g_{k+1}^T d_{k+1} + \beta_{k+1} d_{k+1}^T d_k. \tag{46}$$

Dividing both sides of (46) by $\|g_{k+1}\|^4$ with the help of (13), we get

$$\begin{aligned}\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} &\leq -\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^4} + \frac{\|g_{k+1}\|^2}{\theta\|g_k\|^2}\frac{d_{k+1}^T d_k}{\|g_{k+1}\|^4}\\ &\leq \left|-\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^4}\right| + \frac{\|g_{k+1}\|^2}{\theta\|g_k\|^2}\left|\frac{d_{k+1}^T d_k}{\|g_{k+1}\|^4}\right|.\end{aligned} \tag{47}$$

From Cauchy-Schwartz inequality, we get

$$\begin{aligned}\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} &\leq \frac{\|g_{k+1}\|\|d_{k+1}\|}{\|g_{k+1}\|^4} + \frac{\|d_k\|}{\theta\|g_k\|^2}\frac{\|d_{k+1}\|}{\|g_{k+1}\|^2},\\ &\leq \frac{\|g_{k+1}\|\|d_{k+1}\|}{\|g_{k+1}\|^4} - \frac{1}{2}\left(\frac{\|d_k\|}{\theta\|g_k\|^2} - \frac{\|d_{k+1}\|}{\|g_{k+1}\|^2}\right)^2\\ &\quad + \frac{1}{2}\frac{\|d_k\|^2}{\theta^2\|g_k\|^4} + \frac{1}{2}\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4}, \leq \frac{\|g_{k+1}\|\|d_{k+1}\|}{\|g_{k+1}\|^4}\\ &\quad + \frac{1}{2}\frac{\|d_k\|^2}{\theta^2\|g_k\|^4} + \frac{1}{2}\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4}.\end{aligned} \tag{48}$$

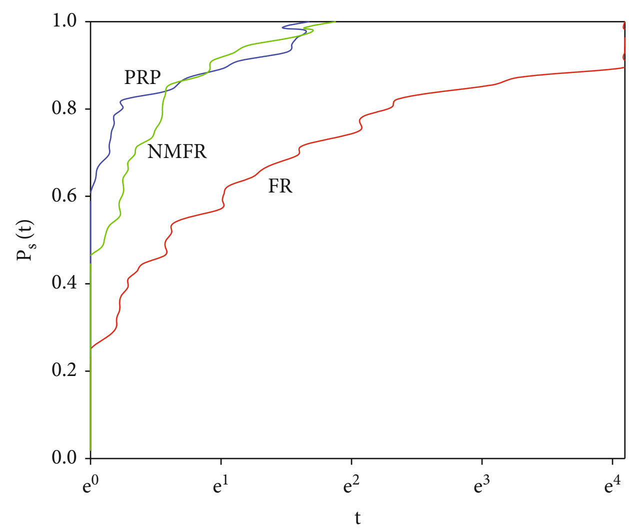

Figure 1: Performance profile according to the iterations number based on exact line search.

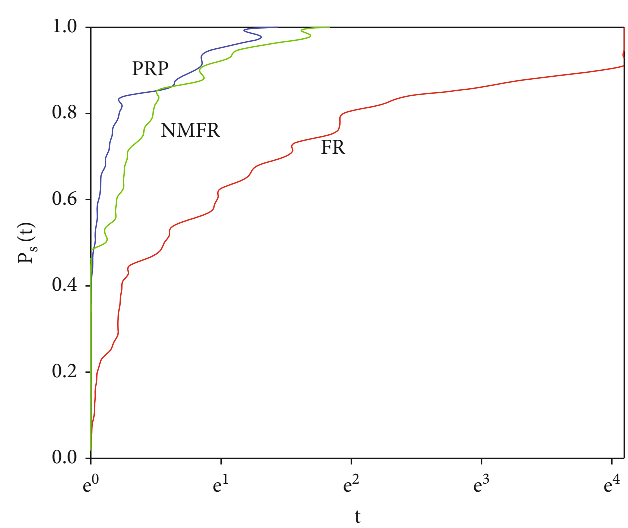

Figure 2: Performance profile based on the CPU time based on exact line search.

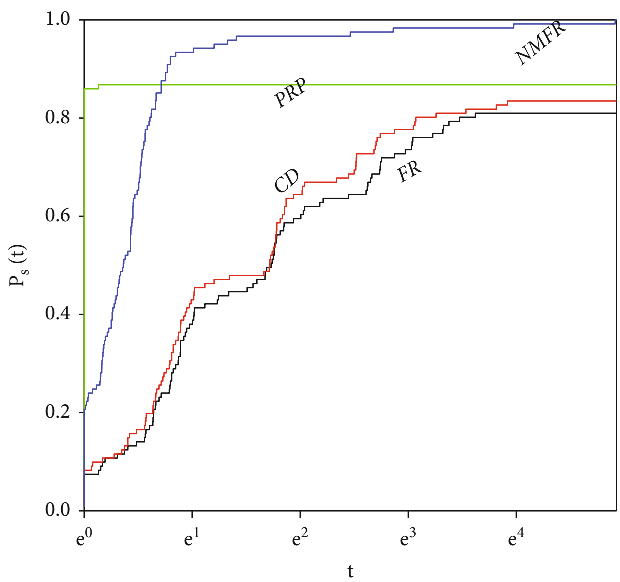

Figure 3: Performance profile according to the iterations number based on strong Wolfe line search.

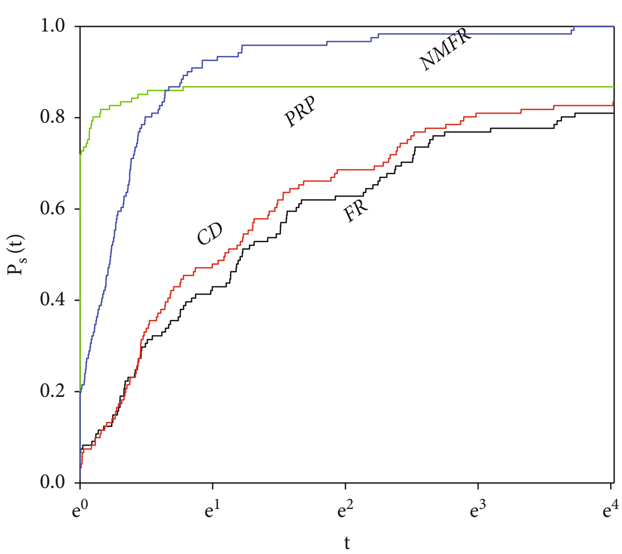

Figure 4: Performance profile based on the CPU time based on strong Wolfe line search.

So we come to

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4}\left(\frac{1}{2}-\frac{\|g_{k+1}\|}{\|d_{k+1}\|}\right) \leq \frac{1}{2}\frac{\|d_k\|^2}{\theta^2\|g_k\|^4}. \tag{49}$$

Referring to (39), using Cauchy-Schwartz inequality, we get

$$c\|g_{k+1}\|^2 \leq -g_{k+1}^T d_{k+1} \leq \|g_{k+1}\|\|d_{k+1}\|. \tag{50}$$

Then,

$$\frac{\|g_{k+1}\|}{\|d_{k+1}\|} \leq \frac{1}{c}. \tag{51}$$

Substituting (51) into (49), we get

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \leq \frac{1}{2}\frac{\|d_k\|^2}{\gamma\theta^2\|g_k\|^4}, \tag{52}$$

where $\gamma = 1/2 - 1/c$.

Recursively using (52) and noticing that $\|d_1\|^2 = \|g_1\|^2$, we have

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \leq \frac{1}{2}\left(\frac{1}{\gamma\theta^2}\right)\sum_{i=0}^{k}\frac{1}{\|g_i\|^2}, \tag{53}$$

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \leq \frac{1}{2}\left(\frac{1}{\gamma\theta^2}\right)\frac{k}{\epsilon^2} = \frac{k}{2\epsilon^2\theta^2\gamma}, \tag{54}$$

$$\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \geq \frac{2\gamma\theta^2\epsilon^2}{k}. \tag{55}$$

As a result of (53) and (44), it can be concluded that

$$\sum_{k=0}^{\infty}\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} = \infty, \tag{56}$$

which contradicts (43); hence, the proof is completed. □

4. Numerical Results and Discussion

Most of the test problems used in this study are taken from Andrei [8], and they are used to evaluate the efficiency of the NMFR method to that of FR and PRP under exact line search and to that of FR, PRP, and CD under strong Wolfe line search. The step-size is computed using the exact and strong Wolfe line search techniques, and numerical results are compared based on the number of iterations and CPU time. For all test problems, the stopping criteria are set to be $\|g_k\| \leq \epsilon$, where $\epsilon = 10^{-6}$; for each test problem, various starting points are used, as suggested by Hillstrom [26]. All runs are performed on a PC ACER (Intel® Core™ i3-3217u CPU @ 1.8 GHZ, with 4.00 GB RAMS, Windows 7 Ultimate). Every problem mentioned in Table 1 is solved using Matlab10 subroutine programming. The performance results are shown in Figures 1–4, respectively, using a performance profile introduced by Dolan and Jorge [13].

In the list of problem functions, *IN* indicates the number of iterations, and *CPU* indicates the CPU time as shown in Table 1. *F* indicates that the test problem function is failing. Also, in some cases, it means that the computation came to a halt when a line search failed to locate a positive step-size, and it was deemed a failure.

We offer the concept of a way of evaluating and examine the effectiveness of set solvers s on a test set p using the performance profile. Assuming there are n_s solvers and n_p

problems; they characterized $t_{p,s}$ as the computing time (computational time, CPU time, or other factors) needed to tackle problems p by solver s. They used the performance ratio $r_{p,s} = t_{p,s}/\min\{t_{p,s} : s \in S\}$ to compare solver s performance on problem p to the best performance by any solver on this problem.

We let $r_M \geq r_{p,s}$ for each chosen p and s such that $r_{p,s} = r_M$, whenever s is not a solution of problem p. The solution for performance s of presented problems has to be reliable, and we wish to have the entire evaluation of solution for performance s. Thus, we define $P(t)_s = (1/n_p)\text{size}\{p \in P : r_{p,s} \leq t\}$, where $P(t)_s$ stands for the probability of solution for performance $s \in S$ and P_s is a cumulative distribution function for $r_{p,s}$. The value $P(1)_s$ is the probability that the solver will win over the rest of the solvers. In all, a solver with high values of $P(t)_s$ or at the top right of the figures is preferable or represents robust solver.

Clearly from Figures 1 and 2, we see that NMFR has better performance, since it solved all the test problems and reached 100% percentage. These CG coefficients could also be divided into three categories: the first category of which consists of NMFR, while the second consists of FR, and the third consists of PRP. It is easy to see that their performance is much better. Although the performance of the third category seems to be much better than NMFR, it could only solve 86% of the problems, whereas the performance of the second category only reached 89%. Hence, we considered NMFR as an efficient performance and robust method with the others because it can solve all the problems.

Also, Figures 3 and 4 show that the curve of NMFR is higher than that of PRP, CD, and FR. This implies that the NMFR approach outperforms the other three methods significantly. Furthermore, the NMFR approach solves all problems; meanwhile, the PRP method solves about 86 percent of problems, and the CD and FR methods solve about 89 percent. As a result, we can infer that NMFR is the preferred approach because it has the highest curve and solves all problems.

5. Conclusion

In this article, we have proposed a new and simple modification for β_k^{FR} that is easy to implement, known as β_k^{NMFR}. Numerical results show that β_k^{NMFR} has efficient performance compared to other standard CG methods. In contrast to β_k^{FR}, we have seen that β_k^{NMFR} shows good numerical performance at each step. We have also proved that β_k^{NMFR} converges globally based on the exact and strong Wolfe line searches.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. H. Dai and Y. Yaxiang, “A nonlinear conjugate gradient method with a strong global convergence property,” *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.

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

[3] R. Fletcher and M. R. Colin, “Function minimization by conjugate gradients,” *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.

[4] Y. Liu and C. Storey, “Efficient generalized conjugate gradient algorithms, part 1: theory,” *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991.

[5] 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. 409–435, 1952.

[6] M. Rivaie, M. Mamat, W. J. Leong, and I. Mod, “A new class of nonlinear conjugate gradient coefficients with global convergence properties,” *Applied Mathematics and Computation*, vol. 218, no. 22, pp. 11323–11332, 2012.

[7] M. Al-Baali, “Descent property and global convergence of the Fletcher—Reeves method with inexact line search,” *IMA Journal of Numerical Analysis*, vol. 5, no. 1, pp. 121–124, 1985.

[8] N. Andrei, “An unconstrained optimization test functions collection,” *Advanced Modeling and Optimization*, vol. 10, no. 1, pp. 147–161, 2008.

[9] N. Andrei, “Open problems in nonlinear conjugate gradient algorithms for unconstrained optimization,” *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 34, no. 2, 2011.

[10] A. Abdelrahman, O. O. O. Yousif, M. Mhammed, and M. K. Elbashir, “Global convergence of nonlinear conjugate gradient coefficients with inexact line search,” *The Scientific Journal of King Faisal University: Basic and Applied Sciences*, vol. 22, no. 2, pp. 86–89, 2021.

[11] Y. H. Dai and Y. Yaxiang, *Nonlinear Conjugate Gradient Methods*, Shanghai Science and Technology Publisher, Shanghai, 2000.

[12] Y. Dai, “Analyses of conjugate gradient methods (Ph.D. thesis),” in *Institute of Computational Mathematics and Scientific/Engineering Computing*, Chinese Academy of Sciences, 1997.

[13] E. D. Dolan and J. M. Jorge, “Benchmarking optimization software with performance profiles,” *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.

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

[15] A. A. Goldstein, “On steepest descent,” *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, vol. 3, no. 1, pp. 147–151, 1965.

[16] Z. J. Shi and G. Jinhua, “A new family of conjugate gradient methods,” *Journal of Computational and Applied Mathematics*, vol. 224, no. 1, pp. 444–457, 2009.

[17] J. K. Liu and S. J. Li, “New hybrid conjugate gradient method for unconstrained optimization,” *Applied Mathematics and Computation*, vol. 245, pp. 36–43, 2014.

[18] L. H. J. Guanghui and H. Yin, “Global convergence of the Fletcher-Reeves algorithm with inexact linesearch,” *Applied Mathematics-A Journal of Chinese Universities*, vol. 10, no. 1, pp. 75–82, 1995.

[19] M. J. D. Powell, “Restart procedures for the conjugate gradient method,” *Mathematical Programming*, vol. 12, no. 1, pp. 241–254, 1977.

[20] M. J. Powell, “Nonconvex minimization calculations and the conjugate gradient method,” in *Numerical Analysis*, pp. 122–141, Springer, Berlin, Heidelberg, 1984.

[21] E. Polak and G. Ribiere, “Note sur la convergence de méthodes de directions conjuguées,” *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, vol. 3, no. 16, pp. 35–43, 1969.

[22] G. Zoutendijk, “Nonlinear programming computational methods,” in *Integer and Nonlinear Programming*, J. Abadie, Ed., pp. 37–86, North Holland, Amsterdam, 1970.

[23] H. Nosratipour and A. Keyvan, “A descent PRP conjugate gradient method for unconstrained optimization,” *TWMS Journal of Applied and Engineering Mathematics*, vol. 9, no. 3, pp. 535–548, 2019.

[24] L. W. Z. Zhang and L. Donghui, “Global convergence of a modified Fletcher–Reeves conjugate gradient method with Armijo-type line search,” *Numerische Mathematik*, vol. 104, no. 4, pp. 561–572, 2006.

[25] M. Rivaie, M. Mamat, and A. Abashar, “A new class of nonlinear conjugate gradient coefficients with exact and inexact line searches,” *Applied Mathematics and Computation*, vol. 268, no. n/a, pp. 1152–1163, 2015.

[26] K. E. Hillstrom, “A simulation test approach to the evaluation of nonlinear optimization algorithms,” *ACM Transactions on Mathematical Software (TOMS)*, vol. 3, no. 4, pp. 305–315, 1977.