Global Convergence of Modified RMIL Parameter in Conjugate Gradient Method

Laily Dwi Retno Wahyuningtias and Salmah

Abstract—**The conjugate gradient method provides a straightforward approach to solve unconstrained optimization problems. In this paper, we propose a modified formula for the conjugate gradient parameter by adapting the RMIL method. The proposed method ensures sufficient descent and global convergence under certain assumptions, utilizing an exact line search. Experimental results validate the efficiency of the modified method.**

Index Terms—**conjugate, gradient, parameter, global, convergent.**

I. Introduction

THE arrangements for different optimization issues are often crucial for researchers in engineering and scientific fields. A straightforward optimization strategy is the conjugate gradient method. For example, Cao et al. [1] defined the image restoration problem as a large-scale optimization issue. In addition, in Abubakar et al. [2], the signal repair problem is solved by formulating an optimization problem. Helmig et al. [3] create an optimization problem to estimate the distance and the number of sensors in the inverse calculation of thermal boundary conditions. Furthermore, Nicolaide [4] derived the formula for solving systems of linear equations with the conjugate gradient method and others (see [5], [6], and [7]). There are two types of optimization problems in their application: constrained and unconstrained.

The method described in this article is the conjugate gradient method. This method is convenient for solving unconstrained optimization problems because it requires less memory and is easier to compute. In this method, the direction of finding the solution of an iteration is determined by the objective function gradient, the conjugate gradient parameter, and the search direction of the preceding iteration. The development of this method is very diverse, especially in terms of modifying the conjugate gradient parameter. If the parameters of the conjugate gradients are different, the conjugate gradient method and its properties will also differ.

The first conjugate gradient parameter introduced is the FR parameter [8]. Powell [9] further investigated the FR parameter and found that this method with exact line search can generate a small step length without significantly improving the optimal solution. Then, Polak and Ribiere [10] introduce the PR parameter. Numerically, the PR parameter has better performance than the FR parameter. Polak and Ribiere [10] have shown that the PR parameter, with an exact line search, gives globally convergent results for convex objective function. However, Powell [11] has also shown that this is not necessarily true for nonconvex functions. Additionally, Powell [11] found that the PR parameter, with an exact line search, can rotate indefinitely and does not approach the solution point. This behavior can occur when β_k is negative, so Powell [11] suggests that the conjugate gradient parameter should not be negative. Therefore, Gilbert and Nocedal [12] modified the PRP parameter to take the maximum value between 0 and the PRP parameter, obtaining global convergence results with an inexact line search.

Rivaie et al. [13] have modified the PR parameter by changing the numerator from the gradient norm of the objective function to the norm of the search direction, which is referred to as the RMIL parameter in the future. Generally, the RMIL parameter is better than the PR parameter because it can solve more test functions in optimization problems. However, the RMIL parameter is not necessarily negative, so Dai [14] modified the parameter so that if the RMIL parameter is negative, its value is set to zero. With this modification, a globally convergent method is obtained. For reliable references to research that have discussed modern conjugate gradient methods and produced meaningful results, please refer to [15], [16], [17] and [18].

This article presents a modified conjugate gradient parameter formula that is simpler and maintains nonnegative parameter values. This parameter combines the FR parameter, which exhibits good global convergence, and the RMIL parameter, which demonstrate good numerical performance. With this modified parameter, the method satisfies the requirements of sufficient descent and achieves global convergence. Additionally, the paper includes numerical computations and comparisons between methods.

The structure of this article is as follows: Section II provides definitions and assumptions related to this research. In Section III, a discussion of the modified conjugate gradient parameter and its algorithm is presented. Section IV contains a convergence analysis, including sufficient descent conditions and global convergence. In Section V, experimental results and comparisons between methods with modified and existing parameters are provided. Section VI concludes this work.

II. Preliminaries

In this study, we consider the unconstrained optimization problem stated as follows.

$$\min_{\mathrm{x}\in\mathbb{R}^n} f(\mathrm{x}), \tag{1}$$

with $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function with $\nabla f(\mathrm{x}) = g(\mathrm{x})$ available. The conjugate gradient method is an iterative

Manuscript received December 31, 2022; revised June 12, 2023.

Laily Dwi Retno Wahyuningtias is a postgraduate student of the Mathematics Department, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia (e-mail: laily.dwi.r@mail.ugm.ac.id)

Salmah is a professor of Mathematics Department, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia (Corresponding author to provide e-mail: syalmah@yahoo.com)

approach to find the answer to the problem (1), which is formulated by the following recursive computational formula:

$$\mathrm{x}_{i+1} = \mathrm{x}_i + \alpha_i \mathrm{d}_i, \quad \text{for } i = 0, 1, 2, ... \tag{2}$$

where x_i represents the solution in the $i-$th iteration, $\alpha_i > 0$ denotes the step length, and d_i represents the direction. The step length is determined through a one-dimensional search known as a line search. The most commonly used line search method is the exact line search, defined as:

$$f(\mathrm{x}_i + \alpha_i \mathrm{d}_i) = \min_{\alpha \geq 0} f(\mathrm{x}_i + \alpha \mathrm{d}_i). \tag{3}$$

The exact line search is employed due to its ability to determine an optimal step length [19]. Recent advancements in technology, such as faster processors and improved equipment, have alleviated the computational burden associated with the exact line search method, as demonstrated by Rivaie et al. in 2012 [13].

The search direction d_i is defined as follows:

$$\mathrm{d}_i = \begin{cases} -g_i, & i = 0 \\ -g_i + \beta_i \mathrm{d}_{i-1}, & i \geq 1, \end{cases} \tag{4}$$

where g_i represents the gradient of the function f at the point x_i and β_i is the conjugate gradient parameter. As mentioned in section I, several well-known conjugate gradient parameter formulas, including the Fletcher and Reeves (FR), Polak and Ribiere (PR), and Rivaie, Mustafa, Ismail, and Leong (RMIL) parameters, are presented below.

$$\beta_i^{FR} = \frac{g_i^T g_i}{\|g_{i-1}\|^2}, \tag{5}$$

$$\beta_i^{PRP} = \frac{g_i^T (g_i - g_{i-1})}{\|g_{i-1}\|^2}, \tag{6}$$

$$\beta_i^{RMIL} = \frac{g_i^T (g_i - g_{i-1})}{\|\mathrm{d}_{i-1}\|^2}. \tag{7}$$

where $\|.\|$ denotes the Euclidean norm. Equation (5)-(7) correspond to the Fletcher and Reeves (FR) in [8], the Polak and Ribiere (PR) in [10], and the Rivaie, Mustafa, Ismail, and Leong (RMIL) parameter, respectively.

In order to analyze the global convergence of the conjugate gradient method, two essential properties need to be considered: sufficient descent and global convergence [17]. The notion of minimizing the objective function is closely tied to the concept of descent steps, which implies that each search step should lead to a reduction in the cost of the function f. Consequently, the following definition establishes a condition that ensures the desired search direction vector corresponds to a descent direction for the function f.

Definition 2.1: An algorithm is considered to have sufficient descend if there exists a positive constant C such that for every search direction vector d_i the inequality

$$g_i^T \mathrm{d}_i \leq -C\|g_i\|^2, \tag{8}$$

holds for all $i \geq 0$.

Additionally, the algorithm's efficacy relies on ensuring the convergence of the sequence $\{\mathrm{x}_i\}$ for all $i = 0, 1, ...$ generated by the algorithm from any initial point to a stationary point of the function f.

Definition 2.2: An algorithm of the conjugate gradient method is said to be globally convergent if it satisfies the condition

$$\lim_{i \to \infty} \inf \|g_i\| = 0. \tag{9}$$

Several assumptions are necessary to investigate the global convergence of this method for the objective function considered below.

Assumption 2.1: The objective function f is bounded below and satisfies continuity and differentiability conditions in a neighborhood $\mathcal{B}$ of the level set $\Omega = \{\mathrm{x} \in \mathbb{R}^n | f(\mathrm{x}) \leq f(\mathrm{x}_0)\}$ at the initial point.

Assumption 2.2: The gradient function g is Lipschitz continuous in $\mathcal{B}$. Specifically, there exists a constant $K > 0$ such that $\|g(\mathrm{x}) - g(\mathrm{y})\| \leq K\|\mathrm{x} - \mathrm{y}\|$ for all $\mathrm{x}, \mathrm{y} \in \mathcal{B}$.

The existence of a lower bound for the objective function f is crucial for a well-defined optimization problem. The Lipschitz continuity of the gradient is a commonly assumed condition and is often satisfied in practice. It is implied by many of the smoothness conditions used in local convergence theorems [16].

Based on Assumption 2.1 and Assumption 2.2, a lemma known as the Zoutendijk condition is established. This lemma has wide-ranging implications. It demonstrates, for example, that the steepest descent algorithm achieves global convergence. It also provides insights into other algorithms by determining the extent to which d_i can deviate from the steepest descent direction while still ensuring global convergence of the iterations. Various line search termination criteria can be employed to establish this result, but in this study, we will focus on the exact line search.

Lemma 2.1: Assume that Assumption 2.1 and 2.2 hold. For any conjugate gradient method with (2)-(4), where α_i is determined by (3). The following condition holds

$$\sum_{i=0}^{\infty} \frac{(g_i^T \mathrm{d}_i)^2}{\|\mathrm{d}_i\|^2} < \infty. \tag{10}$$

This condition plays a crucial role in ensuring the convergence of the conjugate gradient method. The lemma mentioned above was originally proven by Zoutendijk in [23]. The assumptions regarding the sequence x_i and the objective function f are necessary to demonstrate the convergence rate of this method.

Assumption 2.3: Sequence $\{\mathrm{x}_{i+1}\}$, where $\mathrm{x}_{i+1} = \mathrm{x}_i + \alpha_i \mathrm{d}_i, \forall k = 0, 1, 2, 3, ..., \alpha_i$ determined by (3), d_i formulated by (4), for all $i = 0, 1, 2, 3...$ converges to x^*. Additionally, function f is a twice continuously differentiable on the neighborhood of $\mathcal{B}(\mathrm{x}^*, \epsilon_0) = \{\mathrm{x} | \|\mathrm{x} - \mathrm{x}^*\| < \epsilon_0\}$ for some $\epsilon_0 > 0$ and $\nabla^2 f(\mathrm{x}^*)$ is a symmetric and positive definite matrix.

The rate of convergence of an algorithm is a crucial performance indicator. There are two types of convergence rates to consider:

Definition 2.3: Let $\{\mathrm{x}_i\}$ be a sequence in $\mathbb{R}^n$ that converges to x^*. The sequence $\{\mathrm{x}_i\}$ is said to be Q-linear convergence if there exists a constant $r \in (0, 1)$ such that

$$\frac{\|\mathrm{x}_{i+1} - \mathrm{x}^*\|}{\|\mathrm{x}_i - \mathrm{x}^*\|} \leq r \tag{11}$$

$\forall i$ sufficiently large.

This indicates that the distance to the solution x^* decreases by at least a constant factor with each iteration. The ”Q” in Q-linear convergence represents the quotient, as this type of convergence is described in terms of the quotient of successive errors. On the other hand, ”R” denotes a slightly weaker form of convergence, referred to as R-linear convergence or root convergence, which focuses on the overall rate of error reduction throughout the entire algorithm rather than in each individual step.

Definition 2.4: Let $\{x_i\}$ be a sequence in $\mathbb{R}^n$ that converges to x^*. The sequence $\{x_i\}$ is said to be R-linearly convergent if there exists a nonnegative sequence of scalar $\{v_i\}$ such that

$$\|x_i - x^*\| \leq v_i \tag{12}$$

$\forall i$, and the sequence $\{v_i\}$ converges Q-linearly to zero.

The linear convergence rate of conjugate gradient methods is typically demonstrated by establishing the following result, as presented in [13].

Lemma 2.2: Assuming Assumption 2.3 holds true, and considering the angle θ_i between $-g_i$ and d_i, where x_i is determined by exact line search and d_i is a descent direction. If there exists a constant $\lambda > 0$ satisfying

$$\Pi_{k=0}^{i-1} \cos\theta_i \geq \lambda^i,$$

then, there exist constants $\alpha > 0$ and $r \in (0,1)$, such that

$$\|x_{i+1} - x^*\| \leq ar^{i+1}. \tag{13}$$

Hence, x_i converges to x^*, at least in an R-linearly convergent.

III. New Conjugation Gradient Parameter

This section proposes a modified conjugate gradient parameter formula, which combines the RMIL and FR parameters. The modified formula is given as follows:

$$\beta_i^{FRMIL} = \begin{cases} \beta_i^{RMIL}, & 0 \leq \beta_i^{RMIL} \leq \beta_i^{FR} \\ \beta_i^{FR}, & \text{other.} \end{cases} \tag{14}$$

The algorithm can be summarized as follows, incorporating equations (2), (3), (4), and (14)

1) Preparation. Given $x_0 \in \mathbb{R}^n$, $\varepsilon > 0$ and put $i = 0$.
2) Calculate $\|g_i\|$. If $\|g_i\| \leq \varepsilon$, then x_i is a solution point. If $\|g_i\| > \varepsilon$, move to step (3).
3) Calculate β_i based on the modified conjugate parameter formula (14).
4) Calculate d_i based on (4).
5) Perform an exact line search to determine the step length α_i by (3).
6) Update $i = i + 1$, calculate the next step by equation (2), and move to step (2).

IV. Convergence Properties

In order to establish the conditions of sufficient descent and global convergence for the conjugate gradient method with the modified conjugate parameter and exact line search (3), the following properties play a crucial role in simplifying the analysis:

1) Sufficient Descent: It is important to ensure that the algorithm exhibits a descent direction, meaning that the objective function f decreases at each iteration. This property guarantees that the method is making progress towards a solution. The condition of sufficient descent ensures that the search direction d_i leads to a decrease in the function value, as stated in Definition 2.1.
2) Global Convergence: The algorithm's global convergence is essential to ensure that the sequence x_i generated by the method, starting from any initial point, converges to a stationary point of the function f. The definition of global convergence, as stated in Definition 2.2, requires that the norm of the gradient $\|g_i\|$ approaches zero as i tends to infinity.

By establishing the sufficient descent and global convergence properties, the investigation of the conjugate gradient method with the modified conjugate parameter and exact line search becomes more accessible, enabling the development of a practical and high-quality algorithm.

Lemma 4.1: For every $i \geq 0$, the condition

$$0 \leq \beta_i^{FRMIL} \leq \beta_i^{FR} \tag{15}$$

always holds.

Proof: We consider the modified conjugate parameter β_i^{FRMIL} defined in (14). From the definition, we have two cases:

Case 1: If $0 \leq \beta_i^{RMIL} \leq \beta_i^{FR}$, then according to (14), we have

$$\beta_i^{FRMIL} = \beta_i^{RMIL}.$$

Since both β_i^{RMIL} and β_i^{FR} are nonnegative, it follows that

$$0 \leq \beta_i^{FRMIL} \leq \beta_i^{FR}.$$

Case 2: If β_i^{RMIL} does not satisfy the condition $0 \leq \beta_i^{RMIL} \leq \beta_i^{FR}$, then according to (14), we have

$$\beta_i^{FRMIL} = \beta_i^{FR}.$$

Again, since both β_i^{FRMIL} and β_i^{FR} are nonnegative, we have

$$0 \leq \beta_i^{FRMIL} \leq \beta_i^{FR}.$$

Therefore, in both cases, the condition (15) is always satisfied for every $i \geq 0$. This completes the proof of Lemma 4.1. ■

The following theorem proves that the modified parameter fulfills the condition of sufficient descent.

Theorem 4.1: Given a conjugate gradient method with d_i determined by (4) and β_i^{FRMIL} formulated by (14), then

$$\nabla g_i^T d_i \leq -C\|g_i\|^2, \tag{16}$$

with $C > 0$.

Proof: We will prove that the conjugate gradient method with the modified conjugate gradient parameter satisfies (16). Note that

$$d_i = \begin{cases} -g_i, & i = 0 \\ -g_i + \beta_i^{FRMIL} d_{i-1}, & i \geq 1, \end{cases} \tag{17}$$

First, If $i = 0$, by (4) we have

$$g_0^T d_0 = -g_0^T g_0 = -\|g_0\|^2.$$

In other words, the inequality (16) is satisfied for $i = 0$. Furthermore, if $i \geq 1$, then multiplying (17) by g_i^T, we get

$$g_i^T d_i = g_i^T(-g_i + \beta_i^{FRMIL} d_{i-1}) \tag{18}$$

$$= -\|g_i\|^2 + \beta_i^{FRMIL} g_i^T d_{i-1}. \tag{19}$$

Using the exact line search method (3), we get $g_i^T \mathrm{d}_{i-1} = 0$. Consequently,

$$g_i^T d_i = -\|g_i\|^2,$$

means that d_{i+1} is a sufficient descent direction. Therefore, $g_i^T d_i \leq -C\|g_i\|^2$ holds. We have finished the proof. ∎

Next, we will establish the global convergence property of the conjugate gradient method with the FRMIL parameter by employing an exact line search. To analyze the global convergence of this method, we rely on the assumptions outlined in 2.1-2.2 and the use of Lemma 2.1.

Theorem 4.2: Given the conjugate gradient method with the modified conjugate parameter formulation (14) and the exact line search (3), and assuming that Assumption 2.1 and Assumption 2.2 are satisfied, along with the condition of sufficient descent, the following holds:

$$\lim_{i\to\infty} \inf \|g_i\| = 0. \tag{20}$$

Proof: Based on Assumption 2.1 and 2.2, Lemma 2.1 can be applied. We will now proceed to prove this by contradiction. Assuming that (20) is not true, it implies the existence of a positive constant c such that

$$\|g_i\| \geq c, \tag{21}$$

for i is quite large, which means

$$\frac{1}{\|g_i\|^2} \leq \frac{1}{c^2}. \tag{22}$$

Based on (4), for every $i \geq 1$,

$$d_{i+1} + g_{i+1} = \beta_{i+1}^{FRMIL} d_i.$$

Squaring both sides of the equation, for every $i \geq 1$, we get

$$\begin{aligned}
(d_{i+1} + g_{i+1})^2 &= (\beta_{i+1}^{FRMIL}\|d_i\|)^2 \\
\|d_{i+1}\|^2 + 2g_{i+1}^T d_{i+1} + \|g_{i+1}\|^2 &= {\beta_{i+1}^{FRMIL}}^2\|d_i\|^2 \\
\|d_{i+1}\|^2 &= {\beta_{i+1}^{FRMIL}}^2\|d_i\|^2 \\
&\quad -2g_{i+1}^T d_{i+1} \\
&\quad -\|g_{i+1}\|^2.
\end{aligned}$$

By dividing both sides of the equation by $(g_{i+1}^T d_{i+1})^2$, for every $i \geq 1$, we obtain

$$\begin{aligned}
\frac{\|d_{i+1}\|^2}{(g_{i+1}^T d_{i+1})^2} &= \frac{{\beta_{i+1}^{FRMIL}}^2\|d_i\|^2}{(g_{i+1}^T d_{i+1})^2} - \frac{2}{g_{i+1}^T d_{i+1}} \\
&\quad -\frac{\|g_{i+1}\|^2}{(g_{i+1}^T d_{i+1})^2} \\
&= \frac{{\beta_{i+1}^{FRMIL}}^2\|d_i\|^2}{({g_{i+1}}^T d_{i+1})^2} \\
&\quad -\left(\frac{1}{\|g_{i+1}\|} + \frac{\|g_{i+1}\|}{{g_{i+1}}^T d_{i+1}}\right)^2 \\
&\quad +\frac{1}{\|g_{i+1}\|^2} \\
&\leq \frac{{\beta_{i+1}^{FRMIL}}^2\|d_i\|^2}{({g_{i+1}}^T d_{i+1})^2} + \frac{1}{\|g_{i+1}\|^2}.
\end{aligned}$$

By inequality (15), for every $i \geq 1$ we get

$$\begin{aligned}
\frac{\|d_{i+1}\|^2}{({g_{i+1}}^T d_{i+1})^2} &\leq \frac{{\beta_{i+1}^{FR}}^2\|d_i\|^2}{({g_{i+1}}^T d_{i+1})^2} + \frac{1}{\|g_{i+1}\|^2} \\
&= \left(\frac{\|g_{i+1}\|^2}{\|g_i\|^2}\right)^2 \frac{\|d_i\|^2}{({g_{i+1}}^T d_{i+1})^2} \\
&\quad +\frac{1}{\|g_{i+1}\|^2} \\
&= \frac{\|g_{i+1}\|^4}{\|g_i\|^4}\frac{\|d_i\|^2}{\|g_{i+1}\|^4} + \frac{1}{\|g_{i+1}\|^2} \\
&= \frac{\|d_i\|^2}{\|g_i\|^4} + \frac{1}{\|g_{i+1}\|^2}.
\end{aligned} \tag{23}$$

Since $\frac{\|d_0\|^2}{({g_0}^T d_0)^2} = \frac{1}{\|g_0\|^2}$, then with (22), (23), and (4) we get

$$\begin{aligned}
\frac{\|d_{i+1}\|^2}{({g_{i+1}}^T d_{i+1})^2} &= \frac{\|d_{i+1}\|^2}{\|g_{i+1}\|^2} \\
&\leq \frac{\|d_i\|^2}{\|g_i\|^4} + \frac{1}{\|g_{i+1}\|^2} \\
&\leq \frac{\|d_{i-1}\|^2}{\|g_{i-1}\|^4} + \frac{1}{\|g_i\|^2} + \frac{1}{\|g_{i+1}\|^2} \\
&\leq \frac{\|d_{i-2}\|^2}{\|g_{i-2}\|^4} + \frac{1}{\|g_{i-1}\|^2} + \frac{1}{\|g_i\|^2} \\
&\quad +\frac{1}{\|g_{i+1}\|^2} \\
&\quad \vdots \\
&\leq \sum_{k=0}^{i+1} \frac{1}{\|g_k\|^2} \\
&\leq \frac{i+1}{c^2}.
\end{aligned}$$

Thus,

$$\frac{(g_{i+1}^T d_{i+1})^2}{\|d_{i+1}\|^2} \geq \frac{c^2}{i+1}.$$

By taking the sum of both sides, we have

$$\sum_{i=0}^{\infty} \frac{(g_{i+1}^T d_{i+1})^2}{\|d_{i+1}\|^2} \geq c^2 \sum_{i=0}^{\infty} \frac{1}{i+1}.$$

Note that $\sum_{i=0}^{\infty} \frac{1}{i+1}$ is a harmonic series, and harmonic series diverge. Thus,

$$\sum_{i=0}^{\infty} \frac{(g_{i+1}^T d_{i+1})^2}{\|d_{i+1}\|^2} \geq \infty.$$

The statement above contradicts Lemma 2.1. Therefore, the proof is complete. ∎

Now, the corollary below suggests that when the sum of the squares of the search direction norms is zero, the zoutendijk condition holds, indicating a certain property or behavior of the optimization method being analyzed.

Corollary 4.1: If $\sum_{i=0}^{\infty} \|\mathrm{d}_i\|^2 = 0$, then $\sum_{i=0}^{\infty} \frac{(g_i^T \mathrm{d}_i)^2}{\|\mathrm{d}_i\|^2} < \infty$ holds.

Proof: From the exact line search method and the search direction (4), we have

$$\begin{aligned}
\|g_i\|^2 &= g_i^T g_i \\
&= g_i^T(-\mathrm{d}_i + \beta_i \mathrm{d}_{i-1}) \\
&= -g_i^T \mathrm{d}_i \\
&= -(-\mathrm{d}_i^T + \beta_i \mathrm{d}_{i-1})^T \mathrm{d}_i \\
&= \|\mathrm{d}_i\|^2 - \beta_i \mathrm{d}_{i-1}^T \mathrm{d}_i.
\end{aligned}$$

Then,

$$\begin{aligned}\|g_i\|^4 &= (\|g_i\|^2)^2\\ &= (\|\mathrm{d}_i\|^2 - \beta_i \mathrm{d}_{i-1}^T \mathrm{d}_i)^2\\ &= \|\mathrm{d}_i\|^4 - 2\beta_i \|\mathrm{d}_i\|^2 \mathrm{d}_{i-1}\mathrm{d}_i + \beta_i^2 (\mathrm{d}_{i-1}^T \mathrm{d}_i)^2\\ &= \|\mathrm{d}_i\|^4 - 2\beta_i^2 \|\mathrm{d}_i\|^2 \|\mathrm{d}_{i-1}\|^2 + \beta_i^4 \|\mathrm{d}_{i-1}\|^4 \quad (24)\end{aligned}$$

We will prove by contradiction. Assume that $\|g_i\| \geq m$ and $\sum_{i=0}^{\infty} \|\mathrm{d}_i\|^2 = \infty$. For $\|g_i\| \to \infty$, then $\frac{1}{\|g_i\|} \to 0$. Based on (23), we have

$$\frac{\|\mathrm{d}_i\|^2}{(g_i^T \mathrm{d}_i)^2} \leq \frac{\|\mathrm{d}_i\|^2}{\|g_i\|^4},$$

then, from (24), we have

$$\begin{aligned}\frac{\|\mathrm{d}_i\|^2}{(g_i^T \mathrm{d}_i)^2} &\leq \frac{\|\mathrm{d}_i\|^2}{\|g_i\|^4}\\ &= \frac{\|\mathrm{d}_i\|^2}{\|\mathrm{d}_i\|^4 - 2\beta_i^2 \|\mathrm{d}_i\|^2 \|\mathrm{d}_{i-1}\|^2 + \beta_i^4 \|\mathrm{d}_{i-1}\|^4}\\ &\leq \frac{\|\mathrm{d}_i\|^2}{\|\mathrm{d}_i\|^4 - 2\beta_i^2 \|\mathrm{d}_i\|^2 \mathrm{d}_{i-1}^T \mathrm{d}_i}\end{aligned}$$

This will imply that

$$\begin{aligned}\frac{(g_{i+1}^T d_{i+1})}{\|\mathrm{d}_{i+1}\|^2} &\geq \frac{\|\mathrm{d}_i\|^4 - 2\beta_i^2 \|\mathrm{d}_i\|^2 \|\mathrm{d}_{i-1}\|^2}{\|d_i\|^2}\\ &= \|\mathrm{d}_i\|^2 - 2\beta_i^2 \|\mathrm{d}_{i-1}\|^2\end{aligned}$$

Based on Lemma 4.1 and (24), we have

$$\begin{aligned}\frac{(g_{i+1}^T d_{i+1})}{\|\mathrm{d}_{i+1}\|^2} &\geq \|\mathrm{d}_i\|^2 - 2\beta_i^2 \|\mathrm{d}_{i-1}\|^2\\ &\geq \|\mathrm{d}_i\|^2 - 2\frac{\|g_i\|^4}{\|g_{i-1}\|^4} \|\mathrm{d}_{i-1}\|^2\\ &= \|\mathrm{d}_i\|^2 - 2\frac{\|\mathrm{d}_i\|^4}{\|g_{i-1}\|^4} \|\mathrm{d}_{i-1}\|^2\\ &\quad + \frac{4\beta_i^2 \|\mathrm{d}_i\|^2 \|\mathrm{d}_{i-1}\|^4}{\|g_{i-1}\|^4} - \frac{2\beta_i^4 \|\mathrm{d}_{i-1}\|^6}{\|g_{i-1}\|^4}\end{aligned}$$

Since $\frac{1}{\|g_i\|} \to 0$, then

$$\frac{(g_{i+1}^T d_{i+1})}{\|\mathrm{d}_{i+1}\|^2} \geq \|\mathrm{d}_i\|, \tag{25}$$

which leads to

$$\sum_{i=0}^{\infty} \frac{(g_i^T \mathrm{d}_i)^2}{\|\mathrm{d}_i\|^2} \geq \sum_{i=0}^{\infty} \|\mathrm{d}_i\|^2, \tag{26}$$

and

$$\sum_{i=0}^{\infty} \frac{(g_i^T \mathrm{d}_i)^2}{\|\mathrm{d}_i\|^2} \geq \infty. \tag{27}$$

This conclusion contradicts Lemma 2.1. Therefore, the corollary holds. ■

Next, we will demonstrate that the rate of convergence of the conjugate gradient method with the FRMIL parameter is linear.

Theorem 4.3: Suppose Assumption 2.3 is satisfied. If the sequence $\{\mathrm{x}_i\}$ is generated using exact line search, d_i is a descent direction, and β_i is formulated according to (14), then there exists a constant $a > 0$ and $r \in (0, 1)$, such that

$$\|\mathrm{x}_i - \mathrm{x}^*\| \leq a r^i.$$

Thus, x_i converges to x^* R-linearly.

Proof: If Assumption 2.1-2.3 hold true, then we assume $\forall \mathrm{x}_0 \in N(x^*, \varepsilon)$. Hence, from (4) and (15) we have

$$\begin{aligned}\|\mathrm{d}_i\| &\leq \|g_i\| + \beta_i^{FRMIL} \|\mathrm{d}_{i-1}\|\\ &\leq \|g_i\| + \beta_i^{FR} \|\mathrm{d}_{i-1}\|\\ &\leq \|g_i\| + \frac{\|g_i\|^2}{\|g_{i-1}\|^2} \|\mathrm{d}_{i-1}\|\\ &= \big(1 + \frac{\|g_i\| \|\mathrm{d}_{i-1}\|}{\|g_{i-1}\|^2}\big) \|g_i\|.\end{aligned}$$

This will imply that

$$\cos\theta_i = \frac{-g_i^T \mathrm{d}_i}{\|g_i\| . \|\mathrm{d}_i\|} = \frac{\|g_i\|^2}{\|g_i\| \|\mathrm{d}_i\|} \geq \left(1 + \frac{\|g_i\| \|\mathrm{d}_{i-1}\|}{\|g_{i-1}\|^2}\right)^{-1}.$$

By Lemma 2.2 we get Theorem 4.3. The proof is complete. ■

V. Discussions

This section presents the experimental results of the conjugate gradient method with FR, PRP, RMIL, and FRMIL parameters to demonstrate the efficiency of each method. In this study, artificial problems are selected from 128 test functions that range from small to large scale, namely 2, 4, 10, 50, 100, and 500, as provided in Table I. NP is used to symbolize the number of each problem. Several test functions are taken from Andrei [21]. A large-scale problem is included to avoid the algorithm from being biased towards a specific function [21]. A random starting point is assigned for each of the selected test functions to assess the properties of global convergence and the success of the method.

TABLE I: A list of the test function

NP	Function	n	starting point
1	Three-hump	2	(-1,1)
2	Three-hump	2	(1,-1)
3	Three-hump	2	(-2,2)
4	Goldstein-Price	2	(2,-2)
5	Goldstein-Price	2	(9,9)
6	Goldstein-Price	2	(15,15)
7	Zettl	2	(5,5)
8	Zettl	2	(10,10)
9	Zettl	2	(20,20)
10	Rosenbrock	2	(-2,-2)
11	Rosenbrock	2	(2,2)
12	Rosenbrock	2	(11,11)
13	Quartic	4	(2,...,2)
14	Quartic	4	(5,...,5)
15	Quartic	4	(10,...,10)
16	Extended Maratos	2	(8,8)
17	Extended Maratos	2	(22,22)
18	Extended Maratos	2	(44,44)
19	Extended Maratos	4	(8,...,8)
20	Extended Maratos	4	(22,...,22)
21	Extended Maratos	4	(44,...,44)

(Continued on the next page)

TABLE I: *Continued*

NP	Function	n	starting point
22	Extended White and Holst	4	(2,...,2)
23	Extended White and Holst	4	(3,...,3)
24	Extended White and Holst	4	(-2,...,-2)
25	Extended White and Holst	10	(2,...,2)
26	Extended White and Holst	10	(3,...,3)
27	Extended White and Holst	10	(-2,...,-2)
28	Extended Frudenstein & Roth	4	(3,...,3)
29	Extended Frudenstein & Roth	4	(5,...,5)
30	Extended Frudenstein & Roth	4	(10,...,10)
31	Extended Frudenstein & Roth	100	(3,...,3)
32	Extended Frudenstein & Roth	100	(5,...,5)
33	Extended Frudenstein & Roth	100	(10,...,10)
34	Beale	2	(2,2)
35	Beale	2	(4,4)
36	Beale	2	(6,6)
37	Beale	4	(2,...,2)
38	Beale	4	(4,...,4)
39	Beale	4	(6,...,6)
40	Beale	10	(2,...,2)
41	Beale	10	(4,...,4)
42	Beale	10	(6,...,6)
43	Raydan1	2	(-1,-1)
44	Raydan1	2	(2,2)
45	Raydan1	4	(-1,...,-1)
46	Raydan1	4	(1,...,1)
47	Raydan1	4	(2,...,2)
48	Raydan1	10	(-1,...,-1)
49	Raydan1	10	(1,...,1)
50	Raydan1	10	(2,...,2)
51	Liarwhd	2	(3,3)
52	Liarwhd	2	(5,5)
53	Liarwhd	2	(7,7)
54	Liarwhd	4	(3,...,3)
55	Liarwhd	4	(5,...,5)
56	Liarwhd	4	(7,...,7)
57	Liarwhd	10	(3,...,3)
58	Liarwhd	10	(5,...,5)
59	Liarwhd	10	(7,...,7)
60	Fletcher	2	(5,5)
61	Fletcher	2	(10,10)
62	Fletcher	2	(40,40)
63	Fletcher	4	(5,...,5)
64	Fletcher	4	(10,...,10)
65	Fletcher	4	(40,...,40)
66	Fletcher	10	(5,...,5)
67	Fletcher	10	(10,...,10)
68	Fletcher	10	(40,...,40)
69	Edencsch	2	(3,3)

(Continued on the other side of this page)

TABLE I: *Continued*

NP	Function	n	starting point
70	Edencsch	2	(23,23)
71	Edencsch	2	(43,43)
72	Edencsch	4	(3,...,3)
73	Edencsch	4	(23,...,23)
74	Edencsch	4	(43,...,43)
75	Edencsch	10	(3,...,3)
76	Edencsch	10	(23,...,23)
77	Edencsch	10	(43,...,43)
78	Generalized Quartic	2	(1,1)
79	Generalized Quartic	2,	(10,10)
80	Generalized Quartic	2	(20,20)
81	Generalized Quartic	4	(1,...,1)
82	Generalized Quartic	4	(10,...,10)
83	Generalized Quartic	4	(20,...,20)
84	Generalized Quartic	100	(1,...,1)
85	Generalized Quartic	100	(10,...,10)
86	Generalized Quartic	100	(20,...,20)
87	Extended Denschnf	2	(2,2)
88	Extended Denschnf	2	(13,13)
89	Extended Denschnf	4	(2,..,2)
90	Extended Denschnf	4	(50,..,50)
91	Extended Denschnf	100	(2,..,2)
92	Extended Denschnf	100	(13,...,13)
93	Extended Denschnf	100	(50,..,50)
94	Extended Denschnb	2	(4,4)
95	Extended Denschnb	2	(8,8)
96	Extended Denschnb	2	(15,15)
97	Extended Denschnb	4	(4,...,4)
98	Extended Denschnb	4	(8,...,8)
99	Extended Denschnb	4	(15,...,15)
100	Extended Denschnb	100	(4,...,4)
101	Extended Denschnb	100	(8,...,8)
102	Extended Denschnb	100	(15,...,15)
103	Himmelblau	2	(15,15)
104	Himmelblau	2	(25,25)
105	Himmelblau	2	(35,35)
106	Himmelblau	10	(15,...,15)
107	Himmelblau	10	(25,...,25)
108	Himmelblau	10	(35,...,35)
109	Himmelblau	100	(15,...,15)
110	Himmelblau	100	(25,...,25)
111	Himmelblau	100	(35,...,35)
112	Extended Penalty	2	(2,2)
113	Extended Penalty	2	(5,5)
114	Extended Penalty	2	(10,10)
115	Extended Penalty	10	(2,...,2)
116	Extended Penalty	10	(5,...,5)
117	Extended Penalty	10	(10,...,10)

(Continued on the next page)

TABLE I: *Continued*

NP	Function	n	starting point
118	Extended Penalty	100	(2,...,2)
119	Extended Penalty	100	(5,...,5)
120	Extended Penalty	100	(10,...,10)
121	Tridiagonal1	2	(5,5)
122	Tridiagonal1	2	(7,7)
123	Tridiagonal1	10	(5,...,5)
124	Tridiagonal1	10	(7,...,7)
125	Tridiagonal1	10	(15,...,15)
126	Tridiagonal1	500	(5,...,5)
127	Tridiagonal1	500	(7,...,7)
128	Tridiagonal1	500	(15,...,15)

The test functions in Table I were generated by Matlab R2020a and executed on a laptop with the following specifications: Intel(R) Celeron(R) processor, 4.00 GB RAM, 64-bit Windows 10 Operating System Home Single Language. The iteration stopping criterion is based on condition $\|g_i\| \leq \varepsilon$ with ε set to 10^{-6}. If the step length produces nonpositive results, the experiment result is labeled as "fail," as indicated in some of the experimental results in Table II. Experimental results are compared based on the number of iterations (NI) and CPU time (CT), which are summarized in Table III. These experimental results indicate the performance profile based on the number of iterations and CPU time, as shown in Fig. 1 and Fig. 2. The performance representation curve in Fig. 1 and Fig. 2 follows the approach introduced by Dolan and More [22].

We use the performance profile introduced by Dolan and More in [22] to investigate the capability of every method. The performance profile is determined using the following formula. Let S denote the set of n_S methods, and P denote the set of n_P test functions. For every method $s \in S$ and each function $p \in P$, let $t_{p,s}$ represent the number of iterations or CPU time required by method s to solve function p. Then the performance ratio for comparing methods is given by

$$r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in S\}}.$$

Take the number $r_M \geq r_{p,s}$ for every p, s and set $r_{p,s} = r_M$ if method s does not resolve p.

TABLE II: Experiment results

NP	FRMIL		FR		PRP		RMIL	
	NI	CT	NI	CT	NI	CT	NI	CT
1	10	1.02	1590	76.2	6	0.55	13	0.95
2	10	0.86	1590	75.5	13	0.53	6	0.89
3	7	0.77	7	0.63	5	0.52	6	0.58
4	10	1.34	10	1.31	11	1.23	10	1.28
5	12	1.44	12	1.28	14	1.38	18	1.70
6	9	1.02	11	1.14	10	1.06	11	1.20
7	13	0.66	16	0.80	9	0.53	18	0.90

(Continued on the other side of this page)

TABLE II: *Continued*

NP	FRMIL		FR		PRP		RMIL	
	NI	CT	NI	CT	NI	CT	NI	CT
8	12	0.63	10	0.55	11	0.5	19	0.88
9	14	0.77	20	0.95	9	0.53	19	0.90
10	50	3.00	37	1.78	23	1.23	63	3.80
11	37	1.45	fail	fail	fail	fail	52	2.20
12	69	3.60	170	8.36	fail	fail	136	8.63
13	10	0.81	15	1.19	10	0.89	11	0.97
14	11	1.16	20	2.08	9	1.19	9	1.17
15	12	0.89	22	1.55	12	0.94	13	1.00
16	35	1.36	39	1.47	fail	fail	22	0.88
17	34	1.64	18	1.08	15	0.86	22	1.61
18	20	0.84	44	1.60	19	0.83	22	0.88
19	23	1.44	18	1.20	19	1.27	22	1.40
20	26	2.10	44	3.23	19	1.45	22	1.75
21	20	1.30	43	2.58	20	1.30	22	1.60
22	23	1.48	130	7.16	20	1.48	31	2.31
23	33	1.98	104	5.80	24	1.55	73	4.36
24	63	4.09	96	6.17	fail	fail	39	2.83
25	23	2.90	61	6.73	fail	fail	fail	fail
26	48	5.40	80	8.73	25	3.10	91	9.84
27	48	6.86	43	5.36	fail	fail	42	5.64
28	19	1.81	20	1.90	fail	fail	27	2.53
29	12	1.36	fail	fail	fail	fail	16	1.55
30	15	1.5	25	2.36	5	0.70	43	3.89
31	14	26.52	22	39.27	12	24.09	27	4.75
32	12	24.03	17	32.34	11	22.60	fail	fail
33	19	32.52	21	34.30	5	12.75	18	31.44
34	59	3.28	84	4.58	9	0.67	68	3.63
35	12	0.80	15	0.95	7	0.58	19	1.11
36	20	1.23	213	10.98	12	0.83	17	1.10
37	59	5.23	85	7.53	9	1.06	68	6.67
38	12	1.34	15	1.75	7	0.88	19	2.06
39	16	1.69	145	13.95	11	1.27	19	2.28
40	59	100.1	96	149.8	9	18.98	70	118.6
41	17	33.60	24	42.38	8	18.56	23	42.84
42	15	30.97	253	389	10	25.55	20	37.90
43	4	0.63	5	0.,56	4	0.39	4	0.40
44	5	0.59	7	0.55	5	0.92	5	0.72
45	11	1.06	13	1.22	10	1.00	11	1.20
46	20	3.73	19	3.83	19	3.84	19	3.25
47	12	1.14	12	1.16	11	1.19	11	1.13
48	22	4.42	27	4.08	21	3.77	20	4.19
49	20	3.38	19	3.60	19	3.77	19	3.55
50	22	3.76	21	3.75	20	3.31	22	3.63
51	14	0.86	2617	134.55	13	0.88	11	0.77
52	15	1.00	1761	90.53	12	0.77	15	0.95
53	15	1.03	1546	77.97	13	0.81	15	0.95
54	21	2.09	21	2.03	9	1.03	19	2.11

(Continued on the next page)

TABLE II: *Continued*

NP	FRMIL		FR		PRP		RMIL	
	NI	CT	NI	CT	NI	CT	NI	CT
55	18	1.77	27	2.44	13	1.34	16	1.67
56	18	1.88	20	2.19	13	1.34	19	1.88
57	30	48.81	41	61.42	13	24.25	28	47.77
58	23	38.44	30	43.98	fail	fail	26	44.56
59	34	53.00	62	95.67	12	22.44	40	59.55
60	4	0.28	4	0.28	4	0.31	3	0.28
61	8	0.52	8	0.41	fail	fail	4	0.30
62	2	0.25	2	0.23	3	0.28	4	0.36
63	16	1.27	fail	fail	19	1.39	17	1.28
64	19	1.47	23	1.61	14	1.14	fail	fail
65	16	1.17	26	1.88	17	1.30	20	1.56
66	24	3.90	24	3.88	24	3.90	25	4.09
67	30	4.98	31	4.95	34	5.39	34	5.34
68	53	10.78	45	9.31	27	4.73	62	9.56
69	7	0.50	10	0.56	6	0.75	6	0.44
70	8	1.16	11	1.23	8	0.94	8	1.23
71	7	0.48	8	0.48	8	0.59	7	0.50
72	18	0.50	31	0.56	17	0.75	19	0.44
73	25	2.42	30	2.61	27	2.69	32	3.03
74	30	3.70	80	7.09	34	3.19	50	4.58
75	25	5.81	39	8.70	20	4.78	27	6.34
76	29	6.64	42	9.36	35	7.84	36	8.09
77	45	9.47	75	15.25	53	11.09	56	11.66
78	6	0.45	8	0.70	6	0.36	6	0.56
79	8	0.52	15	0.72	8	0.48	8	0.44
80	8	0.41	17	0.69	8	0.45	8	0.44
81	9	0.73	11	0.89	10	0.86	10	1.30
82	12	0.89	22	1.55	12	0.94	13	1.00
83	13	1.08	36	2.42	13	1.03	13	1.05
84	9	14.64	10	15.44	9	14.95	9	15.77
85	9	16.78	12	18.75	10	16.36	10	16.31
86	10	16.62	12	19.89	11	17.58	11	17.53
87	8	0.95	11	1.32	9	0.94	9	1.14
88	9	0.82	9	0.70	10	1.17	9	0.89
89	8	0.94	11	1.22	9	1.08	9	1.03
90	9	1.01	11	1.08	9	0.98	9	0.98
91	9	21.31	12	22.75	9	21.30	9	19.97
92	9	19.88	9	20.45	10	21.14	10	20.88
93	10	19.98	10	20.56	9	18.25	10	19.75
94	7	0.92	12	0.84	7	1.84	8	0.70
95	10	0.62	13	0.77	8	0.49	8	0.61
96	12	0.71	14	0.80	9	0.60	10	0.60
97	7	1.36	12	1.94	7	1.36	8	1.53
98	10	1.78	14	2.28	9	1.58	9	1.61
99	12	1.97	15	2.33	9	1.58	10	1.77
100	7	12.50	13	18.38	8	14.16	8	14.05
101	10	16.23	16	23.31	9	14.47	10	16.42
102	13	18.47	16	21.67	10	18.28	10	15.20

(Continued on the other side of this page)

TABLE II: *Continued*

NP	FRMIL		FR		PRP		RMIL	
	NI	CT	NI	CT	NI	CT	NI	CT
103	7	0.45	10	0.50	6	0.47	6	0.50
104	8	0.44	11	0.41	8	0.38	8	0.41
105	7	0.39	8	0.41	8	0.44	7	0.39
106	8	1.45	8	1.50	7	1.50	8	1.59
107	7	1.39	6	1.38	6	1.38	6	1.45
108	6	1.31	6	1.34	7	1.52	6	1.41
109	8	11.94	9	13.56	9	17.94	8	13.80
110	7	11.42	6	10.81	7	11.42	8	12.41
111	6	10.39	6	10.61	7	11.44	6	10.92
112	12	0.78	12	0.98	6	0.55	17	0.91
113	8	0.55	8	0.83	7	0.48	9	0.60
114	7	0.50	7	0.66	7	0.48	12	0.79
115	16	2.51	16	3.05	7	1.42	21	3.17
116	19	3.19	19	3.63	7	1.41	21	2.98
117	19	16.95	19	21.88	6	14.44	22	22.05
118	20	29.94	36	42.00	10	16.05	18	24.33
119	13	20.92	13	19.97	8	14.34	15	22.64
120	11	16.95	14	21.88	8	14.44	15	22.05
121	15	0.67	31	1.36	7	0.45	26	1.05
122	8	0.45	9	0.48	6	0.39	18	0.94
123	26	4.42	35	5.72	26	4.34	26	4.41
124	27	4.47	37	6.00	26	4.44	27	4.53
125	29	4.86	32	5.31	28	4.70	31	5.17
126	26	212.5	31	248.1	28	227.4	26	229.3
127	26	208.6	31	249.6	28	240.4	26	210.5
128	29	226.3	39	313.3	31	243.0	30	243.9

To obtain an overall assessment of a method's performance, we aim to derive a general score. For this purpose, we introduce the performance function $\rho_s(\tau)$, which is defined as follows:

$$\rho_s(\tau) = \frac{1}{n_p}\text{size}\{p \in P : r_{p,s} \leq \tau\},$$

where $\rho_s(\tau)$ represents the probability that method s achieves a performance ratio $r_{p,s}$ within a factor of $\tau \in \mathbb{R}$ compared to the best possible ratio. In essence, it quantifies the likelihood of method s performing well relative to the best method. Typically, methods with higher values of $\rho(\tau)$, or those located in the top right region of the performance profile figure, are regarded as recommended or indicative of a superior method.

TABLE III: The resume of experiment results

Parameter	Sum of NI	Sum of CPU time	Success Rate
FR	14,281	2,944.6534	98%
PRP	1,881	1,396.4064	93%
RMIL	2,750	1,710.8595	98%
FRMIL	2,397	1,583.5158	100%

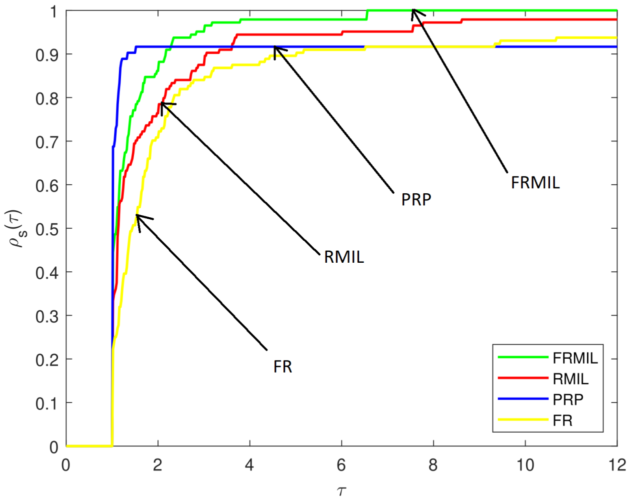

Fig. 1: Performance representation curve based on the number of iterations.

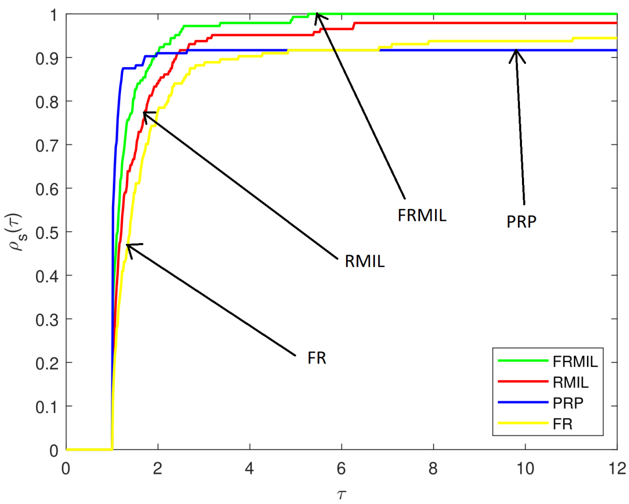

Fig. 2: Performance representation curve based on the CPU time.

Fig. 1 and Fig. 2 represent the performance of all methods based on the number of iterations and the total CPU time required to solve each test function. The curve of the method with the FRMIL parameter is at the top left to right in comparison to the other parameters, indicating that the FRMIL parameter outperforms the FR, PRP, and RMIL parameters. Although the PRP method previously achieved a higher probability value than the modified method, there are several functions that the PRP method was unable to complete, as shown in Table II. According to the summary result in Table III, the modified method successfully completes all test functions, the RMIL method completes the 98% of the test function, and the PRP method completes the 93% of the test function. The FR method can solve 98% of the test function, but its high iteration count hinders its performance compared to the other methods. Therefore, the FRMIL parameter is more efficient than the FR, PRP, and RMIL parameters.

VI. Conclusion

This article introduces a modified formula for the conjugate gradient parameter. The parameter presented in this paper is called the FRMIL parameter, which is obtained by combining the FR parameter, known for its global convergence, with the RMIL parameter, known for its good performance. The FRMIL method satisfies the requirements of sufficient descent and global convergence, especially when used in conjunction with an exact line search. Experimental results demonstrate that the FRMIL parameter outperforms the FR, PRP, and RMIL parameters in terms of efficiency.

References

[1] J. Cao and J. Wu, "A conjugate gradient algorithm and its applications in image restoration," *Applied Numerical Mathematics,* vol. 152, pp. 243-252, 2020.

[2] A. B. Abubakar, P. Kumam, and A. M. Awwal, "Global convergence via descent modified three-term conjugate gradient projection algorithm with applications to signal recovery," *Results in Applied Mathematics,* vol. 4, p. 100069, 2019.

[3] T. Helmig, F. Al-Sibai, and R. Kneer, "Estimating sensor number and spacing for inverse calculation of thermal boundary conditions using the conjugate gradient method," *International Journal of Heat and Mass Transfer,* vol. 153, p. 119638, 2020.

[4] A. Nicolaide, "A General Approach to the Formulae Used in Conjugate Direction Method for Solving Linear Systems of Algebraic Equations," *IAENG Lecture Notes in Engineering and Computer Science: Proceeding of The World Congress on Engineering 2011, WCE 2011, 6-8 July, 2011, London, U.K.,* pp. 205-210.

[5] O. Kardani, A. V. Lyamin, and K. Krabbenhoft, "A Comparative Study of Preconditioning Techniques for Large Sparse Systems Arising in Finite Element Limit Analysis," *IAENG International Journal of Applied Mathematics,* vol. 43, no. 4, pp. 195-203, 2013.

[6] D. Kumar, S. Gupta, dan P. Sehgal, "Improved Training of Predictive ANN with Gradient Techniques," *Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2014,* IMECS 2014, 12-14 March, 2014, Hong Kong, pp. 394-399.

[7] I. A. R. Moghrabi, "A New Preconditioned Conjugate Gradient Method for Optimization," *IAENG International Journal of Applied Mathematics,* vol. 49, no. 1, pp. 29-36, 2019.

[8] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal,* vol. 7, no. 2, pp. 149-154, 1964.

[9] M. J. D. Powell, "Restart Procedures for the Conjugate Gradient Method," *Mathematical Programming,* vol. 12, pp. 241-254, 1977.

[10] E. Polak and G. Ribiere, "Note sur la convergence de m´ethodes de directions conjugu´ees," *ESAIM: Mathematical Modelling and Numerical Analysis-Mod´elisation Math´ematique et Analyse Num´erique,* vol. 3, no. R1, pp. 35-43, 1969.

[11] M. J. D. Powell, "Nonconvex minimization calculations and the conjugate gradient method," *Lecture Notes in Mathematics*, vol. 1066, pp. 122-141, 1984.

[12] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM J. Optimizat*, vol. 2, no. 1, pp. 21-42, 1992.

[13] M. Rivaie, M. Mamat, L. W. June, and I. Mohd, "A new class of nonlinear conjugate gradient coefficient with global convergence properties," *Appl. Math. comput.,* vol. 218, pp. 11323-11332, 2012.

[14] Z. Dai, "Comment on a New Class of Nonlinear Conjugate Gradient Coefficients with Global Convergence Properties," *Applied Mathematics and Computation*, vol. 276, pp. 297-300, 2016.

[15] M. Malik, M. Mamat, S. S. Abas, I. M. Sulaiman and Sukono, "Performance Analysis of New Spectral and Hybrid Conjugate Gradient Methods for Solving Unconstrained Optimization Problem," *IAENG International Journal of Computer Science,* vol. 48, pp. 66-79, 2021.

[16] O. O. O. Yousif, "The convergence properties of RMIL+ conjugate gradient method under the strong Wolfe line search," *Applied Mathematics and Computation,* vol. 367, p. 124777, 2020.

[17] M. Malik, M. Mamat, S. S. Abas, I. M. Sulaiman, and Sukono, "A New Coefficient of the Conjugate Gradient Method with the Sufficient Descent Condition and Global Convergence Properties," *Engineering Letters,* vol. 28, no. 3, pp. 704-714, 2020.

[18] A. M. Awwal, I. M. Sulaiman, M. Malik, M. Mamat, P. Kumam and K. Sitthithakerngkiet, "A Spectral RMIL+ Conjugate Gradient Method for Unconstrained Optimization with Applications in Portofolio Selection and Motion Control," *IEEE Access,* vol. 9, pp. 75398-75414, 2021.

[19] M. A. Hery, M. Mamat, and L. W. June, "BFGS Method: A New Search Direction," *Sains Malaysiana,* vol. 40, no. 10, pp. 1591-1597, 2014.
[20] M. Malik, M. Mamat, S. S. Abas, I. M. Sulaiman, and Sukono, "A New Modification of NPRP Conjugate Gradient Method for Unconstrained Optimization," *Advances in Mathematics: Scientific Journal*, vol. 9 no. 7, pp. 4955-4970, 2020.
[21] N. Andrei,"An unconstrained optimization test functions collection," *Adv. Model. Optim*, vol. 10, no. 1, pp. 147-161, 2008.
[22] J. Dolan and J. J. More, "Benchmarking optimization software with performance profile," *Math. Prog*, vol. 91, pp. 201-213, 2002.
[23] G. Zoutendijk, "Nonlinear programming, computational methods," *Integer and Nonlinear Programming*, pp. 37-86, 1970.