An Improved Version of Rivaie-Mohd-Ismail-Leong Conjugate Gradient Method With Application in Image Restoration

Nur Idalisa, *Member, IAENG,* Mohd Rivaie, Nurul Hafawati Fadhilah,
Ibrahim Mohammed Sulaiman, *Member, IAENG,* Norhaslinda Zullpakkal,
Nurul Akmal Mohamed, *Member, IAENG,* and Norma Alias

***Abstract*—This paper focuses on modifying the existing Conjugate Gradient (CG) method of Rivaie, Mustafa, Ismail and Leong (RMIL). The RMIL technique has been the subject of previous studies to enhance its effectiveness. In this study, a new CG search direction, IRMIL, has been presented. This new variation combines the scaled negative gradient, which acts as an initial direction, and a third-term parameter. This paper proves that the IRMIL satisfies the sufficient descent criteria. The method also exhibits global convergence characteristics for exact and strong Wolfe line searches. The method's efficacy is assessed using two distinct methodologies. The first methodology involved conducting numerical tests on conventional Unconstrained Optimisation (UO) problems. The test shows that, while the IRMIL method performs very similarly to other existing CG methods during exact line search, it excels during strong Wolfe line search and converges more quickly. For the second methodology, the NEWMRIL method is applied to solve issues regarding image restoration. Overall, IRMIL method exhibits excellent theoretical and numerical efficiency potential.**

***Index Terms*—conjugate gradient method, sufficient descent property , global convergence, strong Wolfe line search, unconstrained optimization, image restoration problems.**

I. Introduction

The Conjugate Gradient (CG) method is a notable technique utilised for resolving the Unconstrained Optimisation (UO) problem as follows:

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

Its wide acceptance can be attributed to its capacity for global convergence and relatively low memory requirements. Equation (1) comprises a smooth function, denoted as $f:\mathbb{R}^n\rightarrow\mathbb{R}$, with a computable gradient represented by $g(x)=\nabla f(x)$. Using an initial estimate of $\boldsymbol{x}_0\in\mathbb{R}^n$, the CG method would generate an iterative sequence comprising the following steps:

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\alpha_k\boldsymbol{d}_k, \tag{2}$$

where α_k is the stepsize calculated by any line search techniques along the search direction d_k. Line search techniques, such as the exact line search and the weak or strong Wolfe line search, commonly facilitate the computation of α_k. The exact numerical value of α_k can be determined by solving the following equation:

$$\alpha_k=\arg\min_{\alpha_k\geq 0} f(\boldsymbol{x}_k+\alpha_k\boldsymbol{d}_k). \tag{3}$$

The strong Wolfe line search [1] is computed such that α_k satisfies

$$f(\boldsymbol{x}_k+\alpha_k\boldsymbol{d}_k)\leq f(\boldsymbol{x}_k)+\mu\alpha_k\boldsymbol{g}_k^T\boldsymbol{d}_k, \tag{4}$$

$$\left|\boldsymbol{g}(\boldsymbol{x}_k+\alpha_k\boldsymbol{d}_k)^T\boldsymbol{d}_k\right|\leq-\sigma\boldsymbol{g}_k^T\boldsymbol{d}_k, \tag{5}$$

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

CG's search direction vector $\boldsymbol{d}_k$ is defined as:

$$\boldsymbol{d}_k=\begin{cases}-\boldsymbol{g}_k, & \text{for } k=0\\ -\boldsymbol{g}_k+\beta_k\boldsymbol{d}_{k-1}, & \text{for } k\geq 1.\end{cases} \tag{6}$$

The CG coefficient, abbreviated as β_k, is crucial in determining the theoretical features of the algorithm. The selection of the line search method significantly impacts both the convergence and accuracy of the algorithm, as indicated in reference [2]. Hestenes and Steifel established the CG approach in [3], and Fletcher and Reeves expanded it to address UO issues. After that, they developed the first non-linear CG approach, FR [4]. Besides that, other CG techniques have also been developed throughout the years, including the PRP method [5], [6], the Daniel method [7], the conjugate-descent (CD) method [8], the Liu-Storey (LS) method [9], the Dai-Yuan (DY) method [10]. However, most of these methods lacked global convergence qualities and failed to meet the sufficient descent criterion (SDC) [11].

Manuscript received December 4, 2022; revised July 29, 2023. This work was supported by the Ministry of Higher Education (MOHE) Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2019/STG06/UITM/02/3).

N. Idalisa is a senior lecturer of the Department of Mathematics, College of Computing Informatic and Media, Universiti Teknologi MARA Cawangan Terengganu, Kampus Kuala Terengganu, 21080, Malaysia (phone: +60139955294; e-mail: nuridalisa@uitm.edu.my)

M. Rivaie is a senior lecturer of the Department of Mathematics, College of Computing Informatic and Media, Universiti Teknologi MARA Cawangan Terengganu, Kampus Kuala Terengganu, 21080, Malaysia (e-mail: rivaie75@uitm.edu.my)

N.H. Fadhilah is a PhD candidate of Universiti Teknologi MARA Cawangan Terengganu, Kampus Dungun, 23000, Terengganu, Malaysia (e-mail: nurulhafawatifadhilah@gmail.com)

I.M. Sulaiman is an international senior lecturer attached to the department of Mathematics and Statistics, School of Quantitative Sciences, Universiti Utara Malaysia, Sintok, Kedah, 06010, Malaysia (e-mail: i.mohammed.sulaiman@uum.edu.my)

N. Zullpakkal is a senior lecturer of the Department of Mathematics, College of Computing Informatic and Media, Universiti Teknologi MARA Cawangan Terengganu, Kampus Kuala Terengganu, 21080, Malaysia (e-mail: lindazullpakkal@uitm.edu.my)

N.A. Mohamed is a seniour lecturer of the Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Tanjung Malim, 35900, Perak, Malaysia (e-mail: akmal.mohamed@fsmt.upsi.edu.my)

N. Alias is an associate professor of Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, Skudai, 81310, Johor, Malaysia (e-mail: normaalias@utm.my)

The descent condition can be extended to SDC in the following form:

$$\boldsymbol{g}_k^T \boldsymbol{d}_k \le -c\|\boldsymbol{g}_k\|^2, \ k \ge 0, c > 0. \tag{7}$$

Hager and Zhang published an improved version of the HS technique in 2005 [12]. It is called CG DESCENT and is one of the most effective CG algorithms. This formula provides a working definition of it:

$$\beta_k^{\text{CG DESCENT}} = \max\{\beta_k^N, \eta_k\}, \tag{8}$$

where

$$\beta_k^N = \frac{1}{\boldsymbol{d}_{k-1}^T \boldsymbol{y}_{k-1}} \left(\boldsymbol{y}_{k-1} - 2\boldsymbol{d}_{k-1} \frac{\|\boldsymbol{y}_{k-1}\|^2}{\boldsymbol{d}_{k-1}^T \boldsymbol{y}_{k-1}} \right)^T \boldsymbol{g}_k, \tag{9}$$

$$\eta_k = \frac{-1}{\|\boldsymbol{d}_{k-1}\| \min\{\eta, \|\boldsymbol{g}_{k-1}\|\}}, \tag{10}$$

with the positive constant η, $\boldsymbol{y}_{k-1} = \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$ and $\|.\|$ stands for Euclidean norm. The CG DESCENT technique is unique in consistently generating descent directions when $\boldsymbol{d}_k^T \boldsymbol{y}_k \neq 0$. Due to its excellent performance, the CG DESCENT technique is widely regarded as a benchmark of the CG algorithm. However, its performance varies depending on searched lines [13] .

Rivaie et al. [14] developed the RMIL method, which was described as follows:

$$\beta_k^{\text{RMIL}} = \frac{\boldsymbol{g}_k^T(\boldsymbol{g}_k - \boldsymbol{g}_{k-1})}{\boldsymbol{d}_{k-1}^T(\boldsymbol{d}_{k-1} - \boldsymbol{g}_k)}. \tag{11}$$

The sign of β_k^{RMIL} is determined by the values of $\boldsymbol{d}_{k-1}^T \boldsymbol{g}_k$ and $\boldsymbol{g}_k^T \boldsymbol{g}_{k-1}$, which can be categorised into four cases as presented in Table I.

TABLE I
SEVERAL CASES OF RMIL METHOD

Case	Numerator	Denominator	Sign of β_k^{RMIL}
1	$\|\boldsymbol{g}_k\|^2 > \boldsymbol{g}_k^T \boldsymbol{g}_{k-1}$	$\|\boldsymbol{d}_{k-1}\|^2 > \boldsymbol{d}_{k-1}^T \boldsymbol{g}_k$	Positive
2	$\|\boldsymbol{g}_k\|^2 < \boldsymbol{g}_k^T \boldsymbol{g}_{k-1}$	$\|\boldsymbol{d}_{k-1}\|^2 > \boldsymbol{d}_{k-1}^T \boldsymbol{g}_k$	Negative
3	$\|\boldsymbol{g}_k\|^2 > \boldsymbol{g}_k^T \boldsymbol{g}_{k-1}$	$\|\boldsymbol{d}_{k-1}\|^2 < \boldsymbol{d}_{k-1}^T \boldsymbol{g}_k$	Negative
4	$\|\boldsymbol{g}_k\|^2 < \boldsymbol{g}_k^T \boldsymbol{g}_{k-1}$	$\|\boldsymbol{d}_{k-1}\|^2 < \boldsymbol{d}_{k-1}^T \boldsymbol{g}_k$	Positive

The global convergence of using β_k^{RMIL} was established under exact line search. Given that the orthogonal condition, $\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k = 0$, is satisfied by the exact minimisation rule, Equation (11) can be modified in another variant known as RMIL* [15]. The expression for RMIL* is as follows:

$$\beta_k^{\text{RMIL*}} = \frac{\boldsymbol{g}_k^T(\boldsymbol{g}_k - \boldsymbol{g}_{k-1})}{\|\boldsymbol{d}_{k-1}\|^2}. \tag{12}$$

Under the exact line search, both equations RMIL and RMIL*, derived from Equations (11) and (12), exhibit comparable performance and convergence features. In order to facilitate the analysis, β_k^{RMIL} will be simplified as follows:

$$0 < \beta_k \le \frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{d}_{k-1}\|^2}. \tag{13}$$

As demonstrated in both references [14] and [15], any CG technique that satisfies Equation (13) has the potential for global convergence if an exact or strong Wolfe line search is employed. Based on this, most comparable works, such as [16]–[21], built RMIL variations to satisfy (13).

While RMIL exhibits global convergence properties when exact line search is employed, its numerical output is comparatively slower than the original HS and PRP algorithms. Furthermore, it was argued by [22] that the RMIL technique was incapable of satisfying (13) and was only valid under the condition that the inequality $\|\boldsymbol{g}_k\|^2 \ge \boldsymbol{g}_k^T \boldsymbol{g}_{k-1}$ was fulfilled. In Case 1, [22]'s modified RMIL version, RMIL+, was included in the RMIL procedure. The RMIL+ method was defined in its entirety as follows:

$$\beta_k^{\text{RMIL+}} = \begin{cases} \frac{\boldsymbol{g}_k^T(\boldsymbol{g}_k - \boldsymbol{g}_{k-1})}{\|\boldsymbol{d}_{k-1}\|^2}, & \text{if } 0 \le \boldsymbol{g}_k^T \boldsymbol{g}_{k-1} \le \|\boldsymbol{g}_k\|^2, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

The objective of RMIL+ is to retain the non-negativity of β_k by applying the Case 1 constraint from the RMIL method. This method is comparable to the PRP+ and PRP* techniques described in [23]. However, when the exact line search is used, RMIL+ does not significantly outperform RMIL and is still slower than CG approaches like HS and PRP. Hence, it is intriguing to investigate alternative ways to enhance the performance and global convergence features of RMIL while refraining from imposing the constraint of Equation (13).

Another consideration for improving the RMIL method involves using an inexact line search instead of an exact line search. The strong Wolfe condition is a widespread criterion for inexact line searches; it has two required parameters, μ and σ, where $0 < \mu < \sigma < 1$. Some studies, such as [17], [24], and [25], expanded RMIL to the strong Wolfe line search, with ranges of $\sigma < \frac{1}{2}$, $\sigma < \frac{1}{4}$, $\sigma < \frac{1}{5}$, respectively. Nevertheless, certain CG methods, such as [26] and [27], can accomplish global convergence for any choice of $\sigma < 1$ under the strong Wolfe condition. Motivated by prior research findings, researchers aim to expand the RMIL method using a strong Wolfe line search to any option of $\sigma < 1$.

The three-term CG comes from its general form [28], defined as

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k, & \text{if } k = 0, \\ -\beta_k^{(1)} \boldsymbol{o}_k + \beta_k^{(2)} \boldsymbol{p}_k + \beta_k^{(3)} \boldsymbol{q}_k, & \text{if } k \ge 1, \end{cases} \tag{15}$$

where $\boldsymbol{o}_k, \boldsymbol{p}_k$ and $\boldsymbol{q}_k$ are directions and $\beta_k^{(1)}, \beta_k^{(2)}, \beta_k^{(3)}$ are scalars. The vector $\boldsymbol{p}_k$ is either $\boldsymbol{d}_{k-1}$ or $\boldsymbol{s}_{k-1} = \boldsymbol{x}_k - \boldsymbol{x}_{k-1}$. Considering the two-term CG structure, this research limits the options to $\beta_k^{(1)} = 1$, $\boldsymbol{o}_k = \boldsymbol{g}_k$ and $\boldsymbol{p}_k = \boldsymbol{d}_{k-1}$.

Since there is limited data on the RMIL method's three-term strategy, presently, only five formulations utilise RMIL or a derivative of it to calculate $\beta_k^{(2)}$. Recent literature [29], [30] have established the SDC property by utilising a strong Wolfe line search, with $\beta_k^{(2)}$ to β_k^{RMIL} and $\beta_k^{\text{RMIL+}}$, respectively. Furthermore, without providing a thorough global convergence analysis, [31] established a three-term CG based on β_k^{SMAR} and β_k^{SMARZ}.

On the other hand, [32] has given a fourth technique that employs either $\beta_k^{\text{RMIL*}}$ or β_k^{MRMIL} as $\beta_k^{(2)}$, with global convergence features assessed using standard Wolfe. Finally, a three-term RMIL method was suggested by [33], which utilises $\beta_k^{\text{RMIL*}}$ as $\beta_k^{(2)}$ and $\boldsymbol{q}_k = \boldsymbol{d}_{k-1}$. However, its SDC property and global convergence are shown only under subjected to an exact line search. The three-term RMIL methods are summarised in Table II.

TABLE II
THREE-TERM RMIL METHOD

Reference	$\beta_k^{(2)}$	$\beta_k^{(3)}$	$\boldsymbol{q}_k$
[30]	$\beta_k^{\text{RMIL+}}$	$\frac{\boldsymbol{g}_k^T \boldsymbol{d}_{k-1}}{\|\boldsymbol{d}_{k-1}\|^2}$	$\boldsymbol{g}_{k-1}$
[31]	β_k^{SMAR} or β_k^{SMARZ}	$\boldsymbol{g}_k - \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}\boldsymbol{d}_{k-1}$	$\boldsymbol{y}_{k-1}$
[29]	β_k^{RMIL}	$-\frac{\boldsymbol{g}_k^T \boldsymbol{d}_{k-1}}{\boldsymbol{d}_{k-1}^T(\boldsymbol{d}_{k-1}-\boldsymbol{g}_k)}$	$\boldsymbol{y}_{k-1}$
[32]	$\beta_k^{\text{RMIL*}}$ or β_k^{MRMIL}	$-\frac{\boldsymbol{g}_k^T \boldsymbol{d}_{k-1}}{\|\boldsymbol{d}_{k-1}\|^2}$	$\boldsymbol{y}_{k-1}$
[33]	$\beta_k^{\text{RMIL*}}$	$-\beta_k \frac{\boldsymbol{g}_k^T \boldsymbol{g}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2}$	$\boldsymbol{d}_{k-1}$

II. PROPOSED ALGORITHM

The IRMIL technique was introduced as a new CG approach for UO. An exhaustive explanation of the IRMIL procedure is as follows:

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k + \omega \boldsymbol{g}_k, & \text{if } k = 0, \\ -\boldsymbol{g}_k + \beta_k^{\text{RMIL}} \boldsymbol{d}_{k-1} + \frac{\boldsymbol{g}_k^T \boldsymbol{d}_{k-1}}{\boldsymbol{d}_{k-1}^T(\boldsymbol{d}_{k-1}-\boldsymbol{g}_k)} \boldsymbol{g}_{k-1}, & \text{if } k \geq 1, \end{cases} \tag{16}$$

where $\omega \in (0, 1)$.

Algorithm 1 describes the operation of the proposed method.

Algorithm 1 IRMIL CG method

Require: Set $k = 0$. Initiate the starting point $\boldsymbol{x}_0 \in \mathbb{R}^n$. Specify the convergence tolerance, ϵ and maximum iterations, Maxit.
Set $\boldsymbol{d}_0 = -\boldsymbol{g}_0 + \omega \boldsymbol{g}_0$ for $\omega \in (0, 1)$.
while $\|\boldsymbol{g}\| > \epsilon$ and $k <$ Maxit **do**
  Determine α_k either by (3) or (4)-(5).
  Update $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$.
  Compute $\beta_{k+1}^{\text{RMIL}}$ using (11).
  Find search direction,

$$\boldsymbol{d}_{k+1} = -\boldsymbol{g}_{k+1} + \beta_{k+1}^{\text{RMIL}} \boldsymbol{d}_k + \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\boldsymbol{d}_k^T(\boldsymbol{d}_k - \boldsymbol{g}_{k+1})} \boldsymbol{g}_k.$$

  $k \Leftarrow k + 1$
end while

III. CONVERGENCE ANALYSIS

In convergence analysis, two key assumptions for the objective function are boundedness on the level set and the Lipschitz assumption.

Assumption 1: If the level set $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n\}$ where $f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0)$ is bounded, then there exists a positive constant, $B > 0$ where $\|\boldsymbol{x}\| \leq B$ for all $\boldsymbol{x} \in \Omega$.

Assumption 2: In some neighbourhoods ζ of Ω, f is continuously differentiable, and its gradient $\boldsymbol{g}$ is Lipschitz continuous, so there exists a constant $L > 0$ such that

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{y})\| \leq L\|\boldsymbol{x} - \boldsymbol{y}\| \text{ for any } \boldsymbol{x}, \boldsymbol{y} \in \zeta. \tag{17}$$

The boundedness assumption is necessary to guarantee that the stepsize α_k is well-defined for every k [34]. As a result, Assumptions (1) and (2) imply the existence of a constant $m > 0$ such that

$$\|\boldsymbol{g}_k\| \leq m, \forall \boldsymbol{x} \in \Omega. \tag{18}$$

The sequence of $\boldsymbol{x}_k$ generated by the CG technique is reduced by integrating the sequence $\{f(\boldsymbol{x}_k)\}$ into the level set of Ω using these two assumptions. This feature may ensure the $\boldsymbol{x}_k$'s progress to stationary.

A. Convergence Analysis under Exact Line Search

This section analyses the IRMIL method when executed with an exact line search. It commences by demonstrating the validity of SDC in Theorem 1, followed by establishing the global convergence property in Theorem 2.

Theorem 1: Presume that the IRMIL algorithm generates a CG method with a search direction for all $k > 0$ under an exact line search using Equation (3). The method proposed in Equation (16) is proven to meet the Sufficient Descent Condition (SDC) outlined in Inequality (7), such that

$$\boldsymbol{g}_k^T \boldsymbol{d}_k \leq -c\|\boldsymbol{g}_k\|^2, \; k \geq 0, c > 0.$$

Proof: If $k = 0$, then the initial direction takes $\boldsymbol{d}_0 = -\boldsymbol{g}_0 + \omega \boldsymbol{g}_0 = -(1 - \omega)\boldsymbol{g}_0$ where $\omega \in (0, 1)$. Next, multiply $\boldsymbol{d}_0$ with $\boldsymbol{g}_0^T$,

$$\boldsymbol{g}_0^T \boldsymbol{d}_0 = -(1 - \omega)\|\boldsymbol{g}_0\|^2. \tag{19}$$

Since $(1 - \omega) > 0$, then the SDC is satisfied when $k = 0$. The proof for $k \geq 1$ is established using the induction method. From Equation (16), set $k = k + 1$ where

$$\boldsymbol{d}_{k+1} = -\boldsymbol{g}_{k+1} + \beta_{k+1}^{\text{RMIL}} \boldsymbol{d}_k + \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\boldsymbol{d}_k^T(\boldsymbol{d}_k - \boldsymbol{g}_{k+1})} \boldsymbol{g}_k.$$

Next, multiply $\boldsymbol{d}_{k+1}$ with $\boldsymbol{g}_{k+1}^T$,

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} = -\|\boldsymbol{g}_{k+1}\|^2 + \beta_{k+1} \boldsymbol{g}_{k+1}^T \boldsymbol{d}_k + \boldsymbol{g}_{k+1}^T \left(\frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\boldsymbol{d}_k^T(\boldsymbol{d}_k - \boldsymbol{g}_k)} \right) \boldsymbol{g}_k.$$

For exact line search, this condition hold $\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k = 0$ [35], [36]. Thus, the equation $\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} = -\|\boldsymbol{g}_{k+1}\|^2$ indicates that $\boldsymbol{d}_{k+1}$ is a sufficient descent direction. Hence, $\boldsymbol{g}_k^T \boldsymbol{d}_k \leq -C\|\boldsymbol{g}_k\|^2$ holds and the proof is concluded. ■

The reduction of the proposed three-term CG method to the standard two-term CG method is apparent when an exact line search is employed. In this case, the search direction in Equation (16) can be expressed as:

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k + \omega \boldsymbol{g}_k, & \text{if } k = 0, \\ -\boldsymbol{g}_k + \beta_k^{\text{RMIL}} \boldsymbol{d}_{k-1}, & \text{if } k \geq 1, \end{cases} \tag{20}$$

where $\omega \in (0, 1)$.

Moreover, according to the derivation presented in Equation (12), the IRMIL approach can be expressed as follows:

$$\boldsymbol{d}_k = \begin{cases} -\boldsymbol{g}_k + \omega \boldsymbol{g}_k, & \text{if } k = 0, \\ -\boldsymbol{g}_k + \beta_k^{\text{RMIL*}} \boldsymbol{d}_{k-1}, & \text{if } k \geq 1. \end{cases} \tag{21}$$

The convergence analysis for the IRMIL method is derived from the following relation:

$$\|\boldsymbol{d}_k\| \leq \bar{\kappa}\|\boldsymbol{g}_k\|, \quad \forall k \geq 0, \tag{22}$$

where $\bar{\kappa} \in \mathbb{R}^+$. The condition (22) is crucial in preventing significant steps in areas with a slight gradient, such as around saddle points and at the bottom of valleys [37]. This indicates that the gradient determines the direction.

Lemma 1: Presume that Assumptions 1 and 2 hold; the search direction is defined by Equation (21) with exact line search (Equation 3). Then the Inequality (22) holds.

Proof: It is obvious for $k = 0$, $\|\boldsymbol{d}_0\| = |1-\omega|\|\boldsymbol{g}_0\| < \|\boldsymbol{g}_0\|$. The next step is to establish that $k > 0$. This can be shown by modifying Equation (21) such that $k = k+1$. The new equation can be written as

$$\|\boldsymbol{d}_{k+1}\| = \left\|-\boldsymbol{g}_{k+1} + \beta_{k+1}^{\text{RMIL*}}\boldsymbol{d}_k\right\|.$$

By applying the triangle inequality of the Euclidean norm, it yields

$$\|\boldsymbol{d}_{k+1}\| \le \|\boldsymbol{g}_{k+1}\|\left(1 + \frac{\|\boldsymbol{y}_k\|}{\|\boldsymbol{d}_k\|}\right). \tag{23}$$

By Lipschitz assumption in Assumption (2),

$$\|\boldsymbol{y}_k\| = \|\boldsymbol{g}_{k+1} - \boldsymbol{g}_k\| \le L\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| = |\alpha_{k+1}|L\|\boldsymbol{d}_k\|.$$

Thus, $\frac{\|\boldsymbol{y}_k\|}{\|\boldsymbol{d}_k\|} \le L|\alpha_{k+1}|$ and hence the condition (22) is satisfied. ■

For any k, the required stepsize α_k must be well-defined; hence the boundedness condition is necessary [34]. As a result, Assumptions (1) and (2) entail the existence of a constant $m > 0$ such that

$$\|\boldsymbol{g}_k\| \le m, \forall \boldsymbol{x} \in \Omega. \tag{24}$$

With these two assumptions, the sequence $\boldsymbol{x}_k$ is contained in the level set Ω for the decreasing sequence of $\{f(\boldsymbol{x}_k)\}$ generated by a CG method. This feature might guarantee the $\boldsymbol{x}_k$ to progress to stationary.

In addition, under Assumptions (1) and (2), the following Lemma, proven by [38], is obtained.

Lemma 2: Presume that Assumptions (1) and (2) hold and α_k satisfies the exact minimisation rule (Equation 3). Consider any iteration of Equation (2), where $\boldsymbol{d}_k$ satisfies SDC in Inequality (7). Then, the following condition, known as Zoutendijk condition, holds:

$$\sum_{k\ge 0} \frac{\left(\boldsymbol{g}_k^T\boldsymbol{d}_k\right)^2}{\|\boldsymbol{d}_k\|^2} < +\infty. \tag{25}$$

The proof of this lemma can be acquired from reference [38]. As demonstrated in [39], Lemma 2 also holds to exact minimization and the Goldstein and Wolfe rules.

Using Lemma 2, the following theorem is applied to prove the global convergence of the IRMIL method by employing an exact line search.

Theorem 2: Presume that Assumptions (1) and (2) hold and let the search direction is defined by Equation (21) where α_k is obtained from an exact line search (Equation 3). If the SDC holds, then either

$$\liminf_{k\to\infty} \|\boldsymbol{g}_k\| = 0, \quad \text{or} \quad \sum_{k\ge 0} \frac{\left(\boldsymbol{g}_k^T\boldsymbol{d}_k\right)^4}{\|\boldsymbol{d}_k\|^2} < +\infty.$$

Proof: The proof is established by contradiction to $\liminf_{k\to\infty} \|\boldsymbol{g}_k\| = 0$ and the Equation (24). Then there exists a constant $r > 0$ such that

$$\|\boldsymbol{g}_k\| \ge r. \tag{26}$$

Using Lemma (1), the search direction is modified in Equation (23) by setting ($k + 1 = k$).

$$\|\boldsymbol{d}_k\| \le \|\boldsymbol{g}_k\|\left(1 + L|\alpha_k|\right) \le \bar{\kappa}\|\boldsymbol{g}_k\|. \tag{27}$$

Consequently, Equation (27) is divided with $\|\boldsymbol{g}_k\|^2$ and, taking into account the contradiction statement $\left(\frac{1}{\|\boldsymbol{g}_k\|} < \frac{1}{r}\right)$, gives Equation 28:

$$\frac{\|\boldsymbol{d}_k\|}{\|\boldsymbol{g}_k\|^2} \le \frac{\bar{\kappa}}{\|\boldsymbol{g}_k\|} \le \frac{\bar{\kappa}}{r}. \tag{28}$$

Squaring both sides of (28),

$$\frac{\|\boldsymbol{d}_k\|^2}{\|\boldsymbol{g}_k\|^4} \le \frac{\bar{\kappa}^2}{r^2}, \tag{29}$$

which can be altered to

$$\frac{\|\boldsymbol{g}_k\|^4}{\|\boldsymbol{d}_k\|^2} \ge \frac{r^2}{\bar{\kappa}^2}, \quad \text{for all } k. \tag{30}$$

For $k = 0$,

$$\frac{\|\boldsymbol{g}_0\|^4}{\|\boldsymbol{d}_0\|^2} = \frac{\|\boldsymbol{g}_0\|^4}{(1-\omega)^2\|\boldsymbol{g}_0\|^2} \ge \frac{r^2}{(1-\omega)^2} \ge \frac{r^2}{\bar{\kappa}^2}. \tag{31}$$

Summing up all the terms to obtain,

$$\sum_{k=0}^{\infty} \frac{\|\boldsymbol{g}_k\|^4}{\|\boldsymbol{d}_k\|^2} \ge \sum_{k=0}^{\infty} \frac{r^2}{\bar{\kappa}^2}. \tag{32}$$

Since $\lim_{k\to\infty} \frac{r^2}{\bar{\kappa}^2} \ne 0$, by divergent test, the series $\sum_{k=0}^{\infty} \frac{r^2}{\bar{\kappa}^2}$ diverges and consequently violates the Zoutendijk condition. The proof is then concluded. ■

B. Convergence Analysis under strong Wolfe Line Search

This section presents the proofs of the SDC and the global convergence of IRMIL when applied under the strong Wolfe line search. In a prior study, it was demonstrated by [17] that RMIL* converges under the condition that $\sigma < 1/2$ and the $\frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{d}_k\|} \le \frac{1}{1-\Gamma_u}$ where $0 < \Gamma_u < 1$. Following that, [18] investigates the convergence of LAMR+ by limiting the strong Wolfe parameter such that $\sigma < 1/5$ and the relation $\|\boldsymbol{d}_{k-1}\| > (2/3)\|\boldsymbol{g}_{k-1}\|$. In a recent study, [24] expanded the range of σ to $\sigma < 1/4$. In addition, the convergence of RMIL+ was established under the condition that $\|\boldsymbol{d}_{k-1}\| > (1/2)\|\boldsymbol{g}_{k-1}\|$.

For this study, no restrictions are placed on the parameter σ as long as it falls within the range $(0, 1)$ indicated by the strong Wolfe criteria.

The proof is based on the relation

$$\|\boldsymbol{d}_k\| \ge (1-\Gamma)\|\boldsymbol{g}_k\|, \quad \forall k \ge 0, \tag{33}$$

where $0 < \Gamma < 1$ is a scalar.

Moreover, another assumption is as follows:

$$\|\boldsymbol{d}_k\|^2 - \boldsymbol{d}_k^T\boldsymbol{g}_{k+1} > \|\boldsymbol{d}_k\|^2. \tag{34}$$

Theorem 3: Presume the sequences $\{\boldsymbol{g}_k\}$ and $\{\boldsymbol{d}_k\}$ are generated by the IRMIL method when implemented using the strong Wolfe line search. If $\sigma < (1-\Gamma)$ for $0 < \Gamma < 1$ then the Inequality (33) holds.

Proof: The case is proven by the induction method. For $k = 0$,

$$\|\boldsymbol{d}_0\| = |1-\omega|\|\boldsymbol{g}_0\|.$$

Thus, the Inequality (33) is satisfied since $(1-\omega) \in (0, 1)$. Assuming that condition (33) holds for a value of $k > 1$. Subsequently, the IRMIL algorithm, as expressed in Equation

(16), is executed with a value of $k = k + 1$. The outcome of Equation (16) is then multiplied by $\boldsymbol{g}_{k+1}^T$.

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} = -\|\boldsymbol{g}_{k+1}\|^2 + \beta_{k+1}^{\text{RMIL}} \boldsymbol{g}_{k+1}^T \boldsymbol{d}_k + \boldsymbol{g}_{k+1}^T \left(\frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\boldsymbol{d}_k^T(\boldsymbol{d}_k - \boldsymbol{g}_{k+1})} \right) \boldsymbol{g}_k \tag{35}$$

$$= -\|\boldsymbol{g}_{k+1}\|^2 + \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\boldsymbol{d}_k^T(\boldsymbol{d}_k - \boldsymbol{g}_{k+1})} \|\boldsymbol{g}_{k+1}\|^2. \tag{36}$$

Rewriting Equation (36) ,

$$\|\boldsymbol{g}_{k+1}\|^2 = -\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} + \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\|\boldsymbol{d}_k\|^2 - \boldsymbol{d}_k^T \boldsymbol{g}_{k+1}} \|\boldsymbol{g}_{k+1}\|^2. \tag{37}$$

By triangle inequality, Equation (37) will become

$$\|\boldsymbol{g}_{k+1}\|^2 = \left\| -\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} + \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\boldsymbol{d}_k^T(\boldsymbol{d}_k - \boldsymbol{g}_{k+1})} \|\boldsymbol{g}_{k+1}\|^2 \right\| \tag{38}$$

$$\leq \left| \boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} \right| + \frac{\left| \boldsymbol{g}_{k+1}^T \boldsymbol{d}_k \right|}{\left| \|\boldsymbol{d}_k\|^2 - \boldsymbol{d}_k \boldsymbol{g}_{k+1} \right|} \|\boldsymbol{g}_{k+1}\|^2. \tag{39}$$

Hence using Equation (34), we obtain the following:

$$\|\boldsymbol{g}_{k+1}\|^2 < \left| \boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} \right| + \frac{\left| \boldsymbol{g}_{k+1}^T \boldsymbol{d}_k \right|}{\|\boldsymbol{d}_k\|^2} \|\boldsymbol{g}_{k+1}\|^2 \tag{40}$$

Consequently, with the strong Wolfe condition (refer to Equations 4 and 5),

$$\|\boldsymbol{g}_{k+1}\|^2 \leq \left| \boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} \right| + \frac{\sigma \left| \boldsymbol{g}_k^T \boldsymbol{d}_k \right|}{\|\boldsymbol{d}_k\|^2} \|\boldsymbol{g}_{k+1}\|^2, \tag{41}$$

where $0 < \sigma < 1$.

Then by applying the Cauchy–Schwartz inequality and the induction hypothesis, the inequality (41) becomes

$$\|\boldsymbol{g}_{k+1}\|^2 \leq \|\boldsymbol{g}_{k+1}\| \|\boldsymbol{d}_{k+1}\| + \sigma \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{d}_k\|} \|\boldsymbol{g}_{k+1}\|^2. \tag{42}$$

Then, simplify Equation (42) and divide by $\|\boldsymbol{g}_{k+1}\|^2$ on both sides, yields

$$\frac{\|\boldsymbol{d}_{k+1}\|}{\|\boldsymbol{g}_{k+1}\|} \geq 1 - \sigma \frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{d}_k\|}. \tag{43}$$

From the induction hypothesis, $\frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{d}_k\|} \leq \frac{1}{1-\Gamma}$, implies

$$\frac{\|\boldsymbol{d}_{k+1}\|}{\|\boldsymbol{g}_{k+1}\|} \geq 1 - \frac{\sigma}{1 - \Gamma}. \tag{44}$$

The proof is complete. ■

Next is to prove the sufficient descent property for IRMIL method.

Theorem 4: Consider the IRMIL algorithm with α_k is calculated using strong Wolfe line search in Equations 4 and 5 where $0 < \sigma < (1-\Gamma)$. Then for all $k \geq 1$ the SDC holds such that

$$\boldsymbol{g}_k^T \boldsymbol{d}_k \leq -c\|\boldsymbol{g}_k\|^2, \ k \geq 0, c > \frac{1-\Gamma}{\sigma} \quad \text{and} \quad \Gamma \in (0, 1). \tag{45}$$

Proof: According to (16), $\boldsymbol{d}_0 = -\boldsymbol{g}_0 + \omega \boldsymbol{g}_0 = -(1 - \omega)\boldsymbol{g}_0$ for $k = 0$ where $\omega \in (0, 1)$. Furthermore,

$$\boldsymbol{g}_0^T \boldsymbol{d}_0 = -(1 - \omega)\|\boldsymbol{g}_0\|^2. \tag{46}$$

Since $(1-\omega) > 0$, the SDC is fulfilled for $k = 0$. Assume also that the SDC is fullfilled for some $k > 0$. By using Equation (36), the SDC for $k + 1$ can be shown:

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} = -\|\boldsymbol{g}_{k+1}\|^2 + \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_k}{\boldsymbol{d}_k^T(\boldsymbol{d}_k - \boldsymbol{g}_{k+1})} \|\boldsymbol{g}_{k+1}\|^2. \tag{47}$$

The expression for Equation (47) is modified by the presence of Inequality (34), resulting in

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} \leq -\|\boldsymbol{g}_{k+1}\|^2 + \frac{\left| \boldsymbol{g}_{k+1}^T \boldsymbol{d}_k \right|}{\|\boldsymbol{d}_k\|^2} \|\boldsymbol{g}_{k+1}\|^2. \tag{48}$$

Subsequently, when the strong Wolfe condition is applied (Inequalities (4) and (5)), the induction assumption yields

$$\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1} \leq -\|\boldsymbol{g}_{k+1}\|^2 - \sigma \frac{\left| \boldsymbol{g}_k^T \boldsymbol{d}_k \right|}{\|\boldsymbol{d}_k\|^2} \|\boldsymbol{g}_{k+1}\|^2 \tag{49}$$

$$\leq -\|\boldsymbol{g}_{k+1}\|^2 - c\sigma \frac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{d}_k\|^2} \|\boldsymbol{g}_{k+1}\|^2. \tag{50}$$

Consequently, dividing with $\|\boldsymbol{g}_{k+1}\|^2$ and using Theorem 3 leads to

$$\frac{\boldsymbol{g}_{k+1}^T \boldsymbol{d}_{k+1}}{\|\boldsymbol{g}_{k+1}\|^2} \leq -1 - \frac{c\sigma}{1 - \Gamma}. \tag{51}$$

Thus, the SDC is fulfilled.

The proof is concluded. ■

Given that the suggested approach consistently generates a search direction that meets the SDC, it is necessary to restrict the step length to prove global convergence. Lemma (3) demonstrates that the strong Wolfe line search always provides a lower bound for the step length α_k.

Lemma 3: Let the sequence $\{\boldsymbol{g}_k\}$ and $\{\boldsymbol{d}_k\}$ be generated by IRMIL method, and α_k is computed using strong Wolfe line search, then

$$\alpha_k \geq -\frac{(1-\sigma)\boldsymbol{g}_k^T \boldsymbol{d}_k}{L\|\boldsymbol{d}\|^2}. \tag{52}$$

Proof: Initiate the proof by writing the expression of $(\sigma - 1)\boldsymbol{g}_k^T \boldsymbol{d}_k = \sigma \boldsymbol{g}_k^T \boldsymbol{d}_k - \boldsymbol{g}_k^T \boldsymbol{d}_k$. In accordance to the strong Wolfe rule outlined in Inequations 4 and 5, it can be concluded that,

$$\begin{aligned} \sigma \boldsymbol{g}_k^T \boldsymbol{d}_k - \boldsymbol{g}_k^T \boldsymbol{d}_k &\leq \boldsymbol{g}_{k+1}^T \boldsymbol{d}_k - \boldsymbol{g}_k^T \boldsymbol{d}_k, \\ &\leq \|\boldsymbol{g}_{k+1} - \boldsymbol{g}_k\| \|\boldsymbol{d}_k\|, \\ &\leq L\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| \|\boldsymbol{d}_k\|, \\ &\leq \alpha_k L \|\boldsymbol{d}_k\|^2. \end{aligned}$$

Thus, $(\sigma - 1)\boldsymbol{g}_k^T \boldsymbol{d}_k \leq \alpha_k L\|\boldsymbol{d}_k\|^2$. ■

Lemma 2 and the subsequent theorem are necessary to facilitate the proof of global convergence properties.

Theorem 5: Let the sequences $\{\boldsymbol{x}_k\}$ and $\{\boldsymbol{d}_k\}$ be generated by the IRMIL, and α_k is computed using a strong Wolfe line search for all $k \geq 0$. Then the Zoutendijk condition is deemed fulfilled if either

$$\lim_{k\to\infty} inf \|\boldsymbol{g}_k\| = 0 \quad \text{or} \quad \sum_{k\geq 0} \frac{\|\boldsymbol{g}_k\|^4}{\|\boldsymbol{d}_k\|^2} < +\infty. \tag{53}$$

Proof: Utilising Theorem (4) and Inequalities (4 and 5) from the strong Wolfe line search,

$$f(\boldsymbol{x}_{k+1}) \leq f(\boldsymbol{x}_k) + \mu \alpha_k \boldsymbol{g}_k^T \leq f(\boldsymbol{x}_k) \leq \cdots \leq f(\boldsymbol{x}_0), \tag{54}$$

which implies $\{f(\boldsymbol{x}_k)\}_{k\geq 0}$ is a bounded sequence.

Also, Lemma (3) and Inequalities (4) and (5) yields

$$\frac{\mu(1-\sigma)}{L}\frac{(\boldsymbol{g}_k^T\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2} \leq f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1}). \quad (55)$$

Since $\{f(\boldsymbol{x}_k)\}$ is a bounded sequence produces,

$$\frac{\mu(1-\sigma)}{L}\sum_{k=0}^{\infty}\frac{(\boldsymbol{g}_k^T\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2} \leq (f(\boldsymbol{x}_0)-f(\boldsymbol{x}_1)) + (f(\boldsymbol{x}_1)-f(\boldsymbol{x}_2)) + \ldots = f(\boldsymbol{x}_0) - \lim_{k\to\infty} f(\boldsymbol{x}_k) < +\infty. \quad (56)$$

Because $\frac{\mu(1-\sigma)}{L}$ is a non-negative scalar and the direction is descent, Equation 57 holds.

$$\sum_{k=0}^{\infty}\frac{(\boldsymbol{g}_k^T\boldsymbol{d}_k)^2}{\|\boldsymbol{d}_k\|^2} \leq \sum_{k=0}^{\infty}\frac{\|\boldsymbol{g}_k\|^4}{\|\boldsymbol{d}_k\|^2} < \infty \quad (57)$$

■

In conclusion, it can be stated that the IRMIL algorithm achieves global convergence by fulfilling the Zoutendijk condition.

IV. Numerical Experiments

The MATLAB R2020b subroutine programme was utilised to test all CG methods discussed in this study. The computations were executed on an AMD Ryzen 5 3550H with Radeon Vega Mobile Gfx, operating at a frequency of 2100 Mhz, and equipped with 4 Core(s) and 8 Logical Processor(s). The computer system has a central processing unit (CPU) operating at a frequency of 2.30 GHz and a memory capacity of 8.00 GB. The computer is equipped with the Windows 10 operating system. The calculation of β_k has been modified accordingly. The methods have been executed utilising both exact line search and strong Wolfe line search, with the specific parameters $\sigma = 0.1$ and $\mu = 0.0001$, as documented in reference [40]. Furthermore, the IRMIL method utilised a parameter value of $\omega = 0.3$.

Table III summarises the numerical experiments for each method, encompassing 924 runs across 44 test problems of varying dimensions, including small and large dimensions. The test functions utilised in the study were selected from [41] and [42]. The comprehensive list of test questions can be accessed at https://tinyurl.com/mrxe4cj6 .

TABLE III
Total run for experiment

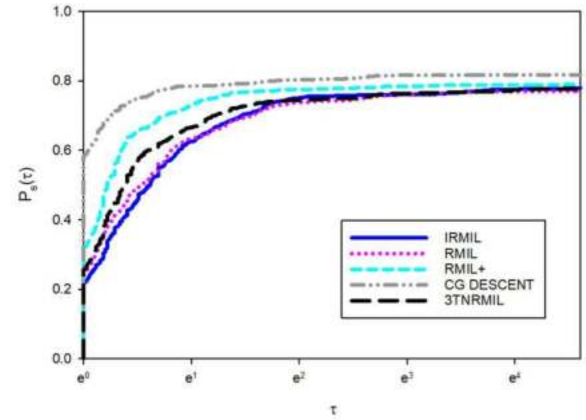
Function	Initial points	Dimension	Total run
44	4	2,4,50,100,500,1000, 10000	924

This experiment has the potential to fail in three distinct situations. The first situation involves the process being terminated if the CPU time hits 500 seconds [43] or the number of iterations (NOI) exceeds 1000. If the situation occurs, the method cannot effectively resolve the corresponding test problems and is classified as "FAIL." The second situation of failure occurs when the final values of the function diverge from the solution. The third instance of failure arises when the line search procedure cannot execute the method due to its inability to determine a positive stepsize. However, among various problem-solving approaches, a particular method can be deemed superior if it demonstrates a faster execution speed and produces a superior output function. The performance of the tested methods was compared using Dolan and Moré's performance profile method [44]. The performance profile plot depicts the algorithms' performance, with the top curve indicating the most outstanding ones. Additionally, the right-hand side of the plot serves as an indicator of the algorithms' reliability.

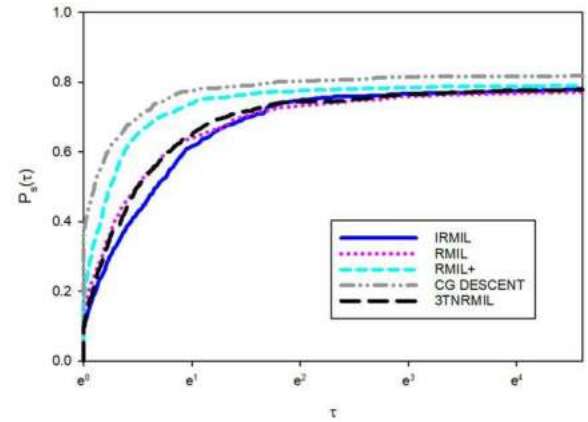
A comprehensive explanation of the numerical findings for the NOI and CPU time can be found at https://tinyurl.com/ms8u59mw and https://tinyurl.com/ychrpcfd , respectively. The performance comparisons are interpreted as follows: i.

1) Performance profile based on exact line search (Figure 1),
2) Performance profile based on strong Wolfe line search (Figure 2).

Figures 1(a) and 1(b) illustrate the comparison of IRMIL to RMIL, RMIL+, CG DESCENT and 3TNRMIL in term of NOI and CPU time, respectively under exact line search.



(a)



(b)

Fig. 1. Performance profile under exact line search due to (a) NOI and (b) CPU time.

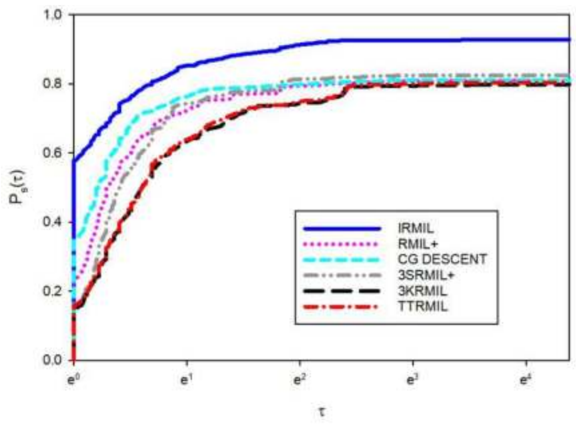
Figure 1(a) and 1(b) depicts the comparison of IRMIL, RMIL, RMIL+, CG DESCENT, and 3TNRMIL regarding the number of iterations (NOI) and central processing unit (CPU) time when the exact line search is employed. Both figures display performance profiles with almost similar shapes. This implies that the findings obtained by each method for NOI and CPU time are very constant. Upon closer examination of the left side of both figures, it is evident that the RMIL and IRMIL techniques are associated with the lowest curve. Hence, these methodologies were considered to exhibit inferior performance. The CG DESCENT method dominates the top curve, so it is determined to be the best performer. The RMIL+ and 3TNRMIL algorithms rank second and third in the exact line search. However, there are no significant distinctions in their respective performances. A closer inspection of the right-hand side of both figures reveals that all methods achieve $P_s(t)$ values in the range of 0.77 to 0.82. This result implies that the methodologies effectively addressed 77% to 82% of the issues. Additionally, all methods have successfully solved 16 out of 44 test problems (27.27%).

Meanwhile, Figures 2(a) and 2(b) depict a NOI and CPU time comparison of IRMIL, RMIL+, CG DESCENT, 3SRMIL+, 3KRMIL, and TTRMIL under a strong Wolfe line search. The superior performance of the IRMIL algorithm is evident in Figures 2(a) and 2(b), as it outperforms all other algorithms across all curves. Thus, it can be concluded that the IRMIL algorithm exhibits superior efficiency compared to RMIL+, CG DESCENT, 3SRMIL+, 3KRMIL, and TTRMIL in terms of both the NOI and CPU time. In addition, 3KRMIL and TTRMIL are shown to have similar performances since they both demonstrate the poorest performance in the figures. It is also revealed that IRMIL is unable to achieve the stationary point for certain functions. Therefore, it effectively resolves approximately 92.8% of the test functions. Following it are 3SRMIL+, CG DESCENT, RMIL+, TTRMIL, and 3KRMIL, which, respectively, solve 82.4%, 81.01%, 80.9%, 80.26%, and 79.82% of test functions. The CG techniques employed for each selected initial point have effectively resolved 18 test problems. At the same time, this study has identified the Rastrigin, NONSCOMP, Extended Freudenstein & Roth, and Extended Beale functions as unsuccessful functions for at least one conjugate gradient (CG) method utilised in this research.

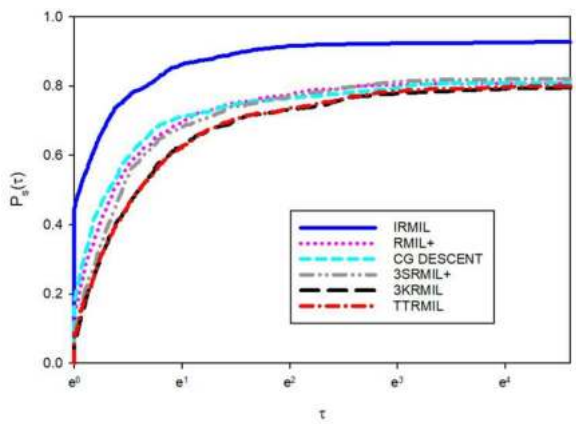
V. Image Restoration Problems

This section employs IRMIL, CG DESCENT, and RMIL+ methods, utilising the strong Wolfe line search to address image restoration issues. The parameter selection process is similar to that outlined in the Section Numerical Experiments. These issues relate to recovering a digital image that has undergone partial damage due to impulse noise. The impulse noise replaces some of the pixels in the original picture, resulting in image corruption. The specifics of the issues at hand may be obtained from [40], [45], [46].

The present study employed a set of standardised test images, namely Cat (256 x 256), Bird (256 x 256), Lena (256 x 256), and Goldhill (256 x 256), each contaminated with salt-and-pepper noise at varying levels of contamination, specifically 30%, 50%, and 90%. The images are then subjected to noise reduction by applying a median filter [46], [47]. The quality of the restored images is assessed by three metrics: CPU time, relative error (Relerr), as described in reference [48]) and peak signal to noise ratio (PSNR), as outlined in reference [49]. To conserve paper space, Figure 3 only depicts the noisy and restored images for IRMIL, RMIL+, and CG DESCENT techniques when the salt-and-paper noise ratio is 90%.



(a)



(b)

Fig. 2. Performance profile under strong Wolfe line search due to (a) NOI and (b) CPU time.

In addition, the numerical findings for each approach are listed in Table IV. The findings in Table IV show a positive correlation between the salt and pepper noise ratio and the CPU time required to restore the corrupted image. Table IV demonstrates that the suggested IRMIL approach, with the exception of a few circumstances, consumes less time while maintaining high PSNR quality. The findings indicate that the IRMIL method effectively restored the given test images with suitable quality and exhibited comparable results to other algorithms.

VI. Conclusion

The IRMIL algorithm is an improved version of RMIL that incorporates a three-term parameter and a scaled initial direction without limiting the sign of the RMIL parameters. In contrast to previous research, this study considered the RMIL CG parameter in its entirety and exhibited the global convergence characteristics of IRMIL. The IRMIL algorithm is designed to satisfy the Sufficient Decrease Condition (SDC) and the strong Wolfe line search criteria, thereby guaranteeing convergence to a stationary point. To thoroughly analyse the IRMIL method, the authors conducted tests with 44 multi-dimensional mathematical test functions of various degrees of complexity. The numerical efficacy of the algorithm was enhanced while preserving its global convergence characteristics, suggesting its viability as an alternative to the RMIL algorithm. Furthermore, IRMIL was also compared

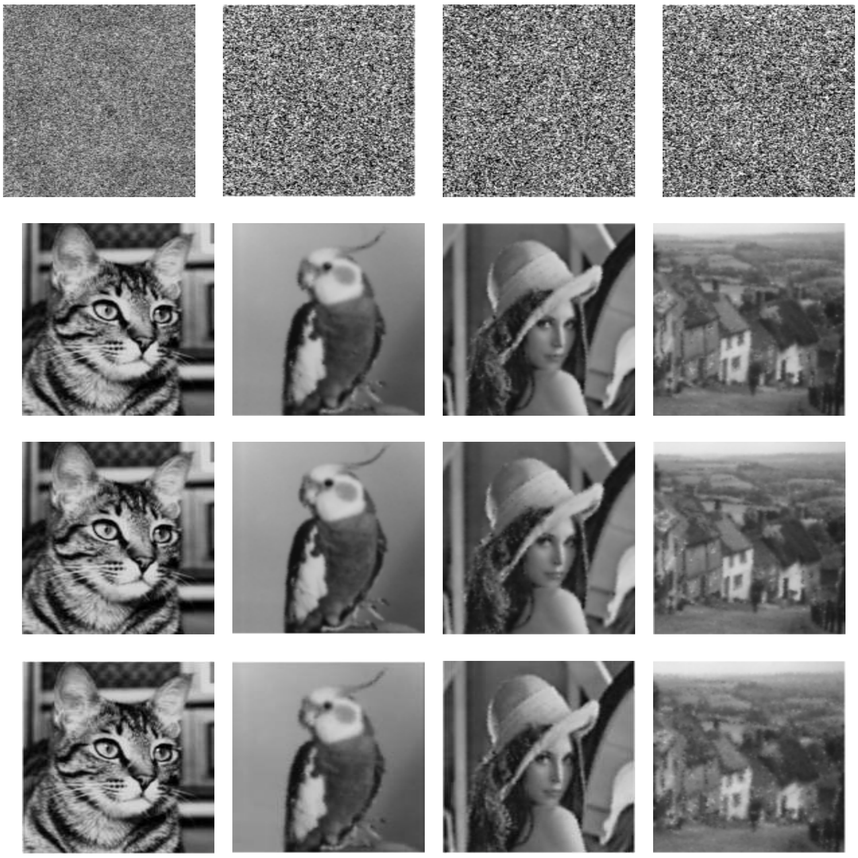


Fig. 3. Restoration of Cat, Bird, Lena and Goldhill. From top to bottom: noisy images with 90% salt-and-pepper noise, restorations obtained by the IRMIL, CG DESCENT, RMIL+ methods, respectively.

TABLE IV
NUMERICAL RESULTS OF IMAGE RESTORATION

Image	Noise Ratio (%)	IRMIL CPU/Relerr/PSNR	CG DESCENT CPU/Relerr/PSNR	RMIL+ CPU/Relerr/PSNR
Cat	0.3	22.84/0.55/28.43	**22.83**/0.58/28.50	22.93/**0.54/28.56**
	0.5	**37.70/0.87**/25.54	38.21/0.89/**25.82**	37.83/ 0.87/25.80
	0.9	**38.73/1.27/24.21**	39.07/1.28/24.19	39.66/ 1.28/24.18
Bird	0.3	8.16/0.39/40.81	8.18/**0.35**/40.93	**8.14/** 0.39/**41.02**
	0.5	**13.74**/0.56/37.42	14.28/0.55/**37.46**	13.80/**0.53**/37.46
	0.9	29.55/1.40/30.62	**23.28/1.15**/32.10	48.93/ 1.69/**30.23**
Lena	0.3	10.86/0.96/**33.57**	**10.76**/1.05/33.24	10.80/**0.90**/33.55
	0.5	17.81/**1.33/30.47**	17.46/1.36/30.23	**17.45**/ 1.38/30.34
	0.9	43.35/2.68/**25.55**	42.59/2.85/25.17	**41.64/2.56**/25.55
Goldhill	0.3	**10.67**/0.89/**32.17**	10.71/0.92/32.13	11.32/**0.88**/32.11
	0.5	**17.89**/1.44/**29.48**	18.18/**1.39**/29.35	18.14/ 1.42/29.33
	0.9	57.95/**2.93/25.68**	**30.39**/2.97/25.57	37.67/ 3.81/23.91

with RMIL, RMIL+, CG DESCENT, 3SRMIL+, 3KRMIL, TTRMIL, and 3TNRMIL. The results prove that the IRMIL algorithm exhibits potential as an optimisation methodology for addressing UO issues.

REFERENCES

[1] P. Wolfe, "Convergence Conditions For Ascent Methods. II: Some Corrections," *SIAM Review*, vol. 13, no. 2, pp. 185–188, 1971.

[2] Z. Zhang and X. Chen, "A Conjugate Gradient Method For Distributed Optimal Control Problems With Nonhomogeneous Helmholtz Equation," *Applied Mathematics and Computation*, vol. 402, p. 126019, 2021.

[3] M. Hestenes and E. Stiefel, "Method of Conjugate Gradients for Solving Linear Systems," *J. Res. Nat. Bur. Standarts*, vol. 49, pp. 409 – 436, 1952.

[4] R. Fletcher and C. M. Reeves, "Function Minimization by Conjugate Gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 01 1964.

[5] 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. R1, pp. 35–43, 1969.

[6] B. T. Polyak, "The Conjugate Gradient Method In Extremal Problems," *USSR Computational Mathematics and Mathematical Physics*, vol. 9, no. 4, pp. 94–112, 1969.

[7] J. W. Daniel, "The Conjugate Gradient Method For Linear And Nonlinear Operator Equations," *SIAM Journal on Numerical Analysis*, vol. 4, no. 1, pp. 10–26, 1967.

[8] R. Fletcher, "Practical Methods of Optimization," *New York*, vol. 80, p. 4, 1987.

[9] 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.

[10] Y. Dai and Y. Yuan, "A Nonlinear Conjugate Gradient Method With A Strong Global Convergence Property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.

[11] W. W. Hager and H. Zhang, "A Survey Of Nonlinear Conjugate Gradient Methods," *Pacific journal of Optimization*, vol. 2, no. 1, pp. 35–58, 2006.

[12] ——, "A New Conjugate Gradient Method With Guaranteed Descent And An Efficient Line Search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170–192, 2005.

[13] Y. Dai and Y. Yuan, "Nonlinear Conjugate Gradient Methods," *Shanghai Science and Technology Publisher, Shanghai*, 2000.

[14] M. Rivaie, M. Mamat, L. W. June, and I. Mohd, "A New Conjugate Gradient Coefficient For Large Scale Nonlinear Unconstrained Optimization," *Int. Journal of Math. Analysis*, vol. 6, no. 23, pp. 1131–1146, 2012.

[15] ——, "A New Class Of Nonlinear Conjugate Gradient Coefficients With Global Convergence Properties," *Applied Mathematics and Computation*, vol. 218, no. 22, pp. 11 323–11 332, 2012.

[16] N. Mohamed, "An Extension of RMIL and Hybrid Conjugate Gradient Method with Global Convergence," Ph.D. dissertation, Universiti Sultan Zainal Abidin, 2017.

[17] 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, pp. 1152–1163, 2015.

[18] N. Zullpakkal, N. 'Aini, N. H. A. Ghani, N. S. Mohamed, N. Idalisa, and M. Rivaie, "Covid–19 Data Modelling Using Hybrid Conjugate Gradient Method," *Journal of Information and Optimization Sciences*, vol. 43, no. 4, pp. 837–853, 2022. [Online]. Available: https://doi.org/10.1080/02522667.2022.2060610

[19] N. S. Mohamed, M. Mamat, M. Rivaie, and S. M. Shaharudin, "A New Hybrid Coefficient Of Conjugate Gradient Method," *Indonesian J Elec Eng & Comp Sci*, vol. 18, no. 3, pp. 1454–1463, 2020.

[20] M. Malik, M. Mamat, S. S. Abas, I. M. Sulaiman *et al.*, "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.

[21] A. V. Mandara, M. Mamat, M. Waziri, and M. A. Mohamed, "A Class Of Conjugate Gradient Parameters And Its Global Convergence For Solving Unconstrained Optimization," *Far East Journal of Mathematical Sciences (FJMS)*, vol. 106, pp. 43–58, 2018.

[22] Z. Dai, "Comments On A New Class Of Nonlinear Conjugate Gradient Coefficients With Global Convergence Properties," *Applied Mathematics and Computation*, vol. 276, pp. 297–300, 2016.

[23] 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.

[24] 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.

[25] N. Zull, "New Conjugate Gradient Methods using Strong Wolfe Line Search for Estimating Dividend Rate," Ph.D. dissertation, Universiti Sultan Zainal Abidin, 2019.

[26] X. Jiang and J. Jian, "Improved Fletcher–Reeves and Dai–Yuan Conjugate Gradient Methods With The Strong Wolfe Line Search," *Journal of Computational and Applied Mathematics*, vol. 348, pp. 525–534, 2019.

[27] L. Zhang, "An Improved Wei–Yao–Liu Nonlinear Conjugate Gradient Method For Optimization Computation," *Applied Mathematics and computation*, vol. 215, no. 6, pp. 2269–2274, 2009.

[28] F. Rahpeymaii, K. Amini, T. Allahviranloo, and M. R. Malkhalifeh, "A New Class Of Conjugate Gradient Methods For Unconstrained Smooth Optimization And Absolute Value Equations," *Calcolo*, vol. 56, no. 1, pp. 1–28, 2019.

[29] W. Khadijah, M. Rivaie, and M. Mamat, "A Three–Term Conjugate Gradient Method Under The Strong-Wolfe Line Search," in *AIP Conference Proceedings*, vol. 1870. AIP Publishing LLC, 2017, p. 040056.

[30] I. M. Sulaiman, M. Mamat, A. E. Owoyemi, P. L. Ghazali, M. Rivaie, and M. Malik, "The Convergence Properties Of Some Descent Conjugate Gradient Algorithms For Optimization Models," *Journal of Mathematics and Computer Science*, vol. 22, no. 3, pp. 204–215, 2021. [Online]. Available: http://dx.doi.org/10.22436/jmcs.022.03.02

[31] I. Sulaiman, M. Mamat, A. Umar, K. Kamfa, and E. N. Madi, "Some Three–Term Conjugate Gradient Algorithms With Descent Condition For Unconstrained Optimization Models," *J. Adv. Res. Dyn. Control Syst*, vol. 12, pp. 2494–2501, 2020.

[32] J. Liu, Y. Feng, and L. Zou, "Some Three–Term Conjugate Gradient Methods With The Inexact Line Search Condition," *Calcolo*, vol. 55, no. 2, pp. 1–16, 2018.

[33] N. H. Fadhilah, M. Rivaie, F. Ishak, and N. Idalisa, "New Three–Term Conjugate Gradient Method with Exact Line Search," *MATEMATIKA: Malaysian Journal of Industrial and Applied Mathematics*, vol. 36, no. 3, pp. 197–207, 2020.

[34] 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.

[35] N. Andrei, *Nonlinear Conjugate Gradient Methods for Unconstrained Optimization*. Berlin, Heidelberg: Springer International Publishing, 2020.

[36] M. Fatemi, "A New Efficient Conjugate Gradient Method for Unconstrained Optimization," *Journal of Computational and Applied Mathematics*, vol. 300, pp. 207–216, 2016.

[37] P. L. Toint, "Non–Monotone Trust–Region Algorithms For Nonlinear Optimization Subject To Convex Constraints," *Mathematical programming*, vol. 77, no. 3, pp. 69–94, 1997.

[38] G. Zoutendijk, "Nonlinear Programming, Computational Methods," *Integer and Nonlinear Programming*, pp. 37–86, 1970.

[39] G. Yuan, S. Lu, Z. Wei *et al.*, "A Line Search Algorithm For Unconstrained Optimization," *Journal of Software Engineering and Applications*, vol. 3, no. 05, p. 503, 2010.

[40] G. Yu, J. Huang, and Y. Zhou, "A Descent Spectral Conjugate Gradient Method for Impulse Noise Removal," *Applied Mathematics Letters*, vol. 23, no. 5, pp. 555–560, 2010.

[41] N. Andrei, "An Unconstrained Optimization Test Functions Collection," *Advanced Modeling and Optimization*, vol. 10, no. 1, pp. 141 – 161, 2008.

[42] M. Jamil and X. Yang, "A Literature Survey Of Benchmark Functions For Global Optimisation Problems," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 4, pp. 150 – 194, 2013.

[43] Y. Narushima, H. Yabe, and J. A. Ford, "A Three–Term Conjugate Gradient Method With Sufficient Descent Property For Unconstrained Optimization," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 212–230, 2011.

[44] E. D. Dolan and J. J. Moré, "Benchmarking Optimization Software With Performance Profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.

[45] R. H. Chan, C.-W. Ho, and M. Nikolova, "Salt–And–Pepper Noise Removal By Median–Type Noise Detectors And Detail–Preserving Regularization," *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1479–1485, 2005.

[46] S. Babaie-Kafaki, N. Mirhoseini, and Z. Aminifard, "A Descent Extension of a Modified Polak–Ribière–Polyak Method with Application in Image Restoration Problem," *Optimization Letters*, pp. 1–17, 2022.

[47] D. Esan, P. A. Owolawi, and C. Tu, "Anomalous Detection in Noisy Image Frames using Cooperative Median Filtering and KNN," *IAENG International Journal of Computer Science*, vol. 49, no. 1, pp. 1–10, 2022.

[48] H. Esmaeili, S. Shabani, and M. Kimiaei, "A New Generalized Shrinkage Conjugate Gradient Method For Sparse Recovery," *Calcolo*, vol. 56, no. 1, pp. 1–38, 2019.

[49] A. Bovik, *Handbook of Image and Video Processing*, ser. Communications, Networking and Multimedia. New York: Elsevier Science, 2010.