Int. J. Appl. Math. Comput. Sci., 2024, Vol. 34, No. 2, 225–233
DOI: 10.61822/amcs-2024-0016

A DESCENT GENERALIZED RMIL SPECTRAL GRADIENT ALGORITHM FOR OPTIMIZATION PROBLEMS

IBRAHIM M. SULAIMAN [a,b], P. KAELO [c], RUZELAN KHALID [a], MOHD KAMAL M. NAWAWI [a,*]

[a]Institute of Strategic Industrial Decision Modelling
Northern University of Malaysia (UUM)
Sintok, 06010, Kedah, Malaysia
e-mail: i.mohammed.sulaiman@uum.edu.my

[b]Faculty of Education and Arts
Sohar University
3111 Al Jamiah Street, Sohar 311, Oman

[c]Department of Mathematics
University of Botswana
4775 Notwane Rd, Private Bag UB00704, Gaborone, Botswana

This study develops a new conjugate gradient (CG) search direction that incorporates a well defined spectral parameter while the step size is required to satisfy the famous strong Wolfe line search (SWP) strategy. The proposed spectral direction is derived based on a recent method available in the literature, and satisfies the sufficient descent condition irrespective of the line search strategy and without imposing any restrictions or conditions. The global convergence results of the new formula are established using the assumption that the gradient of the defined smooth function is Lipschitz continuous. To illustrate the computational efficiency of the new direction, the study presents two sets of experiments on a number of benchmark functions. The first experiment is performed by setting uniform SWP parameter values for all the algorithms considered for comparison. For the second experiment, the study evaluates the performance of all the algorithms by considering the exact SWP parameter values used for the numerical experiments as reported in each work. The idea of these experiments is to study the influence of parameters in the computational efficiency of various CG algorithms. The results obtained demonstrate the effect of the parameter value on the robustness of the algorithms.

Keywords: optimization models, spectral CG algorithm, global convergence, line search strategy.

1. Introduction

Unconstrained optimization involves minimizing or maximizing a function $f(x)$ of many variables over a defined set. These problems are often related to engineering, economics, and sciences (Xia *et al.*, 2015). Generally, a numerical method and mathematical programming software such as Maple or Matlab are often employed to perform all necessary computations that would lead to the solution of these complex problems.

One of the widely used and famous numerical algorithms considered for minimizing a function is the conjugate gradient (CG) formula, which is applied to optimization problems of the form

$$\min f(x), \quad x \in \mathbb{R}^n.$$

Here the gradient $\nabla f(x) = g(x)$ of the differentiable $f : \mathbb{R}^n \to \mathbb{R}$ is available. The CG algorithm is characterized by excellent convergence properties and low memory requirements, and produces an iterative sequence,

$$x_{k+1} = x_k + s_k, \quad s_k = \alpha_k d_k, \quad k \geq 0, \tag{1}$$

with d_k denoting the search direction computed along with the step size α_k. To obtain the step size, a line search scheme, either inexact or exact, is needed. The search direction is usually calculated as

$$d_0 = -g_0, \quad d_k = -g_k + \beta_k d_{k-1}, \quad k \geq 1. \tag{2}$$

*Corresponding author

Here, the CG coefficient β_k is a scalar that makes the CG methods differ.

The convergence result and implementation of the CG formulas are achieved using the line search procedure (Hager and Zhang, 2006). For inexact line searches, α_k is said to satisfy the following standard (weak) Wolfe (WWP) conditions:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \tag{3}$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k, \tag{4}$$

or the strong Wolfe (SWP) conditions (3) and

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

with $0 < \delta < \sigma < 1$. Notice that α_k that satisfies the SWP conditions also satisfies the standard Wolfe conditions.

Due to their simplicity and strong global convergence properties, and the ever-evolving nature of optimization problems, more CG methods have been (and continue to be) developed in the literature in order to effectively solve these optimization problems (see Hager and Zhang, 2006; Hestenes and Stiefel, 1952; Dai *et al.*, 2000; Sabiu *et al.*, 2024; Collignon and Gijzen, 2010; 2004). Most of these modifications are in the form of two term CG formulas (see Hager and Zhang, 2006; Sulaiman *et al.*, 2019), three term CG formulas (see Zhang *et al.*, 2006; Liu *et al.*, 2018; Sulaiman *et al.*, 2022), or spectral CG methods and their variants (see Andrei, 2007a; 2007b; Amini and Faramarzi, 2023; Novkaniza *et al.*, 2022), and are differentiated by the structure of their search direction d_k.

In this study we are interested in the spectral CG algorithm whose direction is given as

$$d_0 = -g_0, \quad d_k = -\theta_k g_k + \beta_k d_{k-1}, \quad k \geq 1,$$

where θ_k is the spectral coefficient. Two important components of any spectral method are the spectral coefficient θ_k and the CG coefficient β_k. These components influence the numerical performance and convergence analysis of the spectral formula (Awwal *et al.*, 2021). The first class of spectral CG formulas were presented by Birgin and Martinez (2001) as

$$\beta_k^1 = \frac{(\theta_k y_{k-1} - s_{k-1})^T g_k}{s_{k-1}^T y_{k-1}},$$

$$\beta_k^2 = \frac{\theta_k}{\alpha_{k-1}\theta_{k-1}} \beta_k^{PRP},$$

where $y_{k-1} = g_k - g_{k-1}$ and $s_{k-1} = x_k - x_{k-1}$. If $\theta_k = 1, \forall k$, then β_k^1 becomes the Perry formula (Perry, 1978). Moreover, if $\theta_k = \theta_{k-1} = 1, \forall k$, and exact minimization criteria are applied, then β_k^2 becomes the classical PRP formula defined as (Polyak, 1969; Hager and Zhang, 2006)

$$\beta_k^{\text{PRP}} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}.$$

The PRP method is among the earliest and widely studied CG algorithms for unconstrained optimization models. This method is characterized by robust and efficient numerical performance. However, the convergence of this formula is not guaranteed using some line search methods (Hager and Zhang, 2006). To overcome this drawback, various variants of the PRP method have been developed; see the works of Babaie-Kafaki and Ghanbari (2017) or Hu *et al.* (2022) and the references therein.

Recently, another conjugate gradient type coefficient was presented by Rivaie *et al.* (2012) with the formula

$$\beta_k^{\text{RMIL}} = \frac{g_k^T y_{k-1}}{\|d_{k-1}\|^2}, \tag{6}$$

where the denominator in β_k^{PRP} is replaced by $\|d_{k-1}\|^2$. This method was shown to satisfy the sufficient descent condition

$$g_k^T d_k \leq -\lambda \|g_k\|^2, \quad \lambda > 0,$$

under the exact line search condition, and numerical results on benchmark functions were presented to demonstrate the efficacy of the scheme. Lately, several researchers have developed different variants of the RMIL method. One of the latest modification was presented by Salihu *et al.* (2023) who extended the idea to define a new spectral CG method that incorporates the RMIL formula (6) and a defined spectral parameter as

$$d_k = -\theta_k g_k + \beta_k^{\text{RMIL}} d_{k-1}, \quad k \geq 1,$$

$$\theta_k = \eta + \frac{\|y_{k-1}\|}{\|d_{k-1}\|},$$

where $\eta > 0$. The convergence result of the formula was discussed using the Wolfe conditions, and numerical results show that the method is promising.

Dai (2016) observed an inequality wrongly used in the convergence result of (6) and defined a modification to suit the convergence results as

$$\beta_k^{\text{RMIL+}} = \begin{cases} \dfrac{g_k^T y_{k-1}}{\|d_{k-1}\|^2} & \text{if } 0 \leq g_k^T g_{k-1} \leq \|g_k\|^2, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

and showed that, with β_k defined as (7), the convergence results of Rivaie *et al.* (2012) hold. An important condition for the convergence of (7) is

$$0 \leq g_k^T g_{k-1} \leq \|g_k\|^2.$$

However, with this condition, the convergence of (7) will not hold for general functions. This drawback has led to numerous studies aimed at tackling this problem.

In this work, we present a new spectral direction for a CG method that incorporates a well defined spectral parameter (Section 2). This direction requires

α_k to satisfy the famous SWP strategy. In Section 3, we show that the defined d_k possesses the sufficient descent property irrespective of the line search procedure, and without imposing any restrictions or conditions. We further discuss the convergence results under suitable assumptions. Section 4 presents the numerical performance findings of the new scheme to support the excellent theoretical properties. The last section summaries the achieved results and contains concluding remarks.

2. New method and its algorithm

Recently, Awwal *et al*. (2023) proposed a modification of the RMIL+ CG formula to ensure that the condition $0 \leq g_k^T g_{k-1} \leq \|g_k\|^2$ holds for general functions. Their modification is presented as

$$\beta_k^{\text{ARMIL+}} = \frac{\min\{|g_k^T y_{k-1}|,\ \|g_k\|^2\}}{\|d_{k-1}\|^2}. \tag{8}$$

This new modification takes the classical RMIL+ (7) as a special case because, for $g_k^T g_{k-1} < 0$, the method will compute $\beta_k^{\text{ARMIL+}} = \frac{\|g_k\|^2}{\|d_{k-1}\|^2}$. However, for $g_k^T g_{k-1} \geq 0$, the method will compute (8) as either $\beta_k^{\text{ARMIL+}} = \frac{|g_k^T y_{k-1}|}{\|d_{k-1}\|^2}$ or $\beta_k^{ARMIL+} = \frac{\|g_k\|^2}{\|d_{k-1}\|^2}$. Another special feature of this formula is that the following property, which is very instrumental in the convergence analysis of RMIL+ (7), also holds for (8):

$$0 \leq \beta_k^{\text{ARMIL+}} \leq \frac{\|g_k\|^2}{\|d_{k-1}\|^2}. \tag{9}$$

Computational findings on some unconstrained optimization and real-life application problems show that ARMIL+ is efficient and superior compared with other variants of RMIL+.

Motivated by the excellent theoretical features of the above discussion and the fact that the condition $0 \leq g_k^T g_{k-1} \leq \|g_k\|^2$ now holds for general functions, we propose a spectral modification of ARMIL+ (Awwal *et al*., 2023) as follows. By multiplying the spectral direction d_k (15) by g_k^T, we have

$$\begin{aligned} g_k^T d_k &= -\theta_k \|g_k\|^2 + \beta_k g_k^T d_{k-1} \\ &= \frac{g_{k-1}^T d_{k-1}}{\|d_{k-1}\|^2} \|g_k\|^2 \psi_k, \end{aligned} \tag{10}$$

where

$$\psi_k = \beta_k \frac{\|d_{k-1}\|^2}{g_{k-1}^T d_{k-1}} \frac{g_k^T d_{k-1}}{\|g_k\|^2} - \frac{\|d_{k-1}\|^2}{g_{k-1}^T d_{k-1}} \theta_k.$$

From (10), we have that

$$\frac{g_k^T d_k}{\|g_k\|^2} = \frac{g_{k-1}^T d_{k-1}}{\|d_{k-1}\|^2} \psi_k. \tag{11}$$

Algorithm 1. SGR algorithm with the strong Wolfe strategy.

Step 1. Initialization: $x_0 \in \mathbb{R}^n$, termination tolerance $\epsilon > 0$.

Step 2. Evaluate g_k. If $\|g_k\| = 0$, terminate the iteration process.

Step 3. If $k = 0$, set $d_0 := -g_0$, otherwise,

$$d_k = -\theta_k g_k + \beta_k d_{k-1},\ k \geq 1, \tag{15}$$

with θ_k and β_k following from (14) and (8), respectively.

Step 4. Determine α_k such that (3) and (5) are satisfied.

Step 5. Calculate the new point via (1).

Step 5. Return to Step 2 with $k := k + 1$.

Choosing $\psi_k = 1$, for every $k \geq 1$, will imply that (10) and (11) give

$$\frac{g_k^T d_k}{\|g_k\|^2} = \frac{g_{k-1}^T d_{k-1}}{\|d_{k-1}\|^2} = \cdots = \frac{g_0^T d_0}{\|d_0\|^2} = -1, \tag{12}$$

and this reduces to

$$g_k^T d_k = -\|g_k\|^2, \quad \forall k \geq 0. \tag{13}$$

This result shows that the choice of θ_k is crucial in determining the descent property of d_k for every spectral method. It is further concluded that the spectral direction d_k will always satisfy the descent property if the parameter θ_k is chosen to satisfy $\psi_k \equiv 1$. It is interesting to note that this condition will hold irrespective of the line search method employed. From the above formulation, we derive the following spectral formula:

$$\theta_k = -\frac{g_{k-1}^T d_{k-1}}{\|d_{k-1}\|^2} + \beta_k \frac{g_k^T d_{k-1}}{\|g_k\|^2}, \tag{14}$$

where $k \geq 1$ and $\psi \equiv 1$.

The execution process of the new formula, denoted as SGR, is described as Algorithm 1.

3. Convergence analysis

This section discusses the theoretical findings of the proposed SGR method. To achieve these results, we need the following important assumptions that are very crucial in establishing the convergence results of most CG formulas.

Assumption A.

1. The level set $\Omega_0 = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$ of $f(x)$ is bounded.

2. For the smooth function f, the gradient $g(x)$ is Lipschitz continuous in some neighborhood N of Ω_0. That is, N is an open convex set containing Ω_0, such that for some constant $L > 0$,

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall x, y \in N. \quad (16)$$

From Assumption A, we have that, there exist positive constants b and γ satisfying

$$\|x - y\| \leq b, \quad \forall x, y \in \Omega_0,$$

and

$$\|g(x_k)\| \leq \gamma, \quad \forall x_k \in \Omega_0. \quad (17)$$

Lemma 1. *Suppose that $\{x_k\}$ is computed via Algorithm 1, where the direction d_k is a descent one and Assumption A holds. If α_k satisfies either the WWP conditions* (3) *and* (4)*, or the SWP conditions* (3) *and* (5)*, then we have*

$$\alpha_k \geq \frac{(1-\sigma)|g_k^T d_k|}{L\|d_k\|^2}. \quad (18)$$

Proof. Subtracting $g_k^T d_k$ from both the sides of (4) and applying the condition of Lipschitz continuity (16) will produce

$$\begin{aligned}(\sigma - 1)g_k^T d_k &\leq (g_{k+1} - g_k)^T d_k \\ &\leq \|g_{k+1} - g_k\|\|d_k\| \\ &\leq L\|x_{k+1} - x_k\|\|d_k\| \\ &= L\alpha_k\|d_k\|^2.\end{aligned}$$

Since d_k is a descent direction and $\sigma < 1$, we have that (18) holds. ∎

We now state the Zoutendijk lemma (Zoutendijk, 1970), which is often needed for the global convergence of most line search formulas.

Lemma 2. *Let Assumption A hold, and d_k be a descent direction. If α_k satisfies the Wolfe conditions* (3) *and* (4)*, or* (3) *and* (5)*, then*

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \quad (19)$$

Proof. From the Wolfe line search condition (3) and (18), we obtain

$$\begin{aligned}f(x_k) - f(x_k + \alpha_k d_k) &\geq -\delta\alpha_k g_k^T d_k \\ &\geq \delta\frac{(1-\sigma)(g_k^T d_k)^2}{L\|d_k\|^2}.\end{aligned}$$

Adding up the above for $k \geq 0$ and using Assumption A, we have that

$$\delta\frac{(1-\sigma)}{L}\sum_{k=0}^{\infty}\frac{(g_k^T d_k)^2}{L\|d_k\|^2} \leq f(x_0) - f^*,$$

where $f^* = \lim_{k\to\infty} f(x_k)$. Thus, the Zoutendijk condition (19) is achieved and, therefore, this completes the proof. ∎

Notice that, from (13) and Lemma 2, the Zoutendijk condition can also be given as

$$\sum_{k=0}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2} \leq +\infty. \quad (20)$$

From the descent condition (13), we conclude that, for all $k \geq 0$,

$$\|d_k\| \geq \|g_k\|. \quad (21)$$

We now present the global convergence result of the new spectral SGR CG formula using the SWP conditions (3) and (5).

Theorem 1. *Let $\{x_k\}$ denote the sequence obtained via Algorithm 1. Then*

$$\liminf_{k\to\infty}\|g_k\| = 0. \quad (22)$$

Proof. We show this by contradiction, that is, we suppose that (22) is not true. Then it follows that there is a positive constant c such that

$$\|g_k\| \geq c, \quad \forall\, k \geq 0. \quad (23)$$

From (15), we have that

$$d_k = -\theta_k g_k + \beta_k d_{k-1},$$

where $\beta_k = \beta_k^{\text{ARMIL+}}$ and θ_k is given by (14). Note here that, using (12), (14) can also be expressed as

$$\theta_k = 1 + \beta_k\frac{g_k^T d_{k-1}}{\|g_k\|^2}. \quad (24)$$

Therefore, from the above, (9), (15), (17), (21), (23), (24), and by using the Cauchy–Schwarz inequality, the following holds:

$$\begin{aligned}\|d_k\| &= \| - \theta_k g_k + \beta_k d_{k-1}\| \\ &= \left\| -\left(1 + \beta_k\frac{g_k^T d_{k-1}}{\|g_k\|^2}\right)g_k + \beta_k d_{k-1}\right\| \\ &\leq \left(1 + \frac{\|g_k\|}{\|d_{k-1}\|}\right)\|g_k\| + \frac{\|g_k\|^2}{\|d_{k-1}\|} \\ &\leq \left(1 + 2\frac{\|g_k\|}{\|g_{k-1}\|}\right)\|g_k\| \\ &\leq \left(1 + \frac{2\gamma}{c}\right)\gamma \\ &= M. \end{aligned} \quad (25)$$

The above inequality (25) implies that

$$\frac{1}{\|d_k\|^2} \geq \frac{1}{M^2},$$

which gives

$$\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} = +\infty. \tag{26}$$

But, from the Zoutendijk condition (20) and from (23), we obtain

$$c^4 \sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} \leq \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty,$$

which contradicts (26). Hence, (22) holds. ■

4. Numerical experiments

To support the excellent theoretical features possessed by the proposed method, this section reports two sets of numerical experiments on a number of unconstrained benchmark functions which can be accessed through the hyperlink `https://acrobat.adobe.com/id/urn:aaid:sc:AP:586acd2b-c42c-4b76-8e9b-cb1a6da57aa2`. Most of these unconstrained benchmark problems used for these experiments are adopted from Andrei (2008) with the dimension (DIM) chosen between 2 and 100,000. The efficiency of the new formula is established by comparing it with the results of other existing algorithms with similar characteristics, and this is done based on some comparison metrics, including the CPU time, the number of function evaluations (NOF) and number of iterations (NOI). For the comparison, we consider the following methods:

- a spectral PRP CG direction (SPRP) by Wan *et al.* (2011) with the following direction and spectral parameter:

  $$d_k = -\theta_k g_k + \beta_k d_{k-1},$$

  where $\beta_k = \beta_k^{\text{PRP}}$ (Polyak, 1969; Polak and Ribiere, 1969) and

  $$\theta_k = \frac{d_{k-1}^T y_{k-1}}{\|g_{k-1}\|^2} - \frac{d_{k-1}^T g_k g_k^T g_{k-1}}{\|g_k\|^2}$$

- a two-term CG method (ARMIL) by Awwal *et al.* (2023) with the direction

  $$d_k = -g_k + \beta_k d_{k-1}, \quad k \geq 1,$$

  with β_k following from (8);

- a sufficiently descent spectral CG method (NSRMIL) by Salihu *et al.* (2023) whose direction follows from (2), with β_k defined as (6), and

  $$\theta_k = \tau + \frac{y_{k-1}^T y_{k-1}}{\|d_{k-1}\|^2}$$

- the classical spectral method (SCG) by Birgin and Martinez (2001).

- an efficient spectral method (sRMILp) by Awwal *et al.* (2021) whose direction follows from (2) with β_k defined as (7).

All algorithms for this experiment are written in MATLAB R2019a, and the computations were performed on an Intel Core i7 PC with 16 GB RAM under SWP conditions. For the stopping condition, we command the program to terminate if $\|g_k\| \leq 10^{-6}$ or if the iteration exceeds 2000, and record this as a failure, which we denote with $(***)$.

For the first experiment, the computation was conducted under the strong Wolfe line search (3) and (5), with the parameter values defined as $\delta = 10^{-4}$ and $\sigma = 0.16$. These same values were used in computing the results of all the other methods, which are based on the NOI, NOF and CPU time. A detailed description of this performance is given in Table 2 which can be accessed through the hyperlink `https://acrobat.adobe.com/id/urn:aaid:sc:AP:91c03434-0b50-4cd6-a74f-b3a6f8e1c3db`. These results demonstrate the performance of the new and existing algorithms. By examining these results, it is easy to see that there was a strong competition between the proposed SGR, ARMIL, and sRMILp algorithms. This might be due to the fact that the three algorithms have similar structures and are in the same category. On the other hand, the classical SCG and SPRP methods recorded the least performance on the discussed set of unconstrained optimization problems. Despite the competition from the other algorithms, the proposed SGR method was still able to outperform all the methods on all the metrics.

For the second experiment, the computations were also performed under the SWP conditions (3) and (5). For the proposed approach, we maintained $\delta = 10^{-4}$ and $\sigma = 0.16$. However, for the other methods, we considered the exact values used for the methods as reported in the respective papers. That is, for ARMIL, we used $\delta = \sigma = 10^{-4}$; for NSRMIL, we set the values to $\delta = \sigma = 10^{-4}$; for sRMILp, the values from the study were given as $\delta = \sigma = 10^{-4}$; for SPRP, we used $\delta = 0.45$ and $\sigma = 0.75$; lastly, for SCG, $\delta = 0.45$ and $\sigma = 0.75$. For a detailed description of the obtained results, we refer the reader to Table 3 that is available through the hyperlink `https://acrobat.adobe.com/id/urn:aaid:sc:AP:9f19ebf7-c01d-4aeb-9c89-0c1c1b2804cc`. From the above results, it is obvious that the exact parameter values do not so much influence the numerical performance of the methods. Also, it can be noted, that in the early iteration process, all methods maintained uniformed performance, however, as the iterations progressed, the proposed method slightly outperformed the other algorithms. On the other hand,

the classical SCG and SPRP algorithms are also the least performers in this result category. This poor performance might be associated with the complexity of the search direction or high parameter values used in the study.

Next, we analyze the numerical performance from Tables 2 and 3, presented in the above hyperlinks using a tool introduced by Dolan and Moré (2002). The authors defined a process that can be used to evaluate and compare the efficiency of a solver (s) on a set of problems (p). For every solver s and problem p, the scheme derives the computational cost (t) by considering n_s solvers and n_p problems as follows: $t_{p,s} =$ the cost of solving problem p by solvers s.

Dolan and Moré (2002) further defined a measure that would be used to compare the efficiency of all the solvers based on the cost of computing $t_{p,s}$. This baseline is called the performance ratio and is evaluated as

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

Additionally, it defines the distribution function as

$$\rho_s(\tau) = \frac{1}{n_p}\text{size}\{p \in P : \log_2(r_{p,s}) \leq \tau\}.$$

The above process generates graphs for all solvers $s \in S$ based on the data used. These graphs, known as the performance profile graphs, can be used to analyze the fraction $\rho_s(\tau)$ of the problems considered for all solvers within the factor of $\tau \geq 0$. Any algorithm that records a higher $\rho_s(\tau)$ value is considered more efficient for the given τ value. This implies that an algorithm whose curve dominates the top spot can be regarded as the most efficient algorithm.

For this study, the performance profile plot for the NOI, NOF and CPU time is based on two experiments. The first set of plots, which includes Figs. 1(a), 2(a) and 3(a), are generated from the numerical results obtained using uniform parameter values, while the second set of figures, including Figs. 1(b), 2(b) and 3(b), are plotted from the outcome obtained using the exact parameter values as reported in the respective study.

These graphs can be further used to analyze the numerical performance discussed above. It is obvious that the figures depict the results earlier discussed. An important point to observe in these figures is that all the algorithms were competing at the initial stage; however, as the factor τ increases, the new method outperformed the other methods on all the metrics. This shows that the new formula is very competitive and robust for the problems considered.

5. Conclusion

In this work, we introduced a spectral conjugate gradient method based on the generalized RMIL conjugate gradient parameter. This presented method satisfies the decent condition independently of any line search used. We established its global convergence under the strong Wolfe (SWP) line search conditions and tested its efficacy on a set of benchmark problems from the literature. The results show that the method is very competitive compared with some other methods from the literature with a similar structure. Considering the ever-evolving nature of optimization problems, further work is needed to come up with more sophisticated conjugate gradient methods that can be applied to solve these kinds of problems.

Acknowledgment

This research was supported by the Ministry of Higher Education (MoHE) of Malaysia through the Fundamental Research Grant Scheme (FRGS/1/2022/STG06/UUM/02/6).

References

Amini, K. and Faramarzi, P. (2023). Global convergence of a modified spectral three-term CG algorithm for nonconvex unconstrained optimization problems, *Journal of Computational and Applied Mathematics* **417**: 114630.

Andrei, N. (2007a). A scaled BFGS preconditioned conjugate gradient algorithm for unconstrained optimization, *Applied Mathematics Letters* **20**(6): 645–650.

Andrei, N. (2007b). Scaled conjugate gradient algorithms for unconstrained optimization, *Computational Optimization and Applications* **38**: 401–416.

Andrei, N. (2008). An unconstrained optimization test functions collection, *Advanced Modeling and Optimization* **10**(1): 147–161.

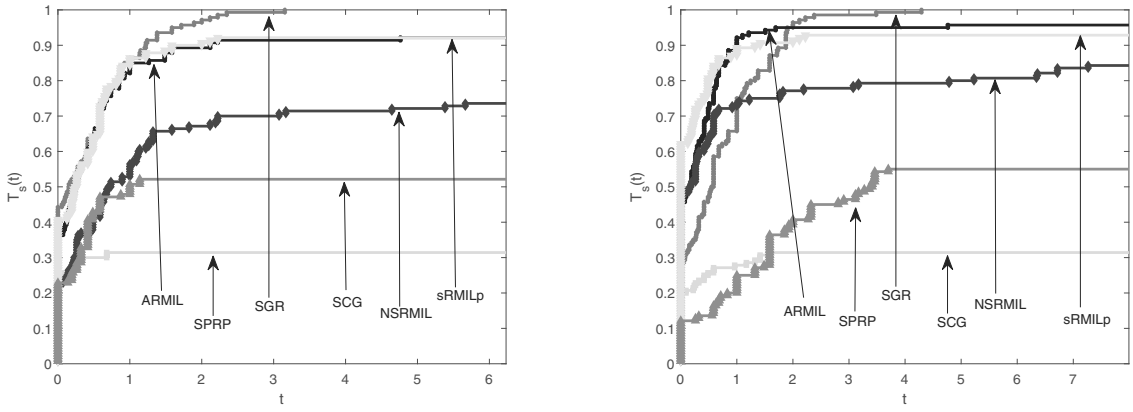
Awwal, A., Sulaiman, M., Maulana, M., Kumam, P. and Sitthithakerngkiet, K. (2021). Spectral RMIL+ conjugate gradient method for unconstrained optimization with applications in portfolio selection and motion control, *IEEE Access* **9**: 75398–75414.

Awwal, A., Wang, L., Kumam, P., Sulaiman, M., Salisu, S., Salihu, N. and Yodjai, P. (2023). Generalized RMIL conjugate gradient method under the strong Wolfe line search with application in image processing, *Mathematical Methods in the Applied Sciences* **2023**: 1–13.

Babaie-Kafaki, S. and Ghanbari, R. (2017). An optimal extension of the Polak–Ribiere–Polyak conjugate gradient method, *Numerical Functional Analysis and Optimization* **38**(9): 1115–1124.

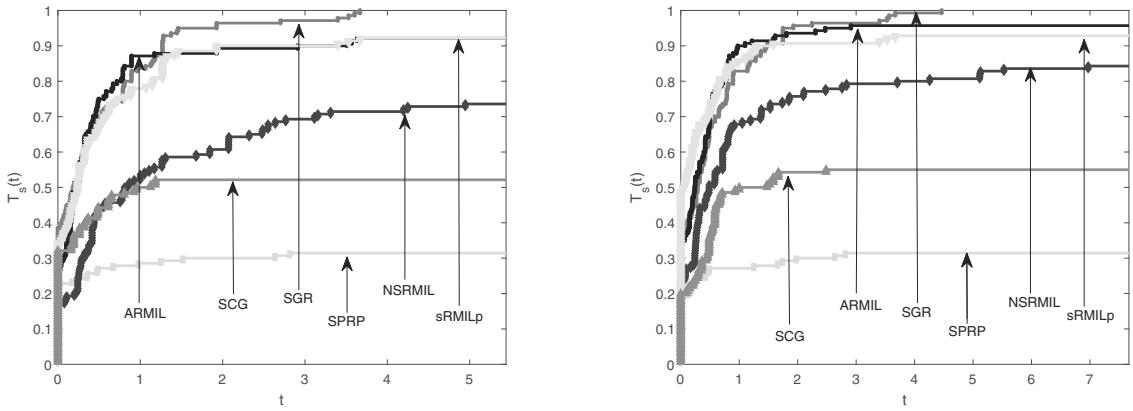
Birgin, E. and Martinez, J. (2001). A spectral conjugate gradient method for unconstrained optimization, *Applied Mathematics and Optimization* **43**: 117–128.

Collignon, T.P. and van Gijzen, M.B. (2010). Two implementations of the preconditioned conjugate gradient method on heterogeneous computing grids, *International Journal of Applied Mathematics and Computer Science* **20**(1): 109–121, DOI: 10.2478/v10006-010-0008-4.



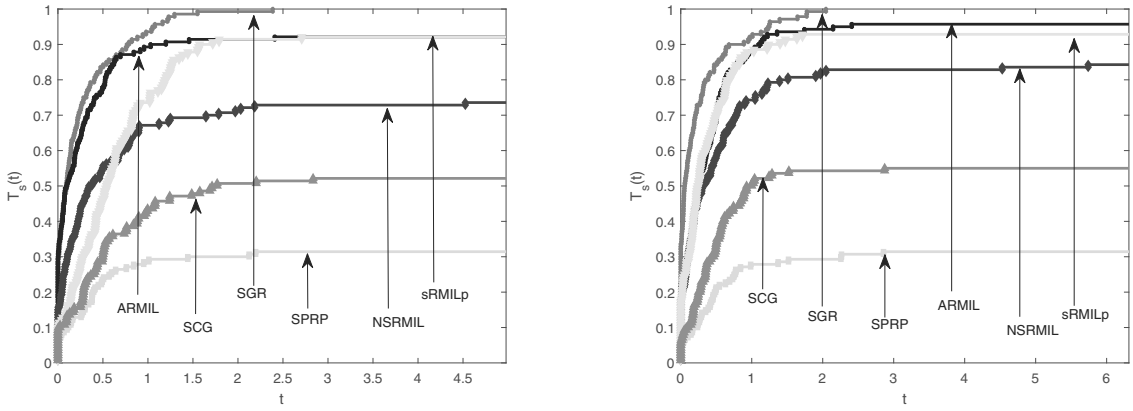
(a) (b)

Fig. 1. Performance based on the number of iterations: using uniform parameter values (a), using exact parameter values reported in each study (b).



(a) (b)

Fig. 2. Performance based on the number of function evaluations: using uniform parameter values (a), using exact parameter values reported in each study (b).



(a) (b)

Fig. 3. Performance based on the CPU time: using uniform parameter values (a), using exact parameter values reported in each study (b).

Dai, Y., Han, J., Liu, G., Sun, D., Yin, H. and Yuan, Y. (2000). Convergence properties of nonlinear conjugate gradient methods, *SIAM Journal on Optimization* **10**(2): 345–358.

Dai, Z. (2016). Comments on a new class of nonlinear conjugate gradient coefficients with global convergence properties, *Applied Mathematics and Computation* **276**: 297–300.

Dolan, E.D. and Moré, J.J. (2002). Benchmarking optimization software with performance profiles, *Mathematical Programming* **91**(2): 201–213.

Hager, W. and Zhang, H. (2006). A survey of nonlinear conjugate gradient methods, *Pacific Journal of Optimization* **2**(1): 35–58.

Hestenes, M. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems, *Journal of Research of the National Institute of Standards and Technology* **49**(6): 409–435.

Hu, Q., Zhang, H. and Chen, Y. (2022). Global convergence of a descent PRP type conjugate gradient method for nonconvex optimization, *Applied Numerical Mathematics* **173**: 38–50.

Koko, J. (2004). Newton's iteration with a conjugate gradient based decomposition method for an elliptic PDE with a nonlinear boundary condition, *International Journal of Applied Mathematics and Computer Science* **14**(1): 13–18.

Liu, J., Feng, Y. and Zou, L. (2018). Some three-term conjugate gradient methods with the inexact line search condition, *Calcolo* **55**(16): 16.

Novkaniza, F., Malik, M., Sulaiman, I. and Aldila, D. (2022). Modified spectral conjugate gradient iterative scheme for unconstrained optimization problems with application on COVID-19 model, *Frontiers in Applied Mathematics and Statistics* **8**: 1014956.

Perry, A. (1978). A modified conjugate gradient algorithm, *Operations Research* **26**(6): 1073–1078.

Polak, B. and Ribiere, G. (1969). Note sur la convergence de methodes de directions conjugues, *Mathematical Modelling and Numerical Analysis* **16**(3): 35–43.

Polyak, B.T. (1969). The conjugate gradient method in extreme problems, *USSR Computational Mathematics and Mathematical Physics* **9**(4): 94–112.

Rivaie, M., Mamat, M., June, L. and Mohd, I. (2012). A new class of nonlinear conjugate gradient coefficients with global convergence properties, *Applied Mathematics and Computation* **218**(22): 11323–11332.

Sabiu, J., Ibrahim, I.S., Kaelo, P., Malik, M. and Saadi, A.K. (2024). An optimal choice Dai–Liao conjugate gradient algorithm for unconstrained optimization and portfolio selection, *AIMS Mathematics* **9**(1): 642–664.

Salihu, N., Kumam, P., Awwal, A., Sulaiman, I. and Seangwattana, T. (2023). The global convergence of spectral RMIL conjugate gradient method for unconstrained optimization with applications to robotic model and image recovery, *PLoS ONE* **18**(3): e0281250.

Sulaiman, I., Malik, M., Awwal, A., Kumam, P., Mamat, M. and Al-Ahmad, S. (2022). On three-term conjugate gradient method for optimization problems with applications on Covid-19 model and robotic motion control, *Advances in Continuous and Discrete Models* **2022**(1): 1–22.

Sulaiman, I., Sukono, S., Supian, S. and Mamat, M. (2019). New class of hybrid conjugate gradient coefficients with guaranteed descent and efficient line search, *7th International Conference on Global Optimization and Its Application (ICoGOIA 2018), Bali, Indonesia*, p. 012021.

Wan, Z., Yang, Z. and Wang, Y. (2011). New spectral PRP conjugate gradient method for unconstrained optimization, *Applied Mathematics Letters* **24**(1): 16–22.

Xia, Z., Wang, X., Sun, X. and Wang, Q. (2015). A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data, *IEEE Transactions on Parallel and Distributed Systems* **27**(2): 340–352.

Zhang, L., Zhou, W. and Li, D.-H. (2006). A descent modified Polak Ribiere Polyak conjugate gradient method and its global convergence, *IMA Journal of Numerical Analysis* **26**(4): 629–640.

Zoutendijk, G. (1970). Nonlinear programming computational methods, *in* J. Abadie (Ed.), *Integer and Nonlinear Programming*, North-Holland, Amsterdam, pp. 37–86.

Ibrahim Mohammed Sulaiman is from Kano State, Nigeria, but is currently an international senior lecturer in Malaysia. He holds a PhD degree in applied mathematics from UniSZA, Malaysia (2018). He is an author of more than 80 research papers included in ISI/Scopus. His main research cover include optimization methods, fuzzy nonlinear problems, and fluids.

P. Kaelo is currently a senior lecturer in the Department of Mathematics, University of Botswana. He holds a PhD in computational mathematics from the University of Witwatersrand, South Africa. His research interests are in optimization, control theory, and numerical analysis. He has published a number of research papers in journals indexed in Scopus and the Web of Science.

Ruzelan Khalid is a senior lecturer of decision science at Universiti Utara Malaysia (UUM), Malaysia. Prior to joining UUM, he had been an associate researcher at the National CAD/CAM Center, SIRIM Sdn. Bhd, Malaysia. His areas of research interests include discrete event simulation, queuing networks, stochastic processes and optimization.

Mohd Kamal Bin Mohd Nawawi is a senior lecturer of operations research at the School of Quantitative Sciences, Universiti Utara Malaysia (UUM). Prior to joining UUM, he had worked as an engineer in industry for more than five years. His research interests include decision analysis, operational research, discrete-event simulation, system dynamics modeling, queuing analysis, lean manufacturing, sustainable manufacturing, and knowledge-based systems. To date he has authored or co-authored nearly 70 publications in national and international journals as well as conference proceedings/seminars.

Received: 3 October 2023
Revised: 14 December 2023
Re-revised: 13 January 2024
Accepted: 14 February 2024