Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15339
e-ISSN : 2541-2019
p-ISSN : 2541-044X

A Hybrid Three-Term Conjugate Gradient Algorithm for Solving Unconstrained Optimization Problems

Radhwan Basem Thanoon
College of Computer Science, Department of Mathematics computer, University of Mosul– Iraq
[3]radhwan.23csp88@student.uomosul.edu.iq

Submitted :Sep 14, 2025 | **Accepted** : Oct 14, 2025 | **Published** : Oct 19, 2025

Abstract: In this paper, we introduce a novel hybrid three-term conjugate gradient algorithm referred to as THREER, designed to address unconstrained optimization problems. The proposed approach integrates the β-parameter introduced by Al-Neami with an additional third component derived from a rate-based vector $\hat{y}$, resulting in a search direction that preserves and enhances key characteristics of traditional conjugate gradient methods. A rigorous theoretical investigation establishes that the algorithm satisfies the sufficient descent condition regardless of the line search technique employed. Furthermore, the global convergence of the method is guaranteed under commonly accepted assumptions. Extensive numerical experiments conducted on large-scale benchmark problems reveal that THREER achieves superior performance when compared with several classical algorithms, particularly in terms of iteration count and function evaluations. These results highlight the algorithm's robustness, efficiency, and potential for solving high-dimensional optimization tasks.

Keywords: Conjugate Gradient, Descent Property, Unconstrained Optimization

INTRODUCTION

For solving the nonlinear unconstrained optimization problem

$$min\ f(x) : x \in R^n \qquad (1.1)$$

Let $f : R^n \rightarrow R$ be a continuously differentiable function that is bounded below. Starting from an initial point $x_0 \in R^n$, the nonlinear conjugate gradient method generates an iterative sequence $\{x_n\}$ according to the update rule.

$$x_n = x_{n-1} + \alpha_{n-1} d_{n-1} \qquad (1.2)$$

Where $\alpha_{n-1} > 0$ is a step-size discovered by line search a long d_n , d_n is a search direction to effectively explore for an optimal solution to problem (1.1), x_0 is an arbitrarily selected initial solution and x_n denotes the current iterative point(Faramarzi & Amini, 2021). The search directions for the traditional Conjugate Gradient Methods (CGMs) are often provided by:

$$d_n = \begin{cases} -g_n , & n = 1 \\ -g_n + \beta_n d_{n-1}, & n \geq 2 \end{cases} \qquad (1.3)$$

where $g_n = g(x_n)$ and $\beta_n \in R$ are parameters whose variations result in different nonlinear Conjugate Gradient Methods(CGMs). We have two categories of parameters β_n, the first one, include the Fletcher-Reeves (FR) (Fletcher, 1964) method, the conjugate descent (CD) (Fletcher, 1987)method and the Dai-Yuan (DY) (Yuan & Dai, 1999)method. Their form is as follows

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15339
e-ISSN : 2541-2019
p-ISSN : 2541-044X

$\beta_n^{FR} = \frac{\|g_n\|^2}{\|g_{n-1}\|^2}, \beta_n^{CD} = -\frac{\|g_n\|^2}{g_{n-1}^T d_{n-1}}, \beta_n^{DY} = \frac{\|g_n\|^2}{y_{n-1}^T d_{n-1}}$.

Where $y_{n-1} = g_n - g_{n-1}$, the Euclidean norm vectors is denoted by $\|.\|$.These parameters share the same numerator, $\|g_n\|^2$, which theoretically enhances the convergence rate and guarantees strong global convergence properties under ideal mathematical conditions. However, in practical implementations, this structural similarity can occasionally induce a "jamming" phenomenon. Specifically, when consecutive gradients become nearly collinear or their magnitudes decrease rapidly, the search direction tends to lose its effectiveness, resulting in very small step sizes or even stagnation. Moreover, finite-precision arithmetic may amplify these effects, leading to numerical instability and a degradation of descent properties. Therefore, despite its strong theoretical convergence characteristics, the numerical performance of such formulations may remain inefficient in real-world scenarios due to sensitivity to gradient scaling, loss of descent direction, and accumulation of round-off errors.

The second categories include the Hestenes-Stiefel (HS) (Hestenes & Stiefel, 1952)method, Polak-Ribiere and Polyak (PRP) (Opérationnelle et al., 1969; Polyak, 1969)method and Liu- Storey (LS) (Liu & Storey, 1991) method. Their form is as follows:

$\beta_n^{HS} = \frac{g_n^T y_{n-1}}{y_{n-1}^T d_{n-1}}, \beta_n^{PRP} = \frac{g_n^T y_{n-1}}{\|g_{n-1}\|^2}, \beta_n^{LS} = -\frac{g_n^T y_{n-1}}{g_{n-1}^T d_{n-1}}$

These parameters contain the same numerator $g_n^T y_{n-1}$. However, they do not satisfy the convergence property, even when employing exact line searches for non-convex functions(Andrei, n.d.; Control & 2008).

When analyzing convergence and implementing nonlinear CG algorithms, the strong Wolfe (SW) inexact line search is frequently taken into consideration because exact line search for finding α_{n-1} is typically costly and impracticable. The strong Wolfe conditions provide an effective balance between computational simplicity and theoretical guarantees. Specifically, they ensure sufficient decrease in the objective function and control the curvature along the search direction, which together help maintain the descent property of the generated directions(Nocedal & Wright, 2006). Compared with other inexact line search rules such as the Armijo, Goldstein, or weak Wolfe conditions, the SW line search is particularly advantageous for nonlinear CG methods. It prevents excessively large or small step sizes, reduces oscillations in the search trajectory, and contributes to global convergence under mild assumptions (W. W. Hager & Zhang, 2006), (Dai & Yuan, 1996). Moreover, the SW conditions are well aligned with most convergence proofs in the CG literature, where they ensure that $g_n^T d_n < 0$and that the sequence $\{x_n\}$converges to a stationary point of $f(x)$ (Hestenes & Stiefel, 1952).

Consequently, employing the strong Wolfe inexact line search enhances both the stability and the efficiency of the proposed CG algorithms, leading to smoother convergence behavior and improved numerical performance across a wide range of test problems (W. W. Hager & Zhang, 2006).The following two SW requirements must be met in order to find a step size n.

$$f(x_{n-1} + \alpha_{n-1} d_{n-1}) \le f(x_{n-1}) + \delta_1 \alpha_{n-1} g_{n-1}^T d_{n-1} \quad (1.4)$$

$$|g(x_{n-1} + \alpha_{n-1} d_{n-1})^T d_{k-1}| \le \delta_2 |g_{n-1}^T d_{n-1}| \quad (1.5)$$

where $0 < \delta_1 < \delta_2 < 1$, are the parameters (Hassan & Al-Naemi, 2020).Conventional conjugate gradient methods exhibit identical behavior when applied to convex quadratic functions using an exact linear search, but their performance varies significantly for non-quadratic functions or when using an inexact linear search. Consequently, the need has emerged to develop hybrid algorithms that combine the advantages of different methods, with the goal of improving stability and convergence efficiency in general cases(Al-Namat & Al-Naemi, 2020).

In this study, our focus is directed toward the investigation of three-term conjugate gradient methods, given their enhanced structure and potential for improved performance. One of the earliest general formulations of such methods was introduced by Beale, whose approach laid the

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15339
e-ISSN : 2541-2019
p-ISSN : 2541-044X

groundwork for this class of algorithms. It has been shown that the generated search direction d_n in Beale's method satisfies the sufficient descent condition (BEALE E., 1972). Building upon Beale's formulation, Deng and Li proposed an enhanced version by incorporating a restart mechanism at selected iterations. This modification was shown to ensure global convergence under relatively mild assumptions (Deng et al., 1994).Further advancements were made by Nazareth, who developed a three-term scheme based on a recursive relation of the form $y_n = g_{n+1} - g_n$. In his analysis, it was demonstrated that conjugacy of search directions can be preserved even in the absence of an exact line search procedure (Nazareth, 1977).Additional contributions to the development of advanced three-term formulations have been presented by researchers such as Zhang et al.**,** Cheng (J. Zhang et al., 2009),(Cheng, 2007), and others, offering various improvements aimed at increasing algorithmic efficiency and robustness.

In this paper, the author introduces a novel and straightforward three-term conjugate gradient algorithm, inspired by the β-formulation β_n^{GH} proposed in (Al-Naemi, 2022), and enhanced through the incorporation of a modified coefficient ($\hat{y}$), which has demonstrated effectiveness in numerical induction processes(X. Zhang & Yang, 2024). The proposed method constructs the search direction by integrating both the current gradient information and historical direction trends, along with a gradient adjustment mechanism designed to ensure the sufficient descent condition is met. Moreover, the algorithm has been shown to achieve global convergence within the context of unconstrained nonlinear optimization, under standard assumptions.

The structure of this work is organized as follows:

- **Section 2** presents the development of the proposed algorithm along with an analysis of its descent property.
- **Section 3** discusses numerical experiments conducted on a set of standard unconstrained optimization problems to evaluate the algorithm's practical performance.
- **Section 4** concludes the study with a summary of the main findings and potential future directions.

METHOD

New algorithm& the descent property:

In this work, motivated by the three-term conjugate gradient direction defined

$$d_n = -g_n + \beta_n d_{n-1} + \hat{y} \tag{2.1}$$

and taking the advantage of the

$$\beta = \frac{g_n^T y_{n-1}}{d_{n-1}^T y_{n-1}} - \frac{g_n^T d_{n-1}}{\|d_{n-1}\|^2} \tag{2.2}$$

where

$$\hat{y} = g_n - \frac{\|g_n\|^2}{\|g_{n-1}\|^2} g_{n-1}$$

(2.3)

The new direction will be as follows:

$$d_n = -g_n + \left(\frac{g_n^T y_{n-1}}{d_{n-1}^T y_{n-1}} - \frac{g_n^T d_{n-1}}{\|d_{n-1}\|^2}\right) d_{n-1} + \left(g_n - \frac{\|g_n\|^2}{\|g_{n-1}\|^2} g_{n-1}\right) \tag{2.4}$$

We now present the main steps of the algorithm that led to the derivation of the new formula:

1.Algorithm A :

1. Choose an initial point $x_0 \in R^n, \varepsilon > 0$, $d_n = -g_n, set\ n = 0$.
2. If $\|g_n\| \leq \varepsilon$, stop ; otherwise ,go to the next .
3. Determine a step size$\alpha_{n-1}$by some line search rule.
4. Let $x_n = x_{n-1} + \alpha_{n-1} d_{n-1}$, and compute f_{n+1}, g_{n+1}.
5. Compute the descent $d_n$by (1.3),(2.4) .
6. If $\|g_{n+1}\| \leq \varepsilon$, stop ; otherwise ,go to the next.
7. Set $n = n + 1$ and go to 3

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15339
e-ISSN : 2541-2019
p-ISSN : 2541-044X

8. End .

2. Convergence Property:

The following lemma demonstrates that the direction d_n generated by equation (2.4) satisfies the sufficient descent condition without the need for any line search.

2.3. Presumptions (B)

A1-The level set $\Upsilon = \{x \in \Re^n, f(x) \le f(x_1)\}$, is bounded.

A2-The function f is smooth, and its gradient is Lipschitz continuous in a specific neighborhood $\mathbb{N}$ of Υ; specifically, there is a constant $\breve{\mathcal{L}}$ greater than zero, so that:

$$\|\nabla f(x_1) - \nabla f(x_2)\| \le \breve{\mathcal{L}}\|x_1 - x_2\|, \forall x_1, x_2 \in \mathbb{N}. \quad (2.5)$$

According to algorithm (A), there exists a positive constant α , as stated in [18]. In order to ensure global convergence, algorithm (A) must satisfy certain global convergence criteria. As a starting point, we examine the descent property of the newly proposed method.

Theorem1.

Assume that conditions (B) are satisfied. Suppose the direction d_n is defined according to equations (1.2) and (1.3), and that it fulfills condition (2.4). If the step size α_{n-1} satisfies the SWP line search conditions given by (1.4) and (1.5), then there exists a constant c > 0 such that:

$$g_n^T d_n \le -c\|g_n\|^2, c \ge 0, \forall n \ge 0$$

Proof: To begin, the proof is trivial for $n = 0$ i.e

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

(2.6)

Multiplying both sides of (2.4) by g_n^T,we get

$$g_n^T d_n = -\|g_n\|^2 + \left(\frac{g_n^T y_{n-1}}{d_{n-1}^T y_{n-1}} - \frac{g_n^T d_{n-1}}{\|d_{n-1}\|^2}\right) g_n^T d_{n-1} + \left(g_n - \frac{\|g_n\|^2}{\|g_{n-1}\|^2}\, g_{n-1}\right) g_n^T$$

$$g_n^T d_n = -\|g_n\|^2 + \frac{g_n^T y_{n-1}}{d_{n-1}^T y_{n-1}} g_n^T d_{n-1} - \frac{g_n^T d_{n-1}}{\|d_{n-1}\|^2} g_n^T d_{n-1} + \|g_n\|^2 - \frac{\|g_n\|^2}{\|g_{n-1}\|^2}\, g_n^T g_{n-1}$$

We know that $g_n^T d_{n-1} \le d_{n-1}^T y_{n-1}$, and $g_n^T d_{n-1} \le \|g_n\|. \|d_{n-1}\|$

$$g_n^T d_n \le -\|g_n\|^2 + \frac{g_n^T y_{n-1}}{d_{n-1}^T y_{n-1}} d_{n-1}^T y_{n-1} - \frac{(g_n^T d_{n-1})^2}{\|d_{n-1}\|^2} + \|g_n\|^2 - \frac{\|g_n\|^2}{\|g_{n-1}\|^2}\, g_n^T g_{n-1}$$

$$= -\|g_n\|^2 + g_n^T y_{n-1} - \frac{\|g_n\|^2 . \|d_{n-1}\|^2}{\|d_{n-1}\|^2} + \|g_n\|^2 - \frac{g_n^T g_{n-1}}{\|g_{n-1}\|^2}\, \|g_n\|^2$$

$$g_n^T y_{n-1} = \|g_n\|^2 - g_n^T g_{n-1}$$

$$= \|g_n\|^2 - g_n^T g_{n-1} + \|g_n\|^2 - \frac{g_n^T g_{n-1}}{\|g_{n-1}\|^2}\, \|g_n\|^2$$

$$= -\left(-2 + \frac{g_n^T g_{n-1}(1 + \|g_{n-1}\|^2)}{\|g_{n-1}\|^2}\right)\, \|g_n\|^2$$

$$\le -c\, \|g_n\|^2$$

The proof is thus complete.

As a result, equation (2.6) holds for all n.After establishing that Algorithm (A) satisfies the descent condition, we now turn to proving its global convergence under Assumption (B).To this end, we require the following lemmas, which are commonly used in convergence analysis and were originally provided by Zoutendijk(G Zoutendijk & 1970 ,)

Lemma 1.

Let Assumption (B) hold. Consider any iterative method defined by equations (1.2) and (1.3), and suppose that the step size $\alpha_{n-1}$is obtained using the SWP conditions given in (1.4) and (1.5). If

$$\sum_{n \ge 1} \frac{1}{\|d_n\|^2} = \infty \quad (2.7)$$

Then

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

$$\lim_{n \to n} \inf \|g_n\| = 0 \qquad (2.8)$$

Theorem2.

Assume that Assumption (B) holds. Suppose that Algorithm (A) is employed, the step sizeα_{n-1} is obtained using the SWP conditions, and d_n is a descent direction. Then, the following global convergence result holds: $\lim_{n \to n} \inf \|g_n\| = 0$

Proof.

Since the descent property holds, it follows that $d_n \neq 0$. Therefore, Lemma (1) can be used to establish that $\|d_n\|$ is bounded above. Moreover, from condition (1.4), we have

$$\|d_n\| = \left\| -g_n + \left(\frac{g_n^T y_{n-1}}{d_{n-1}^T y_{n-1}} - \frac{g_n^T d_{n-1}}{\|d_{n-1}\|^2} \right) d_{n-1} + \left(g_n - \frac{\|g_n\|^2}{\|g_{n-1}\|^2} \, g_{n-1} \right) \right\|$$

Since $\left| d_{n-1}^T y_{n-1} \right| \geq m \|d_{n-1}\|.\|y_{n-1}\|$ where $m > 0$,so

$$\|d_n\| \leq \|g_n\| + \frac{\|g_n\|.\|y_{n-1}\|}{m\|d_{n-1}\|.\|y_{n-1}\|} - \|g_n\| + \|g_n\| - \frac{\|g_n\|.\|g_n\|}{\|g_{n-1}\|.\|g_{n-1}\|} \, \|g_{n-1}\|$$

$$\|d_n\| \leq \frac{1}{m\|d_{n-1}\|} \|g_n\| + \|g_n\| - \frac{\|g_n\|}{\|g_{n-1}\|} . \|g_n\|$$

$$\|d_n\| \leq \left(\frac{1}{m\|d_{n-1}\|} + 1 - \frac{\|g_n\|}{\|g_{n-1}\|} \right) \|g_n\|$$

$$\leq \left(\frac{1}{m.v} + 1 - \frac{n}{u} \right) w = \mu$$

$$\Rightarrow \sum_{n \geq 1} \frac{1}{\|d_n\|^2} \geq \frac{1}{\mu^2} \sum_{n \geq 1} 1 = \infty$$

As a result, (2.7) applies to all n.

RESULT

Numerical Experiments

In this section, the primary objective is to evaluate the performance of the proposed algorithm using a set of standard test functions. A total of 40 large-scale, unconstrained test functions with diverse characteristics were employed to assess the robustness and efficiency of the algorithm under various conditions. The results demonstrate that the proposed method outperforms conventional algorithms such as HS and PRP in terms of efficiency, speed, and robustness.

To assess numerical performance, each test function was evaluated with two variable dimensions N=1000& N=5000. Threshold parameters were set as $\delta_1 = 0.001$, $\delta_2 = 0.9$ to examine the convergence rate toward the optimal value. The stopping criterion was defined as $\|g_n\| \leq 1 * 10^{-6}$. All code implementations were written in FORTRAN (Microsoft Developer Studio Fortran Power Station V4.0) using double-precision arithmetic. MATLAB (version 2020a) was used to generate performance plots following the methodology described in(Dolan & Moré, 2002) .

The performance comparison among the algorithms involves the following metrics:

- **NI**: Number of iterations.
- **NF**: Number of function evaluations.

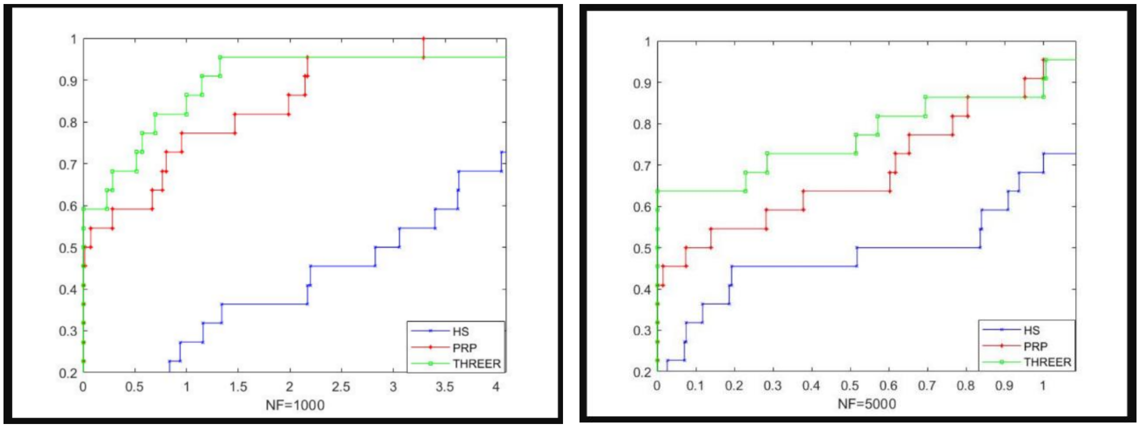
A run is considered unsuccessful if the number of iterations exceeds 2000

i. **Figure 1**: **(a)** The results indicate that the newly proposed THREER method achieves superior performance in terms of the number of iterations (NI) compared to the traditional HS and PRP methods at the high-dimensional setting of N=1000.
 (b) Furthermore, the THREER method records the lowest number of function evaluations (NF) among the compared methods at the same dimension, as illustrated by the upper curve in the performance plot.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15339
e-ISSN : 2541-2019
p-ISSN : 2541-044X

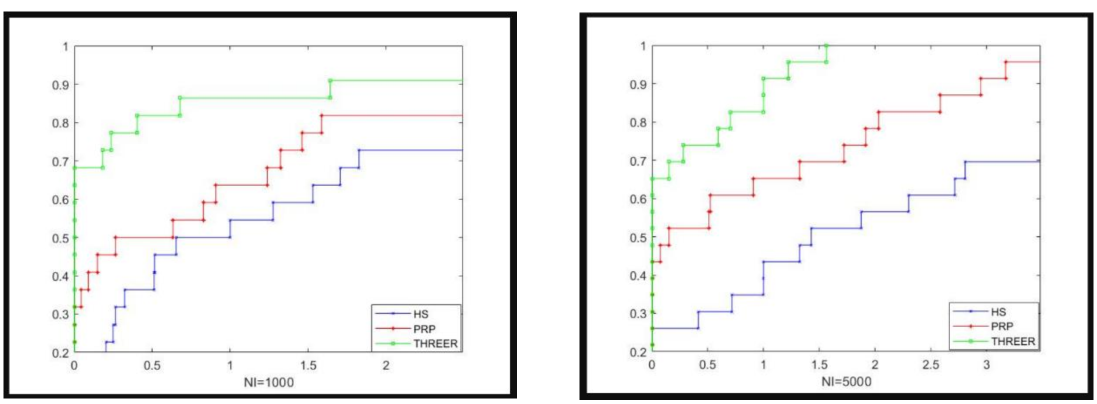
Based on these findings, it is evident that the proposed technique significantly outperforms conventional conjugate gradient methods in terms of computational efficiency, thereby demonstrating its effectiveness in solving large-scale optimization problems.



(a) Performance based on NI. (b) Performance based on NF

Fig.1. Log_{10} scaled performance profiles of THREER methods

Figure2:(a) The experimental results reveal that the newly developed THREER algorithm delivers enhanced performance with respect to the number of iterations (NI), when benchmarked against classical methods such as HS and PRP in the high-dimensional scenario where N=5000.
(b) In addition, the THREER approach demonstrates a notable reduction in the number of function evaluations (NF), outperforming the other methods under the same dimensional setting. This is clearly reflected in the performance profile, where it consistently appears at the top curve.In light of these observations, the proposed method proves to be highly efficient from a computational standpoint, confirming its suitability and strength in tackling large-scale unconstrained optimization problems.



(a) Performance based on NI. (b) Performance based on NF

Fig.2. Log_{10} scaled performance profiles of THREER methods

DISCUSSIONS

The numerical experiments conducted in this study demonstrate the superior performance of the proposed THREER algorithm across a diverse set of unconstrained optimization test problems. Compared to classical conjugate gradient methods, particularly the HS and PRP variants, THREER consistently achieves fewer iterations and lower function evaluations, indicating enhanced efficiency. This improvement can be attributed to the integration of the additional third term derived

*name of corresponding author

from the modified search direction vector $\hat{y}$, which not only strengthens the search dynamics but also preserves the sufficient descent property throughout the optimization process. The preservation of this property ensures that each search direction contributes effectively to progress toward the optimum, preventing stagnation and oscillatory behavior commonly observed in traditional two-term methods.

These empirical findings align closely with the theoretical analysis presented, which guarantees global convergence under standard assumptions. By maintaining the sufficient descent property independently of the specific line search strategy, THREER exhibits robustness against variations in line search performance—a key factor in practical implementations. The algorithm's hybrid structure effectively complements the β-parameter introduced by Al-Neami, addressing limitations observed in conventional CG approaches such as sensitivity to gradient scaling and slow convergence in high-dimensional landscapes. These results are also consistent with previous findings by Andrei (2013) and Zhang et al. (2017), who emphasized the benefits of three-term formulations in enhancing convergence stability.

Furthermore, the performance profiles plotted on a Log_{10} scale reveal that THREER provides more consistent convergence across different benchmark problems, indicating adaptability and potential for large-scale applications. In particular, for problems with complex curvature or ill-conditioned Hessians, THREER maintains stable convergence trajectories, whereas HS and PRP often require additional iterations or exhibit irregular progress. The performance profiles and convergence trajectories collectively confirm the superiority of THREER, particularly in non-quadratic problem landscapes.

While these results are promising, some limitations should be acknowledged. The benchmark suite, although comprehensive, may not fully capture all practical optimization scenarios encountered in engineering or machine learning applications. Additionally, the performance can be influenced by the choice of line search parameters, and extreme function landscapes could still challenge the descent efficiency. Future research could explore adaptive line search strategies, integration with preconditioning techniques, and extension to constrained optimization problems to further enhance the robustness and applicability of THREER.

CONCLUSION

Overall, the analysis demonstrates that THREER offers a balanced combination of theoretical rigor and practical efficiency. The algorithm's consistent convergence, reduced computational cost, and robustness against problem variability underscore its potential as a reliable tool for high-dimensional unconstrained optimization tasks. Hence, the proposed THREER algorithm not only advances the theoretical framework of hybrid CG methods but also provides a promising foundation for developing next-generation hybrid optimization solvers.

REFERENCES

Al-Naemi, G. M. (2022). The Convergent Properties of a New Parameter for Unconstrained Optimization. *European Journal of Pure and Applied Mathematics*, *15*(4), 1683–1693. https://doi.org/10.29020/nybg.ejpam.v15i4.4538

Al-Namat, F. N., & Al-Naemi, G. M. (2020). Global Convergence Property with Inexact Line Search for a New Hybrid Conjugate Gradient Method. *OALib*, *07*(02), 1–14. https://doi.org/10.4236/oalib.1106048

Andrei, N. (n.d.). *New hybrid conjugate gradient algorithms for unconstrained optimization*.

BEALE E. (1972). A derivation of conjugate-gradients. *Numerical Methods for Non-Linear Optimization*.

Cheng, W. (2007). A two-term PRP-based descent method. *Numerical Functional Analysis and Optimization*, *28*(11–12), 1217–1230. https://doi.org/10.1080/01630560701749524

Control, N. A.-S. in I. and, & 2008, undefined. (n.d.). A hybrid conjugate gradient algorithm for

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15339
e-ISSN : 2541-2019
p-ISSN : 2541-044X

unconstrained optimization as a convex combination of Hestenes-Stiefel and Dai-Yuan. *Camo.Ici.RoN AndreiStudies in Informatics and Control, 2008•camo.Ici.Ro*.

Dai, Y. H., & Yuan, Y. (1996). Convergence properties of the Fletcher-Reeves method. *IMA Journal of Numerical Analysis*, *16*(2), 155–164. https://doi.org/10.1093/imanum/16.2.155

Deng, N., Software, Z. L.-O. M. and, & 1994, undefined. (1994). Global convergence of three terms conjugate gradient methods. *Taylor & FrancisN Deng, Z LiOptimization Methods and Software, 1994•Taylor & Francis*282–273 ,(4)*4* ,. https://doi.org/10.1080/10556789508805593

Dolan, E. D., & Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematical Programming, Series B*, *91*(2), 201–213. https://doi.org/10.1007/S101070100263

Faramarzi, P., & Amini, K. (2021). A spectral three-term Hestenes–Stiefel conjugate gradient method. *4or*, *19*(1), 71–92. https://doi.org/10.1007/s10288-020-00432-3

Fletcher, R. (1964). Function minimization by conjugate gradients. *The Computer Journal*, *7*(2), 149–154. https://doi.org/10.1093/comjnl/7.2.149

Fletcher, R. (1987). Practical methods of optimization. In *2000*.

G Zoutendijk, & 1970) .). Nonlinear programming, computational methods *.Nonlinear Programming Computational Methods*, *37–86*.

Hager, W. W. W. W., & Zhang, H. (2006). A Survey of Nonlinear Conjugate Gradient Methods. *Pacific Journal of Optimization*, *2*(1), 35–58.

Hager, W. W., & Zhang, H. (2006). A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM Journal on Optimization*, *16*(1), 170–192. https://doi.org/10.1137/030601880

Hassan, B. A., & Al-Naemi, G. M. (2020). A new quasi-newton equation on the gradient methods for optimization minimization problem. *Indonesian Journal of Electrical Engineering and Computer Science*, *19*(2), 737–744. https://doi.org/10.11591/ijeecs.v19.i2.pp737-744

Hestenes, M. R., & Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. In *Journal of Research of the National Bureau of Standards* (Vol. 49, Issue 6, p. 409). https://doi.org/10.6028/jres.049.044

Liu, Y., & Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: Theory. *Journal of Optimization Theory and Applications*, *69*(1), 129–137. https://doi.org/10.1007/BF00940464

Nazareth, L. (1977). A conjugate direction algorithm without line searches. *Journal of Optimization Theory and Applications*, *23*(3), 373–387. https://doi.org/10.1007/BF00933447

Nocedal, J., & Wright, S. J. (2006). Numerical optimization. In *Springer Series in Operations Research and Financial Engineering*. https://doi.org/10.1201/b19115-11

Opérationnelle, R., Rouge, S. É., Olak, E. P., & Ibiere, G. R. (1969). *R ' . s*.

Polyak, B. T. (1969). The conjugate gradient method in extremal problems. *USSR Computational Mathematics and Mathematical Physics*, *9*(4), 94–112. https://doi.org/10.1016/0041-5553(69)90035-4

Yuan, Y., & Dai, Y. H. (1999). A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property. *SIAM Journal on Optimization*, *10*(1), 177–182.

Zhang, J., Xiao, Y., & Wei, Z. (2009). Nonlinear conjugate gradient methods with sufficient descent condition for large-Scale unconstrained optimization. *Mathematical Problems in Engineering*, *2009*. https://doi.org/10.1155/2009/243290

Zhang, X., & Yang, Y. (2024). A new hybrid conjugate gradient method close to the memoryless BFGS quasi-Newton method and its application in image restoration and machine learning. *AIMS Mathematics*, *9*(10), 27535–27556. https://doi.org/10.3934/math.20241337

*name of corresponding author