# New Hybrid Conjugate Gradient Method Under Exact Line Search

Ain Aqiela Azamuddin[1](✉), Nurul 'Aini[1], Mohd Rivaie[2], and Zuraida Alwadood[3]

[1] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Johor, Kampus Segamat, Malaysia
ainaqiela97@gmail.com
[2] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Terengganu, Kampus Kuala Terengganu, Malaysia
[3] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Malaysia

**Abstract.** Conjugate gradient (CG) method is one of the popular method in solving unconstrained optimization problem. This method is notable for being an intermediate between the steepest descent method and the Newton's method. In this study, a new hybrid CG method is proposed with the main focus on improving Aini-Rivaie-Mustafa (ARM) CG method that were introduced in 2016. The ARM CG method sometimes generates negative CG coefficient that affects the performance of the method. Therefore, the new hybrid CG method is proposed with the intention of eliminating the negative CG coefficient value generated by the ARM CG method. The new hybrid CG method is globally convergent under the exact minimization rules and based on the numerical observation, it shows that it could solve higher number of test problems, as compared to the ARM CG method.

**Keywords:** Hybrid conjugate gradient · Exact line search · Unconstrained optimization

## 1 Introduction

Optimization is a process of decision-making which helps to find the best solution for a problem by making use of a situation. There are two types of optimization, which are constrained and unconstrained optimization. Constrained optimization occurs when there is a certain restriction set to the $x$ which is the subject being observed [14]. Meanwhile, unconstrained optimization provides a general view in solving optimization problem, as it was solved without any boundary or restriction towards $x$, which is the subject observed [14]. The general unconstrained optimization problem is

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

where $f : R^n \to R$ is a continuous function. Unconstrained optimization could be solved by using iterative method, for example, the Conjugate Gradient (CG) method.


N. Wahi et al. (Eds.): ICMSS 2022, ACSR 98, pp. 94–102, 2023.
https://doi.org/10.2991/978-94-6463-014-5_10

CG method is one of the popular methods in solving unconstrained optimization problem. This method is commonly regarded as an intermediate approach between the steepest descent method and the Newton's method [13]. This approach starts with an initial value, $x_0$, that will produce iteration of $[x_k]_{k=1}^{\infty}$ where $k$ is the iteration number that starts with 1, by using the formula in (2),

$$x_{k+1} = x_k + \alpha_k d_k \tag{2}$$

The term $\alpha_k$ is the step length and $d_k$ is the descent direction. The definition of descent direction, $d_k$ is given in (3),

$$d_k = \begin{cases} -g_k & , k = 0 \\ -g_k + \beta_k d_k & , k > 0 \end{cases} \tag{3}$$

where $g_k$ is the gradient and $\beta_k$ is the CG coefficient which varies depending on the CG method used. Next, there are two types of line search method for determining the step length that are, exact and inexact line search. Inexact line search calculates the approximate value of the step length, while exact line search finds the optimal or the exact amount of step length in converging towards the minimizer [13]. The formula for exact line search method is as in (4),

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha_k d_k) \tag{4}$$

A new hybrid CG method is proposed in this paper to further improve the past ARM CG method in terms of the CG coefficient. The next section discusses on the development and the algorithm of the method. Then, the theoretical proof that consists of the sufficient descent condition and the global convergence condition under exact line search are shown. The numerical results are then presented and analysed by using performance profile tool by Dolan and More [9]. This is followed by the discussion on the performance of the new hybrid CG method and lastly, a conclusion is made based on the results obtained.

## 2 New Hybrid CG Method

In 2016, Aini et al. [1] has introduced Aini-Rivaie-Mustafa (ARM) CG method that is a modified CG method with an added scalar, $m_k$. . The ARM CG method is stated in (5),

$$\beta_k^{ARM} = -\frac{m_k \|g_k\|^2 - \left|g_k^T g_{k-1}\right|}{m_k g_{k-1}^T d_{k-1}} \quad where\ m_k = \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|} \tag{5}$$

The ARM CG method has shown good numerical performance in solving unconstrained optimization problems. However, despite the fact that it could outperform other CG methods it was tested with, the ARM CG method sometimes generate negative CG coefficient value. This problem causes this method to sometimes fail to converge to a solution. In addition, Powell [2] stated that CG method could be considered successful if it could generate positive CG coefficient value in any circumstances.

One way to curb this problem is by using the hybrid approach, which involves combining the ARM CG method with another CG method that always produce positive CG coefficient value. This agreed with Hager and Zhang [3] who stated that, by combining different CG methods together, one could produce a hybrid CG method that possess attractive properties as compared to the original method. Furthermore, Kaelo et al. [4] added that the combination of CG methods could help in limiting the drawback from the original method and simultaneously taking advantage of the attractive properties possessed by the original method.

In 2018, Yasir et al. [5] introduced a hybrid CG method of Polak-Ribiere-Polyak and Wei-Yao-Liu which also known as PRP and WYL, respectively as both methods were known to have faster rate of convergence. Despite the fact that WYL is one of PRP variant, it has been proven that it is globally convergent [6]. Henceforth, the purpose to combine them both is to substitute the calculation whenever PRP fail to converge. The CG coefficient is written as (6),

$$\beta_k^{YHM} = \begin{cases} \dfrac{g_k^T(g_k - g_{k-1})}{g_{k-1}^T g_{k-1}}, & if\, 0 \le g_k^T g_{k-1} \le \|g_{k-1}\| \\ \dfrac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2}, & otherwise \end{cases} \tag{6}$$

This method is able to outperform both of its original methods in terms of the iteration number and CPU time, showing that it possesses fast convergence rate from both methods and also from the good convergence properties of the WYL CG method.

Based on these ideas, this study proposes to combine ARM CG method with Mandara-Mamat-Waziri-Usman CG method also known as MMWU CG method proposed by Mandara et al. [7]. The MMWU CG method is chosen due to its good numerical performance and its CG coefficient formula that is simple which assures to always generate positive CG coefficient value. The formula of the MMWU CG method written as in (7).

$$\beta_k^{MMWU} = \frac{\|g_k\|^2}{\|d_{k-1}\|^2} \tag{7}$$

The resulting hybrid CG method proposed by this study is called A-ARM CG method which mainly utilizes the ARM method in general computation. However, when negative CG coefficient is generated, the ARM method will be replaced by MMWU formula. The $\beta_k$ for A-ARM is defined as (8),

$$\beta_k^{A-ARM} = \begin{cases} -\dfrac{m_k\|g_k\|^2 - |g_k^T g_{k-1}|}{m_k g_{k-1}^T d_{k-1}} & , m_k\|g_k\|^2 \ge \left|g_k^T g_{k-1}\right| \\ \dfrac{\|g_k\|^2}{\|d_{k-1}\|^2} & , otherwise \end{cases} \tag{8}$$

By applying $d_k$, $\alpha_k$ and $\beta_k$ from (3), (4) and (8) respectively, the algorithm that will be used is as follows:

**Algorithm 1.** A-ARM Hybrid Conjugate Gradient Method

1. Choose initial point, $x_0$
2. Descent direction, $d_k$ is evaluated based on (3) with (8) for $\beta_k$ .
3. The step length, $\alpha_k$ is evaluated by using exact line search as in (4)
4. Iteration of $x_{k+1}$ is evaluated based on (2).
5. The calculation is terminated either when the stopping criteria of $\|g_k\| \leq 10^{-6}$ or the maximum iteration of 10,000 are reached. Otherwise, return to Step 2 with $k := k+1$

## 3 Convergence Properties

In this section, the theoretical analysis of the new hybrid CG method, A-ARM is described. It involves two parts, namely sufficient descent condition and global convergence condition. Both conditions are required to show that the new hybrid CG method does not only generate descent direction under exact line search, but it could also globally converge towards the minimizer.

### 3.1 Sufficient Descent Condition

By satisfying the sufficient descent condition, the method proposed is shown to be able to descent towards the solution when solving a problem. The condition is defined as

$$g_k^T d_k \leq -c\|g_k\|^2 \text{ for } k \geq 0, c > 0$$

**Theorem 1:** Suppose the calculation of $g_k$ and $d_k$ were done by using Algorithm 1 under exact line search in calculating the step length of $\alpha_k$ and the sufficient descent condition holds for all $k \geq 0$.

**Proof.**

**Case (1):** If $m_k\|g_k\|^2 \geq \left|g_k^T g_{k-1}\right|$, where the $\beta_k$ were calculated by using $\beta_k^{ARM}$. The proof can be referred to Aini et al. [1].
**Case (2):** Otherwise, the calculation will be done by using $\beta_k^{MMWU}$ where the proof can be referred to Mandara et al. [7].

Considering the proofs that have been made, the calculation was done under the exact line search where $g_k^T d_{k-1} = 0$ which implies that $g_k^T d_k = -\|g_k\|^2$. Hence, the condition $g_k d_k \leq -c\|g_k\|^2$ holds.

### 3.2 Global Convergence Condition

The global convergence property helps to ensure that the proposed method could converge towards the minimizer. This condition could be proven by using the following assumptions:

**Assumption 1:**

i. In the neighbourhood of $N$ of $L, f(x)$ is bounded below on level set of $L = \{L \in R^n : f(x) \leq f(x_0)\}$ where $f(x)$ is continuously differentiable and $x_0$ is the initial point.
ii. The gradient is Lipschitz continuous where constant $L > 0$ exist, such that $\|g(x) - g(y)\| = L\|x - y\|, \forall_{x,y \in N}$

**Lemma 1:** Suppose that Assumption 1 holds by considering CG in the form of (4) where the descent direction, $d_k$ and step length, $\alpha_k$ were determined under exact minimization rules. Under the stated assumptions, the lemma by Zoutendijk [8] holds as follows,

$$\sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < \infty,$$

that could be rewritten as $\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$ which can be referred in paper by Kaelo et al. [4].

**Theorem 2:** Suppose that Assumption 1(i) and Theorem 1 hold, consider the CG method in the (2) and (3), where the step length were determined under exact line search. Then,

$$\lim_{k \to \infty} \inf \|g_k\| = 0$$

in which the convergence theorem of the method is obtained by using (8).

**Proof:**

**Case 1:** When $m_k \|g_k\|^2 \geq \left|g_k^T g_{k-1}\right|$, then $\beta_k^{A-ARM} = \beta_k^{ARM}$. The proof can be referred to Aini et al. [1].
**Case 2:** Otherwise, the calculation for $\beta_k^{A-ARM}$ can be done by using $\beta_k^{MMWU}$. The proof can be referred to Mandara et al. [7].

**Table 1.** The list of test functions applied in this study

| No | Function | Number of variables | Initial points |
|---|---|---|---|
| 1 | Three hump | 2 | (5, 5), (15, 15), (100, 100) |
| 2 | Six Hump | 2 | (5, 5), (15, 15), (100, 100) |
| 3 | Biggsb1 | 2 | (5, 5), (15, 15), (100, 100) |
| 4 | Extended wood | 4 | (5,…5), (15,…,15), (100,…,100) |
| 5 | Extended powell | 4, 8, 100, 1000 | (5,…5), (15,…,15), (100,…,100) |
| 6 | Dixon & price | 2, 6,100 | (5,…5), (15,…,15), (100,…,100) |
| 7 | Generalized quartic | 2, 6,100 | (5,…5), (15,…,15), (100,…,100) |
| 8 | FLETCHR (CUTE) | 2, 6,100 | (5,…5), (15,…,15), (100,…,100) |
| 9 | Dixon3dq | 2, 6,100 | (5,…5), (15,…,15), (100,…,100) |
| 10 | Extended rosenbrock | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 11 | Extended himmelblau | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 12 | TRIDIA | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 13 | Extended trigonometric | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 14 | Extended tridiagonal 1 | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 15 | Diagonal 4 | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 16 | Extended DENSCHNB | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 17 | Shallow | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |
| 18 | QUARTC | 2, 6,100,1000 | (5,…5), (15,…,15), (100,…,100) |

## 4 Results and Discussion

In this part, the numerical tests on the hybrid CG method and its original methods are shown in order to analyze the performance of the hybrid CG method. The data obtained from all calculations are tabulated and they are then compared by using performance profile, as introduced by Dolan and More [9].

The stopping criteria applied is $\|g_k\| \leq 10^{-6}$ that was suggested by Hillstrom [10] or when the iteration of 10,000 is met. This is done to avoid any loop or calculation that took too much time to generate results. These stopping criteria are widely used in much recent research of CG method, as in [4, 11] and [12]. Subsequently, to test the new hybrid CG method, 18 test functions with three different initial points and variables ranging from 2 to 1000 are utilized. These are shown in Table 1.

Based on the results obtained, two performance profile graphs comparing the CPU time and the iteration number, which are recorded for all three methods are generated by using the benchmarking tools by Dolan and More [9]. The two graphs are shown in Fig. 1 and Fig. 2, respectively.

Based on the graphs, A-ARM CG method performed better than most of the methods. Based on the two figures, A-ARM CG method surpassed MMWU and FR CG method in terms of the CPU time and iteration number. However, ARM CG method still surpassed

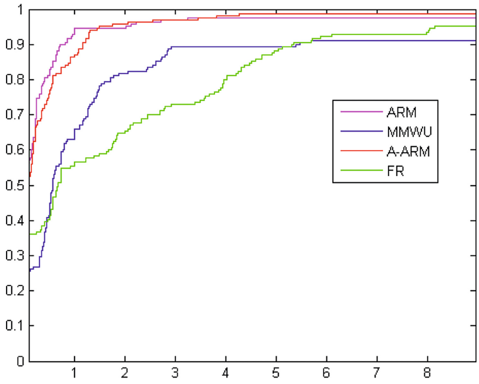


**Fig. 1.** Performance profile based on the CPU time

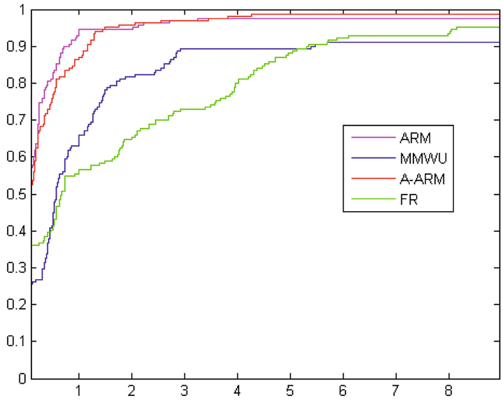


**Fig. 2.** Performance profile based on the iteration number

the A-ARM CG method in terms of the iteration number and CPU time as according to the left side of the graph, ARM CG method performed only slightly better than A-ARM CG method. Despite that, A-ARM CG method could still manage to surpass all of the methods in terms of the problem solved as it could solve 98.83% of the problem, while, ARM, MMWU and FR CG method could only solve 97.66% and 91.23% of the problems, respectively.

## 5 Conclusion

In this study, a new hybrid CG method denoted as A-ARM CG method was introduced by combining ARM with MMWU CG method. This method eliminates the problem of ARM CG method that sometimes produce negative CG coefficient. In general, the A-ARM CG method uses ARM CG method formula to calculate the CG coefficient. However, in the case where a negative value is obtained, the CG coefficient formula will be replaced with that of MMWU method, thus ensuring that all the generated CG coefficient values are positive. This method possesses sufficient descent and global convergence properties under the exact minimization rule. The numerical testing on A-ARM CG method also showed that the hybrid method produced higher number of functions solved as compared to ARM and MMWU CG methods, hence making the hybrid method a better option in contrast to its original methods.

**Acknowledgments.** The authors would like to acknowledge supports by readers and referee that provides motivation and advice in completing this study. The authors huge gratitude also towards the Ministry of Higher Education (MOHE) for the fund under FRGS-RACER (Grant no. RACER/1/2019/STG06/UITM//3) that has been provided.

## References

1. Aini, N., Hajar, N., Rivaie, M., Mamat, M.: A modified conjugate gradient coefficient under exact line search for unconstrained optimization. Int. J. Eng. Technol. (UAE) **7**, 84–88 (2018)
2. Powell, M.J.D.: Nonconvex minimization calculations and the conjugate gradient method. In: Griffiths, D.F. (ed.) Numerical Analysis. LNM, vol. 1066, pp. 122–141. Springer, Heidelberg (1984). https://doi.org/10.1007/BFb0099521
3. Hager, W.W., Zhang, H.: A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim. **16**(1), 170–192 (2005)
4. Kaelo, P., Mtagulwa, P., Thuto, M.V.: A globally convergent hybrid conjugate gradient method with strong Wolfe conditions for unconstrained optimization. Math. Sci. **14**(1), 1–9 (2019). https://doi.org/10.1007/s40096-019-00310-y
5. Salih, Y., Hamoda, M.A., Rivaie, M.: New hybrid conjugate gradient method with global convergence properties for unconstrained optimization. Malay. J. Comput. Appl. Math. **1**(1), 29–38 (2018)
6. Wei, Z., Yao, S., Liu, L.: The convergence properties of some new conjugate gradient methods. Appl. Math. Comput. **183**(2), 1341–1350 (2006)
7. Mandara, A.V., Mamat, M., Waziri, M.Y., Mohamed, M.A., Yakubu, U.A.: A new conjugate gradient coefficient with exact line search for unconstrained optimization. Far East J. Math. Sci. **105**, 193–206 (2018)
8. Zoutendijk, G..: Nonlinear programming, computational methods. Integer Nonlinear Programm., 37–86 (1970)
9. Dolan, E.D., Moré, J.J.: Benchmarking optimization software with performance profiles. Math. Program. **91**(2), 201–213 (2002)
10. Hillstrom, K.E.: A simulation test approach to the evaluation of nonlinear optimization algorithms. ACM Trans. Math. Softw. (TOMS) **3**(4), 305–315 (1977)
11. Ghani, N.H.A., Mohamed, N.S., Zull, N., Shoid, S., Rivaie, M., Mamat, M.: Performance comparison of a new hybrid conjugate gradient method under exact and inexact line searches. In: Journal of Physics: Conference Series, vol. 890, no. 1, p. 012106. IOP Publishing (2017)
12. Liu, M., Ma, G., Yin, J.: Two new conjugate gradient methods for unconstrained optimization. Complexity **2020** (2020)
13. Sun, W., Yuan, Y.X.: Optimization Theory and Methods: Nonlinear Programming, vol. 1. Springer, New York (2006). https://doi.org/10.1007/b106451
14. Nocedal, J., Wright, J.: Numerical optimization. In: Springer Series in Operations Research, New York (2006)