Received: xxx **Accepted:** xxx **Published:** xxx.
DOI. xxxxxxx
Volume xxx, Issue xxx, p.p. 1-23: xxx
Research Article

Open Access

Control and Optimization in Applied Mathematics - COAM

Novel Hybrid Conjugate Gradient Algorithms via Newton-Direction Alignment: Convergence Analysis and Image Restoration Applications

Nabil Sellami[1]✉, **Romaissa Mellal**[1], **Basim A. Hassan**[2]

[1]8th May 1945 University, Guelma, 24000, Algeria
[2]College of Computer Science and Mathematics, University of Mosul, Iraq

✉ **Correspondence:**
Nabil Sellami
E-mail:
sellami.nabil@univ-guelma.dz

How to Cite
Sellami, N., Mellal, R., Hassan. B.A. (2027). "Novel hybrid conjugate gradient algorithms via newton-direction alignment: Convergence analysis and image restoration applications". *Control and Optimization in Applied Mathematics*, 12(-), 1-23. https://doi.org/10.30473/coam.2026.76811.1381

Abstract. This paper introduces two hybrid nonlinear conjugate gradient algorithms, NRB1 and NRB2, for solving unconstrained optimization problems. Both methods are constructed as a convex combination of the AlBayati–AlAssady (BA) and Conjugate Descent methods (CD). The first variant, denoted NRB1, employs an adaptive combination parameter designed to align its search direction more closely with Newton's method, improving curvature approximation and acceleration of convergence. The second variant, NRB2, independently satisfies the conjugacy condition without relying on line search mechanisms, enhancing numerical stability. Both methods guarantee sufficient descent and global convergence properties under the strong Wolfe line search criteria. Extensive numerical experiments, using the performance profile of Dolan and Moré, demonstrate that NRB1 and NRB2 consistently outperform some classical conjugate gradient methods, including BA, CD, and Dai–Yuan, particularly for large-scale problems. Furthermore, NRB1 is applied to image restoration under salt-and-pepper noise, achieving competitive or superior peak signal-to-noise ratios (PSNR) compared to the Fletcher–Reeves (FR) algorithm with significantly fewer iterations, especially at high noise intensities.

Keywords. Nonlinear conjugate gradient, Unconstrained optimization, Global convergence, Image restoration problems.

MSC. 90C30; 65K05; 90C26.

©2026 by the authors. Lisensee PNU, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY4.0) (http:/creativecommons.org/licenses/by/4.0)

1 Introduction

Optimization tasks, particularly those involving image restoration under impulse noise pose significant challenges due to the nonsmooth and complex nature of their objective functions. Traditional gradient-based algorithms are often inadequate for addressing these challenges. However, recent advances in optimization have led to the development of more robust and efficient methods for noise removal, substantially improving the accuracy of restored images [12]. Numerous studies have investigated the performance of conjugate gradient methods (CGMs) for image denoising and restoration [5].

Nonlinear CGMs constitute a prominent class of algorithms for optimization, where the objective is to minimize a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$. This problem is typically formulated as:

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

in which $f : \mathbb{R}^n \mapsto \mathbb{R}$ belongs to class $\mathcal{C}^1$.

Among numerous iterative methods for solving (1), the conjugate gradient methods are considered optimal. They generate an iterative sequence of the following form:

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

where x_k is the current iterate, α_k is the step-length determined through a line search procedure along the search direction defined as:

$$d_k = \begin{cases} -q_k, \text{ for } k = 1, \\ -q_k + \beta_k d_{k-1}, \text{ for } k \geq 2, \end{cases}$$

where q_k stands for the gradient $\nabla f(x_k)$, and the scalar parameter β_k is selected to ensure that d_k constitutes the k^{th} conjugate direction, provided that the objective function is quadratic and the line search is performed exactly [3]. Different versions of this method differ in the way of selecting the parameter β_k.

The CGMs was originally developed by Hestenes-Stiefel (HS) [13] for solving strictly convex quadratic linear systems. It was later extended by Fletcher-Reeves (FR) [10] to address nonlinear unconstrained optimization problems. Since then, numerous nonlinear CG variants have been proposed, including , Fletcher (CD) [9] and AlBayati- AlAssady (BA) [4]. These methods are considered standard CG approaches and coincide with the linear CG method when f is strictly convex quadratic and the step length is chosen as the exact minimizer in the search direction. Their formulas are given by:

$$\beta_k^{HS} = \frac{q_{k+1}^T y_k}{d_k^T y_k}, \qquad \text{(Hestenes-Stiefel)}$$

In Press

$$\beta_k^{FR} = \frac{||q_{k+1}||^2}{||q_k||^2}, \qquad \text{(Fletcher-Reeves)}$$

$$\beta_k^{BA} = \frac{||y_k||^2}{d_k^T y_k}, \qquad \text{(AlBayati- AlAssady)}$$

$$\beta_k^{CD} = \frac{-||q_{k+1}||^2}{d_k^T q_k}, \qquad \text{(Fletcher)}$$

Note that the convergence behavior of all nonlinear CGMs cited above under strong Wolfe line search conditions [19, 20] has been widely studied by many authors [17].

The hybridization of conjugate gradient methods through convex combinations has become a significant area of research in optimization science. Recent research has established strong theoretical foundations for these hybrid methods, demonstrating that they maintain crucial properties such as sufficient descent conditions and global convergence.

The first such formulation, proposed by Andrei [1], adopts the convex structure

$$\beta_k^{\text{hybrid}} = (1-\theta_k)\,\beta_k^{\text{method}_1} + \theta_k\,\beta_k^{\text{method}_2},$$

where $\theta_k \in [0,1]$. In 2008, Andrei [1] introduced a specific hybrid combining the DY and HS methods, defined by

$$\beta_k^{\text{DY-HS}} = \theta_k\,\beta_k^{\text{DY}} + (1-\theta_k)\,\beta_k^{\text{HS}}.$$

Subsequently, in 2009, he proposed a further hybrid conjugate gradient method based on the conjugacy condition [2].

Several hybridization strategies combining the CD method with other formulations have since been developed. In particular, Djordjevic [7] introduced in 2017 a hybrid approach that combines the CD and LS methods, given by

$$\beta_k^{\text{LS-CD}} = \theta_k\,\beta_k^{\text{LS}} + (1-\theta_k)\,\beta_k^{\text{CD}}.$$

A substantial body of work has also addressed BA-based hybrid methods. Notably, Mellal et al. proposed two hybrid conjugate gradient algorithms of this type. The first combines the WYL and BA methods [15]:

$$\beta_k^{\text{RN}} = \theta_k\,\beta_k^{\text{BA}} + (1-\theta_k)\,\beta_k^{\text{WYL}},$$

and the second combines the RMIL and BA methods [16]:

$$\beta_k^{\text{hnBARMIL}} = \theta_k\,\beta_k^{\text{BA}} + (1-\theta_k)\,\beta_k^{\text{RMIL}},$$

where, in both cases, the convex parameter θ_k is chosen so that the resulting conjugate gradient direction aligns with the Newton direction. All of the above formulations preserve the descent property and exhibit strong numerical performance; they also satisfy sufficient descent conditions and global convergence under appropriate assumptions.

In Press

Motivated by these developments, the present paper introduces two hybrid conjugate gradient variants constructed as convex combinations of the BA and CD methods. The main contributions of this work are as follows.

- We propose two new hybrid conjugate gradient algorithms, referred to as NRB1 and NRB2, each based on a convex combination of the BA and CD methods. In NRB1, the combination parameter is determined by requiring the resulting direction to coincide with the Newton direction. In NRB2, the parameter is chosen to satisfy a conjugacy condition that is independent of the line search procedure.
- We establish that both algorithms generate sufficient descent directions and prove their global convergence under the strong Wolfe line search conditions.
- We carry out extensive numerical experiments on a standard set of large-scale unconstrained optimization benchmark problems. Using the Dolan–Moré performance profile framework, we demonstrate that NRB1 is particularly efficient and robust, outperforming the classical BA, CD, and DY methods across a wide range of test cases. All algorithms are implemented in Python 3.13.
- We illustrate the practical applicability of the proposed methods through an image restoration case study, in which NRB1 achieves notably superior performance compared to the FR algorithm, particularly on severely degraded or heavily corrupted images.

The remainder of the paper is organized as follows. Section 2 introduces the proposed hybrid variants and the corresponding algorithmic framework. Section 3 establishes the sufficient descent properties, followed by a rigorous proof of global convergence. Section 4 presents the numerical experiments and performance comparisons. Section 5 concludes the paper with a summary of the main findings and directions for future research.

In Press

2 New Hybrid Conjugate Gradient Method

In this section, we present a newly developed conjugate gradient method. The approach constructs a hybrid CG parameter, β_k^{NRB}, as a convex combination of the AlBayati-Al-Assady (BA) and Conjugate Descent (CD) parameters:

$$\beta_k^{NRB} = \theta_k \beta_k^{BA} + (1 - \theta_k)\beta_k^{CD}, \tag{3}$$

where $\theta_k \in [0, 1]$ is a scalar weighting parameter. The search direction d_{k+1}^{NRB} is defined recursively:

$$d_{k+1}^{NRB} = \begin{cases} -q_{k+1}, & k = 0, \\ -q_{k+1} + \beta_k^{NRB} d_k, & k \geq 1, \end{cases} \tag{4}$$

where $q_k = \nabla f(x_k)$ denotes the gradient. The step size α_k in the iteration $x_{k+1} = x_k + \alpha_k d_k$ is determined using the strong Wolfe conditions:

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq f(x_k) + \rho \alpha_k q_k^T d_k, \\ |q_{k+1}^T d_k| &\leq \sigma |q_k^T d_k|, \end{aligned} \tag{5}$$

with constants $0 < \rho < \sigma < 1$ [19, 20].
It is evident that, if $\theta_k = 0$ then $\beta_k^{NRB} = \beta_k^{CD}$ and if $\theta_k = 1$ then $\beta_k^{NRB} = \beta_k^{BA}$.

We consider two possibilities for selecting the parameter θ_k:
In the first approach, θ_k is chosen so that the direction d_{k+1}^{NRB} from (4) coincides with the Newton direction, i.e.,

$$-q_{k+1} + \left[\theta_k \frac{||y_k||^2}{d_k^T y_k}_{k} + (1 - \theta_k) \frac{-||q_{k+1}||^2}{d_k^T q_k} \right] d_k = -\nabla^2 f(x_{k+1})^{-1} q_{k+1}. \tag{6}$$

Multiplying both sides of the equation (6) by $s_k^T \nabla^2 f(x_{k+1})$, where $s_k = x_{k+1} - x_k$, and using the standard secant equation $s_k^T \nabla^2 f(x_{k+1}) = y_k$, we obtain:

$$\theta_k^{NRB1} = \frac{[y_k^T q_{k+1} - s_k^T q_{k+1}](d_k^T q_k) + ||q_{k+1}||^2 (d_k^T y_k)}{||y_k||^2 (d_k^T q_k) + ||q_{k+1}||^2 (d_k^T y_k)}. \tag{7}$$

In the second approach, θ_k is selected to ensure that the conjugacy condition is satisfied at each iteration, independently of the line search. From the recurrence relation in (4), we have:

$$\begin{aligned} d_{k+1}^{NRB} = -q_{k+1} + \beta_k^{NRB} d_k &= -q_{k+1} + (1 - \theta_k) \beta_k^{CD} d_k + \theta_k \beta_k^{BA} d_k \\ &= -q_{k+1} + (1 - \theta_k) \frac{-||q_{k+1}||^2}{d_k^T q_k} d_k + \theta_k \frac{||y_k||^2}{d_k^T y_k} d_k. \end{aligned}$$

Multiplying both sides by y_k and imposing the conjugacy condition $d_{k+1}^T y_k = 0$ yields:

$$0 = -q_{k+1}^T y_k + (1 - \theta_k) \frac{-||q_{k+1}||^2}{q_k^T d_k} d_k^T y_k + \theta_k \frac{||y_k||^2}{d_k^T y_k} d_k^T y_k,$$

which simplifies to:

$$\theta_k = \frac{\left(-d_k^T q_k\right) q_{k+1}^T y_k - \|q_{k+1}\|^2 d_k^T y_k}{\left(-d_k^T q_k\right) \|y_k\|^2 - \|q_{k+1}\|^2 d_k^T y_k}.$$

Finally, we have

$$\theta_k^{NRB2} = \frac{(d_k^T q_k) q_{k+1}^T y_k + ||q_{k+1}||^2 d_k^T y_k}{(d_k^T q_k) ||y_k||^2 + ||q_{k+1}||^2 d_k^T y_k}. \tag{8}$$

To ensure numerical stability, we regularize θ_k^{NRB} as:

In Press

$$\theta_k^{NRB} = \begin{cases} \theta_k^{NRB1} \text{ or } \theta_k^{NRB2}, & \text{if } 0 < \theta_k^{NRB} < 1, \\ 0, & \text{if } \theta_k^{NRB} \leq 0, \\ 1, & \text{if } \theta_k^{NRB} \geq 1. \end{cases} \tag{9}$$

Algorithm 1 NRB Conjugate Gradient Algorithm

Input: Initial point $x_0 \in \mathbb{R}^n$; parameters $\delta = 0.05$, $\sigma = 0.5$; maximum number of iterations $N_{\max}$. Compute $f(x_0)$ and $q_0 = \nabla f(x_0)$. Set $d_0 = -q_0$.

Output: Approximate solution x^* of problem (1).

1. Set $k \leftarrow 0$.

2. **Repeat** the following steps until $\|q_k\|_\infty \leq 10^{-6}$ or $k > N_{\max}$:

a. *Inexact line search.* Apply the strong Wolfe conditions (5) to compute the step size α_k, and update:

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

Compute $f(x_{k+1})$ and $q_{k+1} = \nabla f(x_{k+1})$.

b. *Parameter update.* Compute $\theta_k = \theta_k^{\text{NRB}}$ via (9) and $\beta_k = \beta_k^{\text{NRB}}$ via (3).

c. *Powell restart criterion.*

If $|q_{k+1}^\top q_k| \geq 0.2\,\|q_{k+1}\|^2$, set

$$d_{k+1} = -q_{k+1}, \qquad \alpha_{k+1} = 1.$$

Else compute the search direction via (4):

$$d_{k+1} = -q_{k+1} + \beta_k d_k.$$

d. Set $k \leftarrow k + 1$.

3. **Return** $x^* = x_k$.

In Press

3 Convergence Analysis

In this section, we will present the theorems which guarantee the convergence of our method under the strong Wolfe line search. Throughout this section, we make the following assumptions:

Assumption 1. The level set $\Gamma = \{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$ is bounded, thus,

$$\exists\, B > 0 : \|x\| \leq B, \qquad \forall x \in \Gamma. \tag{10}$$

Assumption 2. Within a neighborhood N of Γ, the function f belongs to class C^1 with a Lipschitz continuous gradient, specifically:

$$\exists L > 0 : ||q(x) - q(y)|| \leq L||x - y||, \qquad \forall x, y \in \Gamma. \tag{11}$$

Under Assumptions 1 and 2 on f we have:

$$\exists\, M > 0 : ||q(x)|| \leq M, \qquad \forall x, y \in \Gamma \tag{12}$$

Global convergence is established by initially demonstrating the following theorem, which proves the sufficient descent of the NRB direction.

Theorem 1. The search direction generated by NRB conjugate gradient algorithm satisfies the sufficient descent condition:

$$q_k^T d_k^{NRB} \leq -c||q_k||^2, \ \forall k \ \geq 0. \tag{13}$$

Proof. If $k = 0$, then $q_0^T d_0 = -\|q_0\|^2$, so (13) holds.

On the other hand, for $k > 0$, we have

$$\begin{aligned} d_{k+1}^{NRB} &= -q_{k+1} + \beta_k^{NRB} d_k \\ &= -q_{k+1} + \left(\theta_k \beta_k^{BA} + (1-\theta_k)\beta_k^{CD}\right) d_k \\ &= (1-\theta_k)(-q_{k+1} + \beta_k^{CD} d_k) + \theta_k(-q_{k+1} + \beta_k^{BA} d_k). \end{aligned}$$

It follows that

$$d_{k+1}^{NRB} = \theta_k d_{k+1}^{BA} + (1-\theta_k) d_{k+1}^{CD}.$$

Multiplying by q_{k+1}^T from the left, we obtain

$$q_{k+1}^T d_{k+1}^{NRB} = \theta_k q_{k+1}^T d_{k+1}^{BA} + (1-\theta_k) q_{k+1}^T d_{k+1}^{CD}. \tag{14}$$

Firstly, if $\theta_k = 0$, d_k^{NRB} coincides with the CD direction and

$$d_{k+1}^{NRB} = d_{k+1}^{CD} = -q_{k+1} + \beta_k^{CD} d_k.$$

In Press

Taking the inner product with q_{k+1}, we get

$$\begin{aligned} q_{k+1}^T d_{k+1}^{NRB} = q_{k+1}^T d_{k+1}^{CD} &= -q_{k+1}^T q_{k+1} + \beta_k^{CD} q_{k+1}^T d_k \\ &= -\|q_{k+1}\|^2 + \frac{\|q_{k+1}\|^2}{-d_k^T q_k} q_{k+1}^T d_k \\ &\leq -(1-\sigma)\|q_{k+1}\|^2, \end{aligned}$$

which implies that

$$q_{k+1}^T d_{k+1}^{NRB} = q_{k+1}^T d_{k+1}^{CD} \leq -c_1 \|q_{k+1}\|^2, \quad \forall k, \tag{15}$$

where $c_1 = 1 - \sigma$.

Now, if $\theta_k = 1$, then d_{k+1}^{NRB} coincides with the BA direction. In order to bound the BA direction, we adopt the bounded–angle strategy used by Hamdi et al. [11].

The BA search direction is defined as

$$d_{k+1}^{\text{BA}} = -q_{k+1} + \frac{\|y_k\|^2}{d_k^\top y_k} d_k.$$

Multiplying by $q_{k+1}^\top$ from the left and using strong Wolfe condition (5), we get

$$\begin{aligned} q_{k+1}^\top d_{k+1}^{\text{BA}} &= -\|q_{k+1}\|^2 + \frac{\|y_k\|^2}{d_k^\top y_k} q_{k+1}^\top d_k \\ &\leq -\|q_{k+1}\|^2 + \frac{\|y_k\|^2}{d_k^\top y_k}(-\sigma q_k^\top d_k), \end{aligned}$$

Since $d_k^\top y_k = d_k^\top (q_{k+1} - q_k) \geq (\sigma - 1) q_k^\top d_k = (1-\sigma)(-q_k^\top d_k)$, we have

$$\frac{1}{d_k^\top y_k} \leq \frac{1}{(1-\sigma)(-q_k^\top d_k)}.$$

Therefore:

$$\begin{aligned} q_{k+1}^\top d_{k+1}^{\text{BA}} &\leq -\|q_{k+1}\|^2 + \frac{\|y_k\|^2}{(1-\sigma)(-q_k^\top d_k)}(-\sigma q_k^\top d_k) \\ &= -\|q_{k+1}\|^2 + \frac{\sigma}{1-\sigma}\|y_k\|^2. \end{aligned}$$

By the Lipschitz continuity of the gradient (11), we have $\|y_k\| = \|q_{k+1} - q_k\| \leq L\|s_k\|$, where $s_k = x_{k+1} - x_k$. Thus:

$$q_{k+1}^\top d_{k+1}^{\text{BA}} \leq -\|q_{k+1}\|^2 + \frac{\sigma}{1-\sigma} L^2 \|s_k\|^2. \tag{16}$$

As $\|s_k\|^2 \to 0$, we get

$$-\|q_{k+1}\|^2 + \frac{\sigma}{1-\sigma}\|y_k\|^2 \to 0.$$

So there exists ω such that $0 < \omega \leq 1$. Therefore

In Press

$$\frac{\sigma}{1-\sigma} L^2 \|s_k\|^2 \leq \omega \|q_{k+1}\|^2.$$

This gives

$$q_{k+1}^{\top} d_{k+1}^{BA} \leq -\|q_{k+1}\|^2 + \omega \|q_{k+1}\|^2,$$

which yields the sufficient descent property for the BA direction,

$$\begin{aligned} q_{k+1}^{\top} d_{k+1}^{BA} &\leq -(1-\omega) \|q_{k+1}\|^2, \\ &\leq -c_2 \|q_{k+1}\|^2, \qquad \forall k, \end{aligned} \tag{17}$$

Finally, if $0 < \theta_k < 1$, then there exist scalars λ_1 and λ_2 such that

$$0 < \lambda_1 \leq \theta_k \leq \lambda_2 < 1. \tag{18}$$

Using (14) together with (15) and (17), we obtain

$$q_{k+1}^{T} d_{k+1}^{NRB} \leq \lambda_1 \, q_{k+1}^{T} d_{k+1}^{BA} + (1-\lambda_2) \, q_{k+1}^{T} d_{k+1}^{CD}.$$

Let $c = \lambda_1 c_2 + (1-\lambda_2) c_1$; then, from (15)–(18), we finally get

$$q_{k+1}^{T} d_{k+1}^{NRB} \leq -c \|q_{k+1}\|^2,$$

with $c > 0$ independent of k. This proves the sufficient descent condition (13) for all $k \geq 0$. □

In the following, we study the global convergence properties of our proposed conjugate gradient method.

Lemma 1. Under Assumptions 1 and 2, consider any conjugate gradient method defined by (2) and (1), with step length α_k determined through the strong Wolfe line search (5). If

$$\sum_{k \geq 0} \frac{1}{\|d_k\|^2} < \infty,$$

Then

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

Proof. See [3] . □

Lemma 2. Assuming that Assumptions 1 and 2 are satisfied, if d_k provides a descent direction and α_k satisfies

$$q_{k+1}^{T} d_k \geq \sigma q_k^{T} d_k, : \text{where} : 0 < \sigma < 1,$$

then

$$\alpha_k \geq \frac{(1-\sigma) q_k^{T} d_k}{L |d_k|^2}. \tag{19}$$

Proof. For the proof of Lemma 2, refer to the work in [14]. □

In Press

Based on Lemma 2, together with Assumptions 1 and 2 and the strong Wolfe conditions (5), we can establish that $q_k^T d_k \neq 0$ for all $k > 0$. Consequently, $\alpha_k = 0$ cannot satisfy condition (5), indicating that a zero step length is not valid in our new hybrid conjugate gradient algorithm. Furthermore, there exists a positive constant $\lambda > 0$ such that

$$\alpha_k > \lambda, \text{ for all } k \geq 0. \tag{20}$$

The following theorem establishes the global convergence properties of our proposed method.

Theorem 2. Let Assumptions 1 and 2 hold, and consider the iterative method defined by (2), (4), and (5). Then either $q_k = 0$ for some k, or

$$\liminf_{k\to\infty} \|q_k\| = 0. \tag{21}$$

Proof. Suppose that Assumptions 1 and 2, together with condition (10), are satisfied, and assume for contradiction that $q_k \neq 0$ for all $k \geq 0$. We aim to establish (21) by showing that its negation leads to a contradiction.

Suppose (21) does not hold. Then there exists a constant $t > 0$ such that

$$\|q_k\| \geq t \qquad \text{for all } k \geq 0. \tag{22}$$

Combining the strong Wolfe conditions (5) with (13), we obtain

$$-q_k^\top d_k \;\geq\; c(1-\sigma)\,\|q_k\|^2 \;\geq\; c(1-\sigma)\,t^2. \tag{23}$$

Since f satisfies the Lipschitz condition on the level set Γ,

$$\|y_k\| \;=\; \|q_{k+1} - q_k\| \;\leq\; LB, \tag{24}$$

where $B = \max\{\|x - y\| : x, y \in \Gamma\}$ denotes the diameter of Γ.

Recalling the NRB search direction,

$$d_{k+1}^{\text{NRB}} = -q_{k+1} + \big(1-\theta_k\big)\beta_k^{\text{CD}}\, d_k + \theta_k\, \beta_k^{\text{BA}}\, d_k,$$

and using $0 < \theta_k < 1$, the triangle inequality yields

$$\big\|d_{k+1}^{\text{NRB}}\big\| \;\leq\; \|q_{k+1}\| + \big(|\beta_k^{\text{BA}}| + |\beta_k^{\text{CD}}|\big)\|d_k\|.$$

We now bound each conjugate parameter separately. From (23) and (24),

$$\big|\beta_k^{\text{BA}}\big| \;=\; \frac{\|y_k\|^2}{d_k^\top y_k} \;\leq\; \frac{L^2 B^2}{c(1-\sigma)\,t^2}, \tag{25}$$

and

In Press

$$\left|\beta_k^{\mathrm{CD}}\right| \;=\; \frac{\|q_{k+1}\|^2}{\left|-d_k^{\top} q_k\right|} \;\le\; \frac{1}{c}. \tag{26}$$

Substituting (25) and (26) into the bound on $\|d_{k+1}^{\mathrm{NRB}}\|$, and using $\|s_k\| \le B$, gives

$$\begin{aligned}\left\|d_{k+1}^{\mathrm{NRB}}\right\| \;&\le\; M + \left(\frac{L^2B^2}{c(1-\sigma)\,t^2} + \frac{1}{c}\right)\frac{\|s_k\|}{\lambda} \\ &\le\; M + \left(\frac{L^2B^2}{c(1-\sigma)\,t^2} + \frac{1}{c}\right)\frac{B}{\lambda} \;\le\; \zeta \qquad \text{for all } k \ge 0,\end{aligned}$$

where the finite constant ζ is defined by

$$\zeta \;=\; M + \frac{B\left(L^2B^2 + (1-\sigma)\,t^2\right)}{c\,\lambda\,(1-\sigma)\,t^2} \;=\; \frac{Mc\,\lambda\,(1-\sigma)\,t^2 + B\left(B^2L^2 + (1-\sigma)\,t^2\right)}{c\,\lambda\,(1-\sigma)\,t^2}. \tag{27}$$

Since $\|d_{k+1}^{\mathrm{NRB}}\| \le \zeta < \infty$ for all k, it follows that

$$\sum_{k\ge 0} \frac{1}{\left\|d_{k+1}^{\mathrm{NRB}}\right\|^2} \;=\; +\infty. \tag{28}$$

Applying Lemma 1 then gives

$$\liminf_{k\to\infty} \|q_k\| \;=\; 0, \tag{29}$$

which contradicts (22). Hence (21) holds, and the proof is complete. □

4 Numerical Experiments

This section reports the numerical performance of the proposed NRB algorithms across two distinct experimental phases.

4.1 Analysis of Computational Performance

The first phase concerns unconstrained optimization problems of the form (1). To evaluate the effectiveness and convergence behavior of the proposed NRB1 and NRB2 algorithms, we tested them on a representative set of well-known benchmark problems drawn from [6], spanning low, medium, and high dimensions (n ranging from 2 to 1000). The test functions are listed in Table 1. For comparison purposes, we include two established nonlinear conjugate gradient methods, namely the BA and CD methods, both implemented under the strong Wolfe line search conditions.

In Press

Computational environment. All codes were implemented in Python 3.13 and executed on a Lenovo ThinkPad equipped with an AMD Ryzen 7 PRO 5850U processor (Radeon Graphics, 1.90 GHz) and 16.0 GB of RAM, running Windows 10 Professional (64-bit).

Stopping criteria. The algorithm is terminated when either

$$\|q_k\|_\infty \;<\; 10^{-6},$$

or the iteration count reaches 1000, whichever occurs first. A restart along the steepest descent direction is triggered whenever the denominator of β_k vanishes, or the Powell restart criterion [18]

$$\left|q_{k+1}^{\top} q_k\right| \;\geq\; 0.2\,\|q_{k+1}\|^2,$$

is satisfied. The strong Wolfe conditions are enforced with parameters $\rho = 0.05$ and $\sigma = 0.5$.

Performance metrics. Numerical results are compared in terms of the number of iterations (NOI), CPU time (in seconds), and the number of function evaluations (NFEV), as reported in Figure **??**.

Performance profiles. To provide a rigorous and objective comparison among the tested solvers, we adopt the performance profile framework of Dolan and Moré [8]. For a set of n_s solvers and n_p test problems, the performance profile $\sigma_s : \mathbb{R} \longrightarrow [0, 1]$ is defined by

$$\sigma_s(\lambda) \;=\; \frac{1}{n_p}\left|\{p \in P : r_{p,s} \leq \lambda\}\right|, \tag{30}$$

where the performance ratio $r_{p,s}$ is given by

$$r_{p,s} \;=\; \frac{a_{p,s}}{\min\limits_{s\in S} a_{p,s}}, \tag{31}$$

and $a_{p,s}$ denotes either the number of iterations or the CPU time required by solver s to solve problem p. A higher value of $\sigma_s(\lambda)$ indicates superior overall performance of solver s.

4.2 Commentaries

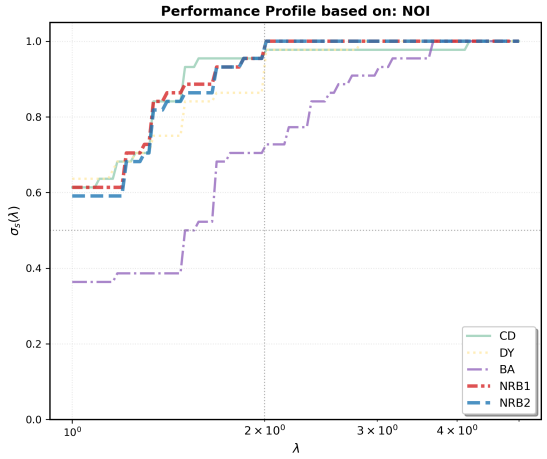
Table 2 presents a comparative evaluation of NRB1 against BA, NRB2, CD, and DY. In terms of actual success rates, NRB1 achieves its strongest performance relative to BA, recording 63.6% in NOI, 45.0% in CPU, and 70.0% in NFEV. The improvement rates further indicate that, whenever NRB1 surpasses a competitor, the margin of gain is considerable, most notably in CPU time, where it leads by 55.0% over BA, 65.0% over NRB2, 60.0% over CD, and 70.0% over

In Press

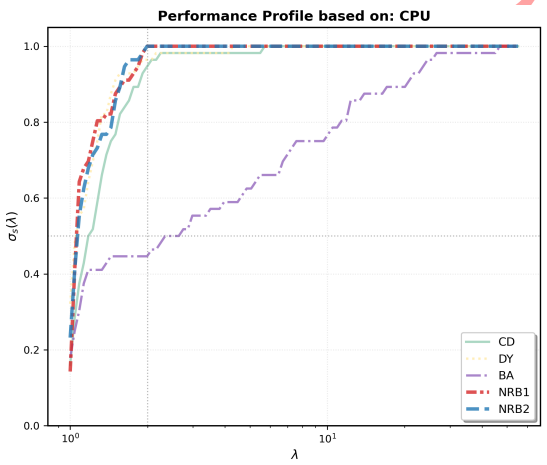
Table 1: Test Functions from CUTEst Collection with Initial Points

No.	Function Name	CUTEst Identifier	Dimensions	Initial Point x_0
1	Rosenbrock	ROSENBR	$10, 100, 500, 1000$	$(-1.2, 1.0, -1.2, 1.0, \ldots)$
2	Sphere	SPHERE	$10, 100, 500, 1000$	$(0.5, 0.5, \ldots, 0.5)$
3	Sum of Squares	SOS	$10, 100, 500, 1000$	$(1, 1, \ldots, 1)$
4	Zakharov	ZAKHAROV	$10, 100, 500, 1000$	$(1, 1, \ldots, 1)$
5	Dixon-Price	DIXMAAN (A-F)	150, 300	$(2, 2, \ldots, 2)$
6	Quadratic QF1	QHILBERT	$10, 100, 500, 1000$	$(1, 1, \ldots, 1)$
7	Raydan 1	RAY1	$10, 100, 500, 1000$	$(0.5, 0.5, \ldots, 0.5)$
8	Raydan 2	RAY2	$10, 100, 500, 1000$	$(1, 1, \ldots, 1)$
9	Extended Rosenbrock	ROSENBR	$10, 100, 500, 1000$	$(-1.2, -1.2, \ldots, -1.2)$
10	Extended DENSCHNF	DENSCHNF	10, 100	$(1, 1, \ldots, 1)$
11	Extended Tridiagonal	TRIDIA	$10, 100, 500, 1000$	$(0, 0, \ldots, 0)$
12	Extended Himmelblau	HIMMELBH	2, 10, 100	$(1, 1, \ldots, 1)$
13	DBVF	DBVF	10, 100	$(0.1, 0.1, \ldots, 0.1)$
14	BRYBND	BRYBND	$10, 100, 500$	$(-1, -1, \ldots, -1)$
15	Perturbed Quadratic	PERTURBEDQUAD	$10, 100, 500, 1000$	$(0.5, 0.5, \ldots, 0.5)$
16	TRIDIA	TRIDIA	$10, 100, 500, 1000$	$(1, 1, \ldots, 1)$
17	Extended Penalty	PENALTY1	$10, 100, 500, 1000$	$(1, 1, \ldots, 1)$
18	BALF	BALF	10	$(0.5, 0.5, \ldots, 0.5)$
19	Diagonal 1	DIAGONAL1	$10, 100, 500$	$(2, 2, \ldots, 2)$
20	Diagonal 2	DIAGONAL2	$10, 100, 500$	$(2, 2, \ldots, 2)$
21	Diagonal 3	DIAGONAL3	$10, 100, 500$	$(0, 0, \ldots, 0)$
22	Diagonal 4	DIAGONAL4	$10, 100, 500$	$(1, 1, \ldots, 1)$
23	Extended Diagonal	DIAGN	100	$(1, 1, \ldots, 1)$
24	Beale	BEALE	2	$(1, 1)$
25	Booth	BOOTH	2	$(0, 0)$
26	Ackley	ACKLEY	$2, 5, 10, 100, 500$	$(2, 2, \ldots, 2)$
27	Rastrigin	RASTRIGIN	$2, 5, 10, 100, 500$	$(1.5, 1.5, \ldots, 1.5)$
28	Griewank	GRIEWANK	$2, 5, 10, 100, 500$	$(10, 10, \ldots, 10)$
29	Matyas	MATYAS	2	$(1, 1)$
30	Schwefel	SCHWEFEL	$2, 5, 10, 100, 500$	$(400, 400, \ldots, 400)$
31	Cosine	COSINE	$2, 5, 10, 100, 500, 1000$	$(1, 1, \ldots, 1)$
32	Dqdrtic	DQDRTIC	$2, 5, 10, 100, 500, 1000$	$(1, 1, \ldots, 1)$

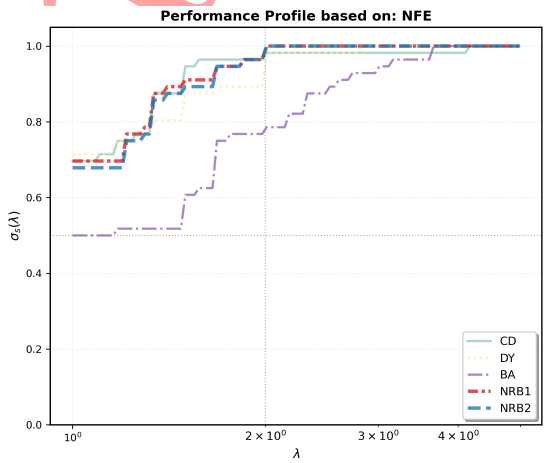
In Press



(a) NOI



(b) CPU



(c) NFEV

Figure 1: Performance Profile of NRB VS Classical CGMs, based on NOI, CPU and NFEV

Table 2: Comparison of global success rates of NRB1 over CG methods.

Methods	Actual Results			Rate of Improvement		
	NOI (%)	CPU (%)	NFEV (%)	NOI (%)	CPU (%)	NFEV (%)
Standard (NRB2, BA, CD, DY)	100.00	100.00	100.00	–	–	–
NRB1 Compared with BA	63.6	45.0	70.0	36.4	55.0	30.0
NRB1 Compared with NRB2	40.9	35.0	45.0	59.1	65.0	55.5
NRB1 Compared with CD	50.0	40.0	50.0	50	60.0	50.0
NRB1 Compared with DY	36.4	30.0	40.0	63.6	70.0	60.0

DY. A closer look at the NRB1 versus NRB2 comparison, yielding 59.1% in NOI and 65.0% in CPU, suggests that incorporating a Newton-direction alignment strategy leads to markedly faster convergence than relying solely on conjugacy conditions. Taken together, these findings underscore the computational efficiency and robustness of NRB1 in the context of large-scale unconstrained optimization.

4.3 Application of the NRB Algorithm to Image Restoration

Image restoration is among the most challenging problems in optimization, aiming to recover an original image from a degraded version corrupted by impulse noise. In this section, we apply the proposed **NRB** algorithm to solve problem (32) and compare its performance against the classical **FR** method.

Let χ be a grayscale image of dimensions $M \times N$, and let $\Delta = \{1, 2, \ldots, M\} \times \{1, 2, \ldots, N\}$ denote its pixel index set. We seek to minimize the functional $\Phi(x)$ defined by

$$\Phi(x) = \sum_{(p,q)\in E} \left(2 \sum_{(m,n)\in \mathcal{V}_{p,q}\setminus E} \varphi(x_{p,q} - y_{m,n}) + \sum_{(m,n)\in \mathcal{V}_{p,q}\cap E} \varphi(x_{p,q} - x_{m,n}) \right), \quad (32)$$

where $x_{p,q}$ and $y_{m,n}$ denote the pixel values of the restored and noisy images, respectively, $E \subset \Delta$ is the index set of noise-corrupted pixels detected in the first phase, and

$$\mathcal{V}_{p,q} = \{(p, q-1),\ (p, q+1),\ (p-1, q),\ (p+1, q)\},$$

is the four-pixel neighborhood of location (p, q) for all $(p, q) \in \Delta$.

Restoration quality is assessed via the Peak Signal-to-Noise Ratio (PSNR), defined as

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\frac{1}{MN} \sum_{p,q} \left(x^k_{p,q} - x^*_{p,q} \right)^2},$$

In Press

where $x_{p,q}^k$ and $x_{p,q}^*$ are the pixel values of the restored and original images, respectively. The algorithm is terminated when the relative change in the objective satisfies

$$\frac{|\Phi(x_k) - \Phi(x_{k-1})|}{\Phi(x_k)} \leq 10^{-4}.$$

Experiments are conducted on four standard test images — Lena, House, Elaine, and Cameraman, each of size 256×256 — corrupted by salt-and-pepper noise at three intensity levels: 50%, 70%, and 90%. Figure 3 illustrates representative restoration results across these noise levels.

Table 3: Numerical results for the image restoration problem under salt-and-pepper noise at three corruption levels.

Image	Noise level (%)	FR			NRB1			NRB2		
		NI	NF	PSNR	NI	NF	PSNR	NI	NF	PSNR
Lena	50	82	153	30.552	58	105	30.402	48	105	30.497
	70	81	155	27.482	59	109	27.302	39	77	27.468
	90	108	211	22.858	63	113	**23.040**	48	98	22.837
House	50	52	53	30.684	40	72	**35.084**	26	54	34.766
	70	63	116	**31.256**	42	76	31.039	35	69	31.334
	90	111	214	25.287	56	75	**25.497**	48	97	24.931
Elaine	50	35	36	33.912	34	53	33.876	25	47	**33.923**
	70	38	39	31.864	36	46	31.784	33	63	**31.871**
	90	65	114	28.201	53	63	28.129	41	80	**28.230**
Cameraman	50	59	87	35.535	35	68	35.282	31	66	**35.598**
	70	78	142	30.625	41	80	**30.733**	31	63	30.653
	90	121	236	24.396	51	99	24.898	47	99	**24.961**

NI: number of iterations; NF: number of function evaluations; PSNR: peak signal-to-noise ratio (dB). Bold values indicate the best PSNR for each setting.

Table 3 summarizes the numerical results. To facilitate visual interpretation of the relatively close PSNR values across methods, the results are additionally presented as a bar chart, which provides a clearer picture of the performance differences.

In Press

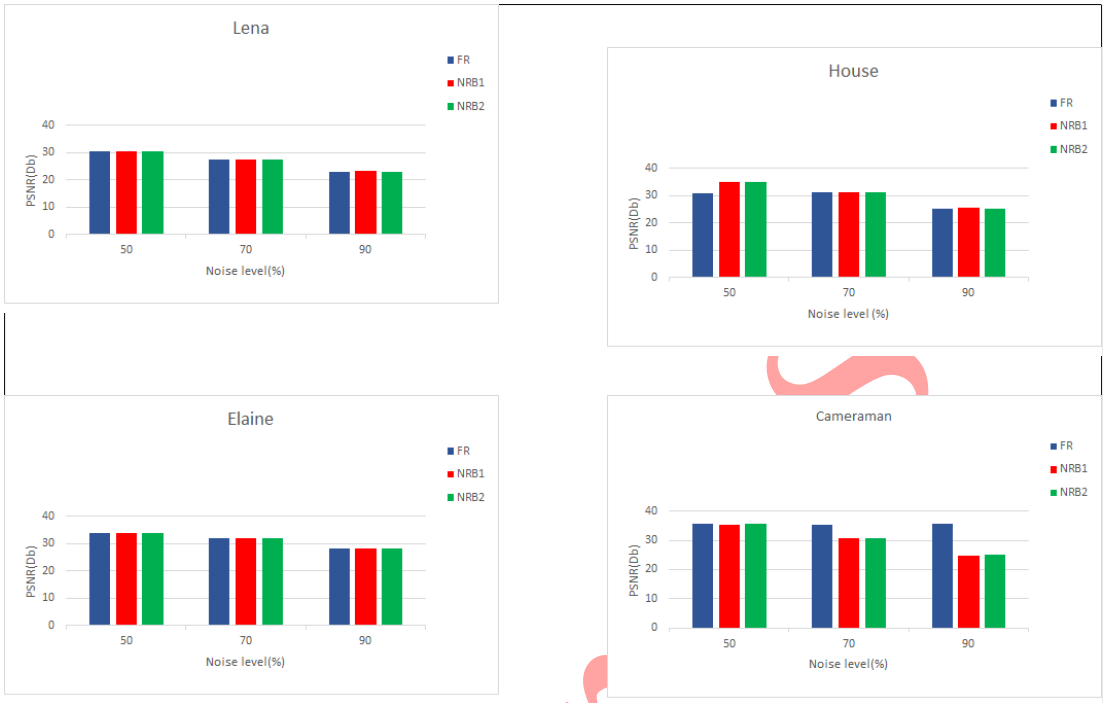


Figure 2: Comparative Analysis of PSNR for (NRB1, NRB2 and FR) methods and images.

4.4 Commentaries

For the second phase, a particularly noteworthy aspect is the application to image restoration problems, where the NRB hybrid method is tested on four standard images (Lena, House, Elaine, Cameraman)corrupted by salt-and-pepper noise at various intensity levels (50%, 70%, and 90%). The results in Table 3 and the accompanying PSNR (Peak Signal-to-Noise Ratio) plots (Figures 3, 4, 5, 6 and 7) indicate that the NRB1 method not only reduces the number of iterations and computational effort in several cases but also maintains or improves the quality of the restored images, especially at higher noise levels and show the superiority of the NRB algorithm compared to NRB2 and FR algorithms.

5 Conclusion

In this work, we have introduced and analyzed two novel hybrid conjugate gradient methods, NRB1 and NRB2, for solving unconstrained optimization problems. The NRB1 method is based on a convex combination of the BA and CD methods, with its parameter specifically

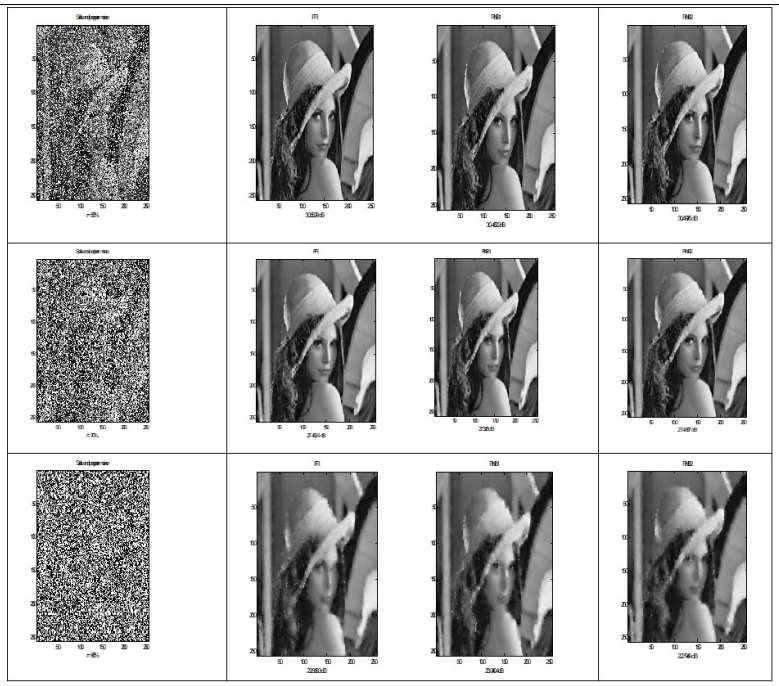


Figure 3: Restoration results for the 256×256 Lena image corrupted by salt-and-pepper noise at 50%, 70%, and 90% intensities, using NRB1, NRB2 and FR methods.

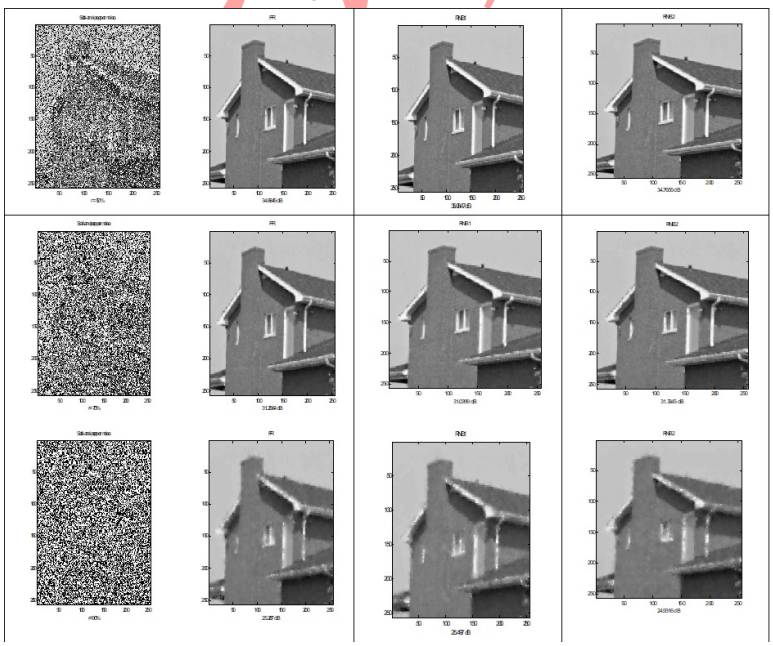


Figure 4: Restoration results for the 256×256 House image corrupted by salt-and-pepper noise at 50%, 70%, and 90% intensities, using NRB1, NRB2 and FR methods.

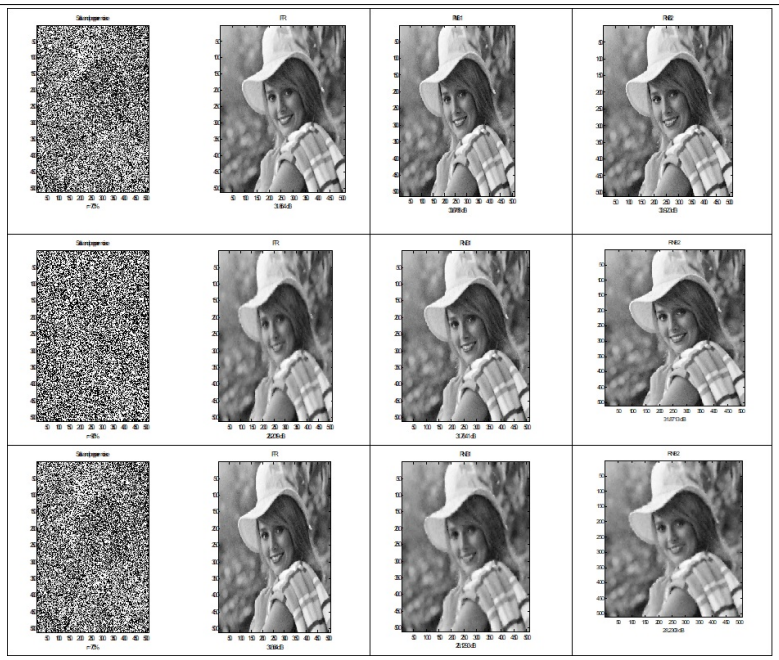


Figure 5: Restoration results for the 256×256 Elaine image corrupted by salt-and-pepper noise at 50%, 70%, and 90% intensities, using NRB1, NRB2 and FR methods..

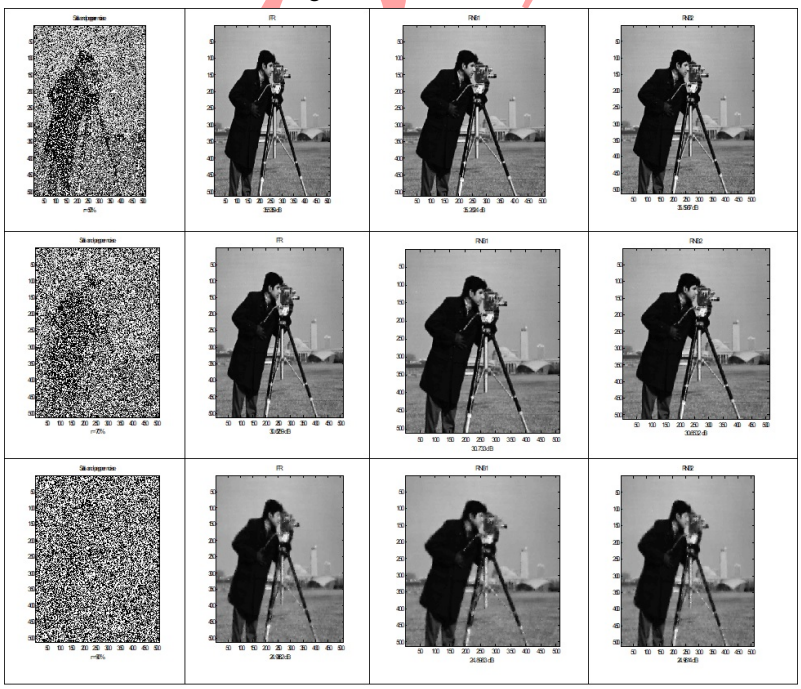


Figure 6: Restoration results for the 256×256 Cameraman image corrupted by salt-and-pepper noise at 50%, 70%, and 90% intensities, using NRB1, NRB2 and FR methods.

chosen to ensure that the conjugate gradient direction aligns with the Newton direction. In contrast, the NRB2 method computes its parameter to satisfy the conjugacy condition independently of the line search. We provided a rigorous convergence analysis under strong Wolfe line search conditions and demonstrated that the proposed algorithms satisfy the sufficient descent property, thereby ensuring global convergence. Extensive numerical experiments on standard test problems from the CUTE library, evaluated via Dolan–Moré performance profiles, confirm that NRB1 consistently outperforms the classical BA, CD, and DY methods as well as NRB2 in terms of iteration count, CPU time, and function evaluations. The application to image restoration under salt-and-pepper noise at intensities of 50%, 70%, and 90% on four standard images further demonstrates the practical utility of NRB1, which achieves competitive PSNR values with substantially fewer iterations than the Fletcher–Reeves and RNB2 algorithms.

Limitations: The numerical assessment carried out in this paper is restricted to problems of size $n \leq 1000$, so the performance of the proposed algorithms for $n \gg 10^4$ remains an open question. Moreover, for some standard test problems (White, Woods, DBVF, BRYBND, ExtendedPenalty, ExtTridiag, QuadraticQF1), the algorithms do not converge when $n \geq 500$, which reduces the scope of the conclusions.

Future Work: A natural extension of this work is to investigate alternative three-term convex combinations of conjugate-gradient parameters. Another promising research direction is to employ more sophisticated step-size strategies, such as advanced line-search procedures or trust-region methods, in order to compute the step length α_k.

Declarations

Availability of Supporting Data

All data generated or analyzed during this study are included in this published article.

Funding

This research was conducted without external funding, grants, or financial support.

Conflict of Interest

The author declares no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author Contributions

Nabil Sellami: Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Writing – Original Draft, Writing – Review & Editing, Project Administration.

In Press

Romaissa Mellal: Conceptualization, Software and Validation, Data Curation, Writing – Review & Editing, Visualization, Supervision. **Basim A. Hassan**: Conceptualization, Software and Validation, Data Curation, Writing – Review & Editing, Visualization.

Artificial Intelligence Statement

Artificial intelligence (AI) tools, including large language models, were used solely for language editing and improving readability. AI tools were not used for generating ideas, performing analyses, interpreting results, or writing the scientific content. All scientific conclusions and intellectual contributions were made exclusively by the authors.

Publisher's Note

The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Andrei, N. (2008). "Another hybrid conjugate gradient algorithm for unconstrained optimization". *Numerical Algorithms*, 47(1), 143–156. https://doi.org/10.1007/s11075-007-9152-9.

[2] Andrei, N. (2009). Another nonlinear conjugate gradient algorithm for unconstrained optimization. *Optimization Methods and Software*, 24(1), 89–104. https://doi.org/10.1080/10556780802393326

[3] Andrei, N. (2020). *Nonlinear conjugate gradient methods for unconstrained optimization*. Springer International Publishing, Cham. (Springer Optimization and Its Applications, Vol. 158). https://doi.org/10.1007/978-3-030-42950-8.

[4] Al-Bayati, A. Y., Al-Assady, N. H. (1986). "Conjugate gradient method". *Technical Report, School of Computer Studies, University of Leeds, UK*.

[5] Babaie–Kafaki, S., Mirhoseini, N., Aminifard, Z. (2023). "A descent extension of a modified Polak–Ribiere–Polyak method with application in image restoration problem". *Optimization Letters, 17(2), 351–367*. https://doi.org/10.1007/s11590-022-01878-6.

[6] Bongartz, I., Conn, A. R., Gould, N. I. M., & Toint, P. L. (1995). "CUTE: constrained and unconstrained testing environments". *ACM Transactions on Mathematical Software, 21, 123–160*. https://doi.org/10.1145/200979.201043.

In Press

[7] Djordjević, S. S. (2017). New hybrid conjugate gradient method as a convex combination of LS and CD methods. *Filomat*, 31(6), 1813–1825. https://doi.org/10.2298/FIL1706813D

[8] Dolan, E. D., & Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2), 201–213. https://doi.org/10.1007/s101070100263

[9] Fletcher, R. (1987). *Practical methods of optimization* (2nd ed.). John Wiley & Sons. https://doi.org/10.1002/9781118723203

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

[11] Hamdi, A., Sellami, B., Belloufi, M. (2020). "A new hybrid conjugate gradient method as a convex combination of BA and DY methods". *Communications in Optimization Theory, 2020 (Article ID 5), 1–9*. https://doi.org/10.23952/cot.2020.5.

[12] Hassan, B. A., Sadiq, H. M. (2022). Efficient new conjugate gradient methods for removing impulse noise images. *European Journal of Pure and Applied Mathematics*, 15(4), 2011–2021. https://doi.org/10.29020/nybg.ejpam.v15i4.4568

[13] Hestenes, M. R., Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6), 409–436. https://doi.org/10.6028/jres.049.044

[14] Liu, J. K., Li, S. J. (2014). "New hybrid conjugate gradient method for unconstrained optimization". *Applied Mathematics and Computation*, *245*, 36–43. https://doi.org/10.1016/j.amc.2014.07.096

[15] Mellal, R., Sellami, N. (2025). A novel hybrid conjugate gradient algorithm for solving unconstrained optimization problems. *Statistics, Optimization and Information Computing*. https://doi.org/10.19139/soic-2310-5070-2807

[16] Mellal, R., Sellami, N., Hassan, B. A. (2025). Novel hybrid conjugate gradient technique based on the Newton direction applied to image restoration problem. *Statistics, Optimization and Information Computing*, 15(3), 1758–1775. https://doi.org/10.19139/soic-2310-5070-2897

[17] Nocedal, J., Wright, S. J. (2006). *Numerical optimization* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-40065-5

[18] Powell, M. J. D. (1977). "Restart procedures of the conjugate gradient method". *Mathematical Programming*, 2, 241–254. https://doi.org/10.1007/BF01593790

In Press

[19] Wolfe, P. (1969). "Convergence conditions for ascent methods". *SIAM Review*, 11, 226–235.

[20] Wolfe, P. (1971). "Convergence conditions for ascent methods. II: Some corrections". *SIAM Review*, 13, 185–188.

Authors Bio-sketches

Nabil Sellami is an associate professor in mathematics at 8th May 1945 university, Guelma, Algeria. His research interests include optimization methods, nonlinear dynamical systems, and differential equations. Corresponding author: Email: sellami.nabil@univ-guelma.dz.

Romaissa Mellal is an associate professor in mathematics at 8th May 1945 university, Guelma, Algeria. Her work focuses on numerical optimization methods, differential equations, and dynamical systems.

Basim A. Hassan is a professor of mathematics at the University of Mosul, Iraq, specializing in numerical optimization. His research focuses on conjugate gradient methods and their applications, particularly in image restoration and large-scale unconstrained optimization.

In Press