Al-Rafidain Journal of Computer Sciences and Mathematics (RJCSM)

www.csmj.uomosul.edu.iq

Development of a Spectral Conjugate Gradient Method for Solving Optimization Problems

Kamal Jameel Murad[1] and Salah Gazi Shareef[2]

[1,2] Department of Mathematics, College of Science, University of Zakho, Kurdistan Region, Iraq
Email: kamal.murad@uoz.edu.krd[1] and salah.shareef@uoz.edu.krd[2]

Article information

Article history:
Received 12 June, 2025
Revised 27 July, 2025
Accepted 13 August, 2025
Published 25 December, 2025

Keywords:
Optimization,
Spectral Conjugate Gradient,
Descent Property,
Global Convergent.

Correspondence:
Kamal Jameel Murad
Email:
kamal.murad@uoz.edu.krd

Abstract

Conjugate gradient methods have been favored to use for their efficiency in solving large-scale unconstrained optimization problems, primarily because of their low memory requirements and exclusive to use the first-order derivative information. In this paper, we introduce a spectral conjugate gradient method that enhances the classical approach by merge a spectral property directly into the determination of the search direction. At the core of our method lies a developed formulation of a spectral search direction and a more precisely adjusted conjugate gradient coefficient, both derived as extensions of established conjugacy condition. To ensure numerical stability, we also include a correction term that accounts for the limitations of machine precision. Our theoretical analysis confirms that the developed method generates search directions satisfying the descent condition, which is critical for ensuring convergence. To assess its real-world effectiveness, we subjected the spectral conjugate gradient method to an extensive set of numerical experiments and benchmarked its performance against that of a standard conjugate gradient method. By using range of test problems, our method consistently delivered superior results, particularly in reducing the number of function evaluations and exhibiting improved scalability in higher-dimensional settings. These findings strongly indicate the spectral conjugate gradient method's potential as a reliable and efficient tool for optimization. Future research may explore further refinements to the method's theoretical foundations, investigate its performance in constrained or stochastic environments, and apply it to practical optimization challenges such as neural network training, signal recovery, structural design, and control system calibration.

DOI: 10.33899/csmj.2025.161414.1201,

1. Introduction

Unconstrained optimization represents a fundamental concept within mathematical modeling, frequently encountered across various scientific and engineering disciplines. These problems focus on minimizing an objective function $f(x)$, defined over $x \in R^n$, without imposing explicit constraints on the decision variable x. Unconstrained optimization problems lie at the heart of both theoretical advancements and the development of practical, efficient optimization algorithms. Their direct formulation and widespread occurrence in real-world applications make them foundational to the field.

Their direct formulation and popular occurrence in real-world applications make them foundational to the field. Proposing reliable and computationally efficient algorithms is particularly pivotal when solving large-scale unconstrained problems. Classical methods often struggle with such challenges because of the excessive memory demands or slow convergence rates. Among the most widely used methods are first-order iterative methods, such as quasi-Newton (QN) and conjugate gradient (CG) algorithms. These methods are chosen for their ability to deal with large datasets while maintaining relatively low memory.

Unconstrained optimization arises in numerous real-world applications practical. For instance, in medical imaging, it plays a key role in reconstructing magnetic resonance (MR) images by formulating the inverse problem

as an optimization task [1]. In finance, such methods are central to portfolio optimization models that aim to balance expected returns with associated risks [2,3]. In robotics, motion control problems are frequently modeled as unconstrained minimization problems [4]. also, in signal processing and computer vision, applications such as compressed sensing and image restoration often rely on solving unconstrained optimization problems to recover accurate signals or images from incomplete or noisy [5,6,7].

Moreover, the fast progress in machine learning and artificial intelligence has significantly high the importance of unconstrained optimization, particularly in the training of deep neural networks. These training problems typically involve highly nonconvex objective functions defined over extremely high-dimensional, demanding optimization techniques that are fast converging and also robust and stable [8,9,10]. Consequently, improving optimization algorithms remains an active and evolving research, with ongoing efforts directed toward improve convergence, enhancing generalization performance, and computational efficiency.

Motivated by these considerations, this paper focuses on unconstrained optimization problems, formally described as the minimization of function defined over R^n, with no restrictions on the domain variables:

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

where the objective function $f : R^n \to R$ possesses continuous partial derivatives, and its gradient, denoted by $\nabla f(x) = g(x)$, is available.

Our goal is to contribute to this field by proposing developed a new method that combines spectral properties with the conjugate gradient framework to enhance convergence and numerical performance.

CG methods constitute a prominent class of iterative algorithms widely employed for solving large-scale optimization problems (1) due to their simple iterative structure, relatively fast convergence properties, and low memory requirements. The iterative steps of the classical CG method are defined by:

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

and the search direction dk is given by:

$$d_{k+1} = \begin{cases} -g_1, & k = 0 \\ -g_{k+1} + \beta_k d_k, & k \geq 1' \end{cases} \tag{3}$$

where $g_k = g(x_k)$, and β_k is the conjugate parameter, and α_k is the step length determined by an appropriate exact or inexact line search. Numerous researchers have investigated the convergence properties of CG methods under various line search conditions, with some employing exact line search (ELS) and others utilizing the strong Wolfe line search (SWL) conditions, defined as:

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \\ g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k \end{cases} \tag{4}$$

or the SWL

$$\begin{cases} f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \\ |g(x_k + \alpha_k d_k)^T d_k| \leq \sigma |g_k^T d_k|, \end{cases} \tag{5}$$

where the parameters $0 < \delta < \sigma < 1$ are typically required. Different choices for the conjugate parameter $\beta_k$lead to different CG methods with varying numerical performance and convergence behavior [17]. Well-known formulas for βk include those proposed by Hestenes and Stiefel (HS) [11], Fletcher and Reeves (FR) [12], Polak, Ribière, and Polyak (PRP) [13, 14], Dai and Yuan (DY) [15], Liu-Storey (LS) [16], Conjugate-Descent (CD) methods [17], given respectively by:

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \tag{6}$$

$$\beta_k^{FR} = \frac{g_{k+1}^T g_k}{g_k^T g_k} \tag{7}$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{g_k^T g_k} \tag{8}$$

$$\beta_k^{DY} = \frac{g_{k+1}^T g_k}{d_k^T y_k} \tag{9}$$

$$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-g_k^T d_k}, \tag{10}$$

$$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-g_k^T d_k}, \tag{11}$$

where $y_k = g_{k+1} - g_k$ and $\|\cdot\|$ denotes the Euclidean norm in R^n. Theoretically, when an exact minimization rule is employed, all these choices of β_k are equivalent for strongly convex quadratic functions. However, for non-quadratic objective functions, each choice of β_k can result in significantly different numerical performance [18].

It is established that the FR and DY CG methods possess favorable global convergence properties. However, their numerical performance in practice is often not optimal. Conversely, the PRP and HS methods typically exhibit excellent performance in practical computations, but establishing their global convergence properties can be challenging. To address these limitations of classical CG methods, considerable research has focused on developing improved CG methods with enhanced theoretical properties and numerical performance, as exemplified by the methods presented in References [19, 20, 21,22,23,24,25,26].

Despite the numerous successes of classical conjugate gradient methods, they face challenges in efficiently handling high-dimensional and strongly nonlinear problems. Achieving convergence often requires accurate line search properties, and the performance of these methods can be sensitive to the choice of the conjugate parameter. Furthermore, ensuring the sufficient descent property, which is fundamental for strong global convergence, is not always guaranteed in practice, especially with inexact line searches.

Given these limitations, spectral gradient methods have emerged as a promising alternative that seeks to incorporate

spectral information from previous iterations with the aim of accelerating convergence and improving numerical performance. These methods leverage eigenvalues or approximations of the Hessian of the objective function to define more effective search directions. However, there remains a need for the development of new spectral conjugate gradient methods that combine practical efficiency with strong theoretical properties, such as guaranteeing the sufficient descent property at each iteration.

In [27, 28], Barzilai, Borwein, and Raydan independently introduced and analyzed spectral gradient methods for unconstrained optimization. Subsequently, drawing inspiration from spectral gradient methods, significant efforts have been directed towards modifying traditional CG methods. Brigin and Martínez [29] proposed a spectral CG method where the search direction is defined as:

$$d_{k+1} = -\theta_k g_{k+1} + \beta_k v_k, d_1 = -\theta_1 g_1, \quad (12)$$

where θ_k and β_k named a spectral parameter and a CG parameter, respectively, are given by

$$\theta_k = \frac{v_k^T v_k}{v_k^T y_k}, \quad (13)$$

$$\beta_K^{BM1} = \frac{(\theta_k y_k - v_k)^T g_{k+1}}{v_k^T y_k}. \quad (14)$$

$$\beta_K^{BM2} = \frac{\theta_k y_k^T g_{k+1}}{\alpha_k \theta_{k-1} g_k^T g_k}, \quad (15)$$

$$\beta_K^{BM3} = \frac{\theta_k g_{k+1}^T g_{k+1}}{\alpha_k \theta_{k-1} g_k^T g_k}, \quad (16)$$

The experimental results obtained using the Wolfe line search strategy on the three CG formulas indicate that the coefficient β_K^{BM1} delivers the most favorable numerical performance. Based on certain reasonable assumptions, Birgin and Martínez [28] demonstrated that their spectral CG method achieves global convergence. Nevertheless, it is important to note that spectral CG approaches do not inherently guarantee descent directions [29]. To address this limitation, Andrei [30] introduced a scaled CG algorithm designed to ensure descent properties under the Wolfe line search conditions. Subsequently, Jiang et al. [31] developed a spectral CG method with sufficient descent properties, building upon the modified CG approach proposed by Zhang et al. [32], where the search direction was defined using the β_K^{PRP}coefficient.

$$\theta_k = \frac{y_{k-1} d_{k-1}^T}{\|g_{k-1}\|^2} - \frac{g_k^T g_k d_{k-1}^T g_{k-1}}{\|g_k\|^2 \|g_{k-1}\|^2}, \quad (17)$$

$$\beta_K^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \quad (18)$$

In this context, the vector $y_{k-1} = g_k - g_{k-1}$,represents the difference between successive gradients. The corresponding algorithm was implemented using a modified Armijo-type line search strategy, and it was later shown to be globally convergent under certain mild assumptions. Building upon this, Liu and Jiang [33] introduced a spectral conjugate gradient method known as SCD, which is derived from the classical CD method. The SCD algorithm is notable for maintaining the sufficient descent property regardless of the line search technique employed, and its global convergence has been established under the strong Wolfe line search conditions. The method is defined by the following expressions:

$$\theta_k = 1 - \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}, \quad (19)$$

$$\beta_K^{CD} = \frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}. \quad (20)$$

Subsequently, Liu et al. [34] proposed another variant of the spectral CG method, which integrates the CD and DY methods. In this formulation, the CG coefficient is computed as:

$$\beta_k = \beta_K^{CD} + \min\{0, \psi_k, \beta_K^{CD}\}, \quad (21)$$

$$\theta_k = 1 - \frac{g_{k-1}^T d_{k-1}}{g_k^T d_{k-1}}, \quad (22)$$

$$\psi_k = -\frac{g_{k-1}^T d_{k-1}}{d_{k-1}(g_{k-1} - g_k)}, \quad (23)$$

In 2010, Andrei [35] introduced another spectral CG method where the search direction is given by:

$$d_{k+1} = -\theta_{k+1} g_{k+1} + \beta_k^N s_k, d_1 = -g_1. \quad (24)$$

with

$$\beta_k^{\mathrm{N}} = \frac{\|g_{k+1}\|^2}{y_k^{\mathrm{T}} s_k} - \frac{\|g_{k+1}\|^2 s_k^{\mathrm{T}} g_{k+1}}{\left(y_k^{\mathrm{T}} s_k\right)^2}, \quad (25)$$

and

$$\theta_{k+1} = \frac{1}{y_k^{\mathrm{T}} g_{k+1}} \left(\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|^2 s_k^{\mathrm{T}} g_{k+1}}{y_k^{\mathrm{T}} s_k} \right). \quad (26)$$

The directions yielded by Equations (24)-(26) possess descent property as follows:

$$g_{k+1}^{\mathrm{T}} d_{k+1} \leqslant -(\theta_{k+1} - 1/4)\|g_{k+1}\|^2. \quad (27)$$

This shows that the direction is descent only in case $\theta_{k+1} > 1/4$. Therefore, to obtain descent in any case, Andrei [35] reset $\theta_{k+1} = 1$ in case $\theta_{k+1} \leqslant 1/4$.

Further examples of spectral CG methods based on the structure of Equation (12) can be found in References [36, 37].

While numerous numerical experiments have demonstrated the superior numerical performance of spectral CG methods compared to traditional CG methods, ensuring the descent property, particularly the sufficient descent property, for the search directions in spectral CG methods remains a challenge. This motivates further research into the development of more robust spectral CG methods.

The primary objective of this study is to develop a novel (SCG) method specifically designed to improve the numerical efficiency of solving large-scale unconstrained optimization problems. The main contributions of this work

can be summarized as follows. First, we propose the development of a new conjugate parameter, derived from a modified form of the Dai–Liao conjugacy condition, which is integrated within the structure of the proposed spectral CG framework. Second, we prove that the resulting search directions satisfy both the (descent and sufficient descent) conditions, and we ensure that the spectral parameter remains bounded an essential aspect for establishing global convergence. Third, we validate the effectiveness of the proposed method through comprehensive numerical experiments, demonstrating its superior performance relative to existing methods on a broad set of standard large-scale benchmark problems.

The structure of the paper is organized as follows. In Section tow introduces the mathematical preliminaries and outlines the assumptions employed in constructing the new method. In Section 3, the derivation of the search directions and provide theoretical results confirming that they satisfy the descent and sufficient descent properties are present. Section 4 is dedicated to analyzing the boundedness of the spectral parameter, which plays a critical role in guaranteeing global convergence. Section 5 reports the results of numerical experiments that compare the proposed method with several state-of-the-art optimization techniques. Finally, Section 6 summarizes the main findings and outlines potential directions for future research.

2. Derivation of a Developed SCG Method and Its Algorithm

In this section, we present the development and derivation of a new SCG method designed to solve optimization problems of the form defined in equation (1). The proposed method introduces a novel spectral search direction defined as:

$$d_{k+1} = -\theta_k g_{k+1} + \beta_k d_k, d_1 = -\theta_1 g_1, \tag{28}$$

where $\theta_k = \frac{v_k^T v_k}{v_k^T y_k}$ is the spectral parameter and $v_k = x_{k+1} - x_k = \alpha_k d_k$. This parameter plays a central role in controlling the direction and efficiency of the iterative optimization process. For the method to be both theoretically sound and practically robust, it is crucial that θ_k remains bounded throughout the iterations. Boundedness ensures stable behavior of the generated directions and supports the convergence analysis.

To enhance the method's performance, particularly in large-scale or ill-conditioned problems, we construct the algorithm to ensure that each search direction satisfies both the (descent and sufficient descent) conditions, which are fundamental for global convergence and numerical stability.

We define a modified gradient-difference vector inspired by [38] as follows:

$$y_k^* = y_k + \frac{(0.2-\rho_k)}{(1-\rho_k)}\left(\frac{\|v_k\|-2\sqrt{\epsilon_m}(1+\|x_{k+1}\|)}{2\sqrt{\epsilon_m}(1+\|x_{k+1}\|)}\right) y_k \ , \tag{29}$$

where $0.2 < \rho_k < 1$ and ϵ_m is error machine used for accuracy which is the smallest positive < 1. Extending the conjugacy condition originally proposed by Dai and Liao [39], we adopt the following modified form:

$$d_{k+1}^T y_k^* = d_{k+1}^T y_k (1 + K_1 K_2) = -t g_{k+1}^T v_k, \tag{30}$$

Based on Equation 30, we arrive at the following result:

$$d_{k+1}^T y_k = -\frac{t g_{k+1}^T v_k}{(1+K_1 K_2)} \tag{31}$$

where $K_1 = \frac{(0.2-\rho_k)}{(1-\rho_k)}$ and $K_2 = \left(\frac{\|v_k\|-2\sqrt{\epsilon_m}(1+\|x_{k+1}\|)}{2\sqrt{\epsilon_m}(1+\|x_{k+1}\|)}\right)$.

By multiplying the search direction defined in Equation (28) by y_k, and using the conjugacy condition given in Equation (31), we derive the following expression for the developed conjugate parameter:

$$\beta_k^{New} = \frac{\theta_k g_{k+1}^T y_k - \frac{t g_{k+1}^T v_k}{(1+K_1 K_2)}}{d_k^T y_k},$$

This yields the following:

$$\beta_k^{New} = \theta_k \beta_k^{HS} - \frac{t g_{k+1}^T v_k}{(1+K_1 K_2)\, d_k^T y_k}. \tag{32}$$

where β_k^{HS} denotes the classical Hestenes–Stiefel parameter defined in (6).

Based on the proposed spectral search direction (28) and the updated conjugate parameter (32), we now outline the algorithmic steps for the new SCG method.

Algorithm: Steps of the Developed New SCG Method

Step; 1 . Given $x_0 \in R^n$

Step; 2 . set $d_0 = -g_0$, $k = 0$. If $\|g_k\| = 0$ stop, otherwise continue.

Step; 3 . Compute the α_k by using minimize $f(x_k + \alpha_k d_k)$.

Step; 4 . Determine $x_{k+1} = x_k + \alpha_k d_k$.

Step; 5 . Compute g_{k+1}, if $\|g_{k+1}\| \leq 10^{-5}$ stop, else continue to Step 6.

Step; 6 . Determine d_{k+1} by using (28) and (32).

Step; 7 . If $\|g_{k+1}\|^2 \leq \frac{|g_k^T g_{k+1}|}{0.2}$ is satisfied go to step 3, else $k = k + 1$ and go to step 3.

3. The Descent Properties of the Developed SCG Method

In this section, we aim to prove that the proposed New SCG method satisfies both the descent condition and the sufficient descent property. To ensure the global convergence of the method, it is essential to demonstrate that the search direction fulfills these two key properties, as they are fundamental to the theory of unconstrained optimization.

Theorem 1: The search direction d_{k+1} of the New SCG

method, generated by (28) where β_k is defined in (32), satisfies the descent condition:

$$g_{k+1}^T d_{k+1} \leq 0 \quad (33)$$

Proof: Multiplying both sides of equation (28) by g_{k+1}^T, we get:

$$g_{k+1}^T d_{k+1} = -\theta_k g_{k+1}^T g_{k+1} + \beta_k g_{k+1}^T d_k, \quad (34)$$

substituting the definition of β_k from equation (32):

$$g_{k+1}^T d_{k+1} = -\theta_k g_{k+1}^T g_{k+1} + \theta_k \frac{g_{k+1}^T y_k}{d_k^T y_k} g_{k+1}^T d_k - \frac{t g_{k+1}^T v_k}{(1+K_1K_2)\, d_k^T y_k} g_{k+1}^T d_k, \quad (35)$$

since the above equation satisfying the descent condition if the search direction is exact, i.e.

$$g_{k+1}^T d_{k+1} = -\theta_k g_{k+1}^T g_{k+1} \leq 0. \quad (36)$$

Because $\theta_k \geq 0$. However, if the search direction (35) is inexact we will prove that (33).

Since in general the inequality $g_{k+1}^T d_k \leq d_k^T y_k$ true and by using the Cauchy-Schwarz inequality $g_{k+1}^T y_k \leq \|g_{k+1}\|\|y_k\|$, we have

$$g_{k+1}^T d_{k+1} \leq -\theta_k \|g_{k+1}\|^2 + \theta_k \|g_{k+1}\|\|y_k\| - \frac{t\, v_k^T y_k}{(1+K_1K_2)}, \quad (37)$$

simplify

$$g_{k+1}^T d_{k+1} \leq -\theta_k \|g_{k+1}\|^2 \left(1 - \frac{\|y_k\|}{\|g_{k+1}\|}\right) - \frac{v_k^T y_k}{(1+K_1K_2)}, \quad (38)$$

since, $1 - \frac{\|y_k\|}{\|g_{k+1}\|} \geq 0$. So, equation (38) can be write as

$$g_{k+1}^T d_{k+1} \leq -\frac{t\, v_k^T y_k}{(1+K_1K_2)}, \quad (39)$$

clearly, $t, v_k^T y_k$ are non-negative and given that $0.2 < \rho_k < 1$, , it follows that $K_1 < 0$. Also, since

$0 < \|v_k\| = \|x_{k+1} - x_k\| \leq \|x_{k+1}\|$,

$\Rightarrow 0 \leq \frac{\|v_k\|}{2\sqrt{\omega}1+\|x_{k+1}\|)} \leq 1$. This implies $K_2 < 0$.

Therefore, since both$K_1 < 0$ and $K_2 < 0$ it follows that: $K_1K_2 > 0$, the right-hand side is non-positive. Thus, the descent condition is satisfied:

$g_{k+1}^T d_{k+1} \leq 0$.

Hence, the proof is complete

4. The Global Convergence Property of the Developed SCG Method

In this section, we aim to establish the global convergence property of the proposed New SCG method. To achieve this, we introduce the following standard assumptions:

Assumption:

I. The level set S is bounded $S = \{x \,|f(x) \leq f(x_0)\}$.

II. The objective function f is continuously differentiable in a neighborhood N of some point S and its gradient is Lipschitz continuous on S with Lipschitz constant $L > 0$, that is,

$$\|g(x) - g(y)\| \leq L\|x - y\| \;\; \forall\, x, y \in S \quad (40)$$

As a consequence of these assumptions, there exists a constant b, such that

$$\|g(x)\| \leq b \quad \forall x \in S. \quad (41)$$

Building upon these conditions, we proceed to establish the global convergence of the New SCG algorithm as follows.

Lemma 1: [39] Assume that conditions (I)–(II) hold. Let the iterative sequence be generated by methods (2) and (10), where the search direction d_{k+1} is a descent and the α_k satisfies the Wolfe conditions. If the condition

$$\sum_{k\geq 1} \frac{1}{\|d_{k+1}\|^2} = \infty. \quad (42)$$

Then,

$$\lim_{k\to\infty} inf\, \|g_{k+1}\| = 0. \quad (43)$$

In light of this result and the previous discussion, the global convergence of the proposed New SCG algorithm is thereby established.

Theorem 3: If assumptions (I)–(II) are true and the corresponding sequences of $\{x_k\}$, $\{d_k\}$, $\{g_k\}$, $\{\alpha_k\}$ are generated by new SCG-Algorithm, then we arrive at the conclusion that

$$\lim_{k\to\infty} inf\, \|g_{k+1}\| = 0. \quad (44)$$

Proof: From the search direction (28), and the new parameter β_k in (32), we have

$$\|d_{k+1}\| \leq \left|\frac{v_k^T v_k}{v_k^T y_k}\right| \|g_{k+1}\| + \left|\frac{v_k^T v_k}{v_k^T y_k}\right| \left|\frac{g_{k+1}^T y_k}{d_k^T y_k}\right| \| d_k\| + \left|\frac{t\, g_{k+1}^T v_k}{(1+M_1M_2)\, d_k^T y_k}\right| \| d_k\|, \quad (45)$$

since $g_{k+1}^T v_k \leq \alpha_k d_k^T y_k$ and by using the Lipschitz Condition $\|y_k\| \leq L\|v_k\|$ along with the fact that $y_k^T v_k \geq \vartheta \|v_k\|^2$, we obtained

$$\|d_{k+1}\| \leq \frac{1}{\vartheta}\|g_{k+1}\| + \frac{L\alpha_k \|g_{k+1}\|}{\vartheta^2 \|v_k\|} + \left|\frac{t\alpha_k}{(1+M_1M_2)}\right| \| d_k\|, \quad (46)$$

from Equation (32), we get that the norm of the gradient is bounded:

$$\|d_{k+1}\| \leq \frac{1}{\vartheta} b + \frac{Lb\alpha_k}{\vartheta^2 \| v_k\|} + \frac{t}{(1+M_1M_2)} \| v_k\|, \quad (47)$$

let $D = max\{\|v_k\| = \|x - x_k\|\}, \forall\, x, x_k \in R\}$.

Hence the inequality in Equation (47), becomes

$$\|d_{k+1}\| \leq \frac{1}{\vartheta} b + \frac{Lb\alpha_k}{\vartheta^2 D} + \frac{t}{(1+M_1M_2)} D = \emptyset, \quad (48)$$

$$\Rightarrow \sum_{k\geq 1} \frac{1}{\|d_{k+1}\|^2} \geq \sum_{k\geq 1} \frac{1}{\emptyset^2} = \infty\,, \quad (49)$$

$$\Rightarrow \sum_{k\geq 1} \frac{1}{\|d_{k+1}\|^2} = \infty. \quad (50)$$

By using lemma (1), we get $\lim_{k\to\infty} inf \|g_{k+1}\| = 0$. which completes the proof.

5. Numerical Results

In this section, we present the numerical evaluation of our proposed method (denoted by New SCG for short) on a set of well-established the optimization test problems [40] of varying dimensions are listed in the appendix. We compare its performance against the Hestenes-Stiefel (HS) method. All algorithms were implemented in FORTRAN 95. The step length for each iteration was determined using a cubic interpolation line search routine that utilized both function and gradient values. In the result tables, the letter "F" indicates that a particular method failed to satisfy the termination criteria within the maximum allowed number of iterations or function evaluations. The results presented in Table 1 and summarized in Table 2 are based on the [number of iterations, the number of function evaluations] write as [NOI, NOF] respectively, required to reach a solution. The experimental results reported in Table 2 suggest that the New SCG algorithm demonstrates superior performance compared to the HS algorithm in terms of both NOI and NOF.

To provide a visual comparison of the algorithms' performance, we employed the performance profile tool introduced by Dolan and Moré [41]. This tool allows for a robust assessment of the relative efficiency of different solvers across a set of test problems. For a given set of problems P with n_p problems and a set of solvers S with ns solvers, and for each problem $p \in P$ and solver $s \in S$, we define $t_{p,s}$ as the value of a specific performance metric (NOI, NOF) required by solver s to solve problem p. The performance ratio for solver s on problem p is then defined as:

$$r_{p,s} = \frac{t_{p,s}}{\min_{s'\in S}\{t_{p,s'}\}}. \quad (51)$$

The performance profile of a solver s is given by the function $\rho_s(\tau)$, which represents the fraction of problems for which the performance ratio of solver s is within a factor τ of the best performance achieved by any solver on that problem:

$$\rho_s(\tau) = \frac{1}{n_p} \text{size}\{p \in \mathcal{P} \mid r_{p,s} \leq \tau\}. \quad (52)$$

In the performance profile plots, the value $\rho_s(\tau)$ on the y - axis represents the percentage of problems solved by algorithm s with a performance within a factor τ (on the x-axis) of the best performing algorithm. An algorithm whose curve appears at the top of the plot is considered to have better overall performance compared to the other algorithms in the comparison for the given metric. Furthermore, the value of $\rho_s(1)$ indicates the percentage of problems for which algorithm s was the most efficient. The value of $\rho_s(\tau)$ as τ increases towards the right shows the robustness of the algorithm, i.e., the percentage of problems solved within a certain tolerance of the best performance.

Figures 1 and 2 illustrate the performance profiles of the compared algorithms based on the NOI and NOF metrics, respectively. In Figure 1, the curves depict the performance of all algorithms with respect to NOI, while Figure 2 displays the performance profile based on the NOF. The algorithm with the curve positioned highest in these figures demonstrates the most favorable performance for the respective metric across the tested problem set.

Table 1. Numerical Comparison of Hs and New Scg Methods on Selected Test Functions.

Method		HS		New SCG	
Test Function	Dimensions	NOI	NOF	NOI	NOF
Wolfe	4	11	24	11	29
	10	32	65	32	71
	100	49	99	43	90
	500	52	105	44	91
	1000	70	141	45	94
	5000	165	348	151	308
G-Central	4	22	159	18	78
	10	22	159	18	78
	100	22	159	19	85
	500	23	171	21	100
	1000	23	171	21	100
	5000	28	248	24	113
Nondiagonal	4	24	64	24	64

	10	26	72	26	72
	100	29	79	29	79
	500	F	F	29	79
	1000	29	79	29	79
	5000	30	81	30	81
Powell	4	37	102	30	120
	10	37	102	30	120
	100	40	117	30	120
	500	44	136	36	130
	1000	44	136	36	130
	5000	44	136	36	130
Rosen	4	30	83	29	79
	10	30	83	29	79
	100	30	83	29	79
	500	30	83	29	79
	1000	30	83	29	79
	5000	30	83	29	79
Miele	4	28	85	28	85
	10	31	102	31	102
	100	33	114	33	114
	500	40	146	40	146
	1000	46	176	46	176
	5000	54	211	54	211
Wood	4	30	68	25	62
	10	30	68	25	62
	100	30	68	27	66
	500	30	68	27	66
	1000	30	68	27	66
	5000	30	68	27	66
Sum	4	3	11	3	11
	10	6	34	6	30
	100	14	81	14	73
	500	21	124	18	88
	1000	23	128	23	105
	5000	31	159	27	124
Edger	4	5	14	5	15
	10	5	14	5	15
	100	5	14	5	15
	500	6	16	5	15
	1000	6	16	5	15
	5000	6	16	5	15
Shallow	4	8	21	8	21

	10	8	21	8	21
	100	8	21	8	21
	500	8	21	8	21
	1000	9	24	9	24
	5000	9	24	9	24
Cubic	4	12	35	12	33
	10	13	37	12	33
	100	13	37	12	33
	500	13	37	13	35
	1000	13	37	13	35
	5000	13	37	13	35
Beale	4	11	28	11	28
	10	11	28	11	28
	100	12	30	12	30
	500	12	30	12	30
	1000	12	30	12	30
	5000	12	30	12	30
Osp.	4	8	45	7	36
	10	13	58	13	52
	100	49	185	50	165
	500	112	353	107	309
	1000	156	473	152	438
	5000	256	774	254	765

Table 2. Overall Performance Comparison of Hs and New Scg Methods.

Method	NOI	NOF
HS	100%	100%
New SCG	91.92 %	87.85 %
Rate of Improvement (%)	8.08 %	12.15 %

From **Table 1**: Overall, the results suggest that the New SCG method demonstrates promising performance compared to the HS method. Several trends can be observed:

- **Iteration Efficiency:** In many test problems and dimensions, the New SCG method required fewer iterations to reach a solution compared to the HS method (as seen in problems like G-Central, Wood, Sum, and Osp. at certain dimensions).
- **Function Evaluation Efficiency:** The superiority of New SCG is more pronounced in the number of function evaluations required. In problems such as G-Central, Wood, and Sum, New SCG consumed significantly fewer function evaluations, indicating higher computational efficiency per iteration or faster convergence.

Table 2 provides a summary of the average performance of the two methods across all tested problems. The performance of the HS method is normalized to 100%, and the performance of the New SCG method is expressed as a percentage thereof. The results indicate that the New SCG method achieved an average improvement of 8.08% in the NOI and 12.15% in the NOF compared to the HS method.

Based on the numerical data provided, it can be academically concluded that the proposed New SCG method

generally exhibits superior performance compared to the HS method on the tested set of problems. The advantage is particularly evident in the reduction of the NOF required, suggesting greater computational efficiency. Furthermore, New SCG demonstrates better robustness in certain cases. While the performance of the two methods was comparable on some problems, the significant improvements observed in others, especially at higher dimensions, support the effectiveness of the proposed new method as a promising alternative to traditional conjugate gradient methods like HS.

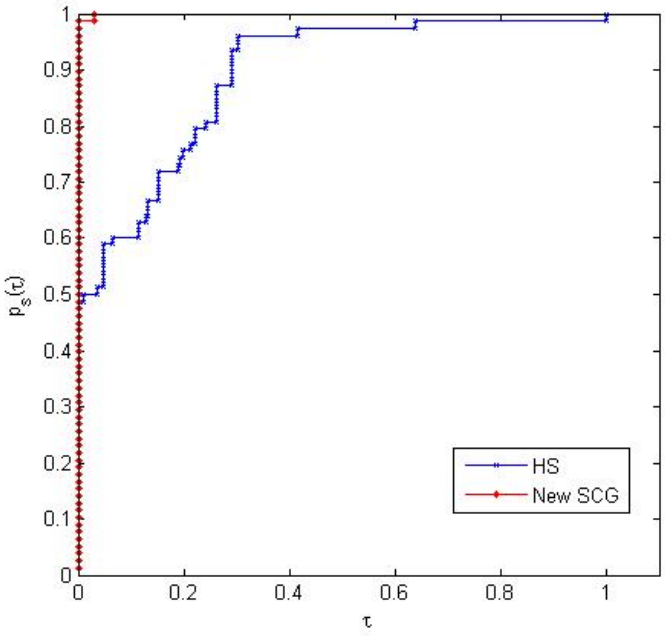

Figure 1. Performance Profile Outputs Based On (Noi).

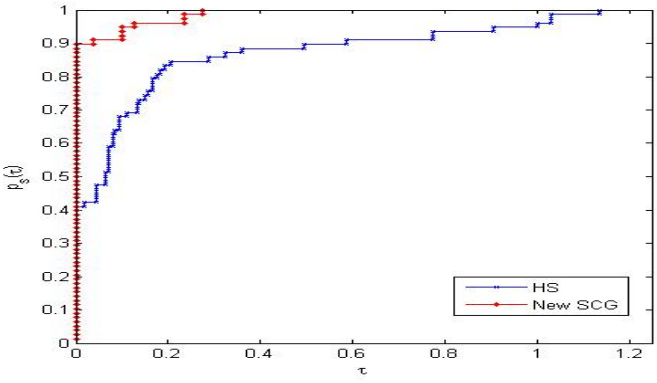

Figure 2. Performance profile outputs based on (NOF).

Figures 1 and 2 illustrate the performance profiles of the HS and New SCG methods with respect to NOI and NOF, respectively. These plots offer a visual and quantitative comparison of the relative efficiency and robustness of both methods across the benchmark test set.

In **Figure 1**, which reflects iteration-based performance:

- The performance curve of the New SCG method lies consistently above that of the HS method across all values of the performance ratio τ.
- The value $\rho_s(1)$ for New SCG is close to 1, indicating that it was the most efficient solver (in terms of iteration count) on nearly all test problems.
- The steep and early rise of the New SCG curve demonstrates its strong consistency and iteration efficiency across a wide range of problem dimensions.
- In contrast, the HS method shows a more gradual increase, reflecting less consistent performance and higher iteration counts on many problems.

In **Figure 2**, which presents the performance based on the NOF:

- The advantage of the New SCG method becomes even more pronounced. Its curve remains well above that of the HS method across all τ\tauτ values.
- The high value of $\rho_s(1)$ again confirms that New SCG was the most function-efficient solver on the vast majority of problems.
- The early and rapid growth of the New SCG profile indicates faster convergence and fewer function evaluations required, highlighting its computational efficiency.

These performance profiles support the conclusion that the New SCG method exhibits superior overall behavior compared to the HS method. It is not only more efficient in terms of iterations and function calls, but also more robust across diverse problem settings. This visual evidence complements the numerical results and reinforces the New SCG method's potential as a competitive and reliable approach for optimization problems.

Conclusion

This paper introduces a novel SCG method developed for solving unconstrained optimization problems. The proposed method features an innovative mechanism for presenting a development search direction, which uses spectral properties in combination with a conjugate coefficient derived from established conjugacy condition to enhance numerical stability. Theoretical analysis confirms that the search directions of the SCG method satisfy the descent condition a fundamental requirement to ensure convergence to a minimum. The practical performance of the proposed method extensive numerical experiments was conducted on a diverse set of standard benchmark functions. The numerical results of SCG method were systematically compared with those of the well-known Hestenes-Stiefel (HS) method. In most test cases, the SCG method consistently outperformed HS, especially in reducing the number of function evaluations while maintaining computational efficiency in high-dimensional optimization

tasks. These result findings highlight the robustness and ability of the proposed method as a reliable solution for large-scale unconstrained optimization problems. Several directions for future research are envisioned. One promising avenue involves extending the SCG framework to deal with the constrained optimization problems. Additionally, merging SCG with stochastic techniques or adaptive strategies could enhance its performance in dynamic or uncertain environments. The real-world applications are a wide range of fields, including signal and image processing (such as denoising and reconstruction), deep learning (e.g., training neural networks), structural optimization, and control engineering. Such applications would further demonstrate the versatility and resilience of the new method. These investigations are expected to deepen theoretical understanding and also to broaden the practical utility of spectral conjugate gradient methods in solving complex real-world challenges.

Conflict of interest

None.

References

[1] Yuan, T., Bi, D., Pan, L., Xie, Y., & Lyu, J. (2018). A compressive hybrid conjugate gradient image recovery approach for radial MRI. Imaging Science Journal, 66(4), 278–288.

[2] Abubakar, A. B., Kumam, P., Malik, M., Chaipunya, P., & Ibrahim, A. H. (2021). A hybrid FR-DY conjugate gradient algorithm for unconstrained optimization in portfolio selection. AIMS Mathematics, 6(6), 6506–6527.

[3] Deepho, J., Abubakur, A. B., Malik, M., & Argyros, I. K. (2022). Solving unconstrained optimization problems via hybrid CD-DY conjugate gradient methods with applications. Journal of Computational and Applied Mathematics, 405, 113823.

[4] Abubakar, A. B., Kumam, P., Malik, M., & Ibrahim, A. H. (2022). A hybrid conjugate gradient based approach for solving unconstrained optimization and motion control problems. Mathematics and Computers in Simulation, 201, 640–657.

[5] Hassan, B. A., & Alashoor, H. A. (2023). On image restoration problems using new conjugate gradient methods. Indonesian Journal of Electrical Engineering and Computer Science, 29(3), 1438–1445.

[6] Ibrahim, A. L., Fathi, B. G., & Abdulrazzaq, M. B. (2025). Conjugate gradient techniques: Enhancing optimization efficiency for large-scale problems and image restoration. Numerical Algebra, Control and Optimization. https://doi.org/10.3934/naco.2025008

[7] 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.

[8] Ibrahim, A. L., Fathi, B. G., & Abdulrazzaq, M. B. (2025). Improving three-term conjugate gradient methods for training artificial neural networks in accurate heart disease prediction. Neural Computing and Applications. https://doi.org/10.1007/s00521-025-11121-9

[9] Omar, D. H., Ibrahim, A. L., Hassan, M. M., Fathi, B. G., & Sulaiman, D. A. (2024). Enhanced conjugate gradient method for unconstrained optimization and its application in neural networks. European Journal of Pure and Applied Mathematics, 17(4), 2692–2705. https://doi.org/10.29020/nybg.ejpam.v17i4.5354

[10] Hassan, B. A., Moghrabi, I. A. R., Ibrahim, A. L., & Jabbar, H. N. (2025). Improved conjugate gradient methods for unconstrained minimization problems and training recurrent neural network. Engineering Reports, 7(1), e70019. https://doi.org/10.1002/eng2.70019

[11] 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), 409–438.

[12] 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

[13] Polak, E., & Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. RAIRO - Modélisation Mathématique et Analyse Numérique, 3(R1), 35–43.

[14] Polyak, B. T. (1969). The conjugate gradient method in extremal problems. USSR Computational Mathematics and Mathematical Physics, 9(6), 94–112. https://doi.org/10.1016/0041-5553(69)90035-4

[15] Dai, Y. H., & Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property. SIAM Journal on Optimization, 10(1), 177–182. https://doi.org/10.1137/S1052623497318992

[16] Liu, Y., & Storey, C. (1991). Efficient generalized conjugate gradient algorithms, Part 1: Theory. Journal of Optimization Theory and Applications, 69, 129-137. https://doi.org/10.1007/BF0094 0464

[17] R. Fletcher, Practical Methods of Optimization, Unconstrained Optimization, vol. 1, John Wiley and Sons, New York, 1987.

[18] Hager, W. W., & Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. Pacific Journal of Optimization, 2(1), 35–58.

[19] Liu, J. K., Feng, Y. M., & Zou, L. M. (2019). A spectral conjugate gradient method for solving large-scale unconstrained optimization. Computers & Mathematics with Applications, 77, 731–739.

[20] Jarjis, M., & Chilmerane, H. (2025). Investigation a Conjugacy Coefficient for Conjugate Gradient Method to Solving Unconstrained Optimization Problems. AL-Rafidain Journal of Computer Science and Mathematics, 19(1), 28–33. https://doi.org/10.33899/csmj.2025.154547.1153

[21] Shareef, S. G., & Ibrahim, A. L. (2016). A new conjugate gradient for unconstrained optimization based on step size of Barzilai and Borwein. Science Journal of University of Zakho, 4(1), 104–114.

[22] Aljarjary, H., Abbo, K., & Nori Mahmood, E. (2024). A modified of Dai and Yuan method for solving unconstrained optimization problems. AL-Rafidain Journal of Computer Science and Mathematics, 18(2), 20–26. https://doi.org/10.33899/csmj.2024.146806.1108

[23] Ibrahim, A. L., Sadiq, M. A., & Shareef, S. G. (2019). A new conjugate gradient coefficient for unconstrained optimization based on Dai–Liao. Science Journal of University of Zakho, 7(1), 34–36. https://doi.org/10.25271/sjuoz.2019.7.1.525

[24] Hassan, B. A., Ibrahim, A. L., & Mehamdia, A. E. (2025). Improved Dai–Liao Conjugate Gradient Methods for Large-Scale Unconstrained Optimization. AL-Rafidain Journal of Computer Science and Mathematics, 19(1), 178–186. https://doi.org/10.33899/csmj.2025.159774.1186

[25] Hassan, B. A., & Alashoor, H. A. (2023). On image restoration problems using new conjugate gradient methods. Indonesian Journal of Electrical Engineering and Computer Science, 29(3), 1438–1445.

[26] Ibrahim, A. L., Sadiq, M. A., & Shareef, S. G. (2019). A new conjugate gradient coefficient for unconstrained optimization based on Dai–Liao. Science Journal of University of Zakho, 7(1), 34–36.

[27] Barzilai, J., & Borwein, J. M. (1988). Two-point step size gradient methods. IMA Journal of Numerical Analysis, 8(1), 141–148.

[28] Raydan, M. (1997). The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. SIAM Journal on Optimization, 7, 26–33.

[29] Birgin, E. G., & Martínez, J. M. (2001). A spectral conjugate gradient method for unconstrained optimization. Applied Mathematics and Optimization, 43, 117–128.4–60, 2014.

[30] N. Andrei. Scaled Conjugate Gradient Algorithms for Unconstrained Optimization. Computational Optimization and Applications, 38(3), 401-416, 2007.

[31] H. Jiang, S. Deng, X. Zheng, and Z. Wan. Global Convergence of a Modified Spectral Conjugate Gradient Method. Journal of Applied Mathematics, 2012.

[32] L. Zhang, W. Zhou, and D. Li. A Descent Modified Polak–Ribière–Polyak Conjugate Gradient Method and Its Global Convergence. IMA Journal of Numerical Analysis, 26(4), 629-640, 2006.

[33] J. Liu, and Y. Jiang. Global Convergence of a Spectral Conjugate Gradient Method for Unconstrained Optimization, Hindawi Publishing Corporation, Abstract and Applied Analysis, Vol.12, Article ID 758287, 2012.

[34] J. K. Liu, X. Du, and K. Wang. A Mixed Spectral CD DY Conjugate Gradient Method. J. Applied Mathematics, 569795-1, 2012.

[35] Andrei, N. (2010). New accelerated conjugate gradient algorithms as a modification of Dai–Yuan's computational scheme for unconstrained optimization. Journal of Computational and Applied Mathematics, 234, 3397–3410.

[36] Andrei, N. (2007). A scaled BFGS preconditioned conjugate gradient algorithm for unconstrained optimization. Applied Mathematics Letters, 20, 645–650.

[37] Andrei, N. (2007). Scaled memoryless BFGS preconditioned conjugate gradient algorithm for unconstrained optimization. Optimization Methods and Software, 22, 561–571.

[38] Powell, M. J. D. (1978). Algorithms for nonlinear constraints that use Lagrangian functions. Mathematical Programming, 14, 224–248. https://doi.org/10.1007/BF01588967

[39] Dai, Y., Han, J., Liu, G., Sun, D., Yin, H., & Yuan, Y.-X. (1999). Convergence properties of nonlinear conjugate gradient methods. SIAM Journal on Optimization, 10, 345–358. https://doi.org/10.1137/S1052623494268443

[40] Andrei, N. (2008). An unconstrained optimization test functions collection. Advances in Modeling and Optimization, 10(1), 147–161.

[41] Dolan, E. D., & Moré, J. J. (2002). Benchmarking optimization software with performance profiles. Mathematical Programming, 91, 201.

Appendix: Benchmark Test Functions

This appendix presents a selected set of classical test functions unconstrained optimization problems used to evaluate the numerical performance of the optimization method. Each function has distinct structural and numerical challenges, making it a esay way to test for assessing convergence behavior, robustness, and scalability.

1. Wolfe: $x_0 = (-1, \dots, -1)^T$,

$$f(x) = \left(-x_1\left(3 - \frac{x_1}{2}\right) + 2x_2 - 1\right)^2 + \sum_{i=1}^{n-1}\left(x_{k-1} - x_k\left(3 - \frac{x_k}{2} + 2x_{k+1} - 1\right)\right)^2 + \left(x_{n-1} - x_n\left(3 - \frac{x_n}{2}\right) - 1\right)^2.$$

2. Central: $x_0 = (1,2,2,2, \dots, 1,2,2,2)^T$,

$$f(x) = \sum_{k=1}^{n/4}(exp(x_{4k-3} + x_{4k-2})^4 + 100((x_{4k2} - x_{4k-1})^6 + arctan((x_{4k2} - x_{4k-1})^4 + x_{4k-3}).$$

3. Non-Diagonal: $x_0(-1, \dots, -1)^T$

$$f(x) = \sum_{k=2}^{n}(100(x_1 - x_k^2)^2 + (1 - x_k)^2).$$

4. Powell: $x_0 = (3, -1,0,1, \dots, 3, -1,0,1)^T$,

$$f(x) = \sum_{k=1}^{n/4}((x_{4k-3} - 10x_{4k-2})^2 + 5(x_{4k-1} - x_{4i})^2 + (x_{4k-2} - 2x_{4k-1})^4 + 10(x_{4k-3} - x_{4k})^4).$$

5. Rosen: $x_0 = (-1.2,1, \dots, -1.2,1)^T$,

$$f(x) = \sum_{k=1}^{n/2}(100(x_{2k} - x_{2k-1}^2)^2 + (1 - x_{2k-1})^2).$$

6. Mile: $x_0 = (1,2,2, \dots, 1,2,2)^T$,

$$f(x) = \sum_{k=1}^{n/4}((e^{x_{4k-3}} + 10x_{4k-2})^2 + 100(x_{4k-2} + x_{4k-1})^6 + (\tan(x_{4k-1} - x_{4k}))^4 + (x_{4k-3})^8 + (x_{4k} - 1)^2).$$

7. Wood: $x_0 = (-3, -1, \dots, -3, -1)^T$

$$f(x) = \sum_{k=1}^{n/4}(100(x_{4k-3}^2 - x_{4i-2})^2 + (x_{4k-3} - 1)^2 + 90(x_{4k-1}^2 - x_{4k})^2 + (1 - x_{4k-1})^2 + 10.1(x_{4k-2} - 1)^2 + (x_{4k} - 1)^2 + 19.8(x_{4k-2} - 1)(x_{4k} - 1)).$$

8. Sum: $x_0 = (1,1, \dots, 1)^T$,

$$f(x) = \sum_{k=1}^{n}(x_k - k)^4.$$

9. Edger: $x_0 = (1,0, \dots, 1,0)^T$,

$$f(x) = \sum_{k=1}^{n/2}((x_{2k-1} - 2)^2 + (x_{2k-1} - 2)^2 x_{2k}^2 + (x_{2k} + 1)^2).$$

10. Shallow: $x_0 = (-2, -2, \dots, -2, -2)^T$,

$$f(x) = \sum_{k=1}^{n/2}([x_{2k-1}^2 - x_{2k}]^2 + (1 - x_{2k-1})^2).$$

11. Cubic: $x_0 = (-1.2,1, \dots, -1.2,1)^T$,

$$f(x) = \sum_{k=1}^{n/2}(100(x_{2k} - x_{2k-1}^3)^2 + (1 - x_{2k})^2).$$

12. Beal: $x_0 = (0,0)^T$,

$$f(x) = ((1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2).$$

13. OSP: $x_0 = (1, \dots, 1)^T$,

$$f(x) = \left(\sum_{k=1}^{n} k(x_k)^2\right)^2.$$