A Modified Conjugate Descent Projection Method for Monotone Nonlinear Equations and Image Restoration

SANI AJI, POOM KUMAM, AUWAL BALA ABUBAKAR, AND MAHMOUD MUHAMMAD YAHAYA

1. INTRODUCTION

This article aims at finding a solution to the nonlinear equation

\[ H(x) = 0, \quad \text{subject to} \quad x \in \Omega \subseteq \mathbb{R}^n, \]

(1)

where \( \Omega \) is nonempty closed and convex, \( H : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is continuous and monotone. Monotone means for all \( x, y \in \mathbb{R}^n \)

\[ \langle H(x) - H(y), x - y \rangle \geq 0. \]

Solving nonlinear systems of monotone equations is crucial due to their occurrence in many applications. They appear as subproblems in proximal algorithms \cite{1}, and also in monotone variational inequality problems \cite{2}, \cite{3}. In addition, monotone nonlinear equations are applicable in \( \ell_1 \) norm regularized optimization problems to solve image restoration problems \cite{4}–\cite{6}. Interested readers can refer to \cite{7}–\cite{9} for further reading on image restoration. Recently, an application of nonlinear equation in financial forecasting is also considered in \cite{10}.

For solving systems of nonlinear equations, a lot of methods have been developed. The prominent and the early method, is the Newton method \cite{11} followed by quasi-Newton methods \cite{12}–\cite{17}. Due to the good numerical performance and the quadratic convergence of the Newton’s method, it was termed as a very good and efficient method for solving nonlinear equations when the operator under consideration is continuously differentiable. However, in this article, equation (1) is Lipschitz and monotone. Hence some assumptions
for which Newton method can be applied are not satisfied. Moreover, Newton method has other shortcomings including computation of Jacobian matrix and solving linear system at each iteration, and when the dimension is very large, the method fails in this aspect.

Linear conjugate gradient methods were first proposed independently by Hestenes and Stiefel to solve linear systems which was the foundation of many studies of conjugate gradient methods. More focus were later given to the study of conjugate gradient methods after the joint work of Hestenes and Stiefel. Conjugate gradient methods and their modified versions [18]–[24] were later extended for solving both unconstrained optimization problems and nonlinear monotone equations. The conjugate gradient methods proposed for solving unconstrained optimization problems can be categorized into two groups; the Fletcher and Reeves (FR) [25], Dai-Yuan (DY) [26] and the CD [27] having the same numerator in the conjugate parameter, while the second category includes the Hestenes and Stiefel (HS) [28], Polak–Ribière–Polyak (PRP) [29] and Liu-Storey method (LS) [30]. The FR, DY, and the CD algorithms are known for their good convergence behavior as opposed to methods in the second category with good numerical performance [31].

On the other hand, Cheng [32] proposed a conjugate gradient algorithm for solving monotone equations. His method was based on the combination of the famous PRP conjugate gradient method [29], [33] and the hyperplane projection method. Xiao and Zhu [6] used the projection technique [34] and the CG parameter of Hager and Zhang [35], to developed a projected conjugate gradient method for solving nonlinear monotone equations. Liu et al. [36] combined two sufficient descent conjugate gradient methods with a hyperplane projection technique to solve nonlinear convex constrained monotone equations. Hu and Wei [37] extended Wei-Yao-Liu conjugate gradient method for solving unconstrained optimization problem to solve nonlinear monotone equations. Liu and Li [38] proposed a CD-like projection method to solve convex constrained nonlinear monotone equations. Papp and Rapajić [39] combine conjugate gradient method and hyperplane projection technique and proposed a new Fletcher and Reeves [25] type directions in the frame of algorithm for solving nonlinear monotone equations. Based on the work of [32] and [39]. Abubakar and Kumam [40] extended the method in [41] and presented an efficient conjugate gradient method for solving nonlinear monotone equation. They achieved this by combining the conjugate gradient method in [41] with a hyperplane projection method. Moreover, Dai and Zhu [42] combined the hyperplane projection method in [34] with the modified HS method in [43] and presented an algorithm for solving monotone nonlinear equations. They proved the global convergence of their method under some assumptions and showed its efficiency numerically in comparison with other methods.

Recently, Yuan et al. [44] proposed a modified LS-like conjugate gradient method for solving nonlinear equations where the function under consideration is continuously differentiable. Their method uses a convex combination of the steepest descent and a modified version of the LS method [30]. This method satisfies both sufficient descent and trust region properties. They proved the global convergence of the method and showed its effectiveness in solving nonlinear equations. However, the method imposed some restrictions in step 3 and 5 of the algorithm.

The motivation of this work is that instead of considering nonlinear equations where the operator under consideration is continuously differentiable as in [44], we are focusing on solving nonlinear equations, where the operator is Lipschitz continuous and monotone. Also, the restrictions in step 3 and 5 of the algorithm in [44] were not considered in our case. In addition, to our knowledge, there are limited CD-like conjugate gradient algorithms for solving nonlinear monotone equations. Therefore, we aim to test how effective a CD-like conjugate gradient algorithm would be in solving monotone nonlinear equations. To achieve this, we consider a suitable line search, and replace the LS parameter in [44] with the CD parameter. Moreover, We proved the convergence of the proposed algorithm, show its strength numerically, and also, its efficiency in restoring blurred images.

The organization of the paper is as follows. We present the proposed algorithm together with its convergence in the second Section. In the third Section, we conduct numerical experiments to depict the strength of the proposed method for finding solutions to (1) and image restoration problems. Finally, we concluded the paper in Section four.

II. ALGORITHM: MOTIVATION AND CONVERGENCE RESULT

In this section, we give the definition of projection map, motivation, and the description of the proposed algorithm.

Definition 1: Let $\Omega \subset \mathbb{R}^n$ be a nonempty, closed and convex set. The projection of any $x \in \mathbb{R}^n$ onto $\Omega$ is

$$P_\Omega(x) = \arg\min\{\|x - y\| : y \in \Omega\}.$$  

It is worth noting that $P_\Omega$ satisfy the inequality

$$\|P_\Omega(x) - P_\Omega(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^n. \quad (2)$$

Conjugate gradient methods for solving (1) generally uses the iterative formula

$$x_{k+1} = x_k + \alpha_k d_k, \quad (3)$$

where $\alpha_k$ represents the step-length along the direction $d_k$ and can be obtained from some line search. $x_k$ and $x_{k+1}$ represent the current and next points of iteration, respectively, and $d_k$ is a search direction defined as

$$d_k = \begin{cases} -H_k, & \text{if } k = 0, \\ -H_k + \beta_k d_{k-1}, & \text{if } k \geq 1, \end{cases} \quad (4)$$

where $\beta_k$ is called the conjugate gradient parameter and $H_k := H(x_k)$.

The search direction usually satisfies the inequality

$$H_k^T d_k \leq -c\|H_k\|^2, \quad (5)$$

where $c > 0$. 

\[VOLUME 8, 2020\] 158657
Different conjugate gradient methods for solving (1) are developed using different search directions. In particular, Yuan et al. [44] proposed a new search direction \( d_k \) by taking a convex combination of the steepest descent with a modified LS method. The direction is given by

\[
d_k = \begin{cases} 
-H_k, & \text{if } k = 0, \\
-L_k H_k & \\
+(1 - L_k) \frac{H_k^T y_{k-1} d_{k-1} - d_{k-1}^T H_k y_{k-1}}{\max\{2\chi \|d_{k-1}\| \|y_{k-1}\|, \|d_{k-1}\| \|H_k\|, \|H_k d_{k-1}\|\}} & \text{if } k \geq 1,
\end{cases}
\]  

(6)

where \( \chi \in (0, 1) \),

\[
L_k = \frac{y_k^T y_k}{y_k^T s_k}
\]  

(7)

and \( \hat{s}_k \) from Li and Fukushima [45], [46] is given as

\[
\hat{s}_k = s_k + (\max\{0, -\frac{y_k^T y_k}{\|s_k\|^2} + 1\}) y_k, \quad s_k = x_{k+1} - x_k
\]

and \( y_k = H_{k+1} - H_k \).

The term \( L_k \) can be found in the work of Birgin and Martínez [47] and has the following property

\[
y_k^T \hat{s}_k \geq \max\{y_k^T s_k, \|y_k\|^2\} \geq \|y_k\|^2 > 0,
\]

which implies \( L_k \in (0, 1) \).

Inspired by the direction (6), we proposed a new search direction corresponding to a new conjugate gradient algorithm for solving (1). Our approach is based on replacing the LS parameter in (6) with the CD parameter, hence we came up with the following search direction

\[
d_k = \begin{cases} 
-H_k, & \text{if } k = 0, \\
-L_k H_k & \\
+(1 - L_k) \frac{H_k^T y_{k-1} d_{k-1} - d_{k-1}^T H_k y_{k-1}}{\max\{2\chi \|d_{k-1}\| \|y_{k-1}\|, \|d_{k-1}\| \|H_k\|, \|H_k d_{k-1}\|\}} & \text{if } k \geq 1,
\end{cases}
\]  

(8)

where \( L_k \) is given as in (7). Now we proceed with the steps of our proposed algorithm for solving (1) as follows.

In order to show the global convergence of the MCDPM algorithm, we need the following assumptions and Lemmas:

\( (G_1) \) The mapping \( H \) is monotone, that is,

\[
(H(x) - H(y))^T (x - y) \geq 0, \quad \forall x, y \in \mathbb{R}^n.
\]

\( (G_2) \) The mapping \( H \) is Lipschitz continuous, that is there exists \( L > 0 \) such that

\[
\|H(x) - H(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n.
\]

\( (G_3) \) The solution set of (1), denoted by \( \hat{\Omega} \), is nonempty.

**Lemma 1:** Let \( \{d_k\} \) be given by (8). Then there exists \( \alpha > 0 \) such that \( \{d_k\} \) holds.

**Proof:** From (8), if \( k = 0, d_0 = -H_0 \) and so we have\( \|H_0 d_0\|^2 = \|H_0\|^2 \). When \( k \neq 0, \)

\[
H_k^T d_k = -L_k \|H_k\|^2
\]

and

\[
+(1 - L_k) \frac{H_k^T y_{k-1} d_{k-1} - d_{k-1}^T H_k y_{k-1}}{\max\{2\chi \|d_{k-1}\| \|y_{k-1}\|, \|d_{k-1}\| \|H_k\|, \|H_k d_{k-1}\|\}}
\]

(8)

The desired result is obtained after solving for \( \alpha \).

**Lemma 3:** Suppose that assumptions \( (G_1)-(G_3) \) hold, then the sequences \( \{x_k\} \) and \( \{z_k\} \) generated by Algorithm 1 are bounded. Moreover, we have

\[
\lim_{k \to \infty} \|x_k - z_k\| = \lim_{k \to \infty} \alpha_k \|d_k\| = 0,
\]  

(12)

and

\[
\lim_{k \to \infty} \|x_{k+1} - x_k\| = 0.
\]  

(13)

**Proof:** To prove \( \{x_k\} \) and \( \{z_k\} \) are bounded, from the non expansive property (2), for any \( \bar{x} \in \hat{\Omega} \) and taking \( x = x_{k+1}, y = \bar{x} \), we have

\[
\|x_{k+1} - \bar{x}\|^2 = \|P_\Omega [x_k - \gamma_\bar{z}_k H(z_{k})] - \bar{x}\|^2
\]

Therefore,

\[
H_k^T d_k = -L_k \|H_k\|^2, \quad L_k \in (0, 1).
\]  

(10)

**Algorithm 1 (MCDPM)**

**Step 0.** Choose initial point \( x_0 \in \Omega, \gamma \in (0, 2), \beta_0 > 0 \), \( \beta, \chi \in (0, 1) \) and Tol > 0. Set \( k := 0 \).

**Step 1.** If \( \|H_k\| \leq \text{Tol} \), stop, otherwise proceed with **Step 2**.

**Step 2.** Compute \( d_k \) from (8).

**Step 3.** Compute \( a_k = \max\{\beta^i : i = 0, 1, 2, \ldots\} \) such that

\[
-H[x_k + \beta^i d_k]^T d_k \geq \alpha beta^i \|d_k\|^2.
\]  

(9)

**Step 4.** Set \( z_k = x_k + \alpha_k d_k \). If \( z_k \in \Omega \), set \( z_k = x_{k+1} \). Else compute

\[
x_{k+1} = P_\Omega [x_k - \gamma z_k H(z_k)],
\]

where

\[
\gamma z_k = H(z_k)^T (x_k - z_k), \|H(z_k)\|^2.
\]

**Step 5.** Let \( k = k + 1 \) and go to **Step 1**.
\[
\begin{align*}
\|x_k - y_c \chi H(z_k) - \bar{x}\|^2 &\leq \frac{\|x_k - \bar{x}\|^2}{2} - 2 \gamma x_k H(z_k) + y^2 \chi^2 \|H(z_k)\|^2 \\
&= \|x_k - \bar{x}\|^2 - 2 \gamma H(z_k)^T (x_k - z_k) + y^2 \chi^2 \|H(z_k)\|^2 \\
&\leq \|x_k - \bar{x}\|^2 - 2 \gamma H(z_k)^T (x_k - z_k) + y^2 \chi^2 \|H(z_k)\|^2 \\
&\leq \|x_k - \bar{x}\|^2 - 2 \gamma (2 - \gamma \chi^2) \|H(z_k)\|^2 \\
&\leq \|x_k - \bar{x}\|^2 - 2 \gamma (2 - \gamma \chi^2) \|H(z_k)\|^2 \\
&\leq \|x_k - \bar{x}\|^2 - 2 \gamma (2 - \gamma \chi^2) \|H(z_k)\|^2 \\
&\leq \|x_k - \bar{x}\|^2 - 2 \gamma (2 - \gamma \chi^2) \|H(z_k)\|^2.
\end{align*}
\]

This shows that the sequence \(\{\|x_k - \bar{x}\|\}\) is a decreasing sequence, and hence \(\{x_k\}\) is bounded. In addition, combining this with assumption \((G_2)\), we can find \(c_1 \geq 0\) such that \(\|H_k\| \leq c_1\). Since \((H_k - H(z_k))^T (x_k - z_k) \geq 0\), using Cauchy-Schwarz inequality and (9), we obtain
\[
\|H_k\| \|x_k - z_k\| \geq \|H(z_k\|^2 (x_k - z_k) \geq -\alpha \|H(z_k\|^2 d_k \geq \sigma \alpha \|d_k\|^2 = \sigma \|x_k - z_k\|^2.
\]

Therefore,
\[
\begin{align*}
\sigma \|x_k - z_k\| &\leq \|H_k\| \leq c_1,
\end{align*}
\]
showing that \(\{z_k\}\) is bounded.

Again, using assumption \((G_2)\), we can find another constant \(c_2 \geq 0\) such that \(\|H(z_k)\| \leq c_2\) for all \(k \geq 0\). This together with (14) give us
\[
\begin{align*}
\gamma (2 - \gamma \chi^2) \|x_k - z_k\|^2 &\leq \|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2, \quad (15)
\end{align*}
\]
adding (15) for \(k = 0, 1, 2, \ldots\), we have
\[
\begin{align*}
\gamma (2 - \gamma \chi^2) \sum_{k=0}^\infty \|x_k - z_k\|^2 &\leq \sum_{k=0}^\infty \|x_k - \bar{x}\|^2 - \|x_{k+1} - \bar{x}\|^2,  \\
\sum_{k=0}^\infty \|x_k - \bar{x}\|^2 &\leq \sum_{k=0}^\infty \|x_{k+1} - \bar{x}\|^2,
\end{align*}
\]
which implies
\[
\lim_{k \to \infty} \|x_k - z_k\| = 0.
\]

**Lemma 4:** Let \(\{d_k\}\) be given by (8), then there is an \(M > 0\) for which \(\|d_k\| \leq M\).

**Proof:** From equation (8), if \(k = 0\), then \(\|d_0\| = \|H_0\| \leq c_1\). If \(k \neq 0\) and since
\[
\max\{2\chi \|d_{k-1}\| \|H_k\|, -d_{k-1}^T H_{k-1}\} \geq 2\chi \|d_{k-1}\| \|H_k\|,
\]
then we have
\[
\|d_k\| = \| -L_k H_k \| + (1 - L_k) \frac{\|H_k\|^2 \|d_{k-1}\| + \|H_k\|^2 \|d_{k-1}\|}{\max\{2\chi \|d_{k-1}\| \|H_k\|, -d_{k-1}^T H_{k-1}\}}.
\]
Applying triangular inequality, we have
\[
\|d_k\| \leq L_k \|H_k\| \|d_{k-1}\| + (1 - L_k) \frac{\|H_k\|^2 \|d_{k-1}\| + \|H_k\|^2 \|d_{k-1}\|}{\max\{2\chi \|d_{k-1}\| \|H_k\|, -d_{k-1}^T H_{k-1}\}}.
\]
By applying Cauchy-Schwarz on \(\|H_k^T d_{k-1}\|\) we get
\[
\|d_k\| \leq L_k \|H_k\| \|d_{k-1}\| + (1 - L_k) \frac{\|H_k\|^2 \|d_{k-1}\| + \|H_k\|^2 \|d_{k-1}\|}{\max\{2\chi \|d_{k-1}\| \|H_k\|, -d_{k-1}^T H_{k-1}\}}.
\]
Letting \(\omega = \min\{L_k + \frac{(1-L_k)}{\chi}, 1\}\) and \(M = \omega c_1\), then
\[
\|d_k\| \leq M. \quad (18)
\]
From (17), we arrived at (18) which show that the direction is bounded. The next theorem is proved by contradiction, and the boundedness of the direction is required to establish the contradiction in (22).

**Theorem 1:** Suppose that assumptions \((G_1)-(G_3)\) hold and let the sequence \(\{x_k\}\) be generated by Algorithm 1, then
\[
\liminf_{k \to \infty} \|H_k\| = 0. \quad (19)
\]
**Proof:** Suppose by contradiction the relation (19) is not satisfied, then there exist a positive constant \(r_1\) such that \(\forall k \geq 0\),
\[
\|H_k\| \geq r_1. \quad (20)
\]
By the relations (10) and (20), we have that \(\forall k \geq 0\),
\[
\|d_k\| \geq r_1. \quad (21)
\]
Multiplying both sides of (11) with \(\|d_k\|\) together with (18) and (20), we get
\[
\alpha_k \|d_k\| \geq \max\left\{ \frac{1}{L_1}, \frac{L_1 \beta \|H_k\|}{(L + \sigma) \|d_k\|^2} \right\} \|d_k\| \geq \max\left\{ r_1, \frac{L_1 \beta \|H_k\|^2}{(L + \sigma) M} \right\}. \quad (22)
\]
The above inequality contradicts with (12). Therefore (19) must hold. □
III. NUMERICAL EXPERIMENTS

In this section, we present some numerical experiments of our algorithm (MCDPM), and compare its performance with some existing algorithms. Using some test problems, we compare the performance of MCDPM with three other algorithms. Specifically, Algorithm 2.1 proposed by Zheng et al. [48], Self-adaptive three-term conjugate gradient method (SATCGM) algorithm by Wang et al. [49] and the NLS algorithm proposed in [44]. In MCDPM, Algorithm 2.1 and SATCGM, Solodov and Svaiter line search [34] is used whereas the NLS algorithm considered a different line search. It is worth noting that all codes are written on Matlab R2019b and executed using a PC of corei3-4005U processor, 4GB RAM and 1.70GHz CPU. We considered seven problems with eight different initial points. All the problems are tested on five different dimensions, starting from \( n = 1000, n = 5000, n = 10, 000, n = 50, 000 \) and \( n = 100, 000 \). We set \( \|H_k\| < 10^{-5} \) as the stopping criteria.

In NLS, Algorithm 2.1 and SATCGM all parameters are maintained as they are in [44], [48], [49] respectively. As for MCDPM, \( \sigma = 0.0001 \) is used whereas for \( \beta \), different values were tested in the interval \((0, 1)\), but \( \beta = 0.6 \) was observed to give the best result. Finally, \( \gamma \) and \( \chi \) are taken as 1.8 and 0.1 respectively.

Shown in the following tables are the results obtained, comparing MCDPM with Algorithm 2.1 and SATCGM. The tables of comparison between MCDPM and NLS can be found in https://documentcloud.adobe.com/link/review?uri=urn:aaid:scds:US:ec7a57a7-4904-433a-911a-e9c0080b8a88#pageNumber=1. We denote ITER to represent the number of iterations, FVAL as the number of function evaluations, TIME as the CPU time measured in seconds, and Norm as the value of norm of the function when the solution is obtained. We used “-“ to indicate when the iteration is bigger than 1000 and the stopping criteria is not satisfied.

Below are the list of the test problems considered in this work, where \( H(x) = (h_1(x), h_2(x), \ldots, h_n(x))^T \).

**Problem 1 ([50]):** Exponential Function.

\[
\begin{align*}
h_1(x) &= e^{x_1} - 1, \\
h_i(x) &= e^{x_i} + x_i - 1, \quad \text{for } i = 2, 3, \ldots, n, \\
\Omega &= \mathbb{R}^n_+.
\end{align*}
\]

**Problem 2 ([50]):** Modified Logarithmic Function.

\[
\begin{align*}
h_i(x) &= \ln(x_i + 1) - \frac{x_i}{n}, \quad \text{for } i = 2, 3, \ldots, n, \\
\Omega &= \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i \leq n, x_i \geq -1, i = 1, 2, \ldots, n\}.
\end{align*}
\]

**Problem 3 ([51]):**

\[
\begin{align*}
h_i(x) &= \min \left( \min(x_i, x_i^2), \max(x_i, x_i^3) \right) \\
\Omega &= \mathbb{R}^n_+.
\end{align*}
\]

**Problem 4 ([50]):** Strictly Convex Function I.

\[
\begin{align*}
h_i(x) &= e^{x_i} - 1, \quad \text{for } i = 1, 2, \ldots, n, \quad \Omega = \mathbb{R}^n_+.
\end{align*}
\]

The obtained results clearly show that MCDPM outperforms Algorithm 2.1, SATCGM and NLS by solving a large proportion of the problems considered with less number of iterations and CPU time. In addition, using the well known Dolan and Moré performance profile [54], we also show graphically (see Figure 1, 2 and 3) the performance of the algorithms in terms of the number of iterations, CPU time and number of function evaluations respectively. In Figure 1 and 2, MCDPM solves 60\% and 55\% of the problems with the least number of iterations and CPU time respectively. In Figure 3, Algorithm 2.1 solves about 46\% of the problems with lesser function evaluations as compared with MCDPM and SATCGM with about 38\% and 20\%, respectively. The NLS algorithm performed poorly in this experiment. This may be due to the fact that it was totally in a different setting with the other algorithms. It is specifically proposed for solving unconstrained nonlinear equations with entirely different line search and assumptions on the function considered. In general, MCDPM proved to be more efficient in solving systems of nonlinear monotone equations. This may be as a result of the direction being a convex combination of a modified CD-like method with the steepest descent.
A. APPLICATIONS IN IMAGE RESTORATION

Solving ill-conditioned linear systems of equations has been a great tool used in different branches of science including signal processing. One of the popular ways involves solving the following problem

\[
\min \frac{1}{2} \| t - Br \|_2^2 + \rho \| r \|_1, \tag{23}
\]

where \( r \in \mathbb{R}^n, t \in \mathbb{R}^m, B \in \mathbb{R}^{m \times n} (m \ll n) \) is a linear operator, \( \rho \geq 0 \), \( \| r \|_2 \) is the Euclidean norm of \( r \) and \( \| r \|_1 = \sum_{i=1}^n |r_i| \) is the \( \ell_1 \)-norm of \( r \). Clearly, problem (23) is a convex
There are several iterative methods to solve (23) (see for example [4], [55]–[59]). The earliest among them was proposed by Figueiredo et al. [4] which is the most popular gradient based projection method used for sparse reconstruction. In this method, (23) is expressed as a quadratic problem by expressing \( r \in \mathbb{R}^n \) into positive and negative parts as follows

\[
    r = x - y, \quad x \geq 0, \quad y \geq 0,
\]

where \( x_i = (r_i)_+ \), \( y_i = (-r_i)_+ \) for all \( i = 1, 2, \ldots, n \), and \( (\cdot)_+ = \max\{0, \cdot\} \). Also, we have \( \|r\|_1 = e^T_n x + e^T_n y \), where \( e_n = (1, 1, \ldots, 1)^T \in \mathbb{R}^n \). So, the problem (23) can now be written as

\[
    \min_{x, y} \frac{1}{2} \|t - B(x - y)\|_2^2 + \rho e^T_n x + \rho e^T_n y, \quad x \geq 0, \quad y \geq 0,
\]

(24)

From [4], equation (24) can be written as

\[
    \min_{z} \frac{1}{2} z^T E z + c^T z, \quad \text{such that} \quad z \geq 0,
\]

(25)

where \( z = \begin{pmatrix} x \\ y \end{pmatrix} \), \( c = \omega e_{2n} + \begin{pmatrix} -b \\ b \end{pmatrix} \), \( b = B^T t \),

\[
    E = \begin{pmatrix} B^T B & -B^T B \\ -B B^T & B^T B \end{pmatrix}.
\]

Problem (25) is a quadratic programming problem since \( E \) is a positive semi-definite.

Later on, (25) was translated to a linear variable problem, equivalently, a linear complementary problem by Xiao and Zhu [6]. Moreover, the variable \( z \) solves the linear complementary problem provided that it solves the nonlinear equation:

\[
    H(z) = \min\{z, E z + c\} = 0.
\]

(26)

where \( H \) is a vector-valued function. The function \( H(z) \) is shown to be continuous and monotone in [5], [60]. This translated problem (23) into problem (1). As a result, the algorithm proposed in our work can be applied to solve it.

Here, we use Lena, Barbara and Aeroplane test images to test the performance of our algorithm in comparison with CGD algorithm [6] to restore blurred images. Each of the two algorithms are tested using metrics that show better restoration. Specifically, we reported the performance of the two algorithms based on signal-to-noise-ratio (SNR) given as

\[
    \text{SNR} = 20 \times \log_{10}\left(\frac{\|\bar{x}\|}{\|x - \bar{x}\|}\right),
\]

For example, the following tables show the numerical results for Algorithm 1 (MCDPM), Algorithm 2.1 and SATCGM for Problem 5, Problem 6, Problem 7, and the efficiency comparison for Algorithm 1 and the CGD Algorithm based on SNR, PSNR and SSIM.
the structural similarity (SSIM) index, which quantify how similar the restored is to the original [61] and the peak signal-to-noise-ratio (PSNR). The Figures 4-6 show the efficiency of each of the algorithms by comparing the original, blurred and the restored images. It is evident from the figures that both our algorithm (MCDPM) and the CGD can restored the blurred images. However, it is clear that our algorithm outperforms the CGD algorithm based on the metrics considered.

**IV. CONCLUSION**

This work proposed a modified conjugate descent projection algorithm for solving system of nonlinear monotone equations. The algorithm is based on modifying the direction considered in [44] by replacing the modified LS-like parameter with a modified CD-like parameter. The efficiency and suitability of the proposed algorithm for large scale problems is as a result of its low storage and lack of Jacobian matrix computation. Numerical results obtained from the problems considered proved that this algorithm is more efficient in comparison with some existing algorithms, specifically, those in [44], [48] and [49]. Moreover, the global convergence properties of the algorithm are proved under some appropriate assumptions. An application of the algorithm in image restoration problems is also considered. It is observed that the proposed algorithm proves to be more efficient in restoring blurred images with higher quality than the CGD algorithm.

**REFERENCES**

[1] N. A. Iusem and V. M. Solodov, “Newton-type methods with generalized distances for constrained optimization,” *Optimization*, vol. 41, no. 3, pp. 257–278, Jan. 1997.

[2] M. Fukushima, “Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems.” *Math. Program.*, vol. 53, nos. 1–3, pp. 99–110, Jan. 1992.

[3] Y.-B. Zhao and D. Li, “Monotonicity of fixed point and normal mappings associated with variational inequality and its application,” *SIAM J. Optim.*, vol. 11, no. 4, pp. 962–973, Jan. 2001.

[4] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, “Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems,” *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.

[5] Y. Xiao, Q. Wang, and Q. Hu, “Non-smooth equations based method for \( \ell_1 \)-norm problems with applications to compressed sensing,” *Nonlinear Anal., Theory, Methods Appl.*, vol. 74, no. 11, pp. 3570–3577, Jul. 2011.

[6] Y. Xiao and H. Zhu, “A conjugate gradient method to solve convex constrained monotone equations with applications in compressive sensing,” *J. Math. Anal. Appl.*, vol. 405, no. 1, pp. 310–319, Sep. 2013.

[7] S. Q. Liu, S. H. Hu, Y. Xiao, and Y. L. An, “Bayesian shearlet shrinkage for SAR image de-noising via sparse representation,” *Multidimensional Syst. Signal Process.*, vol. 25, no. 4, pp. 683–701, Oct. 2014.

[8] S. Liu, M. Liu, P. Li, J. Zhao, Z. Zhu, and X. Wang, “SAR image denoising via sparse representation in shearlet domain based on continuous cycle spinning,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2985–2992, May 2017.

[9] S. Liu, Q. Hu, P. Li, J. Zhao, M. Liu, and Z. Zhu, “Speckle suppression based on weighted nuclear norm minimization and grey theory,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2700–2708, May 2019.

[10] Z. Dai, X. Dong, J. Kang, and L. Hong, “Forecasting stock market returns: New technical indicators and two-step economic constraint method,” *North Amer. J. Econ. Finance*, vol. 53, Jul. 2020, Art. no. 101216.

[11] C. T. Kelley, *Solving Nonlinear Equations With Newton’s Method*, vol. 1, Philadelphia, PA, USA: SIAM, 2003.

[12] G.-Z. Gu, D.-H. Li, L. Qi, and S.-Z. Zhou, “Descent directions of quasi-Newton methods for symmetric nonlinear equations,” *SIAM J. Numer. Anal.*, vol. 40, no. 5, pp. 1763–1774, Jan. 2002.

[13] L. Huang, H. Ji, and Y. Zhu, “Analysis of effective working characteristic of the proportional solenoid,” in *Proc. Int. Conf. Green Energy Appl. (ICGEA)*, Mar. 2017, pp. 35–38.
[14] Y. Li, G. Yuan, and Z. Wei, “A limited-memory BFGS algorithm based on a trust-region quadratic model for large-scale nonlinear equations,” *PLoS ONE*, vol. 10, no. 5, May 2015, Art. no. e0120993.

[15] Y.-Z. Luo, G.-J. Tang, and L.-N. Zhou, “Hybrid approach for solving systems of nonlinear equations using chaos optimization and quasi-Newton methods,” *Appl. Soft Comput.*, vol. 8, pp. 1068–1073, Mar. 2008.

[16] Y. Yueting, Z. Qi, L. Yunlong, and J. Xiaowei, “An inexact quasi-Newton algorithm combined with Jacobian restart technique for nonlinear equation systems,” in *Proc. 5th Int. Joint Conf. Comput. Sci. Optim.*, Jun. 2012, pp. 230–233.

[17] D. Zhu, “Nonmonotone backtracking inexact quasi-Newton algorithms for solving smooth nonlinear equations,” *Appl. Math. Comput.*, vol. 161, no. 3, pp. 875–895, Feb. 2005.

[18] P. Deufhard, “A modified Newton method for the solution of ill-conditioned systems of nonlinear equations with application to multiple shooting,” *Numerische Math.*, vol. 22, no. 4, pp. 289–315, Aug. 1974.

[19] F. Zhang and C. S. Cheng, “A modified Newton method for radial distribution system power flow analysis,” *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 389–397, Feb. 1997.

[20] H. O. Hartley, “The modified Gauss-Newton method for the fitting of nonlinear regression functions by least squares,” *Technometrics*, vol. 3, no. 2, pp. 269–280, May 1961.

[21] X. Li, X. Wang, and X. Duan, “A limited memory BFGS method for solving large-scale symmetric nonlinear equations,” *Abstract Appl. Anal.*, vol. 2014, pp. 1–8, Aug. 2014.

[22] X. Li, X. Wang, L. Cheng, and X. Duan, “A modified conjugate gradient algorithm with backtracking line search technique for large-scale nonlinear equations,” *Int. J. Comput. Math.*, vol. 95, no. 2, pp. 382–395, Feb. 2018.

[23] G. Yuan, Z. Meng, and Y. Li, “A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations,” *J. Theory Optim. Appl.*, vol. 168, no. 1, pp. 129–152, Jan. 2016.

[24] G. Yuan, “Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems,” *Optim. Lett.*, vol. 3, no. 1, pp. 11–21, Jan. 2009.

[25] R. Fletcher and C. M. Reeves, “Function minimization by conjugate gradients,” *Comput. J.*, vol. 7, no. 2, pp. 149–154, Feb. 1964.

[26] Y. H. Dai and Y. Yuan, “A nonlinear conjugate gradient method with a strong global convergence property,” *SIAM J. Optim.*, vol. 10, no. 1, pp. 157–182, Jan. 1999.

[27] R. Fletcher, *Practical Methods of Optimization: Volume 1 Unconstrained Optimization*. Hoboken, NJ, USA: Wiley, 1980.

[28] M. R. Hestenes and E. Stiefel, “Methods of conjugate gradients for solving convex constrained nonlinear monotone equations,” *Math. Program.* vol. 36, no. 1, pp. 54–71, Jul. 2007.

[29] Z. Yu, J. Lin, X. Xiao, L. Liu, and Z. Li, “Spectral projection method for monotone nonlinear equations with convex constraints,” *Calcolo*, vol. 53, no. 2, pp. 133–145, Jun. 2016.

[30] W. La Cruz, J. Martínez, and M. Raydán, “Spectral residual method without gradient information for solving large-scale nonlinear systems of equations,” *Math. Comput.*, vol. 75, no. 255, pp. 1429–1448, 2006.

[31] W. La Cruz, “A spectral algorithm for large-scale systems of nonlinear monotone equations,” *Numer. Algorithms*, vol. 76, no. 4, pp. 1109–1130, Dec. 2017.

[32] Y. Bing and G. Lin, “An efficient implementation of Merrill’s method for sparse or partially separable systems of nonlinear equations,” *SIAM J. Optim.*, vol. 1, no. 2, pp. 206–221, May 1991.

[33] Z. Yu, J. Lin, J. Sun, Y. Xiao, L. Liu, and Z. Li, “Spectral projection method for monotone nonlinear equations with convex constraints,” *Appl. Numer. Math.*, vol. 59, no. 10, pp. 2416–2423, Oct. 2009.

[34] E. D. Dolan and J. J. Moré, “Benchmarking optimization software with performance profiles,” *Math. Program.*, vol. 91, no. 2, pp. 201–213, Jan. 2002.

[35] M. A. T. Figueiredo and R. D. Nowak, “An EM algorithm for wavelet-based image restoration,” *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, Aug. 2003.

[36] E. T. Hale, W. Yin, and Y. Zhang, “A fixed-point continuation method for ℓ1-regularized minimization with applications to compressed sensing,” *CAAM TR07-07, Rice Univ.*, vol. 43, p. 44, Jul. 2007.

[37] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[38] E. van den Berg and M. P. Friedlander, “Probing the Pareto frontier for basis pursuit solutions,” *SIAM J. Sci. Comput.*, vol. 31, no. 2, pp. 890–912, Jan. 2009.

[39] E. G. Birgin and M. J. Martínez, and M. Raydán, “Nonmonotone spectral projected gradient methods on convex sets,” *SIAM J. Optim.*, vol. 10, no. 4, pp. 1196–1211, Jan. 2000.

[40] J.-S. Pang, “Inexact Newton methods for the nonlinear complementarity problem,” *Math. Program.*, vol. 36, no. 1, pp. 54–71, 1986.

[41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
SANI AJI received the B.Sc. degree (Hons.) in mathematics from Gombe State University, Nigeria, in 2015, and the dual master's degrees in interdisciplinary mathematics and mathematical modelling from the University of L'Aquila, Italy, and the University of Silesia in Katowice, Poland, respectively. He is currently pursuing the Ph.D. degree with the King Mongkut's University of Technology Thonburi, Thailand. His research interests include solving nonlinear monotone equations with applications, image processing, machine learning, and computer vision.

POOM KUMAM (Member, IEEE) received the Ph.D. degree in mathematics from Naresuan University, Thailand. He is currently a Full Professor with the Department of Mathematics, King Mongkut's University of Technology Thonburi (KMUTT). He is also the Head of the Fixed Point Theory and Applications Research Group, KMUTT, where he is also with the Theoretical and Computational Science Center (TaCS-Center). He is also the Director of the Computational and Applied Science for Smart Innovation Cluster (CLASSIC Research Cluster), KMUTT. He has over 700 scientific articles and projects either presented or published. His research interests include fixed point theory, variational analysis, random operator theory, optimization theory, approximation theory, fractional differential equations, differential game, entropy and quantum operators, fuzzy soft set, mathematical modeling for fluid dynamics and inverse problems, dynamic games in economics, traffic network equilibria, bandwidth allocation problem, wireless sensor networks, image restoration, signal and image processing, game theory, and cryptology. He has provided and developed many mathematical tools in his fields productively over the past years. Moreover, he is on editorial board for more than 50 journals and also delivers many invited talks on different international conferences every year all around the world.

PUNNARAI SIRICHAROEN received the Ph.D. degree in computing science from Ulster University, U.K. She was a Lecturer with the Department of Mathematics, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand, until 2019. She was a Research Associate with the School of Computing and Information Engineering, Ulster University. She is currently a Lecturer with the Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok. Her research interests include image processing, computer vision, machine learning, and mobile-cloud computing applications.

AUWAL BALA ABUBAKAR received the master's degree in mathematics and the Ph.D. degree in applied mathematics from the King Mongkut's University of Technology Thonburi, Thailand, in 2015. He is currently a Lecturer with the Department of Mathematical Sciences, Faculty of Physical Sciences, Bayero University Kano, Nigeria. He is the author of more than 20 research articles. His main research interest includes methods for solving nonlinear monotone equations with application in signal recovery and image restoration.

MAHMOUD MUHAMMAD YAHAYA received the B.S. degree in pure/applied mathematics from Gombe State University, Gombe, in 2016. He is currently pursuing the master's degree in applied mathematics with the King Mongkut's University of Technology Thonburi, under the supervision of Prof. Poom Kumam. He went on to participate in a one-year mandatory national service called NYSC, from January 2017 to January 2018. His research interests include centred around convex/non-convex optimization, unconstrained optimization methods, and their applications to image processing.

* * *