Inertial Derivative-Free Projection Method for Nonlinear Monotone Operator Equations With Convex Constraints

AUWAL BALA ABUBAKAR, POOM KUMAM, and ABDULKARIM HASSAN IBRAHIM

1Fixed Point Research Laboratory, Fixed Point Theory and Applications Research Group, Center of Excellence in Theoretical and Computational Science (TaCS-CoE), Faculty of Science, King Mongkut’s University of Technology Thonburi (KMUTT), Bangkok 10140, Thailand
2Department of Mathematics and Applied Mathematics, Sefako Makgatho Health Sciences University, Ga-Rankuwa 0204, South Africa
3Department of Mathematics and Applied Mathematics, Faculty of Physical Sciences, Bayero University, Kano, Kano 700241, Nigeria
4Center of Excellence in Theoretical and Computational Science (TaCS-CoE), Faculty of Science, King Mongkut’s University of Technology Thonburi (KMUTT), Bangkok 10140, Thailand
5Departments of Medical Research, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan

Corresponding author: Poom Kumam (poom.kumam@mail.kmutt.ac.th)

This work was supported in part by the Center of Excellence in Theoretical and Computational Science (TaCS-CoE), King Mongkut’s University of Technology Thonburi, and in part by Thailand Science Research and Innovation (TSRI) Basic Research Fund: Fiscal year 2021 under Project 64A306000005.

This work did not involve human subjects or animals in its research.

ABSTRACT In this paper, we propose an inertial derivative-free projection method for solving convex constrained nonlinear monotone operator equations (CNME). The method incorporates the inertial step with an existing method called derivative-free projection (DFPI) method for solving CNME. The reason is to improve the convergence speed of DFPI as it has been shown and reported in several works that indeed the inertial step can speed up convergence. The global convergence of the proposed method is proved under some mild assumptions. Finally, numerical results reported clearly show that the proposed method is more efficient than the DFPI.

INDEX TERMS Monotone nonlinear operator, inertial algorithm, conjugate gradient, projection method.

I. INTRODUCTION

Consider the problem of finding \( y \in E \) such that
\[
T(y) = 0,
\tag{1}
\]
where \( T : \mathbb{R}^n \to \mathbb{R}^n \) is a monotone and Lipschitz continuous operator and \( E \) is a nonempty, closed and convex subset of \( \mathbb{R}^n \). This problem has recently received remarkable attention as it arises in a number of applicable problems. For example, in constrained neural networks [1], nonlinear compressed sensing [2], [3], phase retrieval [4], [5], power flow equations [6], economic and chemical equilibrium problems [7], [8], non-negative matrix factorisation [9], [10], forecasting of financial market, portfolio selection models, price returns [11]–[13] and many more. As such, recently several derivative-free methods such as the conjugate gradient (CG) method have been proposed for solving problem (1). Given an initial point \( y_0 \), the conjugate gradient method computes the next iterate as:
\[
y_{k+1} = y_k + \alpha_k d_k, \quad k = 0, 1, 2, \ldots,
\]
where \( \alpha_k > 0 \) is a step size and \( d_k \) is called the CG direction of search defined as
\[
d_k := \begin{cases} -T(y_k) & \text{if } k = 0, \\ -T(y_k) + \beta_k d_{k-1} & \text{if } k > 0. \end{cases}
\]

The parameter \( \beta_k \) is called the CG parameter. For more on derivative-free methods for solving (1), interested readers can refer to [14]–[35] and references therein.

Recently, several researchers are interested in how to improve the speed of convergence of existing iterative algorithms. One of the approach in this regard is the inertial extrapolation method where a new step called the inertial step is added to the existing step(s) of an iterative method. It has been shown that the inertial step enhance the speed of the existing methods such as methods for solving fixed
point problems, variational inequality problems, equilibrium problems, split feasibility problems, and so on. By choosing two starting points \( y_{-1} \) and \( y_0 \), the inertial term is defined as

\[ v_k = y_k + \theta_k (y_k - y_{k-1}) , \]

where \( \{ \theta_k \}_{k=1}^{\infty} \) is a sequence satisfying certain condition. Inertial extrapolation method has been employed successfully in improving the convergence of the sequence generated by various algorithms. However, to the best of our knowledge, there is no theoretical proof to justify that, indeed, all one can find is numerical justification using some examples. However, the choice of the parameter \( \theta_k \) has an effect on the speed of convergence. For more on iterative methods with inertial extrapolation, the reader is referred to [36]–[41] and references therein.

Inspired by the inertial methods [36]–[41] and the derivative-free projection method proposed by Sun and Liu [17] which is an extension of the work of Cheng [42], we propose an inertial derivative-free projection method for finding solutions to problem (1). The method is based on the work of Sun and Liu [17], where the inertial term is incorporated in order speed up its convergence. The remaining part of this paper is organized as follows: the next section gives some preliminaries and the proposed algorithm, convergence results is provided in the third section, Numerical results in the fourth section and lastly the conclusion.

**Notation.** Unless otherwise stated, the symbol \( \| \cdot \| \) stands for Euclidean norm on \( \mathbb{R}^n \).

II. PROPOSED ALGORITHM

**Definition 2.1:** Let \( \mathbb{R}^n \) be an Euclidean space and \( T : \mathbb{R}^n \to \mathbb{R}^n \) be a mapping. Then \( T \) is

(i) Monotone, if

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

(ii) \( L \)-Lipschitz continuous, if there exists \( L > 0 \) such that

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

**Definition 2.2:** Let \( E \subset \mathbb{R}^n \) be closed and convex, the projection of \( y \in \mathbb{R}^n \) onto \( E \) denoted by \( P_E(y) \), is defined as

\[ P_E(y) = \arg \min \{ \| x - y \| : x \in E \}. \]

**Lemma 2.3 ([43]):** Let \( E \subset \mathbb{R}^n \) be nonempty closed and convex. Then the following inequality hold:

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

**Lemma 2.4 ([44]):** Let \( y, x \in \mathbb{R}^n \). Then the following equality hold:

\[ \| y + x \|^2 = \| y \|^2 + 2 x^T (y + x). \]

**Lemma 2.5 ([45]):** Let \( \{ y_k \} \) and \( \{ x_k \} \) be sequences of nonnegative real number satisfying the following relation

\[ y_{k+1} \leq y_k + x_k, \]

where \( \sum_{k=1}^{\infty} x_k < \infty \), then \( \lim_{k \to \infty} y_k \) exists.

**Lemma 2.6 ([46]):** A point \( y^* \in \text{SOL}(T, E) \) if and only if \( y^* = P_E(y^* - \mu u) \) for some \( u = T(y^*) \) and \( \mu > 0 \).

We make use of the following assumptions.

**Assumption 1:**

(a) The feasible set \( E \) is a nonempty closed and convex subset of the Euclidean space \( \mathbb{R}^n \).

(b) \( T : \mathbb{R}^n \to \mathbb{R}^n \) is monotone and \( L \)-Lipschitz continuous.

(c) The solution set \( \text{SOL}(T, E) \) of (1) is nonempty.

**Assumption 2:** Let \( \{ \theta_k \} \) be a sequence of nonnegative real numbers satisfying the conditions:

\[ \theta_k \in (0, 1), \quad \sum_{k=1}^{\infty} \theta_k \| y_k - y_{k-1} \| < \infty. \]

Based on the Sun and Liu [17] derivative-free projection method for monotone nonlinear equation with convex constraints called DFPI, we present an inertial derivative-free projection method for finding solutions to problem (1).

**Algorithm 2.7** (Inertial Derivative-Free Method (IDFPI))

(S.0) Choose a sequence \( \{ \theta_k \}_{k=1}^{\infty} \) satisfying Assumption 2 and select the parameters: \( Tol > 0, \rho \in (0, 1), \xi > 0, \sigma > 0 \). Select arbitrary points \( y_{-1}, y_0 \in E \). Set \( k := 0 \).

(S.1) Set

\[ v_k = y_k + \theta_k (y_k - y_{k-1}) \]

(S.2) Compute \( T(v_k) \). If \( \| T(v_k) \| \leq Tol \), stop. Otherwise, generate the search direction \( d_k \) by

\[
\begin{align*}
T(v_k) + \beta_k d_k - 1 & \quad \text{if } k = 0, \\
- \frac{\| T(y_k) \|}{\| d_k - 1 \|} & \quad \text{if } k > 0,
\end{align*}
\]

where

\[ \beta_k := 0.01 \frac{\| T(y_k) \|}{\| d_k - 1 \|}. \]

(S.3) Compute a trial point \( x_k = v_k + \alpha_k d_k \).

(S.4) Determine the step-size \( \alpha_k = \xi \rho^i \) where \( i \) is the least nonnegative integer satisfying

\[ -T(v_k + \alpha_k d_k) \geq \sigma \alpha_k \| d_k \|^2. \]

(S.5) If \( x_k \in E \) and \( \| T(x_k) \| \leq Tol \), stop. Otherwise, \( y_{k+1} = P_E [ v_k - \gamma_k T(x_k) ] \).

(S.6) Set \( k = k + 1 \), and go back to (S.1).

**Remark 2.8:** Let \( d_k \) be generated by (2)-(3) in Algorithm 2.7. Then

\[ T(v_k)^T d_k = -\| T(v_k) \|^2. \]
III. CONVERGENCE RESULT

Lemma 3.1: The line search condition (4) is well-defined. That is, for all \( k \geq 0 \), there exists a non-negative integer \( i \) satisfying (4).

Proof: Suppose there is \( k_0 \geq 0 \) for which (4) is not true for any non-negative integer \( i \), i.e.,

\[
-T(v_{k_0} + \xi \rho^i d_{k_0})^T d_{k_0} < \sigma \xi \rho^i ||d_{k_0}||^2.
\]

Using Assumption 1 (b) and allowing \( i \to \infty \), we have that

\[
-T(v_{k_0})^T d_{k_0} \leq 0.
\] (7)

On the other hand, from (6),

\[
-T(v_{k_0})^T d_{k_0} = ||T(v_{k_0})||^2 > 0,
\]

which contradicts (7). Hence, (4) is well defined.

Lemma 3.2: Let \( \{y_k\} \) and \( \{x_k\} \) be generated via Algorithm 2.7. If \( y^* \in \text{SOL}(T, E) \), then under Assumption 1 and 2, it holds that

\[
||y_{k+1} - y^*||^2 \leq ||y_k - y^*|| - \sigma^2 ||y_k - x_k||^4.
\]

Moreover, the sequence \( \{y_k\} \) and \( \{x_k\} \) are bounded and

\[
\lim_{k \to \infty} ||y_k - x_k|| = 0.
\] (8)

Proof: By the monotonicity of the mapping \( T \), we have

\[
T(x_k)^T (v_k - y^*) = T(x_k)^T (v_k - x_k) + T(x_k)^T (x_k - y^*)
\]

\[
\geq T(x_k)^T (v_k - x_k) + T(y^*)^T (x_k - y^*)
\]

\[
= T(x_k)^T (v_k - x_k)
\]

\[
= T(x_k)^T (-\alpha_k d_k)
\]

\[
= \sigma \alpha_k^2 ||d_k||^2
\]

\[
\geq \sigma \||y_k - x_k||^2.
\] (9)

By Lemma 2.3 (iii), (5), (9) and (10), it holds that for any \( y^* \in \text{SOL}(T, E) \),

\[
||y_{k+1} - y^*||^2 = ||P_E(y_k T(x_k)) - y^*||^2
\]

\[
\leq ||y_k - y^*||^2 - 2\gamma_k T(x_k)^T (y_k - y^*)
\]

\[
+ \gamma_k^2 ||T(x_k)||^2
\]

\[
\leq ||y_k - y^*||^2 - 2\gamma_k T(x_k)^T (v_k - x_k)
\]

\[
+ \gamma_k^2 ||T(x_k)||^2
\]

\[
\leq ||y_k - y^*||^2 - \frac{T(x_k)^T (v_k - x_k)^2}{||T(x_k)||^2}
\]

\[
\leq ||y_k - y^*||^2 - \frac{\sigma^2 ||y_k - x_k||^4}{||T(x_k)||^2}.
\] (11)

From equation (11), we can deduce that

\[
||y_{k+1} - y^*|| \leq ||y_k - y^*||
\]

\[
= ||y_k + \theta_k (y_k - y_{k-1}) - y^*||
\]

\[
\leq ||y_k - y^*|| + \theta_k ||y_k - y_{k-1}||.
\] (12)
Thus, we have
\[ \frac{\sigma^2 \|v_k - x_k\|^4}{\|T(x_k)\|^2} \leq \|y_k - y^*\|^2 + 4M\theta_k \|y_k - y_{k-1}\| - \|y_{k+1} - y^*\|^2. \]  
(17)

Adding (17) for \( k = 0, 1, 2, \ldots \) and the fact that \( \{T(x_k)\} \) is bounded, we have
\[ \frac{\sigma^2 \sum_{k=0}^{\infty} \|v_k - x_k\|^4}{N^2} \leq \sum_{k=0}^{\infty} (\|y_k - y^*\|^2 + 4M\theta_k \|y_k - y_{k-1}\| - \|y_{k+1} - y^*\|^2). \]  
(18)

Now, let \( S_k = \sum_{n=0}^{k} (\|y_n - y^*\|^2 - \|y_{n+1} - y^*\|^2) \), then \( S_k = \sum_{n=0}^{k} (\|y_0 - y^*\|^2 - \|y_{k+1} - y^*\|^2) \). As limit of \( \{\|y_k - y^*\|\} \) exists from (12) with limit say \( L_1 \), then
\[ \left( \lim_{k \to \infty} S_k = \|y_0 - y^*\|^2 - L_1 \right) \in \mathbb{R}. \]

So,
\[ \sum_{k=0}^{\infty} (\|y_k - y^*\|^2 - \|y_{k+1} - y^*\|^2) < \infty \]
and \( \sum_{k=0}^{\infty} \theta_k \|y_k - y_{k-1}\| < \infty \).

Using (18) together with the above inequalities, we conclude that
\[ \lim_{k \to \infty} \|v_k - x_k\| = 0. \]

Remark 3.3: By the definition of \( \{x_k\} \) and (8), we have
\[ \lim_{k \to \infty} \alpha_k \|d_k\| = 0. \]

Lemma 3.4: Suppose Assumptions 1-2 hold and the sequence \( \{y_k\} \) and \( \{v_k\} \) are generated by Algorithm 2.7. Then
\[ \lim_{k \to \infty} \|v_k - y_{k+1}\| = 0. \]  
(19)

Proof:
Using definition of \( v_k \),
\[ \|y_k - v_k\| = \|y_k - (y_k + \theta_k(y_k - y_{k-1}))\| = \theta_k \|y_k - y_{k-1}\|. \]
This implies that
\[ \lim_{k \to \infty} \|y_k - v_k\| = 0. \]  
(20)

Also,
\[ \|y_k - x_k\| = \|y_k - v_k + v_k - x_k\| \leq \|y_k - v_k\| + \|v_k - x_k\|. \]
Using (8) and (20), we have
\[ \lim_{k \to \infty} \|y_k - x_k\| = 0. \]

TABLE 1. Starting points.

| SP | \( y_{-1} \) | \( y_0 \) |
|----|-------------|-------------|
| \( y_1 \) | \((0.2, \ldots, 0.2)^T\) | \((0.1, \ldots, 0.1)^T\) |
| \( y_2 \) | \((0.2, \ldots, 0.2)^T\) | \((0.2, \ldots, 0.2)^T\) |
| \( y_3 \) | \((0.5, \ldots, 0.5)^T\) | \((0.5, \ldots, 0.5)^T\) |
| \( y_4 \) | \((1.2, \ldots, 1.2)^T\) | \((1.2, \ldots, 1.2)^T\) |
| \( y_5 \) | \((1.5, \ldots, 1.5)^T\) | \((1.5, \ldots, 1.5)^T\) |
| \( y_6 \) | \((2.2, \ldots, 2)^T\) | \((2.2, \ldots, 2)^T\) |
| \( y_7 \) | \(\operatorname{rand}(n, 1)\) | \(\operatorname{rand}(n, 1)\) |

Note. For DFPI algorithm [17], the starting point is \( y_0 \).

By Lemma 2.3, we have
\[ \|y_{k+1} - y_k\| = \|P_{E}[v_k - y_k T(x_k)] - y_k\| \]
\[ \leq \|v_k - y_k T(x_k) - y_k\| \]
\[ \leq \|v_k - y_k\| + \|y_k E(x_k)\| \]
\[ = \|v_k - y_k\| + \|T(x_k)^T(v_k - x_k)\|/\|T(x_k)\|^2 \]
\[ \leq \|v_k - y_k\| + \|v_k - x_k\|. \]  
(22)

Thus, from (8) and (20), we have
\[ \lim_{k \to \infty} \|y_{k+1} - y_k\| = 0. \]  
(23)

Therefore,
\[ \|y_{k+1} - v_k\| = \|y_{k+1} - (y_k + \theta_k(y_k - y_{k-1}))\| \]
\[ \leq \|y_{k+1} - y_k\| + \theta_k \|y_k - y_{k-1}\|. \]
Using (23) and Assumption 2, the desired equation is obtained.

Theorem 3.5: Let \( \{y_k\} \) be a sequence generated via Algorithm 2.7. Using Assumption 1 and 2, then \( \{y_k\} \) converge to an element of \( \mathbf{SOL}(T, E) \).

Proof: We know that the sequence \( \{y_k\} \) is bounded from (13). This implies that there exists a subsequence \( \{y_{k_j}\} \) of \( \{y_k\} \) such that \( \{y_{k_j}\} \) converge to some point \( \bar{y} \). Also, we have that
\[ \|v_{k_j} - y_{k_j}\| = \theta_{k_j} \|y_{k_j} - y_{k_j-1}\| \to 0, \text{ as } j \to \infty. \]  
(24)

Claim: \( \bar{y} \in \mathbf{SOL}(T, E) \). Suppose on the contrary that \( \bar{y} \notin \mathbf{SOL}(T, E) \). Then from (19) and (24), we have that
\[ \lim_{j \to \infty} y_{k_j+1} = \lim_{j \to \infty} P_E(v_{k_j} - y_{k_j} T(x_{k_j})) = \lim_{j \to \infty} v_{k_j} = \bar{y}. \]  
(25)

Without loss of generality, if \( y_{k_j} \to y^* \) and \( T(x_{k_j}) \to T(x^*) \). Then since \( T \) is continuous, we have \( T(x^*) = T(\bar{y}) \). Therefore, from (25)
\[ P_E(\bar{y} - y^* T(x^*)) = \bar{y}. \]

It then follows from Lemma 2.6 that \( y^* \in \mathbf{SOL}(T, E) \), which is a contradiction. Hence, our claim holds. Substituting \( y^* \) with \( \bar{y} \) in (12), it is easy to see that \( \lim_{k \to \infty} \|y_k - \bar{y}\| \) exists.
IV. NUMERICAL EXAMPLES

By comparing the proposed inertial algorithm (Iner. DFPI) to the DFPI algorithm in [17], we show the numerical efficiency and computational advantage of the proposed inertial algorithm (Iner. DFPI) in this section. The MATLAB implementation of the algorithms was executed on a Windows 10 computer with Intel(R) Core(TM) i7 processor with 8.0GB of RAM and CPU of 2.30GHz using MATLAB R2019b software. The numerical experiment made use of the following test problems to measure the efficiency and robustness of the proposed inertial algorithm (Iner. DFPI).

Problem 1: Modified exponential function [47]
\[ t_i(y) = e^{y_i} - 1 \]
\[ E = \mathbb{R}^n_+ . \]

Problem 2: Logarithmic function [47]
\[ t_i(y) = \log(y_i + 1) - \frac{y_i}{n}, \quad i = 1, 2, \ldots, n, \]
\[ E = \mathbb{R}^n_+ . \]

Problem 3: Nonsmooth function [48]
\[ t_i(y) = 2y_i - \sin(|y_i|), \quad i = 1, 2, \ldots, n, \]
\[ E = \{ y \in \mathbb{R}^n_+ : y \geq 0, \sum_{i=1}^{n} y_i \leq n \} . \]

Problem 4: Strictly convex function I [47]
\[ t_i(y) = e^{y_i} - 1, \quad i = 1, 2, \ldots, n, \]
\[ E = \mathbb{R}^n_+ . \]

Problem 5: Strictly convex function II [47]
\[ t_i(y) = \left( \frac{i}{n} \right) e^{y_i} - 1, \quad i = 1, 2, \ldots, n, \]
\[ E = \mathbb{R}^n_+ . \]
### Table 3: Numerical experiments with different coefficients of $\beta_k$ for problem 6-10 with $n = 1000$. 

| Problem | $\beta_k$ | $n=1000$ | SP | NOI | NFE | CPUT | $\beta_k$ | NOI | NFE | CPUT | $\beta_k$ | NOI | NFE | CPUT | $\beta_k$ | NOI | NFE | CPUT |
|---------|-----------|----------|----|-----|-----|------|-----------|-----|-----|------|-----------|-----|-----|------|-----------|-----|-----|------|
| Problem 6 | 0.01 | 16 | 64 | 0.028171 | 16 | 64 | 0.008079 | 74 | 296 | 0.039752 | 530 | 2120 | 0.27322 | 0.01 | 16 | 64 | 0.006196 | 16 | 64 | 0.007641 | 73 | 292 | 0.038947 | 529 | 2116 | 0.26621 |
|         | 1.0 | 16 | 64 | 0.007024 | 16 | 64 | 0.007142 | 74 | 296 | 0.039977 | 524 | 2096 | 0.24974 | 1.5 | 15 | 60 | 0.007648 | 15 | 60 | 0.007185 | 72 | 288 | 0.031371 | 510 | 2040 | 0.2355 |
|         | 2.0 | 15 | 60 | 0.007387 | 15 | 60 | 0.007385 | 70 | 280 | 0.03291 | 496 | 1984 | 0.24244 | 2.0 | 16 | 60 | 0.00807 | 15 | 60 | 0.007901 | 65 | 260 | 0.031665 | 467 | 1868 | 0.2211 |
|         | 3.0 | 16 | 64 | 0.007314 | 16 | 64 | 0.007423 | 43 | 172 | 0.020312 | 489 | 1956 | 0.22663 |
| Problem 7 | 0.01 | 8 | 32 | 0.024297 | 8 | 32 | 0.00364 | 8 | 32 | 0.004126 | 8 | 32 | 0.00538 | 0.01 | 7 | 28 | 0.00333 | 7 | 28 | 0.002719 | 7 | 28 | 0.002391 | 7 | 28 | 0.003311 |
|         | 1.0 | 8 | 32 | 0.003908 | 8 | 32 | 0.003102 | 9 | 32 | 0.003234 | 9 | 32 | 0.003993 | 1.5 | 8 | 32 | 0.004041 | 8 | 32 | 0.004093 | 9 | 32 | 0.003996 | 8 | 32 | 0.003798 |
|         | 2.0 | 8 | 31 | 0.003253 | 8 | 31 | 0.003851 | 8 | 31 | 0.003888 | 8 | 31 | 0.003131 | 2.0 | 9 | 36 | 0.004528 | 21 | 84 | 0.007788 | 456 | 1824 | 0.16334 | 410 | 1640 | 0.3094 |
|         | 3.0 | 8 | 31 | 0.003030 | 9 | 32 | 0.003161 | 9 | 32 | 0.00347 | 9 | 32 | 0.003267 | 3.0 | 9 | 32 | 0.003586 | 9 | 32 | 0.003489 | 9 | 32 | 0.003431 | 9 | 32 | 0.005727 |
|         | 4.0 | 9 | 32 | 0.003614 | 9 | 32 | 0.002902 | 9 | 32 | 0.00275 | 9 | 32 | 0.004086 | 4.0 | 9 | 31 | 0.002028 | - | - | - | - | - | - | - |
| Problem 8 | 0.01 | 5 | 20 | 0.016133 | 5 | 20 | 0.011863 | 5 | 20 | 0.014187 | 5 | 20 | 0.026229 | 0.01 | 5 | 20 | 0.001592 | 5 | 20 | 0.00164 | 5 | 20 | 0.001465 | 5 | 20 | 0.001785 |
|         | 1.0 | 5 | 20 | 0.002472 | 5 | 20 | 0.001823 | 5 | 20 | 0.001626 | 5 | 20 | 0.002922 | 1.0 | 5 | 24 | 0.001959 | 6 | 24 | 0.001971 | 6 | 24 | 0.002312 | 6 | 24 | 0.002833 |
|         | 2.0 | 5 | 24 | 0.002435 | 6 | 24 | 0.002336 | 6 | 24 | 0.002423 | 6 | 24 | 0.003877 | 2.0 | 6 | 24 | 0.001855 | 6 | 24 | 0.001686 | 6 | 24 | 0.00182 | 6 | 24 | 0.001819 |
|         | 3.0 | 5 | 20 | 0.001481 | 5 | 20 | 0.002265 | 5 | 20 | 0.001523 | 5 | 20 | 0.001975 | 3.0 | 9 | 36 | 0.001723 | 9 | 36 | 0.006713 | 9 | 36 | 0.006291 | 9 | 36 | 0.010543 |
|         | 4.0 | 10 | 40 | 0.00665 | 10 | 40 | 0.006561 | 10 | 40 | 0.007961 | 10 | 40 | 0.009544 | 4.0 | 1 | 3 | 0.000977 | 1 | 3 | 0.001388 | 1 | 3 | 0.001115 | 1 | 3 | 0.002128 |
|         | 5.0 | 1 | 4 | 0.001337 | 1 | 4 | 0.001396 | 1 | 4 | 0.00149 | 1 | 4 | 0.001539 | 5.0 | 1 | 3 | 0.001488 | 1 | 3 | 0.001946 | 1 | 3 | 0.001396 | 1 | 3 | 0.002422 |
|         | 6.0 | 1 | 4 | 0.001899 | 1 | 4 | 0.002218 | 1 | 4 | 0.001877 | 1 | 4 | 0.002951 | 6.0 | 11 | 44 | 0.007231 | 45 | 180 | 0.028201 | 447 | 1788 | 0.24792 | 405 | 1620 | 0.67977 |

**Problem 6:** Tridiagonal exponential function \[ t_i(y) = y_1 - e^{\cos(y_1+y_2)} \]

**Problem 7:** Nonsmooth function II \[ t_i(y) = y_i - e^{\cos(y_{i-1} + y_i + y_{i+1})}, \quad i = 2, \ldots, n - 1, \]

**Problem 8:** Penalty function I \[ t_i(y) = 2c(y_i - y_i^2), \quad c = 10^{-5}, \]

**Problem 9:** Pursuit-Evasion problem \[ t_i(y) = 8^{0.5}y_i - 1, \quad i = 1, 2, \ldots, n, \]

**Problem 10:** Pursuit-Evasion problem \[ t_i(y) = e^{y_i^2} + 3\sin(y_i)\cos(y_i - 1), \quad i = 1, 2, \ldots, n, \]
TABLE 4. Numerical experiments with different sequences \( \{ \theta_k \} \) for problem 1-5 with \( n = 1000 \).

| Problem | \( \theta_k = \frac{1}{2k+1} \) | \( \xi_k = 0.1 \) | \( \rho = 0.7 \) | \( \sigma = 0.01 \) |
|---------|----------------|----------------|----------------|----------------|
| 1       | 36 0.00919 9  | 36 0.19349 3  | 0 0.00001 2  | 36 0.00919 9  |
| 2       | 36 0.00919 14  | 36 0.19349 4  | 0 0.00001 2  | 36 0.00919 14  |
| 3       | 36 0.00919 19  | 36 0.19349 5  | 0 0.00001 2  | 36 0.00919 19  |
| 4       | 36 0.00919 24  | 36 0.19349 6  | 0 0.00001 2  | 36 0.00919 24  |
| 5       | 36 0.00919 29  | 36 0.19349 7  | 0 0.00001 2  | 36 0.00919 29  |

For the compared method (DFPI), its parameters were set as reported in [17]. All iterative procedure terminate when \( \| T(v_k) \| < 10^{-6} \) is fulfilled. If this condition is not satisfied after 1000 iterations, failure is declared.

TABLE 5. Numerical experiments with different sequences \( \{ \theta_k \} \) for problem 6-10 with \( n = 1000 \).

| Problem | \( \theta_k = \frac{1}{2k+1} \) | \( \xi_k = 0.1 \) | \( \rho = 0.7 \) | \( \sigma = 0.01 \) |
|---------|----------------|----------------|----------------|----------------|
| 6       | 36 0.00919 9  | 36 0.19349 3  | 0 0.00001 2  | 36 0.00919 9  |
| 7       | 36 0.00919 14  | 36 0.19349 4  | 0 0.00001 2  | 36 0.00919 14  |
| 8       | 36 0.00919 19  | 36 0.19349 5  | 0 0.00001 2  | 36 0.00919 19  |
| 9       | 36 0.00919 24  | 36 0.19349 6  | 0 0.00001 2  | 36 0.00919 24  |
| 10      | 36 0.00919 29  | 36 0.19349 7  | 0 0.00001 2  | 36 0.00919 29  |

TABLE 6. Numerical results for DFPI and DFPI algorithms on problem 1.

| Problem | \( \theta_k = \frac{1}{2k+1} \) | \( \xi_k = 0.1 \) | \( \rho = 0.7 \) | \( \sigma = 0.01 \) |
|---------|----------------|----------------|----------------|----------------|
| 1       | 36 0.00919 9  | 36 0.19349 3  | 0 0.00001 2  | 36 0.00919 9  |
| 2       | 36 0.00919 14  | 36 0.19349 4  | 0 0.00001 2  | 36 0.00919 14  |
| 3       | 36 0.00919 19  | 36 0.19349 5  | 0 0.00001 2  | 36 0.00919 19  |
| 4       | 36 0.00919 24  | 36 0.19349 6  | 0 0.00001 2  | 36 0.00919 24  |
| 5       | 36 0.00919 29  | 36 0.19349 7  | 0 0.00001 2  | 36 0.00919 29  |

The \( \theta_k \) were chosen for the Iner. DFPI algorithm to obtain the best possible results.
To illustrate in detail the efficiency and robustness of Iner. DFPI, we start by performing some numerical experiments with different coefficients of the parameter $\beta_k$ and the results are reported in Table 2 and 3. It can be observed from the tables that the coefficient $0.01$ is a good choice. In addition, we performed another numerical experiments with different sequences $\{\theta_k\}$ and the results are reported in Table 4 and 5. It can be observed from the tables that the sequence $\theta_k = 1/(2^k + 5)$ is a good choice. We further employ the performance

To be continued...
profile proposed by Dolan and Moré in [50] in order to summarize Table 6-15. The profile is defined as follows:

\[
\rho(\tau) := \frac{1}{|\mathcal{T}_P|} \left| \left\{ \mathbf{t}_p \in \mathcal{T}_P : \log_2 \left( \min_{q \in \mathcal{Q}} (\mathbf{t}_p, q) \right) \leq \tau \right\} \right|
\]

where \(\mathcal{T}_P\) is the test set, \(|\mathcal{T}_P|\) is the number of problems in the test set, \(\mathcal{Q}\) is the set of optimization solvers, and \(\mathbf{t}_p, q\) is the number of iterations (or the number of the function evaluations, or the CPU time (in seconds)) for \(\mathbf{t}_p \in \mathcal{T}_P\) and \(q \in \mathcal{Q}\).
The performance profile tells the percentage of win by each solver. Figures 1, 2 and 3 illustrate the performance of the two solvers (Iner. DFPI and DFPI) where the performance indices are the number of iterations, the number of function evaluations and the CPU time in seconds as reported in Tables 2-11. It can be observed from the figures that Iner. DFPI algorithm performs better with a higher percentage win of at least 90% in all the three metrics, i.e., number of iterations, the number of function evaluations and the CPU time. As a consequence, we can conclude that Iner. DFPI algorithm is an efficient solver. It is worth mentioning that the good numerical performance of the Iner. DFPI algorithm is as a result of the inertial term $v_k$, suitable control parameters such as $\rho$, $\sigma$ and the sequence $\{\theta_k\}$.

A detailed report of our numerical experiments is reported in Table 6-15 in the appendix section. The abbreviations on the tables can be read as follows:
- n: denotes the dimension of the problem
- SP: denotes the starting points
- NOI: denotes the number of iterations
- NFE: denotes the number of function evaluations
- CPUT: denotes the CPU time in seconds
- LNORM: denotes the final norm

V. CONCLUSION
In this paper, we suggested an inertial derivative-free method for solving nonlinear monotone operator equation. Based on the DFPI method, an inertial term was added to it in order to speed up its convergence. We used some mild assumptions to establish the global convergence of the proposed inertial method. To support the theoretical results, we perform some numerical experiments on some benchmark test problems with the proposed method and the DFPI. The results indicate that the proposed inertial method is faster than DFPI.

APPENDIX
See Tables 2–15.

ACKNOWLEDGMENT
The author Auwal Bala Abubakar would like to thank the Postdoctoral Fellowship from King Mongkut’s University of Technology Thonburi (KMUTT), Thailand. He also acknowledge with thanks, the Department of Mathematics and Applied Mathematics at the Sefako Makgatho Health Sciences University.

REFERENCES
[1] J. Chorowski and J. M. Zurada, “Learning understandable neural networks with nonnegative weight constraints,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 1, pp. 62–69, Jan. 2015.
[2] T. Blumensath, “Compressed sensing with nonlinear observations and related nonlinear optimization problems,” *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3466–3474, Jun. 2013.
[3] A. B. Abubakar, P. Kumam, H. Mohammad, and A. M. Awwal, “A Barzilai-Borwein gradient projection method for sparse signal and blurred image restoration,” *J. Franklin Inst.*, vol. 357, no. 11, pp. 7266–7285, Jul. 2020.
[4] E. J. Candes, X. Li, and M. Soltanolkotabi, “Phase retrieval via Wirtinger flow: Theory and algorithms,” *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, Apr. 2015.
A. H. Ibrahim, P. Kumam, A. B. Abubakar, and H. Mohammad, “Spectral conjugate gradient like method for signal reconstruction,” Thai J. Math., vol. 18, no. 4, pp. 2013–2022, 2020.

A. H. Ibrahim and P. Kumam, “Re-modified derivative-free iterative method for nonlinear monotone equations with convex constraints,” Ain Shams Eng. J., vol. 12, no. 2, pp. 2205–2210, Jun. 2021.

H. Mohammad, “Barzilai-Borwein-like method for solving large-scale nonlinear systems of equations,” J. Niger. Math. Soc., vol. 36, no. 1, pp. 71–83, 2017.

A. B. Abubakar and P. Kumam, “A descent Dai-Liao conjugate gradient method for nonlinear equations,” Numer. Algorithms, vol. 81, no. 1, pp. 197–210, May 2020.

A. B. Abubakar and P. Kumam, “An improved three-term derivative-free method for solving nonlinear equations,” Comput. Appl. Math., vol. 37, no. 5, pp. 6760–6773, Nov. 2018.

A. B. Abubakar, P. Kumam, and H. Mohammad, “A note on the spectral gradient projection method for nonlinear monotone equations with applications,” Comput. Appl. Math., vol. 39, no. 2, pp. 1–13, May 2020.

H. Mohammad and A. B. Abubakar, “A descent derivative-free algorithm for nonlinear monotone equations with convex constraints,” RAIRO-Oper. Res., vol. 54, no. 2, pp. 489–505, Mar. 2020.

A. B. Abubakar, K. Muangchoo, A. H. Ibrahim, S. E. Fadugba, K. O. Aremu, and L. O. Jolaoso, “A modified scaled spectral-conjugate gradient-based algorithm for solving monotone operator equations,” J. Math., vol. 2021, pp. 1–9, Apr. 2021.

G. N. Ogwo, C. Izuachuku, and O. T. Mewomo, “Inertial methods for finding minimum-norm solutions of the split variational inequality problem beyond monotonicity,” Numer. Algorithms, 2021, doi: 10.1007/s11075-021-01081-1.

N. Wairojana, H. U. Rehman, N. Pakkarangan, and C. Khanpanuk, “An accelerated Popov’s subgradient extragradient method for strongly pseudomonotone equilibrium problems in a real Hilbert space with applications,” Commun. Math. Appl., vol. 11, no. 4, pp. 513–526, 2020.

N. Wairojana, H. U. Rehman, M. De la Sen, and N. Pakkarangan, “A general inversion-type algorithm for solving equilibrium problem in Hilbert spaces with applications in fixed-point problems,” Axioms, vol. 9, no. 3, p. 101, Aug. 2020.

P. Cholamjiak, D. V. Thong, and Y. J. Cho, “A novel inertial projection and contraction method for solving pseudomonotone variational inequality problems,” Acta Appl. Math., vol. 169, pp. 1–29, Nov. 2019.

D. V. Thong, N. T. Vinh, and Y. J. Cho, “Accelerated subgradient extra-gradient methods for variational inequality problems,” J. Sci. Comput., vol. 80, no. 3, pp. 1438–1462, Sep. 2019.

D. V. Thong, X.-H. Li, Q.-L. Dong, Y. J. Cho, and T. M. Rassias, “An inertial Popov’s method for solving pseudomonotone variational inequalities,” Optim. Lett., vol. 15, pp. 1–21, May 2020.

W. Cheng, “A two-term PRP-based descent method,” Numer. Funct. Anal. Optim., vol. 28, nos. 11–12, pp. 1217–1230, Dec. 2007.

X. Y. Wang, S. J. Li, and X. P. Kou, “A self-adaptive three-term conjugate gradient method for nonlinear monotone equations with convex constraints,” Calcolo, vol. 53, no. 2, pp. 133–145, Jun. 2016.

W. Takahashi, Introduction to Nonlinear and Convex Analysis, Tokyo, Japan: Yokohama Publishers, 2009.

A. Auslender, M. Teboulle, and S. Benslima, “A logarithmic-quadratic proximal method for variational inequalities,” in Computational Optimization, Boston, MA, USA: Springer, 1999, pp. 31–40.

F. Facchinei and J.-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems. New York, NY, USA: Springer-Verlag, 2007.

M. Jaureguy, J.-S. Pang, and R. Womersley, Spectral residual method without gradient information for solving large-scale nonlinear systems: Theory and experiments,” Math. Comput., vol. 75, pp. 1429–1448, Apr. 2006. [Online]. Available: https://www.ime.unicamp.br/martinez/Imrepsort.pdf, doi: 10.1090/S0025-5718-06-01840-0.

Q. Li and D.-H. Li, “A class of derivative-free methods for large-scale nonlinear monotone equations,” IMA J. Numer. Anal., vol. 31, no. 4, pp. 1625–1655, Oct. 2011.

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

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