

RESEARCH

Open Access



New inertial self-adaptive algorithms for the split common null-point problem: application to data classifications

Ratthaprom Promkam¹, Pongsakorn Sunthrayuth^{1*}, Suparat Kesornprom² and Ekapak Tanprayoon¹

*Correspondence:

pongsakorn_su@rmutt.ac.th

¹Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi (RMUTT), Thanyaburi, Pathumthani, 12110, Thailand

Full list of author information is available at the end of the article

Abstract

In this paper, we propose two inertial algorithms with a new self-adaptive step size for approximating a solution of the split common null-point problem in the framework of Banach spaces. The step sizes are adaptively updated over each iteration by a simple process without the prior knowledge of the operator norm of the bounded linear operator. Under suitable conditions, we prove the weak-convergence results for the proposed algorithms in p -uniformly convex and uniformly smooth Banach spaces. Finally, we give several numerical results in both finite- and infinite-dimensional spaces to illustrate the efficiency and advantage of the proposed methods over some existing methods. Also, data classifications of heart diseases and diabetes mellitus are presented as the applications of our methods.

Mathematics Subject Classification: 47H09; 47H10; 47J25; 47J05

Keywords: p -uniformly convex Banach spaces; Weak convergence; Split common null-point problem; Maximal monotone operator; Self-adaptive algorithm

1 Introduction

In this paper, we consider the following *split common null-point problem* [13] (see also [29]): find $z \in H_1$ such that

$$z \in A^{-1}0 \cap T^{-1}(B^{-1}0), \quad (1.1)$$

where $A : H_1 \rightarrow 2^{H_1}$ and $B : H_2 \rightarrow 2^{H_2}$ are set-valued maximal monotone operators, $T : H_1 \rightarrow H_2$ is a bounded linear operator, and H_1 and H_2 are real Hilbert spaces. We denote the solution set of the split common null-point problem (1.1) by Ω . The split common null-point problem can be applied to solving many real-life problems, for instance, in practices as a model in intensity-modulated radiation-therapy treatment planning [14, 15] and in sensor networks in computerized tomography and data compression [19]. In addition, the split common null-point problem also generalizes several split-type problems that is the core the modeling of many inverse problems such as the split feasibility problem, the split equilibrium problem, and the split minimization problem as special cases.

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Byrne et al. [13] introduced the following iterative scheme for solving the split common null-point problem: for given $x_0 \in H_1$ and the sequence $\{x_n\}$ generated iteratively by

$$x_{n+1} = R_\tau(x_n - \lambda T^*(I - Q_\mu)Tx_n), \tag{1.2}$$

where $R_\tau = (I + \tau A)^{-1}$ and $Q_\mu = (I + \mu B)^{-1}$ are the resolvent operators of A for $\tau > 0$ and of B for $\mu > 0$, respectively. They proved the weak-convergence theorem for solving the split common null-point problem provided the step size $\lambda \in (0, \frac{2}{\|T\|^2})$.

Alofi et al. [5] introduced the following iterative scheme based on a modified Halpern’s iteration for solving the split common null-point problem in the case that H_1 is a Hilbert space and F is a uniformly convex and smooth Banach space: for given $x_1 \in H_1$ and the sequence $\{x_n\}$ generated iteratively by

$$x_{n+1} = \beta_n x_n + (1 - \beta_n)(\alpha_n u_n + (1 - \alpha_n)R_{\tau_n}(x_n - \tau_n T^*J(I - Q_{\mu_n})Tx_n)), \tag{1.3}$$

where R_τ is the resolvent of A for $\tau > 0$ and Q_μ is the metric resolvent of B for $\mu > 0$, $\{\tau_n\}, \{\mu_n\} \subset (0, \infty)$, $\{\alpha_n\} \subset (0, 1)$, and $\{\beta_n\} \subset (0, 1)$ that satisfies some appropriate assumptions on the parameters, J is the duality mapping on F , T is the bounded linear operator from H_1 to F , and $\{u_n\}$ is the sequence in H_1 such that $u_n \rightarrow u$. They proved that the sequence $\{x_n\}$ generated by (1.3) converges strongly to a point of Ω provided τ_n satisfies the following inequality:

$$0 < a \leq \tau_n \|T\|^2 \leq b < 2 \tag{1.4}$$

for some $a, b > 0$.

Later, Suantai et al. [39] generalized the result of Alofi et al. [5] in the case that E is a p -uniformly convex and uniformly smooth Banach space, and F is a uniformly convex and smooth Banach space. To be more precise, they introduced the following scheme: for given $x_1 \in E$ and the sequence $\{x_n\}$ generated iteratively by

$$\begin{cases} z_n = J_q^{E^*}(J_p^E(x_n) - \tau_n T^*J_p^F(I - Q_{\mu_n})Tx_n), \\ y_n = J_q^{E^*}(\alpha_n J_p^E(u_n) + (1 - \alpha_n)J_p^E(R_{\tau_n}z_n)), \\ x_{n+1} = J_q^{E^*}(\beta_n J_p^E(x_n) + (1 - \beta_n)J_p^E(y_n)), \end{cases} \tag{1.5}$$

where J_p^E and $J_q^{E^*}$ are the generalized duality mapping of E into E^* and the duality mapping of E^* into E , respectively, where $1 < q \leq 2 \leq p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, and T is the bounded linear operator from E to F . They also proved the strong convergence of the sequence $\{x_n\}$ generated by (1.3) to a point of Ω provided τ_n satisfies the following inequality:

$$0 < a \leq \tau_n \leq b < \left(\frac{q}{\kappa_q \|T\|^q}\right)^{\frac{1}{q-1}} \tag{1.6}$$

for some $a, b > 0$.

However, several iterative methods involve a step size that requires to compute the norm of the bounded linear operator $\|T\|$ prior to choosing τ_n . In general, it may not be easy to compute $\|T\|$. In particular, it makes the algorithms not easily implemented when the

computation of $\|T\|$ is complicated. To overcome this drawback, a new step-size strategy without the prior knowledge of the operator norm of the bounded linear operator has been proposed by López et al. [28]. This method is known as a *self-adaptive method* that was first used to solve the split feasibility problem when the step-size criterion is independent of the operator norm of the bounded linear operator.

In optimization theory, the inertial technique has been widely used to accelerate the rate of convergence of the algorithms. This technique was motivated by the implicit time discretization of second-order dynamical systems (or a heavy ball with friction). Based on the inertial technique, Alvarez and Attouch [7] proposed the following so-called *inertial proximal point algorithm* for finding a zero point of a set-valued maximal monotone operator A : for given $x_0, x_1 \in H$ and the sequence $\{x_n\}$ generated iteratively by

$$x_{n+1} = R_{\tau_n}(x_n + \theta_n(x_n - x_{n-1})), \tag{1.7}$$

where R_{τ_n} is the resolvent of A and $x_n + \theta_n(x_n - x_{n-1})$ is called the *inertial term*. They also proved that the sequence $\{x_n\}$ generated by (1.7) converges weakly to a zero point of A provided $\{\tau_n\}$ is increasing and $\theta_n \in [0, 1)$ is chosen so that $\sum_{n=1}^{\infty} \theta_n \|x_n - x_{n-1}\|^2 < \infty$. In recent years, the inertial method was further studied intensively and it also has been used to solve some other optimization problems (see, for example, [16, 21, 32, 37, 38, 41, 45]).

In 2019, Tang [43] proposed the following inertial algorithm for solving the split common null-point problem in the case that H_1 is a Hilbert space, and F is a 2-uniformly convex and smooth Banach space: for given $x_1 \in H_1$ and $\alpha \in [0, 1)$, choose θ_n such that $0 < \theta_n < \bar{\theta}_n$, where

$$\bar{\theta}_n = \begin{cases} \min\{\alpha, \epsilon_n(\max\{\|x_n - x_{n-1}\|^2, \|x_n - x_{n-1}\|\})^{-1}\}, & x_n \neq x_{n-1}, \\ \alpha, & \text{otherwise,} \end{cases} \tag{1.8}$$

where $\{\epsilon_n\} \subset (0, \infty)$ such that $\sum_{n=1}^{\infty} \epsilon_n < \infty$. Compute the sequence $\{x_n\}$ generated iteratively by

$$\begin{cases} u_n = x_n + \theta_n(x_n - x_{n-1}), \\ x_{n+1} = R_r(I - \tau_n T^* J(I - Q_\mu) T)u_n \end{cases} \tag{1.9}$$

with the step size

$$\tau_n = \rho_n \frac{f(u_n)}{\|F(w_n)\|^2 + \|H(u_n)\|^2}, \tag{1.10}$$

where $\{\rho_n\} \subset (0, 4)$, $f(u_n) = \frac{1}{2} \|J(I - Q_\mu)Tu_n\|^2$, $F(u_n) = T^*J(I - Q_\mu)Tw_n$, and $H(u_n) = (I - R_r)u_n$. The weak convergence of the sequence $\{x_n\}$ is established without the prior knowledge of the operator norm of the bounded linear operator.

Recently, several inertial algorithms for solving the split common null-point problem in Hilbert spaces have been studied by many authors (see, for example, [8, 17, 24, 30, 51]). However, such methods have been studied in Banach spaces by a few authors (see, for example, [43, 44]).

Inspired and motivated by the works mentioned above, in this paper, we introduce two new inertial self-adaptive algorithms that are based on the classical inertial method and relaxed inertial method for finding a solution of the split common null-point problem in Banach spaces. The weak-convergence theorems are proved without the prior knowledge of the operator norm of the bounded linear operator. We provide numerical implementations to show that our algorithms are efficient and competitive with some related algorithms. Our results are new and complement some previous results in the literature.

The contributions of this paper can be summarized as follows:

- (1) The weak-convergence result of iterative scheme (1.9) of Tang [43] is proved in a Hilbert space and a 2-uniformly convex smooth Banach space where this result can only be implemented in ℓ_p for $p \in (1, 2]$ exclude the case of $p > 2$. This is limited in practical applications of such a method. In this paper, our results generalize the weak-convergence result of Tang [43] from between two of those spaces to p -uniformly convex and uniformly smooth Banach spaces, and as a result our results can be implemented in ℓ_p for $p > 1$.
- (2) Even though the step size of the iterative scheme (1.9) of Tang [43] is computed without the prior knowledge of the operator norms it requires calculation of $\|T^*J(I - Q_\mu)u_n\|^2$ and $\|(I - R_\tau)u_n\|^2$ in order to choose the step size τ_n . This could be computationally expensive during implementations, especially in the case where the resolvent of A and the metric resolvent of B are difficult to compute. In this paper, our step size τ_n defined by (3.3), is adaptively updated by a cheap computation without the prior knowledge of the operator norms and only requires us to compute one metric resolvent of B .
- (3) For the iterative scheme (1.5) of Suantai et al. [39], the choice of the sequence of step size depends on the bounded linear operator $\|T\|$, which is a difficult task during the implementation of the algorithm. In this paper, the choice of the sequence of our step size τ_n defined by (3.3), is independent of the operator norm of the bounded linear operator. As a result, we do not require to calculate the norm $\|T\|$ in order to choose the step size τ_n , which is easier to implement than such a method.
- (4) We use the inertial and relaxed inertial techniques to improve the rate of convergence of our algorithms that makes the algorithms converge faster and computationally more efficient for solving the split common null-point problem in Banach spaces. Note that these inertial techniques in this paper are studied outside Hilbert spaces for solving such a problem.
- (5) We present numerical results of our algorithms in Banach spaces to illustrate the efficiency and advantage over iterative scheme (1.5) of Suantai et al. [39] that gives the strong convergence and we also present several numerical results of our algorithms in finite-dimensional spaces. Moreover, we apply our results to data classifications for two datasets of heart diseases and diabetes mellitus.

Our paper is organized as the following four parts. In Sect. 2, we give some of the basic facts and notation that will be used in the paper. In Sect. 3, we propose two new inertial self-adaptive algorithms and prove our convergence results, and finally, in Sect. 4, we present several numerical results to verify the advantages and efficiency of the proposed algorithms.

2 Preliminaries

In this section, we give some definitions and preliminary results that will be used in proving the main results. Throughout this paper, we denote the set of real numbers and the set of positive integers by \mathbb{R} and \mathbb{N} , respectively. Let E be a real Banach space with norm $\|\cdot\|$ with its the dual space E^* . We denote $\langle u, j \rangle$ by the value of a functional j in E^* at $u \in E$, that is, $\langle u, j \rangle = j(u)$ for all $u \in E$. We write $u_n \rightarrow x$ to indicate that a sequence $\{u_n\}$ converges strongly to u . Similarly, $u_n \rightharpoonup u$ and $u_n \rightharpoonup^* u$ will symbolize the weak and weak* convergence, respectively. Let $S_E = \{u \in E : \|u\| = 1\}$ and $B_E = \{u \in E : \|u\| \leq 1\}$ be a unit sphere and unit ball of E , respectively.

Let $1 < q \leq 2 \leq p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. The *modulus of convexity* of E is the function $\delta_E : [0, 2] \rightarrow [0, 1]$ defined by

$$\delta_E(\epsilon) = \inf \left\{ 1 - \frac{\|u + v\|}{2} : u, v \in B_E, \|u - v\| \geq \epsilon \right\}.$$

The *modulus of smoothness* of E is the function $\rho_E : [0, \infty) \rightarrow [0, \infty)$ defined by

$$\rho_E(t) = \sup \left\{ \frac{\|u + tv\| + \|u - tv\|}{2} - 1 : u, v \in S_E \right\}.$$

Definition 2.1 A Banach space E is said to be:

- (1) *strictly convex* if $\frac{\|u+v\|}{2} < 1$ for all $u, v \in S_E$ and $u \neq v$;
- (2) *smooth* if $\lim_{t \rightarrow 0} \frac{\|u+tv\| - \|u\|}{t}$ exists for each $u, v \in S_E$;
- (3) *uniformly convex* if $\delta_E(\epsilon) > 0$ for all $\epsilon \in (0, 2]$;
- (4) *p-uniformly convex* if there is a $\kappa_p > 0$ such that $\delta_E(\epsilon) \geq \kappa_p \epsilon^p$ for all $\epsilon \in (0, 2]$;
- (5) *uniformly smooth* if $\lim_{t \rightarrow 0} \frac{\rho_E(t)}{t} = 0$;
- (6) *q-uniformly smooth* if there exists a $\kappa_q > 0$ such that $\rho_E(t) \leq \kappa_q t^q$ for all $t > 0$.

Remark 2.2 It is known that if E is uniformly convex, then E is reflexive and strictly convex; if E is uniformly smooth, then E is reflexive and smooth (see [2]). From the Definition 2.1, one can see that every p -uniformly convex (q -uniformly smooth) space is a uniformly convex (uniformly smooth) space. Moreover, it is also known that E is p -uniformly convex (q -uniformly smooth) if and only if E^* is q -uniformly smooth (p -uniformly convex) (see [2, 49]).

For the Lebesgue spaces L_p , sequence spaces l_p , and Sobolev spaces W_p^m , it is also known that [23, 50]

$$L_p(l_p) \text{ or } W_p^m \text{ is } \begin{cases} p\text{-uniformly smooth, 2-uniformly convex} & \text{for } 1 < p \leq 2, \\ p\text{-uniformly convex, 2-uniformly smooth} & \text{for } 2 \leq p < \infty. \end{cases} \tag{2.1}$$

For $p > 1$. The mapping $J_p : E \rightarrow 2^{E^*}$ defined by

$$J_p(u) = \{u^* \in E^* : \langle u, u^* \rangle = \|u\|^p, \|u^*\| = \|u\|^{p-1}\}$$

is called the *generalized duality mapping* of E . In particular, $J_2 = J$ is called the *normalized duality mapping* and if E is a Hilbert space, then $J_p = I$, where I is the identity mapping. The

duality mapping J_p of a smooth Banach space E is said to be *weakly sequentially continuous* if for any sequence $\{u_n\} \subset E$ such that $u_n \rightharpoonup u$ implies $J_p(u_n) \rightharpoonup^* J_p(u)$. For the generalized duality mapping, the following facts are known [2, 18, 34]:

- (i) J_p is homogeneous degree $p - 1$, that is, $J_p(\alpha u) = |\alpha|^{p-1} \text{sign}(\alpha) J_p(u)$ for all $u \in E$, $\alpha \in \mathbb{R}$. In particular, $J_p(-u) = -J_p(u)$ for all $u \in E$.
- (ii) If E is smooth, then J_p is monotone, that is, $\langle u - v, J_p(u) - J_p(v) \rangle \geq 0$ for all $u, v \in E$. Moreover, if E is strictly convex, then J_p is strictly monotone.
- (iii) If E is uniformly smooth, then J_p is single valued from E into E^* and it is uniformly continuous on bounded subsets of E .
- (iv) If E is reflexive, smooth, and strictly convex, then the inverse $J_p^{-1} = J_q^*$ is single valued, one-to-one, and surjective, where J_q^* is the duality mapping from E^* into E .

Lemma 2.3 ([49]) *If E is a q -uniformly smooth Banach space, then there is a constant $\kappa_q > 0$ such that*

$$\|u - v\|^q \leq \|u\|^q - q\langle v, J_q(u) \rangle + \kappa_q \|v\|^q, \quad \forall u, v \in E,$$

where κ_q is called the q -uniform smoothness coefficient of E .

Remark 2.4 The exact values of the constant κ_q can be found in [35, 50].

We next recall the definition of Bregman distance. Let E be a real smooth Banach space and f be a convex and Gâteaux differentiable function on E . The bifunction $D_f : E \times E \rightarrow [0, \infty)$ defined by

$$D_f(u, v) = f(u) - f(v) - \langle u - v, \nabla f(v) \rangle$$

is called the *Bregman distance with respect to f* . Note that the Bregman distance is not a metric due to its lack of symmetry and failure to satisfy the triangle inequality. If $f_p(x) = \frac{1}{p} \|x\|^p$ for $p > 1$, then $\nabla f = J_p$. Hence, we have the Bregman distance with respect to $f = f_p$ given by

$$\begin{aligned} D_{f_p}(u, v) &= \frac{1}{p} \|u\|^p - \frac{1}{p} \|v\|^p - \langle u - v, J_p(v) \rangle \\ &= \frac{1}{p} \|u\|^p + \frac{1}{q} \|v\|^p - \langle u, J_p(v) \rangle. \end{aligned}$$

Moreover, if $p = 2$, then $2D_{f_2}(u, v) = \|u\|^2 - \|v\|^2 - 2\langle u, J(v) \rangle = \phi(u, v)$, where ϕ is called the *Lyapunov function* studied in [4, 31]. Also, if E is a Hilbert space, then $\phi(u, v) = \|u - v\|^2$. The following properties of the Bregman distance are well known: for each $u, v, w \in E$,

$$D_{f_p}(u, v) = D_{f_p}(u, w) - D_{f_p}(v, w) + \langle u - v, J_p(w) - J_p(v) \rangle \tag{2.2}$$

and

$$D_{f_p}(u, v) + D_{f_p}(v, u) = \langle u - v, J_p(u) - J_p(v) \rangle. \tag{2.3}$$

For a p -uniformly convex space, it holds that [36]

$$\tau \|u - v\|^p \leq D_{f_p}(u, v) \leq \langle u - v, J_p(u) - J_p(v) \rangle, \tag{2.4}$$

where $\tau > 0$ is some fixed number.

Also, we define a function $V_{f_p} : E \times E^* \rightarrow [0, \infty)$ by

$$V_{f_p}(u, u^*) = \frac{1}{p} \|u\|^p - \langle u, u^* \rangle + \frac{1}{q} \|u^*\|^q$$

for all $u \in E$ and $u^* \in E^*$. Note that V_{f_p} is nonnegative, convex in the second variable and $V_{f_p}(u, u^*) = D_{f_p}(u, J_q(u^*))$ for all $u \in E$ and $u^* \in E^*$. Moreover, the following property is known:

$$V_{f_p}(u, u^*) + \langle J_p^{-1}(u^*) - u, v^* \rangle \leq V_{f_p}(u, u^* + v^*) \tag{2.5}$$

for all $u \in E$ and $u^*, v^* \in E^*$.

Let C be a nonempty, closed, and convex subset of a smooth, strictly convex, and reflexive Banach space. Then, for any $u \in E$, there exists a unique element $w \in C$ such that

$$\|u - w\| = \min_{v \in C} \|u - v\|.$$

The mapping P_C defined by $w = P_C(u)$ is called the *metric projection* of E onto C . We know the following property [36]:

$$\langle v - P_C(u), J_p(u - P_C(u)) \rangle \leq 0, \quad \forall v \in C.$$

Recall that the *Bregman projection with respect to f_p* is defined by

$$\Pi_C^{f_p}(u) = \operatorname{argmin}_{v \in C} D_{f_p}(u, v), \quad \forall u \in E.$$

If $p = 2$, then $\Pi_C^{f_p}$ becomes the generalized projection and denoted by Π_C . Also, in this case, if E is a Hilbert space, then Π_C becomes the metric projection denoted by P_C . We also know the following property [11]:

$$\langle v - \Pi_C^{f_p}(u), J_p(u) - J_p(\Pi_C^{f_p}(u)) \rangle \leq 0, \quad \forall v \in C. \tag{2.6}$$

Let C be a nonempty subset of E and $T : C \rightarrow C$ be a mapping. We denote the fixed-point set of T by $F(T) = \{u \in C : u = Tu\}$. Let $A : E \rightarrow 2^{E^*}$ be a set-valued mapping. The domain of A is denoted by $\mathcal{D}(A) = \{u \in E : Au \neq \emptyset\}$ and the range of A is also denoted by $\mathcal{R}(A) = \bigcup \{Au : u \in \mathcal{D}(A)\}$. The set of zeros of A is defined by $A^{-1}0 = \{u \in \mathcal{D}(A) : 0 \in Au\}$. It is known that $A^{-1}0$ is closed and convex (see [40]). A set-valued mapping A is said to be *monotone* if

$$\langle x - y, u - v \rangle \geq 0, \quad \forall x, y \in \mathcal{D}(A), u \in Ax \text{ and } v \in Ay.$$

A monotone operator A on E is said to be *maximal* if its graph is not properly contained in the graph of any other monotone operator on E .

Let E be a p -uniformly convex and uniformly smooth Banach space and $A : E \rightarrow 2^{E^*}$ be a maximal monotone operator. Following [10], for each $u \in E$ and $\tau > 0$, we define the *resolvent* of A by

$$R_\tau(u) = (J_p + \tau A)^{-1}J_p(u), \quad \forall u \in E.$$

One can see that $A^{-1}0 = F(R_\tau)$ for $\tau > 0$. We also know the following property [26]:

$$D_{J_p}(v, R_\tau(u)) + D_{J_p}(R_\tau(u), u) \leq D_{J_p}(v, u) \tag{2.7}$$

for all $u \in E$ and $v \in A^{-1}0$.

For each $u \in E$ and $\mu > 0$, we define the *metric resolvent* of A for $\mu > 0$ by

$$Q_\mu(u) = (I + \mu J_p^{-1}A)^{-1}(u), \quad \forall u \in E. \tag{2.8}$$

It is clear that in a Hilbert space, the metric resolvent operator is equivalent to the resolvent operator. From (2.8), one can see that $0 \in J_p(Q_\mu(u) - u) + \mu A Q_\mu(u)$ and $A^{-1}0 = F(Q_\mu)$ for $\mu > 0$. The monotonicity of A implies that

$$\langle Q_\mu(u) - Q_\mu(v), J_p(u - Q_\mu(u)) - J_p(v - Q_\mu(v)) \rangle \geq 0 \tag{2.9}$$

for all $u, v \in E$. If $A^{-1}0 \neq \emptyset$, then

$$\langle Q_\mu(u) - v, J_p(u - Q_\mu(u)) \rangle \geq 0 \tag{2.10}$$

for all $u \in E$ and $v \in A^{-1}0$ (see [9]). For any sequence $\{x_n\}$ in E , we see that

$$\begin{aligned} \|x_n - v\| \|x_n - Q_\mu(x_n)\|^{p-1} &\geq \langle x_n - v, J_p(x_n - Q_\mu(x_n)) \rangle \\ &\geq \langle x_n - Q_\mu(x_n), J_p(x_n - Q_\mu(x_n)) \rangle \\ &= \|x_n - Q_\mu(x_n)\|^p. \end{aligned} \tag{2.11}$$

This implies that $\|x_n - Q_\mu(x_n)\| \leq \|x_n - v\|$. If $\{x_n\}$ is bounded, then $\{x_n - Q_\mu(x_n)\}$ is also bounded.

Let E be a p -uniformly convex and uniformly smooth Banach space and $f : E \rightarrow \mathbb{R} \rightarrow (-\infty, +\infty]$ be a proper, convex, and lower semicontinuous function. The *subdifferential* of f at x is defined by

$$\partial f(x) = \{z \in E^* : f(x) + \langle y - x, z \rangle \leq f(y), \forall y \in E\}.$$

Let C be a closed and convex subset of E . The indicator function δ_C of C at x is defined by

$$\delta_C(x) = \begin{cases} 0, & \text{if } x \in C, \\ +\infty, & \text{if } x \notin C. \end{cases}$$

The subdifferentiable $\partial\delta_C$ is a maximal monotone operator since δ_C is a proper, convex, and lower semicontinuous function (see [33]). Moreover, we also know that

$$\partial\delta_C(x) = N_C(x) = \{z \in E^* : \langle y - x, z \rangle \leq 0, \forall y \in C\},$$

where N_C is the normal cone of C . In particular, if we define the resolvent of $\partial\delta_C$ for $\tau > 0$ by $R_\tau(u) = (J_p + \tau\partial\delta_C)^{-1}J_p(u)$ for all $u \in E$, then $R_\tau = \Pi_C^{f_p}$, where $\Pi_C^{f_p}$ is the Bregman projection with respect to f_p (see [48]). Moreover, we also have $(\partial\delta_C)^{-1}0 = C$. Also, if we define the metric resolvent of $\partial\delta_C$ for $\mu > 0$ by $Q_\mu(u) = (I + \mu J_p^{-1}\partial\delta_C)^{-1}(u)$ for all $u \in E$, then

$$\begin{aligned} z = Q_\mu(u) &\Leftrightarrow 0 \in J_p(z - u) + \mu Az \\ &\Leftrightarrow \frac{J_p(u - z)}{\mu} \in Az = \partial\delta_C(z) = N_C(z) \\ &\Leftrightarrow \langle y - z, J_p(u - z) \rangle \leq 0, \quad \forall y \in C \\ &\Leftrightarrow z = P_C(u), \end{aligned}$$

where P_C is the metric projection of E onto C .

Throughout this paper, we adopt the notation $[a]_+ := \max\{a, 0\}$, where $a \in \mathbb{R}$.

Lemma 2.5 ([6]) *Let $\{\varphi_n\}$, $\{\alpha_n\}$, and $\{\beta_n\}$ be three nonnegative real sequences such that*

$$\varphi_{n+1} \leq \varphi_n + \alpha_n(\varphi_n - \varphi_{n-1}) + \beta_n, \quad \forall n \geq 1,$$

with $\sum_{n=1}^\infty \beta_n < \infty$ and there exists a real number α such that $0 \leq \alpha_n \leq \alpha < 1$ for all $n \in \mathbb{N}$. Then, the following results hold:

- (i) $\sum_{n=1}^\infty [\varphi_n - \varphi_{n-1}]_+ < \infty$;
- (ii) There exists $\varphi^* \in [0, \infty)$ such that $\lim_{n \rightarrow \infty} \varphi_n = \varphi^*$.

Lemma 2.6 ([42]) *Assume that $\{s_n\}$ and $\{t_n\}$ are two nonnegative real sequences such that $s_{n+1} \leq s_n + t_n$ for all $n \geq 1$. If $\sum_{n=1}^\infty t_n < \infty$, then $\lim_{n \rightarrow \infty} s_n$ exists.*

3 Main results

In this paper, we propose two weakly convergent inertial self-adaptive algorithms to solve the split common null-point problem in Banach spaces. In what follows, we denote J_p^E and $J_q^{E^*}$ by the generalized duality mapping of E into E^* and the duality mapping of E^* into E , respectively, where $1 < q \leq 2 \leq p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$.

In order to prove the results, the following assumptions are needed in the following.

- (A1) Let E be a p -uniformly convex and uniformly smooth Banach space and F be a uniformly convex and smooth Banach space.
- (A2) Let $A : E \rightarrow 2^{E^*}$ and $B : F \rightarrow 2^{F^*}$ be maximal monotone operators.
- (A3) Let $T : E \rightarrow F$ be a bounded linear operator with $T \neq 0$ and $T^* : F^* \rightarrow E^*$ be the adjoint operator of T .
- (A4) Let R_τ be a resolvent operator associated with A for $\tau > 0$ and Q_μ be a metric resolvent associated with B for $\mu > 0$.
- (A5) The set solution $\Omega := A^{-1}0 \cap T^{-1}(B^{-1}0) \neq \emptyset$.

The following conditions are also assumed:

(C1) Let $\{\alpha_n\} \subset (0, 1]$ with $\liminf_{n \rightarrow \infty} \alpha_n > 0$;

(C2) Let $\{\mu_n\} \subset (0, \infty)$ with $\liminf_{n \rightarrow \infty} \mu_n > 0$.

The first algorithm is stated as follows:

Algorithm 1 (Inertial self-adaptive algorithm for the split common null-point problem)

Step 0. Given $\tau_1 > 0$, $\beta \in (0, 1)$ and $\mu \in (0, \frac{q}{kq})$. Choose $\{s_n\} \subset [0, \infty)$ such that $\sum_{n=1}^{\infty} s_n < \infty$ and $\{\beta_n\} \subset (0, \infty)$ such that $\sum_{n=1}^{\infty} \beta_n < \infty$. Let $x_0, x_1 \in E$ be arbitrary and calculate x_{n+1} as follows:

Step 1. Given the iterates x_{n-1} and x_n ($n \geq 1$). Choose θ_n such that $0 \leq \theta_n \leq \bar{\theta}_n$, where

$$\bar{\theta}_n = \begin{cases} \min\{\beta, \frac{\beta_n}{\|J_p^E(x_n) - J_p^E(x_{n-1})\|^q}, \frac{\beta_n}{D_{f_p}(x_n, x_{n-1})}\}, & \text{if } x_n \neq x_{n-1}, \\ \beta, & \text{otherwise.} \end{cases} \tag{3.1}$$

Step 2. Compute

$$\begin{cases} u_n = J_q^{E*}(J_p^E(x_n) + \theta_n(J_p^E(x_n) - J_p^E(x_{n-1}))), \\ y_n = J_q^{E*}(J_p^E(u_n) - \tau_n T^* J_p^F(I - Q_{\mu_n})Tu_n), \\ x_{n+1} = J_q^{E*}((1 - \alpha_n)J_p^E(u_n) + \alpha_n J_p^E(R_{\tau_n}y_n)), \end{cases} \tag{3.2}$$

where

$$\tau_{n+1} = \begin{cases} \min\{(\frac{\mu\|(I - Q_{\mu_n})Tu_n\|^p}{\|T^* J_p^F(I - Q_{\mu_n})Tu_n\|^q})^{\frac{1}{q-1}}, \tau_n + s_n\}, & \text{if } T^* J_p^F(I - Q_{\mu_n})Tu_n \neq 0, \\ \tau_n + s_n, & \text{otherwise.} \end{cases} \tag{3.3}$$

Remark 3.1 If $x_{n+1} = y_n = u_n$ for some n , then y_n is a solution in Ω . Indeed, if $y_n = u_n$, we see that $y_n = J_q^{E*}(J_p^E(y_n) - \tau_n T^* J_p^F(I - Q_{\mu_n})Ty_n)$ for $\tau_n > 0$. This implies that $(I - Q_{\mu_n})Ty_n = 0$, that is, $Ty_n = Q_{\mu_n}Ty_n$. In addition, if $x_{n+1} = y_n$, then $y_n = J_q^{E*}((1 - \alpha_n)J_p^E(y_n) + \alpha_n J_p^E(R_{\tau_n}y_n))$. This implies that $y_n = R_{\tau_n}y_n$. Now, since $y_n = R_{\tau_n}y_n$ and $Ty_n = Q_{\mu_n}Ty_n$, we have $y_n \in A^{-1}0$ and $y_n \in T^{-1}(B^{-1}0)$. Therefore, $y_n \in \Omega := A^{-1}0 \cap T^{-1}(B^{-1}0)$.

Remark 3.2 From (3.1), we observe that $0 \leq \theta_n \leq \beta < 1$ for all $n \geq 1$. Also, we obtain $\theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q \leq \beta_n$ and $\theta_n D_{f_p}(x_n, x_{n-1}) \leq \beta_n$ for all $n \geq 1$. Since $\sum_{n=1}^{\infty} \beta_n < \infty$, we have

$$\sum_{n=1}^{\infty} \theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \theta_n D_{f_p}(x_n, x_{n-1}) < \infty. \tag{3.4}$$

Lemma 3.3 Let $\{\tau_n\}$ be a sequence generated by (3.3). Then, we have $\lim_{n \rightarrow \infty} \tau_n = \tau$, where

$$\tau \in \left[\min\left\{\left(\frac{\mu}{\|T\|^q}\right)^{\frac{1}{q-1}}, \tau_1\right\}, \tau_1 + s \right] \quad \text{and} \quad s = \sum_{n=1}^{\infty} s_n.$$

Proof In the case of $T^*J_p^F(I - Q_{\mu_n})Tu_n \neq 0$, we see that

$$\begin{aligned} & \frac{\mu \|(I - Q_{\mu_n})Tu_n\|^p}{\|T^*J_p^F(I - Q_{\mu_n})Tu_n\|^q} \\ & \geq \frac{\mu \|(I - Q_{\mu_n})Tu_n\|^p}{\|T^*\|^q \|J_p^F(I - Q_{\mu_n})Tu_n\|^q} \\ & = \frac{\mu \|(I - Q_{\mu_n})Tu_n\|^p}{\|T\|^q \|(I - Q_{\mu_n})Tu_n\|^{(p-1)q}} \\ & = \frac{\mu}{\|T\|^q}. \end{aligned} \tag{3.5}$$

By the definition of τ_n and induction, we have

$$\tau_n \leq \tau_{n-1} + s_{n-1} \leq \tau_{n-2} + s_{n-2} + s_{n-1} \leq \dots \leq \tau_1 + \sum_{i=1}^{n-1} s_i \leq \tau_1 + \sum_{n=1}^{\infty} s_n.$$

Thus, $\tau_n \leq \tau_1 + \sum_{n=1}^{\infty} s_n$ for all $n \geq 1$. From (3.5), we see that

$$\begin{aligned} \tau_{n+1} &= \min \left\{ \left(\frac{\mu \|(I - Q_{\mu_n})Tu_n\|^p}{\|T^*J_p^F(I - Q_{\mu_n})Tu_n\|^q} \right)^{\frac{1}{q-1}}, \tau_n + s_n \right\} \\ &\geq \min \left\{ \left(\frac{\mu}{\|T\|^q} \right)^{\frac{1}{q-1}}, \tau_n + s_n \right\} \\ &\geq \min \left\{ \left(\frac{\mu}{\|T\|^q} \right)^{\frac{1}{q-1}}, \tau_n \right\} \\ &\geq \dots \geq \min \left\{ \left(\frac{\mu}{\|T\|^q} \right)^{\frac{1}{q-1}}, \tau_1 \right\}. \end{aligned}$$

Hence $\tau_n \geq \min\{(\frac{\mu}{\|T\|^q})^{\frac{1}{q-1}}, \tau_1\}$ for all $n \geq 1$. Therefore, $\min\{(\frac{\mu}{\|T\|^q})^{\frac{1}{q-1}}, \tau_1\} \leq \tau_n \leq \tau_1 + \sum_{n=1}^{\infty} s_n$ for all $n \geq 1$. Since $\tau_{n+1} \leq \tau_n + s_n$ for all $n \geq 1$, we have $\lim_{n \rightarrow \infty} \tau_n$ exists by Lemma 2.6. In this case, we denote $\tau = \lim_{n \rightarrow \infty} \tau_n$. Obviously, $\tau \in [\min\{(\frac{\mu}{\|T\|^q})^{\frac{1}{q-1}}, \tau_1\}, \tau_1 + s]$. □

Remark 3.4 The adaptive step size τ_n generated by (3.3) is different from many adaptive step sizes as studied in [43, 44]. Note that τ_n is allowed to increase when the iteration increases. Therefore, it reduces the dependence on the initial step size τ_1 . Since $\sum_{n=1}^{\infty} s_n < \infty$, we have $\lim_{n \rightarrow \infty} s_n = 0$. As a result, τ_n may not increase when n is large.

Lemma 3.5 *Let $\{x_n\}$ be a sequence generated by Algorithm 1. Then, for each $n \geq 1$, the following inequality holds for all $v \in \Omega$:*

$$D_{f_p}(v, x_{n+1}) \leq D_{f_p}(v, x_n) + \theta_n(D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})) + \xi_n(p, q) - \delta_n(p, q),$$

where $\xi_n(p, q) := \theta_n D_{f_p}(x_n, x_{n-1}) + \frac{\kappa q \theta_n^q}{q} \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q$, $\delta_n(p, q) := \alpha_n \tau_n (1 - \frac{\kappa q \mu}{q} (\frac{\tau_n}{\tau_{n+1}})^{q-1}) \|w_n\|^p + \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n)$ and $w_n := Tu_n - Q_{\mu_n} Tu_n$.

Proof Let $v \in \Omega := A^{-1}0 \cap T^{-1}(B^{-1}0)$. From Lemma 2.3, we have

$$\begin{aligned}
 D_{f_p}(v, y_n) &= V_{f_p}(v, J_p^E(u_n) - \tau_n T^* J_p^F(I - Q_{\mu_n})Tu_n) \\
 &= \frac{1}{p} \|v\|^p - \langle v, J_p^E(u_n) - \tau_n T^* J_p^F(w_n) \rangle + \frac{1}{q} \|J_p^E(u_n) - \tau_n T^* J_p^F(w_n)\|^q \\
 &= \frac{1}{p} \|v\|^p - \langle v, J_p^E(u_n) \rangle + \tau_n \langle Tv, J_p^F(w_n) \rangle + \frac{1}{q} \|J_p^E(u_n) - \tau_n T^* J_p^F(w_n)\|^q \\
 &\leq \frac{1}{p} \|v\|^p - \langle v, J_p^E(u_n) \rangle + \tau_n \langle Tv, J_p^F(w_n) \rangle + \frac{1}{q} \|J_p^E(u_n)\|^q - \tau_n \langle u_n, T^* J_p^F(w_n) \rangle \\
 &\quad + \frac{\kappa_q \tau_n^q}{q} \|T^* J_p^F(w_n)\|^q \\
 &= \frac{1}{p} \|v\|^p - \langle v, J_p^E(u_n) \rangle + \frac{1}{q} \|u_n\|^q + \tau_n \langle Tv - Tu_n, J_p^F(w_n) \rangle + \frac{\kappa_q \tau_n^q}{q} \|T^* J_p^F(w_n)\|^q \\
 &= D_{f_p}(v, u_n) + \tau_n \langle Tv - Tu_n, J_p^F(w_n) \rangle + \frac{\kappa_q \tau_n^q}{q} \|T^* J_p^F(w_n)\|^q. \tag{3.6}
 \end{aligned}$$

Note that $w_n := Tu_n - Q_{\mu_n}Tu_n$ and $v \in A^{-1}0 \cap T^{-1}(B^{-1}0)$, we have $v \in A^{-1}0$ and $Tv \in B^{-1}0$. It then follows from (2.10) that

$$\begin{aligned}
 \langle Tv - Tu_n, J_p^F(w_n) \rangle &= \underbrace{\langle Tv - Q_{\mu_n}Tu_n, J_p^F(w_n) \rangle}_{\leq 0} - \langle Tu_n - Q_{\mu_n}Tu_n, J_p^F(w_n) \rangle \\
 &\leq -\|Tu_n - Q_{\mu_n}Tu_n\|^p = -\|w_n\|^p. \tag{3.7}
 \end{aligned}$$

From the definition of τ_{n+1} , we have

$$\|T^* J_p^F(I - Q_{\mu_n})Tu_n\|^q \leq \frac{\mu}{\tau_{n+1}^{q-1}} \|(I - Q_{\mu_n})Tu_n\|^p. \tag{3.8}$$

Combining (3.6), (3.7), and (3.8), we obtain

$$D_{f_p}(v, y_n) \leq D_{f_p}(v, u_n) - \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p. \tag{3.9}$$

Now, we estimate $D_{f_p}(v, u_n)$. From Lemma 2.3, we have

$$\begin{aligned}
 D_{f_p}(v, u_n) &= V_{f_p}(v, J_p^E(x_n) + \theta_n(J_p^E(x_n) - J_p^E(x_{n-1}))) \\
 &= \frac{1}{p} \|v\|^p - \langle v, J_p^E(x_n) + \theta_n(J_p^E(x_n) - J_p^E(x_{n-1})) \rangle \\
 &\quad + \frac{1}{q} \|J_p^E(x_n) + \theta_n(J_p^E(x_n) - J_p^E(x_{n-1}))\|^q \\
 &\leq \frac{1}{p} \|v\|^p - \langle v, J_p^E(x_n) \rangle - \theta_n \langle v, J_p^E(x_n) - J_p^E(x_{n-1}) \rangle \\
 &\quad + \frac{1}{q} \|J_p^E(x_n)\|^q + \theta_n \langle x_n, J_p^E(x_n) - J_p^E(x_{n-1}) \rangle + \frac{\kappa_q}{q} \|\theta_n(J_p^E(x_n) - J_p^E(x_{n-1}))\|^q \\
 &= \frac{1}{p} \|v\|^p - \langle v, J_p^E(x_n) \rangle + \frac{1}{q} \|x_n\|^p + \frac{\kappa_q \theta_n^q}{q} \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q
 \end{aligned}$$

$$\begin{aligned}
 & + \theta_n \langle x_n - v, J_p^E(x_n) - J_p^E(x_{n-1}) \rangle \\
 = & D_{f_p}(v, x_n) + \frac{\kappa_q \theta_n^q}{q} \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q + \theta_n \langle x_n - v, J_p^E(x_n) - J_p^E(x_{n-1}) \rangle. \tag{3.10}
 \end{aligned}$$

We observe that

$$\theta_n \langle x_n - v, J_p^E(x_n) - J_p^E(x_{n-1}) \rangle = \theta_n D_{f_p}(v, x_n) - \theta_n D_{f_p}(v, x_{n-1}) + \theta_n D_{f_p}(x_n, x_{n-1}). \tag{3.11}$$

Combining (3.10) and (3.11), we obtain

$$\begin{aligned}
 D_{f_p}(v, u_n) \leq & D_{f_p}(v, x_n) + \theta_n (D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})) + \theta_n D_{f_p}(x_n, x_{n-1}) \\
 & + \frac{\kappa_q \theta_n^q}{q} \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q. \tag{3.12}
 \end{aligned}$$

Then, from (2.7) and (3.10), we obtain

$$\begin{aligned}
 D_{f_p}(v, x_{n+1}) \leq & (1 - \alpha_n) D_{f_p}(v, u_n) + \alpha_n D_{f_p}(v, R_{\tau_n} y_n) \\
 \leq & (1 - \alpha_n) D_{f_p}(v, u_n) + \alpha_n D_{f_p}(v, y_n) - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \\
 \leq & (1 - \alpha_n) D_{f_p}(v, u_n) + \alpha_n \left[D_{f_p}(v, u_n) - \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p \right] \\
 & - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \\
 = & D_{f_p}(v, u_n) - \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \\
 \leq & D_{f_p}(v, x_n) + \theta_n (D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})) + \theta_n D_{f_p}(x_n, x_{n-1}) \\
 & + \frac{\kappa_q \theta_n^q}{q} \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q \\
 & - \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n). \tag{3.13}
 \end{aligned}$$

From the definitions of $\xi_n(p, q)$ and $\delta_n(p, q)$, then (3.13) can be written in a short form as follows:

$$D_{f_p}(v, x_{n+1}) \leq D_{f_p}(v, x_n) + \theta_n (D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})) + \xi_n(p, q) - \delta_n(p, q).$$

Thus, this lemma is proved. □

Theorem 3.6 *Let $\{x_n\}$ be a sequence generated by Algorithm 1. Suppose, in addition, that J_p^E is weakly sequentially continuous on E . Then, $\{x_n\}$ converges weakly to a point in Ω .*

Proof Using the fact that $\lim_{n \rightarrow \infty} \tau_n$ exists and $\mu \in (0, \frac{q}{\kappa_q})$, we have

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) = \frac{\mu}{q} \left(\frac{q}{\mu} - \kappa_q \right) > 0.$$

Then, there exists $n_0 \in \mathbb{N}$ such that

$$1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} > 0, \quad \forall n \geq n_0,$$

and, in consequence,

$$\delta_n(p, q) := \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p + \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) > 0, \quad \forall n \geq n_0.$$

Then, from Lemma 3.5, we can deduce that

$$D_{f_p}(v, x_{n+1}) \leq D_{f_p}(v, x_n) + \theta_n (D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})) + \xi_n(p, q).$$

Since $\xi_n(p, q) := \theta_n D_{f_p}(x_n, x_{n-1}) + \frac{\kappa_q \theta_n^q}{q} \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q$, it follows from (3.4) that

$$\sum_{n=1}^{\infty} \xi_n(p, q) \leq \sum_{n=1}^{\infty} \left(\theta_n D_{f_p}(x_n, x_{n-1}) + \frac{\kappa_q \theta_n^q}{q} \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q \right) < \infty. \tag{3.14}$$

From (3.14), we also have $\lim_{n \rightarrow \infty} \xi_n(p, q) = 0$. From Lemma 2.5, we can conclude that $\lim_{n \rightarrow \infty} D_{f_p}(v, x_n)$ exists and

$$\sum_{n=1}^{\infty} [D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})]_+ < \infty.$$

Thus, we have $\{D_{f_p}(v, x_n)\}$ is bounded and so $\{x_n\}$ is also bounded by (2.4). Moreover, we obtain

$$\lim_{n \rightarrow \infty} [D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})]_+ = 0. \tag{3.15}$$

From Lemma 3.5, we see that

$$\delta_n(p, q) \leq D_{f_p}(v, x_n) - D_{f_p}(v, x_{n+1}) + \theta_n (D_{f_p}(v, x_n) - D_{f_p}(v, x_{n-1})) + \xi_n(p, q).$$

Since $\lim_{n \rightarrow \infty} D_{f_p}(v, x_n)$ exists and $\lim_{n \rightarrow \infty} \xi_n(p, q) = 0$, we have

$$\lim_{n \rightarrow \infty} \delta_n(p, q) = \lim_{n \rightarrow \infty} \left[\alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p + \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \right] = 0.$$

Consequently,

$$\lim_{n \rightarrow \infty} \|w_n\| = \lim_{n \rightarrow \infty} \|Tu_n - Q_{\mu_n} Tu_n\| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} D_{f_p}(y_n, R_{\tau_n} y_n) = 0.$$

By the continuity of J_p^F , we have

$$\lim_{n \rightarrow \infty} \|J_p^F(w_n)\| = \lim_{n \rightarrow \infty} \|J_p^F(Tu_n - Q_{\mu_n} Tu_n)\| = 0. \tag{3.16}$$

Moreover, we have

$$\lim_{n \rightarrow \infty} \|y_n - R_{\tau_n} y_n\| = 0. \tag{3.17}$$

By the definition of y_n , the continuity of T^* , and from (3.16), we have

$$\lim_{n \rightarrow \infty} \|J_p^E(y_n) - J_p^E(u_n)\| = \lim_{n \rightarrow \infty} \tau_n \|T^* J_p^F(I - Q_{\mu_n}) T u_n\| = 0. \tag{3.18}$$

On the other hand, by the definition of u_n and from (3.4), we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} \|J_p^E(u_n) - J_p^E(x_n)\|^q &= \lim_{n \rightarrow \infty} \theta_n^q \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q \\ &\leq \lim_{n \rightarrow \infty} \beta^{q-1} \theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\|^q \\ &= 0. \end{aligned}$$

Thus,

$$\lim_{n \rightarrow \infty} \|J_p^E(u_n) - J_p^E(x_n)\| = 0. \tag{3.19}$$

From (3.19), we also have $\lim_{n \rightarrow \infty} \|u_n - x_n\| = 0$ by the uniform continuity of $J_q^{E^*}$. Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightarrow w \in E$ and so $u_{n_k} \rightarrow w$. Put $v_n = R_{\tau_n} y_n$ for all $n \in \mathbb{N}$. From (3.17) and (3.18), we see that

$$\begin{aligned} \|J_p^E(u_n) - J_p^E(v_n)\| &= \|J_p^E(u_n) - J_p^E(R_{\tau_n} y_n)\| \\ &\leq \|J_p^E(u_n) - J_p^E(y_n)\| + \|J_p^E(y_n) - J_p^E(R_{\tau_n} y_n)\| \\ &\rightarrow 0. \end{aligned} \tag{3.20}$$

Consequently, $\lim_{n \rightarrow \infty} \|u_n - v_n\| = 0$. Thus,

$$\|v_n - x_n\| \leq \|v_n - u_n\| + \|u_n - x_n\| \rightarrow 0.$$

Using the above inequality, we also obtain $v_{n_k} \rightarrow w$. Since R_{τ_n} is the resolvent of A for $\tau_n > 0$, we have

$$\begin{aligned} v_n = R_{\tau_n} (J_q^{E^*} (J_p^E(u_n) - \tau_n T^* J_p^F(u_n))) &\Leftrightarrow (J_p^E + \tau_n A)^{-1} (J_p^E(u_n) - \tau_n T^* J_p^F(u_n)) \\ &\Leftrightarrow J_p^E(u_n) - \tau_n T^* J_p^F(u_n) \in J_p^E(v_n) + \tau_n A v_n \\ &\Leftrightarrow \frac{1}{\tau_n} (J_p^E(u_n) - J_p^E(v_n) - \tau_n T^* J_p^F(w_n)) \in A v_n. \end{aligned}$$

Replacing n by n_k and using the fact that A is monotone, thus

$$\left\langle s - v_{n_k}, s^* - \frac{1}{\tau_{n_k}} (J_p^E(u_{n_k}) - J_p^E(v_{n_k}) - \tau_{n_k} T^* J_p^F(w_{n_k})) \right\rangle \geq 0$$

for all $(s, s^*) \in A$. Now, T^* is continuous, which is due to the fact that T^* is a bounded and linear operator. Then, from (3.16), (3.20), and $\lim_{k \rightarrow \infty} \tau_{n_k} = \tau > 0$, we obtain $\langle s - w, s^* -$

$0) \geq 0$ for all $(s, s^*) \in A$. Note that A is maximal monotone, thus $w \in A^{-1}0$. On the other hand, we know that T is also continuous. This fact, together with $\|u_{n_k} - x_{n_k}\| \rightarrow 0$ and $\|Tu_{n_k} - Q_{\mu_{n_k}}Tu_{n_k}\| \rightarrow 0$, means that we have $Tu_{n_k} \rightharpoonup Tw$ and $Q_{\mu_{n_k}}u_{n_k} \rightharpoonup Tw$. Since Q_{μ_n} is the metric resolvent of B for $\mu_n > 0$, we have

$$\frac{J_p^E(Tu_n - Q_{\mu_n}Tu_n)}{\mu_n} \in BQ_{\mu_n}Tu_n$$

for all $n \in \mathbb{N}$. Replacing n by n_k , it then follows from the monotonicity of B that

$$\left\langle u - Q_{\mu_{n_k}}Tu_{n_k}, u^* - \frac{J_p^E(Tu_{n_k} - Q_{\mu_{n_k}}Tu_{n_k})}{\mu_{n_k}} \right\rangle \geq 0$$

for all $(u, u^*) \in B$. Then, from (3.16) and $\liminf_{k \rightarrow \infty} \mu_{n_k} > 0$, we obtain $\langle u - Tw, u^* - 0 \rangle \geq 0$ for all $(u, u^*) \in B$. Note that B is maximal monotone, thus $Tw \in B^{-1}0$ and so $w \in T^{-1}(B^{-1}0)$. We thus obtain $w \in \Omega := A^{-1}0 \cap T^{-1}(B^{-1}0)$. In order to prove the weak convergence of the sequence $\{x_n\}$, it is sufficient to show that $\{x_n\}$ has a unique weak limit point in Ω . In this case, we can assume that $\{x_{m_k}\}$ is another subsequence of $\{x_n\}$ such that $x_{m_k} \rightharpoonup w' \in \Omega$. Note that $x_{n_k} \rightharpoonup w \in \Omega$. Indeed, suppose by contradiction that with $w' \neq w$. Since $\lim_{n \rightarrow \infty} D_{f_p}(v, x_n)$ exists for any $v \in \Omega$, it then follows from (2.2) and the weak sequential continuity of J_p^E that

$$\begin{aligned} \lim_{n \rightarrow \infty} D_{f_p}(w, x_n) &= \lim_{k \rightarrow \infty} D_{f_p}(w, x_{m_k}) = \liminf_{k \rightarrow \infty} D_{f_p}(w, x_{m_k}) \\ &= \liminf_{k \rightarrow \infty} (D_{f_p}(w, w') + D_{f_p}(w', x_{m_k}) + \langle w - w', J_p^E(w') - J_p^E(x_{m_k}) \rangle) \\ &\geq \liminf_{k \rightarrow \infty} D_{f_p}(w, w') + \liminf_{k \rightarrow \infty} D_{f_p}(w', x_{m_k}) \\ &\quad + \liminf_{k \rightarrow \infty} \langle w - w', J_p^E(w') - J_p^E(x_{m_k}) \rangle \\ &> \liminf_{k \rightarrow \infty} D_{f_p}(w', x_{m_k}) \\ &= \lim_{n \rightarrow \infty} D_{f_p}(w', x_n). \end{aligned} \tag{3.21}$$

In the same way as above, we can show that

$$\lim_{n \rightarrow \infty} D_{f_p}(w', x_n) > \lim_{n \rightarrow \infty} D_{f_p}(w, x_n),$$

which is a contradiction with (3.21). Hence, $w = w'$ and therefore, the sequence $\{x_n\}$ converges weakly to a point in Ω . This finishes the proof. \square

Next, we present a second algorithm that is slightly different from the first proposed algorithm.

Algorithm 2 (Relaxed inertial self-adaptive algorithm for the split common null-point problem)

Step 0. Given $\tau_1 > 0$, $\beta \in (0, 1)$ and $\mu \in (0, \frac{q}{kq})$. Choose $\{s_n\} \subset [0, \infty)$ such that $\sum_{n=1}^\infty s_n < \infty$ and $\{\beta_n\} \subset (0, \infty)$ such that $\sum_{n=1}^\infty \beta_n < \infty$. Let $x_0, x_1 \in E$ be arbitrary and calculate x_{n+1} as follows:

Step 1. Given the iterates x_{n-1} and x_n ($n \geq 1$). Choose θ_n such that $0 \leq \theta_n \leq \bar{\theta}_n$, where

$$\bar{\theta}_n = \begin{cases} \min\{\beta, \frac{\beta_n}{\|J_p^E(x_n) - J_p^E(x_{n-1})\|}\}, & \text{if } x_n \neq x_{n-1}, \\ \beta, & \text{otherwise.} \end{cases} \tag{3.22}$$

Step 2. Compute

$$\begin{cases} u_n = J_q^{E*}(J_p^E(x_n) + \theta_n(J_p^E(x_{n-1}) - J_p^E(x_n))), \\ y_n = J_q^{E*}(J_p^E(u_n) - \tau_n T^* J_p^E(I - Q_{\mu_n}) T u_n), \\ x_{n+1} = J_q^{E*}((1 - \alpha_n)J_p^E(u_n) + \alpha_n J_p^E(R_{\tau_n} y_n)), \end{cases} \tag{3.23}$$

where τ_n is defined the same as in (3.3).

Remark 3.7 It should be noted that Algorithm 2 is slightly different from Algorithm 1 but $\bar{\theta}_n$ of this algorithm is simpler to compute than $\bar{\theta}_n$ of Algorithm 1, that is, it is chosen without any prior knowledge of the Bregman distance D_{f_p} at two points x_n and x_{n-1} , which is flexible and easy to implement in solving the problem. This is why we call the technique proposed in this case a “relaxed inertial algorithm”.

Remark 3.8 From (3.22), we observe that $0 \leq \theta_n \leq \beta < 1$ for all $n \geq 1$. Also, we obtain $\theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\| \leq \beta_n$ for all $n \geq 1$. Since $\sum_{n=1}^\infty \beta_n < \infty$, we have $\sum_{n=1}^\infty \theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\| < \infty$ and so

$$\lim_{n \rightarrow \infty} \theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\| = 0.$$

Theorem 3.9 Let $\{x_n\}$ be a sequence generated by Algorithm 2. Suppose, in addition, that J_p^E is weakly sequentially continuous on E . Then, $\{x_n\}$ converges weakly to a point in Ω .

Proof Let $v \in \Omega := A^{-1}0 \cap T^{-1}(B^{-1}0)$. Similarly, by using the same argument as in the proof of Theorem 3.6, we have

$$D_{f_p}(v, y_n) \leq D_{f_p}(v, u_n) - \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}}\right)^{q-1}\right) \|w_n\|^p. \tag{3.24}$$

By the definition of u_n in (3.23), we see that

$$\begin{aligned} D_{f_p}(v, u_n) &= D_{f_p}(v, J_q^{E*}((1 - \theta_n)J_p^E(x_n) + \theta_n J_p^E(x_{n-1}))) \\ &\leq (1 - \theta_n)D_{f_p}(v, x_n) + \theta_n D_{f_p}(v, x_{n-1}). \end{aligned} \tag{3.25}$$

Thus, we have

$$\begin{aligned} D_{f_p}(v, x_{n+1}) &\leq (1 - \alpha_n)D_{f_p}(v, u_n) + \alpha_n D_{f_p}(v, y_n) - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \\ &\leq (1 - \alpha_n)D_{f_p}(v, u_n) + \alpha_n \left[D_{f_p}(v, u_n) - \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}}\right)^{q-1}\right) \|w_n\|^p \right] \\ &\quad - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \end{aligned}$$

$$\begin{aligned}
 &= D_{f_p}(v, u_n) - \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \\
 &\leq (1 - \theta_n) D_{f_p}(v, x_n) + \theta_n D_{f_p}(v, x_{n-1}) - \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p \\
 &\quad - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n). \tag{3.26}
 \end{aligned}$$

From Theorem 3.6, we know that

$$\delta_n(p, q) := \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p + \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) > 0, \quad \forall n \geq n_0.$$

Thus, we can deduce that

$$\begin{aligned}
 D_{f_p}(v, x_{n+1}) &\leq (1 - \theta_n) D_{f_p}(v, x_n) + \theta_n D_{f_p}(v, x_{n-1}) \\
 &\leq \max \{ D_{f_p}(v, x_n), D_{f_p}(v, x_{n-1}) \} \\
 &\leq \dots \max \{ D_{f_p}(v, x_{n_0}), D_{f_p}(v, x_{n_0-1}) \}.
 \end{aligned}$$

Hence, $\{D_{f_p}(v, x_n)\}$ is bounded. From (2.4), we also obtain $\{x_n\}$ is bounded. From (3.26), we have

$$\begin{aligned}
 D_{f_p}(v, x_{n+1}) &\leq D_{f_p}(v, x_n) + \theta_n [D_{f_p}(v, x_{n-1}) - D_{f_p}(v, x_n)]_+ \\
 &\quad - \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p \\
 &\quad - \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n), \tag{3.27}
 \end{aligned}$$

which implies that

$$D_{f_p}(v, x_{n+1}) \leq D_{f_p}(v, x_n) + \theta_n [D_{f_p}(v, x_{n-1}) - D_{f_p}(v, x_n)]_+ \tag{3.28}$$

for all $n \geq n_0$. Using (2.2), we see that

$$\begin{aligned}
 \theta_n [D_{f_p}(v, x_{n-1}) - D_{f_p}(v, x_n)]_+ &= -\theta_n D_{f_p}(x_{n-1}, x_n) + \theta_n (v - x_{n-1}, J_p^E(x_n) - J_p^E(x_{n-1})) \\
 &\leq \theta_n (v - x_{n-1}, J_p^E(x_n) - J_p^E(x_{n-1})) \\
 &\leq \theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\| M,
 \end{aligned}$$

where $M := \sup_{n \geq n_0} \{\|x_{n-1} - v\|\}$. From Remark 3.8, we can deduce that

$$\sum_{n=n_0}^{\infty} \theta_n [D_{f_p}(v, x_{n-1}) - D_{f_p}(v, x_n)]_+ \leq \sum_{n=n_0}^{\infty} \theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\| M < \infty. \tag{3.29}$$

Thus, from (3.28) with Lemma 2.6, we have the limit of $\{D_{f_p}(v, x_n)\}$ exists. Note that (3.29) implies that

$$\lim_{n \rightarrow \infty} \theta_n [D_{f_p}(v, x_{n-1}) - D_{f_p}(v, x_n)]_+ = 0. \tag{3.30}$$

From (3.27), we have

$$\begin{aligned} & \alpha_n \tau_n \left(1 - \frac{\kappa_q \mu}{q} \left(\frac{\tau_n}{\tau_{n+1}} \right)^{q-1} \right) \|w_n\|^p + \alpha_n D_{f_p}(y_n, R_{\tau_n} y_n) \\ & \leq D_{f_p}(v, x_n) - D_{f_p}(v, x_{n+1}) + \theta_n [D_{f_p}(v, x_{n-1}) - D_{f_p}(v, x_n)]_+. \end{aligned}$$

This implies by (3.30) that

$$\lim_{n \rightarrow \infty} \|Tu_n - Q_{\mu_n} Tu_n\| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} D_{f_p}(y_n, R_{\tau_n} y_n) = 0.$$

Thus,

$$\lim_{n \rightarrow \infty} \|J_p^E(Tu_n - Q_{\mu_n} Tu_n)\| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \|y_n - R_{\tau_n} y_n\| = 0.$$

By the definition of y_n , we can show that $\lim_{n \rightarrow \infty} \|J_p^E(y_n) - J_p^E(u_n)\| = 0$. Also, by the definition of u_n , we have

$$\lim_{n \rightarrow \infty} \|J_p^E(u_n) - J_p^E(x_n)\| = \lim_{n \rightarrow \infty} \theta_n \|J_p^E(x_n) - J_p^E(x_{n-1})\| = 0.$$

Consequently, $\lim_{n \rightarrow \infty} \|u_n - x_n\| = 0$. Since the rest of the proof is the same as the proof of Theorem 3.6, we omit the details here. □

Next, we apply our algorithms to solve the split feasibility problem in Banach spaces.

Let C and Q be nonempty, closed, and convex subsets of E and F , respectively. Let T be a bounded linear operator with its adjoint operator T^* and $T \neq 0$. We consider the *split feasibility problem* (SFP):

$$\text{find } w \in C \quad \text{such that} \quad Tw \in Q. \tag{3.31}$$

We denote the set of solutions of SFP by $\Gamma := C \cap T^{-1}(Q)$. The SFP was first introduced in 1994 by Censor and Elfving [15] for inverse problems of intensity-modulated radiation therapy (IMRT) in the field of medical care (see [12, 14]).

Setting $A := \partial\delta_C$ and $B := \partial\delta_Q$ in Theorems 3.6 and 3.9, we obtain the following results.

Corollary 3.10 *Let E, F, C, Q, T , and T^* be the same as mentioned above. Let $\tau_1, \beta, \mu, \{s_n\}, \{\beta_n\}, \{\theta_n\}$, and $\{\bar{\theta}_n\}$ be the same as in Algorithm 1, where $\{\bar{\theta}_n\}$ is defined the same as in (3.1). Suppose that $\Gamma \neq \emptyset$. Let $x_0, x_1 \in E$ and $\{x_n\}$ be a sequence generated by*

$$\begin{cases} u_n = J_q^{E^*} (J_p^E(x_n) + \theta_n (J_p^E(x_n) - J_p^E(x_{n-1}))), \\ y_n = J_q^{E^*} (J_p^E(u_n) - \tau_n T^* J_p^F(I - P_Q)Tu_n), \\ x_{n+1} = J_q^{E^*} ((1 - \alpha_n)J_p^E(u_n) + \alpha_n J_p^E(\Pi_C^f y_n)), \\ \tau_{n+1} = \begin{cases} \min\{(\frac{\mu \|(I - P_Q)Tu_n\|^p}{\|T^* J_p^F(I - P_Q)Tu_n\|^q})^{\frac{1}{q-1}}, \tau_n + s_n\}, & \text{if } T^* J_p^F(I - P_Q)Tu_n \neq 0, \\ \tau_n + s_n, & \text{otherwise.} \end{cases} \end{cases}$$

Suppose, in addition, that J_p^E is weakly sequentially continuous on E . Then, the sequence $\{x_n\}$ converges weakly to a point in Γ .

Corollary 3.11 *Let E, F, C, Q, T , and T^* be the same as mentioned above. Let $\tau_1, \beta, \mu, \{s_n\}, \{\beta_n\}, \{\theta_n\}$, and $\{\bar{\theta}_n\}$ be the same as in Algorithm 2, where $\{\bar{\theta}_n\}$ is defined the same as in (3.22). Suppose that $\Gamma \neq \emptyset$. Let $x_0, x_1 \in E$ and $\{x_n\}$ be a sequence generated by*

$$\begin{cases} u_n = J_q^{E^*} (J_p^E(x_n) + \theta_n(J_p^E(x_{n-1}) - J_p^E(x_n))), \\ y_n = J_q^{E^*} (J_p^E(u_n) - \tau_n T^* J_p^E(I - P_Q)Tu_n), \\ x_{n+1} = J_q^{E^*} ((1 - \alpha_n)J_p^E(u_n) + \alpha_n J_p^E(\Pi_C^J y_n)), \\ \tau_{n+1} = \begin{cases} \min\{(\frac{\mu \| (I - P_Q)Tu_n \|^p}{\| T^* J_p^E(I - P_Q)Tu_n \|^q})^{\frac{1}{q-1}}, \tau_n + s_n\}, & \text{if } T^* J_p^E(I - P_Q)Tu_n \neq 0, \\ \tau_n + s_n, & \text{otherwise.} \end{cases} \end{cases}$$

Suppose, in addition, that J_p^E is weakly sequentially continuous on E . Then, the sequence $\{x_n\}$ converges weakly to a point in Γ .

4 Numerical and experiments results

In this section, we apply our Algorithms 1 and 2 to numerically solve some problems in science and engineering and we also compare the numerical performances with the iterative scheme (1.8) proposed by Tang [43] (namely, Tang Algorithm) and the iterative scheme (1.5) proposed by Suantai et al. [39] (namely, Suantai et al. Algorithm).

Problem 4.1 Split feasibility problem in infinite-dimensional Banach spaces

Let $E = F = \ell_p^0(\mathbb{R})$ ($1 < p < \infty, p \neq 2$), where $\ell_p^0(\mathbb{R})$ is the subspace of $\ell_p(\mathbb{R})$, that is,

$$\ell_p^0(\mathbb{R}) = \left\{ x = (x_1, x_2, \dots, x_i, 0, 0, 0, \dots), x_i \in \mathbb{R} \text{ and } \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}$$

with norm $\|x\|_{\ell_p} = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}$ and duality pairing $\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i$ for all $x = (x_1, x_2, \dots, x_i, \dots) \in \ell_p(\mathbb{R})$ and $y = (y_1, y_2, \dots, y_i, \dots) \in \ell_q(\mathbb{R})$, where $\frac{1}{p} + \frac{1}{q} = 1$. The generalized duality mapping $J_p^{\ell_p(\mathbb{R})}$ is computed by the following explicit formula (see [3]):

$$J_p^{\ell_p(\mathbb{R})}(x) = (|x_1|^{p-2}x_1, |x_2|^{p-2}x_2, \dots, |x_i|^{p-2}x_i, \dots), \quad \forall x \in \ell_p(\mathbb{R}).$$

In this example, let $p = 3$, we have $q = \frac{3}{2}$. Then, the smoothness constant $\kappa_q \approx 1.3065$. Let $C = \{x \in \ell_3^0(\mathbb{R}) : \|x\|_{\ell_3} \leq 1\}$ and $Q = \{x \in \ell_3^0(\mathbb{R}) : \langle x, a \rangle \leq 1\}$, where $a := (1, 1, \dots, 1, 0, 0, \dots) \in \ell_{3/2}^0(\mathbb{R})$. Define an operator $Tx = \frac{x}{2}$ with its adjoint $T^* = T$ and $\|T\| = \frac{1}{2}$. In this experiment, we only perform the numerical tests of our Algorithms 1, 2, and Suantai et al. Algorithm [39] since Tang Algorithm [43] cannot be implemented in $\ell_3(\mathbb{R})$. For Algorithms 1 and 2, we set $\tau_1 = 1.99, \beta = 0.75, \mu = 10^{-5}, s_n = \frac{1}{(n+1)^4}, \alpha_n = 0.1$, and $\beta_n = \frac{1}{(n+10)^5}$. For Suantai et al. Algorithm [39], we set $\lambda_n = 10^{-5}, \alpha_n = \frac{1}{n+1}, \beta_n = \frac{n}{n+1}$, and $u_n = (\frac{1}{n^2}, \frac{1}{n^2}, \frac{1}{n^2}, 0, 0, 0, \dots)^T$. The initial points x_0 and x_1 are generated randomly in $\ell_p^0(\mathbb{R})$. We use $E_n = \|x_{n+1} - x_n\|_{l_3} < 10^{-6}$ to terminate iterations for all algorithms. To test the robustness of each algorithm, we run the experiment several times and choose the best four tests of sequences generated by each algorithm. The numerical results are presented in Figs. 1 and 2.

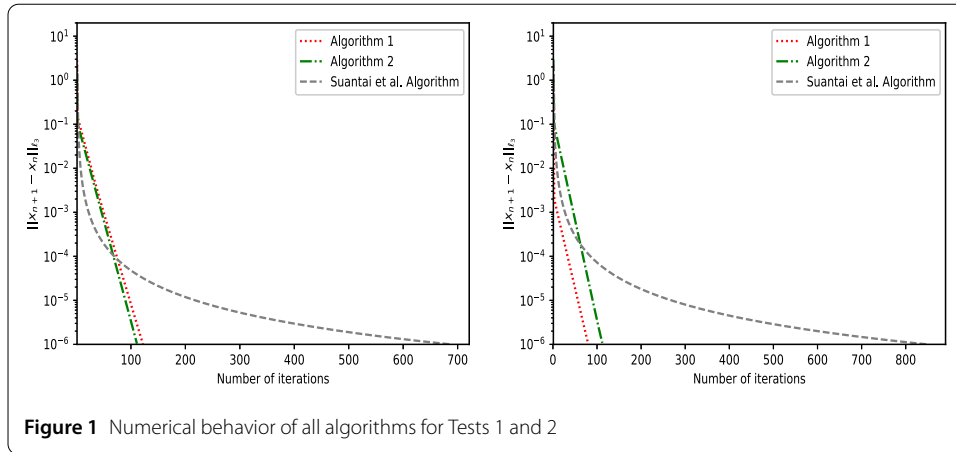


Figure 1 Numerical behavior of all algorithms for Tests 1 and 2

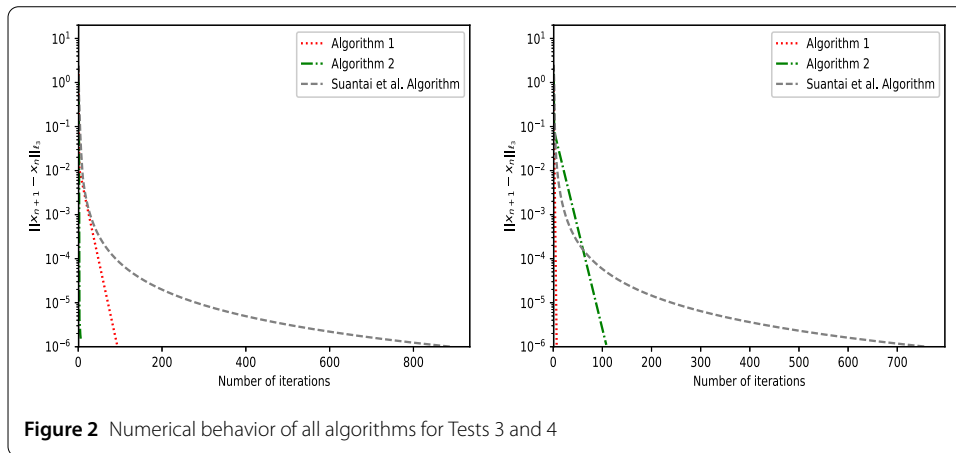


Figure 2 Numerical behavior of all algorithms for Tests 3 and 4

Problem 4.2 [39] Split minimization problem in finite-dimensional spaces

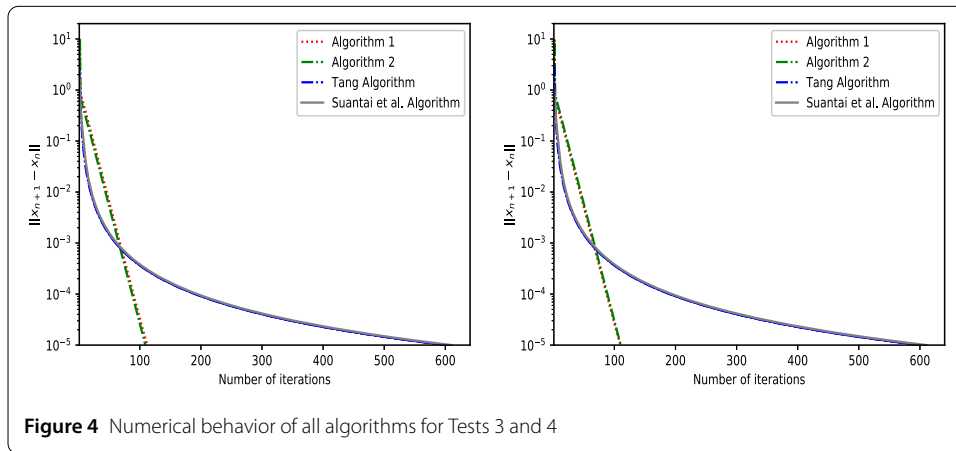
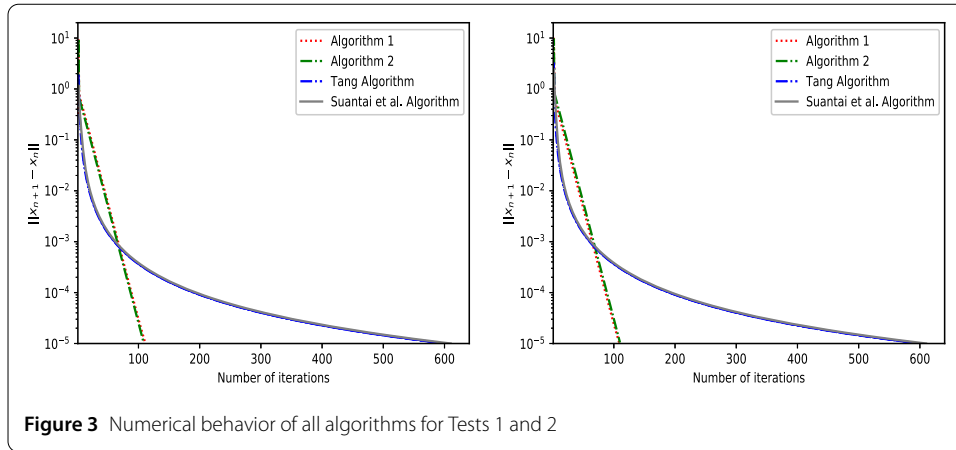
Let $E = F = \mathbb{R}^3$. For each $x \in \mathbb{R}^3$, let $f, g : \mathbb{R}^3 \rightarrow (-\infty, +\infty]$ be defined by

$$f(x) = 20\|x\|^2 + (33, 14, -95)x + 40$$

and

$$g(x) = \frac{1}{2} \|Lx - y\|,$$

where $L = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & 0 \\ 2 & -2 & 4 \end{pmatrix}$ and $y = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$. Let $T = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 5 & 2 \\ 2 & 1 & 0 \end{pmatrix}$. In this case, the split common null-point problem becomes the split minimization problem, that is, find $w \in (\partial f)^{-1}0 \cap T^{-1}(\partial g)^{-1}0$. Note that $T^* = \begin{pmatrix} 1 & 0 & 2 \\ -2 & 5 & 1 \\ 3 & 2 & 0 \end{pmatrix}$ and $\|T\|^2$ is the largest eigenvalue of T^*T . In this experiment, we compare the numerical performances of our Algorithms 1 and 2 with Tang Algorithm [43] and Suantai et al. Algorithm [39]. For Algorithms 1 and 2, we set $\tau_1 = 1.99$, $\beta = 0.75$, $\mu = 0.1$, $\alpha_n = 0.1$, $s_n = \frac{1}{(n+1)^4}$, $\beta_n = \frac{1}{(n+1)^{1.1}}$, and $\tau_n = \mu_n = 0.01$. For Tang Algorithm [43], we set $\alpha = 0.75$, $\rho_n = 3 - \frac{1}{n+1}$, $\epsilon_n = \frac{1}{(n+1)^{1.1}}$, and $r = \mu = 10^{-5}$. For Suantai et al. Algorithm [39], we set $\alpha_n = \frac{1}{n+1}$, $\beta_n = 0.5$, $r_n = 1$, $\lambda_n = 0.01$, and $u_n = (\frac{1}{n^2}, \frac{1}{n^2}, \frac{1}{n^2})^T$. The initial points x_0 and x_1 are generated randomly in \mathbb{R}^3 . We use $E_n = \|x_{n+1} - x_n\| < 10^{-5}$ to terminate iterations for all



algorithms. To test the robustness of each algorithm, we run the experiment several times and choose the best four testes of sequences generated by each algorithm. The numerical results are presented in Figs. 3 and 4.

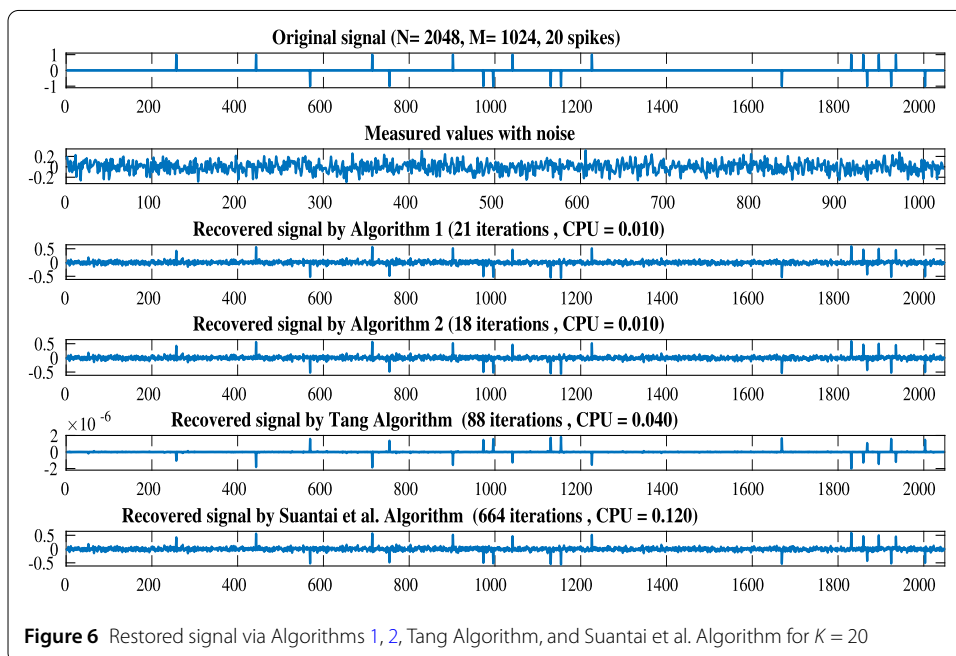
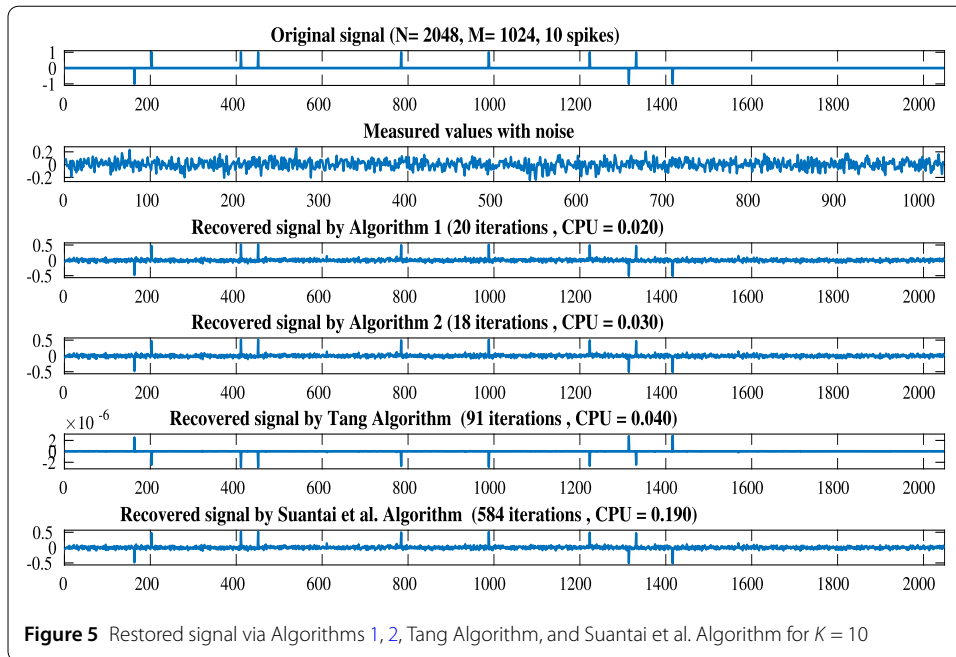
Problem 4.3 Signal-recovery problem

In signal processing, compressed sensing involves the recovery of a “sparse signal” from measured data, aiming to reconstruct the original signal using fewer measurements (see, e.g [27, 47]). In this context, we can model the compressed sensing as the following uncertain linear system:

$$y = Tx + b, \tag{4.1}$$

where $x \in \mathbb{R}^N$ is a K -sparse signal $K (K \ll N)$, to be recovered, $y \in \mathbb{R}^M$ is the observed or measured data with noisy b and $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$ is a bounded linear operator. It is known that the above problem can be seen as the following LASSO problem [46]:

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|Tx - y\|^2 \quad \text{subject to} \quad \|x\|_1 \leq t, \tag{4.2}$$



where $t > 0$ is a given constant and $\|\cdot\|_1$ is the ℓ_1 norm. If $C = \{x \in \mathbb{R}^N : \|x\|_1 \leq t\}$ and $Q = \{y\}$, then (4.2) is a particular case of the SFP (3.31) in the finite-dimensional spaces.

We generated a sparse signal $x \in \mathbb{R}^N$ with K nonzero entries having a length $N = 2048$ and made $M = 1024$ observations. The values of the sparse signal are sampled uniformly from the interval $[-1, 1]$. The observation y is generated from Gaussian noise of variance 10^{-4} . The matrix $T \in \mathbb{R}^{M \times N}$ is generated from a normal distribution with mean zero and one variance. Additionally, the initial signals $x_0 = x_1 = T^*(Tx - y)$. In this experiment, we set $\tau_1 = 1.99$, $\beta = 0.75$, $\mu = 10^{-5}$, $\alpha_n = 0.1$, $s_n = \frac{1}{(n+1)^4}$, $\beta_n = \frac{1}{(n+1)^{1.1}}$, and $\tau_n = \mu_n = 0.01$ in

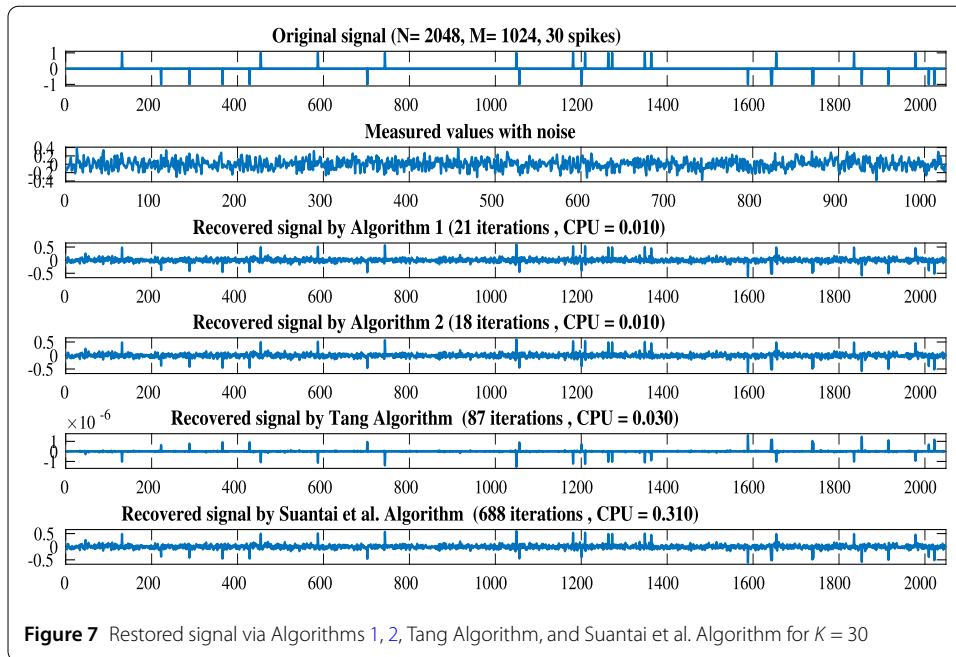


Figure 7 Restored signal via Algorithms 1, 2, Tang Algorithm, and Suantai et al. Algorithm for $K = 30$

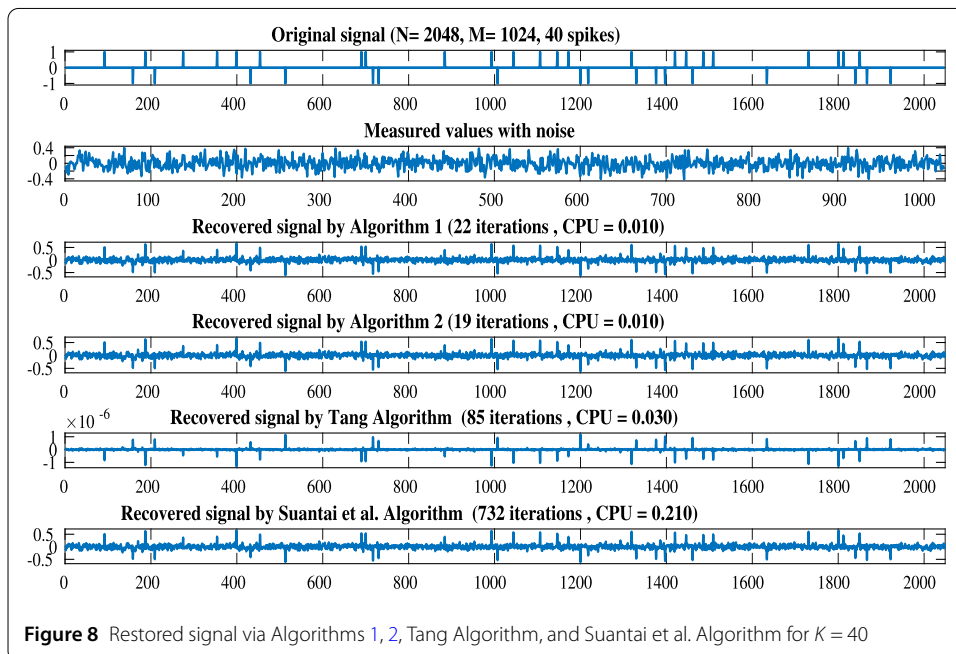
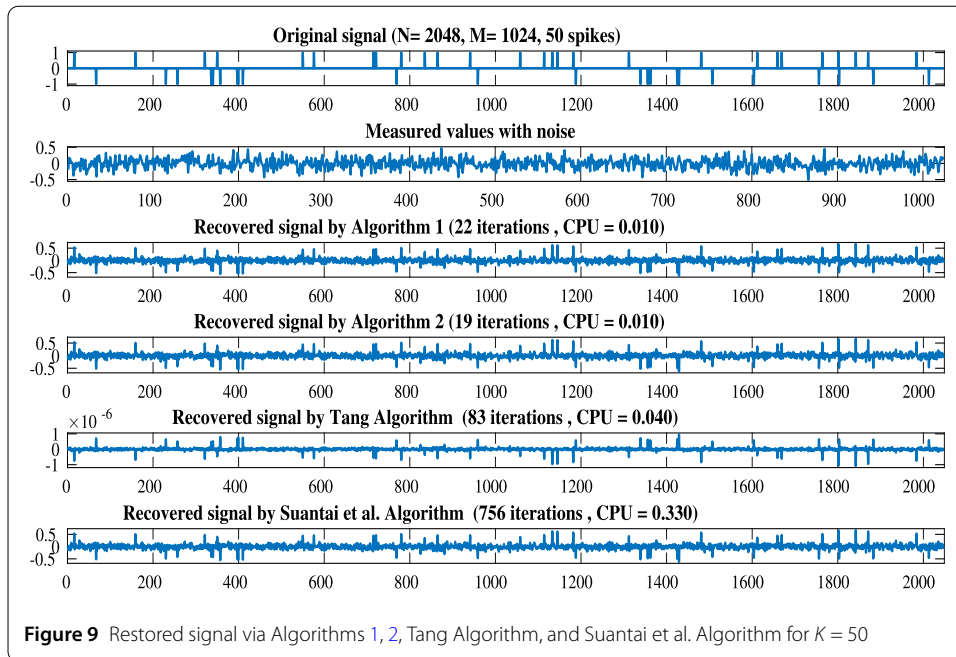


Figure 8 Restored signal via Algorithms 1, 2, Tang Algorithm, and Suantai et al. Algorithm for $K = 40$

Algorithms 1 and 2, we set $\alpha = 0.75$, $\rho_n = 0.5$, $\epsilon_n = \frac{1}{(n+1)^{1.1}}$, and $r = \mu = 0.001$ in Tang Algorithm [43] and we set $\alpha_n = \beta_n = \frac{1}{n+1}$, $u_n = Tx - y$, and $\lambda_n = 0.001$ in Suantai et al. Algorithm [39]. We consider five different tests for the spikes $K \in \{10, 20, 30, 40, 50\}$. Our stopping criterion is $E_n = \|x_{n+1} - x_n\| < 10^{-7}$. The results of the numerical simulations are presented in Figs. 5–9.

Remark 4.4 We observe from the numerical simulations presented in Figs. 5–9 that our proposed Algorithms 1 and 2 outperform Tang Algorithm [43] and Suantai et al. Algorithm [39] in the sense that they satisfy the stopping criteria in fewer iterations and less



computational time in the signal-recovery tests. Furthermore, we observe that while Tang Algorithm [43] requires fewer iterations to satisfy the stopping criteria compared to Suantai et al. Algorithm [39], the reconstructed signal by Tang Algorithm [43] is NOT very close to the original signal compared to that reconstructed via Suantai et al. Algorithm [39].

Problem 4.5 Data classifications

In this example, we apply our algorithms to data-classification problems, which are based on a learning technique called the extreme learning machine (ELM). Let $\mathcal{U} = \{(x_n, y_n) : x_n \in \mathbb{R}^N, y_n \in \mathbb{R}^M, n = 1, 2, 3, \dots, K\}$ be a training set of K distinct samples, x_n is an input training data and y_n is a training target. For the output of ELM with a single hidden layer at the i th hidden node is $h_i(x) = U(\langle a_i, x_n \rangle + b_i)$, where U is an activation function, a_i is the weight at the i th hidden node, and b_i is the bias at the i th hidden node. The output function with L hidden nodes is the single hidden-layer feedforward neural networks (SLFNs)

$$O_n = \sum_{i=1}^L \omega_i h_i(x_n),$$

where ω_i is the optimal output weight at the i th hidden node. The hidden-layer output matrix T is defined by

$$T = \begin{bmatrix} U(\langle a_1, x_1 \rangle + b_1) & \cdots & U(\langle a_L, x_1 \rangle + b_L) \\ \vdots & \ddots & \vdots \\ U(\langle a_1, x_K \rangle + b_1) & \cdots & U(\langle a_L, x_K \rangle + b_L) \end{bmatrix}.$$

The main aim of ELM is to calculate an optimal weight $\omega = (\omega_1, \omega_2, \dots, \omega_L)^T$ such that $T\omega = b$, where $b = (t_1, t_2, \dots, t_K)^T$ is the training target data. A successful model used to

find the solution ω can be translated into the following convex constraint minimization problem:

$$\min_{\omega \in \mathbb{R}^L} \frac{1}{2} \|T\omega - b\|^2 \quad \text{subject to} \quad \|\omega\|_1 \leq \xi, \tag{4.3}$$

where $\xi > 0$ is a given constant. If $C = \{\omega \in \mathbb{R}^L : \|\omega\|_1 \leq \xi\}$ and $Q = \{b\}$, then (4.3) is a particular case of the SFP (3.31) in the finite-dimensional spaces.

The binary crossentropy loss function along with sigmoid activation function for binary classification calculates the loss of an example by computing the following average:

$$\text{Loss} = -\frac{1}{J} \sum_{j=1}^J (y_j \log \hat{y}_j + (1 - y_j) \log(1 - \hat{y}_j)), \tag{4.4}$$

where \hat{y}_j is the j th scalar value in the model output, y_j is the corresponding target value, and J is the number of scalar values in the model output.

The performance evaluation in classification can be justified by precision and recall. The Recall/True Positive Rate can be defined as the level of accuracy of predictions in positive classes and the percentage of the number of predictions that are correct on the positive observations. Then, calculate the accuracy, prediction, and F1-score using the following standard criteria [22]:

- (1) Precision = $\frac{TP}{TP+FP} \times 100\%$;
- (2) Recall = $\frac{TP}{TP+FN} \times 100\%$;
- (3) Accuracy = $\frac{TP+TN}{TP+FP+TN+FN} \times 100\%$;
- (4) F1-score = $\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$,

where a confusion matrix for original and predicted classes is shown in terms of TP := True Positive, TN := True Negative, FP := False Positive, and FN := False Negative.

Next, we consider the following two datasets:

Dataset 1 UCI Machine Learning Heart Disease dataset [20]. This dataset contains 14 attributes and 303 records. This dataset contains the attributes: Age, Gender, CP, Trestbps, Chol, Fbs, Restecg, Thalach, Exang, Oldpeak, Slope, Ca, Thal, and Num (the predicted attribute). The dataset consists of 138 normal instances versus 165 abnormal instances.

Dataset 2 PIMA Indians diabetes dataset [1]. The dataset contains 768 pregnant female patients of which 500 were nondiabetics and 268 that were diabetics. This dataset contains 9 attributes: Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, Age, and Outcome (the predicted attribute).

In particular, we apply our algorithms to the optimized weight parameter in training data for machine learning by using 5-fold crossvalidation [25] in the extreme learning machine (ELM).

For Dataset 1, we start computation by setting the activation function as a sigmoid, hidden nodes $L = 80$, regularization parameter $\xi = 10$, $x_0 = \mathbf{1} := \underbrace{(1, 1, \dots, 1)}_N \in \mathbb{R}^N$, $x_1 = \mathbf{0} := \underbrace{(0, 0, \dots, 0)}_N \in \mathbb{R}^N$, $\tau_1 = 1$, $s_n = 0$, $\beta_n = \frac{1}{(n+1)^{10}}$, and $\alpha_n = \frac{n+2}{2n+1}$. The stopping criteria is the

Table 1 Numerical results of β for Algorithm 1 when $\mu = 0.9$ for Dataset 1

β	Training	Loss	
		Train	Test
0.1	0.0808	0.246457	0.224908
0.2	0.0816	0.246016	0.224521
0.3	0.0841	0.245528	0.224100
0.4	0.0824	0.244935	0.223599
0.5	0.0882	0.244198	0.222994
0.6	0.0839	0.243251	0.222248
0.7	0.0794	0.241977	0.221302
0.8	0.0833	0.240139	0.220054
0.9	0.0875	0.237138	0.218174

Table 2 Numerical results of β for Algorithm 2 when $\mu = 0.9$ for Dataset 1

β	Training	Loss	
		Train	Test
0.1	0.0775	0.247571	0.225920
0.2	0.0774	0.247902	0.226225
0.3	0.0771	0.248156	0.226459
0.4	0.0783	0.248382	0.226669
0.5	0.0755	0.248584	0.226858
0.6	0.0815	0.248767	0.227030
0.7	0.1114	0.248933	0.227185
0.8	0.0784	0.249083	0.227327
0.9	0.0800	0.249221	0.227457

Table 3 Numerical results of μ for Algorithm 1 when $\beta = 0.9$ for Dataset 1

μ	Training	Loss	
		Train	Test
0.1	0.0870	0.245923	0.224461
0.2	0.0846	0.243328	0.222218
0.3	0.0829	0.241784	0.221118
0.4	0.0830	0.240556	0.220319
0.5	0.0838	0.239611	0.219739
0.6	0.0872	0.238817	0.219257
0.7	0.0872	0.238143	0.218838
0.8	0.0870	0.237596	0.218483
0.9	0.0848	0.237138	0.218174

number of iterations 300. We compare the performance of Algorithms 1 and 2 with different parameters β for Dataset 1, as seen in Tables 1 and 2.

From Table 1, we see that β increases from 0.1 to 0.9. The training loss and test loss decrease, it appears that $\beta = 0.9$ performs better for Algorithm 1.

From Table 2, we see that the training loss and test loss increase when β increases, it appears that $\beta = 0.1$ performs better for Algorithm 2. Also, we compare the performance of Algorithms 1 and 2 with different parameters μ for Dataset 1, as seen in Tables 3 and 4.

From Tables 3 and 4, we see that the training loss and test loss decrease when μ increases, it show that $\mu = 0.9$ gives the highly improved the performance of Algorithm 1 and Algorithm 2 for Dataset 1.

Next, we compare the performance of our algorithms with Tang Algorithm [43] and Suantai et al. Algorithm [39] for Dataset 1. For Algorithms 1 and 2, we set $\tau_1 = 1$, $\mu = 0.9$, $s_n = 0$, and $\beta_n = \frac{1}{(n+1)^{10}}$. Moreover, we set $\beta = 0.9$ and $\beta = 0.1$ for Algorithms 1 and 2,

Table 4 Numerical results of μ for Algorithm 2 when $\beta = 0.1$ for Dataset 1

μ	Training	Loss	
		Train	Test
0.1	0.0775	0.250309	0.227553
0.2	0.0802	0.250964	0.229119
0.3	0.0778	0.250372	0.228554
0.4	0.1137	0.249819	0.228024
0.5	0.0883	0.249305	0.227536
0.6	0.0808	0.248827	0.227085
0.7	0.0787	0.248381	0.226668
0.8	0.1336	0.247963	0.226281
0.9	0.0805	0.247571	0.225920

Table 5 The performance of each algorithm for Dataset 1

Algorithms	Iterations	Training time	Precision	Recall	F1-score	Accuracy
Algorithm 1	49	0.0680	87.96	100	93.59	88.21
Algorithm 2	524	0.3763	87.96	100	93.59	88.21
Tang Algorithm [43]	2299	1.4604	87.96	100	93.59	88.21
Suantai et al. Algorithm [39]	1695	0.3028	87.96	100	93.59	88.21

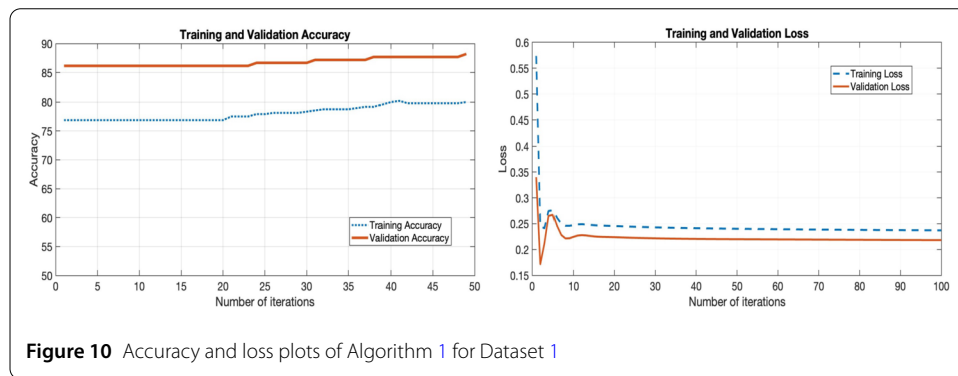


Figure 10 Accuracy and loss plots of Algorithm 1 for Dataset 1

respectively. For Tang Algorithm [43], we set $\alpha = 0.6$, $\rho_n = 3.5$, and $\epsilon_n = \frac{1}{n+1}$. For Suantai et al. Algorithm [39], we set $\alpha_n = \frac{1}{n+1}$, $\beta_n = \frac{n-1}{2n}$, $u_n = \mathbf{1} \in \mathbb{R}^N$, and $\lambda_n = \frac{1}{\|T\|^2}$.

The comparison of all algorithms is presented in Table 5.

From Table 5, we observe that our Algorithms 1 and 2 have fewer iterations than Tang Algorithm [43] and Suantai et al. Algorithm [39] with the same precision, recall, F1-score, and accuracy. This shows that our algorithms have the highest probability of correctly classifying heart disease compared to other algorithms.

Next, we present graphs of the accuracy and loss of training data and testing data for overfitting of Algorithms 1 and 2 to show that our algorithms have no overfitting in the training Dataset 1.

From Figs. 10 and 11, we see that our Algorithms 1 and 2 have suitably learned the training dataset for Dataset 1.

For Dataset 2, we present the comparison of our algorithms with Tang Algorithm [43] and Suantai et al. Algorithm [39]. We start computation by setting the activation function as sigmoid, hidden nodes $L = 160$. For Algorithms 1 and 2, we set $\tau_1 = 1$, $\mu = 0.9$, $s_n = 0$, and $\beta_n = \frac{1}{(n+1)^{10}}$. Moreover, we set $\beta = 0.9$ and $\beta = 0.1$ for Algorithms 1 and 2, respectively. For

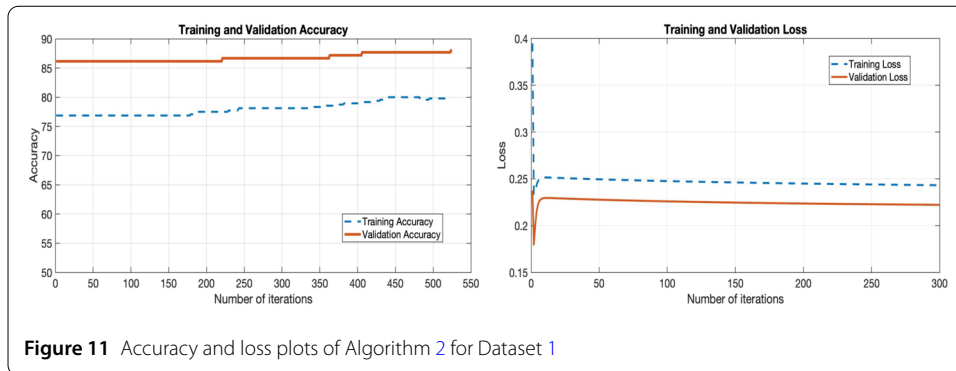


Figure 11 Accuracy and loss plots of Algorithm 2 for Dataset 1

Table 6 The performance of each algorithm for Dataset 2

Algorithms	Iterations	Training time	Precision	Recall	F1-score	Accuracy
Algorithm 1	14	0.0728	80.97	97.50	88.47	80.03
Algorithm 2	14	0.0612	80.97	97.50	88.47	80.03
Tang Algorithm [43]	33	0.0967	80.97	97.50	88.47	80.03
Suantai et al. Algorithm [39]	44	0.0620	80.97	97.50	88.47	80.03

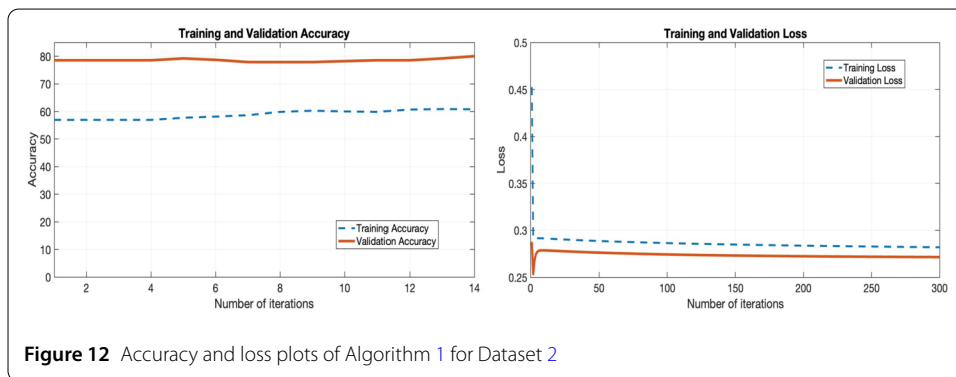


Figure 12 Accuracy and loss plots of Algorithm 1 for Dataset 2

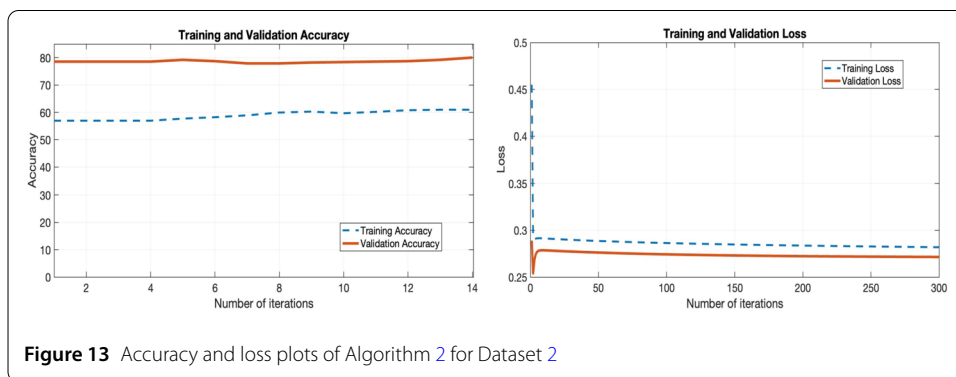


Figure 13 Accuracy and loss plots of Algorithm 2 for Dataset 2

Tang Algorithm [43], we set $\alpha = 0.6$, $\rho_n = 3.5$, and $\epsilon_n = \frac{1}{(n+1)^{10}}$. For Suantai et al. Algorithm [39], we set $\alpha_n = \frac{1}{n+1}$, $\beta_n = \frac{n-1}{2n}$, $u_n = \mathbf{1} \in \mathbb{R}^N$, and $\lambda_n = \frac{1}{\|T\|^2}$.

The comparison of all algorithms is presented in Table 6.

From Table 6, we see that Algorithms 1 and 2 have the most efficiency in precision, recall, F1-score, and accuracy for Dataset 2. This mean that our algorithms have the highest

probability of correctly classifying the PIMA Indians diabetes dataset (Dataset 2) compared to Tang Algorithm [43] and Suantai et al. Algorithm [39].

Next, we present graphs of the accuracy and loss of training data and testing data for overfitting of Algorithms 1 and 2 to show that our algorithms have no overfitting in the training Dataset 2.

From Figs. 12 and 13, we see that Algorithms 1 and 2 have suitably learned the training dataset for Dataset 2.

5 Conclusions

In this paper, we have proposed two inertial self-adaptive algorithms to solve the split common null-point problem for two set-valued mappings in Banach spaces. The step sizes used in our proposed algorithms are adaptively updated without the prior knowledge of the operator norm of the bounded linear operator. We have proved the weak-convergence theorems of the proposed algorithms under suitable conditions in p -uniformly convex, real Banach spaces that are also uniformly smooth. Finally, we have performed experiments to numerically solve some problems in science and engineering, such as, the split feasibility problem, the split minimization problem, signal recovery, and data classifications, and also have compared them with some existing methods to demonstrate the implementability and efficiency of our methods.

Acknowledgements

The authors wish to thank the anonymous referees for their valuable comments and valuable suggestions that have led to considerable improvement of this paper.

Funding

This research was supported by The Science, Research and Innovation Promotion Funding (TSRI) (Grant no. FRB660012/0168). This research block grants was managed under Rajamangala University of Technology Thanyaburi (FRB66E06161.5).

Availability of data and materials

Not applicable.

Declarations

Competing interests

The authors declare no competing interests.

Author contributions

Conceptualization: P.S.; Writing-original draft: R.P., P.S.; Formal analysis: E.K., S.K.; Investigation: R.P., P.S.; Software: S.K., R.P.; Review and editing: P.S., R.P.; Project administration: P.S.; Funding: R.P., E.T. All authors have read and approved final version of the manuscript.

Author details

¹Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi (RMUTT), Thanyaburi, Pathumthani, 12110, Thailand. ²Office of Research Administration, Chiang Mai University, Chiang Mai 50200, Thailand.

Received: 7 July 2023 Accepted: 11 October 2023 Published online: 25 October 2023

References

1. <https://www.kaggle.com/datasets/whenamancodes/predict-diabities>
2. Agarwal, R.P., O'Regan, D., Sahu, D.R.: Fixed Point Theory for Lipschitzian-Type Mappings with Applications. Springer, Berlin (2009)
3. Alber, Y., Ryazantseva, I.: Nonlinear Ill-Posed Problems of Monotone Type. Springer, Dordrecht (2006)
4. Alber, Y.I.: Metric and generalized projection operators in Banach spaces: properties and applications. In: Kartsatos, A.G. (ed.) Theory and Applications of Nonlinear Operator of Accretive and Monotone Type, pp. 15–50. Dekker, New York (1996)
5. Alofi, A.S., Alsulami, S.M., Takahashi, W.: Strongly convergent iterative method for the split common null point problem in Banach spaces. *J. Nonlinear Convex Anal.* **17**, 311–324 (2016)

6. Alvarez, F.: Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert spaces. *SIAM J. Optim.* **14**, 773–782 (2004)
7. Alvarez, F., Attouch, H.: An inertial proximal method for monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Var. Anal.* **9**, 3–11 (2001)
8. Anh, P.K., Thong, D.V., Dung, V.T.: A strongly convergent Mann-type inertial algorithm for solving split variational inclusion problems. *Optim. Eng.* **22**, 159–185 (2021)
9. Aoyama, K., Kohsaka, F., Takahashi, W.: Three generalizations of firmly nonexpansive mappings: their relations and continuity properties. *J. Nonlinear Convex Anal.* **10**, 131–147 (2009)
10. Bauschke, H.H., Borwein, J.M., Combettes, P.L.: Bregman monotone optimization algorithms. *SIAM J. Control Optim.* **42**, 596–636 (2003)
11. Butnariu, D., Resmerita, E.: Bregman distances, totally convex functions and a method for solving operator equations in Banach spaces. *Abstr. Appl. Anal.* **2006**, Article ID 084919 (2006)
12. Byrne, C.: Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Probl.* **18**, 441–453 (2002)
13. Byrne, C., Censor, Y., Gibali, A., Reich, S.: The split common null point problem. *J. Nonlinear Convex Anal.* **13**, 759–775 (2012)
14. Censor, Y., Bortfeld, T., Martin, B., Trofimov, A.: A unified approach for inversion problems in intensity modulated radiation therapy. *Phys. Med. Biol.* **51**, 2353–2365 (2006)
15. Censor, Y., Elfving, T.: A multiprojection algorithm using Bregman projections in product space. *Numer. Algorithms* **8**, 221–239 (1994)
16. Chen, H.Y.: Weak and strong convergence of inertial algorithms for solving split common fixed point problems. *J. Inequal. Appl.* **2021**, 26 (2021)
17. Chuang, C.S.: Hybrid inertial proximal algorithm for the split variational inclusion problem in Hilbert spaces with applications. *Optimization* **66**, 777–792 (2017)
18. Cioranescu, I.: *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*. Kluwer Academic, Dordrecht (1990)
19. Combette, P.L.: The convex feasibility problem in image recovery. *Adv. Imaging Electron Phys.* **95**, 155–270 (1996)
20. Dua, D., Graff, C.: UCI Machine Learning Repository. University of California, School of Information and Computer Science, Irvine (2019). <http://archive.ics.uci.edu/ml>
21. Duan, P., Zhang, Y., Bu, Q.: New inertial proximal gradient methods for unconstrained convex optimization problems. *J. Inequal. Appl.* **2020**, 255 (2020)
22. Han, J., Kamber, M., Pei, J.: *Data Mining: Concepts and Techniques*, vol. 10, 978-1. Kaufman, Waltham (2012)
23. Hanner, O.: On the uniform convexity of L_p and l_p . *Ark. Mat.* **3**, 239–244 (1956)
24. Kesornprom, S., Cholamjiak, K.: Proximal type algorithms involving linesearch and inertial technique for split variational inclusion problem in Hilbert spaces with applications. *Optimization* **68**, 2369–2395 (2019)
25. Kumari, V.A., Chitra, R.: Classification of diabetes disease using support vector machine. *Int. J. Eng. Res. Appl.* **3**, 1797–1801 (2013)
26. Kuo, L.W., Sahu, D.R.: Bregman distance and strong convergence of proximal-type algorithms. *Abstr. Appl. Anal.* **2013**, Article ID 590519 (2006)
27. Kutyniok, G.: Theory and applications of compressed sensing. *GAMM-Mitt.* **36**, 79–101 (2013)
28. López, G., Martín-Marquez, V., Wang, F., Xu, H.K.: Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Probl.* **28**, 085004 (2012)
29. Moudafi, M.: Split monotone variational inclusions. *J. Optim. Theory Appl.* **150**, 275–283 (2011)
30. Ogbuisi, F.U., Shehu, Y., Yao, J.C.: Convergence analysis of new inertial method for the split common null point problem. *Optimization* **71**, 3767–3795 (2022)
31. Reich, S.: A weak convergence theorem for the alternating method with Bregman distances. In: *Theory and Applications of Nonlinear Operators of Accretive and Monotone Type*, pp. 313–318. Dekker, New York (1996)
32. Reich, S., Tuyen, T.M., Sunthrayuth, P., Cholamjiak, P.: Two new inertial algorithms for solving variational inequalities in reflexive Banach spaces. *Numer. Funct. Anal. Optim.* **42**, 1954–1984 (2021)
33. Rockafellar, R.T.: On the maximal monotonicity of subdifferential mappings. *Pac. J. Math.* **33**, 209–216 (1970)
34. Schöpfer, F.: Iterative regularization method for the solution of the split feasibility problem in Banach spaces. Ph.D. thesis, Saabrücken (2007)
35. Schöpfer, F., Louis, A.K., Schuster, T.: Nonlinear iterative methods for linear ill-posed problems in Banach spaces. *Inverse Probl.* **22**, 311–329 (2006)
36. Schöpfer, F., Schuster, T., Louis, A.K.: An iterative regularization method for the solution of the split feasibility problem in Banach spaces. *Inverse Probl.* **24**, 055008 (2008)
37. Shehu, Y., Iyiola, O.S., Ogbuisi, F.U.: Iterative method with inertial terms for nonexpansive mappings: applications to compressed sensing. *Numer. Algorithms* **83**, 1321–1347 (2020)
38. Suantai, S., Pholasa, N., Cholamjiak, P.: Relaxed CQ algorithms involving the inertial technique for multiple-sets split feasibility problems. *Rev. R. Acad. Cienc. Exactas Fís. Nat., Ser. A Mat.* **113**, 1081–1099 (2019)
39. Suantai, S., Shehu, Y., Cholamjiak, P.: Nonlinear iterative methods for solving the split common null point problem in Banach spaces. *Optim. Methods Softw.* **34**, 853–874 (2019)
40. Takahashi, E.: *Nonlinear Functional Analysis*. Yokohama Publishers, Yokohama (2000)
41. Tan, B., Sunthrayuth, P., Cholamjiak, P., Cho, Y.J.: Modified inertial extragradient methods for finding minimum-norm solution of the variational inequality problem with applications to optimal control problem. *Int. J. Comput. Math.* **100**, 525–545 (2023)
42. Tan, K.K., Xu, H.K.: Approximating fixed points of nonexpansive mappings by the Ishikawa iteration process. *J. Math. Anal. Appl.* **178**, 301–308 (1993)
43. Tang, Y.: New inertial algorithm for solving split common null point problem in Banach spaces. *J. Inequal. Appl.* **2019**, 17 (2019)
44. Tang, Y., Sunthrayuth, P.: An iterative algorithm with inertial technique for solving the split common null point problem in Banach spaces. *Asian-Eur. J. Math.* **15**, 2250120 (2022)

45. Tian, M., Jiang, B.N.: Inertial hybrid algorithm for variational inequality problems in Hilbert spaces. *J. Inequal. Appl.* **2020**, 12 (2020)
46. Tibshirani, R.: Regression shrinkage and selection via the LASSO. *J. R. Stat. Soc. Ser. B* **58**, 267–288 (1996)
47. Tropp, J.A.: A mathematical introduction to compressive sensing. *Bull. Am. Math. Soc.* **54**, 151–165 (2017)
48. Tuyen, T.M., Cholamjiak, P., Sunthrayuth, P.: A new self-adaptive method for the multiple-sets split common null point problem in Banach spaces. *Vietnam J. Math.* (2022). <https://doi.org/10.1007/s10013-022-00574-3>
49. Xu, H.K.: Inequalities in Banach spaces with applications. *Nonlinear Anal., Theory Methods Appl.* **16**, 1127–1138 (1991)
50. Xu, Z.B., Roach, G.F.: Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces. *J. Math. Anal. Appl.* **157**, 189–210 (1991)
51. Zhou, Z., Tan, B., Li, S.: Inertial algorithms with adaptive stepsizes for split variational inclusion problems and their applications to signal recovery problem. *Math. Methods Appl. Sci.* (2023). <https://doi.org/10.1002/mma.9436>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
