J. Appl. Numer. Optim. 6 (2024), No. 1, pp. 41-57
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.6.2024.1.02

TSENG-TYPE ALGORITHMS FOR SOLVING VARIATIONAL INEQUALITIES OVER THE SOLUTION SETS OF SPLIT VARIATIONAL INCLUSION PROBLEMS WITH AN APPLICATION TO A BILEVEL OPTIMIZATION PROBLEM

ADEOLU TAIWO[1,*], SIMEON REICH[1], RAVI P. AGARWAL[2]

[1]*Department of Mathematics, The Technion – Israel Institute of Technology, 3200003 Haifa, Israel*
[2]*Department of Mathematics, Texas A&M University-Kingsville, Kingsville, TX 78363, USA*

Dedicated to Professor Yair Censor on the occasion of his 80th birthday

Abstract. We study a new class of variational inclusion problems in the framework of real Hilbert spaces. We propose two Tseng-type algorithms with inertial extrapolation for solving these problems and carry out the convergence analysis of these two methods. We also give an application of our results to solve convex bilevel optimization problems.

Keywords. Bilevel optimization problems; Inertial extrapolation; Inclusion problem; Tseng algorithm.

2020 Mathematics Subject Classification. 65K15, 90C30.

1. Introduction

Let $\mathscr{H}$ and $\mathscr{H}_k$, $k = 1,2,\ldots,K$, be real Hilbert spaces. For each k, let $A_k : \mathscr{H}_k \rightrightarrows \mathscr{H}_k$ be a maximal monotone operator, $B_k : \mathscr{H}_k \to \mathscr{H}_k$ be a monotone and L_k-Lipschitz continuous operator, and $T_k : \mathscr{H} \to \mathscr{H}_k$ be a bounded and linear operator with adjoint $T_k^* : \mathscr{H}_k \to \mathscr{H}$. Let $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ be a strongly monotone and Lipschitz continuous operator.

In this paper, we study the following variational inequality over the solution sets of split variational inclusion problems with multiple output sets (called VI-SM, for short):

$$\text{Find } x \in \mathscr{S} := \cap_{k=1}^{K} T_k^{-1}(A_k + B_k)^{-1}(0) \tag{1.1}$$

such that

$$\langle \mathscr{F}x, y - x\rangle \geq 0 \ \ \forall y \in \mathscr{S}. \tag{1.2}$$

Problem (1.1)–(1.2) encompasses many known problems, such as the split feasibility problem with multiple output sets [1], the split common null point problem with multiple output sets [2], the variational inequality problem over the solution sets of split variational inclusion problems [3], and the split common fixed point problem with multiple output sets [4].

---

*Corresponding author.
E-mail address: taiwo.adeolu@yahoo.com, taiwoa@campus.technion.ac.il (A. Taiwo), sreich@technion.ac.il (S. Reich), ravi.agarwal@tamuk.edu (R. P. Agarwal).

Received 22 October 2023; Accepted 30 November 2023; Published online 5 March 2024.

If $K=1, \mathscr{H}=\mathscr{H}_1$, and $T_1=I$, the identity operator, then (1.1) reduces to the following variational inclusion problem (VI, for short):

$$\text{Find } x\in\mathscr{H} \text{ such that } x\in(A+B)^{-1}(0). \tag{1.3}$$

Under the given data, VI (1.3) is in general ill-posed, but combining it with (1.2) ensures well-posedness. A prominent iterative scheme for solving (1.3) is the forward-backward splitting method (FBSM) [5]: For $x_1\in\mathscr{H}$, let

$$x_{n+1}=(I+\lambda_n A)^{-1}(x_n-\lambda_n Bx_n),\ n\in\mathbb{N}, \tag{1.4}$$

where $\{\lambda_n\}$ is the sequence of step sizes. It is known that the sequences generated by FBSM converge weakly if $\lambda_n\in(0,\frac{2}{L})$, where $\frac{1}{L}$ is the constant of co-coercivity of B. Note that every co-coercive (inverse strongly monotone) operator is monotone and Lipschitz continuous (see Definition 2.1), but the converse is not true in all cases.

As important as the FBSM is, generally, its convergence is not guaranteed when B is monotone, but not co-coercive. This drawback motivates the invention of the forward-backward-forward splitting method (FBFSM) by Tseng [6]: For $x_1\in\mathscr{H}$, let

$$\begin{cases} y_n=(I+\lambda_n A)^{-1}(x_n-\lambda_n Bx_n),\\ x_{n+1}=y_n-\lambda_n(By_n-Bx_n), n\in\mathbb{N},\end{cases} \tag{1.5}$$

where the step size λ_n is chosen by a certain line search technique.

VI-SM has recently been studied by Reich et al. [3] for the case where the B_ks are co-coercive. In order to solve the problem, they proposed an iterative scheme which is a combination of an inertial forward-backward splitting method and a steepest descent method, and they proved a strong convergence theorem for it. The inertial technique first appeared in Polyak [7]. Since then, it has been applied to countless iterative schemes; see, e.g., [8, 9, 10, 11, 12] and the references therein for more details on the inertial technique.

The assumption made in [3] that B_ks are inverse strongly monotone is a strong one. In this study, we weaken this assumption and solve the VI-SM under the assumption that the B_ks are monotone and Lipschitz continuous. However, in view of the drawbacks of the FBSM, one should expect similar drawbacks in trying to apply the method proposed by Reich et al. [1] to solving the VI-SM in the case where the B_ks are merely monotone. Therefore, motivated by the need to devise a more general useful iterative scheme, we propose in this paper two self-adaptive Tseng-type forward-backward-forward splitting methods for solving the VI-SM. We also prove that the sequences generated by each of these methods converge strongly to a solution of the VI-SM under some mild assumptions on the control parameters.

We describe the organization of our paper as follows. Preliminaries are presented in Section 2. Indeed, we present some definitions of key terminologies and state some lemmata that are important to our study. In Section 3, we present our algorithms and the convergence analysis of each of them. We devote Section 4 to an application of our results to a convex bilevel optimization problem. We also present some numerical experiments. The last section, Section 5, contains some conclusions of our work.

2. Preliminaries

Suppose that $\mathscr{H}$ is a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. Let C be a nonempty, closed, and convex subset of $\mathscr{H}$. We denote by $x_n\to x$ and $x_n\rightharpoonup x$ the strong and the weak convergence of a sequence $\{x_n\}$ to a point x, respectively. The identity operator on $\mathscr{H}$ is denoted by I. For simplicity, we let $[K] := \{1,2,\dots,K\}$.

Definition 2.1. [13] A mapping $T : \mathscr{H} \to \mathscr{H}$ is said to be

(i) L-Lipschitz continuous if there exists a constant $L > 0$ such that
$$\|Tx - Ty\| \le L\|x-y\|, \;\; \forall x,y \in \mathscr{H},$$
and a strict contraction if $L \in (0,1)$;

(ii) β-strongly monotone if there exists a constant $\beta > 0$ such that
$$\langle Tx - Ty, x-y\rangle \ge \beta\|x-y\|^2, \;\; \forall x,y \in \mathscr{H};$$

(iii) β-inverse strongly monotone if there exists a constant $\beta > 0$ such that
$$\langle Tx - Ty, x-y\rangle \ge \beta\|Tx-Ty\|^2, \;\; \forall x,y \in \mathscr{H};$$

(iv) monotone if $\langle Tx - Ty, x-y\rangle \ge 0, \forall x,y \in \mathscr{H}$.

Definition 2.2. [13] Let $A : \mathscr{H} \rightrightarrows \mathscr{H}$ be a set-valued operator.

(i) A is called monotone if
$$\langle x-y, u-v\rangle \ge 0 \;\; \forall (x,u),(y,v) \in \mathscr{G}(A) := \{(x,u) \in \mathscr{H}\times\mathscr{H} : u \in Ax\}.$$

(ii) The monotone operator A is called a maximal monotone operator if $\mathscr{G}(A)$ is not properly contained in the graph of any other monotone operator.

(iii) For a maximal monotone operator A and a number $\lambda > 0$, the resolvent of A of parameter λ is the operator given by $J_{\lambda A} := (I+\lambda A)^{-1}$.

Lemma 2.1. [14] *Let $\mathscr{H}$ be a real Hilbert space. Let $A : \mathscr{H} \rightrightarrows \mathscr{H}$ be a maximal monotone operator, and let $B : \mathscr{H} \to \mathscr{H}$ be a monotone and Lipschitz continuous operator. Then the mapping $M = A+B$ is a maximal monotone operator.*

Lemma 2.2. [15] *Let $\mathscr{H}$ be a real Hilbert space. Suppose that $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ is κ-Lipschitz and β-strongly monotone over a closed and convex subset $C \subset \mathscr{H}$. Then the variational inequality problem*
$$\text{Find } \bar{u} \in C \text{ such that } \langle \mathscr{F}\bar{u}, v-\bar{u}\rangle \ge 0 \;\; \forall v \in C$$
has a unique solution $\bar{u} \in C$.

Lemma 2.3. [16] *Let $\{\Psi_n\}$ be a sequence of non-negative real numbers, $\{a_n\}$ be a sequence of real numbers in $(0,1)$ such that $\sum\limits_{n=1}^{\infty} a_n = \infty$, and $\{b_n\}$ be a sequence of real numbers. Assume that $\Psi_{n+1} \le (1-a_n)\Psi_n + a_n b_n$, $n \ge 1$. If $\limsup\limits_{k\to\infty} b_{n_k} \le 0$ for every subsequence $\{\Psi_{n_k}\}$ of $\{\Psi_n\}$ satisfying $\liminf\limits_{k\to\infty}\left(\Psi_{n_k+1} - \Psi_{n_k}\right) \ge 0$, then $\lim\limits_{n\to\infty}\Psi_n = 0$.*

3. Main Results

We are now in a position to present our iterative schemes and carry out their convergence analyses. Below are the assumptions on the control sequences of our iterative schemes.

Assumption 3.1.

(a) The sequence $\{\alpha_m\} \subset (0,1)$ satisfies $\lim\limits_{m\to\infty}\alpha_m = 0$ and $\sum\limits_{m=1}^{\infty}\alpha_m = \infty$;

(b) The sequence $\{\sigma_m\}$ satisfies $0 \le \sigma_m < \sigma$ and $\lim_{m\to\infty} \frac{\sigma_m}{\alpha_m} = 0$ for some $\sigma > 0$;
(c) For each $k \in [K]$, $A_k : \mathscr{H}_k \rightrightarrows \mathscr{H}_k$ is a maximal monotone operator and $B_k : \mathscr{H}_k \to \mathscr{H}_k$ is a monotone and L_k-Lipschitz continuous operator;
(d) The operator $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ is γ-strongly monotone and L-Lipschitz continuous;
(e) For each $k \in [K]$, $T_k : \mathscr{H} \to \mathscr{H}_k$ is a bounded linear operator with adjoint $T_k^* : \mathscr{H}_k \to \mathscr{H}$;
(f) $\mathscr{S} := \{x \in \mathscr{H} : x \in \cap_{k\in[K]} T_k^{-1}(A_k + B_k)^{-1}(0)\} \neq \emptyset$;
(g) $\tau \in (0, \frac{2\gamma}{L^2})$.

Algorithm 3.1. Tseng-type Method 1

Initialization: Let $x_0, x_1 \in \mathscr{H}$, $\mu \in (0,1)$, $\theta \ge 0$, and $\lambda_1 > 0$ be given.
Iterative steps: Calculate x_{m+1} as follows:
Step 1: Given the iterates x_{m-1} and x_m $(m \ge 1)$, choose

$$\theta_m = \begin{cases} \min\left\{\theta, \frac{\sigma_m}{\|x_m - x_{m-1}\|}\right\}, & \text{if } x_m \neq x_{m-1}, \\ \theta, & \text{otherwise.} \end{cases}$$

Compute

$$w_m = x_m + \theta_m(x_m - x_{m-1}). \tag{3.1}$$

Step 2: Compute

$$y_m^k = J_{\lambda_m}^{A_k}(T_k w_m - \lambda_m B_k T_k w_m). \tag{3.2}$$

Step 3: Select k_m such that

$$\|y_m^{k_m} - T_{k_m} w_m\| = \max_k \|y_m^k - T_k w_m\|$$

and compute

$$\begin{aligned} t_m &= y_m^{k_m} - \lambda_m (B_{k_m} y_m^{k_m} - B_{k_m} T_{k_m} w_m), \\ v_m &= w_m - \eta_m T_{k_m}^*(T_{k_m} w_m - t_m), \end{aligned} \tag{3.3}$$

where η_m is chosen so that for some $\varepsilon > 0$ small enough, $\eta_m \in \left[\varepsilon, \frac{\|T_{k_m} w_m - t_m\|^2}{\|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2} - \varepsilon\right]$ when $T_{k_m} w_m - t_m \neq 0$, else, $\eta_m = \eta > 0$.
Step 4: Compute

$$x_{m+1} = (I - \alpha_m \tau \mathscr{F}) v_m. \tag{3.4}$$

Update

$$\lambda_{m+1} = \begin{cases} \min\left\{\lambda_m, \frac{\mu\|y_m^{k_m} - T_{k_m} w_m\|}{\|B_{k_m} y_m^{k_m} - B_{k_m} T_{k_m} w_m\|}\right\}, & \text{if } B_{k_m} y_m^{k_m} \neq B_{k_m} T_{k_m} w_m, \\ \lambda_m, & \text{otherwise.} \end{cases}$$

Set $m := m + 1$ and go back to Step 1.

Remark 3.1. (1) From **Step 1** and Assumption 3.1(b), one can deduce that $\lim_{m\to\infty} \frac{\theta_m}{\alpha_m}\|x_m - x_{m-1}\| = 0$, which implies that there exists a number $M > 0$ such that $\frac{\theta_m}{\alpha_m}\|x_m - x_{m-1}\| \le M$ $\forall m \in \mathbb{N}$.

(2) The sequence $\{\lambda_m\}$ generated by Algorithm 3.1 is decreasing and we have

$$\lim_{m\to\infty}\lambda_m = \lambda \geq \min\{\lambda_1, \frac{\mu}{L}\}$$

(see, for example, [10]).

Proposition 3.1. *The sequence $\{x_n\}$ generated by Algorithm 3.1 is bounded.*

Proof. Let $x \in \mathscr{S}$. Then, for each $k \in [K]$, $0 \in (A_k + B_k)T_k x$. It follows from (3.2) that, for each $k \in [K]$,

$$T_k w_m - \lambda_m B_k T_k w_m - y_m^k + \lambda_m B_k y_m^k \in \lambda_m (A_k + B_k) y_m^k. \tag{3.5}$$

According to Lemma 2.1, the operator $(A_k + B_k)$ is maximal monotone. Therefore, it follows from (3.5) that

$$\langle T_k w_m - y_m^k - \lambda_m (B_k T_k w_m - B_k y_m^k), y_m^k - T_k x\rangle \geq 0. \tag{3.6}$$

Let $t_m := y_m^{k_m} - \lambda_m (B_{k_m} y_m^{k_m} - B_{k_m} T_{k_m} w_m)$. Then we obtain

$$\begin{aligned} \|v_m - x\|^2 &= \|w_m - \eta_m T_{k_m}^*(T_{k_m} - t_m) - x\|^2 \\ &= \|w_m - x\|^2 - 2\eta_m \langle w_m - x, T_{k_m}^*(T_{k_m} w_m - t_m)\rangle + \eta_m^2 \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2 \\ &= \|w_m - x\|^2 - 2\eta_m \langle T_{k_m} w_m - T_{k_m} x, T_{k_m} w_m - t_m\rangle + \eta_m^2 \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2. \end{aligned} \tag{3.7}$$

Moreover, using (3.6) and the identity $\langle a, b\rangle = \frac{1}{2}[\|a\|^2 + \|b\|^2 - \|a-b\|^2]$, we find that

$$\begin{aligned} \langle T_{k_m} w_m - T_{k_m} x, T_{k_m} w_m - t_m\rangle &= \langle T_{k_m} w_m - y_m^{k_m}, T_{k_m} w_m - t_m\rangle + \langle y_m^{k_m} - T_{k_m} x, T_{k_m} w_m - t_m\rangle \\ &\geq \langle T_{k_m} w_m - y_m^{k_m}, T_{k_m} w_m - t_m\rangle \\ &= \frac{1}{2}(\|T_{k_m} w_m - y_m^{k_m}\|^2 + \|T_{k_m} w_m - t_m\|^2 - \|y^{k_m} - t_m\|^2). \end{aligned} \tag{3.8}$$

From the choice of $\{\eta_m\}$ in Step 3, we see that

$$\begin{aligned} \eta_m(\|T_{k_m} w_m - t_m\|^2 - \eta_m \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2) &\geq \varepsilon \eta_m \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2 \\ &\geq \varepsilon^2 \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2. \end{aligned} \tag{3.9}$$

Substituting (3.8) and (3.9) in (3.7), we obtain

$$\begin{aligned} \|v_m - x\|^2 &\leq \|w_m - x\|^2 - \eta_m(\|T_{k_m} w_m - y_m^{k_m}\|^2 + \|T_{k_m} w_m - t_m\|^2 - \|y_m^{k_m} - t_m\|^2) \\ &\quad + \eta_m^2 \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2 \\ &= \|w_m - x\|^2 - \eta_m(\|T_{k_m} w_m - y_m^{k_m}\|^2 - \lambda_m^2 \|B_{k_m} y_m^{k_m} - B_{k_m} T_{k_m} w_m\|^2) \\ &\quad - \eta_m(\|T_{k_m} w_m - t_m\|^2 - \eta_m \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2) \\ &\leq \|w_m - x\|^2 - \eta_m(1 - \frac{\lambda_m^2 \mu^2}{\lambda_{m+1}^2})\|T_{k_m} w_m - y_m^{k_m}\|^2 - \varepsilon^2 \|T_{k_m}^*(T_{k_m} w_m - t_m)\|^2. \end{aligned} \tag{3.10}$$

Also, using [17, Lemma 2], we see that

$$\begin{aligned} \|x_{m+1} - x\| &= \|(I - \alpha_m \tau \mathscr{F})v_m - (I - \alpha_m \tau \mathscr{F})x - \alpha_n \tau \mathscr{F} x\| \\ &\leq \|(I - \alpha_m \tau \mathscr{F})v_m - (I - \alpha_m \tau \mathscr{F})x\| + \alpha_n \tau \|\mathscr{F} x\| \\ &\leq (1 - \alpha_m \nu)\|v_m - x\| + \alpha_m \tau \|\mathscr{F} x\|, \end{aligned} \tag{3.11}$$

where

$$0 < \nu := 1 - \sqrt{1 - \tau(2\gamma - \tau L_1^2)} < 1.$$

Furthermore, it follows from Remark 3.1 (2) that

$$\lim_{m\to\infty}\left(1 - \frac{\lambda_m^2}{\lambda_{m+1}^2}\mu^2\right) = 1 - \mu^2 > 1 - \mu > 0.$$

Using (3.10), we can now find $m_0 \in \mathbb{N}$ such that

$$\|v_m - x\|^2 \leq \|w_m - x\|^2 - \eta_m(1-\mu)\|T_{k_m}w_m - y_m^{k_m}\|^2 - \varepsilon^2\|T_{k_m}^*(T_{k_m}w_m - t_m)\|^2 \ \ \forall m \geq m_0. \quad (3.12)$$

In view of Remark 3.1 (1) and (3.1), we see that

$$\|w_n - x\| \leq \|x_n - x\| + \alpha_m M. \quad (3.13)$$

Therefore it follows from (3.11), (3.12), and (3.13) that

$$\begin{aligned} \|x_{m+1} - x\| &\leq (1-\alpha_n\nu)\|x_m - x\| + \alpha_m\nu\frac{(M+\tau\mathscr{F}x)}{\nu} \\ &\leq \max\left\{\|x_m - x\|, \frac{(M+\tau\mathscr{F}x)}{\nu}\right\} \\ &\ \vdots \\ &\leq \max\left\{\|x_{m_0} - x\|, \frac{(M+\tau\mathscr{F}x)}{\nu}\right\}. \end{aligned}$$

Thus, using the above analysis, we conclude that $\{\|x_m - x\|\}$ is a bounded sequence. Consequently, the sequence $\{x_m\}$ is bounded, as asserted. □

Theorem 3.1. *Under Assumption 3.1, the sequence $\{x_m\}$ converges strongly to $x^\dagger \in \mathscr{S}$, where $x^\dagger$ is the unique solution to the variational inequality problem (VIP) $\langle \mathscr{F}x^\dagger, y - x^\dagger\rangle \geq 0$ for all $y \in \mathscr{S}$.*

Proof. Note that $\mathscr{S}$ is a closed and convex subset of $\mathscr{H}$. According to Lemma 2.2, there exists a unique solution $x^\dagger \in \mathscr{S}$ to the VIP

$$\langle \mathscr{F}x^\dagger, y - x^\dagger\rangle \geq 0 \ \ \forall y \in \mathscr{S}. \quad (3.14)$$

It follows from (3.4) that

$$\begin{aligned} \|x_{m+1} - x^\dagger\|^2 &= \|(I - \alpha_m\tau\mathscr{F})(v_m - x^\dagger) - \alpha_m\tau\mathscr{F}x^\dagger\|^2 \\ &= \|(I - \alpha_m\tau\mathscr{F})(v_m - x^\dagger)\|^2 - 2\alpha_m\tau\langle(1 - \alpha_m\tau\mathscr{F})(v_m - x^\dagger), \mathscr{F}x^\dagger\rangle \\ &\quad + \alpha_m^2\tau^2\|\mathscr{F}x^\dagger\|^2 \\ &= \|(I - \alpha_m\tau\mathscr{F})(v_m - x^\dagger)\|^2 - 2\alpha_m\tau\langle\mathscr{F}x^\dagger, x_{m+1} - x^\dagger\rangle - \alpha_m^2\tau^2\|\mathscr{F}x^\dagger\|^2 \\ &\leq \|(I - \alpha_m\tau\mathscr{F})(v_m - x^\dagger)\|^2 - 2\alpha_m\tau\langle\mathscr{F}x^\dagger, x_{m+1} - x^\dagger\rangle \\ &\leq (1-\alpha_m\nu)\|v_m - x^\dagger\|^2 - 2\alpha_m\tau\langle\mathscr{F}x^\dagger, x_{m+1} - x^\dagger\rangle. \end{aligned} \quad (3.15)$$

Also, from (3.1), we obtain

$$\begin{aligned} \|w_n - x^\dagger\|^2 &= \|x_m + \theta_m(x_m - x_{m-1}) - x^\dagger\|^2 \\ &\leq \|x_m - x^\dagger\|^2 + 2\theta_m\langle x_m - x_{m-1}, w_m - x^\dagger\rangle \\ &\leq \|x_m - x^\dagger\|^2 + 2\sigma_m\|w_m - x^\dagger\|. \end{aligned} \quad (3.16)$$

Thus it follows from (3.12), (3.15), and (3.16) that, $\forall m \geq m_0$,

$$\begin{aligned}\|x_{m+1}-x^\dagger\|^2 &\leq (1-\alpha_m\nu)\|x_m-x^\dagger\|^2-\eta_m(1-\mu)(1-\alpha_m\nu)\|T_{k_m}w_m-y_m^{k_m}\|^2\\ &\quad -\varepsilon^2(1-\alpha_m\nu)\|T_{k_m}^*(T_{k_m}w_m-t_m)\|^2+2\alpha_m\tau\langle\mathscr{F}x^\dagger,x^\dagger-x_{m+1}\rangle\\ &\quad +2\sigma_m\|w_m-x^\dagger\| \qquad (3.17)\\ &\leq (1-\alpha_m\nu)\|x_m-x^\dagger\|^2+\alpha_m\nu\left(\frac{2\tau\langle\mathscr{F}x^\dagger,x^\dagger-x_{m+1}\rangle}{\nu}+2\frac{\sigma_m\|w_m-x^\dagger\|}{\alpha_m\nu}\right).\end{aligned}$$

At this juncture, we define $\Psi_m := \|x_m-x^\dagger\|^2$ and $\Gamma_m := \left(\frac{2\tau\langle\mathscr{F}x^\dagger,x^\dagger-x_{m+1}\rangle}{\nu}+2\frac{\sigma_m\|w_m-x^\dagger\|}{\alpha_m\nu}\right)$. Hence, the last inequality becomes

$$\Psi_{m+1} \leq (1-\alpha_m\nu)\Psi_m+\alpha_m\nu\Gamma_m. \qquad (3.18)$$

By Assumption 3.1 (b) and Proposition 3.1, one sees that $\{\Gamma_m\}$ is bounded. In order to use Lemma 2.3, one assumes that $\{\Psi_{m_j}\}$ is a subsequence of $\{\Psi_m\}$ such that $\liminf_{j\to\infty}(\Psi_{m_{j+1}}-\Psi_{m_j}) \geq 0$. Then it follows from (3.17) that

$$\begin{aligned}\limsup_{j\to\infty}(\eta_{m_j}(1-\mu)\|T_{k_m}w_{m_j}-y_{m_j}^{k_m}\|^2 &+ \varepsilon^2\|T_{k_m}^*(T_{k_m}w_{m_j}-t_{m_j})\|^2)\\ &\leq \limsup_{j\to\infty}((\Psi_{m_j}-\Psi_{m_j+1})+\alpha_{m_j}\nu(\Gamma_{m_j}-\Psi_{m_j})\\ &\leq -\liminf_{j\to\infty}(\Psi_{m_j+1}-\Psi_{m_j})\\ &\leq 0. \qquad (3.19)\end{aligned}$$

Using (3.19), we infer that

$$\lim_{j\to\infty}\|T_{k_m}w_{m_j}-y_{m_j}^{k_m}\|=0 \text{ and } \lim_{j\to\infty}\|T_{k_m}^*(T_{k_m}w_{m_j}-t_{m_j})\|=0. \qquad (3.20)$$

In view of the choice of k_m and (3.20), it is not difficult to see that

$$\lim_{j\to\infty}\|T_kw_{m_j}-y_{m_j}^k\|=0,\ \text{for each } k,$$

and therefore, $\lim_{j\to\infty}\|B_kT_kw_{m_j}-B_ky_{m_j}^k\|=0$. It also follows from (3.20) and (3.3) that

$$\lim_{j\to\infty}\|v_{m_j}-w_{m_j}\|=0. \qquad (3.21)$$

In view of Remark 3.1 (1) and (3.1), it is easy to see that

$$\lim_{j\to\infty}\|w_{m_j}-x_{m_j}\|=0. \qquad (3.22)$$

Combining (3.21), (3.22), and (3.4), we now obtain

$$\lim_{j\to\infty}x_{m_j+1}=\lim_{j\to\infty}v_{m_j}=\lim_{j\to\infty}w_{m_j}=\lim_{j\to\infty}x_{m_j}. \qquad (3.23)$$

Since $\{x_{m_j}\}$ is bounded, there exists a subsequence $\{x_{m_{j_i}}\}$ of $\{x_{m_j}\}$ such that $x_{m_{j_i}} \rightharpoonup \tilde{x}$ as $i\to\infty$ and

$$\begin{aligned}\limsup_{j\to\infty}\langle\mathscr{F}x^\dagger,x^\dagger-x_{m_j+1}\rangle &= \lim_{i\to\infty}\langle\mathscr{F}x^\dagger,x^\dagger-x_{m_{j_i}+1}\rangle\\ &= \langle\mathscr{F}x^\dagger,x^\dagger-\tilde{x}\rangle. \qquad (3.24)\end{aligned}$$

From (3.20) and (3.22), it follows that $y_{m_{j_i}}^k \rightharpoonup T_k\tilde{x}$ for each $k\in[K]$ and as $i\to\infty$.

Now, let $(\tilde{e},\tilde{f}) \in \mathscr{G}(A_k+B_k)$. Invoking (3.2), we obtain, for each $k \in [K]$,

$$T_k w_{m_{j_i}} - \lambda_{m_{j_i}} B_k T_k w_{m_{j_i}} - y^k_{m_{j_i}} + \lambda_{m_{j_i}} B_k y^k_{m_{j_i}} \in \lambda_{m_{j_i}} (A_k+B_k) y^k_{m_{j_i}}.$$

Since (A_k+B_k) is a monotone operator, it follows that

$$\langle \tilde{f} - \frac{T_k w_{m_{j_i}} - \lambda_{m_{j_i}} B_k T_k w_{m_{j_i}} - y^k_{m_{j_i}} + \lambda_{m_{j_i}} B_k y^k_{m_{j_i}}}{\lambda_{m_{j_i}}}, \tilde{e} - y_{m_{j_i}} \rangle \geq 0. \tag{3.25}$$

Taking the limit as $i \to \infty$ in (3.25), we then obtain $\langle \tilde{f}, \tilde{e} - T_k \tilde{x} \rangle \geq 0$. Since each $A_k + B_k$ is a maximal monotone operator, it follows that $\tilde{x} \in \mathscr{S}$. Consequently, it follows from Assumption 3.1 (b), (3.14), and (3.24) that $\limsup_{j\to\infty} \Gamma_{m_j} \leq 0$. Therefore, applying Lemma 2.3 to (3.18), we infer that $x_m \to x^\dagger$ as $m \to \infty$, as asserted. □

In our next method, we consider the possibility of performing regularization and an inertial procedure simultaneously at the first step of the iterative algorithm. This approach was motivated by a recent study of Reich and Taiwo [10].

Algorithm 3.2. Tseng-type Method 2

Initialization: Let $x_0, x_1 \in \mathscr{H}$, $\mu \in (0,1)$, $\theta \geq 0$, and $\lambda_1 > 0$ be given.
Iterative steps: Calculate x_{m+1} as follows:
Step 1: Given the iterates x_{m-1} and x_m $(n \geq 1)$, choose

$$\theta_m = \begin{cases} \min\left\{\theta, \frac{\sigma_m}{\|x_m - x_{m-1}\|}\right\}, & \text{if } x_m \neq x_{m-1}, \\ \theta, & \text{otherwise.} \end{cases}$$

Compute

$$w_m = (I - \alpha_m \tau \mathscr{F}) x_m + \theta_m (x_m - x_{m-1}). \tag{3.26}$$

Step 2: Compute

$$y^k_m = J^{A_k}_{\lambda_m} (T_k w_m - \lambda_m B_k T_k w_m). \tag{3.27}$$

Step 3: Select k_m such that

$$\|y^{k_m}_m - T_{k_m} w_m\| = \max_k \|y^k_m - T_k w_m\|$$

and compute

$$\begin{aligned} t_m &= y^{k_m}_m - \lambda_m (B_{k_m} y^{k_m}_m - B_{k_m} T_{k_m} w_m), \\ x_{m+1} &= w_m - \eta_m T^*_{k_m} (T_{k_m} w_m - t_m), \end{aligned} \tag{3.28}$$

where η_m is chosen so that for some $\varepsilon > 0$ small enough, $\eta_m \in \left[\varepsilon, \frac{\|T_{k_m} w_m - t_m\|^2}{\|T^*_{k_m}(T_{k_m} w_m - t_m)\|^2} - \varepsilon\right]$ when $T_{k_m} w_m - t_m \neq 0$, else, $\eta_m = \eta > 0$.
Update

$$\lambda_{m+1} = \begin{cases} \min\left\{\lambda_m, \frac{\mu \|y^{k_m}_m - T_{k_m} w_m\|}{\|B_{k_m} y^{k_m}_m - B_{k_m} T_{k_m} w_m\|}\right\}, & \text{if } B_{k_m} y^{k_m}_m \neq B_{k_m} T_{k_m} w_m, \\ \lambda_m, & \text{otherwise.} \end{cases}$$

Set $m := m+1$ and go back to Step 1.

Proposition 3.2. *The sequence $\{x_m\}$ generated by Algorithm 3.2 is bounded.*

Proof. Since the proof is similar to the proof of Proposition 3.1, we only sketch it. Let $t_m := y_m^{k_m} - \lambda_m(B_{k_m}y_m^{k_m} - B_{k_m}T_{k_m}w_m)$ and $x \in \mathscr{S}$. Then, using (3.7) – (3.10) and (3.12), we find that

$$\begin{aligned}\|x_{m+1}-x\|^2 &\leq \|w_m-x\|^2 - \eta_m(1-\mu)(1-\alpha_m\nu)\|T_{k_m}w_m - y_m^{k_m}\|^2 \\ &\quad -\varepsilon^2(1-\alpha_m\nu)\|T_{k_m}^*(T_{k_m}w_m - t_m)\|^2 \ \forall m \geq m_0. \end{aligned} \tag{3.29}$$

In addition, employing a similar argument to the one that has led to (3.11), we find that

$$\begin{aligned}\|w_m-x\| &\leq \|(I-\alpha_m\tau\mathscr{F})x_m - x\| + \theta_m\|x_m-x_{m-1}\| \\ &\leq (1-\alpha_m\nu)\|x_m-x\| + \alpha_m\tau\|\mathscr{F}x\| + \theta_m\|x_m-x_{m-1}\|, \end{aligned} \tag{3.30}$$

where

$$0 < \nu := 1-\sqrt{1-\tau(2\gamma-\tau L_1^2)} < 1.$$

Therefore, substituting (3.30) in (3.29), we see that

$$\begin{aligned}\|x_{m+1}-x\| &\leq (1-\alpha_m\nu)\|x_m-x\| + \alpha_m\nu\left(\frac{\tau\|\mathscr{F}x\|}{\nu} + \frac{\theta_m\|x_m-x_{m-1}\|}{\nu\alpha_m}\right) \\ &\leq (1-\alpha_m\nu)\|x_m-x\| + \alpha_m\nu\left(\frac{\tau\|\mathscr{F}x\|}{\nu} + \frac{M}{\nu}\right) \\ &\leq \max\left\{\|x_m-x\|, \frac{\tau\|\mathscr{F}x\|}{\nu} + \frac{M}{\nu}\right\} \\ &\ \vdots \\ &\leq \max\left\{\|x_{m_0}-x\|, \frac{\tau\|\mathscr{F}x\|}{\nu} + \frac{M}{\nu}\right\}. \end{aligned}$$

Therefore, using the above inequality, we infer that $\{\|x_m - x\|\}$ and $\{x_m\}$ are bounded. □

Next, we establish the strong convergence of the sequences generated by Algorithm 3.2.

Theorem 3.2. *Under Assumption 3.1, the sequence $\{x_m\}$ generated by Algorithm 3.2 converges strongly to $x^\dagger \in \mathscr{S}$, where $x^\dagger$ is the unique solution to the variational inequality problem (VIP) $\langle \mathscr{F}x^\dagger, y - x^\dagger\rangle \geq 0$ for all $y \in \mathscr{S}$.*

Proof. Let $x^\dagger \in \mathscr{S}$ be the unique solution to the VIP

$$\langle \mathscr{F}x^\dagger, y - x^\dagger\rangle \geq 0 \ \ \forall y \in \mathscr{S}.$$

Define $b_m := (I - \alpha_m\tau\mathscr{F})x_m$. Using a similar argument to the one that has led to (3.15), we obtain

$$\|b_m - x^\dagger\|^2 \leq (1-\alpha_m\nu)\|x_m - x^\dagger\|^2 - 2\alpha_m\tau\langle \mathscr{F}x^\dagger, b_m - x^\dagger\rangle. \tag{3.31}$$

Moreover,

$$\begin{aligned}\langle b_m - x^\dagger, \theta_m(x_m - x_{m-1})\rangle &= \langle (I-\alpha_m\tau\mathscr{F})x_m - x^\dagger, \theta_m(x_m-x_{m-1})\rangle \\ &= \langle x_m - x^\dagger, \theta_m(x_m-x_{m-1})\rangle - \alpha_m\tau\langle \mathscr{F}x_m, \theta_m(x_m-x_{m-1})\rangle \\ &\leq \theta_m\|x_m-x_{m-1}\|\big(\|x_m-x^\dagger\| + \alpha_m\tau\|\mathscr{F}x_m\|\big) \\ &\leq \sigma_m\big(\|x_m-x^\dagger\| + \alpha_m\tau\|\mathscr{F}x_m\|\big). \end{aligned} \tag{3.32}$$

Furthermore, using (3.31) and (3.32), we have

$$\begin{aligned}\|w_m - x^\dagger\|^2 &= \|b_m - x^\dagger + \theta_m(x_m - x_{m-1})\|^2 \\ &= \|b_m - x^\dagger\|^2 + 2\langle b_m - x^\dagger, \theta_m(x_m - x_{m-1})\rangle + \theta_m^2\|x_m - x_{m-1}\|^2 \\ &\le (1-\alpha_m \nu)\|x_m - x^\dagger\|^2 - 2\alpha_m\tau\langle \mathscr{F}x^\dagger, b_m - x^\dagger\rangle \\ &\quad + 2\sigma_m\big(\|x_m - x^\dagger\| + \alpha_m\tau\|\mathscr{F}x_m\|\big) + \sigma_m^2. \end{aligned} \tag{3.33}$$

Thus, substituting (3.33) in (3.29), we find that

$$\begin{aligned}\|x_{m+1} &- x^\dagger\|^2 \\ &\le (1-\alpha_m\nu)\|x_m - x^\dagger\|^2 - 2\alpha_m\tau\langle\mathscr{F}x^\dagger, b_m - x^\dagger\rangle + 2\sigma_m\big(\|x_m - x^\dagger\| + \alpha_m\tau\|\mathscr{F}x_m\|\big) \\ &\quad + \sigma_m^2 - \eta_m(1-\mu)(1-\alpha_m\nu)\|T_{k_m}w_m - y_m^{k_m}\|^2 - \varepsilon^2(1-\alpha_m\nu)\|T_{k_m}^*(T_{k_m}w_m - t_m)\|^2 \\ &\le (1-\alpha_m\nu)\|x_m - x^\dagger\|^2 \\ &\quad + \alpha_m\nu\left(\frac{2\tau\langle\mathscr{F}x^\dagger, x^\dagger - b_m\rangle}{\nu} + 2\frac{\sigma_m\big(\|x_m - x^\dagger\| + \alpha_m\tau\|\mathscr{F}x_m\|\big)}{\alpha_m\nu} + \frac{\sigma_m^2}{\alpha_m\nu}\right) \\ &\quad ,\forall m \ge m_0. \end{aligned} \tag{3.34}$$

At this point, we define $\Psi_m := \|x_m - x^\dagger\|^2$ and $\tilde{\Gamma}_m := \left(\frac{2\tau\langle\mathscr{F}x^\dagger, x^\dagger - b_m\rangle}{\nu} + 2\frac{\sigma_m\big(\|x_m - x^\dagger\| + \alpha_m\tau\|\mathscr{F}x_m\|\big)}{\alpha_m\nu} + \frac{\sigma_m^2}{\alpha_m\nu}\right)$. Hence, (3.34) becomes

$$\Psi_{m+1} \le (1-\alpha_m\nu)\Psi_m + \alpha_m\nu\tilde{\Gamma}_m \;\; \forall m \ge m_0. \tag{3.35}$$

Therefore, by following similar arguments to those used in the proof of Theorem 3.1, starting from (3.18), we find that $x_m \to x^\dagger$, as asserted. □

The following corollary is an immediate consequence of our main results.

Corollary 3.1. *Suppose that $\mathscr{H}$ is a real Hilbert space, $A : \mathscr{H} \rightrightarrows \mathscr{H}$ is a maximal monotone operator, and $B : \mathscr{H} \to \mathscr{H}$ is a monotone and L_1-Lipschitz continuous operator. Let $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ be a γ-strongly monotone and L-Lipschitz continuous operator, and assume that $\mathscr{S} := \{x \in \mathscr{H} : (A+B)^{-1}(0)\} \neq \emptyset$. Then the sequences generated by Algorithms 3.3 and 3.4 converge strongly to $x^\dagger \in \mathscr{S}$, where $x^\dagger$ is the unique solution to the variational inequality problem (VIP) $\langle\mathscr{F}x^\dagger, y - x^\dagger\rangle \ge 0$ for all $y \in \mathscr{S}$.*

Algorithm 3.3.

Initialization: Let $x_0, x_1 \in \mathscr{H}$, $\mu \in (0,1)$, $\theta \ge 0$, and $\lambda_1 > 0$ be given.
Iterative steps: Calculate x_{m+1} as follows:
Step 1: Given the iterates x_{m-1} and x_m $(m \ge 1)$, choose

$$\theta_m = \begin{cases} \min\left\{\theta, \frac{\sigma_m}{\|x_m - x_{m-1}\|}\right\}, & \text{if } x_m \neq x_{m-1}, \\ \theta, & \text{otherwise.} \end{cases}$$

Compute

$$w_m = x_m + \theta_m(x_m - x_{m-1}).$$

Step 2: Compute

$$y_m = J_{\lambda_m}^A(w_m - \lambda_m B w_m).$$

Step 3: Compute

$$v_m = (1-\eta_m)w_m + \eta_m(y_m - \lambda_m(By_m - Bw_m)),$$

where η_m is chosen so that $\eta_m \in [\varepsilon, 1-\varepsilon]$, for some $\varepsilon > 0$ small enough.
Step 4: Compute

$$x_{m+1} = (I - \alpha_m \tau \mathscr{F})v_m.$$

Update

$$\lambda_{m+1} = \begin{cases} \min\left\{\lambda_m, \frac{\mu\|y_m - w_m\|}{\|By_m - Bw_m\|}\right\}, & \text{if } By_m \neq Bw_m, \\ \lambda_m, & \text{otherwise.} \end{cases}$$

Set $m := m+1$ and go back to Step 1.

Algorithm 3.4.
Initialization: Let $x_0, x_1 \in \mathscr{H}$, $\mu \in (0,1)$, $\theta \geq 0$, and $\lambda_1 > 0$ be given.
Iterative steps: Calculate x_{m+1} as follows:
Step 1: Given the iterates x_{m-1} and x_m $(n \geq 1)$, choose

$$\theta_m = \begin{cases} \min\left\{\theta, \frac{\sigma_m}{\|x_m - x_{m-1}\|}\right\}, & \text{if } x_m \neq x_{m-1}, \\ \theta, & \text{otherwise.} \end{cases}$$

Compute

$$w_m = (I - \alpha_m \tau \mathscr{F})x_m + \theta_m(x_m - x_{m-1}).$$

Step 2: Compute

$$y_m = J^A_{\lambda_m}(w_m - \lambda_m Bw_m).$$

Step 3: Compute

$$x_{m+1} = (1-\eta_m)w_m + \eta_m(y_m - \lambda_m(By_m - Bw_m)),$$

where η_m is chosen so that $\eta_m \in [\varepsilon, 1-\varepsilon]$, for some $\varepsilon > 0$ small enough.
Update

$$\lambda_{m+1} = \begin{cases} \min\left\{\lambda_m, \frac{\mu\|y_m - w_m\|}{\|By_m - Bw_m\|}\right\}, & \text{if } By_m \neq Bw_m, \\ \lambda_m, & \text{otherwise.} \end{cases}$$

Set $m := m+1$ and go back to Step 1.

Proof. Let $k = 1$, $\mathscr{H} = \mathscr{H}_1$ and $T = I$ in Assumption 3.1. Then Algorithms 3.3 and 3.4 reduce to Algorithms 3.1 and 3.1, respectively. Therefore, the proof follows from Theorems 3.1 and 3.2. □

4. An Application and A Numerical Example

4.1. **Composite Bilevel Optimization Problem.** For $k \in [K]$ and $n_k \in \mathbb{N}$, let

$$T_k : \mathbb{R}^m \to \mathbb{R}^{n_k} \text{ be a bounded linear operator with the adjoint } T_k^* : \mathbb{R}^{n_k} \to \mathbb{R}^m,$$

and let $\varphi_k : \mathbb{R}^{n_k} \to \mathbb{R} \cup \{+\infty\}$ be defined by $\varphi_k(x) := f_k(x) + g_k(x)$, where $f_k : \mathbb{R}^{n_k} \to \mathbb{R}$ is a continuously differentiable function with an L_k-Lipschitz continuous gradient ∇f_k and $g_k : \mathbb{R}^{n_k} \to \mathbb{R} \cup \{+\infty\}$ is closed and convex, but not necessarily differentiable. In addition, let $h : \mathbb{R}^m \to \mathbb{R}$ be a γ-strongly convex and differentiable function with an L-Lipschitz continuous gradient ∇h. We consider the following convex bilevel optimization problem (CBOP):

$$\min_{x \in X^* \subset \mathbb{R}^m} h(x), \tag{4.1}$$

where X^* is the solution set of the composite optimization problem

$$\cap_{k \in [K]} T_k^{-1}(\arg \min_{x \in \mathbb{R}^{n_k}} \varphi_k(x)). \tag{4.2}$$

In optimization terminology, (4.1) is called the outer level problem and (4.2) is said to be the inner level problem. This formulation generalizes some bilevel optimization problems studied in the literature. For the case where $K = 1$, $m = n$, and T_k is the identity operator, we refer interested readers to [18, 19, 20] and some of the references therein for studies related to the CBOP (4.1)–(4.2). In our framework, by the optimality condition, finding X^* is equivalent to solving the inclusion problem (see [21, Theorem 3.43])

$$x \in \cap_{k=1}^{K} T_k^{-1}(\nabla f_k + \partial g_k)^{-1}(0). \tag{4.3}$$

Furthermore, solving the outer level problem (4.1) is equivalent to solving the following variational inequality problem (see, for example, [13, Prop. 27.8, p. 501]):

$$\text{Find } x \in X^* \text{ such that } \langle \nabla h(x), y - x \rangle \geq 0 \ \ \forall y \in X^*. \tag{4.4}$$

Thus, by combining (4.3) and (4.4), we find that solving (4.1)–(4.2) is equivalent to solving the following inclusion problem

$$\text{Find } x \in \mathbb{R}^m \text{ such that } x \in X^* = \cap_{k=1}^{K} T_k^{-1}(\nabla f_k + \partial g_k)^{-1}(0) \text{ and } \langle \nabla h(x), y - x \rangle \geq 0 \ \forall y \in X^*. \tag{4.5}$$

Our next theorem is an application of our main results to the composite bilevel optimization problem (4.1)–(4.2).

Theorem 4.1. *Assume that the conditions of Assumption 3.1 hold. Under the data in Section 4.1, suppose $X^* \neq \emptyset$. Then the sequences generated by Algorithms 4.1 and 4.2 converge to a point $x^* \in X^*$, which satisfies*

$$\langle \nabla h(x^*), x - x^* \rangle \geq 0 \, \forall x \in X^*,$$

and x^* is the unique optimal solution to the outer level problem (4.1).

Algorithm 4.1. Tseng-type Method 1 for solving the BOP (4.1)–(4.2).

Initialization: Let $x_0, x_1 \in \mathscr{H}$, $\mu \in (0,1)$, $\theta \geq 0$, and $\lambda_1 > 0$ be given.

Iterative steps: Calculate x_{m+1} as follows:

Step 1: Given the iterates x_{m-1} and x_m $(m \geq 1)$, choose

$$\theta_m = \begin{cases} \min\left\{\theta, \frac{\sigma_m}{\|x_m - x_{m-1}\|}\right\}, & \text{if } x_m \neq x_{m-1}, \\ \theta, & \text{otherwise.} \end{cases}$$

Compute

$$w_m = x_m + \theta_m(x_m - x_{m-1}).$$

Step 2: Compute

$$y_m^k = J_{\lambda_m}^{\partial g_k}(T_k w_m - \lambda_m \nabla f_k T_k w_m).$$

Step 3: Select k_m such that

$$\|y_m^{k_m} - T_{k_m} w_m\| = \max_k \|y_m^k - T_k w_m\|$$

and compute

$$v_m = w_m - \eta_m T_{k_m}^*(T_{k_m} w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m)),$$

where η_m is chosen so that for some $\varepsilon > 0$ small enough,

$$\eta_m \in \left[\varepsilon, \frac{\|T_{k_m} w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m)\|^2}{\|T_{k_m}^*(T_{k_m} w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m))\|^2} - \varepsilon\right]$$

when $T_{k_m} w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m) \neq 0$, else, $\eta_m = \eta > 0$.

Step 4: Compute

$$x_{m+1} = (I - \alpha_m \tau \nabla h) v_m.$$

Update

$$\lambda_{m+1} = \begin{cases} \min\left\{\lambda_m, \frac{\mu\|y_m^{k_m} - T_{k_m} w_m\|}{\|\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m\|}\right\}, & \text{if } \nabla f_{k_m} y_m^{k_m} \neq \nabla f_{k_m} T_{k_m} w_m, \\ \lambda_m, & \text{otherwise.} \end{cases}$$

Set $m := m + 1$ and go back to Step 1.

Algorithm 4.2. Tseng-type Method 2 for solving the BOP (4.1)–(4.2).

Initialization: Let $x_0, x_1 \in \mathscr{H}$, $\mu \in (0,1)$, $\theta \geq 0$, and $\lambda_1 > 0$ be given.
Iterative steps: Calculate x_{m+1} as follows:
Step 1: Given the iterates x_{m-1} and x_m $(n \geq 1)$, choose

$$\theta_m = \begin{cases} \min\left\{\theta, \frac{\sigma_m}{\|x_m - x_{m-1}\|}\right\}, & \text{if } x_m \neq x_{m-1}, \\ \theta, & \text{otherwise.} \end{cases}$$

Compute

$$w_m = (I - \alpha_m \tau \nabla h) x_m + \theta_m(x_m - x_{m-1}).$$

Step 2: Compute

$$y_m^k = J_{\lambda_m}^{\partial g_k}(T_k w_m - \lambda_m \nabla f_k T_k w_m).$$

Step 3: Select k_m such that

$$\|y_m^{k_m} - T_{k_m} w_m\| = \max_k \|y_m^k - T_k w_m\|$$

and compute

$$x_{m+1} = w_m - \eta_m T_{k_m}^*(T_{k_m} w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m)),$$

where η_m is chosen so that for some $\varepsilon > 0$ small enough,

$$\eta_m \in \left[\varepsilon, \frac{\|T_{k_m} w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m)\|^2}{\|T_{k_m}^*(T_{k_m} w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m} y_m^{k_m} - \nabla f_{k_m} T_{k_m} w_m))\|^2} - \varepsilon\right]$$

when $T_{k_m}w_m - y_m^{k_m} + \lambda_m(\nabla f_{k_m}y_m^{k_m} - \nabla f_{k_m}T_{k_m}w_m) \neq 0$, else, $\eta_m = \eta > 0$.
Update

$$\lambda_{m+1} = \begin{cases} \min\left\{\lambda_m, \frac{\mu\|y_m^{k_m} - T_{k_m}w_m\|}{\|\nabla f_{k_m}y_m^{k_m} - \nabla f_{k_m}T_{k_m}w_m\|}\right\}, & \text{if } \nabla f_{k_m}y_m^{k_m} \neq \nabla f_{k_m}T_{k_m}w_m, \\ \lambda_m, & \text{otherwise.} \end{cases}$$

Set $m := m+1$ and go back to Step 1.

Proof. The proof can be derived from the proof of Theorems 3.1 and 3.2 by setting $\mathscr{S} = X^*$, $A_k = \partial g_k$, and $B_k = \nabla f_k$ in Algorithms 3.1 and 3.2, respectively. □

4.2. **Numerical Illustration.** Let $\varphi_1 : \mathbb{R}^2 \to \mathbb{R}$, $\varphi_2 : \mathbb{R}^3 \to \mathbb{R}$, and $\varphi_3 : \mathbb{R}^4 \to \mathbb{R}$ be defined by

$$\varphi_1(a) := \|a\|_2^2 - (3,7)\cdot a + \|a\|_1,\ a \in \mathbb{R}^2,$$

$$\varphi_2(b) := \|b\|_2^2 + (1,-3,-5)\cdot b + 3 + \|b\|_1,\ b \in \mathbb{R}^3,$$

and

$$\varphi_3(c) := \frac{1}{2}\|c\|_2^2 + (0,-1,1,4)\cdot c,\ c \in \mathbb{R}^4,$$

respectively, where '$\cdot$' stands for the scalar dot product, $\|\cdot\|_2$ is the Euclidean norm and $\|\cdot\|_1$ is the ℓ_1-norm. We define the function $h : \mathbb{R}^5 \to \mathbb{R}$ by

$$h(x) := \frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2} + \frac{x_4^2}{2} + \frac{x_5^2}{2} - x_1 + x_2 + 5 - 3x_5.$$

For $k = 1,2,3$, let $T_k : \mathbb{R}^5 \to \mathbb{R}^{k+1}$ be defined by

$$T_1x = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} x,$$

$$T_2x = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 \\ 0 & 1 & 0 & 3 & 1 \end{pmatrix} x,$$

and

$$T_3x = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 \end{pmatrix} x$$

for each $x \in \mathbb{R}^5$. Our aim is to solve the BOP (4.1)–(4.2) for the above data. It can be verified that $(0,1,2) = \arg\min_{b\in\mathbb{R}^3} \varphi_2(b)$ [10, Ex. 5.2]. Similarly, $(1,3) = \arg\min_{a\in\mathbb{R}^2} \varphi_1(a)$, and $(0,1,-1,-4) = \arg\min_{c\in\mathbb{R}^4} \varphi_3(c)$. Furthermore, $(1,-1,0,0,3) = \cap_{k=1}^3 T_k^{-1}(\arg\min_{x\in\mathbb{R}^{k+1}} \varphi_k(x))$ and it turns out that the minimum value of h is attained at $x = (1,-1,0,0,3)$. We apply our Algorithms 3.1 and 3.2 to find the solution to the BOP using the following initial points.
Case a: $x_0 = (-3,7,6,0,1)$ and $x_1 = (1,0,8,2,-3)$;
Case b: $x_0 = (2,2,-1,3,-5)$ and $x_1 = (9,3,-4,1,2)$;
Case c: $x_0 = (0,2,9,15,-5)$ and $x_1 = (23,3,-23,1,6)$;
Case d: $x_0 = (-25,22,-6,1,11)$ and $x_1 = (26,3,-4,7,0)$.

The control parameters for the implementation of our algorithms are chosen as $\alpha_m = \frac{1}{m+2}$, $\sigma_m = \frac{1}{m^2+1}$, $\mu = 0.5$, $\tau = 0.4$, $\eta = 0.3$, and $\lambda_1 = 0.2$. The codes for the algorithms are written in MATLAB 2021b and run on an HP Laptop Windows 10 with Intel(R) Core(TM) i5 CPU and 4GB RAM with the stopping criterion given by $E_n = \|(1,-1,0,0,3) - x_{n+1}\|^2 \le 10^{-6}$. Figure 1 and Table 1 provide the numerical results we obtained for four different choices of the initial values.

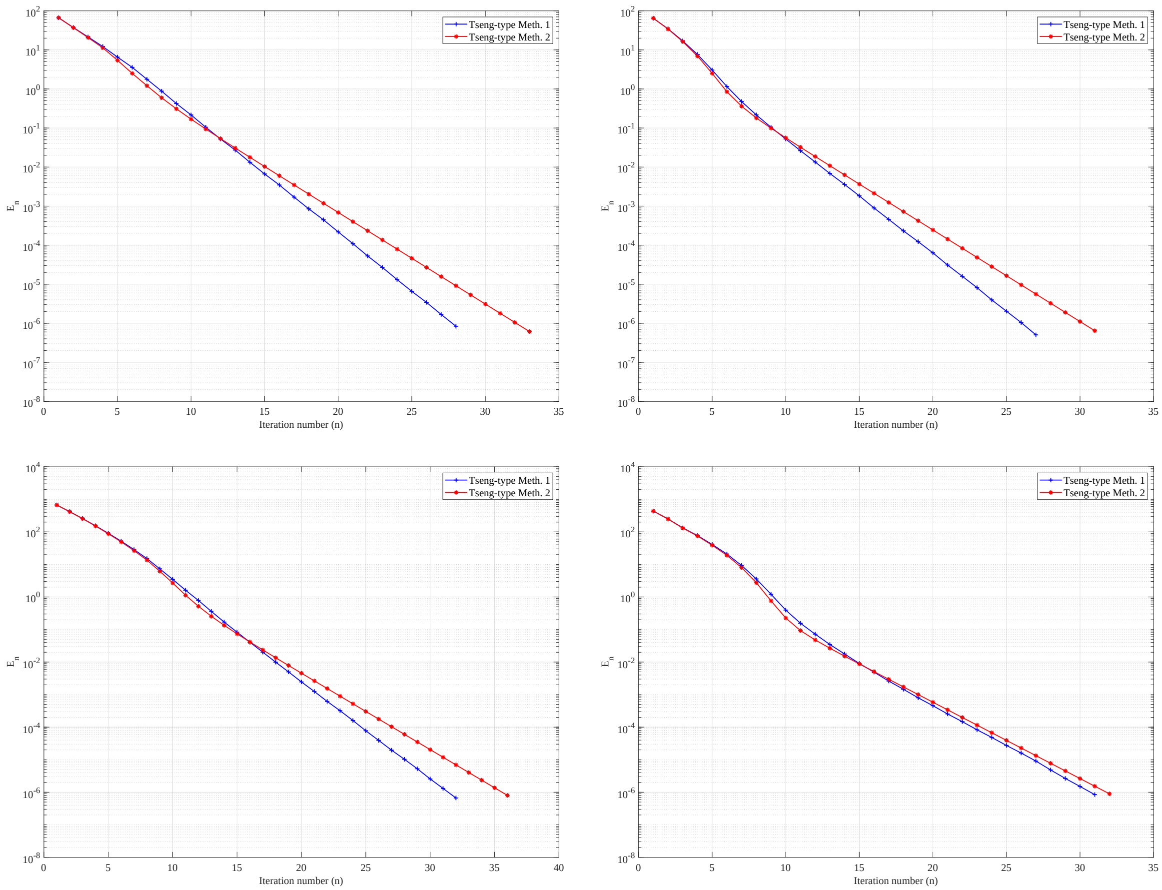

FIGURE 1. Top left: Case a; Top right: Case b; Bottom left: Case c; Bottom right: Case d.

5. CONCLUSIONS

In this paper, we introduced two Tseng-type methods for solving variational inequalities over the solution sets of split variational inclusion problems in the setting of a real Hilbert space. We proved strong convergence theorems for these methods, and provided an application and numerical illustrations of our results. These numerical examples confirm the efficiency and applicability of our methods.

TABLE 1. Numerical results.

		Tseng-type Meth. 1	Tseng-type Meth.2
Case a	CPU time (sec)	0.0428	0.0576
	No. of Iter.	28	33
	Last iterate	(0.9996, -0.9997, 5.3726e-4, -8.6871e-6, 2.9994)	(0.9998, -0.9998, 5.2018e-4, -8.4189e-6, 2.9995)
Case b	CPU time (sec)	0.0029	0.0042
	No. of Iter.	27	31
	Last iterate	(1.0003, -1.0002, -4.4425e-4, -4.8289e-5, 3.0004)	(-1.0000, -1.0000, -5.6853e-4, -9.1023e-5, 3.0006)
Case c	CPU time (sec)	0.0025	0.0031
	No. of Iter.	32	36
	Last iterate	(1.0003, -1.0002, -5.5950e-4, -2.3154e-5, 3.0005)	(1.0000, -1.0000, -7.8448e-4, -4.6932e-5, 3.0004)
Case d	CPU time (sec)	0.0022	0.0023
	No. of Iter.	31	32
	Last iterate	(1.0002, -1.0004, 1.3695e-4, 1.0908e-4, 3.0008)	(1.0004, -1.0004, 5.3389e-4, 2.4871e-5, 3.0005)

Acknowledgments
Adeolu Taiwo is grateful to the Department of Mathematics at the Technion – Israel Institute of Technology for granting him a Postdoctoral Research Fellowship. Simeon Reich was partially supported by the Israel Science Foundation (Grant No. 820/17), by the Fund for the Promotion of Research at the Technion (Grant 2001893), and by the Technion General Research Fund (Grant 2016723).

REFERENCES

[1] S. Reich, M.T. Truong, T.N.H. Mai, The split feasibility problem with multiple output sets in Hilbert spaces, Optim. Lett. 14 (2020), 2335–2353.
[2] S. Reich, T.M. Tuyen, Two new self-adaptive algorithms for solving the split common null point problem with multiple output sets in Hilbert spaces, J. Fixed Point Theory Appl. 23 (2021), Paper No. 16.
[3] S. Reich, T.M. Tuyen, N.S. Ha, Variational inequalities over the solution sets of split variational inclusion problems, Appl. Numer. Math. 192 (2023), 319–336.
[4] A. Taiwo, S. Reich, I. Chinedu, Strong convergence of two regularized relaxed extragradient schemes for solving the split feasibility and fixed point problem with multiple output sets, Appl Anal. 102 (2023), 5132–5156.
[5] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979), 964–979.
[6] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2000), 431–446.
[7] B.T. Polyak, Some methods of speeding up the convergence of iteration methods, USSR Comput. Math. Math. Phys. 4 (1944), 1–17.

[8] C. Izuchukwu, S. Reich, Y. Shehu, A. Taiwo, Strong convergence of forward-reflected-backward splitting methods for solving monotone inclusions with applications to image restoration and optimal control, J. Sci. Comput. 94 (2023), Paper No. 73.
[9] C. Izuchukwu, S. Reich, Y. Shehu, Relaxed inertial methods for solving the split monotone variational inclusion problem beyond co-coerciveness, Optimization 72 (2023), 607–646.
[10] S. Reich, A. Taiwo, Fast hybrid iterative schemes for solving variational inclusion problems, Math. Meth. Appl. Sci. 46 (2023), 17177–17198.
[11] F.U. Ogbuisi, Y. Shehu, A new inertial relaxed Tseng extrgradient method for solving quasi-monotone bilevel variational inequality problems in Hilbert spaces, J. Nonlinear Var. Anal. 7 (2023), 449–464.
[12] X. Li, Q.L. Dong, A. Gibali, PMiCA – Parallel multi-step inertial contracting algorithm for solving common variational inclusions, J. Nonlinear Funct. Anal. 2022 (2022), 1–12.
[13] H.H. Bauschke, P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, CMS Books in Mathematics, Springer, 2011.
[14] H. Brézis, Opérateurs Maximaux Monotones, North-Holland, 1983.
[15] I. Yamada, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. Inherently parallel algorithms in feasibility and optimization and their applications, Stud. Comput. Math. 8 (2001), 473–504.
[16] S. Saejung, P. Yotkaew, Approximation of zeros of inverse strongly monotone operators in Banach spaces, Nonlinear Anal. 75 (2012), 742–750.
[17] P.K. Anh, T.V. Anh, L.D. Muu, On bilevel split pseudomonotone variational inequality problems with applications, Acta Math. Vietnam. 42 (2017), 413–429.
[18] S. Sabach, S. Shtern, A first order method for solving convex bilevel optimization problems, SIAM J. Optim. 27 (2017), 640–660.
[19] Y. Shehu, P.T. Vuong, A. Zemkoho, An inertial extrapolation method for convex simple bilevel optimization, Optim. Meth. Softw. 36 (2021), 1–19.
[20] A. Taiwo, S. Reich, Two regularized inertial Tseng methods for solving inclusion problems with applications to convex bilevel programming, Optimization, in press. doi: 10.1080/02331934.2023.2284970
[21] A. Beck, First-Order Methods in Optimization, Society for Industrial and Applied Mathematics, Philadelphia, 2017.