



# Article A Regularized Tseng Method for Solving Various Variational Inclusion Problems and Its Application to a Statistical Learning Model

Adeolu Taiwo \* and Simeon Reich \*

Department of Mathematics, The Technion—Israel Institute of Technology, 3200003 Haifa, Israel \* Correspondence: taiwoa@campus.technion.ac.il or taiwo.adeolu@yahoo.com (A.T.); sreich@technion.ac.il (S.R.)

Abstract: We study three classes of variational inclusion problems in the framework of a real Hilbert space and propose a simple modification of Tseng's forward-backward-forward splitting method for solving such problems. Our algorithm is obtained via a certain regularization procedure and uses self-adaptive step sizes. We show that the approximating sequences generated by our algorithm converge strongly to a solution of the problems under suitable assumptions on the regularization parameters. Furthermore, we apply our results to an elastic net penalty problem in statistical learning theory and to split feasibility problems. Moreover, we illustrate the usefulness and effectiveness of our algorithm by using numerical examples in comparison with some existing relevant algorithms that can be found in the literature.

Keywords: Hilbert space; inclusion problem; statistical learning; strong convergence; Tseng algorithm

MSC: 46N10; 47H05; 47J25; 62J07



Citation: Taiwo, A.; Reich, S. A Regularized Tseng Method for Solving Various Variational Inclusion Problems and Its Application to a Statistical Learning Model. *Axioms* 2023, *12*, 1037. https://doi.org/ 10.3390/axioms12111037

Academic Editor: Behzad Djafari-Rouhani

Received: 28 September 2023 Revised: 30 October 2023 Accepted: 30 October 2023 Published: 6 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

# 1. Introduction

Variational inclusion problems have widely been studied by researchers because of their many valuable applications and generalizations. It is well known that many problems in applied sciences and engineering, mathematical optimization, machine learning, statistical learning, and optimal control can be modeled as variational inclusion problems. See, for example, [1–3] and references therein. In addition, under some assumptions, such problems involve many important concepts in applied mathematics, such as convex minimization, split feasibility, fixed points, saddle points, and variational inequalities; see, for example, [4–7].

Let  $\mathcal{H}$  be a real Hilbert space, and let  $S : \mathcal{H} \rightrightarrows \mathcal{H}$  and  $T : \mathcal{H} \rightarrow \mathcal{H}$  be maximal monotone and monotone operators, respectively. The variational inclusion problem (VIP) is to

find 
$$u \in \mathcal{H}$$
 such that  $0 \in (T+S)u$ . (1)

A foremost method proposed for solving (1) is the *forward-backward splitting method* (FBSM) [2]. The FBSM operates as follows:

$$x_{n+1} = (I + \lambda S)^{-1} (x_n - \lambda T x_n) \ \forall n \in \mathbb{N},$$

where  $\lambda \in (0, \frac{2}{L})$  and  $I : \mathcal{H} \to \mathcal{H}$  is the identity operator. The FBSM was proved to generate sequences that converge weakly to a solution of (1) under the assumption that T is  $\frac{1}{L}$ -cocoercive (or inverse strongly monotone). Aside from the cocoercivity assumption, convergence is only guaranteed under a similar strong assumption such as the strong monotonicity of S + T [8]. Interested readers could consult [9] for results on finding zeros of sums of maximal monotone operators using a similar *forward-backward* scheme.

As an improvement on the FBSM, Tseng [10] proposed the *modified forward-backward splitting method* (MFBSM) (also called the forward-backward-forward splitting method) for

solving the VIP for a more general case, where *T* is monotone and *L*-Lipschitz continuous. The MFBSM has the following structure:

$$\begin{cases} y_n = (I + \lambda S)^{-1} (x_n - \lambda T x_n) \\ x_{n+1} = y_n - \lambda (Ty_n - Tx_n), \ n \in \mathbb{N} \end{cases}$$

A weak convergence theorem was proved for this algorithm. While the implementation of the FBSM requires the prior knowledge of the Lipschitz constant of *T*, which makes the algorithm a bit restrictive, the MFBSM uses line search techniques to circumvent the onerous task of estimating the Lipschitz constant. Be that as it may, line search techniques are computationally expensive, as they require several extra computations per iteration; see, for example, [11]. For the inexact and the stochastic versions of the MFBSM, respectively, please see [12,13].

Motivated by the need to reduce the computational burden associated with the line search technique, and by the fact that strong convergence is more desirable than weak convergence in infinite dimensional Hilbert spaces and in applications, some authors have incorporated existing hybrid techniques in the MFBSM and have proposed hybrid-like strongly convergent methods with self-adaptive step sizes. See, for instance, [11,14] and references therein.

Let  $S : \mathcal{H} \Rightarrow \mathcal{H}$  and  $T_i : \mathcal{H} \rightarrow \mathcal{H}$  be maximal monotone and monotone operators, respectively, where  $i \in [I] := \{1, 2, ..., I\}$ . We recall the modified variational inclusion problem (MVIP) introduced in [15]:

find 
$$u \in \mathcal{H}$$
 such that  $0 \in \sum_{i \in [I]} (a_i T_i + S) u$ , (2)

where  $a_i \in (0,1)$  and  $\sum_{i \in [I]} a_i = 1$ . Obviously, the MVIP is more general than the VIP in the sense that if  $T_i = T \quad \forall i \in [I]$ , then the MVIP becomes the VIP. It is easy to see that if  $x \in \bigcap_{i \in [I]} (T_i + S)^{-1}(0)$ , then  $0 \in \sum_{i \in [I]} (a_i T_i + S)x$ . However, the converse is not true

in general.

In addition, the MVIP is more general than the following common variational inclusion problem, which has recently been studied in [16]:

find 
$$0 \in \bigcap_{i \in [I]} (T_i + S)u.$$
 (3)

In view of its generality, the MVIP has recently attracted the attention of some authors who studied it and proposed algorithms for solving it. The authors of [17] studied the problem for the case where the  $T_i$ s are inverse strongly monotone. They proposed a Halperninertial forward-backward splitting algorithm for solving the problem and proved a strong convergence theorem. Moreover, the authors of [18] studied the MVIP in the case where the  $T_i$ s are monotone and Lipschitz continuous and designed a modified Tseng method for solving it. By using some symmetry properties, they proved the weak convergence of the method they proposed. Furthermore, they provided a formulation of an image deblurring problem as an MVIP and used their results in order to solve this problem.

On the other hand, motivated by the work [16], the authors of [19] have recently studied the following common variational inclusion problem:

find 
$$0 \in \bigcap_{i \in [I]} (T_i + S_i) u$$
, (4)

where  $S_i : \mathcal{H} \Rightarrow \mathcal{H}$  are maximal monotone operators and  $T_i : \mathcal{H} \rightarrow \mathcal{H}$  are monotone Lipschitz continuous operators. They combined the inertial technique, Tseng's method, and the shrinking projection method to design an iterative algorithm that solves (4).

Motivated by the above studies, in this paper, we propose a unified simple modification of the MFBSM, which we call the *regularized MFBSM* (RMFBSM), for solving (2)–(4) and establish a strong convergence theorem for the sequences it generates. To realize our objectives, we first examine a regularized MVIP and study the solution nets it generates. The novelty of our scheme lies in the fact that, unlike the existing modifications of the MF-BSM in the literature, it yields strong convergence while preserving the two-step structure of the MFBSM. In the case where  $T_i = T$  and  $S_i = S$  for all i = 1, 2, ..., I, we apply our result to a statistical learning model and to split feasibility problems.

The organization of our paper is as follows: In Section 2, we recall some useful definitions and preliminary results, which are needed in our study and in the convergence analysis of our algorithm. In Section 3, we propose the regularized MFBSM and establish a strong convergence theorem for it. In Section 4, we give some applications of our main results. In Section 5, we present numerical examples to illustrate our method and compare it with some existing related algorithms in the literature. We conclude with Section 6.

#### 2. Preliminaries

We start this section by stating some notations and recalling a number of important definitions.

Let *C* be a nonempty, closed, and convex subset of a real Hilbert space  $\mathcal{H}$ , the inner product and induced norm of which are denoted by  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$ , respectively. We denote by  $u_n \rightharpoonup u'$  and  $u_n \rightarrow u'$  the weak and the strong convergence, respectively, of the sequence  $\{u_n\}$  to a point *u*. The following identity is well known:

$$\|u \pm v\|^2 = \|u\|^2 \pm 2\langle u, v \rangle + \|v\|^2 \,\forall u, v \in \mathcal{H}.$$

**Definition 1** ([1,20]). *A mapping*  $T : \mathcal{H} \to \mathcal{H}$  *is said to be* 

(i) L-Lipschitz continuous if there exists a constant L > 0 such that

$$||Tu - Tv|| \leq L||u - v|| \ \forall u, v \in \mathcal{H};$$

(ii)  $\beta$ -strongly monotone if there exists a constant  $\beta > 0$  such that

$$\langle Tu - Tv, u - v \rangle \geq \beta ||u - v||^2 \, \forall u, v \in \mathcal{H};$$

(iii)  $\beta$ -inverse strongly monotone if there exists a constant  $\beta > 0$  such that

$$\langle Tu - Tv, u - v \rangle \geq \beta \|Tu - Tv\|^2 \, \forall u, v \in \mathcal{H};$$

*(iv)* monotone if

$$\langle Tu - Tv, u - v \rangle \geq 0 \ \forall u, v \in \mathcal{H};$$

(v) hemicontinuous if for every  $u, v, w \in \mathcal{H}$ , we have

$$\lim_{\alpha\to 0} \langle T(u+\alpha v), w \rangle = \langle Tu, w \rangle.$$

**Remark 1.** *In view of the above definitions, it is clear that every strongly monotone mapping is monotone. In addition, Lipschitz continuous mappings are hemicontinuous.* 

Let  $S : \mathcal{H} \rightrightarrows \mathcal{H}$  be a set-valued operator. The graph of *S*, denoted by gr(S), is defined by

$$gr(S) := \{(u, v) \in \mathcal{H} \times \mathcal{H} : v \in Su\}.$$

The set-valued operator *S* is called a monotone operator if  $\forall (u, v), (y, z) \in gr(S), \langle u - y, v - z \rangle \ge 0$  and a maximal monotone operator if the graph of *S* is not a proper subset of the graph of any other monotone operator. For a maximal monotone operator *S* and  $\lambda > 0$ , the resolvent of *S* is defined by

$$J_{\lambda S} := (I + \lambda S)^{-1} : \mathcal{H} \to \mathcal{H}.$$

It is well known that  $J_{\lambda S}$  is firmly nonexpansive (in particular, nonexpansive). For each  $u \in \mathcal{H}$ , there exists a unique nearest element, denoted by  $P_C u \in C$  and called the metric projection of  $\mathcal{H}$  onto C at u. That is,

$$|u - P_C u|| \le ||u - v|| \ \forall v \in C.$$

The indicator function of *C*, denoted by  $i_C$ , is defined by

$$i_{\mathcal{C}}(u) := \begin{cases} 0, & \text{if } u \in \mathcal{C}, \\ +\infty, & \text{if } u \notin \mathcal{C}. \end{cases}$$

Recall that the subdifferential  $\partial f$  of a proper convex function f at  $u \in \mathcal{H}$  is defined by

$$\partial f(u) := \{ z \in \mathcal{H} : f(u) + \langle z, y - u \rangle \le f(y), \, \forall y \in \mathcal{H} \}.$$

The normal cone of *C* at the point  $u \in \mathcal{H}$ , denoted by  $N_C(u)$ , is defined by

$$N_{\mathcal{C}}(u) := \begin{cases} \{x \in \mathcal{H} : \langle x, y - u \rangle \le 0 \quad \forall y \in \mathcal{C} \}, & \text{if } u \in \mathcal{C}, \\ \emptyset, & \text{otherwise} \end{cases}$$

We know that  $\partial i_K$  is a maximal monotone operator, and we have  $\partial i_K = N_K$ . Furthermore, for each  $\lambda > 0$ ,

$$P_C u = (I + \lambda \partial i_K)^{-1} u. \tag{5}$$

The following important lemmata are useful in our convergence analysis.

**Lemma 1** ([21]). Let  $\mathcal{H}$  be a real Hilbert space. Let  $S : \mathcal{H} \rightrightarrows \mathcal{H}$  be a maximal monotone operator and let  $T : \mathcal{H} \rightarrow \mathcal{H}$  be a monotone and Lipschitz continuous mapping. Then, the mapping M = S + T is a maximal monotone mapping.

**Lemma 2** ([22]). Let  $\mathcal{H}$  be a real Hilbert space. Let K be a nonempty, closed, and convex subset of  $\mathcal{H}$ , and let  $\mathcal{F} : \mathcal{H} \to \mathcal{H}$  be a hemicontinuous and monotone operator. Then,  $\bar{u}$  is a solution to the variational inequality

find 
$$\bar{u} \in K$$
 such that  $\langle \mathcal{F}\bar{u}, u - \bar{u} \rangle \geq 0 \ \forall u \in K$ 

*if and only if*  $\bar{u}$  *is a solution to the following problem:* 

find  $\bar{u} \in K$  such that  $\langle \mathcal{F}u, u - \bar{u} \rangle \geq 0 \ \forall u \in K$ .

**Lemma 3** ([23]). Let  $\mathcal{H}$  be a real Hilbert space. Suppose that  $\mathcal{F} : \mathcal{H} \to \mathcal{H}$  is  $\kappa$ -Lipschitzian and  $\beta$ -strongly monotone over a closed and convex subset  $K \subset \mathcal{H}$ . Then, the variational inequality problem

find 
$$\bar{u} \in K$$
 such that  $\langle \mathcal{F}\bar{u}, v - \bar{u} \rangle \ge 0 \ \forall v \in K$ 

*has a unique solution*  $\bar{u} \in K$ *.* 

**Lemma 4** ([24]). Let  $\{\Psi_n\}$  be a sequence of non-negative real numbers,  $\{a_n\}$  be a sequence of real numbers in (0,1) satisfying the condition  $\sum_{n=1}^{\infty} a_n = \infty$ , and  $\{b_n\}$  be a sequence of real numbers. Assume that

$$\Psi_{n+1} \leq (1-a_n)\Psi_n + a_n b_n, \ n \geq 1.$$

If  $\limsup_{k\to\infty} b_{n_k} \leq 0$  for every subsequence  $\{\Psi_{n_k}\}$  of  $\{\Psi_n\}$  satisfying  $\liminf_{k\to\infty} (\Psi_{n_k+1} - \Psi_{n_k}) \geq 0$ , then  $\lim_{n\to\infty} \Psi_n = 0$ .

#### 3. Regularized Modified Forward-Backward Splitting Method

Let  $S : \mathcal{H} \Rightarrow \mathcal{H}$  be a maximal monotone operator,  $T_i : \mathcal{H} \rightarrow \mathcal{H}$  be a monotone and  $L_i$ -Lipschitz continuous operator for  $i \in [I] := \{1, 2, ..., I\}$ , and  $a_i \in (0, 1)$  such that  $\sum_{i \in [I]} a_i = 1$ . We denote the solution set of problem (2) by  $\Omega$ , that is,

$$\Omega:=\{u\in\mathcal{H}:0\in(\sum_{i\in[I]}a_iT_i+S)u\},\$$

and assume that  $\Omega \neq \emptyset$ . Let  $\mathcal{F} : \mathcal{H} \to \mathcal{H}$  be a  $\gamma$ -strongly monotone and *L*-Lipschitz continuous operator. We are seeking a solution  $x^* \in \mathcal{H}$  such that

$$x^* \in \Omega \text{ and } \langle \mathcal{F}x^*, u^{\dagger} - x^* \rangle \ge 0 \ \forall u^{\dagger} \in \Omega.$$
 (6)

We denote the solution set of (6) by  $\tilde{\Omega}$ . In this connection, we consider the following regularized modified variational inclusion problem (RMVIP):

find 
$$u \in \mathcal{H}$$
 such that  $0 \in (\sum_{i \in [I]} a_i T_i + S)u + \tau \mathcal{F}u$ , (7)

where  $\tau > 0$  is a regularization parameter. Observe that by Lemma 1,  $\sum_{i \in [I]} a_i T_i + S$  is a maximal monotone operator. In addition,  $(\sum_{i \in [I]} a_i T_i + S) + \tau \mathcal{F}$  is strongly monotone because  $\mathcal{F}$  is strongly monotone. Therefore, for each  $\tau > 0$ , problem (7) possesses a unique solution which we denote by  $u_{\tau}$ . The following result concerning the solution net  $\{u_{\tau}\}_{\tau \in (0,1)}$  can be deduced from [25] of Lemma 2, see also [7] of Proposition 3.1, but we will give the proof for completeness.

**Proposition 1.** Let  $\{u_{\tau}\}_{\tau \in (0,1)}$  be the solution net of (7). Then,

- (a)  $\{u_{\tau}\}_{\tau \in (0,1)}$  is bounded.
- (b) Let  $u_{\tau_j} \in \{u_{\tau}\}_{\tau \in (0,1)}$ , j = 1, 2. Then,  $\|u_{\tau_1} u_{\tau_2}\| \le \frac{|\tau_1 \tau_2|}{\tau_1}M$ , where M > 0 is a constant.
- (c)  $\lim_{\tau \to \pm} u_{\tau}$  exists and belongs to  $\tilde{\Omega}$ .

**Proof.** (a) Let  $u_{\tau}$  be a solution to (7). Then we have

$$-\tau \mathcal{F} u_{\tau} - \sum_{i \in [I]} a_i T_i u_{\tau} \in S u_{\tau}.$$
(8)

Choose  $u^{\dagger} \in \Omega$ . Then,

$$-\sum_{i\in[I]}a_iT_iu^{\dagger}\in Su^{\dagger}.$$
(9)

Using (8) and (9), we obtain

$$\langle -\tau \mathcal{F} u_{\tau} - \sum_{i \in [I]} a_i T_i u_{\tau} + \sum_{i \in [I]} a_i T_i u^{\dagger}, u_{\tau} - u^{\dagger} \rangle \ge 0.$$
<sup>(10)</sup>

Since  $T_i$  is monotone, it follows from (10) that

$$\tau \langle \mathcal{F} u_{\tau}, u^{\dagger} - u_{\tau} \rangle \geq \sum_{i \in [I]} a_i \langle T_i u_{\tau} - T_i u^{\dagger}, u_{\tau} - u^{\dagger} \rangle \geq 0.$$
(11)

Using (11) and the  $\gamma$ -strong monotonicity of  $\mathcal{F}$ , we find that

$$\gamma \|u^{\dagger} - u_{\tau}\|^{2} \leq \langle \mathcal{F}u^{\dagger} - \mathcal{F}u_{\tau}, u^{\dagger} - u_{\tau} \rangle$$

$$\leq \langle \mathcal{F}u^{\dagger}, u^{\dagger} - u_{\tau} \rangle$$

$$\leq \|\mathcal{F}u^{\dagger}\|\|u^{\dagger} - u_{\tau}\|.$$

$$(12)$$

It now follows from (12) that  $||u^{\dagger} - u_{\tau}|| \leq \frac{||\mathcal{F}u^{\dagger}||}{\gamma}$  and hence  $\{u_{\tau}\}_{\tau \in (0,1)}$  is bounded. (b) Let  $u_{\tau_j} \in \{u_{\tau}\}_{\tau \in (0,1)}$ , j = 1, 2. Then,

$$0 \in \left(\sum_{i \in [I]} a_i T_i + S\right) u_{\tau_j} + \tau_j \mathcal{F} u_{\tau_j}.$$
(13)

Using the monotonicity of  $(\sum_{i \in [I]} a_i T_i + S)$  and (13), we find that

$$\langle -\tau_{1}\mathcal{F}u_{\tau_{1}} + \tau_{2}\mathcal{F}u_{\tau_{1}} - \tau_{2}\mathcal{F}u_{\tau_{1}} + \tau_{2}\mathcal{F}u_{\tau_{2}}, u_{\tau_{1}} - u_{\tau_{2}} \rangle = \langle -\tau_{1}\mathcal{F}u_{\tau_{1}} + \tau_{2}\mathcal{F}u_{\tau_{2}}, u_{\tau_{1}} - u_{\tau_{2}} \rangle \ge 0.$$
(14)

Using (14), we see that

$$(\tau_{2} - \tau_{1})\langle \mathcal{F}u_{\tau_{1}}, u_{\tau_{1}} - u_{\tau_{2}} \rangle \geq \tau_{2}\langle \mathcal{F}u_{\tau_{1}} - \mathcal{F}u_{\tau_{2}}, u_{\tau_{1}} - u_{\tau_{2}} \rangle \geq \tau_{2}\gamma ||u_{\tau_{1}} - u_{\tau_{2}}||^{2}.$$
(15)

Therefore, it follows from (15) that

$$\begin{aligned} \tau_{2}\gamma \|u_{\tau_{1}} - u_{\tau_{2}}\|^{2} &\leq (\tau_{2} - \tau_{1})\langle \mathcal{F}u_{\tau_{1}}, u_{\tau_{1}} - u_{\tau_{2}} \rangle \\ &\leq |\tau_{2} - \tau_{1}|||\mathcal{F}u_{\tau_{1}}||||u_{\tau_{1}} - u_{\tau_{2}}||. \end{aligned}$$
(16)

Using (16), we obtain

$$||u_{\tau_1} - u_{\tau_2}|| \le \frac{|\tau_2 - \tau_1| ||\mathcal{F}u_{\tau_1}||}{\tau_2 \gamma} = \frac{|\tau_2 - \tau_1|M}{\tau_2}$$

where  $M = \frac{\|\mathcal{F}u_{\tau_1}\|}{\gamma}$ . (c) The boundedness of the net  $\{u_{\tau}\}_{\tau \in (0,1)}$  implies that there exists a subsequence  $\{u_{\tau_n}\}$ of  $\{u_{\tau}\}$  such that  $u_{\tau_n} \rightharpoonup \bar{u}$  as  $n \rightarrow \infty$  where  $(\tau_n)_{n \in \mathbb{N}}$  is a sequence in (0, 1) such that  $\tau_n \to 0^+$  as  $n \to \infty$ . Note that  $-\tau_n \mathcal{F} u_{\tau_n} \in (\sum_{i \in [I]} a_i T_i + S) u_{\tau_n}$  for each *n*. Since the operator  $(\sum_{i \in [I]} a_i T_i + S)$  is maximal monotone and hence sequentially closed in the weak-strong topology on  $\mathcal{H} \times \mathcal{H}$ , taking  $n \to \infty$ , we infer that  $0 \in (\sum_{i \in [I]} a_i T_i + S)\overline{u}$ . Therefore,  $\overline{u} \in \Omega$ . In addition,  $0 \in (\sum_{i \in [I]} a_i T_i + S) x^*$  and  $\langle \mathcal{F} x^*, u^{\dagger} - x^* \rangle \ge 0 \forall u^{\dagger} \in \Omega$ . It then follows that

$$\langle -\tau_n \mathcal{F} u_{\tau_n} - \sum_{i \in [I]} a_i T_i u_{\tau_n} + \sum_{i \in [I]} a_i T_i x^*, u_{\tau_n} - x^* \rangle \ge 0.$$

$$(17)$$

It now follows from (17) and the monotonicity of  $T_i$  that

$$-\langle \tau_n \mathcal{F} u_{\tau_n}, u_{\tau_n} - x^* \rangle \geq \sum_{i \in [I]} a_i \langle T_i u_{\tau_n} - T_i x^*, u_{\tau_n} - x^* \rangle \geq 0.$$
(18)

Using (18), we conclude that

$$\langle \mathcal{F}u_{\tau_n} - \mathcal{F}x^*, u_{\tau_n} - x^* \rangle + \langle \mathcal{F}x^*, u_{\tau_n} - x^* \rangle = \langle \mathcal{F}u_{\tau_n}, u_{\tau_n} - x^* \rangle \le 0.$$
(19)

Moreover, using (19) and the  $\gamma$ -strong monotonicity of  $\mathcal{F}$ , we see that

$$\gamma \|u_{\tau_n} - x^*\|^2 \leq \langle \mathcal{F}u_{\tau_n} - \mathcal{F}x^*, u_{\tau_n} - x^* \rangle \leq \langle \mathcal{F}x^*, x^* - u_{\tau_n} \rangle.$$
<sup>(20)</sup>

Therefore, using (20), we have

$$\langle \mathcal{F}x^*, x^* - u_{\tau_n} \rangle \ge 0. \tag{21}$$

Passing to the limit in (21) as  $n \to \infty$ , we find that

$$\langle \mathcal{F}x^*, x^* - \bar{u} \rangle \geq 0 \ \forall x^* \in \Omega, \ \bar{u} \in \Omega,$$

and using Lemma 2, we arrive at the conclusion that  $\bar{u}$  solves (6). The uniqueness of the solution to  $\langle \mathcal{F}x^*, u^{\dagger} - x^* \rangle \ge 0$  according to Lemma 3 implies that  $\bar{u} = x^*$ . Thus, passing to the limit as  $n \to \infty$  in (20), we conclude that  $u_{\tau_n} \to x^*$  as  $n \to \infty$ .  $\Box$ 

Following an approach similar to the above analysis, we denote the solution set of problem (4) by  $\Omega_2$  and consider the corresponding regularized common variational inclusion problem:

find 
$$u \in \mathcal{H}$$
 such that  $0 \in (T_i + S_i)u + \tau \mathcal{F}u$  for all *i*. (22)

If  $\{u_{\tau}\}_{\tau \in (0,1)}$  is the solution net of (22), then Proposition 1 (a) and (b) hold, and for (c), we see that  $\lim_{\tau \to 0^+} u_{\tau}$  exists and belongs to  $\tilde{\Omega}_2 := \{x^* \in \Omega_2 : \langle \mathcal{F}x^*, u^* - x^* \rangle \ge 0 \ \forall u^* \in \Omega_2\}$ . We make the following assumptions in connection with our Algorithm 1.

#### Assumption 1.

(A1) 
$$\tau_n \in (0,1)$$
,  $\lim_{n \to \infty} \tau_n = 0$ , and  $\sum_{n=1}^{\infty} \tau_n = \infty$ ;  
(A2)  $\lim_{n \to \infty} \frac{|\tau_{n+1} - \tau_n|}{\tau_n^2} = 0$ ;  
(A3)  $\{\mu_n\}$  and  $\{p_n\}$  are two real sequences satisfying  $\sum_{n=1}^{\infty} \mu_n < \infty$  and  $\sum_{n=1}^{\infty} p_n < \infty$ , respectively.

We next present a unified algorithm for solving the aforementioned problems.

Algorithm 1: Regularized Modified Forward-Backward Splitting Method	ł
(RMFBSM)	
<b>Initialization</b> : Let $u_1 \in \mathcal{H}$ , $\mu \in (0, 1)$ , and $\lambda_1 > 0$ be given.	
<b>Iterative steps</b> : Given $u_n$ , calculate $u_{n+1}$ as follows:	
Step 1: Compute	
$y_n^i = J_{\lambda_n S_i}(u_n - \lambda_n T_i u_n - \lambda_n \tau_n \mathcal{F} u_n).$	(23)
<b>Step 2</b> : Find $i_n := \arg \max_{i \in [I]} \{ \ y_n^i - u_n\  \}.$	
Step 3: Compute	
$u_{n+1} = y_n^{i_n} - \lambda_n (T_{i_n} y_n^{i_n} - T_{i_n} u_n).$	(24)

Update

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n + \rho_n, \frac{(\mu+\mu_n)\|y_n - u_n\|}{\|T_{i_n}y_n^{i_n} - T_{i_n}u_n\|}\right\}, & \text{if } T_{i_n}y_n^{i_n} \neq T_{i_n}u_n\\ \lambda_n + \rho_n, & \text{otherwise.} \end{cases}$$

Set n := n + 1 and go back to Step 1.

**Remark 2.** If  $T_i = T$ ,  $S_i = S$ ,  $\rho_n = 0 = \mu_n$  in Algorithm 1, and  $d_n = 0$  in Algorithm 1 of Hieu et al. [25], then the two algorithms are the same. However, the problems we intend to solve are more general than the problem (VIP) studied by Hieu et al. The sequence of step sizes in Algorithm 1 is convergent, as shown in the next lemma.

**Lemma 5.** The sequence  $\{\lambda_n\}$  generated by Algorithm 1 is bounded, and  $\lambda_n \in [\min\{\frac{\mu}{L_{i_n}}, \lambda_1\}, \lambda_1 + P]$ . Moreover, there exists  $\lambda \in [\min\{\frac{\mu}{L_{i_n}}, \lambda_1\}, \lambda_1 + P]$  such that  $\lim_{n \to \infty} \lambda_n = \lambda$ , where  $P = \sum_{n=1}^{\infty} \rho_n$ .

**Proof.** The proof follows from [26] of Lemma 2.1.  $\Box$ 

**Lemma 6.** Let  $\{u_n\}$  be the sequence generated by Algorithm 1. Then, there exists  $n_0 \in \mathbb{N}$  such that

$$\|u_{n+1} - u_{\tau_n}\|^2 \le (1 - \gamma \lambda_n \tau_n) \|u_n - u_{\tau_n}\|^2 - (1 - \mu) \|y_n^i - u_n\|^2 \quad \forall i \in [I], n \ge n_0,$$

where  $u_{\tau_n}$  solves the RMVIP with  $u_{\tau}$  replaced by  $u_{\tau_n}$ .

**Proof.** Using the definition of  $y_n^i$ , we obtain

$$u_n - \lambda_n T_i u_n - \lambda_n \tau_n \mathcal{F} u_n \in y_n^i + \lambda_n S_i y_n^i,$$

from which it follows that

$$u_n - \lambda_n (T_i u_n - T_i y_n^l) - \lambda_n \tau_n \mathcal{F} u_n - y_n^l \in \lambda_n (T_i + S_i) y_n^l, \text{ for each } i.$$
(25)

Employing (24) and (25) in particular for  $i_n$ , we obtain

$$u_n - u_{n+1} - \lambda_n \tau_n \mathcal{F} u_n \in \lambda_n (T_{i_n} + S_{i_n}) y_n^{\iota_n}.$$
<sup>(26)</sup>

In addition,

$$-\lambda_n \tau_n \mathcal{F} u_{\tau_n} \in \lambda_n (T_{i_n} + S_{i_n}) u_{\tau_n}.$$
<sup>(27)</sup>

Therefore, it follows from (26) and (27) that

$$\langle u_n - u_{n+1}, y_n^{i_n} - u_{\tau_n} \rangle - \lambda_n \tau_n \langle \mathcal{F} u_n - \mathcal{F} u_{\tau_n}, y_n^{i_n} - u_{\tau_n} \rangle \ge 0.$$
<sup>(28)</sup>

Using the fact that  $\mathcal{F}$  is  $\gamma$ -strongly monotone, we infer from (28) that

$$\langle u_n - y_n^{i_n}, y_n^{i_n} - u_{\tau_n} \rangle - \langle u_{n+1} - y_n^{i_n}, y_n^{i_n} - u_{\tau_n} \rangle - \lambda_n \tau_n \langle \mathcal{F}u_n - \mathcal{F}u_{\tau_n}, y_n^{i_n} - u_n \rangle \geq \lambda_n \tau_n \langle \mathcal{F}u_n - \mathcal{F}u_{\tau_n}, u_n - u_{\tau_n} \rangle \\ \geq \lambda_n \tau_n \gamma ||u_n - u_{\tau_n}||^2.$$

$$(29)$$

Note that

$$\langle u_n - y_n^{i_n}, y_n^{i_n} - u_{\tau_n} \rangle = \frac{1}{2} \left( \|u_n - u_{\tau_n}\|^2 - \|u_n - y_n^{i_n}\|^2 - \|y_n^{i_n} - u_{\tau_n}\|^2 \right)$$
(30)

and

$$\langle u_{n+1} - y_n^{i_n}, y_n^{i_n} - u_{\tau_n} \rangle = \frac{1}{2} \big( \|u_{n+1} - u_{\tau_n}\|^2 - \|u_{n+1} - y_n^{i_n}\|^2 - \|y_n^{i_n} - u_{\tau_n}\|^2 \big).$$
(31)

Substituting (30) and (31) in (29) and multiplying throughout by 2, we obtain

$$\begin{aligned} \|u_{n+1} - u_{\tau_n}\|^2 &\leq \|u_n - u_{\tau_n}\|^2 - \|u_n - y_n^{i_n}\|^2 + \|u_{n+1} - y_n^{i_n}\|^2 - 2\lambda_n \tau_n \langle \mathcal{F}u_n - \mathcal{F}u_{\tau_n}, y_n^{i_n} - u_n \rangle \\ &\quad -2\lambda_n \tau_n \gamma \|u_n - u_{\tau_n}\|^2 \\ &= (1 - 2\lambda_n \tau_n \gamma) \|u_n - u_{\tau_n}\|^2 - 2\lambda_n \tau_n \langle \mathcal{F}u_n - \mathcal{F}u_{\tau_n}, y_n^{i_n} - u_n \rangle \\ &\quad - \|u_n - y_n^{i_n}\|^2 + \|u_{n+1} - y_n^{i_n}\|^2. \end{aligned}$$
(32)

Using (24) and the Peter-Paul inequality, as well as (32), for any  $\epsilon_1 > 0$ , we have

$$\begin{aligned} \|u_{n+1} - u_{\tau_n}\|^2 &\leq (1 - 2\lambda_n \tau_n \gamma) \|u_n - u_{\tau_n}\|^2 + \frac{\lambda_n \tau_n L}{\epsilon_1} \|u_n - u_{\tau_n}\|^2 + \lambda_n \tau_n L \epsilon_1 \|y_n^{i_n} - u_n\|^2 \\ &- \|u_n - y_n^{i_n}\|^2 + \lambda_n^2 \|T_{i_n} y_n^{i_n} - T_{i_n} u_n\|^2 \\ &\leq (1 - \lambda_n \tau_n (2\gamma - \frac{L}{\epsilon_1})) \|u_n - u_{\tau_n}\|^2 - \left(1 - \lambda_n \tau_n L \epsilon_1 - \frac{\lambda_n^2 (\mu + \mu_n)^2}{\lambda_{n+1}^2}\right) \|y_n^{i_n} - u_n\|^2. \end{aligned}$$
(33)

9 of 18

Choose  $\epsilon_1 = \frac{L}{\gamma}$ . Observe that from Lemma 5 and the Assumption 1 (A1) and Assumption 1 (A3),

$$\lim_{n \to \infty} \left( 1 - \frac{\lambda_n \tau_n L^2}{\gamma} - \frac{\lambda_n^2 (\mu + \mu_n)^2}{\lambda_{n+1}^2} \right) = 1 - \mu^2 > 1 - \mu > 0.$$

Therefore, there exists  $n_0 \in \mathbb{N}$  such that  $\left(1 - \frac{\lambda_n \tau_n L_1^2}{\gamma} - \frac{\lambda_n^2 \mu^2}{\lambda_{n+1}^2}\right) > (1 - \mu)$  for all  $n \ge n_0$ . Thus, it follows from (33) that

$$\begin{aligned} \|u_{n+1} - u_{\tau_n}\|^2 &\leq (1 - \gamma \lambda_n \tau_n) \|u_n - u_{\tau_n}\|^2 - (1 - \mu) \|y_n^{i_n} - u_n\|^2 \\ &\leq (1 - \gamma \lambda_n \tau_n) \|u_n - u_{\tau_n}\|^2 - (1 - \mu) \|y_n^{i} - u_n\|^2, \ \forall i \in [I], \forall n \geq n_0, \end{aligned}$$
(34)

as asserted.  $\Box$ 

**Theorem 1.** The sequences generated by Algorithm 1 converge strongly to  $x^* \in \tilde{\Omega_2}$ .

**Proof.** Invoking Proposition 1(b) and the Peter-Paul inequality, we see that there exists  $\epsilon_2 > 0$  such that

$$\begin{aligned} \|u_{n+1} - u_{\tau_n}\|^2 &= \|u_{n+1} - u_{\tau_{n+1}} + u_{\tau_{n+1}} - u_{\tau_n}\|^2 \\ &= \|u_{n+1} - u_{\tau_{n+1}}\|^2 + 2\langle u_{n+1} - u_{\tau_{n+1}}, u_{\tau_{n+1}} - u_{\tau_n}\rangle + \|u_{\tau_{n+1}} - u_{\tau_n}\|^2 \\ &\geq \|u_{n+1} - u_{\tau_{n+1}}\|^2 - \epsilon_2 \|u_{n+1} - u_{\tau_{n+1}}\|^2 - \frac{\|u_{\tau_{n+1}} - u_{\tau_n}\|^2}{\epsilon_2} + \|u_{\tau_{n+1}} - u_{\tau_n}\|^2 \\ &= (1 - \epsilon_2) \|u_{n+1} - u_{\tau_{n+1}}\|^2 - (\frac{1 - \epsilon_2}{\epsilon_2}) \|u_{\tau_{n+1}} - u_{\tau_n}\|^2 \\ &\geq (1 - \epsilon_2) \|u_{n+1} - u_{\tau_{n+1}}\|^2 - (\frac{1 - \epsilon_2}{\epsilon_2}) (\frac{\tau_{n+1} - \tau_n}{\tau_n})^2 M^2. \end{aligned}$$
(35)

Next, we define the sequence  $\{\beta_n\}$  by

$$eta_n := egin{cases} au_n, & ext{if } \gamma\lambda_n > 1 \ \gamma\lambda_n au_n, & ext{if } \gamma\lambda_n \leq 1 \end{cases}$$

Then, from Lemma 6 and (35), it follows that  $\forall i \in [I]$  and for  $n \ge n_0$ ,

$$\|u_{n+1} - u_{\tau_{n+1}}\|^{2} \leq \frac{(1 - \beta_{n})}{1 - \epsilon_{2}} \|u_{n} - u_{\tau_{n}}\|^{2} + \frac{1}{\epsilon_{2}} \left(\frac{\tau_{n+1} - \tau_{n}}{\tau_{n}}\right)^{2} M^{2} - \frac{1}{1 - \epsilon_{2}} (1 - \mu) \|y_{n}^{i} - u_{n}\|^{2}.$$
(36)

Let  $\epsilon_2 = 0.5\beta_n$ . Then, from (36), it follows that

$$\|u_{n+1} - u_{\tau_{n+1}}\|^{2} \leq (1 - \frac{\epsilon_{2}}{1 - \epsilon_{2}}) \|u_{n} - u_{\tau_{n}}\|^{2} + \frac{\epsilon_{2}}{1 - \epsilon_{2}} \left(\frac{1 - \epsilon_{2}}{\epsilon_{2}^{2}} \left(\frac{\tau_{n+1} - \tau_{n}}{\tau_{n}}\right)^{2} M^{2}\right) - \frac{1 - \mu}{1 - \epsilon_{2}} \|y_{n}^{i} - u_{n}\|^{2} = (1 - \frac{\epsilon_{2}}{1 - \epsilon_{2}}) \|u_{n} - u_{\tau_{n}}\|^{2} + \frac{\epsilon_{2}}{1 - \epsilon_{2}} \left(\frac{4(1 - \epsilon_{2})}{\beta_{n}^{2}} \left(\frac{\tau_{n+1} - \tau_{n}}{\tau_{n}}\right)^{2} M^{2}\right) - \frac{1 - \mu}{1 - \epsilon_{2}} \|y_{n}^{i} - u_{n}\|^{2}, \ \forall n \ge n_{0}.$$

$$(37)$$

Let  $\Theta_n = ||u_n - u_{\tau_n}||^2$ ,  $\delta_n = \frac{\epsilon_2}{1-\epsilon_2}$  and

$$b_n = \frac{4(1-\epsilon_2)}{\beta_n^2} \left(\frac{\tau_{n+1}-\tau_n}{\tau_n}\right)^2 M^2.$$

We see that  $\delta_n \in (0, 1)$  and  $\sum_{n=1}^{\infty} \delta_n = \infty$ . Then, from (37), we obtain

$$\Theta_{n+1} \leq (1-\delta_n)\Theta_n + \delta_n b_n - \frac{1-\mu}{1-\epsilon_2} \|y_n^i - u_n\|^2, \ \forall n \ge n_0.$$
(38)

It is not difficult to see that the sequence  $\{b_n\}$  is bounded. To complete the proof, we invoke Lemma 4 by showing that  $\limsup_{k\to\infty} b_{n_k} \leq 0$  for every subsequence  $\{\Theta_{n_k}\}$  of  $\{\Theta_n\}$  that satisfies  $\liminf_{k\to\infty} (\Theta_{n_k+1} - \Theta_{n_k}) \geq 0$ . Therefore, let  $\{\Theta_{n_k}\}$  be a subsequence of  $\{\Theta_n\}$  such that  $\liminf_{k\to\infty} (\Theta_{n_k+1} - \Theta_{n_k}) \geq 0$ . It then follows from (38) that

$$\limsup_{k \to \infty} \frac{1-\mu}{(1-\epsilon_2)} \|y_{n_k}^i - u_{n_k}\|^2 \leq \limsup_{k \to \infty} \left( (\Theta_{n_k} - \Theta_{n_k+1}) + \delta_{n_k} (b_{n_k} - \Theta_{n_k}) \right)$$
  
$$\leq -\liminf_{k \to \infty} (\Theta_{n_k+1} - \Theta_{n_k})$$
  
$$\leq 0. \tag{39}$$

From (39), we infer that

$$\lim_{k \to \infty} \|y_{n_k}^i - u_{n_k}\| = 0.$$
(40)

Moreover, from (24) and (40), it follows that

$$\lim_{k\to\infty}\|u_{n_k+1}-u_{n_k}\|=0.$$

Indeed, it immediately follows from conditions (A1) and (A2) that  $\limsup_{k\to\infty} b_k = 0$ . Therefore, using Lemma 4, we obtain that  $\lim_{n\to\infty} \Theta_n = 0$ . Consequently, using Proposition 1 (c) and (40), we conclude that  $\lim_{n\to\infty} u_n = \lim_{n\to\infty} y_n^i = x^*$ .  $\Box$ 

We next present some consequences of the above result to confirm that our Algorithm 1 indeed provides a unified framework for solving the inclusion problems (2)–(4).

**Corollary 1.** Let  $S : \mathcal{H} \rightrightarrows \mathcal{H}$  be a maximal monotone operator and  $T_i : \mathcal{H} \rightarrow \mathcal{H}$  be monotone Lipschitz continuous operators with constants  $L_i$  for  $i \in [I]$ . Let  $\mathcal{F} : \mathcal{H} \rightarrow \mathcal{H}$  be a  $\gamma$ -strongly monotone and Lipschitz continuous operator with constant L. Assume  $\Omega_1 := \{x \in \mathcal{H} : 0 \in \cap_{i \in [I]}(T_i + S)\} \neq \emptyset$ . Then, the sequences generated by Algorithm 2, under Assumption 1, converge strongly to  $x^* \in \Omega_1$ , satisfying  $\langle \mathcal{F}x^*, u^* - x^* \rangle \geq 0 \forall u^* \in \Omega_1$ .

**Proof.** Note that Algorithm 2 is derived from Algorithm 1 by taking  $S_i = S$  for  $i \in [I]$ . Thus, the proof follows from the proof of Theorem 1.  $\Box$ 

**Corollary 2.** Let  $S : \mathcal{H} \rightrightarrows \mathcal{H}$  be a maximal monotone operator and  $T_j : \mathcal{H} \rightarrow \mathcal{H}$  be monotone Lipschitz continuous operators with constants  $L_j$  for  $j \in [J]$ . Let  $\mathcal{F} : \mathcal{H} \rightarrow \mathcal{H}$  be a  $\gamma$ -strongly

monotone and Lipschitz continuous operator with constant L and  $a_j \in (0,1)$  such that  $\sum_{j \in [J]}^{I} a_j = 1$ .

Assume that  $\Omega := \{x \in \mathcal{H} : 0 \in \sum_{j \in [J]} (a_j T_j + S)\} \neq \emptyset$ . Then, the sequences generated by Algorithm 3, under Assumption 1, converge strongly to  $x^* \in \Omega$ , satisfying  $\langle \mathcal{F}x^*, u^* - x^* \rangle \geq 0 \forall u^* \in \Omega$ .

# Algorithm 2: Regularized Modified Forward-Backward Splitting Method (RMFBSM)

**Initialization**: Let  $u_1 \in \mathcal{H}$ ,  $\mu \in (0, 1)$ , and  $\lambda_1 > 0$  be given. **Iterative steps**: Calculate  $u_{n+1}$  as follows: **Step 1**: Compute

$$y_n^i = J_{\lambda_n S}(u_n - \lambda_n T_i u_n - \lambda_n \tau_n \mathcal{F} u_n).$$

**Step 2**: *Find*  $i_n := \arg \max_{i \in [I]} \{ ||y_n^i - u_n|| \}$ . **Step 3**: *Compute* 

$$u_{n+1} = y_n^{i_n} - \lambda_n (T_{i_n} y_n^{i_n} - T_{i_n} u_n).$$

Update

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n + \rho_n, \frac{(\mu + \mu_n) \|y_n - u_n\|}{\|T_{i_n} y_n^{i_n} - T_{i_n} u_n\|}\right\}, & \text{if } T_{i_n} y_n^{i_n} \neq T_{i_n} u_n, \\ \lambda_n + \rho_n, & \text{otherwise.} \end{cases}$$

Set n := n + 1 and go back to Step 1.

**Algorithm 3:** Regularized Modified Forward-Backward Splitting Method (RMFBSM)

**Initialization**: Let  $u_1 \in \mathcal{H}$ ,  $\mu \in (0, 1)$ , and  $\lambda_1 > 0$  be given. **Iterative steps**: Given  $u_n$ , calculate  $u_{n+1}$  as follows: **Step 1**: Compute

$$y_n = J_{\lambda_n S}(u_n - \lambda_n \sum_{j=1}^J a_j T_j u_n - \lambda_n \tau_n \mathcal{F} u_n).$$

Step 2: Compute

$$u_{n+1} = y_n - \lambda_n \sum_{i=1}^{J} a_j (T_j y_n - T_j u_n).$$
(41)

Update

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n + \rho_n, \frac{(\mu + \mu_n) \|y_n - u_n\|}{\|\sum_{j \in [J]} a_j(T_j y_n - T_j u_n)\|}\right\}, & \text{if } \sum_{j \in [J]} a_j T_j y_n \neq \sum_{j \in [J]} a_j T_j u_n, \\ \lambda_n + \rho_n, & \text{otherwise.} \end{cases}$$

Set n := n + 1 and go back to Step 1.

**Proof.** Note that Algorithm 3 is derived from Algorithm 1 by taking I = 1 and setting  $T_1 = T = \sum_{j \in [J]} a_j T_j$ . Note that the new operator *T* is monotone and Lipschitz continuous. Thus, the proof follows from the proof of Theorem 1.  $\Box$ 

**Remark 3.** On the other hand, we note that the structure of some steps in Algorithm 1 differs from that of Algorithm 1 of [18]. For instance, the authors of [18] needed to solve an optimization problem to compute  $u_{n+1}$ . Our iterative scheme 3 provides an alternative and bypasses such difficulties. However, we do not consider accelerated schemes in the present study because such results can easily be obtained from the results in a recent work of ours [27].

#### 4. Applications

In this section, in the cases where I = 1, we will consider two applications of our main results.

#### 4.1. Split Feasibility Problems

Let *C* and *Q* be nonempty, closed, and convex subsets of the real Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively, and let  $A : \mathcal{H}_1 \to \mathcal{H}_2$  be a bounded linear operator, the adjoint of which is denoted by  $A^* : \mathcal{H}_2 \to \mathcal{H}_1$ . We now recall the split feasibility problems (SFP):

find 
$$u \in C$$
 such that  $Au \in Q$ . (42)

We denote the set of solutions of (42) by S and assume that  $S \neq \emptyset$ . The concept of SFP was introduced by Censor and Elfving [28] in 1994 in the framework of Euclidean spaces. It has been successfully applied to model some inverse problems in medical image reconstruction, phase retrievals, gene regulatory network inference, and intensity modulation radiation therapy; see, for example, [29,30]. From its conception to date, the SFP have been widely studied by several authors who also proposed various iterative schemes for solving them. Interested readers should consult [31–38] and references therein for recent studies and generalizations of this problem.

One can verify that the solution set of SFP (42) coincides with the solution set of the following constrained minimization problem; see, for example, [36]:

$$\min_{u \in C} f(u) := \frac{1}{2} \|Au - P_Q Au\|^2.$$
(43)

However, it is to be observed that the minimization problem (43) is ill-posed, in general, and therefore calls for regularization. We consider the following Tikhonov regularization [38]:

$$\min_{u \in C} f^{\kappa}(u) := \frac{1}{2} \|Au - P_Q Au\|^2 + \frac{1}{2} \kappa \|u\|^2.$$
(44)

Equivalently, (44) can be written as the following unconstrained minimization problem:

$$\min_{u \in \mathcal{H}_1} F^{\kappa}(u) := \frac{1}{2} \|Au - P_Q Au\|^2 + \frac{1}{2} \kappa \|u\|^2 + i_C(u).$$
(45)

Note that the function  $f^{\kappa}$  is differentiable and its gradient  $\nabla f^{\kappa}(u) = A^*(Au - P_QAu) + \kappa u$ . Problem (45) is equivalent to the following inclusion problem:

find 
$$u \in \mathcal{H}_1$$
 such that  $0 \in \nabla f^{\kappa}(u) + \partial i_C(u)$ . (46)

It is not difficult to see that  $\nabla f^{\kappa}$  is monotone and  $(||A||^2 + \kappa)$  Lipschitz continuous. Therefore, since  $\partial i_C$  is a maximal monotone operator, (46) assures us that we can apply our result to solving the SFP (42). Thence, our next result.

**Theorem 2.** Let C and Q be nonempty, closed, and convex subsets of the real Hilbert spaces  $\mathcal{H}_1$ and  $\mathcal{H}_2$ , respectively. Suppose that  $A : \mathcal{H}_1 \to \mathcal{H}_2$  is a bounded linear operator and let the operator  $A^* : \mathcal{H}_2 \to \mathcal{H}_1$  be its adjoint operator. Assume that  $S := \{u \in \mathcal{H}_1 : u \in C \text{ and } Au^* \in Q\} \neq \emptyset$ and that Assumption 1 holds. Then, the sequence  $\{u_n\}$  generated by Algorithm 4 converges strongly to a point  $u^* \in S$ , satisfying  $\langle \mathcal{F}u^*, y - u^* \rangle \ge 0 \ \forall y \in S$ .

**Proof.** Let  $S = \partial i_C$  and  $T_i = T = \nabla f^{\kappa}$  for all i = 1, 2, ..., N in Algorithm 1. Note that by (5), we have  $J_{\lambda_n S} = P_C$ . Therefore, the proof can be obtained by following the steps of the proof of Theorem 1.  $\Box$ 

# Algorithm 4: RMFBSM Method for Solving SFP

**Initialization**: Let  $u_1 \in \mathcal{H}_1$ ,  $\mu \in (0, 1)$ , and  $\lambda_1 > 0$  be given. **Iterative steps**: Calculate  $u_{n+1}$  as follows: **Step 1**: Compute

$$y_n = P_C(u_n - \lambda_n \nabla f^{\kappa} u_n - \lambda_n \tau_n \mathcal{F} u_n).$$

Step 2: Compute

$$u_{n+1} = y_n - \lambda_n (\nabla f^{\kappa} y_n - \nabla f^{\kappa} u_n).$$

Update

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \frac{\mu \|y_n - u_n\|}{\|\nabla f^{\kappa} y_n - \nabla f^{\kappa} u_n\|}\right\}, & \text{if } \nabla f^{\kappa} y_n \neq \nabla f^{\kappa} u_n \\ \lambda_n, & \text{otherwise.} \end{cases}$$

Set n := n + 1 and go back to Step 1.

#### 4.2. Elastic Net Penalty Problem

We consider the following linear regression model used in statistical learning:

$$y = A_1 \bar{u}_1 + A_2 \bar{u}_2 + \dots + A_N \bar{u}_N + \epsilon, \tag{47}$$

where *y* is the response predicted by the *N* predictors  $A_1, A_2, \dots, A_N$  and  $\epsilon$  is a random error term. We assume that the model is sparse with a limited number of nonzero coefficients. The model can be written in matrix format as follows:

$$y = Au + \epsilon, \tag{48}$$

where  $y \in \mathbb{R}^M$ ,  $A \in \mathbb{R}^{M \times N}$  is the predictor matrix and  $u = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_N) \in \mathbb{R}^N$ . In the statistics community, one solves the model by recasting (48) as the following penalized optimization problem:

$$\min_{u \in \mathbb{R}^N} \|Au - y\|_2^2 + \operatorname{pen}_{\sigma}(u), \tag{49}$$

where  $pen_{\sigma}(u), \sigma \ge 0$ , is a penalty function of *u*.

u

Penalized regression and variable selection are two key topics in linear regression analysis. They have recently attracted the attention of many authors who have proposed and analyzed various penalty functions (see [39] and references therein). A very popular penalized regression model for variable selection is LASSO (least absolute shrinkage and selection operator), which was proposed by Tibshirani [40]. It is an  $\ell_1$ -norm regularization least square model given by

$$\min_{u \in \mathbb{R}^N} \frac{1}{2} \|Au - y\|_2^2 + \sigma_1 \|u\|_1,$$
(50)

where  $\sigma_1 \geq 0$  is a nonnegative regularization parameter. The  $\ell_1$  penalty makes LASSO perform both continuous shrinkage and automatic variable selection simultaneously [41]. However, its conditions are invalid when applied to a group of highly correlated variables [42]. In addition to LASSO, other approaches for variable selection with penalty functions more general than the  $\ell_1$  penalty have been proposed (please see [43] and references therein for more details on other penalty functions and applications to SVM).

In this subsection, we focus on the penalty function, which was proposed by Zou and Hastie [41] and called the elastic net penalty function. The elastic net is a linear combination of  $\ell_1$  and  $\ell_2$  penalties and is defined by

$$pen_{\sigma}(u) := \sigma_1 \|u\|_1 + \sigma_2 \|u\|_2^2, \ \sigma_1, \sigma_2 > 0.$$

Therefore, we will consider the following elastic net penalty problem:

$$\min_{u \in \mathbb{R}^N} \frac{1}{2} \|Au - y\|_2^2 + \sigma_1 \|u\|_1 + \sigma_2 \|u\|_2^2.$$
(51)

Let

$$f(u) = \frac{1}{2} \|Au - y\|_2^2 + \sigma_2 \|u\|_2^2$$

and

$$g(u) = \sigma_1 \|u\|_1.$$

Then,  $\nabla f(u) = A^{\top}(Au - y) + 2\sigma_2 u$ , where  $A^{\top}$  stands for the transpose of A. Thus,  $\nabla f$  is monotone and Lipschitz continuous with constant  $L + 2\sigma_2$ , L being the largest eigenvalue of the matrix  $A^{\top}A$ . Note that g is a proper lower semicontinuous convex function and  $\partial g$  is maximal monotone. The resolvent of the maximal monotone operator  $\partial g$  is given as (see [20])

$$J_{\sigma\partial g}v = \arg\min_{u\in\mathbb{R}^N} \{g(u) + \frac{1}{2\sigma} \|u - v\|_2^2\}.$$

By the optimality condition, the penalty problem (51) is equivalent to the following inclusion problem:

$$0 \in \nabla f(u^*) + \partial g(u^*). \tag{52}$$

We apply our Algorithm 1 to solve the elastic net penalty problem (51), taking  $S = \partial g$ and  $T_i = T = \nabla f$  for each i = 1, 2, ..., N, and then compare the effectiveness of our methods with some existing methods in the literature. To this end, let  $u \in \mathbb{R}^N$  be a sparse randomly generated  $N \times 1$  matrix and let  $A \in \mathbb{R}^{M \times N}$  and  $\epsilon \in \mathbb{R}^M$  be randomly generated matrices, the entries of which are normally distributed with mean 0 and variance 1. In our experiments, we choose different values for  $u, u_1, A$ , and  $\epsilon$ , as follows:

**Case A:**  $A = \operatorname{randn}(50, 10), \ \epsilon = \operatorname{randn}(50, 1), \ u = \operatorname{sprandn}(10, 1, 0.3), \ u_1 = \operatorname{sprandn}(10, 1, 0.2).$  **Case B:**  $A = \operatorname{randn}(1000, 200), \ \epsilon = \operatorname{randn}(100, 1), \ u = \operatorname{sprandn}(200, 1, 0.1), \ u_1 = \operatorname{sprandn}(200, 1, 0.2).$  **Case C:**  $A = \operatorname{randn}(200, 1200), \ \epsilon = \operatorname{randn}(200, 1), \ u = \operatorname{sprandn}(1200, 1, \frac{1}{60}), \ u_1 = \operatorname{sprandn}(1200, 1, 0.02).$ **Case D:**  $A = \operatorname{randn}(120, 512), \ \epsilon = \operatorname{randn}(120, 1), \ u = \operatorname{sprandn}(512, 1, \frac{1}{64}), \ u_1 = \operatorname{sprandn}(512, 1, \frac{1}{64}).$ 

We also choose  $\sigma_1 = 0.6$ ,  $\sigma_2 = 0.4$ ,  $\mu = 0.02$ ,  $\tau_n = \frac{1}{\sqrt{n+2}}$ , and  $\mathcal{F}u = 10u$  for all  $u \in \mathcal{H}$ . Using MATLAB 2021b, we compare the performance of our Algorithm 1 (RMFBSM) with MFBSM and Algorithm 2 in [14] (VTM). Table 1 and Figure 1 illustrate the outcome of our computations. The stopping criterion is  $e_n = \frac{1}{n} ||u_n - u||_2^2 \le 10^{-3}$ .

Table	1.	Nume	rical	results.
-------	----	------	-------	----------

		RMFBSM	MFBSM	VTM
Case A	CPU time (sec)	0.0131	0.0200	0.0448
	No. of Iter.	590	936	2818
Case B	CPU time (sec)	0.5782	0.6199	2.6694
	No. of Iter.	2441	2734	10,564
Case C	CPU time (sec)	3.8084	5.3986	5.2824
	No. of Iter.	10,710	15,901	15,345
Case D	CPU time (sec)	0.8748	1.1264	0.9293
	No. of Iter.	5412	6833	5530

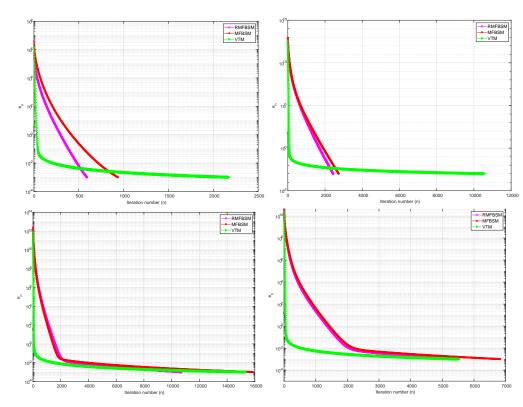


Figure 1. Top left: Case A; top right: Case B; bottom left: Case C; bottom right: Case D.

### 5. Numerical Example

We present the following numerical example to further illustrate the effectiveness of our proposed Algorithm 1 (RMFBSM). The codes are written in MATLAB 2021b and run on an HP Laptop Windows 10 with Intel(R) Core(TM) i5 CPU and 4 GB RAM.

**Example 1.** Consider  $\ell_2(\mathbb{R}) := \{u = (u_1, u_2, \dots, u_j, \dots), u_j \in \mathbb{R} : \sum_{j=1}^{\infty} |u_j|^2 < \infty\}, \|u\| = (\sum_{j=1}^{\infty} |u_j|^2)^{\frac{1}{2}} \forall u \in \ell_2(\mathbb{R}). Let \mathcal{H} = \ell_2(\mathbb{R}), and let S : \mathcal{H} \Rightarrow \mathcal{H}, T_i : \mathcal{H} \to \mathcal{H}, i = 1, 2, 3, and \mathcal{F} : \mathcal{H} \to \mathcal{H}$  be defined by  $Su = (2u_1, 2u_2, \dots, 2u_i, \dots), T_1x = (u_1, \frac{u_2}{2}, \dots, \frac{u_j}{2}, \dots), T_2u = \frac{u}{2} + (1, 0, 0, \dots), T_3u = \frac{u}{3} + (2, 0, 0, \dots), and \mathcal{F}u = (7u_1, 7u_2, \dots, 7u_j, \dots), respectively \quad \forall u \in \mathcal{H}.$  Choose  $a_1 = \frac{1}{2}, a_2 = \frac{1}{5}, and a_3 = \frac{3}{10}$ . One can see that S is maximal monotone,  $T_i$ s are Lipschitz continuous and monotone, and  $\mathcal{F}$  is strongly monotone and Lipschitz continuous. In our computations, we choose, for each  $n \in \mathbb{N}, \tau_n = \frac{1}{\sqrt{n+1}}, \lambda_1 = 0.3, and \mu = 0.1$ .

The following choices of the initial values  $u_1$  are considered: Case a:  $u_1 = (-1, \frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, ...);$ Case b:  $u_1 = (\frac{2}{3}, \frac{1}{9}, \frac{1}{54}, \frac{1}{324}, ...);$ Case c:  $u_1 = (100, 10, 1, 0.1, ...);$ Case d:  $u_1 = (9, 3\sqrt{3}, 3, \sqrt{3}, ...).$ 

We compare the performance of our Algorithm (1) (RMFBSM) with MFBSM and VTM. We choose  $\alpha_n = \frac{1}{n+2}$  and  $f(u) = \frac{u}{2}$  for VTM. The stopping criterion used for our computations is  $e_n = ||u_n - J_{\lambda_n S}(I - \lambda_n T)u_n|| < 10^{-8}$ . Observe that  $e_n = 0$  implies that  $u_n$  is a solution to the inclusion problem of finding  $u \in \mathcal{H}$  such that  $0 \in (S + T)u$ . We plot the graphs of errors against the number of iterations in each case. The figures and numerical results are shown in Figure 2 and Table 2, respectively.

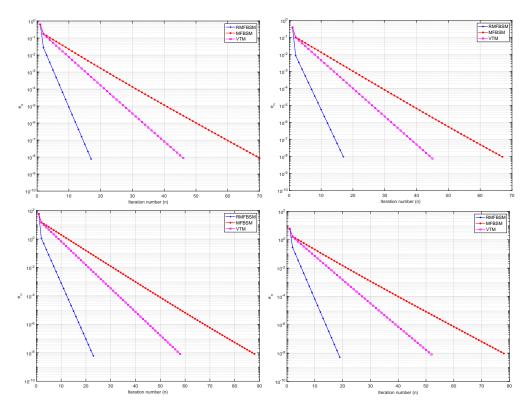


Figure 2. Top left: Case a; top right: Case b; bottom left: Case c; bottom right: Case d.

		RMFBSM	MFBSM	VTM
Case a	CPU time (sec)	0.0145	0.0155	0.0154
	No of Iter.	17	70	46
Case b	CPU time (sec)	0.0040	0.0163	0.0050
	No. of Iter.	17	67	45
Case c	CPU time (sec)	0.0019	0.0038	0.0048
	No of Iter.	23	88	58
Case d	CPU time (sec)	0.0019	0.0027	0.0359
	No of Iter.	19	78	52

Table 2.Numerical results.

**Remark 4.** The numerical results show that our algorithm performs better than the two algorithms with which we compared it with respect to both the computation time taken and the number of iterations. Nevertheless, its performance can be further improved by incorporating inertial extrapolation terms into the algorithm. A possible future direction of research concerns estimating and comparing the rate of convergence of our proposed algorithm with the pertinent algorithms in the literature.

# 6. Conclusions

We study three classes of variational inclusion problems of finding a zero of the sum of monotone operators in real Hilbert space. We propose a unified simple structure, which combines a modification of the Tseng method with self-adaptive step sizes, and prove that the sequences it generates converge strongly to a solution of the problems. We apply our results to the elastic net penalty problem in statistical learning and to split feasibility problems. Numerical experiments are presented in order to illustrate the usefulness and effectiveness of our new method. Results related to accelerated versions of our method can be derived from the results of our recent work [27]. **Author Contributions:** Writing—original draft, A.T. and S.R. Both authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** 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).

Data Availability Statement: Not applicable.

Acknowledgments: Both authors are grateful to four anonymous referees for their helpful comments and useful suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

- 1. Alber, Y.; Ryazantseva, I. Nonlinear Ill-Posed Problems of Monotone Type; Springer: Dordrecht, The Netherlands, 2006; p. xiv+410; ISBN 978-1-4020-4395-6/1-4020-4395-3.
- Lions, P.L.; Mercier, B. Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. 1979, 16, 964–979. [CrossRef]
- 3. Taiwo, A.; Alakoya, T.O.; Mewomo, O.T. Halpern-type iterative process for solving split common fixed point and monotone variational inclusion problem between Banach spaces. *Numer. Algorithms* **2020**, *86*, 1359–1389. [CrossRef]
- 4. Bello, A.U.; Yusuf, H.; Djitte, N. Single-step algorithm for variational inequality problems in 2-uniformly convex banach spaces. *Rend. Circ. Mat. Palermo, II. Ser.* **2023**, *72*, 1463–1481. [CrossRef]
- 5. Rehman, H.; Kumam, P.; Ozdemir, M.; Yildirim, I.; Kumam, W. A class of strongly convergent subgradient extragradient methods for solving quasimonotone variational inequalities. *Dem. Math.* **2023**, *56*, 20220202. [CrossRef]
- 6. Reich, S.; Taiwo, A. Fast hybrid iterative schemes for solving variational inclusion problems. *Math. Meth. Appl. Sci.* 2023, 46, 17177–17198. [CrossRef]
- 7. Taiwo, A.; Reich, S. Bounded perturbation resilience of a regularized forward-reflected-backward splitting method for solving variational inclusion problems with applications. *Optimization* **2023**. [CrossRef]
- 8. Chen, G.H.; Rockafellar, R.T. Convergence rates in forward-backward splitting. SIAM J. Optim. 1997, 7, 421-444. [CrossRef]
- Passty, G.B. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. Appl. 1979, 72, 383–390. [CrossRef]
- 10. Tseng, P. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.* **2000**, *38* 431–446. [CrossRef]
- 11. Cholamjiak, P.; Hieu, D.V.; Cho, Y.J. Relaxed Forward–Backward splitting methods for solving variational inclusions and applications. J. Sci. Comput. 2021, 88, 85. [CrossRef]
- Briceño-Arias, L.M.; Combettes, P.L. A monotone+skew splitting model for composite monotone inclusions in duality. SIAM J. Optim. 2011, 21, 1230–1250. [CrossRef]
- 13. Vũ, B.C. Almost sure convergence of the forward-backward-forward splitting algorithm. *Optim. Lett.* 2016, 10, 781–803.
- 14. Gibali, A.; Thong, D.V. Tseng type methods for solving inclusion problems and its applications. Calcolo 2018, 55, 49. [CrossRef]
- 15. Khuangsatung, W.; Kangtunyakarn, A. Algorithm of a new variational inclusion problem and strictly pseudononspreading mapping with application. *Fixed Point Theory Appl.* **2014**, 2014, 209. [CrossRef]
- Cholamjiak, W.; Khan, S.A.; Yambangwal, D.; Kazmi, K.R. Strong convergence analysis of common variational inclusion problems involving an inertial parallel monotone hybrid method for a novel application to image restoration. *RACSAM* 2020, *114*, 99. [CrossRef]
- 17. Sombut, K.; Sitthithakerngkiet, K.; Arunchai, A.; Seangwattana, T. An inertial forward-backward splitting method for solving modified variational inclusion problems and its applications. *Mathematics* **2023**, *11*, 2017. [CrossRef]
- 18. Seangwattana, T.; Sombut, K.; Arunchai, A.; Sitthithakerngkiet, K. A modified Tseng's method for solving the modified variational inclusion problems and its applications. *Symmetry* **2021**, *13*, 2250. [CrossRef]
- Suparatulatorna, R.; Chaichana, K. A strongly convergent algorithm for solving common variational inclusion with application to image recovery problems. *Appl. Numer. Math.* 2022, 173, 239–248. [CrossRef]
- 20. Bauschke, H.H.; Combettes, P.L. Convex Analysis and Monotone Operator Theory in Hilbert Spaces, CMS Books in Mathematics; Springer: Berlin/Heidelberg, Germany, 2011.
- 21. Brézis, H. Operateurs Maximaux Monotones; North-Holland Publishing Company: Amsterdam, The Netherlands, 1983; 183p.
- 22. Cottle, R.W.; Yao, J.C. Pseudo-monotone complementarity problems in Hilbert space. J. Optim. Theory Appl. 1992, 75, 281–295. [CrossRef]
- Yamada, I. 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.* 2001, *8*, 473–504.

- 24. Saejung, S.; Yotkaew, P. Approximation of zeros of inverse strongly monotone operators in Banach spaces. *Nonlinear Anal.* **2012**, 75, 742–750. [CrossRef]
- 25. Hieu, D.V.; Anh, P.K.; Ha, N.H. Regularization proximal method for monotone variational inclusions. *Netw. Spat. Econ.* **2021**, *21*, 905–932. [CrossRef]
- 26. Wang, Z.; Lei, Z.; Long, X.; Chen, Z. Tseng splitting method with double inertial steps for solving monotone inclusion problems. *arXiv* **2022**, arXiv:2209.11989v1.
- 27. Taiwo, A.; Reich, S. Two regularized inertial Tseng methods for solving inclusion problems with applications to convex bilevel programming. *Optimization* **2023**, *under review*.
- Censor, Y.; Elfving, T. A multiprojection algorithm using Bregman projections in a product space. *Numer. Algorithms* 1994, *8*, 221–239. [CrossRef]
- Censor, Y.; Elfving, T.; Kopf, N.; Bortfeld, T. The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Probl.* 2005, 21, 2071–2084. [CrossRef]
- Wang, J.; Hu, Y.; Li, C.; Yao, J.-C. Linear convergence of CQ algorithms and applications in gene regulatory network inference. *Inverse Probl.* 2017, 33, 5. [CrossRef]
- 31. Byrne, C. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Probl.* **2004**, *20*, 103–120. [CrossRef]
- 32. Lu, T.; Zhao, L.; He, H. A multi-view on the CQ algorithm for split feasibility problems: From optimization lens. *J. Appl. Numer. Optim.* **2020**, *2*, 387–399.
- Moudafi, A. Byrne's extended CQ-algorithms in the light of Moreau-Yosida regularization. *Appl. Set-Valued Anal. Optim.* 2021, 3, 21–26.
- 34. Reich, S.; Taiwo, A. A one-step Tikhonov regularization iterative scheme for solving split feasibility and fixed point problems. *Minimax Theory Appl.* **2023**, *accepted for publication*.
- 35. Reich, S.; Truong, M.T.; Mai, T.N.H. The split feasibility problem with multiple output sets in Hilbert spaces. *Optim. Lett.* **2020**, *14*, 2335–2353. [CrossRef]
- 36. Taiwo, A.; Reich, S.; Chinedu, I. Strong convergence of two regularized relaxed extragradient schemes for solving the split feasibility and fixed point problem with multiple output sets. *Appl. Anal.* **2023**. [CrossRef]
- 37. Takahashi, W. The split feasibility problem and the shrinking projection method in Banach spaces. *J. Nonlinear Convex Anal.* **2015**, *16*, 1449–1459.
- Xu, H.-K. Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces. *Inverse Probl.* 2010, 26, 10. [CrossRef]
- 39. Zeng, L.; Xie, J. Group variable selection via SCAD L<sub>2</sub>. Statistics **2014**, 48, 49–66. [CrossRef]
- 40. Tibshirani, R. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B 1996, 58, 267–288. [CrossRef]
- 41. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B 2005, 67, 301–320. [CrossRef]
- 42. Zhang, C.; Huang, J. The sparsity and bias of the lasso selection in high-dimensional linear regression. *Ann. Stat.* 2008, *36*, 1567–1594. [CrossRef]
- 43. Becker, N.; Toedt, G.; Lichter, P.; Benner, A. Elastic SCAD as a novel penalization method for SVM classification tasks in high-dimensional data. *BMC Bioinform.* **2011**, *12*, 138. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.