manuscript No.
(will be inserted by the editor)

On regularized forward-backward dynamical systems associated with structured monotone inclusions

Pham Ky Anh · Trinh Ngoc Hai

Received: date / Accepted: date

Abstract Recently, a regularized forward-backward dynamical system associated with additively structured monotone inclusions involving a multi-valued maximally monotone operator $\mathcal{A}$ and a single-valued co-coercive operator $\mathcal{B}$ has been studied in Adv. Nonlinear Anal. 2021; 10: 450-476. In this work, we establish strong convergence of the generated trajectories to a solution of the original monotone inclusion under a weaker assumption on the operator $\mathcal{B}$, namely $\mathcal{B}$ is Lipschitz continuous and such that the sum $\mathcal{S} := \mathcal{A} + \mathcal{B}$ is maximally monotone. It is well known that the co-coerciveness of $\mathcal{B}$ implies its monotonicity and Lipschitz continuity, which in turn infers the maximal monotonicity of $\mathcal{S}$. If the operator $\mathcal{A} + \mathcal{B}$ is maximally monotone and strongly pseudomonotone, we obtain a convergence estimate. A time discretization of the dynamical system provides an iterative regularization forward-backward method with relaxation parameters. The performance of the regularized dynamical system approach is illustrated by numerical experiments.

Keywords Structured (composite) monotone inclusions · Dynamical system · Forward-backward method · Iterative regularization method

Mathematics Subject Classification (2010) 34G25 · 47H05 · 65K15 · 65Y05 · 90C25

1 Introduction

In this paper we focus our attention on solving the following additively structured monotone inclusion

$$\text{Find } u^* \in \mathcal{H} \text{ such that } 0 \in (\mathcal{A} + \mathcal{B})u^*, \tag{SMI}$$

where $\mathcal{H}$ is a real Hilbert space, $\mathcal{A} : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximally monotone operator and $\mathcal{B} : \mathcal{H} \to \mathcal{H}$ is a Lipschitz continuous operator, such that the sum $\mathcal{S} := \mathcal{A} + \mathcal{B}$ is maximally monotone.
Throughout this paper, we assume that the solution set of the (SMI), denoted by $\Omega := \mathrm{Zer}(\mathcal{A} + \mathcal{B})$ is nonempty.
When $\mathcal{A} = N_C$, the normal cone of a nonempty closed convex subset C of $\mathcal{H}$, the structured monotone inclusion (SMI) is reduced to the variational inequality problem (VIP):

$$\text{Find } u^* \in C \text{ such that } \langle \mathcal{B}u^*, u - u^* \rangle \geq 0, \ \forall u \in C. \tag{VIP}$$

As is well known, solutions to monotone inclusions in general are not unique and do not depend continuously on the input data. Besides, approximate methods can in general provide only weak convergence to a solution.

Pham Ky Anh
Department of Mathematics, Vietnam National University, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam
E-mail: anhpk@vnu.edu.vn

Trinh Ngoc Hai
School of Applied Mathematics and Informatics, Hanoi University of Science and Technology, 1 Dai Co Viet, Hai Ba Trung, Hanoi, Vietnam
E-mail: hai.trinhngoc@hust.edu.vn

For solving (SMI) in a stable manner, we will study the so-called regularized monotone inclusion

$$\text{Find } u \in \mathcal{H} \text{ such that } 0 \in (\mathcal{A}+\mathcal{B})u + \alpha \mathcal{F}u, \tag{RMI}$$

where $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ is $\gamma-$strongly monotone and $K-$Lipschitz continuous, and $\alpha > 0$ is a regularization parameter. The regularized monotone inclusion (RMI) is well-posed in the sense that for each $\alpha > 0$, it possesses a unique solution u_α.
Among all the solutions to (SMI), we will find an "optimal" one by solving the variational inequality problem on the solution set Ω of the (SMI):

$$\text{Find } u^\dagger \in \Omega \text{ such that } \left\langle \mathcal{F}u^\dagger, u^* - u^\dagger \right\rangle \geq 0, \ \forall u^* \in \Omega. \tag{1}$$

Under the assumptions that $\mathcal{A}$ is maximally monotone, $\mathcal{B}$ is monotone and Lipschitz continuous, and the parameters $\alpha_n, \ \lambda_n$ are chosen properly, the so-called iterative regularization forward-backward method [19]

$$u_{n+1} = J_{\lambda_n \mathcal{A}}(u_n - \lambda_n(\mathcal{B}u_n + \alpha_n \mathcal{F}u_n))$$

strongly converges to the optimal solution $u^\dagger$. Here, $J_{\lambda\mathcal{A}}$ denotes the resolvent of the operator $\mathcal{A}$.
A time-continuous counterpart of the iterative regularization forward-backward method is the Tikhonov regularized forward-backward dynamical system, proposed in [15]

$$\begin{cases} \dot{u}(t) = \mu(t)\left(J_{\lambda(t)\mathcal{A}}\left(u(t) - \lambda(t)(\mathcal{B}u(t) + \alpha(t)\mathcal{F}u(t))\right) - u(t)\right), \\ u(0) = u_0. \end{cases} \tag{2}$$

If $\mathcal{A}$ is a maximally monotone operator, $\mathcal{B}$ is a co-coercive operator, and $\mathcal{F} = I$-the identity operator, while the functions $\alpha(t), \ \lambda(t), \ \mu(t)$ are chosen suitably, the trajectories of (2) strongly converge to the minimum-norm solution of (SMI).
In the last decades, dynamical systems governed by maximally monotone operators have attracted much attention of researchers. Without intending to review the huge related literature, we just mention [1,2,4–6,9–15,18,22] and references therein.
Further, we refer the reader to [3,8,17] for more background, and to [7,15,19] for a treatment adapted to the present setting.
The paper is organized as follows. In Section 2, we recall some notions and concepts for subsequent needs. In Section 3, we establish the existence and uniqueness of the global solution to (2) and prove its strong convergence to $u^\dagger$. Further, we obtain a convergence rate under the maximal monotonicity and strong pseudo-monotonicity assumption on $\mathcal{A}+\mathcal{B}$. In Sections 4, we provide an iterative regularization forward-backward method with relaxation parameters by discretizing the corresponding dynamical system. Finally, in Section 5, we give several numerical experiments to illustrate the performance of the proposed method.

2 Preliminaries

We begin by recalling some notations and concepts from variational analysis. Let $\mathcal{A} : \mathcal{H} \to 2^{\mathcal{H}}$ be a multi-valued operator acting in a real Hilbert space $\mathcal{H}$. The graph of $\mathcal{A}$ is defined by

$$\text{Graph}(\mathcal{A}) = \{(x, u) : x \in \mathcal{H}, u \in \mathcal{A}x\}.$$

A multi-valued operator $\mathcal{A} : \mathcal{H} \to 2^{\mathcal{H}}$ is called: (i) *monotone*, if $\langle u - v, x - y\rangle \geq 0$ for all $x, \, y \in \mathcal{H}$ and $u \in \mathcal{A}x, \, v \in \mathcal{A}y$; (ii) γ - *strongly monotone*, if there exists $\gamma > 0$ such that $\langle u - v, x - y\rangle \geq \gamma ||x - y||^2$ for all $x, \, y \in \mathcal{H}$ and $u \in \mathcal{A}x, \, v \in \mathcal{A}y$; (iii) ρ-*strongly pseudo-monotone* with $\rho > 0$, if for every $x, \, y \in \mathcal{H}, \, u \in \mathcal{A}x, \, v \in \mathcal{A}y$ it holds

$$\langle u, y - x\rangle \geq 0 \Rightarrow \langle v, y - x\rangle \geq \rho \|x - y\|^2;$$

(iv) *maximally monotone*, if $\mathcal{A}$ is monotone and its graph is not properly contained in the graph of any other monotone operator.
The resolvent $J_{\lambda\mathcal{A}} := (I + \lambda\mathcal{A})^{-1}$ of the maximal operator $\lambda\mathcal{A}$ for $\lambda > 0$ is a single-valued operator, defined on the whole space $\mathcal{H}$ and it is firmly nonexpansive, i.e.,

$$\langle J_{\lambda\mathcal{A}}u - J_{\lambda\mathcal{A}}v, u - v\rangle \geq \|J_{\lambda\mathcal{A}}u - J_{\lambda\mathcal{A}}v\|^2, \ \forall u, v \in \mathcal{H}.$$

A single-valued operator $\mathcal{B} : \mathcal{H} \to \mathcal{H}$ is called: (i) *Lipschitz (L−Lipschitz) continuous*, if there exists $L > 0$, such that $\|\mathcal{B}x - \mathcal{B}y\| \leq L\|x - y\|$ for all $x, y \in \mathcal{H}$; (ii) *monotone*, if $\langle \mathcal{B}x - \mathcal{B}y, x - y\rangle \geq 0$ for all $x, y \in \mathcal{H}$; (iii) *strongly (γ− strongly) monotone*, if there exists a constant $\gamma > 0$, such that $\langle \mathcal{B}x - \mathcal{B}y, x - y\rangle \geq \gamma\|x - y\|^2$ for all $x, y \in \mathcal{H}$; (iv) *co-coercive (β-co-coercive)* if there exists $\beta > 0$, such that$\langle \mathcal{B}x - \mathcal{B}y, x - y\rangle \geq \beta\|\mathcal{A}x - \mathcal{A}y\|^2$ for all $x, y \in \mathcal{H}$.
If $\mathcal{A} : \mathcal{A} \to 2^{\mathcal{H}}$ is a maximally monotone operator and $\mathcal{B} : \mathcal{H} \to \mathcal{H}$ is a Lipschitz continuous and monotone operator, then the sum $\mathcal{S} := \mathcal{A} + \mathcal{B}$ is a maximally monotone operator, see [17, Lemma 2.4].
Thus, if $\mathcal{A}$ is maximally monotone and $\mathcal{B}$ is co-coercive, then $\mathcal{B}$ is monotone and Lipschitz continuous, hence the sum $\mathcal{S}$ is maximally monotone.
The converse need not be true. Indeed, let $\mathcal{B} \neq 0$ be a bounded linear operator, satisfying $\langle \mathcal{B}x, x\rangle = 0$ for all $x \in \mathcal{H}$. Then $\mathcal{B}$ is monotone and Lipschitz continuous, however, it is not co-coercive (cf. Example 4.1 below). Further, let $\mathcal{B}$ be a nonexpansive operator, which is not monotone and let $\mathcal{A} = N_C + I$, where N_C is the normal cone of a nonempty closed convex subset $C \subset \mathcal{H}$ and I is the identity operator. According to [8, Example 20.26], $I + \mathcal{B}$ is monotone and Lipschitz continuous, hence, by [17, Lemma 2.4], the operators $\mathcal{A} = N_C + I$ and $\mathcal{A} + \mathcal{B} = N_C + I + \mathcal{B}$ are maximally monotone.
Assume that the operator $\mathcal{A}+\mathcal{B}$ is maximally monotone and for each $\alpha > 0$, $\alpha\mathcal{F}$ is Lipschitz continuous and strongly monotone, then the operator $\mathcal{A} + \mathcal{B} + \alpha\mathcal{F}$ is maximally monotone and strongly monotone, hence its solution set $\text{Zer}\,(\mathcal{A} + \mathcal{B} + \alpha\mathcal{F}) = \{u_\alpha\}$ is a singleton (see [8, Corollary 23.37]).
For analyzing the convergence of trajectories of dynamical systems, we will use the well-known Minty Lemma and a technical lemma on a sequence of nonnegative real numbers.

Lemma 2.1 [16, Lemma 2.1 (Minty)] *Let $C \subset \mathcal{H}$ be a nonempty, closed, convex set, and $F : C \to C$ be a continuous, pseudomonotone mapping. Then,*

$$\{x^* \in C : \langle Fx^*, y - x^*\rangle \geq 0\ \forall y \in C\} = \{x^* \in C : \langle Fy, y - x^*\rangle \geq 0\ \forall y \in C\}.$$

Lemma 2.2 [23] *Let $\{a_k\}$, $\{\theta_k\} \subset (0, \infty)$, $\{\zeta_k\} \in (0, 1)$ be sequences satisfying*

$$\begin{cases} a_{k+1} \leq (1 - \zeta_k)a_k + \theta_k\ \forall k \geq 0 \\ \lim_{k\to\infty} \zeta_k = 0;\ \sum_{k=0}^{\infty} \zeta_k = \infty;\ \lim_{k\to\infty} \frac{\theta_k}{\zeta_k} = 0. \end{cases}$$

Then, $a_k \to 0$ as $k \to \infty$.

Finally, we will denote by $AC_{\text{loc}}([0, +\infty), \mathcal{H})$, $L^1_{\text{loc}}([0, +\infty), \mathcal{H})$ the spaces of locally absolutely continuous functions and locally integrable functions, respectively. For more details as well as for unexplained terminologies and notations we refer to [7, 15, 19].

3 The regularized forward-backward dynamics

3.1 Existence and uniqueness of global solutions

Consider the dynamical system of equations (2), where, λ and α are positive and continuous functions. The continuity of these functions is assumed just for the sake of simplicity. Besides, for the same reason we put $\mu(t) \equiv 1$.
Thus, we focus on the following system

$$\begin{cases} \dot{u}(t) = J_{\lambda(t)\mathcal{A}}\,(u(t) - \lambda(t)(\mathcal{B}u(t) + \alpha(t)\mathcal{F}u(t))) - u(t), \\ u(0) = u_0, \end{cases} \tag{3}$$

which sometimes is rewritten in the form of a differential-algebraic equation

$$\begin{cases} z(t) = J_{\lambda(t)\mathcal{A}}\,(u(t) - \lambda(t)(\mathcal{B}u(t) + \alpha(t)\mathcal{F}u(t))) \\ \dot{u}(t) = z(t) - u(t), \\ u(0) = u_0, \end{cases} \tag{4}$$

Following [7], we call $u : [0, +\infty) \to \mathcal{H}$ a strong global solution of (3) if the following properties hold:

i) $u \in AC_{\text{loc}}([0,+\infty),\mathcal{H})$, i.e., $u(t)$ is absolutely continuous on each interval $[0,T]$, for any $0 < T < +\infty$;
ii) For almost everywhere $t \in [0,+\infty)$ equation (3) holds;
iii) $u(0) = u_0$.

Define the function $f : (0,+\infty) \times (0,+\infty) \times \mathcal{H} \to \mathcal{H}$ as

$$f(\alpha,\lambda,u) := J_{\lambda\mathcal{A}}\left(u - \lambda(\mathcal{B}u + \alpha\mathcal{F}u)\right) - u.$$

Then (3) is reduced to an initial-value problem for the non-autonomous differential equation

$$\begin{cases} \dot{u}(t) = f(\alpha(t),\lambda(t),u(t)) \\ u(0) = u_0. \end{cases} \tag{5}$$

The following result is established in [15, Theorem 2.7], however we provide a straightforward proof for the reader's ease.

Theorem 3.1 *Let $\alpha : [0,+\infty) \to (0,\alpha^*) \subset (0,+\infty)$ and $\lambda : [0,+\infty) \to (0,\lambda^*) \subset (0,+\infty)$ be two continuous functions. Then for each $u_0 \in \mathcal{H}$, there exists a unique global solution $u \in AC_{loc}([0,+\infty),\mathcal{H})$, satisfying equation (5) for almost every $t \in [0,+\infty)$, and $u(0) = u_0$.*

Proof For applying the Cauchy-Lipschitz-Picard theorem on the existence and uniqueness of the global solution to (5), see, [20, Prop. 6.2.1], we need to verify the following conditions:

(i) $\forall u \in \mathcal{H},\ f(\cdot,\cdot,u) \in L^1_{\text{loc}}\left([0,+\infty),\mathcal{H}\right);$
(ii) $\forall t \in [0,+\infty),\ f(\alpha(t),\lambda(t),\cdot) : \mathcal{H} \to \mathcal{H}$ is continuous, moreover, $\forall u,v \in \mathcal{H},\ \|f(\alpha(t),\lambda(t),u) - f(\alpha(t),\lambda(t),v)\| \le \omega(t,\|u\|+\|v\|)\|u-v\|$, where $\forall r > 0,\ \omega(t,r) \in L^1_{\text{loc}}[0,+\infty)$.
(iii) $\forall t \in [0,+\infty),\ \|f(\alpha(t),\lambda(t),u)\| \le \sigma(t)(1+\|u\|)$, where $\sigma \in L^1_{\text{loc}}[0,+\infty)$.

For better readability, we denote $J := J_{\lambda\mathcal{A}}$, $\mathcal{C} := I - \lambda\mathcal{B} - \lambda\alpha\mathcal{F}$ and rewrite the right-hand-side of (5) as $f(\alpha,\lambda,u) = (J \circ \mathcal{C} - I)\,u$.
First, observe that the function $\alpha \longmapsto f(\alpha,\lambda,u)$ is continuous on $[0,+\infty)$. Further, due to [7, Lemma 1], the function $\lambda \longmapsto f(\alpha,\lambda,u)$ is continuous on $(0,+\infty)$ and $\lim_{\lambda\downarrow 0} f(\alpha,\lambda,u) = 0$, for every $u \in \mathrm{Dom}\mathcal{A}$. Since $u^\dagger \in \mathrm{Dom}\mathcal{A}$ is the unique solution to the bilevel problem (1), the function $\lambda \longmapsto f(\alpha,\lambda,u^\dagger)$ can be extended continuously on the interval $[0,+\infty)$, hence the function

$$t \longmapsto \varphi(t) := \|f(\alpha(t),\lambda(t),u^\dagger)\|$$

is continuous on $[0,+\infty)$.
Next we show that the function $f(\alpha,\lambda,u)$ is globally Lipschitz continuous w.r.t. the third variable. Indeed, for all $u,v \in \mathcal{H}$, by the firm non-expansiveness of J and the Lipschitz continuity of $\mathcal{B}$ and $\mathcal{F}$, we get:
$\|f(\alpha,\lambda,u) - f(\alpha,\lambda,v)\| = \|J\circ\mathcal{C}u - J\circ\mathcal{C}v - (u-v)\| \le (2+\lambda L+\alpha\lambda K)\|u-v\| \le M^*\|u-v\|$, where $M^* := 2 + \lambda^*L + \alpha^*\lambda^*K$. Thus, condition (ii) is satisfied.
Since the function $\varphi(t)$ is continuous, the function $\sigma(t) := \max\{\varphi(t) + M^*\|u^\dagger\|, M^*\}$ is also continuous on $[0,+\infty)$. We find $\|f(\alpha(t),\lambda(t),u)\| \le \|f(\alpha(t),\lambda(t),u^\dagger)\| + \|f(\alpha(t),\lambda(t),u) - f(\alpha(t),\lambda(t),u^\dagger)\| \le \varphi(t) + M^*\|u-u^\dagger\| \le \varphi(t) + M^*\|u^\dagger\| + M^*\|u\| \le \sigma(t)(1+\|u\|)$, hence condition (iii) holds.
Finally, recalling that for each $u \in \mathcal{H}$, the function $\alpha \longmapsto f(\alpha,\lambda,u)$ is continuous on $[0,+\infty)$, while the function $\lambda \longmapsto f(\alpha,\lambda,u)$ is continuous on $(0,+\infty)$ we can conclude that the function $t \longmapsto f(\alpha(t),\lambda(t),u)$ is measurable on $[0,+\infty)$. Condition (iii) ensures the local integrability of $f(\cdot,\cdot,u)$ for each $u \in \mathcal{H}$, which means Condition (i).
The proof of Theorem 3.1 is complete. ■

3.2 Strong convergence of trajectories

Let $\alpha : [0,+\infty) \to (0,+\infty)$ be a time-varying continuous function. For each $t \in [0,+\infty)$, there exists a unique solution $u_{\alpha(t)}$ to the following time-dependent regularized monotone inclusion:

$$\text{Find } u \in \mathcal{H} \text{ such that } 0 \in (\mathcal{A}+\mathcal{B})u + \alpha(t)\mathcal{F}u. \tag{RMIt}$$

The following result slightly extends [15, Lemma 2.3] and Lemmas 6, 7, 8 in [19].

Lemma 3.1 *Suppose $\mathcal{A}$ is maximally monotone, $\mathcal{B}$ is L-Lipschitz continuous, such that $\mathcal{A}+\mathcal{B}$ is maximally monotone, and $\mathcal{F}$ is K-Lipschitz continuous and γ-strongly monotone. Further, assume that $\alpha : [0,+\infty) \to (0,+\infty)$ is a continuous function. Then it holds:*

(i) *The set $\{u_{\alpha(t)}\}$ is uniformly bounded on the interval $[0,+\infty)$;*
(ii) *There exists a constant $N > 0$, such that for all $t, s \in [0,+\infty)$;*

$$\|u_{\alpha(t)} - u_{\alpha(s)}\| \leq N\frac{|\alpha(t)-\alpha(s)|}{\alpha(s)};$$

(iii) *If $\lim_{t\to+\infty}\alpha(t) = 0$ then $\lim_{t\to+\infty} u_{\alpha(t)} = u^\dagger$.*

Proof (i) We have

$$-\alpha(t)\mathcal{F}u_{\alpha(t)} \in (\mathcal{A}+\mathcal{B})u_{\alpha(t)} \tag{6}$$

and

$$0 \in (\mathcal{A}+\mathcal{B})\bar{u} \quad \forall \bar{u} \in Zer(\mathcal{A}+\mathcal{B}).$$

Using the monotonicity of $\mathcal{A}+\mathcal{B}$, we get

$$\left\langle \mathcal{F}u_{\alpha(t)}, \bar{u} - u_{\alpha(t)} \right\rangle \geq 0.$$

Since $\mathcal{F}$ is γ-strongly monotone, it implies that

$$\left\langle \mathcal{F}\bar{u}, \bar{u} - u_{\alpha(t)} \right\rangle \geq \gamma\|\bar{u} - u_{\alpha(t)}\|^2. \tag{7}$$

Consequently,

$$\|\bar{u} - u_{\alpha(t)}\| \leq \frac{1}{\gamma}\|\mathcal{F}\bar{u}\| \quad \forall t \geq 0,$$

and hence $\{u_{\alpha(t)}\}$ is bounded on the interval $[0,+\infty)$;

(ii) From (6), we get

$$\left\langle \alpha(s)\mathcal{F}u_{\alpha(s)} - \alpha(t)\mathcal{F}u_{\alpha(t)}, u_{\alpha(t)} - u_{\alpha(s)} \right\rangle \geq 0.$$

Combining this with the γ-strong monotonicity of $\mathcal{F}$, we infer that

$$|\alpha(s)-\alpha(t)| \left\langle \mathcal{F}u_{\alpha(t)}, u_{\alpha(t)} - u_{\alpha(s)} \right\rangle \geq \alpha(s)\gamma\|u_{\alpha(t)} - u_{\alpha(s)}\|^2.$$

Thus,

$$\|u_{\alpha(t)} - u_{\alpha(s)}\| \leq N\frac{|\alpha(t)-\alpha(s)|}{\alpha(s)},$$

where $N = \sup\left\{\frac{\|\mathcal{F}u_{\alpha(t)}\|}{\gamma} : t \geq 0\right\} < \infty$.

(iii) Take $u \in \mathcal{H}$, $v \in (\mathcal{A}+\mathcal{B})u$ arbitrarily. From the monotonicity of $\mathcal{A}+\mathcal{B}$, we have

$$\left\langle v + \alpha(t)\mathcal{F}u_{\alpha(t)}, u - u_{\alpha(t)} \right\rangle \geq 0. \tag{8}$$

Let $\hat{u}$ be a weak cluster point of $\{u_{\alpha(t)}\}$. From (8), it implies that

$$\langle v, u - \hat{u} \rangle \geq 0. \tag{9}$$

Since $\mathcal{A}+\mathcal{B}$ is maximally monotone and (9) holds for all $(u,v) \in Graph(\mathcal{A}+\mathcal{B})$, we infer that $\hat{u} \in Zer(\mathcal{A}+\mathcal{B})$. Moreover, from (7) and Lemma 2.1, it implies that $\{u_{\alpha(t)}\}$ has a unique weak cluster point $u^\dagger$, which is the unique solution of problem (1). Thus, $u_{\alpha(t)} \rightharpoonup u^\dagger$. Finally, in (7), let $\bar{u} = u^\dagger$, we get

$$\|u^\dagger - u_{\alpha(t)}\|^2 \leq \frac{1}{\gamma}\left\langle \mathcal{F}u^\dagger, u^\dagger - u_{\alpha(t)} \right\rangle \to 0.$$

Theorem 3.2 *Let $\mathcal{A}$ be a maximally monotone operator, $\mathcal{B}$ be an L-Lipschitz continuous operator, such that the sum $\mathcal{S} := \mathcal{A}+\mathcal{B}$ is maximally monotone. Further, let $\alpha(t)$ be a positive, strictly decreasing and continuously differentiable function and $\lambda(t)$ be a positive, continuous function satisfying the following conditions*

(A1) $\lim_{t\to+\infty}\alpha(t) = 0$; $\lim_{t\to+\infty}\frac{\lambda(t)}{\alpha(t)} = 0$;
(A2) $\int_0^\infty \alpha(t)\lambda(t)\mathrm{d}t = +\infty$;

(A3) $\lim_{t\to+\infty} \frac{\dot{\alpha}(t)}{\alpha^2(t)\lambda(t)} = 0.$

Then the trajectory $u(t)$ defined by (4) converges strongly to the unique solution $u^\dagger$ of problem (1).

Proof Firstly, observe that Conditions (A1) and (A2) imply that $\lim_{t\to+\infty}\lambda(t) = 0$ and $\int_0^\infty \lambda(t)\mathrm{dt} = +\infty$. Moreover, Condition (A3) implies that $\lim_{t\to+\infty}\dot{\alpha}(t) = 0$, which ensures $|\dot{\alpha}(t)| \le C,\ \forall t \in [0,+\infty)$ for some positive constant C. Hence, $\alpha(t)$ is locally absolutely continuous. Further, there are constants $\alpha^* > 0$ and $\lambda^* > 0$, such that all the conditions of Theorem 3.1 are fulfilled, hence there exists a unique global solution $u(t)$ of the initial-value problem (4).

From the relation $\|u_{\alpha(t+\Delta)} - u_{\alpha(t)}\| \le N\frac{|\alpha(t+\Delta)-\alpha(t)|}{\alpha(t+\Delta)}$, we conclude that $u_{\alpha(t)}$ is locally absolutely continuous, i.e., $u_\alpha \in AC_{\mathrm{loc}}[0,+\infty)$. Moreover,

$$\left\|\frac{\mathrm{d}}{\mathrm{dt}}u_{\alpha(t)}\right\| \le N\frac{|\dot{\alpha}(t)|}{\alpha(t)}. \tag{10}$$

Now let us consider the Lyapunov function $V(t) := \frac{1}{2}\|u(t) - u_{\alpha(t)}\|^2$. From (10), we have

$$\begin{aligned}\dot{V}(t) &= \langle \dot{u}(t) - \dot{u}_{\alpha(t)}, u(t) - u_{\alpha(t)}\rangle \\ &\le \left\langle -u(t) + z(t), z(t) - u_{\alpha(t)}\right\rangle - \|u(t) - z(t)\|^2 + N\frac{|\dot{\alpha}(t)|}{\alpha(t)}\|u(t) - u_{\alpha(t)}\|. \end{aligned}\tag{11}$$

From the definition of $z(t)$ in (4), we get

$$u(t) - \lambda(t)\left[\mathcal{B}u(t) + \alpha(t)\mathcal{F}u(t)\right] \in z(t) + \lambda(t)\mathcal{A}z(t),$$

or equivalently,

$$u(t) - z(t) - \lambda(t)\left[\mathcal{B}u(t) - \mathcal{B}z(t) + \alpha(t)\mathcal{F}u(t)\right] \in \lambda(t)(\mathcal{A}+\mathcal{B})z(t). \tag{12}$$

Since $u_{\alpha(t)}$ is a solution of (RMIt), it follows that

$$-\lambda(t)\alpha(t)\mathcal{F}u_{\alpha(t)} \in \lambda(t)(\mathcal{A}+\mathcal{B})u_{\alpha(t)}. \tag{13}$$

Using (12), (13) and the monotonicity of $(\mathcal{A}+\mathcal{B})$, we find

$$\left\langle u(t) - z(t) - \lambda(t)\left[\mathcal{B}u(t) - \mathcal{B}z(t)\right] - \lambda(t)\alpha(t)\left[\mathcal{F}u(t) - \mathcal{F}u_{\alpha(t)}\right], z(t) - u_{\alpha(t)}\right\rangle \ge 0. \tag{14}$$

Combining (14) with the γ-strong monotonicity, the K-Lipschitz continuity of $\mathcal{F}$ and the L-Lipschitz continuity of $\mathcal{B}$, we obtain

$$\begin{aligned}&\left\langle u(t) - z(t), z(t) - u_{\alpha(t)}\right\rangle \ge \\ &\ge \lambda(t)\alpha(t)\left\langle \mathcal{F}u(t) - \mathcal{F}u_{\alpha(t)}, z(t) - u_{\alpha(t)}\right\rangle + \lambda(t)\left\langle \mathcal{B}u(t) - \mathcal{B}z(t), z(t) - u_{\alpha(t)}\right\rangle \\ &\ge -K\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|\|z(t) - u(t)\| + \gamma\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|^2 \\ &\quad - \lambda(t)L\|u(t) - z(t)\|\|z(t) - u_{\alpha(t)}\|. \end{aligned}\tag{15}$$

From (11) and (15), we have

$$\begin{aligned}\dot{V}(t) &\le K\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|\|z(t) - u(t)\| - \gamma\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|^2 + \\ &+ N\frac{|\dot{\alpha}(t)|}{\alpha(t)}\|u(t) - u_{\alpha(t)}\| - \|u(t) - z(t)\|^2 + \lambda(t)L\|u(t) - z(t)\|\|z(t) - u_{\alpha(t)}\|. \end{aligned}\tag{16}$$

It is straightforward to check that

$$K\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|\|z(t) - u(t)\| - \frac{\gamma}{4}\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|^2 - \frac{1}{3}\|u(t) - z(t)\|^2 \le 0 \tag{17}$$

if

$$K^2 - \frac{\gamma}{3\alpha(t)\lambda(t)} \le 0$$

or equivalently

$$\alpha(t) \le \frac{\gamma}{3K^2\lambda(t)}. \tag{18}$$

Since $\lambda(t)\alpha(t) \to 0$ as $t \to \infty$, without loss of generality, we may assume that (18), and hence (17), hold for all $t \ge 0$. Applying the inequality of arithmetic and geometric means, we have

$$-\frac{\gamma}{4}\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|^2 + N\frac{|\dot{\alpha}(t)|}{\alpha(t)}\|u(t) - u_{\alpha(t)}\| \le \frac{N^2\dot{\alpha}^2(t)}{\alpha^3(t)\gamma\lambda(t)} \tag{19}$$

and

$$\begin{aligned}\lambda(t)L\|u(t) - z(t)\|\|z(t) - u_{\alpha(t)}\| &\le \frac{1}{4}\|u(t) - z(t)\|^2 + (\lambda(t)L)^2\|z(t) - u_{\alpha(t)}\|^2 \\ &\le \left(\frac{1}{4} + 2(\lambda(t)L)^2\right)\|u(t) - z(t)\|^2 + 2(\lambda(t)L)^2\|u(t) - u_{\alpha(t)}\|^2.\end{aligned} \tag{20}$$

Since $\lim_{t\to\infty}\frac{\lambda(t)}{\alpha(t)} = 0$ and $\lim_{t\to\infty}\lambda(t) = 0$, without loss of generality we may assume that

$$2(\lambda(t)L)^2\|u(t) - u_{\alpha(t)}\|^2 \le \frac{\gamma}{4}\lambda(t)\alpha(t)\|u(t) - u_{\alpha(t)}\|^2 \ \forall t \ge 0 \tag{21}$$

and

$$2(\lambda(t)L)^2 \le \frac{5}{12} \ \forall t \ge 0. \tag{22}$$

Combining (16), (17), (19), (20), (21) and (22) we have

$$\dot{V}(t) + \frac{1}{2}\gamma\lambda(t)\alpha(t)V(t) \le \frac{N^2\dot{\alpha}^2(t)}{\alpha^3(t)\gamma\lambda(t)}.$$

The last inequality can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{dt}}\left(V(t)e^{\frac{1}{2}\int_0^t \gamma\lambda(u)\alpha(u)du}\right) \le \frac{\mathrm{d}}{\mathrm{dt}}\left(\int_0^t e^{\frac{1}{2}\int_0^u \gamma\lambda(s)\alpha(s)ds}\frac{N^2\dot{\alpha}^2(u)}{\alpha^3(u)\gamma\lambda(u)}du\right)$$

or

$$\frac{\mathrm{d}}{\mathrm{dt}}\left(V(t)e^{\frac{1}{2}\int_0^t \gamma\lambda(u)\alpha(u)du} - \int_0^t e^{\frac{1}{2}\int_0^u \gamma\lambda(s)\alpha(s)ds}\frac{N^2\dot{\alpha}^2(u)}{\alpha^3(u)\gamma\lambda(u)}du\right) \le 0.$$

It implies that the function

$$h(t) := V(t)e^{\frac{1}{2}\int_0^t \gamma\lambda(u)\alpha(u)du} - \int_0^t e^{\frac{1}{2}\int_0^u \gamma\lambda(s)\alpha(s)ds}\frac{N^2\dot{\alpha}^2(u)}{\alpha^3(u)\gamma\lambda(u)}du$$

is decreasing and hence, $h(t) \le h(0) = V(0)$ for all $t \ge 0$. We obtain

$$V(t) \le e^{-\frac{1}{2}\int_0^t \gamma\lambda(u)\alpha(u)du}\left(\int_0^t e^{\frac{1}{2}\int_0^u \gamma\lambda(s)\alpha(s)ds}\frac{N^2\dot{\alpha}^2(u)}{\alpha^3(u)\gamma\lambda(u)}du + V(0)\right) \tag{23}$$

Due to assumption (A2), we have

$$\lim_{t\to\infty} e^{\frac{1}{2}\int_0^t \gamma\lambda(u)\alpha(u)du} = \infty.$$

If

$$\int_0^\infty e^{\frac{1}{2}\int_0^u \gamma\lambda(s)\alpha(s)ds}\frac{N^2\dot{\alpha}^2(u)}{\alpha^3(u)\gamma\lambda(u)}du < \infty,$$

then from (23) it implies that $V(t) \to 0$. In the opposite case, applying l'Hospital's rule, we have

$$\begin{aligned}\lim_{t\to\infty} V(t) &= \lim_{t\to\infty}\frac{e^{\frac{1}{2}\int_0^t \gamma\lambda(s)\alpha(s)ds}\frac{N^2\dot{\alpha}^2(t)}{\alpha^3(t)\gamma\lambda(t)}}{\frac{1}{2}e^{\frac{1}{2}\int_0^t \gamma\lambda(u)\alpha(u)du}\gamma\lambda(t)\alpha(t)} \\ &= \lim_{t\to\infty}\frac{2N^2\dot{\alpha}^2(t)}{\alpha^4(t)\gamma^2\lambda^2(t)} \\ &= 0.\end{aligned}$$

The last equality comes from condition (A3). Taking into account Lemma 3.1-(iii), we obtain the desired result. ∎

Remark 3.1 An example of $\alpha(t)$, $\lambda(t)$ satisfying conditions (A1)-(A3) in Theorem 3.2 is $\lambda(t) = \frac{1}{(t+1)^p}$, $\alpha(t) = \frac{1}{(t+1)^q}$, where $0 < q < p$, $q + p < 1$.

Remark 3.2 Note that conditions imposed on the operator $\mathcal{B}$ in Theorem 3.2 are also necessary. Indeed, Let $\mathcal{A} = 0$, then being monotone and continuous, $\mathcal{A}$ is maximally monotone. Now let $\mathcal{B} = -I$, and $\mathcal{F} = I$. The sum $\mathcal{A} + \mathcal{B} = -I$ is not monotone.
Since $J_{\lambda\mathcal{A}} = (I + \lambda\mathcal{A})^{-1} = I$, our dynamical system is of the form:

$$\begin{cases} \dot{u}(t) = \lambda(t)\left(1 - \alpha(t)\right)u(t) \\ u(0) = u_0. \end{cases}$$

Thus $u(t) = u(0)e^{\int_0^t \lambda(s)(1-\alpha(s))\mathrm{d}s}$. Since $\alpha(t) \to 0$ as $t \to +\infty$, without loss of generality, we can assume that $0 < \alpha(t) < \frac{1}{2}$ for all $t \geq 0$. We have

$$\int_0^t \lambda(s)(1 - \alpha(s))\mathrm{d}s = \int_0^t \lambda(s)\alpha(s)\frac{1 - \alpha(s)}{\alpha(s)}\mathrm{d}s \geq \int_0^t \lambda(s)\alpha(s)\mathrm{d}s \to +\infty$$

as $t \to +\infty$. Therefore, $\|u(t)\| \to +\infty$ as $t \to +\infty$.

If in (4), instead of regularizing $\mathcal{B}$, we regularize $\mathcal{A}$, then the dynamical system becomes:

$$\begin{cases} z(t) = J_{\lambda(t)(\mathcal{A}+\alpha(t)\mathcal{F})}\left(u(t) - \lambda(t)\mathcal{B}u(t)\right) \\ \dot{u}(t) = -u(t) + z(t) \\ u(0) = u_0 \in \mathcal{H}, \end{cases} \tag{24}$$

The existence and the uniqueness of global solutions of (24) can be obtained similarly as in Theorem 3.1. Meanwhile, the convergence of the trajectory $u(t)$ can be proved in almost the same way as the proof of Theorem 3.2. Note that from the definition of $z(t)$ in (24), we have

$$u(t) - \lambda(t)\mathcal{B}u(t) \in z(t) + \lambda(t)(\mathcal{A} + \alpha(t)\mathcal{F})z(t),$$

which is equivalent to (12).

Corollary 3.1 *In Theorem 3.2, if in addition, $(\mathcal{A}+\mathcal{B})$ is ρ-strongly pseudomonotone, then there exist $\tau, \zeta > 0$ such that*

$$\|u(t) - u^\dagger\|^2 \leq \tau\left(\alpha^2(t) + e^{-\zeta\int_0^t \lambda(u)du}\right) \quad \forall t \geq 0.$$

Proof From the definitions of $z(t)$ and $u^\dagger$, we have

$$\begin{aligned} u(t) - z(t) - \lambda(t)\left(\mathcal{B}u(t) - \mathcal{B}z(t) + \alpha(t)\mathcal{F}u(t)\right) &\in \lambda(t)(\mathcal{A} + \mathcal{B})z(t) \\ 0 &\in \lambda(t)(\mathcal{A} + \mathcal{B})u^\dagger. \end{aligned}$$

Using the strong pseudomonotonicity of $\mathcal{A} + \mathcal{B}$, we get

$$\left\langle u(t) - z(t) - \lambda(t)\left(\mathcal{B}u(t) - \mathcal{B}z(t) + \alpha(t)\mathcal{F}u(t)\right), z - u^\dagger\right\rangle \geq \lambda(t)\rho\|z(t) - u^\dagger\|^2. \tag{25}$$

From the proof of Theorem 3.2, we infer that the function $u(t)$ is bounded, and hence, so is $\mathcal{F}u(t)$. There exists a constant $\kappa > 0$ such that $\|\mathcal{F}u(t)\| \leq \kappa$ for all $t \geq 0$. Let $G(t) := \frac{1}{2}\|u(t) - u^\dagger\|^2$, (25) implies

$$\begin{aligned} \dot{G}(t) &= -(z(t) - u(t))^2 + \left\langle z(t) - u(t), z(t) - u^\dagger\right\rangle \\ &\leq -(z(t) - u(t))^2 - \lambda(t)\rho\|z(t) - u^\dagger\|^2 + L\lambda(t)\|u(t) - z(t)\|\|z - u^\dagger\| + \kappa\lambda(t)\alpha(t)\|z - u^\dagger\|. \end{aligned} \tag{26}$$

We observe that

$$L\lambda(t)\|u(t) - z(t)\|\|z - u^\dagger\| \leq \frac{1}{3}\lambda(t)\rho\|z(t) - u^\dagger\|^2 + \frac{3\lambda(t)L^2}{4\rho}\|u(t) - z(t)\|^2, \tag{27}$$

$$\kappa\lambda(t)\alpha(t)\|z - u^\dagger\| \leq \frac{1}{3}\lambda(t)\rho\|z(t) - u^\dagger\|^2 + \frac{3\kappa^2\lambda(t)\alpha(t)^2}{4\rho}. \tag{28}$$

Combining (26), (27) and (28), we obtain

$$\begin{aligned}\dot{G}(t) &\leq -\left(1-\frac{3\lambda(t)L^2}{4\rho}\right)\|u(t)-z(t)\|^2-\frac{1}{3}\lambda(t)\rho\|z(t)-u^\dagger\|^2+\frac{3\kappa^2\lambda(t)\alpha(t)^2}{4\rho}\\ &\leq -\left(1-\frac{3\lambda(t)L^2}{4\rho}-\frac{1}{3}\lambda(t)\rho\right)\|u(t)-z(t)\|^2-\frac{1}{6}\lambda(t)\rho\|u(t)-u^\dagger\|^2+\frac{3\kappa^2\lambda(t)\alpha(t)^2}{4\rho}.\end{aligned}$$

Since $\lim_{t\to\infty}\lambda(t)=0$, without loss of generality we may assume that $1-\frac{3\lambda(t)L^2}{4\rho}-\frac{1}{3}\lambda(t)\rho\geq 0$ for all $t\geq 0$. Thus,

$$\dot{G}(t)+\frac{1}{3}\lambda(t)\rho G(t)\leq\frac{3\kappa^2\lambda(t)\alpha(t)^2}{4\rho}.$$

Consequently,

$$G(t)\leq e^{-\frac{1}{3}\int_0^t\lambda(u)\rho du}\left(\int_0^t e^{\frac{1}{3}\int_0^u\lambda(s)\rho ds}\frac{3\kappa^2\lambda(u)\alpha(u)^2}{4\rho}du+G(0)\right).$$

If $\int_0^\infty e^{\frac{1}{3}\int_0^u\lambda(s)\rho ds}\lambda(u)\alpha(u)^2du<\infty$, the assertion is established immediately. In the opposite case, we will show that $e^{-\frac{1}{3}\int_0^t\lambda(u)\rho du}\int_0^t e^{\frac{1}{3}\int_0^u\lambda(s)\rho ds}\lambda(u)\alpha(u)^2du=\mathcal{O}\left(\alpha^2(t)\right)$ as $t\to\infty$. Applying the L'Hospital's rule, we have

$$\lim_{t\to\infty}\frac{\int_0^t e^{\frac{1}{3}\int_0^u\lambda(s)\rho ds}\lambda(u)\alpha(u)^2du}{\alpha^2(t)e^{\frac{1}{3}\int_0^t\lambda(u)\rho du}}=\lim_{t\to\infty}\frac{e^{\frac{1}{3}\int_0^t\lambda(s)\rho ds}\lambda(t)\alpha(t)^2}{e^{\frac{1}{3}\int_0^t\lambda(u)\rho du}\left(2\alpha(t)\dot{\alpha}(t)+\frac{1}{3}\rho\lambda(t)\alpha^2(t)\right)}=\frac{3}{\rho}.$$

The last equality is implied from Condition (A3). The proof is complete.■

3.3 Time discretization of the dynamical system

Using the approximation $\dot{u}(t_k)\approx\frac{u(t_{k+1})-u(t_k)}{h_k}$, we propose a discrete-time version of the dynamical system (4):

$$\begin{cases} u^0\in\mathcal{H},\\ z^k=J_{\lambda_k\mathcal{A}}\left(u^k-\lambda_k(\mathcal{B}u^k+\alpha_k\mathcal{F}u^k)\right),\\ u^{k+1}=(1-h_k)u^k+h_kz^k.\end{cases}\tag{29}$$

Theorem 3.3 *Assume that the parameters in (29) satisfy the following conditions:*

(B1) $\lambda_k,\alpha_k\in(0,\infty)$ *for all* $k\geq 0$*;*
(B2) $\lim_{k\to\infty}\alpha_k=0$*;* $\lim_{k\to\infty}\frac{\lambda_k}{\alpha_k}=0$
(B3) $\lim_{k\to\infty}\frac{|\alpha_{k+1}-\alpha_k|}{\lambda_k\alpha_k^2}=0$*;*
(B4) $\sum_{k=0}^\infty\lambda_k\alpha_k=\infty$*;*
(B5) $h_k\in[a,b]\subset(0,2)$ *for all* $k\geq 0$.

Then, the sequence $\{u^k\}$ *generated by (29) converges strongly to the unique solution* $u^\dagger$ *of (1).*

Proof Denote $u_{\alpha_k}:=(\mathcal{A}+\mathcal{B}+\alpha_k\mathcal{F})^{-1}0$. Lemma 3.1-(iii) implies $u_{\alpha_k}\to u^\dagger$ as $k\to\infty$. We will prove that $\lim_{k\to\infty}\|u^k-u_{\alpha_k}\|=0$. From Lemma 3.1-(ii), we obtain

$$\begin{aligned}\left\|u^{k+1}-u_{\alpha_{k+1}}\right\|^2&=\left\|u^k+h_k\left(-u^k+z^k\right)-u_{\alpha_{k+1}}\right\|^2\\ &\leq\|u^k-u_{\alpha_k}\|^2+h_k^2\|u^k-z^k\|^2+N^2\frac{|\alpha_k-\alpha_{k+1}|^2}{\alpha_k^2}+\\ &\quad+2N\frac{|\alpha_k-\alpha_{k+1}|}{\alpha_k}\|u^k-u_{\alpha_k}\|+2Nh_k\frac{|\alpha_k-\alpha_{k+1}|}{\alpha_k}\|u^k-z^k\|+\\ &\quad+2h_k\left\langle -u^k+z^k,z^k-u_{\alpha_k}\right\rangle-2h_k\|u^k-z^k\|^2.\end{aligned}\tag{30}$$

Applying arguments similar to those of (15), we have

$$\begin{aligned}\left\langle -u^k + z^k, z^k - u_{\alpha_k}\right\rangle \le K\lambda_k\alpha_k\|u^k - u_{\alpha_k}\|\|z^k - u^k\| - \gamma\lambda_k\alpha_k\|u^k - u_{\alpha_k}\|^2 \\ + \lambda_k L\|u_k - z_k\|\|z_k - u_{\alpha_k}\|.\end{aligned} \tag{31}$$

Using (30) and (31), we arrive at

$$\begin{aligned}\left\|u^{k+1} - u_{\alpha_{k+1}}\right\|^2 \le{}& (1 - 2h_k\gamma\lambda_k\alpha_k)\|u^k - u_{\alpha_k}\|^2 - (2h_k - h_k^2)\|u^k - z^k\|^2 + \\ &+ 2N\frac{|\alpha_k - \alpha_{k+1}|}{\alpha_k}\|u^k - u_{\alpha_k}\| + 2Nh_k\frac{|\alpha_k - \alpha_{k+1}|}{\alpha_k}\|u^k - z^k\| + \\ &+ 2h_kK\lambda_k\alpha_k\|u^k - u_{\alpha_k}\|\|z^k - u^k\| + N^2\frac{|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^2} + \\ &+ 2h_k\lambda_k L\|u_k - z_k\|\|z_k - u_{\alpha_k}\|.\end{aligned} \tag{32}$$

Taking any interval $[a, b] \subset (0, 2)$ and defining two continuous functions $a(\xi) := 1 - \sqrt{1 - 3\xi}$, $b(\xi) := 1 + \sqrt{1 - 3\xi}$, where $\xi \in (0, \frac{1}{3})$, we find that $\lim_{\xi\to 0^+} a(\xi) = 0$, and $\lim_{\xi\to 0^+} b(\xi) = 2$. Then for sufficiently small ξ, we get $0 < a(\xi) < a < b < b(\xi) < 2$. Moreover, with the defined above ξ, for any $h \in [a, b]$, there holds the relation $h^2 - 2h + 3\xi < 0$. Thus, since $h_k \in [a, b] \subset (0, 2)$, we get

$$h_k^2 - 2h_k \le -3\xi < 0 \ \forall k \ge 0. \tag{33}$$

It is easy seen that

$$2Nh_k\frac{|\alpha_k - \alpha_{k+1}|}{\alpha_k}\|u^k - z^k\| \le \xi\|u^k - z^k\|^2 + \frac{N^2h_k^2}{\xi}\frac{|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^2}, \tag{34}$$

$$2h_kK\lambda_k\alpha_k\|u^k - u_{\alpha_k}\|\|z^k - u^k\| \le \xi\|u^k - z^k\|^2 + \frac{(h_kK\lambda_k\alpha_k)^2}{\xi}\|u^k - u_{\alpha_k}\|^2, \tag{35}$$

$$2N\frac{|\alpha_k - \alpha_{k+1}|}{\alpha_k}\|u^k - u_{\alpha_k}\| \le \frac{1}{2}h_k\gamma\lambda_k\alpha_k\|u^k - u_{\alpha_k}\|^2 + \frac{2N^2|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^3h_k\gamma\lambda_k} \tag{36}$$

and

$$\begin{aligned}2h_k\lambda_k L\|u_k - z_k\|\|z_k - u_{\alpha_k}\| &\le 2h_k\lambda_k L\|u_k - z_k\|\left(\|z_k - u_k\| + \|u_k - u_{\alpha_k}\|\right) \\ &\le (2h_k\lambda_k L)\|u_k - z_k\|^2 + 2h_k\lambda_k L\|z_k - u_k\|\|u_k - u_{\alpha_k}\| \\ &\le \left(2h_k\lambda_k L + \frac{2L^2h_k\lambda_k}{\gamma\alpha_k}\right)\|u_k - z_k\|^2 + \frac{1}{2}h_k\gamma\lambda_k\alpha_k\|u_k - u_{\alpha_k}\|^2.\end{aligned}$$

Since $\lim_{k\to\infty}\lambda_k = \lim_{k\to\infty}\frac{\lambda_k}{\alpha_k} = 0$, without loss of generality we may assume that $2h_k\lambda_k L + \frac{2L^2h_k\lambda_k}{\gamma\alpha_k} \le \xi$ for all $k \ge 0$. Hence,

$$2h_k\lambda_k L\|u_k - z_k\|\|z_k - u_{\alpha_k}\| \le \xi\|u_k - z_k\|^2 + \frac{1}{2}h_k\gamma\lambda_k\alpha_k\|u_k - u_{\alpha_k}\|^2. \tag{37}$$

From (32), (33), (34), (35), (36) and (37), we get

$$\begin{aligned}\left\|u^{k+1} - u_{\alpha_{k+1}}\right\|^2 \le{}& \left(1 - h_k\gamma\lambda_k\alpha_k + \frac{(h_kK\lambda_k\alpha_k)^2}{\xi}\right)\|u^k - u_{\alpha_k}\|^2 + \frac{2N^2|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^3h_k\gamma\lambda_k} + \\ &+ \left[\frac{N^2h_k^2}{\xi} + N^2\right]\frac{|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^2}.\end{aligned} \tag{38}$$

Since $\lambda_k, \alpha_k \to 0$, without loss of generality we can assume that $\zeta_k := h_k\gamma\lambda_k\alpha_k - \frac{(h_kK\lambda_k\alpha_k)^2}{\xi} \in (0, 1)$ for all $k \ge 0$. Moreover, the condition $\sum_{k=0}^{\infty}\lambda_k\alpha_k = \infty$ implies $\sum_{k=0}^{\infty}\zeta_k = \infty$. Denote

$$\theta_k := \frac{2N^2|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^3h_k\gamma\lambda_k} + \left[\frac{N^2h_k^2}{\xi} + N^2\right]\frac{|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^2}.$$

We have

$$\lim_{k\to\infty}\frac{\theta_k}{\zeta_k} = \lim_{k\to\infty}\frac{2N^2|\alpha_k - \alpha_{k+1}|^2}{\alpha_k^4h_k^2\gamma^2\lambda_k^2} = 0.$$

The last equality is implied from Condition (B3). Using Lemma 2.2 and (38), we get $\|u^k - u_{\alpha_k}\| \to 0$. The proof is complete. ■

Remark 3.3 An example of α_k, λ_k satisfying the conditions in Theorem 3.3 is $\lambda_k = \frac{1}{(k+1)^p}$, $\alpha_k = \frac{1}{(k+1)^q}$, where $0 < q < p$, $q + p < 1$.

Corollary 3.2 *Suppose that the parameters λ_k and α_k satisfy all the conditions in Theorem 3.3. Then, the sequence $\{u^k\}$ generated by*

$$\begin{cases} u^0 \in \mathcal{H}, \\ z^k = J_{\lambda_k(\mathcal{A}+\alpha_k\mathcal{F})}\left(u^k - \lambda_k \mathcal{B}u^k\right), \\ u^{k+1} = (1-h_k)u^k + h_k z^k. \end{cases} \tag{39}$$

converges strongly to the unique solution $u^\dagger$ of (1).

The proof of Corollary 3.2 is similar to that of Theorem 3.3, and hence, is omitted.

Remark 3.4 Algorithms (29) and (39) can be regarded as iterative regularization forward-backward methods with relaxation parameters. The first one recovers [19, Algor. 1], when $h_k \equiv 1$.

The following corollary is a direct consequence of Theorem 3.3 and Corollary 3.2 when $\mathcal{A} = N_C$.

Corollary 3.3 *Let C be a nonempty, closed and convex set in $\mathcal{H}$, and let $\mathcal{B}, \mathcal{F} : C \to C$ be Lipschitz continuous and monotone (strongly monotone, respectively) mappings on C. Suppose that the solution set $Sol(\mathcal{B}, C)$ of (VIP) is not empty, the parameters λ_k, α_k satisfy all the conditions in Theorem 3.3. Then,*

- *The sequence $\{u^k\}$ generated by*

$$\begin{cases} u^0 \in \mathcal{H}, \\ z^k = P_C\left(u^k - \lambda_k \mathcal{B}u^k - \lambda_k\alpha_k\mathcal{F}u^k\right), \\ u^{k+1} = (1-h_k)u^k + h_k z^k. \end{cases} \tag{40}$$

  converges strongly to the unique solution $u^\dagger$ of the following bilevel variational inequality:

$$\text{Find } u^\dagger \in Sol(\mathcal{B}, C) \text{ such that } \left\langle \mathcal{F}u^\dagger, u - u^\dagger \right\rangle \geq 0 \quad \forall u \in Sol(\mathcal{B}, C).$$

- *The sequence $\{u^k\}$ generated by*

$$\begin{cases} u^0 \in \mathcal{H}, \\ z^k = P_C\left(\frac{1}{1+\lambda_k\alpha_k}\left[u^k - \lambda_k \mathcal{B}u^k\right]\right), \\ u^{k+1} = (1-h_k)u^k + h_k z^k. \end{cases} \tag{41}$$

  converges strongly to the minimum-norm solution $u^\dagger$ of (VIP).

Proof It remains to show that $J_{\lambda_k(N_C+\alpha_k I)}(x) = P_C\left(\frac{1}{1+\lambda_k\alpha_k}x\right)$ for all $x \in \mathcal{H}$. Indeed, let $z = J_{\lambda_k(N_C+\alpha_k I)}(x)$. We have $x \in z + \lambda_k N_C z + \alpha_k\lambda_k z$. This is equivalent to $\frac{1}{1+\lambda_k\alpha_k}x - z \in N_C z$ or $z = P_C\left(\frac{1}{1+\lambda_k\alpha_k}x\right)$. We obtain the desired result. ■

4 Numerical experiments

Example 4.1 Let $\mathcal{H} = l^2$, $\mathcal{A} = N_C$, where $C := \{u \in l^2 : u_3 = 0\}$,

$$\mathcal{B}u = (-u_2, u_1, 0, 0, \dots)$$

for all $u = (u_1, u_2, u_3, \dots) \in l^2$. It is easy seen that $\mathcal{A}$ is maximally monotone, $\mathcal{B}$ is a bounded linear operator satisfying $\langle \mathcal{B}u, u\rangle = 0$ for all $u \in l^2$ and $\|\mathcal{B}\| = 1$. Thus, $\mathcal{B}$ is a monotone, Lipschitz continuous operator. Besides, it is not co-coercive.

We apply Algorithm (4) with $\mathcal{F} = I$-the identity mapping. It holds that $J_{\lambda\mathcal{A}}(u) = P_C(u) = (u_1, u_2, 0, u_4, u_5, \ldots)$. The dynamical system (4) now becomes

$$\begin{cases} \dot{u}_1(t) = -\alpha(t)\lambda(t)u_1(t) + \lambda(t)u_2(t), \\ \dot{u}_2(t) = -\lambda(t)u_1(t) - \alpha(t)\lambda(t)u_2(t), \\ \dot{u}_3(t) = -u_3(t), \\ \dot{u}_i(t) = -\alpha(t)\lambda(t)u_i(t), \ \forall i \geq 4, \\ u(0) \in \mathcal{H}. \end{cases}$$

Firstly, we have

$$\begin{cases} u_3(t) = u_3(0)e^{-t}, \\ u_i(t) = u_i(0)e^{-\int_0^t \alpha(s)\lambda(s)\mathrm{d}s}, \ \forall i \geq 4. \end{cases}$$

Setting $v_i(t) := e^{\int_0^t \alpha(s)\lambda(s)\mathrm{d}s}u_i(t), \ i = 1, 2$, from the first two equations, we find

$$\begin{cases} \dot{v}_1(t) = \lambda(t)v_2(t), \\ \dot{v}_2(t) = -\lambda(t)v_1(t), \\ v_1(0) = u_1(0), \ v_2(0) = u_2(0). \end{cases} \tag{42}$$

Further, from (42), we obtain $\frac{\mathrm{d}v_2(t)}{\mathrm{d}v_1(t)} = -\frac{v_1(t)}{v_2(t)}$, or equivalently, $v_2(t)\mathrm{d}v_2(t) + v_1(t)\mathrm{d}v_1(t) = 0$, which follows that $v_1^2(t) + v_2^2(t) = \text{const} = u_1^2(0) + u_2^2(0)$. Thus,

$$u_1^2(t) + u_2^2(t) = (u_1^2(0) + u_2^2(0))e^{-2\int_0^t \alpha(s)\lambda(s)\mathrm{d}s}.$$

Under the assumption that $\int_0^\infty \alpha(t)\lambda(t)\mathrm{d}t = \infty$, we have $u_i(t) \to 0$ as $t \to \infty$ for all $i \geq 1$. Moreover, $\|u(t) - u^\dagger\|^2 = \left(\|u(0)\|^2 - u_3^2(0)\right) e^{-2\int_0^t \alpha(s)\lambda(s)\mathrm{d}s} + u_3^2(0)e^{-2t}$, hence, $u(t)$ converges strongly to the minimum-norm solution $u^\dagger := (0, 0, 0, \ldots)$ of the problem.

Example 4.2 We compare Algorithm (4) with the Boţ's algorithm [15, Algorithm (5.2)] using the example (6.2) in [15]. Let $C = \{x \in \mathbb{R}^3 : 3x_1 - x_2 + x_3 = 0\}$, $\mathcal{A} = N_C$, $\mathcal{B} = Bx$, where

$$B = \begin{pmatrix} 0 & 0.1 & 0.5 \\ -0.1 & 0 & -0.4 \\ -0.5 & 0.4 & 0 \end{pmatrix}.$$

Clearly, $\mathcal{B}$ is not a co-coercive operator, hence, [15, Theorem 4.3] does not work. The parameters of the Boţ's algorithm are the same as in [15], i.e., $\lambda(t) = 0.5$, $\alpha(t) = \frac{1}{(t+1)^\beta}$ for all $t \geq 0$, $\beta = 0.1, 0.5, 0.9$. In our algorithm, according to Conditions (A1)-(A3), we choose $\lambda(t) = \frac{1}{(t+1)^{0.3}}$, $\alpha(t) = \frac{1}{(t+1)^{0.1}}$. The both algorithms use the same starting point $x(0) = u(0) = (1, 1, 1)^T$. We obtain the comparison results in Figure 1. We can see that our algorithm gives better results than the Boţ's algorithm in all three cases $\beta = 0.1, 0.5, 0.9$.

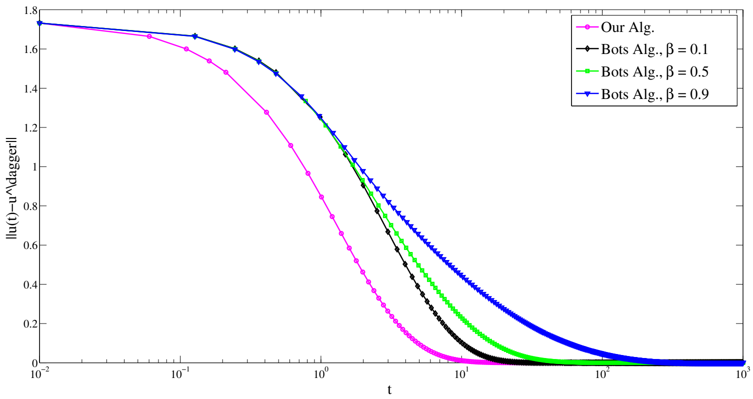

Fig. 1 Comparisons of Algorithm (4) with the Boţ's algorithm [15, Algorithm (5.2)] in Example 4.2

Example 4.3 In this example, we compare algorithm (29) with the Iterative regularization forward-backward method introduced by Hieu et al. [19, Algorithm 1]. Let $\mathcal{A} = N_C$, where

$$C = \{x \in \mathbb{R}^m, -5 \le x_i \le 5,\ \forall i = 1, \dots, m\},$$

and let $\mathcal{B} = Bx$ for all $x \in \mathbb{R}^m$, where $B = (b_{ij})_{m\times m}$,

$$b_{ij} = \begin{cases} -1 & \text{if } j = m+1-i,\ j > i \\ 1 & \text{if } j = m+1-i,\ j < i \\ 0 & \text{otherwise.} \end{cases}$$

It is easy seen that $\mathcal{A}$ is maximally monotone, $\mathcal{B}$ is monotone and 1-Lipschitz continuous on $\mathbb{R}^m$, $Zer(A+B) = \{u^\dagger := (0, \dots, 0)^T\}$. The two algorithms use the same regularization mapping $\mathcal{F} = I$ and the same starting point x^0, which is randomly generated.

- In Hieu's algorithm, as chosen in [19], we have $\lambda_k = \frac{2\gamma - \min(\gamma,1)}{(L+k)^2(k+1)^{0.5}}$, $\alpha_k = \frac{1}{(k+1)^{0.2}}$. Here $L = k = 1$ is the Lipschitz constants of $\mathcal{B}$ and $\mathcal{F}$, $\gamma = 1$ is the strong monotonicity constant of $\mathcal{F}$.
- In our algorithm, we choose $\alpha_k = \frac{1}{(k+1)^{0.3}}$, $\lambda_k = \frac{1}{(k+1)^{0.5}}$ and $h_k = 1.1$ for all $k \ge 0$.

We test the algorithms with different m. The comparison results are presented in Figure 2. As we can see, our algorithm gives a better behavior in term of computational time.

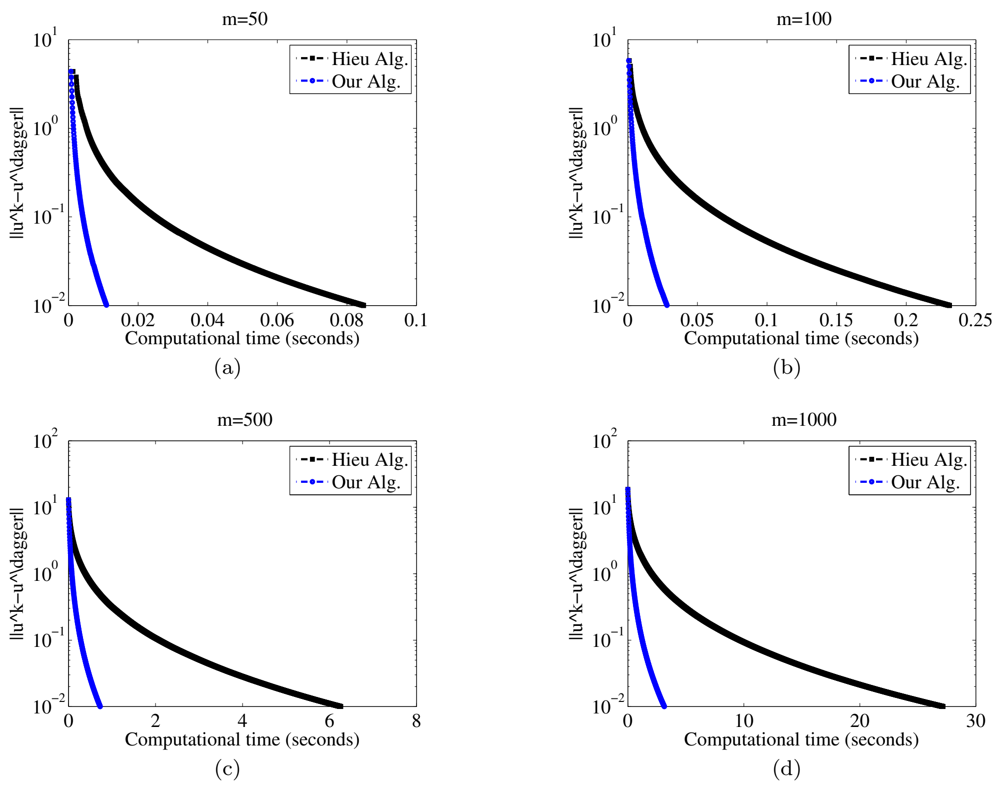

Fig. 2 Comparisons of the two algorithms in Example 4.3

Example 4.4 Now, we compare algorithm (29) with the Gradient projection method (GPM) [21, Algorithm 3.1] in solving a strongly monotone variational inequality. Let $H = \mathbb{R}^m$, $C = \{x \in \mathbb{R}^m : \|x\| \le 1\}$, $\mathcal{B}x = Bx$ for all $x \in \mathbb{R}^m$, where B is the diagonal matrix whose diagonal entries are $10^2, 10^{-2}, 1, \dots, 1$. Note that $\mathcal{B}$ is 10^2-Lipschitz continuous and 10^{-2}-strongly monotone on $\mathbb{R}^m$. The Gradient projection method has the following form:

$$\begin{cases} u^0 \in C, \\ u^{k+1} = P_C\left(u^k - \lambda_k B u^k\right). \end{cases} \tag{GPM}$$

Under the condition $\lambda_k \in [a,b] \subset \left(0, \frac{2\gamma}{L^2}\right)$, (GPM) converges linearly to the unique solution $u^\dagger$ of the problem. Following [21, Remark 4.2], the optimal value of λ_k is $\lambda_k = \frac{\gamma}{L^2}$ for all $k \geq 0$. In our test, we will choose λ_k according to this recommendation for (GPM). The parameters in our algorithm are the same as in the Example 4.3, i.e., $\alpha_k = \frac{1}{(k+1)^{0.3}}$, $\lambda_k = \frac{1}{(k+1)^{0.5}}$ and $h_k = 1.1$ for all $k \geq 0$, $\mathcal{F}$ is the identity mapping. The both algorithms use the same starting point, which is randomly generated. The comparison results are presented in Table 1. Interestingly, in this example, our algorithm converges faster than (GPM) - the algorithm designed specifically for strongly monotone and Lipschitz continuous variational inequalities. This happens because in (GPM), the stepsize λ_k needs to satisfy the condition: $\lambda_k \in [a,b] \subset \left(0, \frac{2\gamma}{L^2}\right)$. We have $L = 100$ and $\gamma = 0.01$, hence $\lambda_k < 2.10^{-6}$. Since the step-size λ_k is too small, the algorithm (GPM) converges very slowly.

	Algorithm (29)			(PGM)		
	`Times(s)`	`Iter.`	$\|u^k - u^\dagger\|$	`Times(s)`	`Iter.`	$\|u^k - u^\dagger\|$
m=100	0.052837	3270	$9.4284.10^{-5}$	0.74027	45449	0.95278
m=200	0.10129	3261	$1.8797.10^{-4}$	1.4186	48505	0.94912
m=500	1.1987	3253	$3.3966.10^{-4}$	13.187	36250	0.96432
m=1000	8.0661	3248	$4.8657\ .10^{-4}$	72.597	29601	0.97067
m=2000	30.396	3169	$4.9176\ .10^{-2}$	273.57	28495	0.9716

Table 1 Comparison of the two algorithms in Example 4.4

5 Conclusion

In this paper, we established strong convergence of the regularized forward-backward dynamics to a solution of an additively structured monotone inclusion under a weak assumption on the single-valued operator $\mathcal{B}$. A convergence estimate is obtained if the composite operator $\mathcal{A} + \mathcal{B}$ is maximally monotone and strongly pseudo-monotone. Time discretization of the continuous dynamics provides an iterative regularization forward-backward method with relaxation parameters. Some simple numerical examples were given to illustrate the agreement between analytical and numerical results.

6 Acknowledgements

The authors gratefully acknowledge support by the Vietnam Institute for Advanced Study in Mathematics. The research of PKA is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant No. 101.01-2017.315.

References

1. B. Abbas, H. Attouch, Dynamical systems and forward-backward algorithms associated with the sum of a convex subdifferential and a monotone cocoercive operator, Optimization, 64(10) (2015) 2223-2252
2. B. Abbas, H. Attouch, B. F. Svaiter, Newton-like dynamics and forward backward methods for structured monotone inclusions in Hilbert spaces, J. Optim. Theory Appl. 161(2) (2014) 331-360
3. Y. Alber, D. Butnariu, I. Ryazantseva, Regularization methods for ill-posed inclusions and variational inequalities with domain perturbations, J. Nonlinear Convex Anal. 2 (2001) 53-79
4. P.K. Anh, T.N. Hai, Dynamical system for solving bilevel variational inequalities, J. Global Optim. 80 (4) (2021) 945-963
5. H. Attouch, F. Alvarez, The heavy ball with friction dynamical system for convex constrained minimization problems. In: Nguyen V.H., Strodiot JJ., Tossings P. (eds) Optimization. Lecture Notes in Economics and Mathematical Systems, Vol. 481 (2000) Springer, Berlin, Heidelberg.
6. H. Attouch, B.F. Svaiter, A continuous dynamical Newton-like approach to solving monotone inclusions, SIAM J. Control Optim. 49(2) (2011), 574-598
7. S. Banert, R.I. Boţ, A forward-backward-forward differential equation and its asymptotic properties. J. Convex Anal. 25(2) (2018) 371-388
8. H.H. Bauschke, P. L. Combettes, Convex analysis and monotone operator theory in hilbert spaces, CMS Books in Mathematics, Springer, (2011)

9. J. Bolte, Continuous gradient projection method in Hilbert spaces, J. Optim. Theory Appl., 119(2) (2003) 235-259
10. R.I. Boţ, E.R. Csetnek, A dynamical system associated with the fixed points set of a nonexpansive operator, J. Dynam. Differential Equations, 64(10) (2015) 155-168
11. R.I. Boţ, E.R. Csetnek, Approaching the solving of constrained variational inequalities via penalty term-based dynamical systems, J. Math. Anal. Appl., 435(2) (2016) 1688-1700
12. R.I. Boţ, E.R. Csetnek, Second order forward-backward dynamical systems for monotone inclusion problems, SIAM J. Control Optim., 54(3) (2016) 1423-1443
13. R.I Boţ, E.R. Csetnek, P.T. Vuong, The forward-backward-forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in Hilbert spaces, European J. Oper. Res. 287(1) (2020) 49-60
14. R.I. Boţ, C. Hendrich, Convergence analysis for a primal-dual monotone skew splitting algorithm with applications to total variation minimization, J. Math. Imaging Vision 49(3) (2014), 551–568
15. R.I. Boţ, S.M. Grad, D. Meier, M. Staudigl, Inducing strong convergence of trajectories in dynamical systems associated to monotone inclusions with composite structure, Adv. Nonlinear Anal. 10 (2021) 450-476
16. R.W. Cottle, J.C. Yao, Pseudo-monotone complementarity problems in Hilbert space, J. Optim. Theory Appl. 75 (1992) 281–295
17. H. Brézis, I.I. Chapitre, Operateurs maximaux monotones, North-Holland Math. Stud. 5 (1973) 19-51
18. T.N. Hai, Dynamical systems for solving variational inequalities, J. Dyn. Control Syst., (2021), DOI 10.1007/s10883-021-09531-8
19. D.V. Hieu, P.K. Anh, L.D. Muu, J.J. Strodiot, Iterative regularization methods with new stepsize rules for solving variational inclusions, J. Appl. Math. Comput. (2021), DOI 10.1007/s12190-021-01534-9
20. A. Haraux, Systémes dynamiques dissipatifs et applications, Recherches en Mathématiques Appliquées, Masson, (1991)
21. P.D. Khanh, P.T. Vuong, Modified projection method for strongly pseudomonotone variational inequalities, J. Global Optim. 58(2) (2014), 341-350
22. J. Peypouquet, S. Sorin, Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time, J. Convex Anal. 17 (2010) 1113-1163
23. H. Xu, Another control condition in an iterative method for nonexpansive mappings. Bull. Austral. Math. Soc. 65 (2002) 109–113