



Third Order Dynamical Systems for the Sum of Two Generalized Monotone Operators

Pham Viet Hai¹ · Phan Tu Vuong²

Received: 25 May 2023 / Accepted: 4 April 2024
© The Author(s) 2024

Abstract

In this paper, we propose and analyze a third-order dynamical system for finding zeros of the sum of two generalized operators in a Hilbert space \mathcal{H} . We establish the existence and uniqueness of the trajectories generated by the system under appropriate continuity conditions, and prove exponential convergence to the unique zero when the sum of the operators is strongly monotone. Additionally, we derive an explicit discretization of the dynamical system, which results in a forward–backward algorithm with double inertial effects and larger range of stepsize. We establish the linear convergence of the iterates to the unique solution using this algorithm. Furthermore, we provide convergence analysis for the class of strongly pseudo-monotone variational inequalities. We illustrate the effectiveness of our approach by applying it to structured optimization and pseudo-convex optimization problems.

Keywords Monotone inclusion · Dynamical system · Generalized monotonicity · Variational inequality · Exponential convergence · Linear convergence

Mathematics Subject Classification 47J20 · 49J40 · 90C30 · 90C52

Dedicate to Professor Pham Ky Anh (Vietnam National University) on the occasion of his 75th birthday.

Communicated by Boris S. Mordukhovich.

✉ Phan Tu Vuong
t.v.phan@soton.ac.uk

Pham Viet Hai
hai.phamviet@hust.edu.vn

¹ Faculty of Mathematics and Informatics, Hanoi University of Science and Technology, Khoa Toan-Tin, Dai hoc Bach khoa Hanoi, 1 Dai Co Viet, Hanoi, Vietnam

² School of Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, UK

1 Introduction

In practical applications, many nonlinear phenomena can be represented as finding a zero of a monotone operator. This problem arises in various contexts, such as solving variational inequalities related to monotone operators, minimizing convex functions, finding fixed points of nonexpansive mappings, and more. One of the most widely used methods for solving this problem is the proximal point algorithm, which was originally proposed by Martinet and systematically studied by Rockafellar [39] in the context of Hilbert spaces.

Another important problem is to find a zero of the sum of two maximally monotone operators

$$\text{find } x_* \in \mathcal{H} \text{ such that } 0 \in A(x_*) + B(x_*). \quad (1.1)$$

Problem (1.1) arises in a wide range of applications such as convex optimization, image processing, and signal processing. A crucial special case of Problem (1.1) is the following variational inequality (VI) problem

$$\text{find } x_* \in \mathcal{H} \text{ such that } 0 \in N_C(x_*) + B(x_*), \quad (1.2)$$

where C is a nonempty closed convex subset of \mathcal{H} and $N_C(x_*)$ is the normal cone of C at x_* . When B is single-valued, the VI problem (1.2) is equivalent to finding a point $x_* \in C$ such that

$$\langle B(x_*), y - x_* \rangle \geq 0 \quad \forall y \in C. \quad (1.3)$$

The Douglas-Rachford splitting algorithm [22], presented by Lions and Mercier [32], is a fundamental method to solve such problems. Under additional assumptions on the involved operators, linear rates of convergence for the algorithm are possible. Some other splitting methods are derived from the Douglas-Rachford algorithm (such as the primal-dual hybrid gradient method [35], Alternating Direction Method of Multiplier (ADMM) [23], and Spingarn's method of partial inverses [23]). There are many other methods for solving Problem (1.1), especially when one of the operators is single-valued. A popular method for solving this problem is the forward-backward algorithm, which consists of a forward step with one operator and a backward step with another. The algorithm generates a sequence of iterates that converges to a solution under suitable assumptions on the operators. The forward-backward algorithm has been widely studied and applied in both finite-dimensional and infinite-dimensional settings [11, 24, 33, 38].

Nowadays, there is a growing interest in connecting and integrating optimization with other fields. This research direction has become increasingly attractive as it can provide new insights into optimization results and lead to interesting findings. Among the emerging research directions, there is a line of works that uses ordinary differential equations (ODEs) to design algorithms for optimization problems [2, 3, 9, 14], variational inequalities [18, 27, 34, 43], monotone inclusions [1, 5, 6], fixed point problems [15, 17] and equilibrium problems [20, 36, 42, 45]. Using ODE interpretation not

only provides a better understanding of Nesterov’s scheme, but also helps design new schemes with similar convergence rates. The readers can refer to [9, 10, 18, 40] and references therein for more examples.

1.1 Some Historical Aspects

The Heavy Ball with Friction method is a popular optimization algorithm based on inertial dynamics. The algorithm was proposed by Polyak to accelerate the gradient method in optimization [37]. It introduces an inertial system with a fixed viscous damping coefficient

$$x^{(2)}(t) + \gamma x^{(1)}(t) + \nabla f(x(t)) = 0, \tag{1.4}$$

for minimizing a convex and differentiable function f . Note that, when f has a Lipschitz continuous gradient then ∇f is a co-coercive operator (see definition in Sect. 2). Attouch and Alvarez extended the heavy ball dynamical system (1.4) for constrained optimization as well as co-coercive operator in [4]. Recently, Boğ and Csetnek [15] studied the second order dynamical system with variable viscous damping coefficient

$$x^{(2)}(t) + \gamma(t)x^{(1)}(t) + \lambda(t)B(x(t)) = 0,$$

for finding a zero of a co-coercive operator B . The results were applied to second order forward–backward dynamical systems for monotone inclusion problems (1.1)

$$x^{(2)}(t) + \gamma(t)x^{(1)}(t) + \lambda(t)[(x(t)) - \mathcal{J}_A(x(t) - \eta B(x(t)))] = 0, \tag{1.5}$$

where A is maximal monotone and B is co-coercive. Here $\mathcal{J}_A \triangleq (I + A)^{-1}$ is the *resolvent* of an operator A with I stands for the identity operator. When the operator is only merely monotone but not co-coercive, a second order forward–backward–forward dynamical system and its discretization have been recently proposed and investigated in [19]. In particular, when the operator $A + B$ is strongly monotone, the exponential convergence rate of the second order dynamical system (1.5) was obtained in [16]. Under suitable conditions on parameters, the authors established the convergence rate of $O(e^{-t})$ for the trajectories.

Attouch, Chbani and Riahi are the first authors who studied third order dynamical system for minimizing a convex and differentiable function in Hilbert spaces [7, 8]. They proposed and studied the (TOGES) dynamical system [7]

$$x^{(3)}(t) + \frac{\alpha}{t}x^{(2)}(t) + \frac{2\alpha - 6}{t^2}x^{(1)}(t) + \nabla f(x(t) + tx^{(1)}(t)) = 0. \tag{1.6}$$

Using the temporal scaling techniques, the third order dynamical system (1.6) was reformulated as a second order dynamical system and the convergence analysis was obtained using Lyapunov’s energy function techniques developed for second order dynamical system. The authors showed a convergence rate of the values of the order $\frac{1}{t^3}$, i.e. $f(x(t) + tx^{(1)}(t)) - \inf_{\mathcal{H}} f \leq \frac{C}{t^3}$ for some constant $C > 0$ and obtained the

convergence of the trajectories towards optimal solutions of $\min_{x \in \mathcal{H}} f(x)$. When the objective function f is strongly convex, the authors established an exponential rate of convergence. Proximal-based algorithms obtained by temporal discretization of (TOGES) was also investigated. Nevertheless, the rate of values of $f(x(t))$ in (TOGES) is only of order $\frac{1}{t}$, i.e. $f(x(t)) - \inf_{\mathcal{H}} f \leq \frac{C}{t}$, which is not completely satisfactory from the point of view of fast optimization. Hence, very recently, an improved version of (TOGES), called (TOGES-V) has been proposed and investigated by the same authors in [8]

$$x^{(3)}(t) + \frac{\alpha + 7}{t}x^{(2)}(t) + \frac{5(\alpha + 1)}{t^2}x^{(1)}(t) + \nabla f \left(x(t) + \frac{1}{4}tx^{(1)}(t) \right) = 0,$$

where they obtained the rate $\mathcal{O}\left(\frac{1}{t^3}\right)$ for $f(x(t)) - \inf_{\mathcal{H}} f$.

1.2 Our Contributions

In this paper, we propose for the first time a third order dynamical system for the monotone inclusion (1.1) and investigate its convergence properties in both continuous time and discrete time settings. The motivation of considering third (or higher) order dynamical system comes from the fact that it can potentially provide faster convergence rate, as seen in optimization problems [7, 8]. This will be also the case of monotone inclusion obtained in this paper. Indeed, we derive the convergence rate of $\mathcal{O}(e^{-\varepsilon t})$ for some $\varepsilon > 1$ (in particular for $\varepsilon = 2$) under suitable choices of parameters, which is significantly faster than the classical results obtained in [16] for second order dynamical systems. In discrete setting, the third order dynamical system provides a new forward backward algorithm with double momentum and a larger range of stepsize.

In contrast to the classical monotone inclusion problem, where each individual operator A and B is required to be (maximally) monotone, we only require A and B to be generalized monotone (see Definitions in Sect. 2). This approach allows us to handle not only the classical monotone inclusion problem but also the problem of finding zeros of the sum of a weakly monotone operator and a strongly monotone operator, and the pseudo-monotone variational inequalities. Applications of these models include minimizing the sum of a weakly convex function and a strongly convex function [21, 26] or minimizing a pseudo-convex function. The convergence analysis developed in this paper is purely relied on Lyapunov's energy function techniques, in contrast to the temporal scaling technique using in [7, 8]. In summary, our contributions are as follows:

- Propose a third order dynamical system for the sum of two generalized monotone operators.
- Establish the existence and uniqueness of the trajectories generated by the proposed dynamical system.
- Provide the exponential convergence analysis of the trajectories to the unique solution of the inclusion, and show that it is faster than classical results.

- Investigate the temporal discretization of the system and prove the linear convergence of the corresponding forward–backward algorithm with double inertial effects.
- Study the third order dynamical system for strongly pseudo-monotone variational inequalities.

The paper is structured as follows. In Sect. 2, we introduce some terminologies and results that are necessary for the analysis presented in the subsequent sections. In Sects. 3 and 4, we focus on solving Problem (1.1) under the assumption of generalized monotonicity of the operators involved. In Sect. 3, we propose a third-order dynamical system and establish its exponential convergence to the zero of Problem (1.1). The explicit discretization of this system leads to a new forward backward algorithm studied in Sect. 4. In Sect. 5, motivated by the third-order dynamical system, we find the solution of Problem (1.2) under the assumption of strong pseudo-monotonicity of the operator B .

2 Preliminaries

We start the section with listing the notations used. The set of integers is denoted by \mathbb{Z} and the set of real numbers is denoted by \mathbb{R} . Let $\mathbb{Z}_{\geq 1} = \{j \in \mathbb{Z} : j \geq 1\}$ and $\mathbb{R}_{\geq 0} = \{t \in \mathbb{R} : t \geq 0\}$. The symbol $g^{(k)}$ stands for the k -th derivative of the function g .

Throughout this work \mathcal{H} is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. We use the notation $A : \mathcal{H} \rightrightarrows \mathcal{H}$ to indicate that A is a set-valued operator defined on \mathcal{H} , and $A : \mathcal{H} \rightarrow \mathcal{H}$ to indicate that A is a single-valued operator on \mathcal{H} .

Let A be an operator on \mathcal{H} . The *graph* of A is $\text{Gra}A = \{(x, u) \in \mathcal{H} \times \mathcal{H}, u \in A(x)\}$. The *inverse* of A , denoted by A^{-1} , is the operator with graph $\text{Gra}A^{-1} = \{(u, x) \in \mathcal{H} \times \mathcal{H}, u \in A(x)\}$.

2.1 Generalized Monotone Operators

We first recall some generalized versions of monotone operator defined and studied in [21, 29].

Definition 2.1 The operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is called γ_A -monotone if there exists a scalar $\gamma_A \in \mathbb{R}$ such that

$$\langle u - v, x - y \rangle \geq \gamma_A \|x - y\|^2 \quad \forall (x, u), (y, v) \in \text{Gra}A.$$

The constant γ_A is referred to the monotonicity modulus of A . We also say that A is *maximally* γ_A -monotone if it is γ_A -monotone and there is no γ_A -monotone operator whose graph strictly contains $\text{Gra}A$.

Remark 2.2 Note that in the definition of *generalized monotonicity*, γ_A can be negative. If $\gamma_A = 0$, then the generalized monotonicity reduces to the classical monotonicity.

If $\gamma_A > 0$, then A is *strongly monotone*. Finally, if $\gamma_A < 0$ then A is called *weakly-monotone*. For more detailed discussion on (maximally) monotone operators and the connection to optimization problems, we refer the readers to [12, 13, 21].

Definition 2.3 The single-valued operator $T : \mathcal{H} \rightarrow \mathcal{H}$ is called

1. γ_T -strongly pseudo-monotone if $\gamma_T > 0$ and

$$\langle T(x), y - x \rangle \geq 0 \implies \langle T(y), y - x \rangle \geq \gamma_T \|x - y\|^2$$

for all $x, y \in \mathcal{H}$.

2. γ_T -cocoercive if $\gamma_T > 0$ and

$$\langle T(x) - T(y), x - y \rangle \geq \gamma_T \|T(x) - T(y)\|^2 \quad \forall x, y \in \mathcal{H}.$$

3. L_T -Lipschitz continuous if $L_T > 0$ and

$$\|T(x) - T(y)\| \leq L_T \|x - y\| \quad \forall x, y \in \mathcal{H}.$$

Remark 2.4 It is clear from the Cauchy–Schwartz inequality that if T is γ_T -co-coercive then it is $1/\gamma_T$ -Lipschitz continuous.

The *resolvent* of an operator A is denoted as $\mathcal{J}_A \triangleq (I + A)^{-1}$, where I is the *identity operator*. We will need the following properties of resolvent operator.

Lemma 2.5 [21] Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be an γ_A -monotone operator and let $\omega > 0$ such that $1 + \omega\gamma_A > 0$. Then the followings hold:

1. $\mathcal{J}_{\omega A}$ is single-valued;
2. $\mathcal{J}_{\omega A}$ is $(1 + \omega\gamma_A)$ -co-coercive;
3. $\text{dom } \mathcal{J}_{\omega A} = \mathcal{H}$ if and only if A is maximally γ_A -monotone.

2.2 Absolutely Continuous Functions

Definition 2.6 A function $h : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$ is called *locally absolutely continuous* if it is absolutely continuous on every compact interval, which means that for each interval $[t_0, t_1]$ there exists an integrable function $g : [t_0, t_1] \rightarrow \mathbb{R}^d$ such that

$$h(t) = h(t_0) + \int_{t_0}^t g(s) ds \quad \forall t \in [t_0, t_1].$$

Remark 2.7 If $h : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$ is a locally absolutely continuous function, then it is differentiable almost everywhere and its derivative agrees with its distributional derivative almost everywhere.

Proposition 2.8 For $s, u \geq 0$ and $m \in \mathbb{Z}_{\geq 1}$, it holds

$$\int_u^s e^{t\varepsilon} g^{(m)}(t) dt = e^{s\varepsilon} \left(\sum_{j=0}^{m-1} (-\varepsilon)^{m-1-j} g^{(j)}(s) \right) + (-\varepsilon)^m \int_u^s e^{t\varepsilon} g(t) dt - e^{u\varepsilon} \left(\sum_{j=0}^{m-1} (-\varepsilon)^{m-1-j} g^{(j)}(u) \right).$$

Proof The case when $m = 1$ is done by using integration by parts. Now we suppose that the conclusion holds for m and prove the case $m + 1$. Indeed, we have

$$\int_u^s e^{t\varepsilon} g^{(m+1)}(t) dt = \int_u^s e^{t\varepsilon} dg^{(m)}(t) = e^{t\varepsilon} g^{(m)}(t) \Big|_u^s - \varepsilon \int_u^s e^{t\varepsilon} g^{(m)}(t) dt,$$

which together with the induction assumption completes the proof. □

2.3 A Third Order Dynamical System

In this paper, we propose the following dynamical system for Problem (1.1).

$$y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t) + \alpha_0 [y(t) - \mathcal{J}_{\omega A}(y(t) - \omega B(y(t)))] = 0, \tag{2.1}$$

where $\alpha_2, \alpha_1, \alpha_0, \omega > 0$ and $y^{(j)}(t_0) = v_j, j \in \{0, 1, 2\}$.

The solution of dynamical system (2.1) is understood in the following sense.

Definition 2.9 A function $y(\cdot)$ is called a *strong global solution* of Eq. (2.1) if it holds:

1. For every $j \in \{0, 1, 2, 3\}$, $y^{(j)} : [t_0, +\infty) \rightarrow \mathcal{H}$ is locally absolutely continuous; in other words, absolutely continuous on each interval $[\delta, \eta]$ for $\eta > \delta > t_0$.
2. $y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t) + \alpha_0 [y(t) - \mathcal{J}_{\omega A}(y(t) - \omega B(y(t)))] = 0$ for almost every $t \geq t_0$.
3. $y^{(j)}(t_0) = v_j, j \in \{0, 1, 2\}$.

Proposition 2.10 (Equivalent form) Equation (2.1) is equivalent to the system $x^{(1)}(t) = G(x(t))$, where $G : \mathcal{H} \times \mathcal{H} \times \mathcal{H} \rightarrow \mathcal{H} \times \mathcal{H} \times \mathcal{H}$ is defined by

$$G(x_1, x_2, x_3) = (x_2, x_3, -\alpha_1 x_2 - \alpha_2 x_3 - \alpha_0 [x_1 - \mathcal{J}_{\omega A}(x_1 - \omega B(x_1))]),$$

where $(x_1, x_2, x_3) \in \mathcal{H} \times \mathcal{H} \times \mathcal{H}$.

Proof The conclusion follows from doing the change of variables

$$(x_1(t), x_2(t), x_3(t)) = (y(t), y^{(1)}(t), y^{(2)}(t)).$$

□

Theorem 2.11 (Existence and uniqueness of a solution) *Consider dynamical system (2.1), where $\alpha_0, \alpha_1, \alpha_2, \omega > 0$ and the operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is γ_A -maximally monotone, $B : \mathcal{H} \rightarrow \mathcal{H}$ is γ_B -monotone and L -Lipschitz such that $1 + \omega\gamma_A > 0$. Then for each $v_0, v_1, v_2 \in \mathcal{H}$ there exists a unique strong global solution of (2.1).*

Proof We endow $\mathcal{H} \times \mathcal{H} \times \mathcal{H}$ with scalar product

$$\langle (y_1, y_2, y_3), (z_1, z_2, z_3) \rangle_{\mathcal{H} \times \mathcal{H} \times \mathcal{H}} = \langle y_1, z_1 \rangle + \langle y_2, z_2 \rangle + \langle y_3, z_3 \rangle.$$

We show that the operator G is Lipschitz. Indeed, let $y = (y_1, y_2, y_3), z = (z_1, z_2, z_3) \in \mathcal{H} \times \mathcal{H} \times \mathcal{H}$. We have

$$\begin{aligned} & \| \mathcal{J}_{\omega A}(y_1 - \omega B(y_1)) - \mathcal{J}_{\omega A}(z_1 - \omega B(z_1)) \|^2 \\ & \leq \frac{1}{(1 + \omega\gamma_A)^2} \|y_1 - z_1 - \omega(B(y_1) - B(z_1))\|^2 \\ & \leq \frac{(1 + \omega L)^2}{(1 + \omega\gamma_A)^2} \|y_1 - z_1\|^2, \end{aligned}$$

and so

$$\begin{aligned} \|G(y) - G(z)\|_{\mathcal{H} \times \mathcal{H} \times \mathcal{H}}^2 & \leq \|y_2 - z_2\|^2 + \|y_3 - z_3\|^2 \\ & \quad + (\alpha_1^2 + \alpha_2^2 + 2\alpha_0^2) \left[\|y_2 - z_2\|^2 + \|y_3 - z_3\|^2 + \|y_1 - z_1\|^2 \right] \\ & \quad + \| \mathcal{J}_{\omega A}(y_1 - \omega B(y_1)) - \mathcal{J}_{\omega A}(z_1 - \omega B(z_1)) \|^2 \\ & \leq \|y_2 - z_2\|^2 + \|y_3 - z_3\|^2 \\ & \quad + (\alpha_1^2 + \alpha_2^2 + 2\alpha_0^2) \left[1 + \frac{(1 + \omega L)^2}{(1 + \omega\gamma_A)^2} \right] \|y - z\|_{\mathcal{H} \times \mathcal{H} \times \mathcal{H}}^2 \\ & \leq \left[1 + (\alpha_1^2 + \alpha_2^2 + 2\alpha_0^2) \left(1 + \frac{(1 + \omega L)^2}{(1 + \omega\gamma_A)^2} \right) \right] \cdot \|y - z\|_{\mathcal{H} \times \mathcal{H} \times \mathcal{H}}^2. \end{aligned}$$

By using the Cauchy–Picard theorem (see, for example, [28, Proposition 6.2.1]), we get the existence and uniqueness of a strong global solution. □

2.4 Difference Operators

In the section, we give the discrete counterpart of the dynamical system (2.1). To that aim, we recall the operation of forward difference and its properties used in the convergence analysis. For $z : \mathbb{Z} \rightarrow \mathcal{H}$ and $\kappa \in \mathbb{Z}_{\geq 1}$, we denote

$$z^{\Delta(\kappa+1)} \triangleq (z^{\Delta(\kappa)})^\Delta, \quad \text{where } z^\Delta(n) \triangleq z(n + 1) - z(n).$$

Remark 2.12 Let $f, g, h : \mathbb{Z} \rightarrow \mathcal{H}$ and $\theta \in \mathbb{R}$. It can be proven that

$$\langle h, g \rangle^\Delta(n) = \langle h^\Delta(n), g^\Delta(n) \rangle + \langle h^\Delta(n), g(n) \rangle + \langle h(n), g^\Delta(n) \rangle,$$

and consequently

$$\begin{aligned} \theta^{n+1}g^\Delta(n) &= (\theta^n g)^\Delta(n) + (1 - \theta)\theta^n g(n), \\ (\|f\|^2)^\Delta(n) &= \|f^\Delta(n)\|^2 + 2\langle f^\Delta(n), f(n) \rangle. \end{aligned}$$

Consider the difference equation, which is the discrete version of (2.1):

$$z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^\Delta(n) + \alpha_0 [z(n) - \mathcal{J}_{\omega A}(z(n) - \omega B(z(n)))] = 0, \tag{2.2}$$

where $\alpha_2, \alpha_1, \alpha_0, \omega > 0$.

Proposition 2.13 (Equivalent form) *Equation (2.2) has an equivalent form*

$$\begin{aligned} z(n + 3) &= (3 - \alpha_2)z(n + 2) + (2\alpha_2 - \alpha_1 - 3)z(n + 1) \\ &+ (\alpha_1 + 1 - \alpha_2)z(n) - \alpha_0 [z(n) - \mathcal{J}_{\omega A}(z(n) - \omega B(z(n)))]. \end{aligned} \tag{2.3}$$

Proof The proof makes use of the facts that

$$\begin{aligned} z^{\Delta(2)}(n) &= z(n + 2) - 2z(n + 1) + z(n), \\ z^{\Delta(3)}(n) &= z(n + 3) - 3z(n + 2) + 3z(n + 1) - z(n). \end{aligned}$$

□

Remark 2.14 The numerical scheme (2.3) can be re-written as

$$\begin{aligned} z(n + 3) &= z(n + 2) + (2 - \alpha_2)(z(n + 2) - z(n + 1)) \\ &+ (\alpha_2 - \alpha_1 - 1)(z(n + 1) - z(n)) \\ &- \alpha_0 [z(n) - \mathcal{J}_{\omega A}(z(n) - \omega B(z(n)))], \end{aligned} \tag{2.4}$$

which is a forward–backward algorithm with double momentum.

3 Continuous Time Dynamical System

In this section, we will establish the exponential convergence of dynamical system (2.1) under the following assumption and notations.

Assumption 3.1 (i) The coefficients $\alpha_0, \alpha_1, \alpha_2 > 0$.

(ii) The operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally γ_A -monotone, $B : \mathcal{H} \rightarrow \mathcal{H}$ is γ_B -monotone and L -Lipschitz continuous such that

$$\gamma \triangleq \gamma_A + \gamma_B > 0. \tag{3.1}$$

(iii) The parameter $\omega > 0$ satisfies

$$1 + \omega\gamma_A > 0, \tag{3.2}$$

$$\frac{1}{\omega} > \frac{L^2}{4\gamma} + L - \gamma. \tag{3.3}$$

Remark 3.1 Condition (3.1) implies that the sum operator $A + B$ is strongly monotone but not the individual operator A, B . A similar condition was studied in [21]. An direct application of this model is to minimize a sum of a weakly convex function and a strongly convex function. Condition (3.2) is imposed to ensure that the resolvent operator $\mathcal{J}_{\omega A}$ is single valued. Finally, condition (3.3) means that the stepsize ω must be bounded from above. Note that condition (3.3) gives

$$1 > \frac{L^2}{4\gamma} \cdot \frac{1}{\frac{1}{\omega} + \gamma - L}.$$

Hence, we can find $\theta > 0$ such that

$$1 > \theta > \frac{L^2}{4\gamma} \cdot \frac{1}{\frac{1}{\omega} + \gamma - L}. \tag{3.4}$$

The following notations are used.

$$\ell \triangleq \frac{2\omega}{2\omega\gamma + 1} \left(\frac{1}{\omega} + \gamma - L - \frac{L^2}{4\theta\gamma} \right), \quad \delta \triangleq \frac{2\omega\gamma(1 - \theta)}{2\omega\gamma + 1}. \tag{3.5}$$

3.1 Global Exponential Convergence

First, we consider the dynamical system (2.1) whose the global convergence relates to the following parameters

$$\begin{cases} A_2 \triangleq \frac{\ell\alpha_1}{\alpha_0}, \\ A_1 \triangleq \frac{\ell\alpha_2\alpha_1}{\alpha_0} - 3, \\ A_0 \triangleq \frac{\ell\alpha_1^2}{\alpha_0} - 2\alpha_2, \end{cases} \quad \begin{cases} B_1 \triangleq \frac{\ell\alpha_2}{\alpha_0}, \\ B_0 \triangleq \frac{\ell}{\alpha_0}(\alpha_2^2 - 2\alpha_1), \end{cases} \quad C_0 \triangleq \frac{\ell}{\alpha_0}. \tag{3.6}$$

We denote the functions

$$a(t) \triangleq \|y(t) - x_*\|^2, \quad b_k(t) \triangleq \|y^{(k)}(t)\|^2. \tag{3.7}$$

Theorem 3.2 *Suppose that the operators A and B satisfy Assumption 3.1. Let x_* be the unique solution of Problem (1.1). Let θ satisfy (3.4) and denote the parameters as*

in (3.5)–(3.6). Assume that there exists $\varepsilon > 0$ such that the following conditions hold

$$-\varepsilon^3 + \alpha_2\varepsilon^2 - \alpha_1\varepsilon + \delta\alpha_0 \geq 0, \tag{3.8}$$

$$A_2\varepsilon^2 - A_1\varepsilon + A_0 \geq 0, \tag{3.9}$$

$$3\varepsilon^2 - 2\alpha_2\varepsilon + \alpha_1 \geq 0, \tag{3.10}$$

$$-2A_2\varepsilon + A_1 \geq 0, \tag{3.11}$$

$$-B_1\varepsilon + B_0 \geq 0, \tag{3.12}$$

$$\alpha_2 > 2\varepsilon. \tag{3.13}$$

Then the trajectories $y(\cdot)$ generated by dynamical system (2.1) converges exponentially to x_* , i.e., there exist positive numbers μ, η such that

$$\|y(t) - x^*\| \leq \mu \|y(t_0) - x^*\| e^{-\eta t} \quad \forall t \geq t_0.$$

Proof In the next arguments, we often use the identities:

$$b_1^{(1)}(t) = 2 \left\langle y^{(2)}(t), y^{(1)}(t) \right\rangle,$$

$$b_1^{(2)}(t) = 2 \left\langle y^{(3)}(t), y^{(1)}(t) \right\rangle + 2\|y^{(2)}(t)\|^2 = 2 \left\langle y^{(3)}(t), y^{(1)}(t) \right\rangle + 2b_2(t).$$

Since

$$\begin{aligned} a^{(1)}(t) &= 2 \left\langle y^{(1)}(t), y(t) - x_* \right\rangle, \\ a^{(2)}(t) &= 2 \left\langle y^{(2)}(t), y(t) - x_* \right\rangle + 2b_1(t), \\ a^{(3)}(t) &= 2 \left\langle y^{(3)}(t), y(t) - x_* \right\rangle + 3b_1^{(1)}(t), \end{aligned}$$

we have

$$\begin{aligned} &2 \left\langle y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t), y(t) - x_* \right\rangle \\ &= a^{(3)}(t) + \alpha_2 a^{(2)}(t) + \alpha_1 a^{(1)}(t) - 3b_1^{(1)}(t) - 2\alpha_2 b_1(t). \end{aligned} \tag{3.14}$$

We observe

$$\begin{aligned} &\|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\|^2 \\ &= \alpha_1 b_1^{(2)}(t) + \alpha_2 \alpha_1 b_1^{(1)}(t) + \alpha_1^2 b_1(t) + \alpha_2 b_2^{(1)}(t) + (\alpha_2^2 - 2\alpha_1) b_2(t) + b_3(t). \end{aligned} \tag{3.15}$$

Using the definition of resolvent, equation (2.1) gives the following

$$B\left(\frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t)\right) - B(y(t)) \\ - \frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\omega \alpha_0} \\ \in (A + B)\left(\frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t)\right),$$

which combined with $0 \in (A + B)(x_*)$ and the γ -monotonicity of $A + B$ implies

$$\gamma \left\| \frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t) - x_* \right\|^2 \\ \leq \left\langle B\left(\frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t)\right) \right. \\ \left. - B(y(t)) - \frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\omega \alpha_0}, \right. \\ \left. \frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t) - x_* \right\rangle.$$

Since the operator B is L -Lipschitz, we can estimate the right hand side of the inequality above and then

$$\gamma \left\| \frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t) - x_* \right\|^2 \\ \leq \frac{1}{\alpha_0^2} \left(L - \frac{1}{\omega} \right) \|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\|^2 \\ + \left\langle B\left(\frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t)\right) - B(y(t)), y(t) - x_* \right\rangle \\ - \frac{1}{\omega \alpha_0} \left\langle y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t), y(t) - x_* \right\rangle.$$

Note that by the Cauchy–Schwarz inequality

$$\begin{aligned} & \left\langle B \left(\frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t) \right) - B(y(t)), y(t) - x_* \right\rangle \\ & \leq \frac{L}{\alpha_0} \|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\| \cdot \|y(t) - x_*\| \\ & \leq \frac{L^2}{4\alpha_0^2 \theta \gamma} \|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\|^2 + \theta \gamma a(t). \end{aligned}$$

Thus, we get

$$\begin{aligned} & \gamma \left\| \frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t) - x_* \right\|^2 \\ & \leq \frac{1}{\alpha_0^2} \left(L - \frac{1}{\omega} + \frac{L^2}{4\theta \gamma} \right) \|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\|^2 \\ & \quad - \frac{1}{\omega \alpha_0} \left\langle y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t), y(t) - x_* \right\rangle + \theta \gamma a(t). \quad (3.16) \end{aligned}$$

Note that

$$\begin{aligned} & \gamma \left\| \frac{y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)}{\alpha_0} + y(t) - x_* \right\|^2 \\ & = \frac{\gamma}{\alpha_0^2} \|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\|^2 + \gamma a(t) \\ & \quad + \frac{2\gamma}{\alpha_0} \left\langle y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t), y(t) - x_* \right\rangle. \end{aligned}$$

Inserting the equality above into (3.16), we obtain

$$\begin{aligned} & \frac{\ell}{\alpha_0} \|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\|^2 + \delta \alpha_0 a(t) \\ & \quad + 2 \left\langle y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t), y(t) - x_* \right\rangle \leq 0, \quad (3.17) \end{aligned}$$

which implies, by (3.14) and (3.15), that

$$\begin{aligned} & a^{(3)}(t) + \alpha_2 a^{(2)}(t) + \alpha_1 a^{(1)}(t) + \delta \alpha_0 a(t) \\ & \quad + A_2 b_1^{(2)}(t) + A_1 b_1^{(1)}(t) + A_0 b_1(t) + B_1 b_2^{(1)}(t) + B_0 b_2(t) + C_0 b_3(t) \leq 0. \end{aligned}$$

By (3.4), we have $\ell > 0$ and so is C_0 . Thus, we can write

$$\begin{aligned} & a^{(3)}(t) + \alpha_2 a^{(2)}(t) + \alpha_1 a^{(1)}(t) + \delta \alpha_0 a(t) \\ & \quad + A_2 b_1^{(2)}(t) + A_1 b_1^{(1)}(t) + A_0 b_1(t) + B_1 b_2^{(1)}(t) + B_0 b_2(t) \leq 0. \end{aligned}$$

Multiplying both sides by $e^{\varepsilon(t-t_0)}$ and then using Proposition 2.8, we get

$$\begin{aligned}
 & e^{\varepsilon(s-t_0)}[(\varepsilon^2 - \alpha_2\varepsilon + \alpha_1)a(s) + (\alpha_2 - \varepsilon)a^{(1)}(s) + a^{(2)}(s)] \\
 & + (-\varepsilon^3 + \alpha_2\varepsilon^2 - \alpha_1\varepsilon + \delta\alpha_0) \int_{t_0}^s e^{\varepsilon(t-t_0)}a(t)dt \\
 & + e^{\varepsilon(s-t_0)}[(-A_2\varepsilon + A_1)b_1(s) + A_2b_1^{(1)}(s)] \\
 & + (A_2\varepsilon^2 - A_1\varepsilon + A_0) \int_{t_0}^s e^{\varepsilon(t-t_0)}b_1(t)dt \\
 & + B_1e^{\varepsilon(s-t_0)}b_2(s) + (-B_1\varepsilon + B_0) \int_{t_0}^s e^{\varepsilon(t-t_0)}b_2(t)dt \leq D_1,
 \end{aligned}$$

for some constant D_1 , which implies, after using (3.8), (3.9), (3.12) and (3.4), that

$$\begin{aligned}
 & e^{\varepsilon(s-t_0)}[(\varepsilon^2 - \alpha_2\varepsilon + \alpha_1)a(s) + (\alpha_2 - \varepsilon)a^{(1)}(s) + a^{(2)}(s)] \\
 & + e^{\varepsilon(s-t_0)}[(-A_2\varepsilon + A_1)b_1(s) + A_2b_1^{(1)}(s)] + B_1e^{\varepsilon(s-t_0)}b_2(s) \leq D_1,
 \end{aligned}$$

Integrating the above inequality with respect to the variable $s \in [t_0; t]$ we deduce

$$\begin{aligned}
 & e^{\varepsilon(t-t_0)}a^{(1)}(t) + (\alpha_2 - 2\varepsilon)e^{\varepsilon(t-t_0)}a(t) + (3\varepsilon^2 - 2\alpha_2\varepsilon + \alpha_1) \int_{t_0}^t e^{\varepsilon(s-t_0)}a(s)ds \\
 & + A_2e^{\varepsilon(t-t_0)}b_1(t) + (-2A_2\varepsilon + A_1) \int_{t_0}^t e^{\varepsilon(s-t_0)}b_1(s)ds \leq D_1t + D_2.
 \end{aligned}$$

for some constant D_2 . Using (3.10), (3.11), (3.4), we get

$$e^{\varepsilon(t-t_0)}a^{(1)}(t) + (\alpha_2 - 2\varepsilon)e^{\varepsilon(t-t_0)}a(t) \leq D_1t + D_2. \tag{3.18}$$

Note that equation (3.18) reduces to the following

$$\begin{aligned}
 a(t) & \leq e^{-(\alpha_2-2\varepsilon)(t-t_0)}D_3 \\
 & + e^{-(\alpha_2-2\varepsilon)(t-t_0)} \int_{t_0}^t e^{(\alpha_2-3\varepsilon)(s-t_0)}(D_1s + D_2) ds.
 \end{aligned}$$

for some constant D_3 .

– If $\alpha_2 \geq 3\varepsilon$, then $e^{(\alpha_2-3\varepsilon)(s-t_0)} \leq e^{(\alpha_2-3\varepsilon)(t-t_0)}$, and so

$$a(t) \leq e^{-(\alpha_2-2\varepsilon)(t-t_0)} D_3 + e^{-\varepsilon(t-t_0)} \int_{t_0}^t (D_1 s + D_2) ds. \tag{3.19}$$

– If $2\varepsilon < \alpha_2 < 3\varepsilon$, then $e^{(\alpha_2-3\varepsilon)(s-t_0)} \leq 1$, and so

$$a(t) \leq e^{-(\alpha_2-2\varepsilon)(t-t_0)} \left(D_3 + \int_{t_0}^t (D_1 s + D_2) ds \right). \tag{3.20}$$

The arguments above show that $y(\cdot)$ converges exponentially to x_* . □

Remark 3.3 It follows from (3.19) that $a(t)$ converges to 0 with the rate of $O((Pt^2 + Qt + Rt)e^{-\varepsilon t})$ for some constants P, Q, R , while the rate obtained from (3.20) is $O((Pt^2 + Qt + Rt)e^{-(\alpha_2-2\varepsilon)t})$. With the suitable choice of ε and α_2 , these rates can be controlled so that they are faster than the rate $O(e^{-t})$ of the second order dynamical systems established in [16].

3.2 Parameters Choices

We now discuss the question “how to find ε ?”. It can be seen from (3.19) that the larger ε implies the faster rate. Finding the maximal value of ε is cumbersome as it depends on many other parameters. However, we will discuss how to find a "good enough" ε in this section. The following remark offers a way which concerns the coefficients.

Remark 3.4 If A_0, A_1, B_0 satisfy

$$A_0, A_1, B_0 > 0, \tag{3.21}$$

then conditions (3.8)–(3.13) can be obtained by letting $\varepsilon \rightarrow 0^+$.

In the following result, we simplify the assumption (3.21) in algebraic terms of the coefficients $\alpha_0, \alpha_1, \alpha_2$.

Corollary 3.5 Consider equation (2.1), under Assumption 3.1. Let x_* be the unique solution of Problem (1.1). Let θ satisfy (3.4). Denote (3.5)–(3.6). Then $y(\cdot)$ converges exponentially to x_* provided that coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy the following conditions

$$\alpha_1 < \frac{\alpha_2^2}{2}, \tag{3.22}$$

$$\alpha_0 < \ell \cdot \min \left\{ \frac{\alpha_1 \alpha_2}{3}, \frac{\alpha_1^2}{2\alpha_2} \right\}. \tag{3.23}$$

Let us first examine Theorem 3.2 when $\varepsilon = 1$, for which it matches the rate obtained for the second order dynamical system established in [16].

Theorem 3.6 *Suppose that the operators A and B satisfy Assumption 3.1. Let x_* be the unique solution of Problem (1.1). Let θ satisfy (3.4) and denote the parameters as in (3.5)–(3.6) and*

$$\varphi \triangleq \frac{1}{\ell\delta}. \tag{3.24}$$

Then $y(\cdot)$ converges exponentially to x_* provided that coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy

$$\alpha_2 > \max\{3, 3\varphi + 2, 4\varphi\}, \tag{3.25}$$

$$\underline{\beta} \triangleq \max\{2\alpha_2 - 3, \varphi(2\alpha_2 - 3)\} < \alpha_1 < \bar{\beta} \triangleq 0.5\alpha_2(\alpha_2 - 1), \tag{3.26}$$

$$\underline{q} \triangleq \frac{\alpha_1 - \alpha_2 + 1}{\delta} < \alpha_0 < \bar{p} \triangleq \ell \cdot \min\left\{\frac{\alpha_1(\alpha_2 - 2)}{3}, \frac{\alpha_1(\alpha_1 - \alpha_2 + 1)}{2\alpha_2 - 3}\right\}. \tag{3.27}$$

Proof First, we show that (3.25) ensures the validity of (3.26); that is $\underline{\beta} < \bar{\beta}$. Indeed, it follows from (3.25) that $\alpha_2 > 3$ and so $2\alpha_2 - 3 < \frac{\alpha_2(\alpha_2 - 1)}{2}$. Also from (3.25), we have $\alpha_2 > 4\varphi$ and then

$$\frac{\alpha_2(\alpha_2 - 1)}{2} > \frac{\alpha_2(2\alpha_2 - 3)}{4} > \varphi(2\alpha_2 - 3).$$

Next, we show that (3.25)–(3.26) ensure the validity of (3.27); that is $\underline{q} < \bar{p}$. It results from (3.25) that $\alpha_2 > 3\varphi + 2$ and so

$$\frac{\alpha_1 - \alpha_2 + 1}{\delta} < \frac{\alpha_1}{\delta} < \ell \cdot \frac{\alpha_1(\alpha_2 - 2)}{3}.$$

Meanwhile, by (3.26), we have $\alpha_1 > \varphi(2\alpha_2 - 3)$, which gives

$$\frac{\alpha_1 - \alpha_2 + 1}{\delta} < \ell \cdot \frac{\alpha_1(\alpha_1 - \alpha_2 + 1)}{2\alpha_2 - 3}.$$

Now we can obtain the exponential convergence of $y(\cdot)$ by using Theorem 3.2 for $\varepsilon = 1$. □

Now let us examine Theorem 3.2 when $\varepsilon = 2$. In this case, we will obtain from (3.19) that the convergence rate of $a(t)$ is

$$O((Pt^2 + Qt + R)e^{-2t}),$$

which is faster than the rate $O(e^{-t})$ obtained in [16] for the second order dynamical system.

Theorem 3.7 *Suppose that the operators A and B satisfy Assumption 3.1. Let x_* be the unique solution of Problem (1.1). Let θ satisfy (3.4) and denote the parameters as in (3.5)–(3.6) and (3.24). Then $y(\cdot)$ converges exponentially to x_* provided that coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy*

$$\alpha_2 > \max\{8\varphi, 6, 6\varphi + 4\}, \tag{3.28}$$

$$\underline{\beta} \triangleq \max\{4(\alpha_2 - 3), 4\varphi(\alpha_2 - 3)\} < \alpha_1 < \bar{\beta} \triangleq \frac{1}{2}\alpha_2(\alpha_2 - 2), \tag{3.29}$$

$$\underline{q} \triangleq \frac{2}{\delta}(\alpha_1 - 2\alpha_2 + 4) < \alpha_0 < \bar{p} \triangleq \ell\alpha_1 \cdot \min\left\{\frac{\alpha_1 - 2\alpha_2 + 4}{2(\alpha_2 - 3)}, \frac{1}{3}(\alpha_2 - 4)\right\}. \tag{3.30}$$

Proof Like in Theorem 3.6, we must check $\underline{\beta} < \bar{\beta}$. Indeed, we have

$$\frac{1}{2}\alpha_2(\alpha_2 - 2) > \frac{1}{2}\alpha_2(\alpha_2 - 3) > 4\varphi(\alpha_2 - 3).$$

Next is to prove $\underline{q} < \bar{p}$. It follows from (3.29) that

$$\alpha_1 > 4(\alpha_2 - 3) > 2(\alpha_2 - 2),$$

which gives

$$\alpha_1 - 2\alpha_2 + 4 > 0. \tag{3.31}$$

Again using (3.29), we get $\alpha_1 > 4\varphi(\alpha_2 - 3)$, which gives

$$\frac{\ell\alpha_1}{2(\alpha_2 - 3)} > \frac{2}{\delta}. \tag{3.32}$$

From (3.31)–(3.32), we obtain

$$\ell\alpha_1 \cdot \frac{\alpha_1 - 2\alpha_2 + 4}{2(\alpha_2 - 3)} > \frac{2}{\delta}(\alpha_1 - 2\alpha_2 + 4). \tag{3.33}$$

We observe $\alpha_1(\alpha_2 - 6\varphi - 4) + 12\varphi(\alpha_2 - 2) > 0$, which is equivalent to saying that

$$\ell\alpha_1 \cdot \frac{1}{3}(\alpha_2 - 4) > \frac{2}{\delta}(\alpha_1 - 2\alpha_2 + 4). \tag{3.34}$$

Hence, the inequality $\underline{q} < \bar{p}$ follows from (3.33)–(3.34). We left the reader to checking (3.8)–(3.13). Thus, $y(\cdot)$ converges exponentially to x_* . Moreover, it follows from (3.19) that the convergence rate is

$$O((Pt^2 + Qt + R)e^{-2t}),$$

for some constants P, Q, R . □

4 Discrete Time Dynamical System

In this section, we establish the linear convergence of the numerical scheme (2.3) for solving (1.1) under the following additional assumption.

Assumption 4.1 The coefficients $\alpha_0, \alpha_1, \alpha_2$ satisfy

$$\frac{\ell}{\alpha_0} (1 - \alpha_2 + \alpha_1) > 1, \quad (4.1)$$

$$\frac{\ell}{\alpha_0} (2\alpha_1 - \alpha_2) < 3, \quad (4.2)$$

$$\frac{\ell\alpha_1}{\alpha_0} > 3, \quad (4.3)$$

where ℓ is defined in (3.5).

We denote the following parameters

$$\begin{cases} D_2 \triangleq \frac{\ell\alpha_1}{\alpha_0} - 3, \\ D_1 \triangleq \frac{\ell\alpha_2\alpha_1}{\alpha_0} - 2\alpha_2 - 3, \\ D_0 \triangleq \frac{\ell\alpha_1^2}{\alpha_0} - 2\alpha_2 - \alpha_1, \end{cases} \quad \begin{cases} E_1 \triangleq \frac{\ell}{\alpha_0} (\alpha_2 - 2\alpha_1) + 3, \\ E_0 \triangleq \frac{\ell}{\alpha_0} (\alpha_2^2 - 2\alpha_1 - \alpha_2\alpha_1) + \alpha_2 + 3, \end{cases} \quad (4.4)$$

$$F_0 \triangleq \frac{\ell}{\alpha_0} (1 - \alpha_2 + \alpha_1) - 1 \quad (4.5)$$

and

$$u(n) \triangleq \|z(n) - x_*\|^2, \quad c_k(t) \triangleq \|z^{\Delta(k)}(n)\|^2. \quad (4.6)$$

Remark 4.1 Under Assumption 4.1, we have $F_0, E_1, D_2 > 0$. Note also that under Assumption 3.1, the stepsize ω must be bounded from above, i.e.

$$\omega < \frac{4\gamma}{L^2 + 4L\gamma - 4\gamma^2}.$$

This upper bound of ω is larger than that of the classical forward–backward algorithm, which is $\omega < \frac{2\gamma}{L^2}$ (see e.g. [13, Proposition 25.9]) when A is maximal monotone and B is γ -strongly monotone and L -Lipschitz continuous.

4.1 Global Linear Convergence

Theorem 4.2 Suppose that the operators A and B satisfy Assumption 3.1. Let x_* be the unique solution of Problem (1.1). Let θ satisfy (3.4) and denote the parameters as in (3.5), (4.4), (4.5) and Assumption 4.1 holds. Assume that there exists $\xi > 0, \xi \neq 1$

such that the following conditions hold

$$-\xi^3 + \alpha_2 \xi^2 - \alpha_1 \xi + \delta \alpha_0 \geq 0, \tag{4.7}$$

$$D_2 \xi^2 - D_1 \xi + D_0 \geq 0, \tag{4.8}$$

$$3\xi^2 - 2\alpha_2 \xi + \alpha_1 \geq 0, \tag{4.9}$$

$$-2D_2 \xi + D_1 \geq 0, \tag{4.10}$$

$$-E_1 \xi + E_0 \geq 0, \tag{4.11}$$

$$\alpha_2 > 3\xi. \tag{4.12}$$

Then $z(n)$ converges linearly to x_* , i.e. there exist $M > 0$ and $q \in (0, 1)$ such that

$$\|z(n) - x_*\| \leq Mq^n \quad \forall n.$$

Proof Since

$$\begin{aligned} u^\Delta(n) &= 2 \langle z^\Delta(n), z(n) - x_* \rangle + c_1(n), \\ u^{\Delta(2)}(n) &= 2 \langle z^{\Delta(2)}(n), z(n) - x_* \rangle + 2c_1^\Delta(n) + 2c_1(n) - c_2(n), \\ u^{\Delta(3)}(n) &= 2 \langle z^{\Delta(3)}(n), z(n) - x_* \rangle \\ &\quad + 3c_1^{\Delta(2)}(n) + 3c_1^\Delta(n) - 3c_2^\Delta(n) - 3c_2(n) + c_3(n), \end{aligned}$$

we have

$$\begin{aligned} &2 \langle z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^\Delta(n), z(n) - x_* \rangle \\ &= u^{\Delta(3)}(n) + \alpha_2 u^{\Delta(2)}(n) + \alpha_1 u^\Delta(n) \\ &\quad - 3c_1^{\Delta(2)}(n) - (2\alpha_2 + 3)c_1^\Delta(n) - (2\alpha_2 + \alpha_1)c_1(n) \\ &\quad + 3c_2^\Delta(n) + (\alpha_2 + 3)c_2(n) - c_3(n). \end{aligned} \tag{4.13}$$

We observe

$$\begin{aligned} &\|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^\Delta(n)\|^2 \\ &= \alpha_1 c_1^{\Delta(2)}(n) + \alpha_2 \alpha_1 c_1^\Delta(n) + \alpha_1^2 c_1(n) \\ &\quad + (\alpha_2 - 2\alpha_1)c_2^\Delta(n) + (\alpha_2^2 - 2\alpha_1 - \alpha_2 \alpha_1)c_2(n) + (1 - \alpha_2 + \alpha_1)c_3(n). \end{aligned} \tag{4.14}$$

Using the definition of resolvent, equation (2.2) gives

$$\begin{aligned} & B \left(\frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) \right) \\ & - B(z(n)) - \frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\omega \alpha_0} \\ & \in (A + B) \left(\frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) \right), \end{aligned}$$

which combined with $0 \in (A + B)(x_*)$ and the γ -monotonicity of $A + B$ implies

$$\begin{aligned} & \gamma \left\| \frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) - x_* \right\|^2 \\ & \leq \left\langle B \left(\frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) \right) \right. \\ & \quad \left. - B(z(n)) - \frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\omega \alpha_0}, \right. \\ & \quad \left. \frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) - x_* \right\rangle. \end{aligned}$$

Since the operator B is L -Lipschitz, we can estimate the right hand side of the inequality above and then

$$\begin{aligned} & \gamma \left\| \frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) - x_* \right\|^2 \\ & \leq \left\langle B \left(\frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) \right) - B(z(n)), z(n) - x_* \right\rangle \\ & \quad + \frac{1}{\alpha_0^2} \left(L - \frac{1}{\omega} \right) \|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)\|^2 \\ & \quad - \frac{1}{\omega \alpha_0} \left\langle z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n), z(n) - x_* \right\rangle. \end{aligned}$$

Note that by the Cauchy–Schwarz inequality

$$\begin{aligned} & \left\langle B \left(\frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) \right) - B(z(n)), z(n) - x_* \right\rangle \\ & \leq \frac{L}{\alpha_0} \|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)\| \cdot \|z(n) - x_*\| \\ & \leq \frac{L^2}{4\theta\gamma\alpha_0^2} \|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)\|^2 + \theta\gamma u(n). \end{aligned}$$

Thus, we get

$$\begin{aligned} & \gamma \left\| \frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) - x_* \right\|^2 \\ & \leq \frac{1}{\alpha_0^2} \left(L - \frac{1}{\omega} + \frac{L^2}{4\theta\gamma} \right) \|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)\|^2 + \theta\gamma u(n) \\ & \quad - \frac{1}{\omega\alpha_0} \left\langle z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n), z(n) - x_* \right\rangle. \end{aligned} \tag{4.15}$$

Note that

$$\begin{aligned} & \gamma \left\| \frac{z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)}{\alpha_0} + z(n) - x_* \right\|^2 \\ & = \frac{\gamma}{\alpha_0^2} \|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)\|^2 + \gamma u(n) \\ & \quad + \frac{2\gamma}{\alpha_0} \left\langle z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n), z(n) - x_* \right\rangle. \end{aligned}$$

Inserting the equality above into (4.15), we get

$$\begin{aligned} & \frac{\ell}{\alpha_0} \|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)\|^2 + \delta\alpha_0 u(n) \\ & \quad + 2 \left\langle z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n), z(n) - x_* \right\rangle \leq 0, \end{aligned}$$

which implies, by (4.13) and (4.14), that

$$\begin{aligned} & u^{\Delta(3)}(n) + \alpha_2 u^{\Delta(2)}(n) + \alpha_1 u^{\Delta}(n) + \delta\alpha_0 u(n) \\ & \quad + D_2 c_1^{\Delta(2)}(n) + D_1 c_1^{\Delta(1)}(n) + D_0 c_1(n) + E_1 c_2^{\Delta(1)}(n) + E_0 c_2(n) + F_0 c_3(n) \leq 0. \end{aligned}$$

By (4.1), the inequality above gives

$$\begin{aligned} & u^{\Delta(3)}(n) + \alpha_2 u^{\Delta(2)}(n) + \alpha_1 u^{\Delta}(n) + \delta\alpha_0 u(n) \\ & \quad + D_2 c_1^{\Delta(2)}(n) + D_1 c_1^{\Delta}(n) + D_0 c_1(n) + E_1 c_2^{\Delta}(n) + E_0 c_2(n) \leq 0. \end{aligned}$$

Setting

$$\varepsilon \triangleq \frac{1}{1 - \xi}.$$

Then $\varepsilon > 1$ and conditions (4.7)–(4.12) can be written as

$$\delta\alpha_0\varepsilon^3 + \alpha_1\varepsilon^2(1 - \varepsilon) + (\alpha_2\varepsilon + 1 - \varepsilon)(1 - \varepsilon)^2 \geq 0, \tag{4.16}$$

$$D_0\varepsilon^2 + D_1(1 - \varepsilon)\varepsilon + D_2(1 - \varepsilon)^2 \geq 0, \tag{4.17}$$

$$\alpha_1\varepsilon^2 + 2\alpha_2\varepsilon(1 - \varepsilon) + 3(1 - \varepsilon)^2 \geq 0, \tag{4.18}$$

$$D_1\varepsilon + 2D_2(1 - \varepsilon) \geq 0, \tag{4.19}$$

$$E_0\varepsilon + E_1(1 - \varepsilon) \geq 0, \tag{4.20}$$

$$\varepsilon\alpha_2 + 3(1 - \varepsilon) > 0. \tag{4.21}$$

Through multiplying both sides by ε^{n+3} and then using Remark 2.12,

$$\begin{aligned} &(\varepsilon^{n+2}u^{\Delta(2)})^\Delta(n) + (\alpha_2\varepsilon + 1 - \varepsilon)(\varepsilon^{n+1}u^\Delta)^\Delta(n) \\ &+ [\alpha_1\varepsilon^2 + (\alpha_2\varepsilon + 1 - \varepsilon)(1 - \varepsilon)](\varepsilon^n u)^\Delta(n) \\ &+ \underbrace{[\delta\alpha_0\varepsilon^3 + \alpha_1\varepsilon^2(1 - \varepsilon) + (\alpha_2\varepsilon + 1 - \varepsilon)(1 - \varepsilon)^2]}_{\geq 0 \text{ (by (4.16))}} \varepsilon^n u(n) \\ &+ D_2(\varepsilon^{n+2}c_1^\Delta)^\Delta(n) + [D_1\varepsilon - D_2(\varepsilon - 1)](\varepsilon^{n+1}c_1)^\Delta(n) \\ &+ \underbrace{[D_0\varepsilon^2 + D_1(1 - \varepsilon)\varepsilon + D_2(1 - \varepsilon)^2]}_{\geq 0 \text{ (by (4.17))}} \varepsilon^{n+1}c_1(n) \\ &+ E_1(\varepsilon^{n+2}c_2)^\Delta(n) + \underbrace{[-E_1(\varepsilon - 1) + E_0\varepsilon]}_{\geq 0 \text{ (by (4.20))}} \varepsilon^{n+2}c_2(n) \leq 0. \end{aligned}$$

Let $m \in \mathbb{Z}_{\geq 1}$. After summing from $n = 0$ to $n = m - 1$,

$$\begin{aligned} &\varepsilon^{m+2}u^{\Delta(2)}(m) + (\alpha_2\varepsilon + 1 - \varepsilon)\varepsilon^{m+1}u^\Delta(m) + [\alpha_1\varepsilon^2 + (\alpha_2\varepsilon + 1 - \varepsilon)(1 - \varepsilon)]\varepsilon^m u(m) \\ &+ D_2\varepsilon^{m+2}c_1^\Delta(m) + [D_1\varepsilon - D_2(\varepsilon - 1)]\varepsilon^{m+1}c_1(m) + \underbrace{E_1\varepsilon^{m+2}c_2(m)}_{\geq 0 \text{ (by (4.2))}} \leq M_1, \end{aligned}$$

where M_1 is some positive constant. Again using Remark 2.12,

$$\begin{aligned} &(\varepsilon^{m+1}u^\Delta)^\Delta(m) + [\alpha_2\varepsilon + 2(1 - \varepsilon)](\varepsilon^m u)^\Delta(m) \\ &+ \underbrace{[\alpha_1\varepsilon^2 + 2\alpha_2\varepsilon(1 - \varepsilon) + 3(1 - \varepsilon)^2]}_{\geq 0 \text{ (by (4.18))}} \varepsilon^m u(m) \\ &+ D_2(\varepsilon^{m+1}c_1)^\Delta(m) + \underbrace{[D_1\varepsilon - 2D_2(\varepsilon - 1)]}_{\geq 0 \text{ (by (4.19))}} \varepsilon^{m+1}c_1(m) \leq M_1. \end{aligned}$$

Let $\kappa \in \mathbb{Z}_{\geq 2}$. After summing from $m = 1$ to $m = \kappa - 1$,

$$\varepsilon^{\kappa+1} u^\Delta(\kappa) + [\alpha_2 \varepsilon + 2(1 - \varepsilon)] \varepsilon^\kappa u(\kappa) + \underbrace{D_2 \varepsilon^{\kappa+1} c_1(\kappa)}_{\geq 0 \text{ (by (4.3))}} \leq M_1 \kappa + M_2,$$

where M_2 is some positive constant. Again using Remark 2.12,

$$(\varepsilon^\kappa u)^\Delta(\kappa) + \underbrace{[\alpha_2 \varepsilon + 3(1 - \varepsilon)] \varepsilon^\kappa u(\kappa)}_{\geq 0 \text{ (by (4.21))}} \leq M_1 \kappa + M_2,$$

which implies, after summing from $\kappa = 2$ to $\kappa = n - 1$, that

$$\varepsilon^n u(n) \leq M_1 n^2 + M_2 n + M_3 \leq M_4 n^2.$$

Here $n \in \mathbb{Z}_{\geq 3}$ and M_3, M_4 are some positive constants. Let q such that $1 < q < \varepsilon$. We have

$$u(n) \leq \frac{M_4 n^2}{\varepsilon^n} = \left(\frac{q}{\varepsilon}\right)^n \cdot \frac{M_4 n^2}{q^n} \leq M_5 \left(\frac{q}{\varepsilon}\right)^n,$$

where M_5 is some constant. The inequality above means that $z(n)$ converges linearly to x_* . □

4.2 Parameters Choices

Let us discuss now how to choose the parameters fulfilling all Assumptions in Theorem 4.2. Note that if D_0, D_1, E_0 satisfy

$$D_0, D_1, E_0 > 0, \tag{4.22}$$

then conditions (4.16)–(4.21) hold by letting $\xi \rightarrow 0^+$.

The following result simplifies the assumption (4.22) in algebraic terms of the coefficients $\alpha_0, \alpha_1, \alpha_2$.

Corollary 4.3 *Suppose that the operators A and B satisfy Assumption 3.1. Let x_* be the unique solution of Problem (1.1). Let θ satisfy (3.4) and denote the parameters as in (3.5) (4.4), (4.5). Then $z(n)$ converges linearly to x_* provided that coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy*

$$\alpha_2 < 2, \tag{4.23}$$

$$\max\{0, \alpha_2 - 1\} < \alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2}, \tag{4.24}$$

$$\alpha_0 < \ell \cdot \min \left\{ \frac{\alpha_1^2}{\alpha_1 + 2\alpha_2}, 1 - \alpha_2 + \alpha_1 \right\}. \tag{4.25}$$

Proof Since $\alpha_1 < \frac{\alpha_2^2}{\alpha_2+2}$, we have $E_0 > \alpha_2 + 3 > 0$. Also using $\alpha_1 < \frac{\alpha_2^2}{\alpha_2+2}$ and the fact that $\alpha_2 < 2$, we get

$$\alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2} < \frac{\alpha_2}{2}, \tag{4.26}$$

which gives (4.2). It follows from (4.26) that $\alpha_1 < \alpha_2$ and so

$$\frac{\ell\alpha_1}{3} > \frac{\ell\alpha_1^2}{\alpha_1 + 2\alpha_2} > \alpha_0.$$

The last inequality proves (4.3). Thus, Assumption 4.1 holds. Note that

$$\alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2} < \frac{2\alpha_2^2}{\alpha_2 + 3},$$

which gives

$$\frac{\ell\alpha_1\alpha_2}{2\alpha_2 + 3} > \frac{\ell\alpha_1^2}{\alpha_1 + 2\alpha_2} > \alpha_0$$

and then $D_1 > 0$. □

Remark 4.4 Note that there are common choices of parameters satisfied both Corollary 3.5 (as $\varepsilon \rightarrow 0$) and Corollary 4.3 (as $\xi \rightarrow 0$). The reader can check the following selection

$$\begin{aligned} \alpha_2 &< 1, \\ \alpha_1 &< \frac{\alpha_2^2}{\alpha_2 + 2}, \\ \alpha_0 &< \ell \cdot \min \left\{ \frac{1}{3}\alpha_1\alpha_2, \frac{\alpha_1^2}{\alpha_1 + 2\alpha_2} \right\}. \end{aligned}$$

Remark 4.5 An important application of the monotone inclusion (1.1) is the following important optimization problem

$$\min_{x \in \mathcal{H}} f(x) + g(x), \tag{4.27}$$

where $f : \mathcal{H} \rightarrow \mathbb{R}$ is a differentiable function with L -Lipschitz continuous gradient for some $L > 0$ and $g : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ is a proper and lower semicontinuous function.

Recall that the Fréchet subdifferential of g at x is defined by

$$\hat{\partial}g(x) := \left\{ u \in \mathcal{H}, \liminf_{y \rightarrow x} \frac{f(y) - f(x) - \langle u, y - x \rangle}{\|y - x\|} \geq 0 \right\}.$$

It is well known that if g is differentiable at x , then $\hat{\partial}g(x) = \{\nabla g(x)\}$. When g is a convex function, the Fréchet subdifferential coincides with the classical convex subdifferential, i.e.

$$\hat{\partial}g(x) = \partial g(x) = \{u \in \mathcal{H} : g(y) \geq g(x) + \langle u, y - x \rangle \ \forall y \in \mathcal{H}\}.$$

We notice that, if g is proper, γ_g -convex and lower semicontinuous then $\hat{\partial}g$ is maximally generalized γ_g -monotone. We assume that f and g are respectively γ_f and γ_g convex functions such that $\gamma = \gamma_f + \gamma_g > 0$. Then the set of minimizers of (4.27) coincides with the solution set of the following monotone inclusion problem

$$\text{find } x^* \in \mathcal{H} \text{ such that } 0 \in \nabla f(x^*) + \hat{\partial}g(x^*), \tag{4.28}$$

for which the results obtained from previous Sections can be applied.

5 Strongly Pseudo-monotone Variational Inequality

Let C be a nonempty and closed convex subset of \mathcal{H} . The normal cone of C at x is defined as

$$N_C(x) = \{u \in \mathcal{H}, \langle u, y - x \rangle \leq 0, \ \forall y \in C\},$$

which is maximally monotone [13]. In this section, we focus on the restrictive category of Problem (1.1) of the form

$$\text{find } x_* \in \mathcal{H} \text{ such that } 0 \in A(x_*) + N_C(x_*). \tag{5.1}$$

Note that, if A is γ_A -monotone, then (5.1) is a special case of (1.1). Indeed, the sum of two monotone operators is still monotone [13]. However, this is not the case if A is non-monotone (e.g. only pseudo-monotone). For example, the operator

$$A(x_1, x_2) := (x_1^2 + x_2^2)(-x_2, x_1)^T$$

is pseudo-monotone but $A + \epsilon I$ is not (pseudo)-monotone for any $\epsilon > 0$ (see [41, Counterexample 2.1]).

In this section, we will consider the case when A is γ -strongly pseudo-monotone and hence the results obtained in the previous Sections cannot be directly applied. Problem (5.1) is equivalent to the variational inequality $VI(A, C)$: find $x_* \in C$ such that

$$\langle A(x_*), y - x_* \rangle \geq 0 \ \forall y \in C. \tag{5.2}$$

For each $x \in \mathcal{H}$, there exists a unique point in C (see, e.g., [31]), denoted by $P_C(x)$, such that

$$\|x - P_C(x)\| \leq \|x - y\| \ \forall y \in C.$$

Some well-known properties of the metric projection $P_C : \mathcal{H} \rightarrow C$ are given in the following lemma [25, 31].

Lemma 5.1 *Assume that the set C is a closed convex subset of \mathcal{H} . Then we have the following:*

(a) $P_C(\cdot)$ is a nonexpansive operator, i.e., for all $x, y \in \mathcal{H}$, it holds that

$$\|P_C(x) - P_C(y)\| \leq \|x - y\|.$$

(b) For any $x \in \mathcal{H}$ and $y \in C$, it holds that

$$\langle x - P_C(x), y - P_C(x) \rangle \leq 0.$$

Assumption 5.1 (i) The coefficients $\alpha_0, \alpha_1, \alpha_2 > 0$.

(ii) The operator $A : \mathcal{H} \rightarrow \mathcal{H}$ is γ -strongly pseudo-monotone and L -Lipschitz continuous.

(iii) The parameter $\omega > 0$ satisfies

$$\omega < \frac{4\gamma}{L^2}. \tag{5.3}$$

Remark 5.2 Under Assumption 5.1 (ii) and (iii), the problem $VI(A, C)$ has a unique solution [30].

We will need the following important estimate and error bounds.

Proposition 5.3 [44] *Let $C \subset \mathcal{H}$ be a nonempty closed convex subset. Let A be an operator that is γ -strongly pseudo-monotone and L -Lipschitz on C . Let x_* be the unique solution of Problem (5.2). For every $\omega > 0$ and $x \in \mathcal{H}$, we have*

$$\langle x - P_C(x - \omega A(x)), x - x_* \rangle \geq \left(1 - \frac{\omega L^2}{4\gamma}\right) \|x - P_C(x - \omega A(x))\|^2 \tag{5.4}$$

and

$$\|x - x_*\| \leq \frac{1 + \omega\gamma + \omega L}{\omega\gamma} \|x - P_C(x - \omega A(x))\|. \tag{5.5}$$

In the whole section, we denote

$$\mu \triangleq 1 - \frac{\omega L^2}{4\gamma}, \quad \eta \triangleq \left(\frac{\omega\gamma}{1 + \omega\gamma + \omega L}\right)^2. \tag{5.6}$$

5.1 Continuous Time

In this case, we consider

$$y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t) + \alpha_0 [y(t) - P_C(y(t) - \omega A(y(t)))] = 0, \tag{5.7}$$

where $y^{(j)}(t_0) = v_j, j \in \{0, 1, 2\}$.

Denote

$$\begin{cases} G_2 \triangleq \frac{\mu\alpha_1}{\alpha_0}, \\ G_1 \triangleq \frac{\mu\alpha_2\alpha_1}{\alpha_0^2} - 3, \\ G_0 \triangleq \frac{\mu\alpha_1^2}{\alpha_0} - 2\alpha_2, \end{cases} \quad \begin{cases} H_1 \triangleq \frac{\mu\alpha_2}{\alpha_0}, \\ H_0 \triangleq \frac{\mu}{\alpha_0}(\alpha_2^2 - 2\alpha_1), \end{cases} \quad K_0 \triangleq \frac{\mu}{\alpha_0}. \tag{5.8}$$

5.1.1 Global Exponential Convergence

Theorem 5.4 *Suppose that Assumption 5.1 is satisfied. Let x_* be the unique solution of Problem (5.2). Let the parameters be denoted by (5.6) and (5.8). Assume that there exists $\varepsilon > 0$ such that the following conditions hold*

$$-\varepsilon^3 + \alpha_2\varepsilon^2 - \alpha_1\varepsilon + \mu\eta\alpha_0 \geq 0, \tag{5.9}$$

$$G_2\varepsilon^2 - G_1\varepsilon + G_0 \geq 0, \tag{5.10}$$

$$3\varepsilon^2 - 2\alpha_2\varepsilon + \alpha_1 \geq 0, \tag{5.11}$$

$$-2G_2\varepsilon + G_1 \geq 0, \tag{5.12}$$

$$-H_1\varepsilon + H_0 \geq 0, \tag{5.13}$$

$$\alpha_2 > 2\varepsilon. \tag{5.14}$$

Then the trajectory $y(\cdot)$ generated by dynamical system (5.7) converges exponentially to x_* .

Proof Consider the functions in (3.7). Similarly as (3.14), we also have

$$\begin{aligned} & a^{(3)}(t) + \alpha_2 a^{(2)}(t) + \alpha_1 a^{(1)}(t) - 3b_1^{(1)}(t) - 2\alpha_2 b_1(t) \\ &= 2 \left\langle y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t), y(t) - x_* \right\rangle \\ &= 2\alpha_0 \langle [-y(t) + P_C(y(t) - \omega A(y(t)))], y(t) - x_* \rangle. \end{aligned} \tag{5.15}$$

On one hand, by (5.4), we can estimate

$$\begin{aligned} & \alpha_0 \langle [-y(t) + P_C(y(t) - \omega A(y(t)))], y(t) - x_* \rangle \\ & \leq -\alpha_0 \mu \|y(t) - P_C(y(t) - \omega A(y(t)))\|^2 = -\frac{\mu}{\alpha_0} \|y^{(3)}(t) + \alpha_2 y^{(2)}(t) + \alpha_1 y^{(1)}(t)\|^2 \\ & = -\frac{\mu}{\alpha_0} [\alpha_1 b_1^{(2)}(t) + \alpha_2 \alpha_1 b_1^{(1)}(t) + \alpha_1^2 b_1(t) + \alpha_2 b_2^{(1)}(t) + (\alpha_2^2 - 2\alpha_1) b_2(t) + b_3(t)]. \end{aligned} \tag{5.16}$$

On the other hand, by (5.4) and (5.5), we get

$$\alpha_0 \langle [-y(t) + P_C(y(t) - \omega A(y(t)))], y(t) - x_* \rangle \leq -\alpha_0 \mu \eta a(t). \tag{5.17}$$

Thus, using (5.16) and (5.17), we estimate (5.15) as follows

$$\begin{aligned} & a^{(3)}(t) + \alpha_2 a^{(2)}(t) + \alpha_1 a^{(1)}(t) + \alpha_0 \mu \eta a(t) \\ & + G_2 b_1^{(2)}(t) + G_1 b_1^{(1)}(t) + G_0 b_1(t) + H_1 b_2^{(1)}(t) + H_0 b_2(t) + K_0 b_3(t) \leq 0. \end{aligned} \tag{5.18}$$

By arguments similar to those used in Theorem 3.2 but now applied to (5.18); meaning, do integrating after three times, we get the exponential convergence of $y(\cdot)$. \square

5.1.2 Parameters Choices

Remark 5.5 If G_0, G_1, H_0 satisfy

$$G_0, G_1, H_0 > 0, \tag{5.19}$$

then conditions (5.9)–(5.14) can be obtained by letting $\varepsilon \rightarrow 0^+$.

In the following result, we simplify the assumption (5.19) in the term of the upper and lower bounds of the coefficients $\alpha_0, \alpha_1, \alpha_2$.

Corollary 5.6 *Suppose that Assumption 5.1 is satisfied. Let x_* be the unique solution of Problem (5.2). Let the parameters be denoted by (5.6) and (5.8). Then the trajectory $y(\cdot)$ generated by dynamical system (5.7) converges exponentially to x_* provided that coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy the following conditions*

$$\begin{aligned} \alpha_1 &< \frac{\alpha_2^2}{2}, \\ \alpha_0 &< \mu \cdot \min \left\{ \frac{\alpha_1 \alpha_2}{3}, \frac{\alpha_1^2}{2\alpha_2} \right\}. \end{aligned}$$

Now we examine Theorem 5.4 when $\varepsilon = 1$.

Corollary 5.7 *Suppose that Assumption 5.1 is satisfied. Let x_* be the unique solution of Problem (5.2). Let the parameters be denoted by (5.6) and (5.8) and*

$$\psi \triangleq \frac{1}{\mu^2 \eta}. \tag{5.20}$$

Then the trajectory $y(\cdot)$ generated by dynamical system (5.7) converges exponentially to x_* provided that coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy the following conditions

$$\alpha_2 > \max\{3, 3\psi + 2, 4\psi\}, \tag{5.21}$$

$$\underline{\beta} \triangleq \max\{2\alpha_2 - 3, \psi(2\alpha_2 - 3)\} < \alpha_1 < \bar{\beta} \triangleq 0.5\alpha_2(\alpha_2 - 1), \tag{5.22}$$

$$\underline{q} \triangleq \frac{\alpha_1 - \alpha_2 + 1}{\mu\eta} < \alpha_0 < \bar{p} \triangleq \mu \cdot \min\left\{\frac{\alpha_1(\alpha_2 - 2)}{3}, \frac{\alpha_1(\alpha_1 - \alpha_2 + 1)}{2\alpha_2 - 3}\right\}. \tag{5.23}$$

Proof First, we show that (5.21) ensures the validity of (5.22); that is $\underline{\beta} < \bar{\beta}$. Indeed, it follows from (5.21) that $\alpha_2 > 3$ and so $2\alpha_2 - 3 \leq \frac{\alpha_2(\alpha_2 - 1)}{2}$. Also from (5.21), we have $\alpha_2 > 4\psi$ and then

$$\frac{\alpha_2(\alpha_2 - 1)}{2} > \frac{\alpha_2(2\alpha_2 - 3)}{4} > \psi(2\alpha_2 - 3).$$

Next, we show that (5.21)–(5.22) ensure the validity of (5.23); that is $\underline{q} < \bar{p}$. It results from (5.21) that $\alpha_2 > 3\psi + 2$ and so

$$\frac{\alpha_1 - \alpha_2 + 1}{\mu\eta} < \frac{\alpha_1}{\mu\eta} < \mu \cdot \frac{\alpha_1(\alpha_2 - 2)}{3}.$$

Meanwhile, by (5.22), we have $\alpha_1 > \psi(2\alpha_2 - 3)$, which gives

$$\frac{\alpha_1 - \alpha_2 + 1}{\mu\eta} < \mu \cdot \frac{\alpha_1(\alpha_1 - \alpha_2 + 1)}{2\alpha_2 - 3}.$$

Now we can prove the exponential convergence of $y(\cdot)$ by using Theorem 5.4 for $\varepsilon = 1$. □

5.2 Discrete Time

We consider the difference equation

$$z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n) + \alpha_0 [z(n) - P_C(z(n) - \omega A(z(n)))] = 0, \tag{5.24}$$

where $\alpha_2, \alpha_1, \alpha_0, \omega > 0$.

Denote

$$\begin{cases} S_2 \triangleq \frac{\mu\alpha_1}{\alpha_0} - 3, \\ S_1 \triangleq \frac{\mu\alpha_2\alpha_1}{\alpha_0} - 2\alpha_2 - 3, \\ S_0 \triangleq \frac{\mu\alpha_1^2}{\alpha_0} - 2\alpha_2 - \alpha_1, \end{cases} \quad \begin{cases} T_1 \triangleq \frac{\mu}{\alpha_0}(\alpha_2 - 2\alpha_1) + 3, \\ T_0 \triangleq \frac{\mu}{\alpha_0}(\alpha_2^2 - 2\alpha_1 - \alpha_2\alpha_1) + \alpha_2 + 3, \end{cases} \tag{5.25}$$

$$R_0 \triangleq \frac{\mu}{\alpha_0}(1 - \alpha_2 + \alpha_1) - 1. \tag{5.26}$$

5.2.1 Global Exponential Convergence

Assumption 5.2 The coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy

$$\frac{\mu}{\alpha_0}(1 - \alpha_2 + \alpha_1) > 1, \quad (5.27)$$

$$\frac{\mu}{\alpha_0}(2\alpha_1 - \alpha_2) < 3, \quad (5.28)$$

$$\frac{\mu\alpha_1}{\alpha_0} > 3, \quad (5.29)$$

where μ is defined in (5.6).

Remark 5.8 Under Assumption 5.2, we have $R_0, T_1, S_2 \geq 0$.

Theorem 5.9 Suppose that Assumptions 5.1 and 5.2 are satisfied. Let x_* be the unique solution of Problem (5.2). Let the parameters be denoted by (5.6) and (5.25)–(5.26). Assume that there exists $\xi > 0, \xi \neq 1$ such that the following conditions hold

$$-\xi^3 + \alpha_2\xi^2 - \alpha_1\xi + \mu\eta\alpha_0 \geq 0, \quad (5.30)$$

$$S_2\xi^2 - S_1\xi + S_0 \geq 0, \quad (5.31)$$

$$3\xi^2 - 2\alpha_2\xi + \alpha_1 \geq 0, \quad (5.32)$$

$$-2S_2\xi + S_1 \geq 0, \quad (5.33)$$

$$-T_1\xi + T_0 \geq 0, \quad (5.34)$$

$$\alpha_2 > 3\xi. \quad (5.35)$$

Then the sequence $z(\cdot)$ generated by (5.24) converges linearly to x_* .

Proof Consider the functions (4.6). Similarly as (4.13), we also have

$$\begin{aligned} & u^{\Delta(3)}(n) + \alpha_2 u^{\Delta(2)}(n) + \alpha_1 u^{\Delta}(n) \\ & \quad - 3c_1^{\Delta(2)}(n) - (2\alpha_2 + 3)c_1^{\Delta}(n) - (2\alpha_2 + \alpha_1)c_1(n) \\ & \quad + 3c_2^{\Delta}(n) + (\alpha_2 + 3)c_2(n) - c_3(n) \\ & = 2 \left\langle z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n), z(n) - x_* \right\rangle \\ & = 2\alpha_0 \langle -z(n) + P_C(z(n) - \omega A(z(n))), z(n) - x_* \rangle. \end{aligned} \quad (5.36)$$

On one hand, by (5.4), we can estimate

$$\begin{aligned}
 & \alpha_0 \langle -z(n) + P_C(z(n) - \omega A(z(n))), z(n) - x_* \rangle \\
 & \leq -\alpha_0 \mu \|z(n) - P_C(z(n) - \omega A(z(n)))\|^2 \\
 & = -\frac{\mu}{\alpha_0} \|z^{\Delta(3)}(n) + \alpha_2 z^{\Delta(2)}(n) + \alpha_1 z^{\Delta}(n)\|^2 \\
 & = -\frac{\mu}{\alpha_0} [\alpha_1 c_1^{\Delta(2)}(n) + \alpha_2 \alpha_1 c_1^{\Delta}(n) + \alpha_1^2 c_1(n) \\
 & \quad + (\alpha_2 - 2\alpha_1) c_2^{\Delta}(n) + (\alpha_2^2 - 2\alpha_1 - \alpha_2 \alpha_1) c_2(n) + (1 - \alpha_2 + \alpha_1) c_3(n)].
 \end{aligned} \tag{5.37}$$

On the other hand, by (5.4) and (5.5), we get

$$\alpha_0 \langle -z(n) + P_C(z(n) - \omega A(z(n))), z(n) - x_* \rangle \leq -\alpha_0 \mu \eta u(n). \tag{5.38}$$

Thus, using (5.37) and (5.38), we estimate (5.36) as follows

$$\begin{aligned}
 & u^{\Delta(3)}(n) + \alpha_2 u^{\Delta(2)}(n) + \alpha_1 u^{\Delta}(n) + \alpha_0 \mu \eta u(n) \\
 & + S_2 c_1^{\Delta(2)}(n) + S_1 c_1^{\Delta}(n) + S_0 c_1(n) + T_1 c_2^{\Delta}(n) + T_0 c_2(n) + R_0 c_3(n) \leq 0.
 \end{aligned} \tag{5.39}$$

Setting

$$\varepsilon \triangleq \frac{1}{1 - \xi},$$

then $\varepsilon > 1$ and conditions can be written as

$$\begin{aligned}
 & \delta \alpha_0 \varepsilon^3 + \alpha_1 \varepsilon^2 (1 - \varepsilon) + (\alpha_2 \varepsilon + 1 - \varepsilon) (1 - \varepsilon)^2 \geq 0, \\
 & S_0 \varepsilon^2 + S_1 (1 - \varepsilon) \varepsilon + S_2 (1 - \varepsilon)^2 \geq 0, \\
 & \alpha_1 \varepsilon^2 + 2\alpha_2 \varepsilon (1 - \varepsilon) + 3(1 - \varepsilon)^2 \geq 0, \\
 & S_1 \varepsilon + 2S_2 (1 - \varepsilon) \geq 0, \\
 & T_0 \varepsilon + T_1 (1 - \varepsilon) \geq 0, \\
 & \varepsilon \alpha_2 + 3(1 - \varepsilon) > 0.
 \end{aligned}$$

By arguments similar to those used in Theorem 4.2 but now applied to (5.39); meaning, do summing after three times, we get the exponential convergence of $z(\cdot)$. \square

5.2.2 Parameters Choices

Remark 5.10 If S_0, S_1, T_0 satisfy

$$S_0, S_1, T_0 > 0, \tag{5.40}$$

then we can get conditions (5.2.1)–(5.2.1) by letting $\xi \rightarrow 0$.

The following result simplifies condition (5.40) in the term of the lower and upper bounds of the coefficients $\alpha_0, \alpha_1, \alpha_2$. There are common choices of parameters satisfied both Corollary 5.6 and Corollary 5.11 below.

Corollary 5.11 *Suppose that Assumptions 5.1 and 5.2 are satisfied. Let x_* be the unique solution of Problem (5.2). Let the parameters be denoted by (5.6) and (5.25)–(5.26). Then $z(\cdot)$ converges linearly to x_* provided that coefficients $\alpha_0, \alpha_1, \alpha_2, \omega$ satisfy*

$$\alpha_2 < 2, \tag{5.41}$$

$$\max\{0, \alpha_2 - 1\} < \alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2}, \tag{5.42}$$

$$\alpha_0 < \mu \cdot \min \left\{ \frac{\alpha_1^2}{\alpha_1 + 2\alpha_2}, 1 - \alpha_2 + \alpha_1 \right\}. \tag{5.43}$$

Proof Since $\alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2}$, we have $T_0 > \alpha_2 + 3 > 0$. Also using $\alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2}$ and the fact that $\alpha_2 < 2$, we get

$$\alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2} < \frac{\alpha_2}{2}, \tag{5.44}$$

which gives (5.28). It follows from (5.44) that $\alpha_1 < \alpha_2$ and so

$$\frac{\mu\alpha_1}{3} > \frac{\mu\alpha_1^2}{\alpha_1 + 2\alpha_2} > \alpha_0.$$

The last inequality proves (5.29). Thus, Assumption 5.2 holds. Note that

$$\alpha_1 < \frac{\alpha_2^2}{\alpha_2 + 2} < \frac{2\alpha_2^2}{\alpha_2 + 3},$$

which gives

$$\frac{\mu\alpha_1\alpha_2}{2\alpha_2 + 3} > \frac{\mu\alpha_1^2}{\alpha_1 + 2\alpha_2} > \alpha_0$$

and then $S_1 > 0$. □

Remark 5.12 We consider the following optimization problem

$$\min_{x \in C} f(x), \tag{5.45}$$

where C is a nonempty and closed subset of \mathcal{H} , $f : \mathcal{H} \rightarrow \mathbb{R}$ is a γ -strongly pseudoconvex on C and differentiable function with L -Lipschitz continuous gradient for

some $L > 0$. Recall that the differentiable function f is called γ -strongly pseudo convex if there exists $\gamma > 0$ such that

$$\langle \nabla f(x), y - x \rangle \geq 0 \implies \langle \nabla f(y), y - x \rangle \geq \gamma \|x - y\|^2$$

for all $x, y \in C$. For more details on generalized convexity functions and their characterization, the readers are referred to [29]. The optimization problem (5.45) is equivalent to the following strongly pseudo-monotone variational inequality

$$\langle \nabla f(x_*), y - x_* \rangle \geq 0 \quad \forall y \in C. \quad (5.46)$$

As a consequence, all the results presented in this section can be applied directly to the pseudo-convex optimization problem (5.45).

Acknowledgements The authors are grateful to two anonymous reviewers for their constructive comments, which help improving significantly the presentation of the paper. Additionally, the authors would like to extend their appreciation to the Vietnam Institute for Advanced Study in Mathematics (VIASM) for organizing the International Conference "New Trends in Numerical Optimization and Applications" in December, 2021. It was during this event that the authors first met and initiated fruitful discussions that ultimately led to this research project. P. T. Vuong thanks the London Mathematical Society (LMS) for supporting his visit to P.V. Hai at the Hanoi University of Science and Technology in February 2024, when the authors finished the final version of this paper.

Funding P.V. Hai is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.02-2021.24.

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

References

1. Alvarez, F., Attouch, H.: An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Anal.* **9**, 3–11 (2001)
2. Antipin, A.S.: Continuous and iterative processes with projection operators and projection like operators. AN SSSR, Scientific Counsel on the Complex Problem Cybernetics, Moscow, pp. 5–43 (1989)
3. Antipin, A.S.: Minimization of convex functions on convex sets by means of differential equations. (Russian) *Differentsial' nye Uravneniya* **30**(9), 1475–1486 (1994); translation in *Differential Equations* **30**, 1365–1375 (1994)
4. Attouch, H., Alvarez, F.: The heavy ball with friction dynamical system for convex constrained minimization problems. In: *Optimization* (Namur, 1998), *Lecture Notes in Economics and Mathematical Systems* 481. Springer, Berlin, pp. 25–35 (2000)
5. Attouch, H., Cabot, A.: Convergence of a relaxed inertial proximal algorithm for maximally monotone operators. *Math. Program.* **184**, 243–287 (2020)
6. Attouch, H., Cabot, A.: Convergence of a relaxed inertial forward–backward algorithm for structured monotone inclusions. *Appl. Math. Optim.* **80**, 547–598 (2019)
7. Attouch, H., Chbani, Z., Riahi, H.: Fast convex optimization via a third-order in time evolution equation. *Optimization* **71**, 1275–1304 (2022)

8. Attouch, H., Chbani, Z., Riahi, H.: Fast convex optimization via a third-order in time evolution equation: TOGES-V an improved version of TOGES. *Optimization* (2022). <https://doi.org/10.1080/02331934.2022.2119084>
9. Attouch, H., Peypouquet, J., Redont, P.: A dynamical approach to an inertial forward–backward algorithm for convex minimization. *SIAM J. Optim.* **24**, 232–256 (2014)
10. Attouch, H., Goudou, X., Redont, P.: The heavy ball with friction method. I. The continuous dynamical system: global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system. *Commun. Contemp. Math.* **2**, 1–34 (2000)
11. Attouch, H., Peypouquet, J., Redont, P.: Backward–forward algorithms for structured monotone inclusions in Hilbert spaces. *J. Math. Anal. Appl.* **457**, 1095–1117 (2018)
12. Avriel, M., Diewert, W.E., Schaible, S., Zang, I.: *Generalized Concavity*. Society for Industrial and Applied Mathematics (2010)
13. Bauschke, H., Combettes, P.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer, Berlin (2011)
14. Bolte, J.: Continuous gradient projection method in Hilbert spaces. *J. Optim. Theory Appl.* **119**, 235–259 (2003)
15. Boţ, R.I., Csetnek, E.R.: Second order forward–backward dynamical systems for monotone inclusion problems. *SIAM J. Control Optim.* **54**, 1423–1443 (2016)
16. Boţ, R.I., Csetnek, E.R.: Convergence rates for forward–backward dynamical systems associated with strongly monotone inclusions. *J. Math. Anal. Appl.* **457**, 1135–1152 (2018)
17. Boţ, R.I., Csetnek, E.R.: A dynamical system associated with the fixed points set of a nonexpansive operator. *J. Dyn. Differ. Equat.* **29**, 155–168 (2017)
18. Boţ, R.I., Csetnek, E.R., Vuong, P.T.: The forward–backward–forward method from discrete and continuous perspective for pseudo-monotone variational inequalities in Hilbert spaces. *Eur. J. Oper. Res.* **287**, 49–60 (2020)
19. Boţ, R.I., Sedlmayer, M., Vuong, P.T.: A relaxed inertial forward–backward–forward algorithm for solving monotone inclusions with application to GANs. *J. Mach. Learn. Res.* **24**, 1–37 (2023)
20. Cavazzuti, E., Pappalardo, P., Passacantando, M.: Nash equilibria, variational inequalities, and dynamical systems. *J. Optim. Theory Appl.* **114**, 491–506 (2002)
21. Dao, M., Phan, H.: Adaptive Douglas–Rachford splitting algorithm for the sum of two operators. *SIAM J. Optim.* **29**, 2697–2724 (2019)
22. Douglas, J., Rachford, H.: On the numerical solution of heat conduction problems in two and three space variables. *Trans. Am. Math. Soc.* **82**, 421–439 (1956)
23. Eckstein, J., Bertsekas, D.: On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.* **55**, 293–318 (1992)
24. Facchinei, F., Pang, J.-S.: *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vols. I and II. Springer, New York (2003)
25. Goebel, K., Reich, S.: *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*. Marcel Dekker, New York (1984)
26. Guo, K., Han, D., Yuan, X.: Convergence analysis of Douglas–Rachford splitting method for strongly+ weakly convex programming. *SIAM J. Numer. Anal.* **55**, 1549–1577 (2017)
27. Ha, N.T.T., Strodiot, J.J., Vuong, P.T.: On the global exponential stability of a projected dynamical system for strongly pseudomonotone variational inequalities. *Opt. Lett.* **12**, 1625–1638 (2018)
28. Haraux, A.: *Systemes Dynamiques Dissipatifs et Applications*, Recherches en Mathematiques Appliquees 17. Masson, Paris (1991)
29. Karamardian, S., Schaible, S.: Seven kinds of monotone maps. *J. Optim. Theory Appl.* **66**, 37–46 (1990)
30. Kim, D.S., Vuong, P.T., Khanh, P.D.: Qualitative properties of strongly pseudomonotone variational inequalities. *Opt. Lett.* **10**, 1669–1679 (2016)
31. Kinderlehrer, D., Stampacchia, G.: *An Introduction to Variational Inequalities and Their Applications*. Academic, New York (1980)
32. Lions, P., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16**, 964–979 (1979)
33. Lorenz, D.A., Pock, T.: An inertial forward–backward algorithm for monotone inclusions. *J. Math. Imaging Vis.* **51**, 311–325 (2015)
34. Nagurney, A., Zhang, D.: *Projected Dynamical Systems and Variational Inequalities with Applications*. Kluwer Academic, Norwell (1996)

35. O'Connor, D., Vandenberghe, L.: On the equivalence of the primal–dual hybrid gradient method and Douglas–Rachford splitting. *Math. Program.* **179**, 85–108 (2020)
36. Pappalardo, M., Passacantando, M.: Stability for equilibrium problems: from variational inequalities to dynamical systems. *J. Optim. Theory Appl.* **113**, 567–582 (2002)
37. Polyak, B.T.: Some methods of speeding up the convergence of iteration methods. *USSR Comput. Math. Math. Phys.* **4**, 1–17 (1964)
38. Passty, G.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.* **72**, 383–390 (1979)
39. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* **14**, 877–898 (1976)
40. Su, W., Boyd, S., Candes, E.J.: A differential equation for modeling Nesterov's accelerated gradient method: theory and insights. In: *Advances in Neural Information Processing Systems (NIPS) 27* (2014)
41. Tam, N.N., Yao, J.C., Yen, N.D.: Solution methods for pseudomonotone variational inequalities. *J. Optim. Theory Appl.* **138**, 253–273 (2008)
42. Vinh, L.V., Tran, V.N., Vuong, P.T.: A second-order dynamical system for equilibrium problems. *Numer. Algorithms* **91**, 327–351 (2022)
43. Vuong, P.T.: The global exponential stability of a dynamical system for solving variational inequalities. *Netw. Spat. Econ.* **22**, 395–407 (2022)
44. Vuong, P.T.: A second order dynamical system and its discretization for strongly pseudo-monotone variational inequalities. *SIAM J. Control Optim.* **59**, 2875–2897 (2021)
45. Vuong, P.T., Strodiot, J.J.: A dynamical system for strongly pseudo-monotone equilibrium problems. *J. Optim. Theory Appl.* **185**, 767–784 (2020)

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