

UNIVERSIDAD DE CONCEPCIÓN



CENTRO DE INVESTIGACIÓN EN INGENIERÍA MATEMÁTICA (CI²MA)



Nonlinear forward-backward algorithm for solving
non-monotone+Lipschitz inclusions with applications to adjoint
mismatch problems

JEAN-CHRISTOPHE PESQUET, FERNANDO ROLDÁN

PREPRINT 2026-27

SERIE DE PRE-PUBLICACIONES

NONLINEAR FORWARD-BACKWARD ALGORITHM FOR SOLVING NON-MONOTONE+LIPSCHITZ INCLUSIONS WITH APPLICATIONS TO ADJOINT MISMATCH PROBLEMS

JEAN-CHRISTOPHE PESQUET¹ AND FERNANDO ROLDÁN²

ABSTRACT. This article presents a novel algorithmic framework for solving a broad class of nonlinear inclusion problems involving non-monotone and Lipschitz continuous operators. Our primary motivation originates from the study of adjoint mismatch problems, which frequently arise in inverse problems and data science when the adjoint of a linear measurement operator is replaced by an approximation. Because this approximation may inherently destroy classical monotonicity properties, standard convergence guarantees for splitting methods may no longer apply. To address this, we consider a Nonlinear Forward-Backward algorithm and rigorously establish its convergence without assuming monotonicity. By leveraging a warped resolvent formulation and semimonotonicity assumptions, we derive explicit conditions ensuring both weak convergence and, under stronger assumptions, R-linear convergence. Furthermore, our theoretical analysis yields new convergence guarantees for several popular methods, including the Condat-Vũ and Forward-Half-Reflected-Backward algorithms. Finally, we illustrate our theoretical findings and demonstrate the applicability of our approach through numerical experiments in signal recovery.

Keywords. *Splitting algorithms, convergence analysis, fixed point theory, convex optimization, adjoint mismatch*

2020 *Mathematics Subject Classification.* 47H05, 47H10, 65K05, 90C25.

1. INTRODUCTION

The main objective of this paper is to develop numerical methods for solving the following problem.

Problem 1.1. *Let $\mathcal{H}$ and $\mathcal{G}$ be real Hilbert spaces. Let $L: \mathcal{H} \rightarrow \mathcal{G}$ be a bounded linear operator with adjoint $L^*: \mathcal{G} \rightarrow \mathcal{H}$. Let $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ and $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$ be set-valued operators, let $C: \mathcal{H} \rightarrow \mathcal{H}$ be a β -cocoercive operator for $\beta \in]0, +\infty[$, and let $D: \mathcal{H} \rightarrow \mathcal{H}$ be a ϑ -Lipschitzian operator for $\vartheta \in [0, +\infty[$. The problem is to*

$$\text{find } (x, u) \in \mathcal{H} \times \mathcal{G} \text{ such that } \begin{cases} 0 & \in (A + C + D)x + L^*u \\ u & \in BLx, \end{cases} \quad (1.1)$$

under the assumption that its solution set, denoted by $\mathbf{Z}$, is nonempty.

This problem encompasses several scenarios in convex programming [18], game theory [2], data science [23], and image recovery [6, 8], among others. Note that, given $(\hat{x}, \hat{u}) \in \mathbf{Z}$, $\hat{x}$ is a solution to the primal inclusion

$$\text{find } x \in \mathcal{H} \text{ such that } 0 \in (A + L^*BL + C + D)x, \quad (1.2)$$

¹UNIVERSITÉ PARIS-SACLAY, CENTRALESUPÉLEC, CVN, INRIA, GIF-SUR-YVETTE 91190, FRANCE.

² DEPARTAMENTO DE INGENIERÍA MATEMÁTICA AND CI²MA, UNIVERSIDAD DE CONCEPCIÓN, CONCEPCIÓN, CHILE.

E-mail addresses: jean-christophe.pesquet@centralesupelec.fr, fernandoroldan@udec.cl (corresponding author).

and $\hat{u}$ is a solution to the dual inclusion

$$\text{find } u \in \mathcal{G} \text{ such that } 0 \in B^{-1}u - L(A + C + D)^{-1}(-L^*u). \quad (1.3)$$

An instance of Problem 1.1 is the following optimization problem.

Problem 1.2. Let $(\mu_f, \mu_g) \in \mathbb{R}^2$ and $(\beta, \vartheta_d) \in]0, +\infty[\times [0, +\infty[$. Let $\mathcal{H}$, $\mathcal{G}$, and $\mathcal{K}$ be real Hilbert spaces. Let $L: \mathcal{H} \rightarrow \mathcal{G}$ and $T: \mathcal{H} \rightarrow \mathcal{K}$ be bounded linear operators. Let $f: \mathcal{H} \rightarrow]-\infty, +\infty]$ be a proper, lower-semicontinuous, μ_f -convex function, let $g: \mathcal{G} \rightarrow]-\infty, +\infty]$ be a proper, lower-semicontinuous, μ_g -convex function, let $h: \mathcal{H} \rightarrow \mathbb{R}$ be a differentiable convex function with a $1/\beta$ -Lipschitzian gradient, and let $d: \mathcal{K} \rightarrow \mathbb{R}$ be a differentiable convex function with a ϑ_d -Lipschitzian gradient. The problem is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) + h(x) + d(Tx), \quad (1.4)$$

under the assumption that the set of solutions is nonempty.

Indeed, under standard qualification conditions, Problem 1.2 can be written as (1.2) by considering $A = \partial f$, $B = \partial g$, $C = \nabla h$, $D = T^* \circ \nabla d \circ T$, and $\vartheta = \|T\|^2 \vartheta_d$, where ∂f (resp. ∂g) denotes the Fréchet subdifferential of f (resp. g). In turn, replacing T^* in the definition of D by a bounded linear operator $K: \mathcal{K} \rightarrow \mathcal{H}$ amounts to introducing an *adjoint mismatch*. In this case, D is no longer guaranteed to be a cocoercive operator, but it is ϑ -Lipschitzian with $\vartheta = \|K\| \|T\| \vartheta_d$. Even though A , B , and C are monotone operators, the monotonicity of the overall inclusion is not guaranteed.

If the operators A , B , and D are monotone, then (1.1) defines a monotone inclusion, which can be solved by the Primal-Dual Method with Block Triangular Resolvent (PDBTR) proposed in [48]. This method is derived from the Nonlinear Forward-Backward algorithm under suitable metrics and Lipschitzian operators (see also the related warped-resolvent framework in [10, 19, 33, 52, 51]). In particular, when $D = 0$, PDBTR reduces to the Condat–Vũ (CV) algorithm [24, 59] and to Chambolle–Pock (CP) if $C = 0$ [12]. On the other hand, when $L = 0$ and $B = 0$, it reduces to the Forward-Half-Reflected-Backward (FHRB) algorithm proposed in [46]. Moreover, if $L = 0$ and $D = 0$, the Forward-Backward (FB) algorithm is recovered [50]. If $D = 0$, $\mathcal{H} = \mathcal{G}$, and $L = \text{Id}$, Problem 1.1 can be solved by the Three-Operator Splitting (TOS) proposed in [25], which further reduces to Douglas–Rachford (DR) splitting [28, 29, 41] if $C = 0$. On the other hand, if $B = 0$ and $L = 0$, Problem 1.1 can be solved by the Forward-Backward-Half-Forward (FBHF) algorithm proposed in [9] which reduces to FB when $D = 0$ and to Forward-Backward-Forward (FBF) when $C = 0$ [58]. Other methods for solving this problem are available in [3, 11, 17, 20, 53].

The adjoint mismatch problem arises when the adjoint T^* of a linear measurement operator T must be approximated because of physical or computational limitations, or efficiency requirements. It is particularly common in inverse problems such as computed tomography, where its importance was first highlighted in [60], and has since been widely studied [14, 15, 16, 27, 30, 42, 43, 55]. Algebraic analyses of mismatched quadratic problems corresponding to (1.4) with $f = g = h = L = 0$ and $d = \tilde{d} := \alpha \|\cdot\|^2/2$, $\alpha \in [0, +\infty[$, were provided in [27, 30], while a mismatched Kaczmarz method was studied in [42]. Mismatched and preconditioned versions of FB were proposed in [15, 55]. For $h = 0$ and $d = \tilde{d}$, mismatched versions of the CV, Loris–Verhoeven [44], and Combettes–Pesquet [22] algorithms were developed in [14], allowing iteration-dependent approximations $K_n: \mathcal{K} \rightarrow \mathcal{H}$ of K . A mismatched CP method was introduced in [43], whereas mismatched FBHF and Forward–Douglas–Rachford–Forward methods [54] were proposed in [16] for Problem 1.1 with $B = L = 0$.

To address scenarios beyond those covered by the preceding methods, we consider inclusion problems outside the monotone setting. Over the years, several authors have investigated

algorithms that relax the monotonicity assumption on the operators. For instance, in [21, 37, 38], the proximal point algorithm (PPA) [47] was studied for μ -comonotone operators, including the cohypomonotone case $\mu < 0$. Moreover, the convergence of the PPA was analyzed under metric subregularity and submonotonicity in [45], for (strongly) quasiconvex functions in [36, 39], and for prox-convex functions in [34, 35]. Extragradient-type algorithms for ρ -comonotone operators were studied in [7, 40, 56] and in [26] under a weak Minty condition. The convergence of the Douglas–Rachford algorithm and the Chambolle–Pock algorithm under a weak Minty condition was established in [31, 32]. In [1], the authors studied the convergence of Halpern and Krasnosel’skiĭ–Mann iterations for solving minimax problems under a weak Minty condition. Warped-resolvent algorithms involving μ -comonotone operators have been studied in [49, 57]. Note that in these works the notions of comonotonicity and Lipschitz continuity are defined with respect to the operator inducing the warped resolvent. Finally, in [5], the authors investigated properties of μ -monotone and ρ -comonotone operators and their resolvents.

In this article, we propose a primal-dual method for solving Problem 1.1. Its convergence analysis is based on a nonlinear forward-backward algorithm for comonotone inclusions. In particular, we extend the convergence analysis of the Nonlinear Forward-Backward algorithm with momentum from [48] to comonotone operators. We analyze conditions on operators A , B , C , and D to ensure that the primal-dual operators associated with our method are comonotone, thereby recovering, as a special case, the conditions established in [31] for CP beyond the monotone setting. Moreover, our conditions apply to CV, FHRB, and FB in the comonotone setting.

The structure of the paper is as follows. In Section 2, we present the notation and necessary mathematical background. In Section 3, we introduce an auxiliary inclusion, present the Nonlinear Forward-Backward algorithm, and study its convergence for comonotone operators. In Section 4, we provide conditions guaranteeing the comonotonicity of the associated primal-dual operators and establish the convergence of the proposed methods for solving Problem 1.1. Finally, in Section 5, we present an application to adjoint mismatch problems and numerical experiments in signal recovery.

2. NOTATION AND PRELIMINARIES

Throughout this paper, $\mathcal{H}$ and $\mathcal{G}$ are real Hilbert spaces with inner product $\langle \cdot | \cdot \rangle$ and associated norm $\| \cdot \|$. The symbols $\rightharpoonup$ and $\rightarrow$ denote the weak and strong convergence, respectively. The identity operator is denoted by Id and the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$ is denoted by $\mathcal{B}(\mathcal{H}, \mathcal{G})$. Given a linear operator $M \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, we denote its adjoint by $M^* \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ and, when $\mathcal{G} = \mathcal{H}$ define $\langle \cdot | \cdot \rangle_M = \langle \cdot | M \cdot \rangle$. If $S \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ is a self-adjoint strongly monotone operator, i.e, there exists $\mu \in]0, +\infty[$ such that, for every $x \in \mathcal{H}$, $\langle Sx | x \rangle \geq \mu \|x\|^2$, then, $\langle \cdot | \cdot \rangle_S$ is an inner product on $\mathcal{H}$. The norm of $x \in \mathcal{H}$ induced by S is denoted by $\|x\|_S = \sqrt{\langle x | x \rangle_S}$. Let $\mathcal{E} \subset \mathcal{H}$ and $R: \mathcal{E} \rightarrow \mathcal{H}$ and let $\beta \in]0, +\infty[$. The operator R is β -cocoercive with respect to S if

$$(\forall x \in \mathcal{E})(\forall y \in \mathcal{E}) \quad \langle x - y | Rx - Ry \rangle \geq \beta \|Rx - Ry\|_{S^{-1}}^2. \quad (2.1)$$

Note that this definition is equivalent to the cocoercivity of $S^{-1} \circ R$ in the Hilbert space $(\mathcal{H}, \langle \cdot | \cdot \rangle_S)$ (see [48, Section 1.2]). Similarly, R is said to be $1/\beta$ -Lipschitzian with respect to S if

$$(\forall x \in \mathcal{E})(\forall y \in \mathcal{E}) \quad \beta \|Rx - Ry\|_{S^{-1}} \leq \|x - y\|_S. \quad (2.2)$$

Obviously, (2.1) implies (2.2). We say that R is nonexpansive with respect to S if it is 1-Lipschitzian with respect to S . The set of fixed points of R is $\text{Fix } R = \{x \in \mathcal{E} \mid x = Rx\}$. Let $2^{\mathcal{H}}$ be the power set of $\mathcal{H}$ and let $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be a set-valued operator. The domain, range, set of zeros, and graph of A are $\text{dom } A = \{x \in \mathcal{H} \mid Ax \neq \emptyset\}$, $\text{ran } A = \{u \in \mathcal{H} \mid (\exists x \in \mathcal{H}) u \in Ax\}$,

$\text{zer } A = \{x \in \mathcal{H} \mid 0 \in Ax\}$, and $\text{gra } A = \{(x, u) \in \mathcal{H} \times \mathcal{H} \mid u \in Ax\}$, respectively. Moreover, the inverse of A is the operator $A^{-1}: \mathcal{H} \rightarrow 2^{\mathcal{H}}: u \mapsto \{x \in \mathcal{H} \mid u \in Ax\}$ and its resolvent is $J_A = (\text{Id} + A)^{-1}$. Let $M \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ and $N \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be self-adjoint (not necessarily strongly monotone) operators. The following extension of monotonicity was considered in [31]. The operator A is (M, N) -semimonotone on $\mathcal{S} \subset \mathcal{H}^2$ if

$$(\forall (y, v) \in \mathcal{S})(\forall (x, u) \in \text{gra } A) \quad \langle x - y \mid u - v \rangle \geq \langle x - y \mid M(x - y) \rangle + \langle u - v \mid N(u - v) \rangle. \quad (2.3)$$

This is also equivalent to the fact that A^{-1} is (N, M) -semimonotone on $\tilde{\mathcal{S}} = \{(u, x) \in \mathcal{H}^2 \mid (x, u) \in \mathcal{S}\}$. If $\mathcal{S} = \text{gra } A$, the operator A is called (M, N) -semimonotone. When $N = 0$ (resp. $M = 0$), we refer to M -monotonicity (resp. N -comonotonicity). When $M = 0$ and $N = \beta S^{-1}$, this reduces to β -cocoercivity with respect to S . Additionally, A is maximally (M, N) -semimonotone if it is (M, N) -semimonotone and its graph is maximal with respect to inclusion among the graphs of (M, N) -semimonotone operators. Given $(\mu, \rho) \in \mathbb{R}^2$, if $M = \mu \text{Id}$ and $N = \rho \text{Id}$, for convenience, we simply say that A is (μ, ρ) -semimonotone on $\mathcal{S}$. The following table summarizes various notions of monotonicity which are particular cases of (M, N) -semimonotonicity. Note that [49] also introduces the notion of $(\rho M, W)$ -hypomonotonicity

M	N	Property	Reference
μId	ρId	(μ, ρ) -semimonotone	[32]
μM^2	0	(μM) -hypomonotone	[49, 57]
μId	0	μ -monotone	[5]
0	ρId	ρ -comonotone	[5]
$- \mu \text{Id}$	0	$ \mu $ -hypomonotone	[21]
0	$- \rho \text{Id}$	$ \rho $ -cohypomonotone	[21]
$ \mu \text{Id}$	0	$ \mu $ -strongly monotone	[4]
0	$ \rho \text{Id}$	$ \rho $ -cocoercive	[4]
0	0	monotone	[4]

TABLE 1. Special cases of (M, N) -semimonotonicity. Here, $(\mu, \rho) \in \mathbb{R}^2$.

where $W: \mathcal{H} \rightarrow \mathcal{H}$ is a single-valued operator (not necessarily linear). Let A be a maximally (M, N) -semimonotone operator. If $M = \mu \text{Id}$, $\mu > -1$, $N = \rho \text{Id}$, and $\rho \geq 0$, J_A is everywhere defined and single-valued on $\mathcal{H}$. Since $J_A = \text{Id} - J_{A^{-1}}$, if $M = \mu \text{Id}$, $\mu \geq 0$, $N = \rho \text{Id}$, and $\rho > -1$, J_A is also everywhere defined and single-valued on $\mathcal{H}$ [5]. We say that $\text{gra } A$ is sequentially weak-strong closed if, for every sequence $(x_n, u_n)_{n \in \mathbb{N}}$ in $\text{gra } A$ such that $x_n \rightharpoonup x$ and $u_n \rightarrow u$, we have $u \in Ax$. An operator $T: \mathcal{E} \subset \mathcal{H} \rightarrow \mathcal{H}$ is demiclosed at $y \in \mathcal{H}$ if, for every $x \in \mathcal{E}$ and every sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathcal{E}$ such that $x_n \rightharpoonup x$ and $Tx_n \rightarrow y$, we have $Tx = y$. The operator T is demiclosed if it is demiclosed at every point of $\mathcal{H}$.

We denote by $\Gamma_0(\mathcal{H})$ the class of proper, lower semicontinuous, convex functions $f: \mathcal{H} \rightarrow]-\infty, +\infty]$. Let $f \in \Gamma_0(\mathcal{H})$. The Fenchel conjugate of f is defined by $f^*: u \mapsto \sup_{x \in \mathcal{H}} (\langle x \mid u \rangle - f(x))$ and we have $f^* \in \Gamma_0(\mathcal{H})$. The Moreau subdifferential of f is the maximally monotone operator $\partial f: x \mapsto \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) f(x) + \langle y - x \mid u \rangle \leq f(y)\}$. In addition, $(\partial f)^{-1} = \partial f^*$ and the set of minimizers of f is $\text{zer } \partial f = \arg \min_{x \in \mathcal{H}} f$.

Examples of semimonotone operators can be found in [5, Section 6], where properties and examples of hypoconvex functions, as well as the semimonotonicity properties of their *generalized subdifferentials* (Clarke–Rockafellar, Mordukhovich, and Fréchet), are presented. In addition, several examples of semimonotone problems can be found in [31, 32, 49].

For further properties of monotone operators, nonexpansive mappings, and convex analysis, the reader is referred to [4].

We conclude this section by presenting two lemmas, which will be useful for dealing with sums of semimonotone operators.

Lemma 2.1. *Let $M \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be a self-adjoint monotone operator and let $(\rho, \beta) \in \mathbb{R}^2$ be such that $\rho + \beta > 0$. The following inequality holds:*

$$(\forall (x, y) \in \mathcal{H}^2) \quad \rho \|x\|_M^2 + \beta \|y\|_M^2 \geq \frac{\beta \rho}{\rho + \beta} \|x + y\|_M^2. \quad (2.4)$$

Proof. We have

$$\left\| \frac{\rho x}{\rho + \beta} + \frac{\beta(-y)}{\rho + \beta} \right\|_M^2 + \frac{\rho \beta}{(\rho + \beta)^2} \|x + y\|_M^2 = \frac{\rho}{\rho + \beta} \|x\|_M^2 + \frac{\beta}{\rho + \beta} \|y\|_M^2.$$

Multiplying this identity by $\rho + \beta > 0$ yields the result. $\square$

Lemma 2.2. *Let $S \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be self-adjoint and strongly monotone, let $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$, and let $C: \mathcal{H} \rightarrow 2^{\mathcal{H}}$. Let $(\nu_1, \nu_2, \rho_1, \rho_2) \in \mathbb{R}^4$ be such that $\rho_1 + \rho_2 > 0$. Suppose that A is $(\nu_1 S, \rho_1 S^{-1})$ -semimonotone and C is $(\nu_2 S, \rho_2 S^{-1})$ -semimonotone. Then, $A + C$ is $((\nu_1 + \nu_2)S, \frac{\rho_1 \rho_2}{\rho_1 + \rho_2} S^{-1})$ -semimonotone.*

Proof. For every $((x, u_1), (y, v_1)) \in (\text{gra } A)^2$ and every $((x, u_2), (y, v_2)) \in (\text{gra } C)^2$, we have

$$\begin{aligned} \langle u_1 + u_2 - (v_1 + v_2) \mid x - y \rangle &= \langle u_1 - v_1 \mid x - y \rangle + \langle u_2 - v_2 \mid x - y \rangle \\ &\geq \nu_1 \|x - y\|_S^2 + \rho_1 \|u_1 - v_1\|_{S^{-1}}^2 + \nu_2 \|x - y\|_S^2 + \rho_2 \|u_2 - v_2\|_{S^{-1}}^2 \\ &\geq (\nu_1 + \nu_2) \|x - y\|_S^2 + \frac{\rho_1 \rho_2}{\rho_1 + \rho_2} \|(u_1 + u_2) - (v_1 + v_2)\|_{S^{-1}}^2, \end{aligned}$$

where the last inequality stems from Lemma 2.1. $\square$

Lemma 2.3. *Let $S \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be self-adjoint and strongly monotone, let $A: \mathcal{H} \rightarrow \mathcal{H}$, and let $C: \mathcal{H} \rightarrow \mathcal{H}$. Let $(\rho_1, \rho_2) \in]0, +\infty[\times \mathbb{R}$ be such that $\rho_1 + \rho_2 > 0$. Suppose that A is ρ_1 -cocoercive with respect to S , and C is $\rho_2 S^{-1}$ -comonotone and weakly continuous. Then, $\text{gra}(A + C)$ is sequentially weak-strong closed.*

Proof. Let $(x, u) \in \mathcal{H}^2$ and let $(x_n)_{n \in \mathbb{N}}$ be a sequence of $\mathcal{H}$ such that $x_n \rightharpoonup x$ and $(A + C)x_n \rightarrow u$. For every $(n, m) \in \mathbb{N}^2$,

$$\begin{aligned} &\langle (A + C)x_n - (A + C)x_m \mid x_n - x_m \rangle \\ &= \langle Ax_n - Ax_m \mid x_n - x_m \rangle + \langle Cx_n - Cx_m \mid x_n - x_m \rangle \\ &\geq \rho_1 \|Ax_n - Ax_m\|_{S^{-1}}^2 + \rho_2 \|Cx_n - Cx_m\|_{S^{-1}}^2 \\ &= \rho_1 (\|(A + C)x_n - (A + C)x_m\|_{S^{-1}}^2 - 2 \langle S^{-1}((A + C)x_n - (A + C)x_m) \mid Cx_n - Cx_m \rangle) \\ &\quad + (\rho_1 + \rho_2) \|Cx_n - Cx_m\|_{S^{-1}}^2. \end{aligned} \quad (2.5)$$

The weak convergence of $(x_n)_{n \in \mathbb{N}}$ implies that this sequence is bounded, whereas the strong convergence of $((A + C)x_n)_{n \in \mathbb{N}}$ implies that $(A + C)x_n - (A + C)x_m \rightarrow 0$, as $n \rightarrow +\infty$ and $m \rightarrow +\infty$. Consequently,

$$\langle (A + C)x_n - (A + C)x_m \mid x_n - x_m \rangle \rightarrow 0 \quad \text{and} \quad \|(A + C)x_n - (A + C)x_m\|_{S^{-1}} \rightarrow 0.$$

In addition, since C is weakly continuous, $Cx_n \rightharpoonup Cx$ and

$$\begin{aligned} &|\langle S^{-1}((A + C)x_n - (A + C)x_m) \mid Cx_n - Cx_m \rangle| \\ &\leq \|S^{-1}\| \|(A + C)x_n - (A + C)x_m\| \|Cx_n - Cx_m\| \rightarrow 0. \end{aligned}$$

Using the positivity of $\rho_1 + \rho_2$, it follows from (2.5) that $\|Cx_n - Cx_m\|_{S^{-1}} \rightarrow 0$. Because the norm induced by S^{-1} is equivalent to the original norm, $(Cx_n)_{n \in \mathbb{N}}$ is a Cauchy sequence.

Therefore, it converges strongly and since $Cx_n \rightharpoonup Cx$, its strong limit is Cx . It follows that $Ax_n \rightarrow u - Cx$. By using now the cocoercivity of A ,

$$\rho_1 \|Ax_n - Ax\|_{S^{-1}}^2 \leq \langle Ax_n - Ax \mid x_n - x \rangle. \quad (2.6)$$

Since $Ax_n - Ax \rightarrow u - Cx - Ax$ and $x_n \rightharpoonup x$, $\langle Ax_n - Ax \mid x_n - x \rangle \rightarrow 0$, which implies that $Ax_n \rightarrow Ax$. In conclusion $u = (A + C)x$, which proves that $\text{gra}(A + C)$ is sequentially weak-strong closed. $\square$

3. NONLINEAR FORWARD-BACKWARD FOR COMONOTONE OPERATORS

We first focus on the case when $B = 0$ and $D = 0$ in Problem 1.1. In this specific case, the dual variable in Problem 1.1 vanishes. Hence, the problem reduces to solving a primal inclusion and it can be addressed efficiently, for example, by the algorithms in [10, 33, 41, 48, 50]. We investigate this primal problem under the assumption that the sum of the operators involved is comonotone. For notational convenience, and for reasons that will become clear later, we use bold symbols for the variables in this section.

Let us list our standing assumptions on the involved operators.

Assumption 3.1. *Let $\mathcal{H}$ be the underlying real Hilbert space. Let $\mathbf{S} \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be a self-adjoint and strongly monotone operator, let $\mathbf{A}: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be a set-valued operator, let $\mathbf{C}: \mathcal{H} \rightarrow \mathcal{H}$ be a β -cocoercive operator with respect to $\mathbf{S}$, for some $\beta \in]0, +\infty[$, and let $\mathbf{M}: \mathcal{H} \rightarrow \mathcal{H}$ be a single-valued operator. We suppose that*

- (i) $(\mathbf{M} + \mathbf{A})^{-1}$ is single-valued and has full domain;
- (ii) $\text{gra}(\mathbf{A} + \mathbf{C})$ is sequentially weak-strong closed;
- (iii) $\mathbf{A}$ is $(\rho\mathbf{S}^{-1})$ -comonotone on $\mathbf{S} = \{(x, -Cx) \in \mathcal{H}^2 \mid x \in \text{zer}(\mathbf{A} + \mathbf{C})\}$ for some $\rho \in]-\beta, +\infty[$.
- (iv) There exist $\zeta \in [0, 1/2[$ and $\tau \in]\tau_\zeta, +\infty[$ such that $\tau\mathbf{M} - \mathbf{S}$ is ζ -Lipschitzian with respect to $\mathbf{S}$ where

$$\hat{\rho} = \min\{\rho, 0\} \quad \text{and} \quad \tau_\zeta = -\frac{\zeta\hat{\rho}}{\beta + \hat{\rho}}.$$

Let us make some comments on these assumptions.

Remark 3.2.

- (i) Since $\mathbf{S}$ is nonexpansive with respect to $\mathbf{S}$, it follows from Assumption 3.1(iv) that $\mathbf{M} = (\tau\mathbf{M} - \mathbf{S} + \mathbf{S})/\tau$ is $\zeta_{\mathbf{M}}$ -Lipschitzian with respect to $\mathbf{S}$, its Lipschitz modulus being $\zeta_{\mathbf{M}} = \xi/\tau$ and $\xi \in [1 - \zeta, 1 + \zeta]$.
- (ii) Under Assumption 3.1(iv), if $\mathbf{A}$ is maximally monotone, then $(\mathbf{M} + \mathbf{A})^{-1}$ is single-valued and has full domain (see [48, Proposition 3.1]). Weaker conditions guaranteeing this assumption can be found in [10, Proposition 3.8 & Proposition 3.9].
- (iii) Let $\mathbf{A} = \tilde{\mathbf{A}} + \tilde{\mathbf{B}}$ with $\tilde{\mathbf{A}}: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ maximally $(-\rho)$ -cohyppomonotone for $\rho < 0$, $\tilde{\mathbf{B}}: \mathcal{H} \rightarrow \mathcal{H}$ single-valued, and $\mathbf{M} = \text{Id}/\tau - \tilde{\mathbf{B}}$. Then, the operator $(\mathbf{M} + \mathbf{A})^{-1} = (\text{Id}/\tau + \tilde{\mathbf{A}})^{-1}$ is single-valued with full domain when $\rho/\tau > -1$, i.e., $\tau > -\rho$. In this situation, by considering $\mathbf{S} = \text{Id}$, we have $\tau\mathbf{M} - \mathbf{S} = -\tau\tilde{\mathbf{B}}$. Hence, if $\tilde{\mathbf{B}}$ is $\zeta_{\tilde{\mathbf{B}}}$ -Lipschitzian, where $\zeta_{\tilde{\mathbf{B}}} \in]0, -\rho^{-1} \min\{\beta + \rho, \frac{1}{2}\} [$, then Assumptions 3.1(i) and 3.1(iv) hold by choosing $\tau \in]-\rho, (2\zeta_{\tilde{\mathbf{B}}})^{-1} [$.
- (iv) If $\mathbf{C} = \mathbf{0}$, Assumption 3.1(iii) reduces to the so-called weak Minty condition for $\mathbf{A}$ at every point of $\text{zer}\mathbf{A}$ (see for instance [1, 13, 26, 31, 32]).

The following proposition provides conditions on $\mathbf{A}$ and $\mathbf{C}$ for which Assumption 3.1(ii) holds.

Proposition 3.3. *Let $\mathbf{S}$ be a self-adjoint and strongly monotone linear operator, let $\mathbf{A}: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be a $\rho\mathbf{S}^{-1}$ -comonotone operator for some $\rho \in \mathbb{R}$ and let $\mathbf{C}: \mathcal{H} \rightarrow \mathcal{H}$ be a β -cocoercive operator with respect to $\mathbf{S}$ for some $\beta \in]0, +\infty[$. The following statements hold:*

- (i) *If $\rho\|\mathbf{S}^{-1}\| > -1$, $\mathbf{T} := \mathbf{Id} - J_{\mathbf{A}}$ is single-valued on $\text{ran}(\mathbf{Id} + \mathbf{A})$. Moreover, if $\text{ran}(\mathbf{Id} + \mathbf{A})$ is weakly sequentially closed, $\mathbf{T}$ is demiclosed.*
- (ii) *If $\beta + \rho > 0$, then $(\mathbf{A} + \mathbf{C})$ is $(\frac{\beta\rho}{\beta+\rho}\mathbf{S}^{-1})$ -comonotone.*
- (iii) *Suppose that $\beta + \rho > -\beta\rho\|\mathbf{S}^{-1}\|$ and $\text{ran}(\mathbf{Id} + \mathbf{A} + \mathbf{C})$ is weakly sequentially closed. Then, $\text{gra}(\mathbf{A} + \mathbf{C})$ is sequentially weak-strong closed.*

Proof. (i) If $\rho \geq 0$, the result follows from [4, Proposition 23.8 & Theorem 4.27]. If $\rho < 0$, since $(\forall \mathbf{u} \in \mathcal{H}) \rho\|\mathbf{u}\|_{\mathbf{S}^{-1}}^2 \geq \rho\|\mathbf{S}^{-1}\|\|\mathbf{u}\|^2$, $\mathbf{A}$ is $\rho\|\mathbf{S}^{-1}\|$ -comonotone. Since $\rho\|\mathbf{S}^{-1}\| > -1$, $J_{\mathbf{A}}$ is single-valued on $\text{ran}(\mathbf{Id} + \mathbf{A})$ and α -conically nonexpansive for $\alpha = \frac{1}{2(\rho\|\mathbf{S}^{-1}\|+1)}$ [5, Proposition 3.6]. Let $\mathbf{N}: \text{ran}(\mathbf{Id} + \mathbf{A}) \rightarrow \mathcal{H}$ be a nonexpansive operator such that $J_{\mathbf{A}} = (1 - \alpha)\mathbf{Id} + \alpha\mathbf{N}$, i.e., $\mathbf{T} = \alpha(\mathbf{Id} - \mathbf{N})$. Let $(\mathbf{x}_n)_{n \in \mathbb{N}}$ be a sequence in $\text{ran}(\mathbf{Id} + \mathbf{A})$ such that $\mathbf{x}_n \rightharpoonup \mathbf{x}$ and $\mathbf{T}\mathbf{x}_n \rightarrow \mathbf{u}$. Since $\text{ran}(\mathbf{Id} + \mathbf{A})$ is weakly sequentially closed and $\mathbf{x}_n \rightharpoonup \mathbf{x}$, we have $\mathbf{x} \in \text{ran}(\mathbf{Id} + \mathbf{A})$. Thus, it follows from the nonexpansiveness of $\mathbf{N}$ that

$$\begin{aligned}
& \|\mathbf{T}\mathbf{x} - \mathbf{u}\|^2 \\
&= \|\alpha(\mathbf{Id} - \mathbf{N})\mathbf{x} - \mathbf{u}\|^2 \\
&= \|\alpha(\mathbf{x}_n - \mathbf{N}\mathbf{x}) - \mathbf{u}\|^2 - 2\alpha\langle \mathbf{x}_n - \mathbf{x} \mid \alpha(\mathbf{x} - \mathbf{N}\mathbf{x}) - \mathbf{u} \rangle - \alpha^2\|\mathbf{x}_n - \mathbf{x}\|^2 \\
&= \|\mathbf{T}\mathbf{x}_n - \mathbf{u}\|^2 + \alpha^2\|\mathbf{N}\mathbf{x}_n - \mathbf{N}\mathbf{x}\|^2 + 2\alpha\langle \mathbf{N}\mathbf{x}_n - \mathbf{N}\mathbf{x} \mid \mathbf{T}\mathbf{x}_n - \mathbf{u} \rangle \\
&\quad - 2\alpha\langle \mathbf{x}_n - \mathbf{x} \mid \mathbf{T}\mathbf{x} - \mathbf{u} \rangle - \alpha^2\|\mathbf{x}_n - \mathbf{x}\|^2 \\
&\leq \|\mathbf{T}\mathbf{x}_n - \mathbf{u}\|^2 + 2\alpha\langle \mathbf{N}\mathbf{x}_n - \mathbf{N}\mathbf{x} \mid \mathbf{T}\mathbf{x}_n - \mathbf{u} \rangle - 2\alpha\langle \mathbf{x}_n - \mathbf{x} \mid \mathbf{T}\mathbf{x} - \mathbf{u} \rangle \\
&\leq \|\mathbf{T}\mathbf{x}_n - \mathbf{u}\|^2 + 2\alpha\|\mathbf{N}\mathbf{x}_n - \mathbf{N}\mathbf{x}\|\|\mathbf{T}\mathbf{x}_n - \mathbf{u}\| - 2\alpha\langle \mathbf{x}_n - \mathbf{x} \mid \mathbf{T}\mathbf{x} - \mathbf{u} \rangle \\
&\leq \|\mathbf{T}\mathbf{x}_n - \mathbf{u}\|^2 + 2\alpha\|\mathbf{x}_n - \mathbf{x}\|\|\mathbf{T}\mathbf{x}_n - \mathbf{u}\| - 2\alpha\langle \mathbf{x}_n - \mathbf{x} \mid \mathbf{T}\mathbf{x} - \mathbf{u} \rangle.
\end{aligned}$$

By taking $n \rightarrow +\infty$ and by using the fact that $(\mathbf{x}_n)_{n \in \mathbb{N}}$ is bounded, we conclude that $\mathbf{T}\mathbf{x} = \mathbf{u}$.

- (ii) Suppose that $\beta + \rho > 0$. Since $\mathbf{A}$ is $(0, \rho\mathbf{S}^{-1})$ -semimonotone and $\mathbf{C}$ is $(0, \beta\mathbf{S}^{-1})$ -semimonotone, the result follows directly from Lemma 2.2.
- (iii) We first note that $\beta + \rho > 0$. Indeed, this is immediate when $\rho \geq 0$, whereas, when $\rho < 0$, it follows from $\beta + \rho > -\beta\rho\|\mathbf{S}^{-1}\|$. Let $\bar{\rho} = \beta\rho/(\beta + \rho)$. By (ii), $\mathbf{A} + \mathbf{C}$ is $\bar{\rho}\mathbf{S}^{-1}$ -comonotone. The assumption $\beta + \rho > -\beta\rho\|\mathbf{S}^{-1}\|$ being equivalent to $\bar{\rho}\|\mathbf{S}^{-1}\| > -1$, it follows from (i) that $\mathbf{Id} - J_{\mathbf{A}+\mathbf{C}}$ is demiclosed. Let $(\mathbf{x}_n, \mathbf{u}_n)_{n \in \mathbb{N}}$ be a sequence in $\text{gra}(\mathbf{A} + \mathbf{C})$ such that $\mathbf{x}_n \rightharpoonup \mathbf{x}$ and $\mathbf{u}_n \rightarrow \mathbf{u}$. Note that

$$\mathbf{u}_n \in (\mathbf{A} + \mathbf{C})\mathbf{x}_n \Leftrightarrow (\mathbf{Id} - J_{\mathbf{A}+\mathbf{C}})(\mathbf{x}_n + \mathbf{u}_n) = \mathbf{u}_n. \quad (3.1)$$

Since $\mathbf{x}_n + \mathbf{u}_n \rightharpoonup \mathbf{x} + \mathbf{u}$, we conclude that $(\mathbf{Id} - J_{\mathbf{A}+\mathbf{C}})(\mathbf{x} + \mathbf{u}) = \mathbf{u}$ and $\mathbf{u} \in (\mathbf{A} + \mathbf{C})\mathbf{x}$. $\square$

Remark 3.4.

- (i) *Suppose that, in the standard metric $\mathbf{S} = \mathbf{Id}$, $\mathbf{A}$ is ρ -comonotone for $\rho > -\beta/(1 + \beta)$. It follows from [5, Theorem 2.17] that $\mathbf{A}$ is maximally ρ -comonotone if and only if $\text{ran}(\mathbf{Id} + \mathbf{A}) = \mathcal{H}$. Therefore, if $\mathbf{A} + \mathbf{C}$ is maximally $\bar{\rho}$ -comonotone for $\bar{\rho} \in]-1, \frac{\beta\rho}{\beta+\rho}]$, $\text{ran}(\mathbf{Id} + \mathbf{A} + \mathbf{C})$ is weakly sequentially closed and $\text{gra}(\mathbf{A} + \mathbf{C})$ is sequentially weak-strong closed.*

- (ii) *Additional conditions ensuring the single-valuedness of $(\mathbf{M} + \mathbf{A})^{-1}$ and the sequential weak-strong closedness of $\text{gra } \mathbf{A}$ for comonotone operators can be found in [49, Lemma 1 & Proposition 3].*

We now focus our attention on the following problem.

Problem 3.5. *Suppose that Assumption 3.1 holds. We want to*

$$\text{find } \mathbf{x} \in \mathcal{H} \text{ such that } \mathbf{0} \in (\mathbf{A} + \mathbf{C})\mathbf{x}, \quad (3.2)$$

under the assumption that its solution set, denoted by $\tilde{\mathbf{Z}}$, is nonempty.

The next proposition provides a sufficient condition for the solution to Problem 3.5 to be unique.

Proposition 3.6. *In the context of Problem 3.5, suppose additionally that $\mathbf{A}$ is $(\mu\mathbf{S}, \rho\mathbf{S}^{-1})$ -semimonotone on $\mathbf{S}$ for some $\mu \in]0, +\infty[$. Then, the solution set of Problem 3.5 is a singleton.*

Proof. Let $(\mathbf{x}, \mathbf{y}) \in \tilde{\mathbf{Z}}^2$. We have that $-\mathbf{C}\mathbf{x} \in \mathbf{A}\mathbf{x}$ and $-\mathbf{C}\mathbf{y} \in \mathbf{A}\mathbf{y}$. Then, by the $(\mu\mathbf{S}, \rho\mathbf{S}^{-1})$ -semimonotonicity of $\mathbf{A}$ on $\mathbf{S}$, we have

$$-\langle \mathbf{x} - \mathbf{y} \mid \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{y} \rangle \geq \mu \|\mathbf{x} - \mathbf{y}\|_{\mathbf{S}}^2 + \rho \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{y}\|_{\mathbf{S}^{-1}}^2. \quad (3.3)$$

Moreover, since $\mathbf{C}$ is β -cocoercive with respect to $\mathbf{S}$, we have that

$$\langle \mathbf{x} - \mathbf{y} \mid \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{y} \rangle \geq \beta \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{y}\|_{\mathbf{S}^{-1}}^2. \quad (3.4)$$

Summing up (3.3) and (3.4), we obtain

$$0 \geq \mu \|\mathbf{x} - \mathbf{y}\|_{\mathbf{S}}^2 + (\rho + \beta) \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{y}\|_{\mathbf{S}^{-1}}^2. \quad (3.5)$$

Since $\rho + \beta > 0$, we deduce that $\mathbf{x} = \mathbf{y}$. $\square$

Now, we introduce the algorithm studied in this section for solving Problem 3.5. This algorithm was initially proposed in [48] for solving monotone problems. We investigate its convergence in the absence of monotonicity by using Assumption 3.1(i) to define a warped resolvent of $\mathbf{A}$.

Algorithm 3.7. *In the context of Problem 3.5 and Assumption 3.1, let $(\mathbf{x}_0, \mathbf{u}_0) \in \mathcal{H}^2$, let $\theta \in]0, 2[$, and consider the following recurrence:*

$$\begin{aligned} (\forall n \in \mathbb{N}) \quad & \begin{cases} \mathbf{p}_{n+1} = (\mathbf{M} + \mathbf{A})^{-1}(\mathbf{M}\mathbf{x}_n - \mathbf{C}\mathbf{x}_n + \mathbf{u}_n/\tau) & (3.6a) \\ \mathbf{u}_{n+1} = (\tau\mathbf{M} - \mathbf{S})\mathbf{p}_{n+1} - (\tau\mathbf{M} - \mathbf{S})\mathbf{x}_n & (3.6b) \\ \mathbf{x}_{n+1} = (1 - \theta)\mathbf{x}_n + \theta\mathbf{p}_{n+1} & (3.6c) \end{cases} \end{aligned}$$

The next proposition provides a key estimate for deriving the convergence of Algorithm 3.7.

Proposition 3.8. *In the context of Problem 3.5, let $\zeta_{\mathbf{M}} \in]0, +\infty[$ be the Lipschitz modulus of $\mathbf{M}$ with respect to $\mathbf{S}$, let $(\mathbf{x}_0, \mathbf{u}_0) \in \mathcal{H}^2$, let $(\mathbf{p}_n, \mathbf{u}_n, \mathbf{x}_n)_{n \geq 1}$ be the sequence generated by Algorithm 3.7. Let $\mathbf{x} \in \text{zer}(\mathbf{A} + \mathbf{C})$, and define*

$$\eta = \begin{cases} 1, & \text{if } \theta \in [1, 2[\text{ and } (\mathbf{S} - \tau\mathbf{M}) \text{ is monotone,} \\ 1 + |1 - \theta|, & \text{otherwise,} \end{cases} \quad (3.7)$$

$$\lambda = 2 - \theta - 2\eta\zeta - \frac{(\tau + 2\hat{\rho})^2}{2\tau(\beta + \hat{\rho}(1 + \tau^{-1}\zeta))} + 2\hat{\rho}\tau \left((\tau^{-1}\zeta + \zeta_{\mathbf{M}})^2 + 4\tau^{-1}\zeta \right), \quad (3.8)$$

$$\lambda_1 = \zeta(2\eta - 1 - 2\hat{\rho}(\tau^{-1}\zeta + \zeta_{\mathbf{M}} + 2)), \quad (3.9)$$

$$(\forall n \in \mathbb{N}) \quad a_{n+1} = \|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 + 2\theta \langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle + \theta\lambda_1 \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2. \quad (3.10)$$

Then, the following assertions hold:

(i) For every $n \in \mathbb{N} \setminus \{0\}$,

$$a_{n+1} \leq a_n - \theta \lambda \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2. \quad (3.11)$$

(ii) $(a_n)_{n \geq 1}$ is a nonnegative sequence such that

$$(\forall n \in \mathbb{N} \setminus \{0\}) \quad a_n \geq (1 - \zeta \theta) \|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2. \quad (3.12)$$

Proof. First note that

$$\tau > \tau_\zeta = -\frac{\zeta \hat{\rho}}{\beta + \hat{\rho}} \Leftrightarrow \beta \tau + \hat{\rho}(\tau + \zeta) > 0, \quad (3.13)$$

which ensures that λ is well defined. In addition, if $\zeta = 0$, then, for every $n \in \mathbb{N}$, $\mathbf{u}_{n+1} = (\tau \mathbf{M} - \mathbf{S})\mathbf{p}_{n+1} - (\tau \mathbf{M} - \mathbf{S})\mathbf{x}_n = \mathbf{0}$ and $\lambda_1 = 0$. The decay property (3.11) then follows directly from the arguments below after removing all terms involving $\mathbf{u}_n$. Moreover, for every $n \in \mathbb{N} \setminus \{0\}$, $a_n = \|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2$, so (3.12) holds immediately. In the remainder of the proof, we therefore assume that $\zeta > 0$.

(i) Fix $n \in \mathbb{N} \setminus \{0\}$. It follows from (3.6a) that

$$\tau \mathbf{M} \mathbf{x}_n - \tau \mathbf{C} \mathbf{x}_n + \mathbf{u}_n - \tau \mathbf{M} \mathbf{p}_{n+1} \in \tau \mathbf{A} \mathbf{p}_{n+1}, \quad (3.14)$$

which, by using (3.6b), is equivalent to

$$\mathbf{S} \mathbf{x}_n + \mathbf{u}_n - (\mathbf{S} \mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau \mathbf{C} \mathbf{x}_n \in \tau \mathbf{A} \mathbf{p}_{n+1}. \quad (3.15)$$

Since $\mathbf{x} \in \text{zer}(\mathbf{A} + \mathbf{C})$, $-\mathbf{C} \mathbf{x} \in \mathbf{A} \mathbf{x}$. By the $\rho \mathbf{S}^{-1}$ -comonotonicity of $\mathbf{A}$ on $\mathbf{S}$ (Assumption 3.1(iii)), we deduce that

$$\begin{aligned} 0 &\leq 2 \langle \mathbf{S} \mathbf{x}_n + \mathbf{u}_n - (\mathbf{S} \mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau \mathbf{C} \mathbf{x}_n + \tau \mathbf{C} \mathbf{x} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\ &\quad - \frac{2\rho}{\tau} \|\mathbf{S} \mathbf{x}_n + \mathbf{u}_n - (\mathbf{S} \mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau \mathbf{C} \mathbf{x}_n + \tau \mathbf{C} \mathbf{x}\|_{\mathbf{S}^{-1}}^2, \end{aligned} \quad (3.16)$$

that is

$$\begin{aligned} 0 &\leq 2 \langle \mathbf{x}_n - \mathbf{p}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle_{\mathbf{S}} + 2 \langle \mathbf{u}_n - \mathbf{u}_{n+1} - \tau \mathbf{C} \mathbf{x}_n + \tau \mathbf{C} \mathbf{x} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\ &\quad - \frac{2\rho}{\tau} \|\mathbf{S} \mathbf{x}_n + \mathbf{u}_n - (\mathbf{S} \mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau \mathbf{C} \mathbf{x}_n + \tau \mathbf{C} \mathbf{x}\|_{\mathbf{S}^{-1}}^2. \end{aligned} \quad (3.17)$$

Note that

$$2 \langle \mathbf{x}_n - \mathbf{p}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle_{\mathbf{S}} = \|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2 - \|\mathbf{x}_n - \mathbf{p}_{n+1}\|_{\mathbf{S}}^2 - \|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2. \quad (3.18)$$

Moreover, it follows from (3.6c) that

$$\begin{aligned} &2 \langle \mathbf{u}_n - \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\ &= -2 \langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle + 2 \langle \mathbf{u}_n \mid \mathbf{p}_n - \mathbf{x} \rangle + 2 \langle \mathbf{u}_n \mid \mathbf{p}_{n+1} - \mathbf{x}_n \rangle + 2 \langle \mathbf{u}_n \mid \mathbf{x}_n - \mathbf{p}_n \rangle \\ &= -2 \langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle + 2 \langle \mathbf{u}_n \mid \mathbf{p}_n - \mathbf{x} \rangle + 2 \langle \mathbf{u}_n \mid \mathbf{p}_{n+1} - \mathbf{x}_n \rangle + 2(1 - \theta) \langle \mathbf{u}_n \mid \mathbf{x}_{n-1} - \mathbf{p}_n \rangle. \end{aligned} \quad (3.19)$$

From the ζ -Lipschitz continuity of $\tau \mathbf{M} - \mathbf{S}$ with respect to $\mathbf{S}$, we deduce that

$$\begin{aligned} \|\mathbf{u}_n\|_{\mathbf{S}^{-1}} &= \|(\tau \mathbf{M} - \mathbf{S})\mathbf{p}_n - (\tau \mathbf{M} - \mathbf{S})\mathbf{x}_{n-1}\|_{\mathbf{S}^{-1}} \\ &\leq \zeta \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathbf{S}}. \end{aligned} \quad (3.20)$$

Hence, the third term on the right-hand side of equality (3.19) can be upper bounded as

$$\begin{aligned} 2\langle \mathbf{u}_n \mid \mathbf{p}_{n+1} - \mathbf{x}_n \rangle &\leq \frac{1}{\zeta} \|\mathbf{u}_n\|_{\mathcal{S}^{-1}}^2 + \zeta \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \\ &\leq \zeta \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}}^2 + \zeta \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2. \end{aligned} \quad (3.21)$$

For the last term on the right-hand side of (3.19), it follows from (3.7) and (3.20) that

$$\begin{aligned} 2(1-\theta)\langle \mathbf{u}_n \mid \mathbf{x}_{n-1} - \mathbf{p}_n \rangle &= 2(\theta-1)\langle (\tau\mathbf{M} - \mathbf{S})\mathbf{p}_n - (\tau\mathbf{M} - \mathbf{S})\mathbf{x}_{n-1} \mid \mathbf{p}_n - \mathbf{x}_{n-1} \rangle \\ &\leq 2(\eta-1)\zeta \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}}^2. \end{aligned} \quad (3.22)$$

Since $\hat{\rho} \leq \rho$, combining (3.17)-(3.22), we obtain

$$\begin{aligned} &\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 2\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\ &\leq \|\mathbf{x}_n - \mathbf{x}\|_{\mathcal{S}}^2 + 2\langle \mathbf{u}_n \mid \mathbf{p}_n - \mathbf{x} \rangle + (2\eta-1)\zeta \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}}^2 - (1-\zeta) \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \\ &\quad - 2\tau \langle \mathbf{p}_{n+1} - \mathbf{x} \mid \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \rangle - \frac{2\hat{\rho}}{\tau} \|\mathbf{S}\mathbf{x}_n + \mathbf{u}_n - (\mathbf{S}\mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau\mathbf{C}\mathbf{x}_n + \tau\mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2. \end{aligned} \quad (3.23)$$

Let us bound the last term in (3.23). First note that

$$\begin{aligned} &\|\mathbf{S}\mathbf{x}_n + \mathbf{u}_n - (\mathbf{S}\mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau\mathbf{C}\mathbf{x}_n + \tau\mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2 \\ &= \|\mathbf{S}(\mathbf{x}_n - \mathbf{p}_{n+1}) + \mathbf{u}_n - \mathbf{u}_{n+1}\|_{\mathcal{S}^{-1}}^2 - 2\tau \langle \mathbf{x}_n - \mathbf{p}_{n+1} \mid \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \rangle \\ &\quad - 2\tau \langle \mathbf{u}_n - \mathbf{u}_{n+1} \mid \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \rangle_{\mathcal{S}^{-1}} + \tau^2 \|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2. \end{aligned} \quad (3.24)$$

Equation (3.6b) yields

$$\begin{aligned} &\|\mathbf{S}(\mathbf{x}_n - \mathbf{p}_{n+1}) + \mathbf{u}_n - \mathbf{u}_{n+1}\|_{\mathcal{S}^{-1}}^2 \\ &= \|(\tau\mathbf{M}\mathbf{p}_{n+1} - \tau\mathbf{M}\mathbf{x}_n) - ((\tau\mathbf{M} - \mathbf{S})\mathbf{p}_n - (\tau\mathbf{M} - \mathbf{S})\mathbf{x}_{n-1})\|_{\mathcal{S}^{-1}}^2 \\ &\leq (\tau\zeta_M \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}} + \zeta \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}})^2 \\ &\leq \tau^2 \zeta_M^2 \left(1 + \frac{\zeta}{\zeta_M \tau}\right) \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 + \zeta^2 \left(1 + \frac{\zeta_M \tau}{\zeta}\right) \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}}^2 \\ &= \tau\zeta_M (\tau\zeta_M + \zeta) \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 + \zeta (\zeta + \zeta_M \tau) \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}}^2. \end{aligned} \quad (3.25)$$

In addition, by the Cauchy-Schwarz and Young inequalities, (3.20) leads to

$$\begin{aligned} -2\tau \langle \mathbf{u}_n - \mathbf{u}_{n+1} \mid \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \rangle_{\mathcal{S}^{-1}} &\leq 2\tau \|\mathbf{u}_n - \mathbf{u}_{n+1}\|_{\mathcal{S}^{-1}} \|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}} \\ &\leq \frac{\tau}{\zeta} \|\mathbf{u}_n - \mathbf{u}_{n+1}\|_{\mathcal{S}^{-1}}^2 + \zeta \tau \|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2 \\ &\leq \frac{2\tau}{\zeta} (\|\mathbf{u}_n\|_{\mathcal{S}^{-1}}^2 + \|\mathbf{u}_{n+1}\|_{\mathcal{S}^{-1}}^2) + \zeta \tau \|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2 \\ &\leq 2\zeta \tau (\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 + \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}}^2) + \zeta \tau \|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2. \end{aligned} \quad (3.26)$$

Therefore, (3.24)-(3.26) allow us to deduce from (3.23) that

$$\begin{aligned} &\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 2\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\ &\leq \|\mathbf{x}_n - \mathbf{x}\|_{\mathcal{S}}^2 + 2\langle \mathbf{u}_n \mid \mathbf{p}_n - \mathbf{x} \rangle + \lambda_1 \|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathcal{S}}^2 \\ &\quad - (1 - \zeta + 2\hat{\rho}(2\zeta + \zeta_M(\tau\zeta_M + \zeta))) \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 - 2\hat{\rho}(\tau + \zeta) \|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2 \\ &\quad - 2\tau \langle \mathbf{p}_{n+1} - \mathbf{x} \mid \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \rangle + 4\hat{\rho} \langle \mathbf{x}_n - \mathbf{p}_{n+1} \mid \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \rangle. \end{aligned} \quad (3.27)$$

Using the β -cocoercivity of $\mathbf{C}$, the last two terms of (3.27) can be upper-bounded as follows:

$$\begin{aligned}
& -2\tau\langle \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle - 4\hat{\rho}\langle \mathbf{p}_{n+1} - \mathbf{x}_n \mid \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \rangle \\
& = -2(\tau + 2\hat{\rho})\langle \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \mid \mathbf{p}_{n+1} - \mathbf{x}_n \rangle - 2\tau\langle \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \mid \mathbf{x}_n - \mathbf{x} \rangle \\
& \leq 2|\tau + 2\hat{\rho}|\|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathbf{S}^{-1}}\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}} - 2\tau\langle \mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x} \mid \mathbf{x}_n - \mathbf{x} \rangle \\
& \leq \varepsilon|\tau + 2\hat{\rho}|\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 + \frac{|\tau + 2\hat{\rho}|}{\varepsilon}\|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathbf{S}^{-1}}^2 - 2\beta\tau\|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathbf{S}^{-1}}^2 \\
& = \varepsilon|\tau + 2\hat{\rho}|\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 + \frac{1}{\varepsilon}(|\tau + 2\hat{\rho}| - 2\beta\tau\varepsilon)\|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathbf{S}^{-1}}^2. \tag{3.28}
\end{aligned}$$

Then, combining (3.27) and (3.28) leads to

$$\begin{aligned}
& \|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 + 2\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\
& \leq \|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2 + 2\langle \mathbf{u}_n \mid \mathbf{p}_n - \mathbf{x} \rangle + \lambda_1\|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathbf{S}}^2 \\
& \quad - (1 - \varepsilon|\tau + 2\hat{\rho}| - \zeta + 2\hat{\rho}(2\zeta + \zeta_M(\tau\zeta_M + \zeta)))\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
& \quad + \frac{1}{\varepsilon}(|\tau + 2\hat{\rho}| - 2\varepsilon(\beta\tau + \hat{\rho}(\tau + \zeta)))\|\mathbf{C}\mathbf{x}_n - \mathbf{C}\mathbf{x}\|_{\mathbf{S}^{-1}}^2. \tag{3.29}
\end{aligned}$$

According to Assumption 3.1(iv), $\tau > \tau_{\zeta} \Rightarrow \beta\tau + \hat{\rho}(\tau + \zeta) > 0$. Thus, to eliminate the dependence on $\|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}_n\|_{\mathbf{S}^{-1}}^2$, we set $\varepsilon = \frac{|\tau + 2\hat{\rho}|}{2\beta\tau + 2\hat{\rho}(\tau + \zeta)}$ in the resulting bound when $\tau + 2\hat{\rho} \neq 0$, leading to

$$\begin{aligned}
& \|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 + 2\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\
& \leq \|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2 + 2\langle \mathbf{u}_n \mid \mathbf{p}_n - \mathbf{x} \rangle + \lambda_1\|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathbf{S}}^2 \\
& \quad - \left(1 - \frac{(\tau + 2\hat{\rho})^2}{2\beta\tau + 2\hat{\rho}(\tau + \zeta)} - \zeta + 2\hat{\rho}(2\zeta + \zeta_M(\tau\zeta_M + \zeta))\right)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2. \tag{3.30}
\end{aligned}$$

It follows from (3.29) that (3.30) also holds if $\tau + 2\hat{\rho} = 0$. In addition, (3.6c) yields

$$\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 = (1 - \theta)\|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2 + \theta\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 - \theta(1 - \theta)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2. \tag{3.31}$$

Therefore, using (3.30) and (3.31), and adding $\theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2$, we conclude that

$$\begin{aligned}
& \|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 + 2\theta\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
& \leq \|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2 + 2\theta\langle \mathbf{u}_n \mid \mathbf{p}_n - \mathbf{x} \rangle + \theta\lambda_1\|\mathbf{p}_n - \mathbf{x}_{n-1}\|_{\mathbf{S}}^2 \\
& \quad - \theta\left(2 - \theta - 2\eta\zeta - \frac{(\tau + 2\hat{\rho})^2}{2\beta\tau + 2\hat{\rho}(\tau + \zeta)} + 2\hat{\rho}(2\zeta + \zeta_M(\tau\zeta_M + \zeta))\right)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \tag{3.32}
\end{aligned}$$

and the result follows.

(ii) Let us now prove the second assertion. We have

$$\begin{aligned}
a_{n+1} &= \|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 + 2\theta\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
&= \|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 + 2\theta\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x}_{n+1} \rangle + 2\theta\langle \mathbf{u}_{n+1} \mid \mathbf{x}_{n+1} - \mathbf{x} \rangle + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
&= \|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 + 2\theta(1-\theta)\langle \mathbf{u}_{n+1} \mid \mathbf{p}_{n+1} - \mathbf{x}_n \rangle + 2\theta\langle \mathbf{u}_{n+1} \mid \mathbf{x}_{n+1} - \mathbf{x} \rangle + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
&\geq \|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 - 2\theta(\eta-1)\zeta\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 - \theta(\zeta^{-1}\|\mathbf{u}_{n+1}\|_{\mathbf{S}^{-1}}^2 + \zeta\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2) \\
&\quad + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
&\geq (1-\theta\zeta)\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 - \theta((2\eta-1)\zeta - \lambda_1)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
&= (1-\theta\zeta)\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 - 2\theta\hat{\rho}\zeta(\tau^{-1}\zeta + \zeta_{\mathbf{M}} + 2)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 \\
&\geq (1-\theta\zeta)\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathbf{S}}^2 \geq 0,
\end{aligned}$$

where the third equality stems from (3.6c), the first inequality from (3.22), and the second one from (3.20). The positivity of $(1-\theta\zeta)$ is ensured since $\zeta < 1/2$ and $\theta < 2$. $\square$

Remark 3.9. When $\theta \in [1, 2[$ and $\mathbf{S} - \tau\mathbf{M}$ is monotone, one can take $\eta = 1$ instead of $\eta = \theta$. This increases λ by $2(\theta-1)\zeta$, thereby making the condition $\lambda > 0$ less restrictive.

The theorem below establishes the convergence of Algorithm 3.7 to a solution to Problem 3.5.

Theorem 3.10. In the context of Problem 3.5, let $(\mathbf{x}_0, \mathbf{u}_0) \in \mathcal{H}^2$ and consider the sequence $(\mathbf{x}_n, \mathbf{u}_n)_{n \geq 1}$ generated by Algorithm 3.7. Suppose that λ defined in (3.8) is positive. Then the following hold.

- (i) $(\mathbf{x}_n)_{n \in \mathbb{N}}$ converges weakly to a solution to Problem 3.5.
- (ii) If $\mathbf{A}$ is $(\mu\mathbf{S}, \rho\mathbf{S}^{-1})$ -semimonotone on $\mathbf{S}$ for some $\mu \in]0, +\infty[$, then $(\mathbf{x}_n)_{n \in \mathbb{N}}$ converges R -linearly to the unique solution to Problem 3.5.

Proof. (i) Let us use the same notation as in Proposition 3.8. Since $(a_n)_{n \in \mathbb{N}}$, defined in (3.10), is nonnegative (Proposition 3.8(ii)) and λ defined in (3.8) is assumed to be positive, Proposition 3.8(i) shows that $(a_n)_{n \in \mathbb{N}}$ is nonincreasing, convergent, and $\sum_{n \in \mathbb{N}} \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}^2 < +\infty$. In particular, we deduce from (3.12) that $(\mathbf{x}_n)_{n \in \mathbb{N}}$ is bounded. Since $\mathbf{p}_n - \mathbf{x}_{n-1} \rightarrow \mathbf{0}$, $(\mathbf{p}_n)_{n \in \mathbb{N}}$ is also bounded and it follows from (3.20) that $\mathbf{u}_n \rightarrow \mathbf{0}$. We deduce from these properties and (3.10) that $\lim_{n \rightarrow +\infty} \|\mathbf{x}_n - \mathbf{x}\|_{\mathbf{S}}^2 = \lim_{n \rightarrow +\infty} a_n$ is finite. Additionally, by the Lipschitz property of $(\tau\mathbf{M} - \mathbf{S})$ with respect to $\mathbf{S}$, we have

$$\begin{aligned}
\|\mathbf{S}^{-1}\tau(\mathbf{M}\mathbf{p}_{n+1} - \mathbf{M}\mathbf{x}_n)\|_{\mathbf{S}} &= \|\tau\mathbf{M}\mathbf{p}_{n+1} - \tau\mathbf{M}\mathbf{x}_n\|_{\mathbf{S}^{-1}} \\
&\leq \|(\tau\mathbf{M} - \mathbf{S})\mathbf{p}_{n+1} - (\tau\mathbf{M} - \mathbf{S})\mathbf{x}_n\|_{\mathbf{S}^{-1}} + \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}} \\
&\leq (\zeta + 1)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}.
\end{aligned} \tag{3.33}$$

Since $\tau > 0$ and $\mathbf{S}$ is bounded, $\mathbf{M}\mathbf{p}_{n+1} - \mathbf{M}\mathbf{x}_n \rightarrow \mathbf{0}$. Moreover, the cocoercivity of $\mathbf{C}$ with respect to $\mathbf{S}$ yields

$$\|\mathbf{C}\mathbf{p}_{n+1} - \mathbf{C}\mathbf{x}_n\|_{\mathbf{S}^{-1}} \leq \frac{1}{\beta}\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathbf{S}}. \tag{3.34}$$

Therefore, it follows from (3.33), (3.34) and the facts that $\mathbf{u}_n \rightarrow \mathbf{0}$ and $\mathbf{p}_{n+1} - \mathbf{x}_n \rightarrow \mathbf{0}$ that

$$\mathbf{w}_n := \mathbf{M}\mathbf{x}_n - \mathbf{M}\mathbf{p}_{n+1} + \mathbf{u}_n/\tau + (\mathbf{C}\mathbf{p}_{n+1} - \mathbf{C}\mathbf{x}_n) \rightarrow \mathbf{0}. \tag{3.35}$$

Now, let $(\mathbf{x}_{n_k})_{k \in \mathbb{N}}$ be a weakly convergent subsequence of $(\mathbf{x}_n)_{n \in \mathbb{N}}$. Since $\mathbf{p}_{n+1} - \mathbf{x}_n \rightarrow \mathbf{0}$, $(\mathbf{p}_{n_k+1})_{k \in \mathbb{N}}$ is also weakly convergent and has the same limit. It follows from (3.35) that $\mathbf{w}_{n_k} \rightarrow \mathbf{0}$ and from (3.6a) that

$$(\forall k \in \mathbb{N}) \quad \mathbf{w}_{n_k} \in (\mathbf{A} + \mathbf{C})\mathbf{p}_{n_k+1}. \quad (3.36)$$

Since $\text{gra}(\mathbf{A} + \mathbf{C})$ is sequentially weak-strong closed (Assumption 3.1(ii)), we conclude that every weak cluster point of $(\mathbf{x}_n)_{n \in \mathbb{N}}$ belongs to $\text{zer}(\mathbf{A} + \mathbf{C})$. The result follows from Opial's lemma [4, Lemma 2.47] applied in the Hilbert space $(\mathcal{H}, \langle \cdot | \cdot \rangle_{\mathcal{S}})$.

- (ii) By Proposition 3.6, $\tilde{\mathbf{Z}}$ is a singleton. Now, the $(\mu\mathbf{S}, \rho\mathbf{S}^{-1})$ -semimonotonicity of $\mathbf{A}$ allows us to refine (3.11) as shown below. First, (3.17) becomes

$$\begin{aligned} \tau\mu\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 &\leq \langle \mathbf{S}\mathbf{x}_n + \mathbf{u}_n - (\mathbf{S}\mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau\mathbf{C}\mathbf{x}_n + \tau\mathbf{C}\mathbf{x} \mid \mathbf{p}_{n+1} - \mathbf{x} \rangle \\ &\quad - \frac{\hat{\rho}}{\tau}\|\mathbf{S}\mathbf{x}_n + \mathbf{u}_n - (\mathbf{S}\mathbf{p}_{n+1} + \mathbf{u}_{n+1}) - \tau\mathbf{C}\mathbf{x}_n + \tau\mathbf{C}\mathbf{x}\|_{\mathcal{S}^{-1}}^2. \end{aligned} \quad (3.37)$$

Hence, by proceeding similarly to the proof of Proposition 3.8(i), we obtain

$$a_n \geq 2\tau\mu\theta\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + \theta\lambda\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 + a_{n+1}, \quad (3.38)$$

where a_n is defined in (3.10). The parameter λ_1 defined in (3.9) is nonnegative since $\eta \geq 1$, $\zeta \geq 0$, and $\hat{\rho} \leq 0$. Let us set $\vartheta_1 = \theta \min\{\tau\mu, \frac{\lambda}{3}\}$, and $\vartheta_2 = \vartheta_1 \min\{1, \frac{1}{\lambda_1}\}$ if $\lambda_1 > 0$ and $\vartheta_2 = \vartheta_1$ otherwise. Since $\zeta \in [0, 1/2[$ and $\theta \in]0, 2[$, it follows from (3.6c) and (3.38) that, for every $n \in \mathbb{N}$,

$$\begin{aligned} a_n - a_{n+1} &\geq \vartheta_1 \left(2\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 3\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \right) \\ &\geq \vartheta_1 \left(\frac{1}{2}\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + \|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 3\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 - \|\mathbf{p}_{n+1} - \mathbf{x}_{n+1}\|_{\mathcal{S}}^2 \right) \\ &= \frac{\vartheta_1}{2} \left(\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 2\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 2(3 - (1 - \theta)^2)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \right) \\ &= \frac{\vartheta_1}{2} \left(\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 2(\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2) + 2(1 + 2\theta - \theta^2)\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \right) \\ &\geq \frac{\vartheta_2}{2} \left(\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 2(\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + \|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2) + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \right), \end{aligned} \quad (3.39)$$

where we have used the fact that $2(1 + 2\theta - \theta^2) - \theta = (2 - \theta)(2\theta + 1) > 0$. When $\zeta > 0$, using (3.20) and noting that $\theta\zeta \in]0, 1[$ yield

$$\begin{aligned} a_n - a_{n+1} &\geq \frac{\vartheta_2}{2} \left(\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + \theta\zeta(\|\mathbf{p}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + \zeta^{-2}\|\mathbf{u}_{n+1}\|_{\mathcal{S}^{-1}}^2) + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \right) \\ &\geq \frac{\vartheta_2}{2} \left(\|\mathbf{x}_{n+1} - \mathbf{x}\|_{\mathcal{S}}^2 + 2\theta\langle \mathbf{p}_{n+1} - \mathbf{x} \mid \mathbf{u}_{n+1} \rangle + \theta\lambda_1\|\mathbf{p}_{n+1} - \mathbf{x}_n\|_{\mathcal{S}}^2 \right) \\ &= \frac{\vartheta_2}{2} a_{n+1}. \end{aligned} \quad (3.40)$$

When $\zeta = 0$, $\lambda_1 = 0$ and the inequality $a_n - a_{n+1} \geq (\vartheta_2 a_{n+1})/2$ follows directly from (3.39). Consequently, $(a_n)_{n \in \mathbb{N}}$ converges Q-linearly to 0 with rate $1/(1 + \vartheta_2/2)$ and the R-linear convergence of $(\mathbf{x}_n)_{n \in \mathbb{N}}$ to $\mathbf{x}$ follows from (3.12). $\square$

We now provide further insight into the choice of the step size τ ensuring the convergence of Algorithm 3.7.

Proposition 3.11. *In the context of Problem 3.5, with the notation of Proposition 3.8, the following properties hold.*

- (i) *For fixed ζ , ξ , and η , the convergence condition $\lambda > 0$ is equivalent to the following cubic polynomial inequality:*

$$P(\tau) = -\tau^3 + a_2\tau^2 + a_1\tau + a_0 > 0, \quad (3.41)$$

where

$$a_0 = 4\hat{\rho}^2\zeta(\zeta + \xi)^2, \quad a_1 = 2c\hat{\rho}\zeta - 4\hat{\rho}^2 + 4\hat{\rho}(\beta + \hat{\rho})(\zeta + \xi)^2, \quad a_2 = 2c(\beta + \hat{\rho}) - 4\hat{\rho},$$

with $c = 2 - \theta - 2(\eta - 4\hat{\rho})\zeta$ and $\xi \in [1 - \zeta, 1 + \zeta]$ as defined in Remark 3.2(i).

- (ii) *If (3.41) holds, then $\zeta \in [0, (2 - \theta)/(2(\eta - 4\hat{\rho}))]$.*

- (iii) *If $\theta \in]0, \frac{2}{1 + \sqrt{-\hat{\rho}/\beta}}[$, there exists*

$$\bar{\zeta} \in]0, \min \left\{ \frac{2 - \theta}{2(\eta - 4\hat{\rho})}, \frac{(\beta + \hat{\rho})((2 - \theta)(\beta + \hat{\rho}) - 4\hat{\rho})}{2(\eta - 4\hat{\rho})(\beta + \hat{\rho})^2 - 2\hat{\rho}} \right\} [\quad (3.42)$$

such that, for every $\zeta \in [0, \bar{\zeta}[$ and $\tau \in]\tau_{1,\zeta}, \tau_{2,\zeta}[$, (3.41) holds, where $(\tau_{1,\zeta}, \tau_{2,\zeta}) \in]\tau_{\zeta}, +\infty[^2$ with $\tau_{1,\zeta} < \tau_{2,\zeta}$.¹

Proof. (i) According to (3.8) and Remark 3.2(i), λ can be written as

$$\lambda = \frac{P(\tau)}{2\tau((\beta + \hat{\rho})\tau + \hat{\rho}\zeta)}, \quad (3.43)$$

where

$$\begin{aligned} P(\tau) &= 2\left((2 - \theta - 2\eta\zeta)\tau + 2\hat{\rho}((\zeta + \xi)^2 + 4\tau\zeta)\right)((\beta + \hat{\rho})\tau + \hat{\rho}\zeta) - \tau(\tau + 2\hat{\rho})^2 \\ &= 2\left(\tau c + 2\hat{\rho}(\zeta + \xi)^2\right)((\beta + \hat{\rho})\tau + \hat{\rho}\zeta) - \tau(\tau + 2\hat{\rho})^2 \\ &= -\tau^3 + a_2\tau^2 + a_1\tau + a_0. \end{aligned} \quad (3.44)$$

In view of Assumption 3.1(iv), $(\beta + \hat{\rho})\tau + \hat{\rho}\zeta > 0$, and since $\tau > 0$, the denominator in (3.43) is positive. Hence, $\lambda > 0$ if and only if $P(\tau) > 0$.

- (ii) Assume that $\lambda > 0$. We have $2 - \theta - 2\eta\zeta + \frac{2\hat{\rho}}{\tau}((\zeta + \xi)^2 + 4\tau\zeta) \geq \lambda > 0$, hence $\tau(2 - \theta - 2\eta\zeta + 8\hat{\rho}\zeta) > -2\hat{\rho}(\zeta + \xi)^2$, which implies that $c > 0$. This is equivalent to $\zeta \in [0, (2 - \theta)/(2(\eta - 4\hat{\rho}))]$.
- (iii) Assume that $c > 0$. We have $a_0 \geq 0$, $a_2 > 0$, and $a_1 \leq 0$. Consequently, for every $\tau \in]-\infty, 0[$, $P(\tau) > 0$, which means that all the real roots of P are nonnegative. Note also that $P(\tau_{\zeta}) \leq 0$. Indeed, $\tau_{\zeta} = -\hat{\rho}\zeta/(\beta + \hat{\rho}) \Rightarrow P(\tau_{\zeta}) = -\tau_{\zeta}(\tau_{\zeta} + 2\hat{\rho})^2 \leq 0$. Since $P(0) = a_0 \geq 0$ and $P(\tau_{\zeta}) \leq 0$, P has a root $\tau_{0,\zeta} \in [0, \tau_{\zeta}]$. Two cases may arise: either I) P is nonpositive on $]\tau_{\zeta}, +\infty[$ (in particular, if it is negative over $]\tau_{\zeta}, +\infty[$, then P has all its real roots in $[0, \tau_{\zeta}]$) or II) there exists $\tau \in]\tau_{\zeta}, +\infty[$ such that $P(\tau) > 0$ and P has two roots $\tau_{1,\zeta} \geq \tau_{\zeta}$ and $\tau_{2,\zeta} > \tau_{1,\zeta}$. In the latter case, for every $\tau \in]\tau_{1,\zeta}, \tau_{2,\zeta}[$, $P(\tau) > 0$.

We have, for every $\tau \in \mathbb{R}$,

$$P(\tau) = -\tau^3 + a_2\tau^2 + 2\hat{\rho}(c\zeta - 2\hat{\rho})\tau + 4\hat{\rho}(\zeta + \xi)^2((\beta + \hat{\rho})\tau + \hat{\rho}\zeta). \quad (3.45)$$

Let

$$\tilde{\tau}_{\zeta} = a_2/2 = c(\beta + \hat{\rho}) - 2\hat{\rho}. \quad (3.46)$$

¹The quantities $\tau_{1,\zeta}$ and $\tau_{2,\zeta}$ also depend on ξ and η , but, for notational simplicity, we omit this dependence.

Now, from (3.45),

$$\begin{aligned} P(\tilde{\tau}_\zeta) &= \tilde{\tau}_\zeta^3 + 2\hat{\rho}(c\zeta - 2\hat{\rho})\tilde{\tau}_\zeta + 4\hat{\rho}(\zeta + \xi)^2((\beta + \hat{\rho})\tilde{\tau}_\zeta + \hat{\rho}\zeta) \\ &= \tilde{\tau}_\zeta(\tilde{\tau}_\zeta^2 + 2\hat{\rho}(c\zeta - 2\hat{\rho})) + 4\hat{\rho}(\zeta + \xi)^2(\beta + \hat{\rho})(\tilde{\tau}_\zeta - \underline{\tau}_\zeta) \\ &= \tilde{\tau}_\zeta c(\beta + \hat{\rho})(\tilde{\tau}_\zeta - 2\hat{\rho} - 2\underline{\tau}_\zeta) + 4\hat{\rho}(\zeta + \xi)^2(\beta + \hat{\rho})(\tilde{\tau}_\zeta - \underline{\tau}_\zeta) \\ &= (\beta + \hat{\rho})(\tilde{\tau}_\zeta c(\tilde{\tau}_\zeta - 2\hat{\rho} - 2\underline{\tau}_\zeta) + 4\hat{\rho}(\zeta + \xi)^2(\tilde{\tau}_\zeta - \underline{\tau}_\zeta)). \end{aligned} \quad (3.47)$$

$$(3.48)$$

Suppose that

$$\tilde{\tau}_\zeta > 2(\underline{\tau}_\zeta + \hat{\rho}). \quad (3.49)$$

From the expressions of $\tilde{\tau}_\zeta$ and $\underline{\tau}_\zeta$, the latter inequality is satisfied if and only if

$$\zeta < \zeta_1 = \frac{(\beta + \hat{\rho})((2 - \theta)(\beta + \hat{\rho}) - 4\hat{\rho})}{2(\eta - 4\hat{\rho})(\beta + \hat{\rho})^2 - 2\hat{\rho}}, \quad (3.50)$$

where the upper bound is positive. Note that $\tilde{\tau}_\zeta > \underline{\tau}_\zeta$. Indeed, if $-2\hat{\rho} \geq \underline{\tau}_\zeta$, then $\tilde{\tau}_\zeta = c(\beta + \hat{\rho}) - 2\hat{\rho} > \underline{\tau}_\zeta$. If $-2\hat{\rho} < \underline{\tau}_\zeta$, then, the condition $\tilde{\tau}_\zeta > 2(\underline{\tau}_\zeta + \hat{\rho})$ implies that $\tilde{\tau}_\zeta > \underline{\tau}_\zeta$. Due to this inequality, the last additive term in (3.47) is nonpositive. Thus, (3.49) is a necessary condition for $P(\tilde{\tau}_\zeta)$ to be positive.

Let us now look at what happens when ζ is small enough. Since $\xi \in [1 - \zeta, 1 + \zeta]$, $\xi \rightarrow 1$ as $\zeta \rightarrow 0$. In addition, as $\zeta \rightarrow 0$, $\underline{\tau}_\zeta \rightarrow 0$, $c \rightarrow 2 - \theta$, $\tilde{\tau}_\zeta \rightarrow (2 - \theta)\beta - \theta\hat{\rho} > 0$. Thus, we have

$$\tilde{\tau}_\zeta c(\tilde{\tau}_\zeta - 2\hat{\rho} - 2\underline{\tau}_\zeta) + 4\hat{\rho}(\zeta + \xi)^2(\tilde{\tau}_\zeta - \underline{\tau}_\zeta) \rightarrow ((2 - \theta)\beta - \theta\hat{\rho})((2 - \theta)^2\beta + \theta^2\hat{\rho}).$$

The limit is strictly positive when $\theta \in]0, \frac{2}{1 + \sqrt{-\hat{\rho}/\beta}}[$ (note that $\frac{2}{1 + \sqrt{-\hat{\rho}/\beta}} \in]1, 2]$).

Thus, it follows from (3.47) that there exists $\bar{\zeta} \in]0, +\infty[$, independent of ξ , such that, for every $\zeta < \bar{\zeta}$, $P(\tilde{\tau}_\zeta) > 0$, which ensures that we are in case II) mentioned above. In addition to the condition $\bar{\zeta} < \zeta_1$, it follows from the initial assumption $c > 0$ and (ii) that $\bar{\zeta} \in [0, (2 - \theta)/(2(\eta - 4\hat{\rho}))]$. $\square$

Let us analyze more precisely some noteworthy special cases.

Corollary 3.12. *In the context of Problem 3.5, let η be defined by (3.7), let ξ be defined as Remark 3.2(i), let $(\mathbf{x}_0, \mathbf{u}_0) \in \mathcal{H}^2$ and consider the sequence $(\mathbf{x}_n, \mathbf{u}_n)_{n \geq 1}$ generated by Algorithm 3.7. Consider one of the following settings:*

(i) (standard forward-backward algorithm) $\tau \mathbf{M} = \mathbf{S}$ (i.e., $\zeta = 0$),

$$\bar{\theta} = \frac{2}{1 + \sqrt{-\hat{\rho}/\beta}}, \quad (3.51)$$

$$\tau_{1,0} = (2 - \theta)(\beta + \hat{\rho}) - 2\hat{\rho} - \sqrt{((2 - \theta)(\beta + \hat{\rho}) - 2\hat{\rho})^2 + 4\hat{\rho}\beta}, \quad (3.52)$$

$$\tau_{2,0} = (2 - \theta)(\beta + \hat{\rho}) - 2\hat{\rho} + \sqrt{((2 - \theta)(\beta + \hat{\rho}) - 2\hat{\rho})^2 + 4\hat{\rho}\beta}; \quad (3.53)$$

(ii) (comonotone inclusion) $\mathbf{C} = \mathbf{0}$, $\bar{\theta} = 2$, $\zeta < \bar{\zeta} = \frac{2 - \theta}{2(\eta - 4\hat{\rho})}$, $\tau_{1,\zeta} = -2\hat{\rho}(\zeta + \xi)^2/(2 - \theta - 2(\eta - 4\hat{\rho})\zeta)$ and $\tau_{2,\zeta} = +\infty$;

(iii) (monotone+cocoercive operator splitting) $\mathbf{A}$ is monotone, $\bar{\theta} = 2$, $\zeta < \bar{\zeta} = \frac{2 - \theta}{2\eta}$, $\tau_{1,\zeta} = 0$, and $\tau_{2,\zeta} = 2\beta(2 - \theta - 2\eta\zeta)$.

Then, if $\theta \in]0, \bar{\theta}[$ and $\tau \in]\tau_{1,\zeta}, \tau_{2,\zeta}[$, the same conclusions as in Theorem 3.10 hold. In particular, the R-linear convergence property holds whenever the additional assumption in Theorem 3.10(ii) is satisfied.

Proof. The result follows from Theorem 3.10 and Proposition 3.11(i). Let us use the same notation as in Proposition 3.11.

- (i) Assume that $\tau\mathbf{M} = \mathbf{S}$. Then $\zeta = 0$, $\zeta_M = 1/\tau$, $\xi = 1$, and

$$P(\tau) = \tau(-\tau^2 + 2((2 - \theta)(\beta + \hat{\rho}) - 2\hat{\rho})\tau + 4\hat{\rho}\beta). \quad (3.54)$$

The quadratic factor has two distinct nonnegative roots $\tau_{1,0}$ and $\tau_{2,0}$ when $(2 - \theta)(\beta + \hat{\rho}) - 2\hat{\rho} > 2\sqrt{-\hat{\rho}\beta}$, that is $\theta < \bar{\theta}$. Then $P(\tau)$ is positive if $\tau \in]\tau_{1,0}, \tau_{2,0}[$.

- (ii) Since the zero operator is β -cocoercive for every $\beta > 0$, we may let $\beta \rightarrow +\infty$ in the expression of λ . Thus we deduce from (3.8) that $\lambda > 0$ if and only if

$$\tau c + 2\hat{\rho}(\zeta + \xi)^2 > 0, \quad (3.55)$$

which yields the intervals for ζ and τ ensuring convergence.

- (iii) In the case when $\mathbf{A}$ is monotone, that is $\hat{\rho} = 0$, we have $P(\tau) = \tau^2(2c\beta - \tau)$, which yields the result. □

Remark 3.13.

- (i) The necessary condition for $\lambda > 0$ in Proposition 3.11(ii) defines a subinterval of the one initially considered in Assumption 3.1(iv).
- (ii) Suppose that $\mathbf{S} = \mathbf{Id}$, $\tau\mathbf{M} = \mathbf{Id}$, $\mathbf{u}_0 = \mathbf{0}$, and $\theta = 1$. Then Algorithm 3.7 reduces to the standard unrelaxed FB algorithm. From our standing assumptions, $\tau\mathbf{A}$ is $(-\hat{\rho}/\tau)$ -cohyppomonotone on $\mathbb{S}$. Let us assume that the cohyppomonotonicity property extends to $\mathcal{H}$ to provide a simple interpretations of the conditions in Corollary 3.12(i). The condition $\tau > \tau_{1,0} = -2\hat{\rho}$ guarantees that $J_{\tau\mathbf{A}}$ is an averaged operator [5, Corollary 3.10(iii)], whereas $\tau < \tau_{2,0} = 2\beta$ is the classical condition ensuring the averagedness of $\mathbf{Id} - \tau\mathbf{C}$.
- (iii) In Corollary 3.12(iii), the obtained condition coincides with the assumption made in [48, Theorem 3.1].
- (iv) It follows from (3.42) that, for fixed $\hat{\rho} < 0$, when $\beta \rightarrow -\hat{\rho}$, then $\bar{\zeta} \rightarrow 0$.
- (v) Let $\lambda = \lambda(\tau, \theta, \beta, \hat{\rho}, \zeta)$ be defined in (3.8). The supremum value of $\zeta > 0$ for which there exists $\tau \in]\tau_{\zeta}(\beta, \hat{\rho}), +\infty[$ such that $\lambda(\tau, \theta, \beta, \hat{\rho}, \zeta) > 0$ depends on β , $\hat{\rho}$, and θ . With a slight abuse of notation, this supremum will be denoted by $\bar{\zeta}(\beta, \hat{\rho}, \theta)$. This value is non-decreasing with respect to β , that is, if $\beta_2 > \beta_1$, then $\bar{\zeta}(\beta_2, \hat{\rho}, \theta) \geq \bar{\zeta}(\beta_1, \hat{\rho}, \theta)$. Indeed, if $\zeta > 0$ and $\beta_2 > \beta_1$, then $\tau_{\zeta}(\beta_2, \hat{\rho}) < \tau_{\zeta}(\beta_1, \hat{\rho})$ and $\lambda(\tau, \theta, \beta_2, \hat{\rho}, \zeta) \geq \lambda(\tau, \theta, \beta_1, \hat{\rho}, \zeta)$. Hence, every ζ feasible for β_1 is feasible for β_2 .

Analogously, it can be proved that $\bar{\zeta}(\beta, \hat{\rho}, \theta)$ is a nondecreasing function of $\hat{\rho}$ since λ is increasing with respect to $\hat{\rho}$. Indeed,

$$\frac{\partial \lambda}{\partial \hat{\rho}} = \frac{2}{\tau}((\zeta + \tau\zeta_M)^2 + 4\tau\zeta) - \frac{(\tau + 2\hat{\rho})(4\beta\tau + (2\hat{\rho} - \tau)(\tau + \zeta))}{2(\beta\tau + \hat{\rho}(\tau + \zeta))^2}.$$

Moreover, since $\tau\zeta_M \geq 1 - \zeta$,

$$\begin{aligned} \left. \frac{\partial \lambda}{\partial \hat{\rho}} \right|_{\hat{\rho}=0} &= \frac{2}{\tau}((\zeta + \tau\zeta_M)^2 + 4\tau\zeta) - \frac{4\beta - \tau - \zeta}{2\beta^2} \\ &\geq \frac{2}{\tau}(1 + 4\tau\zeta) - \frac{4\beta - \tau - \zeta}{2\beta^2} \\ &= \frac{(2\beta - \tau)^2 + \tau\zeta(16\beta^2 + 1)}{2\beta^2\tau} > 0. \end{aligned}$$

Besides,

$$\frac{\partial^2 \lambda}{\partial \hat{\rho}^2} = -\frac{\tau^2(2\beta - \tau - \zeta)^2}{(\beta\tau + \hat{\rho}(\tau + \zeta))^3} \leq 0.$$

Then, $\frac{\partial \lambda}{\partial \hat{\rho}}$ is nonincreasing as a function of $\hat{\rho}$ and it is positive at $\hat{\rho} = 0$. We conclude that $\frac{\partial \lambda}{\partial \hat{\rho}}$ is positive on the admissible domain $] -\beta\tau/(\tau + \zeta), 0]$ and that λ is increasing with respect to $\hat{\rho}$ on this domain. In addition, $\tau(\beta, \hat{\rho})$ decreases as a function of $\hat{\rho}$.

On the other hand, for $\theta \neq 1$, we have

$$\frac{\partial \lambda}{\partial \theta} = \begin{cases} -1, & \text{if } \theta \in [1, 2[\text{ and } (\mathbf{S} - \tau \mathbf{M}) \text{ is monotone,} \\ -1 - 2\text{sgn}(\theta - 1)\zeta, & \text{otherwise.} \end{cases}$$

Since $\zeta \in]0, 1/2[$, it follows that $\frac{\partial \lambda}{\partial \theta} < 0$ in all cases. Since λ is continuous at $\theta = 1$, λ is strictly decreasing with respect to θ and $\bar{\zeta}(\beta, \hat{\rho}, \theta)$ is nonincreasing with respect to θ . Hence, smaller values of θ lead to less restrictive conditions for the existence of an admissible step size τ . Figure 1 shows examples of $\bar{\zeta}(\beta, \hat{\rho}, \theta)$ as a function of $\hat{\rho}$ for selected values of β and θ .

4. SOLUTION OF PROBLEM 1.1

In this section, we present our main algorithm for solving Problem 1.1. First, we introduce some notation and results that are useful to establish the convergence of the proposed method.

4.1. Preliminary results. We first introduce some additional notation and assumptions.

Notation 4.1. In the context of Problem 1.1, consider the following operators defined on the product space $\mathcal{H} = \mathcal{H} \times \mathcal{G}$:

$$\begin{aligned} \mathbf{B}: \mathcal{H} &\rightarrow 2^{\mathcal{H}}: (z, v) \mapsto (A + D)z \times B^{-1}v \\ \mathbf{L}: \mathcal{H} &\rightarrow \mathcal{H}: (z, v) \mapsto (L^*v, -Lz) \\ \mathbf{A}: \mathcal{H} &\rightarrow 2^{\mathcal{H}}: (z, v) \mapsto (\mathbf{B} + \mathbf{L})(z, v) \\ \mathbf{C}: \mathcal{H} &\rightarrow \mathcal{H}: (z, v) \mapsto (Cz, 0) \end{aligned} \tag{4.1}$$

and define

$$\mathbf{S} = \{(x^*, -Cx^*) \mid x^* \in \text{zer}(\mathbf{A} + \mathbf{C})\}. \tag{4.2}$$

Note that, based on the above definitions, $\text{zer}(\mathbf{A} + \mathbf{C}) = \mathbf{Z}$.

Assumption 4.2. In the context of Problem 1.1, let $(\mu_A, \mu_B, \rho_A, \rho_B) \in \mathbb{R}^4$ and $(\nu_A, \nu_B) \in [0, +\infty]^2$. Suppose that

- (i) $\rho_A + \rho_B \geq 0$, $\mu_A + \mu_B \geq 0$, $(\rho_A + \rho_B = 0 \Rightarrow \rho_A = \rho_B = 0)$, and $(\mu_A + \mu_B = 0 \Rightarrow \mu_A = \mu_B = 0)$;
- (ii) $A + D$ is $(\nu_A \text{Id} + \mu_A L^* L, \rho_A)$ -semimonotone on $\mathbf{S}_1 = \{(x^*, -Cx^* - L^*u^*) \mid (x^*, u^*) \in \mathbf{Z}\}$;
- (iii) B is $(\mu_B, \nu_B \text{Id} + \rho_B L L^*)$ -semimonotone on $\mathbf{S}_2 = \{(Lx^*, u^*) \mid (x^*, u^*) \in \mathbf{Z}\}$;
- (iv) there exist nonempty sets $\Omega_A \subset]0, +\infty[$ and $\Omega_B \subset]0, +\infty[$ such that, for every $(\tau, \sigma) \in \Omega_A \times \Omega_B$, $J_{\tau A}$ and $J_{\sigma B^{-1}}$ are single-valued and have full domain;
- (v) $\text{gra}(\mathbf{A} + \mathbf{C})$ is sequentially weak-strong closed.

The following proposition showcases some situations where Assumption 4.2(ii) holds, whenever the resulting value of ν_A is nonnegative.

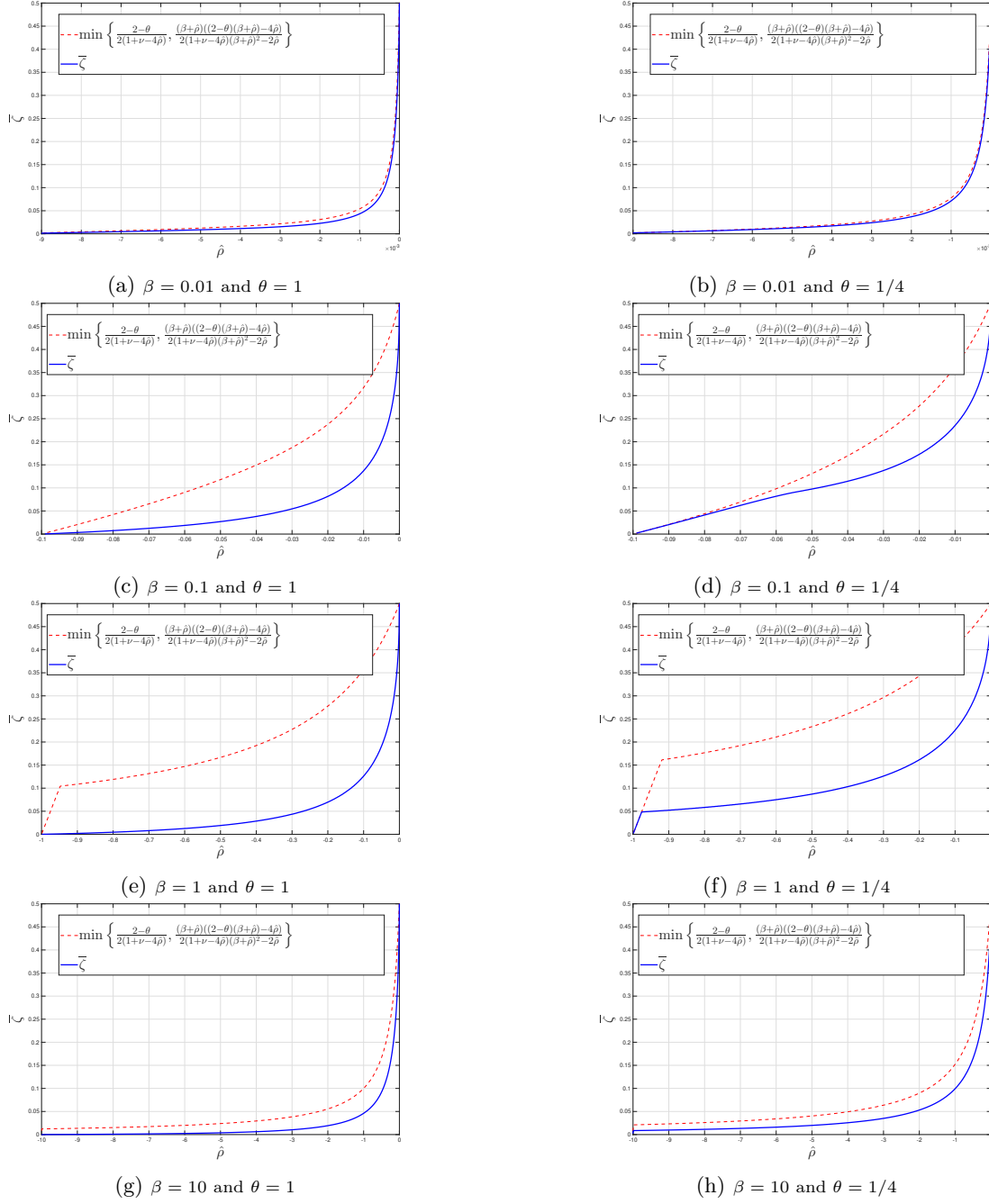


FIGURE 1. $\bar{\zeta}(\beta, \cdot, \theta)$ for $\beta \in \{0.01, 0.1, 1, 10\}$ and $\theta \in \{1, 1/2\}$. In each case $\xi = 1 + \zeta$.

Proposition 4.3. *Let $(\nu, \mu, \rho, \tilde{\rho}) \in \mathbb{R}^4$, let T be a bounded linear operator from $\mathcal{H}$ to a real Hilbert space $\mathcal{K}$, let $K \neq 0$ be a linear bounded operator from $\mathcal{K}$ to $\mathcal{H}$, and let $\tilde{D}: \mathcal{K} \rightarrow \mathcal{K}$ be a $(\vartheta_{\tilde{D}})^{-1}$ -cocoercive operator with $\vartheta_{\tilde{D}} \in]0, +\infty[$. Assume that A is $(\nu \text{Id} + \mu L^* L, \rho)$ -semimonotone. Then, $A + D$ is $(\nu_A \text{Id} + \mu L^* L, \rho_A)$ -semimonotone if*

- (i) $D = K \circ \tilde{D} \circ T$, $1 + \vartheta_{\tilde{D}} \rho \|K\|^2 > 0$, and $(\nu_A, \rho_A) = (\nu - \vartheta_{T^* - K}, \rho / (1 + \vartheta_{\tilde{D}} \rho \|K\|^2))$, where $\vartheta_{T^* - K}$ is the Lipschitz modulus of $(T^* - K) \circ \tilde{D} \circ T$.
- (ii) D is a bounded linear operator, $D - \tilde{\rho} D^* D$ is a monotone operator with $\tilde{\rho} + \rho > 0$, and $(\nu_A, \rho_A) = (\nu, \rho \tilde{\rho} / (\rho + \tilde{\rho}))$.
- (iii) $D = 0$ and $(\nu_A, \rho_A) = (\nu, \rho)$.

Proof.

- (i) For every $(x, y) \in \mathcal{H}^2$, we have

$$\begin{aligned} \langle x - y \mid K \tilde{D} T x - K \tilde{D} T y \rangle &= \langle T x - T y \mid \tilde{D} T x - \tilde{D} T y \rangle + \langle x - y \mid (K - T^*)(\tilde{D} T x - \tilde{D} T y) \rangle \\ &\geq \vartheta_{\tilde{D}}^{-1} \|\tilde{D} T x - \tilde{D} T y\|^2 - \vartheta_{K - T^*} \|x - y\|^2 \\ &\geq \frac{1}{\vartheta_{\tilde{D}} \|K\|^2} \|K \tilde{D} T x - K \tilde{D} T y\|^2 - \vartheta_{K - T^*} \|x - y\|^2. \end{aligned}$$

In view of Lemma 2.2, if $1 + \vartheta_{\tilde{D}} \rho \|K\|^2 > 0$, then $A + D$ is $((\nu - \vartheta_{T^* - K}) \text{Id} + \mu L^* L, \rho / (1 + \vartheta_{\tilde{D}} \rho \|K\|^2))$ -semimonotone.

- (ii) If D is linear, D is $\tilde{\rho}$ -comonotone for $\tilde{\rho} \in \mathbb{R}$ if and only if

$$(\forall x \in \mathcal{H}) \quad \langle x \mid D x \rangle \geq \tilde{\rho} \|D x\|^2, \quad (4.3)$$

that is

$$(\forall x \in \mathcal{H}) \quad \langle x \mid D x \rangle - \tilde{\rho} \langle x \mid D^* D x \rangle \geq 0. \quad (4.4)$$

Hence, D is $\tilde{\rho}$ -comonotone if and only if $D - \tilde{\rho} D^* D$ is monotone. One can thus apply Lemma 2.2 to claim that $A + D$ is $(\nu \text{Id} + \mu L^* L, \rho \tilde{\rho} / (\rho + \tilde{\rho}))$ -semimonotone.

- (iii) The last case is straightforward. □

The following remark provides more insight on Assumptions 4.2(i)-(iii) when semimonotonicity holds globally.

Remark 4.4.

- (i) If $\rho_A = \rho_B = 0$ and $\mu_A + \mu_B > 0$, then $A + D + L^* \circ B \circ L$ is monotone, which means that (1.2) is a monotone inclusion problem. In turn, if $\mu_A = \mu_B = 0$ and $\rho_A + \rho_B > 0$, then $B^{-1} - L \circ (A + D)^{-1} \circ (-L^*)$ is monotone, which implies that (1.3) is a monotone inclusion problem when $C = 0$. However, when $A + D$ or B are not monotone, we will show that $\mathbf{A}$ is cohypomonotone.
- (ii) If $A + D$ is $\mu_A L^* L$ -monotone with $\mu_A \in]-\infty, 0]$, then it is $\mu_A \|L\|^2$ -monotone. Indeed, let $((z, v), (x, u)) \in (\text{gra}(A + D))^2$. We have

$$\langle z - x \mid v - u \rangle \geq \mu_A \langle z - x \mid L^* L(z - x) \rangle = \mu_A \|L(z - x)\|^2 \geq \mu_A \|L\|^2 \|z - x\|^2.$$

Similarly, if B is $\rho_B L L^*$ -comonotone for $\rho_B \in]-\infty, 0]$, then B is $\rho_B \|L\|^2$ -comonotone.

- (iii) Assume that $L \neq 0$. If $A + D$ is ν_A -strongly monotone for $\nu_A \in]0, +\infty[$, then it is $\mu_A L^* L$ -monotone with $\mu_A = \nu_A / \|L\|^2$. Indeed, for every $((z, v), (x, u)) \in (\text{gra}(A + D))^2$,

$$\langle z - x \mid v - u \rangle \geq \nu_A \|z - x\|^2 \geq \frac{\nu_A}{\|L\|^2} \|L(z - x)\|^2 = \frac{\nu_A}{\|L\|^2} \|z - x\|_{L^* L}^2.$$

Similarly, if B is ν_B -cocoercive for $\nu_B > 0$, B is $\nu_B \|L\|^{-2} LL^*$ -comonotone.

The next proposition provides conditions going beyond maximal monotonicity, ensuring that Assumption 4.2(iv) is fulfilled.

Proposition 4.5. *Let $(\nu, \mu, \rho, \nu_B, \mu_B, \rho_B) \in \mathbb{R}^6$ and let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Assume that A is $(\nu \text{Id} + \mu L^* L, \rho)$ -semimonotone and B is $(\mu_B, \nu_B \text{Id} + \rho_B LL^*)$ -semimonotone. Let $\tilde{\Omega}_A \subset]0, +\infty[$, $\tilde{\Omega}_B \subset]0, +\infty[$, and*

$$\Omega_A = \{\tau \in \tilde{\Omega}_A \mid \text{ran}(\text{Id} + \tau A) = \mathcal{H}\}, \quad (4.5)$$

$$\Omega_B = \{\sigma \in \tilde{\Omega}_B \mid \text{ran}(\text{Id} + \sigma B^{-1}) = \mathcal{G}\}. \quad (4.6)$$

Then Assumption 4.2(iv) holds if $\Omega_A \neq \emptyset$, $\Omega_B \neq \emptyset$, one of the following two conditions holds:

- (i) $(\nu, \mu) \in [0, +\infty[^2$ and $\tilde{\Omega}_A =]\max\{0, -\rho\}, +\infty[$,
- (ii) $\rho \in [0, +\infty[$ and

$$\tilde{\Omega}_A = \begin{cases}]0, +\infty[, & \text{if } \nu + \min\{\mu, 0\} \|L\|^2 \geq 0, \\]0, -1/(\nu + \min\{\mu, 0\} \|L\|^2)[& \text{otherwise,} \end{cases} \quad (4.7)$$

and one of the following two conditions holds:

- (iii) $(\nu_B, \rho_B) \in [0, +\infty[^2$ and $\tilde{\Omega}_B =]\max\{0, -\mu_B\}, +\infty[$,
- (iv) $\mu_B \in [0, +\infty[$ and

$$\tilde{\Omega}_B = \begin{cases}]0, +\infty[, & \text{if } \nu_B + \min\{\rho_B, 0\} \|L\|^2 \geq 0, \\]0, -1/(\nu_B + \min\{\rho_B, 0\} \|L\|^2)[& \text{otherwise.} \end{cases} \quad (4.8)$$

In addition, assume that A is maximally $(\nu \text{Id} + \mu L^* L, \rho)$ -semimonotone and B is maximally $(\mu_B, \nu_B \text{Id} + \rho_B LL^*)$ -semimonotone. If $\nu = \mu = 0$ or $\rho = 0$, then $\Omega_A = \tilde{\Omega}_A$ and, if $\mu_B = 0$ or $\nu_B = \rho_B = 0$, then $\Omega_B = \tilde{\Omega}_B$.

Proof. If A is $(\nu \text{Id} + \mu L^* L, \rho)$ -semimonotone with $(\nu, \mu) \in [0, +\infty[^2$, then it is ρ -comonotone. If $\tau > -\rho$, then $J_{\tau A}$ is single-valued on $\text{ran}(\text{Id} + \tau A)$ [5, Proposition 2.12]. Let us set $\tilde{\Omega}_A =]0, +\infty[$ if $\rho \geq 0$, and $\tilde{\Omega}_A =]-\rho, +\infty[$ otherwise. For every $\tau \in \tilde{\Omega}_A$, $\text{ran}(\text{Id} + \tau A) = \mathcal{H}$. Therefore, $J_{\tau A}$ is single-valued and has full domain, for every $\tau \in \tilde{\Omega}_A$.

On the other hand, if A is $(\nu \text{Id} + \mu L^* L, \rho)$ -semimonotone with $\rho \in [0, +\infty[$ it follows from Remark 4.4(ii) that it is $(\nu + \min\{\mu, 0\} \|L\|^2)$ -monotone. Then $J_{\tau A}$ is single-valued on its domain when $\tau \in \tilde{\Omega}_A$ with $\tilde{\Omega}_A =]0, +\infty[$ if $\nu + \min\{\mu, 0\} \|L\|^2 \geq 0$, and $\tilde{\Omega}_A =]0, -1/(\nu + \min\{\mu, 0\} \|L\|^2)[$ otherwise.

Similarly, if B is $(\mu_B, \nu_B \text{Id} + \rho_B LL^*)$ -semimonotone with $(\nu_B, \rho_B) \in [0, +\infty[^2$, B is μ_B -monotone, that is B^{-1} is μ_B -comonotone. Then, $J_{\sigma B^{-1}}$ is single-valued on $\text{ran}(\text{Id} + \sigma B^{-1})$, if $\sigma \in \tilde{\Omega}_B$, where $\tilde{\Omega}_B =]0, +\infty[$ if $\mu_B \geq 0$ and $\tilde{\Omega}_B =]-\mu_B, +\infty[$ otherwise.

On the other hand, if B is $(\mu_B, \nu_B \text{Id} + \rho_B LL^*)$ -semimonotone with $\mu_B \in [0, +\infty[$ then it follows from Remark 4.4(ii) that B^{-1} is $(\nu_B + \min\{\rho_B, 0\} \|L\|^2)$ -monotone. Thus, we can set $\tilde{\Omega}_B$ as indicated in (4.8).

Suppose now that A and B are maximally semimonotone with respect to the moduli stated in the proposition. We prove the additional range conclusions by considering separately the cases $\nu = \mu = 0$, $\rho = 0$, $\nu_B = \rho_B = 0$, and $\mu_B = 0$.

If $\nu = \mu = 0$, then τA is maximally ρ/τ -comonotone and, by [5, Corollary 2.11], $\text{ran}(\text{Id} + \tau A) = \mathcal{H}$, for every $\tau \in]\max\{0, -\rho\}, +\infty[$.

If $\rho = 0$, then A is $\nu \text{Id} + \mu L^* L$ -monotone, that is $\tilde{A} = A - \nu \text{Id} - \mu L^* L$ is monotone. The graph of A is mapped to the graph of $\tilde{A}$ by $(x, u) \mapsto (x, u - \nu x - \mu L^* Lx)$. Since this transformation

is bijective, it preserves maximality. For every $\tau \in]0, +\infty[$,

$$\text{Id} + \tau A = \tau \tilde{A} + Q_A, \quad Q_A = (1 + \tau \nu) \text{Id} + \tau \mu L^* L. \quad (4.9)$$

We have $Q_A \succeq (1 + \tau(\nu + \min\{\mu, 0\}\|L\|^2)) \text{Id}$. If $\nu + \min\{\mu, 0\}\|L\|^2 \geq 0$, then Q_A is strongly monotone and maximally monotone, for every $\tau \in]0, +\infty[$. Maximal strong monotonicity of Q_A also holds if $\nu + \min\{\mu, 0\}\|L\|^2 < 0$ and $1 + \tau(\nu + \min\{\mu, 0\}\|L\|^2) > 0$, that is $\tau \in]0, -1/(\nu + \min\{\mu, 0\}\|L\|^2)[$. Under these conditions, since $\tilde{A}$ is maximally monotone and Q_A is strongly monotone and maximally monotone with full domain, $\tau \tilde{A} + Q_A$ is strongly monotone and maximally monotone [4, Corollary 25.5]. Hence its range is $\mathcal{H}$ [4, Corollary 23.37].

The $(\mu_B, \nu_B \text{Id} + \rho_B L L^*)$ -semimonotonicity of B is equivalent to the monotonicity of $\tilde{B} = B^{-1} - \nu_B \text{Id} - \rho_B L L^*$, if $\mu_B = 0$, and to the μ_B -comonotonicity of B^{-1} , if $\nu_B = \rho_B = 0$. By arguments analogous to those used above, if $(\nu_B = \rho_B = 0$ or $\mu_B = 0)$ and $\sigma \in \tilde{\Omega}_B$, then $\text{ran}(\text{Id} + \sigma B^{-1}) = \mathcal{G}$. $\square$

We now state a number of results which will be useful in the convergence proofs of the next section. In particular, the next result highlights cohypomonotonicity properties of $\mathbf{A}$.

Proposition 4.6. *In the context of Notation 4.1 and Assumption 4.2, let $(\tau, \sigma) \in]0, +\infty[^2$, let $\mathbf{S} \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be a self adjoint and strongly monotone, and suppose that*

$$(\forall (z, v) \in \mathcal{H}) \quad \underline{\chi} \|(z, v)\|_{\mathbf{S}}^2 \leq \|z\|^2 + \frac{\tau}{\sigma} \|v\|^2 \quad \text{and} \quad \|z\|^2 + \frac{\sigma}{\tau} \|v\|^2 \leq \bar{\chi} \|(z, v)\|_{\mathbf{S}^{-1}}^2, \quad (4.10)$$

where $(\underline{\chi}, \bar{\chi}) \in]0, +\infty[^2$. In addition, define the parameters

$$\hat{\mu} = \min \left\{ \nu_A, \frac{\sigma \nu_B}{\tau} \right\}, \quad \hat{\rho} = \begin{cases} \min \{0, \rho_{A,B}, \frac{\tau}{\sigma} \mu_{A,B}\}, & \text{if } \min\{\rho_A + \rho_B, \mu_A + \mu_B\} > 0 \\ \min \{0, \frac{\tau}{\sigma} \mu_{A,B}\}, & \text{if } \rho_A = \rho_B = 0 \text{ and } \mu_B + \mu_A > 0 \\ \min \{0, \rho_{A,B}\}, & \text{if } \mu_A = \mu_B = 0 \text{ and } \rho_A + \rho_B > 0 \\ 0, & \text{if } \mu_A = \mu_B = \rho_B = \rho_A = 0 \end{cases} \quad (4.11)$$

with $\rho_{A,B} = \rho_A \rho_B / (\rho_A + \rho_B)$ if $\rho_A + \rho_B > 0$, and $\mu_{A,B} = \mu_A \mu_B / (\mu_A + \mu_B)$ if $\mu_A + \mu_B > 0$. Then, $\mathbf{A}$ is $(\underline{\chi} \hat{\mu} \mathbf{S}, \bar{\chi} \hat{\rho} \mathbf{S}^{-1})$ -semimonotone on $\mathbf{S}$.

Proof. Let $(\mathbf{x}, \mathbf{u}) = ((z, v), (x, u)) \in \text{gra } \mathbf{A}$ and let $\mathbf{x}^* = (z^*, v^*) \in \text{zer}(\mathbf{A} + \mathbf{C})$, thus $(\mathbf{x}^*, -\mathbf{C}\mathbf{x}^*) \in \mathbf{S}$. We have then $x - L^*v \in (A + D)z$, $u + Lz \in B^{-1}v$, $-Cz^* - L^*v^* \in (A + D)z^*$, and

$Lz^* \in B^{-1}v^*$. First, suppose that $\min\{\rho_A + \rho_B, \mu_B + \mu_A\} > 0$. It follows from Assumption 4.2(ii)-(iii), Lemma 2.1, and (4.10) that

$$\begin{aligned}
\langle \mathbf{x} - \mathbf{x}^* \mid \mathbf{u} + \mathbf{C}\mathbf{x}^* \rangle &= \langle (z - z^*, v - v^*) \mid (x + Cz^*, u) \rangle \\
&= \langle z - z^* \mid x + Cz^* \rangle + \langle v - v^* \mid u \rangle \\
&= \langle z - z^* \mid x - L^*v + Cz^* + L^*v^* \rangle + \langle v - v^* \mid u + Lz - Lz^* \rangle \\
&\geq \nu_A \|z - z^*\|^2 + \mu_A \|z - z^*\|_{L^*L}^2 + \rho_A \|x - L^*v + Cz^* + L^*v^*\|^2 \\
&\quad + \nu_B \|v - v^*\|^2 + \rho_B \|v - v^*\|_{LL^*}^2 + \mu_B \|u + Lz - Lz^*\|^2 \\
&\geq \nu_A \|z - z^*\|^2 + \nu_B \|v - v^*\|^2 \\
&\quad + \frac{\rho_A \rho_B}{\rho_A + \rho_B} \|x + Cz^*\|^2 + \frac{\mu_A \mu_B}{\mu_B + \mu_A} \|u\|^2 \\
&\geq \min\{\nu_A, \sigma \nu_B / \tau\} \left(\|z - z^*\|^2 + \frac{\tau}{\sigma} \|v - v^*\|^2 \right) \\
&\quad + \min \left\{ 0, \frac{\rho_A \rho_B}{\rho_A + \rho_B}, \frac{\tau \mu_A \mu_B}{\sigma(\mu_B + \mu_A)} \right\} \left(\|x + Cz^*\|^2 + \frac{\sigma}{\tau} \|u\|^2 \right) \\
&\geq \underline{\chi} \hat{\mu} \|\mathbf{x} - \mathbf{x}^*\|_{\mathbf{S}}^2 + \bar{\chi} \hat{\rho} \|\mathbf{u} + \mathbf{C}\mathbf{x}^*\|_{\mathbf{S}^{-1}}^2.
\end{aligned}$$

The remaining cases are treated analogously. $\square$

Proposition 4.7. *In the context of Notation 4.1 and Assumption 4.2, let $(\sigma, \tau) \in]0, +\infty[^2$ be such that $\sigma\tau\|L\|^2 < 1$, and consider the operators*

$$\begin{aligned}
\mathbf{S}: \mathcal{H} &\rightarrow \mathcal{H}: (z, v) \mapsto (z - \tau L^*v, -\tau Lz + \tau v / \sigma) \\
T_1: \mathcal{H} &\rightarrow \mathcal{H}: z \mapsto (\text{Id} - \sigma\tau L^*L)z \\
T_2: \mathcal{G} &\rightarrow \mathcal{G}: v \mapsto (\text{Id} - \sigma\tau LL^*)v.
\end{aligned} \tag{4.12}$$

The following assertions hold:

- (i) T_1 , T_2 , and $\mathbf{S}$ are positive definite, self-adjoint, and strongly monotone linear operators. Furthermore,

$$(\forall (z, v) \in \mathcal{H}) \quad \mathbf{S}^{-1}(z, v) = (T_1^{-1}(z + \sigma L^*v), \sigma T_2^{-1}(Lz + v / \tau)). \tag{4.13}$$

- (ii) $L \circ T_1^{-1} = T_2^{-1} \circ L$ and $L^* \circ T_2^{-1} = T_1^{-1} \circ L^*$.
- (iii) For every $(z, v) \in \mathcal{H}$,

$$\|(z, v)\|_{\mathbf{S}^{-1}}^2 = \|z + \sigma L^*v\|_{T_1^{-1}}^2 + \frac{\sigma}{\tau} \|v\|^2 = \|z\|^2 + \frac{\sigma}{\tau} \|v + \tau Lz\|_{T_2^{-1}}^2. \tag{4.14}$$

In addition,

$$\|z\|^2 + \frac{\sigma}{\tau} \|v\|^2 \leq (1 + \sqrt{\sigma\tau}\|L\|) \|(z, v)\|_{\mathbf{S}^{-1}}^2. \tag{4.15}$$

- (iv) For every $(z, v) \in \mathcal{H}$,

$$\|(z, v)\|_{\mathbf{S}}^2 \leq (1 + \sqrt{\sigma\tau}\|L\|) \left(\|z\|^2 + \frac{\tau}{\sigma} \|v\|^2 \right). \tag{4.16}$$

- (v) $\mathbf{A}$ is $(\underline{\chi} \hat{\mu} \mathbf{S}, \bar{\chi} \hat{\rho} \mathbf{S}^{-1})$ -semimonotone on $\mathbf{S}$, where $\hat{\mu}$ and $\hat{\rho}$ are defined in (4.11) with $\bar{\chi} = 1 + \sqrt{\sigma\tau}\|L\|$ and $\underline{\chi} = \bar{\chi}^{-1}$.

Proof.

- (i) The first assertion follows directly from existing results, see, e.g., [48, Lemma 6.1].
- (ii) Note that $LT_1 = T_2L$, thus,

$$(\forall z \in \mathcal{H}) \quad Lz = LT_1T_1^{-1}z = T_2LT_1^{-1}z.$$

Then, we conclude that $T_2^{-1}Lz = LT_1^{-1}z$. The second identity is deduced similarly.

(iii) For every $(z, v) \in \mathcal{H}$,

$$\begin{aligned}
\|(z, v)\|_{\mathcal{S}^{-1}}^2 &= \langle \mathcal{S}^{-1}(z, v) \mid (z, v) \rangle \\
&= \langle T_1^{-1}(z + \sigma L^* v) \mid z \rangle + \frac{\sigma}{\tau} \langle T_2^{-1}(v + \tau Lz) \mid v \rangle \\
&= \langle T_1^{-1}(z + \sigma L^* v) \mid z + \sigma L^* v \rangle - \sigma \langle T_1^{-1}(z + \sigma L^* v) \mid L^* v \rangle + \frac{\sigma}{\tau} \langle T_2^{-1}(v + \tau Lz) \mid v \rangle \\
&= \|z + \sigma L^* v\|_{T_1^{-1}}^2 - \sigma \langle T_1^{-1}(z + \sigma L^* v) \mid L^* v \rangle + \frac{\sigma}{\tau} \langle T_2^{-1}(v + \tau Lz) \mid v \rangle \\
&\stackrel{(ii)}{=} \|z + \sigma L^* v\|_{T_1^{-1}}^2 - \sigma \langle T_2^{-1}L(z + \sigma L^* v) \mid v \rangle + \frac{\sigma}{\tau} \langle T_2^{-1}(v + \tau Lz) \mid v \rangle \\
&= \|z + \sigma L^* v\|_{T_1^{-1}}^2 + \frac{\sigma}{\tau} \langle T_2^{-1}(v - \sigma \tau L L^* v) \mid v \rangle \\
&\stackrel{(4.12)}{=} \|z + \sigma L^* v\|_{T_1^{-1}}^2 + \frac{\sigma}{\tau} \|v\|^2.
\end{aligned}$$

The identity $\|(z, v)\|_{\mathcal{S}^{-1}}^2 = \|z\|^2 + \frac{\sigma}{\tau} \|v + \tau Lz\|_{T_2^{-1}}^2$ is deduced in a similar manner.

Let $(z', v') = \mathcal{S}^{-1}(z, v)$. Set $\omega = \sqrt{\sigma\tau}\|L\|$ and $u = \sqrt{\tau/\sigma}v'$. We have then

$$\begin{aligned}
\|z\|^2 + \frac{\sigma}{\tau} \|v\|^2 &= \|z' - \tau L^* v'\|^2 + \sigma \tau \left\| -Lz' + \frac{v'}{\sigma} \right\|^2 \\
&= \|z'\|^2 - 2\tau \langle Lz' \mid v' \rangle + \tau^2 \|L^* v'\|^2 + \sigma \tau \|Lz'\|^2 - 2\tau \langle Lz' \mid v' \rangle + \frac{\tau}{\sigma} \|v'\|^2 \\
&\leq (1 + \omega^2) \|z'\|^2 - 4\sqrt{\sigma\tau} \langle Lz' \mid u \rangle + (1 + \omega^2) \|u\|^2
\end{aligned}$$

and

$$\|(z, v)\|_{\mathcal{S}^{-1}}^2 = \|(z', v')\|_{\mathcal{S}}^2 = \|z'\|^2 - 2\sqrt{\sigma\tau} \langle Lz' \mid u \rangle + \|u\|^2.$$

Using Cauchy-Schwarz inequality yields

$$\begin{aligned}
(1 + \omega) \|(z, v)\|_{\mathcal{S}^{-1}}^2 - \|z\|^2 - \frac{\sigma}{\tau} \|v\|^2 &\geq (1 - \omega) (\omega \|z'\|^2 + 2\sqrt{\sigma\tau} \langle Lz' \mid u \rangle + \omega \|u\|^2) \\
&\geq (1 - \omega) (\omega \|z'\|^2 - 2\sqrt{\sigma\tau} \|Lz'\| \|u\| + \omega \|u\|^2) \\
&\geq (1 - \omega) \omega (\|z'\| - \|u\|)^2.
\end{aligned}$$

Since $\omega \in [0, 1[$, we deduce that

$$\|z\|^2 + \frac{\sigma}{\tau} \|v\|^2 \leq (1 + \omega) \|(z, v)\|_{\mathcal{S}^{-1}}^2. \quad (4.17)$$

(iv) Assume that $L \neq 0$. For every $(z, v) \in \mathcal{H}$ and $\delta \in]0, +\infty[$,

$$\begin{aligned}
\|(z, v)\|_{\mathcal{S}}^2 &= \langle (z - \tau L^* v, \tau v/\sigma - \tau Lz) \mid (z, v) \rangle \\
&= \|z\|^2 - 2\tau \langle Lz \mid v \rangle + \frac{\tau}{\sigma} \|v\|^2 \\
&\leq \|z\|^2 + 2\tau \|L\| \|z\| \|v\| + \frac{\tau}{\sigma} \|v\|^2 \\
&\leq \|z\|^2 + \delta \|z\|^2 + \frac{\tau^2 \|L\|^2}{\delta} \|v\|^2 + \frac{\tau}{\sigma} \|v\|^2 \\
&\leq \max \left\{ 1 + \delta, 1 + \frac{\sigma\tau \|L\|^2}{\delta} \right\} \left(\|z\|^2 + \frac{\tau}{\sigma} \|v\|^2 \right).
\end{aligned}$$

The tightest value of the bound is obtained by setting $\delta = \sqrt{\sigma\tau}\|L\|$, which yields

$$\|(z, v)\|_{\mathcal{S}}^2 \leq (1 + \sqrt{\sigma\tau}\|L\|) \left(\|z\|^2 + \frac{\tau}{\sigma} \|v\|^2 \right). \quad (4.18)$$

This inequality remains valid when $L = 0$.

(v) The last assertion follows directly from (iii), (iv), and Proposition 4.6. $\square$

Proposition 4.8. *In the context of Notation 4.1 and Assumption 4.2, let $(\tau, \sigma) \in \Omega_A \times \Omega_B$. Define*

$$\mathbf{M}: \mathcal{H} \rightarrow \mathcal{H}: (z, v) \mapsto (z/\tau - Dz - L^*v, -Lz + v/\sigma). \quad (4.19)$$

Then $(\mathbf{M} + \mathbf{A})^{-1}$ is single-valued and has full domain. Moreover,

$$(\forall (x, w) \in \mathcal{H}) \quad (\mathbf{M} + \mathbf{A})^{-1}(x, w) = (J_{\tau A}(\tau x), J_{\sigma B^{-1}}(\sigma(w + 2LJ_{\tau A}(\tau x)))) . \quad (4.20)$$

Proof. It follows from (4.1) and (4.19) that

$$(\forall (z, v) \in \mathcal{H}) \quad (\mathbf{M} + \mathbf{A})(z, v) = (Az + z/\tau) \times (B^{-1}v + v/\sigma - 2Lz).$$

Therefore, given $(z, v) \in \mathcal{H}$ and $(x, w) \in \mathcal{H}$, $(z, v) \in (\mathbf{M} + \mathbf{A})^{-1}(x, w)$ if and only if $x \in (\tau^{-1}\text{Id} + A)z$ and $w + 2Lz \in (\sigma^{-1}\text{Id} + B^{-1})v$. Since $J_{\tau A}$ and $J_{\sigma B^{-1}}$ are single-valued, we conclude that $z = J_{\tau A}(\tau x)$ and $v = J_{\sigma B^{-1}}(\sigma(w + 2LJ_{\tau A}(\tau x)))$. Since $J_{\tau A}$ and $J_{\sigma B^{-1}}$ are single-valued and have full domain, (4.20) shows that $(\mathbf{M} + \mathbf{A})^{-1}$ is also single-valued and has full domain. $\square$

4.2. Algorithm and convergence analysis. In this subsection, we will show that the following primal-dual algorithm provides a numerical method for solving Problem 1.1.

Algorithm 4.9. *In the context of Problem 1.1 and Assumption 4.2, let $(\tau, \sigma) \in \Omega_A \times \Omega_B$, let $\theta \in]0, 2[$, let $(z_{-1}, z_0, p_0, v_0) \in \mathcal{H}^3 \times \mathcal{G}$, and set $y_{-1} = Dz_{-1}$. Consider the iteration:*

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = Dz_n \\ x_n = z_n - \tau(y_n - y_{n-1} + Dp_n + Cz_n + L^*v_n) \\ p_{n+1} = J_{\tau A}x_n \\ w_n = v_n + \sigma L(2p_{n+1} - z_n) \\ q_{n+1} = J_{\sigma B^{-1}}w_n \\ (z_{n+1}, v_{n+1}) = (1 - \theta)(z_n, v_n) + \theta(p_{n+1}, q_{n+1}). \end{cases} \quad (4.21)$$

Our main result concerning the convergence of the above algorithm is stated next.

Theorem 4.10. *In the context of Problem 1.1 and Assumption 4.2, consider the sequence $(z_n, v_n)_{n \in \mathbb{N}}$ generated by Algorithm 4.9 where $(\tau, \sigma) \in \Omega_A \times \Omega_B$ is such that $\kappa = 1 - \sigma\tau\|L\|^2 > 0$. Let $\hat{\rho} \leq 0$ be defined by (4.11) and let $\bar{\chi} = 1 + \sqrt{\sigma\tau}\|L\|$. Moreover define*

$$\eta_\theta = \begin{cases} 1, & \text{if } \theta \in [1, 2] \text{ and } D \text{ is monotone,} \\ 1 + |1 - \theta|, & \text{otherwise,} \end{cases} \quad (4.22)$$

Suppose that $\beta\kappa + \bar{\chi}\hat{\rho}(1 + \vartheta/\kappa) > 0$ and

$$\tau \left(2 - \theta - 2\tau \frac{\eta_\theta \vartheta}{\kappa} \right) > \frac{(\tau + 2\bar{\chi}\hat{\rho})^2}{2(\beta\kappa + \bar{\chi}\hat{\rho}(1 + \vartheta/\kappa))} - 2\bar{\chi}\hat{\rho} \left(\left(2\tau \frac{\vartheta}{\kappa} + 1 \right)^2 + 4\tau^2 \frac{\vartheta}{\kappa} \right). \quad (4.23)$$

Then the following hold.

- (i) The sequence $(z_n, v_n)_{n \in \mathbb{N}}$ converges weakly to a solution to Problem 1.1.
- (ii) If $\nu_A > 0$ and $\nu_B > 0$, then $(z_n, v_n)_{n \in \mathbb{N}}$ converges R -linearly to the unique solution to Problem 1.1.

Proof. Let $(\tau, \sigma) \in \Omega_A \times \Omega_B$ and consider the operators defined in (4.1), operator $\mathbf{S}$ defined in (4.12), and operator $\mathbf{M}$ defined in (4.19). Let, for every $n \in \mathbb{N}$, $\mathbf{x}_n = (z_n, v_n)$, $\mathbf{p}_{n+1} = (p_{n+1}, q_{n+1})$, and $\mathbf{u}_{n+1} = (\tau\mathbf{M} - \mathbf{S})\mathbf{p}_{n+1} - (\tau\mathbf{M} - \mathbf{S})\mathbf{x}_n$, where $\tau\mathbf{M} - \mathbf{S}: (z, v) \mapsto (-\tau Dz, 0)$. It follows from (4.21) and (4.20) that

$$(p_{n+1}, q_{n+1}) = (J_{\tau A}x_n, J_{\sigma B^{-1}}w_n) \Leftrightarrow \mathbf{p}_{n+1} = (\mathbf{M} + \mathbf{A})^{-1}(x_n/\tau, w_n/\sigma - 2Lp_{n+1}).$$

Moreover,

$$\begin{cases} x_n/\tau = z_n/\tau - Dz_n - Dp_n + Dz_{n-1} - Cz_n - L^*v_n \\ w_n/\sigma - 2Lp_{n+1} = v_n/\sigma - Lz_n \end{cases} \\ \Leftrightarrow (x_n/\tau, w_n/\sigma - 2Lp_{n+1}) = \mathbf{M}\mathbf{x}_n - \mathbf{C}\mathbf{x}_n + \mathbf{u}_n/\tau.$$

Therefore, by setting $\mathbf{u}_0 = (-\tau Dp_0 + \tau Dz_{-1}, 0)$, we conclude that

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \mathbf{p}_{n+1} = (\mathbf{M} + \mathbf{A})^{-1}(\mathbf{M}\mathbf{x}_n - \mathbf{C}\mathbf{x}_n + \mathbf{u}_n/\tau) \\ \mathbf{u}_{n+1} = (\tau\mathbf{M} - \mathbf{S})\mathbf{p}_{n+1} - (\tau\mathbf{M} - \mathbf{S})\mathbf{x}_n \\ \mathbf{x}_{n+1} = (1 - \theta)\mathbf{x}_n + \theta\mathbf{p}_{n+1}, \end{cases} \quad (4.24)$$

which is of the form of Algorithm 3.7. In addition, $(\mathbf{M} + \mathbf{A})^{-1}$ is single-valued and has full domain by Proposition 4.8 and $\text{gra}(\mathbf{A} + \mathbf{C})$ is sequentially weak-strong closed by Assumption 4.2. According to Proposition 4.7, $\mathbf{A}$ is $(\mu\mathbf{S}, \rho\mathbf{S}^{-1})$ -semimonotone on $\mathbf{S}$ with $(\mu, \rho) = (\underline{\chi}\hat{\mu}, \bar{\chi}\hat{\rho}) \in [0, +\infty[\times] -\infty, 0]$ with $\underline{\chi} = (1 + \sqrt{\sigma\tau}\|L\|)^{-1}$. Moreover, it follows from the proof of [48, Corollary 6.1] that $\mathbf{C}$ is $(\beta\kappa)$ -cocoercive with respect to $\mathbf{S}$ and $\tau\mathbf{M} - \mathbf{S}$ is ζ -Lipschitzian with respect to $\mathbf{S}$ with $\zeta = \tau\vartheta/\kappa$. In addition, η_θ coincides with η defined in (3.7). Finally, the condition $\beta\kappa + \bar{\chi}\hat{\rho}(1 + \vartheta/\kappa) > 0$ is equivalent to $\tau > -\zeta\rho/(\rho + \beta\kappa) = \underline{\tau}_\zeta$, and implies that $\rho > -\beta\kappa$. In summary, Assumption 3.1 is satisfied provided that $\zeta < 1/2$. In addition, according to Remark 3.2(i), $\mathbf{M}$ is $\zeta_{\mathbf{M}}$ -Lipschitzian with respect to $\mathbf{S}$. The convergence results then follow from Theorem 3.10 by noticing that, in the considered setting, $\text{zer}(\mathbf{A} + \mathbf{C}) = \mathbf{Z}$ and the condition $\lambda > 0$ reads

$$\lambda = 2 - \theta - \frac{2\eta_\theta\tau\vartheta}{\kappa} - \frac{(\tau + 2\rho)^2}{2\tau(\beta\kappa + \rho(1 + \vartheta/\kappa))} + 2\rho\tau \left(\left(\frac{\vartheta}{\kappa} + \zeta_{\mathbf{M}} \right)^2 + \frac{4\vartheta}{\kappa} \right) > 0. \quad (4.25)$$

By virtue of Remark 3.13(i), this condition ensures that $\zeta < 1/2$. Majorizing $\zeta_{\mathbf{M}}$ by $(1 + \zeta)/\tau$, condition (4.23) implies (4.25).

The R-linear convergence is deduced from Theorem 3.10 when $\hat{\mu}$, as defined by (4.11), is positive. $\square$

Remark 4.11. Suppose that $L \neq 0$ and set $\kappa \in]0, 1[$. For any $\tau \in]0, +\infty[$ we can find $\sigma \in]0, +\infty[$ such that $\kappa = 1 - \sigma\tau\|L\|^2$. Then, $\sigma = (1 - \kappa)/(\tau\|L\|^2)$, $\underline{\chi} = (1 + \sqrt{1 - \kappa})^{-1}$, and $\bar{\chi} = 1 + \sqrt{1 - \kappa}$.

(i) When $\hat{\rho} = 0$ and $\theta = 1$, (4.23) reduces to

$$\tau < \frac{\kappa}{\frac{1}{2\beta} + 2\vartheta}, \quad (4.26)$$

which is the condition in [48, Corollary 6.1] guaranteeing convergence of Primal-Dual Method with Block Triangular Resolvent splitting in the monotone case.

(ii) When $\hat{\rho} \neq 0$ and $\rho_A = \rho_B = 0$, it follows from (4.11) that $\hat{\rho} = \tilde{\rho}\tau^2$ where $\tilde{\rho} = \|L\|^2\mu_{A,B}/(1 - \kappa) < 0$. Thus, (4.23) reads

$$\frac{(1 + 2\bar{\chi}\tilde{\rho}\tau)^2}{4(\beta\kappa + \bar{\chi}\tilde{\rho}(1 + \vartheta/\kappa)\tau^2)} - \bar{\chi}\tilde{\rho} \left(\left(\frac{2\tau\vartheta}{\kappa} + 1 \right)^2 + \frac{4\tau^2\vartheta}{\kappa} \right) + \frac{\eta_\theta\vartheta}{\kappa} < \frac{2 - \theta}{2\tau}, \quad (4.27)$$

which holds for τ small enough. This is consistent with the choice of $\tilde{\Omega}_A$ (resp. $\tilde{\Omega}_B$) in Condition (ii) (resp. (iii)) of Proposition 4.5.

- (iii) Situations where $\hat{\rho} \neq 0$ and $\mu_A = \mu_B = 0$ will be studied in Section 4.3.
- (iv) The authors of [48] included an extrapolation step in the iterates of the Primal–Dual Method with Block Triangular Resolvent. Such an extrapolation step could also be incorporated into our analysis, but we omit it for the sake of simplicity. In the same article, the authors also proposed the Primal–Dual Algorithm with Resolvent-Corrected Kernel, which could likewise be extended to our setting.

4.3. Specific cases. We now present iterative methods derived from Algorithm 4.9 for specific instances of Problem 1.1.

4.3.1. Case $D = 0$. The previous analysis allows us to establish convergence results for the Condat–Vũ algorithm [24, 59] in the semimonotone case.

Proposition 4.12. *In the context of Problem 1.1 and Assumption 4.2, suppose that $D = 0$. Let $(\tau, \sigma) \in \Omega_A \times \Omega_B$, let $\theta \in]0, 2[$, and let $\hat{\rho} \leq 0$ be defined by (4.11). Moreover, assume that*

$$\sqrt{\sigma\tau}\|L\| < \frac{\beta + \hat{\rho}}{\beta}, \quad (4.28)$$

$$\begin{aligned} & - (1 + 2\beta\sigma\|L\|^2(2 - \theta))\tau^2 - 2\hat{\rho}\sqrt{\sigma}\|L\|(2\beta\sigma\|L\|^2 + \theta)\tau^{3/2} \\ & + 2(\beta(2 - \theta) - \hat{\rho}(2\beta\sigma\|L\|^2 + \theta))\tau + 4\beta\hat{\rho}\sqrt{\sigma}\|L\|\tau^{1/2} + 4\beta\hat{\rho} > 0. \end{aligned} \quad (4.29)$$

Let $(z_0, v_0) \in \mathcal{H} \times \mathcal{G}$ and consider the sequence $(z_n, v_n)_{n \geq 1}$ generated by the following iteration:

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_{n+1} = J_{\tau A}(z_n - \tau(Cz_n + L^*v_n)) \\ q_{n+1} = J_{\sigma B^{-1}}(v_n + \sigma L(2p_{n+1} - z_n)) \\ (z_{n+1}, v_{n+1}) = (1 - \theta)(z_n, v_n) + \theta(p_{n+1}, q_{n+1}). \end{cases} \quad (4.30)$$

Then the following hold.

- (i) $(z_n, v_n)_{n \in \mathbb{N}}$ converges weakly to a solution to Problem 1.1.
- (ii) If $\nu_A > 0$ and $\nu_B > 0$, then $(z_n, v_n)_{n \in \mathbb{N}}$ converges R -linearly to the unique solution to Problem 1.1.

Proof. Let $\kappa = 1 - \sigma\tau\|L\|^2$ and $\bar{\chi} = 1 + \sqrt{\sigma\tau}\|L\|$. The result follows directly from Theorem 4.10 by noticing that, when $D = 0$, $\vartheta = 0$. The condition $\beta\kappa + \bar{\chi}\hat{\rho} > 0$ implies that $\kappa > 0$ and it is equivalent to (4.28). In addition, (4.23) can be reexpressed as

$$2(\beta\kappa + \bar{\chi}\hat{\rho})((2 - \theta)\tau + 2\bar{\chi}\hat{\rho}) > (\tau + 2\bar{\chi}\hat{\rho})^2 \quad (4.31)$$

$$\Leftrightarrow 2(2 - \theta)\tau\beta\kappa + 2(2\beta\kappa - \theta\tau)\bar{\chi}\hat{\rho} > \tau^2. \quad (4.32)$$

After simplification, this yields the algebraic inequality (4.29), which is a fourth-degree polynomial inequality in $\sqrt{\tau}$ when $\hat{\rho}$ is independent of τ . $\square$

Remark 4.13. When $\theta = 1$, under (4.28), (4.31) shows that (4.29) is equivalent to

$$-\hat{\rho}(1 + \sqrt{\sigma\tau}\|L\|) < \frac{\tau}{2} < \beta(1 - \sigma\tau\|L\|^2). \quad (4.33)$$

As shown below, the previous proposition recovers standard results in the monotone case, while also extending to settings where the operator A or B is hypomonotone or cohypomonotone.

Corollary 4.14. *In the context of Problem 1.1 and Assumption 4.2(v), suppose that $D = 0$, let $(z_0, v_0) \in \mathcal{H} \times \mathcal{G}$, and consider the sequence $(z_n, v_n)_{n \geq 1}$ generated by iteration (4.30) with $(\tau, \sigma) \in]0, +\infty[^2$ and $\theta \in]0, 2[$. Then, $(z_n, v_n)_{n \in \mathbb{N}}$ converges weakly to a solution to Problem 1.1 if one of the following conditions holds:*

- (i) A and B are maximally monotone and $\tau < 2\beta(2 - \theta)(1 - \sigma\tau\|L\|^2)$,
- (ii) $\rho_A < 0$, $\rho_B > -\rho_A$, A is maximally ρ_A -comonotone, B is maximally $\rho_B LL^*$ -comonotone, and (4.28)-(4.29) hold with $\hat{\rho} = \frac{\rho_A \rho_B}{\rho_A + \rho_B}$.
- (iii) $\rho_B < 0$, $\rho_A > -\rho_B$, $L \neq 0$, A is maximally ρ_A -comonotone, B is maximally $\rho_B LL^*$ -comonotone, (4.28)-(4.29) hold with $\hat{\rho} = \frac{\rho_A \rho_B}{\rho_A + \rho_B}$, and $\sigma < -\frac{1}{\rho_B \|L\|^2}$,
- (iv) $\mu_A < 0$, $\mu_B > -\mu_A$, $L \neq 0$, A is maximally $\mu_A L^* L$ -monotone, B is maximally μ_B -monotone, (4.28)-(4.29) hold with $\hat{\rho} = \frac{\tau}{\sigma} \frac{\mu_A \mu_B}{\mu_A + \mu_B}$, and $\tau < -\frac{1}{\mu_A \|L\|^2}$.
- (v) $\mu_B < 0$, $\mu_A > -\mu_B$, A is maximally $\mu_A L^* L$ -monotone, B is maximally μ_B -monotone, (4.28)-(4.29) hold with $\hat{\rho} = \frac{\tau}{\sigma} \frac{\mu_A \mu_B}{\mu_A + \mu_B}$, and $\sigma > -\mu_B$.

Proof. This result is deduced from Proposition 4.12 by noticing that, in each of the considered cases, Assumptions 4.2(i)-(iv) are satisfied.

- (i) Suppose that $\mu_A = \mu_B = \rho_A = \rho_B = 0$. We have then $\hat{\rho} = 0$.
- (ii) Suppose that $\nu_A = \nu_B = \mu_A = \mu_B = 0$ and $\rho_A + \rho_B > 0$. It follows from Proposition 4.5(i)&(iii) that $\Omega_A =]-\rho_A, +\infty[$ and $\Omega_B =]0, +\infty[$. From (4.11), $\hat{\rho} = \rho_{A,B} = \frac{\rho_A \rho_B}{\rho_A + \rho_B}$. In addition, it is deduced from (4.32) that $\tau > -2\bar{\chi}\hat{\rho}/(2 - \theta)$. The latter lower bound is larger than $-\rho_A$ since $-\hat{\rho} > -\rho_A$ and $2\bar{\chi}/(2 - \theta) > 1$. Thus, condition (4.29) implies that $\tau > -\rho_A$.
- (iii) Similarly to the previous case, we deduce from Proposition 4.5(i)&(iv) that $\Omega_A =]0, +\infty[$ and $\Omega_B =]0, -1/(\rho_B \|L\|^2)[$.
- (iv) Suppose that $\nu_A = \nu_B = \rho_A = \rho_B = 0$, and $\mu_A + \mu_B > 0$. It follows from Proposition 4.5(ii)&(iii) that $\Omega_A =]0, -1/(\mu_A \|L\|^2)[$ and $\Omega_B =]0, +\infty[$. In addition, from (4.11), $\hat{\rho} = \frac{\tau}{\sigma} \mu_{A,B} = \frac{\tau}{\sigma} \frac{\mu_A \mu_B}{\mu_A + \mu_B}$.
- (v) Similarly to the previous case, we deduce from Proposition 4.5(ii)&(iii) that $\Omega_A =]0, +\infty[$ and $\Omega_B =]-\mu_B, +\infty[$.

□

Remark 4.15. *In the particular case when $C = 0$, (4.30) corresponds to the primal-dual algorithm proposed in [12]. Then, by letting $\beta \rightarrow +\infty$, we obtain convergence conditions for this iterative method for semimonotone operators. Similar results were established in [31, Corollary 5.2].*

4.3.2. Case $B = 0$ and $L = 0$. The following result is a specialization of Theorem 3.10. It provides a comonotone version of the Forward-Half-Reflected-Backward algorithm proposed in [46] for finding a zero of $A + C + D$. We will make the following assumption on the involved operators.

Assumption 4.16. *Let $\rho_A \in \mathbb{R}$ and $\nu_A \in [0, +\infty[$, and suppose that*

- (i) $A + D$ is (ν_A, ρ_A) -semimonotone on $\{(x^*, -Cx^*) \mid -Cx^* \in (A + D)x^*\}$;
- (ii) there exists $\Omega_A \subset]0, +\infty[$ such that, for every $\tau \in \Omega_A$, $J_{\tau A}$ is single-valued and has full domain;
- (iii) $\text{gra}(A + D + C)$ is sequentially weak-strong closed.

Proposition 4.17. *Consider Problem 1.1 where $B = 0$ and $L = 0$. Let $\rho_A \in \mathbb{R}$ and $\nu_A \in [0, +\infty[$, and suppose that Assumption 4.16 holds. Let $\tau \in \Omega_A$, let $\theta \in]0, 2[$, and let $\hat{\rho} = \min(0, \rho_A)$. Let $(z_0, p_0, z_{-1}) \in \mathcal{H}^3$, set $y_{-1} = Dz_{-1}$, and consider the sequence $(z_n)_{n \geq 1}$*

generated by the following recurrence:

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = Dz_n \\ x_n = z_n - \tau(y_n - y_{n-1} + Dp_n + Cz_n) \\ p_{n+1} = J_{\tau A} x_n \\ z_{n+1} = (1 - \theta)z_n + \theta p_{n+1}. \end{cases} \quad (4.34)$$

Set η_θ as in (4.22). Moreover, assume that

$$\gamma = \beta + \hat{\rho}(1 + \vartheta) > 0, \quad (4.35)$$

$$b_2\tau^2 + b_1\tau + b_0 > 0 \quad (4.36)$$

with

$$b_0 = 4\hat{\rho}(\gamma - \hat{\rho}), \quad b_1 = 2(\gamma(2 - \theta + 8\hat{\rho}\vartheta) - 2\hat{\rho}), \quad b_2 = 16\gamma\hat{\rho}\vartheta(1 + \vartheta) - 4\gamma\vartheta\eta_\theta - 1. \quad (4.37)$$

Then, the following hold.

- (i) $(z_n)_{n \in \mathbb{N}}$ converges weakly to $z \in \text{zer}(A + C + D)$.
- (ii) If $\nu_A > 0$, then $(z_n)_{n \in \mathbb{N}}$ converges R -linearly to the unique $z \in \text{zer}(A + C + D)$.

Proof. In the considered setting, Algorithm 3.7 reduces to algorithm (4.34). We apply Theorem 4.10 with $\kappa = 1$, $\bar{\chi} = 1$, $\mu_B = 0$, and choosing ν_B and ρ_B arbitrarily large. In this context, (4.23) reduces to

$$(2 - \theta)\tau > 2(\eta_\theta - 4\hat{\rho})\tau^2\vartheta + \frac{1}{2\gamma}(\tau + 2\hat{\rho})^2 - 2\hat{\rho}(2\tau\vartheta + 1)^2 \quad (4.38)$$

After some algebra, this condition can be rewritten as (4.36). $\square$

Remark 4.18. Consider the setting of Proposition 4.17.

(i) Define

$$R(\tau) = b_2\tau^2 + b_1\tau + b_0, \quad (4.39)$$

$$\Delta = b_1^2 - 4b_2b_0. \quad (4.40)$$

Since $\hat{\rho} \leq 0$ and $\gamma > 0$, we have $b_0 \leq 0$ and $b_2 < 0$. Hence, if $b_1 \leq 0$ or $\Delta \leq 0$, $R(\tau) \leq 0$ for every $\tau > 0$. Otherwise, if $b_1 > 0$ and $\Delta > 0$, $R(\tau) > 0$ for $\tau \in]\tau_0, \tau_1[$. Because of the signs of (b_0, b_1, b_2) , it follows that $\tau_0 \geq 0$ ($\tau_0 > 0$ if $\hat{\rho} \neq 0$). Moreover,

$$\begin{aligned} \frac{\Delta}{4\gamma} &= ((2 - \theta)^2 + 16\hat{\rho}\vartheta(2 - \theta + \eta_\theta - 4\hat{\rho}))\gamma \\ &\quad + 4\hat{\rho}(\theta - 1 - 4\vartheta(2 + \eta_\theta)\hat{\rho} + 16\vartheta(1 + \vartheta)\hat{\rho}^2). \end{aligned} \quad (4.41)$$

Suppose that $\theta \in [1, 2[$ and

$$-\hat{\rho}\vartheta < \frac{(2 - \theta)^2}{16(2 - \theta + \eta_\theta - 4\hat{\rho})}, \quad \beta > -\hat{\rho} \frac{\theta^2 + \vartheta(2 - \theta)^2 + 16\vartheta\hat{\rho}(\vartheta(2 - \theta + \eta_\theta) - \theta)}{(2 - \theta)^2 + 16\hat{\rho}\vartheta(2 - \theta + \eta_\theta - 4\hat{\rho})}. \quad (4.42)$$

Then (4.35)-(4.36) hold for $\tau \in]\tau_0, \tau_1[$ with

$$0 \leq \tau_0 = \frac{-b_1 + \sqrt{\Delta}}{2b_2} < \tau_1 = \frac{-b_1 - \sqrt{\Delta}}{2b_2}. \quad (4.43)$$

Indeed, since $\theta - 1 - 4\vartheta(2 + \eta_\theta)\hat{\rho} + 16\vartheta(1 + \vartheta)\hat{\rho}^2 \geq 0$, Δ is positive if and only if

$$(2 - \theta)^2 + 16\hat{\rho}\vartheta(2 - \theta + \eta_\theta - 4\hat{\rho}) > 0 \quad \Leftrightarrow \quad -\hat{\rho}\vartheta < \frac{(2 - \theta)^2}{16(2 - \theta + \eta_\theta - 4\hat{\rho})} \quad (4.44)$$

and

$$\gamma > -4\hat{\rho} \frac{\theta - 1 + 4\vartheta\hat{\rho}(4(1 + \vartheta)\hat{\rho} - 2 - \eta_\theta)}{(2 - \theta)^2 + 16\hat{\rho}\vartheta(2 - \theta + \eta_\theta - 4\hat{\rho})}. \quad (4.45)$$

Note that (4.44) implies that

$$2 - \theta + 8\hat{\rho}\vartheta > 2 - \theta - \frac{(2 - \theta)^2}{2(2 - \theta + \eta_\theta - 4\hat{\rho})} = (2 - \theta) \frac{2 + 2\eta_\theta - \theta - 8\hat{\rho}}{2(2 - \theta + \eta_\theta - 4\hat{\rho})} > 0, \quad (4.46)$$

hence $b_1 > 0$. Besides, (4.45) implies that $\gamma > 0$, i.e., (4.35) holds. Furthermore, (4.45) is equivalent to

$$\beta > -4\hat{\rho} \frac{\theta - 1 + 4\vartheta\hat{\rho}(4(1 + \vartheta)\hat{\rho} - 2 - \eta_\theta)}{(2 - \theta)^2 + 16\hat{\rho}\vartheta(2 - \theta + \eta_\theta - 4\hat{\rho})} - \hat{\rho}(1 + \vartheta), \quad (4.47)$$

which yields the second inequality in (4.42).

(ii) In the case when $C = 0$ and $D \neq 0$, (4.34) reduces to

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = Dz_n \\ x_n = z_n - \tau(y_n - y_{n-1} + Dp_n) \\ p_{n+1} = J_{\tau A} x_n \\ z_{n+1} = (1 - \theta)z_n + \theta p_{n+1}. \end{cases} \quad (4.48)$$

This recurrence coincides with the Forward-Reflected-Backward algorithm [46] when $\theta = 1$ and $p_0 = z_0$. The corresponding convergence conditions are obtained by letting $\beta \rightarrow +\infty$. Hence, by assuming $\theta \in [1, 2]$, the convergence conditions reduce to $(2 - \theta)^2 + 16\hat{\rho}\vartheta(2 - \theta + \eta_\theta - 4\hat{\rho}) > 0$ and $\tau \in [\tau_0, \tau_1] \cap \Omega_A$, where

$$\tau_i = \frac{-(2 - \theta) - 8\hat{\rho}\vartheta + (1 - 2i)\sqrt{(2 - \theta)^2 + 16\hat{\rho}\vartheta(2 - \theta + \eta_\theta - 4\hat{\rho})}}{4\vartheta(4\hat{\rho}(1 + \vartheta) - \eta_\theta)}, \quad i \in \{0, 1\}. \quad (4.49)$$

In [46], it is further assumed that A is maximally monotone and D is monotone. In this case, we have $\hat{\rho} = 0$, $\Omega_A =]0, +\infty[$, and $(\tau_0, \tau_1) = (0, (2 - \theta)/(2\vartheta))$.

(iii) When $D = 0$, (4.34) reduces to the Forward-Backward algorithm and we recover the convergence conditions in Section 3 (see Corollary 3.12(i)).

5. NUMERICAL EXPERIMENTS

In this section, we present a numerical illustration of our theoretical results. First, we introduce the adjoint mismatch problem. Next, we provide numerical experiments in the context of signal recovery in the presence of adjoint mismatch.

5.1. Application to adjoint mismatch problems. Let $F: \mathcal{H} \rightarrow \mathcal{H}$ be a ρ -cocoercive operator for $\rho \in]0, +\infty[$ and let $T: \mathcal{H} \rightarrow \mathcal{G}$ and $K: \mathcal{G} \rightarrow \mathcal{H}$ be bounded linear operators such that $KT \neq 0$. We assume that $KT - \tilde{\rho}T^*K^*KT$ is a monotone operator with $\tilde{\rho} \in \mathbb{R}$ such that $\tilde{\rho} + \rho > 0$. We want to

$$\text{find } x \in \mathcal{H} \text{ such that } 0 = Fx + KTx. \quad (5.1)$$

By setting $D = KT$, it follows from Proposition 4.3(ii) that $F + D$ is $\rho_{F,D}$ -comonotone with $\rho_{F,D} = \rho\tilde{\rho}/(\rho + \tilde{\rho})$. The problem can be solved by using algorithms (4.34) and (4.48) under the following two splittings.

- Set $A = F$ and $C = 0$. Assumption 4.16(i) holds with $\nu_A = 0$ and $\rho_A = \rho_{F,D}$. In addition, since A is maximally monotone, Assumption 4.16(ii) is satisfied for $\Omega_A =]0, +\infty[$. Since F is ρ -cocoercive, D is $\tilde{\rho}$ -comonotone, $\rho + \tilde{\rho} > 0$, and the bounded linear operator D is weakly continuous, Lemma 2.3 shows that $\text{gra}(F + D)$ is sequentially weak-strong closed. Thus, Assumption 4.16(iii) also holds. Let $\hat{\rho} = \min\{0, \rho\tilde{\rho}/(\rho + \tilde{\rho})\}$,

let $(\tau_i)_{0 \leq i \leq 1}$ be defined by (4.49) with $\vartheta = \|KT\|$. As a consequence of Remark 4.18(ii), if i) $\theta \in [1, 2[$, ii) $(2 - \theta)^2 + 16\hat{\rho}\|KT\|(2 - \theta + \eta_\theta - 4\hat{\rho}) > 0$ and iii) $\tau \in]\tau_0, \tau_1[$, then algorithm (4.48) converges weakly to a solution to (5.1).

- Set $A = 0$ and $C = F$. Note that $0 + D$ is ρ_A -comonotone with $\rho_A = \tilde{\rho}$ and Assumption 4.16 holds with $\Omega_A =]0, +\infty[$. Let $\theta \in [1, 2[$, and suppose that condition (4.42) holds with $\hat{\rho} = \min\{0, \tilde{\rho}\}$, $\beta = \rho$, and $\vartheta = \|KT\|$. Then, according to Proposition 4.17 and Remark 4.18(i), inclusion (5.1) can be solved by algorithm (4.34) when $\tau \in]\tau_0, \tau_1[$ and (τ_0, τ_1) are the bounds given in Remark 4.18(i).

5.2. Signal recovery example. Let $N \in \mathbb{N} \setminus \{0\}$ and $\bar{x} \in \mathcal{H} = \mathbb{R}^N$ be a signal to be recovered from an observation $r \in \mathcal{G} = \mathbb{R}^M$ which is modelled as follows:

$$r = T\bar{x} + e, \quad (5.2)$$

where $M \in \mathbb{N} \setminus \{0\}$, $T \in \mathbb{R}^{M \times N}$, and e is a vector modelling an additive zero-mean white Gaussian noise. To recover the original signal from the observation r , we solve the following optimization problem:

$$\min_{x \in \mathbb{R}^N} \lambda H_\delta(Wx) + \frac{1}{2} \|Tx - r\|_2^2, \quad (5.3)$$

where H_δ is the Huber function, $\delta \in]0, +\infty[$, $\lambda \in]0, +\infty[$ is a regularization parameter, and $W \in \mathbb{R}^{N \times N}$ is an orthonormal wavelet decomposition. The Huber penalty is defined as

$$H_\delta: \mathcal{H} \rightarrow \mathbb{R}: (x_i)_{1 \leq i \leq N} \mapsto \sum_{i \in \mathbb{I}} h_\delta(x_i),$$

where

$$h_\delta: \mathbb{R} \rightarrow \mathbb{R}: t \mapsto \begin{cases} \frac{t^2}{2\delta}, & \text{if } |t| \leq \delta, \\ |t| - \frac{\delta}{2}, & \text{if } |t| > \delta. \end{cases}$$

and $\mathbb{I} \subset \{1, \dots, N\}$ is the index set of wavelet (detail) coefficients. By Fermat's rule, the considered optimization problem is equivalent to

$$\text{find } x \in \mathcal{H} \text{ such that } 0 = \lambda W^* \nabla H_\delta(Wx) + T^*(Tx - r). \quad (5.4)$$

In the case of an adjoint mismatch, T^* is not available and needs to be approximated by $K \in \mathbb{R}^{N \times M}$, and the inclusion becomes

$$\text{find } x \in \mathcal{H} \text{ such that } 0 = \lambda W^* \nabla H_\delta(Wx) + K(Tx - r), \quad (5.5)$$

which is an instance of (5.1) for $F = \lambda W^* \circ \nabla H_\delta \circ W - Kr$. Moreover, F is ρ -cocoercive for $\rho = \frac{\delta}{\lambda}$, thus, problem (5.5) can be solved by algorithms (4.48) or (4.34) under the scenarios analysed in the previous subsection.

To test these algorithms, we define T as a random normalized matrix of size $M \times N = 256 \times 256$ and a mismatched adjoint $K = T^* + sE$ where E is a rank-one perturbation matrix with unit Frobenius norm, and $s \in \mathbb{R}$. Additionally, we consider $\delta = 10^{-1}$, W is a decomposition onto an orthonormal Daubechies wavelet basis using filters of length 4 and 3 resolution levels. We choose $\tilde{\rho}$ such that $KT - \tilde{\rho}T^*K^*KT$ is monotone, in particular, the minimum eigenvalue of $Q := (KT + T^*K^*)/2$ is -2×10^{-3} and that of $Q - \tilde{\rho}T^*K^*KT$ is 4.982×10^{-4} with $\tilde{\rho} = -0.01$. Note that, in this situation, the affine data fidelity term in (5.5) is not monotone. We consider two test signals plotted in Figures 2(a) and 4(a). The components of the additive Gaussian noise e have a standard deviation of 10^{-2} . For algorithm (4.34) we consider the step size $\tau = -0.99(b_1 + \sqrt{b_1^2 - 4b_0b_2})/(2b_2)$, where b_0 , b_1 , and b_2 are defined in (4.37). For algorithm (4.48) we consider $\tau = 0.99\tau_1$, where τ_1 is defined in (4.49). To compare the behaviour of the algorithms in the monotone ($K = T^*$) and non-monotone case ($K \neq T^*$), we also solve the inclusion in (5.4) with algorithms (4.48) and (4.34) for $\hat{\rho} = 0$. As stopping criterion, we use

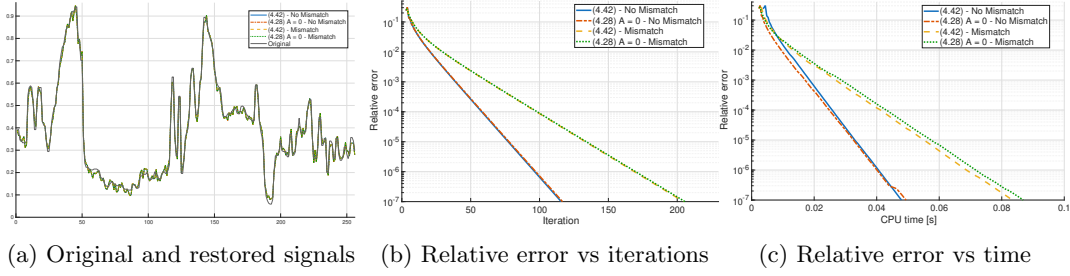


FIGURE 2. $\lambda = 0.005$ and $\theta = 1$. The PSNR values of all restored signals are approximately 35.23 dB.

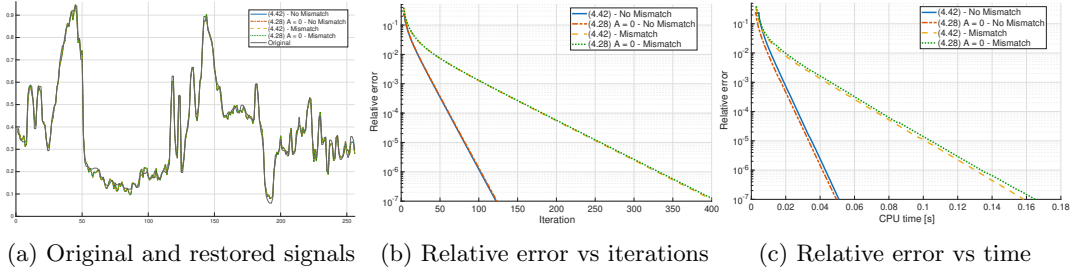


FIGURE 3. $\lambda = 0.005$ and $\theta = 1.3$. The PSNR values of all restored signals are approximately 35.23 dB.

the relative error expressed as $\|x_{n+1} - x_n\|/\|x_n\|$ with tolerance 10^{-7} . All algorithms were initialized with $z_{-1} = z_0 = p_0 = 0$. The results obtained for the first signal are presented in Figures 2 and 3 for $\theta = 1$ and $\theta = 1.3$, respectively. Figures 4 and 5 show the results for the second signal for $\theta = 1$ and $\theta = 1.2$, respectively. From Figures 2(a), 3(a), 4(a), and 5(a) we observe that the algorithm with and without mismatch recover similar signals. In Figures 2(b), 3(b), 4(b), and 5(b), we observe that both algorithms (4.34) and (4.48) present a similar decay of the relative error versus the iteration number or computational time for $\lambda = 0.005$. On the other hand, for $\lambda = 0.5$, from Figures 4(b), 5(b), 4(c) and 5(c), we observe that algorithm (4.48) presents a faster decay of the relative error. In all cases the algorithms with no mismatch are faster in terms of number of iterations or computational time. This is related to the fact that, in the considered scenarios, the negative comonotonicity parameter induced by the mismatch leads to smaller admissible step sizes, which explains the slower convergence we observe. Note that, in our experiments, increasing θ leads to more iterations to reach the stopping criterion. This can also be explained by the fact that larger values of θ impose smaller step-sizes.

6. CONCLUSION

In this article, we developed a nonlinear Forward-Backward algorithm for solving non-monotone inclusions involving Lipschitz continuous operators. We established the convergence of the method under a comonotonicity assumption and analyzed the conditions on the algorithm parameters in detail. In addition, we provided explicit conditions ensuring the comonotonicity of sums of operators and primal-dual operators. As a consequence, we established convergence guarantees for several methods from the literature for solving comonotone inclusions, including Forward-Backward, Chambolle-Pock, Condat-Vũ, and Forward-Reflected-Backward. Finally,

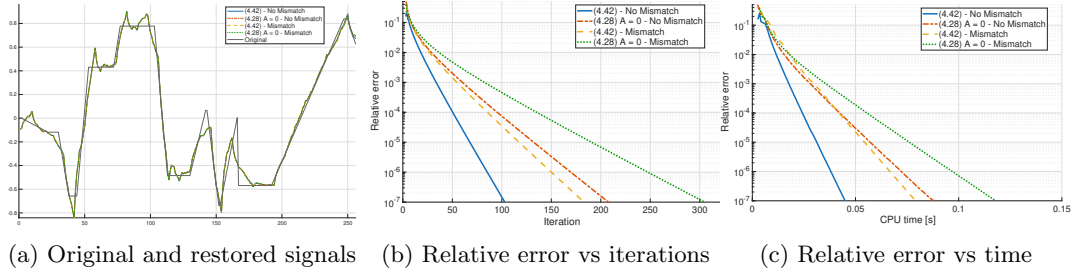


FIGURE 4. $\lambda = 0.5$ and $\theta = 1$. The PSNR values of all restored signals are approximately 22.92 dB.

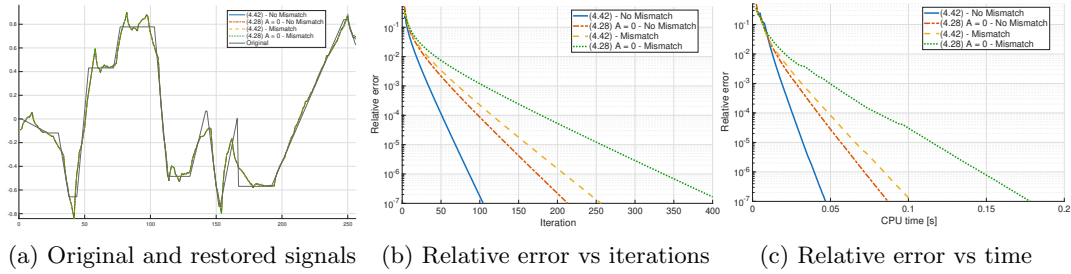


FIGURE 5. $\lambda = 0.5$ and $\theta = 1.2$. The PSNR values of all restored signals are approximately 22.92 dB.

we presented an application to adjoint mismatch problems and illustrated our theoretical results through a signal recovery example.

ACKNOWLEDGMENT

The second author was partially supported by ANID through FONDECYT Iniciación Grant 11250164.

REFERENCES

- [1] Alacaoglu, A., Kim, D., Wright, S.J.: Revisiting inexact fixed-point iterations for min-max problems: stochasticity and structured nonconvexity. In: Proceedings of the 41st International Conference on Machine Learning, ICML'24. JMLR.org (2024)
- [2] Briceño Arias, L.M., Combettes, P.L.: Monotone operator methods for Nash equilibria in non-potential games. In: Computational and analytical mathematics, *Springer Proc. Math. Stat.*, vol. 50, pp. 143–159. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-7621-4_9
- [3] Attouch, H., Briceño-Arias, L.M., Combettes, P.L.: A parallel splitting method for coupled monotone inclusions. *SIAM J. Control Optim.* **48**(5), 3246–3270 (2009/10). <https://doi.org/10.1137/090754297>
- [4] Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces, second edn. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer (2017). <https://doi.org/10.1007/978-3-319-48311-5>
- [5] Bauschke, H.H., Moursi, W.M., Wang, X.: Generalized monotone operators and their averaged resolvents. *Math. Program.* **189**(1-2, Ser. B), 55–74 (2021). <https://doi.org/10.1007/s10107-020-01500-6>
- [6] Boţ, R.I., Hendrich, C.: Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization. *J. Math. Imaging Vision* **49**(3), 551–568 (2014). <https://doi.org/10.1007/s10851-013-0486-8>
- [7] Böhm, A.: Solving nonconvex-nonconcave min-max problems exhibiting weak minty solutions. *Transactions on Machine Learning Research* (3) (2023). <https://openreview.net/forum?id=Gp0pHyUyrb>

- [8] Briceño-Arias, L.M., Combettes, P.L., Pesquet, J.C., Pustelnik, N.: Proximal algorithms for multicomponent image recovery problems. *J. Math. Imaging Vision* **41**(1-2), 3–22 (2011). <https://doi.org/10.1007/s10851-010-0243-1>
- [9] Briceño-Arias, L.M., Davis, D.: Forward-backward-half forward algorithm for solving monotone inclusions. *SIAM J. Optim.* **28**(4), 2839–2871 (2018). <https://doi.org/10.1137/17M1120099>
- [10] Bui, M.N., Combettes, P.L.: Warped proximal iterations for monotone inclusions. *J. Math. Anal. Appl.* **491**(1), 124,315, 21 (2020). <https://doi.org/10.1016/j.jmaa.2020.124315>
- [11] Bui, M.N., Combettes, P.L.: Multivariate monotone inclusions in saddle form. *Math. Oper. Res.* **47**(2), 1082–1109 (2022). <https://doi.org/10.1287/moor.2021.1161>
- [12] Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision* **40**(1), 120–145 (2011). <https://doi.org/10.1007/s10851-010-0251-1>
- [13] Chorobura, F., Necoara, I., Pesquet, J.C.: An adaptive forward-backward-forward splitting algorithm for solving pseudo-monotone inclusions. *Mathematics of Operations Research* (2026). <https://doi.org/10.1287/moor.2025.0931>
- [14] Chouzenoux, E., Contreras, A., Pesquet, J.C., Savanier, M.: Convergence results for primal-dual algorithms in the presence of adjoint mismatch. *SIAM J. Imaging Sci.* **16**(1), 1–34 (2023). <https://doi.org/10.1137/22M1490223>
- [15] Chouzenoux, E., Pesquet, J.C., Riddell, C., Savanier, M., Trouset, Y.: Convergence of proximal gradient algorithm in the presence of adjoint mismatch. *Inverse Problems* **37**(6), Paper No. 065,009, 29 (2021). <https://doi.org/10.1088/1361-6420/abd85c>
- [16] Chouzenoux, E., Pesquet, J.C., Roldán, F.: Solution of mismatched monotone+Lipschitz inclusion problems. *SIAM J. Optim.* **34**(4), 3564–3591 (2024). <https://doi.org/10.1137/23M1609166>
- [17] Combettes, P.L.: Systems of structured monotone inclusions: duality, algorithms, and applications. *SIAM J. Optim.* **23**(4), 2420–2447 (2013). <https://doi.org/10.1137/130904160>
- [18] Combettes, P.L.: Monotone operator theory in convex optimization. *Math. Program.* **170**(1), 177–206 (2018). <https://doi.org/10.1007/s10107-018-1303-3>
- [19] Combettes, P.L.: The geometry of monotone operator splitting methods. *Acta Numer.* **33**, 487–632 (2024). <https://doi.org/10.1017/S0962492923000065>
- [20] Combettes, P.L., Eckstein, J.: Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions. *Math. Program.* **168**(1-2, Ser. B), 645–672 (2018). <https://doi.org/10.1007/s10107-016-1044-0>
- [21] Combettes, P.L., Pennanen, T.: Proximal methods for cohypomonotone operators. *SIAM J. Control Optim.* **43**(2), 731–742 (2004). <https://doi.org/10.1137/S0363012903427336>
- [22] Combettes, P.L., Pesquet, J.C.: Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators. *Set-Valued Var. Anal.* **20**(2), 307–330 (2012). <https://doi.org/10.1007/s11228-011-0191-y>
- [23] Combettes, P.L., Pesquet, J.C.: Fixed point strategies in data science. *IEEE Transactions on Signal Processing* **69**, 3878–3905 (2021). <https://doi.org/10.1109/TSP.2021.3069677>
- [24] Condat, L.: A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.* **158**(2), 460–479 (2013). <https://doi.org/10.1007/s10957-012-0245-9>
- [25] Davis, D., Yin, W.: A three-operator splitting scheme and its optimization applications. *Set-Valued Var. Anal.* **25**(4), 829–858 (2017). <https://doi.org/10.1007/s11228-017-0421-z>
- [26] Diakonikolas, J., Daskalakis, C., Jordan, M.I.: Efficient methods for structured nonconvex-nonconcave min-max optimization. In: *International Conference on Artificial Intelligence and Statistics* (2020)
- [27] Dong, Y., Hansen, P.C., Hochstenbach, M.E., Brogaard Riis, N.A.: Fixing nonconvergence of algebraic iterative reconstruction with an unmatched backprojector. *SIAM Journal on Scientific Computing* **41**(3), A1822–A1839 (2019). <https://doi.org/10.1137/18M1206448>
- [28] Douglas Jr., J., Rachford Jr., H.H.: On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.* **82**, 421–439 (1956). <https://doi.org/10.2307/1993056>
- [29] Eckstein, J., Bertsekas, D.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.* **55**(1), 293–318 (1992)
- [30] Elfving, T., Hansen, P.C.: Unmatched projector/backprojector pairs: Perturbation and convergence analysis. *SIAM Journal on Scientific Computing* **40**(1), A573–A591 (2018). <https://doi.org/10.1137/17M1133828>
- [31] Evens, B., Latafat, P., Patrinos, P.: Convergence of the Chambolle-Pock algorithm in the absence of monotonicity. *J. Optim. Theory Appl.* **206**(1), Paper No. 7, 45 (2025). <https://doi.org/10.1007/s10957-025-02680-x>

- [32] Evens, B., Pas, P., Latafat, P., Patrinos, P.: Convergence of the preconditioned proximal point method and Douglas–Rachford splitting in the absence of monotonicity. *Math. Program.* **214**(1-2), 247–301 (2025). <https://doi.org/10.1007/s10107-024-02182-0>
- [33] Giselsson, P.: Nonlinear forward-backward splitting with projection correction. *SIAM J. Optim.* **31**(3), 2199–2226 (2021). <https://doi.org/10.1137/20M1345062>
- [34] Grad, S.M., Lara, F.: An extension of the proximal point algorithm beyond convexity. *J. Global Optim.* **82**(2), 313–329 (2022). <https://doi.org/10.1007/s10898-021-01081-4>
- [35] Grad, S.M., Lara, F., Marcavillaca, R.T.: Proximal point type algorithms with relaxed and inertial effects beyond convexity. *Optimization* **73**(11), 3393–3410 (2024). <https://doi.org/10.1080/02331934.2024.2329779>
- [36] Iusem, A., Lara, F.: Proximal point algorithms for quasiconvex pseudomonotone equilibrium problems. *J. Optim. Theory Appl.* **193**(1-3), 443–461 (2022). <https://doi.org/10.1007/s10957-021-01951-7>
- [37] Iusem, A.N., Pennanen, T., Svaiter, B.F.: Inexact variants of the proximal point algorithm without monotonicity. *SIAM J. Optim.* **13**(4), 1080–1097 (2003). <https://doi.org/10.1137/S1052623401399587>
- [38] Kohlenbach, U.: On the proximal point algorithm and its Halpern-type variant for generalized monotone operators in Hilbert space. *Optim. Lett.* **16**(2), 611–621 (2022). <https://doi.org/10.1007/s11590-021-01738-9>
- [39] Lara, F.: On strongly quasiconvex functions: existence results and proximal point algorithms. *J. Optim. Theory Appl.* **192**(3), 891–911 (2022). <https://doi.org/10.1007/s10957-021-01996-8>
- [40] Lee, S., Kim, D.: Fast extra gradient methods for smooth structured nonconvex-nonconcave minimax problems. In: *Neural Information Processing Systems* (2021)
- [41] Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16**(6), 964–979 (1979). <https://doi.org/10.1137/0716071>
- [42] Lorenz, D.A., Rose, S., Schöpfer, F.: The randomized Kaczmarz method with mismatched adjoint. *BIT* **58**(4), 1079–1098 (2018). <https://doi.org/10.1007/s10543-018-0717-x>
- [43] Lorenz, D.A., Schneppe, F.: Chambolle-Pock’s primal-dual method with mismatched adjoint. *Appl. Math. Optim.* **87**(2), Paper No. 22, 26 (2023). <https://doi.org/10.1007/s00245-022-09933-5>
- [44] Loris, I., Verhoeven, C.: On a generalization of the iterative soft-thresholding algorithm for the case of non-separable penalty. *Inverse Problems* **27**(12), 125,007, 15 (2011). <https://doi.org/10.1088/0266-5611/27/12/125007>
- [45] Luke, D.R., Tam, M.K.: Generalized monotonicity and the proximal point algorithm. *Mathematics of Operations Research* **0**(0), null (0). <https://doi.org/10.1287/moor.2025.0863>
- [46] Malitsky, Y., Tam, M.K.: A forward-backward splitting method for monotone inclusions without cocoercivity. *SIAM J. Optim.* **30**(2), 1451–1472 (2020). <https://doi.org/10.1137/18M1207260>
- [47] Martinet, B.: Régularisation d’inéquations variationnelles par approximations successives. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique* **4**(R3), 154–158 (1970)
- [48] Morin, M., Banert, S., Giselsson, P.: Nonlinear Forward-Backward Splitting with Momentum Correction. *Set-Valued Var. Anal.* **31**(4), 37 (2023). <https://doi.org/10.1007/s11228-023-00700-4>
- [49] Papadimitriou, D., Vu, B.C.: On the convergence of warped proximal iterations for solving nonmonotone inclusions and applications. In: *OPT 2023: Optimization for Machine Learning* (2023). <https://openreview.net/forum?id=6tFG0XyykI>
- [50] Passty, G.: Ergodic convergence to a zero of the sum of monotone operators in hilbert space. *J. Math. Anal. Appl.* **72**(2), 383–390 (1979)
- [51] Qin, L., Gibali, A., Zhang, C., Tang, Y.: Nonlinear three-operator splitting algorithms with momentum for monotone inclusions (2025). <https://arxiv.org/abs/2511.14050>
- [52] Qin, L., Tang, Y., Peng, J.: Nonlinear forward-backward-half forward splitting with momentum for monotone inclusions (2025). <https://arxiv.org/abs/2510.24489>
- [53] Roldán, F.: Forward-primal-dual-half-forward algorithm for splitting four operators. *J. Optim. Theory Appl.* **204**(1), Paper No. 11, 26 (2025). <https://doi.org/10.1007/s10957-024-02560-w>
- [54] Ryu, E.K., Vũ, B.C.: Finding the forward-Douglas-Rachford-forward method. *J. Optim. Theory Appl.* **184**(3), 858–876 (2020). <https://doi.org/10.1007/s10957-019-01601-z>
- [55] Savanier, M., Chouzenoux, E., Pesquet, J.C., Riddell, C.: Unmatched preconditioning of the proximal gradient algorithm. *IEEE Signal Processing Letters* **29**, 1122–1126 (2022). <https://doi.org/10.1109/LSP.2022.3169088>
- [56] Tran-Dinh, Q.: Extragradient-type methods with $\mathcal{O}(1/k)$ last-iterate convergence rates for co-hypomonotone inclusions. *J. Global Optim.* **89**(1), 197–221 (2024). <https://doi.org/10.1007/s10898-023-01347-z>

- [57] Trang, T.T., Nguyen, V.D.: Fast warped iteration for nonmonotone inclusions with a convergence rate of $o(\frac{1}{k})$. *Optim. Methods Softw.* **41**(1), 256–275 (2026). <https://doi.org/10.1080/10556788.2026.2617626>
- [58] Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.* **38**(2), 431–446 (2000). <https://doi.org/10.1137/S0363012998338806>
- [59] Vũ, B.C.: A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comput. Math.* **38**(3), 667–681 (2013). <https://doi.org/10.1007/s10444-011-9254-8>
- [60] Zeng, G., Gullberg, G.: Unmatched projector/backprojector pairs in an iterative reconstruction algorithm. *IEEE Transactions on Medical Imaging* **19**(5), 548–555 (2000). <https://doi.org/10.1109/42.870265>

Centro de Investigación en Ingeniería Matemática (CI²MA)

PRE-PUBLICACIONES 2026

- 2026-16 GABRIEL N. GATICA, SALIM MEDDAHI, KEVIN W. PUCHA-ATAN, RICARDO RUIZ-BAIER: *A Banach spaces-based fully mixed finite element method for the thermo-electro-hydrodynamic Boussinesq problem*
- 2026-17 TOMÁS BARRIOS, EDWIN BEHRENS, ROMMEL BUSTINZA, JOSE M. CASCON: *A stabilized displacement - stress formulation for a linear elasticity problema with mixed boundary conditions*
- 2026-18 GONZALO A. BENAVIDES, SERGIO CAUCAO, GABRIEL N. GATICA, YURI D. SOBRAL: *New Banach spaces-based mixed finite element methods for steady-state flows of magnetic fluids*
- 2026-19 FAHIM ASLAM, MARCELO CAVALCANTI, ZAYD HAJJEJ, YOU HAIYANG LIN BO, MAURICIO SEPÚLVEDA: *Blow-up of a Lamé system with fractional damping and infinite memory: theoretical and numerical study*
- 2026-20 GABRIEL N. GATICA, SALIM MEDDAHI: *A hybridizable discontinuous Galerkin method for the elasto-acoustic transmission problem*
- 2026-21 IDULFO ARROCHA, RAIMUND BÜRGER, MARÍA GÓMEZ, YOLANDA VÁSQUEZ: *Modeling and numerical simulation of the colonization of plant growth-promoting bacteria in the rice rhizosphere*
- 2026-22 ABRAHAM J. ARENAS, JUAN BARAJAS-CALONGE, GILBERTO GONZÁLEZ-PARRA, LUIS M. VILLADA: *Convergence analysis of a nonstandard finite volume method for a diffusion-production-destruction model*
- 2026-23 JUAN BARAJAS-CALONGE, RAIMUND BÜRGER, STEFAN DIEHL, LUIS M. VILLADA: *A network of conservation laws with discontinuous fluxes and source terms modeling countercurrent decantation*
- 2026-24 RAIMUND BÜRGER, BAKHTIYOR KHUZHAYOROV, USMONALI SAYDULLAEV, IRODA BEKNAZAROVA: *Numerical simulation of filtration of a suspension through a radial filter with given flow rate*
- 2026-25 FERNANDO MUÑOZ GARCÍA, FERNANDO ROLDÁN: *Forward-reflected-backward algorithm with linesearch*
- 2026-26 EIDER ALDANA, RICARDO OYARZÚA: *A Lagrange multiplier approach for fluid-flow problem with inhomogeneous mixed boundary conditions*
- 2026-27 JEAN-CHRISTOPHE PESQUET, FERNANDO ROLDÁN: *Nonlinear forward-backward algorithm for solving non-monotone+Lipschitz inclusions with applications to adjoint mismatch problems*

Para obtener copias de las Pre-Publicaciones, escribir o llamar a: DIRECTOR, CENTRO DE INVESTIGACIÓN EN INGENIERÍA MATEMÁTICA, UNIVERSIDAD DE CONCEPCIÓN, CASILLA 160-C, CONCEPCIÓN, CHILE, TEL.: 41-2661324, o bien, visitar la página web del centro: <http://www.ci2ma.udec.cl>



**CENTRO DE INVESTIGACIÓN EN
INGENIERÍA MATEMÁTICA (CI²MA)
Universidad de Concepción**



Casilla 160-C, Concepción, Chile
Tel.: 56-41-2661324/2661554/2661316
<http://www.ci2ma.udec.cl>

