Warped Proximal Iterations for Monotone Inclusions*

Minh N. Bùi and Patrick L. Combettes

North Carolina State University, Department of Mathematics, Raleigh, NC 27695-8205, USA
mnbui@ncsu.edu and plc@math.ncsu.edu

Abstract. Resolvents of set-valued operators play a central role in various branches of mathematics and in particular in the design and the analysis of splitting algorithms for solving monotone inclusions. We propose a generalization of this notion, called warped resolvent, which is constructed with the help of an auxiliary operator. The properties of warped resolvents are investigated and connections are made with existing notions. Abstract weak and strong convergence principles based on warped resolvents are proposed and shown to not only provide a synthetic view of splitting algorithms but to also constitute an effective device to produce new solution methods for challenging inclusion problems.

Keywords. Monotone inclusion, operator splitting, strong convergence, warped resolvent, warped proximal iterations.

2010 Mathematics Subject Classification: 47J25, 47N10, 47H05, 90C25.

1 Introduction

A generic problem in nonlinear analysis and optimization is to find a zero of a maximally monotone operator $M\colon \mathcal{X} \to 2^{\mathcal{X}}$, where $\mathcal{X}$ is a real Hilbert space. The most elementary method designed for this task is the proximal point algorithm [34]

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma_n M} x_n, \quad \text{where} \quad \gamma_n \in \left]0,+\infty\right[\quad \text{and} \quad J_{\gamma_n M} = (\mathrm{Id} + \gamma_n M)^{-1}. \tag{1.1}$$

In practice, the execution of (1.1) may be hindered by the difficulty of evaluating the resolvents $(J_{\gamma_n M})_{n\in\mathbb{N}}$. Thus, even in the simple case when M is the sum of two monotone operators A and B, there is no mechanism to express conveniently the resolvent of M in terms of operators involving A and B separately. To address this issue, various splitting strategies have been proposed to handle increasingly complex formulations in which M is a composite operator assembled from several elementary blocks that can be linear operators and monotone operators [5, 7, 9, 10, 11, 12, 17, 18, 21, 22, 30, 37]. In the present paper, we explore a different path by placing at the core of our analysis the following extension of the classical notion of a resolvent.

Definition 1.1 (Warped resolvent) Let $\mathcal{X}$ be a reflexive real Banach space with topological dual $\mathcal{X}^*$, let D be a nonempty subset of $\mathcal{X}$, let $K\colon D \to \mathcal{X}^*$, and let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be such that $\operatorname{ran} K \subset \operatorname{ran}(K+M)$ and $K+M$ is injective (see Definition 2.1). The warped resolvent of M with kernel K is $J_M^K = (K+M)^{-1} \circ K$.

*Contact author: P. L. Combettes, plc@math.ncsu.edu, phone: +1 (919) 515 2671. The work of M. N. Bùi was supported by the National Science Foundation under grant CCF-1715671 and the work of P. L. Combettes was supported by the National Science Foundation under grant DMS-1818946.

A main motivation for introducing warped resolvents is that, through judicious choices of a kernel K tailored to the structure of an inclusion problem, one can create simple patterns to design and analyze new, flexible, and modular splitting algorithms. At the same time, the theory required to analyze the static properties of warped resolvents as nonlinear operators, as well as the dynamics of algorithms using them, needs to be developed as it cannot be extrapolated from the classical case, where K is simply the identity operator. In the present paper, this task is undertaken and we illustrate the pertinence of warped iteration methods through applications to challenging monotone inclusion problems.

The paper is organized as follows. Section 2 is dedicated to notation and background. In Section 3, we provide important illustrations of Definition 1.1 and make connections with constructions found in the literature. The properties of warped resolvents are also discussed in that section. Weakly and strongly convergent warped proximal iteration methods are introduced and analyzed in Section 4. Besides the use of kernels varying at each iteration, our framework also features evaluations of warped resolvents at points that may not be the current iterate, which adds considerable flexibility and models in particular inertial phenomena and other perturbations. New splitting algorithms resulting from the proposed warped iteration constructs are devised in Section 5 to solve monotone inclusions.

2 Notation and background

Throughout the paper, $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are reflexive real Banach spaces. We denote the canonical pairing between $\mathcal{X}$ and its topological dual $\mathcal{X}^*$ by $\langle\cdot,\cdot\rangle$, and by Id the identity operator. The weak convergence of a sequence $(x_n)_{n\in\mathbb{N}}$ to x is denoted by $x_n \rightharpoonup x$, while $x_n \to x$ denotes its strong convergence. The space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ is denoted by $\mathcal{B}(\mathcal{X},\mathcal{Y})$, and we set $\mathcal{B}(\mathcal{X}) = \mathcal{B}(\mathcal{X},\mathcal{X})$.

Let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. We denote by $\operatorname{gra} M = \big\{(x,x^*) \in \mathcal{X}\times\mathcal{X}^* \mid x^* \in Mx\big\}$ the graph of M, by $\operatorname{dom} M = \big\{x \in \mathcal{X} \mid Mx \neq \varnothing\big\}$ the domain of M, by $\operatorname{ran} M = \big\{x^* \in \mathcal{X}^* \mid (\exists\, x \in \mathcal{X})\; x^* \in Mx\big\}$ the range of M, by $\operatorname{zer} M = \big\{x \in \mathcal{X} \mid 0 \in Mx\big\}$ the set of zeros of M, and by M^{-1} the inverse of M, i.e., $\operatorname{gra} M^{-1} = \big\{(x^*,x) \in \mathcal{X}^*\times\mathcal{X} \mid x^* \in Mx\big\}$. Further, M is monotone if

$$\big(\forall (x,x^*) \in \operatorname{gra} M\big)\big(\forall (y,y^*) \in \operatorname{gra} M\big) \quad \langle x-y, x^*-y^*\rangle \geqslant 0, \tag{2.1}$$

and maximally monotone if, in addition, there exists no monotone operator $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ such that $\operatorname{gra} M \subset \operatorname{gra} A \neq \operatorname{gra} M$. We say that M is uniformly monotone with modulus $\phi\colon\; [0,+\infty[\, \to [0,+\infty]$ if ϕ is increasing, vanishes only at 0, and

$$\big(\forall (x,x^*) \in \operatorname{gra} M\big)\big(\forall (y,y^*) \in \operatorname{gra} M\big) \quad \langle x-y, x^*-y^*\rangle \geqslant \phi\big(\|x-y\|\big). \tag{2.2}$$

In particular, M is strongly monotone with constant $\alpha \in\,]0,+\infty[$ if it is uniformly monotone with modulus $\phi = \alpha|\cdot|^2$.

Definition 2.1 An operator $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ is injective if $(\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\; Mx \cap My \neq \varnothing \Rightarrow x = y$.

The following lemma, which concerns a type of duality for monotone inclusions studied in [20, 29, 32], will be instrumental.

Lemma 2.2 *Let $A\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ and $B\colon \mathcal{Z} \to 2^{\mathcal{Z}^*}$ be maximally monotone, let $L \in \mathcal{B}(\mathcal{Y},\mathcal{Z})$, let $s^* \in \mathcal{Y}^*$, and let $r \in \mathcal{Z}$. Suppose that $\mathcal{X} = \mathcal{Y}\times\mathcal{Z}\times\mathcal{Z}^*$ (hence $\mathcal{X}^* = \mathcal{Y}^*\times\mathcal{Z}^*\times\mathcal{Z}$), define*

$$M\colon \mathcal{X} \to 2^{\mathcal{X}^*}\colon (x,y,v^*) \mapsto (-s^* + Ax + L^*v^*) \times (By - v^*) \times \{r - Lx + y\}, \tag{2.3}$$

and set $Z = \big\{(x,v^*) \in \mathcal{Y} \times \mathcal{Z}^* \mid s^* - L^*v^* \in Ax \text{ and } Lx - r \in B^{-1}v^*\big\}$*. In addition, denote by* $\mathscr{P}$ *the set of solutions to the primal problem*

$$\text{find } x \in \mathcal{Y} \text{ such that } s^* \in Ax + L^*\big(B(Lx-r)\big), \tag{2.4}$$

and by $\mathscr{D}$ *the set of solutions to the dual problem*

$$\text{find } v^* \in \mathcal{Z}^* \text{ such that } -r \in -L\big(A^{-1}(s^*-L^*v^*)\big) + B^{-1}v^*. \tag{2.5}$$

Then the following hold:

(i) Z *is a closed convex subset of* $\mathscr{P} \times \mathscr{D}$.

(ii) M *is maximally monotone.*

(iii) *Suppose that* $(x,y,v^*) \in \operatorname{zer} M$. *Then* $(x,v^*) \in Z$, $x \in \mathscr{P}$, *and* $v^* \in \mathscr{D}$.

(iv) $\mathscr{P} \neq \varnothing \Leftrightarrow \mathscr{D} \neq \varnothing \Leftrightarrow Z \neq \varnothing \Leftrightarrow \operatorname{zer} M \neq \varnothing$.

Proof. (i): [20, Proposition 2.1(i)(a)].

(ii): Define

$$\begin{cases} C\colon \mathcal{X} \to 2^{\mathcal{X}^*}\colon (x,y,v^*) \mapsto (-s^*+Ax) \times By \times \{r\} \\ S\colon \mathcal{X} \to \mathcal{X}^*\colon (x,y,v^*) \mapsto (L^*v^*, -v^*, -Lx+y). \end{cases} \tag{2.6}$$

It follows from the maximal monotonicity of A and B that C is maximally monotone. On the other hand, S is linear and bounded, and

$$\big(\forall (x,y,v^*) \in \mathcal{X}\big) \quad \big\langle (x,y,v^*), S(x,y,v^*)\big\rangle = \langle x, L^*v^*\rangle - \langle y, v^*\rangle + \langle y - Lx, v^*\rangle = 0. \tag{2.7}$$

Thus, we derive from [35, Section 17] that S is maximally monotone with $\operatorname{dom} S = \mathcal{X}$. In turn, [35, Theorem 24.1(a)] asserts that $M = C + S$ is maximally monotone.

(iii): We deduce from (2.3) that $s^* \in Ax + L^*v^*$, $v^* \in By$, and $y = Lx - r$; hence $v^* \in B(Lx-r)$. Consequently, $s^* - L^*v^* \in Ax$ and $Lx - r \in B^{-1}v^*$, which yields $(x,v^*) \in Z$. Finally, (i) entails that $x \in \mathscr{P}$ and $v^* \in \mathscr{D}$.

(iv): By [20, Proposition 2.1(i)(c)], $\mathscr{P} \neq \varnothing \Leftrightarrow \mathscr{D} \neq \varnothing \Leftrightarrow Z \neq \varnothing$. In addition, in view of (iii), $\operatorname{zer} M \neq \varnothing \Rightarrow Z \neq \varnothing$. Suppose that $(x,v^*) \in Z$ and set $y = Lx - r$. Then $y = Lx - r \in B^{-1}v^*$ and $s^* \in Ax + L^*v^*$. Hence $0 \in By - v^*$ and $0 \in -s^* + Ax + L^*v^*$. Altogether, $0 \in (-s^* + Ax + L^*v^*) \times (By - v^*) \times \{r - Lx + y\} = M(x,y,v^*)$, i.e., $(x,y,v^*) \in \operatorname{zer} M$. ☐

Now suppose that $\mathcal{X}$ is a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$. An operator $T\colon \mathcal{X} \to \mathcal{X}$ is nonexpansive if it is 1-Lipschitzian, α-averaged with $\alpha \in \left]0,1\right[$ if $\operatorname{Id} + (1/\alpha)(T - \operatorname{Id})$ is nonexpansive, firmly nonexpansive if it is $1/2$-averaged, and β-cocoercive with $\beta \in \left]0,+\infty\right[$ if βT is firmly nonexpansive. Averaged operators were introduced in [4]. The projection operator onto a nonempty closed convex subset C of $\mathcal{X}$ is denoted by proj_C. The resolvent of $M\colon \mathcal{X} \to 2^{\mathcal{X}}$ is $J_M = (\operatorname{Id} + M)^{-1}$.

3 Warped resolvents

We provide illustrations of Definition 1.1 and then study the properties of warped resolvents.

Our first example is the warped resolvent of a subdifferential. This leads to the following notion, which extends Moreau's classical proximity operator in Hilbert spaces [28].

Example 3.1 (Warped proximity operator) Let D be a nonempty subset of $\mathcal{X}$, let $K\colon D \to \mathcal{X}^*$, and let $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be a proper lower semicontinuous convex function such that $\operatorname{ran} K \subset \operatorname{ran}(K+\partial\varphi)$ and $K+\partial\varphi$ is injective. The warped proximity operator of φ with kernel K is $\operatorname{prox}_\varphi^K = (K+\partial\varphi)^{-1} \circ K$. It is characterized by the variational inequality

$$\big(\forall (x,p) \in \mathcal{X}\times\mathcal{X}\big) \quad p = \operatorname{prox}_\varphi^K x \quad\Leftrightarrow\quad (\forall y \in \mathcal{X}) \quad \langle y-p, Kx-Kp\rangle + \varphi(p) \leqslant \varphi(y). \tag{3.1}$$

In particular, in the case of normal cones, we arrive at the following definition (see Figure 1).

Example 3.2 (Warped projection operator) Let D be a nonempty subset of $\mathcal{X}$, let $K\colon D \to \mathcal{X}^*$, and let C be a nonempty closed convex subset of $\mathcal{X}$ with normal cone operator N_C such that $\operatorname{ran} K \subset \operatorname{ran}(K+N_C)$ and $K+N_C$ is injective. The warped projection operator onto C with kernel K is $\operatorname{proj}_C^K = (K+N_C)^{-1} \circ K$. The warped projection onto C is characterized by

$$\big(\forall (x,p) \in \mathcal{X}\times\mathcal{X}\big) \quad p = \operatorname{proj}_C^K x \quad\Leftrightarrow\quad \big[\, p \in C \quad \text{and} \quad (\forall y \in C) \quad \langle y-p, Kx-Kp\rangle \leqslant 0 \,\big]. \tag{3.2}$$

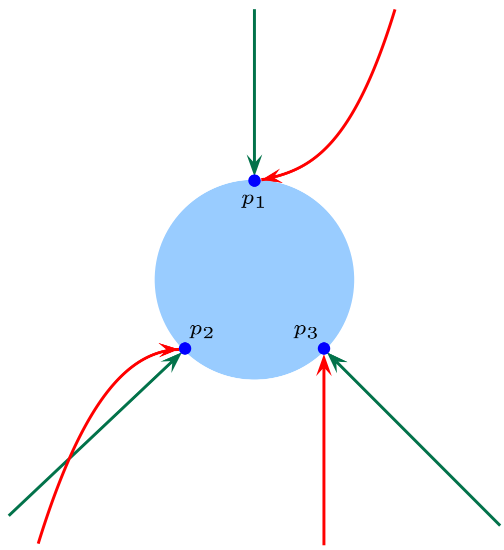

Fig. 1. Warped projections onto the closed unit ball C centered at the origin in the Euclidean plane. Sets of points projecting onto p_1, p_2, and p_3 for the kernels $K_1 = \mathrm{Id}$ (in green) and $K_2\colon (\xi_1,\xi_2) \mapsto (\xi_1^3/2 + \xi_1/5 - \xi_2, \xi_1+\xi_2)$ (in red). Note that K_2 is not a gradient.

Example 3.3 Suppose that $\mathcal{X}$ is strictly convex, let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone, and let K be the normalized duality mapping of $\mathcal{X}$. Then J_M^K is a well-defined warped resolvent which was introduced in [26].

Example 3.4 Let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone such that $\operatorname{zer} M \neq \varnothing$, let $f\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be a Legendre function [6] such that $\operatorname{dom} M \subset \operatorname{int}\operatorname{dom} f$, and set $K = \nabla f$. Then it follows from [6, Corollary 3.14(ii)] that J_M^K is a well-defined warped resolvent, called the D-resolvent of M in [6].

Example 3.5 Let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone and let $K\colon \mathcal{X} \to \mathcal{X}^*$ be strictly monotone, surjective, and 3^* monotone in the sense that [39, Definition 32.40(c)]

$$(\forall x \in \mathcal{X})(\forall x^* \in \operatorname{ran} K) \quad \sup_{y\in\mathcal{X}} \langle x-y, Ky-x^*\rangle < +\infty. \tag{3.3}$$

Then it follows from [8, Theorem 2.3] that J_M^K is a well-defined warped resolvent, called the K-resolvent of M in [8].

Example 3.6 Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and $B\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone, and let $f\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be a proper lower semicontinuous convex function which is essentially smooth [6]. Suppose that $D = (\operatorname{int}\operatorname{dom} f) \cap \operatorname{dom} A$ is a nonempty subset of $\operatorname{int}\operatorname{dom} B$, that B is single-valued on $\operatorname{int}\operatorname{dom} B$, that ∇f is strictly monotone on D, and that $(\nabla f - B)(D) \subset \operatorname{ran}(\nabla f + A)$. Set $M = A + B$ and $K\colon D \to \mathcal{X}^*\colon x \mapsto \nabla f(x) - Bx$. Then the warped resolvent J_M^K is well defined and coincides with the Bregman forward-backward operator $(\nabla f + A)^{-1} \circ (\nabla f - B)$ investigated in [13], where it is shown to capture a construction found in [31].

Example 3.7 Consider the setting of Lemma 2.2. For simplicity (more general kernels can be considered), take $s^* = 0$, $r = 0$, and assume that $\mathcal{Y}$ and $\mathcal{Z}^*$ are strictly convex, with normalized duality mapping $K_{\mathcal{Y}}$ and $K_{\mathcal{Z}^*}$. As seen in (i), finding a zero of the Kuhn–Tucker operator $U\colon \mathcal{Y} \times \mathcal{Z}^* \to 2^{\mathcal{Y}^* \times \mathcal{Z}}\colon (x, v^*) \mapsto (Ax + L^*v^*) \times (B^{-1}v^* - Lx)$ provides a solution to the primal-dual problem (2.4)–(2.5). Now set $K\colon (x, v^*) \mapsto (K_{\mathcal{Y}}x - L^*v^*, Lx + K_{\mathcal{Z}^*}v^*)$. Then the warped resolvent J_U^K is well defined and

$$J_U^K\colon (x, v^*) \mapsto \big((K_{\mathcal{Y}} + A)^{-1}(K_{\mathcal{Y}}x - L^*v^*), (K_{\mathcal{Z}^*} + B^{-1})^{-1}(Lx + K_{\mathcal{Z}^*}v^*)\big). \tag{3.4}$$

For instance, in a Hilbertian setting, $J_U^K\colon (x, v^*) \mapsto (J_A(x - L^*v^*), J_{B^{-1}}(Lx + v^*))$, whereas J_U is intractable; note also that the kernel K is a non-Hermitian bounded linear operator.

Further examples will appear in Section 5. Let us turn our attention to the properties of warped resolvents.

Proposition 3.8 (viability) *Let D be a nonempty subset of $\mathcal{X}$, let $K\colon D \to \mathcal{X}^*$, and let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be such that $\operatorname{ran} K \subset \operatorname{ran}(K + M)$ and $K + M$ is injective. Then $J_M^K\colon D \to D$.*

Proof. By assumption, $\operatorname{dom} J_M^K = \operatorname{dom}((K + M)^{-1} \circ K) = \big\{x \in \operatorname{dom} K \mid Kx \in \operatorname{dom}(K + M)^{-1}\big\} = \big\{x \in D \mid Kx \in \operatorname{ran}(K + M)\big\} = D$. Next, observe that

$$\operatorname{ran} J_M^K = \operatorname{ran}\big((K + M)^{-1} \circ K\big) \subset \operatorname{ran}(K + M)^{-1} = \operatorname{dom}(K + M) \subset \operatorname{dom} K = D. \tag{3.5}$$

Finally, to show that $(K + M)^{-1}$ is at most single-valued, suppose that $(x^*, x_1) \in \operatorname{gra}(K + M)^{-1}$ and $(x^*, x_2) \in \operatorname{gra}(K + M)^{-1}$. Then $\{x^*\} \subset (K + M)x_1 \cap (K + M)x_2$ and, since $K + M$ is injective, it follows that $x_1 = x_2$. ☐

Sufficient conditions that guarantee that warped resolvents are well-defined are made explicit below.

Proposition 3.9 *Let D be a nonempty subset of $\mathcal{X}$, let $K\colon D \to \mathcal{X}^*$, and let $M\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Then the following hold:*

(i) *Suppose that one of the following holds:*

  [a] *$K + M$ is surjective.*

  [b] *$K + M$ is maximally monotone and $D \cap \operatorname{dom} M$ is bounded.*

  [c] *$K + M$ is maximally monotone, $K + M$ is uniformly monotone with modulus ϕ, and $\phi(t)/t \to +\infty$ as $t \to +\infty$.*

[d] *$K+M$ is maximally monotone and strongly monotone.*

[e] *M is maximally monotone, $D=\mathcal{X}$, and K is maximally monotone, strictly monotone, 3^* monotone, and surjective.*

[f] *K is maximally monotone and there exists a lower semicontinuous coercive convex function $\varphi\colon\mathcal{X}\to\mathbb{R}$ such that $M=\partial\varphi$.*

Then $\operatorname{ran}K\subset\operatorname{ran}(K+M)$.

(ii) *Suppose that one of the following holds:*

[a] *$K+M$ is strictly monotone.*

[b] *M is monotone and K is strictly monotone on* $\operatorname{dom}M$.

[c] *K is monotone and M is strictly monotone.*

[d] *$-(K+M)$ is strictly monotone.*

Then $K+M$ is injective.

Proof. Set $A=K+M$.

(i): Item [a] is clear. We prove the remaining ones as follows.

[b]: It follows from [39, Theorem 32.G] that $\operatorname{ran}A=\mathcal{X}\supset\operatorname{ran}K$.

[c]&[d]: Since [20, Lemma 2.7(ii)] and [39, Corollary 32.35] assert that A is surjective, the claim follows from (i)[a].

[e]: See [8, Theorem 2.3].

[f]: Take $z\in D$ and set $B=A(\cdot+z)-Kz$. By coercivity of φ, there exists $\rho\in\,]0,+\infty[$ such that

$$(\forall x\in\mathcal{X})\quad \|x\|\geqslant\rho\quad\Rightarrow\quad \inf\langle x,\partial\varphi(x+z)\rangle\geqslant\varphi(x+z)-\varphi(z)\geqslant 0. \tag{3.6}$$

Now take $(x,x^*)\in\operatorname{gra}B$ and suppose that $\|x\|\geqslant\rho$. Then $x^*+Kz-K(x+z)\in\partial\varphi(x+z)$ and it follows from (3.6) and the monotonicity of K that

$$0\leqslant\langle x,x^*+Kz-K(x+z)\rangle=\langle x,x^*\rangle-\langle(x+z)-z,K(x+z)-Kz\rangle\leqslant\langle x,x^*\rangle. \tag{3.7}$$

On the other hand, since $\operatorname{dom}\partial\varphi=\mathcal{X}$ [38, Theorems 2.2.20(b) and 2.4.12], A is maximally monotone [35, Theorem 24.1(a)], and so is B. Altogether, [33, Proposition 2] asserts that there exists $\overline{x}\in\mathcal{X}$ such that $0\in B\overline{x}$. Consequently, $Kz\in A(\overline{x}+z)\subset\operatorname{ran}(K+M)$.

(ii): We need to prove only [a] since [b] and [c] are special cases of it, and [d] is similar. To this end, let $(x_1,x_2)\in\mathcal{X}^2$ and suppose that $Ax_1\cap Ax_2\neq\varnothing$. We must show that $x_1=x_2$. Take $x^*\in Ax_1\cap Ax_2$. Then (x_1,x^*) and (x_2,x^*) lie in $\operatorname{gra}A$. In turn, since A is strictly monotone and $\langle x_1-x_2,x^*-x^*\rangle=0$, we obtain $x_1=x_2$. ☐

Proposition 3.10 *Let $M\colon\mathcal{X}\to 2^{\mathcal{X}^*}$, let $\gamma\in\,]0,+\infty[$, and let $K\colon\mathcal{X}\to\mathcal{X}^*$ be such that* $\operatorname{ran}K\subset\operatorname{ran}(K+\gamma M)$ *and $K+\gamma M$ is injective. Then the following hold:*

(i) $\operatorname{Fix}J^K_{\gamma M}=\operatorname{zer}M$.

(ii) *Let $x\in\mathcal{X}$ and $p\in\mathcal{X}$. Then $p=J^K_{\gamma M}x\Leftrightarrow(p,\gamma^{-1}(Kx-Kp))\in\operatorname{gra}M$.*

(iii) *Suppose that M is monotone. Let $x\in\mathcal{X}$ and $y\in\mathcal{X}$, and set $p=J^K_{\gamma M}x$ and $q=J^K_{\gamma M}y$. Then $\langle p-q,Kx-Ky\rangle\geqslant\langle p-q,Kp-Kq\rangle$.*

(iv) *Suppose that M is monotone, that K is uniformly continuous and ϕ-uniformly monotone, and that $\psi\colon t \mapsto \phi(t)/t$ is real-valued on $]0,\xi[$ for some $\xi \in]0,+\infty[$ and strictly increasing. Then $J^K_{\gamma M}$ is uniformly continuous.*

(v) *Suppose that M is monotone and that K is β-Lipschitzian and α-strongly monotone for some $\alpha \in]0,+\infty[$ and $\beta \in]0,+\infty[$. Then $J^K_{\gamma M}$ is (β/α)-Lipschitzian.*

(vi) *Suppose that M is monotone. Let $x \in \mathcal{X}$, and set $y = J^K_{\gamma M}x$ and $y^* = \gamma^{-1}(Kx - Ky)$. Then $\operatorname{zer} M \subset \big\{z \in \mathcal{X} \mid \langle z - y, y^*\rangle \leqslant 0\big\}$.*

Proof. (i): We derive from Proposition 3.8 that $(\forall x \in \mathcal{X})\ x \in \operatorname{zer} M \Leftrightarrow Kx \in Kx + \gamma Mx \Leftrightarrow x = J^K_{\gamma M}x \Leftrightarrow x \in \operatorname{Fix} J^K_{\gamma M}$.

(ii): We have $p = J^K_{\gamma M}x \Leftrightarrow p = (K + \gamma M)^{-1}(Kx) \Leftrightarrow Kx \in Kp + \gamma Mp \Leftrightarrow Kx - Kp \in \gamma Mp \Leftrightarrow (p, \gamma^{-1}(Kx - Kp)) \in \operatorname{gra} M$.

(iii): This follows from (ii) and the monotonicity of M.

(iv): Let x and y be in $\mathcal{X}$, and set $p = J^K_{\gamma M}x$ and $q = J^K_{\gamma M}y$. Then we deduce from (iii) that

$$\phi(\|p - q\|) \leqslant \langle p - q, Kp - Kq\rangle \leqslant \langle p - q, Kx - Ky\rangle \leqslant \|p - q\|\,\|Kx - Ky\|. \tag{3.8}$$

Now fix $\varepsilon \in]0,\xi[$ and let $\eta \in]0,\psi(\varepsilon)]$. By uniform continuity of K, there exists $\delta \in]0,+\infty[$ such that $\|x - y\| \leqslant \delta \Rightarrow \|Kx - Ky\| \leqslant \eta$. Without loss of generality, suppose that $p \neq q$. Then, if $\|x - y\| \leqslant \delta$, we derive from (3.8) that $\psi(\|p - q\|) \leqslant \|Kx - Ky\| \leqslant \eta \leqslant \psi(\varepsilon)$. Consequently, since ψ is strictly increasing, $\|p - q\| \leqslant \varepsilon$.

(v): Let x and y be in $\mathcal{X}$ and set $p = J^K_{\gamma M}x$ and $q = J^K_{\gamma M}y$. Then we deduce from (iii) that

$$\alpha\|p-q\|^2 \leqslant \langle p - q, Kp - Kq\rangle \leqslant \langle p - q, Kx - Ky\rangle \leqslant \|p-q\|\,\|Kx-Ky\| \leqslant \beta\|p-q\|\,\|x-y\|. \tag{3.9}$$

In turn, $\|p - q\| \leqslant (\beta/\alpha)\|x - y\|$.

(vi): Suppose that $z \in \operatorname{zer} M$. Then $(z, 0) \in \operatorname{gra} M$. On the other hand, we derive from (ii) that $(y, y^*) \in \operatorname{gra} M$. Hence, by monotonicity of M, $\langle y - z, y^*\rangle \geqslant 0$. ☐

In Hilbert spaces, standard resolvents are firmly nonexpansive, hence $1/2$-averaged. A related property for warped resolvents is the following.

Proposition 3.11 *Suppose that $\mathcal{X}$ is a Hilbert space. Let $M\colon \mathcal{X} \to 2^{\mathcal{X}}$ be maximally monotone and let $K\colon \mathcal{X} \to \mathcal{X}$ be averaged with constant $\alpha \in]0,1[$. Suppose that $K + M$ is 1-strongly monotone. Then J^K_M is averaged with constant $1/(2-\alpha)$.*

Proof. Since K is nonexpansive by virtue of [7, Remark 4.34(i)], it follows from the Cauchy–Schwarz inequality that

$$\begin{aligned}(\forall x \in \mathcal{X})(\forall y \in \mathcal{X})\quad \langle x - y \mid (2\,\mathrm{Id}+K)x - (2\,\mathrm{Id}+K)y\rangle &= 2\|x - y\|^2 + \langle x - y \mid Kx - Ky\rangle\\ &\geqslant 2\|x - y\|^2 - \|x - y\|^2\\ &= \|x - y\|^2\end{aligned} \tag{3.10}$$

and therefore, by continuity of $2\,\mathrm{Id}+K$, that $2\,\mathrm{Id}+K$ is maximally monotone [7, Corollary 20.28]. Thus, in the light of [7, Corollary 25.5(i)], $2\,\mathrm{Id}+K + M$ is maximally monotone. In turn, since $2\,\mathrm{Id}+K+M$ is strongly monotone by (3.10), [7, Proposition 22.11(ii)] entails that $\operatorname{ran}(3\,\mathrm{Id}+K+M-\mathrm{Id}) = \operatorname{ran}(2\,\mathrm{Id}+K+M) = \mathcal{X}$, which yields $\operatorname{ran}(\mathrm{Id}+(K+M-\mathrm{Id})/3) = \mathcal{X}$. Hence, by monotonicity of

$K+M-\mathrm{Id}$ and Minty's theorem [7, Theorem 21.1], we infer that $K+M-\mathrm{Id}$ is maximally monotone. Thus, in view of [7, Corollary 23.9], $(K+M)^{-1} = (\mathrm{Id}+K+M-\mathrm{Id})^{-1}$ is averaged with constant $1/2$. Consequently, we infer from [7, Proposition 4.44] that $J_M^K = (K+M)^{-1}\circ K$ is averaged with constant $1/(2-\alpha)$. ☐

4 Warped proximal iterations

Throughout this section, $\mathcal{X}$ is a real Hilbert space identified with its dual. We start with an abstract principle for the basic problem of finding a zero of a maximally monotone operator.

Proposition 4.1 *Let $M\colon \mathcal{X}\to 2^{\mathcal{X}}$ be a maximally monotone operator such that $Z = \operatorname{zer} M \neq \varnothing$, let $x_0\in\mathcal{X}$, let $\varepsilon\in\left]0,1\right[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2-\varepsilon]$, and let $(y_n,y_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{gra} M$. Set*

$$(\forall n\in\mathbb{N})\quad x_{n+1} = \begin{cases} x_n + \dfrac{\lambda_n\langle y_n - x_n \mid y_n^*\rangle}{\|y_n^*\|^2}\, y_n^*, & \text{if } \langle y_n - x_n \mid y_n^*\rangle < 0;\\ x_n, & \text{otherwise.}\end{cases} \tag{4.1}$$

Then the following hold:

(i) $\sum_{n\in\mathbb{N}} \|x_{n+1}-x_n\|^2 < +\infty$.

(ii) *Suppose that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in Z. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in Z.*

Proof. By [7, Proposition 23.39], Z is a nonempty closed convex subset of $\mathcal{X}$. Set $(\forall n\in\mathbb{N})$ $H_n = \big\{z\in\mathcal{X} \mid \langle z - y_n \mid y_n^*\rangle \leqslant 0\big\}$. For every $z\in Z$ and every $n\in\mathbb{N}$, since $(z,0)$ and (y_n,y_n^*) lie in $\operatorname{gra} M$, the monotonicity of M forces $\langle y_n - z\mid y_n^*\rangle \geqslant 0$. Thus $Z\subset\bigcap_{n\in\mathbb{N}} H_n$. In addition, [7, Example 29.20] asserts that

$$(\forall n\in\mathbb{N})\quad \operatorname{proj}_{H_n} x_n = \begin{cases} x_n + \dfrac{\langle y_n - x_n \mid y_n^*\rangle}{\|y_n^*\|^2}\, y_n^*, & \text{if } \langle y_n - x_n \mid y_n^*\rangle < 0;\\ x_n, & \text{otherwise.}\end{cases} \tag{4.2}$$

Hence, we derive from (4.1) that

$$(\forall n\in\mathbb{N})\quad x_{n+1} = x_n + \lambda_n(\operatorname{proj}_{H_n} x_n - x_n). \tag{4.3}$$

Therefore (i) follows from [16, Equation (10)] and (ii) follows from [16, Proposition 6i)]. ☐

To implement the conceptual principle outlined in Proposition 4.1, one is required to construct points in the graph of the underlying monotone operator. Towards this end, our strategy is to use Proposition 3.10(ii). We shall then seamlessly obtain in Section 5 a broad class of algorithms to solve a variety of monotone inclusions. It will be convenient to use the notation

$$(\forall y^*\in\mathcal{Y}^*)\quad (y^*)^\sharp = \begin{cases} \dfrac{y^*}{\|y^*\|}, & \text{if } y^*\neq 0;\\ 0, & \text{if } y^* = 0.\end{cases} \tag{4.4}$$

Our first method employs, at iteration n, a warped resolvent based on a different kernel, and this warped resolvent is applied at a point $\widetilde{x}_n$ that may not be the current iterate x_n.

Theorem 4.2 *Let $M\colon \mathcal{X} \to 2^{\mathcal{X}}$ be a maximally monotone operator such that $Z = \operatorname{zer} M \neq \varnothing$, let $x_0 \in \mathcal{X}$, let $\varepsilon \in \left]0,1\right[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2-\varepsilon]$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $\left[\varepsilon, +\infty\right[$. Further, for every $n \in \mathbb{N}$, let $\widetilde{x}_n \in \mathcal{X}$ and let $K_n\colon \mathcal{X} \to \mathcal{X}$ be a monotone operator such that $\operatorname{ran} K_n \subset \operatorname{ran}(K_n + \gamma_n M)$ and $K_n + \gamma_n M$ is injective. Iterate*

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
y_n = J^{K_n}_{\gamma_n M}\widetilde{x}_n\\
y_n^* = \gamma_n^{-1}(K_n\widetilde{x}_n - K_n y_n)\\
\text{if } \langle y_n - x_n \mid y_n^*\rangle < 0\\
\left\lfloor\; x_{n+1} = x_n + \dfrac{\lambda_n\langle y_n - x_n \mid y_n^*\rangle}{\|y_n^*\|^2}\, y_n^* \right.\\
\text{else}\\
\left\lfloor\; x_{n+1} = x_n. \right.
\end{array}
\right.
\end{array}
\tag{4.5}
$$

Then the following hold:

(i) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$.

(ii) *Suppose that the following are satisfied:*

[a] $\widetilde{x}_n - x_n \to 0$.

[b] $\left\langle \widetilde{x}_n - y_n \mid (K_n\widetilde{x}_n - K_n y_n)^\sharp \right\rangle \to 0 \quad\Rightarrow\quad \begin{cases} \widetilde{x}_n - y_n \rightharpoonup 0\\ K_n\widetilde{x}_n - K_n y_n \to 0. \end{cases}$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in Z.

Proof. By Proposition 3.10(ii),

$$
(\forall n \in \mathbb{N})\quad (y_n, y_n^*) \in \operatorname{gra} M. \tag{4.6}
$$

Therefore, (i) follows from Proposition 4.1(i). It remains to prove (ii). To this end, take a strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ and a point $x \in \mathcal{X}$ such that $x_{k_n} \rightharpoonup x$. In view of Proposition 4.1(ii), we must show that $x \in Z$. We infer from (ii)[a] that

$$
\widetilde{x}_{k_n} \rightharpoonup x. \tag{4.7}
$$

Next, by (4.4) and (4.5), for every $n \in \mathbb{N}$, if $\langle x_n - y_n \mid y_n^*\rangle > 0$, then $y_n^* \neq 0$ and

$$
\left\langle x_n - y_n \mid (y_n^*)^\sharp \right\rangle = \frac{\langle x_n - y_n \mid y_n^*\rangle}{\|y_n^*\|} = \lambda_n^{-1}\|x_{n+1} - x_n\| \leqslant \varepsilon^{-1}\|x_{n+1} - x_n\|; \tag{4.8}
$$

otherwise, $\langle x_n - y_n \mid y_n^*\rangle \leqslant 0$ and it thus results from (4.4) that

$$
\begin{aligned}
\left\langle x_n - y_n \mid (y_n^*)^\sharp \right\rangle &= \begin{cases} 0, & \text{if } y_n^* = 0;\\ \dfrac{\langle x_n - y_n \mid y_n^*\rangle}{\|y_n^*\|}, & \text{otherwise} \end{cases}\\
&\leqslant 0\\
&= \varepsilon^{-1}\|x_{n+1} - x_n\|.
\end{aligned} \tag{4.9}
$$

Therefore, using (i) and the monotonicity of $(K_n)_{n\in\mathbb{N}}$, we obtain

$$
0 \leftarrow \varepsilon^{-1}\|x_{n+1} - x_n\|
$$

$$
\begin{aligned}
&\geqslant \langle x_n - y_n \mid (y_n^*)^\sharp\rangle \\
&= \big\langle x_n - \widetilde{x}_n \mid (K_n\widetilde{x}_n - K_n y_n)^\sharp\big\rangle + \big\langle \widetilde{x}_n - y_n \mid (K_n\widetilde{x}_n - K_n y_n)^\sharp\big\rangle \\
&\geqslant \big\langle x_n - \widetilde{x}_n \mid (K_n\widetilde{x}_n - K_n y_n)^\sharp\big\rangle.
\end{aligned}
\tag{4.10}
$$

However, by the Cauchy–Schwarz inequality and (ii)[a],

$$
\big|\big\langle x_n - \widetilde{x}_n \mid (K_n\widetilde{x}_n - K_n y_n)^\sharp\big\rangle\big| \leqslant \|x_n - \widetilde{x}_n\| \to 0. \tag{4.11}
$$

Hence, (4.10) implies that $\langle \widetilde{x}_n - y_n \mid (K_n\widetilde{x}_n - K_n y_n)^\sharp\rangle \to 0$. In turn, we deduce from (ii)[b] that $\widetilde{x}_n - y_n \rightharpoonup 0$ and $K_n\widetilde{x}_n - K_n y_n \to 0$. Altogether, since $\sup_{n\in\mathbb{N}} \gamma_n^{-1} \leqslant \varepsilon^{-1}$, it follows from (4.6) and (4.7) that

$$
y_{k_n} = \widetilde{x}_{k_n} + (y_{k_n} - \widetilde{x}_{k_n}) \rightharpoonup x \tag{4.12}
$$

and

$$
My_{k_n} \ni y_{k_n}^* = \gamma_{k_n}^{-1}(K_{k_n}\widetilde{x}_{k_n} - K_{k_n} y_{k_n}) \to 0. \tag{4.13}
$$

Appealing to the maximal monotonicity of M, [7, Proposition 20.38(ii)] allows us to conclude that $x \in Z$. ☐

Remark 4.3 Condition (ii)[b] in Theorem 4.2 is satisfied in particular when there exist α and β in $]0,+\infty[$ such that the kernels $(K_n)_{n\in\mathbb{N}}$ are α-strongly monotone and β-Lipschitzian.

Remark 4.4 The auxiliary sequence $(\widetilde{x}_n)_{n\in\mathbb{N}}$ in Theorem 4.2 can serve several purposes. In general, it provides the flexibility of not applying the warped resolvent to the current iterate. Here are some noteworthy candidates.

(i) At iteration n, $\widetilde{x}_n$ can model an additive perturbation of x_n, say $\widetilde{x}_n = x_n + e_n$. Here the error sequence $(e_n)_{n\in\mathbb{N}}$ need only satisfy $\|e_n\| \to 0$ and not the usual summability condition $\sum_{n\in\mathbb{N}} \|e_n\| < +\infty$ required in many methods, e.g., [11, 17, 21, 37].

(ii) Mimicking the behavior of so-called inertial methods [3, 19], let $(\alpha_n)_{n\in\mathbb{N}}$ be a bounded sequence in $\mathbb{R}$ and set $(\forall n \in \mathbb{N}\smallsetminus\{0\})\ \widetilde{x}_n = x_n + \alpha_n(x_n - x_{n-1})$. Then Theorem 4.2(i) yields $\|\widetilde{x}_n - x_n\| = |\alpha_n|\,\|x_n - x_{n-1}\| \to 0$ and therefore assumption (ii)[a] holds in Theorem 4.2. More generally, weak convergence results can be derived from Theorem 4.2 for iterations with memory, that is,

$$
(\forall n \in \mathbb{N})\quad \widetilde{x}_n = \sum_{j=0}^{n} \mu_{n,j} x_j, \quad \text{where} \quad (\mu_{n,j})_{0\leqslant j\leqslant n} \in \mathbb{R}^{n+1} \quad \text{and} \quad \sum_{j=0}^{n} \mu_{n,j} = 1. \tag{4.14}
$$

Here condition (ii)[a] holds if $(1-\mu_{n,n})x_n - \sum_{j=0}^{n-1} \mu_{n,j} x_j \to 0$. In the case of standard inertial methods, weak convergence requires more stringent conditions on the weights $(\mu_{n,j})_{n\in\mathbb{N},0\leqslant j\leqslant n}$ [19].

(iii) Nonlinear perturbations can also be considered. For instance, at iteration n, $\widetilde{x}_n = \operatorname{proj}_{C_n} x_n$ is an approximation to x_n from some suitable closed convex set $C_n \subset \mathcal{X}$.

Remark 4.5 The independent work [23] was posted on arXiv at the same time as the report [14] from which our paper is derived. The former also uses a notion of resolvents subsumed by Definition 1.1 to explore the application of an algorithm similar to (4.5) with no perturbation (i.e., for every $n \in \mathbb{N}$, $\widetilde{x}_n = x_n$). The work [23] nicely complements ours in the sense that it proposes applications to splitting schemes not discussed here, which further attests to the versatility and effectiveness of the notion of warped proximal iterations.

We now turn our attention to a variant of Theorem 4.2 that guarantees strong convergence of the iterates to a best approximation. In the spirit of Haugazeau's algorithm (see [24, Théorème 3-2] and [7, Corollary 30.15]), it involves outer approximations consisting of the intersection of two half-spaces. For convenience, given $(x,y,z) \in \mathcal{X}^3$, we set

$$H(x,y) = \big\{u \in \mathcal{X} \mid \langle u-y \mid x-y\rangle \leqslant 0\big\} \tag{4.15}$$

and, if $R = H(x,y) \cap H(y,z) \neq \varnothing$, $Q(x,y,z) = \operatorname{proj}_R x$. The latter can be computed explicitly as follows (see [24, Théorème 3-1] or [7, Corollary 29.25]).

Lemma 4.6 *Let $(x,y,z) \in \mathcal{X}^3$. Set $R = H(x,y) \cap H(y,z)$, $\chi = \langle x-y \mid y-z\rangle$, $\mu = \|x-y\|^2$, $\nu = \|y-z\|^2$, and $\rho = \mu\nu - \chi^2$. Then exactly one of the following holds:*

(i) *$\rho = 0$ and $\chi < 0$, in which case $R = \varnothing$.*

(ii) *$[\,\rho = 0$ and $\chi \geqslant 0\,]$ or $\rho > 0$, in which case $R \neq \varnothing$ and*

$$Q(x,y,z) = \begin{cases} z, & \text{if } \rho = 0 \text{ and } \chi \geqslant 0;\\ x + (1+\chi/\nu)(z-y), & \text{if } \rho > 0 \text{ and } \chi\nu \geqslant \rho;\\ y + (\nu/\rho)\big(\chi(x-y) + \mu(z-y)\big), & \text{if } \rho > 0 \text{ and } \chi\nu < \rho. \end{cases} \tag{4.16}$$

Our second abstract convergence principle can now be stated.

Proposition 4.7 *Let $M\colon \mathcal{X} \to 2^{\mathcal{X}}$ be a maximally monotone operator such that $Z = \operatorname{zer} M \neq \varnothing$, let $x_0 \in \mathcal{X}$, and let $(y_n, y_n^*)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{gra} M$. For every $n \in \mathbb{N}$, set*

$$x_{n+1/2} = \begin{cases} x_n + \dfrac{\langle y_n - x_n \mid y_n^*\rangle}{\|y_n^*\|^2}\, y_n^*, & \text{if } \langle y_n - x_n \mid y_n^*\rangle < 0;\\ x_n, & \text{otherwise} \end{cases} \quad \text{and} \quad x_{n+1} = Q\big(x_0, x_n, x_{n+1/2}\big). \tag{4.17}$$

Then the following hold:

(i) *$\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$ and $\sum_{n\in\mathbb{N}} \|x_{n+1/2} - x_n\|^2 < +\infty$.*

(ii) *Suppose that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in Z. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to $\operatorname{proj}_Z x_0$.*

Proof. Set $(\forall n \in \mathbb{N})$ $H_n = \big\{z \in \mathcal{X} \mid \langle z - y_n \mid y_n^*\rangle \leqslant 0\big\}$. Then, as in the proof of Proposition 4.1, Z is a nonempty closed convex subset of $\mathcal{X}$ and $Z \subset \bigcap_{n\in\mathbb{N}} H_n$. On the one hand,

$$(\forall n \in \mathbb{N}) \quad x_{n+1/2} = \operatorname{proj}_{H_n} x_n \quad \text{and} \quad x_{n+1} = Q\big(x_0, x_n, x_{n+1/2}\big). \tag{4.18}$$

On the other hand, by (4.15),

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad H\big(x_n, x_{n+1/2}\big) &= \begin{cases} \mathcal{X}, & \text{if } x \in H_n;\\ H_n, & \text{otherwise} \end{cases}\\ &\supset Z. \end{aligned} \tag{4.19}$$

The claims therefore follow from [2, Proposition 2.1]. ☐

Theorem 4.8 *Let $M\colon \mathcal{X} \to 2^{\mathcal{X}}$ be a maximally monotone operator such that $Z = \operatorname{zer} M \neq \varnothing$, let $x_0 \in \mathcal{X}$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0,+\infty[$ such that $\inf_{n\in\mathbb{N}} \gamma_n > 0$. For every $n \in \mathbb{N}$, let $\widetilde{x}_n \in \mathcal{X}$ and let $K_n\colon \mathcal{X} \to \mathcal{X}$ be a monotone operator such that $\operatorname{ran} K_n \subset \operatorname{ran}(K_n + \gamma_n M)$ and $K_n + \gamma_n M$ is injective. Iterate*

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
y_n = J^{K_n}_{\gamma_n M}\widetilde{x}_n\\
y_n^* = \gamma_n^{-1}(K_n\widetilde{x}_n - K_n y_n)\\
\text{if } \langle y_n - x_n \mid y_n^*\rangle < 0\\
\left\lfloor
\begin{array}{l}
x_{n+1/2} = x_n + \dfrac{\langle y_n - x_n \mid y_n^*\rangle}{\|y_n^*\|^2}\, y_n^*
\end{array}
\right.\\
\text{else}\\
\left\lfloor
\begin{array}{l}
x_{n+1/2} = x_n
\end{array}
\right.\\
x_{n+1} = Q(x_0, x_n, x_{n+1/2}).
\end{array}
\right.
\end{array}
\tag{4.20}
$$

Then the following hold:

(i) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$ *and* $\sum_{n\in\mathbb{N}} \|x_{n+1/2} - x_n\|^2 < +\infty$.

(ii) *Suppose that the following are satisfied:*

[a] $\widetilde{x}_n - x_n \to 0$.

[b] $\big\langle \widetilde{x}_n - y_n \mid (K_n\widetilde{x}_n - K_n y_n)^\sharp \big\rangle \to 0 \quad \Rightarrow \quad \begin{cases} \widetilde{x}_n - y_n \rightharpoonup 0\\ K_n\widetilde{x}_n - K_n y_n \to 0. \end{cases}$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to $\operatorname{proj}_Z x_0$.

Proof. Proposition 3.10(ii) asserts that $(\forall n \in \mathbb{N})$ $(y_n, y_n^*) \in \operatorname{gra} M$. Thus, we obtain (i) from Proposition 4.7(i). In the light of Proposition 4.7(ii), to establish (ii), we need to show that every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is a zero of M. Since (i) asserts that $x_{n+1/2} - x_n \to 0$, this is done as in the proof of Theorem 4.2(ii). ☐

We complete this section with the following remarks.

Remark 4.9 Suppose that $\mathcal{Y}$ and $\mathcal{Z}$ are real Hilbert spaces and that $\mathcal{X} = \mathcal{Y} \times \mathcal{Z}$. Let $A\colon \mathcal{Y} \to 2^{\mathcal{Y}}$ and $B\colon \mathcal{Z} \to 2^{\mathcal{Z}}$ be maximally monotone, and let $L \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$. Define

$$M\colon \mathcal{X} \to 2^{\mathcal{X}}\colon (x, v^*) \mapsto (Ax + L^*v^*) \times (-Lx + B^{-1}v^*). \tag{4.21}$$

In [1, 2, 18] the problem of finding a zero of M (and hence a solution to the monotone inclusion $0 \in Ax + L^*(B(Lx)))$ is approached by generating, at each iteration n, points $(a_n, a_n^*) \in \operatorname{gra} A$ and $(b_n, b_n^*) \in \operatorname{gra} B$. This does provide a point $(y_n, y_n^*) = ((a_n, b_n^*), (a_n^* + L^*b_n^*, -La_n + b_n)) \in \operatorname{gra} M$, which shows that the algorithms proposed in [1, 2, 18] are actually instances of the conceptual principles laid out in Propositions 4.1 and 4.7. In particular, the primal-dual framework of [1] corresponds to applying Theorem 4.2 to the operator M of (4.21) with kernels

$$(\forall n \in \mathbb{N}) \quad K_n\colon \mathcal{X} \to \mathcal{X}\colon (x, v^*) \mapsto \big(\gamma_n^{-1}x - L^*v^*, Lx + \mu_n v^*\big). \tag{4.22}$$

Likewise, that of [2] corresponds to the application of Theorem 4.8 to this setting.

Remark 4.10 In Theorems 4.2 and 4.8, the algorithms operate by using a single point (y_n, y_n^*) in $\operatorname{gra} M$ at iteration n. It may be advantageous to use a finite family $(y_{i,n}, y_{i,n}^*)_{i\in I_n}$ of points in $\operatorname{gra} M$, say

$$(\forall i \in I_n) \quad (y_{i,n}, y_{i,n}^*) = \Big(J_{\gamma_{i,n}M}^{K_{i,n}} \widetilde{x}_{i,n}, \gamma_{i,n}^{-1}(K_{i,n}\widetilde{x}_{i,n} - K_{i,n}y_{i,n})\Big). \tag{4.23}$$

By monotonicity of M, $(\forall i \in I_n)(\forall z \in \operatorname{zer} M)$ $\langle z \mid y_{i,n}^*\rangle \leqslant \langle y_{i,n} \mid y_{i,n}^*\rangle$. Therefore, using ideas found in the area of convex feasibility algorithms [15, 27], at every iteration n, given strictly positive weights $(\omega_{i,n})_{i\in I_n}$ adding up to 1, we average these inequalities to create a new half-space H_n containing $\operatorname{zer} M$, namely

$$\operatorname{zer} M \subset H_n = \big\{z \in \mathcal{X} \mid \langle z \mid y_n^*\rangle \leqslant \eta_n\big\}, \quad \text{where} \quad \begin{cases} y_n^* = \sum_{i\in I_n} \omega_{i,n} y_{i,n}^* \\ \eta_n = \sum_{i\in I_n} \omega_{i,n}\langle y_{i,n} \mid y_{i,n}^*\rangle. \end{cases} \tag{4.24}$$

Now set

$$\Lambda_n = \begin{cases} \dfrac{\sum_{i\in I_n} \omega_{i,n}\langle y_{i,n} - x_n \mid y_{i,n}^*\rangle}{\big\| \sum_{i\in I_n} \omega_{i,n} y_{i,n}^*\big\|^2}, & \text{if } \sum_{i\in I_n} \omega_{i,n}\langle x_n - y_{i,n} \mid y_{i,n}^*\rangle > 0;\\ 0, & \text{otherwise.} \end{cases} \tag{4.25}$$

Then, employing $\operatorname{proj}_{H_n} x_n = x_n + \Lambda_n \sum_{i\in I_n} \omega_{i,n} y_{i,n}^*$ as the point x_{n+1} in (4.5) and as the point $x_{n+1/2}$ in (4.20) results in multi-point extensions of Theorems 4.2 and 4.8.

5 Applications

We now apply Theorem 4.2 to design new algorithms to solve complex monotone inclusion problems in a real Hilbert space $\mathcal{X}$. We do not mention explicitly minimization problems as they follow, with usual constraint qualification conditions, by considering monotone inclusions involving subdifferentials as maximally monotone operators [7, 17]. For brevity, we do not mention either the strongly convergent counterparts of each of the corollaries below that can be systematically obtained using Theorem 4.8.

Let us note that the most basic instantiation of Theorem 4.2 is obtained by setting $(\forall n \in \mathbb{N})$ $K_n = \operatorname{Id}$, $\widetilde{x}_n = x_n$, and $\lambda_n = 1$. In this case, the warped proximal algorithm (4.5) reduces to the basic proximal point algorithm (1.1).

In connection with Remark 4.4, let us first investigate the convergence of a novel perturbed forward-backward-forward algorithm with memory. This will require the following fact.

Lemma 5.1 *Let $B\colon \mathcal{X} \to \mathcal{X}$ be Lipschitzian with constant $\beta \in \left]0, +\infty\right[$, let $W\colon \mathcal{X} \to \mathcal{X}$ be strongly monotone with constant $\alpha \in \left]0, +\infty\right[$, let $\varepsilon \in \left]0, \alpha\right[$, let $\gamma \in \left]0, (\alpha - \varepsilon)/\beta\right]$, and set $K = W - \gamma B$. Then the following hold:*

(i) *K is ε-strongly monotone.*

(ii) *Suppose that $\alpha = 1$ and $W = \operatorname{Id}$. Then K is cocoercive with constant $1/(2 - \varepsilon)$.*

Proof. (i): By the Cauchy–Schwarz inequality

$$(\forall x \in \mathcal{X})(\forall y \in \mathcal{X}) \quad \langle x - y \mid Kx - Ky\rangle = \langle x - y \mid Wx - Wy\rangle - \gamma\langle x - y \mid Bx - By\rangle$$

$$
\begin{aligned}
&\geqslant \alpha\|x-y\|^2 - \gamma\|x-y\|\,\|Bx-By\| \\
&\geqslant \alpha\|x-y\|^2 - \gamma\beta\|x-y\|^2 \\
&\geqslant \varepsilon\|x-y\|^2. \qquad (5.1)
\end{aligned}
$$

(ii): Since γB is $(1-\varepsilon)$-Lipschitzian, [7, Proposition 4.38] entails that γB is averaged with constant $(2-\varepsilon)/2$. Hence, since $\gamma B = \mathrm{Id} - K$, [7, Proposition 4.39] implies that K is cocoercive with constant $1/(2-\varepsilon)$. ☐

Corollary 5.2 *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}}$ be maximally monotone, let $B\colon \mathcal{X} \to \mathcal{X}$ be monotone and β-Lipschitzian for some $\beta \in \left]0,+\infty\right[$, let $(\alpha,\chi) \in \left]0,+\infty\right[^2$, and let $\varepsilon \in \left]0, \alpha/(\beta+1)\right[$. For every $n \in \mathbb{N}$, let $W_n\colon \mathcal{X} \to \mathcal{X}$ be α-strongly monotone and χ-Lipschitzian, and let $\gamma_n \in [\varepsilon, (\alpha-\varepsilon)/\beta]$. Take $x_0 \in \mathcal{X}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\left]0,2\right[$ such that $0 < \inf_{n\in\mathbb{N}} \lambda_n \leqslant \sup_{n\in\mathbb{N}} \lambda_n < 2$, and let $(e_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{X}$ such that $e_n \to 0$. Furthermore, let $m \in \mathbb{N} \smallsetminus \{0\}$ and let $(\mu_{n,j})_{n\in\mathbb{N}, 0\leqslant j\leqslant n}$ be a real array that satisfies the following:*

[a] *For every integer $n > m$ and every integer $j \in [0, n-m-1]$, $\mu_{n,j} = 0$.*

[b] *For every $n \in \mathbb{N}$, $\sum_{j=0}^{n} \mu_{n,j} = 1$.*

[c] $\sup_{n\in\mathbb{N}} \max_{0\leqslant j\leqslant n} |\mu_{n,j}| < +\infty$.

Iterate

$$
\begin{array}{l}
\textit{for } n = 0,1,\ldots \\
\left\lfloor
\begin{array}{l}
\widetilde{x}_n = e_n + \sum_{j=0}^{n} \mu_{n,j} x_j \\
v_n^* = W_n \widetilde{x}_n - \gamma_n B \widetilde{x}_n \\
y_n = (W_n + \gamma_n A)^{-1} v_n^* \\
y_n^* = \gamma_n^{-1}(v_n^* - W_n y_n) + B y_n \\
\textit{if } \langle y_n - x_n \mid y_n^* \rangle < 0 \\
\left\lfloor\; x_{n+1} = x_n + \dfrac{\lambda_n \langle y_n - x_n \mid y_n^* \rangle}{\|y_n^*\|^2}\, y_n^* \right. \\
\textit{else} \\
\left\lfloor\; x_{n+1} = x_n. \right.
\end{array}
\right.
\end{array}
\qquad (5.2)
$$

Suppose that $\mathrm{zer}(A+B) \neq \varnothing$. Then the following hold:

(i) $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|^2 < +\infty$.

(ii) *$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\mathrm{zer}(A+B)$.*

Proof. We apply Theorem 4.2 with $M = A + B$ and $(\forall n \in \mathbb{N})$ $K_n = W_n - \gamma_n B$. First, [7, Corollary 20.28] asserts that B is maximally monotone. Therefore, M is maximally monotone by virtue of [7, Corollary 25.5(i)]. Next, in view of Lemma 5.1(i), the kernels $(K_n)_{n\in\mathbb{N}}$ are ε-strongly monotone. Furthermore, the kernels $(K_n)_{n\in\mathbb{N}}$ are Lipschitzian with constant $\alpha + \chi$ since

$$
\begin{aligned}
(\forall x \in \mathcal{X})(\forall y \in \mathcal{X}) \quad \|K_n x - K_n y\| &\leqslant \|W_n x - W_n y\| + \gamma_n \|Bx - By\| \\
&\leqslant \chi\|x-y\| + \frac{\alpha-\varepsilon}{\beta}\beta\|x-y\| \\
&\leqslant (\alpha+\chi)\|x-y\|. \qquad (5.3)
\end{aligned}
$$

Therefore, for every $n \in \mathbb{N}$, since $K_n + \gamma_n M$ is maximally monotone, Proposition 3.9(i)[d]&(ii)[b] entail that $\operatorname{ran} K_n \subset \operatorname{ran}(K_n + \gamma_n M)$ and $K_n + \gamma_n M$ is injective. Let us also observe that (5.2) is a special case of (4.5).

(i): This follows from Theorem 4.2(i).

(ii): Set $\mu = \sup_{n\in\mathbb{N}} \max_{0\leqslant j\leqslant n} |\mu_{n,j}|$. For every integer $n > m$, it results from [a] and [b] that

$$
\begin{aligned}
\|\widetilde{x}_n - x_n\| &= \left\| e_n + \sum_{j=n-m}^{n} \mu_{n,j}(x_j - x_n) \right\| \\
&\leqslant \|e_n\| + \sum_{j=n-m}^{n} |\mu_{n,j}| \|x_j - x_n\| \\
&\leqslant \|e_n\| + \mu \sum_{j=n-m}^{n} \|x_j - x_n\| \\
&= \|e_n\| + \mu \sum_{j=0}^{m} \|x_n - x_{n-j}\|.
\end{aligned}
\tag{5.4}
$$

Therefore, (i) and [c] imply that $\widetilde{x}_n - x_n \to 0$. On the other hand, it follows from Remark 4.3 that condition (ii)[b] in Theorem 4.2 is satisfied. Hence, the conclusion follows from Theorem 4.2(ii). ☐

Next, we recover Tseng's forward-backward-forward algorithm [7, 36].

Corollary 5.3 *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}}$ be maximally monotone, let $B\colon \mathcal{X} \to \mathcal{X}$ be monotone and β-Lipschitzian for some $\beta \in \left]0, +\infty\right[$. Suppose that $\operatorname{zer}(A+B) \neq \varnothing$, take $x_0 \in \mathcal{X}$, let $\varepsilon \in \left]0, 1/(\beta+1)\right[$, and let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$. Iterate*

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
v_n^* = \gamma_n B x_n \\
y_n = J_{\gamma_n A}(x_n - v_n^*) \\
x_{n+1} = y_n - \gamma_n B y_n + v_n^*.
\end{array}
\right.
\end{array}
\tag{5.5}
$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{zer}(A+B)$.

Proof. We apply Theorem 4.2 with $M = A + B$ and $(\forall n \in \mathbb{N})$ $K_n = \operatorname{Id} - \gamma_n B$ and $\widetilde{x}_n = x_n$. Note that the kernels $(K_n)_{n\in\mathbb{N}}$ are cocoercive with constant $1/(2-\varepsilon)$ by virtue of Lemma 5.1(ii). Moreover, using Lemma 5.1(i), we deduce that the kernels $(K_n)_{n\in\mathbb{N}}$ are strongly monotone with constant ε. Thus, for every $n \in \mathbb{N}$, since $K_n + \gamma_n M = \operatorname{Id} + \gamma_n A$ is maximally monotone, Proposition 3.9(i)[d]&(ii)[b] assert that $\operatorname{ran} K_n \subset \operatorname{ran}(K_n + \gamma_n M)$ and $K_n + \gamma_n M$ is injective. Now set

$$
(\forall n \in \mathbb{N}) \quad y_n^* = \gamma_n^{-1}(K_n x_n - K_n y_n) \quad \text{and} \quad \lambda_n =
\begin{cases}
\dfrac{\gamma_n \|y_n^*\|^2}{\langle x_n - y_n \mid y_n^* \rangle}, & \text{if } \langle x_n - y_n \mid y_n^* \rangle > 0; \\
\varepsilon, & \text{otherwise.}
\end{cases}
\tag{5.6}
$$

Fix $n \in \mathbb{N}$. Then, by strong monotonicity of K_n and the Cauchy–Schwarz inequality,

$$
\varepsilon \|x_n - y_n\|^2 \leqslant \langle x_n - y_n \mid K_n x_n - K_n y_n \rangle \leqslant \|x_n - y_n\| \, \|K_n x_n - K_n y_n\|.
\tag{5.7}
$$

This implies that $\langle x_n - y_n \mid y_n^* \rangle = \gamma_n^{-1} \langle x_n - y_n \mid K_n x_n - K_n y_n \rangle \leqslant \gamma_n^{-1} \|x_n - y_n\| \, \|K_n x_n - K_n y_n\| \leqslant (\varepsilon\gamma_n)^{-1} \|K_n x_n - K_n y_n\|^2 = \varepsilon^{-1} \gamma_n \|y_n^*\|^2$ and therefore that $\lambda_n \geqslant \varepsilon$. In addition, by cocoercivity of K_n,

$\gamma_n\|y_n^*\|^2 = \gamma_n^{-1}\|K_n x_n - K_n y_n\|^2 \leqslant (2-\varepsilon)\gamma_n^{-1}\langle x_n - y_n \mid K_n x_n - K_n y_n\rangle = (2-\varepsilon)\langle x_n - y_n \mid y_n^*\rangle$ and thus $\lambda_n \leqslant 2-\varepsilon$. Next, we derive from (5.5) that $y_n = J_{\gamma_n M}^{K_n} x_n$. If $\langle x_n - y_n \mid y_n^*\rangle > 0$, then (5.5) and (5.6) yield $x_{n+1} = x_n - \gamma_n y_n^* = x_n + \lambda_n\langle y_n - x_n \mid y_n^*\rangle y_n^*/\|y_n^*\|^2$. Otherwise, $\langle x_n - y_n \mid y_n^*\rangle \leqslant 0$ and the cocoercivity of K_n yields $\|y_n^*\|^2 = \gamma_n^{-2}\|K_n x_n - K_n y_n\|^2 \leqslant (2-\varepsilon)\gamma_n^{-2}\langle x_n - y_n \mid K_n x_n - K_n y_n\rangle \leqslant 0$. Hence, $y_n^* = 0$ and we therefore deduce from (5.5) that $x_{n+1} = x_n$. Thus (5.5) is an instance of (4.5). Next, condition (ii)[a] in Theorem 4.2 is trivially satisfied and, in view of Remark 4.3, condition (ii)[b] in Theorem 4.2 is also fulfilled. ☐

We conclude this section by further illustrating the effectiveness of warped resolvent iterations by designing a new method to solve an intricate system of monotone inclusions and its dual. We are not aware of a splitting method that could handle such a formulation with a comparable level of flexibility. Special cases of this system appear in [1, 10, 18, 25].

Problem 5.4 Let $(\mathcal{Y}_i)_{i\in I}$ and $(\mathcal{Z}_j)_{j\in J}$ be finite families of real Hilbert spaces. For every $i \in I$ and $j \in J$, let $A_i\colon \mathcal{Y}_i \to 2^{\mathcal{Y}_i}$ and $B_j\colon \mathcal{Z}_j \to 2^{\mathcal{Z}_j}$ be maximally monotone, let $C_i\colon \mathcal{Y}_i \to \mathcal{Y}_i$ be monotone and μ_i-Lipschitzian for some $\mu_i \in \left]0,+\infty\right[$, let $D_j\colon \mathcal{Z}_j \to \mathcal{Z}_j$ be monotone and ν_j-Lipschitzian for some $\nu_j \in \left]0,+\infty\right[$, let $L_{ji} \in \mathcal{B}(\mathcal{Y}_i,\mathcal{Z}_j)$, let $s_i^* \in \mathcal{Y}_i$, and let $r_j \in \mathcal{Z}_j$. Consider the system of coupled inclusions

$$\text{find } (x_i)_{i\in I} \in \bigtimes_{i\in I}\mathcal{Y}_i \text{ such that}$$
$$(\forall i\in I)\quad s_i^* \in A_i x_i + \sum_{j\in J} L_{ji}^*\bigg((B_j + D_j)\bigg(\sum_{k\in I} L_{jk}x_k - r_j\bigg)\bigg) + C_i x_i, \tag{5.8}$$

its dual problem

$$\text{find } (v_j^*)_{j\in J} \in \bigtimes_{j\in J}\mathcal{Z}_j \text{ such that}$$
$$\bigg(\exists\,(x_i)_{i\in I} \in \bigtimes_{i\in I}\mathcal{Y}_i\bigg)(\forall i\in I)(\forall j\in J)\quad \begin{cases} s_i^* - \sum_{k\in J} L_{ki}^* v_k^* \in A_i x_i + C_i x_i\\ v_j^* \in (B_j + D_j)\bigg(\sum_{k\in I} L_{jk}x_k - r_j\bigg), \end{cases} \tag{5.9}$$

and the associated Kuhn–Tucker set

$$Z = \bigg\{\big((x_i)_{i\in I},(v_j^*)_{j\in J}\big) \;\bigg|\; (\forall i\in I)\;\; x_i \in \mathcal{Y}_i \text{ and } s_i^* - \sum_{k\in J} L_{ki}^* v_k^* \in A_i x_i + C_i x_i,$$
$$\text{and } (\forall j\in J)\;\; v_j^* \in \mathcal{Z}_j \text{ and } \sum_{k\in I} L_{jk}x_k - r_j \in (B_j + D_j)^{-1} v_j^*\bigg\}. \tag{5.10}$$

We denote by $\mathscr{P}$ and $\mathscr{D}$ the sets of solutions to (5.8) and (5.9), respectively. The problem is to find a point in Z.

Corollary 5.5 *Consider the setting of Problem 5.4. For every $i \in I$ and every $j \in J$, let $(\alpha_i,\chi_i,\beta_j,\kappa_j) \in \left]0,+\infty\right[^4$, let $\varepsilon_i \in \left]0,\alpha_i/(\mu_i+1)\right[$, let $\delta_j \in \left]0,\beta_j/(\nu_j+1)\right[$, let $(F_{i,n})_{n\in\mathbb{N}}$ be operators from $\mathcal{Y}_i$ to $\mathcal{Y}_i$ that are α_i-strongly monotone and χ_i-Lipschitzian, let $(W_{j,n})_{n\in\mathbb{N}}$ be operators from $\mathcal{Z}_j$ to $\mathcal{Z}_j$ that are β_j-strongly monotone and κ_j-Lipschitzian; in addition, let $(\gamma_{i,n})_{n\in\mathbb{N}}$ and $(\tau_{j,n})_{n\in\mathbb{N}}$ be sequences in $[\varepsilon_i,(\alpha_i-\varepsilon_i)/\mu_i]$ and $[\delta_j,(\beta_j-\delta_j)/\nu_j]$, respectively. Suppose that $Z \neq \varnothing$ and that*

$$\mathcal{Y} = \bigtimes_{i\in I}\mathcal{Y}_i,\quad \mathcal{Z} = \bigtimes_{j\in J}\mathcal{Z}_j,\quad\text{and}\quad \mathcal{X} = \mathcal{Y}\times\mathcal{Z}\times\mathcal{Z}. \tag{5.11}$$

Let $((x_{i,0})_{i\in I},(y_{j,0})_{j\in J},(v^*_{j,0})_{j\in J})$ *and* $((\widetilde{x}_{i,n})_{i\in I},(\widetilde{y}_{j,n})_{j\in J},(\widetilde{v}^*_{j,n})_{j\in J})_{n\in\mathbb{N}}$ *be in* $\mathcal{X}$*, and let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,2[$ *such that* $0<\inf_{n\in\mathbb{N}}\lambda_n\leqslant\sup_{n\in\mathbb{N}}\lambda_n<2$*. Iterate*

$$
\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\text{for every } i\in I\\
\left\lfloor
\begin{array}{l}
l^*_{i,n}=F_{i,n}\widetilde{x}_{i,n}-\gamma_{i,n}C_i\widetilde{x}_{i,n}-\gamma_{i,n}\sum_{j\in J}L^*_{ji}\widetilde{v}^*_{j,n}\\
a_{i,n}=\big(F_{i,n}+\gamma_{i,n}A_i\big)^{-1}(l^*_{i,n}+\gamma_{i,n}s^*_i)\\
o^*_{i,n}=\gamma_{i,n}^{-1}(l^*_{i,n}-F_{i,n}a_{i,n})+C_ia_{i,n}
\end{array}
\right.\\
\text{for every } j\in J\\
\left\lfloor
\begin{array}{l}
t^*_{j,n}=W_{j,n}\widetilde{y}_{j,n}-\tau_{j,n}D_j\widetilde{y}_{j,n}+\tau_{j,n}\widetilde{v}^*_{j,n}\\
b_{j,n}=\big(W_{j,n}+\tau_{j,n}B_j\big)^{-1}t^*_{j,n}\\
f^*_{j,n}=\tau_{j,n}^{-1}(t^*_{j,n}-W_{j,n}b_{j,n})+D_jb_{j,n}\\
c_{j,n}=\sum_{i\in I}L_{ji}\widetilde{x}_{i,n}-\widetilde{y}_{j,n}+\widetilde{v}^*_{j,n}-r_j
\end{array}
\right.\\
\text{for every } i\in I\\
\left\lfloor\; a^*_{i,n}=o^*_{i,n}+\sum_{j\in J}L^*_{ji}c_{j,n}\right.\\
\text{for every } j\in J\\
\left\lfloor
\begin{array}{l}
b^*_{j,n}=f^*_{j,n}-c_{j,n}\\
c^*_{j,n}=r_j+b_{j,n}-\sum_{i\in I}L_{ji}a_{i,n}
\end{array}
\right.\\
\sigma_n=\sum_{i\in I}\|a^*_{i,n}\|^2+\sum_{j\in J}\big(\|b^*_{j,n}\|^2+\|c^*_{j,n}\|^2\big)\\
\theta_n=\sum_{i\in I}\langle a_{i,n}-x_{i,n}\mid a^*_{i,n}\rangle+\sum_{j\in J}\big(\langle b_{j,n}-y_{j,n}\mid b^*_{j,n}\rangle+\langle c_{j,n}-v^*_{j,n}\mid c^*_{j,n}\rangle\big)\\
\text{if } \theta_n<0\\
\left\lfloor\; \rho_n=\lambda_n\theta_n/\sigma_n\right.\\
\text{else}\\
\left\lfloor\; \rho_n=0\right.\\
\text{for every } i\in I\\
\left\lfloor\; x_{i,n+1}=x_{i,n}+\rho_na^*_{i,n}\right.\\
\text{for every } j\in J\\
\left\lfloor
\begin{array}{l}
y_{j,n+1}=y_{j,n}+\rho_nb^*_{j,n}\\
v^*_{j,n+1}=v^*_{j,n}+\rho_nc^*_{j,n}.
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{5.12}
$$

Suppose that

$$(\forall i\in I)(\forall j\in J)\quad \widetilde{x}_{i,n}-x_{i,n}\to0,\quad \widetilde{y}_{j,n}-y_{j,n}\to0,\quad \textit{and}\quad \widetilde{v}^*_{j,n}-v^*_{j,n}\to0. \tag{5.13}$$

Set $(\forall n\in\mathbb{N})$ $x_n=(x_{i,n})_{i\in I}$ *and* $v^*_n=(v^*_{j,n})_{j\in J}$*. Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point* $\overline{x}\in\mathscr{P}$*,* $(v^*_n)_{n\in\mathbb{N}}$ *converges weakly to a point* $\overline{v}^*\in\mathscr{D}$*, and* $(\overline{x},\overline{v}^*)\in Z$*.*

Proof. Define

$$
\begin{cases}
A\colon \mathcal{Y}\to 2^{\mathcal{Y}}\colon (x_i)_{i\in I}\mapsto \bigtimes_{i\in I}(A_ix_i+C_ix_i)\\
B\colon \mathcal{Z}\to 2^{\mathcal{Z}}\colon (y_j)_{j\in J}\mapsto \bigtimes_{j\in J}(B_jy_j+D_jy_j)\\
L\colon \mathcal{Y}\to\mathcal{Z}\colon (x_i)_{i\in I}\mapsto \bigg(\displaystyle\sum_{i\in I}L_{ji}x_i\bigg)_{j\in J}\\
s^*=(s^*_i)_{i\in I}\quad\text{and}\quad r=(r_j)_{j\in J}.
\end{cases}
\tag{5.14}
$$

We observe that

$$L^*\colon \mathcal{Z} \to \mathcal{Y}\colon (v_j^*)_{j\in J} \mapsto \left(\sum_{j\in J} L_{ji}^* v_j^*\right)_{i\in I}. \tag{5.15}$$

In the light of [7, Proposition 20.23], A and B are maximally monotone. On the other hand, we deduce from (5.10), (5.14), and (5.15) that

$$Z = \big\{(x, v^*) \in \mathcal{Y} \times \mathcal{Z} \mid s^* - L^*v^* \in Ax \text{ and } Lx - r \in B^{-1}v^*\big\}. \tag{5.16}$$

Define

$$M\colon \mathcal{X} \to 2^{\mathcal{X}}\colon (x, y, v^*) \mapsto (-s^* + Ax + L^*v^*) \times (By - v^*) \times \{r - Lx + y\}. \tag{5.17}$$

Lemma 2.2(ii) entails that M is maximally monotone. Furthermore, since $Z \neq \varnothing$, Lemma 2.2(iv) yields $\operatorname{zer} M \neq \varnothing$. Next, set

$$S\colon \mathcal{X} \to \mathcal{X}\colon (x, y, v^*) \mapsto (-L^*v^*, v^*, Lx - y) \tag{5.18}$$

and, for every $n \in \mathbb{N}$,

$$K_n\colon \mathcal{X} \to \mathcal{X}\colon (x, y, v^*) \mapsto \Big(\big(\gamma_{i,n}^{-1} F_{i,n} x_i - C_i x_i\big)_{i\in I} - L^*v^*, \big(\tau_{j,n}^{-1} W_{j,n} y_j - D_j y_j\big)_{j\in J} + v^*, Lx - y + v^*\Big) \tag{5.19}$$

and

$$T_n\colon \mathcal{X} \to \mathcal{X}\colon (x, y, v^*) \mapsto \Big(\big(\gamma_{i,n}^{-1} F_{i,n} x_i - C_i x_i\big)_{i\in I}, \big(\tau_{j,n}^{-1} W_{j,n} y_j - D_j y_j\big)_{j\in J}, v^*\Big). \tag{5.20}$$

For every $i \in I$ and every $n \in \mathbb{N}$, using the facts that C_i is μ_i-Lipschitzian, that $F_{i,n}$ is α_i-strongly monotone, and that $\gamma_{i,n} \in [\varepsilon_i, (\alpha_i - \varepsilon_i)/\mu_i]$, Lemma 5.1(i) implies that $F_{i,n} - \gamma_{i,n} C_i$ is ε_i-strongly monotone and therefore, since $\gamma_{i,n}^{-1} \geqslant \mu_i/(\alpha_i - \varepsilon_i)$, it follows that $\gamma_{i,n}^{-1} F_{i,n} - C_i$ is strongly monotone with constant $\varepsilon_i\mu_i/(\alpha_i - \varepsilon_i)$. Likewise, for every $j \in J$ and every $n \in \mathbb{N}$, $\tau_{j,n}^{-1} W_{j,n} - D_j$ is strongly monotone with constant $\delta_j\nu_j/(\beta_j - \delta_j)$. Thus, upon setting

$$\vartheta = \min\left\{\min_{i\in I} \frac{\varepsilon_i\mu_i}{\alpha_i - \varepsilon_i}, \min_{j\in J} \frac{\delta_j\nu_j}{\beta_j - \delta_j}, 1\right\}, \tag{5.21}$$

we get

$$\begin{aligned}
&(\forall n \in \mathbb{N})\big(\forall (x, y, v^*) \in \mathcal{X}\big)\big(\forall (a, b, c^*) \in \mathcal{X}\big)\\
&\qquad \big\langle (x, y, v^*) - (a, b, c^*) \mid T_n(x, y, v^*) - T_n(a, b, c^*)\big\rangle\\
&\qquad\qquad = \sum_{i\in I} \big\langle x_i - a_i \mid \big(\gamma_{i,n}^{-1} F_{i,n} x_i - C_i x_i\big) - \big(\gamma_{i,n}^{-1} F_{i,n} a_i - C_i a_i\big)\big\rangle\\
&\qquad\qquad\quad + \sum_{j\in J} \big\langle y_j - b_j \mid \big(\tau_{j,n}^{-1} W_{j,n} y_j - D_j y_j\big) - \big(\tau_{j,n}^{-1} W_{j,n} b_j - D_j b_j\big)\big\rangle + \|v^* - c^*\|^2\\
&\qquad\qquad \geqslant \vartheta \sum_{i\in I} \|x_i - a_i\|^2 + \vartheta \sum_{j\in J} \|y_j - b_j\|^2 + \vartheta \|v^* - c^*\|^2\\
&\qquad\qquad = \vartheta \|(x, y, v^*) - (a, b, c^*)\|^2.
\end{aligned} \tag{5.22}$$

Hence, the operators $(T_n)_{n\in\mathbb{N}}$ are ϑ-strongly monotone. However, S is linear, bounded, and $S^*=-S$. It follows that the kernels $(K_n)_{n\in\mathbb{N}}=(T_n+S)_{n\in\mathbb{N}}$ are ϑ-strongly monotone. Now, for every $i\in I$ and every $n\in\mathbb{N}$, since $\gamma_{i,n}^{-1}F_{i,n}$ is Lipschitzian with constant χ_i/ε_i, we deduce that $\gamma_{i,n}^{-1}F_{i,n}-C_i$ is Lipschitzian with constant $\chi_i/\varepsilon_i+\mu_i$. Likewise, for every $j\in J$ and every $n\in\mathbb{N}$, $\tau_{j,n}^{-1}W_{j,n}-D_j$ is Lipschitzian with constant $\kappa_j/\delta_j+\nu_j$. Hence, upon setting

$$\eta=\max\Big\{\max_{i\in I}\{\chi_i/\varepsilon_i+\mu_i\},\max_{j\in J}\{\kappa_j/\delta_j+\nu_j\},1\Big\}, \tag{5.23}$$

we obtain

$$\begin{aligned}
(\forall n\in\mathbb{N})\big(\forall(x,y,v^*)\in\mathcal{X}\big)\big(\forall(a,b,c^*)\in\mathcal{X}\big)\quad &\|T_n(x,y,v^*)-T_n(a,b,c^*)\|^2\\
&=\sum_{i\in I}\big\|\big(\gamma_{i,n}^{-1}F_{i,n}x_i-C_ix_i\big)-\big(\gamma_{i,n}^{-1}F_{i,n}a_i-C_ia_i\big)\big\|^2\\
&\quad+\sum_{j\in J}\big\|\big(\tau_{j,n}^{-1}W_{j,n}y_j-D_jy_j\big)-\big(\tau_{j,n}^{-1}W_{j,n}b_j-D_jb_j\big)\big\|^2+\|v^*-c^*\|^2\\
&\leqslant\eta^2\sum_{i\in I}\|x_i-a_i\|^2+\eta^2\sum_{j\in J}\|y_j-b_j\|^2+\eta^2\|v^*-c^*\|^2\\
&=\eta^2\|(x,y,v^*)-(a,b,c^*)\|^2.
\end{aligned} \tag{5.24}$$

This implies that the operators $(T_n)_{n\in\mathbb{N}}$ are η-Lipschitzian. On the other hand, S is Lipschitzian with constant $\|S\|$. Altogether, the kernels $(K_n)_{n\in\mathbb{N}}$ are Lipschitzian with constant $\eta+\|S\|$. In turn, using Proposition 3.9(i)[d]&(ii)[b], we infer that, for every $n\in\mathbb{N}$, $\operatorname{ran}K_n\subset\operatorname{ran}(K_n+M)$ and K_n+M is injective. Now set

$$\begin{aligned}
(\forall n\in\mathbb{N})\quad p_n&=\big((x_{i,n})_{i\in I},(y_{j,n})_{j\in J},(v_{j,n}^*)_{j\in J}\big),\quad \widetilde{p}_n=\big((\widetilde{x}_{i,n})_{i\in I},(\widetilde{y}_{j,n})_{j\in J},(\widetilde{v}_{j,n}^*)_{j\in J}\big),\\
q_n&=\big((a_{i,n})_{i\in I},(b_{j,n})_{j\in J},(c_{j,n})_{j\in J}\big),\quad\text{and}\quad q_n^*=\big((a_{i,n}^*)_{i\in I},(b_{j,n}^*)_{j\in J},(c_{j,n}^*)_{j\in J}\big).
\end{aligned} \tag{5.25}$$

In view of (5.19), (5.17), (5.14), and (5.15), we deduce that (5.12) assumes the form

$$\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
q_n=J_M^{K_n}\widetilde{p}_n\\
q_n^*=K_n\widetilde{p}_n-K_nq_n\\
\text{if } \langle q_n-p_n\mid q_n^*\rangle<0\\
\left\lfloor\begin{array}{l} p_{n+1}=p_n+\dfrac{\lambda_n\langle q_n-p_n\mid q_n^*\rangle}{\|q_n^*\|^2}\,q_n^*\end{array}\right.\\
\text{else}\\
\left\lfloor\begin{array}{l} p_{n+1}=p_n.\end{array}\right.
\end{array}\right.
\end{array} \tag{5.26}$$

In addition, (5.13) implies that $\widetilde{p}_n-p_n\to0$. Altogether, in the light of Theorem 4.2 and Remark 4.3, there exists $(\overline{x},\overline{y},\overline{v}^*)\in\operatorname{zer}M$ such that $p_n\rightharpoonup(\overline{x},\overline{y},\overline{v}^*)$. It follows that $x_n\rightharpoonup\overline{x}$ and $v_n^*\rightharpoonup\overline{v}^*$. Further, we conclude by using Lemma 2.2(iii) that $\overline{x}\in\mathscr{P}$, $\overline{v}^*\in\mathscr{D}$, and $(\overline{x},\overline{v}^*)\in Z$. ☐

References

[1] A. Alotaibi, P. L. Combettes, and N. Shahzad, Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set, *SIAM J. Optim.*, vol. 24, pp. 2076–2095, 2014.

[2] A. Alotaibi, P. L. Combettes, and N. Shahzad, Best approximation from the Kuhn-Tucker set of composite monotone inclusions, *Numer. Funct. Anal. Optim.*, vol. 36, pp. 1513–1532, 2015.

[3] H. Attouch and A. Cabot, Convergence of a relaxed inertial proximal algorithm for maximally monotone operators, *Math. Program. A*, published online 2019-06-29.

[4] J.-B. Baillon, R. E. Bruck, and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, *Houston J. Math.*, vol. 4, pp. 1–9, 1978.

[5] S. Banert, A. Ringh, J. Adler, J. Karlsson, and O. Öktem, Data-driven nonsmooth optimization, *SIAM J. Optim.,* vol. 30, pp. 102–131, 2020.

[6] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, pp. 596–636, 2003.

[7] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed., correct. printing. Springer, New York, 2019.

[8] H. H. Bauschke, X. Wang, and L. Yao, General resolvents for monotone operators: Characterization and extension, in: *Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning and Inverse Problems*, (Y. Censor, M. Jiang, and G. Wang, eds.), pp. 57–74. Medical Physics Publishing, Madison, WI, 2010.

[9] R. I. Boţ and E. R. Csetnek, ADMM for monotone operators: Convergence analysis and rates, *Adv. Comput. Math.*, vol. 45, pp. 327–359, 2019.

[10] R. I. Boţ, E. R. Csetnek, and A. Heinrich, A primal-dual splitting algorithm for finding zeros of sums of maximal monotone operators, *SIAM J. Optim.*, vol. 23, pp. 2011–2036, 2013.

[11] R. I. Boţ and C. Hendrich, A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators, *SIAM J. Optim.*, vol. 23, pp. 2541–2565, 2013.

[12] L. M. Briceño-Arias, Forward-partial inverse-forward splitting for solving monotone inclusions, *J. Optim. Theory Appl.*, vol. 166, pp. 391–413, 2015.

[13] M. N. Bùi and P. L. Combettes, Bregman forward-backward operator splitting, 2019-09-13.
https://arxiv.org/pdf/1908.03878

[14] M. N. Bùi and P. L. Combettes, Warped proximal iterations for monotone inclusions, 2019-08-19.
https://arxiv.org/pdf/1908.07077v1

[15] P. L. Combettes, Construction d'un point fixe commun à une famille de contractions fermes, *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 320, pp. 1385–1390, 1995.

[16] P. L. Combettes, Fejér-monotonicity in convex optimization, in: *Encyclopedia of Optimization*, (C. A. Floudas and P. M. Pardalos, Eds.), vol. 2, Springer-Verlag, New York, 2001, pp. 106–114. (Also available in 2nd ed., pp. 1016–1024, 2009.)

[17] P. L. Combettes, Systems of structured monotone inclusions: Duality, algorithms, and applications, *SIAM J. Optim.*, vol. 23, pp. 2420–2447, 2013.

[18] P. L. Combettes and J. Eckstein, Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions, *Math. Program.*, vol. B168, pp. 645–672, 2018.

[19] P. L. Combettes and L. E. Glaudin, Quasinonexpansive iterations on the affine hull of orbits: From Mann's mean value algorithm to inertial methods, *SIAM J. Optim.*, vol. 27, pp. 2356–2380, 2017.

[20] P. L. Combettes and Q. V. Nguyen, Solving composite monotone inclusions in reflexive Banach spaces by constructing best Bregman approximations from their Kuhn-Tucker set, *J. Convex Anal.*, vol. 23, pp. 481–510, 2016.

[21] P. L. Combettes and J.-C. Pesquet, Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators, *Set-Valued Var. Anal.*, vol. 20, pp. 307–330, 2012.

[22] L. Condat, A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, *J. Optim. Theory Appl.*, vol. 158, pp. 460–479, 2013.

[23] P. Giselsson, Nonlinear forward-backward splitting with projection correction, 2019-08-20. https://arxiv.org/pdf/1908.07449v1

[24] Y. Haugazeau, *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes.* Thèse, Université de Paris, Paris, France, 1968.

[25] P. R. Johnstone and J. Eckstein, Projective splitting with forward steps: Asynchronous and block-iterative operator splitting. https://arxiv.org/pdf/1803.07043.pdf

[26] G. Kassay, The proximal points algorithm for reflexive Banach spaces, *Studia Univ. Babeş-Bolyai Math.*, vol. 30, pp. 9–17, 1985.

[27] K. C. Kiwiel and B. Łopuch, Surrogate projection methods for finding fixed points of firmly nonexpansive mappings, *SIAM J. Optim.*, vol. 7, pp. 1084–1102, 1997.

[28] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A*, vol. 255, pp. 2897–2899, 1962.

[29] T. Pennanen, Dualization of generalized equations of maximal monotone type, *SIAM J. Optim.*, vol. 10, pp. 809–835, 2000.

[30] H. Raguet, A note on the forward-Douglas–Rachford splitting for monotone inclusion and convex optimization, *Optim. Lett.*, vol. 13, pp. 717–740, 2019.

[31] A. Renaud and G. Cohen, An extension of the auxiliary problem principle to nonsymmetric auxiliary operators, *ESAIM Control Optim. Calc. Var.*, vol. 2, pp. 281–306, 1997.

[32] S. M. Robinson, Composition duality and maximal monotonicity, *Math. Program.*, vol. 85, pp. 1–13, 1999.

[33] R. T. Rockafellar, On the maximality of sums of nonlinear monotone operators, *Trans. Amer. Math. Soc.*, vol. 149, no. 1, pp. 75–88, 1970.

[34] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.*, vol. 14, pp. 877–898, 1976.

[35] S. Simons, *From Hahn-Banach to Monotonicity*, Lecture Notes in Math. 1693, Springer-Verlag, New York, 2008.

[36] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.*, vol. 38, pp. 431–446, 2000.

[37] B. C. Vũ, A splitting algorithm for dual monotone inclusions involving cocoercive operators, *Adv. Comput. Math.*, vol. 38, pp. 667–681, 2013.

[38] C. Zălinescu, *Convex Analysis in General Vector Spaces.* World Scientific Publishing, River Edge, NJ, 2002.

[39] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B – Nonlinear Monotone Operators,* Springer-Verlag, New York, 1990.