A Monotone + Skew Splitting Model for Composite Monotone Inclusions in Duality*

Luis M. Briceño-Arias[1,2] and Patrick L. Combettes[1]

[1]UPMC Université Paris 06
Laboratoire Jacques-Louis Lions – UMR CNRS 7598
75005 Paris, France

[2]UPMC Université Paris 06
Équipe Combinatoire et Optimisation – UMR CNRS 7090
75005 Paris, France

Abstract

The principle underlying this paper is the basic observation that the problem of simultaneously solving a large class of composite monotone inclusions and their duals can be reduced to that of finding a zero of the sum of a maximally monotone operator and a linear skew-adjoint operator. An algorithmic framework is developed for solving this generic problem in a Hilbert space setting. New primal-dual splitting algorithms are derived from this framework for inclusions involving composite monotone operators, and convergence results are established. These algorithms draw their simplicity and efficacy from the fact that they operate in a fully decomposed fashion in the sense that the monotone operators and the linear transformations involved are activated separately at each iteration. Comparisons with existing methods are made and applications to composite variational problems are demonstrated.

2000 Mathematics Subject Classification: Primary 47H05; Secondary 47J25, 65K05, 90C25.

Keywords: composite operator, convex optimization, decomposition, duality, Fenchel-Rockafellar duality, minimization algorithm, monotone inclusion, monotone operator, operator splitting

---

*Contact author: P. L. Combettes, `plc@math.jussieu.fr`, phone: +33 1 4427 6319, fax: +33 1 4427 7200. This work was supported by the Agence Nationale de la Recherche under grant ANR-08-BLAN-0294-02.

1 Introduction

A wide range of problems in areas such as optimization, variational inequalities, partial differential equations, mechanics, economics, signal and image processing, or traffic theory can be reduced to solving inclusions involving monotone set-valued operators in a Hilbert space $\mathcal{H}$, say

$$\text{find } x \in \mathcal{H} \quad \text{such that} \quad z \in Mx, \tag{1.1}$$

where $M\colon \mathcal{H} \to 2^{\mathcal{H}}$ is monotone and $z \in \mathcal{H}$, e.g., [12, 13, 18, 19, 22, 25, 33, 37, 38, 41]. In many formulations of this type, the operator M can be expressed as the sum of two monotone operators, one of which is the composition of a monotone operator with a linear transformation and its adjoint. In such situations, it is often desirable to also solve an associated dual inclusion [1, 3, 4, 15, 21, 26, 27, 28, 30, 31, 32]. The present paper is concerned with the numerical solution of such composite inclusion problems in duality. More formally, the basic problem we consider is the following.

Problem 1.1 Let $\mathcal{H}$ and $\mathcal{G}$ be two real Hilbert spaces, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone, let $L\colon \mathcal{H} \to \mathcal{G}$ be linear and bounded, let $z \in \mathcal{H}$, and let $r \in \mathcal{G}$. The problem is to solve the primal inclusion

$$\text{find } x \in \mathcal{H} \quad \text{such that} \quad z \in Ax + L^*B(Lx - r) \tag{1.2}$$

together with the dual inclusion

$$\text{find } v \in \mathcal{G} \quad \text{such that} \quad -r \in -LA^{-1}(z - L^*v) + B^{-1}v. \tag{1.3}$$

The set of solutions to (1.2) is denoted by $\mathcal{P}$ and the set of solutions to (1.3) by $\mathcal{D}$.

A classical instance of the duality scheme described in Problem 1.1 is the Fenchel-Rockafellar framework [32] which, under suitable constraint qualification, corresponds to letting A and B be subdifferentials of proper lower semicontinuous convex functions $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ and $g\colon \mathcal{G} \to \left]-\infty, +\infty\right]$, respectively. In this scenario, the problems in duality are

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + g(Lx - r) - \langle x \mid z \rangle \tag{1.4}$$

and

$$\underset{v \in \mathcal{G}}{\text{minimize}} \;\; f^*(z - L^*v) + g^*(v) + \langle v \mid r \rangle\,. \tag{1.5}$$

Extensions of the Fenchel-Rockafellar framework to variational inequalities were considered in [1, 17, 21, 27], while extensions to saddle function problems were proposed in [24]. On the other hand, general monotone operators were investigated in [3, 4, 7, 26] in the case when $\mathcal{G} = \mathcal{H}$ and $L = \mathrm{Id}$. The general duality setting described in Problem 1.1 appears in [15, 28, 30].

Our objective is to devise an algorithm which solves (1.2) and (1.3) simultaneously, and which uses the operators A, B, and L separately. In the literature, several splitting algorithms are available for solving the primal problem (1.2), but they are restricted by stringent hypotheses. Let us set

$$A_1\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto -z + Ax \quad \text{and} \quad A_2\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto L^*B(Lx - r), \tag{1.6}$$

and observe that solving (1.2) is equivalent to finding a zero of $A_1 + A_2$. If B is single-valued and cocoercive (its inverse is strongly monotone), then so is A_2, and (1.2) can be solved by the forward-backward algorithm [10, 25, 38]. If B is merely Lipschitzian, or even just continuous, so is A_2, and

(1.2) can then be solved via the algorithm proposed in [39]. These algorithms employ the resolvent of A_1, which is easily derived from that of A, and explicit applications of A_2, i.e., of B and L. They are however limited in scope by the fact that B must be single-valued and smooth. The main splitting algorithm to find a zero of $A_1 + A_2$ when both operators are set-valued is the Douglas-Rachford algorithm [11, 14, 23, 36]. This algorithm requires that both operators be maximally monotone and that their resolvents be computable to within some quantifiable error. Unfortunately, these conditions are seldom met in the present setting since A_2 may not be maximally monotone [28, 30] and, more importantly, since there is no convenient rule to compute the resolvent of A_2 in terms of L and the resolvent of B unless stringent conditions are imposed on L (see [6, Proposition 23.23] and [19]).

Our approach is motivated by the classical Kuhn-Tucker theory [35], which asserts that points $\overline{x} \in \mathcal{H}$ and $\overline{v} \in \mathcal{G}$ satisfying the conditions

$$(0,0) \in \big(-z + \partial f(\overline{x}) + L^*\overline{v},\, r + \partial g^*(\overline{v}) - L\overline{x} \big) \tag{1.7}$$

are solutions to (1.4) and (1.5), respectively. By analogy, it is natural to consider the following problem in conjunction with Problem 1.1.

Problem 1.2 In the setting of Problem 1.1, let $\boldsymbol{\mathcal{K}} = \mathcal{H} \oplus \mathcal{G}$ and set

$$\boldsymbol{M} \colon \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}} \colon (x,v) \mapsto (-z + Ax) \times (r + B^{-1}v) \quad \text{and} \quad \boldsymbol{S} \colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}} \colon (x,v) \mapsto (L^*v, -Lx). \tag{1.8}$$

The problem is to

$$\text{find } \boldsymbol{x} \in \boldsymbol{\mathcal{K}} \quad \text{such that} \quad \boldsymbol{0} \in \boldsymbol{M}\boldsymbol{x} + \boldsymbol{S}\boldsymbol{x}. \tag{1.9}$$

The investigation of this companion problem may have various purposes [1, 15, 28, 30]. Ours is to exploit its simple structure to derive a new splitting algorithm to solve efficiently Problem 1.1. The crux of our approach is the simple observation that (1.9) reduces the original primal- dual problem (1.2)–(1.3) to that of finding a zero of the sum of a maximally monotone operator $\boldsymbol{M}$ and a bounded linear skew-adjoint transformation $\boldsymbol{S}$. In Section 2 we establish the convergence of an inexact splitting algorithm proposed in its original form in [39]. Each iteration of this forward-backward-forward scheme performs successively an explicit step on $\boldsymbol{S}$, an implicit step on $\boldsymbol{M}$, and another explicit step on $\boldsymbol{S}$. We then review the tight connections existing between Problem 1.1 and Problem 1.2 and, in particular, the fact that solving the latter provides a solution to the former. In Section 3, we apply the forward-backward-forward algorithm to the monotone+skew Problem 1.2 and obtain a new type of splitting algorithm for solving (1.2) and (1.3) simultaneously. The main feature of this scheme, that distinguishes it from existing techniques, is that at each iteration it employs the operators A, B, and L separately without requiring any additional assumption to those stated above except, naturally, existence of solutions. Using a product space technique, we then obtain a parallel splitting method for solving the m-term inclusion

$$\text{find } x \in \mathcal{H} \quad \text{such that} \quad z \in \sum_{i=1}^{m} L_i^* B_i(L_i x - r_i), \tag{1.10}$$

where each maximally monotone operator B_i acts on a Hilbert space $\mathcal{G}_i$, $r_i \in \mathcal{G}_i$, and $L_i \colon \mathcal{H} \to \mathcal{G}_i$ is linear and bounded. Applications to variational problems are discussed in Section 4, where we provide a proximal splitting scheme for solving the primal dual problem (1.4)–(1.5), as well as one for minimizing the sum of m composite functions.

Notation. We denote the scalar products of $\mathcal{H}$ and $\mathcal{G}$ by $\langle \cdot \mid \cdot \rangle$ and the associated norms by $\|\cdot\|$. $\mathcal{B}(\mathcal{H},\mathcal{G})$ is the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$, $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H},\mathcal{H})$, and the symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence. Moreover, $\mathcal{H}\oplus\mathcal{G}$ denotes the Hilbert direct sum of $\mathcal{H}$ and $\mathcal{G}$. The projector onto a nonempty closed convex set $C \subset \mathcal{H}$ is denoted by P_C, and its normal cone operator by N_C, i.e.,

$$N_C\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \begin{cases} \big\{u \in \mathcal{H} \mid (\forall y \in C) \quad \langle y - x \mid u\rangle \leq 0\big\}, & \text{if } x \in C;\\ \varnothing, & \text{otherwise.}\end{cases} \tag{1.11}$$

Let $M\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. We denote by $\operatorname{ran} M = \big\{u \in \mathcal{H} \mid (\exists\, x \in \mathcal{H})\ u \in Mx\big\}$ the range of M, by $\operatorname{dom} M = \big\{x \in \mathcal{H} \mid Mx \neq \varnothing\big\}$ its domain, by $\operatorname{zer} M = \big\{x \in \mathcal{H} \mid 0 \in Mx\big\}$ its set of zeros, by $\operatorname{gra} M = \big\{(x,u) \in \mathcal{H}\times\mathcal{H} \mid u \in Mx\big\}$ its graph, and by M^{-1} its inverse, i.e., the operator with graph $\big\{(u,x) \in \mathcal{H}\times\mathcal{H} \mid u \in Mx\big\}$. The resolvent of M is $J_M = (\operatorname{Id} + M)^{-1}$. Moreover, M is monotone if

$$(\forall (x,y) \in \mathcal{H}\times\mathcal{H})(\forall (u,v) \in Mx \times My) \quad \langle x - y \mid u - v\rangle \geq 0, \tag{1.12}$$

and maximally so if there exists no monotone operator $\widetilde{M}\colon \mathcal{H} \to 2^{\mathcal{H}}$ such that $\operatorname{gra} M \subset \operatorname{gra} \widetilde{M} \neq \operatorname{gra} M$. In this case, J_M is a nonexpansive operator defined everywhere in $\mathcal{H}$. For background on convex analysis and monotone operator theory, the reader is referred to [6, 40].

2 Preliminary results

2.1 Technical facts

The following results will be needed subsequently.

Lemma 2.1 [9, Lemma 3.1] *Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence in $[0,+\infty[$, let $(\beta_n)_{n\in\mathbb{N}}$ be a sequence in $[0,+\infty[$, and let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a summable sequence in $[0,+\infty[$ such that $(\forall n \in \mathbb{N})$ $\alpha_{n+1} \leq \alpha_n - \beta_n + \varepsilon_n$. Then $(\alpha_n)_{n\in\mathbb{N}}$ converges and $(\beta_n)_{n\in\mathbb{N}}$ is summable.*

Lemma 2.2 [9, Theorem 3.8] *Let C be a nonempty subset of $\mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. Suppose that, for every $x \in C$, there exists a summable sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ in $[0,+\infty[$ such that*

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 + \varepsilon_n, \tag{2.1}$$

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

Definition 2.3 [2, Definition 2.3] *An operator $M\colon \mathcal{H} \to 2^{\mathcal{H}}$ is* demiregular *at $x \in \operatorname{dom} M$ if, for every sequence $((x_n,u_n))_{n\in\mathbb{N}}$ in $\operatorname{gra} M$ and every $u \in Mx$ such that $x_n \rightharpoonup x$ and $u_n \to u$, we have $x_n \to x$.*

Lemma 2.4 [2, Proposition 2.4] *Let $M\colon \mathcal{H} \to 2^{\mathcal{H}}$ and let $x \in \operatorname{dom} M$. Then M is demiregular at x in each of the following cases.*

(i) *M is uniformly monotone at x, i.e., there exists an increasing function $\phi\colon\ [0,+\infty[\ \to [0,+\infty]$ that vanishes only at 0 such that $(\forall u \in Mx)(\forall (y,v) \in \operatorname{gra} M)$ $\langle x - y \mid u - v\rangle \geq \phi(\|x-y\|)$. In particular, M is uniformly monotone, i.e., these inequalities hold for every $x \in \operatorname{dom} M$ and, a fortiori, M is α-strongly monotone, i.e., $M - \alpha\operatorname{Id}$ is monotone for some $\alpha \in\ ]0,+\infty[$.*

(ii) *J_M is compact, i.e., for every bounded set $C \subset \mathcal{H}$, the closure of $J_M(C)$ is compact. In particular,* $\operatorname{dom} M$ *is boundedly relatively compact, i.e., the intersection of its closure with every closed ball is compact.*

(iii) *$M\colon \mathcal{H} \to \mathcal{H}$ is single-valued with a single-valued continuous inverse.*

(iv) *M is single-valued on* $\operatorname{dom} M$ *and* $\operatorname{Id} - M$*, i.e., for every bounded sequence $(x_n)_{n\in\mathbb{N}}$ in* $\operatorname{dom} M$ *such that $(Mx_n)_{n\in\mathbb{N}}$ converges strongly, $(x_n)_{n\in\mathbb{N}}$ admits a strong cluster point.*

2.2 An inexact forward-backward-forward algorithm

Our algorithmic framework will hinge on the following splitting algorithm, which was proposed in the error-free case in [39]. We provide an analysis of the asymptotic behavior of a inexact version of this method which is of interest in its own right.

Theorem 2.5 *Let $\boldsymbol{\mathcal{H}}$ be a real Hilbert space, let $\boldsymbol{A}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ be maximally monotone, and let $\boldsymbol{B}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}$ be monotone. Suppose that $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) \neq \varnothing$ and that $\boldsymbol{B}$ is β-Lipschitzian for some $\beta \in \left]0,+\infty\right[$. Let $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ be sequences in $\boldsymbol{\mathcal{H}}$ such that*

$$\sum_{n\in\mathbb{N}} \|\boldsymbol{a}_n\| < +\infty, \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{b}_n\| < +\infty, \quad \textit{and} \quad \sum_{n\in\mathbb{N}} \|\boldsymbol{c}_n\| < +\infty, \tag{2.2}$$

let $\boldsymbol{x}_0 \in \boldsymbol{\mathcal{H}}$, let $\varepsilon \in \left]0, 1/(\beta+1)\right[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{a}_n) \\ \boldsymbol{p}_n = J_{\gamma_n \boldsymbol{A}}\, \boldsymbol{y}_n + \boldsymbol{b}_n \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n. \end{array} \right. \tag{2.3}$$

Then the following hold for some $\overline{\boldsymbol{x}} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$.

(i) *$\sum_{n\in\mathbb{N}} \|\boldsymbol{x}_n - \boldsymbol{p}_n\|^2 < +\infty$ and $\sum_{n\in\mathbb{N}} \|\boldsymbol{y}_n - \boldsymbol{q}_n\|^2 < +\infty$.*

(ii) *$\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ and $\boldsymbol{p}_n \rightharpoonup \overline{\boldsymbol{x}}$.*

(iii) *Suppose that one of the following is satisfied.*

  (a) *$\boldsymbol{A}+\boldsymbol{B}$ is demiregular at $\overline{\boldsymbol{x}}$.*

  (b) *$\boldsymbol{A}$ or $\boldsymbol{B}$ is uniformly monotone at $\overline{\boldsymbol{x}}$.*

  (c) $\operatorname{int}\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) \neq \varnothing$.

  Then $\boldsymbol{x}_n \to \overline{\boldsymbol{x}}$ and $\boldsymbol{p}_n \to \overline{\boldsymbol{x}}$.

Proof. Let us set

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \widetilde{\boldsymbol{y}}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{B}\boldsymbol{x}_n \\ \widetilde{\boldsymbol{p}}_n = J_{\gamma_n \boldsymbol{A}}\, \widetilde{\boldsymbol{y}}_n \\ \widetilde{\boldsymbol{q}}_n = \widetilde{\boldsymbol{p}}_n - \gamma_n \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{u}_n = \gamma_n^{-1}(\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n) + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n - \boldsymbol{B}\boldsymbol{x}_n \\ \boldsymbol{e}_n = \boldsymbol{y}_n - \boldsymbol{q}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n. \end{cases} \tag{2.4}$$

Then

$$(\forall n \in \mathbb{N}) \quad \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n \in \gamma_n \boldsymbol{A}\widetilde{\boldsymbol{p}}_n \quad \text{and} \quad \boldsymbol{u}_n = \gamma_n^{-1}(\widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n) + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \in \boldsymbol{A}\widetilde{\boldsymbol{p}}_n + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n. \tag{2.5}$$

Now let $\boldsymbol{x} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ and let $n \in \mathbb{N}$. We first note that $(\boldsymbol{x}, -\gamma_n \boldsymbol{B}\boldsymbol{x}) \in \operatorname{gra} \gamma_n \boldsymbol{A}$. On the other hand, (2.5) yields $(\widetilde{\boldsymbol{p}}_n, \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n) \in \operatorname{gra} \gamma_n \boldsymbol{A}$. Hence, by monotonicity of $\gamma_n \boldsymbol{A}$, $\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n \boldsymbol{B}\boldsymbol{x} \rangle \leq 0$. However, by monotonicity of $\boldsymbol{B}$, $\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n \boldsymbol{B}\boldsymbol{x} - \gamma_n \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \rangle \leq 0$. Upon adding these two inequalities, we obtain $\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \rangle \leq 0$. In turn, we derive from (2.4) that

$$\begin{aligned}
2\gamma_n \langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \rangle &= 2\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \rangle + 2\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n \boldsymbol{B}\boldsymbol{x}_n + \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n \rangle \\
&\leq 2\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n \boldsymbol{B}\boldsymbol{x}_n + \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n \rangle \\
&= 2\langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n \rangle \\
&= \|\boldsymbol{x}_n - \boldsymbol{x}\|^2 - \|\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}\|^2 - \|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2
\end{aligned} \tag{2.6}$$

and, therefore, using the Lipschitz continuity of $\boldsymbol{B}$, that

$$\begin{aligned}
\|\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x}\|^2 &= \|(\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}) + \gamma_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n)\|^2 \\
&= \|\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}\|^2 + 2\gamma_n \langle \widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \rangle + \gamma_n^2 \|\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n\|^2 \\
&\leq \|\boldsymbol{x}_n - \boldsymbol{x}\|^2 - \|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2 + \gamma_n^2 \|\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n\|^2 \\
&\leq \|\boldsymbol{x}_n - \boldsymbol{x}\|^2 - (1 - \gamma_n^2\beta^2)\|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2 \\
&\leq \|\boldsymbol{x}_n - \boldsymbol{x}\|^2 - \varepsilon^2 \|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2.
\end{aligned} \tag{2.7}$$

We also derive from (2.3) and (2.4) the following inequalities. First,

$$\|\widetilde{\boldsymbol{y}}_n - \boldsymbol{y}_n\| = \gamma_n \|\boldsymbol{a}_n\| \leq \|\boldsymbol{a}_n\|/\beta. \tag{2.8}$$

Hence, since $J_{\gamma_n \boldsymbol{A}}$ is nonexpansive,

$$\begin{aligned}
\|\widetilde{\boldsymbol{p}}_n - \boldsymbol{p}_n\| &= \|J_{\gamma_n \boldsymbol{A}} \widetilde{\boldsymbol{y}}_n - J_{\gamma_n \boldsymbol{A}} \boldsymbol{y}_n - \boldsymbol{b}_n\| \\
&\leq \|J_{\gamma_n \boldsymbol{A}} \widetilde{\boldsymbol{y}}_n - J_{\gamma_n \boldsymbol{A}} \boldsymbol{y}_n\| + \|\boldsymbol{b}_n\| \\
&\leq \|\widetilde{\boldsymbol{y}}_n - \boldsymbol{y}_n\| + \|\boldsymbol{b}_n\| \\
&\leq \|\boldsymbol{a}_n\|/\beta + \|\boldsymbol{b}_n\|.
\end{aligned} \tag{2.9}$$

In turn, we get

$$\begin{aligned}
\|\widetilde{\boldsymbol{q}}_n - \boldsymbol{q}_n\| &= \|\widetilde{\boldsymbol{p}}_n - \gamma_n \boldsymbol{B}\widetilde{\boldsymbol{p}}_n - \boldsymbol{p}_n + \gamma_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n)\| \\
&\leq \|\widetilde{\boldsymbol{p}}_n - \boldsymbol{p}_n\| + \gamma_n \|\boldsymbol{B}\widetilde{\boldsymbol{p}}_n - \boldsymbol{B}\boldsymbol{p}_n\| + \gamma_n \|\boldsymbol{c}_n\| \\
&\leq (1 + \gamma_n \beta)\|\widetilde{\boldsymbol{p}}_n - \boldsymbol{p}_n\| + \gamma_n \|\boldsymbol{c}_n\| \\
&\leq 2(\|\boldsymbol{a}_n\|/\beta + \|\boldsymbol{b}_n\|) + \|\boldsymbol{c}_n\|/\beta.
\end{aligned} \tag{2.10}$$

Combining (2.4), (2.8), and (2.10) yields $\|\boldsymbol{e}_n\| \leq \|\widetilde{\boldsymbol{y}}_n - \boldsymbol{y}_n\| + \|\widetilde{\boldsymbol{q}}_n - \boldsymbol{q}_n\| \leq 3\|\boldsymbol{a}_n\|/\beta + 2\|\boldsymbol{b}_n\| + \|\boldsymbol{c}_n\|/\beta$ and, in view of (2.2), it follows that

$$\sum_{k \in \mathbb{N}} \|\boldsymbol{e}_k\| < +\infty. \tag{2.11}$$

Furthermore, (2.3), (2.4), and (2.7) imply that

$$\|\boldsymbol{x}_{n+1} - \boldsymbol{x}\| = \|\boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n - \boldsymbol{x}\| \leq \|\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x}\| + \|\boldsymbol{e}_n\| \leq \|\boldsymbol{x}_n - \boldsymbol{x}\| + \|\boldsymbol{e}_n\|. \tag{2.12}$$

Thus, it follows from (2.11) and Lemma 2.1 that $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is bounded, and we deduce from (2.4) that, since the operators $\boldsymbol{B}$ and $(J_{\gamma_k \boldsymbol{A}})_{k\in\mathbb{N}}$ are Lipschitzian, $(\widetilde{\boldsymbol{y}}_k)_{k\in\mathbb{N}}$, $(\widetilde{\boldsymbol{p}}_k)_{k\in\mathbb{N}}$, and $(\widetilde{\boldsymbol{q}}_k)_{k\in\mathbb{N}}$ are bounded. Consequently, $\mu = \sup_{k\in\mathbb{N}} \|\boldsymbol{x}_k - \widetilde{\boldsymbol{y}}_k + \widetilde{\boldsymbol{q}}_k - \boldsymbol{x}\| < +\infty$ and, using (2.3), (2.4), and (2.7), we obtain

$$\begin{aligned}\|\boldsymbol{x}_{n+1} - \boldsymbol{x}\|^2 &= \|\boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n - \boldsymbol{x}\|^2 \\ &= \|\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x} + \boldsymbol{e}_n\|^2 \\ &= \|\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x}\|^2 + 2\left\langle \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x} \mid \boldsymbol{e}_n\right\rangle + \|\boldsymbol{e}_n\|^2 \\ &\leq \|\boldsymbol{x}_n - \boldsymbol{x}\|^2 - \varepsilon^2 \|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2 + \varepsilon_n, \quad \text{where} \quad \varepsilon_n = 2\mu\|\boldsymbol{e}_n\| + \|\boldsymbol{e}_n\|^2. \end{aligned} \tag{2.13}$$

(i): It follows from (2.11), (2.13), and Lemma 2.1 that

$$\sum_{n\in\mathbb{N}} \|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2 < +\infty. \tag{2.14}$$

Hence, since (2.2) and (2.9) imply that $\sum_{n\in\mathbb{N}} \|\widetilde{\boldsymbol{p}}_n - \boldsymbol{p}_n\| < +\infty$, we have $\sum_{n\in\mathbb{N}} \|\widetilde{\boldsymbol{p}}_n - \boldsymbol{p}_n\|^2 < +\infty$. We therefore infer that $\sum_{n\in\mathbb{N}} \|\boldsymbol{x}_n - \boldsymbol{p}_n\|^2 < +\infty$. Furthermore, since (2.4) yields

$$\begin{aligned}(\forall n\in\mathbb{N})\quad \|\boldsymbol{y}_n - \boldsymbol{q}_n\|^2 &= \|\widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{q}}_n + \boldsymbol{e}_n\|^2 \\ &= \|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n - \gamma_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n) + \boldsymbol{e}_n\|^2 \\ &\leq 3\big(\|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2 + \gamma_n^2\beta^2\,\|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2 + \|\boldsymbol{e}_n\|^2\big) \\ &\leq 6\|\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\|^2 + 3\|\boldsymbol{e}_n\|^2, \end{aligned} \tag{2.15}$$

we derive from (2.11) that $\sum_{n\in\mathbb{N}} \|\boldsymbol{y}_n - \boldsymbol{q}_n\|^2 < +\infty$.

(ii): It follows from (2.14), the Lipschitz continuity of $\boldsymbol{B}$, and (2.4) that

$$\boldsymbol{B}\widetilde{\boldsymbol{p}}_n - \boldsymbol{B}\boldsymbol{x}_n \to 0 \quad \text{and} \quad \boldsymbol{u}_n \to 0. \tag{2.16}$$

Now, let $\boldsymbol{w}$ be a weak sequential cluster point of $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, say $\boldsymbol{x}_{k_n} \rightharpoonup \boldsymbol{w}$. It follows from (2.5) that $(\widetilde{\boldsymbol{p}}_{k_n}, \boldsymbol{u}_{k_n})_{n\in\mathbb{N}}$ lies in $\operatorname{gra}(\boldsymbol{A}+\boldsymbol{B})$, and from (2.14) and (2.16) that

$$\widetilde{\boldsymbol{p}}_{k_n} \rightharpoonup \boldsymbol{w} \quad \text{and} \quad \boldsymbol{u}_{k_n} \to \boldsymbol{0}. \tag{2.17}$$

Since $\boldsymbol{B}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}$ is monotone and continuous, it is maximally monotone [6, Example 20.29]. Furthermore, since $\operatorname{dom}\boldsymbol{B} = \boldsymbol{\mathcal{H}}$, $\boldsymbol{A}+\boldsymbol{B}$ is maximally monotone [6, Corollary 24.4(i)] and its graph is therefore sequentially closed in $\boldsymbol{\mathcal{H}}^{\text{weak}} \times \boldsymbol{\mathcal{H}}^{\text{strong}}$ [6, Proposition 20.33(ii)]. Therefore, $(\boldsymbol{w}, \boldsymbol{0}) \in \operatorname{gra}(\boldsymbol{A}+\boldsymbol{B})$. Using (2.13), (2.11), and Lemma 2.2, we conclude that there exists $\overline{\boldsymbol{x}} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ such that $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$. Finally, in view of (i), $\boldsymbol{p}_n \rightharpoonup \overline{\boldsymbol{x}}$.

(iii)(a): As shown in (ii), $\boldsymbol{p}_n \rightharpoonup \overline{\boldsymbol{x}}$. In turn, it follows from (2.2) that $\widetilde{\boldsymbol{p}}_n = \boldsymbol{p}_n + \boldsymbol{b}_n \rightharpoonup \overline{\boldsymbol{x}}$. Moreover, (2.16) yields $\boldsymbol{u}_n \to \boldsymbol{0}$ and (2.5) yields $(\forall n\in\mathbb{N})$ $(\widetilde{\boldsymbol{p}}_n, \boldsymbol{u}_n) \in \operatorname{gra}(\boldsymbol{A}+\boldsymbol{B})$. Altogether, Definition 2.3 implies that $\widetilde{\boldsymbol{p}}_n \to \overline{\boldsymbol{x}}$ and, therefore, that $\boldsymbol{p}_n = \widetilde{\boldsymbol{p}}_n - \boldsymbol{b}_n \to \overline{\boldsymbol{x}}$. Finally, it results from (i) that $\boldsymbol{x}_n \to \overline{\boldsymbol{x}}$.

(iii)(b)⇒(iii)(a): The assumptions imply that $\boldsymbol{A}+\boldsymbol{B}$ is uniformly monotone at $\overline{\boldsymbol{x}}$. Hence, the result follows from Lemma 2.4(i).

(iii)(c): It follows from (2.13), (2.11), (ii), and [9, Proposition 3.10] that $\boldsymbol{x}_n \to \overline{\boldsymbol{x}}$. In turn, (i) yields $\boldsymbol{p}_n \to \overline{\boldsymbol{x}}$. □

Remark 2.6 The sequence $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ in (2.3) model errors in the implementation of the operators. In the error-free setting, the weak convergence of $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ to a zero of $\boldsymbol{A}+\boldsymbol{B}$ in Theorem 2.5(ii) follows from [39, Theorem 3.4(b)].

2.3 The monotone+skew model

Let us start with some elementary facts about the operators $\boldsymbol{M}$ and $\boldsymbol{S}$ appearing in Problem 1.2.

Proposition 2.7 *Consider the setting of Problem 1.1 and Problem 1.2. Then the following hold.*

(i) $\boldsymbol{M}$ *is maximally monotone.*

(ii) $\boldsymbol{S} \in \mathcal{B}(\boldsymbol{\mathcal{K}})$, $\boldsymbol{S}^* = -\boldsymbol{S}$, *and* $\|\boldsymbol{S}\| = \|L\|$.

(iii) $\boldsymbol{M} + \boldsymbol{S}$ *is maximally monotone.*

(iv) $(\forall \gamma \in \,]0, +\infty[)(\forall x \in \mathcal{H})(\forall v \in \mathcal{G})\ J_{\gamma \boldsymbol{M}}(x, v) = \big(J_{\gamma A}(x + \gamma z)\,,\, J_{\gamma B^{-1}}(v - \gamma r)\big)$.

(v) $(\forall \gamma \in \,]0, +\infty[)(\forall x \in \mathcal{H})(\forall v \in \mathcal{G})$

$$J_{\gamma \boldsymbol{S}}(x, v) = \big((\mathrm{Id} + \gamma^2 L^* L)^{-1}(x - \gamma L^* v)\,,\, (\mathrm{Id} + \gamma^2 L L^*)^{-1}(v + \gamma L x)\big).$$

Proof. (i): Since A and B are maximally monotone, it follows from [6, Propositions 20.22 and 20.23] that $A \times B^{-1}$ is likewise. In turn, $\boldsymbol{M}$ is maximally monotone.

(ii): The first two assertions are clear. Now let $(x, v) \in \boldsymbol{\mathcal{K}}$. Then $\|\boldsymbol{S}(x, v)\|^2 = \|(L^* v, -Lx)\|^2 = \|L^* v\|^2 + \|Lx\|^2 \leq \|L\|^2(\|v\|^2 + \|x\|^2) = \|L\|^2 \|(x, v)\|^2$. Thus, $\|\boldsymbol{S}\| \leq \|L\|$. Conversely, $\|x\| \leq 1 \Rightarrow \|(x, 0)\| \leq 1 \Rightarrow \|Lx\| = \|\boldsymbol{S}(x, 0)\| \leq \|\boldsymbol{S}\|$. Hence $\|L\| \leq \|\boldsymbol{S}\|$.

(iii): By (i), $\boldsymbol{M}$ is maximally monotone. On the other hand, it follows from (ii) that $\boldsymbol{S}$ is monotone and continuous, hence maximally monotone [6, Example 20.29]. Altogether, since $\operatorname{dom} \boldsymbol{S} = \boldsymbol{\mathcal{K}}$, it follows from [6, Corollary 24.4] that $\boldsymbol{M} + \boldsymbol{S}$ is maximally monotone.

(iv): This follows from [6, Proposition 23.16].

(v): Let $(x, v) \in \boldsymbol{\mathcal{K}}$ and set $(p, q) = J_{\gamma \boldsymbol{S}}(x, v)$. Then $(x, v) = (p, q) + \gamma \boldsymbol{S}(p, q)$ and hence $x = p + \gamma L^* q$ and $v = q - \gamma L p$. Hence, $Lx = Lp + \gamma L L^* q$ and $L^* v = L^* q - \gamma L^* L p$. Thus, $x = p + \gamma L^* v + \gamma^2 L^* L p$ and therefore $p = (\mathrm{Id} + \gamma^2 L^* L)^{-1}(x - \gamma L^* v)$. Likewise, $v = q - \gamma L x + \gamma^2 L L^* q$, and therefore $q = (\mathrm{Id} + \gamma^2 L L^*)^{-1}(v + \gamma L x)$. □

The next proposition makes the tight interplay between Problem 1.1 and Problem 1.2 explicit. An alternate proof of the equivalence (iii)⇔(iv)⇔(v) can be found in [28] (see also [3, 15, 26, 30] for partial results); we provide a direct argument for completeness.

Proposition 2.8 *Consider the setting of Problem 1.1 and Problem 1.2. Then*

(i) $\operatorname{zer}(\boldsymbol{M} + \boldsymbol{S})$ *is a closed convex subset of* $\mathcal{P} \times \mathcal{D}$.

Furthermore, the following are equivalent.

(ii) $z \in \operatorname{ran}(A + L^* \circ B \circ (L \cdot - r))$.

(iii) $\mathcal{P} \neq \varnothing$.

(iv) $\operatorname{zer}(\boldsymbol{M} + \boldsymbol{S}) \neq \varnothing$.

(v) $\mathcal{D} \neq \varnothing$.

(vi) $-r \in \operatorname{ran}(-L \circ A^{-1} \circ (z - L^* \cdot) + B^{-1})$.

Proof. The equivalences (ii)⇔(iii) and (v)⇔(vi) are clear. Now let $(x, v) \in \boldsymbol{\mathcal{K}}$.

(i): We derive from from (1.8) that $(x,v) \in \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}) \Leftrightarrow 0 \in -z+Ax+L^*v$ and $0 \in r+B^{-1}v-Lx$ $\Leftrightarrow$ ($z - L^*v \in Ax$ and $Lx - r \in B^{-1}v$) $\Leftrightarrow$ ($z - L^*v \in Ax$ and $v \in B(Lx - r)$) $\Rightarrow$ ($z - L^*v \in Ax$ and $L^*v \in L^*(B(Lx - r))$) $\Rightarrow z \in Ax + L^*(B(Lx - r)) \Leftrightarrow x \in \mathcal{P}$. Similarly, ($z - L^*v \in Ax$ and $Lx - r \in B^{-1}v$) $\Leftrightarrow$ ($x \in A^{-1}(z - L^*v)$ and $r - Lx \in -B^{-1}v$) $\Rightarrow$ ($Lx \in L(A^{-1}(z - L^*v))$ and $r - Lx \in -B^{-1}v$) $\Rightarrow r \in L(A^{-1}(z - L^*v)) - B^{-1}v \Leftrightarrow v \in \mathcal{D}$. Finally, since $\boldsymbol{M} + \boldsymbol{S}$ is maximally monotone by Proposition 2.7(iii), $\operatorname{zer}(\boldsymbol{M} + \boldsymbol{S})$ is closed and convex [6, Proposition 23.39].

(iii)⇒(iv): In view of (1.8), $x \in \mathcal{P} \Leftrightarrow z \in Ax+L^*(B(Lx-r)) \Leftrightarrow (\exists\, w \in \mathcal{G})\ \big(z-L^*w \in Ax$ and $w \in B(Lx-r)\big) \Leftrightarrow \big((\exists\, w \in \mathcal{G})\ z \in Ax + L^*w$ and $-r \in B^{-1}w - Lx\big) \Leftrightarrow (\exists\, w \in \mathcal{G})\ (x, w) \in \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S})$.

(iv)⇒(iii) and (iv)⇒(v): These follow from (i).

(v)⇒(iv): $v \in \mathcal{D} \Leftrightarrow r \in LA^{-1}(z - L^*v) - B^{-1}v \Leftrightarrow (\exists\, y \in \mathcal{H})$ ($y \in A^{-1}(z - L^*v)$ and $r \in Ly - B^{-1}v$) $\Leftrightarrow (\exists\, y \in \mathcal{H})$ ($0 \in -z + Ay + L^*v$ and $0 \in r + B^{-1}v - Ly$) $\Leftrightarrow (\exists\, y \in \mathcal{H})\ (y, v) \in \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S})$. □

Remark 2.9 Suppose that $z \in \operatorname{ran}(A + L^*B(L \cdot -r))$. Then Proposition 2.8 assert that solutions to (1.2) and (1.3) can be found as zeros of $\boldsymbol{M} + \boldsymbol{S}$. In principle, this can be achieved via the Douglas-Rachford algorithm applied to (1.9): let $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ be sequences in $\boldsymbol{\mathcal{K}}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0, 2[$ such that $\boldsymbol{b}_n \rightharpoonup \boldsymbol{0}$, $\sum_{n\in\mathbb{N}} \lambda_n(\|\boldsymbol{a}_n\| + \|\boldsymbol{b}_n\|) < +\infty$, and $\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$, let $\boldsymbol{y}_0 \in \boldsymbol{\mathcal{K}}$, let $\gamma \in\ ]0, +\infty[$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{x}_n = J_{\gamma \boldsymbol{S}}\, \boldsymbol{y}_n + \boldsymbol{b}_n \\ \boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \lambda_n\big(J_{\gamma \boldsymbol{M}}(2\boldsymbol{x}_n - \boldsymbol{y}_n) + \boldsymbol{a}_n - \boldsymbol{x}_n\big). \end{array} \right. \tag{2.18}$$

Then it follows from Proposition 2.7(i)–(iii) and [11, Theorem 2.1(i)(c)] that $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{zer}(\boldsymbol{M} + \boldsymbol{S})$. Now set $(\forall n \in \mathbb{N})$ $\boldsymbol{x}_n = (x_n, v_n)$, $\boldsymbol{y}_n = (y_{1,n}, y_{2,n})$, $\boldsymbol{a}_n = (a_{1,n}, a_{2,n})$, and $\boldsymbol{b}_n = (b_{1,n}, b_{2,n})$. Then, using Proposition 2.7(iv)&(v), (2.18) becomes

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} x_n = (\mathrm{Id} + \gamma^2 L^*L)^{-1}(y_{1,n} - \gamma L^* y_{2,n}) + b_{1,n} \\ v_n = (\mathrm{Id} + \gamma^2 LL^*)^{-1}(y_{2,n} + \gamma L y_{1,n}) + b_{2,n} \\ y_{1,n+1} = y_{1,n} + \lambda_n\big(J_{\gamma A}(2x_n - y_{1,n} + \gamma z) + a_{1,n} - x_n\big) \\ y_{2,n+1} = y_{2,n} + \lambda_n\big(J_{\gamma B^{-1}}(2v_n - y_{2,n} - \gamma r) + a_{2,n} - v_n\big). \end{array} \right. \tag{2.19}$$

Moreover, $(x_n)_{n\in\mathbb{N}}$ converges weakly to a solution $\overline{x}$ to (1.2) and $(v_n)_{n\in\mathbb{N}}$ to a solution $\overline{v}$ to (1.3) such that $z - L^*\overline{v} \in A\overline{x}$ and $\overline{v} \in B(L\overline{x} - r)$. However, a practical limitation of (2.19) is that it necessitates the inversion of two operators at each iteration, which may be quite demanding numerically.

Remark 2.10 It follows from (2.12) that the error-free version of the forward-backward-forward algorithm (2.3) is Fejér-monotone with respect to $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$, i.e., for every $n \in \mathbb{N}$ and every $\boldsymbol{x} \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$, $\|\boldsymbol{x}_{n+1} - \boldsymbol{x}\| \leq \|\boldsymbol{x}_n - \boldsymbol{x}\|$. Now let $n \in \mathbb{N}$. Then it follows from [5, Section 2] that there exist $\lambda_n \in [0, 2]$ and a closed affine halfspace $\boldsymbol{H}_n \subset \boldsymbol{\mathcal{H}}$ containing $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$ such that

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n(P_{\boldsymbol{H}_n}\boldsymbol{x}_n - \boldsymbol{x}_n). \tag{2.20}$$

In the setting of Problem 1.2, $\boldsymbol{H}_n$ and λ_n can be determined easily. To see this, consider Theorem 2.5 with $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{K}}$, $\boldsymbol{A} = \boldsymbol{M}$, and $\boldsymbol{B} = \boldsymbol{S}$. Let $\overline{\boldsymbol{x}} \in \operatorname{zer}(\boldsymbol{M} + \boldsymbol{S})$ and suppose that $\boldsymbol{q}_n \neq \boldsymbol{y}_n$ (otherwise, we trivially have $\boldsymbol{H}_n = \boldsymbol{\mathcal{K}}$). In view of (2.3), $\boldsymbol{y}_n - \boldsymbol{p}_n \in \gamma_n \boldsymbol{M}\boldsymbol{p}_n$ and $-\gamma_n \boldsymbol{S}\overline{\boldsymbol{x}} \in \gamma_n \boldsymbol{M}\overline{\boldsymbol{x}}$. Hence, using the monotonicity of $\gamma_n \boldsymbol{M}$ and Proposition 2.7(ii), we get $0 \le \langle \boldsymbol{p}_n - \overline{\boldsymbol{x}} \mid \boldsymbol{y}_n - \boldsymbol{p}_n + \gamma_n \boldsymbol{S}\overline{\boldsymbol{x}}\rangle = \langle \boldsymbol{p}_n \mid \boldsymbol{y}_n - \boldsymbol{p}_n\rangle - \langle \boldsymbol{x} \mid \boldsymbol{y}_n - \boldsymbol{p}_n\rangle + \gamma_n \langle \boldsymbol{S}^*\boldsymbol{p}_n \mid \overline{\boldsymbol{x}}\rangle = \langle \boldsymbol{p}_n \mid \boldsymbol{y}_n - \boldsymbol{p}_n\rangle - \langle \overline{\boldsymbol{x}} \mid \boldsymbol{y}_n - \boldsymbol{p}_n + \gamma_n \boldsymbol{S}\boldsymbol{p}_n\rangle$. Therefore, we deduce from (2.3) that $\langle \overline{\boldsymbol{x}} \mid \boldsymbol{y}_n - \boldsymbol{q}_n\rangle \le \langle \boldsymbol{p}_n \mid \boldsymbol{y}_n - \boldsymbol{p}_n\rangle = \langle \boldsymbol{p}_n \mid \boldsymbol{y}_n - \boldsymbol{q}_n\rangle$. Now set

$$\boldsymbol{H}_n = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{K}} \mid \langle \boldsymbol{x} \mid \boldsymbol{y}_n - \boldsymbol{q}_n\rangle \le \langle \boldsymbol{p}_n \mid \boldsymbol{y}_n - \boldsymbol{q}_n\rangle\big\} \quad \text{and} \quad \lambda_n = 1 + \gamma_n^2 \frac{\|\boldsymbol{S}(\boldsymbol{p}_n - \boldsymbol{x}_n)\|^2}{\|\boldsymbol{p}_n - \boldsymbol{x}_n\|^2}. \tag{2.21}$$

Then $\operatorname{zer}(\boldsymbol{M} + \boldsymbol{S}) \subset \boldsymbol{H}_n$ and $\lambda_n \le 1 + \gamma_n^2 \|\boldsymbol{S}\|^2 < 2$. Altogether, it follows from (2.3) and the skew-adjointness of $\boldsymbol{S}$ that

$$\begin{aligned}
\boldsymbol{x}_n + \lambda_n(\boldsymbol{P}_{\boldsymbol{H}_n}\boldsymbol{x}_n - \boldsymbol{x}_n) &= \boldsymbol{x}_n + \lambda_n\bigg(\frac{\langle \boldsymbol{p}_n - \boldsymbol{x}_n \mid \boldsymbol{y}_n - \boldsymbol{q}_n\rangle}{\|\boldsymbol{y}_n - \boldsymbol{q}_n\|^2}\bigg)(\boldsymbol{y}_n - \boldsymbol{q}_n)\\
&= \boldsymbol{x}_n + \lambda_n\bigg(\frac{\langle \boldsymbol{p}_n - \boldsymbol{x}_n \mid \boldsymbol{x}_n - \boldsymbol{p}_n + \gamma_n \boldsymbol{S}(\boldsymbol{p}_n - \boldsymbol{x}_n)\rangle}{\|\boldsymbol{x}_n - \boldsymbol{p}_n + \gamma_n \boldsymbol{S}(\boldsymbol{p}_n - \boldsymbol{x}_n)\|^2}\bigg)(\boldsymbol{y}_n - \boldsymbol{q}_n)\\
&= \boldsymbol{x}_n + \lambda_n\bigg(\frac{\|\boldsymbol{x}_n - \boldsymbol{p}_n\|^2}{\|\boldsymbol{x}_n - \boldsymbol{p}_n\|^2 + \gamma_n^2\|\boldsymbol{S}(\boldsymbol{p}_n - \boldsymbol{x}_n)\|^2}\bigg)(\boldsymbol{q}_n - \boldsymbol{y}_n)\\
&= \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n = \boldsymbol{x}_{n+1}.
\end{aligned} \tag{2.22}$$

Thus, the updating rule of algorithm of Theorem 2.5 applied to $\boldsymbol{M}$ and $\boldsymbol{S}$ is given by (2.20)–(2.21). In turn, using results from [5], this iteration process can easily be modified to become strongly convergent.

3 Main results

The main result of the paper can now be presented. It consists of an application of Theorem 2.5 to find solutions to Problem 1.2, and thus obtain solutions to Problem 1.1. The resulting algorithm employs the operators A, B, and L separately. Moreover, the operators A and B can be activated in parallel and all the steps involving L are explicit.

Theorem 3.1 *In Problem 1.1, suppose that $L \neq 0$ and that $z \in \operatorname{ran}\big(A + L^* \circ B \circ (L \cdot - r)\big)$. Let $(a_{1,n})_{n\in\mathbb{N}}$, $(b_{1,n})_{n\in\mathbb{N}}$, and $(c_{1,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, and let $(a_{2,n})_{n\in\mathbb{N}}$, $(b_{2,n})_{n\in\mathbb{N}}$, and $(c_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}$. Furthermore, let $x_0 \in \mathcal{H}$, let $v_0 \in \mathcal{G}$, let $\varepsilon \in \left]0, 1/(\|L\|+1)\right[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\|L\|\,]$, and set*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{aligned}
y_{1,n} &= x_n - \gamma_n(L^* v_n + a_{1,n})\\
y_{2,n} &= v_n + \gamma_n(L x_n + a_{2,n})\\
p_{1,n} &= J_{\gamma_n A}(y_{1,n} + \gamma_n z) + b_{1,n}\\
p_{2,n} &= J_{\gamma_n B^{-1}}(y_{2,n} - \gamma_n r) + b_{2,n}\\
q_{1,n} &= p_{1,n} - \gamma_n(L^* p_{2,n} + c_{1,n})\\
q_{2,n} &= p_{2,n} + \gamma_n(L p_{1,n} + c_{2,n})\\
x_{n+1} &= x_n - y_{1,n} + q_{1,n}\\
v_{n+1} &= v_n - y_{2,n} + q_{2,n}.
\end{aligned}\right. \tag{3.1}$$

Then the following hold for some solution $\overline{x}$ to (1.2) and some solution $\overline{v}$ to (1.3) such that $z - L^*\overline{v} \in A\overline{x}$ and $\overline{v} \in B(L\overline{x} - r)$.

(i) *$x_n - p_{1,n} \to 0$ and $v_n - p_{2,n} \to 0$.*

(ii) $x_n \rightharpoonup \overline{x}$, $p_{1,n} \rightharpoonup \overline{x}$, $v_n \rightharpoonup \overline{v}$, *and* $p_{2,n} \rightharpoonup \overline{v}$.

(iii) *Suppose that* A *is uniformly monotone at* $\overline{x}$. *Then* $x_n \to \overline{x}$ *and* $p_{1,n} \to \overline{x}$.

(iv) *Suppose that* B^{-1} *is uniformly monotone at* $\overline{v}$. *Then* $v_n \to \overline{v}$ *and* $p_{2,n} \to \overline{v}$.

Proof. Consider the setting of Problem 1.2. As seen in Proposition 2.7, $\boldsymbol{M}$ is maximally monotone, and $\boldsymbol{S} \in \mathcal{B}(\boldsymbol{\mathcal{K}})$ is monotone and Lipschitzian with constant $\|L\|$. Moreover, Proposition 2.8 yields

$$\varnothing \neq \operatorname{zer}(\boldsymbol{M}+\boldsymbol{S}) \subset \mathcal{P} \times \mathcal{D}. \tag{3.2}$$

Now set

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \boldsymbol{x}_n = (x_n, v_n) \\ \boldsymbol{y}_n = (y_{1,n}, y_{2,n}) \\ \boldsymbol{p}_n = (p_{1,n}, p_{2,n}) \\ \widetilde{\boldsymbol{p}}_n = (\widetilde{p}_{1,n}, \widetilde{p}_{2,n}) \\ \boldsymbol{q}_n = (q_{1,n}, q_{2,n}). \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{a}_n = (a_{1,n}, a_{2,n}) \\ \boldsymbol{b}_n = (b_{1,n}, b_{2,n}) \\ \boldsymbol{c}_n = (c_{1,n}, c_{2,n}). \end{cases} \tag{3.3}$$

Then, using (1.8) and Proposition 2.7(iv), (3.1) we can written as (2.3) in $\boldsymbol{\mathcal{K}}$. Moreover, our assumptions imply that (2.2) is satisfied. Hence, using (2.9), we obtain

$$p_{1,n} - \widetilde{p}_{1,n} \to 0 \quad \text{and} \quad p_{2,n} - \widetilde{p}_{2,n} \to 0. \tag{3.4}$$

Furthermore, we derive from (2.4) and (1.8) that

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \gamma_n^{-1}(x_n - \widetilde{p}_{1,n}) + L^*\widetilde{p}_{2,n} - L^*v_n \in A\widetilde{p}_{1,n} - z + L^*\widetilde{p}_{2,n} \\ \gamma_n^{-1}(v_n - \widetilde{p}_{2,n}) - L\widetilde{p}_{1,n} + Lx_n \in B^{-1}\widetilde{p}_{2,n} + r - L\widetilde{p}_{1,n}. \end{cases} \tag{3.5}$$

These observations allow us to establish the following.

(i)&(ii): These follows from Theorem 2.5(i)&(ii) applied to $\boldsymbol{M}$ and $\boldsymbol{S}$ in $\boldsymbol{\mathcal{K}}$.

(iii): Since $\overline{x}$ solves (1.2), there exist $u \in \mathcal{H}$ and $v \in \mathcal{G}$ such that

$$u \in A\overline{x}, \quad v \in B(L\overline{x} - r), \quad \text{and} \quad z = u + L^*v. \tag{3.6}$$

Now let $n \in \mathbb{N}$. We derive from (3.5) that

$$\gamma_n^{-1}(x_n - \widetilde{p}_{1,n}) - L^*v_n + z \in A\widetilde{p}_{1,n} \quad \text{and} \quad \gamma_n^{-1}(v_n - \widetilde{p}_{2,n}) + Lx_n - r \in B^{-1}\widetilde{p}_{2,n}, \tag{3.7}$$

which yields

$$\gamma_n^{-1}(x_n - \widetilde{p}_{1,n}) - L^*v_n + z \in A\widetilde{p}_{1,n} \quad \text{and} \quad \widetilde{p}_{2,n} \in B\big(\gamma_n^{-1}(v_n - \widetilde{p}_{2,n}) + Lx_n - r\big). \tag{3.8}$$

Now set

$$\alpha_n = \|x_n - \widetilde{p}_{1,n}\|\big(\varepsilon^{-1}\|\widetilde{p}_{1,n} - \overline{x}\| + \|L\|\,\|v_n - v\|\big) \quad \text{and} \quad \beta_n = \varepsilon^{-1}\|v_n - \widetilde{p}_{2,n}\|\,\|\widetilde{p}_{2,n} - v\|. \tag{3.9}$$

It follows from (3.6), (3.8), and the uniform monotonicity of A that there exists an increasing function $\phi\colon [0,+\infty[\,\to[0,+\infty]$ that vanishes only at 0 such that

$$\begin{aligned}
&\alpha_n+\langle x_n-\overline{x}\mid L^*v-L^*v_n\rangle\\
&\qquad\geq \varepsilon^{-1}\|\widetilde{p}_{1,n}-\overline{x}\|\,\|x_n-\widetilde{p}_{1,n}\|+\langle \widetilde{p}_{1,n}-x_n\mid L^*v-L^*v_n\rangle+\langle x_n-\overline{x}\mid L^*v-L^*v_n\rangle\\
&\qquad= \varepsilon^{-1}\|\widetilde{p}_{1,n}-\overline{x}\|\,\|x_n-\widetilde{p}_{1,n}\|+\langle \widetilde{p}_{1,n}-\overline{x}\mid L^*v-L^*v_n\rangle\\
&\qquad\geq \big\langle \widetilde{p}_{1,n}-\overline{x}\mid \gamma_n^{-1}(x_n-\widetilde{p}_{1,n})-L^*v_n+L^*v\big\rangle\\
&\qquad= \big\langle \widetilde{p}_{1,n}-\overline{x}\mid \gamma_n^{-1}(x_n-\widetilde{p}_{1,n})-L^*v_n+z-u\big\rangle\\
&\qquad\geq \phi(\|\widetilde{p}_{1,n}-\overline{x}\|).
\end{aligned}\tag{3.10}$$

On the other hand, since B is monotone, (3.9), (3.6), and (3.8) yield

$$\begin{aligned}
\beta_n+\langle x_n-\overline{x}\mid L^*\widetilde{p}_{2,n}-L^*v\rangle &\geq \big\langle \gamma_n^{-1}(v_n-\widetilde{p}_{2,n})+L(x_n-\overline{x})\mid \widetilde{p}_{2,n}-v\big\rangle\\
&= \big\langle (\gamma_n^{-1}(v_n-\widetilde{p}_{2,n})+Lx_n-r)-(L\overline{x}-r)\mid \widetilde{p}_{2,n}-v\big\rangle\\
&\geq 0.
\end{aligned}\tag{3.11}$$

Upon adding these two inequalities, we obtain

$$\alpha_n+\beta_n+\|x_n-\overline{x}\|\,\|L\|\,\|\widetilde{p}_{2,n}-v_n\|\geq \alpha_n+\beta_n+\langle x_n-\overline{x}\mid L^*(\widetilde{p}_{2,n}-v_n)\rangle\geq \phi(\|\widetilde{p}_{1,n}-\overline{x}\|).\tag{3.12}$$

Hence, since (ii), (i), and (3.4) imply that the sequences $(x_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$, $(\widetilde{p}_{1,n})_{n\in\mathbb{N}}$, and $(\widetilde{p}_{2,n})_{n\in\mathbb{N}}$ are bounded, it follows from (3.9), (3.4), and (i) that $\phi(\|\widetilde{p}_{1,n}-\overline{x}\|)\to 0$, from which we infer that $\widetilde{p}_{1,n}\to\overline{x}$ and, by (3.4), that $p_{1,n}\to\overline{x}$. In turn, (i) yields $x_n\to\overline{x}$.

(iv): Proceed as in (iii), using the dual objects. □

Remark 3.2 Using a well-known resolvent identity, the computation of $p_{2,n}$ in (3.1) can be performed in terms of the resolvent of B via the identity $J_{\gamma_nB^{-1}}y=y-\gamma_nJ_{\gamma_n^{-1}B}(\gamma_n^{-1}y)$.

Remark 3.3 Set $\boldsymbol{\mathcal{Z}}=\big\{(x,v)\in\mathcal{P}\times\mathcal{D}\mid z-L^*v\in Ax\ \text{ and }\ v\in B(Lx-r)\big\}$. Since Theorem 3.1 is an application of Theorem 2.5 in $\boldsymbol{\mathcal{K}}$, we deduce from Remark 2.10 that the updating process for (x_n,v_n) in (3.1) results from a relaxed projection onto a closed affine halfspace $\boldsymbol{H}_n$ containing $\boldsymbol{\mathcal{Z}}$, namely

$$(x_{n+1},v_{n+1})=(x_n,v_n)+\lambda_n\big(P_{\boldsymbol{H}_n}(x_n,v_n)-(x_n,v_n)\big),\tag{3.13}$$

where

$$\begin{aligned}
\boldsymbol{H}_n=\big\{(x,v)\in\boldsymbol{\mathcal{K}}\mid \langle x\mid y_{1,n}-q_{1,n}\rangle+\langle v\mid y_{2,n}-q_{2,n}\rangle\leq\langle p_{1,n}\mid y_{1,n}-q_{1,n}\rangle+\langle p_{2,n}\mid y_{2,n}-q_{2,n}\rangle\big\}\\
\text{and}\quad \lambda_n=1+\gamma_n^2\frac{\|L(p_{1,n}-x_n)\|^2+\|L^*(p_{2,n}-v_n)\|^2}{\|p_{1,n}-x_n\|^2+\|p_{2,n}-v_n\|^2}.
\end{aligned}\tag{3.14}$$

In the special case when $\mathcal{G}=\mathcal{H}$ and $L=\mathrm{Id}$, an analysis of such outer projection methods is provided in [16].

Corollary 3.4 *Let $A_1\colon\mathcal{H}\to 2^{\mathcal{H}}$ and $A_2\colon\mathcal{H}\to 2^{\mathcal{H}}$ be maximally monotone operators such that $\mathrm{zer}(A_1+A_2)\neq\varnothing$. Let $(b_{1,n})_{n\in\mathbb{N}}$ and $(b_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, let x_0 and v_0*

be in $\mathcal{H}$, let $\varepsilon \in]0, 1/2[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1-\varepsilon]$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} p_{1,n} = J_{\gamma_n A_1}(x_n - \gamma_n v_n) + b_{1,n} \\ p_{2,n} = J_{\gamma_n A_2^{-1}}(v_n + \gamma_n x_n) + b_{2,n} \\ x_{n+1} = p_{1,n} + \gamma_n(v_n - p_{2,n}) \\ v_{n+1} = p_{2,n} + \gamma_n(p_{1,n} - x_n). \end{array} \right. \tag{3.15}$$

Then the following hold for some $\overline{x} \in \mathrm{zer}(A_1 + A_2)$ and some $\overline{v} \in \mathrm{zer}(-A_1^{-1} \circ (-\mathrm{Id}) + A_2^{-1})$ such that $-\overline{v} \in A_1\overline{x}$ and $\overline{v} \in A_2\overline{x}$.

(i) *$x_n \rightharpoonup \overline{x}$ and $v_n \rightharpoonup \overline{v}$.*

(ii) *Suppose that A_1 is uniformly monotone at $\overline{x}$. Then $x_n \to \overline{x}$.*

(iii) *Suppose that A_2^{-1} is uniformly monotone at $\overline{v}$. Then $v_n \to \overline{v}$.*

Proof. Apply Theorem 3.1 with $\mathcal{G} = \mathcal{H}$, $L = \mathrm{Id}$, $A = A_1$, $B = A_2$, $r = 0$, and $z = 0$. ☐

Remark 3.5 The most popular algorithm to find a zero of the sum of two maximally monotone operators is the Douglas-Rachford algorithm [11, 14, 23, 36] (see (2.18)). Corollary 3.4 provides an alternative scheme which is also based on evaluations of the resolvents of the two operators.

Corollary 3.6 *In Problem 1.1, suppose that $L \neq 0$ and that $\mathrm{zer}(L^*BL) \neq \varnothing$. Let $(a_{1,n})_{n\in\mathbb{N}}$ and $(c_{1,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, and let $(a_{2,n})_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, and $(c_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}$. Let $x_0 \in \mathcal{H}$, let $v_0 \in \mathcal{G}$, let $\varepsilon \in]0, 1/(\|L\|+1)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\|L\|\,]$, and set*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} s_n = \gamma_n(L^* v_n + a_{1,n}) \\ y_n = v_n + \gamma_n(L x_n + a_{2,n}) \\ p_n = J_{\gamma_n B^{-1}} y_n + b_n \\ x_{n+1} = x_n - \gamma_n(L^* p_n + c_{1,n}) \\ v_{n+1} = p_n - \gamma_n(L s_n + c_{2,n}). \end{array} \right. \tag{3.16}$$

Then the following hold for some $\overline{x} \in \mathrm{zer}(L^*BL)$ and some $\overline{v} \in (\mathrm{ran}\, L)^\perp \cap B(L\overline{x})$.

(i) *$x_n \rightharpoonup \overline{x}$ and $v_n \rightharpoonup \overline{v}$.*

(ii) *Suppose that B^{-1} is uniformly monotone at $\overline{v}$. Then $v_n \to \overline{v}$.*

Proof. Apply Theorem 3.1 with $A = 0$, $r = 0$, and $z = 0$. ☐

Remark 3.7 In connection with Corollary 3.6, a weakly convergent splitting method was proposed in [29] for finding a zero of L^*BL. This method requires the additional assumption that $\mathrm{ran}\, L$ be closed. In addition, unlike the algorithm described in (3.16), it requires the exact implementation of the generalized inverse of L at each iteration, which is challenging task.

Next, we extend (1.2) to the problem of solving an inclusion involving the sum of m composite monotone operators. We obtain an algorithm in which the operators $(B_i)_{1\leq i\leq m}$ can be activated in parallel, and independently from the transformations $(L_i)_{1\leq i\leq m}$.

Theorem 3.8 *Let $z \in \mathcal{H}$ and let $(\omega_i)_{1\leq i\leq m}$ be reals in $]0,1]$ such that $\sum_{i=1}^m \omega_i = 1$. For every $i \in \{1,\ldots,m\}$, let $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone, and suppose that $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Moreover, assume that*

$$z \in \operatorname{ran} \sum_{i=1}^m \omega_i L_i^* \circ B_i \circ (L_i \cdot - r_i). \tag{3.17}$$

Consider the problem

$$\text{find} \quad x \in \mathcal{H} \ \text{ such that } \quad z \in \sum_{i=1}^m \omega_i L_i^* B_i(L_i x - r_i), \tag{3.18}$$

and the problem

$$\text{find } v_1 \in \mathcal{G}_1,\ldots,v_m \in \mathcal{G}_m \ \text{ such that } \sum_{i=1}^m \omega_i L_i^* v_i = z \ \text{ and } \ (\exists\, x \in \mathcal{H}) \begin{cases} v_1 \in B_1(L_1 x - r_1) \\ \quad\vdots \\ v_m \in B_m(L_m x - r_m). \end{cases} \tag{3.19}$$

Now, for every $i \in \{1,\ldots,m\}$, let $(a_{1,i,n})_{n\in\mathbb{N}}$ and $(c_{1,i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, let $(a_{2,i,n})_{n\in\mathbb{N}}$, $(b_{i,n})_{n\in\mathbb{N}}$, and $(c_{2,i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$, let $x_{i,0} \in \mathcal{H}$, and let $v_{i,0} \in \mathcal{G}_i$. Furthermore, set $\beta = \max_{1\leq i\leq m} \|L_i\|$, let $\varepsilon \in\,]0, 1/(\beta+1)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} x_n = \sum_{i=1}^m \omega_i x_{i,n} \\ \text{For } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} y_{1,i,n} = x_{i,n} - \gamma_n(L_i^* v_{i,n} + a_{1,i,n}) \\ y_{2,i,n} = v_{i,n} + \gamma_n(L_i x_{i,n} + a_{2,i,n}) \end{array} \right. \\ p_{1,n} = \sum_{i=1}^m \omega_i y_{1,i,n} + \gamma_n z \\ \text{For } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} p_{2,i,n} = J_{\gamma_n B_i^{-1}}(y_{2,i,n} - \gamma_n r_i) + b_{i,n} \\ q_{1,i,n} = p_{1,n} - \gamma_n(L_i^* p_{2,i,n} + c_{1,i,n}) \\ q_{2,i,n} = p_{2,i,n} + \gamma_n(L_i p_{1,n} + c_{2,i,n}) \\ x_{i,n+1} = x_{i,n} - y_{1,i,n} + q_{1,i,n} \\ v_{i,n+1} = v_{i,n} - y_{2,i,n} + q_{2,i,n}. \end{array} \right. \end{array} \right. \tag{3.20}$$

Then the following hold for some solution $\overline{x}$ to (3.18) and some solution $(\overline{v}_i)_{1\leq i\leq m}$ to (3.19) such that, for every $i \in \{1,\ldots,m\}$, $\overline{v}_i \in B_i(L_i\overline{x} - r_i)$.

(i) *$x_n \rightharpoonup \overline{x}$ and, for every $i \in \{1,\ldots,m\}$, $v_{i,n} \rightharpoonup \overline{v}_i$.*

(ii) *Suppose that, for every $i \in \{1,\ldots,m\}$, B_i^{-1} is strongly monotone at $\overline{v}_i$. Then, for every $i \in \{1,\ldots,m\}$, $v_{i,n} \to \overline{v}_i$.*

Proof. Let $\boldsymbol{\mathcal{H}}$ be the real Hilbert space obtained by endowing the Cartesian product $\mathcal{H}^m$ with the scalar product $\langle \cdot \mid \cdot \rangle_{\boldsymbol{\mathcal{H}}} \colon (\boldsymbol{x},\boldsymbol{y}) \mapsto \sum_{i=1}^m \omega_i \langle x_i \mid y_i \rangle$, where $\boldsymbol{x} = (x_i)_{1\leq i\leq m}$ and $\boldsymbol{y} = (y_i)_{1\leq i\leq m}$ denote generic elements in $\boldsymbol{\mathcal{H}}$. The associated norm is $\|\cdot\|_{\boldsymbol{\mathcal{H}}}\colon \boldsymbol{x} \mapsto \sqrt{\sum_{i=1}^m \omega_i \|x_i\|^2}$. Likewise, let $\boldsymbol{\mathcal{G}}$ denote the real Hilbert space obtained by endowing $\mathcal{G}_1 \times \cdots \times \mathcal{G}_m$ with the scalar product and the associated norm respectively defined by

$$\langle \cdot \mid \cdot \rangle_{\boldsymbol{\mathcal{G}}} : (\boldsymbol{y}, \boldsymbol{z}) \mapsto \sum_{i=1}^m \omega_i \langle y_i \mid z_i \rangle_{\mathcal{G}_i} \quad \text{and} \quad \| \cdot \|_{\boldsymbol{\mathcal{G}}} : \boldsymbol{y} \mapsto \sqrt{\sum_{i=1}^m \omega_i \|y_i\|^2_{\mathcal{G}_i}}. \tag{3.21}$$

Define

$$\boldsymbol{V} = \big\{(x, \ldots, x) \in \boldsymbol{\mathcal{H}} \mid x \in \mathcal{H}\big\} \quad \text{and} \quad \boldsymbol{j} \colon \mathcal{H} \to \boldsymbol{V} \colon x \mapsto (x, \ldots, x). \tag{3.22}$$

In view of (1.11), the normal cone operator of $\boldsymbol{V}$ is

$$N_{\boldsymbol{V}} \colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}} \colon \boldsymbol{x} \mapsto \begin{cases} \boldsymbol{V}^\perp = \big\{\boldsymbol{u} \in \boldsymbol{\mathcal{H}} \mid \sum_{i=1}^m \omega_i u_i = 0\big\}, & \text{if } \boldsymbol{x} \in \boldsymbol{V}; \\ \varnothing, & \text{otherwise.} \end{cases} \tag{3.23}$$

Now set

$$\boldsymbol{A} = N_{\boldsymbol{V}}, \ \boldsymbol{B} \colon \boldsymbol{\mathcal{G}} \to 2^{\boldsymbol{\mathcal{G}}} \colon \boldsymbol{y} \mapsto \bigtimes_{i=1}^m B_i y_i, \ \boldsymbol{L} \colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{G}} \colon \boldsymbol{x} \mapsto (L_i x_i)_{1 \le i \le m}, \text{ and } \boldsymbol{r} = (r_i)_{1 \le i \le m}. \tag{3.24}$$

It is easily checked that $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximally monotone with resolvents

$$(\forall \gamma \in \,]0, +\infty[) \quad J_{\gamma \boldsymbol{A}} \colon \boldsymbol{x} \mapsto P_{\boldsymbol{V}} \boldsymbol{x} = \boldsymbol{j}\bigg(\sum_{i=1}^m \omega_i x_i\bigg) \quad \text{and} \quad J_{\gamma \boldsymbol{B}^{-1}} \colon \boldsymbol{y} \mapsto \big(J_{\gamma B_i^{-1}} y_i\big)_{1 \le i \le m}. \tag{3.25}$$

Moreover, $\boldsymbol{L} \in \mathcal{B}(\boldsymbol{\mathcal{H}}, \boldsymbol{\mathcal{G}})$ and

$$\boldsymbol{L}^* \colon \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{H}} \colon \boldsymbol{v} \mapsto \big(L_i^* v_i\big)_{1 \le i \le m}. \tag{3.26}$$

Now, set

$$\begin{cases} \boldsymbol{\mathcal{P}} = \big\{\boldsymbol{x} \in \boldsymbol{\mathcal{H}} \mid \boldsymbol{j}(z) \in \boldsymbol{A}\boldsymbol{x} + \boldsymbol{L}^* \boldsymbol{B}(\boldsymbol{L}\boldsymbol{x} - \boldsymbol{r})\big\} \\ \boldsymbol{\mathcal{D}} = \big\{\boldsymbol{v} \in \boldsymbol{\mathcal{G}} \mid -\boldsymbol{r} \in -\boldsymbol{L}\boldsymbol{A}^{-1}(\boldsymbol{j}(z) - \boldsymbol{L}^* \boldsymbol{v}) + \boldsymbol{B}^{-1} \boldsymbol{v}\big\}. \end{cases} \tag{3.27}$$

Then, for every $x \in \mathcal{H}$,

$$\begin{aligned} x \text{ solves (3.18)} \quad &\Leftrightarrow \quad z \in \sum_{i=1}^m \omega_i L_i^* \big(B_i(L_i x - r_i)\big) \\ &\Leftrightarrow \quad \bigg(\exists\, (v_i)_{1 \le i \le m} \in \bigtimes_{i=1}^m B_i(L_i x - r_i)\bigg) \quad z = \sum_{i=1}^m \omega_i L_i^* v_i \\ &\Leftrightarrow \quad \bigg(\exists\, (v_i)_{1 \le i \le m} \in \bigtimes_{i=1}^m B_i(L_i x - r_i)\bigg) \quad \sum_{i=1}^m \omega_i (z - L_i^* v_i) = 0 \\ &\Leftrightarrow \quad \big(\exists\, \boldsymbol{v} \in \boldsymbol{B}(\boldsymbol{L}\boldsymbol{j}(x) - \boldsymbol{r})\big) \quad \boldsymbol{j}(z) - \boldsymbol{L}^* \boldsymbol{v} \in \boldsymbol{V}^\perp = \boldsymbol{A}\boldsymbol{j}(x) \\ &\Leftrightarrow \quad \boldsymbol{j}(z) \in \boldsymbol{A}\boldsymbol{j}(x) + \boldsymbol{L}^* \boldsymbol{B}(\boldsymbol{L}\boldsymbol{j}(x) - \boldsymbol{r}) \\ &\Leftrightarrow \quad \boldsymbol{j}(x) \in \boldsymbol{\mathcal{P}} \subset \boldsymbol{V}. \end{aligned} \tag{3.28}$$

Moreover, for every $\boldsymbol{v} \in \boldsymbol{\mathcal{G}}$,

$$\begin{aligned}
\boldsymbol{v} \text{ solves (3.19)} \quad &\Leftrightarrow \quad \sum_{i=1}^m \omega_i(z - L_i^* v_i) = 0 \quad \text{and} \quad (\exists\, x \in \mathcal{H}) \quad (v_i)_{1\leq i\leq m} \in \bigtimes_{i=1}^m B_i(L_i x - r_i) \\
&\Leftrightarrow \quad (\exists\, x \in \mathcal{H}) \quad \boldsymbol{j}(z) - \boldsymbol{L}^*\boldsymbol{v} \in \boldsymbol{V}^\perp = \boldsymbol{A}\boldsymbol{j}(x) \quad \text{and} \quad \boldsymbol{v} \in \boldsymbol{B}(\boldsymbol{L}\boldsymbol{j}(x) - \boldsymbol{r}) \\
&\Leftrightarrow \quad (\exists\, x \in \mathcal{H}) \quad \boldsymbol{j}(x) \in \boldsymbol{A}^{-1}\big(\boldsymbol{j}(z) - \boldsymbol{L}^*\boldsymbol{v}\big) \quad \text{and} \quad \boldsymbol{L}\boldsymbol{j}(x) - \boldsymbol{r} \in \boldsymbol{B}^{-1}\boldsymbol{v} \\
&\Leftrightarrow \quad (\exists\, \boldsymbol{x} \in \boldsymbol{V} = \operatorname{dom} \boldsymbol{A}) \quad \boldsymbol{x} \in \boldsymbol{A}^{-1}\big(\boldsymbol{j}(z) - \boldsymbol{L}^*\boldsymbol{v}\big) \quad \text{and} \quad \boldsymbol{L}\boldsymbol{x} - \boldsymbol{r} \in \boldsymbol{B}^{-1}\boldsymbol{v} \\
&\Leftrightarrow \quad -\boldsymbol{r} \in -\boldsymbol{L}\boldsymbol{A}^{-1}\big(\boldsymbol{j}(z) - \boldsymbol{L}^*\boldsymbol{v}\big) + \boldsymbol{B}^{-1}\boldsymbol{v} \\
&\Leftrightarrow \quad \boldsymbol{v} \in \boldsymbol{\mathcal{D}}.
\end{aligned} \tag{3.29}$$

Altogether, solving the inclusion (3.18) in $\mathcal{H}$ is equivalent to solving the inclusion $\boldsymbol{j}(z) \in \boldsymbol{A}\boldsymbol{x} + \boldsymbol{L}^*\boldsymbol{B}(\boldsymbol{L}\boldsymbol{x} - \boldsymbol{r})$ in $\boldsymbol{\mathcal{H}}$ and solving (3.19) in $\boldsymbol{\mathcal{G}}$ is equivalent to solving $-\boldsymbol{r} \in \boldsymbol{B}^{-1}\boldsymbol{v} - \boldsymbol{L}\boldsymbol{A}^{-1}\big(\boldsymbol{j}(z) - \boldsymbol{L}^*\boldsymbol{v}\big)$ in $\boldsymbol{\mathcal{G}}$. Next, let us show that the algorithm described in (3.20) is a particular case of the algorithm described in (3.1) in Theorem 3.1. To this end define, for every $n \in \mathbb{N}$, $\boldsymbol{x}_n = (x_{i,n})_{1\leq i\leq m}$, $\boldsymbol{v}_n = (v_{i,n})_{1\leq i\leq m}$, $\boldsymbol{y}_{1,n} = (y_{1,i,n})_{1\leq i\leq m}$, $\boldsymbol{y}_{2,n} = (y_{2,i,n})_{1\leq i\leq m}$, $\boldsymbol{p}_{1,n} = \boldsymbol{j}(p_{1,n})$, $\boldsymbol{p}_{2,n} = (p_{2,i,n})_{1\leq i\leq m}$, $\boldsymbol{q}_{1,n} = (q_{1,i,n})_{1\leq i\leq m}$, $\boldsymbol{q}_{2,n} = (q_{2,i,n})_{1\leq i\leq m}$, $\boldsymbol{a}_{1,n} = (a_{1,i,n})_{1\leq i\leq m}$, $\boldsymbol{a}_{2,n} = (a_{2,i,n})_{1\leq i\leq m}$, $\boldsymbol{b}_{2,n} = (b_{i,n})_{1\leq i\leq m}$, $\boldsymbol{c}_{1,n} = (c_{1,i,n})_{1\leq i\leq m}$, and $\boldsymbol{c}_{2,n} = (c_{2,i,n})_{1\leq i\leq m}$. Then we deduce from (3.24), (3.25), and (3.26) that, in terms of these new variables, (3.20) can be rewritten as

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l}
\boldsymbol{y}_{1,n} = \boldsymbol{x}_n - \gamma_n(\boldsymbol{L}^*\boldsymbol{v}_n + \boldsymbol{a}_{1,n}) \\
\boldsymbol{y}_{2,n} = \boldsymbol{v}_n + \gamma_n(\boldsymbol{L}\boldsymbol{x}_n + \boldsymbol{a}_{2,n}) \\
\boldsymbol{p}_{1,n} = J_{\gamma_n \boldsymbol{A}}(\boldsymbol{y}_{1,n} + \gamma_n z) \\
\boldsymbol{p}_{2,n} = J_{\gamma_n \boldsymbol{B}^{-1}}(\boldsymbol{y}_{2,n} - \gamma_n r) + \boldsymbol{b}_{2,n} \\
\boldsymbol{q}_{1,n} = \boldsymbol{p}_{1,n} - \gamma_n(\boldsymbol{L}^*\boldsymbol{p}_{2,n} + \boldsymbol{c}_{1,n}) \\
\boldsymbol{q}_{2,n} = \boldsymbol{p}_{2,n} + \gamma_n(\boldsymbol{L}\boldsymbol{p}_{1,n} + \boldsymbol{c}_{2,n}) \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_{1,n} + \boldsymbol{q}_{1,n} \\
\boldsymbol{v}_{n+1} = \boldsymbol{v}_n - \boldsymbol{y}_{2,n} + \boldsymbol{q}_{2,n}.
\end{array}\right. \tag{3.30}$$

Moreover, $\|\boldsymbol{L}\| \leq \max_{1\leq i\leq m}\|L_i\| = \beta$, and our assumptions imply that the sequences $(\boldsymbol{a}_{1,n})_{n\in\mathbb{N}}$, $(\boldsymbol{c}_{1,n})_{n\in\mathbb{N}}$, $(\boldsymbol{a}_{2,n})_{n\in\mathbb{N}}$, $(\boldsymbol{b}_{2,n})_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_{2,n})_{n\in\mathbb{N}}$ are absolutely summable. Furthermore, (3.17) and (3.28) assert that $\boldsymbol{j}(z) \in \operatorname{ran}(\boldsymbol{A} + \boldsymbol{L}^* \circ \boldsymbol{B} \circ (\boldsymbol{L}\cdot - \boldsymbol{r}))$.

(i): It follows from Theorem 3.1(ii) that there exists $\overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{P}}$ and $(\overline{v}_i)_{1\leq i\leq m} = \overline{\boldsymbol{v}} \in \boldsymbol{\mathcal{D}}$ such that $\boldsymbol{j}(z) - \boldsymbol{L}^*\overline{\boldsymbol{v}} \in \boldsymbol{A}\overline{\boldsymbol{x}}$, $\overline{\boldsymbol{v}} \in \boldsymbol{B}(\boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r})$, $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$, and $\boldsymbol{v}_n \rightharpoonup \overline{\boldsymbol{v}}$. Hence, $\boldsymbol{j}(x_n) = P_{\boldsymbol{V}}\boldsymbol{x}_n \rightharpoonup P_{\boldsymbol{V}}\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}}$. Since (3.28) asserts that there exists a solution $\overline{x}$ to (3.18) such that $\overline{\boldsymbol{x}} = \boldsymbol{j}(\overline{x})$, we obtain that $x_n = \boldsymbol{j}^{-1}(P_{\boldsymbol{V}}\boldsymbol{x}_n) \rightharpoonup \boldsymbol{j}^{-1}(\overline{\boldsymbol{x}}) = \overline{x}$. Altogether, by (3.29), for every $i \in \{1, \ldots, m\}$, $v_{i,n} \rightharpoonup \overline{v}_i$, where $(\overline{v}_i)_{1\leq i\leq m}$ solves (3.19).

(ii): Let $(\boldsymbol{w}_1, \boldsymbol{y}_1)$ and $(\boldsymbol{w}_2, \boldsymbol{y}_2)$ in $\operatorname{gra} \boldsymbol{B}^{-1}$. We derive from (3.24) that $(\forall i \in \{1, \ldots, m\})$ $y_{1,i} \in B_i^{-1}w_{1,i}$ and $y_{2,i} \in B_i^{-1}w_{2,i}$. Hence, since the operators $(B_i^{-1})_{1\leq i\leq m}$ are strongly monotone, there exist constants $(\rho_i)_{1\leq i\leq m}$ in $]0, +\infty[$ such that $\langle \boldsymbol{y}_1 - \boldsymbol{y}_2 \mid \boldsymbol{w}_1 - \boldsymbol{w}_2\rangle_{\boldsymbol{\mathcal{G}}} = \sum_{i=1}^m \omega_i \langle y_{1,i} - y_{2,i} \mid w_{1,i} - w_{2,i}\rangle_{\mathcal{G}_i} \geq \sum_{i=1}^m \omega_i\rho_i\|w_{1,i} - w_{2,i}\|^2_{\mathcal{G}_i} \geq \rho\|\boldsymbol{w}_1 - \boldsymbol{w}_2\|^2_{\boldsymbol{\mathcal{G}}}$, where $\rho = \min_{1\leq i\leq m}\rho_i \in \,]0, +\infty[$. Therefore, $\boldsymbol{B}^{-1}$ is strongly monotone and hence uniformly monotone. Thus, the result follows from Theorem 3.1(iv). ☐

4 Variational problems

We apply the results of the previous sections to minimization problems. Let us first recall some standard notation and results [6, 40]. We denote by $\Gamma_0(\mathcal{H})$ the class of lower semicontinuous convex functions $f\colon \mathcal{H} \to\,]-\infty,+\infty]$ such that $\operatorname{dom} f = \big\{x \in \mathcal{H} \mid f(x) < +\infty\big\} \neq \varnothing$. Now let $f \in \Gamma_0(\mathcal{H})$. The conjugate of f is the function $f^* \in \Gamma_0(\mathcal{H})$ defined by $f^*\colon u \mapsto \sup_{x\in\mathcal{H}}(\langle x \mid u\rangle - f(x))$. Moreover, for every $x \in \mathcal{H}$, $f + \|x - \cdot\|^2/2$ possesses a unique minimizer, which is denoted by $\operatorname{prox}_f x$. Alternatively,

$$\operatorname{prox}_f = (\mathrm{Id} + \partial f)^{-1} = J_{\partial f}, \tag{4.1}$$

where $\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \ \langle y - x \mid u\rangle + f(x) \leq f(y)\big\}$ is the subdifferential of f, which is a maximally monotone operator. Finally, let C be a convex subset of $\mathcal{H}$. The indicator function of C is denoted by ι_C, its support function by σ_C, and its strong relative interior (the set of points in $x \in C$ such that the cone generated by $-x + C$ is a closed vector subspace of $\mathcal{H}$) by $\operatorname{sri} C$. The following facts will also be required.

Proposition 4.1 *Let $f \in \Gamma_0(\mathcal{H})$, let $g \in \Gamma_0(\mathcal{G})$, let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$, let $z \in \mathcal{H}$, and let $r \in \mathcal{G}$. Then the following hold.*

(i) $\operatorname{zer}(-z + \partial f + L^* \circ (\partial g) \circ (L \cdot - r)) \subset \operatorname{Argmin}(f - \langle \cdot \mid z\rangle + g \circ (L \cdot - r))$.

(ii) $\operatorname{zer}(r - (L \circ (\partial f^*) \circ (z - L^* \cdot)) + \partial g^*) \subset \operatorname{Argmin}(f^*(z - L^*\cdot) + g^* + \langle r \mid \cdot\rangle)$.

(iii) *Suppose that one of the following is satisfied.*

- (a) $\operatorname{Argmin}(f + g \circ (L \cdot - r) - \langle \cdot \mid z\rangle) \neq \varnothing$ *and* $r \in \operatorname{sri}(L(\operatorname{dom} f) - \operatorname{dom} g)$.
- (b) $\operatorname{Argmin}(f + g \circ (L \cdot - r) - \langle \cdot \mid z\rangle) \subset \operatorname{Argmin}(f - \langle \cdot \mid z\rangle) \cap \operatorname{Argmin} g \circ (L \cdot - r) \neq \varnothing$ *and* $r \in \operatorname{sri}(\operatorname{ran} L - \operatorname{dom} g)$.
- (c) $f = \iota_C$ *and* $g = \iota_D$, $z = 0$, *where C and D are closed convex subset of $\mathcal{H}$ and $\mathcal{G}$, respectively, such that* $C \cap L^{-1}(r + D) \neq \varnothing$ *and* $r \in \operatorname{sri}(\operatorname{ran} L - D)$.

Then $z \in \operatorname{ran}(\partial f + L^* \circ (\partial g) \circ (L \cdot - r))$.

Proof. (i)&(ii): By [6, Proposition 16.5(ii) and Theorem 16.2], $\operatorname{zer}(-z + \partial f + L^* \circ (\partial g) \circ (L \cdot - r)) \subset \operatorname{zer}(\partial(f - \langle \cdot \mid z\rangle + g \circ (L \cdot - r))) = \operatorname{Argmin}(f - \langle \cdot \mid z\rangle + g \circ (L \cdot - r))$. We obtain (ii) similarly.

(iii)(a): By [6, Theorem 16.2 and Theorem 16.37(i)], we have

$$\begin{aligned}\varnothing \neq \operatorname{Argmin}(f + g \circ (L \cdot - r) - \langle \cdot \mid z\rangle) &= \operatorname{zer} \partial(f + g \circ (L \cdot - r) - \langle \cdot \mid z\rangle)\\ &= \operatorname{zer}(-z + \partial f + L^* \circ (\partial g) \circ (L \cdot - r)).\end{aligned} \tag{4.2}$$

(iii)(b): Since $r \in \operatorname{sri}(\operatorname{ran} L - \operatorname{dom} g)$, using (i) and standard convex analysis, we obtain

$$\begin{aligned}\operatorname{Argmin}(f - \langle \cdot \mid z\rangle) \cap \operatorname{Argmin}(g \circ (L \cdot - r)) &= \operatorname{zer}(-z + \partial f) \cap \operatorname{zer} \partial(g \circ (L \cdot - r))\\ &= \operatorname{zer}(-z + \partial f) \cap \operatorname{zer}(L^* \circ (\partial g) \circ (L \cdot - r))\\ &\subset \operatorname{zer}(-z + \partial f + L^* \circ (\partial g) \circ (L \cdot - r))\\ &\subset \operatorname{Argmin}(f + g \circ (L \cdot - r) - \langle \cdot \mid z\rangle).\end{aligned} \tag{4.3}$$

Therefore, the hypotheses yield $\operatorname{zer}(-z+\partial f+L^*\circ(\partial g)\circ(L\cdot -r)) = \operatorname{Argmin}(f-\langle\cdot\mid z\rangle)\cap\operatorname{Argmin}(g\circ(L\cdot -r))\neq\varnothing$.

(iii)(c): Since $\operatorname{dom}(\iota_C+\iota_D(L\cdot -r)) = C\cap L^{-1}(r+D)$,

$$\begin{aligned}\operatorname{Argmin}(\iota_C+\iota_D\circ(L\cdot -r)) &= \operatorname{Argmin}\iota_{C\cap L^{-1}(r+D)}\\ &= C\cap L^{-1}(r+D)\\ &= \operatorname{Argmin}\iota_C\cap\operatorname{Argmin}(\iota_D\circ(L\cdot -r))\neq\varnothing.\end{aligned}\tag{4.4}$$

In view of (ii) applied to $f=\iota_C$, $g=\iota_D$, and $z=0$, the proof is complete. ◻

Our first result is a new splitting method for the Fenchel-Rockafellar duality framework (1.4)–(1.5).

Proposition 4.2 *Let $f\in\Gamma_0(\mathcal{H})$, let $g\in\Gamma_0(\mathcal{G})$, let $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$, let $z\in\mathcal{H}$, and let $r\in\mathcal{G}$. Suppose that $L\neq 0$ and that*

$$z\in\operatorname{ran}\big(\partial f+L^*\circ(\partial g)\circ(L\cdot -r)\big).\tag{4.5}$$

Consider the primal problem

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; f(x)+g(Lx-r)-\langle x\mid z\rangle,\tag{4.6}$$

and the dual problem

$$\underset{v\in\mathcal{G}}{\text{minimize}}\;\; f^*(z-L^*v)+g^*(v)+\langle v\mid r\rangle.\tag{4.7}$$

Let $(a_{1,n})_{n\in\mathbb{N}}$, $(b_{1,n})_{n\in\mathbb{N}}$, and $(c_{1,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, and let $(a_{2,n})_{n\in\mathbb{N}}$, $(b_{2,n})_{n\in\mathbb{N}}$, and $(c_{2,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}$. Furthermore, let $x_0\in\mathcal{H}$, let $v_0\in\mathcal{G}$, let $\varepsilon\in\,]0,1/(\|L\|+1)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,(1-\varepsilon)/\|L\|\,]$, and set

$$(\forall n\in\mathbb{N})\quad\left\lfloor\begin{array}{l} y_{1,n}=x_n-\gamma_n(L^*v_n+a_{1,n})\\ y_{2,n}=v_n+\gamma_n(Lx_n+a_{2,n})\\ p_{1,n}=\operatorname{prox}_{\gamma_n f}(y_{1,n}+\gamma_n z)+b_{1,n}\\ p_{2,n}=\operatorname{prox}_{\gamma_n g^*}(y_{2,n}-\gamma_n r)+b_{2,n}\\ q_{1,n}=p_{1,n}-\gamma_n(L^*p_{2,n}+c_{1,n})\\ q_{2,n}=p_{2,n}+\gamma_n(Lp_{1,n}+c_{2,n})\\ x_{n+1}=x_n-y_{1,n}+q_{1,n}\\ v_{n+1}=v_n-y_{2,n}+q_{2,n}.\end{array}\right.\tag{4.8}$$

Then the following hold for some solution $\overline{x}$ to (4.6) and some solution $\overline{v}$ to (4.7) such that $z-L^*\overline{v}\in\partial f(\overline{x})$ and $\overline{v}\in\partial g(L\overline{x}-r)$.

(i) *$x_n-p_{1,n}\to 0$ and $v_n-p_{2,n}\to 0$.*

(ii) *$x_n\rightharpoonup\overline{x}$, $p_{1,n}\rightharpoonup\overline{x}$, $v_n\rightharpoonup\overline{v}$, and $p_{2,n}\rightharpoonup\overline{v}$.*

(iii) *Suppose that f is uniformly convex at $\overline{x}$. Then $x_n\to\overline{x}$ and $p_{1,n}\to\overline{x}$.*

(iv) *Suppose that g^* is uniformly convex at $\overline{v}$. Then $v_n\to\overline{v}$ and $p_{2,n}\to\overline{v}$.*

Proof. Suppose that $A=\partial f$ and $B=\partial g$ in Problem 1.1. Then, since $A^{-1}=\partial f^*$ and $B^{-1}=\partial g^*$, we derive from Proposition 4.1(i)&(ii) that the solutions to (1.2) and (1.3) are solutions to (4.6) and (4.7), respectively. Moreover, (4.1) implies that (4.8) is a special case of (3.1). Finally, the uniform convexity of a function $\varphi\in\Gamma_0(\mathcal{H})$ at a point of the domain of $\partial\varphi$ implies the uniform monotonicity of $\partial\varphi$ at that point [40, Section 3.4]. Altogether, the results follow from Theorem 3.1. ◻

Remark 4.3 Here are some comments on Proposition 4.2.

(i) Sufficient conditions for (4.5) to hold are provided in Proposition 4.1.

(ii) As in Remark 3.2, if the proximity operator of g is simpler to implement than that of g^*, $p_{2,n}$ in (4.8) can be computed via the identity $\operatorname{prox}_{\gamma_n g^*} y = y - \gamma_n \operatorname{prox}_{\gamma_n^{-1} g}(\gamma_n^{-1} y)$.

(iii) In the special case when $\mathcal{H}$ and $\mathcal{G}$ are Euclidean spaces, an alternative primal-dual algorithm is proposed in [8], which also uses the proximity operators of f and g, and the operator L in separate steps. This method is derived there in the spirit of the proximal [34] and alternating direction (see [20] and the references therein) methods of multipliers.

We now turn our attention to problems involving the sum of m composite functions.

Proposition 4.4 *Let $z \in \mathcal{H}$ and let $(\omega_i)_{1\leq i\leq m}$ be reals in $]0,1]$ such that $\sum_{i=1}^m \omega_i = 1$. For every $i \in \{1,\ldots,m\}$, let $(\mathcal{G}_i, \|\cdot\|_{\mathcal{G}_i})$ be a real Hilbert space, let $r_i \in \mathcal{G}_i$, let $g_i \in \Gamma_0(\mathcal{G}_i)$, and suppose that $0 \neq L_i \in \mathcal{B}(\mathcal{H},\mathcal{G}_i)$. Moreover, assume that*

$$z \in \operatorname{ran} \sum_{i=1}^m \omega_i L_i^* \circ (\partial g_i) \circ (L_i \cdot - r_i). \tag{4.9}$$

Consider the problem

$$\underset{x\in\mathcal{H}}{\text{minimize}} \quad \sum_{i=1}^m \omega_i\, g_i(L_i x - r_i) - \langle x \mid z\rangle, \tag{4.10}$$

and the problem

$$\underset{\substack{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m\\ \sum_{i=1}^m \omega_i L_i^* v_i = z}}{\text{minimize}} \quad \sum_{i=1}^m \omega_i \big(g_i^*(v_i) + \langle v_i \mid r_i\rangle\big). \tag{4.11}$$

For every $i \in \{1,\ldots,m\}$, let $(a_{1,i,n})_{n\in\mathbb{N}}$ and $(c_{1,i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{H}$, let $(a_{2,i,n})_{n\in\mathbb{N}}$, $(b_{i,n})_{n\in\mathbb{N}}$, and $(c_{2,i,n})_{n\in\mathbb{N}}$ be absolutely summable sequences in $\mathcal{G}_i$, let $x_{i,0} \in \mathcal{H}$, and let $v_{i,0} \in \mathcal{G}_i$. Furthermore, set $\beta = \max_{1\leq i\leq m} \|L_i\|$, let $\varepsilon \in]0, 1/(\beta+1)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$, and set

$$(\forall n\in\mathbb{N}) \quad \left\lfloor \begin{array}{l} x_n = \sum_{i=1}^m \omega_i x_{i,n} \\ \text{For } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} y_{1,i,n} = x_{i,n} - \gamma_n(L_i^* v_{i,n} + a_{1,i,n}) \\ y_{2,i,n} = v_{i,n} + \gamma_n(L_i x_{i,n} + a_{2,i,n}) \end{array}\right. \\ p_{1,n} = \sum_{i=1}^m \omega_i y_{1,i,n} + \gamma_n z \\ \text{For } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} p_{2,i,n} = \operatorname{prox}_{\gamma_n g_i^*}(y_{2,i,n} - \gamma_n r_i) + b_{i,n} \\ q_{1,i,n} = p_{1,n} - \gamma_n(L_i^* p_{2,i,n} + c_{1,i,n}) \\ q_{2,i,n} = p_{2,i,n} + \gamma_n(L_i p_{1,n} + c_{2,i,n}) \\ x_{i,n+1} = x_{i,n} - y_{1,i,n} + q_{1,i,n} \\ v_{i,n+1} = v_{i,n} - y_{2,i,n} + q_{2,i,n}. \end{array}\right. \end{array}\right. \tag{4.12}$$

Then the following hold for some solution $\overline{x}$ to (4.10) and some solution $(\overline{v}_i)_{1\leq i\leq m}$ to (4.11) such that, for every $i \in \{1,\ldots,m\}$, $\overline{v}_i \in \partial g_i(L_i\overline{x} - r_i)$.

(i) $x_n \rightharpoonup \overline{x}$ *and, for every* $i \in \{1,\ldots,m\}$, $v_{i,n} \rightharpoonup \overline{v}_i$.

(ii) *Suppose that, for every* $i \in \{1,\ldots,m\}$, g_i^* *is strongly convex at* $\overline{v}_i$. *Then, for every* $i \in \{1,\ldots,m\}$, $v_{i,n} \to \overline{v}_i$.

Proof. Define $\boldsymbol{\mathcal{H}}$, $\boldsymbol{\mathcal{G}}$, $\boldsymbol{L}$, $\boldsymbol{V}$, $\boldsymbol{r}$, and $\boldsymbol{j}\colon \mathcal{H} \to \boldsymbol{V}$ as in the proof of Theorem 3.8. Moreover, set $\boldsymbol{f} = \iota_{\boldsymbol{V}}$ and $\boldsymbol{g}\colon \boldsymbol{\mathcal{G}} \to \left]-\infty,+\infty\right] : \boldsymbol{y} \mapsto \sum_{i=1}^m \omega_i\, g_i(y_i)$. Then, $\boldsymbol{f} \in \Gamma_0(\boldsymbol{\mathcal{H}})$, $\boldsymbol{g} \in \Gamma_0(\boldsymbol{\mathcal{G}})$, $\boldsymbol{f}^* = \iota_{\boldsymbol{V}^\perp}$, and $\boldsymbol{g}^*\colon \boldsymbol{v} \mapsto \sum_{i=1}^m \omega_i\, g_i^*(v_i)$. Therefore, (4.9) is equivalent to

$$\boldsymbol{j}(z) \in \operatorname{ran}\big(\partial \boldsymbol{f} + \boldsymbol{L}^* \circ (\partial \boldsymbol{g}) \circ (\boldsymbol{L}\cdot - \boldsymbol{r})\big). \tag{4.13}$$

Furthermore, (4.10) and (4.11) are equivalent to

$$\underset{\boldsymbol{x}\in\boldsymbol{\mathcal{H}}}{\text{minimize}}\;\; \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{L}\boldsymbol{x} - \boldsymbol{r}) - \langle \boldsymbol{x} \mid \boldsymbol{j}(z)\rangle_{\boldsymbol{\mathcal{H}}} \tag{4.14}$$

and

$$\underset{\boldsymbol{v}\in\boldsymbol{\mathcal{G}}}{\text{minimize}}\;\; \boldsymbol{f}^*(\boldsymbol{j}(z) - \boldsymbol{L}^*\boldsymbol{v}) + \boldsymbol{g}^*(\boldsymbol{v}) + \langle \boldsymbol{r} \mid \boldsymbol{v}\rangle_{\boldsymbol{\mathcal{G}}}\,, \tag{4.15}$$

respectively. On the other hand since, for every $\gamma \in \left]0,+\infty\right[$, $\operatorname{prox}_{\gamma\boldsymbol{f}}\colon \boldsymbol{x} \mapsto \boldsymbol{j}(\sum_{i=1}^m \omega_i x_i)$ and $\operatorname{prox}_{\gamma\boldsymbol{g}^*} = (\operatorname{prox}_{\gamma g_i^*})_{1\leq i\leq m}$, (4.12) is a particular case of (4.8). Finally, in (ii), $\boldsymbol{g}^*$ is strongly, hence uniformly, convex at $\overline{\boldsymbol{v}}$. Altogether, the results follow from Proposition 4.2. ☐

Remark 4.5 Suppose that (4.10) has a solution and that

$$(r_1,\ldots,r_m) \in \operatorname{sri}\big\{(L_1x - y_1,\ldots,L_mx - y_m) \mid x \in \mathcal{H},\, y_1 \in \operatorname{dom} g_1,\ldots,\, y_m \in \operatorname{dom} g_m\big\}. \tag{4.16}$$

Then, with the notation of the proof of Proposition 4.4, (4.16) is equivalent to $\boldsymbol{r} \in \operatorname{sri}(\boldsymbol{L}(\boldsymbol{V}) - \operatorname{dom}\boldsymbol{g}) = \operatorname{sri}(\boldsymbol{L}(\operatorname{dom}\boldsymbol{f}) - \operatorname{dom}\boldsymbol{g})$. Thus, Proposition 4.1(iii)(a) asserts that (4.9) holds.

References

[1] G. Alduncin, Composition duality principles for mixed variational inequalities, *Math. Comput. Modelling,* vol. 41, pp. 639–654, 2005.

[2] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

[3] H. Attouch and M. Théra, A general duality principle for the sum of two operators, *J. Convex Anal.,* vol. 3, pp. 1–24, 1996.

[4] H. Attouch and M. Théra, A duality proof of the Hille-Yosida theorem, in: *Progress in Partial Differential Equations: the Metz Surveys,* vol. 4, pp. 18–35, 1996.

[5] H. H. Bauschke and P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Math. Oper. Res.,* vol. 26, pp. 248–264, 2001.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer-Verlag, New York, 2011.

[7] H. H. Bauschke, P. L. Combettes, and S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear Anal.*, vol. 60, pp. 283–301, 2005.

[8] G. Chen and M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Math. Programming,* vol. 64, pp. 81–101, 1994.

[9] P. L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in: *Inherently Parallel Algorithms for Feasibility and Optimization*, D. Butnariu, Y. Censor, S. Reich (eds.). Elsevier, Amsterdam, pp. 115–152, 2001.

[10] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization,* vol. 53, pp. 475–504, 2004.

[11] P. L. Combettes, Iterative construction of the resolvent of a sum of maximal monotone operators, *J. Convex Anal.,* vol. 16, pp. 727–748, 2009.

[12] P. L. Combettes and J.-C. Pesquet, Proximal splitting methods in signal processing, in: *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* H. H. Bauschke *et al.* (eds.). Springer-Verlag, New York, 2011.

[13] P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.*, vol. 4, pp. 1168–1200, 2005.

[14] J. Eckstein and D. P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Math. Programming,* vol. 55, pp. 293–318, 1992.

[15] J. Eckstein and M. C. Ferris, Smooth methods of multipliers for complementarity problems, *Math. Programming,* vol. 86, pp. 65–90, 1999.

[16] J. Eckstein and B. F. Svaiter, A family of projective splitting methods for the sum of two maximal monotone operators, *Math. Programming,* vol. 111, pp. 173–199, 2008.

[17] I. Ekeland and R. Temam, *Analyse Convexe et Problèmes Variationnels,* Dunod, Paris, 1974; *Convex Analysis and Variational Problems,* SIAM, Philadelphia, PA, 1999.

[18] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer-Verlag, New York, 2003.

[19] M. Fukushima, The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with applications to the traffic equilibrium problem, *Math. Programming,* vol. 72, pp. 1–15, 1996.

[20] M. Fortin and R. Glowinski (eds.), *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary Value Problems.* North-Holland, Amsterdam, 1983.

[21] D. Gabay, Applications of the method of multipliers to variational inequalities, in: M. Fortin and R. Glowinski (eds.), *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary Value Problems,* pp. 299–331. North-Holland, Amsterdam, 1983.

[22] R. Glowinski and P. Le Tallec (eds.), *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics.* SIAM, Philadelphia, 1989.

[23] P.-L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.*, vol. 16, pp. 964–979, 1979.

[24] L. McLinden, An extension of Fenchel's duality theorem to saddle functions and dual minimax problems, *Pacific J. Math.,* vol. 50, pp. 135–158, 1974.

[25] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[26] B. Mercier, *Inéquations Variationnelles de la Mécanique* (Publications Mathématiques d'Orsay, no. 80.01). Université de Paris-XI, Orsay, France, 1980.

[27] U. Mosco, Dual variational inequalities, *J. Math. Anal. Appl.,* vol. 40, pp. 202–206, 1972.

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

[29] T. Pennanen, A splitting method for composite mappings, *Numer. Funct. Anal. Optim.*, vol. 23, pp. 875–890, 2002.

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

[31] S. M. Robinson, Generalized duality in variational analysis, in: N. Hadjisavvas and P. M. Pardalos (eds.), *Advances in Convex Analysis and Global Optimization,* pp. 205–219. Dordrecht, The Netherlands, Kluwer, 2001.

[32] R. T. Rockafellar, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.,* vol. 21, pp. 167–187, 1967.

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

[34] R. T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.,* vol. 1, 97–116, 1976.

[35] R. T. Rockafellar, *Conjugate Duality and Optimization.* SIAM, Philadelphia, PA, 1974.

[36] B. F. Svaiter, Weak convergence on Douglas-Rachford method, *SIAM J. Control Optim.,* to appear.

[37] P. Tseng, Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming, *Math. Programming,* vol. 48, no. 2, pp. 249–263, 1990.

[38] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM J. Control Optim.,* vol. 29, pp. 119–138, 1991.

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

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

[41] E. Zeidler, *Nonlinear Functional Analysis and Its Applications,* vols. I–V. Springer-Verlag, New York, 1985–1993.