Iterative construction of the resolvent of a sum of maximal monotone operators

Patrick L. Combettes

UPMC Université Paris 06
Laboratoire Jacques-Louis Lions – UMR 7598
75005 Paris, France
plc@math.jussieu.fr

Abstract

We propose two inexact parallel splitting algorithms for computing the resolvent of a weighted sum of maximal monotone operators in a Hilbert space and show their strong convergence. We start by establishing new results on the asymptotic behavior of the Douglas-Rachford splitting algorithm for the sum of two operators. These results serve as a basis for the first algorithm. The second algorithm is based on an extension of a recent Dykstra-like method for computing the resolvent of the sum of two maximal monotone operators. Under standard qualification conditions, these two algorithms provide a means for computing the proximity operator of a weighted sum of lower semicontinuous convex functions. We show that a version of the second algorithm performs the same task without requiring any qualification condition. In turn, this provides a parallel splitting algorithm for qualification-free strongly convex programming.

2000 Mathematics Subject Classification: Primary 47H05; Secondary 47J25, 49M29, 65K05, 90C25.

Keywords: Dykstra's algorithm, Douglas-Rachford algorithm, maximal monotone operator, method of partial inverses, operator splitting, proximity operator, resolvent

1 Introduction and notation

Let $\mathcal{H}$ be a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$ and norm $\| \cdot \|$, and let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator, i.e.,

$$(\forall (x,y) \in \mathcal{H} \times \mathcal{H})(\forall (u,v) \in Ax \times Ay) \quad \langle x - y \mid u - v \rangle \geq 0. \tag{1}$$

We denote by $\operatorname{ran} A = \big\{ u \in \mathcal{H} \mid (\exists\, x \in \mathcal{H})\ u \in Ax \big\}$ the range of A, by $\operatorname{zer} A = \big\{ x \in \mathcal{H} \mid 0 \in Ax \big\}$ its set of zeros, by $\operatorname{gra} A = \big\{ (x,u) \in \mathcal{H} \times \mathcal{H} \mid u \in Ax \big\}$ its graph, and by A^{-1} its inverse, i.e., the operator with graph $\{(u,x)\in\mathcal{H}\times\mathcal{H}\mid u\in Ax\}$. The resolvent of A is $J_A=(\mathrm{Id}+A)^{-1}$. This operator enjoys many important properties that make it a central tool in monotone operator theory and its applications [3, 10, 44, 45, 52]. In particular, it is single-valued, firmly nonexpansive in the sense that

$$(\forall x\in\operatorname{ran}(\mathrm{Id}+A))(\forall y\in\operatorname{ran}(\mathrm{Id}+A))\quad \|J_Ax-J_Ay\|^2\leq\langle x-y\mid J_Ax-J_Ay\rangle\,,\tag{2}$$

and Minty's theorem states that it is defined everywhere in $\mathcal{H}$, i.e., $\operatorname{ran}(\mathrm{Id}+A)=\mathcal{H}$, if and only if A is maximal monotone in the sense that, if $B\colon\mathcal{H}\to 2^{\mathcal{H}}$ is monotone and $\operatorname{gra}A\subset\operatorname{gra}B$, then $B=A$. Moreover, (2) implies that the reflection operator $R_A=2J_A-\mathrm{Id}$ is nonexpansive, that is,

$$(\forall x\in\operatorname{ran}(\mathrm{Id}+A))(\forall y\in\operatorname{ran}(\mathrm{Id}+A))\quad \|R_Ax-R_Ay\|\leq\|x-y\|.\tag{3}$$

Finally, the set $\operatorname{Fix}J_A=\{x\in\mathcal{H}\mid J_Ax=x\}$ of fixed points of J_A coincides with $\operatorname{zer}A$.

The goal of this paper is to propose two strongly convergent splitting methods for computing the resolvent of a monotone operator $A\colon\mathcal{H}\to 2^{\mathcal{H}}$ which can be decomposed as a weighted sum of maximal monotone operators $(A_i)_{1\leq i\leq m}$, say

$$A=\sum_{i=1}^m\omega_iA_i,\quad\text{where}\quad\{\omega_i\}_{1\leq i\leq m}\subset\,]0,1[\quad\text{and}\quad\sum_{i=1}^m\omega_i=1,\tag{4}$$

where the individual resolvents $(J_{A_i})_{1\leq i\leq m}$ can be implemented relatively easily. Both methods proceed by splitting in the sense that, at each iteration, they employ these resolvents separately. In addition, only approximate evaluations of the resolvents are needed. Note that since computing J_Ar for some $r\in\mathcal{H}$ is equivalent to solving

$$r\in x+\sum_{i=1}^m\omega_iA_ix,\tag{5}$$

the proposed algorithms can be viewed as splitting methods for solving strongly monotone inclusions (recall that A is said to be α-strongly monotone for some $\alpha\in\,]0,+\infty[$ if $A-\alpha\,\mathrm{Id}$ is monotone).

The first method is discussed in Section 2. It is based on new results that we establish on the asymptotic behavior of the Douglas-Rachford splitting method for the sum of two maximal monotone operators. This section also contains new results on the convergence of a splitting method for the (not necessarily strongly monotone) sum of m maximal monotone operators. In Section 3, we present an alternative method, which finds its roots in a recent extension of Dykstra's best approximation algorithm to the construction of the resolvent of the sum of two monotone operators. In Section 4, we turn our attention to the problem of constructing the proximity operator of functions that can be decomposed as weighted sums of m proper lower semicontinuous convex functions. This problem can naturally be tackled by restricting the algorithms of Sections 2 and 3 to subdifferentials, but at the expense of imposing qualification conditions. Instead, we exploit a recent extension of Dykstra's projection method to the construction of the proximity operator of the sum of two convex functions to obtain a splitting method that requires no qualification condition. Connections with projection methods, as well as applications to qualification-free strongly convex programming and signal denoising are also discussed.

In addition to the notation already introduced above, we shall also need the following. The symbols $\rightharpoonup$ and $\to$ denote respectively weak and strong convergence. The projector onto a nonempty closed convex set $C \subset \mathcal{H}$ is denoted by P_C, its indicator function by ι_C, and its normal cone operator by N_C, i.e.,

$$N_C \colon \mathcal{H} \to 2^{\mathcal{H}} \colon x \mapsto \begin{cases} \big\{u \in \mathcal{H} \mid (\forall y \in C) \quad \langle y - x \mid u \rangle \leq 0\big\}, & \text{if } x \in C; \\ \varnothing, & \text{otherwise.} \end{cases} \tag{6}$$

Finally, a monotone operator $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ is uniformly monotone on $C \subset \mathcal{H}$ if there exists an increasing function $\phi \colon \left[0, +\infty\right[\to \left[0, +\infty\right[$ vanishing only at 0 such that

$$(\forall (x, y) \in C \times C)(\forall (u, v) \in Ax \times Ay) \quad \langle x - y \mid u - v \rangle \geq \phi(\|x - y\|). \tag{7}$$

In particular, if $\phi \colon t \mapsto \alpha t^2$ for some $\alpha \in \left]0, +\infty\right[$, then A is α-strongly monotone on C.

2 Douglas-Rachford splitting for the resolvent of the sum

In the context of monotone operator theory, what is known as the Douglas-Rachford algorithm is a splitting scheme initially proposed in [34] for finding a zero of the sum of two maximal monotone operators (we refer the reader to [17] for connections with the original work of Douglas and Rachford [24]). In Section 2.1, we present new convergence results for this algorithm. In Section 2.2, these results are utilized to obtain weak and strong convergence conditions for a splitting scheme devised to find a zero of the weighted sum of m maximal monotone operators. This scheme is shown to be closely related to an algorithm originally designed by Spingarn [48, 49] with different tools. The application of the results of Section 2.2 to the construction of the resolvent of the sum of m maximal monotone operators is discussed in Section 2.3.

2.1 Asymptotic behavior of the Douglas-Rachford algorithm

We revisit an algorithm which was proposed in its initial form by Lions and Mercier in [34].

Theorem 2.1 *Let $(\boldsymbol{\mathcal{H}}, |||\cdot|||)$ be a real Hilbert space, let $\boldsymbol{A}$ and $\boldsymbol{B}$ be maximal monotone operators from $\boldsymbol{\mathcal{H}}$ to $2^{\boldsymbol{\mathcal{H}}}$ such that $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \neq \varnothing$, let $\gamma \in \left]0, +\infty\right[$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $\left]0, 2\right]$, and let $(\boldsymbol{a}_n)_{n \in \mathbb{N}}$ and $(\boldsymbol{b}_n)_{n \in \mathbb{N}}$ be sequences in $\boldsymbol{\mathcal{H}}$. Furthermore, let $(\boldsymbol{y}_n)_{n \in \mathbb{N}}$ and $(\boldsymbol{z}_n)_{n \in \mathbb{N}}$ be the sequences generated by the following routine.*

$$\begin{array}{l} \text{Initialization} \\ \left\lfloor \; \boldsymbol{z}_0 \in \boldsymbol{\mathcal{H}} \right. \\ \text{For } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{ll} \boldsymbol{y}_n & = J_{\gamma \boldsymbol{B}} \boldsymbol{z}_n + \boldsymbol{b}_n \\ \boldsymbol{z}_{n+1} & = \boldsymbol{z}_n + \lambda_n \big(J_{\gamma \boldsymbol{A}}(2\boldsymbol{y}_n - \boldsymbol{z}_n) + \boldsymbol{a}_n - \boldsymbol{y}_n \big). \end{array} \right. \end{array} \tag{8}$$

Then the following hold.

(i) *Suppose that* $\sum_{n\in\mathbb{N}}\lambda_n(|||\boldsymbol{a}_n|||+|||\boldsymbol{b}_n|||)<+\infty$*, that* $\sum_{n\in\mathbb{N}}\lambda_n(2-\lambda_n)=+\infty$*, and that* $(\forall n\in\mathbb{N})$ $\lambda_n<2$*. Then the following hold.*

   (a) $(\boldsymbol{z}_n)_{n\in\mathbb{N}}$ *converges weakly to a point* $\boldsymbol{z}\in\operatorname{Fix}(R_{\gamma\boldsymbol{A}}\circ R_{\gamma\boldsymbol{B}})$ *and* $J_{\gamma\boldsymbol{B}}\boldsymbol{z}$ *is a zero of* $\boldsymbol{A}+\boldsymbol{B}$*.*

   (b) $\big(R_{\gamma\boldsymbol{A}}(R_{\gamma\boldsymbol{B}}\boldsymbol{z}_n)-\boldsymbol{z}_n\big)_{n\in\mathbb{N}}$ *converges strongly to* 0*.*

   (c) *Suppose that* $J_{\gamma\boldsymbol{B}}$ *is weakly sequentially continuous and that* $\boldsymbol{b}_n\rightharpoonup 0$*. Then* $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$ *converges weakly to a zero of* $\boldsymbol{A}+\boldsymbol{B}$*.*

   (d) *Suppose that* $\boldsymbol{\mathcal{H}}$ *is finite dimensional. Then* $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$ *converges to a zero of* $\boldsymbol{A}+\boldsymbol{B}$*.*

   (e) *Suppose that* $\boldsymbol{A}=N_{\boldsymbol{D}}$*, where* $\boldsymbol{D}$ *is a closed affine subspace of* $\boldsymbol{\mathcal{H}}$*. Then* $(J_{\gamma\boldsymbol{A}}\boldsymbol{z}_n)_{n\in\mathbb{N}}$ *converges weakly to a zero of* $\boldsymbol{A}+\boldsymbol{B}$*.*

   (f) *Suppose that* $\boldsymbol{A}=N_{\boldsymbol{D}}$*, where* $\boldsymbol{D}$ *is a closed vector subspace of* $\boldsymbol{\mathcal{H}}$*, and that* $\boldsymbol{b}_n\rightharpoonup 0$*. Then* $(J_{\gamma\boldsymbol{A}}\boldsymbol{y}_n)_{n\in\mathbb{N}}$ *converges weakly to a zero of* $\boldsymbol{A}+\boldsymbol{B}$*.*

(ii) *Suppose that* $\sum_{n\in\mathbb{N}}|||\boldsymbol{a}_n|||<+\infty$*, that* $\sum_{n\in\mathbb{N}}|||\boldsymbol{b}_n|||<+\infty$*, and that* $\inf_{n\in\mathbb{N}}\lambda_n>0$*. Then the following hold.*

   (a) *Suppose that* $\operatorname{int}\operatorname{Fix}(R_{\gamma\boldsymbol{A}}\circ R_{\gamma\boldsymbol{B}})\neq\varnothing$*. Then* $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$ *converges strongly to a zero of* $\boldsymbol{A}+\boldsymbol{B}$*.*

   (b) *Suppose that* $\boldsymbol{B}$ *is uniformly monotone on the bounded subsets of* $\boldsymbol{\mathcal{H}}$*. Then* $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$ *converges strongly to the unique zero of* $\boldsymbol{A}+\boldsymbol{B}$*.*

Proof. Denote the scalar product of $\boldsymbol{\mathcal{H}}$ by $\langle\langle\cdot\mid\cdot\rangle\rangle$ and set $T=R_{\gamma\boldsymbol{A}}\circ R_{\gamma\boldsymbol{B}}$. Then it follows from (3) that T is nonexpansive. Moreover, [17, Lemma 2.6] (see also [34]) asserts that

$$T=2J_{\gamma\boldsymbol{A}}\circ\big(2J_{\gamma\boldsymbol{B}}-\mathbf{Id}\big)+\mathbf{Id}-2J_{\gamma\boldsymbol{B}} \tag{9}$$

and that

$$\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})=J_{\gamma\boldsymbol{B}}(\operatorname{Fix}T). \tag{10}$$

(i)(a): [17, Corollary 5.2] and its proof.

(i)(b): See the proofs of [17, Corollary 5.2] and [17, Lemma 5.1].

(i)(c): It follows from (i)(a) that $\boldsymbol{y}_n=J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n+\boldsymbol{b}_n\rightharpoonup J_{\gamma\boldsymbol{B}}\boldsymbol{z}\in\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$.

(i)(c)⇒(i)(d): Clear by continuity of $J_{\gamma\boldsymbol{B}}$.

(i)(e): By assumption $J_{\gamma\boldsymbol{A}}=P_{\boldsymbol{D}}$ is the projector onto $\boldsymbol{D}$ and it is therefore continuous and affine. As seen in (i)(a), $\boldsymbol{z}_n\rightharpoonup\boldsymbol{z}\in\operatorname{Fix}T$ and $J_{\gamma\boldsymbol{B}}\boldsymbol{z}\in\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$. Hence, since $P_{\boldsymbol{D}}$ is weakly continuous, $P_{\boldsymbol{D}}\boldsymbol{z}_n\rightharpoonup P_{\boldsymbol{D}}\boldsymbol{z}$. However, it follows from (9) that

$$\begin{aligned}\boldsymbol{z}\in\operatorname{Fix}T\quad&\Leftrightarrow\quad\boldsymbol{z}=2P_{\boldsymbol{D}}(2J_{\gamma\boldsymbol{B}}\boldsymbol{z}+(1-2)\boldsymbol{z})+\boldsymbol{z}-2J_{\gamma\boldsymbol{B}}\boldsymbol{z}\\&\Leftrightarrow\quad J_{\gamma\boldsymbol{B}}\boldsymbol{z}=2P_{\boldsymbol{D}}(J_{\gamma\boldsymbol{B}}\boldsymbol{z})+(1-2)P_{\boldsymbol{D}}\boldsymbol{z}\in\boldsymbol{D}\\&\Leftrightarrow\quad P_{\boldsymbol{D}}(J_{\gamma\boldsymbol{B}}\boldsymbol{z})=J_{\gamma\boldsymbol{B}}\boldsymbol{z}=2P_{\boldsymbol{D}}(J_{\gamma\boldsymbol{B}}\boldsymbol{z})-P_{\boldsymbol{D}}\boldsymbol{z}\\&\Leftrightarrow\quad P_{\boldsymbol{D}}\boldsymbol{z}=J_{\gamma\boldsymbol{B}}\boldsymbol{z}.\end{aligned} \tag{11}$$

Altogether, $J_{\gamma\boldsymbol{A}}\boldsymbol{z}_n=P_{\boldsymbol{D}}\boldsymbol{z}_n\rightharpoonup P_{\boldsymbol{D}}\boldsymbol{z}=J_{\gamma\boldsymbol{B}}\boldsymbol{z}\in\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$.

(i)(f): By assumption $J_{\gamma \boldsymbol{A}} = P_{\boldsymbol{D}}$ is linear and nonexpansive. Hence, we derive from (9) and (i)(b) that

$$\begin{aligned}|||P_{\boldsymbol{D}}(J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n) - P_{\boldsymbol{D}}\boldsymbol{z}_n||| &= |||P_{\boldsymbol{D}}(P_{\boldsymbol{D}}(2J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n - \boldsymbol{z}_n) - J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n)||| \\ &\leq |||P_{\boldsymbol{D}}(2J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n - \boldsymbol{z}_n) - J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n||| \\ &= \frac{1}{2}|||T\boldsymbol{z}_n - \boldsymbol{z}_n||| \\ &\to 0. \end{aligned} \tag{12}$$

On the other hand, it follows from (i)(e) that there exists $\boldsymbol{y} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ such that $P_{\boldsymbol{D}}\boldsymbol{z}_n \rightharpoonup \boldsymbol{y}$. Thus, (12) yields $P_{\boldsymbol{D}}(J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n) \rightharpoonup \boldsymbol{y}$. Since, by weak continuity of $P_{\boldsymbol{D}}$, $P_{\boldsymbol{D}}\boldsymbol{b}_n \rightharpoonup 0$, we conclude that $J_{\gamma \boldsymbol{A}}\boldsymbol{y}_n = P_{\boldsymbol{D}}(J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n) + P_{\boldsymbol{D}}\boldsymbol{b}_n \rightharpoonup \boldsymbol{y}$.

(ii): Set $\lambda = \inf_{n\in\mathbb{N}} \lambda_n$, let $\boldsymbol{z} \in \operatorname{Fix} T$, and set

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{c}_n = \mu_n\big(2\boldsymbol{a}_n + R_{\gamma \boldsymbol{A}}(R_{\gamma \boldsymbol{B}}\boldsymbol{z}_n + 2\boldsymbol{b}_n) - R_{\gamma \boldsymbol{A}}(R_{\gamma \boldsymbol{B}}\boldsymbol{z}_n)\big), \quad \text{where} \quad \mu_n = \frac{\lambda_n}{2}. \tag{13}$$

Using (9) and straightforward manipulations, we derive from (8) that

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{z}_{n+1} = \boldsymbol{x}_n + \boldsymbol{c}_n, \quad \text{where} \quad \boldsymbol{x}_n = \boldsymbol{z}_n + \mu_n(T\boldsymbol{z}_n - \boldsymbol{z}_n). \tag{14}$$

Hence, since T is nonexpansive,

$$(\forall n \in \mathbb{N}) \quad |||\boldsymbol{z}_{n+1} - \boldsymbol{z}||| \leq |||\boldsymbol{x}_n - \boldsymbol{z}||| + |||\boldsymbol{c}_n||| \tag{15}$$

$$\leq (1-\mu_n)|||\boldsymbol{z}_n - \boldsymbol{z}||| + \mu_n|||T\boldsymbol{z}_n - T\boldsymbol{z}||| + |||\boldsymbol{c}_n|||$$

$$\leq |||\boldsymbol{z}_n - \boldsymbol{z}||| + |||\boldsymbol{c}_n|||. \tag{16}$$

Moreover, since $R_{\gamma \boldsymbol{A}}$ is nonexpansive and $\sup_{n\in\mathbb{N}} \mu_n \leq 1$, (13) yields

$$\begin{aligned}\sum_{n\in\mathbb{N}} |||\boldsymbol{c}_n||| &\leq 2\sum_{n\in\mathbb{N}} |||\boldsymbol{a}_n||| + \sum_{n\in\mathbb{N}} |||R_{\gamma \boldsymbol{A}}(R_{\gamma \boldsymbol{B}}\boldsymbol{z}_n + 2\boldsymbol{b}_n) - R_{\gamma \boldsymbol{A}}(R_{\gamma \boldsymbol{B}}\boldsymbol{z}_n)||| \\ &\leq 2\sum_{n\in\mathbb{N}} (|||\boldsymbol{a}_n||| + |||\boldsymbol{b}_n|||) \\ &< +\infty. \end{aligned} \tag{17}$$

In turn, we derive from (16), (17), and [42, Lemma 2.2.2] that

$$(|||\boldsymbol{z}_n - \boldsymbol{z}|||)_{n\in\mathbb{N}} \text{ converges.} \tag{18}$$

(ii)(a): It follows from (16), (17), and [16, Proposition 3.10] that there exists $\boldsymbol{x} \in \boldsymbol{\mathcal{H}}$ such that $\boldsymbol{z}_n \to \boldsymbol{x}$. Hence, by continuity of T, $T\boldsymbol{z}_n - \boldsymbol{z}_n \to T\boldsymbol{x} - \boldsymbol{x}$. On the other hand, (14) and (17) yield

$$|||T\boldsymbol{z}_n - \boldsymbol{z}_n||| = \frac{1}{\mu_n}|||\boldsymbol{z}_{n+1} - \boldsymbol{z}_n - \boldsymbol{c}_n||| \leq \frac{2}{\lambda}(|||\boldsymbol{z}_{n+1} - \boldsymbol{x}||| + |||\boldsymbol{z}_n - \boldsymbol{x}||| + |||\boldsymbol{c}_n|||) \to 0. \tag{19}$$

As a result, $T\boldsymbol{x} - \boldsymbol{x} = 0$, i.e., $\boldsymbol{x} \in \operatorname{Fix} T$. Appealing to (10), we conclude that $\boldsymbol{y}_n = J_{\gamma \boldsymbol{B}}\boldsymbol{z}_n + \boldsymbol{b}_n \to J_{\gamma \boldsymbol{B}}\boldsymbol{x} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$.

(ii)(b): By assumption, $\boldsymbol{B}$ is strictly monotone, and so is therefore $\boldsymbol{A}+\boldsymbol{B}$. Hence $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ is a singleton. Next, in view of (18) and of the nonexpansivity of $J_{\gamma\boldsymbol{B}}$, there exists a bounded set $\boldsymbol{C}\subset\boldsymbol{\mathcal{H}}$ that contains $(J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n)_{n\in\mathbb{N}}$ and $J_{\gamma\boldsymbol{B}}\boldsymbol{z}$. On the other hand, since $\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n \in \gamma\boldsymbol{B}(J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n)$ and $\boldsymbol{z}-J_{\gamma\boldsymbol{B}}\boldsymbol{z}\in\gamma\boldsymbol{B}(J_{\gamma\boldsymbol{B}}\boldsymbol{z})$, (7) yields

$$(\forall n\in\mathbb{N})\quad \langle\langle J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z} \mid \boldsymbol{z}_n-\boldsymbol{z}\rangle\rangle \geq |||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||^2 + \gamma\phi(|||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||), \tag{20}$$

for some increasing function $\phi\colon\,[0,+\infty[\;\to[0,+\infty[$ that vanishes only at 0. Hence, since $R_{\gamma\boldsymbol{A}}$ is nonexpansive,

$$\begin{aligned}(\forall n\in\mathbb{N})\quad |||T\boldsymbol{z}_n-\boldsymbol{z}|||^2 &= |||R_{\gamma\boldsymbol{A}}(R_{\gamma\boldsymbol{B}}\boldsymbol{z}_n) - R_{\gamma\boldsymbol{A}}(R_{\gamma\boldsymbol{B}}\boldsymbol{z})|||^2\\ &\leq |||R_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - R_{\gamma\boldsymbol{B}}\boldsymbol{z}|||^2\\ &= |||\boldsymbol{z}_n-\boldsymbol{z}|||^2 - 4\langle\langle J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}\mid \boldsymbol{z}_n-\boldsymbol{z}\rangle\rangle + 4|||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||^2\\ &\leq |||\boldsymbol{z}_n-\boldsymbol{z}|||^2 - 4\gamma\phi(|||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||).\end{aligned} \tag{21}$$

Using (14) and (21), we obtain

$$\begin{aligned}(\forall n\in\mathbb{N})\quad |||\boldsymbol{x}_n-\boldsymbol{z}|||^2 &\leq (1-\mu_n)|||\boldsymbol{z}_n-\boldsymbol{z}|||^2 + \mu_n|||T\boldsymbol{z}_n-\boldsymbol{z}|||^2\\ &\leq |||\boldsymbol{z}_n-\boldsymbol{z}|||^2 - 4\mu_n\gamma\phi(|||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||)\\ &\leq |||\boldsymbol{z}_n-\boldsymbol{z}|||^2 - 2\lambda\gamma\phi(|||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||) \qquad (22)\\ &\leq |||\boldsymbol{z}_n-\boldsymbol{z}|||^2. \qquad (23)\end{aligned}$$

Now set $\nu = 2\sup_{k\in\mathbb{N}}|||\boldsymbol{x}_k-\boldsymbol{z}||| + \sup_{k\in\mathbb{N}}|||\boldsymbol{c}_k|||$. It follows from (17), (18), and (23) that $\nu<+\infty$. Furthermore, we derive from (15) and (22) that

$$\begin{aligned}(\forall n\in\mathbb{N})\quad |||\boldsymbol{z}_{n+1}-\boldsymbol{z}|||^2 &\leq |||\boldsymbol{x}_n-\boldsymbol{z}|||^2 + (2|||\boldsymbol{x}_n-\boldsymbol{z}||| + |||\boldsymbol{c}_n|||)|||\boldsymbol{c}_n|||\\ &\leq |||\boldsymbol{z}_n-\boldsymbol{z}|||^2 - 2\lambda\gamma\phi(|||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||) + \nu|||\boldsymbol{c}_n|||.\end{aligned} \tag{24}$$

Thus, (17) and (18) yield $\phi(|||J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n - J_{\gamma\boldsymbol{B}}\boldsymbol{z}|||)\to 0$ and, in turn, $J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n\to J_{\gamma\boldsymbol{B}}\boldsymbol{z}$. Hence, we get $\boldsymbol{y}_n = J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n + \boldsymbol{b}_n \to J_{\gamma\boldsymbol{B}}\boldsymbol{z}$ and, in view of (10), the proof is complete. ☐

Remark 2.2 Let us make a few commentaries about Theorem 2.1 and its connections to results available in the literature.

(i) Special cases of Theorem 2.1(i)(a) are [15, Proposition 12], [26, Theorem 7], and the original Lions and Mercier result [34, Theorem 1]. Let us note that, at this level of generality, there is no weak or strong convergence result available for the sequences $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$, $(J_{\gamma\boldsymbol{A}}\,\boldsymbol{y}_n)_{n\in\mathbb{N}}$, $(J_{\gamma\boldsymbol{A}}\,\boldsymbol{z}_n)_{n\in\mathbb{N}}$, and $(J_{\gamma\boldsymbol{B}}\,\boldsymbol{z}_n)_{n\in\mathbb{N}}$.

(ii) In numerical applications [20, 28, 35], the scaling parameter γ has been experienced to impact the speed of convergence of the Douglas-Rachford algorithm.

(iii) The conditions $\sum_{n\in\mathbb{N}}\lambda_n(2-\lambda_n)=+\infty$ and $\sum_{n\in\mathbb{N}}\lambda_n|||\boldsymbol{b}_n|||<+\infty$ used in Theorem 2.1(i) do not imply that $\boldsymbol{b}_n\rightharpoonup 0$. Indeed, set

$$(\forall n\in\mathbb{N})\quad \lambda_n = \begin{cases}1, & \text{if } n=0;\\ 1/n, & \text{if } n\geq 1.\end{cases} \tag{25}$$

Then $\sum_{n\in\mathbb{N}}\lambda_n(2-\lambda_n) \geq \sum_{n\in\mathbb{N}}\lambda_n = +\infty$. Now let $(\boldsymbol{e}_n)_{n\in\mathbb{N}}$ be a sequence of unit norm vectors in $\boldsymbol{\mathcal{H}}$ and set

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{b}_n = \begin{cases} \boldsymbol{e}_0, & \text{if } n \in \mathbb{S}; \\ \boldsymbol{e}_n/n, & \text{if } n \notin \mathbb{S}, \end{cases} \tag{26}$$

where $\mathbb{S} = \big\{n \in \mathbb{N} \mid (\exists\, k \in \mathbb{N})\ n = k^2\big\}$. Then clearly $\boldsymbol{b}_n \not\to 0$. However,

$$\begin{aligned} \sum_{n\in\mathbb{N}} \lambda_n |||\boldsymbol{b}_n||| &= 1 + \sum_{n\in\mathbb{S}\smallsetminus\{0\}} \frac{1}{n} + \sum_{n\in\mathbb{N}\smallsetminus\mathbb{S}} \frac{1}{n^2} \\ &= 1 + \sum_{n\in\mathbb{N}\smallsetminus\{0\}} \frac{1}{n^2} + \sum_{n\in\mathbb{N}\smallsetminus\mathbb{S}} \frac{1}{n^2} \\ &< +\infty. \end{aligned} \tag{27}$$

It is noteworthy that our framework allows for non summable error sequences: in the above example, we actually have $\overline{\lim}\,|||\boldsymbol{b}_n||| = 1$.

(iv) If we set $\lambda_n \equiv 2$ in Theorem 2.1(ii), we obtain strong convergence conditions for an inexact version of the Peaceman-Rachford algorithm [17, 34]. In general, the sequences $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{z}_n)_{n\in\mathbb{N}}$ produced by the Peaceman-Rachford algorithm do not converge, even weakly.

(v) In [7], the asymptotic behavior of algorithm (8) when $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) = \varnothing$ is investigated in the special case when $\boldsymbol{A}$ and $\boldsymbol{B}$ are the normal cone operators of closed convex sets and when $\lambda_n \equiv 1$, $\boldsymbol{a}_n \equiv 0$, and $\boldsymbol{b}_n \equiv 0$.

(vi) Suppose that $\boldsymbol{B} = 0$ and that $\boldsymbol{b}_n \equiv 0$ in Theorem 2.1(i)(c). Then $(\boldsymbol{z}_n)_{n\in\mathbb{N}} = (\boldsymbol{y}_n)_{n\in\mathbb{N}}$ and we obtain the weak convergence to a zero of $\boldsymbol{A}$ of the proximal point iteration

$$\boldsymbol{z}_0 \in \boldsymbol{\mathcal{H}} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \boldsymbol{z}_{n+1} = \boldsymbol{z}_n + \lambda_n\big(J_{\gamma\boldsymbol{A}}\boldsymbol{z}_n + \boldsymbol{a}_n - \boldsymbol{z}_n\big) \tag{28}$$

provided that $\sum_{n\in\mathbb{N}}\lambda_n|||\boldsymbol{a}_n||| < +\infty$, $\sum_{n\in\mathbb{N}}\lambda_n(2-\lambda_n) = +\infty$, and $(\forall n \in \mathbb{N})\ \lambda_n < 2$. Alternate convergence results for the proximal point algorithm can be found in [17] and the references therein, in particular in the classical papers [11, 44].

Corollary 2.3 *Let $(\boldsymbol{\mathcal{H}}, |||\cdot|||)$ be a real Hilbert space, let $\boldsymbol{D}$ be a closed affine subspace of $\boldsymbol{\mathcal{H}}$, let $\boldsymbol{B}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ be a maximal monotone operator such that $\operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B}) \neq \varnothing$, let $\gamma \in \,]0,+\infty[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0,2]$, and let $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ be a sequence in $\boldsymbol{\mathcal{H}}$. Furthermore, let $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ be the sequences generated by the following routine.*

$$\begin{array}{l} \text{Initialization} \\ \left\lfloor \; \boldsymbol{z}_0 \in \boldsymbol{\mathcal{H}} \right. \\ \text{For } n = 0,1,\ldots \\ \left\lfloor \begin{array}{l} \boldsymbol{y}_n \;\;\; = J_{\gamma\boldsymbol{B}}\boldsymbol{z}_n + \boldsymbol{b}_n \\ \boldsymbol{x}_n \;\;\; = P_{\boldsymbol{D}}\boldsymbol{y}_n \\ \boldsymbol{p}_n \;\;\; = P_{\boldsymbol{D}}\boldsymbol{z}_n \\ \boldsymbol{z}_{n+1} = \boldsymbol{z}_n + \lambda_n(2\boldsymbol{x}_n - \boldsymbol{p}_n - \boldsymbol{y}_n). \end{array} \right. \end{array} \tag{29}$$

Then the following hold.

(i) *Suppose that* $\sum_{n\in\mathbb{N}} \lambda_n |||\boldsymbol{b}_n||| < +\infty$, *that* $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$, *and that* $(\forall n \in \mathbb{N})\ \lambda_n < 2$. *Then the following hold.*

(a) $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ *converges weakly to a zero of* $N_{\boldsymbol{D}} + \boldsymbol{B}$.

(b) *Suppose that* $\boldsymbol{D}$ *is a closed vector subspace of* $\boldsymbol{\mathcal{H}}$ *and that* $\boldsymbol{b}_n \rightharpoonup 0$. *Then* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly to a zero of* $N_{\boldsymbol{D}} + \boldsymbol{B}$.

(ii) *Suppose that* $\sum_{n\in\mathbb{N}} |||\boldsymbol{b}_n||| < +\infty$, *that* $\inf_{n\in\mathbb{N}} \lambda_n > 0$, *and that* $\boldsymbol{B}$ *is uniformly monotone on the bounded subsets of* $\boldsymbol{\mathcal{H}}$. *Then* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges strongly to the unique zero of* $N_{\boldsymbol{D}} + \boldsymbol{B}$.

Proof. Set $\boldsymbol{A} = N_{\boldsymbol{D}}$ and $\boldsymbol{a}_n \equiv 0$ in Theorem 2.1. Then $J_{\gamma\boldsymbol{A}} = P_{\boldsymbol{D}}$ is an affine operator and (8) can therefore be written as (29). Consequently, we can draw the following conclusions.

(i)(a): Theorem 2.1(i)(e) asserts that there exists $\boldsymbol{y} \in \operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B})$ such that $\boldsymbol{p}_n = P_{\boldsymbol{D}}\boldsymbol{z}_n \rightharpoonup \boldsymbol{y}$.

(i)(b): Theorem 2.1(i)(f) asserts that there exists $\boldsymbol{y} \in \operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B})$ such that $\boldsymbol{x}_n = P_{\boldsymbol{D}}\boldsymbol{y}_n \rightharpoonup \boldsymbol{y}$.

(ii): Theorem 2.1(ii)(b) asserts that $\boldsymbol{y}_n \to \boldsymbol{y}$, where $\{\boldsymbol{y}\} = \operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B}) \subset \boldsymbol{D}$. Since $P_{\boldsymbol{D}}$ is continuous and $\boldsymbol{y} \in \boldsymbol{D}$, we conclude that $\boldsymbol{x}_n = P_{\boldsymbol{D}}\boldsymbol{y}_n \to P_{\boldsymbol{D}}\boldsymbol{y} = \boldsymbol{y}$. □

Remark 2.4 Let $\boldsymbol{\mathcal{H}}$ be a real Hilbert space, let $\boldsymbol{D}$ be a closed vector subspace of $\boldsymbol{\mathcal{H}}$, and let $\boldsymbol{B}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}$ be a maximal monotone operator such that $\operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B}) \neq \varnothing$. It follows from Corollary 2.3(i)(b) with $\gamma = 1$, $\lambda_n \equiv 1$, and $\boldsymbol{b}_n \equiv 0$ that a point $\boldsymbol{s} \in \operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B})$ can be constructed by the basic Douglas-Rachford algorithm

$$\boldsymbol{z}_0 \in \boldsymbol{\mathcal{H}} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{ll} \boldsymbol{y}_n & = J_{\boldsymbol{B}}\boldsymbol{z}_n \\ (\boldsymbol{x}_n, \boldsymbol{r}_n) & = (P_{\boldsymbol{D}}\boldsymbol{y}_n, P_{\boldsymbol{D}^\perp}\boldsymbol{z}_n) \\ \boldsymbol{z}_{n+1} & = \boldsymbol{r}_n + 2\boldsymbol{x}_n - \boldsymbol{y}_n. \end{array} \right. \tag{30}$$

On the other hand, (6) yields $\operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B}) = \big\{\boldsymbol{s} \in \boldsymbol{D} \mid (\exists\, \boldsymbol{v} \in \boldsymbol{D}^\perp)\ \boldsymbol{v} \in \boldsymbol{B}\boldsymbol{s}\big\}$. In [48], Spingarn considered the problem

$$\text{find} \quad (\boldsymbol{s}, \boldsymbol{v}) \in \boldsymbol{D} \times \boldsymbol{D}^\perp \quad \text{such that} \quad \boldsymbol{v} \in \boldsymbol{B}\boldsymbol{s} \tag{31}$$

and proposed the "method of partial inverses"

$$(\boldsymbol{s}_0, \boldsymbol{v}_0) \in \boldsymbol{D} \times \boldsymbol{D}^\perp \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \text{find } (\boldsymbol{y}_n, \boldsymbol{u}_n) \in \operatorname{gra} \boldsymbol{B} \text{ such that } \boldsymbol{y}_n + \boldsymbol{u}_n = \boldsymbol{s}_n + \boldsymbol{v}_n \\ (\boldsymbol{s}_{n+1}, \boldsymbol{v}_{n+1}) = (P_{\boldsymbol{D}}\boldsymbol{y}_n, P_{\boldsymbol{D}^\perp}\boldsymbol{u}_n) \end{array} \right. \tag{32}$$

to solve it. Strong connections between (30) and (32) were established in [31, Section 1] (see also [26, Section 5] and [35]).

2.2 Splitting for the sum of maximal monotone operators

The following result concerns an algorithm for finding a zero of the sum of m maximal monotone operators. Its proof revolves around a 2-operator product space reformulation of the original m-operator problem. Such a strategy can be traced back to the work of Pierra [40, 41], who introduced it for solving convex feasibility, best approximation, and constrained optimization problems (see also [14] for its use in inconsistent convex feasibility problems, [5, 12] for its use in Bregman projection algorithms, and [18] for its use in visco-penalization problems). It is also instrumental in the operator splitting method proposed by Spingarn [48] (see also [7, 32]).

Theorem 2.5 *Let* $(B_i)_{1\leq i\leq m}$ *be* $m \geq 2$ *maximal monotone operators from* $\mathcal{H}$ *to* $2^{\mathcal{H}}$*, and set*

$$B = \sum_{i=1}^{m} \omega_i B_i, \quad \text{where} \quad \{\omega_i\}_{1\leq i\leq m} \subset \,]0,1[\quad \text{and} \quad \sum_{i=1}^{m} \omega_i = 1. \tag{33}$$

Let $\gamma \in \,]0,+\infty[$*, let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,2]$*, and, for every* $i \in \{1,\ldots,m\}$*, let* $(b_{i,n})_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}$*. Furthermore, suppose that* $\operatorname{zer} B \neq \varnothing$*, and let* $(x_n)_{n\in\mathbb{N}}$ *and* $(p_n)_{n\in\mathbb{N}}$ *be the sequences generated by the following routine.*

$$
\begin{array}{l}
\text{Initialization}\\
\left\lfloor
\begin{array}{l}
\text{For } i = 1,\ldots,m\\
\left\lfloor\ z_{i,0} \in \mathcal{H}\right.
\end{array}
\right.\\[2mm]
\text{For } n = 0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\text{For } i = 1,\ldots,m\\
\left\lfloor\ y_{i,n} = J_{\gamma B_i} z_{i,n} + b_{i,n}\right.\\
x_n = \sum_{i=1}^{m} \omega_i y_{i,n}\\
p_n = \sum_{i=1}^{m} \omega_i z_{i,n}\\
\text{For } i = 1,\ldots,m\\
\left\lfloor\ z_{i,n+1} = z_{i,n} + \lambda_n\big(2x_n - p_n - y_{i,n}\big).\right.
\end{array}
\right.
\end{array}
\tag{34}
$$

Then the following hold.

(i) *Suppose that* $\max_{1\leq i\leq m} \sum_{n\in\mathbb{N}} \lambda_n \|b_{i,n}\| < +\infty$*, that* $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$*, and that* $(\forall n \in \mathbb{N})$ $\lambda_n < 2$*. Then the following hold.*

   (a) $(p_n)_{n\in\mathbb{N}}$ *converges weakly to a zero of* B*.*

   (b) *Suppose that* $(\forall i \in \{1,,\ldots,m\})$ $b_{i,n} \rightharpoonup 0$*. Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a zero of* B*.*

(ii) *Suppose that* $\max_{1\leq i\leq m} \sum_{n\in\mathbb{N}} \|b_{i,n}\| < +\infty$*, that* $\inf_{n\in\mathbb{N}} \lambda_n > 0$*, and that the operators* $(B_i)_{1\leq i\leq m}$ *are* α*-strongly monotone on* $\mathcal{H}$ *for some* $\alpha \in \,]0,+\infty[$*. Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly to the unique zero of* B*.*

Proof. Let $\boldsymbol{\mathcal{H}}$ be the real Hilbert space obtained by endowing the Cartesian product $\mathcal{H}^m$ with the scalar product $(\boldsymbol{x},\boldsymbol{y}) \mapsto \sum_{i=1}^{m} \omega_i \langle x_i \mid y_i\rangle$, 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

$$|||\cdot|||\colon \boldsymbol{x} \mapsto \sqrt{\sum_{i=1}^{m} \omega_i \|x_i\|^2}. \tag{35}$$

Define

$$\boldsymbol{D} = \big\{(x,\ldots,x) \in \boldsymbol{\mathcal{H}} \mid x \in \mathcal{H}\big\}. \tag{36}$$

In view of (6), we have

$$N_{\boldsymbol{D}} \colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}} \colon \boldsymbol{x} \mapsto \begin{cases} \boldsymbol{D}^{\perp} = \big\{\boldsymbol{u} \in \boldsymbol{\mathcal{H}} \mid \sum_{i=1}^m \omega_i u_i = 0\big\}, & \text{if } \boldsymbol{x} \in \boldsymbol{D};\\ \varnothing, & \text{otherwise.}\end{cases} \tag{37}$$

We also introduce the canonical isometry

$$\boldsymbol{j} \colon \mathcal{H} \to \boldsymbol{D} \colon x \mapsto (x,\ldots,x). \tag{38}$$

Now set

$$\boldsymbol{A} = N_{\boldsymbol{D}} \quad \text{and} \quad \boldsymbol{B} \colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}} \colon \boldsymbol{x} \mapsto \bigtimes_{i=1}^m B_i x_i. \tag{39}$$

It is an easy matter to check that $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximal monotone with resolvents

$$(\forall \boldsymbol{x} \in \boldsymbol{\mathcal{H}}) \quad J_{\gamma \boldsymbol{A}}\boldsymbol{x} = P_{\boldsymbol{D}}\boldsymbol{x} = \boldsymbol{j}\bigg(\sum_{i=1}^m \omega_i x_i\bigg) \quad \text{and} \quad J_{\gamma \boldsymbol{B}}\boldsymbol{x} = \big(J_{\gamma B_i} x_i\big)_{1\leq i\leq m}. \tag{40}$$

Moreover, for every $y \in \mathcal{H}$, (33) and (37) yield $y \in \operatorname{zer} B \Leftrightarrow 0 \in \sum_{i=1}^m \omega_i B_i y \Leftrightarrow (\exists\, (u_i)_{1\leq i\leq m} \in \bigtimes_{i=1}^m B_i y)\ \sum_{i=1}^m \omega_i u_i = 0 \Leftrightarrow (\exists\, \boldsymbol{u} \in \boldsymbol{B}\boldsymbol{j}(y))\ {-\boldsymbol{u}} \in \boldsymbol{D}^{\perp} = N_{\boldsymbol{D}}\, \boldsymbol{j}(y) \Leftrightarrow \boldsymbol{j}(y) \in \operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B}) \subset \boldsymbol{D}$. Thus,

$$\boldsymbol{j}(\operatorname{zer} B) = \operatorname{zer}(N_{\boldsymbol{D}} + \boldsymbol{B}). \tag{41}$$

Now set $(\forall n \in \mathbb{N})$ $\boldsymbol{z}_n = (z_{i,n})_{1\leq i\leq m}$, $\boldsymbol{y}_n = (y_{i,n})_{1\leq i\leq m}$, $\boldsymbol{b}_n = (b_{i,n})_{1\leq i\leq m}$, $\boldsymbol{x}_n = \boldsymbol{j}(x_n)$, and $\boldsymbol{p}_n = \boldsymbol{j}(p_n)$. Then it follows from (34) and (40) that the sequences thus defined are precisely those appearing in (29).

(i): In view of (35),

$$\sum_{n\in\mathbb{N}} \lambda_n |||\boldsymbol{b}_n||| = \sum_{n\in\mathbb{N}} \lambda_n \sqrt{\sum_{i=1}^m \omega_i \|b_{i,n}\|^2} \leq \sum_{i=1}^m \sum_{n\in\mathbb{N}} \lambda_n \|b_{i,n}\| < +\infty. \tag{42}$$

(i)(a): Corollary 2.3(i)(a) and (41) imply that $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ converges weakly to a point $\boldsymbol{j}(y)$, where $y \in \operatorname{zer} B$. Hence, $p_n = \boldsymbol{j}^{-1}(\boldsymbol{p}_n) \rightharpoonup y$.

(i)(b): The assumptions imply that $\boldsymbol{b}_n \rightharpoonup 0$. Hence, it results from Corollary 2.3(i)(b) and (41) that $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly to a point $\boldsymbol{j}(y)$, where $y \in \operatorname{zer} B$. Hence, $x_n = \boldsymbol{j}^{-1}(\boldsymbol{x}_n) \rightharpoonup y$.

(ii): By assumption, $\boldsymbol{B}$ is α-strongly monotone, hence uniformly monotone, on $\boldsymbol{\mathcal{H}}$. On the other hand, proceeding as in (42), we obtain $\sum_{n\in\mathbb{N}} |||\boldsymbol{b}_n||| < +\infty$. Hence, Corollary 2.3(ii) and (41) imply that $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges strongly to $\boldsymbol{j}(y)$, where $\{y\} = \operatorname{zer} B$. Thus, $x_n = \boldsymbol{j}^{-1}(\boldsymbol{x}_n) \to y$. □

We have obtained Theorem 2.5 as a corollary to Theorem 2.1 on the asymptotic behavior of the Douglas-Rachford algorithm. In [48, 49], Spingarn proposed a splitting method for m monotone operators based on a product space implementation of the method of partial inverses (32). Since, as mentioned in Remark 2.4, connections exist between the Douglas-Rachford algorithm and (32), we naturally obtain a connection between the product space transpositions of these algorithms. In the next corollary, we exploit this connection to derive from Theorem 2.5(i)(b) the convergence of Spingarn's m-operator splitting method (see also [7, Section 4] for the product space behavior of the Douglas-Rachford algorithm in the special case of convex feasibility problems and its connection to Spingarn's parallel projection method [48, Section 6], [50]).

Corollary 2.6 [48, Corollary 5.1(i)] *Let $(B_i)_{1\le i\le m}$ be $m\ge 2$ maximal monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, and set*

$$B=\sum_{i=1}^m \omega_i B_i, \quad \textit{where} \quad \{\omega_i\}_{1\le i\le m}\subset\,]0,1[\quad \textit{and} \quad \sum_{i=1}^m \omega_i=1. \tag{43}$$

Suppose that $\operatorname{zer} B\neq\varnothing$ *and let $(s_n)_{n\in\mathbb{N}}$ be the sequence generated by the following routine.*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l} s_0\in\mathcal{H}\\ (v_{i,0})_{1\le i\le m}\in\mathcal{H}^m \ \textit{satisfy}\ \sum_{i=1}^m \omega_i v_{i,0}=0\end{array}\right.\\
\text{For } n=0,1,\ldots\\
\left\lfloor\begin{array}{l}
\text{For } i=1,\ldots,m\\
\left\lfloor\begin{array}{l}\text{find } (y_{i,n},u_{i,n})\in\operatorname{gra}B_i \text{ such that } y_{i,n}+u_{i,n}=s_n+v_{i,n}\end{array}\right.\\
s_{n+1}=\sum_{i=1}^m \omega_i y_{i,n}\\
q_n\ \ \ =\sum_{i=1}^m \omega_i u_{i,n}\\
\text{For } i=1,\ldots,m\\
\left\lfloor\begin{array}{l} v_{i,n+1}=u_{i,n}-q_n.\end{array}\right.
\end{array}\right.
\end{array} \tag{44}$$

Then $(s_n)_{n\in\mathbb{N}}$ converges weakly to a zero of B.

Proof. Fix temporarily $n\in\mathbb{N}$. For every $i\in\{1,\ldots,m\}$, the conditions defining $(y_{i,n},u_{i,n})$ in (44) can be expressed as $s_n+v_{i,n}-y_{i,n}\in B_i y_{i,n}$ and $u_{i,n}=s_n+v_{i,n}-y_{i,n}$, that is, $y_{i,n}=J_{B_i}(s_n+v_{i,n})$ and $u_{i,n}=s_n+v_{i,n}-y_{i,n}$. Now set $(\forall i\in\{1,\ldots,m\})$ $z_{i,n}=s_n+v_{i,n}$ and $x_n=s_{n+1}$. Upon eliminating the variables $(v_{i,n})_{1\le i\le m}$, the loop on n in (44) can be rewritten as

$$\left\lfloor\begin{array}{l}
\text{For } i=1,\ldots,m\\
\left\lfloor\begin{array}{l} y_{i,n}=J_{B_i}z_{i,n}\\ u_{i,n}=z_{i,n}-y_{i,n}\end{array}\right.\\
x_n=\sum_{i=1}^m \omega_i y_{i,n}\\
q_n=\sum_{i=1}^m \omega_i u_{i,n}\\
\text{For } i=1,\ldots,m\\
\left\lfloor\begin{array}{l} z_{i,n+1}-x_n=u_{i,n}-q_n.\end{array}\right.
\end{array}\right. \tag{45}$$

Now set $p_n=\sum_{i=1}^m \omega_i z_{i,n}$. Then $q_n=\sum_{i=1}^m \omega_i z_{i,n}-\sum_{i=1}^m \omega_i y_{i,n}=p_n-x_n$ and hence, for every $i\in\{1,\ldots,m\}$, $u_{i,n}-q_n=z_{i,n}-y_{i,n}-p_n+x_n$. Therefore, upon eliminating $(u_{i,n})_{1\le i\le m}$, u_n, and

q_n, an introducing p_n, (45) can be reduced to

$$\left\lfloor \begin{array}{l} \text{For } i = 1, \ldots, m \\ \left\lfloor \; y_{i,n} = J_{B_i} z_{i,n} \right. \\ x_n = \sum_{i=1}^m \omega_i y_{i,n} \\ p_n = \sum_{i=1}^m \omega_i z_{i,n} \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \; z_{i,n+1} = z_{i,n} + 2x_n - p_n - y_{i,n}, \right. \end{array} \right. \tag{46}$$

which coincides with the loop on n in (34) in the special case when it is implemented with $\gamma = 1$, $\lambda_n = 1$, and $(\forall i \in \{1, \ldots, m\})\ b_{i,n} = 0$. Since Theorem 2.5(i)(b) asserts that in this case $(x_n)_{n\in\mathbb{N}}$ converges weakly to a zero of B, so does $(s_n)_{n\in\mathbb{N}}$. □

Remark 2.7

- Another angle on the problem of finding a zero of the sum of m maximal monotone operators is the ergodic method proposed by Passty [39]. This approach, however, requires that the sum be itself maximal monotone, which imposes additional restrictions; see [2] and [47, Section 32]. In addition, it involves finely tuned vanishing parameters, which leads to numerical instabilities (see also [33]).
- In the case when the operators $(B_i)_{1\leq i\leq m}$ are subdifferentials, applications of Theorem 2.5(i)(a) in the area of inverse problems can be found in [21].

2.3 Splitting for the resolvent of the sum of maximal monotone operators

In this section, we apply Theorem 2.5(ii) to our initial problem of devising a strongly convergent splitting method for computing the resolvent of a sum of maximal monotone operators.

Theorem 2.8 *Let $(A_i)_{1\leq i\leq m}$ be $m \geq 2$ maximal monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, and set*

$$A = \sum_{i=1}^m \omega_i A_i, \quad \textit{where} \quad \{\omega_i\}_{1\leq i\leq m} \subset \left]0,1\right[\quad \textit{and} \quad \sum_{i=1}^m \omega_i = 1. \tag{47}$$

Let $\gamma \in \left]0,+\infty\right[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\left]0,2\right]$ such that $\inf_{n\in\mathbb{N}} \lambda_n > 0$, and, for every $i \in \{1, \ldots, m\}$, let $(a_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|a_{i,n}\| < +\infty$. Furthermore, let $r \in$

ran(Id $+A$) *and let* $(x_n)_{n\in\mathbb{N}}$ *be the sequence generated by the following routine.*

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor \begin{array}{l}
\text{For } i = 1, \ldots, m \\
\left\lfloor\ z_{i,0} \in \mathcal{H} \right.
\end{array}\right. \\[2mm]
\text{For } n = 0, 1, \ldots \\
\left\lfloor \begin{array}{l}
\text{For } i = 1, \ldots, m \\
\left\lfloor\ y_{i,n} = J_{\frac{\gamma}{\gamma+1}A_i}\left(\dfrac{z_{i,n}+\gamma r}{\gamma+1}\right) + a_{i,n} \right. \\
x_n = \sum_{i=1}^m \omega_i y_{i,n} \\
p_n = \sum_{i=1}^m \omega_i z_{i,n} \\
\text{For } i = 1, \ldots, m \\
\left\lfloor\ z_{i,n+1} = z_{i,n} + \lambda_n\big(2x_n - p_n - y_{i,n}\big). \right.
\end{array}\right.
\end{array}
\tag{48}
$$

Then $x_n \to J_A\, r$.

Proof. Set

$$
(\forall i \in \{1, \ldots, m\}) \quad B_i \colon \mathcal{H} \to 2^{\mathcal{H}} \colon y \mapsto -r + y + A_i y \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad b_{i,n} = a_{i,n}. \tag{49}
$$

The operators $(B_i)_{1\le i\le m}$ are maximal monotone and 1-strongly monotone. In addition,

$$
\begin{aligned}
(\forall i \in \{1, \ldots, m\})(\forall y \in \mathcal{H})(\forall z \in \mathcal{H}) \quad y = J_{\gamma B_i} z \quad &\Leftrightarrow \quad z \in y + \gamma B_i y \\
&\Leftrightarrow \quad z + \gamma r \in (\gamma+1) y + \gamma A_i y \\
&\Leftrightarrow \quad y = J_{\frac{\gamma}{\gamma+1}A_i}\big((z+\gamma r)/(\gamma+1)\big).
\end{aligned} \tag{50}
$$

Thus, (48) coincides with (34). Now set $B = \sum_{i=1}^m \omega_i B_i$. Then (47) and (49) yield $B = -r + \mathrm{Id} + A$ and, since $r \in \operatorname{ran}(\mathrm{Id} + A)$, we obtain $\operatorname{zer} B = \{J_A\, r\}$. Appealing to Theorem 2.5(ii), we conclude that $x_n \to J_A\, r$. ☐

3 Dykstra-like splitting for the resolvent of the sum

In [6, Theorem 2.4], Dykstra's method for computing the projection onto the intersection of two closed convex sets [9, 23, 25] was extended to a method for computing the resolvent of the sum of two maximal monotone operators. In Proposition 3.2, we establish the convergence of an inexact version of this method. This result is then used to obtain Theorem 3.3, where we introduce an alternative splitting method for computing the resolvent of $m \ge 2$ maximal monotone operators. The following fact will be needed.

Lemma 3.1 *Let* $(\boldsymbol{\mathcal{H}}, |||\cdot|||)$ *be a real Hilbert space, let* T_1 *and* T_2 *be firmly nonexpansive operators from* $\boldsymbol{\mathcal{H}}$ *to* $\boldsymbol{\mathcal{H}}$ *such that* $\mathrm{Fix}(T_1 \circ T_2) \neq \varnothing$, *and let* $(\boldsymbol{e}_{1,n})$ *and* $(\boldsymbol{e}_{2,n})$ *be sequences in* $\boldsymbol{\mathcal{H}}$ *such that*

$\sum_{n\in\mathbb{N}} |||\boldsymbol{e}_{1,n}||| < +\infty$ *and* $\sum_{n\in\mathbb{N}} |||\boldsymbol{e}_{2,n}||| < +\infty$. *Furthermore, let* $(\boldsymbol{u}_n)_{n\in\mathbb{N}}$ *be the sequence resulting from the iteration*

$$\boldsymbol{u}_0 \in \boldsymbol{\mathcal{H}} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \boldsymbol{u}_{n+1} = T_1\big(T_2\boldsymbol{u}_n + \boldsymbol{e}_{2,n}\big) + \boldsymbol{e}_{1,n}. \tag{51}$$

Then there exists $\boldsymbol{u} \in \operatorname{Fix}(T_1 \circ T_2)$ *such that* $\boldsymbol{u_n} \rightharpoonup \boldsymbol{u}$. *Moreover,* $T_2\boldsymbol{u}_n - \boldsymbol{u}_n \to T_2\boldsymbol{u} - \boldsymbol{u}$.

Proof. See the statement and the proof of [36, Théorème 5.5.2] or, from a more general perspective, those of [17, Corollary 7.1]. □

Proposition 3.2 *Let* $(\boldsymbol{\mathcal{H}}, |||\cdot|||)$ *be a real Hilbert space, let* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *be maximal monotone operators from* $\boldsymbol{\mathcal{H}}$ *to* $2^{\boldsymbol{\mathcal{H}}}$, *and let* $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ *and* $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ *be sequences in* $\boldsymbol{\mathcal{H}}$ *such that*

$$\sum_{n\in\mathbb{N}} |||\boldsymbol{a}_n||| < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} |||\boldsymbol{b}_n||| < +\infty. \tag{52}$$

Furthermore, let $\boldsymbol{r} \in \operatorname{ran}(\mathbf{Id} + \boldsymbol{A} + \boldsymbol{B})$ *and let* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *be the sequence generated by the following routine.*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}
\boldsymbol{y}_0 = \boldsymbol{r}\\
\boldsymbol{q}_0 = 0\\
\boldsymbol{p}_0 = 0
\end{array}\right.\\
\text{For } n = 0, 1, \ldots\\
\left\lfloor\begin{array}{l}
\boldsymbol{x}_n \quad = J_{\boldsymbol{B}}(\boldsymbol{y}_n + \boldsymbol{q}_n) + \boldsymbol{b}_n\\
\boldsymbol{q}_{n+1} = \boldsymbol{y}_n + \boldsymbol{q}_n - \boldsymbol{x}_n\\
\boldsymbol{y}_{n+1} = J_{\boldsymbol{A}}(\boldsymbol{x}_n + \boldsymbol{p}_n) + \boldsymbol{a}_n\\
\boldsymbol{p}_{n+1} = \boldsymbol{x}_n + \boldsymbol{p}_n - \boldsymbol{y}_{n+1}.
\end{array}\right.
\end{array} \tag{53}$$

Then $\boldsymbol{x}_n \to J_{\boldsymbol{A}+\boldsymbol{B}}\,\boldsymbol{r}$.

Proof. The first part of the proof is closely patterned after that of [6, Theorem 2.4], where $(\forall n \in \mathbb{N})$ $\boldsymbol{a}_n = 0$ and $\boldsymbol{b}_n = 0$. We first derive from (53) that $(\forall n \in \mathbb{N})$ $(\boldsymbol{q}_{n+1} + \boldsymbol{x}_n) + \boldsymbol{p}_n = \boldsymbol{y}_n + \boldsymbol{q}_n + \boldsymbol{p}_n$. On the other hand, a simple induction argument yields $(\forall n \in \mathbb{N})$ $\boldsymbol{q}_n + \boldsymbol{p}_n = \boldsymbol{r} - \boldsymbol{y}_n$. Thus,

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{r} = \boldsymbol{y}_n + \boldsymbol{q}_n + \boldsymbol{p}_n = \boldsymbol{q}_{n+1} + \boldsymbol{p}_n + \boldsymbol{x}_n, \tag{54}$$

so that (53) can be rewritten as

$$\left\lfloor\begin{array}{l}
\boldsymbol{y}_0 = \boldsymbol{r}\\
\boldsymbol{q}_0 = 0\\
\boldsymbol{p}_0 = 0
\end{array}\right. \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor\begin{array}{l}
\boldsymbol{x}_n \quad = J_{\boldsymbol{B}}(\boldsymbol{r} - \boldsymbol{p}_n) + \boldsymbol{b}_n\\
\boldsymbol{q}_{n+1} = \boldsymbol{r} - \boldsymbol{p}_n - \boldsymbol{x}_n\\
\boldsymbol{y}_{n+1} = J_{\boldsymbol{A}}(\boldsymbol{r} - \boldsymbol{q}_{n+1}) + \boldsymbol{a}_n\\
\boldsymbol{p}_{n+1} = \boldsymbol{r} - \boldsymbol{q}_{n+1} - \boldsymbol{y}_{n+1}.
\end{array}\right. \tag{55}$$

Now set $\boldsymbol{u}_0 = -\boldsymbol{r}$ and $(\forall n \in \mathbb{N})$ $\boldsymbol{u}_n = \boldsymbol{p}_n - \boldsymbol{r}$ and $\boldsymbol{v}_n = -\boldsymbol{q}_{n+1}$. Then it follows from (54) that

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{v}_n - \boldsymbol{u}_n = \boldsymbol{x}_n \quad \text{and} \quad \boldsymbol{v}_n - \boldsymbol{u}_{n+1} = \boldsymbol{y}_{n+1}, \tag{56}$$

and that, in conjunction with (55),

$$(\forall n \in \mathbb{N}) \quad \left\lfloor\begin{array}{l}
\boldsymbol{v}_n = \boldsymbol{p}_n - \boldsymbol{r} + \boldsymbol{x}_n = \boldsymbol{u}_n + J_{\boldsymbol{B}}(-\boldsymbol{u}_n) + \boldsymbol{b}_n\\
\boldsymbol{u}_{n+1} = \boldsymbol{p}_{n+1} - \boldsymbol{r} = -\boldsymbol{q}_{n+1} - \boldsymbol{y}_{n+1} = \boldsymbol{v}_n - J_{\boldsymbol{A}}(\boldsymbol{v}_n + \boldsymbol{r}) - \boldsymbol{a}_n.
\end{array}\right. \tag{57}$$

Now set $\boldsymbol{C}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}\colon \boldsymbol{v} \mapsto \boldsymbol{A}^{-1}(\boldsymbol{v}+\boldsymbol{r})$ and $\boldsymbol{D} = \boldsymbol{B}^{\sim}$, where we use the notation $\boldsymbol{B}^{\sim} = (-\mathbf{Id}) \circ \boldsymbol{B}^{-1} \circ (-\mathbf{Id})$. Then $\boldsymbol{C}$ and $\boldsymbol{D}$ are maximal monotone, and

$$\boldsymbol{C}^{-1} = -\boldsymbol{r} + \boldsymbol{A}, \quad \boldsymbol{D}^{\sim} = \boldsymbol{B}, \quad J_{\boldsymbol{C}} = \mathbf{Id} - J_{\boldsymbol{A}}(\cdot + \boldsymbol{r}), \quad \text{and} \quad J_{\boldsymbol{D}} = \mathbf{Id} + \big(J_{\boldsymbol{B}} \circ (-\mathbf{Id})\big). \tag{58}$$

Thus, the sequence $(\boldsymbol{u}_n)_{n\in\mathbb{N}}$ is generated by the algorithm

$$\boldsymbol{u}_0 = -\boldsymbol{r} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{ll} \boldsymbol{v}_n & = J_{\boldsymbol{D}}\boldsymbol{u}_n + \boldsymbol{b}_n \\ \boldsymbol{u}_{n+1} & = J_{\boldsymbol{C}}\boldsymbol{v}_n - \boldsymbol{a}_n. \end{array} \right. \tag{59}$$

To complete the proof, we invoke successively (58), [8, Equation (8)], [8, Fact 2.1], and [8, Proposition 3.2(i)] to get

$$\begin{aligned} \boldsymbol{r} \in \operatorname{ran}(\mathbf{Id} + \boldsymbol{A} + \boldsymbol{B}) \quad &\Leftrightarrow \quad \operatorname{zer}(-\boldsymbol{r} + \boldsymbol{A} + \mathbf{Id} + \boldsymbol{B}) \neq \varnothing \\ &\Leftrightarrow \quad \operatorname{zer}(\boldsymbol{C}^{-1} + \mathbf{Id} + \boldsymbol{D}^{\sim}) \neq \varnothing \\ &\Leftrightarrow \quad \operatorname{zer}(\boldsymbol{C}^{-1} + (\mathbf{Id} - J_{\boldsymbol{D}})^{\sim}) \neq \varnothing \\ &\Leftrightarrow \quad \operatorname{zer}(\boldsymbol{C} + \mathbf{Id} - J_{\boldsymbol{D}}) \neq \varnothing \\ &\Leftrightarrow \quad \operatorname{Fix}(J_{\boldsymbol{C}} \circ J_{\boldsymbol{D}}) \neq \varnothing. \end{aligned} \tag{60}$$

Hence, since $J_{\boldsymbol{C}}$ and $J_{\boldsymbol{D}}$ are firmly nonexpansive, we derive from (59), (52), (56), and Lemma 3.1 that there exists $\boldsymbol{u} \in \operatorname{Fix}(J_{\boldsymbol{C}} \circ J_{\boldsymbol{D}})$ such that $\boldsymbol{x}_n = \boldsymbol{v}_n - \boldsymbol{u}_n = \boldsymbol{b}_n + J_{\boldsymbol{D}}\boldsymbol{u}_n - \boldsymbol{u}_n \to J_{\boldsymbol{D}}\boldsymbol{u} - \boldsymbol{u}$. However, since [8, Proposition 3.2] asserts that $J_{\boldsymbol{D}}\boldsymbol{u} - \boldsymbol{u} = J_{\boldsymbol{C}^{-1}+\boldsymbol{D}^{\sim}}\, 0 = J_{\boldsymbol{A}+\boldsymbol{B}}\, \boldsymbol{r}$, the proof is complete. ☐

By transcribing the above result in a product space, we obtain a parallel splitting method for computing the resolvent of the weighted sum of an arbitrary number of operators.

Theorem 3.3 *Let $(A_i)_{1\leq i\leq m}$ be $m \geq 2$ maximal monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, and set*

$$A = \sum_{i=1}^{m} \omega_i A_i, \quad \textit{where} \quad \{\omega_i\}_{1\leq i\leq m} \subset \,]0,1[\quad \textit{and} \quad \sum_{i=1}^{m} \omega_i = 1. \tag{61}$$

For every $i \in \{1,\ldots,m\}$, let $(a_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|a_{i,n}\| < +\infty$. Furthermore, let $r \in \operatorname{ran}(\operatorname{Id} + A)$ and let $(x_n)_{n\in\mathbb{N}}$ be the sequence generated by the following routine.

$$\begin{array}{l} \text{Initialization} \\ \left\lfloor \begin{array}{l} x_0 = r \\ \text{For } i = 1,\ldots,m \\ \lfloor \; z_{i,0} = x_0 \end{array} \right. \\ \\ \text{For } n = 0,1,\ldots \\ \left\lfloor \begin{array}{l} \text{For } i = 1,\ldots,m \\ \lfloor \; y_{i,n} = J_{A_i} z_{i,n} + a_{i,n} \\ x_{n+1} = \sum_{i=1}^{m} \omega_i y_{i,n} \\ \text{For } i = 1,\ldots,m \\ \lfloor \; z_{i,n+1} = x_{n+1} + z_{i,n} - y_{i,n}. \end{array} \right. \end{array} \tag{62}$$

Then $x_n \to J_A\, r$.

Proof. Let $\boldsymbol{\mathcal{H}}$ be as in the proof of Theorem 2.5, let $\boldsymbol{D}$ be as in (36), and let $\boldsymbol{j}$ be as in (38). Set

$$\boldsymbol{A}\colon \boldsymbol{\mathcal{H}} \to 2^{\boldsymbol{\mathcal{H}}}\colon \boldsymbol{x} \mapsto \bigtimes_{i=1}^{m} A_i x_i \quad \text{and} \quad \boldsymbol{B} = N_{\boldsymbol{D}}. \tag{63}$$

As in (40), (35) yields

$$J_{\boldsymbol{A}}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}}\colon \boldsymbol{x} \mapsto \big(J_{A_i} x_i\big)_{1\leq i\leq m} \quad \text{and} \quad J_{\boldsymbol{B}} = P_{\boldsymbol{D}}\colon \boldsymbol{\mathcal{H}} \to \boldsymbol{D}\colon \boldsymbol{x} \mapsto \boldsymbol{j}\bigg(\sum_{i=1}^{m} \omega_i x_i\bigg). \tag{64}$$

Since by assumption $r \in \operatorname{ran}(\operatorname{Id} + A)$, $J_A\, r$ is well defined. Moreover, we derive from (61), (63), and (37) that, for every $x \in \mathcal{H}$,

$$\begin{aligned}
x = J_A\, r \quad &\Leftrightarrow \quad r - x \in Ax = \sum_{i=1}^{m} \omega_i A_i x\\
&\Leftrightarrow \quad \Big(\exists\, (u_i)_{1\leq i\leq m} \in \bigtimes_{i=1}^{m} A_i x\Big) \quad \sum_{i=1}^{m} \omega_i (r - x - u_i) = 0\\
&\Leftrightarrow \quad (\exists\, \boldsymbol{u} \in \boldsymbol{A}\boldsymbol{j}(x)) \quad \boldsymbol{j}(r) - \boldsymbol{j}(x) - \boldsymbol{u} \in \boldsymbol{D}^{\perp}\\
&\Leftrightarrow \quad \boldsymbol{j}(r) - \boldsymbol{j}(x) \in \boldsymbol{A}\boldsymbol{j}(x) + \boldsymbol{B}\boldsymbol{j}(x)\\
&\Leftrightarrow \quad \boldsymbol{j}(x) = J_{\boldsymbol{A}+\boldsymbol{B}}\, \boldsymbol{j}(r).
\end{aligned} \tag{65}$$

This shows that

$$\boldsymbol{j}(J_A\, r) = J_{\boldsymbol{A}+\boldsymbol{B}}\, \boldsymbol{j}(r). \tag{66}$$

To construct $J_{\boldsymbol{A}+\boldsymbol{B}}\, \boldsymbol{j}(r)$, we can use Proposition 3.2. Let $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{q}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ be the sequences generated by algorithm (53), where we set $\boldsymbol{r} = \boldsymbol{j}(r)$ and $(\forall n \in \mathbb{N})$ $\boldsymbol{b}_n = 0$. Proposition 3.2 asserts that

$$\sum_{n\in\mathbb{N}} |||\boldsymbol{a}_n||| < +\infty \quad \Rightarrow \quad \boldsymbol{x}_n \to J_{\boldsymbol{A}+\boldsymbol{B}}\, \boldsymbol{j}(r). \tag{67}$$

On the other hand, since $J_{\boldsymbol{B}} = P_{\boldsymbol{D}}$, it follows from (53) that, for every $n \in \mathbb{N}$, $\boldsymbol{q}_n \in \boldsymbol{D}^{\perp}$ and therefore $\boldsymbol{x}_n = P_{\boldsymbol{D}}(\boldsymbol{y}_n + \boldsymbol{q}_n) = P_{\boldsymbol{D}}\boldsymbol{y}_n$. Thus, the sequence $(\boldsymbol{q}_n)_{n\in\mathbb{N}}$ plays no role in (53), which can therefore be simplified to

$$\left\lfloor \begin{array}{l} \boldsymbol{y}_0 = \boldsymbol{j}(r)\\ \boldsymbol{p}_0 = 0 \end{array}\right. \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{ll} \boldsymbol{x}_n & = P_{\boldsymbol{D}}\, \boldsymbol{y}_n\\ \boldsymbol{y}_{n+1} & = J_{\boldsymbol{A}}(\boldsymbol{x}_n + \boldsymbol{p}_n) + \boldsymbol{a}_n\\ \boldsymbol{p}_{n+1} & = \boldsymbol{x}_n + \boldsymbol{p}_n - \boldsymbol{y}_{n+1}. \end{array}\right. \tag{68}$$

After reordering the computations, we can rewrite (68) as

$$\left\lfloor \begin{array}{l} \boldsymbol{x}_0 = \boldsymbol{j}(r)\\ \boldsymbol{p}_0 = 0 \end{array}\right. \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{ll} \boldsymbol{y}_n & = J_{\boldsymbol{A}}(\boldsymbol{x}_n + \boldsymbol{p}_n) + \boldsymbol{a}_n\\ \boldsymbol{p}_{n+1} & = \boldsymbol{x}_n + \boldsymbol{p}_n - \boldsymbol{y}_n\\ \boldsymbol{x}_{n+1} & = P_{\boldsymbol{D}}\, \boldsymbol{y}_n. \end{array}\right. \tag{69}$$

To further simplify the algorithm, let us set $(\forall n \in \mathbb{N})$ $\boldsymbol{z}_n = \boldsymbol{x}_n + \boldsymbol{p}_n$. Then (69) becomes

$$\left\lfloor \begin{array}{l} \boldsymbol{x}_0 = \boldsymbol{j}(r)\\ \boldsymbol{z}_0 = \boldsymbol{x}_0 \end{array}\right. \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{ll} \boldsymbol{y}_n & = J_{\boldsymbol{A}}\boldsymbol{z}_n + \boldsymbol{a}_n\\ \boldsymbol{x}_{n+1} & = P_{\boldsymbol{D}}\, \boldsymbol{y}_n\\ \boldsymbol{z}_{n+1} & = \boldsymbol{x}_{n+1} + \boldsymbol{z}_n - \boldsymbol{y}_n. \end{array}\right. \tag{70}$$

In view of (70), (64), and (62), we can write $(\forall n \in \mathbb{N})$ $\boldsymbol{x}_n = \boldsymbol{j}(x_n)$, $\boldsymbol{a}_n = (a_{i,n})_{1\leq i\leq m}$, $\boldsymbol{y}_n = (y_{i,n})_{1\leq i\leq m}$, and $\boldsymbol{z}_n = (z_{i,n})_{1\leq i\leq m}$. Moreover, since

$$\sum_{n\in\mathbb{N}} |||\boldsymbol{a}_n||| = \sum_{n\in\mathbb{N}} \sqrt{\sum_{i=1}^m \omega_i \|a_{i,n}\|^2} \leq \sum_{i=1}^m \sum_{n\in\mathbb{N}} \|a_{i,n}\| < +\infty, \tag{71}$$

(67) and (66) yield

$$x_n = \boldsymbol{j}^{-1}(\boldsymbol{x}_n) \to \boldsymbol{j}^{-1}(J_{\boldsymbol{A}+\boldsymbol{B}}\, \boldsymbol{j}(r)) = J_A\, r, \tag{72}$$

which completes the proof. □

Remark 3.4 Theorems 2.8 and 3.3 provide two strongly convergent iterative methods for constructing the resolvent of a weighted sum of maximal monotone operators at a given point. Algorithms (48) and (62) share similar structures, computational costs, and storage requirements. At iteration n, they both involve a parallel step at which the resolvents of the operators $(A_i)_{1\leq i\leq m}$ are evaluated individually (and possibly simultaneously) with some tolerances $(a_{i,n})_{1\leq i\leq m}$. This step is followed by a coordination step at which the resolvents are averaged. The last step is a parallel step at which the auxiliary variables $(z_{i,n})_{1\leq i\leq m}$ are updated. In terms of convergence speed, the behavior of the algorithms will be compared through numerical experiments in future work.

4 Dykstra-like splitting for the proximity operator of the sum

We denote by $\operatorname{dom} f = \big\{x \in \mathcal{H} \mid f(x) < +\infty\big\}$ the domain of a function $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$, and by $\Gamma_0(\mathcal{H})$ the class of lower semicontinuous convex functions from $\mathcal{H}$ to $\left]-\infty,+\infty\right]$ with nonempty domain. Moreau [37] observed that, if $f \in \Gamma_0(\mathcal{H})$ then, for every $r \in \mathcal{H}$, the function $f + \|r - \cdot\|^2/2$ admits a unique minimizer, which he denoted by $\operatorname{prox}_f r$, i.e.,

$$\operatorname{prox}_f r = \underset{y\in\mathcal{H}}{\operatorname{argmin}}\; f(y) + \frac{1}{2}\|r - y\|^2. \tag{73}$$

Alternatively, the proximity operator of f thus defined can be expressed as [38]

$$\operatorname{prox}_f = J_{\partial f}, \tag{74}$$

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 maximal monotone operator [47, Theorem 18.7].

Let $f \in \Gamma_0(\mathcal{H})$ and let $r \in \mathcal{H}$. In this section, we address the problem of computing $\operatorname{prox}_f r$ when f can be decomposed into a weighted sum of functions $(f_i)_{1\leq i\leq m}$ in $\Gamma_0(\mathcal{H})$, say

$$f = \sum_{i=1}^m \omega_i f_i, \quad \text{where} \quad \{\omega_i\}_{1\leq i\leq m} \subset \left]0,1\right[\quad \text{and} \quad \sum_{i=1}^m \omega_i = 1, \tag{75}$$

for which the proximity operators $(\operatorname{prox}_{f_i})_{1\leq i\leq m}$ can be implemented easily (we refer to [13, 19, 22, 38] for the closed-form expressions of a variety of proximity operators). In this context, $\operatorname{prox}_f r$ is simply the solution to the strongly convex program

$$\underset{y\in\mathcal{H}}{\operatorname{minimize}}\;\; \sum_{i=1}^m \omega_i f_i(y) + \frac{1}{2}\|r - y\|^2. \tag{76}$$

For instance, such formulations arise naturally in the area of signal denoising, where $r = x + w$ is an observation of an ideal signal $x \in \mathcal{H}$ which is corrupted by a realization w of a noise process. The quadratic term $\|r - \cdot\|^2/2$ promotes a least-squares data fit, while the potentials $(f_i)_{1\le i\le m}$ model various priors on the original signal x, e.g., [20, 22, 46, 51]. The state-of-the art in such applications is limited to at most two nonsmooth potentials. By contrast, the results of this section provide a strongly convergent splitting algorithm that can handle $m > 2$ nonsmooth potentials.

In view of (74), a first approach to construct $\operatorname{prox}_f r = J_{\partial f}\, r$ is to make the additional assumption that $\sum_{i=1}^m \omega_i \partial f_i$ is maximal monotone, i.e.,

$$\partial\bigg(\sum_{i=1}^m \omega_i f_i\bigg) = \sum_{i=1}^m \omega_i \partial f_i. \tag{77}$$

In this case, we apply Theorem 2.8 or Theorem 3.3 with, for every $i \in \{1,\ldots,m\}$, $A_i = \partial f_i$, which amounts to replacing each resolvent by the corresponding proximity operator in algorithms (48) and (62) (Passty's method [39] is also applicable in this case but, as discussed in Remark 2.7, it has numerical limitations). A shortcoming of this approach is of course that (77) does not come for free and requires that so-called qualification conditions be imposed; see [1] and [47, Section 18]. We adopt a different strategy, which will lead to a qualification-free method. Our starting point is the following result of [6] on the proximity operator of the sum of two functions, which itself relies on Fenchel duality arguments developed in [8].

Proposition 4.1 [6, Theorem 3.3] *Let $\boldsymbol{\mathcal{H}}$ be a real Hilbert space, and let $\boldsymbol{f}$ and $\boldsymbol{g}$ be functions in $\Gamma_0(\boldsymbol{\mathcal{H}})$ such that $\operatorname{dom}\boldsymbol{f} \cap \operatorname{dom}\boldsymbol{g} \neq \varnothing$. Furthermore, let $\boldsymbol{r} \in \boldsymbol{\mathcal{H}}$ and let $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ be the sequence generated by the following routine.*

$$\begin{array}{l}
\text{Initialization}\\
\left\lfloor\begin{array}{l}
\boldsymbol{y}_0 = \boldsymbol{r}\\
\boldsymbol{q}_0 = 0\\
\boldsymbol{p}_0 = 0
\end{array}\right.\\
\text{For } n = 0,1,\ldots\\
\left\lfloor\begin{array}{l}
\boldsymbol{x}_n \;\;\; = \operatorname{prox}_{\boldsymbol{g}}(\boldsymbol{y}_n + \boldsymbol{q}_n)\\
\boldsymbol{q}_{n+1} = \boldsymbol{y}_n + \boldsymbol{q}_n - \boldsymbol{x}_n\\
\boldsymbol{y}_{n+1} = \operatorname{prox}_{\boldsymbol{f}}(\boldsymbol{x}_n + \boldsymbol{p}_n)\\
\boldsymbol{p}_{n+1} = \boldsymbol{x}_n + \boldsymbol{p}_n - \boldsymbol{y}_{n+1}.
\end{array}\right.
\end{array} \tag{78}$$

Then $\boldsymbol{x}_n \to \operatorname{prox}_{\boldsymbol{f}+\boldsymbol{g}}\, \boldsymbol{r}$.

The main result of this section is the following.

Theorem 4.2 *Let $(f_i)_{1\le i\le m}$ be $m \ge 2$ functions in $\Gamma_0(\mathcal{H})$ such that*

$$\bigcap_{i=1}^m \operatorname{dom} f_i \neq \varnothing, \tag{79}$$

and set

$$f = \sum_{i=1}^m \omega_i f_i, \quad \textit{where} \quad \{\omega_i\}_{1\le i\le m} \subset\,]0,1[\quad \textit{and} \quad \sum_{i=1}^m \omega_i = 1. \tag{80}$$

Furthermore, let $r \in \mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be the sequence generated by the following routine.

$$
\begin{array}{l}
\text{Initialization} \\
\left\lfloor
\begin{array}{l}
x_0 = r \\
\text{For } i = 1, \ldots, m \\
\left\lfloor\; z_{i,0} = x_0 \right.
\end{array}
\right. \\
\text{For } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
x_{n+1} = \sum_{i=1}^m \omega_i \operatorname{prox}_{f_i} z_{i,n} \\
\text{For } i = 1, \ldots, m \\
\left\lfloor\; z_{i,n+1} = x_{n+1} + z_{i,n} - \operatorname{prox}_{f_i} z_{i,n}. \right.
\end{array}
\right.
\end{array}
\tag{81}
$$

Then $x_n \to \operatorname{prox}_f r$.

Proof. Let us first observe that (79) and (80) imply that $f \in \Gamma_0(\mathcal{H})$. As a result, $\operatorname{prox}_f r$ is well defined. Let us define $\boldsymbol{\mathcal{H}}$ as in the proof of Theorem 2.5 and denote its norm as in (35). We also define $\boldsymbol{D}$ as in (36) and $\boldsymbol{j}$ as in (38), and set

$$
\boldsymbol{f} : \boldsymbol{\mathcal{H}} \to \left]-\infty, +\infty\right] : \boldsymbol{x} \mapsto \sum_{i=1}^m \omega_i f_i(x_i) \quad \text{and} \quad \boldsymbol{g} = \iota_{\boldsymbol{D}}. \tag{82}
$$

Then $\boldsymbol{f}$ and $\boldsymbol{g}$ are in $\Gamma_0(\boldsymbol{\mathcal{H}})$, $\partial \boldsymbol{f} = \times_{i=1}^m \partial f_i$, and $\operatorname{prox}_{\boldsymbol{g}} = P_{\boldsymbol{D}}$. Therefore, (74) and (64) yield

$$
P_{\boldsymbol{D}} \circ \operatorname{prox}_{\boldsymbol{f}} = P_{\boldsymbol{D}} \circ J_{\partial \boldsymbol{f}} : \boldsymbol{\mathcal{H}} \to \boldsymbol{D} : \boldsymbol{x} \mapsto \boldsymbol{j}\Bigg(\sum_{i=1}^m \omega_i J_{\partial f_i} x_i\Bigg) = \boldsymbol{j}\Bigg(\sum_{i=1}^m \omega_i \operatorname{prox}_{f_i} x_i\Bigg). \tag{83}
$$

Moreover, since (79) implies that $\operatorname{dom} \boldsymbol{f} \cap \boldsymbol{D} = \operatorname{dom} \boldsymbol{f} \cap \operatorname{dom} \boldsymbol{g} \neq \varnothing$, we have $\boldsymbol{f} + \boldsymbol{g} \in \Gamma_0(\boldsymbol{\mathcal{H}})$. We derive from (73), (80), (82), (38), and (35) that, for every $x \in \mathcal{H}$,

$$
\begin{aligned}
x = \operatorname{prox}_f r \quad &\Leftrightarrow \quad x = \underset{y\in\mathcal{H}}{\operatorname{argmin}}\; f(y) + \frac{1}{2}\|r - y\|^2 \\
&\Leftrightarrow \quad x = \underset{y\in\mathcal{H}}{\operatorname{argmin}}\; (\boldsymbol{f} \circ \boldsymbol{j})(y) + \frac{1}{2}|||\boldsymbol{j}(r) - \boldsymbol{j}(y)|||^2 \\
&\Leftrightarrow \quad \boldsymbol{j}(x) = \underset{\boldsymbol{y}\in\boldsymbol{D}}{\operatorname{argmin}}\; \boldsymbol{f}(\boldsymbol{y}) + \frac{1}{2}|||\boldsymbol{j}(r) - \boldsymbol{y}|||^2 \\
&\Leftrightarrow \quad \boldsymbol{j}(x) = \underset{\boldsymbol{y}\in\boldsymbol{\mathcal{H}}}{\operatorname{argmin}}\; (\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{y}) + \frac{1}{2}|||\boldsymbol{j}(r) - \boldsymbol{y}|||^2 \\
&\Leftrightarrow \quad \boldsymbol{j}(x) = \operatorname{prox}_{\boldsymbol{f}+\boldsymbol{g}} \boldsymbol{j}(r).
\end{aligned} \tag{84}
$$

Thus,

$$
\boldsymbol{j}(\operatorname{prox}_f r) = \operatorname{prox}_{\boldsymbol{f}+\boldsymbol{g}} \boldsymbol{j}(r). \tag{85}
$$

Now, let $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ be the sequence generated by algorithm (78) with $\boldsymbol{r} = \boldsymbol{j}(r)$. In view of Proposition 4.1,

$$
\boldsymbol{x}_n \to \operatorname{prox}_{\boldsymbol{f}+\boldsymbol{g}} \boldsymbol{j}(r). \tag{86}
$$

On the other hand, it follows from (74) that (78) is a specialization of (53) to the case when $\boldsymbol{A} = \partial \boldsymbol{f}$, $\boldsymbol{B} = \partial \boldsymbol{g}$, and $(\forall n \in \mathbb{N})$ $\boldsymbol{a}_n = 0$ and $\boldsymbol{b}_n = 0$. In addition, $\operatorname{prox}_{\boldsymbol{g}} = P_{\boldsymbol{D}}$. Therefore, just as we reduced (53) to (70) in the proof of Theorem 3.3, we can reduce (78) to

$$\left\lfloor \begin{array}{l} \boldsymbol{x}_0 = \boldsymbol{j}(r) \\ \boldsymbol{z}_0 = \boldsymbol{x}_0 \end{array} \right. \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{x}_{n+1} = P_{\boldsymbol{D}}\big(\operatorname{prox}_{\boldsymbol{f}} \boldsymbol{z}_n\big) \\ \boldsymbol{z}_{n+1} = \boldsymbol{x}_{n+1} + \boldsymbol{z}_n - \operatorname{prox}_{\boldsymbol{f}} \boldsymbol{z}_n. \end{array} \right. \tag{87}$$

Upon inspecting (87) and (81) in the light of (83), it becomes apparent that $(\forall n \in \mathbb{N})$ $\boldsymbol{x}_n = \boldsymbol{j}(x_n)$ and $\boldsymbol{z}_n = (z_{i,n})_{1 \le i \le m}$. Consequently, it follows from (86) and (85) that $x_n = \boldsymbol{j}^{-1}(\boldsymbol{x}_n) \to \boldsymbol{j}^{-1}(\operatorname{prox}_{\boldsymbol{f}+\boldsymbol{g}} \boldsymbol{r}) = \operatorname{prox}_f r$. □

As a corollary, we recover a parallel projection method to project onto the intersection of closed convex sets. The following result first appeared in [27, Section 6] (see also [4] and [29] for further analysis).

Corollary 4.3 *Let $(C_i)_{1 \le i \le m}$ be $m \ge 2$ closed convex subsets of $\mathcal{H}$ such that $C = \bigcap_{i=1}^m C_i \neq \varnothing$, and let $\{\omega_i\}_{1 \le i \le m} \subset \left]0,1\right[$ be such that $\sum_{i=1}^m \omega_i = 1$. Furthermore, let $r \in \mathcal{H}$ and let $(x_n)_{n \in \mathbb{N}}$ be the sequence generated by the following routine.*

$$\begin{array}{l} \text{Initialization} \\ \left\lfloor \begin{array}{l} x_0 = r \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \; z_{i,0} = x_0 \right. \end{array} \right. \\ \text{For } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} x_{n+1} = \sum_{i=1}^m \omega_i P_{C_i} z_{i,n} \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \; z_{i,n+1} = x_{n+1} + z_{i,n} - P_{C_i} z_{i,n}. \right. \end{array} \right. \end{array} \tag{88}$$

Then $x_n \to P_C\, r$.

Proof. Apply Theorem 4.2 with $(\forall i \in \{1, \ldots, m\})$ $f_i = \iota_{C_i}$. □

Remark 4.4 Suppose that the sets $(C_i)_{1 \le i \le m}$ are closed vector subspaces in Corollary 4.3. By orthogonality, the update rule in (88) reduces to $x_{n+1} = \sum_{i=1}^m \omega_i P_{C_i} x_n$ and we obtain $(\sum_{i=1}^m \omega_i P_{C_i})^n \to P_C$. This result can be found in [30, Proposition 26] (see also [43, Corollary 2.6]).

References

[1] H. Attouch, H. Brézis: Duality for the sum of convex functions in general Banach spaces, in: Aspects of Mathematics and its Applications, J. A. Barroso (ed.), North-Holland Math. Library, vol. 34, 125–133. North-Holland, Amsterdam, The Netherlands, 1986.

[2] H. Attouch, H. Riahi, M. Théra: Somme ponctuelle d'opérateurs maximaux monotones, Serdica Math. J. 22 (1996) 267–292.

[3] J.-P. Aubin, H. Frankowska: Set-Valued Analysis, Birkhäuser, Boston, MA, 1990.

[4] H. H. Bauschke, J. M. Borwein: Dykstra's alternating projection algorithm for two sets, J. Approx. Theory 79 (1994) 418–443.

[5] H. H. Bauschke, J. M. Borwein: Legendre functions and the method of random Bregman projections, *J Convex Anal.* 4 (1997) 27–67.

[6] H. H. Bauschke, P. L. Combettes: A Dykstra-like algorithm for two monotone operators, Pacific J. Optim. 4 (2008) 383–391.

[7] H. H. Bauschke, P. L. Combettes, D. R. Luke: Finding best approximation pairs relative to two closed convex sets in Hilbert spaces, J. Approx. Theory 127 (2004) 178–192.

[8] H. H. Bauschke, P. L. Combettes, S. Reich: The asymptotic behavior of the composition of two resolvents, Nonlinear Anal. 60 (2005) 283–301.

[9] J. P. Boyle, R. L. Dykstra: A method for finding projections onto the intersection of convex sets in Hilbert spaces, Lecture Notes in Statistics 37 (1986) 28–47.

[10] H. Brézis: Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert, North-Holland/Elsevier, New York, 1973.

[11] H. Brézis, P.-L. Lions: Produits infinis de résolvantes, Israel J. Math. 29 (1978) 329–345.

[12] Y. Censor, T. Elfving: A multiprojection algorithm using Bregman projections in a product space, Numerical Algorithms 8 (1994) 221–239.

[13] C. Chaux, P. L. Combettes, J.-C. Pesquet, V. R. Wajs: A variational formulation for frame-based inverse problems, Inverse Problems 23 (2007) 1495–1518.

[14] P. L. Combettes: Inconsistent signal feasibility problems: Least-squares solutions in a product space, IEEE Trans. Signal Process. 42 (1994) 2955–2966.

[15] P. L. Combettes: Fejér monotonicity in convex optimization, in Encyclopedia of Optimization, C. A. Floudas, P. M. Pardalos (eds.), vol. 2, Kluwer, Boston, MA (2001) 106–114.

[16] 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, The Netherlands (2001) 115–152.

[17] P. L. Combettes: Solving monotone inclusions via compositions of nonexpansive averaged operators, Optimization 53 (2004) 475–504.

[18] P. L. Combettes, S. A. Hirstoaga: Visco-penalization of the sum of two monotone operators, Nonlinear Anal. 69 (2008) 579–591.

[19] P. L. Combettes, J.-C. Pesquet: Proximal thresholding algorithm for minimization over orthonormal bases, SIAM J. Optim. 18 (2007) 1351–1376.

[20] P. L. Combettes, J.-C. Pesquet: A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery, IEEE J. Selected Topics Signal Process. 1 (2007) 564–574.

[21] P. L. Combettes, J.-C. Pesquet: A proximal decomposition method for solving convex variational inverse problems, *Inverse Problems* 24 (2008), Article ID 065014.

[22] P. L. Combettes, V. R. Wajs: Signal recovery by proximal forward-backward splitting, Multiscale Model. Simul. 4 (2005) 1168–1200.

[23] F. Deutsch: Best Approximation in Inner Product Spaces, Springer-Verlag, New York, 2001.

[24] J. Douglas, H. H. Rachford: On the numerical solution of heat conduction problems in two or three space variables, Trans. Amer. Math. Soc. 82 (1956) 421–439.

[25] R. L. Dykstra: An algorithm for restricted least squares regression, J. Amer. Stat. Assoc. 78 (1983) 837–842.

[26] J. Eckstein, D. P. Bertsekas: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Programming 55 (1992) 293–318.

[27] N. Gaffke, R. Mathar: A cyclic projection algorithm via duality, Metrika 36 (1989) 29–54.

[28] H. Idrissi, O. Lefebvre, C. Michelot: Applications and numerical convergence of the partial inverse method, Lecture Notes in Mathematics 1405 (1989) 39–54.

[29] A. N. Iusem, A. R. de Pierro: On the convergence of Han's method for convex programming with quadratic objective, Math. Programming 52 (1991) 265–284.

[30] M. L. Lapidus: Generalization of the Trotter-Lie formula, Integral Equations Operator Theory 4 (1981) 366–415.

[31] J. Lawrence, J. E. Spingarn: On fixed points of non-expansive piecewise isometric mappings, Proc. London Math. Soc. 55 (1987) 605–624.

[32] B. Lemaire: The proximal algorithm, in: New Methods in Optimization and Their Industrial Uses, (J.-P. Penot, Ed.), International Series of Numerical Mathematics 87 (1989) 73–87, Birkhäuser, Boston, MA.

[33] P.-L. Lions: Une méthode itérative de résolution d'une inéquation variationnelle, Israel J. Math. 31 (1978) 204–208.

[34] P.-L. Lions, B. Mercier: Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979) 964–979.

[35] Ph. Mahey, S. Oualibouch, Pham Dinh Tao: Proximal decomposition on the graph of a maximal monotone operator, SIAM J. Optim. 5 (1995) 454–466.

[36] B. Martinet: Algorithmes pour la Résolution de Problèmes d'Optimisation et de Minimax, Thèse, Université de Grenoble, France, 1972.

[37] J.-J. Moreau: Fonctions convexes duales et points proximaux dans un espace hilbertien, C. R. Acad. Sci. Paris Sér. A Math. 255 (1962) 2897–2899.

[38] J.-J. Moreau: Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France 93 (1965) 273–299.

[39] G. B. Passty: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, J. Math. Anal. Appl. 72 (1979) 383–390.

[40] G. Pierra: Éclatement de contraintes en parallèle pour la minimisation d'une forme quadratique, Lecture Notes in Computer Science 41 (1976) 200–218.

[41] G. Pierra: Decomposition through formalization in a product space, Math. Programming 28 (1984) 96–115.

[42] B. T. Polyak: Introduction to Optimization, Optimization Software Inc., New York, 1987.

[43] S. Reich: A limit theorem for projections, Linear and Multilinear Algebra 13 (1983) 281–290.

[44] R. T. Rockafellar: Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877–898.

[45] R. T. Rockafellar, R. J. B. Wets: Variational Analysis, Springer-Verlag, New York, 1998.

[46] L. I. Rudin, S. Osher, E. Fatemi: Nonlinear total variation based noise removal algorithms, Physica D 60 (1992) 259–268.

[47] S. Simons: From Hahn-Banach to Monotonicity, Lecture Notes in Mathematics 1693, Springer-Verlag, New York, 2008.

[48] J. E. Spingarn: Partial inverse of a monotone operator, Appl. Math. Optim. 10 (1983) 247–265.

[49] J. E. Spingarn: Applications of the method of partial inverses to convex programming: Decomposition, Math. Programming 32 (1985) 199–223.

[50] J. E. Spingarn: A projection method for least-squares solutions to overdetermined systems of linear inequalities, Linear Algebra Appl. 86 (1987) 211–236.

[51] J. A. Tropp: Just relax: Convex programming methods for identifying sparse signals in noise, IEEE Trans. Inform. Theory 52 (2006) 1030–1051.

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