Optimization
Vol. 00, No. 00, Month 200x, 1–22

RESEARCH ARTICLE

Forward-Douglas-Rachford splitting and forward-partial inverse method for solving monotone inclusions

Luis M. Briceño-Arias[1] †

[1] Departamento de Matemática
Universidad Técnica Federico Santa María
Santiago, Chile

(*Received 00 Month 200x; in final form 00 Month 200x*)

We provide two weakly convergent algorithms for finding a zero of the sum of a maximally monotone operator, a cocoercive operator, and the normal cone to a closed vector subspace of a real Hilbert space. The methods exploit the intrinsic structure of the problem by activating explicitly the cocoercive operator in the first step, and taking advantage of a vector space decomposition in the second step. The second step of the first method is a Douglas-Rachford iteration involving the maximally monotone operator and the normal cone. In the second method it is a proximal step involving the partial inverse of the maximally monotone operator with respect to the vector subspace. Connections between the proposed methods and other methods in the literature are provided. Applications to monotone inclusions with finitely many maximally monotone operators and optimization problems are examined.

Keywords: composite operator, convex optimization, minimization algorithm, monotone inclusion, partial inverse, splitting methods.

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

1. Introduction

This paper is concerned with the numerical resolution of the following problem.

Problem 1.1: Let $\mathcal{H}$ be a real Hilbert space and let V be a closed vector subspace of $\mathcal{H}$. The normal cone to V is denoted by N_V. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and let $B\colon \mathcal{H} \to \mathcal{H}$ be a β–cocoercive operator in V, i.e., it satisfies

$$(\forall x \in V)(\forall y \in V) \quad \langle x - y \mid Bx - By\rangle \geq \beta\|Bx - By\|^2. \tag{1}$$

The problem is to

$$\text{find } x \in \mathcal{H} \quad \text{such that} \quad 0 \in Ax + Bx + N_V x, \tag{2}$$

under the assumption that the set of solutions Z of (2) is nonempty.

† Email: L. M. Briceño-Arias, `luis.briceno@usm.cl`, phone: +56 2 432 6662.This work was partially supported by CONICYT under grant FONDECYT 3120054 and Anillo ACT1106, by "Programa de financiamiento basal" from the Center for Mathematical Modeling, Universidad de Chile, and by Project Math-Amsud N 13MATH01.

ISSN: 0233-1934 print/ISSN 1029-4945 online

DOI: 10.1080/0233193YYxxxxxxxx
http://www.informaworld.com

Problem 1.1 arises in a wide range of areas such as optimization [17, 42], variational inequalities [30, 44, 45], monotone operator theory [22, 31, 36, 41], partial differential equations [25, 26, 30, 33, 48], economics [28, 35], signal and image processing [2, 12, 13, 20, 21, 38], evolution inclusions [1, 27, 40], and traffic theory [7, 8, 24, 37, 39], among others.

In the particular case when $B \equiv 0$, (2) becomes

$$\text{find} \;\; x \in V \quad \text{such that} \quad (\exists\, y \in V^{\perp}) \quad y \in Ax, \tag{3}$$

where $V^{\perp}$ stands for the orthogonal complement of V. Problem (3) has been studied in [41] and it is solved with the method of partial inverses. On the other hand, when $V = \mathcal{H}$, (2) reduces to

$$\text{find} \;\; x \in \mathcal{H} \quad \text{such that} \quad 0 \in Ax + Bx, \tag{4}$$

which can be solved by the forward-backward splitting (see [15] and the references therein). In the general case, Problem 1.1 can be solved by several algorithms, but any of them exploits the intrinsic structure of the problem. The forward-backward splitting [15] can solve Problem 1.1 by an explicit computation of B and an implicit computation involving the resolvent of $A + N_V$. The disadvantage of this method is that this resolvent is not always easy to compute. It is preferable, hence, to activate separately A and N_V. In [34] an ergodic method involving the resolvents of each maximally monotone operator separately is proposed, and weak convergence to a solution to Problem 1.1 is obtained. However, the method includes vanishing parameters which leads to numerical instabilities and, moreover, it involves the computation of $(\mathrm{Id} + \gamma B)^{-1}$ for some positive constant γ, which is not always easy to compute explicitly. The methods proposed in [10, 16, 18, 41] for finding a zero of the sum of finitely many maximally monotone operators overcomes the problem caused by the vanishing parameters in [34], but it still needs to compute $(\mathrm{Id} + \gamma B)^{-1}$. The primal-dual method proposed in [46] overcomes the disadvantages of previous algorithms by computing explicit steps of B. However, the method does not take advantage of the vector subspace involved and, as a consequence, it needs to store several auxiliary variables at each iteration, which can be difficult for high dimensional problems.

In this paper we propose two methods for solving Problem 1.1 that exploit all the intrinsic properties of the problem. The first algorithm computes an explicit step on B followed by a Douglas-Rachford step [31, 43] involving A and N_V. The second method computes an explicit step on B followed by an implicit step involving the partial inverse of A with respect to V. The latter method generalizes the partial inverse method [41] and the forward-backward splitting [15] in the particular cases (3) and (4), respectively. We also provide connections between the proposed methods, we study some relations with other methods in the literature, and we illustrate the flexibility of this framework via some applications.

The paper is organized as follows. In Section 2 we provide the notation and some preliminaries. In Section 3 we provide a new version of the Krasnosel'skiĭ-Mann iteration for the composition of averaged operators. In Section 4 the latter method is applied to particular averaged operators for obtaining the forward-Douglas-Rachford splitting and in Section 5 the forward-partial inverse algorithm is proposed. We also provide connections with other algorithms in the literature. Finally, in Section 6 we examine an application for finding a zero of a sum of m maximally monotone operators and a cocoercive operator and an application to optimization problems.

2. Notation and preliminaries

Throughout this paper, $\mathcal{H}$ is a real Hilbert space with scalar product denoted by $\langle \cdot \mid \cdot \rangle$ and associated norm $\| \cdot \|$. The symbols $\rightharpoonup$ and $\to$ denote, respectively, weak and strong convergence and Id denotes the identity operator. The indicator function of a subset C of $\mathcal{H}$ is

$$\iota_C \colon x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C, \end{cases} \tag{5}$$

if C is nonempty, closed, and convex, the projection of x onto C, denoted by $P_C x$, is the unique point in $\operatorname{Argmin}_{y \in C} \|x - y\|$, and the normal cone to C is the maximally monotone operator

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

An operator $R\colon \mathcal{H} \to \mathcal{H}$ is nonexpansive if

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|Rx - Ry\| \leq \|x - y\| \tag{7}$$

and $\operatorname{Fix} R = \big\{x \in \mathcal{H} \mid Rx = x\big\}$ is the set of fixed points of R. An operator $T\colon \mathcal{H} \to \mathcal{H}$ is α–averaged for some $\alpha \in]0,1[$ if

$$T = (1 - \alpha)\operatorname{Id} + \alpha R \tag{8}$$

for some nonexpansive operator R, or, equivalently,

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1-\alpha}{\alpha}\|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2, \tag{9}$$

or

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad 2(1-\alpha)\,\langle x - y \mid Tx - Ty\rangle \geq \|Tx - Ty\|^2 + (1 - 2\alpha)\|x - y\|^2. \tag{10}$$

On the other hand, T is β–cocoercive for some $\beta \in \,]0, +\infty[$ if

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Tx - Ty\rangle \geq \beta\|Tx - Ty\|^2. \tag{11}$$

We say that T is firmly nonexpansive if T is 1/2–averaged, or equivalently, if T is 1–cocoercive.

We denote by $\operatorname{gra} A = \big\{(x,u) \in \mathcal{H} \times \mathcal{H} \mid u \in Ax\big\}$ the graph of a set-valued operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, by $\operatorname{dom} A = \big\{x \in \mathcal{H} \mid Ax \neq \varnothing\big\}$ its domain, by $\operatorname{zer} A = \big\{x \in \mathcal{H} \mid 0 \in Ax\big\}$ its set of zeros, and by $J_A = (\operatorname{Id} + A)^{-1}$ its resolvent. If A is monotone, then J_A is single-valued and nonexpansive and, furthermore, if A is maximally monotone, then $\operatorname{dom} J_A = \mathcal{H}$. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone. The reflection operator of A is $R_A = 2J_A - \operatorname{Id}$, which is nonexpansive. The partial inverse of A with respect to a vector subspace V of $\mathcal{H}$, denoted by A_V, is defined by

$$(\forall (x,y) \in \mathcal{H}^2) \quad y \in A_V x \quad \Leftrightarrow \quad (P_V y + P_{V^\perp} x) \in A(P_V x + P_{V^\perp} y). \tag{12}$$

For complements and further background on monotone operator theory, see [1, 6, 41, 47, 48].

3. Krasnosel'skiĭ–Mann iterations for the composition of averaged operators

The following result will be useful for obtaining the convergence of the first method. It provides the weak convergence of the iterates generated by the Krasnosel'skiĭ–Mann iteration [15, 29, 32] applied to the composition of finitely many averaged operators to a common fixed point. In [6, Corollary 5.15] a similar method is proposed with guaranteed convergence, but without including errors in the computation of the operators involved. On the other hand, in [15] another algorithm involving inexactitudes in the computation of the averaged operators is studied in the case when such operators may vary at each iteration. However, the relaxation parameters in this case are forced to be in $]0,1]$. We propose a new method which includes summable errors in the computation of the averaged operators and allows for a larger choice for the relaxation parameters. First, for every strictly positive integer i and a family of averaged operators $(T_j)_{1\leq j\leq m}$, let us define

$$\prod_{j=i}^{m} T_j = \begin{cases} T_i\circ T_{i+1}\circ\cdots\circ T_m, & \text{if } i\leq m;\\ \mathrm{Id}, & \text{otherwise.}\end{cases} \tag{13}$$

Proposition 3.1: *Let $m\geq 1$, for every $i\in\{1,\ldots,m\}$, let $\alpha_i\in\,]0,1[$, let T_i be an α_i-averaged operator, and let $(e_{i,n})_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. In addition, set*

$$\alpha = \frac{m\,\max\{\alpha_1,\ldots,\alpha_m\}}{1+(m-1)\max\{\alpha_1,\ldots,\alpha_m\}}, \tag{14}$$

let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0,1/\alpha[$, suppose that $\mathrm{Fix}(T_1\circ\cdots\circ T_m)\neq\varnothing$, and suppose that

$$\sum_{n\in\mathbb{N}}\lambda_n(1-\alpha\lambda_n)=+\infty \quad \textit{and} \quad (\forall i\in\{1,\ldots,m\}) \quad \sum_{n\in\mathbb{N}}\lambda_n\|e_{i,n}\|<+\infty. \tag{15}$$

Moreover, let $z_0\in\mathcal{H}$ and set, for every $n\in\mathbb{N}$,

$$z_{n+1}=z_n+\lambda_n\Big(T_1\big(T_2(\cdots T_{m-1}(T_m z_n+e_{m,n})+e_{m-1,n}\cdots)+e_{2,n}\big)+e_{1,n}-z_n\Big). \tag{16}$$

Then the following hold for some $\overline{z}\in\mathrm{Fix}(T_1\circ\cdots\circ T_m)$.

(i) $z_n \rightharpoonup \overline{z}$.

(ii) $\displaystyle\sum_{n\in\mathbb{N}}\lambda_n(1-\alpha\lambda_n)\Big\|\prod_{j=1}^{m}T_j z_n - z_n\Big\|^2<+\infty$.

(iii) $\displaystyle\prod_{j=1}^{m}T_j z_n - z_n\to 0$.

(iv) $z_{n+1}-z_n\to 0$.

(v) $\displaystyle\max_{1\leq i\leq m}\sum_{n\in\mathbb{N}}\lambda_n\Big\|(\mathrm{Id}-T_i)\prod_{j=i+1}^{m}T_j z_n-(\mathrm{Id}-T_i)\prod_{j=i+1}^{m}T_j\overline{z}\Big\|^2<+\infty$.

Proof: (i): First note that (16) can be written equivalently as

$$(\forall n\in\mathbb{N})\quad z_{n+1}=z_n+\lambda_n(Tz_n+e_n-z_n), \tag{17}$$

where

$$\begin{cases} T = T_1 \circ T_2 \circ \cdots \circ T_m = \prod_{j=1}^{m} T_j \\ e_n = T_1\big(T_2(\cdots T_{m-1}(T_m z_n + e_{m,n}) + e_{m-1,n} \cdots) + e_{2,n}\big) + e_{1,n} - T z_n. \end{cases} \tag{18}$$

It follows from [15, Lemma 2.2(iii)] that T is α-averaged with α defined in (14), and, using the nonexpansivity of $(T_i)_{1\leq i\leq m}$, we obtain, for every $n \in \mathbb{N}$,

$$\begin{aligned} \|e_n\| &\leq \|T_1\big(T_2(\cdots T_{m-1}(T_m z_n + e_{m,n}) + e_{m-1,n} \cdots) + e_{2,n}\big) \\ &\qquad - T_1\big(T_2(\cdots T_{m-1}(T_m z_n) \cdots)\big)\| + \|e_{1,n}\| \\ &\leq \|T_2\big(T_3(\cdots T_{m-1}(T_m z_n + e_{m,n}) + e_{m-1,n} \cdots) + e_{3,n}\big) \\ &\qquad - T_2\big(T_3(\cdots T_{m-1}(T_m z_n) \cdots)\big)\| + \|e_{2,n}\| + \|e_{1,n}\| \\ &\leq \\ &\ \vdots \\ &\leq \|T_{m-1}(T_m z_n + e_{m,n}) - T_{m-1}(T_m z_n)\| + \|e_{m-1,n}\| + \cdots + \|e_{2,n}\| + \|e_{1,n}\| \\ &\leq \sum_{i=1}^{m} \|e_{i,n}\|. \end{aligned} \tag{19}$$

Hence, it follows from (15) that

$$\sum_{n\in\mathbb{N}} \lambda_n \|e_n\| \leq \sum_{n\in\mathbb{N}} \lambda_n \sum_{i=1}^{m} \|e_{i,n}\| = \sum_{i=1}^{m} \sum_{n\in\mathbb{N}} \lambda_n \|e_{i,n}\| < +\infty. \tag{20}$$

Now, set $R = (1 - 1/\alpha)\,\mathrm{Id} + (1/\alpha)T$ and, for every $n \in \mathbb{N}$, set $\mu_n = \alpha\lambda_n$. Then it follows from the firm nonexpansiveness of T and (8) that R is a nonexpansive operator, $\operatorname{Fix} R = \operatorname{Fix} T$, and (17) is equivalent to

$$(\forall n \in \mathbb{N}) \quad z_{n+1} = z_n + \mu_n(R z_n + c_n - z_n), \tag{21}$$

where $c_n = e_n/\alpha$. Since $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in $]0,1[$ and (15) and (20) yields $\sum_{n\in\mathbb{N}} \mu_n(1-\mu_n) = +\infty$ and $\sum_{n\in\mathbb{N}} \mu_n \|c_n\| < +\infty$, the result follows from [15, Lemma 5.1].

(ii): Fix $n \in \mathbb{N}$. It follows from (17), Cauchy-Schwartz inequality, and [6, Lemma 2.13(ii)] that

$$\begin{aligned} \|z_{n+1} - \overline{z}\|^2 &= \|(1-\lambda_n)(z_n - \overline{z}) + \lambda_n(T z_n - T\overline{z} + e_n)\|^2 \\ &\leq \|(1-\lambda_n)(z_n - \overline{z}) + \lambda_n(T z_n - T\overline{z})\|^2 + \varepsilon_n \\ &= (1-\lambda_n)\|z_n - \overline{z}\|^2 + \lambda_n\|T z_n - T\overline{z}\|^2 - \lambda_n(1-\lambda_n)\|T z_n - z_n\|^2 + \varepsilon_n, \end{aligned} \tag{22}$$

where,

$$(\forall k \in \mathbb{N}) \quad \varepsilon_k = \lambda_k^2 \|e_k\|^2 + 2\lambda_k \|(1-\lambda_k)(z_k - \overline{z}) + \lambda_k(T z_k - T\overline{z})\| \|e_k\|. \tag{23}$$

Note that the convexity of $\|\cdot\|$, the nonexpansivity of T, and (i) yield

$$\begin{aligned}\sum_{k\in\mathbb{N}}\varepsilon_k &= \sum_{k\in\mathbb{N}}\lambda_k^2\|e_k\|^2 + 2\sum_{k\in\mathbb{N}}\lambda_k\|(1-\lambda_k)(z_k-\overline{z})+\lambda_k(Tz_k-T\overline{z})\|\|e_k\|\\ &\leq \Big(\sum_{k\in\mathbb{N}}\lambda_k\|e_k\|\Big)^2 + 2\sum_{k\in\mathbb{N}}\lambda_k\big((1-\lambda_k)\|z_k-\overline{z}\|+\lambda_k\|Tz_k-T\overline{z}\|\big)\|e_k\|\\ &\leq \Big(\sum_{k\in\mathbb{N}}\lambda_k\|e_k\|\Big)^2 + 2\Big(\sup_{k\in\mathbb{N}}\|z_k-\overline{z}\|\Big)\sum_{k\in\mathbb{N}}\lambda_k\|e_k\| < +\infty. \qquad (24)\end{aligned}$$

On one hand, since T is α-averaged, it follows from (22) and (9) that

$$\begin{aligned}\|z_{n+1}-\overline{z}\|^2 &\leq (1-\lambda_n)\|z_n-\overline{z}\|^2 + \lambda_n\Big(\|z_n-\overline{z}\|^2 - \frac{(1-\alpha)}{\alpha}\|Tz_n-z_n\|^2\Big)\\ &\qquad - \lambda_n(1-\lambda_n)\|Tz_n-z_n\|^2+\varepsilon_n\\ &\leq \|z_n-\overline{z}\|^2 - \frac{\lambda_n(1-\alpha\lambda_n)}{\alpha}\|Tz_n-z_n\|^2+\varepsilon_n, \qquad (25)\end{aligned}$$

and, hence, the result is deduced from [14, Lemma 3.1(iii)].

(iii): It follows from (15) and (ii) that $\underline{\lim}\,\|Tz_n-z_n\|=0$. Moreover, it follows from (17) that

$$\begin{aligned}(\forall n\in\mathbb{N})\quad \|Tz_{n+1}-z_{n+1}\| &\leq \|Tz_{n+1}-Tz_n\|+(1-\lambda_n)\|Tz_n-z_n\|+\lambda_n\|e_n\|\\ &\leq \|z_{n+1}-z_n\|+(1-\lambda_n)\|Tz_n-z_n\|+\lambda_n\|e_n\|\\ &\leq \|Tz_n-z_n\|+2\lambda_n\|e_n\|. \qquad (26)\end{aligned}$$

Hence, from (20) and [14, Lemma 3.1] we deduce that $(\|Tz_n-z_n\|)_{n\in\mathbb{N}}$ converges, and therefore, $Tz_n-z_n\to 0$.

(iv): From (17), (18), (iii), and (20) we obtain

$$\|z_{n+1}-z_n\|\leq\lambda_n\|Tz_n-z_n\|+\lambda_n\|e_n\|\leq(1/\alpha)\|Tz_n-z_n\|+\lambda_n\|e_n\|\to 0. \qquad (27)$$

(v): Since $(T_i)_{1\leq i\leq m}$ are averaged operators, we have from (18) and (9) that

$$\begin{aligned}\|Tz_n-T\overline{z}\|^2 &\leq \Big\|\prod_{j=2}^m T_jz_n-\prod_{j=2}^m T_j\overline{z}\Big\|^2-\frac{1-\alpha_1}{\alpha_1}\Big\|(\operatorname{Id}-T_1)\prod_{j=2}^m T_jz_n-(\operatorname{Id}-T_1)\prod_{j=2}^m T_j\overline{z}\Big\|^2\\ &\leq \Big\|\prod_{j=3}^m T_jz_n-\prod_{j=3}^m T_j\overline{z}\Big\|^2-\frac{1-\alpha_2}{\alpha_2}\Big\|(\operatorname{Id}-T_2)\prod_{j=3}^m T_jz_n-(\operatorname{Id}-T_2)\prod_{j=3}^m T_j\overline{z}\Big\|^2\\ &\qquad -\frac{1-\alpha_1}{\alpha_1}\Big\|(\operatorname{Id}-T_1)\prod_{j=2}^m T_jz_n-(\operatorname{Id}-T_1)\prod_{j=2}^m T_j\overline{z}\Big\|^2\\ &\ \ \vdots\\ &\leq \|z_n-\overline{z}\|^2-\sum_{i=1}^m\frac{1-\alpha_i}{\alpha_i}\Big\|(\operatorname{Id}-T_i)\prod_{j=i+1}^m T_jz_n-(\operatorname{Id}-T_i)\prod_{j=i+1}^m T_j\overline{z}\Big\|^2. \qquad (28)\end{aligned}$$

Hence, from (22) we deduce

$$\|z_{n+1}-\overline{z}\|^2 \leq \|z_n-\overline{z}\|^2-\lambda_n \sum_{i=1}^{m} \frac{1-\alpha_i}{\alpha_i} \left\|(\operatorname{Id}-T_i)\prod_{j=i+1}^{m} T_j z_n-(\operatorname{Id}-T_i)\prod_{j=i+1}^{m} T_j\overline{z}\right\|^2+\varepsilon_n. \tag{29}$$

Therefore, it follows from [14, Lemma 3.1(iii)] that

$$\begin{aligned}\sum_{i=1}^{m} \frac{1-\alpha_i}{\alpha_i} \sum_{n\in\mathbb{N}} \lambda_n \left\|(\operatorname{Id}-T_i)\prod_{j=i+1}^{m} T_j z_n-(\operatorname{Id}-T_i)\prod_{j=i+1}^{m} T_j\overline{z}\right\|^2 \\ = \sum_{n\in\mathbb{N}} \lambda_n \sum_{i=1}^{m} \frac{1-\alpha_i}{\alpha_i} \left\|(\operatorname{Id}-T_i)\prod_{j=i+1}^{m} T_j z_n-(\operatorname{Id}-T_i)\prod_{j=i+1}^{m} T_j\overline{z}\right\|^2 < +\infty,\end{aligned} \tag{30}$$

which yields the result. □

Remark 3.1: In the particular case when $m = 1$, Proposition 3.1 provides the weak convergence of the iterates generated by the classical Krasnosel'skiĭ-Mann iteration [15, 29, 32] in the case of averaged operators. This result is interesting in this own right since it generalizes [6, Proposition 5.15] by considering errors on the computation of the involved operator and provides a larger choice of relaxation parameters than in the nonexpansive case (see, e.g.,[15, 29, 32]).

4. Forward-Douglas-Rachford splitting

In this section we provide the first method for solving Problem 1.1. We provide a characterization of the solutions to Problem 1.1, then the algorithm is proposed and its weak convergence to a solution to Problem 1.1 is proved.

4.1. Characterization

Let us start with a characterization of the solutions to Problem 1.1.

Proposition 4.1: *Let $\gamma \in]0, 2\beta[$ and $\mathcal{H}$, V, A, B, and Z be as in Problem 1.1. Define*

$$\begin{cases} T_\gamma = \frac{1}{2}(\operatorname{Id}+R_{\gamma A}\circ R_{N_V})\colon \mathcal{H}\to\mathcal{H} \\ S_\gamma = \operatorname{Id}-\gamma P_V\circ B\circ P_V\colon \mathcal{H}\to\mathcal{H}. \end{cases} \tag{31}$$

Then the following hold.

(i) *T_γ is firmly nonexpansive.*
(ii) *S_γ is $\gamma/(2\beta)$–averaged.*
(iii) *Let $x \in \mathcal{H}$. Then $x \in Z$ if and only if $x \in V$ and*

$$(\exists\, y \in V^\perp\cap(Ax+Bx)) \quad \text{such that} \quad x-\gamma(y-P_{V^\perp}Bx) \in \operatorname{Fix}(T_\gamma\circ S_\gamma). \tag{32}$$

Proof: (i): Since γA is maximally monotone $J_{\gamma A}$ is firmly nonexpansive and $R_{\gamma A} = 2J_{\gamma A} - \operatorname{Id}$ is nonexpansive. An analogous argument yields the nonexpansivity of $R_{N_V} = 2P_V - \operatorname{Id}$. Hence, $R_{\gamma A} \circ R_{N_V}$ is nonexpansive and the result follows from the definition of a firmly nonexpansive operator in (8).

(ii): Since V is a closed vector subspace of $\mathcal{H}$ we have that P_V is linear and $P_V^* = P_V$. Hence, the cocoercivity of B in V yields, for every $(z,w) \in \mathcal{H}^2$ and $\eta \in \,]0,+\infty[$,

$$\begin{aligned}\big\langle z-w \mid \eta P_V\big(B(P_Vz)-B(P_Vw)\big)\big\rangle &= \eta\,\langle P_Vz-P_Vw \mid B(P_Vz)-B(P_Vw)\rangle\\ &\geq \eta\beta\|B(P_Vz)-B(P_Vw)\|^2\\ &\geq (\beta/\eta)\|\eta P_V\big(B(P_Vz)\big)-\eta P_V\big(B(P_Vw)\big)\|^2.\end{aligned} \tag{33}$$

In the particular case when $\eta=\beta$, (33) implies that the operator $z \mapsto \beta P_V\circ B\circ P_V z$ is firmly nonexpansive (1/2-averaged). Since $\gamma \in \,]0,2\beta[$ the result follows from [15, Lemma 2.3].

(iii): Let $x \in \mathcal{H}$ be a solution to Problem 1.1. We have $x \in V$ and there exists $y \in V^\perp = N_V x$ such that $y \in Ax+Bx$. Set $z = x-\gamma(y-P_{V^\perp}Bx)$. Note that $R_{N_V}z = 2P_Vz - z = x+\gamma(y-P_{V^\perp}Bx)$ and $P_Vz = x$. Hence, since B is single valued and, for every $w\in V$, $R_Vw=w$, it follows from the linearity of P_V that

$$x+\gamma y-\gamma Bx = x+\gamma(y-P_{V^\perp}Bx)-\gamma P_VBx = R_{N_V}z-\gamma P_VBx = R_{N_V}(z-\gamma P_VBP_Vz), \tag{34}$$

and, therefore,

$$\begin{aligned} y\in Ax+Bx \quad&\Leftrightarrow\quad x+\gamma y-\gamma Bx \in x+\gamma Ax\\ &\Leftrightarrow\quad x = J_{\gamma A}(x+\gamma y-\gamma Bx) = J_{\gamma A}\big(R_{N_V}(z-\gamma P_VBP_Vz)\big)\\ &\Leftrightarrow\quad x = \frac{1}{2}\Big(2J_{\gamma A}\big(R_{N_V}(z-\gamma P_VBP_Vz)\big)\\ &\qquad\qquad - R_{N_V}(z-\gamma P_VBP_Vz)+x+\gamma y-\gamma Bx\Big)\\ &\Leftrightarrow\quad x = \frac{1}{2}\Big(R_{\gamma A}\big(R_{N_V}(z-\gamma P_VBP_Vz)\big)+x+\gamma y-\gamma Bx\Big)\\ &\Leftrightarrow\quad x = \frac{1}{2}\Big(R_{\gamma A}\big(R_{N_V}(z-\gamma P_VBP_Vz)\big)+z-\gamma P_VBP_Vz\Big)\\ &\qquad\qquad +\gamma(y-P_{V^\perp}Bx)\\ &\Leftrightarrow\quad z = T_\gamma\circ S_\gamma z, \end{aligned} \tag{35}$$

which yields the result. ☐

4.2. Algorithm and convergence

In the following result we propose our first algorithm and we prove its convergence to a solution to Problem 1.1. The method is inspired from the characterization provided in Proposition 4.1 and Proposition 3.1.

Theorem 4.2: *Let $\mathcal{H}$, V, A, B, and Z be as in Problem 1.1, let $\gamma\in\,]0,2\beta[$, let $\alpha = \max\{2/3, 2\gamma/(\gamma+2\beta)\}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0,1/\alpha[$, let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$, and suppose that*

$$\sum_{n\in\mathbb{N}}\lambda_n(1-\alpha\lambda_n)=+\infty \quad \textit{and} \quad \sum_{n\in\mathbb{N}}\lambda_n(\|a_n\|+\|b_n\|)<+\infty. \tag{36}$$

Moreover, let $z_0 \in \mathcal{H}$ and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} x_n = P_V z_n \\ y_n = (x_n - z_n)/\gamma \\ s_n = x_n - \gamma P_V\big(Bx_n + a_n\big) + \gamma y_n \\ p_n = J_{\gamma A} s_n + b_n \\ z_{n+1} = z_n + \lambda_n(p_n - x_n). \end{array} \right. \tag{37}$$

Then the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are in V and $V^\perp$, respectively, and the following hold for some $\overline{x} \in Z$ and some $\overline{y} \in V^\perp \cap \big(A\overline{x} + P_V B\overline{x}\big)$.

(i) $x_n \rightharpoonup \overline{x}$ *and* $y_n \rightharpoonup \overline{y}$.
(ii) $x_{n+1} - x_n \to 0$ *and* $y_{n+1} - y_n \to 0$.
(iii) $\sum_{n\in\mathbb{N}} \lambda_n \|P_V(Bx_n - B\overline{x})\|^2 < +\infty$.

Proof: First note that (37) can be written equivalently as

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} x_n = P_V z_n \\ y_n = -P_{V^\perp} z_n/\gamma \\ z_{n+1} = z_n + \lambda_n\big(T_\gamma(S_\gamma z_n + c_n) + b_n - z_n\big), \end{array} \right. \tag{38}$$

where T_γ and S_γ are defined in (31) and, for every $n \in \mathbb{N}$, $c_n = -\gamma P_V a_n$. We have from (36) that

$$\sum_{n\in\mathbb{N}} \lambda_n(\|b_n\| + \|c_n\|) \leq \sum_{n\in\mathbb{N}} \lambda_n(\|b_n\| + \gamma\|a_n\|) \leq \max\{1,\gamma\} \sum_{n\in\mathbb{N}} \lambda_n(\|a_n\| + \|b_n\|) < +\infty. \tag{39}$$

Moreover, it follows from Proposition 4.1(i)&(ii) that T_γ is 1/2-averaged and S_γ is $\gamma/(2\beta)$-averaged. Altogether, by setting $m = 2$, $T_1 = T_\gamma$, $T_2 = S_\gamma$, $\alpha_1 = 1/2$, $\alpha_2 = \gamma/(2\beta)$, $e_{1,n} = b_n$, $e_{2,n} = c_n$, and noting that

$$\frac{2\max\{1/2, \gamma/(2\beta)\}}{1 + \max\{1/2, \gamma/(2\beta)\}} = \max\{2/3, 2\gamma/(\gamma + 2\beta)\} = \alpha, \tag{40}$$

it follows from Proposition 3.1 that there exists $\overline{z} \in \operatorname{Fix}(T_\gamma \circ S_\gamma)$ such that

$$z_n \rightharpoonup \overline{z} \tag{41}$$

$$z_{n+1} - z_n \to 0 \tag{42}$$

$$\sum_{n\in\mathbb{N}} \lambda_n \|(\operatorname{Id} - S_\gamma)z_n - (\operatorname{Id} - S_\gamma)\overline{z}\|^2 < +\infty. \tag{43}$$

Now set $\overline{x} = P_V \overline{z}$ and $\overline{y} = -P_{V^\perp}\overline{z}/\gamma$. It follows from Proposition 4.1(iii) that $\overline{x}$ is solution to Problem 1.1 and $\overline{y} = y - P_{V^\perp} B\overline{x}$ for some $y \in V^\perp \cap (A\overline{x} + B\overline{x})$. Then, $\overline{y} \in V^\perp \cap (A\overline{x} + P_V B\overline{x})$.

(i): It is clear from (38) and (41) that $x_n \rightharpoonup \overline{x}$ and $y_n \rightharpoonup \overline{y}$.

(ii): It is a consequence of (42) and

$$(\forall n \in \mathbb{N}) \quad \|z_{n+1} - z_n\|^2 = \|x_{n+1} - x_n\|^2 + \gamma^2 \|y_{n+1} - y_n\|^2. \tag{44}$$

(iii): It follows from (31) that

$$(\forall n \in \mathbb{N}) \quad \|(\operatorname{Id} - S_\gamma)z_n - (\operatorname{Id} - S_\gamma)\overline{z}\|^2 = \gamma^2 \|P_V(Bx_n) - P_V(B\overline{x})\|^2. \tag{45}$$

Hence, the result follows from (43). □

Remark 4.1: Note that, if $\underline{\lim}\,\lambda_n > 0$, then Theorem 4.2(iii) implies $P_V(Bx_n) \to P_V(B\overline{x})$.

5. Forward-partial-inverse splitting

We provide a second characterization of solutions to Problem 1.1 via the partial inverse operator introduced in [41]. This characterization motivates a second algorithm, whose convergence to a solution to Problem 1.1 is proved. The proposed method generalizes the partial inverse method proposed in [41] and the forward-backward splitting [15].

5.1. Characterization

Proposition 5.1: *Let $\gamma \in \left]0,+\infty\right[$ and $\mathcal{H}$, A, B, and V be as in Problem 1.1. Define*

$$\begin{cases} \mathcal{A}_\gamma = (\gamma A)_V \colon \mathcal{H} \to 2^{\mathcal{H}} \\ \mathcal{B}_\gamma = \gamma P_V \circ B \circ P_V \colon \mathcal{H} \to V. \end{cases} \tag{46}$$

Then the following hold.

(i) *$\mathcal{A}_\gamma$ is maximally monotone.*
(ii) *$\mathcal{B}_\gamma$ is β/γ–cocoercive.*
(iii) *Let $x \in \mathcal{H}$. Then x is a solution to Problem 1.1 if and only if $x \in V$ and*

$$(\exists\, y \in V^{\perp} \cap (Ax+Bx)) \quad \text{such that} \quad x+\gamma(y-P_{V^\perp}Bx) \in \operatorname{zer}(\mathcal{A}_\gamma+\mathcal{B}_\gamma). \tag{47}$$

Proof: (i): Since γA is maximally monotone, the result follows from [41, Proposition 2.1]. (ii): It is a direct consequence of (33). (iii): Let $x \in \mathcal{H}$ be a solution to Problem 1.1. We have $x \in V$ and there exists $y \in V^\perp = N_V x$ such that $y \in Ax + Bx$. Since B is single valued and P_V is linear, it follows from (12) that

$$\begin{aligned} y \in Ax + Bx \quad &\Leftrightarrow \quad \gamma y - \gamma Bx \in \gamma Ax \\ &\Leftrightarrow \quad -\gamma P_V(Bx) \in (\gamma A)_V\big(x + \gamma(y - P_{V^\perp}Bx)\big) \\ &\Leftrightarrow \quad 0 \in (\gamma A)_V(x + \gamma(y - P_{V^\perp}Bx)) \\ &\qquad\qquad + \gamma P_V\big(B\big(P_V(x + \gamma(y - P_{V^\perp}Bx))\big)\big) \\ &\Leftrightarrow \quad x + \gamma(y - P_{V^\perp}Bx) \in \operatorname{zer}(\mathcal{A}_\gamma + \mathcal{B}_\gamma), \end{aligned} \tag{48}$$

which yields the result. □

Remark 5.1: Note that the characterizations provided in Proposition 4.1 y Proposition 5.1 are related. Indeed, Proposition 4.1(iii) and Proposition 5.1(iii) yield

$$Z = P_V(\operatorname{Fix}(T_\gamma \circ S_\gamma)) = P_V(\operatorname{zer}(\mathcal{A}_\gamma + \mathcal{B}_\gamma)) \text{ and } R_{N_V}(\operatorname{Fix}(T_\gamma \circ S_\gamma)) = \operatorname{zer}(\mathcal{A}_\gamma + \mathcal{B}_\gamma). \tag{49}$$

5.2. Algorithm and convergence

Theorem 5.2: *Let $\mathcal{H}$, V, A, B, and Z be as in Problem 1.1, let $\gamma \in \left]0,+\infty\right[$, let $\varepsilon \in \left]0, \max\{1, \beta/\gamma\}\right[$, let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (2\beta/\gamma) - \varepsilon]$, and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$. Moreover, let $x_0 \in V$, let $y_0 \in V^\perp$, and, for every $n \in \mathbb{N}$, consider the following routine.*

Step 1. Find $(p_n, q_n) \in \mathcal{H}^2$ such that $x_n - \delta_n\gamma P_V Bx_n + \gamma y_n = p_n + \gamma q_n$

$$\text{and } \frac{P_V q_n}{\delta_n} + P_{V^\perp} q_n \in A\Big(P_V p_n + \frac{P_{V^\perp} p_n}{\delta_n}\Big). \tag{50}$$

Step 2. Set $x_{n+1} = x_n + \lambda_n(P_V p_n - x_n)$ and $y_{n+1} = y_n + \lambda_n(P_{V^\perp} q_n - y_n)$.

Go to *Step* 1.

Then, the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are in V and $V^\perp$, respectively, and the following hold for some $\overline{x} \in Z$ and $\overline{y} \in V^\perp \cap (A\overline{x} + P_V B\overline{x})$.

(i) $x_n \rightharpoonup \overline{x}$ *and* $y_n \rightharpoonup \overline{y}$.
(ii) $x_{n+1} - x_n \to 0$ *and* $y_{n+1} - y_n \to 0$.
(iii) $P_V Bx_n \to P_V B\overline{x}$.

Proof: Since $x_0 \in V$ and $y_0 \in V^\perp$, (50) yields $(x_n)_{n\in\mathbb{N}} \subset V$ and $(y_n)_{n\in\mathbb{N}} \subset V^\perp$. Thus, for every $n \in \mathbb{N}$, it follows from (50) and the linearity of P_V and $P_{V^\perp}$ that

$$P_V p_n + \gamma P_V q_n = P_V(p_n + \gamma q_n) = P_V(x_n - \delta_n\gamma P_V Bx_n + \gamma y_n) = x_n - \delta_n\gamma P_V Bx_n \tag{51}$$

and

$$P_{V^\perp} p_n + \gamma P_{V^\perp} q_n = P_{V^\perp}(p_n + \gamma q_n) = P_{V^\perp}(x_n - \delta_n\gamma P_V Bx_n + \gamma y_n) = \gamma y_n, \tag{52}$$

which yield

$$\begin{cases} P_V q_n &= (x_n - \delta_n\gamma P_V Bx_n - P_V p_n)/\gamma = (x_n - x_{n+1})/(\gamma\lambda_n) - \delta_n P_V Bx_n \\ P_{V^\perp} p_n &= \gamma(y_n - P_{V^\perp} q_n) = \gamma(y_n - y_{n+1})/\lambda_n. \end{cases} \tag{53}$$

On the other hand, from (50) we obtain

$$P_V p_n = x_n + \frac{x_{n+1} - x_n}{\lambda_n} \quad \text{and} \quad P_{V^\perp} q_n = y_n + \frac{y_{n+1} - y_n}{\lambda_n}. \tag{54}$$

Hence, it follows from (53) and (50) that

$$\frac{(x_n - x_{n+1})}{\lambda_n\delta_n\gamma} - P_V Bx_n + y_n + \frac{y_{n+1} - y_n}{\lambda_n} \in A\Big(x_n + \frac{x_{n+1} - x_n}{\lambda_n} + \frac{\gamma(y_n - y_{n+1})}{\lambda_n\delta_n}\Big), \tag{55}$$

or equivalently,

$$\frac{(x_n - x_{n+1})}{\lambda_n\delta_n} - \gamma P_V Bx_n + \gamma y_n + \frac{\gamma(y_{n+1} - y_n)}{\lambda_n} \in \gamma A\Big(x_n + \frac{x_{n+1} - x_n}{\lambda_n} + \frac{\gamma(y_n - y_{n+1})}{\lambda_n\delta_n}\Big). \tag{56}$$

Thus, by using the definition of partial inverse (12) we obtain

$$\frac{(x_n - x_{n+1})}{\lambda_n\delta_n} - \gamma P_V B x_n + \frac{\gamma(y_n - y_{n+1})}{\lambda_n\delta_n} \in (\gamma A)_V\Big(x_n + \gamma y_n + \frac{x_{n+1} - x_n + \gamma(y_{n+1} - y_n)}{\lambda_n}\Big), \quad (57)$$

which can be written equivalently as

$$x_n + \gamma y_n - \delta_n\gamma P_V B x_n - \Big(x_n + \gamma y_n + \frac{x_{n+1} - x_n + \gamma(y_{n+1} - y_n)}{\lambda_n}\Big) \in \delta_n(\gamma A)_V\Big(x_n + \gamma y_n + \frac{x_{n+1} - x_n + \gamma(y_{n+1} - y_n)}{\lambda_n}\Big). \quad (58)$$

Hence, we have

$$x_n + \gamma y_n + \frac{x_{n+1} - x_n + \gamma(y_{n+1} - y_n)}{\lambda_n} = J_{\delta_n(\gamma A)_V}(x_n + \gamma y_n - \delta_n\gamma P_V B x_n), \quad (59)$$

or equivalently,

$$x_{n+1} + \gamma y_{n+1} = x_n + \gamma y_n + \lambda_n\Big(J_{\delta_n(\gamma A)_V}(x_n + \gamma y_n - \delta_n\gamma P_V B x_n) - x_n + \gamma y_n\Big). \quad (60)$$

If, for every $n \in \mathbb{N}$, we denote $r_n = x_n + \gamma y_n$, from (60) and (46) we obtain

$$r_{n+1} = r_n + \lambda_n\big(J_{\delta_n\mathcal{A}_\gamma}(r_n - \delta_n\mathcal{B}_\gamma r_n) - r_n\big). \quad (61)$$

Since $(\delta_n)_{n\in\mathbb{N}} \subset [\varepsilon, 2(\beta/\gamma) - \varepsilon]$, it follows from Proposition 5.1(i)&(ii) and [3, Theorem 2.8] that there exists $\overline{r} \in \operatorname{zer}(\mathcal{A}_\gamma + \mathcal{B}_\gamma)$ such that $r_n \rightharpoonup \overline{r}$, $\mathcal{B}_\gamma r_n \to \mathcal{B}_\gamma\overline{r}$, $r_n - r_{n+1} = \lambda_n(r_n - J_{\delta_n\mathcal{A}_\gamma}(r_n - \delta_n\mathcal{B}_\gamma r_n)) \to 0$. Hence, by taking $\overline{x} = P_V\overline{r}$ and $\overline{y} = P_{V^\perp}\overline{r}/\gamma$, Proposition 5.1(iii) asserts that $\overline{x} \in Z$, $\overline{y} \in V^\perp \cap (A\overline{x} + P_V B\overline{x})$, and the results follow from

$$(\forall (x, y) \in \mathcal{H}^2) \quad \langle x \mid y\rangle = \langle P_V x \mid P_V y\rangle + \langle P_{V^\perp} x \mid P_{V^\perp} y\rangle \quad (62)$$

and the definition of $\mathcal{B}_\gamma$. □

Remark 5.2:

(i) It is known that the forward–backward splitting admits errors in the computations of the operators involved [15]. In our algorithm these inexactitudes have not been considered for simplicity.

(ii) In the particular case when $\gamma < 2\beta$, $\lambda_n \equiv 1$, and $B \equiv 0$, the forward-partial-inverse method reduces to the partial inverse method proposed in [41] for solving (3).

The sequence $(\delta_n)_{n\in\mathbb{N}}$ in Theorem 5.2 can be manipulated in order to accelerate the algorithm. However, as in [41], *Step* 1 in Theorem 5.2 is not always easy to compute. The following result show us a particular case of our method in which *Step* 1 can be obtained explicitly when the resolvent of A is computable.

Corollary 5.3: *Let $\gamma \in]0, 2\beta[$, let $x_0 \in V$, let $y_0 \in V^\perp$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$, and consider the following routine.*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} s_n = x_n - \gamma P_V B x_n + \gamma y_n \\ p_n = J_{\gamma A} s_n \\ y_{n+1} = y_n + (\lambda_n/\gamma)(P_V p_n - p_n) \\ x_{n+1} = x_n + \lambda_n (P_V p_n - x_n). \end{array} \right. \tag{63}$$

Then, the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are in V and $V^\perp$, respectively, and the following hold for some $\overline{x} \in Z$ and $\overline{y} \in V^\perp \cap (A\overline{x} + P_V B\overline{x})$.

(i) $x_n \rightharpoonup \overline{x}$ *and* $y_n \rightharpoonup \overline{y}$.
(ii) $x_{n+1} - x_n \to 0$ *and* $y_{n+1} - y_n \to 0$.
(iii) $P_V B x_n \to P_V B\overline{x}$

Proof: For every $n \in \mathbb{N}$, set $q_n = (s_n - p_n)/\gamma$. It follows from (63) that

$$\begin{cases} \gamma q_n = s_n - p_n \in \gamma A p_n \\ s_n = p_n + \gamma q_n, \end{cases} \tag{64}$$

which yield $x_n - \delta_n \gamma P_V B x_n + \gamma y_n = p_n + \gamma q_n$, $p_n - P_V p_n = P_{V^\perp} p_n = \gamma(y_n - P_{V^\perp} q_n)$, and $q_n \in A p_n$. Therefore, (63) is a particular case of (50) when $\delta_n \equiv 1 \in]0, 2(\beta/\gamma)[$ and the results follow from Theorem 5.2. □

Remark 5.3: Note that, when $V = \mathcal{H}$, (63) reduces to

$$x_{n+1} = x_n + \lambda_n \big(J_{\gamma A}(x_n - \gamma B x_n) - x_n \big), \tag{65}$$

which is the forward–backward splitting with constant step size (see [15] and the references therein).

Remark 5.4: Set $a_n \equiv b_n \equiv 0$ in Theorem 4.2, set $\gamma \in]0, 2\beta[$ and $\delta_n \equiv 1$ in Theorem 5.2, and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$ for some $\varepsilon \in]0, 1[$. Moreover denote by $(x_n^1, y_n^1)_{n\in\mathbb{N}}$ the sequence in $V \times V^\perp$ generated by Theorem 4.2 and by $(x_n^2, y_n^2)_{n\in\mathbb{N}}$ the sequence in $V \times V^\perp$ generated by Theorem 5.2 when $x_0^1 = x_0^2 = x_0 \in V$ and $y_0^1 = y_0^2 = y_0 \in V^\perp$. Then, for every $n \in \mathbb{N}$, $x_n^1 = x_n^2$ and $y_n^1 = y_n^2$. Indeed, $x_0^1 = x_0^2$ and $y_0^1 = y_0^2$ by assumption. Proceeding by mathematical induction, suppose that $x_n^1 = x_n^2 = x_n$ and $y_n^1 = y_n^2 = y_n$. Hence, we deduce from (37), $a_n \equiv b_n \equiv 0$, and (63) that

$$\begin{aligned} x_{n+1}^1 &= x_n^1 + \lambda_n (P_V J_{\gamma A}(x_n^1 - \gamma P_V B x_n^1 + \gamma y_n^1) - x_n^1) \\ &= x_n^2 + \lambda_n (P_V J_{\gamma A}(x_n^2 - \gamma P_V B x_n^2 + \gamma y_n^2) - x_n^2) \\ &= x_{n+1}^2. \end{aligned} \tag{66}$$

Moreover, since $P_{V^\perp} = \mathrm{Id} - P_V$, we obtain

$$\begin{aligned} y_{n+1}^1 &= y_n^1 - (\lambda_n/\gamma) P_{V^\perp} J_{\gamma A}(x_n^1 - \gamma P_V B x_n^1 + \gamma y_n^1) \\ &= y_n^2 - (\lambda_n/\gamma) P_{V^\perp} J_{\gamma A}(x_n^2 - \gamma P_V B x_n^2 + \gamma y_n^2) \\ &= y_{n+1}^2, \end{aligned} \tag{67}$$

which yields the result. Therefore, both algorithms are the same in this case. However, even if both methods are very similar, they can be used differently depending on the nature of each problem. Indeed, the algorithm proposed in Theorem 4.2 allows for explicit errors in the computation of the operators involved in the general case and the relaxation parameters $(\lambda_n)_{n\in\mathbb{N}}$ are allowed to be greater than those of the method in Theorem 5.2. On the other hand, the method in Theorem 5.2 allows for a dynamic step size δ_n in the general case, which is not permitted in the algorithm proposed in Theorem 4.2.

6. Applications

In this section we study two applications of our algorithms. First we study the problem of finding a zero of the sum of m maximally monotone operators and a cocoercive operator and, next, we study the variational case. Connections with other methods in this framework are also provided.

6.1. Inclusion involving the sum of m monotone operators

Let us consider the following problem.

Problem 6.1: Let $(\mathsf{H}, |\cdot|)$ be a real Hilbert space, for every $i \in \{1,\ldots,m\}$, let $\mathsf{A}_i\colon \mathsf{H} \to 2^{\mathsf{H}}$ be a maximally monotone operator, and let $\mathsf{B}\colon \mathsf{H} \to \mathsf{H}$ be a β–cocoercive operator. The problem is to

$$\text{find} \quad \mathsf{x} \in \mathsf{H} \quad \text{such that} \quad \mathsf{0} \in \sum_{i=1}^{m} \mathsf{A}_i\mathsf{x} + \mathsf{B}\mathsf{x}, \tag{68}$$

under the assumption that such a solution exists.

Problem 6.1 has several applications in image processing, principally in the variational setting (see, e.g., [17, 23] and the references therein), variational inequalities [44, 45], partial differential equations [33], and economics [28, 35], among others. In [23, 46] two different methods for solving Problem 6.1 are proposed. In [46] auxiliary variables are included for solving a more general problem including linear transformations and additional strongly monotone operators. This generality does not exploits the intrinsic properties of Problem 6.1 and it restricts the choice of the parameters involved. On the other hand, the method in [23] takes into advantage the structure of the problem, but involves restricting relaxation parameters and errors. We provide an alternative version to the latter method, which allows for a wider class of errors and relaxation parameters. The method is obtained as a consequence of Theorem 4.2 and the version obtained from Theorem 5.2 is also examined.

Let us provide a connection between Problem 6.1 and Problem 1.1 via product space techniques. Let $(\omega_i)_{1\leq i\leq m}$ be real numbers in $]0,1[$ such that $\sum_{i=1}^{m}\omega_i = 1$, let $\mathcal{H}$ be the real Hilbert space obtained by endowing the Cartesian product H^m with the scalar product and associated norm respectively defined by

$$\langle\cdot \mid \cdot\rangle : (x,y) \mapsto \sum_{i=1}^{m} \omega_i \langle \mathsf{x}_i \mid \mathsf{y}_i\rangle \quad \text{and} \quad \|\cdot\|\colon x \mapsto \sqrt{\sum_{i=1}^{m} \omega_i|\mathsf{x}_i|^2}, \tag{69}$$

where $x = (\mathsf{x}_i)_{1\le i\le m}$ is a generic element of $\mathcal{H}$. Define

$$\begin{cases} V = \big\{x = (\mathsf{x}_i)_{1\le i\le m} \in \mathcal{H} \mid \mathsf{x}_1 = \cdots = \mathsf{x}_m\big\} \\ j\colon \mathsf{H} \to V \subset \mathcal{H}\colon \mathsf{x} \mapsto (\mathsf{x},\ldots,\mathsf{x}) \\ A\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \frac{1}{\omega_1}\mathsf{A}_1\mathsf{x}_1 \times \cdots \times \frac{1}{\omega_m}\mathsf{A}_m\mathsf{x}_m \\ B\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto (\mathsf{B}\mathsf{x}_1,\ldots,\mathsf{B}\mathsf{x}_m). \end{cases} \tag{70}$$

Proposition 6.2: *Let* H, $(\mathsf{A}_i)_{1\le i\le m}$, *and* B *be as in Problem 6.1, and let* V, j, A, *and* B *be as in* (70). *Then the following hold.*

(i) V *is a closed vector subspace of* $\mathcal{H}$, $P_V\colon (\mathsf{x}_i)_{1\le i\le m} \mapsto j(\sum_{i=1}^m \omega_i\mathsf{x}_i)$, *and*

$$\begin{aligned} N_V\colon\ \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\mapsto \begin{cases} V^\perp = \big\{x = (\mathsf{x}_i)_{1\le i\le m} \in \mathcal{H} \mid \sum_{i=1}^m \omega_i\mathsf{x}_i = 0\big\}, & \text{if } x \in V; \\ \varnothing, & \text{otherwise.} \end{cases} \end{aligned} \tag{71}$$

(ii) $j\colon \mathsf{H} \to V$ *is a bijective isometry and* $j^{-1}\colon (\mathsf{x},\ldots,\mathsf{x}) \mapsto \mathsf{x}$.

(iii) A *is a maximally monotone operator and, for every* $\gamma \in \left]0,+\infty\right[$, $J_{\gamma A}\colon (\mathsf{x}_i)_{1\le i\le m} \mapsto (J_{\gamma\mathsf{A}_i/\omega_i}\mathsf{x}_i)$.

(iv) B *is* β*–cocoercive,* $B(j(\mathsf{x})) = j(\mathsf{B}\mathsf{x})$, *and* $B(V) \subset V$.

(v) *For every* $\mathsf{x} \in \mathsf{H}$, x *is a solution to Problem 6.1 if and only if* $j(\mathsf{x}) \in \operatorname{zer}(A + B + N_V)$.

Proof: (i)&(ii): They follow from (6) and easy computations. (iii): See [6, Proposition 23.16]. (iv): Let $x = (\mathsf{x}_i)_{1\le i\le m}$ and $y = (\mathsf{y}_i)_{1\le i\le m}$ be in $\mathcal{H}$. Then, it follows from (70) and the β–cocoercivity of B that

$$\langle Bx - By \mid x - y\rangle = \sum_{i=1}^m \omega_i\langle \mathsf{B}\mathsf{x}_i - \mathsf{B}\mathsf{y}_i \mid \mathsf{x}_i - \mathsf{y}_i\rangle \ge \beta\sum_{i=1}^m \omega_i|\mathsf{B}\mathsf{x}_i - \mathsf{B}\mathsf{y}_i|^2 = \beta\|Bx - By\|^2, \tag{72}$$

which yields the cocoercivity of B. The other results are clear from the definition.
(v): Let $\mathsf{x} \in \mathsf{H}$. We have

$$\begin{aligned} 0 \in \sum_{i=1}^m \mathsf{A}_i\mathsf{x} + \mathsf{B}\mathsf{x} \quad &\Leftrightarrow \quad \Big(\exists\, (\mathsf{y}_i)_{1\le i\le m} \in \bigtimes_{i=1}^m \mathsf{A}_i\mathsf{x}\Big) \quad 0 = \sum_{i=1}^m \mathsf{y}_i + \mathsf{B}\mathsf{x} \\ &\Leftrightarrow \quad \Big(\exists\, (\mathsf{y}_i)_{1\le i\le m} \in \bigtimes_{i=1}^m \mathsf{A}_i\mathsf{x}\Big) \quad 0 = \sum_{i=1}^m \omega_i(-\mathsf{y}_i/\omega_i - \mathsf{B}\mathsf{x}) \\ &\Leftrightarrow \quad \Big(\exists\, (\mathsf{y}_i)_{1\le i\le m} \in \bigtimes_{i=1}^m \mathsf{A}_i\mathsf{x}\Big) \\ &\qquad\qquad -(\mathsf{y}_1/\omega_1,\ldots,\mathsf{y}_m/\omega_m) - j(\mathsf{B}\mathsf{x}) \in V^\perp = N_V(j(\mathsf{x})) \\ &\Leftrightarrow \quad 0 \in A(j(\mathsf{x})) + B(j(\mathsf{x})) + N_V(j(\mathsf{x})) \\ &\Leftrightarrow \quad j(\mathsf{x}) \in \operatorname{zer}(A + B + N_V), \end{aligned} \tag{73}$$

which yields the result. □

The following algorithm solves Problem 6.1 and is a direct consequence of Theorem 4.2.

Proposition 6.3: *Let $\gamma \in]0, 2\beta[$, let $\alpha = \max\{2/3, 2\gamma/(\gamma + 2\beta)\}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0, 1/\alpha[$, for every $i \in \{1, \ldots, m\}$, let $(\mathsf{a}_n)_{n\in\mathbb{N}}$ and $(\mathsf{b}_{i,n})_{n\in\mathbb{N}}$ be sequences in H, and suppose that*

$$\sum_{n\in\mathbb{N}} \lambda_n(1 - \alpha\lambda_n) = +\infty \quad \text{and} \quad \max_{1\leq i\leq m} \sum_{n\in\mathbb{N}} \lambda_n\Big(|\mathsf{a}_n| + |\mathsf{b}_{i,n}|\Big) < +\infty. \tag{74}$$

Moreover let $(\mathsf{z}_{i,0})_{1\leq i\leq m} \in \mathsf{H}^m$ and consider the following routine.

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \mathsf{x}_n = \sum_{i=1}^m \omega_i \mathsf{z}_{i,n} \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} \mathsf{s}_{i,n} = 2\mathsf{x}_n - \mathsf{z}_{i,n} - \gamma(\mathsf{B}\mathsf{x}_n + \mathsf{a}_n) \\ \mathsf{p}_{i,n} = J_{\gamma \mathsf{A}_i/\omega_i} \mathsf{s}_{i,n} + \mathsf{b}_{i,n} \\ \mathsf{z}_{i,n+1} = \mathsf{z}_{i,n} + \lambda_n(\mathsf{p}_{i,n} - \mathsf{x}_n). \end{array} \right. \end{array} \right. \tag{75}$$

Then, the following hold for some solution $\overline{\mathsf{x}}$ to Problem 6.1.

(i) $\mathsf{x}_n \rightharpoonup \overline{\mathsf{x}}$.
(ii) $\mathsf{B}\mathsf{x}_n \to \mathsf{B}\overline{\mathsf{x}}$.
(iii) $\mathsf{x}_{n+1} - \mathsf{x}_n \to 0$.

Proof: Set, for every $n \in \mathbb{N}$, $x_n = j(\mathsf{x}_n)$, $a_n = j(\mathsf{a}_n)$, $b_n = (\mathsf{b}_{i,n})_{1\leq i\leq m}$, $y_n = (\mathsf{y}_{i,n})_{1\leq i\leq m}$, $z_n = (\mathsf{z}_{i,n})_{1\leq i\leq m}$, $p_n = (\mathsf{p}_{i,n})_{1\leq i\leq m}$, and $q_n = (\mathsf{q}_{i,n})_{1\leq i\leq m}$. It follows from Proposition 6.2(i) and (75) that, for every $n \in \mathbb{N}$, $x_n = P_V z_n$. Hence, it follows from (70) and Proposition 6.2 that (75) can be written equivalently as

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} x_n = P_V z_n \\ y_n = (x_n - z_n)/\gamma \\ s_n = x_n - \gamma P_V\big(Bx_n + a_n\big) + \gamma y_n \\ p_n = J_{\gamma A} s_n + b_n \\ z_{n+1} = z_n + \lambda_n(p_n - x_n). \end{array} \right. \tag{76}$$

Moreover, it follows from (69) that

$$\|a_n\| + \|b_n\| = |\mathsf{a}_n| + \sqrt{\sum_{i=1}^m \omega_i |\mathsf{b}_{i,n}|^2} \leq |\mathsf{a}_n| + \sum_{i=1}^m |\mathsf{b}_{i,n}| \tag{77}$$

and, hence, (74) yields $\sum_{n\in\mathbb{N}} \lambda_n(\|a_n\| + \|b_n\|) < +\infty$. Altogether, Theorem 4.2 and Proposition 6.2(v) yield the results. □

Remark 6.1:

(i) In the particular case when $(\lambda_n)_{n\in\mathbb{N}}$ is such that $0 < \underline{\lim}\,\lambda_n \leq \overline{\lim}\,\lambda_n < 1/\alpha$ and the errors are summable, the algorithm (75) reduces to the method in [23]. Condition (74) allows for a larger class of errors and relaxation parameters.
(ii) Set $\mathsf{a}_n \equiv 0$, for every $i \in \{1, \ldots, m\}$, set $\mathsf{b}_{i,n} \equiv 0$, let $\gamma \in]0, 2\beta[$, and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$ for some $\varepsilon \in]0, 1[$. Then it follows from Remark 5.4 that the algorithm in Proposition 6.3 coincides with the routine:

let $\mathsf{x}_0 \in \mathsf{H}$, let $(\mathsf{y}_{i,0})_{1\leq i\leq m} \in \mathsf{H}^m$ such that $\sum_{i=1}^m \omega_i \mathsf{y}_{i,0} = 0$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} \mathsf{s}_{i,n} = \mathsf{x}_n - \gamma \mathsf{B}\mathsf{x}_n + \gamma \mathsf{y}_{i,n} \\ \mathsf{p}_{i,n} = J_{\gamma \mathsf{A}_i/\omega_i} \mathsf{s}_{i,n} \\ \mathsf{y}_{i,n+1} = \mathsf{y}_{i,n} + (\lambda_n/\gamma)(\sum_{i=1}^m \omega_i \mathsf{p}_{i,n} - \mathsf{p}_{i,n}) \end{array} \right. \\ \mathsf{x}_{n+1} = \mathsf{x}_n + \lambda_n(\sum_{i=1}^m \omega_i \mathsf{p}_{i,n} - \mathsf{x}_n) \end{array} \right. \tag{78}$$

which is the method proposed in Corollary 5.3 applied to Problem 6.1. In the particular case when $\mathsf{B} = 0$, $\gamma = 1$, and $\lambda_n \equiv 1$, (78) reduces to [16, Corollary 2.6].

(iii) It follows from (38) that, in the case when $B = 0$, the method proposed in Proposition 6.3 follows from the iteration

$$(\forall n \in \mathbb{N}) \quad z_{n+1} = z_n + \lambda_n(T_\gamma z_n + b_n - z_n) \tag{79}$$

where A and V are defined in (70). This method is very similar to the algorithm proposed in [16, Theorem 2.5]. Indeed the only difference is that instead of the operator $T_\gamma = (\mathrm{Id} + R_{\gamma A} R_{N_V})/2$ used in Proposition 6.3, in [16, Theorem 2.5] is used the operator $(\mathrm{Id} + R_{N_V} R_{\gamma A})/2$.

Corollary 6.4: *Let $\gamma \in \,]0, +\infty[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0, 3/2[$, for every $i \in \{1, \ldots, m\}$, let $(\mathsf{b}_{i,n})_{n\in\mathbb{N}}$ be sequences in H, and suppose that*

$$\sum_{n\in\mathbb{N}} \lambda_n(3 - 2\lambda_n) = +\infty \quad \textit{and} \quad \max_{1\leq i\leq m} \sum_{n\in\mathbb{N}} \lambda_n \|\mathsf{b}_{i,n}\| < +\infty. \tag{80}$$

Moreover, let $(\mathsf{z}_{1,0}, \mathsf{z}_{2,0}) \in \mathsf{H}^2$ and consider the following routine.

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \mathsf{x}_n = (\mathsf{z}_{1,n} + \mathsf{z}_{2,n})/2 \\ \mathsf{p}_{1,n} = J_{2\gamma \mathsf{A}_1}(\mathsf{z}_{2,n}) + \mathsf{b}_{1,n} \\ \mathsf{p}_{2,n} = J_{2\gamma \mathsf{A}_2}(\mathsf{z}_{1,n}) + \mathsf{b}_{2,n} \\ \mathsf{z}_{1,n+1} = \mathsf{z}_{1,n} + \lambda_n(\mathsf{p}_{1,n} - \mathsf{x}_n) \\ \mathsf{z}_{2,n+1} = \mathsf{z}_{2,n} + \lambda_n(\mathsf{p}_{2,n} - \mathsf{x}_n). \end{array} \right. \tag{81}$$

Then, the following hold for some solution $\overline{\mathsf{x}} \in \mathrm{zer}(\mathsf{A}_1 + \mathsf{A}_2)$.

(i) $\mathsf{x}_n \rightharpoonup \overline{\mathsf{x}}$.
(ii) $\mathsf{x}_{n+1} - \mathsf{x}_n \to 0$.

Proof: Is a direct consequence of Proposition 6.3 in the particular case when $m = 2$, $B = 0$, $\alpha = 2/3$, and $\omega_1 = \omega_2 = 1/2$. □

Remark 6.2:

(i) The most popular method for finding a zero of the sum of two maximally monotone operators is the Douglas–Rachford splitting [31, 43], in which the resolvents of the operators involved are computed sequentially. In the case when these resolvents are hard to compute, Corollary 6.4 provides an alternative method which computes in parallel both resolvents. This method is different to the parallel algorithm proposed in [10, Corollary 3.4].
(ii) For every $i \in \{1, \ldots, m\}$, set $\mathsf{b}_{i,n} \equiv 0$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$ for some $\varepsilon \in \,]0, 1[$. Then it follows from Remark 5.4 that the algorithm in

Corollary 6.4 coincides with the routine: let $\mathsf{x}_0 \in \mathsf{H}$, let $\mathsf{v}_0 \in \mathsf{H}$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \mathsf{s}_{1,n} = \mathsf{x}_n + \gamma \mathsf{v}_n \\ \mathsf{s}_{2,n} = \mathsf{x}_n - \gamma \mathsf{v}_n \\ \mathsf{p}_{1,n} = J_{2\gamma \mathsf{A}_1} \mathsf{s}_{1,n} \\ \mathsf{p}_{2,n} = J_{2\gamma \mathsf{A}_2} \mathsf{s}_{2,n} \\ \mathsf{v}_{n+1} = \mathsf{v}_n + (\lambda_n/(2\gamma))(\mathsf{p}_{2,n} - \mathsf{p}_{1,n}) \\ \mathsf{x}_{n+1} = (1-\lambda_n)\mathsf{x}_n + (\lambda_n/2)(\mathsf{p}_{1,n} + \mathsf{p}_{2,n}), \end{array} \right. \tag{82}$$

which is the method proposed in (78) applied to find a zero of $\mathsf{A}_1 + \mathsf{A}_2$ when $\omega_1 = \omega_2 = 1/2$ and $\mathsf{y}_{1,n} = -\mathsf{y}_{2,n} = \mathsf{v}_n$.

6.2. Variational case

We apply the results of the previous sections to minimization problems. Let us first recall some standard notation and results [6, 47]. We denote by $\Gamma_0(\mathcal{H})$ be the class of lower semicontinuous convex functions $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ such that $\operatorname{dom} f = \big\{x \in \mathcal{H} \mid f(x) < +\infty\big\} \neq \varnothing$. Let $f \in \Gamma_0(\mathcal{H})$. The function $f + \|\cdot - z\|^2/2$ possesses a unique minimizer, which is denoted by $\operatorname{prox}_f z$. Alternatively,

$$\operatorname{prox}_f = (\operatorname{Id} + \partial f)^{-1} = J_{\partial f}, \tag{83}$$

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 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 6.5: *Let V be a closed vector subspace of $\mathcal{H}$, let $f \in \Gamma_0(\mathcal{H})$ be such that $V \cap \operatorname{dom} f \neq \varnothing$, let $g\colon \mathcal{H} \to \mathbb{R}$ be differentiable and convex. Then the following hold.*

(i) $\operatorname{zer}(\partial f + \nabla g + N_V) \subset \operatorname{Argmin}(f + g + \iota_V)$.
(ii) *Suppose that one of the following is satisfied.*
  (a) $\operatorname{Argmin}(f + g + \iota_V) \neq \varnothing$ *and* $0 \in \operatorname{sri}(\operatorname{dom} f - V)$.
  (b) $\operatorname{Argmin}(f + g + \iota_V) \subset \operatorname{Argmin} f \cap \operatorname{Argmin}(g + \iota_V) \neq \varnothing$.
  Then $\operatorname{zer}(\partial f + \nabla g + N_V) \neq \varnothing$.

Proof: (i): Since $\operatorname{dom} g = \mathcal{H}$, it follows from [6, Corollary 16.38(iii)] that $\partial(f + g) = \partial f + \nabla g$. Hence, it follows from $V \cap \operatorname{dom} f \neq \varnothing$, [6, Proposition 16.5(ii)], and [6, Theorem 16.2] that $\operatorname{zer}(\partial f + \nabla g + N_V) = \operatorname{zer}(\partial(f+g) + N_V) \subset \operatorname{zer}(\partial(f + g + \iota_V)) = \operatorname{Argmin}(f + g + \iota_V)$.

(ii)(a): Since $\operatorname{dom} g = \mathcal{H}$ yields $\operatorname{dom}(f+g) = \operatorname{dom} f$, $\operatorname{sri}(\operatorname{dom} f - V) = \operatorname{sri}(\operatorname{dom}(f+g) - \operatorname{dom} \iota_V)$. Therefore, it follows from Fermat's rule ([6, Theorem 16.2]) and [6, Theorem 16.37(i)] that, for every $x \in \mathcal{H}$,

$$\begin{aligned} \varnothing \neq \operatorname{Argmin}(f + g + \iota_V) &= \operatorname{zer} \partial\big(f + g + \iota_V\big) \\ &= \operatorname{zer}\big(\partial(f+g) + N_V\big) \\ &= \operatorname{zer}\big(\partial f + \nabla g + N_V\big). \end{aligned} \tag{84}$$

(ii)(b): Using [6, Corollary 16.38(iii)] and (i), from standard convex analysis we have

$$\begin{aligned}\operatorname{Argmin} f \cap \operatorname{Argmin}(g+\iota_V) &= \operatorname{zer}\partial f \cap \operatorname{zer}\partial(g+\iota_V)\\ &= \operatorname{zer}\partial f \cap \operatorname{zer}(\nabla g + N_V)\\ &\subset \operatorname{zer}(\partial f + \nabla g + N_V)\\ &\subset \operatorname{Argmin}(f+g+\iota_V). \end{aligned} \tag{85}$$

Therefore, the hypothesis yields $\operatorname{zer}(\partial f+\nabla g+N_V) = \operatorname{Argmin} f \cap \operatorname{Argmin}(g+\iota_V) \neq \varnothing$. □

The problem under consideration in this section is the following.

Problem 6.6: Let V be a closed vector subspace of $\mathcal{H}$, let $f \in \Gamma_0(\mathcal{H})$, and let $g\colon \mathcal{H} \to \mathbb{R}$ be a differentiable convex function such that ∇g is β^{-1}–Lipschitzian. The problem is to

$$\underset{x\in V}{\operatorname{minimize}} \;\; f(x)+g(x). \tag{86}$$

Problem 6.6 has several applications in partial differential equations [33, Section 3], signal and image processing [2, 12, 13, 17, 20, 21], and traffic theory [3, 39] among other fields.

In the particular case when $V = \mathcal{H}$, Problem 6.6 has been widely studied, the forward-backward splitting can solve it (see [3, 15] and the references therein), and several applications to multicomponent image processing can be found in [9] and [11]. In the case when $g \equiv 0$, the partial inverse method in [42] solves Problem 6.6 with some applications to convex programming. In the general setting, Problem 6.6 can be solved by methods developed in [10, 17, 23] but without exploiting the structure of the problem. Indeed, in the algorithms presented in [10, 17] it is necessary to compute $\operatorname{prox}_g = (\operatorname{Id} + \nabla g)^{-1}$ and, hence, they do not exploit the fact that ∇g is single-valued. In [23] the method proposed computes explicitly ∇g, however, it generates auxiliary variables for obtaining P_V via product space techniques, which may be numerically costly in problems with a big number of variables. This is because this method does not exploit the vector subspace properties of V. The following result provides a method which exploit the whole structure of the problem and it follows from Proposition 5.3 applied to optimization problems.

Proposition 6.7: *Let $\mathcal{H}$, V, f, and g be as in Problem 6.6, let $\gamma \in \left]0, 2\beta\right[$, let $\alpha = \max\{2/3, 2\gamma/(\gamma+2\beta)\}$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\left]0, 1/\alpha\right[$, let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$, and suppose that*

$$\sum_{n\in\mathbb{N}} \lambda_n(1-\alpha\lambda_n) = +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \lambda_n(\|a_n\| + \|b_n\|) < +\infty \tag{87}$$

and that

$$\operatorname{zer}(\partial f + \nabla g + N_V) \neq \varnothing. \tag{88}$$

Moreover let $z_0 \in \mathcal{H}$ and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} x_n = P_V z_n \\ y_n = (x_n - z_n)/\gamma \\ s_n = x_n - \gamma P_V\big(\nabla g(x_n) + a_n\big) + \gamma y_n \\ p_n = \operatorname{prox}_{\gamma f} s_n + b_n \\ z_{n+1} = z_n + \lambda_n(p_n - x_n). \end{array} \right. \tag{89}$$

Then, the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are in V and $V^\perp$, respectively, and the following hold for some solution $\overline{x}$ to Problem 6.6 and some $\overline{y} \in V^\perp \cap \big(\partial f(\overline{x}) + P_V \nabla g(\overline{x})\big)$.

(i) $x_n \rightharpoonup \overline{x}$ *and* $y_n \rightharpoonup \overline{y}$.
(ii) $x_{n+1} - x_n \to 0$ *and* $y_{n+1} - y_n \to 0$.
(iii) $\sum_{n\in\mathbb{N}} \lambda_n \|P_V\big(\nabla g(x_n) - \nabla g(\overline{x})\big)\|^2 < +\infty$.

Proof: It follows from Baillon–Haddad theorem [4] (see also [5]) that ∇g is β–cocoercive and, in addition, ∂f is maximally monotone. Therefore, the results follow from Theorem 4.2, Proposition 6.5(i), and (83) by taking $A = \partial f$ and $B = \nabla g$. □

Remark 6.3:

(i) Conditions for assuring condition (88) are provided in Proposition 6.5(ii).
(ii) Set $a_n \equiv 0$ and $b_n \equiv 0$, let $\gamma \in \left]0, 2\beta\right[$, and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 1]$ for some $\varepsilon \in \left]0,1\right[$. Then it follows from Remark 5.4 that the algorithm in Proposition 6.7 coincides with the routine: let $x_0 \in V$, let $y_0 \in V^\perp$, and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} s_n = x_n - \gamma P_V \nabla g(x_n) + \gamma y_n \\ p_n = \operatorname{prox}_{\gamma f} s_n \\ y_{n+1} = y_n + (\lambda_n/\gamma)(P_V p_n - p_n) \\ x_{n+1} = x_n + \lambda_n(P_V p_n - x_n), \end{array} \right. \tag{90}$$

which is the method proposed in Corollary 5.3 applied to Problem 6.6.
(iii) Recently in [19] an algorithm is proposed for solving simultaneously

$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; f(x) + g(x) + h(Lx), \tag{91}$$

and its dual, where $\mathcal{G}$ is a real Hilbert space, $h \in \Gamma_0(\mathcal{G})$, and $L\colon \mathcal{H} \to \mathcal{G}$ is linear and bounded. In the particular case when $\mathcal{G} = \mathcal{H}$, $h = \iota_V$, and $L = \mathrm{Id}$, (91) reduces to Problem 6.6. In this case, the method is different to (89) and, additionally, it needs a more restrictive condition on the proximity parameter and the gradient step when the constants involved are equal.
(iv) Consider the problem involving N convex functions

$$\underset{\mathsf{x}\in\mathsf{V}}{\text{minimize}} \;\; \sum_{i=1}^{N} \mathsf{f}_i(\mathsf{x}) + \mathsf{g}(\mathsf{x}), \tag{92}$$

where H is a real Hilbert space, V is a closed vector subspace of H, $(\mathsf{f}_i)_{1\le i\le N}$ are functions in $\Gamma_0(\mathsf{H})$, and g is convex differentiable with Lipschitz gradient. Under qualification conditions, (92) can be reduced to Problem 6.1 with $m = N+1$, for every $i \in \{1,\ldots,N\}$, $\mathsf{A}_i = \partial \mathsf{f}_i$, $\mathsf{A}_{N+1} = N_{\mathsf{V}}$, and $\mathsf{B} = \nabla \mathsf{g}$. Hence, Proposition 6.3 provides an algorithm that solves (92), which generalizes the method in [23] in this context by allowing a larger class of relaxation parameters and errors.

Acknowledgements

The author wish to thank Patrick Combettes for helpful discussions. This work was partially supported by CONICYT under grant FONDECYT 3120054 and Anillo ACT 1106, by "Programa de financiamiento basal" from the Center for Mathematical Modeling, Universidad de Chile, and by Project Math-Amsud N 13MATH01.

References

[1] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis,* Birkhäuser, Boston, MA, 1990.
[2] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, *Structure-texture image decomposition – Modeling, algorithms, and parameter selection*, Int. J. Comput. Vis. 67 (2006), pp. 111–136.
[3] H. Attouch, L.M. Briceño-Arias, and P.L. Combettes, *A parallel splitting method for coupled monotone inclusions*, SIAM J. Control Optim. 48 (2010), pp. 3246–3270.
[4] J.-B. Baillon and G. Haddad, *Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones*, Israel J. Math. 26 (1977), pp. 137–150.
[5] H.H. Bauschke and P.L. Combettes, *The Baillon-Haddad theorem revisited*, J. Convex Anal. 17 (2010), pp. 781–787.
[6] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer-Verlag, New York, 2011.
[7] M. Beckmann, C. McGuire, and C. Winsten, *Studies in Economics of Transportation,* Yale University Press, New Haven, 1956.
[8] D.P. Bertsekas and E.M. Gafni, *Projection methods for variational inequalities with application to the traffic assignment problem*, Math. Programming Stud. 17 (1982), pp. 139–159.
[9] L.M. Briceño-Arias and P.L. Combettes, *Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery*, Numer. Math. Theory Methods Appl. 2 (2009), pp. 485–508.
[10] L.M. Briceño-Arias and P.L. Combettes, *A monotone+skew splitting model for composite monotone inclusions in duality*, SIAM J. Optim. 21 (2011), pp. 1230–1250.
[11] L.M. Briceño-Arias, P.L. Combettes, J.-C. Pesquet, and N. Pustelnik, *Proximal algorithms for multicomponent image processing*, J. Math. Imaging Vision 41 (2011), pp. 3–22.
[12] C.L. Byrne, *Iterative oblique projection onto convex sets and the split feasibility problem*, Inverse Problems 18 (2002), pp. 441–453.
[13] A. Chambolle and P.L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math. 76 (1997), pp. 167–188.
[14] P.L. Combettes, *Quasi-Fejérian analysis of some optimization algorithms*, in: *Inherently Parallel Algorithms for Feasibility and Optimization*, D. Butnariu, Y. Censor, and S. Reich, eds., Elsevier, Amsterdam, 2001, pp. 115–152.
[15] P.L. Combettes, *Solving monotone inclusions via compositions of nonexpansive averaged operators*, Optimization 53 (2004), pp. 475–504.
[16] P.L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators*, J. Convex Anal. 16 (2009), pp. 727–748.
[17] P.L. Combettes and J.-C. Pesquet, *A proximal decomposition method for solving convex variational inverse problems*, Inverse Problems 24(6) (2008), 065014, 27 pp.
[18] P.L. Combettes and J.-C. Pesquet, *Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators*, Set-Valued Var. Anal. 20 (2012), pp. 307–330.
[19] L. Condat, *A generic first-order primal-dual method for convex optimization involving Lipschitzian, proximable and linear composite terms*, http://hal.archives-ouvertes.fr/hal-00609728/en/.
[20] I. Daubechies, M. Defrise, and C. De Mol, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math. 57 (2004), pp. 1413–1457.
[21] B. Eicke, *Iteration methods for convexly constrained ill-posed problems in Hilbert space*, Numer. Funct. Anal. Optim. 13 (1992), pp. 413–429.

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

[23] H. Raguet, J. Fadili, and G. Peyré, *Generalized forward-backward splitting*, http://arxiv.org/abs/1108.4404.

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

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

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

[27] A. Haraux, *Nonlinear Evolution Equations: Global Behavior of Solutions*, Lecture Notes in Math., vol. 841, Springer-Verlag, New York, 1981.

[28] A. Jofré, R.T. Rockafellar, and R.J.-B. Wets, *Variational inequalities and economic equilibrium*, Math. Oper. Res. 32 (2007), pp. 32–50.

[29] M.A. Krasnosel'skiĭ, *Two remarks on the method of successive approximations*, Uspehi Mat. Nauk (N.S.) 10 (1955), pp. 123–127.

[30] J.-L. Lions and G. Stampacchia, *Variational inequalities*, Comm. Pure Appl. Math. 20 (1967), pp. 493–519.

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

[32] W.R. Mann, *Mean value methods in iteration*, Proc. Amer. Math. Soc. 4 (1953), pp. 506–510.

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

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

[35] T. Pennanen, *Introduction to convex optimization in financial markets*, Math. Program. 134 (2012), pp. 157–186.

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

[37] R.T. Rockafellar, *Network Flows and Monotropic Optimization,* Wiley, New York, 1984.

[38] A. Sabharwal and L.C. Potter, *Convexly constrained linear inverse problems: Iterative least-squares and regularization*, IEEE Trans. Signal Process. 46 (1998), pp. 2345–2352.

[39] Y. Sheffi, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods,* Prentice-Hall, Englewood Cliffs, NJ, 1985.

[40] R.E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Mathematical Surveys and Monographs, vol. 49, Amer. Math. Soc., Providence, RI, 1997.

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

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

[43] B.F. Svaiter, *On weak convergence of the Douglas-Rachford method*, SIAM J. Control Optim. 49 (2011), pp. 280–287.

[44] P. Tseng, *Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming*, Math. Programming 48 (1990), pp. 249–263.

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

[46] B.C. Vũ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Adv. Comput. Math. DOI: 10.1007/s10444-011-9254-8, 2011.

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

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