SOLVING MONOTONE INCLUSIONS VIA COMPOSITIONS OF NONEXPANSIVE AVERAGED OPERATORS

Patrick L. Combettes
Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie – Paris 6
75005 Paris, France
E-mail: plc@math.jussieu.fr

Abstract

A unified fixed point theoretic framework is proposed to investigate the asymptotic behavior of algorithms for finding solutions to monotone inclusion problems. The basic iterative scheme under consideration involves nonstationary compositions of perturbed averaged nonexpansive operators. The analysis covers proximal methods for common zero problems as well as various splitting methods for finding a zero of the sum of monotone operators.

Keywords: Averaged operator; Douglas-Rachford method; Forward-backward method; Monotone inclusion; Monotone operator; Proximal point algorithm.

1 Introduction

Let $\mathcal{H}$ be a real Hilbert space, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone operator, and let $J_A = (\mathrm{Id} + A)^{-1}$ denote its resolvent. A basic problem that arises in several branches of applied mathematics (see for instance [20, 29, 33, 51, 52, 54, 60] and the references therein) is to

$$\text{Find } x \in \mathcal{H} \quad \text{such that} \quad 0 \in Ax. \tag{1.1}$$

In this synthetic formulation, the operator A can often be decomposed as a sum of two or more maximal monotone operators $(A_i)_{i \in I}$ [2, 26, 33, 38, 43, 48, 53, 55, 56], which leads to problems of the form

$$\text{Find } x \in \mathcal{H} \quad \text{such that} \quad 0 \in \sum_{i \in I} A_i x. \tag{1.2}$$

In other applications, the decomposition of A assumes the form of an intersection [11, 18, 31, 32, 50] and the problem is therefore

$$\text{Find } x \in \mathcal{H} \quad \text{such that} \quad 0 \in \bigcap_{i \in I} A_i x. \tag{1.3}$$

There is a vast literature on the topic of solving the above monotone inclusion problems. In the present paper, we propose a fixed point setting that unifies and extends a large number of approaches and convergence results. The operators under consideration will be averaged nonexpansive operators.

Definition 1.1 [4] *Let $\alpha \in]0,1[$. An operator $T\colon \operatorname{dom} T = \mathcal{H} \to \mathcal{H}$ is nonexpansive if*

$$(\forall (x,y) \in \mathcal{H}^2) \ \ \|Tx - Ty\| \leq \|x - y\| \tag{1.4}$$

and α-averaged if $T = (1-\alpha)\operatorname{Id} + \alpha R$ for some nonexpansive operator $R\colon \operatorname{dom} R = \mathcal{H} \to \mathcal{H}$. The class of α-averaged operators on $\mathcal{H}$ is denoted by $\mathcal{A}(\alpha)$. In particular, $\mathcal{A}(\frac{1}{2})$ is the class of firmly nonexpansive operators.

Firmly nonexpansive operators have a very natural connection with the basic problem (1.1). Indeed, an operator $T\colon \operatorname{dom} T = \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive if and only if it is the resolvent of a maximal monotone operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, i.e., $T = J_A$ (this fact appears implicitly in Minty's classical paper [44] and it is stated more explicitly in [15, 26, 43, 45]). On the other hand, it is an easy matter to see that (1.1) is equivalent to the problem of finding a fixed point of J_A. Since for firmly nonexpansive operators the successive approximation method converges weakly to a fixed point [14], it can be used to solve (1.1). The weak convergence to a zero of A of the sequence $(x_n)_{n\in\mathbb{N}}$ constructed as

$$x_{n+1} = Tx_n \ \text{ where } \ T = J_A, \tag{1.5}$$

was thus established in [41] in the case when A is the subdifferential of a lower semicontinuous convex function.

Let us now turn to the sum problem (1.2) in the case of two maximal monotone operators $A, B\colon \mathcal{H} \to 2^{\mathcal{H}}$, i.e.,

$$\text{Find } x \in \mathcal{H} \quad \text{such that} \quad 0 \in Ax + Bx. \tag{1.6}$$

An elementary form of this problem is to solve the equation $u = Ax + Bx$ in $\mathbb{R}^N$, where A and B are positive definite matrices. In the 1950s, several implicit decomposition methods have been proposed to solve this problem in connection with the numerical solution of partial differential equations [57, 58] and some of them have served as a basis to develop algorithms for solving the monotone inclusion (1.6). The Douglas-Rachford algorithm [24] for $u = Ax + Bx$ is described by the recursion

$$\begin{cases} y_{n+\frac{1}{2}} - y_n + Ay_{n+\frac{1}{2}} + By_n = u \\ y_{n+1} - y_{n+\frac{1}{2}} - By_n + By_{n+1} = 0, \end{cases} \tag{1.7}$$

the Peaceman-Rachford algorithm [47] by

$$\begin{cases} y_{n+\frac{1}{2}} - y_n + Ay_{n+\frac{1}{2}} + By_n = u \\ y_{n+1} - y_{n+\frac{1}{2}} + Ay_{n+\frac{1}{2}} + By_{n+1} = u, \end{cases} \tag{1.8}$$

and the fractional steps method [36] by

$$\begin{cases} y_{n+\frac{1}{2}} - y_n + Ay_{n+\frac{1}{2}} = u \\ y_{n+1} - y_n + Ay_{n+\frac{1}{2}} + By_{n+1} = u. \end{cases} \tag{1.9}$$

After eliminating the intermediate variable $y_{n+\frac{1}{2}}$ in the Douglas-Rachford algorithm (1.7), we obtain

$$y_{n+1} = (\mathrm{Id} + B)^{-1}\big((\mathrm{Id} + A)^{-1}(\mathrm{Id} - B) + B + u\big)y_n = J_B\big(J_A(\mathrm{Id} - B + u) + B\big)y_n. \tag{1.10}$$

In [38], it was observed that with the change of variable $x_n = (\mathrm{Id} + B)y_n$, the identities $J_A - AJ_A = 2J_A - \mathrm{Id}$ and $J_B - BJ_B = 2J_B - \mathrm{Id}$ make it possible to rewrite (1.10) for $u = 0$ as

$$x_{n+1} = \big(J_A(J_B - BJ_B) + BJ_B\big)x_n = \big(J_A(2J_B - \mathrm{Id}) + \mathrm{Id} - J_B\big)x_n. \tag{1.11}$$

It was shown there that, for general maximal monotone operators A and B, the operator $J_A(2J_B - \mathrm{Id}) + \mathrm{Id} - J_B$ is firmly nonexpansive and the iteration (1.11) converges weakly to some point x such that J_Bx solves (1.6). Let us note that the recursion (1.11) can also be obtained with the same procedure from the iteration

$$\begin{cases} y_{n+\frac{1}{2}} - y_n + Ay_{n+\frac{1}{2}} + By_n = u \\ y_{n+1} - y_n + Ay_{n+\frac{1}{2}} + By_{n+1} = u, \end{cases} \tag{1.12}$$

which was studied in [36, section V-II] for single-valued monotone operators in $\mathbb{R}^N$. In the case of the Peaceman-Rachford algorithm (1.8), proceeding as above, we arrive at the iteration

$$x_{n+1} = (\mathrm{Id} - A)J_A(\mathrm{Id} - B)J_Bx_n = (2J_A - \mathrm{Id})(2J_B - \mathrm{Id})x_n, \tag{1.13}$$

which was investigated in [38] for general maximal monotone operators. Let us add that for the fractional steps method (1.9), this same procedure leads to what is known as the backward-backward method, namely

$$x_{n+1} = J_AJ_Bx_n. \tag{1.14}$$

Another splitting method of interest is the so-called forward-backward algorithm

$$x_{n+1} = J_A(\mathrm{Id} - B)x_n, \tag{1.15}$$

which is also meaningful for the general problem (1.6) as long as B is single-valued. Formally, it can be obtained by iterating directly the first equation of (1.7), (1.8), or (1.12) with $u = 0$, $x_n = y_n$ and $x_{n+1} = y_{n+\frac{1}{2}}$, i.e., $x_{n+1} - x_n + Ax_{n+1} + Bx_n = 0$. Here the words "forward" and "backward" refer respectively to the standard notions of a forward difference (explicit) step and of a backward difference (implicit) step in numerical analysis.

Just like the above methods, algorithms for solving the common zero problem (1.3) also draw their inspiration from classical linear numerical analysis. Consider the simple realization of (1.3) consisting of solving a linear system of m equation in $\mathbb{R}^m$. The classical Kaczmarz' algorithm [28] iterates $x_{n+1} = P_1 \cdots P_m x_n$, where P_i is the projection operator onto the hyperplane defined by the ith equation. Replacing P_i by more general nonlinear resolvents, we obtain the iteration [18, 25]

$$x_{n+1} = J_{A_1} \cdots J_{A_m} x_n, \tag{1.16}$$

which converges weakly to a solution to (1.3) under the provision that such a point exists; the same is true for the iteration [18, 32, 50]

$$x_{n+1} = \frac{1}{m} \sum_{i=1}^{m} J_{A_i} x_n, \tag{1.17}$$

which is directly inspired by Cimmino's method [16] for solving systems of linear equations in $\mathbb{R}^m$.

Over the years, the algorithms mentioned above have undergone various improvements to gain more flexibility, improve convergence patterns, or incorporate numerical errors. For instance, the basic proximal point algorithm (1.5) has now evolved to [21, 26]

$$x_{n+1} = x_n + \lambda_n (T_n x_n + a_n - x_n), \quad \text{where} \quad T_n = J_{\gamma_n A}. \tag{1.18}$$

Here $\lambda_n \in \,]0, +\infty[$ is a relaxation parameter, $\gamma_n \in \,]0, +\infty[$, and $a_n \in \mathcal{H}$ is an error term that models the inexact computation of $J_{\gamma_n A} x_n$. In [11, 21], a fixed point theoretic framework was developed to study the asymptotic behavior of iterations of type (1.18). This framework, however, fails to cover other algorithms such as the nonstationary version of the forward-backward method (1.15) proposed in [56] (see also [35] for a perturbed model), namely

$$x_{n+1} = T_{1,n} T_{2,n} x_n, \quad \text{where} \quad \begin{cases} T_{1,n} = J_{\gamma_n A}, \\ T_{2,n} = \operatorname{Id} - \gamma_n B. \end{cases} \tag{1.19}$$

On the other hand, the fixed point analysis of this algorithm proposed in [34, 35] is not applicable to some algorithms covered in [11, 21]. In order to study and generalize the above algorithms in a unified framework, we therefore need to introduce a flexible iteration scheme involving a sufficiently broad class of operators. The analysis presented in this paper will revolve around the following algorithm.

Algorithm 1.2 Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set

$$x_{n+1} = x_n + \lambda_n \bigg(T_{1,n} \Big(T_{2,n} \big(\cdots T_{m-1,n} (T_{m,n} x_n + e_{m,n}) + e_{m-1,n} \cdots \big) + e_{2,n} \Big) + e_{1,n} - x_n \bigg), \tag{1.20}$$

where $(T_{i,n})_{1 \le i \le m} \in \bigtimes_{i=1}^{m} \mathcal{A}(\alpha_{i,n})$ with $(\alpha_{i,n})_{1 \le i \le m} \in \,]0,1[^m$, $(e_{i,n})_{1 \le i \le m} \in \mathcal{H}^m$, and $\lambda_n \in \,]0,1]$.

The remainder of the paper is organized as follows. In section 2, we introduce our notation and provide preliminary results. Section 3 is devoted to the convergence analysis of Algorithm 1.2. These results, which are of interest in their own right in constructive fixed point theory, are applied in subsequent sections to study and generalize a number of monotone inclusion algorithms and establish their convergence properties. Section 4 focuses on proximal methods for solving the common zero problem (1.3) when it is feasible. The Douglas-Rachford and Peaceman-Rachford algorithms for the sum problem (1.6) are investigated in section 5. In section 6, we study the forward-backward method for (1.6) and apply it in particular to infeasible common zero problems. Further applications are discussed in section 7.

2 Preliminary results

2.1 Notation

Throughout $\mathbb{N}$ is the set of nonnegative integers and $\mathcal{H}$ is a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$, norm $\|\cdot\|$, and distance d. Id denotes the identity operator on $\mathcal{H}$. The expressions $x_n \rightharpoonup x$ and $x_n \to x$ denote respectively the weak and strong convergence to x of a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$, and $\mathfrak{W}(x_n)_{n\in\mathbb{N}}$ its set of weak cluster points. The subdifferential of a proper function $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ is the set-valued operator

$$\partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \quad \langle y - x \mid u\rangle + f(x) \leq f(y)\big\}. \tag{2.1}$$

$\Gamma_0(\mathcal{H})$ denotes the class of proper, lower semicontinuous convex functions from $\mathcal{H}$ to $\left]-\infty,+\infty\right]$. If $f \in \Gamma_0(\mathcal{H})$, then $\operatorname{prox}_f = J_{\partial f}$ is Moreau's proximity operator [45]; moreover, the Moreau envelope of index $\gamma \in \left]0,+\infty\right[$ of f is the function ${}^{\gamma}f\colon x \mapsto \min_{y\in\mathcal{H}} f(y)+\frac{1}{2\gamma}\|x-y\|^2$. Now let C be a subset of $\mathcal{H}$. Then d_C is the distance function to C, $\operatorname{int} C$ its interior, $\overline{C}$ its closure, and ι_C its indicator function, which takes the value 0 on C and $+\infty$ on its complement. If C is nonempty, closed, and convex, then P_C is the projector onto C and $N_C = \partial\iota_C$ its normal cone operator. Now let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. The sets $\operatorname{dom} A = \{x \in \mathcal{H} \mid Ax \neq \varnothing\}$, $\operatorname{ran} A = \{u \in \mathcal{H} \mid (\exists\, x \in \mathcal{H})\ u \in Ax\}$, and $\operatorname{gr} A = \{(x,u) \in \mathcal{H}^2 \mid u \in Ax\}$ are the domain, the range, and the graph of A, respectively. The inverse A^{-1} of A is the set-valued operator with graph $\{(u,x) \in \mathcal{H}^2 \mid u \in Ax\}$. The resolvent of A is $J_A = (\operatorname{Id} + A)^{-1}$ and its Yosida approximation of index $\gamma \in \left]0,+\infty\right[$ is

$${}^{\gamma}A = \frac{\operatorname{Id} - J_{\gamma A}}{\gamma} = (\gamma \operatorname{Id} + A^{-1})^{-1}. \tag{2.2}$$

It will also be convenient to introduce the "reflection" operator

$$R_A = 2J_A - \operatorname{Id}. \tag{2.3}$$

$\operatorname{Fix} T = \{x \in \mathcal{H} \mid Tx = x\}$ denotes the set of fixed points of an operator $T\colon \mathcal{H} \to \mathcal{H}$. Given operators $(T_k)_{1\leq k\leq m}$ from $\mathcal{H}$ to $\mathcal{H}$ and a strictly positive integer i, we define (the directed composition product)

$$\prod_{k=i}^{m} T_k = \begin{cases} T_i T_{i+1} \cdots T_m, & \text{if } i \leq m;\\ \operatorname{Id}, & \text{otherwise.}\end{cases} \tag{2.4}$$

2.2 Averaged nonexpansive operators

In the case of firmly nonexpansive operators, i.e., $\alpha = \frac{1}{2}$ in Definition 1.1, the following characterizations go back to [59].

Lemma 2.1 *Take $T\colon \mathcal{H} \to \mathcal{H}$ and $\alpha \in \left]0,1\right[$. Then the following properties are equivalent.*

(i) $T \in \mathcal{A}(\alpha)$.

(ii) $(\forall (x,y) \in \mathcal{H}^2)\ \ \|Tx - Ty\|^2 \leq \|x-y\|^2 - \dfrac{1-\alpha}{\alpha}\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2$.

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

Proof. (i) $\Leftrightarrow$ (ii): Set $R = (1-1/\alpha)\,\mathrm{Id} + T/\alpha$ and fix $(x,y) \in \mathcal{H}^2$. Then

$$\|Rx - Ry\|^2 = \left(1 - \frac{1}{\alpha}\right)\|x-y\|^2 + \frac{1}{\alpha}\|Tx-Ty\|^2 - \frac{1}{\alpha}\left(1-\frac{1}{\alpha}\right)\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2. \tag{2.5}$$

In other words,

$$\alpha\big(\|x-y\|^2 - \|Rx-Ry\|^2\big) = \|x-y\|^2 - \|Tx-Ty\|^2 - \frac{1-\alpha}{\alpha}\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2. \tag{2.6}$$

Now observe that (i) $\Leftrightarrow$ R is nonexpansive $\Leftrightarrow$ the left-hand side of (2.6) is nonnegative $\Leftrightarrow$ (ii). (ii) $\Leftrightarrow$ (iii): Write $\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 = \|x-y\|^2 + \|Tx-Ty\|^2 - 2\left\langle x-y \mid Tx-Ty\right\rangle$ in (ii). ☐

As we now show, averaged operators are closed under relaxations, convex combinations, and compositions.

Lemma 2.2 *Let $(T_i)_{1\leq i\leq m}$ be a finite family of operators from $\mathcal{H}$ to $\mathcal{H}$, let $(\omega_i)_{1\leq i\leq m}$ be real numbers in $\left]0,1\right]$ adding up to 1, and let $(\alpha_i)_{1\leq i\leq m}$ be real numbers in $\left]0,1\right[$ such that, for every $i \in \{1,\ldots,m\}$, $T_i \in \mathcal{A}(\alpha_i)$. Then:*

(i) $(\forall i \in \{1,\ldots,m\})(\forall \lambda \in \left]0, 1/\alpha_i\right[)\ \ \mathrm{Id} + \lambda(T_i - \mathrm{Id}) \in \mathcal{A}(\lambda\alpha_i)$.

(ii) $\sum_{i=1}^m \omega_i T_i \in \mathcal{A}(\alpha)$, *with* $\alpha = \max_{1\leq i\leq m} \alpha_i$.

(iii) $T_1 \cdots T_m \in \mathcal{A}(\alpha)$, *with*

$$\alpha = \frac{m}{m - 1 + \dfrac{1}{\max\limits_{1\leq i\leq m} \alpha_i}}. \tag{2.7}$$

(iv) *If* $\bigcap_{i=1}^m \operatorname{Fix} T_i \neq \varnothing$, *then* $\bigcap_{i=1}^m \operatorname{Fix} T_i = \operatorname{Fix} T_1 \cdots T_m = \operatorname{Fix} \sum_{i=1}^m \omega_i T_i$.

Proof. (i): Fix $i \in \{1,\dots,m\}$ and $\lambda \in \left]0, 1/\alpha_i\right[$. Then, $T_i = (1-\alpha_i)\operatorname{Id} + \alpha_i R_i$ for some nonexpansive operator $R_i\colon \mathcal{H} \to \mathcal{H}$. Hence $\operatorname{Id} + \lambda(T_i - \operatorname{Id}) = (1-\lambda\alpha_i)\operatorname{Id} + \lambda\alpha_i R_i \in \mathcal{A}(\lambda\alpha_i)$. (ii): Set $T = \sum_{i=1}^m \omega_i T_i$ and fix $(x,y) \in \mathcal{H}^2$. Since $\alpha = \max_{1\leq i\leq m}\alpha_i$, Lemma 2.1(ii) yields

$$(\forall i \in \{1,\dots,m\})\ \ \|T_i x - T_i y\|^2 + \frac{1-\alpha_i}{\alpha_i}\|(\operatorname{Id} - T_i)x - (\operatorname{Id} - T_i)y\|^2 \leq \|x-y\|^2. \tag{2.8}$$

Hence, by convexity of $\|\cdot\|^2$,

$$\begin{aligned}
&\|Tx - Ty\|^2 + \frac{1-\alpha}{\alpha}\|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2 \\
&\quad = \left\|\sum_{i=1}^m \omega_i T_i x - \sum_{i=1}^m \omega_i T_i y\right\|^2 + \frac{1-\alpha}{\alpha}\left\|\sum_{i=1}^m \omega_i(\operatorname{Id} - T_i)x - \sum_{i=1}^m \omega_i(\operatorname{Id} - T_i)y\right\|^2 \\
&\quad \leq \sum_{i=1}^m \omega_i\|T_i x - T_i y\|^2 + \sum_{i=1}^m \frac{1-\alpha_i}{\alpha_i}\omega_i\|(\operatorname{Id} - T_i)x - (\operatorname{Id} - T_i)y\|^2 \\
&\quad \leq \|x-y\|^2.
\end{aligned} \tag{2.9}$$

(iii): Set $T = T_1 \cdots T_m$, $(\forall i \in \{1,\dots,m\})$ $\kappa_i = \alpha_i/(1-\alpha_i)$, and $\kappa = \max_{1\leq i\leq m}\kappa_i$. In addition, fix $(x,y) \in \mathcal{H}^2$. Then we derive from the convexity of $\|\cdot\|^2$ and Lemma 2.1(ii) that

$$\begin{aligned}
\|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2/m &= \|(x-y) - (T_m x - T_m y) + (T_m x - T_m y) \\
&\quad - (T_{m-1}T_m x - T_{m-1}T_m y) + (T_{m-1}T_m x - T_{m-1}T_m y) - \cdots \\
&\quad - (T_2\cdots T_m x - T_2\cdots T_m y) + (T_2\cdots T_m x - T_2\cdots T_m y) \\
&\quad - (T_1\cdots T_m x - T_1\cdots T_m y)\|^2/m \\
&= \|(\operatorname{Id} - T_m)x - (\operatorname{Id} - T_m)y \\
&\quad + (\operatorname{Id} - T_{m-1})T_m x - (\operatorname{Id} - T_{m-1})T_m y + \cdots \\
&\quad + (\operatorname{Id} - T_1)T_2\cdots T_m x - (\operatorname{Id} - T_1)T_2\cdots T_m y\|^2/m \\
&\leq \|(\operatorname{Id} - T_m)x - (\operatorname{Id} - T_m)y\|^2 \\
&\quad + \|(\operatorname{Id} - T_{m-1})T_m x - (\operatorname{Id} - T_{m-1})T_m y\|^2 + \cdots \\
&\quad + \|(\operatorname{Id} - T_1)T_2\cdots T_m x - (\operatorname{Id} - T_1)T_2\cdots T_m y\|^2 \\
&\leq \kappa_m\big(\|x-y\|^2 - \|T_m x - T_m y\|^2\big) \\
&\quad + \kappa_{m-1}\big(\|T_m x - T_m y\|^2 - \|T_{m-1}T_m x - T_{m-1}T_m y\|^2\big) + \cdots \\
&\quad + \kappa_1\big(\|T_2\cdots T_m x - T_2\cdots T_m y\|^2 - \|T_1\cdots T_m x - T_1\cdots T_m y\|^2\big) \\
&\leq \kappa\big(\|x-y\|^2 - \|Tx - Ty\|^2\big).
\end{aligned} \tag{2.10}$$

Consequently, Lemma 2.1 asserts that $T \in \mathcal{A}(\alpha)$, with $\alpha = m/(m+1/\kappa)$. This is precisely the expression provided in (2.7). (iv): Fix $i \in \{1,\dots,m\}$, $x \in \mathcal{H} \smallsetminus \operatorname{Fix} T_i$, and $y \in \operatorname{Fix} T_i$. Then it follows from Lemma 2.1(ii) that $\|T_i x - y\| < \|x - y\|$, i.e., T_i is attracting in the sense of [9, Definition 2.1]. The two identities therefore follow from [9, Proposition 2.10(i)] and [9, Proposition 2.12(i)]. ☐

Lemma 2.3 *Suppose that* $B\colon \mathcal{H} \to \mathcal{H}$ *and* $\beta \in \left]0,+\infty\right[$ *satisfy* $\beta B \in \mathcal{A}(\frac{1}{2})$, *and let* $\gamma \in \left]0, 2\beta\right[$. *Then,* $\operatorname{Id} - \gamma B \in \mathcal{A}(\frac{\gamma}{2\beta})$.

Proof. Since $\beta B \in \mathcal{A}(\frac{1}{2})$, there exists a nonexpansive operator $R\colon \mathcal{H} \to \mathcal{H}$ such that $B = (\mathrm{Id} + R)/(2\beta)$. In turn,

$$\mathrm{Id} - \gamma B = \left(1 - \frac{\gamma}{2\beta}\right) \mathrm{Id} + \frac{\gamma}{2\beta}(-R) \in \mathcal{A}\left(\frac{\gamma}{2\beta}\right). \tag{2.11}$$

□

2.3 Monotone operators

A set-valued operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is monotone if

$$(\forall (x,u) \in \operatorname{gr} A)(\forall (y,v) \in \operatorname{gr} A) \quad \langle x - y \mid u - v \rangle \geq 0, \tag{2.12}$$

and maximal monotone if, furthermore, $\operatorname{gr} A$ is not properly contained in the graph of any monotone operator $B\colon \mathcal{H} \to 2^{\mathcal{H}}$.

Lemma 2.4 [15, 44] *Let $T\colon \mathcal{H} \to \mathcal{H}$. Then $T \in \mathcal{A}(\frac{1}{2})$ if and only if $T = J_A$ for some maximal monotone operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$.*

Lemma 2.5 *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone operator and let $\gamma \in \left]0, +\infty\right[$. Then*

(i) $\gamma({}^{\gamma}A) \in \mathcal{A}(\frac{1}{2})$.

(ii) *The set*

$$\operatorname{Fix} J_{\gamma A} = A^{-1}(0) = \left({}^{\gamma}A\right)^{-1}(0) \tag{2.13}$$

is closed and convex.

(iii) $\operatorname{gr} A$ *is sequentially weakly-strongly closed in* $\mathcal{H} \times \mathcal{H}$.

(iv) $(\forall z \in A^{-1}(0))(\forall x \in \mathcal{H}) \ \ \|J_A x - x\|^2 \leq \langle z - x \mid J_A x - x \rangle$.

Proof. (i): It follows from Lemma 2.4 that $J_{\gamma A} \in \mathcal{A}(\frac{1}{2})$. However, in view of Lemma 2.1(ii), $J_{\gamma A} \in \mathcal{A}(\frac{1}{2}) \Leftrightarrow \gamma({}^{\gamma}A) = \mathrm{Id} - J_{\gamma A} \in \mathcal{A}(\frac{1}{2})$. (ii): [3, Proposition 3.5.6.1]. (iii): [3, Proposition 3.5.6.2]. (iv): Fix $z \in A^{-1}(0)$, $x \in \mathcal{H}$, and set $T = J_A$. Then (2.13) yields $z = Tz$. Hence, we deduce from Lemma 2.4 and Lemma 2.1(iii) that $\|Tx - z\|^2 \leq \langle Tx - z \mid x - z \rangle$. Hence, $\langle Tx - z \mid Tx - x \rangle \leq 0$ and, in turn, $\|Tx - x\|^2 \leq \langle z - x \mid Tx - x \rangle$. □

Our analysis will also exploit the following properties, which involve the reflection operators of (2.3).

Lemma 2.6 *Let $A, B\colon \mathcal{H} \to 2^{\mathcal{H}}$ be two maximal monotone operators, let $\gamma \in \left]0, +\infty\right[$, and set $T = R_{\gamma A} R_{\gamma B}$. Then*

(i) *T is nonexpansive.*

(ii) $\frac{1}{2}(T + \mathrm{Id}) = J_{\gamma A}(2J_{\gamma B} - \mathrm{Id}) - J_{\gamma B} + \mathrm{Id}$.

(iii) $(A + B)^{-1}(0) = J_{\gamma B}(\mathrm{Fix}\, T)$.

Proof. (i): Lemma 2.4 asserts that $J_{\gamma A}$ and $J_{\gamma B}$ belong to $\mathcal{A}(\frac{1}{2})$. Therefore, $R_{\gamma A}$ and $R_{\gamma B}$ are nonexpansive and it follows that $R_{\gamma A}R_{\gamma B}$ is nonexpansive as the composition of two nonexpansive operators. (ii): $T + \mathrm{Id} = 2J_{\gamma A}(2J_{\gamma B} - \mathrm{Id}) - (2J_{\gamma B} - \mathrm{Id}) + \mathrm{Id} = 2\big(J_{\gamma A}(2J_{\gamma B} - \mathrm{Id}) - J_{\gamma B} + \mathrm{Id}\big)$. (iii): For every $y \in \mathcal{H}$

$$\begin{aligned} 0 \in Ay + By \;&\Leftrightarrow\; (\exists\, x \in \mathcal{H})\;\; y - x \in \gamma Ay \;\text{ and }\; x - y \in \gamma By \\ &\Leftrightarrow\; (\exists\, x \in \mathcal{H})\;\; 2y - x \in (\mathrm{Id} + \gamma A)y \;\text{ and }\; y = J_{\gamma B}x \\ &\Leftrightarrow\; (\exists\, x \in \mathcal{H})\;\; y = J_{\gamma A}(R_{\gamma B}x) \;\text{ and }\; y = J_{\gamma B}x \\ &\Leftrightarrow\; (\exists\, x \in \mathcal{H})\;\; x = 2y - R_{\gamma B}x = R_{\gamma A}(R_{\gamma B}x) \;\text{ and }\; y = J_{\gamma B}x \\ &\Leftrightarrow\; (\exists\, x \in \mathrm{Fix}\, T)\;\; y = J_{\gamma B}x \\ &\Leftrightarrow\; y \in J_{\gamma B}(\mathrm{Fix}\, T). \end{aligned} \tag{2.14}$$

□

2.4 Quasi-Fejér sequences

The subsequent convergence analyses will be greatly simplified by the following facts.

Lemma 2.7 [49, Lemma 2.2.2] *Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence in $[0, +\infty[$, let $(\beta_n)_{n\in\mathbb{N}}$ be a summable sequence in $[0, +\infty[$, and let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a summable sequence in $[0, +\infty[$ such that $(\forall n \in \mathbb{N})\;\; \alpha_{n+1} \leq (1 + \beta_n)\alpha_n + \varepsilon_n$. Then $(\alpha_n)_{n\in\mathbb{N}}$ converges.*

Lemma 2.8 *Let C be a nonempty closed subset of $\mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ which is quasi-Fejér monotone with respect to C, i.e., there exists a summable sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ in $[0, +\infty[$ such that*

$$(\forall x \in C)(\forall n \in \mathbb{N})\;\; \|x_{n+1} - x\| \leq \|x_n - x\| + \varepsilon_n. \tag{2.15}$$

Then:

(i) *The sequence $(x_n)_{n\in\mathbb{N}}$ is bounded.*

(ii) *The sequence $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in C if and only if $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset C$.*

(iii) *The sequence $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in C if and only if $\underline{\lim}\, d_C(x_n) = 0$.*

(iv) *If $\mathrm{int}\, C \neq \varnothing$, then the sequence $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in $\mathcal{H}$.*

Proof. (i): Lemma 2.7. (ii): [21, Proposition 3.2(i) & Theorem 3.8]. (iii): [21, Theorem 3.11(iv)]. (iv): [21, Proposition 3.10]. □

3 Convergence of Algorithm 1.2

Theorem 3.1 *Let* $(x_n)_{n\in\mathbb{N}}$ *be an arbitrary orbit of Algorithm 1.2. Suppose that*

$$G = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_{1,n}\cdots T_{m,n} \neq \varnothing \tag{3.1}$$

and

$$(\forall i \in \{1,\ldots,m\}) \quad \sum_{n\in\mathbb{N}} \lambda_n \|e_{i,n}\| < +\infty. \tag{3.2}$$

Then:

(i) *The sequence* $(x_n)_{n\in\mathbb{N}}$ *is quasi-Fejér monotone with respect to* G.

(ii) $(\forall x \in G)\ \max\limits_{1\le i\le m} \sum\limits_{n\in\mathbb{N}} \lambda_n \dfrac{1-\alpha_{i,n}}{\alpha_{i,n}} \left\| (\operatorname{Id} - T_{i,n}) \prod\limits_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod\limits_{k=i+1}^{m} T_{k,n} x \right\|^2 < +\infty.$

(iii) $\sum\limits_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) \left\| \prod\limits_{k=1}^{m} T_{k,n} x_n - x_n \right\|^2 < +\infty.$

Proof. Let $n \in \mathbb{N}$ and fix $x \in G$. Then we can rewrite (1.20) as

$$x_{n+1} = z_n + e_n, \tag{3.3}$$

where

$$\begin{cases} z_n = x_n + \lambda_n(y_n - x_n) \\ y_n = T_{1,n}\cdots T_{m,n} x_n \\ e_n = \lambda_n \left(T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n + e_{m,n}) + e_{m-1,n} \cdots \big) + e_{2,n} \Big) + e_{1,n} - T_{1,n}\cdots T_{m,n} x_n \right). \end{cases} \tag{3.4}$$

Since $x \in \operatorname{Fix} T_{1,n}\cdots T_{m,n}$ and the operators $(T_{i,n})_{1\le i\le m}$ are nonexpansive, we have

$$\begin{aligned} \|x_{n+1} - x\| &\le \|z_n - x\| + \|e_n\| && (3.5)\\ &= \|(1-\lambda_n)(x_n - x) + \lambda_n(y_n - x)\| + \|e_n\| \\ &\le (1-\lambda_n)\|x_n - x\| + \lambda_n \|T_{1,n}\cdots T_{m,n}x_n - T_{1,n}\cdots T_{m,n}x\| + \|e_n\| \\ &\le \|x_n - x\| + \|e_n\|. && (3.6) \end{aligned}$$

It also follows from the nonexpansivity of the operators $(T_{i,n})_{1\leq i\leq m}$ that

$$\begin{aligned}
\|e_n\|/\lambda_n &\leq \|e_{1,n}\| + \\
&\quad \left\| T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\Big) - T_{1,n}\cdots T_{m,n}x_n \right\| \\
&\leq \|e_{1,n}\| + \\
&\quad \left\| T_{2,n}\Big(T_{3,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{3,n}\Big) + e_{2,n} - T_{2,n}\cdots T_{m,n}x_n \right\| \\
&\leq \|e_{1,n}\| + \|e_{2,n}\| + \\
&\quad \left\| T_{3,n}\Big(T_{4,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{4,n}\Big) + e_{3,n} - T_{3,n}\cdots T_{m,n}x_n \right\| \\
&\leq \sum_{i=1}^{m} \|e_{i,n}\|. \qquad (3.7)
\end{aligned}$$

Accordingly, we deduce from (3.2) that

$$\sum_{n\in\mathbb{N}} \|e_n\| < +\infty \qquad (3.8)$$

and, thereby, that (i) holds.

We now turn to (ii) and (iii). We first observe that (i) and Lemma 2.8(i) imply that

$$\zeta = \sup_{n\in\mathbb{N}} \|x_n - x\| < +\infty. \qquad (3.9)$$

On the other hand, it follows from (3.5) and (3.4) that

$$\begin{aligned}
\|x_{n+1} - x\|^2 &\leq \|z_n - x\|^2 + (2\|z_n - x\| + \|e_n\|)\|e_n\| \\
&\leq \|(1-\lambda_n)(x_n - x) + \lambda_n(y_n - x)\|^2 + \nu\|e_n\| \\
&= (1-\lambda_n)\|x_n - x\|^2 + \lambda_n\|y_n - x\|^2 \\
&\quad -\lambda_n(1-\lambda_n)\|y_n - x_n\|^2 + \nu\|e_n\|, \qquad (3.10)
\end{aligned}$$

where $\nu = 2\zeta + \sup_{n\in\mathbb{N}}\|e_n\| < +\infty$. Next, we derive from Lemma 2.1 that

$$(\forall i \in \{1,\ldots,m\})(\forall (u,v) \in \mathcal{H}^2)$$
$$\|T_{i,n}u - T_{i,n}v\|^2 \leq \|u - v\|^2 - \frac{1-\alpha_{i,n}}{\alpha_{i,n}}\|(\mathrm{Id} - T_{i,n})u - (\mathrm{Id} - T_{i,n})v\|^2. \qquad (3.11)$$

Repeated applications of (3.11) yield

$$\begin{aligned}
\|y_n - x\|^2 &= \left\| \prod_{k=1}^{m} T_{k,n} x_n - \prod_{k=1}^{m} T_{k,n} x \right\|^2 \\
&\leq \left\| \prod_{k=2}^{m} T_{k,n} x_n - \prod_{k=2}^{m} T_{k,n} x \right\|^2 \\
&\quad - \frac{1-\alpha_{1,n}}{\alpha_{1,n}} \left\| (\operatorname{Id} - T_{1,n}) \prod_{k=2}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{1,n}) \prod_{k=2}^{m} T_{k,n} x \right\|^2 \\
&\leq \left\| \prod_{k=3}^{m} T_{k,n} x_n - \prod_{k=3}^{m} T_{k,n} x \right\|^2 \\
&\quad - \frac{1-\alpha_{2,n}}{\alpha_{2,n}} \left\| (\operatorname{Id} - T_{2,n}) \prod_{k=3}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{2,n}) \prod_{k=3}^{m} T_{k,n} x \right\|^2 \\
&\quad - \frac{1-\alpha_{1,n}}{\alpha_{1,n}} \left\| (\operatorname{Id} - T_{1,n}) \prod_{k=2}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{1,n}) \prod_{k=2}^{m} T_{k,n} x \right\|^2 \\
&\leq \|x_n - x\|^2 - \sum_{i=1}^{m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 . \quad (3.12)
\end{aligned}$$

Combining (3.10) and (3.12), we obtain

$$\begin{aligned}
\|x_{n+1} - x\|^2 &\leq \|x_n - x\|^2 - \lambda_n \sum_{i=1}^{m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 \\
&\quad - \lambda_n (1-\lambda_n) \|y_n - x_n\|^2 + \nu \|e_n\|. \quad (3.13)
\end{aligned}$$

Consequently, for every $N \in \mathbb{N}$,

$$\begin{aligned}
&\sum_{n=0}^{N} \lambda_n \sum_{i=1}^{m} \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 \\
&\quad + \sum_{n=0}^{N} \lambda_n (1-\lambda_n) \|y_n - x_n\|^2 \leq \|x_0 - x\|^2 - \|x_{N+1} - x\|^2 + \nu \sum_{n=0}^{N} \|e_n\|. \quad (3.14)
\end{aligned}$$

In view of (3.8), taking the limit as $N \to +\infty$ yields

$$\max_{1 \leq i \leq m} \sum_{n \in \mathbb{N}} \lambda_n \frac{1-\alpha_{i,n}}{\alpha_{i,n}} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 < +\infty \quad (3.15)$$

and

$$\sum_{n \in \mathbb{N}} \lambda_n (1-\lambda_n) \left\| \prod_{k=1}^{m} T_{k,n} x_n - x_n \right\|^2 < +\infty. \quad (3.16)$$

We have thus proven (ii) and (iii). ◻

If we combine Theorem 3.1 and Lemma 2.8(ii), we obtain our main convergence result.

Theorem 3.2 *Suppose that the following conditions are satisfied.*

(i) $G = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_{1,n} \cdots T_{m,n} \neq \varnothing$.

(ii) *For every subsequence* $(x_{k_n})_{n\in\mathbb{N}}$ *of an orbit* $(x_n)_{n\in\mathbb{N}}$ *generated by Algorithm 1.2, we have*

$$\begin{cases} (\forall x \in G) \ \max_{1\le i\le m} \sum_{n\in\mathbb{N}} \lambda_n \dfrac{1-\alpha_{i,n}}{\alpha_{i,n}} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 < +\infty \\ \sum_{n\in\mathbb{N}} \lambda_n (1-\lambda_n) \left\| \prod_{k=1}^{m} T_{k,n} x_n - x_n \right\|^2 < +\infty \\ x_{k_n} \rightharpoonup y \end{cases}$$

$$\Rightarrow \ y \in G. \tag{3.17}$$

(iii) $(\forall i \in \{1,\ldots,m\}) \ \ \sum_{n\in\mathbb{N}} \lambda_n \|e_{i,n}\| < +\infty$.

Then every orbit of Algorithm 1.2 converges weakly to a point in G.

Proof. For every $n \in \mathbb{N}$, $T_{1,n} \cdots T_{m,n}$ is nonexpansive as a composition of nonexpansive operators and $\operatorname{Fix} T_{1,n} \cdots T_{m,n}$ is therefore closed. In turn, G is closed and the claim therefore follows from Theorem 3.1 and Lemma 2.8(ii). ◻

Likewise, we derive from Theorem 3.1 and Lemma 2.8(iii)–(iv) the following strong convergence statements.

Theorem 3.3 *Suppose that the following conditions are satisfied.*

(i) $G = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_{1,n} \cdots T_{m,n} \neq \varnothing$.

(ii) *For every orbit* $(x_n)_{n\in\mathbb{N}}$ *generated by Algorithm 1.2, we have*

$$\begin{cases} (\forall x \in G) \ \max_{1\le i\le m} \sum_{n\in\mathbb{N}} \lambda_n \dfrac{1-\alpha_{i,n}}{\alpha_{i,n}} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 < +\infty \\ \sum_{n\in\mathbb{N}} \lambda_n (1-\lambda_n) \left\| \prod_{k=1}^{m} T_{k,n} x_n - x_n \right\|^2 < +\infty \end{cases}$$

$$\Rightarrow \ \underline{\lim}\, d_G(x_n) = 0. \tag{3.18}$$

(iii) $(\forall i \in \{1,\ldots,m\})\ \ \sum_{n\in\mathbb{N}} \lambda_n \|e_{i,n}\| < +\infty$.

Then every orbit of Algorithm 1.2 converges strongly to a point in G. This is true in particular if $\operatorname{int} G \neq \varnothing$ *and condition* (ii) *in Theorem 3.2 holds.*

Remark 3.4 A special case of interest is when

$$\underline{\lim}\, \lambda_n > 0 \quad \text{and} \quad (\forall i \in \{1,\ldots,m\})\ \ \overline{\lim}\, \alpha_{i,n} < 1. \tag{3.19}$$

First of all, in this setting, (ii) in Theorem 3.1 yields

$$(\forall x \in G)\ \ \max_{1\le i\le m} \sum_{n\in\mathbb{N}} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 < +\infty. \tag{3.20}$$

Now, fix $x \in G$. Then, recalling that $G = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} \prod_{k=1}^{m} T_{k,n}$ and invoking the convexity of $\|\cdot\|^2$, we obtain, for every $n \in \mathbb{N}$,

$$\begin{aligned}
\left\| \prod_{k=1}^{m} T_{k,n} x_n - x_n \right\|^2 &= \left\| \left(\operatorname{Id} - \prod_{k=1}^{m} T_{k,n} \right) x_n - \left(\operatorname{Id} - \prod_{k=1}^{m} T_{k,n} \right) x \right\|^2 \\
&= \left\| \sum_{i=1}^{m} (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - \sum_{i=1}^{m} (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 \\
&\le m \sum_{i=1}^{m} \left\| (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x_n - (\operatorname{Id} - T_{i,n}) \prod_{k=i+1}^{m} T_{k,n} x \right\|^2 .
\end{aligned} \tag{3.21}$$

It therefore follows from (3.20) that (iii) in Theorem 3.1 can be replaced by

$$\sum_{n\in\mathbb{N}} \left\| \prod_{k=1}^{m} T_{k,n} x_n - x_n \right\|^2 < +\infty. \tag{3.22}$$

In turn, (3.17) and (3.18) can be modified accordingly.

4 Common zero problem

We consider the common zero problem (1.3), where $(A_i)_{i\in I}$ is a countable family of maximal monotone operators. Its set of solutions is $S = \bigcap_{i\in I} A_i^{-1}(0)$.

For clarity, we first restate Algorithm 1.2 and Theorem 3.2 in the case when $m = 1$.

Algorithm 4.1 Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set

$$x_{n+1} = x_n + \lambda_n\big(T_{1,n}x_n + e_{1,n} - x_n\big), \tag{4.1}$$

where $T_{1,n} \in \mathcal{A}(\alpha_{1,n})$ with $\alpha_{1,n} \in \,]0,1[$, $e_{1,n} \in \mathcal{H}$, and $\lambda_n \in \,]0,1]$.

Theorem 4.2 *Suppose that the following conditions are satisfied.*

(i) $G = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_{1,n} \neq \varnothing$.

(ii) *For every subsequence* $(x_{k_n})_{n\in\mathbb{N}}$ *of an orbit* $(x_n)_{n\in\mathbb{N}}$ *generated by Algorithm 4.1, we have*

$$\begin{cases} \sum_{n\in\mathbb{N}} \lambda_n \dfrac{1-\alpha_{1,n}}{\alpha_{1,n}} \|T_{1,n}x_n - x_n\|^2 < +\infty \\ \sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n)\|T_{1,n}x_n - x_n\|^2 < +\infty \\ x_{k_n} \rightharpoonup y \end{cases} \quad\Rightarrow\quad y \in G. \tag{4.2}$$

(iii) $\sum_{n\in\mathbb{N}} \lambda_n \|e_{1,n}\| < +\infty$.

Then every orbit of Algorithm 4.1 converges weakly to a point in G.

Our first application of Theorem 4.2 is the following result on the convergence of a parallel block-iterative proximal method for solving (1.3).

Corollary 4.3 *Suppose that* $S \neq \varnothing$ *and that the following conditions are satisfied:*

(i) *For every* $n \in \mathbb{N}$, I_n *is a nonempty finite subset of* I. *Moreover, there exist strictly positive integers* $(M_i)_{i\in I}$ *such that* $(\forall (i,n) \in I \times \mathbb{N})\;\; i \in \bigcup_{k=n}^{n+M_i-1} I_k$.

(ii) *For every* $i \in I$, $(\gamma_{i,n})_{n\in\mathbb{N}}$ *is a sequence in* $]0,+\infty[$ *such that, for every strictly increasing sequence* $(k_n)_{n\in\mathbb{N}}$ *in* $\mathbb{N}$ *such that* $i \in \bigcap_{n\in\mathbb{N}} I_{k_n}$, $\inf_{n\in\mathbb{N}} \gamma_{i,k_n} > 0$.

(iii) $(\mu_n)_{n\in\mathbb{N}}$ *lies in* $]0,2[$ *and* $0 < \underline{\lim}\, \mu_n \leq \overline{\lim}\, \mu_n < 2$.

(iv) $$(\exists\, \delta \in \,]0,1[)(\forall n \in \mathbb{N}) \begin{cases} (\forall i \in I_n)\;\; \omega_{i,n} \in \,]0,1]\,, \\ \sum_{i\in I_n} \omega_{i,n} = 1, \\ (\exists\, j \in I_n) \begin{cases} \|J_{\gamma_{j,n}A_j}x_n - x_n\| = \max\limits_{i\in I_n} \|J_{\gamma_{i,n}A_i}x_n - x_n\|, \\ \omega_{j,n} \geq \delta. \end{cases} \end{cases}$$

(v) $\sum_{n\in\mathbb{N}} \|\sum_{i\in I_n} \omega_{i,n} a_{i,n}\| < +\infty$.

Take $x_0 \in \mathcal{H}$ and set

$$(\forall n \in \mathbb{N})\ \ x_{n+1} = x_n + \mu_n\bigg(\sum_{i\in I_n} \omega_{i,n}\big(J_{\gamma_{i,n}A_i}x_n + a_{i,n}\big) - x_n\bigg). \tag{4.3}$$

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

Proof. For every $n \in \mathbb{N}$, set

$$T_{1,n} = \operatorname{Id} + \mu_n\left(\sum_{i\in I_n} \omega_{i,n} J_{\gamma_{i,n}A_i} - \operatorname{Id}\right),\ \ \lambda_n = 1,\ \ \alpha_{1,n} = \mu_n/2,\ \text{ and }\ e_{1,n} = \mu_n \sum_{i\in I_n} \omega_{i,n}a_{i,n}. \tag{4.4}$$

Lemma 2.4 yields $(\forall i \in I_n)\ J_{\gamma_{i,n}A_i} \in \mathcal{A}(\frac{1}{2})$. Hence, it follows from (iv) and Lemma 2.2(ii) that $\sum_{i\in I_n} \omega_{i,n}J_{\gamma_{i,n}A_i} \in \mathcal{A}(\frac{1}{2})$ and, in turn, from Lemma 2.2(i) that $T_{1,n} \in \mathcal{A}(\alpha_{1,n})$. Thus, in view of (4.4), (4.3) is a special case of the recursion (4.1) governing Algorithm 4.1. It now remains to verify the assumptions of Theorem 4.2. First, since $S \neq \varnothing$, it results from Lemma 2.2(iv) and (2.13) that

$$(\forall n \in \mathbb{N})\ \ \operatorname{Fix} T_{1,n} = \operatorname{Fix} \sum_{i\in I_n} \omega_{i,n}J_{\gamma_{i,n}A_i} = \bigcap_{i\in I_n} \operatorname{Fix} J_{\gamma_{i,n}A_i} = \bigcap_{i\in I_n} A_i^{-1}(0). \tag{4.5}$$

Hence, it follows from (i) that $G = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_{1,n} = \bigcap_{i\in I} A_i^{-1}(0) = S \neq \varnothing$, which supplies item (i) in Theorem 4.2. Next, we derive from (4.4), (iii), and (v) that

$$\sum_{n\in\mathbb{N}} \lambda_n\|e_{1,n}\| = \sum_{n\in\mathbb{N}} \|e_{1,n}\| \leq 2\sum_{n\in\mathbb{N}} \bigg\|\sum_{i\in I_n} \omega_{i,n}a_{i,n}\bigg\| < +\infty, \tag{4.6}$$

which establishes item (iii) in Theorem 4.2. Finally, fix $j \in I$ and suppose that $x_{k_n} \rightharpoonup y$. We have $G = S$ and $\sum_{n\in\mathbb{N}} \lambda_n(1-\alpha_{1,n})\,\|T_{1,n}x_n - x_n\|^2/\alpha_{1,n} = \sum_{n\in\mathbb{N}}(2-\mu_n)\,\|T_{1,n}x_n - x_n\|^2/\mu_n$. Hence, in view of (iii), it suffices to check that $T_{1,n}x_n - x_n \to 0 \Rightarrow 0 \in A_jy$ to verify item (ii) in Theorem 4.2. So suppose $T_{1,n}x_n - x_n \to 0$. We first deduce from (4.1) and (4.6) that

$$\|x_{n+1} - x_n\| \leq \|T_{1,n}x_n - x_n\| + \|e_{1,n}\| \to 0. \tag{4.7}$$

On the other hand, in view of (i), there exists a sequence $(p_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$(\forall n \in \mathbb{N})\ \ k_n \leq p_n \leq k_n + M_j - 1 < k_{n+1} \leq p_{n+1} \quad \text{and} \quad j \in I_{p_n}. \tag{4.8}$$

Now set

$$(\forall n \in \mathbb{N})\ \ y_n = J_{\gamma_{j,p_n}A_j}x_{p_n}\ \text{ and }\ u_n = \frac{x_{p_n} - y_n}{\gamma_{j,p_n}}. \tag{4.9}$$

By (4.7), $\|x_{p_n} - x_{k_n}\| \leq \sum_{l=k_n}^{k_n+M_j-2} \|x_{l+1} - x_l\| \leq (M_j - 1)\max_{k_n\leq l\leq k_n+M_j-2} \|x_{l+1} - x_l\| \to 0$. Hence, $x_{p_n} - x_{k_n} \to 0$ and, in turn, $x_{p_n} \rightharpoonup y$. Now fix $z \in S$ and set $\gamma = \inf_{n\in\mathbb{N}} \gamma_{j,p_n}$ (> 0 by (ii)), $\zeta = \sup_{n\in\mathbb{N}} \|z - x_n\|$ ($< +\infty$ by (3.9)), and $\varepsilon = \underline{\lim}\, \mu_n/2$ (> 0 by (iii)). Then (4.9), (iv), Lemma 2.5(iv), and the Cauchy-Schwarz inequality imply that, for n large enough,

$$\begin{aligned}
\delta\gamma^2\|u_n\|^2 &\leq \delta\|y_n - x_{p_n}\|^2 \\
&\leq \delta \max_{i\in I_{p_n}} \|J_{\gamma_{i,p_n}A_i} x_{p_n} - x_{p_n}\|^2 \\
&\leq \sum_{i\in I_{p_n}} \omega_{i,p_n} \|J_{\gamma_{i,p_n}A_i} x_{p_n} - x_{p_n}\|^2 \\
&\leq \left\langle z - x_{p_n} \,\middle|\, \sum_{i\in I_{p_n}} \omega_{i,p_n} J_{\gamma_{i,p_n}A_i} x_{p_n} - x_{p_n} \right\rangle \\
&\leq \zeta \|T_{1,p_n} x_{p_n} - x_{p_n}\|/\varepsilon.
\end{aligned} \tag{4.10}$$

Altogether, $u_n \to 0$ and $y_n - x_{p_n} \to 0$. Therefore $y_n \rightharpoonup y$, while (4.9) gives $Ay_n \ni u_n \to 0$. In view of Lemma 2.5(iii), we conclude that $0 \in A_j y$. ☐

Remark 4.4 (Strong convergence) Using Theorem 3.3, we infer immediately that the convergence is strong in Corollary 4.3 if $\operatorname{int} S \neq \varnothing$. Another sufficient condition is that some operator A_j in $(A_i)_{i\in I}$ have a boundedly relatively compact domain (the intersection of its closure with any closed ball is compact). Indeed, we already have $x_n \rightharpoonup y \in S$. Now extract a subsequence $(x_{p_n})_{n\in\mathbb{N}}$ such that $j \in \bigcap_{n\in\mathbb{N}} I_{p_n}$ and define $(y_n)_{n\in\mathbb{N}}$ as in (4.9). It remains to check (3.18) with $G = S$. As above, we assume $T_{1,n}x_n - x_n \to 0$ and obtain $y_n - x_{p_n} \to 0$ and $y_n \rightharpoonup y$. At the same time, for every $n \in \mathbb{N}$, $y_n \in \operatorname{ran} J_{\gamma_{j,p_n}A_j} = \operatorname{dom}(\operatorname{Id} + \gamma_{j,p_n}A_j) \subset \overline{\operatorname{dom} A_j}$. Accordingly, $y_n \to y$ and, in turn, $x_{p_n} \to y \in S$, whence $\underline{\lim}\, d_S(x_n) = 0$.

Corollary 4.3 covers and extends several known results. For instance, if $a_{i,n} \equiv 0$ and each I_n reduces to a singleton, then Corollary 4.3 reduces to [11, Corollary 6.1(i)]. On the other hand, when $\gamma_{i,n} \equiv \gamma_i$ and $a_{i,n} \equiv 0$, we recover the results of [18] and, in particular, those of [32, section 4] if we further assume $\omega_{i,n} \equiv \omega_i$ and $\mu_n \equiv 1$. In another direction, if we now take each A_i to be the normal cone operator to a nonempty closed convex set S_i, then the operator $J_{\gamma_{i,n}A_i}$ is the projector P_i onto S_i and Corollary 4.3 and Remark 4.4 capture various convergence results for projection methods for solving convex feasibility problems, see [9, 19] and the references therein. In particular, if $I = \{1, \ldots, m\}$ is a finite index set, we recover the classical results of [27] for the cyclic projection method

$$x_{n+1} = x_n + \mu_n\big(P_{n\,(\text{modulo } m)+1} x_n - x_n\big), \quad \text{where } \varepsilon \leq \mu_n \leq 2 - \varepsilon. \tag{4.11}$$

Another special case of interest is when a single operator is involved. Then (1.3) reduces to (1.1), (4.3) reduces to the standard proximal point algorithm (1.18), and Corollary 4.3 reduces to [26, Theorem 3] and, in particular, to [52, Theorem 1] for $\lambda_n \equiv 1$. In these results, the parameters $(\gamma_n)_{n\in\mathbb{N}}$ must be bounded away from zero. An alternative use of Theorem 4.2 leads to the following corollary, in which this condition is weakened.

Corollary 4.5 *Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0, +\infty[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0, 1]$. Suppose that $0 \in \operatorname{ran} A$, $\sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty$, $\underline{\lim}\, \lambda_n > 0$, and $\sum_{n\in\mathbb{N}} (1 - \lambda_n)\gamma_n/\gamma_{n+1} < +\infty$. Take $x_0 \in \mathcal{H}$ and set*

$$(\forall n \in \mathbb{N}) \;\; x_{n+1} = x_n + \lambda_n\big(J_{\gamma_n A} x_n - x_n\big). \tag{4.12}$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $A^{-1}(0)$.

Proof. Let $n \in \mathbb{N}$ and set $y_n = J_{\gamma_n A}x_n$ and $u_n = (x_n - y_n)/\gamma_n$. Then $u_n \in Ay_n$ and $y_n - y_{n+1} = \gamma_{n+1}u_{n+1} + y_n - x_{n+1} = \gamma_{n+1}u_{n+1} - (1-\lambda_n)\gamma_n u_n$. Hence, by monotonicity,

$$\begin{aligned} 0 &\leq \langle y_n - y_{n+1} \mid u_n - u_{n+1}\rangle / \gamma_{n+1} \\ &= \langle u_{n+1} - \beta_n u_n \mid u_n - u_{n+1}\rangle \\ &= (1+\beta_n)\langle u_{n+1} \mid u_n\rangle - \|u_{n+1}\|^2 - \beta_n\|u_n\|^2 \\ &\leq (1+\beta_n)\langle u_{n+1} \mid u_n\rangle - \|u_{n+1}\|^2, \end{aligned} \tag{4.13}$$

where $\beta_n = (1-\lambda_n)\gamma_n/\gamma_{n+1}$. Hence, it follows from Cauchy-Schwarz that $\|u_{n+1}\| \leq (1+\beta_n)\|u_n\|$ and, in turn, from Lemma 2.7 that $(\|u_n\|)_{n\in\mathbb{N}}$ converges. Now set $T_{1,n} = J_{\gamma_n A}$ (hence $\alpha_{1,n} = \frac{1}{2}$) and $e_{1,n} = 0$. Then (4.12) is a special instance of (4.1) and the claim will follow from Theorem 4.2 by establishing (4.2). To this end, it is enough to suppose that $\sum_{n\in\mathbb{N}}\|y_n - x_n\|^2 < +\infty$ and that $x_{k_n} \rightharpoonup y$, and to show that $0 \in Ay$. We therefore have $\sum_{n\in\mathbb{N}}\gamma_n^2\|u_n\|^2 < +\infty$ and, since $\sum_{n\in\mathbb{N}}\gamma_n^2 = +\infty$, we obtain $\underline{\lim}\,\|u_n\| = 0$. Accordingly, $u_n \to 0$ since $(\|u_n\|)_{n\in\mathbb{N}}$ converges. Thus $Ay_n \ni u_n \to 0$ and $y_{k_n} \rightharpoonup y$ since $y_n - x_n \to 0$. Lemma 2.5(iii) then yields $0 \in Ay$. ☐

In particular, for $\lambda_n \equiv 1$, Corollary 4.5 coincides with [13, Proposition 8].

5 Douglas-Rachford and Peaceman-Rachford splitting

We turn our attention to the sum problem (1.6) for two maximal monotone operators $A, B\colon \mathcal{H} \to 2^{\mathcal{H}}$. The Douglas-Rachford and Peaceman-Rachford algorithms proposed in [38] for solving this problem are defined by (1.11) and (1.13), respectively. In this section, we shall investigate a more general form of these algorithms. It will be assumed that the problem is feasible, i.e., $0 \in \operatorname{ran}(A+B)$ (in the case of normal cone operators, the Douglas-Rachford algorithm in the infeasible case is studied in [12]).

Our convergence result for the Douglas-Rachford algorithm will be derived from Theorem 4.2 via the following lemma.

Lemma 5.1 *Let* $T\colon \operatorname{dom} T = \mathcal{H} \to \mathcal{H}$ *be a nonexpansive operator, let* $(\mu_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,1[$*, and let* $(c_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}$*. Suppose that* $\operatorname{Fix} T \neq \varnothing$, $\sum_{n\in\mathbb{N}}\mu_n(1-\mu_n) = +\infty$*, and* $\sum_{n\in\mathbb{N}}\mu_n\|c_n\| < +\infty$*. Take* $x_0 \in \mathcal{H}$ *and set*

$$(\forall n \in \mathbb{N})\;\; x_{n+1} = x_n + \mu_n(Tx_n + c_n - x_n). \tag{5.1}$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\operatorname{Fix} T$.

Proof. The recursion (5.1) is a specialization of (4.1) with

$$(\forall n \in \mathbb{N})\;\; T_{1,n} = \operatorname{Id} + \mu_n(T - \operatorname{Id}) \in \mathcal{A}(\mu_n),\;\; \lambda_n = 1,\;\; \alpha_{1,n} = \mu_n,\;\; \text{and } e_{1,n} = \mu_n c_n. \tag{5.2}$$

It is clear that conditions (i) and (iii) are satisfied in Theorem 4.2. In view of (4.2) and (5.2), to check (ii) it is enough to verify that for an arbitrary suborbit $(x_{k_n})_{n\in\mathbb{N}}$ we have

$$\begin{cases}\sum_{n\in\mathbb{N}}\mu_n(1-\mu_n)\|Tx_n-x_n\|^2<+\infty\\ x_{k_n}\rightharpoonup y\end{cases}\quad\Rightarrow\quad Ty=y. \tag{5.3}$$

To this end, suppose that $\sum_{n\in\mathbb{N}}\mu_n(1-\mu_n)\|Tx_n-x_n\|^2<+\infty$. Since $\sum_{n\in\mathbb{N}}\mu_n(1-\mu_n)=+\infty$, we get $\underline{\lim}\,\|Tx_n-x_n\|=0$. However, it follows from (5.1) that

$$\begin{aligned}(\forall n\in\mathbb{N})\ \ \|Tx_{n+1}-x_{n+1}\| &\leq \|Tx_{n+1}-Tx_n\|+(1-\mu_n)\|Tx_n-x_n\|+\mu_n\|c_n\|\\ &\leq \|x_{n+1}-x_n\|+(1-\mu_n)\|Tx_n-x_n\|+\mu_n\|c_n\|\\ &\leq \|Tx_n-x_n\|+2\mu_n\|c_n\|. \end{aligned}\tag{5.4}$$

Since $\sum_{n\in\mathbb{N}}\mu_n\|c_n\|<+\infty$, the sequence $(\|Tx_n-x_n\|)_{n\in\mathbb{N}}$ converges and therefore $Tx_n-x_n\to 0$. If, in addition, $x_{k_n}\rightharpoonup y$, then it follows at once from the demiclosed principle for nonexpansive operators [14, Lemma 4] that $Ty=y$. ☐

We now establish results on the asymptotic behavior of a perturbed, relaxed extension of the Douglas-Rachford algorithm (1.11).

Corollary 5.2 *Let $\gamma\in\,]0,+\infty[$, let $(\nu_n)_{n\in\mathbb{N}}$ be a sequence in $]0,2[$, and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$. Suppose that $0\in\operatorname{ran}(A+B)$, $\sum_{n\in\mathbb{N}}\nu_n(2-\nu_n)=+\infty$, and $\sum_{n\in\mathbb{N}}\nu_n(\|a_n\|+\|b_n\|)<+\infty$. Take $x_0\in\mathcal{H}$ and set*

$$(\forall n\in\mathbb{N})\ \ x_{n+1}=x_n+\nu_n\Big(J_{\gamma A}\big(2(J_{\gamma B}x_n+b_n)-x_n\big)+a_n-(J_{\gamma B}x_n+b_n)\Big). \tag{5.5}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to some point $x\in\mathcal{H}$ and $J_{\gamma B}x\in(A+B)^{-1}(0)$.

Proof. Recall the notation (2.3) and set

$$(\forall n\in\mathbb{N})\ \ \mu_n=\frac{\nu_n}{2}\ \text{ and }\ c_n=2a_n+R_{\gamma A}(R_{\gamma B}x_n+2b_n)-R_{\gamma A}(R_{\gamma B}x_n), \tag{5.6}$$

and define $T=R_{\gamma A}R_{\gamma B}$. Then it follows from Lemma 2.6(ii) and straightforward manipulations that we can rewrite the updating rule in (5.5) as $x_{n+1}=x_n+\mu_n\big(Tx_n+c_n-x_n\big)$. Since $R_{\gamma A}$ is nonexpansive,

$$\begin{aligned}\sum_{n\in\mathbb{N}}\mu_n\|c_n\| &\leq \sum_{n\in\mathbb{N}}\nu_n\|a_n\|+\sum_{n\in\mathbb{N}}\nu_n\|R_{\gamma A}(R_{\gamma B}x_n+2b_n)-R_{\gamma A}(R_{\gamma B}x_n)\|/2\\ &\leq \sum_{n\in\mathbb{N}}\nu_n(\|a_n\|+\|b_n\|)\ <\ +\infty.\end{aligned}\tag{5.7}$$

On the other hand, $\sum_{n\in\mathbb{N}}\mu_n(1-\mu_n)=\sum_{n\in\mathbb{N}}\nu_n(2-\nu_n)/4=+\infty$. Moreover, Lemma 2.6(iii) and the assumption $0\in\operatorname{ran}(A+B)$ imply $\operatorname{Fix}T\neq\varnothing$. It therefore follows from Lemma 2.6(i) and Lemma 5.1 that $(x_n)_{n\in\mathbb{N}}$ converges weakly to some point $x\in\operatorname{Fix}T$. In view of Lemma 2.6(iii), the proof is complete. ☐

The above Corollary improves, on the one hand, upon [20, Proposition 12], where the additional assumptions $a_n \equiv 0$ and $b_n \equiv 0$ are made and, on the other hand, upon [26, Theorem 7], where the additional assumptions $0 < \underline{\lim}\, \nu_n \le \overline{\lim}\, \nu_n < 2$, $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, and $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$ are made. The classical Lions and Mercier result [38, Theorem 1] is recovered when $\nu_n \equiv 1$, $a_n \equiv 0$, and $b_n \equiv 0$.

Let us now consider the Peaceman-Rachford algorithm. In view of (1.13), this algorithm can be rewritten as

$$x_{n+1} = Rx_n, \quad \text{where} \quad R = R_{\gamma A} R_{\gamma B}. \tag{5.8}$$

Let us note that since R is merely nonexpansive, this iteration does not converge even weakly in general. We now prove that strong convergence is achieved for a perturbed extension of this algorithm under a Slater condition.

Corollary 5.3 *Let $\gamma \in\ ]0, +\infty[$ and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$. Suppose that $\operatorname{int}(A + B)^{-1}(0) \neq \varnothing$ and that $\sum_{n\in\mathbb{N}}(\|a_n\| + \|b_n\|) < +\infty$. Take $x_0 \in \mathcal{H}$ and set*

$$(\forall n \in \mathbb{N})\ \ x_{n+1} = 2\Big(J_{\gamma A}\big(2(J_{\gamma B}x_n + b_n) - x_n\big) + a_n\Big) - 2(J_{\gamma B}x_n + b_n) + x_n. \tag{5.9}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point $x \in \mathcal{H}$ such that $J_{\gamma B}x \in (A+B)^{-1}(0)$ and $(J_{\gamma B}x_n)_{n\in\mathbb{N}}$ converges strongly to $J_{\gamma B}x$.

Proof. Set $T = R_{\gamma A}R_{\gamma B}$ and define $(c_n)_{n\in\mathbb{N}}$ as in (5.6). Then it follows from Lemma 2.6(ii) that (5.9) can be rewritten as $(\forall n \in \mathbb{N})\ \ x_{n+1} = Tx_n + c_n$. Now fix $y \in \operatorname{Fix} T$, which is nonempty by Lemma 2.6(iii). Then $(\forall n \in \mathbb{N})\ \ \|x_{n+1} - y\| \le \|Tx_n - y\| + \|c_n\| \le \|x_n - y\| + \|c_n\|$. Hence, since by nonexpansivity of $R_{\gamma A}$ (5.6) yields $\sum_{n\in\mathbb{N}} \|c_n\| \le 2\sum_{n\in\mathbb{N}}(\|a_n\| + \|b_n\|) < +\infty$, $(x_n)_{n\in\mathbb{N}}$ is a quasi-Fejér sequence with respect to $\operatorname{Fix} T$. Since $\operatorname{int} \operatorname{Fix} T \neq \varnothing$, it follows from Lemma 2.8(iv) that $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point $x \in \mathcal{H}$. Hence, by continuity of T, $Tx_n \to Tx$ and, since $c_n \to 0$, we obtain $x \leftarrow x_{n+1} = Tx_n + c_n \to Tx$. In turn, this yields $x \in \operatorname{Fix} T$ and, via Lemma 2.6(iii), $J_{\gamma B}x \in (A + B)^{-1}(0)$. The continuity of $J_{\gamma B}$ allows us to conclude that $J_{\gamma B}x_n \to J_{\gamma B}x$. ☐

We conclude this section by observing that the Peaceman-Rachford recursion (5.9) is the limiting case of the Douglas-Rachford recursion (5.5) as $\nu_n \to 2$.

6 Forward-backward splitting

In this section we revisit the inclusion (1.6) under the following assumption.

Assumption 6.1 $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B \colon \mathcal{H} \to \mathcal{H}$ are maximal monotone and $\beta B \in \mathcal{A}(\frac{1}{2})$ for some $\beta \in\ ]0, +\infty[$.

This set of assumptions is clearly more demanding on the operator B than those in section 5. However, it leads to an algorithmic framework in which only one implicit (backward) step is required at each iteration, as opposed to two in the Douglas-Rachford and Peaceman-Rachford methods.

6.1 Preliminaries

For convenience, we specialize Algorithm 1.2 and Theorem 3.2 to the case when $m = 2$ (Theorem 3.3 can be rephrased in a like manner).

Algorithm 6.2 Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set

$$x_{n+1} = x_n + \lambda_n\big(T_{1,n}\big(T_{2,n}x_n + e_{2,n}\big) + e_{1,n} - x_n\big), \tag{6.1}$$

where $T_{1,n} \in \mathcal{A}(\alpha_{1,n})$ and $T_{2,n} \in \mathcal{A}(\alpha_{2,n})$, with $(\alpha_{1,n}, \alpha_{2,n}) \in \,]0,1[^2$, $(e_{1,n}, e_{2,n}) \in \mathcal{H}^2$, and $\lambda_n \in \,]0,1]$.

We now state Theorem 3.2 is the setting described in Remark 3.4.

Theorem 6.3 *Suppose that the following conditions are satisfied.*

(i) $G = \bigcap_{n\in\mathbb{N}} \operatorname{Fix}\big(T_{1,n}T_{2,n}\big) \neq \varnothing$.

(ii) $\underline{\lim}\, \lambda_n > 0$, $\overline{\lim}\, \alpha_{1,n} < 1$, *and* $\overline{\lim}\, \alpha_{2,n} < 1$.

(iii) *For every subsequence* $(x_{k_n})_{n\in\mathbb{N}}$ *of an orbit* $(x_n)_{n\in\mathbb{N}}$ *generated by Algorithm 6.2, we have*

$$\begin{cases} (\forall x \in G) \ \displaystyle\sum_{n\in\mathbb{N}} \|(\operatorname{Id} - T_{1,n})T_{2,n}x_n + (\operatorname{Id} - T_{2,n})x\|^2 < +\infty \\ (\forall x \in G) \ \displaystyle\sum_{n\in\mathbb{N}} \|(\operatorname{Id} - T_{2,n})x_n - (\operatorname{Id} - T_{2,n})x\|^2 < +\infty \\ \displaystyle\sum_{n\in\mathbb{N}} \|T_{1,n}T_{2,n}x_n - x_n\|^2 < +\infty \\ x_{k_n} \rightharpoonup y \end{cases} \Rightarrow \ y \in G. \tag{6.2}$$

(iv) $\sum_{n\in\mathbb{N}} \|e_{1,n}\| < +\infty$ *and* $\sum_{n\in\mathbb{N}} \|e_{2,n}\| < +\infty$.

Then every orbit of Algorithm 6.2 converges weakly to a point in G.

6.2 Main result

We investigate the following nonstationary form of the forward-backward method (1.15) with relaxations and errors.

Algorithm 6.4 Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set

$$x_{n+1} = x_n + \lambda_n\Big(J_{\gamma_n A}\big(x_n - \gamma_n(Bx_n + b_n)\big) + a_n - x_n\Big), \tag{6.3}$$

where $\gamma_n \in \,]0, 2\beta[$, $(a_n, b_n) \in \mathcal{H}^2$, and $\lambda_n \in \,]0,1]$.

Corollary 6.5 *Suppose that Assumption 6.1 is in force and that the following conditions are satisfied.*

(i) $0 \in \operatorname{ran}(A+B)$.

(ii) $\underline{\lim}\, \lambda_n > 0$ *and* $0 < \underline{\lim}\, \gamma_n \leq \overline{\lim}\, \gamma_n < 2\beta$.

(iii) $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$ *and* $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$.

Then every orbit of Algorithm 6.4 converges weakly to a zero of $A+B$.

Proof. We shall show that this result is a special case of Theorem 6.3. Indeed set

$$(\forall n \in \mathbb{N}) \ \ T_{1,n} = J_{\gamma_n A} \quad \text{and} \quad T_{2,n} = \operatorname{Id} - \gamma_n B. \tag{6.4}$$

Then $(T_{1,n})_{n\in\mathbb{N}}$ lies in $\mathcal{A}(\frac{1}{2})$ by Assumption 6.1 and Lemma 2.4. On the other hand, since $\beta B \in \mathcal{A}(\frac{1}{2})$ by Assumption 6.1, it follows from Lemma 2.3 that $(\forall n \in \mathbb{N})\ T_{2,n} \in \mathcal{A}(\frac{\gamma_n}{2\beta})$. Altogether, Algorithm 6.4 is a special case of Algorithm 6.2 with $\alpha_{1,n} = 1/2$, $\alpha_{2,n} = \gamma_n/(2\beta)$, $e_{1,n} = a_n$, and $e_{2,n} = -\gamma_n b_n$. Furthermore, since B is single-valued,

$$(\forall n \in \mathbb{N})(\forall x \in \mathcal{H}) \ \ x \in (A+B)^{-1}(0) \quad \Leftrightarrow \quad x - \gamma_n Bx \in x + \gamma_n Ax \quad \Leftrightarrow \quad x \in \operatorname{Fix} T_{1,n}T_{2,n}. \tag{6.5}$$

Hence, $G = (A+B)^{-1}(0)$ and items (i), (ii), and (iv) in Theorem 6.3 are implied by (i)–(iii) above. It remains to check item (iii) in Theorem 6.3. To this end, let us fix a suborbit $(x_{k_n})_{n\in\mathbb{N}}$ of Algorithm 6.4, $x \in (A+B)^{-1}(0)$, and set

$$(\forall n \in \mathbb{N}) \ \ y_n = J_{\gamma_n A}(x_n - \gamma_n Bx_n) \ \text{ and } \ u_n = \frac{x_n - y_n}{\gamma_n} - Bx_n. \tag{6.6}$$

Then, in view of (6.4) and item (ii) above, (6.2) holds if

$$\begin{cases} u_n \to -Bx \\ Bx_n \to Bx \\ y_n - x_n \to 0 \\ x_{k_n} \rightharpoonup y \end{cases} \quad \Rightarrow \quad 0 \in Ay + By. \tag{6.7}$$

To show this implication, note that the above bracketed conditions imply that $y_{k_n} \rightharpoonup y$. In addition, B is continuous and monotone on $\mathcal{H}$, hence maximal monotone [3, Proposition 3.5.7]. Therefore, by Lemma 2.5(iii), the conditions $x_{k_n} \rightharpoonup y$ and $Bx_{k_n} \to Bx$ force $Bx = By$. Thus, we get $y_{k_n} \rightharpoonup y$, $u_{k_n} \to -By$, and, since by (6.6) $\big((y_{k_n}, u_{k_n})\big)_{n\in\mathbb{N}}$ lies in $\operatorname{gr} A$, Lemma 2.5(iii) yields $-By \in Ay$, i.e., $0 \in Ay + By$. ☐

Remark 6.6 (Strong convergence) We have shown that $x_n \rightharpoonup y$ for some $y \in (A+B)^{-1}(0)$. Strong convergence conditions can be derived easily from Theorem 3.3. For instance, we obtain at once $x_n \to y$ if $\operatorname{int}(A+B)^{-1}(0) \neq \varnothing$. To get other conditions, it suffices to check (3.18) or, arguing as above, simply that

$$\begin{cases} u_n \to -By \\ Bx_n \to By \\ y_n - x_n \to 0 \end{cases} \quad \Rightarrow \quad \underline{\lim}\, d_{(A+B)^{-1}(0)}(x_n) = 0. \tag{6.8}$$

Thus, we obtain strong convergence when B is uniformly monotone on bounded sets, i.e., for every bounded set $C \subset \mathcal{H}$ there exists a strictly increasing function $c\colon\ [0,+\infty[\ \to [0,+\infty[$ with $c(0)=0$ such that [60, section 25.3]

$$(\forall (x,z) \in C^2) \quad \langle x-z \mid Bx - Bz\rangle \geq \|x-z\| \cdot c(\|x-z\|). \tag{6.9}$$

Indeed, (6.9) and Cauchy-Schwarz yield $(\forall n \in \mathbb{N})\ \|Bx_n - By\| \geq c(\|x_n - y\|)$. Hence $Bx_n \to By \Rightarrow x_n \to y$. We also get strong convergence when $\operatorname{dom} A$ is boundedly relatively compact. To see this, we use the condition $y_n - x_n \to 0$ and the same argument as in Remark 4.4 since (6.6) implies that $(y_n)_{n\in\mathbb{N}}$ lies in $\overline{\operatorname{dom} A}$.

Corollary 6.5 captures and extends several known results that were obtained using different approaches. The following example is an unrelaxed method that involves a specific model for the errors associated with the operators A and B in (6.3).

Corollary 6.7 [35, Proposition 3.2] *Suppose that Assumption 6.1 is in force. Take $x_0 \in \mathcal{H}$, $\varepsilon \in\ ]0,\beta]$, $(c_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$, $(\gamma_n)_{n\in\mathbb{N}}$ in $[\varepsilon, 2\beta-\varepsilon]$, and set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma_n A_n}\big(x_n - \gamma_n(B+B_n)x_n\big) + c_n, \tag{6.10}$$

where $(A_n)_{n\in\mathbb{N}}$ is a sequence of maximal monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$ and $(B_n)_{n\in\mathbb{N}}$ is a sequence of operators from $\mathcal{H}$ to $\mathcal{H}$. Suppose further that

(i) $0 \in \operatorname{ran}(A+B)$.

(ii) $(\forall \rho \in [0,+\infty[)\ \ \sum_{n\in\mathbb{N}} \sup_{\|y\|\leq\rho} \|J_{\gamma_n A} y - J_{\gamma_n A_n} y\| < +\infty$.

(iii) $(\exists\, z \in \mathcal{H})(\forall n \in \mathbb{N})\ \ B_n z = 0$.

(iv) *For every $n \in \mathbb{N}$, $B_n\colon \mathcal{H} \to \mathcal{H}$ is Lipschitz-continuous with constant $\kappa_n \in\ ]0,+\infty[$.*

(v) $\sum_{n\in\mathbb{N}} \kappa_n < +\infty$.

(vi) $\sum_{n\in\mathbb{N}} \|c_n\| < +\infty$.

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

Proof. The recursion (6.10) is a special case of (6.3), where

$$(\forall n\in\mathbb{N})\quad \begin{cases} a_n = J_{\gamma_n A_n}\big(x_n-\gamma_n(B+B_n)x_n\big) - J_{\gamma_n A}\big(x_n-\gamma_n(B+B_n)x_n\big)+c_n,\\ b_n = B_n x_n,\\ \lambda_n = 1. \end{cases} \tag{6.11}$$

Therefore, in view of Corollary 6.5, it remains to show that $\sum_{n\in\mathbb{N}}\|a_n\|<+\infty$ and $\sum_{n\in\mathbb{N}}\|b_n\|<+\infty$. To this end, let us fix $x\in(A+B)^{-1}(0)$. We first observe that (iii) and (iv) yield

$$(\forall n\in\mathbb{N})\quad \|b_n\|\leq\|B_nx_n-B_nx\|+\|B_nx-B_nz\|\leq\kappa_n\big(\|x_n-x\|+\|x-z\|\big). \tag{6.12}$$

On the other hand, since, for every $n\in\mathbb{N}$, the operators $J_{\gamma_nA_n}$ and $T_{2,n}=\operatorname{Id}-\gamma_nB$ are nonexpansive and $x\in\operatorname{Fix}J_{\gamma_nA}T_{2,n}$, we derive from (6.10) and (6.12) that

$$\begin{aligned}
\|x_{n+1}-x\| &\leq \|J_{\gamma_nA_n}\big(T_{2,n}x_n-\gamma_nb_n\big)-x\|+\|c_n\|\\
&\leq \|J_{\gamma_nA_n}\big(T_{2,n}x_n-\gamma_nb_n\big)-J_{\gamma_nA_n}(T_{2,n}x)\|+\|J_{\gamma_nA_n}(T_{2,n}x)-J_{\gamma_nA}(T_{2,n}x)\|+\|c_n\|\\
&\leq \|T_{2,n}x_n-\gamma_nb_n-T_{2,n}x\|+\|J_{\gamma_nA_n}(T_{2,n}x)-J_{\gamma_nA}(T_{2,n}x)\|+\|c_n\|\\
&\leq \|x_n-x\|+2\beta\|b_n\|+\|J_{\gamma_nA_n}(T_{2,n}x)-J_{\gamma_nA}(T_{2,n}x)\|+\|c_n\|\\
&\leq (1+2\beta\kappa_n)\|x_n-x\|+\varepsilon_n,
\end{aligned} \tag{6.13}$$

where

$$\varepsilon_n=2\beta\kappa_n\|x-z\|+\|J_{\gamma_nA_n}(T_{2,n}x)-J_{\gamma_nA}(T_{2,n}x)\|+\|c_n\|. \tag{6.14}$$

Now let $\rho=\|x\|+2\beta\|Bx\|$. Then

$$\sup_{n\in\mathbb{N}}\|T_{2,n}x\|\leq\rho \tag{6.15}$$

and it follows from (ii), (v), and (vi) that $\sum_{n\in\mathbb{N}}\varepsilon_n<+\infty$. We therefore derive from (6.13), (v), and Lemma 2.7 that $\zeta=\sup_{n\in\mathbb{N}}\|x_n-x\|<+\infty$ and, in turn, from (6.12) that $\sum_{n\in\mathbb{N}}\|b_n\|<+\infty$. Consequently, (6.15) yields

$$\begin{aligned}
\sup_{n\in\mathbb{N}}\|T_{2,n}x_n-\gamma_nb_n\| &\leq \sup_{n\in\mathbb{N}}\|T_{2,n}x_n-T_{2,n}x\|+\|T_{2,n}x\|+\gamma_n\|b_n\|\\
&\leq \zeta+\rho+2\beta\sup_{n\in\mathbb{N}}\|b_n\| \;<\; +\infty
\end{aligned} \tag{6.16}$$

and we conclude from (6.11), (ii), and (vi) that

$$\sum_{n\in\mathbb{N}}\|a_n\|\leq\sum_{n\in\mathbb{N}}\|J_{\gamma_nA_n}\big(T_{2,n}x_n-\gamma_nb_n\big)-J_{\gamma_nA}\big(T_{2,n}x_n-\gamma_nb_n\big)\|+\sum_{n\in\mathbb{N}}\|c_n\|<+\infty. \tag{6.17}$$

□

Let us note that in the special case when $A_n\equiv A$, $B_n\equiv 0$, and $c_n=0$ above (i.e., $a_n=b_n=0$ and $\lambda_n\equiv 1$ in Corollary 6.5), we recover [34, Proposition 3.1] and [56, Proposition 1(c)]. If we further assume that $\gamma_n\equiv\gamma$, we recover [43, Remarque 3.1], which seems to be the first weak convergence result of this type for the forward-backward method. The perturbation model (ii) above goes back to [54].

Now, take $\varphi \in \Gamma_0(\mathcal{H})$ and set $A = \partial\varphi$. Then $J_A = \operatorname{prox}_\varphi$ and (1.6) reduces to the variational inequality problem [37]

$$\text{Find } x \in \mathcal{H} \text{ such that } (\forall y \in \mathcal{H}) \ \ \langle x - y \mid Bx\rangle + \varphi(x) \leq \varphi(y). \tag{6.18}$$

Moreover, Corollary 6.5 gives conditions for the weak convergence of the iteration

$$x_{n+1} = x_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n\varphi}\big(x_n - \gamma_n(Bx_n + b_n)\big) + a_n - x_n\Big) \tag{6.19}$$

to a solution to this problem. Now set $\varphi = \iota_C$, where C is a nonempty closed convex subset of $\mathcal{H}$. Then, (6.18) turns into the classical variational inequality problem

$$\text{Find } x \in C \text{ such that } (\forall y \in C) \ \ \langle x - y \mid Bx\rangle \leq 0. \tag{6.20}$$

Furthermore, for $\lambda_n \equiv 1$ and $a_n \equiv 1$, (6.19) becomes $x_{n+1} = P_C\big(x_n - \gamma_n(Bx_n + b_n)\big)$. The strong convergence of this method was established in [6] under conditions akin to some of those discussed in Remark 6.6. If we further assume that $\gamma_n \equiv \gamma$ and $b_n \equiv 0$, Corollary 6.5 furnishes the weak convergence of the iteration $x_{n+1} = P_C\big(x_n - \gamma Bx_n\big)$ to a solution to (6.20). This result was obtained in [42, Theorem 10]. Another special case of interest, is the following result that pertains to the projected gradient method.

Corollary 6.8 *Suppose that C is a closed convex subset of $\mathcal{H}$, that $f\colon \mathcal{H} \to \mathbb{R}$ is convex and differentiable with a $1/\beta$-Lipschitz-continuous gradient, and that the following conditions are satisfied.*

(i) *f achieves its infimum on C.*

(ii) $\underline{\lim}\,\lambda_n > 0$ *and* $0 < \underline{\lim}\,\gamma_n \leq \overline{\lim}\,\gamma_n < 2\beta$.

(iii) $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$ *and* $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$.

Take $x_0 \in \mathcal{H}$ and set

$$(\forall n \in \mathbb{N}) \ \ x_{n+1} = x_n + \lambda_n\Big(P_C\big(x_n - \gamma_n(\nabla f(x_n) + b_n)\big) + a_n - x_n\Big). \tag{6.21}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a minimizer of f on C.

Proof. If follows from the Baillon-Haddad theorem [5, Corollaire 10] that $\beta\nabla f \in \mathcal{A}(\frac{1}{2})$. Hence the result is a direct application of Corollary 6.5, where $A = N_C$ and $B = \nabla f$. □

6.3 Partial Yosida approximation of monotone inclusions

In this section, $I = \{0, \ldots, m\}$ is a finite index set and $(A_i)_{i\in I}$ is a family of maximal monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$. We apply the framework of section 6.2 to extend certain results on the numerical solution of infeasible convex feasibility problems which arise in particular in signal theory (see [17, 22] and the references therein).

In section 4 we have examined the common zero problem (1.3) under the premise that it was feasible, i.e., that its set of solutions

$$S = \bigcap_{i=0}^{m} A_i^{-1}(0) \tag{6.22}$$

was nonempty. In practical situations, however, (1.3) may turn out to be inconsistent. In such instances, it is natural to approximate it by a more general problem, which exhibits more regularity properties and is solvable. In this connection, we shall investigate the following extension of (1.3), which assumes the form of the sum problem (1.2).

Definition 6.9 *Fix parameters* $(\rho_i)_{1\leq i\leq m}$ *in* $]0,+\infty[$. *The partial Yosida approximation to problem* (1.3) *is*

$$\text{Find } x \in \mathcal{H} \quad \text{such that} \quad 0 \in A_0 x + \sum_{i=1}^{m} {}^{\rho_i}\!A_i x \tag{6.23}$$

and its set of solutions is denoted by G, *i.e.,*

$$G = \left(A_0 + \sum_{i=1}^{m} {}^{\rho_i}\!A_i\right)^{-1}(0). \tag{6.24}$$

In this sum reformulation of the common zero problem (1.3), the operators $(A_i)_{1\leq i\leq m}$ are replaced by their Yosida approximation (2.2), while A_0 is not regularized. In the case when $m = 1$, this type of regularization is quite standard, e.g., [39, 43, 46]. Note, however, that the objectives and methodologies of these papers are different from ours since there (1.3) is assumed to have solutions and the problem is to approach a particular solution by regularization as $\rho_i \to 0$.

Problem (6.23) is a special case of (1.6) in which

$$A = A_0 \quad \text{and} \quad B = \sum_{i=1}^{m} {}^{\rho_i}\!A_i = \frac{1}{\beta}\left(\mathrm{Id} - \sum_{i=1}^{m} \omega_i J_{\rho_i A_i}\right), \tag{6.25}$$

where

$$\frac{1}{\beta} = \sum_{i=1}^{m} \frac{1}{\rho_i} \quad \text{and} \quad (\forall i \in \{1,\ldots,m\}) \quad \omega_i = \frac{\beta}{\rho_i}. \tag{6.26}$$

On the other hand, (6.23) is an extension of (1.3) in the following sense.

Proposition 6.10 *Suppose that* $S \neq \emptyset$. *Then* $G = S$.

Proof. Lemma 2.5(i) asserts that the operators $\big(\rho_i({}^{\rho_i}\!A_i)\big)_{1\leq i\leq m}$ lie in $\mathcal{A}(\frac{1}{2})$. It therefore follows from (6.25), (6.26), and Lemma 2.2(ii) that $\beta B = \sum_{i=1}^{m} \omega_i \rho_i({}^{\rho_i}\!A_i) \in \mathcal{A}(\frac{1}{2})$. Now set $T_1 = J_{\beta A}$ and

$T_2 = \mathrm{Id} - \beta B$. Then Lemma 2.3 yields $T_2 \in \mathcal{A}(\frac{1}{2})$ and we derive from (6.22), (2.13), Lemma 2.4, Lemma 2.2(iv), and (6.25) that

$$\varnothing \neq S \subset \bigcap_{i=1}^{m} A_i^{-1}(0) = \bigcap_{i=1}^{m} \operatorname{Fix} J_{\rho_i A_i} = \operatorname{Fix} \sum_{i=1}^{m} \omega_i J_{\rho_i A_i} = \operatorname{Fix} T_2. \tag{6.27}$$

Thus, using (6.22), (2.13), Lemma 2.2(iv), (6.5), (6.25), and (6.24), we obtain

$$\varnothing \neq S = A_0^{-1}(0) \cap \bigcap_{i=1}^{m} A_i^{-1}(0) = \operatorname{Fix} T_1 \cap \operatorname{Fix} T_2 = \operatorname{Fix} T_1 T_2 = (A+B)^{-1}(0) = G. \tag{6.28}$$

□

In view of (6.25), allowing for an error $b_{i,n}$ in the evaluation of $J_{\rho_i A_i} x_n$ leads to the following implementation of Algorithm 6.4.

Algorithm 6.11 Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set

$$x_{n+1} = x_n + \lambda_n \left(J_{\beta \mu_n A_0} \left(x_n + \mu_n \Bigg(\sum_{i=1}^{m} \omega_i \big(J_{\rho_i A_i} x_n + b_{i,n} \big) - x_n \Bigg) \right) + a_n - x_n \right) \tag{6.29}$$

where $\mu_n \in \left]0, 2\right[$, $(a_n, b_{1,n}, \ldots, b_{m,n}) \in \mathcal{H}^{m+1}$, and $\lambda_n \in \left]0, 1\right]$.

Corollary 6.12 *Suppose that the following conditions are satisfied.*

(i) $G \neq \varnothing$.

(ii) $\varliminf \lambda_n > 0$ *and* $0 < \varliminf \mu_n \leq \varlimsup \mu_n < 2$.

(iii) $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ *and* $\max_{1 \leq i \leq m} \sum_{n \in \mathbb{N}} \|b_{i,n}\| < +\infty$.

Then every orbit of Algorithm 6.11 converges weakly to a point in G.

Proof. The claim is a consequence of Corollary 6.5 with A and B defined in (6.25)–(6.26) and $(\forall n \in \mathbb{N})$ $b_n = -\sum_{i=1}^{m} \omega_i b_{i,n}/\beta$ and $\mu_n = \gamma_n/\beta$. □

Remark 6.13 (Backward-backward splitting) Suppose that $m = 1$ and set $\lambda_n \equiv 1$, $\mu_n \equiv 1$, $a_n \equiv 0$, and $b_{1,n} \equiv 0$. Then (6.29) reduces to the backward-backward method (1.14), more specifically to $x_{n+1} = J_{\rho_1 A_0} J_{\rho_1 A_1} x_n$. Corollary 6.12 states that this iteration converges weakly to a zero of $A_0 + {}^{\rho_1}A_1$ if such a point exists. In particular, if φ and ψ are two functions in $\Gamma_0(\mathcal{H})$ and we set $\rho_1 = 1$, $A_0 = \partial\varphi$, and $A_1 = \partial\psi$, the backward-backward iterative process becomes $x_{n+1} = \operatorname{prox}_\varphi \operatorname{prox}_\psi x_n$. This method was studied in [1] in connection with the problem of minimizing $\varphi + {}^{1}\psi$.

As an illustration of the above result, let us consider the problem of solving the convex inequality system

$$\text{Find } x \in C_0 \text{ such that } \max_{1\le i\le m} f_i(x) \le 0, \tag{6.30}$$

where $(f_i)_{1\le i\le m}$ is a family of functions in $\Gamma_0(\mathcal{H})$ and C_0 is a closed convex set in $\mathcal{H}$ playing the role of a hard constraint. This problem fits the general format (1.3), where $A_0 = N_{C_0}$ and, for every $i \in \{1,\ldots,m\}$, $A_i = \partial\varphi_i$ with $\varphi_i = \max\{0, f_i\}^2$. When it has no solution, Problem (6.30) can therefore be replaced by (6.23) and solved by (6.29), which becomes

$$x_{n+1} = x_n + \lambda_n\left(P_0\left(x_n + \mu_n\Big(\sum_{i=1}^m \omega_i\big(\operatorname{prox}_{\rho_i\varphi_i} x_n + b_{i,n}\big) - x_n\Big)\right) + a_n - x_n\right), \tag{6.31}$$

where P_0 is the projector onto C_0. In this case, it follows from [45, Proposition 7.d] and elementary convex calculus that (6.23) can be formulated as the problem of minimizing $\varphi = \sum_{i=1}^m {}^{\rho_i}\varphi_i$ over C_0. In particular, let $(f_i)_{1\le i\le m}$ be the indicator functions of nonempty closed convex sets $(C_i)_{1\le i\le m}$ with projectors $(P_i)_{1\le i\le m}$. Then (6.30) reduces to the basic convex feasibility problem

$$\text{Find } x \in \bigcap_{i=0}^m C_i \tag{6.32}$$

and (6.23) amounts to approximating it by the problem of minimizing $\varphi = \frac{1}{2}\sum_{i=1}^m d_{C_i}^2/\rho_i$ over C_0. The recursion (6.31) then assumes the form

$$x_{n+1} = x_n + \lambda_n\left(P_0\left(x_n + \mu_n\Big(\sum_{i=1}^m \omega_i\big(P_i x_n + b_{i,n}\big) - x_n\Big)\right) + a_n - x_n\right). \tag{6.33}$$

In this setting Corollary 6.12 extends various convergence results for projection methods. For example, the case $\mu_n \equiv \mu$, $a_n \equiv 0$, and $b_{i,n} \equiv 0$ was considered in [22] (in particular in [17] with $C_0 = \mathcal{H}$ and in [8, 23] with the additional hypothesis $\lambda_n \equiv 1$).

7 Stationary iteration

The following corollary of Theorem 3.2 involves an iteration process which is stationary in the sense that the operators involved do not vary with n.

Corollary 7.1 *For every $i \in \{1,\ldots,m\}$, let $T_i \in \mathcal{A}(\alpha_i)$, where $\alpha_i \in \,]0,1[$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n\Big(T_1\Big(T_2\big(\cdots T_{m-1}(T_m x_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\Big) + e_{1,n} - x_n\Big), \tag{7.1}$$

where $(e_{i,n})_{1\le i\le m} \in \mathcal{H}^m$ and $\lambda_n \in \,]0,1]$. Suppose that the following conditions are satisfied.

(i) $\operatorname{Fix} T_1\cdots T_m \neq \varnothing$.

(ii) $\underline{\lim}\,\lambda_n > 0$.

(iii) $(\forall i \in \{1,\ldots,m\})\ \ \sum_{n\in\mathbb{N}} \|e_{i,n}\| < +\infty$.

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point* y *in* $\operatorname{Fix} T_1\cdots T_m$*. Moreover,*

$$\big(T_1\cdots T_m x_n, T_2\cdots T_m x_n, \ldots, T_m x_n\big) \rightharpoonup \big(T_1\cdots T_m y, T_2\cdots T_m y, \ldots, T_m y\big). \tag{7.2}$$

Proof. Let $T = T_1\cdots T_m$ and let $(x_{k_n})_{n\in\mathbb{N}}$ be a subsequence such that $x_{k_n} \rightharpoonup y$ for some $y \in \mathcal{H}$. In view of (i)–(iii), Theorem 3.2, and Remark 3.4, it is enough to show that

$$(\forall x \in \operatorname{Fix} T)\ \max_{1\le j\le m} \sum_{n\in\mathbb{N}} \|(\operatorname{Id} - T_j)T_{j+1}\cdots T_m x_n - (\operatorname{Id} - T_j)T_{j+1}\cdots T_m x\|^2 < +\infty \tag{7.3}$$

implies that $y \in \operatorname{Fix} T$ to establish the first claim. First, we derive from (3.22) that

$$Tx_n - x_n \to 0. \tag{7.4}$$

Hence, since T is nonexpansive, it follows from the demiclosed principle [14, Lemma 4] that $y \in \operatorname{Fix} T$. Therefore, we get $x_n \rightharpoonup y \in \operatorname{Fix} T$. Let us now prove the second claim by induction. For $i = 1$, (7.4) yields $T_i\cdots T_m x_n = (Tx_n - x_n) + x_n \rightharpoonup y = T_i\cdots T_m y$. Now suppose that, for some $i \in \{1,\ldots,m-1\}$, $T_i\cdots T_m x_n \rightharpoonup T_i\cdots T_m y$. Then, since (7.3) yields $T_{i+1}\cdots T_m x_n - T_i\cdots T_m x_n \to T_{i+1}\cdots T_m y - T_i\cdots T_m y$, we conclude that $T_{i+1}\cdots T_m x_n \rightharpoonup T_{i+1}\cdots T_m y$. ☐

In particular, Corollary 7.1 asserts that if $(T_i)_{1\le i\le m}$ are averaged operators whose composition has a fixed point, the iterates $x_{n+1} = T_1\cdots T_m x_n$ converge weakly to such a point. This result can also be deduced from [15] (combine Proposition 1.3, Proposition 1.1, and Corollary 1.3 in that paper) and, in the special case of firmly nonexpansive operators, it appears in [40, Théorème 5.5.2]. If we take each T_i to be the resolvent of a maximal monotone operator $A_i\colon \mathcal{H} \to 2^{\mathcal{H}}$, then Corollary 7.1 provides information on the asymptotic behavior of a relaxed, inexact version of the m-step backward-backward method (1.16) (see also Remark 6.13) when the inclusion (1.3) is infeasible.

For an alternative interpretation, let us call a cycle an m-tuple $(y_i)_{1\le i\le m} \in \mathcal{H}^m$ such that

$$y_m = T_m y_1 \ \text{ and } \ (\forall i \in \{1,\ldots,m-1\})\ \ y_i = T_i y_{i+1}, \tag{7.5}$$

where the notation and assumptions are as in Corollary 7.1. Then Corollary 7.1 states that $\big((x_n, T_2\cdots T_m x_n, T_3\cdots T_m x_n, \ldots, T_m x_n)\big)_{n\in\mathbb{N}}$ converges weakly to a cycle in $\mathcal{H}^m$. In particular, if each T_i is the projector P_i onto a nonempty closed convex set $S_i \subset \mathcal{H}$, $\operatorname{Fix} P_1\cdots P_m \neq \varnothing$ (e.g., one of the sets is bounded), $\lambda_n \equiv 1$, and $e_{i,n} \equiv 0$, we obtain the weak convergence of $\big((x_n, P_2\cdots P_m x_n, P_3\cdots P_m x_n, \ldots, P_m x_n)\big)_{n\in\mathbb{N}}$ to a cycle $(y_i)_{1\le i\le m} \in \bigtimes_{i=1}^m S_i$. This classical result was obtained in [27, Theorem 2] (see also [10] for more information on cyclic projection methods for inconsistent feasibility problems and [7] for the case when $\operatorname{Fix} P_1\cdots P_m = \varnothing$).

References

[1] F. Acker and M. A. Prestel (1980). Convergence d'un schéma de minimisation alternée. *Annales de la Faculté des Sciences de Toulouse – Série 5*, **2**, 1–9.

[2] H. Attouch and M. Théra (1996). A general duality principle for the sum of two operators. *Journal of Convex Analysis,* **3**, 1–24.

[3] J. P. Aubin and H. Frankowska (1990). *Set-Valued Analysis.* Birkhäuser, Boston, MA.

[4] J. B. Baillon, R. E. Bruck, and S. Reich (1978). On the asymptotic behavior of nonexpansive mappings and semigroups. *Houston Journal of Mathematics*, **4**, 1–9.

[5] J. B. Baillon and G. Haddad (1977). Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. *Israel Journal of Mathematics,* **26**, 137–150.

[6] A. B. Bakušinskiĭ and B. T. Polyak (1974). The solution of variational inequalities. *Soviet Mathematics – Doklady,* **15**, 1705–1710.

[7] H. H. Bauschke (2003). The composition of finitely many projections onto closed convex sets in Hilbert space is asymptotically regular. *Proceedings of the American Mathematical Society,* **131**, 141–146.

[8] H. H. Bauschke and J. M. Borwein (1994). Dykstra's alternating projection algorithm for two sets. *Journal of Approximation Theory,* **79**, 418–443.

[9] H. H. Bauschke and J. M. Borwein (1996). On projection algorithms for solving convex feasibility problems. *SIAM Review*, **38**, 367–426.

[10] H. H. Bauschke, J. M. Borwein, and A. S. Lewis (1997). The method of cyclic projections for closed convex sets in Hilbert space. *Contemporary Mathematics,* **204**, 1–38.

[11] H. H. Bauschke and P. L. Combettes (2001). A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. *Mathematics of Operations Research,* **26**, 248–264.

[12] H. H. Bauschke, P. L. Combettes, and D. R. Luke (2004). Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *Journal of Approximation Theory,* **127**, 178–192.

[13] H. Brézis and P. L. Lions (1978). Produits infinis de résolvantes. *Israel Journal of Mathematics,* **29**, 329–345.

[14] F. E. Browder (1967). Convergence theorems for sequences of nonlinear operators in Banach spaces. *Mathematische Zeitschrift,* **100**, 201–225.

[15] R. E. Bruck and S. Reich (1977). Nonexpansive projections and resolvents of accretive operators in Banach spaces. *Houston Journal of Mathematics*, **3**, 459–470.

[16] G. Cimmino (1938). Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari. *La Ricerca Scientifica (Roma),* **1**, 326–333.

[17] P. L. Combettes (1994). Inconsistent signal feasibility problems: Least-squares solutions in a product space. *IEEE Transactions on Signal Processing,* **42**, 2955–2966.

[18] P. L. Combettes (1995). Construction d'un point fixe commun à une famille de contractions fermes. *Comptes Rendus de l'Académie des Sciences de Paris, Série I (Mathématique)*, **320**, 1385–1390.

[19] P. L. Combettes (1997). Hilbertian convex feasibility problem: Convergence of projection methods. *Applied Mathematics and Optimization,* **35**, 311–330.

[20] P. L. Combettes (2001). Fejér monotonicity in convex optimization. In: C. A. Floudas and P. M. Pardalos (Eds.), *Encyclopedia of Optimization*, **2**, 106–114. Kluwer, Boston, MA.

[21] P. L. Combettes (2001). Quasi-Fejérian analysis of some optimization algorithms. In: D. Butnariu, Y. Censor, and S. Reich (Eds.), *Inherently Parallel Algorithms for Feasibility and Optimization*, 115–152. Elsevier, New York.

[22] P. L. Combettes and P. Bondon (1999). Hard-constrained inconsistent signal feasibility problems. *IEEE Transactions on Signal Processing,* **47**, 2460–2468.

[23] A. R. De Pierro and A. N. Iusem (1985). A parallel projection method for finding a common point of a family of convex sets. *Pesquisa Operacional,* **5**, 1–20.

[24] J. Douglas and H. H. Rachford (1956). On the numerical solution of heat conduction problems in two or three space variables. *Transactions of the American Mathematical Society*, **82**, 421–439.

[25] J. M. Dye and S. Reich (1992). Unrestricted iterations of nonexpansive mappings in Hilbert space. *Nonlinear Analysis, Theory, Methods, and Applications,* **18**, 199–207.

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

[27] L. G. Gubin, B. T. Polyak, and E. V. Raik (1967). The method of projections for finding the common point of convex sets. *USSR Computational Mathematics and Mathematical Physics,* **7**, 1–24.

[28] S. Kaczmarz (1937). Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin de l'Académie des Sciences de Pologne,* **A35**, 355–357.

[29] A. Kaplan and R. Tichatschke (2001). A general view on proximal point methods to variational inequalities in Hilbert spaces – Iterative regularization and approximation. *Journal of Nonlinear and Convex Analysis,* **2**, 305–332.

[30] R. B. Kellogg (1969). A nonlinear alternating direction method. *Mathematics of Computation*, **23**, 23–27.

[31] K. C. Kiwiel and B. Łopuch (1997). Surrogate projection methods for finding fixed points of firmly nonexpansive mappings. *SIAM Journal on Optimization,* **7**, 1084–1102.

[32] N. Lehdili and B. Lemaire (1999). The barycentric proximal method. *Communications on Applied Nonlinear Analysis,* **6**, 29–47.

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

[34] B. Lemaire (1996). Stability of the iteration method for nonexpansive mappings. *Serdica Mathematical Journal*, **22**, 331–340.

[35] B. Lemaire (1997). Which fixed point does the iteration method select? *Lecture Notes in Economics and Mathematical Systems*, **452**, 154–167. Springer-Verlag, New York.

[36] J. Lieutaud (1969). *Approximation d'Opérateurs par des Méthodes de Décomposition*. Thèse, Université de Paris.

[37] J. L. Lions (1969). *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires.* Dunod, Paris.

[38] P. L. Lions and B. Mercier (1979). Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis,* **16**, 964-979.

[39] P. Mahey and D. T. Pham (1993). Partial regularization of the sum of two maximal monotone operators. *RAIRO Modélisation Mathématique et Analyse Numérique,* **27**, 375–392.

[40] B. Martinet (1972). *Algorithmes pour la Résolution de Problèmes d'Optimisation et de Minimax.* Thèse, Université de Grenoble, France.

[41] B. Martinet (1972). Détermination approchée d'un point fixe d'une application pseudo-contractante. Cas de l'application prox. *Comptes Rendus de l'Académie des Sciences de Paris,* **A274**, 163–165.

[42] B. Mercier (1979). *Lectures on Topics in Finite Element Solution of Elliptic Problems*. Lectures on Mathematics and Physics, **63**, Tata Institute of Fundamental Research, Bombay.

[43] B. Mercier (1980). *Inéquations Variationnelles de la Mécanique,* Publications Mathématiques d'Orsay, no. 80.01. Université Paris 11, Orsay, France.

[44] G. J. Minty (1962). Monotone (nonlinear) operators in Hilbert space. *Duke Mathematical Journal,* **29**, 341-346.

[45] J.-J. Moreau (1965). Proximité et dualité dans un espace hilbertien. *Bulletin de la Société Mathématique de France,* **93**, 273–299.

[46] A. Moudafi (2000). On the regularization of the sum of two maximal monotone operators. *Nonlinear Analysis,* **42**, 1203–1208.

[47] D. W. Peaceman and H. H. Rachford (1955). The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, **3**, 28–41.

[48] T. Pennanen (2002). A splitting method for composite mappings. *Numerical Functional Analysis and Optimization,* **23**, 875–890.

[49] B. T. Polyak (1987). *Introduction to Optimization.* Optimization Software Inc., New York.

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

[51] R. T. Rockafellar (1976). Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research,* **1**, 97–116.

[52] R. T. Rockafellar (1976). Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization,* **14**, 877–898.

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

[54] P. Tossings (1994). The perturbed proximal point algorithm and some of its applications. *Applied Mathematics and Optimization,* **29**, 125–159.

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

[56] P. Tseng (1991). Applications of a splitting algorithm to decomposition in convex programming and variational inequalities. *SIAM Journal on Control and Optimization,* **29**, 119–138.

[57] R. S. Varga (2000). *Matrix Iterative Analysis*, 2nd edition. Springer-Verlag, New York.

[58] N. N. Yanenko (1968). *Méthode à Pas Fractionnaires.* Armand Colin, Paris (see also *The Method of Fractional Steps* (1971). Springer-Verlag, New York).

[59] E. H. Zarantonello (1971). Projections on convex sets in Hilbert space and spectral theory. In: E. H. Zarantonello (Ed.), *Contributions to Nonlinear Functional Analysis,* 237–424. Academic Press, New York.

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