Extrapolation algorithm for affine-convex feasibility problems

Heinz H. Bauschke[*], Patrick L. Combettes[†], and Serge G. Kruk[‡]

October 5, 2005

Abstract

The convex feasibility problem under consideration is to find a common point of a countable family of closed affine subspaces and convex sets in a Hilbert space. To solve such problems, we propose a general parallel block-iterative algorithmic framework in which the affine subspaces are exploited to introduce extrapolated over-relaxations. This framework encompasses a wide range of projection, subgradient projection, proximal, and fixed point methods encountered in various branches of applied mathematics. The asymptotic behavior of the method is investigated and numerical experiments are provided to illustrate the benefits of the extrapolations.

1   Introduction

Let $(S_i)_{i \in I}$ be a countable family of intersecting closed convex sets in a real Hilbert space $\mathcal{H}$. The associated convex feasibility problem is to

$$\text{find } \ x \in \bigcap_{i \in I} S_i. \tag{1.1}$$

This problem has a long and rich history in applied mathematics, going back at least to the nineteenth century [13]. We refer the reader to [3, 14, 16, 18, 20, 29] for surveys and background, and to [12] for recent developments.

The early methods by Cimmino [17] and by Kaczmarz [41] on systems of linear equations relied on projections. For every $i \in I$, let P_i denote the projection operator onto S_i. A sequential projection method generates a sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathcal{H}$ according to the recursion

$$x_{n+1} = x_n + \lambda_n(P_{\mathsf{i}(n)}x_n - x_n), \quad \text{where} \quad \mathsf{i}(n) \in I \ \text{ and } \ 0 < \lambda_n < 2. \tag{1.2}$$

[*]Mathematics, Irving K. Barber School, UBC Okanagan, Kelowna, B.C. V1V 1V7, Canada. E-mail: heinz.bauschke@ubc.ca.

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

[‡]Department of Mathematics and Statistics, Oakland University, Rochester, Michigan 48309-4485, USA. E-mail: sgkruk@acm.org.

This scheme covers various methods by means of different index mappings i: $\mathbb{N} \to I$, which may select indices cyclically [9], via the remotest set strategy [32, 40], or according to some other sweeping rules [10]. In contrast, fully parallel methods activate all the sets simultaneously and the iteration assumes the form [1, 26, 27, 28, 51]

$$x_{n+1} = x_n + \lambda_n \left(\sum_{i \in I} \omega_i P_i x_n - x_n \right), \quad \text{where} \quad \begin{cases} 0 < \lambda_n < 2, \\ \{\omega_i\}_{i \in I} \subset \,]0, 1], \\ \sum_{i \in I} \omega_i = 1. \end{cases} \tag{1.3}$$

Since these schemes can be slow, Pierra introduced in [48, 49] an extrapolated variant of (1.3), namely (for convenience, we define $0/0 = 1$ for the rest of this section)

$$x_{n+1} = x_n + L_n \left(\sum_{i \in I} \omega_i P_i x_n - x_n \right), \quad \text{where} \quad L_n = \frac{\sum_{i \in I} \omega_i \|P_i x_n - x_n\|^2}{\| \sum_{i \in I} \omega_i P_i x_n - x_n \|^2}. \tag{1.4}$$

In numerical experiments, Pierra observed that the extrapolation parameter L_n can be much larger than 2 and that the sequence $(x_n)_{n \in \mathbb{N}}$ can converge much faster than the sequence generated by (1.3). In the case of affine half-spaces in $\mathbb{R}^N$, this scheme had previously been investigated by Merzlyakov [45]. Over the years, projection algorithms have been extended to more flexible block-iterative methods in which only a block of sets $(S_i)_{i \in I_n \subset I}$ is activated at iteration n [11, 22, 33, 34, 35, 38, 47]. Following a different track, some researchers observed that projections can be replaced by alternative operators that may be easier to compute, such as subgradient projectors, resolvents of monotone operators, resolvents of bifunctions, proximity operators, or firmly nonexpansive operators [3, 15, 19, 21, 25, 31, 42, 43, 44, 53, 55]. In [23], these approaches were unified through the iteration method

$$x_{n+1} = x_n + \lambda_n \left(\sum_{i \in I_n} \omega_{i,n} T_{i,n} x_n - x_n \right), \quad \text{where} \quad \begin{cases} I_n \subset I, \\ \{\omega_{i,n}\}_{i \in I_n} \subset [0, 1], \\ \sum_{i \in I_n} \omega_{i,n} = 1, \\ L_n = \dfrac{\sum_{i \in I_n} \omega_{i,n} \|T_{i,n} x_n - x_n\|^2}{\| \sum_{i \in I_n} \omega_{i,n} T_{i,n} x_n - x_n \|^2}, \\ 0 < \lambda_n < 2 L_n. \end{cases} \tag{1.5}$$

In this method, $T_{i,n}$ is a so-called $\mathfrak{T}$-class operator (see section 2) chosen so that its fixed point set coincides with S_i.

Let A be a closed vector subspace of $\mathcal{H}$ and let B be a nonempty closed convex subset of $\mathcal{H}$. Denote the projectors onto A and B by P_A and P_B, respectively. In [48, 49], Algorithm (1.4) was derived by investigating — in a suitable product space setting — the two-set iteration

$$x_{n+1} = x_n + K_n(P_A P_B x_n - x_n), \quad \text{where} \quad K_n = \frac{\|P_B x_n - x_n\|^2}{\|P_A P_B x_n - x_n\|^2}. \tag{1.6}$$

This two-set relaxation strategy was originally used in [40] in order to potentially accelerate the convergence of the basic alternating projection method $x_{n+1} = P_A P_B x_n$. It should be noted that

(1.6) is not covered by (1.5). Indeed, when $(I_n)_{n\in\mathbb{N}}$ reduces to a sequence of singletons $(\{i(n)\})_{n\in\mathbb{N}}$, then (1.5) reduces to (1.2), where the relaxations $(\lambda_n)_{n\in\mathbb{N}}$ cannot exceed 2. However, as will be illustrated in Figure 2 and Figure 6, the sequence of extrapolation parameters $(K_n)_{n\in\mathbb{N}}$ from (1.6) can include terms that are much larger than 2.

The objective of this paper is to introduce and analyze a general algorithmic framework that captures and extends both (1.5) and (1.6). This investigation will unify existing convergence results and provide new algorithms for solving convex feasibility problems. The crux of the proposed algorithm is to insert in (1.5) extrapolation steps of type (1.6) involving affine subspaces. More specifically, let I' be a subset of I containing indices corresponding to affine subspaces. Then our iterative scheme is based upon the updating rule

$$x_{n+1} = x_n + \rho_n \left(P_{\mathrm{i}(n)}x_n + \chi_n L_n \left(\sum_{i\in I_n} \omega_{i,n} P_{\mathrm{i}(n)} T_{i,n} P_{\mathrm{i}(n)} x_n - P_{\mathrm{i}(n)} x_n \right) - x_n \right), \qquad (1.7)$$

where

$$\begin{cases} \mathrm{i}(n) \in I', \\ I_n \subset I, \\ \{T_{i,n}\}_{i\in I_n} \subset \mathfrak{T}, \\ \{\omega_{i,n}\}_{i\in I_n} \subset [0,1], \quad \sum_{i\in I_n} \omega_{i,n} = 1, \\ L_n = \dfrac{\sum_{i\in I_n} \omega_{i,n}\|T_{i,n}P_{\mathrm{i}(n)}x_n - P_{\mathrm{i}(n)}x_n\|^2}{\|\sum_{i\in I_n} \omega_{i,n}T_{i,n}P_{\mathrm{i}(n)}x_n - P_{\mathrm{i}(n)}x_n\|^2}, \\ 0 < \chi_n \le K_n, \\ 0 < \rho_n < 2, \end{cases} \qquad (1.8)$$

and K_n is an extrapolation factor in $[1, +\infty[$ similar to that defined in (1.6). The practical importance of the insertion of these extrapolation steps is the acceleration of block-parallel methods of type (1.5). The method obtained is also much more flexible than algorithm (1.6).

An important tool in our analysis will be the notion of a $\mathfrak{T}$-class operator. In section 2, we review known results concerning these operators and establish new properties that will be used in subsequent sections. The new extrapolation algorithm is presented and analyzed in section 3. In particular, the convergence of the method is established and connections with existing results are made. In section 4, these results are specialized to the two-set case, which allows for significant refinements. Numerical simulation results that confirm the expected acceleration are also presented.

2 $\mathfrak{T}$-class operators

This paper hinges to a large extent on the following notion.

Definition 2.1 [4, Definition 2.2] $\mathfrak{T}$ *is the class of all operators* $T \colon \operatorname{dom} T = \mathcal{H} \to \mathcal{H}$ *such that*

$$(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T) \ \langle y - Tx \mid x - Tx \rangle \leq 0. \tag{2.1}$$

As the examples below show, the class $\mathfrak{T}$ contains various operators encountered in nonlinear analysis and applied mathematics.

Proposition 2.2 *Let* $T \colon \mathcal{H} \to \mathcal{H}$ *and consider the following statements:*

(i) T *is the projector onto a nonempty closed convex subset of* $\mathcal{H}$.

(ii) T *is the resolvent of a maximal monotone operator* $M \colon \mathcal{H} \to 2^{\mathcal{H}}$, *that is,* $T = (\operatorname{Id} + M)^{-1}$.

(iii) T *is the resolvent of a bifunction* $F \colon K \times K \to \mathbb{R}$, *where* K *is a nonempty closed convex subset of* $\mathcal{H}$ *and* F *satisfies*

    (a) $(\forall x \in K) \ F(x, x) = 0$.
    (b) $(\forall x \in K)(\forall y \in K) \ F(x, y) + F(y, x) \leq 0$.
    (c) *For every* $x \in K$, $F(x, \cdot) \colon K \to \mathbb{R}$ *is lower semicontinuous and convex.*
    (d) $(\forall x \in K)(\forall y \in K)(\forall z \in K) \ \varlimsup_{\varepsilon \to 0^+} F\big((1 - \varepsilon)x + \varepsilon z, y\big) \leq F(x, y)$.

    In other words, for every $x \in \mathcal{H}$, Tx *is the unique point in* K *that satisfies*

$$(\forall y \in K) \ F(Tx, y) + \langle Tx - x \mid y - Tx \rangle \geq 0. \tag{2.2}$$

(iv) $\operatorname{dom} T = \mathcal{H}$ *and* T *is firmly nonexpansive, that is,*

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \ \|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2. \tag{2.3}$$

(v) T *is a subgradient projector onto the lower level set* $\{x \in \mathcal{H} \mid f(x) \leq 0\}$, *that is,*

$$(\forall x \in \mathcal{H}) \ Tx = \begin{cases} x - \dfrac{f(x)}{\|g(x)\|^2} g(x), & \text{if } f(x) > 0; \\ x, & \text{if } f(x) \leq 0, \end{cases} \tag{2.4}$$

    where $f \colon \mathcal{H} \to \mathbb{R}$ *is a continuous convex function such that* $\{x \in \mathcal{H} \mid f(x) \leq 0\} \neq \emptyset$ *and where* g *is a selection of the subdifferential* ∂f *of* f.

(vi) $\operatorname{dom} T = \mathcal{H}$ *and* $R = 2T - \operatorname{Id}$ *is quasi-nonexpansive, that is,*

$$(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} R) \ \|Rx - y\| \leq \|x - y\|. \tag{2.5}$$

4

(vii) $T \in \mathfrak{T}$.

Then

$$
\begin{array}{ccccccc}
\text{(i)} & \Rightarrow & \text{(ii)} & \Leftrightarrow & \text{(iv)} & \Leftarrow & \text{(iii)} \\
\Downarrow & & & & \Downarrow & & \\
\text{(v)} & \Rightarrow & \text{(vi)} & \Leftrightarrow & \text{(vii)}, & &
\end{array}
\tag{2.6}
$$

where all the one-sided implications are not reversible in general.

Proof. See [25, Lemma 2.13(i)&(ii)] for (iii) $\Rightarrow$ (iv), and [4, Proposition 2.3] for the rest. $\square$

We now state some basic properties of $\mathfrak{T}$-class operators, starting with straightforward reformulations of (2.1).

Proposition 2.3 *Let $T\colon \mathcal{H} \to \mathcal{H}$ be an operator with $\operatorname{dom} T = \mathcal{H}$. Then the following statements are equivalent.*

(i) $T \in \mathfrak{T}$.

(ii) $(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T)\ \ \|Tx - x\|^2 \leq \langle y - x \mid Tx - x \rangle$.

(iii) $(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T)\ \ \|Tx - y\|^2 \leq \|x - y\|^2 - \|Tx - x\|^2$.

Proposition 2.4 *Let $T \in \mathfrak{T}$. Then*

(i) *Set $R = \operatorname{Id} + \kappa(T - \operatorname{Id})$, where $\kappa \in [0, 2]$. Then*

$$
(\forall x \in \mathcal{H})(\forall y \in \operatorname{Fix} T)\ \ \|Rx - y\|^2 \leq \|x - y\|^2 - \kappa(2 - \kappa)\|Tx - x\|^2.
\tag{2.7}
$$

(ii) $(\forall x \in \mathcal{H} \smallsetminus \operatorname{Fix} T)(\forall y \in \operatorname{Fix} T)\ \ \|Tx - y\| < \|x - y\|$.

(iii) *T is quasi-nonexpansive.*

(iv) *$\operatorname{Fix} T$ is closed and convex.*

Proof. (i): [23, Proposition 2.3(ii)]. (ii): Take $x \in \mathcal{H} \smallsetminus \operatorname{Fix} T$ and $y \in \operatorname{Fix} T$. It follows from Proposition 2.3 that $\|Tx - y\|^2 \leq \|x - y\|^2 - \|Tx - x\|^2 < \|x - y\|^2$. (iii): An immediate consequence of (ii). (iv): [4, Proposition 2.6(ii)]. $\square$

Proposition 2.5 *Let $(T_i)_{1 \leq i \leq m}$ be a finite family of operators in $\mathfrak{T}$ such that $\bigcap_{i=1}^{m} \operatorname{Fix} T_i \neq \emptyset$. Then $\operatorname{Fix} T_m \cdots T_1 = \bigcap_{i=1}^{m} \operatorname{Fix} T_i$.*

Proof. We proceed by induction. If $m = 1$, the property clearly holds. Now suppose that $m > 1$ and that, for some $p \in \{1, \ldots, m-1\}$, $\operatorname{Fix} T_p \cdots T_1 = \bigcap_{i=1}^{p} \operatorname{Fix} T_i$. Set $R_1 = T_p \cdots T_1$ and $R_2 = T_{p+1}$. Then $\varnothing \neq \bigcap_{i=1}^{m} \operatorname{Fix} T_i \subset \bigcap_{i=1}^{p+1} \operatorname{Fix} T_i = \operatorname{Fix} R_2 \cap \operatorname{Fix} R_1 \subset \operatorname{Fix} R_2 R_1$. Conversely, take $x \in \operatorname{Fix} R_2 R_1$. Then it is enough to show that $x \in \operatorname{Fix} R_1$ as this will imply that $x \in \operatorname{Fix} R_2$. Suppose to the contrary that $x \notin \operatorname{Fix} R_1$. Then $R_1 x \notin \operatorname{Fix} R_2$ since otherwise we would have $R_1 x = R_2 R_1 x = x \neq R_1 x$. Now let $y \in \operatorname{Fix} R_2 \cap \operatorname{Fix} R_1$. The operators $(T_i)_{1 \leq i \leq p}$ satisfy (2.5) by Proposition 2.4(iii) and $y \in \operatorname{Fix} R_1 = \bigcap_{i=1}^{p} \operatorname{Fix} T_i$. Therefore,

$$\|R_1 x - y\| = \|T_p \cdots T_1 x - y\| \leq \|T_{p-1} \cdots T_1 x - y\| \leq \cdots \leq \|T_1 x - y\| \leq \|x - y\|. \tag{2.8}$$

On the other hand, since R_2 satisfies property (ii) in Proposition 2.4, we obtain

$$\|x - y\| = \|R_2(R_1 x) - y\| < \|R_1 x - y\| \leq \|x - y\|, \tag{2.9}$$

which is absurd. Thus, $\operatorname{Fix} R_2 R_1 = \operatorname{Fix} R_2 \cap \operatorname{Fix} R_1$, i.e., $\operatorname{Fix} T_{p+1} \cdots T_1 = \bigcap_{i=1}^{p+1} \operatorname{Fix} T_i$. $\square$

We now introduce a new example of a $\mathfrak{T}$-class operator that will play a key role in the proposed algorithm. First, we recall some elementary properties of the projector onto an affine subspace. Henceforth, P_C designates the projector onto a nonempty closed convex subset C of $\mathcal{H}$.

Proposition 2.6 *Let A be a closed affine subspace of $\mathcal{H}$ and let $U = A - A$. Then*

(i) $(\forall x \in \mathcal{H})(\forall y \in A)(\forall z \in A) \ \langle x - P_A x \mid y - z \rangle = 0$.

(ii) $(\forall x \in \mathcal{H})(\forall y \in A) \ \|x - y\|^2 = \|x - P_A x\|^2 + \|P_A x - y\|^2$.

(iii) $(\forall x \in \mathcal{H})(\forall y \in A) \ P_A x = P_U x + P_{U^\perp} y$.

(iv) $(\forall x \in \mathcal{H})(\forall y \in A) \ \|P_A x - x\|^2 = \langle y - x \mid P_A x - x \rangle$.

Proposition 2.7 *Let A be a closed affine subspace of $\mathcal{H}$ and let T be an operator in $\mathfrak{T}$ such that $A \cap \operatorname{Fix} T \neq \varnothing$. Then*

$$(\forall x \in A)(\forall y \in A \cap \operatorname{Fix} T) \ \|Tx - x\|^2 \leq \langle y - x \mid P_A T x - x \rangle. \tag{2.10}$$

Proof. Take $x \in A$ and $y \in A \cap \operatorname{Fix} T$. Then Proposition 2.3 and Proposition 2.6(i) yield $\|Tx - x\|^2 \leq \langle y - x \mid Tx - x \rangle = \langle y - x \mid P_A T x - Tx \rangle + \langle y - x \mid Tx - x \rangle = \langle y - x \mid P_A T x - x \rangle$. $\square$

Theorem 2.8 *Let A be a closed affine subspace of $\mathcal{H}$ and let T be an operator in $\mathfrak{T}$ such that $A \cap \operatorname{Fix} T \neq \varnothing$. Define*

$$(\forall x \in \mathcal{H}) \ K(x, A, T) = \begin{cases} \dfrac{\|T P_A x - P_A x\|^2}{\|P_A T P_A x - P_A x\|^2}, & \text{if } P_A x \notin \operatorname{Fix} T; \\ 1, & \text{otherwise}, \end{cases} \tag{2.11}$$

and

$$R\colon \mathcal{H} \to A\colon x \mapsto P_A x + \kappa(x)\big(P_A T P_A x - P_A x\big), \quad \text{where } 0 < \kappa(x) \le K(x, A, T). \tag{2.12}$$

Now fix $x \in \mathcal{H}$. Then

(i) $K(x, A, T)$ *is a well defined number in $[1, +\infty[$; moreover,*

$$K(x, A, T) = 1 \quad \Leftrightarrow \quad T P_A x \in A \quad \Leftrightarrow \quad P_A T P_A x = T P_A x \tag{2.13}$$

and

$$P_A x \in \operatorname{Fix} T \quad \Leftrightarrow \quad P_A x \in A \cap \operatorname{Fix} T \quad \Rightarrow \quad T P_A x \in A \cap \operatorname{Fix} T \quad \Rightarrow \quad T P_A x \in A. \tag{2.14}$$

(ii) $K(x, A, T) \| P_A T P_A x - P_A x \|^2 = \| T P_A x - P_A x \|^2.$

(iii) $\operatorname{Fix} R = A \cap \operatorname{Fix} T.$

(iv) $R \in \mathfrak{T}.$

Proof. For brevity, we write $K(x)$ instead of $K(x, A, T)$. (i): It follows from Proposition 2.2(i) that $P_A \in \mathfrak{T}$. Consequently, since $A \cap \operatorname{Fix} T \neq \varnothing$, we deduce from Proposition 2.5 that

$$\operatorname{Fix} P_A T = \operatorname{Fix} P_A \cap \operatorname{Fix} T = A \cap \operatorname{Fix} T = \operatorname{Fix} T P_A. \tag{2.15}$$

Thus,

$$\begin{aligned}
T P_A x = P_A x \quad &\Leftrightarrow \quad P_A x \in \operatorname{Fix} T \\
&\Leftrightarrow \quad P_A x \in A \cap \operatorname{Fix} T \\
&\Leftrightarrow \quad P_A x \in \operatorname{Fix} P_A T \\
&\Leftrightarrow \quad P_A T P_A x = P_A x.
\end{aligned} \tag{2.16}$$

As a result, $K(x)$ is a well defined number in $\mathbb{R}$. Since A is an affine subspace, range $R \subset A$. Now, suppose that $P_A x \notin \operatorname{Fix} T$. Then (2.11) and Proposition 2.6(ii) yield

$$\begin{aligned}
K(x) &= \frac{\| T P_A x - P_A x \|^2}{\| P_A T P_A x - P_A x \|^2} \\
&= \frac{\| T P_A x - P_A T P_A x \|^2 + \| P_A T P_A x - P_A x \|^2}{\| P_A T P_A x - P_A x \|^2} \\
&= \frac{\| T P_A x - P_A T P_A x \|^2}{\| P_A T P_A x - P_A x \|^2} + 1.
\end{aligned} \tag{2.17}$$

Therefore $K(x) \ge 1$ and $K(x) = 1 \Leftrightarrow T P_A x = P_A T P_A x \Leftrightarrow T P_A x \in A$. Next, suppose that $P_A x \in \operatorname{Fix} T$. Then $T P_A x = P_A x \in A$. On the other hand, (2.11) yields $K(x) = 1$. Altogether,

7

we have verified (2.13). We now turn to (2.14) and observe that the equivalence is clear since range $P_A \subset A$. Next, if $P_A x \in A \cap \operatorname{Fix} T$, then $T P_A x = P_A x \in A \cap \operatorname{Fix} T$, which verifies the middle implication. The rightmost implication is obvious.

(ii): If $P_A x \notin \operatorname{Fix} T$, then the conclusion follows directly from (2.11). If $P_A x \in \operatorname{Fix} T$, then (2.11) and (2.13) imply that $K(x) = 1$ and that $P_A T P_A x = T P_A x = P_A x$. The conclusion is now clear.

(iii): Since A is a closed affine subspace and $\{P_A x, P_A T P_A x\} \subset A$, we have $x = P_A x \Leftrightarrow x \in A$ and

$$
\begin{aligned}
x \in \operatorname{Fix} R \quad &\Leftrightarrow \quad x = \big(1 - \kappa(x)\big) P_A x + \kappa(x) P_A T P_A x \in A \\
&\Leftrightarrow \quad x = \big(1 - \kappa(x)\big) x + \kappa(x) P_A T x \in A \\
&\Leftrightarrow \quad x = P_A T x \\
&\Leftrightarrow \quad x \in \operatorname{Fix} P_A T \\
&\Leftrightarrow \quad x \in A \cap \operatorname{Fix} T,
\end{aligned}
\tag{2.18}
$$

where (2.18) follows from (2.15).

(iv): Item (i) implies that $\operatorname{dom} R = \mathcal{H}$ and item (iii) states that $\operatorname{Fix} R = A \cap \operatorname{Fix} T$. Now let $y \in \operatorname{Fix} R$. In view of Proposition 2.3, it suffices to show that $\|Rx - x\|^2 \leq \langle y - x \mid Rx - x \rangle$. Since A is a closed affine subspace, we derive from Proposition 2.6(i) and from item (ii) that

$$
\begin{aligned}
\|Rx - x\|^2 &= \|P_A x - x + \kappa(x)\big(P_A T P_A x - P_A x\big)\|^2 \\
&= \|P_A x - x\|^2 + \kappa(x)^2 \|P_A T P_A x - P_A x\|^2 + 2\kappa(x)\langle P_A x - x \mid P_A T P_A x - P_A x \rangle \\
&\leq \|P_A x - x\|^2 + \kappa(x) K(x) \|P_A T P_A x - P_A x\|^2 \\
&= \|P_A x - x\|^2 + \kappa(x) \|T P_A x - P_A x\|^2.
\end{aligned}
\tag{2.19}
$$

Moreover, since $y \in A$, Proposition 2.6(iv) yields

$$
\|P_A x - x\|^2 = \langle y - x \mid P_A x - x \rangle.
\tag{2.20}
$$

On the other hand, it follows from Proposition 2.7 and Proposition 2.6(i) that

$$
\begin{aligned}
\|T P_A x - P_A x\|^2 &\leq \langle y - P_A x \mid P_A T P_A x - P_A x \rangle \\
&= \langle y - x \mid P_A T P_A x - P_A x \rangle + \langle x - P_A x \mid P_A T P_A x - P_A x \rangle \\
&= \langle y - x \mid P_A T P_A x - P_A x \rangle.
\end{aligned}
\tag{2.21}
$$

Upon combining (2.19), (2.20), and (2.21) we obtain

$$
\|Rx - x\|^2 \leq \langle y - x \mid P_A x - x \rangle + \kappa(x)\langle y - x \mid P_A T P_A x - P_A x \rangle = \langle y - x \mid Rx - x \rangle,
\tag{2.22}
$$

which is the desired inequality. $\square$

Remark 2.9 Several comments on Theorem 2.8 are in order.

(i) As will be shown in Example 4.9, we can have $K(x, A, T) = 1$ and $P_A x \notin \operatorname{Fix} T$ in (2.11).

(ii) Let $A = \mathcal{H}$, $T = P_{\{0\}}$, and $x \in \mathcal{H} \smallsetminus \{0\}$. Then $P_A x = x \neq 0 = T P_A x \in \{0\} = A \cap \operatorname{Fix} T$. This demonstrates that the implication $P_A x \in A \cap \operatorname{Fix} T \Rightarrow T P_A x \in A \cap \operatorname{Fix} T$ in (2.14) is not reversible in general.

(iii) Suppose that $T = P_B$ is the projector onto some closed convex set B such that $A \cap B \neq \emptyset$. If $T P_A x \in A$, then $T P_A x \in A \cap \operatorname{Fix} T$ since $\operatorname{range} T = \operatorname{range} P_B = B = \operatorname{Fix} T$. Thus the rightmost implication in (2.14) is actually an equivalence for projectors.

(iv) If T is not a projector, then the rightmost implication in (2.14) is not reversible in general. For instance, let $A = \mathcal{H}$ and $T = \frac{1}{2} \operatorname{Id}$. Then $T \in \mathfrak{T}$ and $\operatorname{Fix} T = \{0\}$. If $x \in \mathcal{H} \smallsetminus \{0\}$, then $T P_A x = \frac{1}{2} x \in \mathcal{H} = A$, yet $T P_A x \notin \operatorname{Fix} T$.

(v) As will be shown in Remark 4.10(ii), it is not possible to extend the range of the relaxation parameter $\kappa(x)$ in (2.12) beyond $K(x, A, T)$ without destroying the $\mathfrak{T}$-class property.

Corollary 2.10 *Let A be a closed affine subspace of $\mathcal{H}$ and let T be an operator in $\mathfrak{T}$ such that $A \cap \operatorname{Fix} T \neq \emptyset$. Then $\operatorname{Fix} P_A T P_A = A \cap \operatorname{Fix} T$ and $P_A T P_A \in \mathfrak{T}$.*

Proof. Set $\kappa(x) \equiv 1$ in Theorem 2.8. $\square$

In the next two corollaries of Theorem 2.8, we recover two results previously established with different tools.

Corollary 2.11 [4, Proposition 2.6(iii)] *Let $T \in \mathfrak{T}$. Then $(\forall \mu \in \,]0, 1])$ $\operatorname{Id} + \mu(T - \operatorname{Id}) \in \mathfrak{T}$.*

Proof. Set $A = \mathcal{H}$ and $\kappa(x) \equiv \mu$ in Theorem 2.8. $\square$

Corollary 2.12 [23, Proposition 2.4] *Let I be a finite ordered index set, let $(T_i)_{i \in I}$ be a family of operators in $\mathfrak{T}$ such that $C = \bigcap_{i \in I} \operatorname{Fix} T_i \neq \emptyset$, and let $(\omega_i)_{i \in I}$ be real numbers in $]0, 1]$ such that $\sum_{i \in I} \omega_i = 1$. Define*

$$(\forall x \in \mathcal{H}) \quad L\big(x, (T_i)_{i \in I}, (\omega_i)_{i \in I}\big) = \begin{cases} \dfrac{\sum_{i \in I} \omega_i \|T_i x - x\|^2}{\|\sum_{i \in I} \omega_i T_i x - x\|^2}, & \text{if } x \notin C; \\ 1, & \text{otherwise,} \end{cases} \tag{2.23}$$

and

$$R \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto x + \lambda(x) \left(\sum_{i \in I} \omega_i T_i x - x \right) \quad \text{where } 0 < \lambda(x) \leq L\big(x, (T_i)_{i \in I}, (\omega_i)_{i \in I}\big). \tag{2.24}$$

Then

(i) *For every $x \in \mathcal{H}$, $L\big(x, (T_i)_{i \in I}, (\omega_i)_{i \in I}\big)$ is a well defined number in $[1, +\infty[$.*

9

(ii) $\operatorname{Fix} R = C$.

(iii) $R \in \mathfrak{T}$.

Proof. For brevity we write $L(x)$ instead of $L\big(x, (T_i)_{i \in I}, (\omega_i)_{i \in I}\big)$. We use a product space technique devised in [48, 49] in the context of projection methods. Let $\mathbf{x} = (x_i)_{i \in I}$ denote a generic element in the product space $\boldsymbol{\mathcal{H}} = \bigtimes_{i \in I} \mathcal{H}$. The space $\boldsymbol{\mathcal{H}}$ endowed with the scalar product $\langle \cdot \mid \cdot \rangle \colon (\mathbf{x}, \mathbf{y}) \mapsto \sum_{i \in I} \omega_i \langle x_i \mid y_i \rangle$ is a Hilbert space. The associated norm is $\|\cdot\| \colon \mathbf{x} \mapsto \sqrt{\sum_{i \in I} \omega_i \|x_i\|^2}$. Let us introduce the closed vector subspace

$$\mathbf{D} = \big\{(x, \ldots, x) \in \boldsymbol{\mathcal{H}} \mid x \in \mathcal{H}\big\} \tag{2.25}$$

and the operator

$$\mathbf{T} \colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}} \colon \mathbf{x} \mapsto (T_i x_i)_{i \in I}. \tag{2.26}$$

The projector $\mathbf{P_D}$ onto $\mathbf{D}$ is given by

$$\mathbf{P_D} \colon \mathbf{x} \mapsto \left(\sum_{i \in I} \omega_i x_i, \ldots, \sum_{i \in I} \omega_i x_i \right). \tag{2.27}$$

It is easily verified that $\operatorname{Fix} \mathbf{T} = \bigtimes_{i \in I} \operatorname{Fix} T_i$ and that $\mathbf{T}$ belongs to the $\mathfrak{T}$-class relative to $\boldsymbol{\mathcal{H}}$. Now, let us define

$$(\forall \mathbf{x} \in \boldsymbol{\mathcal{H}}) \quad \mathbf{K}(\mathbf{x}) = \begin{cases} \dfrac{\|\mathbf{T P_D x} - \mathbf{P_D x}\|^2}{\|\mathbf{P_D T P_D x} - \mathbf{P_D x}\|^2}, & \text{if } \mathbf{P_D x} \notin \operatorname{Fix} \mathbf{T}; \\ 1, & \text{otherwise,} \end{cases} \tag{2.28}$$

and

$$\mathbf{R} \colon \boldsymbol{\mathcal{H}} \to \mathbf{D} \colon \mathbf{x} \mapsto \mathbf{P_D x} + \boldsymbol{\kappa}(\mathbf{x})\big(\mathbf{P_D T P_D x} - \mathbf{P_D x}\big), \quad \text{where } 0 < \boldsymbol{\kappa}(\mathbf{x}) \le \mathbf{K}(\mathbf{x}). \tag{2.29}$$

Then it follows from Theorem 2.8 that $\mathbf{R}$ is a $\mathfrak{T}$-class operator on $\boldsymbol{\mathcal{H}}$ with

$$\operatorname{Fix} \mathbf{R} = \mathbf{D} \cap \operatorname{Fix} \mathbf{T} = \big\{(x, \ldots, x) \mid x \in C\big\} \ne \varnothing. \tag{2.30}$$

Let us fix $x \in \mathcal{H}$ and set $\mathbf{x} = (x, \ldots, x) \in \mathbf{D}$. We deduce from (2.28), (2.26), (2.27), and (2.30) that

$$\mathbf{K}(\mathbf{x}) = \begin{cases} \dfrac{\|\mathbf{T x} - \mathbf{x}\|^2}{\|\mathbf{P_D T x} - \mathbf{x}\|^2} = \dfrac{\sum_{i \in I} \omega_i \|T_i x - x\|^2}{\|\sum_{i \in I} \omega_i T_i x - x\|^2}, & \text{if } x \notin C; \\ 1, & \text{otherwise} \end{cases}$$

$$= L(x) \tag{2.31}$$

and then from (2.29), (2.26), and (2.27) that

$$\mathbf{R x} = \mathbf{x} + \boldsymbol{\kappa}(\mathbf{x})\big(\mathbf{P_D T x} - \mathbf{x}\big) = \big(Rx, \ldots, Rx\big), \tag{2.32}$$

where R is as in (2.24), with $\lambda(x) = \boldsymbol{\kappa}(\mathbf{x})$. Hence, we obtain (i) and we derive from (2.30) that $\operatorname{Fix} R = C$, which establishes (ii). To show (iii), let us further fix $y \in C$ and set $\mathbf{y} = (y, \ldots, y) \in \mathbf{D} \cap \operatorname{Fix} \mathbf{R}$. Then $(\mathbf{x}, \mathbf{y}) \in \mathbf{D} \times \operatorname{Fix} \mathbf{R}$ and, since $\mathbf{R}$ is a $\mathfrak{T}$-class operator on $\boldsymbol{\mathcal{H}}$, we obtain

$$\big\langle y - Rx \mid x - Rx \big\rangle = \sum_{i \in I} \omega_i \big\langle y - Rx \mid x - Rx \big\rangle = \big\langle \mathbf{y} - \mathbf{R x} \mid \mathbf{x} - \mathbf{R x} \big\rangle \le 0. \tag{2.33}$$

Therefore, $R \in \mathfrak{T}$. $\square$

3 Extrapolation algorithm for affine-convex feasibility

Let $(S_i)_{i\in I}$ be a countable (finite or countably infinite) family of closed convex subsets of $\mathcal{H}$ such that

$$S = \bigcap_{i\in I} S_i \neq \varnothing. \tag{3.1}$$

Denote by I' a subset of I such that $(S_i)_{i\in I'}$ are closed affine subspaces onto which projections are easily computed (see [20] for concrete examples). It is assumed, without loss of generality, that $I' \neq \varnothing$. Indeed, one can always add the whole space $\mathcal{H}$ to the family $(S_i)_{i\in I}$ if necessary. For every $i \in I'$, we denote by P_i the projector onto the closed affine subspace S_i.

In this section, we present and analyze a flexible algorithm for finding a point in S that exploits the presence of the affine subspaces $(S_i)_{i\in I'}$ through extrapolated relaxations.

3.1 Extrapolation algorithm

The proposed algorithm involves a mix of $\mathfrak{T}$-class operators which are constructed by utilizing Theorem 2.8 and Corollary 2.12. Recall that the operators K and L are defined in (2.11) and (2.23), respectively.

Algorithm 3.1 Fix $\delta \in]0,1[$ and $x_0 \in \mathcal{H}$. At iteration $n \in \mathbb{N}$, x_n designates the current iterate and the update x_{n+1} is constructed according to the following steps.

① $i(n) \in I'$.

② $\varnothing \neq I_n \subset I$, where I_n is finite.

③ $(\forall i \in I_n)\ T_{i,n} \in \mathfrak{T}$, $\mathrm{Fix}\, T_{i,n} = S_i$, and, if $i \in I'$, then $T_{i,n} = P_i$.

④ $(\forall i \in I_n)\ \omega_{i,n} \in [0,1]$, $\sum_{i\in I_n} \omega_{i,n} = 1$, and

$$(\exists\, j \in I_n) \begin{cases} \|T_{j,n}P_{i(n)}x_n - P_{i(n)}x_n\| = \max_{i\in I_n} \|T_{i,n}P_{i(n)}x_n - P_{i(n)}x_n\| \\ \omega_{j,n} \geq \delta. \end{cases} \tag{3.2}$$

⑤ $I_n^+ = \big\{ i \in I_n \mid \omega_{i,n} > 0 \big\}$.

⑥ $R_n = \mathrm{Id} + L\big(\cdot, (T_{i,n})_{i\in I_n^+}, (\omega_{i,n})_{i\in I_n^+}\big)\big(\sum_{i\in I_n^+} \omega_{i,n} T_{i,n} - \mathrm{Id}\big)$.

⑦ $T_n \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto P_{i(n)}x + \kappa_n(x)\big(P_{i(n)}R_nP_{i(n)}x - P_{i(n)}x\big)$, where $0 < \kappa_n(x) \leq K(x, S_{i(n)}, R_n)$.

⑧ $x_{n+1} = x_n + \rho_n(T_nx_n - x_n)$, where $0 < \rho_n < 2$.

In view of Corollary 2.12 and Theorem 2.8(iv), the operators R_n and T_n defined in Step ⑥ and Step ⑦ are $\mathfrak{T}$-class operators.

3.2 Convergence analysis

We use the standard notation $x_n \rightharpoonup x$ and $x_n \to x$ to denote, respectively, the weak and strong convergence to x of a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$.

Condition 3.2

(i) There exists $\varepsilon \in \,]0,1[$ such that, for every $n \in \mathbb{N}$, $\varepsilon \leq \rho_n \leq 2 - \varepsilon$ and $\kappa_n(x_n) \geq \varepsilon$.

(ii) There exist strictly positive integers $(M_i)_{i\in I}$ such that

$$(\forall i \in I)(\forall n \in \mathbb{N}) \quad i \in \bigcup_{k=n}^{n+M_i-1} \{\mathrm{i}(k)\} \cup I_k. \tag{3.3}$$

(iii) For every index $i \in I \smallsetminus I'$, every orbit $(x_n)_{n\in\mathbb{N}}$ of Algorithm 3.1, every $y \in \mathcal{H}$, and every strictly increasing sequence $(p_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$, the implication

$$\left.\begin{array}{r} i \in \bigcap_{n\in\mathbb{N}} I_{p_n}, \\ P_{\mathrm{i}(p_n)} x_{p_n} \rightharpoonup y, \\ T_{i,p_n} P_{\mathrm{i}(p_n)} x_{p_n} - P_{\mathrm{i}(p_n)} x_{p_n} \to 0 \end{array}\right\} \quad \Rightarrow \quad y \in S_i \tag{3.4}$$

holds.

We are now ready to present our main result.

Theorem 3.3 *Let $(x_n)_{n\in\mathbb{N}}$ be an arbitrary orbit of Algorithm 3.1. Then*

(i) *$(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to S, that is,*

$$(\forall y \in S)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - y\| \leq \|x_n - y\|. \tag{3.5}$$

(ii) $\sum_{n\in\mathbb{N}} \rho_n(2 - \rho_n)\|P_{\mathrm{i}(n)}x_n - x_n\|^2 < +\infty.$

(iii) $\sum_{n\in\mathbb{N}} \rho_n\kappa_n(x_n)\left(2 - \rho_n\kappa_n(x_n)/K_n(x_n, S_{\mathrm{i}(n)}, R_n)\right)\|R_nP_{\mathrm{i}(n)}x_n - P_{\mathrm{i}(n)}x_n\|^2 < +\infty.$

(iv) $\sum_{n\in\mathbb{N}} \rho_n(2 - \rho_n)\|T_nx_n - x_n\|^2 < +\infty.$

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

12

(vi) *If Condition 3.2 is satisfied, then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in S.*

Proof. Let us fix $y \in S$ and, for brevity, let us set

$$(\forall n \in \mathbb{N}) \quad \begin{cases} z_n = P_{\mathrm{i}(n)}x_n, \\ L_n = L\big(z_n, (T_{i,n})_{i\in I_n^+}, (\omega_{i,n})_{i\in I_n^+}\big), \\ K_n = K(x_n, S_{\mathrm{i}(n)}, R_n), \\ \chi_n = \kappa_n(x_n). \end{cases} \tag{3.6}$$

For every $n \in \mathbb{N}$, the weights $(\omega_{i,n})_{i\in I_n^+}$ are strictly positive and $\sum_{i\in I_n^+}\omega_{i,n} = 1$. Therefore, ③, ⑥, (3.1), and Corollary 2.12 yield

$$(\forall n \in \mathbb{N}) \quad R_n \in \mathfrak{T} \quad \text{and} \quad \operatorname{Fix} R_n = \bigcap_{i\in I_n^+} S_i. \tag{3.7}$$

It follows from ⑧ that

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - y\|^2 = \|x_n - y + \rho_n(T_n x_n - x_n)\|^2$$
$$= \|x_n - y\|^2 + 2\rho_n\langle x_n - y \mid T_n x_n - x_n\rangle + \rho_n^2\|T_n x_n - x_n\|^2. \tag{3.8}$$

We observe that, for every $n \in \mathbb{N}$, the points y, z_n, and $P_{\mathrm{i}(n)}R_n z_n$ belong to the closed affine subspace $S_{\mathrm{i}(n)}$. Therefore, ⑦, Proposition 2.6(i), (3.7), and Proposition 2.7 yield

$$(\forall n \in \mathbb{N}) \quad \langle x_n - y \mid T_n x_n - x_n\rangle = \big\langle x_n - y \mid z_n - x_n + \chi_n(P_{\mathrm{i}(n)}R_n z_n - z_n)\big\rangle$$
$$= \big\langle x_n - z_n \mid z_n - x_n + \chi_n(P_{\mathrm{i}(n)}R_n z_n - z_n)\big\rangle$$
$$+ \big\langle z_n - y \mid z_n - x_n + \chi_n(P_{\mathrm{i}(n)}R_n z_n - z_n)\big\rangle$$
$$= -\|z_n - x_n\|^2 + \chi_n\langle z_n - y \mid P_{\mathrm{i}(n)}R_n z_n - z_n\rangle$$
$$\leq -\|z_n - x_n\|^2 - \chi_n\|R_n z_n - z_n\|^2. \tag{3.9}$$

On the other hand, by ⑦, Proposition 2.6(i), and Theorem 2.8(ii),

$$(\forall n \in \mathbb{N}) \quad \|T_n x_n - x_n\|^2 = \|z_n - x_n + \chi_n(P_{\mathrm{i}(n)}R_n z_n - z_n)\|^2$$
$$= \|z_n - x_n\|^2 + \chi_n^2\|P_{\mathrm{i}(n)}R_n z_n - z_n\|^2$$
$$= \|z_n - x_n\|^2 + \frac{\chi_n^2}{K_n}\|R_n z_n - z_n\|^2. \tag{3.10}$$

Altogether, (3.8), (3.9), and (3.10) result in

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - y\|^2 \leq \|x_n - y\|^2 - 2\rho_n\big(\|z_n - x_n\|^2 + \chi_n\|R_n z_n - z_n\|^2\big)$$
$$+ \rho_n^2\left(\|z_n - x_n\|^2 + \frac{\chi_n^2}{K_n}\|R_n z_n - z_n\|^2\right)$$
$$= \|x_n - y\|^2 - \rho_n(2 - \rho_n)\|z_n - x_n\|^2$$
$$- \rho_n\chi_n\left(2 - \frac{\rho_n\chi_n}{K_n}\right)\|R_n z_n - z_n\|^2. \tag{3.11}$$

13

Observe that the conditions imposed on the relaxation parameters in ⑦ and ⑧ yield

$$(\forall n \in \mathbb{N}) \quad \rho_n(2 - \rho_n) > 0 \quad \text{and} \quad \rho_n\chi_n\left(2 - \frac{\rho_n\chi_n}{K_n}\right) > 0. \tag{3.12}$$

Let us now prove items (i)–(vi).

(i) follows at once from (3.11) and (3.12). As a result, we obtain

$$\sum_{n \in \mathbb{N}} \left(\|x_n - y\|^2 - \|x_{n+1} - y\|^2\right) \leq \|x_0 - y\|^2 < +\infty. \tag{3.13}$$

(ii)&(iii): We deduce from (3.11) that

$$(\forall n \in \mathbb{N}) \quad \rho_n(2-\rho_n)\|z_n - x_n\|^2 + \rho_n\chi_n\left(2 - \frac{\rho_n\chi_n}{K_n}\right)\|R_n z_n - z_n\|^2 \leq \|x_n - y\|^2 - \|x_{n+1} - y\|^2. \tag{3.14}$$

Hence the claims follow from (3.12) and (3.13).

(iv): By ⑦,

$$(\forall n \in \mathbb{N}) \quad 0 < \frac{\chi_n}{K_n} \leq 1, \tag{3.15}$$

and therefore

$$(\forall n \in \mathbb{N}) \quad \rho_n(2 - \rho_n)\frac{\chi_n^2}{K_n} = \rho_n\chi_n\left(\frac{2\chi_n}{K_n} - \frac{\rho_n\chi_n}{K_n}\right) \leq \rho_n\chi_n\left(2 - \frac{\rho_n\chi_n}{K_n}\right). \tag{3.16}$$

Thus, (3.12), (3.10), (3.16), and (3.14) yield

$$(\forall n \in \mathbb{N}) \quad 0 \leq \rho_n(2 - \rho_n)\|T_n x_n - x_n\|^2$$
$$= \rho_n(2 - \rho_n)\|z_n - x_n\|^2 + \rho_n(2 - \rho_n)\frac{\chi_n^2}{K_n}\|R_n z_n - z_n\|^2$$
$$\leq \rho_n(2 - \rho_n)\|z_n - x_n\|^2 + \rho_n\chi_n\left(2 - \frac{\rho_n\chi_n}{K_n}\right)\|R_n z_n - z_n\|^2$$
$$\leq \|x_n - y\|^2 - \|x_{n+1} - y\|^2. \tag{3.17}$$

Hence the assertion follows from (3.13).

(v): In view of ⑧,

$$(\forall n \in \mathbb{N}) \quad (2 - \rho_n)\|x_{n+1} - x_n\|^2 = \rho_n^2(2 - \rho_n)\|T_n x_n - x_n\|^2 \leq 2\rho_n(2 - \rho_n)\|T_n x_n - x_n\|^2, \tag{3.18}$$

and the claim follows from (iv).

(vi): In view of (i), to show that $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in S, it is enough to show (see, e.g., [4, Proposition 2.1]) that each of its weak cluster points is in S. To this end, fix a weak

14

cluster point z, say $x_{k_n} \rightharpoonup z$, and an index $j \in I$. Then it suffices to show that $z \in S_j$. Using successively ⑥, the inequality $L_n \geq 1$ (see Corollary 2.12(i)), (2.23), and ④, we obtain

$$
\begin{aligned}
(\forall n \in \mathbb{N}) \ \|R_n z_n - z_n\|^2 = L_n^2 &\left\| \sum_{i \in I_n^+} \omega_{i,n} T_{i,n} z_n - z_n \right\|^2 \\
\geq L_n &\left\| \sum_{i \in I_n^+} \omega_{i,n} T_{i,n} z_n - z_n \right\|^2 \\
= &\sum_{i \in I_n^+} \omega_{i,n} \|T_{i,n} z_n - z_n\|^2 \\
\geq &\, \delta \max_{i \in I_n} \|T_{i,n} z_n - z_n\|^2.
\end{aligned}
\tag{3.19}
$$

On the other hand, Condition 3.2(i), (3.15) and (iii) yield

$$
\begin{aligned}
\varepsilon^3 \sum_{n \in \mathbb{N}} \|R_n z_n - z_n\|^2 &\leq \sum_{n \in \mathbb{N}} \rho_n \chi_n (2 - \rho_n) \|R_n z_n - z_n\|^2 \\
&\leq \sum_{n \in \mathbb{N}} \rho_n \chi_n (2 - \rho_n \chi_n / K_n) \|R_n z_n - z_n\|^2 \\
&< +\infty.
\end{aligned}
\tag{3.20}
$$

It follows from (3.19) and (3.20) that

$$
\max_{i \in I_n} \|T_{i,n} z_n - z_n\| \to 0.
\tag{3.21}
$$

After passing to a subsequence of $(x_{k_n})_{n \in \mathbb{N}}$ if necessary, we assume that, for every $n \in \mathbb{N}$, $k_{n+1} \geq k_n + M_j$. Then Condition 3.2(ii) asserts that there exists a sequence $(p_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$ such that

$$
(\forall n \in \mathbb{N}) \ k_n \leq p_n \leq k_n + M_j - 1 < k_{n+1} \leq p_{n+1} \quad \text{and} \quad j \in \{\mathrm{i}(p_n)\} \cup I_{p_n}.
\tag{3.22}
$$

The Cauchy-Schwarz inequality then yields

$$
(\forall n \in \mathbb{N}) \ \|x_{p_n} - x_{k_n}\| \leq \sum_{l=k_n}^{k_n + M_j - 2} \|x_{l+1} - x_l\| \leq \sqrt{M_j - 1} \sqrt{\sum_{l \geq k_n} \|x_{l+1} - x_l\|^2}.
\tag{3.23}
$$

Hence (v) and Condition 3.2(i) imply that $x_{p_n} - x_{k_n} \to 0$, and thus $x_{p_n} \rightharpoonup z$. Since (ii) and Condition 3.2(i) result in $z_{p_n} - x_{p_n} \to 0$, we therefore obtain

$$
z_{p_n} \rightharpoonup z.
\tag{3.24}
$$

In view of (3.22), there exists a subsequence of $(p_n)_{n \in \mathbb{N}}$, which we still denote by $(p_n)_{n \in \mathbb{N}}$ for convenience, such that

$$
(\forall n \in \mathbb{N}) \ j = \mathrm{i}(p_n)
\tag{3.25}
$$

15

or

$$j \in \bigcap_{n\in\mathbb{N}} I_{p_n}. \tag{3.26}$$

Assume first that (3.25) holds. Using (3.24), we see that $S_j \ni P_j x_{p_n} = z_{p_n} \rightharpoonup z$. Hence $z \in S_j$, as S_j is weakly closed.

Now assume that (3.26) holds and observe that (3.21) and (3.24) yield

$$T_{j,p_n} z_{p_n} - z_{p_n} \to 0 \quad \text{and} \quad z_{p_n} \rightharpoonup z. \tag{3.27}$$

In turn, we obtain $T_{j,p_n} z_{p_n} \rightharpoonup z$. If $j \in I'$, then ③ gives $T_{j,p_n} \equiv P_j$ and therefore $S_j \ni P_j z_{p_n} \rightharpoonup z$, hence $z \in S_j$ as S_j is weakly closed. On the other hand, if $j \notin I'$, then we deduce at once from (3.27) and Condition 3.2(iii) that $z \in S_j$.

We have thus shown that $z \in S_j$ in all cases, which completes the proof. $\square$

Remark 3.4 In the setting of Theorem 3.3, we observe that (3.7) and Theorem 2.8 imply that

$$(\forall n \in \mathbb{N}) \ \ T_n \in \mathfrak{T} \ \text{ and } \ \text{Fix}\, T_n = S_{\mathrm{i}(n)} \cap \text{Fix}\, R_n = S_{\mathrm{i}(n)} \cap \bigcap_{i\in I_n^+} S_i. \tag{3.28}$$

Therefore, ⑧ and the results of [23, Section 4] imply directly that $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to S, and that items (iv) and (v) hold in Theorem 3.3. However, under our general working hypotheses, it does not seem possible to derive the crucial item (vi) of Theorem 3.3 from the framework of [23].

3.3   Implementation and special cases

Algorithm 3.1 has been presented above in a form that suits the theoretical convergence analysis of section 3.2 well. From a practical viewpoint, however, the operation of the algorithm at step n can be more conveniently broken up as follows. One first selects the index i(n) of an affine subspace onto which the projection $P_{\mathrm{i}(n)} x_n$ is computed. Then one selects a finite block of indices I_n and, for every index $i \in I_n$, one chooses an operator $T_{i,n} \in \mathfrak{T}$ such that $\text{Fix}\, T_{i,n} = S_i$. The choice of these operators is tailored to a form in which the set S_i can most conveniently be expressed as a fixed point set; some examples are given in Proposition 2.2. Next, one evaluates the vectors $(T_{i,n} P_{\mathrm{i}(n)} x_n)_{i\in I_n}$. This step can be distributed over parallel processors and, through a judicious choice of I_n, it may be possible to match the computational load with the distributed computer resources at hand. The next step is a coordination phase in which the convex combination $\sum_{i\in I_n^+} \omega_{i,n} T_{i,n} P_{\mathrm{i}(n)} x_n$ is formed.

Using the fact that $P_{\mathrm{i}(n)}$ is affine, we see that the update can then be computed as

$$x_{n+1} = x_n + \rho_n\big(T_n x_n - x_n\big)$$

$$= x_n + \rho_n\Big(P_{\mathrm{i}(n)} x_n + \chi_n\big(P_{\mathrm{i}(n)} R_n P_{\mathrm{i}(n)} x_n - P_{\mathrm{i}(n)} x_n\big) - x_n\Big)$$

$$= x_n + \rho_n\left(P_{\mathrm{i}(n)} x_n + \chi_n L_n\left(P_{\mathrm{i}(n)}\left(\sum_{i\in I_n^+} \omega_{i,n} T_{i,n} P_{\mathrm{i}(n)} x_n\right) - P_{\mathrm{i}(n)} x_n\right) - x_n\right), \qquad (3.29)$$

where $\rho_n \in]0,2[$, $\chi_n = \kappa_n(x_n)$, and $L_n = L\big(P_{\mathrm{i}(n)} x_n, (T_{i,n})_{i\in I_n^+}, (\omega_{i,n})_{i\in I_n^+}\big)$. In (3.29), the range of the parameter χ_n is $]0, K(x_n, S_{\mathrm{i}(n)}, R_n)]$. Furthermore, Corollary 2.12(ii) and ③ imply $\operatorname{Fix} R_n = \bigcap_{i\in I_n^+} \operatorname{Fix} T_{i,n} = \bigcap_{i\in I_n^+} S_i$. Thus, using (2.11) and the fact that $P_{\mathrm{i}(n)}$ is an affine operator, we obtain

$$0 < \chi_n \le \begin{cases} \dfrac{\|\sum_{i\in I_n^+} \omega_{i,n} T_{i,n} P_{\mathrm{i}(n)} x_n - P_{\mathrm{i}(n)} x_n\|^2}{\|P_{\mathrm{i}(n)}\big(\sum_{i\in I_n^+} \omega_{i,n} T_{i,n} P_{\mathrm{i}(n)} x_n\big) - P_{\mathrm{i}(n)} x_n\|^2}, & \text{if } P_{\mathrm{i}(n)} x_n \notin \bigcap_{i\in I_n^+} S_i; \\ 1, & \text{otherwise.} \end{cases} \qquad (3.30)$$

Likewise, we derive from (2.23) that

$$L_n = \begin{cases} \dfrac{\sum_{i\in I_n^+} \omega_{i,n}\|T_{i,n} P_{\mathrm{i}(n)} x_n - P_{\mathrm{i}(n)} x_n\|^2}{\|\sum_{i\in I_n^+} \omega_{i,n} T_{i,n} P_{\mathrm{i}(n)} x_n - P_{\mathrm{i}(n)} x_n\|^2}, & \text{if } P_{\mathrm{i}(n)} x_n \notin \bigcap_{i\in I_n^+} S_i; \\ 1, & \text{otherwise.} \end{cases} \qquad (3.31)$$

Altogether, the parameter $\chi_n L_n$ satisfies

$$0 < \chi_n L_n \le \begin{cases} \dfrac{\sum_{i\in I_n^+} \omega_{i,n}\|T_{i,n} P_{\mathrm{i}(n)} x_n - P_{\mathrm{i}(n)} x_n\|^2}{\|P_{\mathrm{i}(n)}\big(\sum_{i\in I_n^+} \omega_{i,n} T_{i,n} P_{\mathrm{i}(n)} x_n\big) - P_{\mathrm{i}(n)} x_n\|^2}, & \text{if } P_{\mathrm{i}(n)} x_n \notin \bigcap_{i\in I_n^+} S_i; \\ 1, & \text{otherwise.} \end{cases} \qquad (3.32)$$

We thus arrive at the following practical implementation of Algorithm 3.1.

Algorithm 3.5 Fix $\delta \in]0,1[$ and $x_0 \in \mathcal{H}$. At iteration $n \in \mathbb{N}$, x_n designates the current iterate and the update x_{n+1} is constructed by executing the following steps.

① $\mathrm{i}(n) \in I'$ and $z_n = P_{\mathrm{i}(n)} x_n$.

② $\varnothing \neq I_n \subset I$, where I_n is finite.

③ $(\forall i \in I_n)$ $t_{i,n} = T_{i,n} z_n$, where $T_{i,n} \in \mathfrak{T}$, $\operatorname{Fix} T_{i,n} = S_i$, and, if $i \in I'$, then $t_{i,n} = P_i z_n$.

④ $(\forall i \in I_n)$ $\omega_{i,n} \in [0,1]$, $\sum_{i\in I_n} \omega_{i,n} = 1$, and

$$(\exists j \in I_n) \begin{cases} \|t_{j,n} - z_n\| = \max_{i\in I_n} \|t_{i,n} - z_n\| \\ \omega_{j,n} \ge \delta. \end{cases} \qquad (3.33)$$

17

⑤ $I_n^+ = \{i \in I_n \mid \omega_{i,n} > 0\}$.

⑥ $d_n = \sum_{i \in I_n^+} \omega_{i,n} t_{i,n}$ and $p_n = P_{i(n)} d_n$.

⑦ $0 < \rho_n < 2$ and $0 < \mu_n \leq \begin{cases} \dfrac{\sum_{i \in I_n^+} \omega_{i,n} \|t_{i,n} - z_n\|^2}{\|p_n - z_n\|^2}, & \text{if } z_n \notin \bigcap_{i \in I_n^+} S_i; \\ 1, & \text{otherwise.} \end{cases}$

⑧ $x_{n+1} = x_n + \rho_n (z_n + \mu_n (p_n - z_n) - x_n)$.

Remark 3.6

(i) Consider the special case when $\{S_i\}_{i \in I'} = \{\mathcal{H}\}$. Then, for every $n \in \mathbb{N}$, $P_{i(n)} = P_{\mathcal{H}} = \mathrm{Id}$, $z_n = x_n$, and $p_n = d_n$. Consequently, the above iteration becomes

$$\begin{aligned} x_{n+1} &= x_n + \rho_n (z_n + \mu_n (p_n - z_n) - x_n) \\ &= x_n + \rho_n \mu_n (d_n - x_n) \\ &= x_n + \lambda_n \left(\sum_{i \in I_n^+} \omega_{i,n} T_{i,n} x_n - x_n \right), \end{aligned} \tag{3.34}$$

where

$$0 < \lambda_n < \begin{cases} 2 \dfrac{\sum_{i \in I_n^+} \omega_{i,n} \|T_{i,n} x_n - x_n\|^2}{\|\sum_{i \in I_n^+} \omega_{i,n} T_{i,n} x_n - x_n\|^2}, & \text{if } x_n \notin \bigcap_{i \in I_n^+} S_i; \\ 2, & \text{otherwise.} \end{cases} \tag{3.35}$$

This is precisely the parallel block-iterative algorithm discussed in [23] which, in turn, covers the projection methods of [1, 2, 3, 9, 22, 26, 32, 33, 34, 40, 47, 49], the firmly nonexpansive operator methods of [10, 19, 43, 53], the subgradient projection methods of [3, 15, 21, 31, 50, 54, 55], the proximal point algorithms of [4, 44, 52], and the equilibrium programming algorithm of [25].

(ii) Another special case of interest, leading in particular to (1.6), arises when $\{S_i\}_{i \in I'} = \{A\}$ and when, for every $n \in \mathbb{N}$, the family $\{S_i\}_{i \in I_n}$ reduces to a single set $\{B\}$. This setting will be discussed in detail in Section 4.

(iii) Suppose that the solution set S in (3.1) is a closed vector subspace. Then it follows from Theorem 3.3(i) and [5, Fact 2.2] that Theorem 3.3(vi) can be strengthened to "If Condition 3.2 is satisfied, then $(x_n)_{n \in \mathbb{N}}$ converges weakly to $P_S x_0$." In particular, let A and B be closed vector subspaces of $\mathcal{H}$ and consider the problem of finding the projection of a given point $x \in \mathcal{H}$ onto $A \cap B$. Specializing Gubin et al.'s extrapolation algorithm [40] to this setting leads to the iteration

$$x_0 = P_A P_B x \quad \text{and} \quad (\forall n \in \mathbb{N}) \ x_{n+1} = x_n + \mu_n (P_A P_B x_n - x_n), \tag{3.36}$$

18

where

$$\mu_n = \begin{cases} \dfrac{\langle x_n \mid x_n - P_A P_B x_n \rangle}{\|P_A P_B x_n - x_n\|^2}, & \text{if } x_n \notin A \cap B; \\ 1, & \text{otherwise.} \end{cases} \tag{3.37}$$

Since $(x_n)_{n \in \mathbb{N}}$ lies in A, it follows from the fact that P_A is self-adjoint and Proposition 2.6(iv) that

$$(\forall n \in \mathbb{N}) \ \langle x_n \mid x_n - P_A P_B x_n \rangle = \langle x_n \mid P_A x_n - P_A P_B x_n \rangle$$
$$= \langle x_n \mid x_n - P_B x_n \rangle$$
$$= \|P_B x_n - x_n\|^2. \tag{3.38}$$

Consequently, for every $n \in \mathbb{N}$, $\mu_n = K(x_n, A, P_B)$. It should be noted that, in this specific linear setting, the recent results on the iteration (3.36)–(3.37) provided in [6] (see also [5]) guarantee strong convergence.

3.4   On Condition 3.2

The purpose of this section is to give explicit examples of scenarios in which Condition 3.2 holds.

Condition 3.2(i) is rather standard in this type of iterative methods (see [3, 12, 23] and the references therein) and it simply imposes hard bounds on the sequences $(\rho_n)_{n \in \mathbb{N}}$ and $(\kappa_n(x_n))_{n \in \mathbb{N}}$.

Condition 3.2(ii) requires a certain regularity on the order in which the indices in I are used, namely, that every index i be used at least once over M_i consecutive iterations. In the most general case, both I' and $I'' = I \smallsetminus I'$ are countable sets, say $I' = \{1, 2, 3, \cdots\}$ and $I'' = -I'$. For every $i \in I$, set $M_i = 2^{|i|}$. Then an example of control rules that satisfy (3.3) is obtained by setting

$$i(0) = 1, \quad I_0 = -\{1, 2\}, \tag{3.39}$$

and

$$\left.\begin{array}{c} n \in \mathbb{N} \smallsetminus \{0\}, \\ m \in \mathbb{N} \smallsetminus \{0\}, \\ n = 2^{m-1} \ (\text{modulo } 2^m) \end{array}\right\} \quad \Rightarrow \quad \left\{\begin{array}{c} i(n) = m, \\ I_n = -\{2m - 1, 2m\}, \end{array}\right. \tag{3.40}$$

i.e., by

$$\begin{aligned} \big(i(n)\big)_{n \in \mathbb{N}} = \big(&1, 1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1, 5, 1, 2, 1, 3, 1, 2, 1, 4, \\ &1, 2, 1, 3, 1, 2, 1, 6, 1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1, 5, \\ &1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1, 7, 1, 2, 1, 3, 1, 2, 1, 4, \\ &1, 2, 1, 3, 1, 2, 1, 5, 1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1, 6, \\ &1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, \dots \big) \end{aligned} \tag{3.41}$$

and

$$\left(I_n\right)_{n\in\mathbb{N}} = -\big(\{1,2\}, \{1,2\}, \{3,4\}, \{1,2\}, \{5,6\}, \{1,2\}, \{3,4\}, \{1,2\}, \{7,8\},$$
$$\{1,2\}, \{3,4\}, \{1,2\}, \{5,6\}, \{1,2\}, \{3,4\}, \{1,2\}, \{9,10\},$$
$$\{1,2\}, \{3,4\}, \{1,2\}, \{5,6\}, \{1,2\}, \{3,4\}, \{1,2\}, \{7,8\},$$
$$\{1,2\}, \{3,4\}, \{1,2\}, \{5,6\}, \{1,2\}, \{3,4\}, \{1,2\}, \{11,12\},$$
$$\{1,2\}, \{3,4\}, \{1,2\}, \{5,6\}, \{1,2\}, \{3,4\}, \{1,2\}, \{7,8\}, \dots\big). \tag{3.42}$$

In practice, $\left(\mathrm{i}(n)\right)_{n\in\mathbb{N}}$ and $\left(I_n\right)_{n\in\mathbb{N}}$ can be precomputed in a fashion similar to the determination of prime numbers by the sieve of Eratosthenes.

Let us now turn to Condition 3.2(iii) and the choice of the operators $(T_{i,n})_{i\in I_n}$ at Step ③ of Algorithm 3.1. Let us fix an index $i \in I$ and let $(p_n)_{n\in\mathbb{N}}$ be the sequence of all nonnegative integers such that S_i is activated at Step ③, i.e., such that $i \in \bigcap_{n\in\mathbb{N}} I_{p_n}$. If $i \in I'$, then Step ③ imposes $T_{i,p_n} \equiv P_i$. We therefore consider the case when $i \notin I'$. We must construct a sequence $(T_{i,p_n})_{n\in\mathbb{N}}$ in $\mathfrak{T}$ such that $\operatorname{Fix} T_{i,p_n} \equiv S_i$ and (3.4) holds. These requirements will be met if the conditions

$$(\forall n \in \mathbb{N}) \quad T_{i,p_n} \in \mathfrak{T} \quad \text{and} \quad \operatorname{Fix} T_{i,p_n} = S_i \tag{3.43}$$

and

$$\left.\begin{array}{c} (y_n)_{n\in\mathbb{N}} \text{ in } \mathcal{H}, \\ y_n \rightharpoonup y, \\ T_{i,p_n} y_n - y_n \to 0 \end{array}\right\} \quad \Rightarrow \quad y \in S_i \tag{3.44}$$

are satisfied. As we now illustrate, (3.43) and (3.44) hold in important concrete situations.

Proposition 3.7 *Suppose that $S_i \neq \varnothing$. Then (3.43) and (3.44) are satisfied in each of the following cases.*

(i) *For every $n \in \mathbb{N}$, $T_{i,p_n} = P_i$ is the projector onto S_i.*

(ii) *$S_i = A_i^{-1}(0)$, where $A_i \colon \mathcal{H} \to 2^{\mathcal{H}}$ is a maximal monotone operator and, for every $n \in \mathbb{N}$, T_{i,p_n} is the resolvent of $\gamma_{i,p_n} A_i$, where $(\gamma_{i,p_n})_{n\in\mathbb{N}}$ is a sequence in $]0,+\infty[$ such that $\inf_{n\in\mathbb{N}} \gamma_{i,p_n} > 0$.*

(iii) *$S_i = \big\{x \in \mathcal{H} \mid (\forall y \in K)\ F_i(x,y) \geq 0\big\}$, where K is a nonempty closed convex subset of $\mathcal{H}$ and $F_i \colon K \times K \to \mathbb{R}$ satisfies*

    (a) *$(\forall x \in K)\ F_i(x,x) = 0$,*

    (b) *$(\forall x \in K)(\forall y \in K)\ F_i(x,y) + F_i(y,x) \leq 0$,*

    (c) *For every $x \in K$, $F_i(x,\cdot)\colon K \to \mathbb{R}$ is lower semicontinuous and convex,*

    (d) *$(\forall x \in K)(\forall y \in K)(\forall z \in K)\ \varlimsup_{\varepsilon\to 0^+} F_i\big((1-\varepsilon)x + \varepsilon z, y\big) \leq F_i(x,y)$,*

    and, for every $n \in \mathbb{N}$, T_{i,p_n} is the resolvent of $\gamma_{i,p_n} F_i$, where $(\gamma_{i,p_n})_{n\in\mathbb{N}}$ is a sequence in $]0,+\infty[$ such that $\inf_{n\in\mathbb{N}} \gamma_{i,p_n} > 0$.

(iv) $S_i = \operatorname{Fix} R_i$, *where* $R_i \colon \operatorname{dom} R_i = \mathcal{H} \to \mathcal{H}$ *is a firmly nonexpansive operator and, for every* $n \in \mathbb{N}$, $T_{i,p_n} = R_i$.

(v) $S_i = \{x \in \mathcal{H} \mid f_i(x) \leq 0\}$, *where* $f_i \colon \mathcal{H} \to \mathbb{R}$ *is a continuous convex function which is bounded on bounded sets and, for every* $n \in \mathbb{N}$, T_{i,p_n} *is a subgradient projector onto* S_i *(see* (2.4)*).*

Proof. (i): See Proposition 2.2(i) and [3, Example 3.8]. (ii): See Proposition 2.2(ii) and [23, Section 6.3]. (iii): See Proposition 2.2(iii) and [25, Section 3]. (iv): See Proposition 2.2(iv) and [3, Example 3.8]. (v): See Proposition 2.2(v) and [3, Theorem 7.7]. □

In practice, item (i) corresponds to the case when it is relatively easy to compute the best approximation to x from S_i (see [3, 20, 30] for examples); item (ii) corresponds to the monotone inclusion problem $0 \in A_i x$, which arises in many applied mathematics problems [24, 52, 56]; item (iii) corresponds to equilibrium problems [8, 25, 39, 46]; item (iv) corresponds to (firmly) nonexpansive fixed point problems [36, 37] (recall that T is firmly nonexpansive if and only if $T' = 2T - \operatorname{Id}$ is nonexpansive [37, Theorem 12.1], while $\operatorname{Fix} T = \operatorname{Fix} T'$); finally, item (v) corresponds to the inequality $f_i(x) \leq 0$, which arises in convex inequality systems [21, 55] (note that, if $\dim \mathcal{H} < +\infty$, then bounded sets are relatively compact and thus the boundedness condition on the function is always satisfied).

4 Affine-convex pair

In this section, we specialize problem (1.1) to the case when $\{S_i\}_{i \in I} = \{A, B\}$, where A is a closed affine subspace of $\mathcal{H}$ and where B is a closed convex subset B of $\mathcal{H}$. The problem is thus to

$$\text{find } x \in A \cap B \tag{4.1}$$

and, as in (3.1), we assume that

$$A \cap B \neq \varnothing. \tag{4.2}$$

This setting allows for significant refinements of our general results.

4.1 Algorithm and convergence

Algorithm 4.1 Fix a starting point $x_0 \in A$ and construct the sequence $(x_n)_{n \in \mathbb{N}}$ recursively by

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(P_A R_n x_n - x_n), \tag{4.3}$$

where

$$(\forall n \in \mathbb{N}) \quad R_n \in \mathfrak{T}, \ \operatorname{Fix} R_n = B, \ \text{ and } \ \lambda_n \in \ ]0, 2K(x_n, A, R_n)[\,. \tag{4.4}$$

We observe that in this algorithm, since A is an affine subspace and $x_0 \in A$, we have

$$(\forall n \in \mathbb{N}) \quad x_n \in A. \tag{4.5}$$

Condition 4.2

(i) There exists $\varepsilon \in {]0,1[}$ such that, for every $n \in \mathbb{N}$, $\varepsilon \le \lambda_n \le (2 - \varepsilon)K(x_n, A, R_n)$.

(ii) For every orbit $(x_n)_{n \in \mathbb{N}}$ of Algorithm 4.1, every $y \in \mathcal{H}$, and every strictly increasing sequence $(p_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$, the implication

$$\left. \begin{array}{c} x_{p_n} \rightharpoonup y, \\ R_{p_n} x_{p_n} - x_{p_n} \to 0 \end{array} \right\} \quad \Rightarrow \quad y \in B \tag{4.6}$$

holds.

Let us now state a specialization of Theorem 3.3(vi) to the present setting.

Corollary 4.3 *Suppose that Condition 4.2 is satisfied. Then every orbit of Algorithm 4.1 converges weakly to a point in $A \cap B$.*

Proof. We observe that Problem (4.1) is a special case of Problem (1.1) with $I = \{1, 2\}$, $S_1 = A$, and $S_2 = B$. Next, let us verify that Algorithm 4.1 under Condition 4.2 is a special case of Algorithm 3.1 under Condition 3.2. In Algorithm 3.1, set $I' = \{1\}$, $I_n \equiv \{2\}$, $i(n) \equiv 1$, $\omega_{2,n} \equiv 1$, and $(\forall n \in \mathbb{N})$ $T_{2,n} = R_n$. Thus, using (2.23), we see that R_n is the same as in Step ⑥. Next, set

$$(\forall n \in \mathbb{N}) \qquad K_n = K(x_n, A, R_n), \quad \rho_n = \begin{cases} \varepsilon, & \text{if } \lambda_n/K_n < \varepsilon; \\ \lambda_n/K_n, & \text{otherwise,} \end{cases} \tag{4.7}$$

and, furthermore,

$$(\forall n \in \mathbb{N}) \quad \kappa_n \colon \mathcal{H} \to \mathbb{R} \colon x \mapsto \begin{cases} \lambda_n/\varepsilon, & \text{if } \lambda_n/K_n < \varepsilon \text{ and } x = x_n; \\ K_n, & \text{if } \lambda_n/K_n \ge \varepsilon \text{ and } x = x_n; \\ K(x, A, R_n), & \text{otherwise.} \end{cases} \tag{4.8}$$

Fix $n \in \mathbb{N}$. Then $0 < \kappa_n \le K(\,\cdot\,, A, R_n)$ and, at Step ⑦ of Algorithm 3.1, we have

$$T_n \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto P_A x + \kappa_n(x)(P_A R_n P_A x - P_A x). \tag{4.9}$$

On the other hand, (4.4) and (4.7) imply that $\rho_n \in {]0, 2[}$. Next, in view of (4.7), (4.8), (4.5), and (4.9), we rewrite the update rule (4.3) in Algorithm 4.1 as

$$\begin{aligned} x_{n+1} &= x_n + \lambda_n(P_A R_n x_n - x_n) \\ &= x_n + \rho_n \kappa_n(x_n)(P_A R_n x_n - x_n) \\ &= x_n + \rho_n\big(x_n + \kappa_n(x_n)(P_A R_n x_n - x_n) - x_n\big) \\ &= x_n + \rho_n\big(P_A x_n + \kappa_n(x_n)(P_A R_n P_A x_n - P_A x_n) - x_n\big) \\ &= x_n + \rho_n\big(T_n x_n - x_n\big), \end{aligned} \tag{4.10}$$

22

which is precisely Step ⑧ in Algorithm 3.1. Hence Algorithm 4.1 is an instance of Algorithm 3.1. It is readily verified that Condition 4.2(i) implies Condition 3.2(i). Moreover, Condition 3.2(ii) holds with $M_1 = M_2 = 1$, and Condition 4.2(ii) and (4.5) imply Condition 3.2(iii). The conclusion is therefore a consequence of Theorem 3.3(vi). □

Remark 4.4

(i) Condition 4.2(ii) is satisfied if $R_n \equiv R$ and $(R - \mathrm{Id})|_A$ is demiclosed at 0, that is, for every sequence $(y_n)_{n \in \mathbb{N}}$ in A and every $y \in A$, the conditions $y_n \rightharpoonup y$ and $Ry_n - y_n \to 0$ imply $y \in \mathrm{Fix}\, R$. In particular, this is true when R is nonexpansive [10, Lemma 4].

(ii) Corollary 4.3 is related to [21, Theorem 1], where A is a closed vector subspace and R_n the projector onto a closed convex superset B_n of B such that $x_n \notin B_n$. The special case when $R_n = P_B$ and $\lambda_n = K(x_n, A, P_B)$ is considered in [49, Section 1].

The next result provides information about the behavior of the sequence $(P_A R_n x_n)_{n \in \mathbb{N}}$ in Algorithm 4.1.

Proposition 4.5 *Suppose that Condition 4.2 holds and that $(x_n)_{n \in \mathbb{N}}$ is an arbitrary orbit of Algorithm 4.1 generated with $\inf_{n \in \mathbb{N}} \lambda_n \geq 1$. Then*

$$(\forall y \in A \cap B)(\forall n \in \mathbb{N}) \quad \|P_A R_{n+1} x_{n+1} - y\|^2 \leq \|P_A R_n x_n - y\|^2 \\ - (\lambda_n - 1)\big(2K(x_n, A, R_n) - 1 - \lambda_n\big)\|P_A R_n x_n - x_n\|^2. \quad (4.11)$$

Therefore, if $\inf_{n \in \mathbb{N}} 2K(x_n, A, R_n) - \lambda_n \geq 1$, then $(P_A R_n x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $A \cap B$, that is,

$$(\forall y \in A \cap B)(\forall n \in \mathbb{N}) \quad \|P_A R_{n+1} x_{n+1} - y\| \leq \|P_A R_n x_n - y\|. \quad (4.12)$$

Proof. Corollary 2.10 results in

$$(\forall n \in \mathbb{N}) \qquad P_A R_n P_A \in \mathfrak{T} \quad \text{and} \quad \mathrm{Fix}\, P_A R_n P_A = A \cap \mathrm{Fix}\, R_n = A \cap B. \quad (4.13)$$

Now let $n \in \mathbb{N}$ and $y \in A \cap B$. Then $y \in \mathrm{Fix}\, R_n$, and it follows from Proposition 2.7 that

$$\langle P_A R_n x_n - x_n \mid x_n - y \rangle \leq -\|R_n x_n - x_n\|^2. \quad (4.14)$$

Using (4.5), (4.13), and Proposition 2.4(iii), we obtain

$$\|P_A R_{n+1} x_{n+1} - y\|^2 = \|P_A R_{n+1} P_A x_{n+1} - y\|^2 \leq \|x_{n+1} - y\|^2. \quad (4.15)$$

On the other hand,

$$
\begin{aligned}
\|x_{n+1} - y\|^2 &= \|(1 - \lambda_n)(x_n - y) + \lambda_n(P_A R_n x_n - y)\|^2 \\
&= (1 - \lambda_n)\|x_n - y\|^2 + \lambda_n\|P_A R_n x_n - y\|^2 - \lambda_n(1 - \lambda_n)\|P_A R_n x_n - x_n\|^2 \\
&= \|P_A R_n x_n - y\|^2 - (\lambda_n - 1)\big(\|x_n - y\|^2 - \|P_A R_n x_n - y\|^2 - \lambda_n\|P_A R_n x_n - x_n\|^2\big) \\
&= (\lambda_n - 1)\big(\|P_A R_n x_n - x_n\|^2 + 2\langle P_A R_n x_n - x_n \mid x_n - y\rangle + \lambda_n\|P_A R_n x_n - x_n\|^2\big) \\
&\quad + \|P_A R_n x_n - y\|^2 \\
&= (\lambda_n - 1)\big((\lambda_n + 1)\|P_A R_n x_n - x_n\|^2 + 2\langle P_A R_n x_n - x_n \mid x_n - y\rangle\big) \\
&\quad + \|P_A R_n x_n - y\|^2 \\
&\leq (\lambda_n - 1)\big((\lambda_n + 1)\|P_A R_n x_n - x_n\|^2 - 2\|R_n x_n - x_n\|^2\big) \tag{4.16} \\
&\quad + \|P_A R_n x_n - y\|^2 \\
&= (\lambda_n - 1)\big(\lambda_n + 1 - 2K(x_n, A, R_n)\big)\|P_A R_n x_n - x_n\|^2 + \|P_A R_n x_n - y\|^2, \tag{4.17}
\end{aligned}
$$

where (4.16) follows from (4.14), and where (4.17) is a consequence of (4.5) and Theorem 2.8(ii). Combining this with (4.15), we therefore obtain (4.11) and, in turn, (4.12). $\square$

4.2 Computation of the extrapolation parameter

This section illustrates the fact that, under certain circumstances, the function $K(\cdot, A, R_n)$ takes on only two values, which simplifies the implementation of Algorithm 4.1. Let us introduce the closed vector subspace

$$
U = A - A, \tag{4.18}
$$

and let us assume that the operators $(R_n)_{n \in \mathbb{N}}$ in Algorithm 4.1 satisfy

$$
(\forall n \in \mathbb{N})(\forall a \in A) \quad R_n a = a + \alpha_n(a)v, \quad \text{where} \quad v \in \mathcal{H} \quad \text{and} \quad \alpha_n \colon A \to \mathbb{R}. \tag{4.19}
$$

We also recall from (4.4) that

$$
(\forall n \in \mathbb{N}) \qquad R_n \in \mathfrak{T} \quad \text{and} \quad \operatorname{Fix} R_n = B. \tag{4.20}
$$

The following result shows that, if $P_U v = 0$, then solving (4.1) amounts simply to selecting a point in A.

Proposition 4.6 *If $P_U v = 0$, then $A \subset B$.*

Proof. Let $x \in \mathcal{H}$, $n \in \mathbb{N}$, and $a \in A$. Proposition 2.6 and (4.19) yield

$$
\begin{aligned}
P_A R_n P_A x - P_A x &= P_U(R_n P_A x) + P_{U^\perp} a - \big(P_U x + P_{U^\perp} a\big) \\
&= P_U(R_n P_A x - P_A x) - P_U(x - P_A x) \\
&= P_U\big(\alpha_n(P_A x)v\big) \\
&= \alpha_n(P_A x)P_U v \tag{4.21}
\end{aligned}
$$

and thus

$$\|P_A R_n P_A x - P_A x\|^2 = \big(\alpha_n(P_A x)\big)^2 \|P_U v\|^2. \tag{4.22}$$

Therefore, the equality $P_U v = 0$ implies that $P_A R_n P_A x = P_A x$. It then follows from Proposition 2.5 that $P_A x \in \mathrm{Fix}(P_A R_n) = A \cap \mathrm{Fix}\, R_n = A \cap B$ and, in turn, that $A \subset B$. $\square$

Proposition 4.7 $(\forall n \in \mathbb{N})(\forall x \in \mathcal{H})\ K(x, A, R_n) = \begin{cases} \dfrac{\|v\|^2}{\|P_U v\|^2}, & \text{if } P_A x \notin B; \\ 1, & \text{otherwise.} \end{cases}$

Proof. Let $x \in \mathcal{H}$ and $n \in \mathbb{N}$. Then (4.19) implies that $R_n P_A x - P_A x = \alpha_n(P_A x) v$ and hence

$$\|R_n P_A x - P_A x\|^2 = \big(\alpha_n(P_A x)\big)^2 \|v\|^2. \tag{4.23}$$

Now suppose that $P_A x \notin B = \mathrm{Fix}\, R_n$. In view of (2.11), (4.22), and (4.23), we obtain

$$K(x, A, R_n) = \frac{\|R_n P_A x - P_A x\|^2}{\|P_A R_n P_A x - P_A x\|^2} = \frac{\|v\|^2}{\|P_U v\|^2}. \tag{4.24}$$

On the other hand, if $P_A x \in B = \mathrm{Fix}\, R_n$, then (2.11) yields $K(x, A, R_n) = 1$. $\square$

Note that an operator R_n assumes the form described in (4.19) if it is the projector onto a closed affine hyperplane, a closed halfspace, or a closed hyperslab (i.e., the intersection of two closed halfspaces with linearly dependent normal vectors). It is then possible to construct explicitly a point in $A \cap B$, as we now illustrate in the case of a hyperplane.

Proposition 4.8 *Let B be a hyperplane with normal vector v and suppose that E is the fully extrapolated operator in (2.12) (with $T = P_B$), that is,*

$$E \colon \mathcal{H} \to A \colon x \mapsto P_A x + K(x, A, P_B)\big(P_A P_B P_A x - P_A x\big). \tag{4.25}$$

Then, for every $x \in \mathcal{H}$, the point Ex lies in $A \cap B$ and hence solves (4.1).

Proof. There exists $\beta \in \mathbb{R}$ such that $B = \big\{ x \in \mathcal{H} \mid \langle x \mid v \rangle = \beta \big\}$ and therefore [30, Theorem 6.17] implies that

$$(\forall x \in \mathcal{H})\ P_B x = x + \frac{\beta - \langle x \mid v \rangle}{\|v\|^2} v. \tag{4.26}$$

Now let $x \in \mathcal{H}$. Since range $E \subset A$, it is enough to show that $Ex \in B$. It follows from Proposition 4.7 that

$$Ex = \begin{cases} P_A x + \dfrac{\|v\|^2}{\|P_U v\|^2}\big(P_A P_B P_A x - P_A x\big), & \text{if } P_A x \notin B; \\ P_A P_B P_A x, & \text{otherwise.} \end{cases} \tag{4.27}$$

Therefore, if $P_A x \in B$, then $Ex = P_A x \in A \cap B$. Now assume that $P_A x \notin B$. Then, since $\|P_U v\|^2 = \langle P_U v \mid v \rangle$, we have

$$\langle Ex \mid v \rangle = \langle P_A x \mid v \rangle + \frac{\|v\|^2}{\langle P_U v \mid v \rangle}\langle P_A P_B P_A x - P_A x \mid v \rangle. \tag{4.28}$$

25

On the other hand, (4.21) and (4.26) yield

$$P_A P_B P_A x - P_A x = \frac{\beta - \langle P_A x \mid v \rangle}{\|v\|^2} P_U v, \tag{4.29}$$

and thus

$$\langle P_A P_B P_A x - P_A x \mid v \rangle = \big(\beta - \langle P_A x \mid v \rangle\big)\langle P_U v \mid v \rangle / \|v\|^2. \tag{4.30}$$

Combining (4.28) and (4.30), we obtain $\langle Ex \mid v \rangle = \beta$ and therefore $Ex \in B$. $\square$

Example 4.9 Let $\mathcal{H} = A = \mathbb{R}^2$, $v = (0,1)$, $R_n \equiv P_B$, where $B = \{x \in \mathcal{H} \mid \langle x \mid v \rangle = 0\} = \mathbb{R} \times \{0\}$. Then $A \cap B = B$ and

$$(\forall x \in \mathcal{H}) \quad K(x, A, P_B) = 1. \tag{4.31}$$

Suppose also that $\delta \colon \mathcal{H} \to \mathbb{R}$ and set

$$E_\delta \colon \mathcal{H} \to A \colon x \mapsto P_A x + \big(K(x, A, P_B) + \delta(x)\big)\big(P_A P_B P_A x - P_A x\big) = P_B x - \delta(x) P_{B^\perp} x, \tag{4.32}$$

so that $E_0 = P_B$ corresponds to the fully extrapolated operator E of Proposition 4.8. Then:

(i) If there exists a point $x \in \mathcal{H} \smallsetminus B$ such that $\delta(x) \neq 0$, then $E_\delta x \notin B$. Consequently, if range $E_\delta \subset B = A \cap B$, then $E_\delta = E_0$.

(ii) If there exists a point $x \in \mathcal{H} \smallsetminus B$ such that $\delta(x) > 0$, then $E_\delta \notin \mathfrak{T}$.

Proof. Note that (4.31) follows from Proposition 4.7 since $U = A - A = \mathcal{H}$. We readily obtain (4.32) which, in turn, implies (i). (ii): Suppose that $x \in \mathcal{H} \smallsetminus B$ satisfies $\delta(x) > 0$. Then (4.32) yields $x - E_\delta x = (1 + \delta(x)) P_{B^\perp} x$ and hence $\|x - E_\delta x\| = (1 + \delta(x))\|P_{B^\perp} x\| > \|P_{B^\perp} x\| = \|x - P_B x\|$. Thus

$$\|x - P_B x\| < \|x - E_\delta x\|. \tag{4.33}$$

Using (4.32), we also observe that $P_B x \in B \subset \text{Fix } E_\delta$. Let us now argue by contradiction by assuming that $E_\delta \in \mathfrak{T}$. Then Proposition 2.3 and the Cauchy-Schwarz inequality yield

$$\|E_\delta x - x\|^2 \leq \langle P_B x - x \mid E_\delta x - x \rangle \leq \|P_B x - x\|\|E_\delta x - x\|. \tag{4.34}$$

Hence $\|E_\delta x - x\| \leq \|P_B x - x\|$, which contradicts (4.33). Therefore $E_\delta \notin \mathfrak{T}$. $\square$

Remark 4.10 Example 4.9 provides limiting cases in the following two senses.

(i) Example 4.9(i) illustrates that, with respect to the parameter range, only the fully extrapolated operator of Proposition 4.8 provides a point in $A \cap B$.

(ii) Example 4.9(ii) shows that in (2.12) it is in general impossible to extend the range of $\kappa(x)$ beyond $K(x, A, P_B)$ without destroying the $\mathfrak{T}$-class property.

4.3 Extrapolated alternating projection method: numerical experiments

We shall henceforth focus on the following projection setting.

Corollary 4.11 (extrapolated alternating projection method (EAPM)) *Let* $(x_n)_{n \in \mathbb{N}}$ *be generated by the recursion* $x_0 \in A$ *and*

$$(\forall n \in \mathbb{N}) \ \ x_{n+1} = x_n + \rho K(x_n, A, P_B)(P_A P_B x_n - x_n), \ \ where \ \rho \in \,]0, 2[\,. \tag{4.35}$$

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

Proof. This is a special case of Corollary 4.3 with $R_n \equiv P_B$. Indeed, (4.35) is a specialization of Algorithm 4.1 for which Condition 4.2(i) is trivially satisfied. On the other hand, Condition 4.2(ii) follows from Proposition 3.7(i). □

We shall compare EAPM to the following three methods.

- **Projections Onto Convex Sets (POCS) method:** Setting $R_n \equiv P_B$ and $\lambda_n \equiv 1$ in Corollary 4.3, we obtain the weak convergence to a point in $A \cap B$ of the POCS iterations [9, 18], with starting point $x_0 \in A$ and

$$(\forall n \in \mathbb{N}) \ \ x_{n+1} = P_A P_B x_n. \tag{4.36}$$

- **Reflection-projection method (RPM):** This method, which generates a sequence via $x_0 \in A$ and
$$(\forall n \in \mathbb{N}) \ \ x_{n+1} = P_A(2 P_B x_n - x_n), \tag{4.37}$$
was shown in [7] to converge weakly to a point in $A \cap B$ provided that B is a closed convex cone that is obtuse, i.e., B contains its dual cone ([7] treats also the case when $A \cap B = \varnothing$). Since P_A is an affine operator in our present setting, the update rule can be rewritten as $x_{n+1} = x_n + 2(P_A P_B x_n - x_n)$, which formally corresponds to the case when $\rho K(x_n, A, P_B) \equiv 2$ in (4.35). This case is not necessarily covered by the results in the present paper. Indeed, if $\big(K(x_n, A, P_B)\big)_{n \in \mathbb{N}}$ is not bounded away from 1, then we cannot have $2 \leq \inf_{n \in \mathbb{N}} \rho K(x_n, A, P_B)$ for any $\rho \in \,]0, 2[$.

- **Extrapolated parallel projection method (EPPM):** Pierra's extrapolated parallel projection method [49], specialized to the present setup, is $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \ \ x_{n+1} = x_n + \frac{\|P_A x_n - x_n\|^2 + \|P_B x_n - x_n\|^2}{\|P_A x_n + P_B x_n - 2 x_n\|^2} (P_A x_n + P_B x_n - 2 x_n), \tag{4.38}$$

which corresponds to Algorithm 3.1 with $I = \{1, 2, 3\}$, $S_1 = A$, $S_2 = B$, $S_3 = \mathcal{H}$, $I' = \{3\}$, $I_n \equiv \{1, 2\}$, $T_{1,n} \equiv P_A$, $T_{2,n} \equiv P_B$, $\omega_{1,n} \equiv \omega_{2,n} \equiv 1/2$, $\kappa_n(\cdot) \equiv 1$, and $\rho_n \equiv 1$.

We shall consider the experimental framework of [7, Section 5], where RPM compared favorably with POCS. In

$$\mathcal{H} = \mathbb{R}^{450}, \tag{4.39}$$

we generate random instances of A, where

$$A \text{ is a closed affine subspace of } \mathcal{H} \text{ such that } \dim(A - A) = 300. \tag{4.40}$$

We assume further that

$$B = [0, +\infty[^{450}, \tag{4.41}$$

and that $A \cap B \neq \varnothing$. As in [21], the performance of the algorithms is measured by the decibel (dB) values of the relative proximity function evaluated at the nth iterate x_n, i.e., by

$$10 \log_{10} \left(\frac{\|P_A x_n - x_n\|^2 + \|P_B x_n - x_n\|^2}{\|P_A x_0 - x_0\|^2 + \|P_B x_0 - x_0\|^2} \right). \tag{4.42}$$

In the first experiment, we create 5 random instances of A such that $A \cap B \neq \varnothing$. From a randomly chosen starting point $x_0 \in A$, we generated orbits $(x_n)_{n \in \mathbb{N}}$ of the four algorithms EAPM (with $\rho = 1$ in (4.35)), POCS, RPM, and EPPM. The proximity function values (4.42), averaged over all instances, are plotted in Figure 1. The experiment shows that EAPM outperforms the other
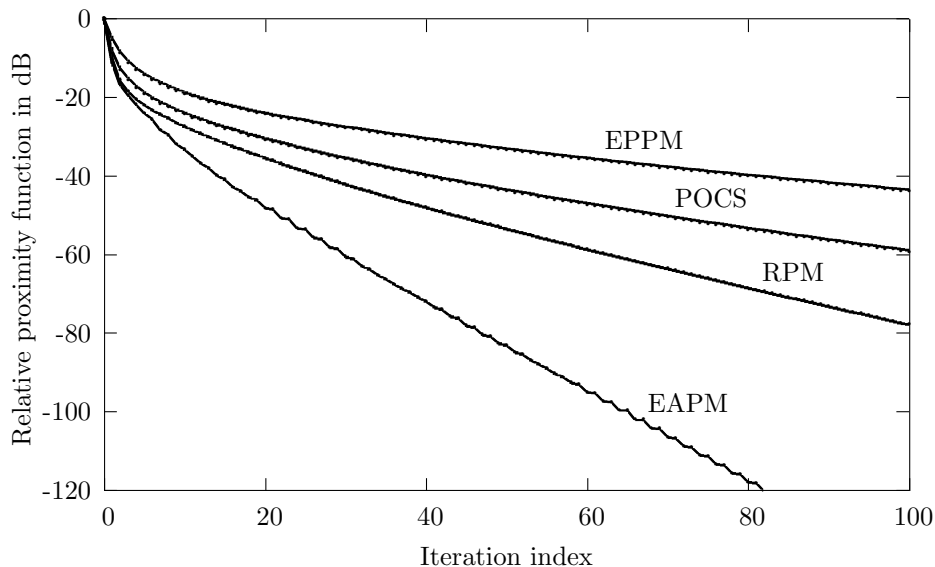
Figure 1: Values of the relative proximity function.

methods. The better performance of EAPM may be attributed to the large step-sizes allowed by the algorithm; Figure 2 shows that $K(x_n, A, P_B)$ can indeed be substantially larger than 2. We now explore the effect of the relaxed step-size, via the relaxation parameter ρ in (4.35), on the performance of EAPM. Figures 3 and 4 depict the behavior of the relative proximity measure as a function of ρ in the case of under- and over-relaxation, respectively.
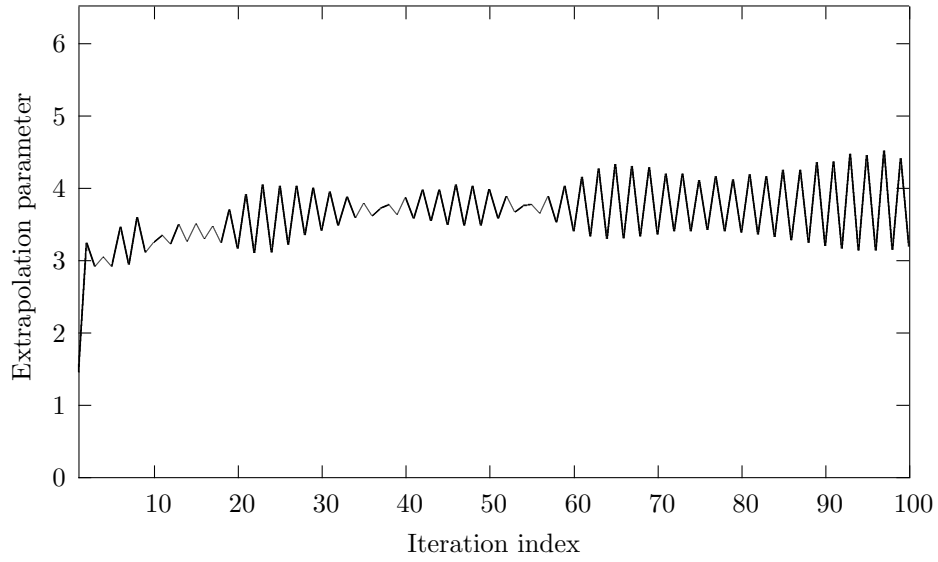
28

Figure 2: Values of the extrapolation parameter $K(x_n, A, P_B)$ in the experiment of Figure 1.
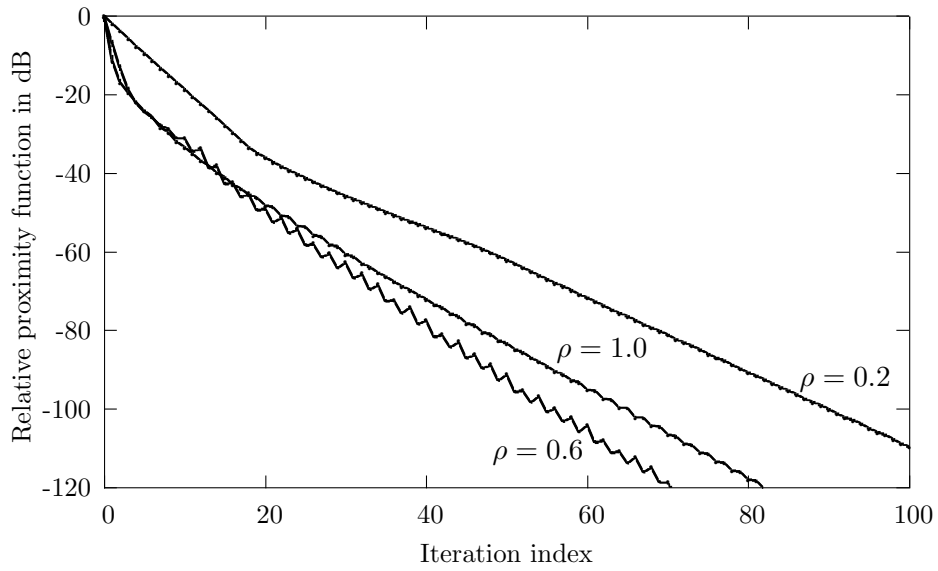

Figure 3: Values of the relative proximity function with relaxation parameter $\rho \leq 1$.
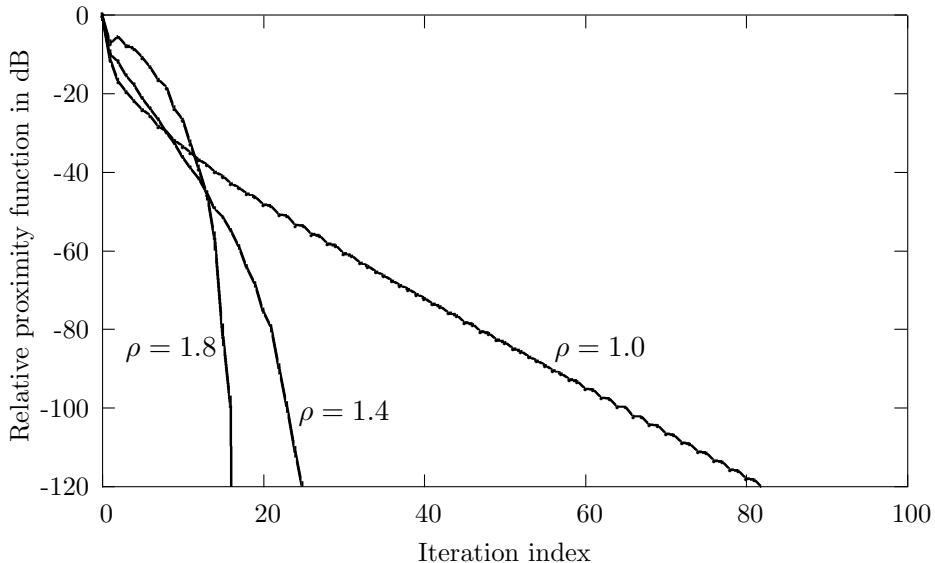
Figure 4: Values of the relative proximity function with relaxation parameter $\rho \geq 1$.

It is well known that periodic "centering" of the iterates can significantly improve the convergence properties of EPPM and related methods [21, 48, 49]. This technique amounts to periodically halving the relaxation parameter. In the present setting, we replace the relaxation parameter $\lambda_n = \rho K(x_n, A, R_n)$ in (4.35) by (convergence is still guaranteed by Corollary 4.3)

$$\lambda_n = \begin{cases} \rho K(x_n, A, R_n)/2, & \text{if } n = 2 \text{ (modulo 3)}; \\ \rho K(x_n, A, R_n), & \text{otherwise.} \end{cases} \tag{4.43}$$

We repeat the experiment with this strategy and report the numerical results in Figures 5–8, which correspond to Figures 1–4, respectively. Let us conclude by noting that EAPM also benefits from centering to the extent that it still outperforms the other methods.
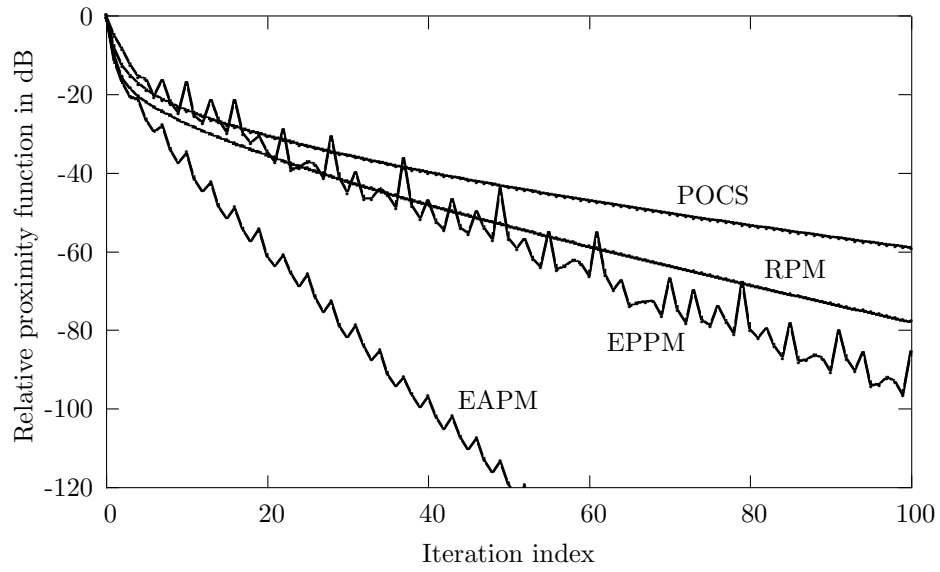
Figure 5: Values of the relative proximity function when EPPM and EAPM are centered.
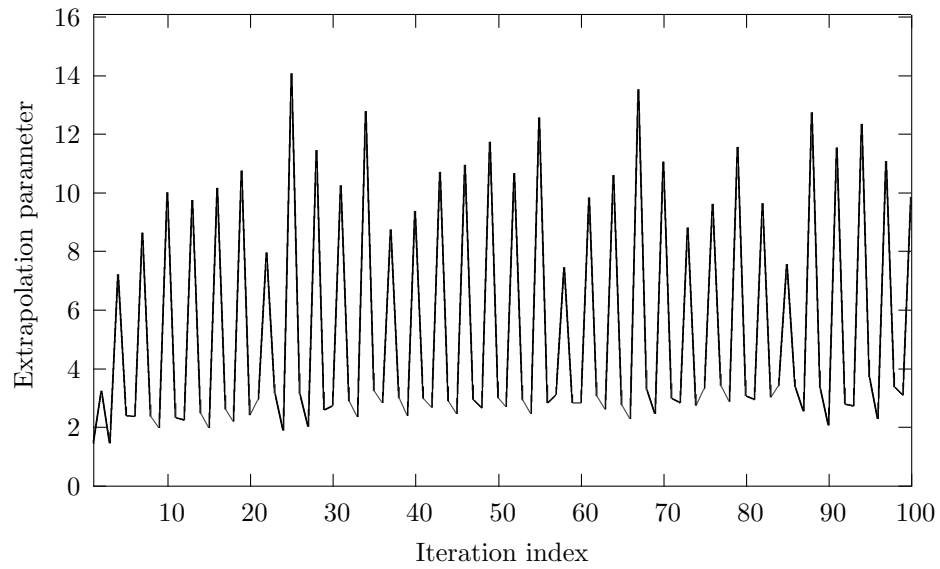

Figure 6: Values of the extrapolation parameter $K(x_n, A, P_B)$ in the experiment of Figure 5.
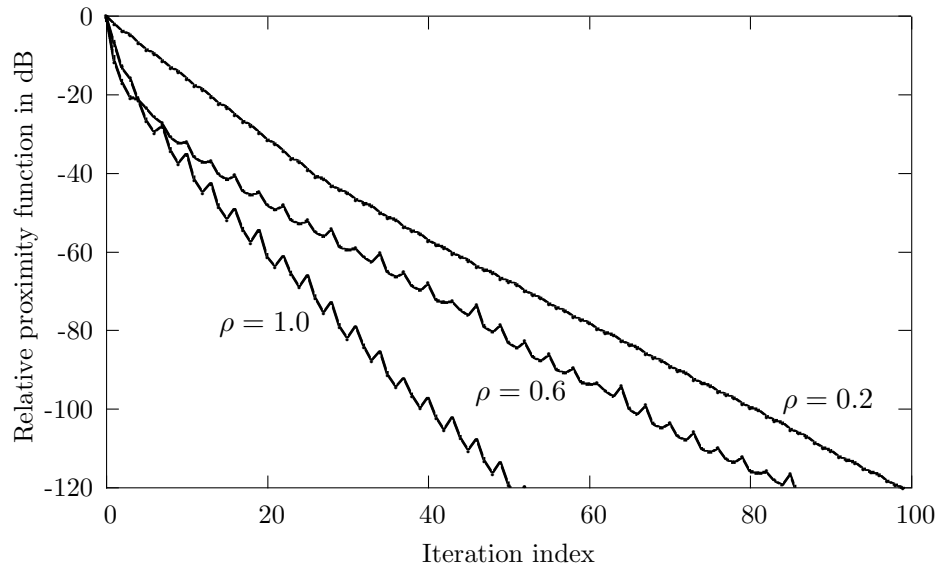
Figure 7: Values of the relative proximity function with relaxation parameter $\rho \leq 1$, when EAPM and EPPM are centered.
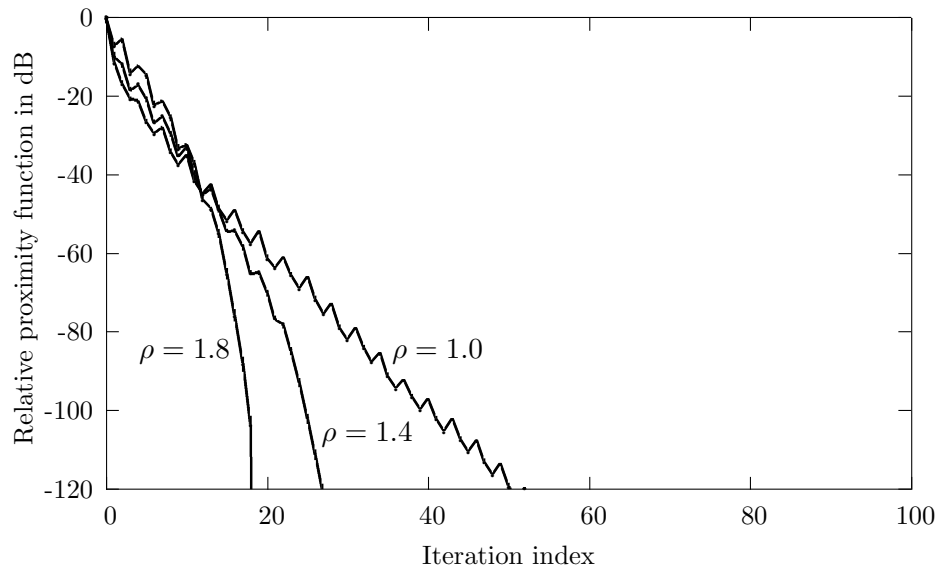

Figure 8: Values of the relative proximity function with relaxation parameter $\rho \geq 1$, when EAPM and EPPM are centered.

Acknowledgment

References

[1] A. AUSLENDER, *Méthodes Numériques pour la Résolution des Problèmes d'Optimisation avec Contraintes*, Thèse, Faculté des Sciences, Grenoble, France, 1969.

[2] H. H. BAUSCHKE, *Projection Algorithms and Monotone Operators*, Ph.D. Thesis, Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada, August 1996. Available at `http://www.cecm.sfu.ca/preprints/1996pp.html`

[3] H. H. BAUSCHKE AND J. M. BORWEIN, *On projection algorithms for solving convex feasibility problems*, SIAM Rev., 38 (1996), pp. 367–426.

[4] H. H. BAUSCHKE AND P. L. COMBETTES, *A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces*, Math. Oper. Res., 26 (2001), pp. 248–264.

[5] H. H. BAUSCHKE, F. DEUTSCH, H. HUNDAL, AND S.-H. PARK, *Fejér monotonicity and weak convergence of an accelerated method of projections,* M. Théra, ed., in Constructive, Experimental, and Nonlinear Analysis, CMS Conference Proceedings 27, 2000, pp. 1–6.

[6] H. H. BAUSCHKE, F. DEUTSCH, H. HUNDAL, AND S.-H. PARK, *Accelerating the convergence of the method of alternating projections*, Trans. Amer. Math. Soc., 355 (2003), pp. 3433–3461.

[7] H. H. BAUSCHKE AND S. G. KRUK, *Reflection-projection method for convex feasibility problems with an obtuse cone*, J. Optim. Theory Appl., 120 (2004), pp. 503–531.

[8] E. BLUM AND W. OETTLI, *From optimization and variational inequalities to equilibrium problems,* Math. Student, 63 (1994), 123–145.

[9] L. M. BREGMAN, *The method of successive projection for finding a common point of convex sets*, Soviet Math. Dokl., 6 (1965), pp. 688–692.

[10] F. E. BROWDER, *Convergence theorems for sequences of nonlinear operators in Banach spaces*, Math. Z., 100 (1967), pp. 201–225.

[11] D. BUTNARIU AND Y. CENSOR, *On the behavior of a block-iterative projection method for solving convex feasibility problems*, Int. J. Computer Math., 34 (1990), pp. 79–94.

[12] D. BUTNARIU, Y. CENSOR, AND S. REICH, Eds., *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications,* Elsevier, New York, 2001.

[13] A. CAUCHY, *Méthode générale pour la résolution des systèmes d'équations simultanées,* C. R. Acad. Sci. Paris, 25 (1847), pp. 536–538.

[14] Y. CENSOR *Iterative methods for the convex feasibility problem*, Ann. Discrete Math., 20 (1984), pp. 83–91.

[15] Y. CENSOR AND A. LENT, *Cyclic subgradient projections*, Math. Programming, 24 (1982), pp. 233–235, 1982.

[16] Y. Censor and S. A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications,* Oxford University Press, New York, 1997.

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

[18] P. L. Combettes, *The foundations of set theoretic estimation,* Proc. IEEE, 81 (1993), pp. 182–208.

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

[20] P. L. Combettes, *The convex feasibility problem in image recovery,* in Advances in Imaging and Electron Physics, P. Hawkes, ed., Vol. 95, Academic Press, New York, 1996, pp. 155–270.

[21] P. L. Combettes, *Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections,* IEEE Trans. Image Process., 6 (1997), pp. 493–506.

[22] P. L. Combettes, *Hilbertian convex feasibility problem: Convergence of projection methods,* Appl. Math. Optim., 35 (1997), pp. 311–330.

[23] P. L. Combettes, *Quasi-Fejérian analysis of some optimization algorithms,* in Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, D. Butnariu, Y. Censor, and S. Reich, Eds., Elsevier, New York, 2001, pp. 115–152.

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

[25] P. L. Combettes and S. A. Hirstoaga, *Equilibrium programming in Hilbert spaces,* J. Nonlinear Convex Anal., 6 (2005), pp. 117–136.

[26] G. Crombez, *Image recovery by convex combinations of projections,* J. Math. Anal. Appl., 155 (1991), pp. 413–419.

[27] G. Crombez, *Viewing parallel projection methods as sequential ones in convex feasibility problems,* Trans. Amer. Math. Soc., 347 (1995), pp. 2575–2583.

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

[29] F. Deutsch, *The method of alternating orthogonal projections,* in Approximation Theory, Spline Functions and Applications, S. P. Singh, ed., Kluwer, The Netherlands, 1992, pp. 105–121.

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

[31] L. T. Dos Santos, *A parallel subgradient projections method for the convex feasibility problem,* J. Comput. Appl. Math., 18 (1987), pp. 307–320.

[32] I. I. Eremin, *Generalization of the relaxation method of Motzkin-Agmon,* Uspekhi Mat. Nauk, 20 (1965), pp. 183–187.

[33] S. D. Flåm and J. Zowe, *Relaxed outer projections, weighted averages, and convex feasibility,* BIT, 30 (1990), pp. 289–300.

[34] U. García-Palomares, *Parallel-projected aggregation methods for solving the convex feasibility problem,* SIAM J. Optim., 3 (1993), pp. 882–900.

[35] U. M. García-Palomares and F. J. González-Castaño, *Incomplete projection algorithms for solving the convex feasibility problem,* Numer. Algorithms, 18 (1998), pp. 177–193.

[36] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings.* Marcel Dekker, New York, 1984.

[37] K. Goebel and W. A. Kirk, *Topics in Metric Fixed Point Theory,* Cambridge University Press, Cambridge, 1990.

[38] F. J. González-Castaño, U. M. García-Palomares, J. L. Alba-Castro, and J. M. Pousada-Carballo, *Fast image recovery using dynamic load balancing in parallel architectures, by means of incomplete projections*, IEEE Trans. Image Process., 10 (2001), pp. 493–499.

[39] A. Göpfert, H. Riahi, C. Tammer, and C. Zălinescu, *Variational Methods in Partially Ordered Spaces,* Springer-Verlag, New York, 2003.

[40] L. G. Gubin, B. T. Polyak, and E. V. Raik, *The method of projections for finding the common point of convex sets*, USSR Comput. Math. Math. Phys., 7 (1967), pp. 1–24.

[41] S. Kaczmarz, *Angenäherte Auflösung von Systemen linearer Gleichungen,* Bull. Acad. Sci. Pologne, A35 (1937), pp. 355–357.

[42] K. C. Kiwiel, *Block-iterative surrogate projection methods for convex feasibility problems*, Linear Algebra Appl., 215 (1995), pp. 225–259.

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

[44] N. Lehdili and B. Lemaire, *The barycentric proximal method*, Comm. Appl. Nonlinear Anal., 6 (1999), pp. 29–47.

[45] Y. I. Merzlyakov, *On a relaxation method of solving systems of linear inequalities*, USSR Comput. Math. Math. Phys., 2 (1963), pp. 504–510.

[46] W. Oettli, *A remark on vector-valued equilibria and generalized monotonicity,* Acta Math. Vietnamica, 22 (1997), pp. 215–221.

[47] N. Ottavy, *Strong convergence of projection-like methods in Hilbert spaces*, J. Optim. Theory Appl., 56 (1988), pp. 433–461.

[48] G. Pierra, *Éclatement de contraintes en parallèle pour la minimisation d'une forme quadratique*, in Lecture Notes in Computer Science, Vol. 41, Springer-Verlag, New York, 1976, pp. 200–218.

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

[50] B. T. Polyak, *Minimization of unsmooth functionals,* USSR Comput. Math. Math. Phys., 9 (1969), pp. 14–29.

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

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

[53] P. Tseng, *On the convergence of products of firmly nonexpansive mappings*, SIAM J. Optim., 2 (1992), pp. 425–434.

[54] I. Yamada, K. Slavakis, and K. Yamada, *An efficient robust adaptive filtering algorithm based on parallel subgradient projection techniques*, IEEE Trans. Signal Process., 50 (2002), pp. 1091–1101.

[55] I. Yamada and N. Ogura, *Adaptive projected subgradient method for asymptotic minimization of sequence of nonnegative convex functions,* Numer. Funct. Anal. Optim., 25 (2004), pp. 593–617.

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