Extended forward–backward algorithm

CrossMark

Marc Lassonde *, Ludovic Nagesseur

LAMIA, Université des Antilles et de la Guyane, 97159 Pointe à Pitre, France

A R T I C L E   I N F O

A B S T R A C T

We propose an extended forward–backward algorithm for approximating a zero of a maximal monotone operator which can be split as the extended sum of two maximal monotone operators. We establish the weak convergence in average of the sequence generated by the algorithm under assumptions similar to those used in classical forward–backward algorithms. This provides as a special case an algorithm for solving convex constrained minimization problems without qualification condition.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Let $\mathcal{H}$ be a real Hilbert space and let $T : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator. A fundamental problem is that of designing efficient algorithms for constructing a zero of T, i.e. a solution of the inclusion

$$0 \in Tx. \tag{1}$$

Proximal point algorithms have been introduced by Martinet [8,9] to solve (1) in specific situations such as when T is the subdifferential ∂f of a proper convex lower semicontinuous function f: in this case the inclusion (1) is just the minimization of f over $\mathcal{H}$. The general situation of an arbitrary maximal monotone operator T was then considered by Rockafellar [16] and Brezis–Lions [1]. See, e.g., Peypouquet–Sorin [13] for a survey on the topic.

When the maximal monotone operator T admits a decomposition of the form $T = A + B$ with A and B maximal monotone on $\mathcal{H}$, it is more efficient to use splitting methods which combine a step for A and a step for B: see, e.g., Combettes [4] for a unified presentation of these methods. A prominent example is the forward–backward algorithm considered by Passty [12], which consists of a forward step on B and a backward (proximal) step on A. A special case of this method is the projected subgradient algorithm aimed at solving constrained minimization problems.

There are important examples, however, where the problem to be solved involves two maximal monotone operators A and B whose sum $A + B$ is not maximal monotone so that the usual splitting methods do not apply. A simple instance of such a situation is the problem

$$\text{Find } x \in C \text{ such that } f(x) = \min f(C) \tag{2}$$

of minimizing a proper lower semicontinuous convex function $f : \mathcal{H} \rightarrow]-\infty, +\infty]$ over a nonempty closed convex subset C of $\mathcal{H}$. This problem can be rewritten as

$$\text{Find } x \in \mathcal{H} \text{ such that } 0 \in \partial(f + \delta_C)(x), \tag{3}$$

---

* Corresponding author.
  E-mail addresses: marc.lassonde@univ-ag.fr, marc.lassonde@gmail.com (M. Lassonde), ludovic.nagesseur@univ-ag.fr (L. Nagesseur).

where δ_C is the indicator of the set C. When a qualification condition is satisfied, e.g. $0 \in \mathrm{cor}\,(\mathrm{dom}\,f - C)$, then $\partial(f + \delta_C) = \partial f + \partial\delta_C$, so (3) is equivalent to

$$\text{Find } x \in \mathcal{H} \text{ such that } 0 \in (\partial f + \partial\delta_C)(x), \tag{4}$$

where $\partial f + \partial\delta_C$ is maximal monotone: in this case, standard splitting methods apply. But without a qualification condition, such a decomposition of $\partial(f + \delta_C)$ as the sum $\partial f + \partial\delta_C$ is not admissible. However, it is always true that $\partial(f + \delta_C)$ can be split as the *extended sum* $\partial f +_e \partial\delta_C$ of the two maximal monotone operators ∂f and $\partial\delta_C$. Hence, (3) is always equivalent to

$$\text{Find } x \in \mathcal{H} \text{ such that } 0 \in (\partial f +_e \partial\delta_C)(x), \tag{5}$$

where $\partial f +_e \partial\delta_C$ is maximal monotone.

The aim of this paper is to present a modified forward–backward splitting algorithm to solve the inclusion (1) for a maximal monotone operator T which can be written as the extended sum $A +_e B$ of two maximal monotone operators. This includes (2) as a special case. We establish the weak convergence in average of the sequence generated by the algorithm to a zero of T under hypotheses on the step sizes similar to those used in the classical splitting algorithms in Bruck [2], Lions [7] or Passty [12] where the maximal monotone operator T is equal to the usual sum $A + B$ of two maximal monotone operators.

Previous extensions of splitting algorithms to treat the case of the extended sum of two maximal monotone operators rely heavily on the assumption that the ε-enlargement of one of the operators is bounded on the solution set as in Moudafi–Théra [11]. Such a strong assumption is avoided here.

2. Tools

This section presents the main concepts and results we need in the following. Most of these tools are valid in arbitrary Banach spaces, but we restrict ourselves to the Hilbert space setting which suffices for our purpose.

Let $\mathcal{H}$ be a real Hilbert space with inner product $\langle .,. \rangle$. A set-valued operator $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is identified with its graph $T \subset \mathcal{H} \times \mathcal{H}$, and its domain is the set $D(T) = \{x \in \mathcal{H} : Tx \neq \emptyset\}$. A set-valued operator $T : \mathcal{H} \rightrightarrows \mathcal{H}$, or a subset $T \subset \mathcal{H} \times \mathcal{H}$, is said to be *monotone* provided

$$\langle v - u, y - x \rangle \geq 0, \quad \forall (x, u), (y, v) \in T,$$

and *maximal monotone* provided it is monotone and maximal (under set inclusion) in the family of all monotone sets contained in $\mathcal{H} \times \mathcal{H}$.

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a convex lower semicontinuous function which is proper, that is, its effective domain $\mathrm{dom}\,f = \{x \in \mathcal{H} : f(x) < +\infty\}$ is nonempty. For $\varepsilon \geq 0$, the ε-*subdifferential* $\partial_\varepsilon f$ of f at $x \in \mathcal{H}$ is the set:

$$\partial_\varepsilon f(x) := \{u \in \mathcal{H} : f(x) + \langle u, y - x \rangle \leq f(y) + \varepsilon, \ \forall y \in \mathcal{H}\}.$$

The case $\varepsilon = 0$ gives the *subdifferential* of f at x, denoted by $\partial f(x)$. For every $\varepsilon > 0$ and $x \in \mathrm{dom}\,f$, the set $\partial_\varepsilon f(x)$ is nonempty. On the contrary, the set $\partial f(x)$ may be empty at some points $x \in \mathrm{dom}\,f$. The subdifferential ∂f is a basic example of a maximal monotone operator on $\mathcal{H}$ (see Moreau [10]).

The usual (Minkowski) sum $\partial f + \partial g$ of the subdifferentials of two proper convex lower semicontinuous functions $f, g : \mathcal{H} \to]-\infty, +\infty]$ is a monotone operator which is not maximal monotone in general. Said differently, $\partial f + \partial g$ is a monotone subset of $\partial(f + g)$ and the inclusion is proper in general without additional assumptions. However, an exact formula for the subdifferential of the sum can be given in terms of the sum of the ε-subdifferentials of each function, namely

$$\partial(f + g)(x) = \bigcap_{\varepsilon > 0} \overline{\partial_\varepsilon f(x) + \partial_\varepsilon g(x)}, \quad \forall x \in \mathcal{H}. \tag{6}$$

This formula is due to Hiriart-Urruty and Phelps [6]. It shows that the operator on the right hand side is maximal monotone.

Let now $T : \mathcal{H} \rightrightarrows \mathcal{H}$ be an arbitrary monotone operator. Given $\varepsilon \geq 0$, the ε-*enlargement* of T is the operator $T^\varepsilon : \mathcal{H} \rightrightarrows \mathcal{H}$ defined as follows:

$$T^\varepsilon x := \{u \in \mathcal{H} : \langle v - u, y - x \rangle \geq -\varepsilon, \ \forall (y, v) \in T\}.$$

For a proper convex lower semicontinuous function f, one has $\partial_\varepsilon f(x) \subset (\partial f)^\varepsilon(x)$. For more details on this concept, see, e.g., [3,5,14,15] and the references therein. The following useful inequality is established in [3]

Proposition 2.1. *Let* $T : \mathcal{H} \rightrightarrows \mathcal{H}$ *be maximal monotone and let* $\varepsilon_1 \geq 0, \ \varepsilon_2 \geq 0$. *Then*

$$\langle v - u, y - x \rangle \geq -(\sqrt{\varepsilon_1} + \sqrt{\varepsilon_2})^2, \quad \forall (x, u) \in T^{\varepsilon_1}, \ (y, v) \in T^{\varepsilon_2}.$$

As for the subdifferentials, the usual sum $(A + B)(x) = Ax + Bx$, $x \in \mathcal{H}$, of two maximal monotone operators A, B defines a monotone operator which is not maximal monotone in general. Thus, motivated by the above formula (6), a concept of *extended sum*, based on the notion of enlargement of a monotone operator, was proposed in [14,15], namely:

$$(A +_e B)(x) := \bigcap_{\varepsilon > 0} \overline{A^\varepsilon x + B^\varepsilon x}, \quad \forall x \in \mathcal{H}. \tag{7}$$

The following result asserts that the extended sum $A +_e B$ is indeed a monotone extension of the usual sum $A + B$ (see [5]):

Proposition 2.2. *If $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ are two maximal monotone operators, the extended sum $A +_e B$ is a monotone operator containing the usual sum $A + B$.*

It follows from Proposition 2.2 and formula (6) that for any proper lower semicontinuous convex functions $f, g : \mathcal{H} \rightarrow]-\infty, +\infty]$ we have

$$\partial(f + g) = \partial f +_e \partial g. \tag{8}$$

This formula is due to Revalski and Théra [15] (see also [5]). It shows that the extended sum $\partial f +_e \partial g$ is a maximal monotone operator.

3. Extended forward–backward algorithm

In what follows, $\mathcal{H}$ is a real Hilbert space with inner product $\langle ., . \rangle$ and unit ball $B_{\mathcal{H}}$, and $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximal monotone operator which can be split as the extended sum of two maximal monotone operators A and B:

$$Tx := \bigcap_{\varepsilon > 0} \overline{A^\varepsilon x + B^\varepsilon x}, \quad \text{for all } x \in \mathcal{H}. \tag{9}$$

Our aim is to take advantage of this splitting to design a forward–backward algorithm for constructing a zero of T, i.e. a solution $x \in \mathcal{H}$ of the inclusion (1). The set of solutions of (1) is supposed to be nonempty and is denoted by $\mathcal{S}$.

Let (λ_n) be a sequence of positive real numbers defining the step sizes of a discretization scheme, and let $\sigma_n := \sum_{k=1}^{n} \lambda_k$ denote the step length at stage n. As usual, σ_n is supposed to grow to $+\infty$ as $n \to \infty$. Given a sequence (x_n) in $\mathcal{H}$ together with a sequence of step sizes (λ_n), we define the sequence $(\bar{x}_n)$ of weighted averages by:

$$\bar{x}_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \lambda_k x_k. \tag{10}$$

Given a sequence of step sizes (λ_n) and a sequence of extension parameters (ε_n), we study the *extended forward–backward* algorithm defined by the iteration:

$$\begin{cases} x_0 \in D(B^{\varepsilon_0}) \\ x_{n+1} = (I + \lambda_n A)^{-1}(x_n - \lambda_n u_n) \quad \text{with } u_n \in B^{\varepsilon_n} x_n, \ \forall n \in \mathbb{N}. \end{cases} \tag{11}$$

Our main result is as follows:

Theorem 3.1. *Assume $D(A) \subset \cap_{\varepsilon > 0} D(B^\varepsilon)$ and let (x_n) be the sequence generated by (11), with (λ_n) and (ε_n) any sequences of positive real numbers verifying*

$$\sum \lambda_n = \infty, \qquad \sum \lambda_n^{4/3} < \infty, \qquad \varepsilon_n = \lambda_n^{1/3}.$$

Assume further that:

(H1) *The sequence $(\varepsilon_n u_n)$ is bounded,*
(H2) *For every $x \in \mathcal{S}$ there is $M > 0$ such that $0 \in \bigcap_n \overline{A^{\varepsilon_n} x + B^{\varepsilon_n} x \cap (M/\varepsilon_n) B_{\mathcal{H}}}$.*

Then, the sequence $(\bar{x}_n)$ of weighted averages given by (10) weakly converges to a point in $\mathcal{S}$.

Proof. The argument relies on Passty's variant of Opial's lemma (see [12,13]):

Opial–Passty's Lemma. *Let (λ_n) be a sequence of positive numbers such that $\sum \lambda_n = \infty$, (x_n) a sequence in a Hilbert space $\mathcal{H}$ and $(\bar{x}_n)$ its associated sequence of weighted averages given by (10). Let $\mathcal{S} \subset \mathcal{H}$ be a nonempty closed subset. Assume*

 (i) *Every weak sequential cluster point of $(\bar{x}_n)$ belongs to $\mathcal{S}$,*
(ii) *$\lim \|x_n - x\|$ exists for every $x \in \mathcal{S}$.*

Then, the sequence $(\bar{x}_n)$ converges to a point in $\mathcal{S}$.

The theorem is proved by showing that (i) and (ii) hold.

(i) Let $(x, y) \in T$, that is, $y \in \overline{A^\varepsilon x + B^\varepsilon x}$ for every $\varepsilon > 0$. Fix $\varepsilon > 0$, and take $y_\varepsilon = y_{1,\varepsilon} + y_{2,\varepsilon}$ with $y_{1,\varepsilon} \in A^\varepsilon x$, $y_{2,\varepsilon} \in B^\varepsilon x$ and $\|y - y_\varepsilon\| \leq \varepsilon$. As $u_n \in B^{\varepsilon_n} x_n$ and $y_{2,\varepsilon} \in B^\varepsilon x$, Proposition 2.1 yields:

$$\langle u_n - y_{2,\varepsilon}, x_n - x \rangle \geq -(\sqrt{\varepsilon_n} + \sqrt{\varepsilon})^2 \geq -2(\varepsilon_n + \varepsilon). \tag{12}$$

Since $x_n - \lambda_n u_n \in x_{n+1} + \lambda_n A x_{n+1}$ by (11) and $y_{1,\varepsilon} \in A^\varepsilon x$, we get

$$\left\langle \frac{x_n - \lambda_n u_n - x_{n+1}}{\lambda_n} - y_{1,\varepsilon}, x_{n+1} - x \right\rangle \geq -\varepsilon,$$

equivalently:

$$2\langle x_n - x_{n+1}, x_{n+1} - x \rangle \geq 2\lambda_n \langle u_n + y_{1,\varepsilon}, x_{n+1} - x \rangle - 2\lambda_n \varepsilon. \tag{13}$$

By using the equality, valid for all $a, b, c \in \mathcal{H}$:

$$\|a - c\|^2 = \|a - b\|^2 + \|b - c\|^2 + 2\langle a - b, b - c \rangle,$$

we see that (13) is equivalent to

$$\|x_n - x\|^2 - \|x_{n+1} - x\|^2 \geq \|x_n - x_{n+1}\|^2 + 2\lambda_n \langle u_n + y_{1,\varepsilon}, x_{n+1} - x \rangle - 2\lambda_n \varepsilon. \tag{14}$$

Now, combining the facts that:

$$2\lambda_n \langle u_n + y_{1,\varepsilon}, x_{n+1} - x_n \rangle \geq -\|x_{n+1} - x_n\|^2 - \lambda_n^2 \|u_n + y_{1,\varepsilon}\|^2,$$

and

$$\langle u_n + y_{1,\varepsilon}, x_n - x \rangle = \langle u_n - y_{2,\varepsilon}, x_n - x \rangle + \langle y_\varepsilon, x_n - x \rangle,$$

we derive from (14) that:

$$\begin{aligned}
\|x_n - x\|^2 - \|x_{n+1} - x\|^2 \geq{}& \|x_n - x_{n+1}\|^2 - \|x_{n+1} - x_n\|^2 - \lambda_n^2 \|u_n + y_{1,\varepsilon}\|^2 \\
&+ 2\lambda_n \langle u_n - y_{2,\varepsilon}, x_n - x \rangle + 2\lambda_n \langle y_\varepsilon, x_n - x \rangle - 2\lambda_n \varepsilon.
\end{aligned} \tag{15}$$

Then, injecting (12) into (15) we obtain

$$2\lambda_n \langle y_\varepsilon, x_n - x \rangle \leq \|x_n - x\|^2 - \|x_{n+1} - x\|^2 + \lambda_n^2 \|u_n + y_{1,\varepsilon}\|^2 + 6\lambda_n \varepsilon + 4\lambda_n \varepsilon_n. \tag{16}$$

Summing up these inequalities for n going from 1 to $k \in \mathbb{N}^*$, and then dividing by $\sigma_k = \sum_{n=1}^{k} \lambda_n$, we get:

$$2\langle y_\varepsilon, \bar{x}_k - x \rangle \leq \frac{\|x_1 - x\|^2}{\sigma_k} + \frac{\sum_{n=1}^{k} \lambda_n^2 \|u_n + y_{1,\varepsilon}\|^2}{\sigma_k} + \frac{\sum_{n=1}^{k} \lambda_n (6\varepsilon + 4\varepsilon_n)}{\sigma_k}. \tag{17}$$

It follows from $\varepsilon_n = \lambda_n^{1/3}$ and (H1) that there exists a constant $M > 0$ such that

$$\lambda_n^2 \|u_n\|^2 = \lambda_n^{4/3} \varepsilon_n^2 \|u_n\|^2 \leq \lambda_n^{4/3} M^2,$$

hence, since $\sum \lambda_n^{4/3} < \infty$,

$$\lim_{k \to \infty} \sum_{n=1}^{k} \lambda_n^2 \|u_n + y_{1,\varepsilon}\|^2 \leq 2 \lim_{k \to \infty} \left(\sum_{n=1}^{k} \lambda_n^{4/3} M^2 + \sum_{n=1}^{k} \lambda_n^2 \|y_{1,\varepsilon}\|^2 \right) < +\infty.$$

Finally, since $\sigma_k \to +\infty$, we conclude that

$$\lim_{k \to \infty} \frac{\sum_{n=1}^{k} \lambda_n^2 \|u_n + y_{1,\varepsilon}\|^2}{\sigma_k} = 0.$$

On the other hand, since $\varepsilon_n \searrow 0$, there is $N \in \mathbb{N}$ such that $\varepsilon_n < \varepsilon$ for $n > N$, hence, for k large enough:

$$\frac{\sum_{n=1}^{k} \lambda_n (6\varepsilon + 4\varepsilon_n)}{\sigma_k} \leq \frac{\sum_{n=1}^{N} \lambda_n (6\varepsilon + 4\varepsilon_n)}{\sigma_k} + 10\varepsilon,$$

from which we derive that

$$\limsup_{k \to \infty} \frac{\sum_{n=1}^{k} \lambda_n (6\varepsilon + 4\varepsilon_n)}{\sigma_k} \leq 10\varepsilon.$$

Consequently, if $\bar{x}$ is any weak sequential cluster point of the sequence $(\bar{x}_n)$, letting $k \to \infty$ on both sides of (17) yields:

$$\langle y_\varepsilon, \bar{x} - x \rangle \leq 5\varepsilon, \tag{18}$$

and then letting $\varepsilon \to 0$ gives:

$$\langle y, x - \bar{x} \rangle \geq 0. \tag{19}$$

Thus, every weak sequential cluster point $\bar{x}$ of $(\bar{x}_n)$ satisfies (19) for every $(x, y) \in T$. The maximal monotonicity of T then implies that $0 \in T\bar{x}$, that is, $\bar{x} \in \mathcal{S}$.

(ii) Let $x \in \mathcal{S}$. It follows from (H2) that there exist $M > 0$ and sequences $y_{1,n} \in A^{\varepsilon_n}x$ and $y_{2,n} \in B^{\varepsilon_n}x$ with $\varepsilon_n\|y_{2,n}\| < M$ and $\|y_{1,n} + y_{2,n}\| < \varepsilon_n/(1 + \|x_n - x\|)$. Without loss of generality, we may also assume that $\varepsilon_n\|y_{1,n}\| < M$ and, by (H1), that $\varepsilon_n\|u_n\| < M$ for all n.

Proceeding as above with $y = 0$, $\varepsilon = \varepsilon_n$ and $y_{\varepsilon_n} = y_{1,n} + y_{2,n}$, we arrive at (16), hence

$$2\lambda_n\langle y_{\varepsilon_n}, x_n - x\rangle \leq \|x_n - x\|^2 - \|x_{n+1} - x\|^2 + 2\lambda_n^2(\|u_n\|^2 + \|y_{1,n}\|^2) + 10\lambda_n\varepsilon_n. \tag{20}$$

Then, using $\varepsilon_n = \lambda_n^{1/3}$ and taking the above upper bounds into account, we get:

$$2\lambda_n\langle y_{\varepsilon_n}, x_n - x\rangle \leq \|x_n - x\|^2 - \|x_{n+1} - x\|^2 + 4\lambda_n^{4/3}M^2 + 10\lambda_n^{4/3}, \tag{21}$$

and, since $\|y_{\varepsilon_n}\| \leq \varepsilon_n/(1 + \|x_n - x\|)$,

$$-2\lambda_n^{4/3} = -2\lambda_n\varepsilon_n \leq -2\lambda_n\|y_{\varepsilon_n}\|\,\|x_n - x\| \leq 2\lambda_n\langle y_{\varepsilon_n}, x_n - x\rangle, \tag{22}$$

so, putting (21) and (22) together,

$$\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 + \lambda_n^{4/3}(4M^2 + 12). \tag{23}$$

Since $\sum \lambda_n^{4/3} < \infty$, we derive from (23) that $\lim \|x_n - x\|^2$ exists, hence also $\lim \|x_n - x\|$. $\quad\square$

A useful special case of Theorem 3.1 is as follows:

Corollary 3.2. *Assume* $D(A) \subset \cap_{\varepsilon > 0} D(B^\varepsilon)$ *and let* (x_n) *be the sequence generated by* (11), *with* (λ_n) *and* (ε_n) *any sequences of positive real numbers verifying*

$$\sum \lambda_n = \infty, \qquad \sum \lambda_n^{4/3} < \infty, \qquad \varepsilon_n = \lambda_n^{1/3}.$$

Assume further that:

(H1) *The sequence* $(\varepsilon_n u_n)$ *is bounded,*
(H2′) $\mathcal{S} = (A + B)^{-1}(0)$.

Then, the sequence $(\bar{x}_n)$ *of weighted averages given by* (10) *weakly converges to a point in* $\mathcal{S}$.

Proof. We claim that assumption (H2′) implies assumption (H2). Indeed, let $x \in \mathcal{S}$. Then, we can find $v \in Ax \cap (-Bx)$, so for every small $\varepsilon_n > 0$, $0 \in Ax + Bx \cap (\|v\|/\varepsilon_n)B_{\mathcal{H}}$. Hence (H2) holds. The result therefore follows from Theorem 3.1. $\quad\square$

Remark 3.3. Examples of sequences (λ_n) and (ε_n) verifying the assumptions in the above results are $\lambda_n = 1/n$ and $\varepsilon_n = (1/n)^{1/3}$. As was pointed out by the referee, the proof remains valid for any sequences (λ_n) and (ε_n) verifying the relations

$$\sum \lambda_n = \infty, \qquad \sum (\lambda_n/\varepsilon_n)^2 < \infty, \qquad \sum \lambda_n\varepsilon_n < \infty, \quad \varepsilon_n \searrow 0.$$

Clearly, the optimal combination for these relations to be satisfied is when $(\lambda_n/\varepsilon_n)^2 = \lambda_n\varepsilon_n$, that is, $\varepsilon_n = \lambda_n^{1/3}$.

Remark 3.4. Assumption (H1) is of course satisfied if, as in [12, Theorem 2], the sequence (u_n) is supposed to be bounded. Simple examples show that (H1) may hold while (u_n) is not bounded, see Example 3.7 below.

Remark 3.5. Assumption (H2′) is of course satisfied if, as in Passty's result [12, Theorem 2], the usual sum $A + B$ is maximal monotone. In the aforementioned theorem, the assumptions are

$$\sum \lambda_n = \infty, \qquad \sum \lambda_n^2 < \infty, \quad u_n \in Bx_n, \ (u_n) \text{ bounded}.$$

Passty's result and our results are therefore not comparable. However, simple examples exist where (H2) is satisfied whereas (H2′) is not, see Example 3.7 below.

Remark 3.6. Assumption (H2′) is satisfied if, as in [11, Theorem 2] (dealing with the double-backward scheme), for every $x \in \mathcal{S}$ there exists $\varepsilon > 0$ such that $A^\varepsilon x$ is bounded. Indeed, let $x \in \mathcal{S}$. Since

$$0 \in \bigcap_n \overline{A^{\varepsilon_n}x + B^{\varepsilon_n}x},$$

there is a sequence $(y_{1,n}, y_{2,n}) \in A^{\varepsilon_n}x \times B^{\varepsilon_n}x$ such that $y_{1,n} + y_{2,n} \to 0$. From the boundedness of $A^\varepsilon x$ and $A^{\varepsilon_n}x \subset A^\varepsilon x$ for large n, we derive that $(y_{1,n})$ is bounded, hence we may assume that $(y_{1,n})$ weakly converges to some point y and $(y_{2,n})$ weakly converges to $-y$. Since the sets $A^{\varepsilon_n}x$ and $B^{\varepsilon_n}x$ are weakly closed, we derive that $y \in A^{\varepsilon_n}x$ and $-y \in B^{\varepsilon_n}x$ for every n. From the maximal monotonicity of A and B, we conclude that $y \in Ax \cap (-Bx)$, that is, $x \in (A + B)^{-1}(0)$. Thus, $\mathcal{S} = (A + B)^{-1}(0)$, hence (H2′) holds.

Example 3.7. In $\mathcal{H} = \mathbb{R}$, let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = -\sqrt{x}$ for $x \geq 0$, $f(x) = +\infty$ otherwise, and let $C = \{0\}$. We consider the problem (2) of minimizing the proper lower semicontinuous convex function f over the nonempty closed convex set C. As already seen in the introduction, this problem amounts to the inclusion (1) where $T = \partial\delta_C +_e \partial f$ is maximal monotone by (8). Clearly, the set of solutions of this problem is the nonempty set $\mathcal{S} = \{0\}$. Putting $A = \partial\delta_C$ and $B = \partial f$, the algorithm (11) becomes the projected approximate subgradient algorithm, noting that $(I+\partial\delta_C)^{-1}$ is the projection P_C onto C:

$$\begin{cases} x_0 \in \mathrm{dom}f \\ x_{n+1} = P_C(x_n - \lambda_n u_n) \quad \text{with } u_n \in (\partial f)^{\varepsilon_n}(x_n), \; \forall n \in \mathbb{N}. \end{cases} \tag{24}$$

We show that the assumptions of Theorem 3.1 are satisfied.

First, note that $\mathrm{D}(A) = \{0\}$ and for every $\varepsilon > 0$, $\mathrm{D}(\partial_\varepsilon f) = \mathrm{dom}f = [0, +\infty[$, hence $\mathrm{D}(A) \subset \bigcap_{\varepsilon>0} \mathrm{D}(B^\varepsilon)$.

Next, we have to check the assumptions (H1) and (H2). Easy computations show that for every $\varepsilon > 0$, $\partial_\varepsilon \delta_C(0) =]-\infty, +\infty[$ and $\partial_\varepsilon f(0) =]-\infty, -1/4\varepsilon]$.

For (H1), observe that $x_n = 0$ for every $n = 1, 2, \ldots$, since $C = \{0\}$. We may therefore choose $u_n = -1/4\varepsilon_n \in \partial_{\varepsilon_n} f(0)$ so that $\varepsilon_n u_n = -1/4$. Assumption (H1) is satisfied.

For (H2), observe that the only point in $\mathcal{S}$ is $x = 0$. Since $\partial_\varepsilon f(0) \cap [-1/4\varepsilon, 1/4\varepsilon] = \{-1/4\varepsilon\}$, we have

$$0 = 1/4\varepsilon - 1/4\varepsilon \in \partial_\varepsilon \delta_C(0) + \partial_\varepsilon f(0) \cap (1/4\varepsilon)B_\mathbb{R}$$

for every $\varepsilon > 0$. Assumption (H2) is satisfied.

Thus, we may apply Theorem 3.1 to derive the convergence in average of the sequence generated by the algorithm (24) to the point minimizing f over C. We point out that no sequence $u_n \in \partial_{\varepsilon_n} f(x_n)$ is bounded, the set $\partial f(x_n) = \partial f(0)$ is empty, hence the operator $\partial\delta_C + \partial f$ is empty, *a fortiori* the set $(\partial\delta_C + \partial f)^{-1}(0)$ is also empty, the sets $\partial_\varepsilon \delta_C(0)$ and $\partial_\varepsilon f(0)$ are not bounded. So none of the previous algorithms in [2,7,12,11] is applicable.

References

[1] H. Brezis, P.-L. Lions, Produits infinis de résolvantes, Israel J. Math. 29 (4) (1978) 329–345.
[2] R.E. Bruck Jr., On the weak convergence of an ergodic iteration for the solution of variational inequalities for monotone operators in Hilbert space, J. Math. Anal. Appl. 61 (1) (1977) 159–164.
[3] R.S. Burachik, B.F. Svaiter, ϵ-enlargements of maximal monotone operators in Banach spaces, Set-Valued Anal. 7 (2) (1999) 117–132.
[4] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, Optimization 53 (5–6) (2004) 475–504.
[5] Y. García, M. Lassonde, J.P. Revalski, Extended sums and extended compositions of monotone operators, J. Convex Anal. 13 (3–4) (2006) 721–738.
[6] J.-B. Hiriart-Urruty, R.R. Phelps, Subdifferential calculus using ϵ-subdifferentials, J. Funct. Anal. 118 (1) (1993) 154–166.
[7] P.-L. Lions, Une méthode itérative de résolution d'une inéquation variationnelle, Israel J. Math. 31 (2) (1978) 204–208.
[8] B. Martinet, Régularisation d'inéquations variationnelles par approximations successives, Rev. Française Informat. Recherche Opérationnelle 4 (Ser. R-3) (1970) 154–158.
[9] B. Martinet, Détermination approchée d'un point fixe d'une application pseudo-contractante. Cas de l'application prox, C. R. Acad. Sci. Paris Sér. A–B 274 (1972) A163–A165.
[10] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France 93 (1965) 273–299.
[11] A. Moudafi, M. Théra, Ergodic convergence to a zero of the extended sum of two maximal monotone operators, in: Inherently Parallel Algorithms in Feasibility and Optimization and their Applications (Haifa, 2000), in: Stud. Comput. Math., vol. 8, North-Holland, Amsterdam, 2001, pp. 369–379.
[12] G.B. Passty, Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, J. Math. Anal. Appl. 72 (2) (1979) 383–390.
[13] J. Peypouquet, S. Sorin, Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time, J. Convex Anal. 17 (3–4) (2010) 1113–1163.
[14] J.P. Revalski, M. Théra, Generalized sums of monotone operators, C. R. Acad. Sci. Paris Sér. I 329 (11) (1999) 979–984.
[15] J.P. Revalski, M. Théra, Enlargements and sums of monotone operators, Nonlinear Anal. Ser. A: Theory Methods 48 (4) (2002) 505–519.
[16] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (5) (1976) 877–898.