Ergodic convergence of the forward-backward algorithm to a zero of the extended sum of two maximal monotone operators

Ludovic NAGESSEUR

University of French West Indies and Guiana
Laboratory LAMIA
97159, Pointe-à-Pitre, Guadeloupe, France
Framework, notations and definitions

- $\mathcal{H}$ denotes a Hilbert space;
- All operators considered are set-valued from $\mathcal{H}$ to $\mathcal{H}$;
- $A^{-1}(0)$ is the set of zeros of the operator $A$:

$$x \in A^{-1}(0) \iff 0 \in A(x).$$

- $A$ is said to be pointwise bounded (resp. locally bounded) at $x$ if the set $A(x)$ is bounded (resp. if there exists a neighborhood $U$ of $x$ such that $A(U)$ is bounded).
Let $X$ be a Banach space, $X^*$ its dual space, and $\langle ., . \rangle : X^* \times X \to \mathbb{R}$ denotes the inner product.  
A set-valued operator $T : X \to 2^{X^*}$ is

- **monotone**, if
  \[
  \forall (x, x^*) \in \text{Gr}(T), \quad \langle x^* - y^*, x - y \rangle \geq 0;
  \]

- **maximal monotone** if it is monotone and its graph is not contained in the graph of another monotone operator.

### Proposition

A monotone operator $T$ is maximal monotone iff
\[
\forall (x, x^*) \in X \times X^* \quad \forall (y, y^*) \in \text{Gr}(T) : \quad \langle x^* - y^*, x - y \rangle \geq 0 \quad \Rightarrow \quad (x, x^*) \in \text{Gr}(T).
\]

### Example

$f : X \to \mathbb{R} \cup \{+\infty\}$ convex, $\text{dom} f := \{x \in X | f(x) < +\infty\} \neq \emptyset$. The subdifferential of $f$ is the set-valued mapping denoted $\partial f : X \to 2^{X^*}$ defined by:
\[
\partial f(x) := \{x^* \in X | \langle x^*, y - x \rangle + f(x) \leq f(y) \quad \forall y \in X\}.
\]

$f$ proper, convex, lower semi-continuous (l.s.c.) $\Rightarrow \partial f$ is maximal monotone [6, Rockafellar, 1970].
$f$ proper, convex, then $0 \in \partial f(x) \iff x \in \text{argmin}_X f$. 

References
Resolvent and Yosida regularization of a monotone operator $A$

Let $\lambda > 0$ and $A$ be a monotone operator.

**Resolvent**

$$J_{\lambda A} := (I + \lambda A)^{-1}.$$  

This operator is a contraction; it has full domain iff $A$ is maximal monotone.

**Yosida regularization**

$$A_{\lambda} := \frac{I - J_{\lambda A}}{\lambda}.$$  

This operator is $\frac{1}{\lambda}$-lipschitz; it has full domain iff $A$ is maximal monotone. Moreover, $A_{\lambda}$ verifies:

$$\forall x \in \mathcal{H}, \quad A_{\lambda}(x) \in A(J_{\lambda A}(x)).$$
Objectives

**Goal:** We want to solve the problem:

"Find \( x \in \mathcal{H} \) such that \( 0 \in \left( A + B \right)_{\text{ext}}(x) \)," \hspace{1cm} (1)

where \( A \) and \( B \) are maximal monotone operators defined on \( \mathcal{H} \).

**Method:** Construct a sequence \( \{z_n\} \) converging to a solution of the problem (1).
The sequence \( \{z_n\} \) we will construct is based on the Forward-Backward (F-B) algorithm, given by \( x_0 \in D(B) \) and :

\[
x_{n+1} = J_{\lambda_n A}(x_n - \lambda_n y_n) \quad \text{avec} \quad y_n \in B(x_n), \quad \forall n \in \mathbb{N},
\]

where \( \{\lambda_n\}_{n \in \mathbb{N}} \) is a sequence of positive reals. The F-B iteration \( \{x_n\} \) is given equivalently by :

\[
(2) \quad \Leftrightarrow \quad x_{n+1} \in J_{\lambda_n A}(x_n - \lambda_n B(x_n))
\]

\[
\Leftrightarrow \quad (I - \lambda_n B)(x_n) \ni (I + \lambda_n A)(x_{n+1}).
\]

The F-B algorithm is generally used to solve the problem :

"Find \( x \in \mathcal{H} \) such that \( 0 \in (A + B)(x) \)."
The sequence \( \{z_n\} \) we will study is given by:

\[
z_n = \frac{\sum_{k=1}^{n} \lambda_k x_k}{\sum_{k=1}^{n} \lambda_k}, \quad \forall n \in \mathbb{N}.
\]

It is the sequence of the weighted average of the F-B iteration \( \{x_n\} \). The convergence of such a sequence is called **ergodic convergence**. We will prove the weak convergence of \( \{z_n\} \).
Motivations

We know that the **usual sum** of two maximal monotone operators $A$ and $B$ is monotone, but in general, it is not maximal monotone.

The **extended sum** is an extension of the usual sum, in the sense of graph inclusion, and it can be maximal monotone in cases where the usual sum it is not.
Plan

1. $\epsilon$-Enlargement of monotone operators

2. The extended sum of two monotone operators
   - Definition of the extended sum of two monotone operators
   - Some properties of the extended sum

3. Convergence of the backward-backward and barycentric-proximal algorithms to a zero of the extended sum of two maximal monotone operators (Moudafi et Théra)

4. Ergodic convergence of the forward-backward algorithm to a zero of the extended sum of two maximal monotone operators
Plan

1. $\epsilon$-Enlargement of monotone operators

2. The extended sum of two monotone operators
   - Definition of the extended sum of two monotone operators
   - Some properties of the extended sum

3. Convergence of the backward-backward and barycentric-proximal algorithms to a zero of the extended sum of two maximal monotone operators (Moudafi et Théra)

4. Ergodic convergence of the forward-backward algorithm to a zero of the extended sum of two maximal monotone operators
**Definition**

$T : X \rightarrow 2^{X^*}$, monotone. We define the $\epsilon$-enlargement of $T$ for any $\epsilon \geq 0$ and all $x \in X$, by:

$$T^\epsilon(x) := \{x^* \in X^*| \langle y^* - x^*, y - x\rangle \geq -\epsilon \text{ for all } (y, y^*) \in \text{Gr}(T)\}.$$

**Properties**

- If $T$ is maximal monotone: $T^0(x) = T(x) = \bigcap_{\epsilon > 0} T^\epsilon(x)$, for all $x \in X$;
- $0 < \epsilon_1 < \epsilon_2 \Rightarrow T(x) \subset T^\epsilon_1(x) \subset T^\epsilon_2(x)$;
1. $\epsilon$-Enlargement of monotone operators

2. The extended sum of two monotone operators
   - Definition of the extended sum of two monotone operators
   - Some properties of the extended sum

3. Convergence of the backward-backward and barycentric-proximal algorithms to a zero of the extended sum of two maximal monotone operators (Moudafi et Théra)

4. Ergodic convergence of the forward-backward algorithm to a zero of the extended sum of two maximal monotone operators
The extended sum of two monotone operators $A$ and $B$ is a notion introduced by J. P. Revalski and M. Théra in 2002 in [5]. It is the operator defined by:

$$
(A + B)^\text{ext}(x) := \bigcap_{\epsilon > 0} A^\epsilon(x) + B^\epsilon(x)^{w^*}, \text{ for all } x \in \mathcal{H}.
$$

As weak $w$ and weak star $w^*$ topologies coincide on a reflexive Banach space (it is the case of $\mathcal{H}$), it follows that:

$$
(A + B)^\text{ext}(x) = \bigcap_{\epsilon > 0} A^\epsilon(x) + B^\epsilon(x)^{w}, \text{ for all } x \in \mathcal{H}.
$$

$$
\text{Gr}(A + B) \subset \text{Gr}
$$

$$
(A + B)^\text{ext}.
$$
Theorem

Let $X$ be a Banach space, $X^*$ its continuous dual, and $A, B : X \rightarrow 2^{X^*}$ two maximal monotone operators. Then:

(i) [1, Proposition 3.4, 2006] the extended sum $A + B$ is a monotone operator;

(ii) [5, Corollary 4.2, 2002] if $A + B$ is maximal monotone, then

$$(A + B)(x) = \left( A + B \right)_{\text{ext}} (x)$$

for all $x \in X$;

(iii) [5, Theorem 4.4, 2002] let $f, g : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be two proper lower semicontinuous convex functions; according to Rockafellar [6], the subdifferentials $\partial f$ and $\partial g$ are maximal monotone. If $\text{dom}f \cap \text{dom}g \neq \emptyset$, then:

$$\partial(f + g)(x) = \left( \partial f + \partial g \right)_{\text{ext}} (x), \text{ for all } x \in X,$$

while $\partial f(x) + \partial g(x) \neq \partial(f + g)(x)$ in general.
Plan

1. $\epsilon$-Enlargement of monotone operators

2. The extended sum of two monotone operators
   - Definition of the extended sum of two monotone operators
   - Some properties of the extended sum

3. Convergence of the backward-backward and barycentric-proximal algorithms to a zero of the extended sum of two maximal monotone operators (Moudafi et Théra)

4. Ergodic convergence of the forward-backward algorithm to a zero of the extended sum of two maximal monotone operators
In 2001, in order to solve the inclusion (1), Moudafi and Théra in [3], have proposed two splitting methods: the backward-backward and the barycentric-proximal algorithms, respectively given by:

$$x_n = J_{\lambda_n}A J_{\lambda_n}B x_{n-1} \quad \forall n \in \mathbb{N}^*,$$

(3)

$$\tilde{x}_n = \frac{\mu_n}{\lambda_n + \mu_n} J_{\lambda_n}A \tilde{x}_{n-1} + \frac{\lambda_n}{\lambda_n + \mu_n} J_{\lambda_n}B \tilde{x}_{n-1} \quad \forall n \in \mathbb{N}^*,$$

(4)

where \{\lambda_n\} and \{\mu_n\} are two sequences of positive real numbers.

Originally, those algorithms were devoted to solve the more classical problem of finding a zero of the sum \(A + B\). Thus, in Passty [4, 1978], and Lehdili and Lemaire [2, 1999], it has been proved their weak \textit{ergodic} convergence.
This is the result of Moudafi and Théra:

**Theorem**

[3, Theorem 2] Suppose that $D(A) \cap D(B) \neq \emptyset$ (to guarantee the existence of the iterates), and that $A + B$ is maximal monotone. Further, suppose that the problem (1) has a solution. Let $\{x_n\}$ (resp. $\{\tilde{x}_n\}$) be a sequence generated by (3) (resp. by (4)) and $\{z_n\}$ (resp. $\{\tilde{z}_n\}$) be the associated weighted average. If

$$\sum_{n=1}^{\infty} \lambda_n^2 < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \lambda_n = \infty,$$

then, any weak limit point of a subsequence of $\{z_n\}$ (resp. $\{\tilde{z}_n\}$) is a zero of the extended sum. Moreover, if $A^\epsilon$ is locally bounded on $\left(A + B \right)_{\text{ext}}^{-1}(0)$, then the whole sequence $\{z_n\}_{n \in \mathbb{N}}$ (resp. $\{\tilde{z}_n\}$) weakly converges to some solution of (1).
Plan

1. \(\epsilon\)-Enlargement of monotone operators

2. The extended sum of two monotone operators
   - Definition of the extended sum of two monotone operators
   - Some properties of the extended sum

3. Convergence of the backward-backward and barycentric-proximal algorithms to a zero of the extended sum of two maximal monotone operators (Moudafi et Théra)

4. Ergodic convergence of the forward-backward algorithm to a zero of the extended sum of two maximal monotone operators
The forward-backward algorithm is given by the iteration:

\[ x_{n+1} = J_{\lambda_n A}(x_n - \lambda_n y_n) \quad \text{with} \quad y_n \in B(x_n), \quad \forall n \in \mathbb{N}, \quad (5) \]

and with \( D(A) \subset D(B) \) so that the iteration be well defined.
\textbf{Lemma 1}

Let \( \{\lambda_n\}_{n \in \mathbb{N}} \) be a sequence of positive reals such that \( \sum \lambda_n = +\infty \), and let \( \{x_n\}_{n \in \mathbb{N}} \) be a sequence of real numbers to which we associate the weighted average sequence \( \{z_n\}_{n \in \mathbb{N}} \) defined by \( z_n = \frac{\sum_{k=1}^{n} \lambda_k x_k}{\sum_{k=1}^{n} \lambda_k} \). If \( \lim_{n \to \infty} x_n = a \), then \( \lim_{n \to \infty} z_n = a \).

\textbf{Lemma 2 (Opial-Passty)}

Let \( \{\lambda_n\}_{n \in \mathbb{N}} \) be a sequence of positive reals such that \( \sum \lambda_n = +\infty \), and let \( \{x_n\}_{n \in \mathbb{N}} \) be a sequence of \( \mathcal{H} \) to which we associate the weighted average sequence \( \{z_n\}_{n \in \mathbb{N}} \) defined by:

\[
z_n = \frac{\sum_{k=1}^{n} \lambda_k x_k}{\sum_{k=1}^{n} \lambda_k}.
\]

Let us assume there exists a nonempty set \( \mathcal{S} \subset \mathcal{H} \) such that:

1. any weak sequential cluster point of \( \{z_n\}_{n \in \mathbb{N}} \) is in \( \mathcal{S} \);  
2. for all \( a \in \mathcal{S} \), the limit \( \lim_{n \to +\infty} \|x_n - a\| \) exists.

Then, \( \{z_n\}_{n \in \mathbb{N}} \) weakly converges to an element of \( \mathcal{S} \).
Lemma 3

Let \( \{\mu_n\}_{n \in \mathbb{N}} \) and \( \{a_n\}_{n \in \mathbb{N}} \) be two sequences of positive reals such that:

\[
a_{n+1} \leq a_n + \mu_n \quad \text{and} \quad \sum \mu_n < +\infty.
\]

Then, the limit \( \lim_{n \to +\infty} a_n \) exists.
Main result

Let $A$ and $B$ be two maximal monotone operators on a Hilbert space $\mathcal{H}$, with $D(A) \subset D(B)$. Assume that $A + B$ is maximal monotone and that

$$\left(A + B\right)^{-1}_{\text{ext}}(0) \neq \emptyset.$$ 

Let $\{x_n\}_{n \in \mathbb{N}}$ be the sequence generated by the F-B iteration:

$$x_{n+1} = J_{\lambda_n A}(x_n - \lambda_n y_n) \quad \text{with} \quad y_n \in B(x_n), \quad \forall n \in \mathbb{N}$$

and let $\{z_n\}_{n \in \mathbb{N}}$ be the corresponding weighted average sequence:

$$z_n = \frac{\sum_{k=1}^{n} \lambda_k x_k}{\sum_{k=1}^{n} \lambda_k},$$

with any starting point $x_0 \in D(B)$, and with $\{\lambda_n\}$, a sequence of positive reals satisfying $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$ and $\sum_{n=1}^{\infty} \lambda_n = \infty$. Assume moreover that $\{y_n\}$ is bounded. Then, any weak sequential cluster point of $\{z_n\}_{n \in \mathbb{N}}$ is a zero of the extended sum.

Moreover, if there exists $\epsilon_0 > 0$ such that $A^{\epsilon_0}$ is pointwise bounded on $\left(A + B\right)^{-1}_{\text{ext}}(0)$, then the whole sequence $\{z_n\}_{n \in \mathbb{N}}$ weakly converges to some zero of

$$\left(A + B\right)^{-1}_{\text{ext}}(0).$$
Idea of the proof

We show that the hypothesis of Lemme 2 are verified for
\[ \mathcal{J} = \left( A + B \right)_{\text{ext}}^{-1}(0). \]

\[ \mathcal{J} \neq \emptyset \] by hypothesis.

Let \( (x, y) \in \text{Gr} \left( A + B \right)_{\text{ext}} \); by definition of the extended sum, one has:
\[ y \in A^\epsilon(x) + B^\epsilon(x)^w \] for all \( \epsilon > 0 \).

\[ \Rightarrow \quad \forall \epsilon > 0, \quad \exists \{ y_{p, \epsilon} \}_p : \]
\[ y_{p, \epsilon} = y_{1, p, \epsilon} + y_{2, p, \epsilon} \quad \text{and} \quad y_{p, \epsilon} \rightharpoonup y. \]
\[
A^\epsilon(x) \quad \quad B^\epsilon(x)
\]
Let us fix $\epsilon > 0$. By definition of $B^\epsilon$, one has:

$$y_n \in B(x_n)$$
$$y_{2,p,\epsilon} \in B^\epsilon(x)$$

$\Rightarrow \langle y_n - y_{2,p,\epsilon}, x_n - x \rangle \geq -\epsilon.$  \hspace{1cm} (6)

As $A_{\lambda_n}(x_n - \lambda_n y_n) \in A(J_{\lambda_n}A(x_n - \lambda_n y_n)) = A(x_{n+1})$, we have, by definition of $A^\epsilon$:

$$\left\langle \frac{(x_n - \lambda_n y_n) - x_{n+1}}{\lambda_n} - y_{1,p,\epsilon}, x_{n+1} - x \right\rangle \geq -\epsilon,$$

which can be rewritten as:

$$2\langle x_n - x_{n+1}, x_{n+1} - x \rangle \geq 2\lambda_n \langle y_n + y_{1,p,\epsilon}, x_{n+1} - x \rangle - 2\lambda_n \epsilon.$$  \hspace{1cm} (7)
By using the equality, true for all \( a, b, c \in \mathcal{H} \):

\[
2\langle a - b, b - c \rangle = \| a - c \|^2 - \| a - b \|^2 - \| b - c \|^2,
\]

we obtain:

\[
(7) \Leftrightarrow \| x_n - x \|^2 - \| x_n - x_{n+1} \|^2 \geq 2\lambda_n \langle y_n + y_1, p, \epsilon, x_{n+1} - x \rangle - 2\lambda_n \epsilon
\]

By combining the fact that:

\[
\langle y_n + y_1, p, \epsilon, x_{n+1} - x_n \rangle \geq -\frac{1}{2} \| x_{n+1} - x_n \|^2 - \frac{1}{2} \| y_n + y_1, p, \epsilon \|^2,
\]

and:

\[
\langle y_n + y_1, p, \epsilon, x_n - x \rangle = \langle y_n - y_2, p, \epsilon, x_n - x \rangle + \langle y_p, \epsilon, x_n - x \rangle,
\]

it follows that:

\[
\| x_n - x \|^2 - \| x_{n+1} - x \|^2 \geq \| x_n - x_{n+1} \|^2 - \| x_{n+1} - x_n \|^2 - \lambda_n^2 \| y_n + y_1, p, \epsilon \|^2
\]

\[
+ 2\lambda_n \langle y_n - y_2, p, \epsilon, x_n - x \rangle + 2\lambda_n \langle y_p, \epsilon, x_n - x \rangle - 2\lambda_n \epsilon.
\]
The inequality (6) allows us to obtain:

\[ \| x_n - x \|^2 - \| x_{n+1} - x \|^2 \geq -\lambda_n^2 \| y_n + y_{1,\rho,\epsilon} \|^2 + 2\lambda_n \langle y_{\rho,\epsilon}, x_n - x \rangle - 4\lambda_n \epsilon, \]

which is equivalent to:

\[ 2\lambda_n \langle y_{\rho,\epsilon}, x_n - x \rangle \leq \| x_n - x \|^2 - \| x_{n+1} - x \|^2 + \lambda_n^2 \| y_n + y_{1,\rho,\epsilon} \|^2 + 4\lambda_n \epsilon. \quad (8) \]

By using the hypothesis \{y_n\} bounded, there exists a constant \( M_{\rho,\epsilon} \), independent of \( n \) such that:

\[ 2\lambda_n \langle y_{\rho,\epsilon}, x_n - x \rangle \leq \| x_n - x \|^2 - \| x_{n+1} - x \|^2 + M_{\rho,\epsilon} \lambda_n^2 + 4\lambda_n \epsilon. \]

By summing this equality for \( n \) going from 1 to \( k \in \mathbb{N}^* \), we obtain:

\[ 2 \left\langle y_{\rho,\epsilon}, \sum_{n=1}^{k} \lambda_n (x_n - x) \right\rangle \leq \| x_1 - x \|^2 - \| x_{k+1} - x \|^2 + M_{\rho,\epsilon} \sum_{n=1}^{k} \lambda_n^2 + 4\epsilon \sum_{n=1}^{k} \lambda_n \]

\[ \leq \| x_1 - x \|^2 + M_{\rho,\epsilon} \sum_{n=1}^{k} \lambda_n^2 + 4\epsilon \sum_{n=1}^{k} \lambda_n. \quad (9) \]
Then, by dividing the inequality (9) by $\sum_{n=1}^{k} \lambda_n$, one has:

$$2\langle y_p, \epsilon, z_k - x \rangle \leq \frac{\|x_1 - x\|^2}{\sum_{n=1}^{k} \lambda_n} + \frac{M_p, \epsilon \sum_{n=1}^{k} \lambda_n^2}{\sum_{n=1}^{k} \lambda_n} + 4\epsilon.$$

Let $\bar{z}$ be a weak sequential cluster point of $\{z_n\}$. By passing to the limit $k \to +\infty$ for a subsequence of $\{z_n\}$ converging to $\bar{z}$ in the last inequality, one has:

$$\langle y_p, \epsilon, \bar{z} - x \rangle \leq 2\epsilon.$$

For $p \to +\infty$ and $\epsilon \to 0$, we obtain:

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

$$\forall (x, y) \in \text{Gr} \left(A + B \right)_{\text{ext}}, \; \langle y, x - \bar{z} \rangle \geq 0 \quad \Rightarrow \quad 0 \in \left(A + B \right)_{\text{ext}}(\bar{z}).$$

We have shown that any weak sequential cluster point of $\{z_n\}$ is a zero of $A + B$. The first hypothesis of Opial-Passty’s Lemma is then satisfied, and we prove also the first part of the theorem.
Let us put now \( x \in \left( A + B \right)^{\text{ext}}_{-1}(0) \) in (8) (we take \( y = 0 \)). Let \( 0 < \epsilon < \epsilon_0 \). The inequality (8) can be rewritten as:

\[
\| x_{n+1} - x \|^2 \leq \| x_n - x \|^2 - 2\lambda_n \langle y_{p,\epsilon}, x_n - x \rangle + \lambda_n^2 \| y_n + y_{1,p,\epsilon} \|^2 + 4\lambda_n \epsilon. \tag{10}
\]

We have:

\[
\begin{align*}
&x \in \left( A + B \right)^{\text{ext}}_{-1}(0) \\
&\forall p, y_{1,p,\epsilon} \in A^\epsilon(x) \subset A^{\epsilon_0}(x) \Rightarrow \{ y_{1,p,\epsilon} \}_{p} \text{ is bounded.}
\end{align*}
\]

\( A^{\epsilon_0} \) is pointwise bounded on \( \left( A + B \right)^{\text{ext}}_{-1}(0) \).
Combining that with \( \{y_n\} \) bounded, there exists a constant \( C > 0 \) such that:

\[
\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - 2\lambda_n \langle y_p, \epsilon, x_n - x \rangle + \lambda_n^2 C + 4\lambda_n \epsilon. \tag{11}
\]

By reminding that \( y_{p, \epsilon} \to 0 \) when \( p \to +\infty \), by doing \( p \to +\infty \) and \( \epsilon \to 0 \) in the last inequality, one obtains:

\[
\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 + \lambda_n^2 C.
\]

Since \( \sum \lambda_n^2 < +\infty \), we can apply the Lemma 3 with \( a_n = \|x_n - x\|^2 \) and \( \mu_n = C\lambda_n^2 \) for all \( n \in \mathbb{N} \), to obtain that \( \lim_{n \to +\infty} \|x_n - x\|^2 \) exists.

Consequently, the second hypothesis of Lemma 2 is also verified. The Lemma 2 permits us to establish the weak convergence of the sequence \( \{z_n\} \) to a zero of the extended sum of \( A \) and \( B \).
Case $A + B$ maximal monotone

**Corollary 1**

Let $A$ and $B$ be two maximal monotone operators on a Hilbert space $\mathcal{H}$, with $D(A) \subset D(B)$. Assume that $A + B$ is maximal monotone, and that $(A + B)^{-1}(0) \neq \emptyset$. Let $\{x_n\}_{n \in \mathbb{N}}$ be the sequence generated by the iteration:

$$x_{n+1} = J_{\lambda_n A}(x_n - \lambda_n y_n) \quad \text{with} \quad y_n \in B(x_n), \quad \forall n \in \mathbb{N},$$

and let $\{z_n\}_{n \in \mathbb{N}}$ be the corresponding weighted average sequence:

$$z_n = \frac{\sum_{k=1}^{n} \lambda_k x_k}{\sum_{k=1}^{n} \lambda_k},$$

with any starting point $x_0 \in D(B)$, and with $\{\lambda_n\}$, a sequence of positive reals satisfying $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$ and $\sum_{n=1}^{\infty} \lambda_n = \infty$. Assume also that $\{y_n\}$ is bounded. Then, $\{z_n\}_{n \in \mathbb{N}}$ is bounded and any weak sequential cluster point of $\{z_n\}_{n \in \mathbb{N}}$ is a zero of $A + B$. Moreover, assume that there exists $\epsilon_0 > 0$ such that $A^{\epsilon_0}$ is pointwise bounded on $(A + B)^{-1}(0)$. Then, $\{z_n\}$ weakly converges to an element $\bar{z} \in (A + B)^{-1}(0)$. 
Corollary 2

Let $f$ and $g$ be two convex, proper and lower semicontinuous functions on a Hilbert space $\mathcal{H}$, with $D(\partial f) \subset D(\partial g)$. Assume that $\partial f + \partial g$ is maximal monotone (that is $\partial f + \partial g = \partial(f + g)$), and that

$$ \text{argmin}_{\mathcal{H}} (f + g) := \{ z \in \mathcal{H} : (f + g)(z) \leq (f + g)(y) \ \forall y \in \mathcal{H} \} $$

is nonempty. Let $\{x_n\}_{n \in \mathbb{N}}$ be the sequence generated by the iteration

$$ x_{n+1} = J_{\lambda_n} \partial f (x_n - \lambda_n y_n) \quad \text{with} \quad y_n \in \partial g(x_n), \quad \forall n \in \mathbb{N}, $$

and let $\{z_n\}_{n \in \mathbb{N}}$ be the corresponding weighted average sequence:

$$ z_n = \frac{\sum_{k=1}^{n} \lambda_k x_k}{\sum_{k=1}^{n} \lambda_k}, $$

with any starting point $x_0 \in D(\partial g)$, and with $\{\lambda_n\}$, a sequence of positive reals satisfying $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$ and $\sum_{n=1}^{\infty} \lambda_n = \infty$. Assume also that $\{y_n\}$ is bounded. Then, $\{z_n\}_{n \in \mathbb{N}}$ is bounded and any weak sequential cluster point of $\{z_n\}_{n \in \mathbb{N}}$ is a minimum of $f + g$. Moreover, assume that there exists $\epsilon_0 > 0$, $(\partial f)^{\epsilon_0}$ is pointwise bounded on $\text{argmin}_{\mathcal{H}} (f + g)$. Then $\{z_n\}$ weakly converges to $\bar{z} \in \text{argmin}_{\mathcal{H}} (f + g)$. 

Ergodic convergence of the forward-backward algorithm to a zero of the extended sum of two maximal monotone operators


Julian P. Revalski and Michel Théra. 
Enlargements and sums of monotone operators. 

R. T. Rockafellar. 
On the maximal monotonicity of subdifferential mappings. 
Thank you for your attention.