Compositions and Convex Combinations of Averaged Nonexpansive Operators*

Patrick L. Combettes[1] and Isao Yamada[2]

[1]Sorbonne Universités – UPMC Univ. Paris 06
UMR 7598, Laboratoire Jacques-Louis Lions
F-75005 Paris, France
plc@ljll.math.upmc.fr

[2]Tokyo Institute of Technology
Department of Communications and Computer Engineering
Tokyo 152-8550, Japan
isao@sp.ce.titech.ac.jp

Abstract

Properties of compositions and convex combinations of averaged nonexpansive operators are investigated and applied to the design of new fixed point algorithms in Hilbert spaces. An extended version of the forward-backward splitting algorithm for finding a zero of the sum of two monotone operators is obtained.

Keywords. averaged operator · fixed-point algorithm · forward-backward splitting · monotone operator · nonexpansive operator

*Contact author: P. L. Combettes, plc@ljll.math.upmc.fr, phone: +33 1 4427 6319, fax: +33 1 4427 7200.

1 Introduction

Since their introduction in [3], averaged nonexpansive operators have proved to be very useful in the analysis and the numerical solution of problems arising in nonlinear analysis and its applications; see, e.g., [2, 4, 5, 6, 7, 8, 11, 14, 15, 16, 18, 19, 20, 21].

Definition 1.1 Let $\mathcal{H}$ be a real Hilbert space, let D be a nonempty subset of $\mathcal{H}$, let $\alpha \in \left]0,1\right[$, and let $T\colon D \to \mathcal{H}$ be a nonexpansive (i.e., 1-Lipschitz) operator. Then T is averaged with constant α, or α-averaged, if there exists a nonexpansive operator $R\colon D \to \mathcal{H}$ such that $T = (1-\alpha)\mathrm{Id} + \alpha R$.

As discussed in [6, 11, 16], averaged operators are stable under compositions and convex combinations and such operations form basic building blocks in various composite fixed point algorithms. The averagedness constants resulting from such operations determine the range of the step sizes and other parameters in such algorithms. It is therefore important that they be tight since these parameters have a significant impact on the speed of convergence.

In this paper, we discuss averagedness constants for compositions and convex combinations of averaged operators and construct novel fixed point algorithms based on these constants. In particular, we obtain a new version of the forward-backward algorithm with an extended relaxation range and iteration-dependent step sizes.

Throughout the paper, $\mathcal{H}$ is a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$ and associated norm $\|\cdot\|$. We denote by Id the identity operator on $\mathcal{H}$ and by d_S the distance function to a set $S \subset \mathcal{H}$; $\rightharpoonup$ and $\to$ denote, respectively, weak and strong convergence in $\mathcal{H}$.

2 Compositions and convex combinations of averaged operators

We first recall some characterizations of averaged operators (see [11, Lemma 2.1] or [6, Proposition 4.25]).

Proposition 2.1 *Let D be a nonempty subset of $\mathcal{H}$, let $T\colon D \to \mathcal{H}$ be nonexpansive, and let $\alpha \in \left]0,1\right[$. Then the following are equivalent:*

(i) *T is α-averaged.*

(ii) *$(1-1/\alpha)\mathrm{Id} + (1/\alpha)T$ is nonexpansive.*

(iii) $(\forall x \in D)(\forall y \in D)\ \|Tx - Ty\|^2 \leqslant \|x-y\|^2 - \dfrac{1-\alpha}{\alpha}\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2$.

(iv) $(\forall x \in D)(\forall y \in D)\ \|Tx - Ty\|^2 + (1-2\alpha)\|x-y\|^2 \leqslant 2(1-\alpha)\langle x - y \mid Tx - Ty\rangle$.

The next result concerns the averagedness of a convex combination of averaged operators.

Proposition 2.2 *Let D be a nonempty subset of $\mathcal{H}$, let $(T_i)_{i\in I}$ be a finite family of nonexpansive operators from D to $\mathcal{H}$, let $(\alpha_i)_{i\in I}$ be a family in $\left]0,1\right[$, and let $(\omega_i)_{i\in I}$ be a family in $\left]0,1\right]$ such that $\sum_{i\in I}\omega_i = 1$. Suppose that, for every $i \in I$, T_i is α_i-averaged, and set $T = \sum_{i\in I}\omega_i T_i$ and $\alpha = \sum_{i\in I}\omega_i\alpha_i$. Then T is α-averaged.*

Proof. For every $i \in I$, there exists a nonexpansive operator $R_i \colon D \to \mathcal{H}$ such that $T_i = (1-\alpha_i)\mathrm{Id} + \alpha_i R_i$. Now set $R = (1/\alpha)\sum_{i\in I}\omega_i\alpha_i R_i$. Then R is nonexpansive and

$$\sum_{i\in I}\omega_i T_i = \sum_{i\in I}\omega_i(1-\alpha_i)\mathrm{Id} + \sum_{i\in I}\omega_i\alpha_i R_i = (1-\alpha)\mathrm{Id} + \alpha R. \tag{2.1}$$

We conclude that T is α-averaged. ☐

Remark 2.3 In view of [8, Corollary 2.2.17], Proposition 2.2 is equivalent to [8, Theorem 2.2.35], and it improves the averagedness constant of [11, Lemma 2.2(ii)] which was $\alpha = \max_{i\in I}\alpha_i$. In the case of two operators, Proposition 2.2 can be found in [16, Theorem 3(a)].

Next, we turn our attention to compositions of averaged operators, starting with the following result, which was obtained in [16, Theorem 3(b)] with a different proof.

Proposition 2.4 *Let D be a nonempty subset of $\mathcal{H}$, let $(\alpha_1,\alpha_2) \in \left]0,1\right[^2$, let $T_1 \colon D \to D$ be α_1-averaged, and let $T_2 \colon D \to D$ be α_2-averaged. Set*

$$T = T_1T_2 \quad \text{and} \quad \alpha = \frac{\alpha_1+\alpha_2-2\alpha_1\alpha_2}{1-\alpha_1\alpha_2}. \tag{2.2}$$

Then $\alpha \in \left]0,1\right[$ and T is α-averaged.

Proof. Since $\alpha_1(1-\alpha_2) < (1-\alpha_2)$, we have $\alpha_1+\alpha_2 < 1+\alpha_1\alpha_2$ and, therefore, $\alpha \in \left]0,1\right[$. Now let $x \in D$, let $y \in D$, and set

$$\tau = \frac{1-\alpha_1}{\alpha_1} + \frac{1-\alpha_2}{\alpha_2}. \tag{2.3}$$

It follows from Proposition 2.1 that

$$\begin{aligned}\|T_1T_2x - T_1T_2y\|^2 &\leqslant \|T_2x - T_2y\|^2 - \frac{1-\alpha_1}{\alpha_1}\|(\mathrm{Id}-T_1)T_2x - (\mathrm{Id}-T_1)T_2y\|^2 \\ &\leqslant \|x-y\|^2 - \frac{1-\alpha_2}{\alpha_2}\|(\mathrm{Id}-T_2)x - (\mathrm{Id}-T_2)y\|^2 \\ &\quad - \frac{1-\alpha_1}{\alpha_1}\|(\mathrm{Id}-T_1)T_2x - (\mathrm{Id}-T_1)T_2y\|^2. \end{aligned}\tag{2.4}$$

Moreover, by [6, Corollary 2.14], we have

$$\begin{aligned}&\frac{1-\alpha_1}{\tau\alpha_1}\|(\mathrm{Id}-T_1)T_2x - (\mathrm{Id}-T_1)T_2y\|^2 + \frac{1-\alpha_2}{\tau\alpha_2}\|(\mathrm{Id}-T_2)x - (\mathrm{Id}-T_2)y\|^2 \\ &\quad = \left\|\frac{1-\alpha_1}{\tau\alpha_1}\big((\mathrm{Id}-T_1)T_2x - (\mathrm{Id}-T_1)T_2y\big) - \frac{1-\alpha_2}{\tau\alpha_2}\big((\mathrm{Id}-T_2)x - (\mathrm{Id}-T_2)y\big)\right\|^2 \\ &\qquad + \frac{(1-\alpha_1)(1-\alpha_2)}{\tau^2\alpha_1\alpha_2}\|(x-y) - (T_1T_2x - T_1T_2y)\|^2 \\ &\quad \geqslant \frac{(1-\alpha_1)(1-\alpha_2)}{\tau^2\alpha_1\alpha_2}\|(\mathrm{Id}-T_1T_2)x - (\mathrm{Id}-T_1T_2)y\|^2. \end{aligned}\tag{2.5}$$

Combining (2.4), (2.5), and (2.2) yields

$$\begin{aligned}\|T_1T_2x - T_1T_2y\|^2 &\leqslant \|x-y\|^2 - \frac{(1-\alpha_1)(1-\alpha_2)}{\tau\alpha_1\alpha_2}\,\|(\mathrm{Id}\, - T_1T_2)\,x - (\mathrm{Id}\, - T_1T_2)\,y\|^2 \\ &= \|x-y\|^2 - \frac{1-\alpha_1-\alpha_2+\alpha_1\alpha_2}{\alpha_1+\alpha_2-2\alpha_1\alpha_2}\,\|(\mathrm{Id}\, - T_1T_2)\,x - (\mathrm{Id}\, - T_1T_2)\,y\|^2 \\ &= \|x-y\|^2 - \frac{1-\alpha}{\alpha}\,\|(\mathrm{Id}\, - T_1T_2)\,x - (\mathrm{Id}\, - T_1T_2)\,y\|^2 . \end{aligned} \tag{2.6}$$

In view of Proposition 2.1, we conclude that T is α-averaged. ☐

In [8, Theorem 2.2.37], the averagedness constant of (2.2) was written as

$$\alpha = \frac{1}{1+\dfrac{1}{\dfrac{\alpha_1}{1-\alpha_1}+\dfrac{\alpha_2}{1-\alpha_2}}}. \tag{2.7}$$

By induction, it leads to the following result for the composition of m averaged operators, which was obtained in [8] (combine [8, Theorem 2.2.42] and [8, Corollary 2.2.17]).

Proposition 2.5 *Let D be a nonempty subset of $\mathcal{H}$, let $m \geqslant 2$ be an integer, and set*

$$\phi\colon\ ]0,1[^m \to]0,1[\,:(\alpha_1,\dots,\alpha_m) \mapsto \frac{1}{1+\dfrac{1}{\displaystyle\sum_{i=1}^{m}\frac{\alpha_i}{1-\alpha_i}}}. \tag{2.8}$$

For every $i \in \{1,\dots,m\}$, let $\alpha_i \in]0,1[$ and let $T_i\colon D \to D$ be α_i-averaged. Set

$$T = T_1\cdots T_m \quad \text{and} \quad \alpha = \phi(\alpha_1,\dots,\alpha_m). \tag{2.9}$$

Then T is α-averaged.

Proof. We proceed by induction on $k \in \{2,\dots,m\}$. To this end, let us set $(\forall k \in \{2,\dots,m\})$ $\beta_k = [1+[\sum_{i=1}^k \alpha_i/(1-\alpha_i)]^{-1}]^{-1}$. By Proposition 2.4 and (2.7), the claim is true for $k=2$. Now assume that, for some $k \in \{2,\dots,m-1\}$, $T_1\cdots T_k$ is β_k-averaged. Then we deduce from Proposition 2.4 and (2.7) that the averagedness constant of $(T_1\cdots T_k)T_{k+1}$ is

$$\frac{1}{1+\dfrac{1}{\dfrac{1}{\beta_k^{-1}-1}+\dfrac{\alpha_{k+1}}{1-\alpha_{k+1}}}} = \frac{1}{1+\dfrac{1}{\left(\displaystyle\sum_{i=1}^{k}\frac{\alpha_i}{1-\alpha_i}\right)+\dfrac{\alpha_{k+1}}{1-\alpha_{k+1}}}} = \beta_{k+1}, \tag{2.10}$$

which concludes the induction argument. ☐

Remark 2.6 Let $m \geqslant 2$ be an integer, let ϕ be as in (2.8), let $(\alpha_i)_{1\leqslant i\leqslant m} \in]0,1[^m$, and let $(\sigma_j)_{1\leqslant j\leqslant m}$ the elementary symmetric polynomials in the variables $(\alpha_i)_{1\leqslant i\leqslant m}$, i.e.,

$$(\forall j \in \{1,\dots,m\}) \quad \sigma_j = \sum_{1\leqslant i_1<\cdots<i_j\leqslant m}\ \prod_{l=1}^{j}\alpha_{i_l}. \tag{2.11}$$

Then one shows by induction that $\phi(\alpha_1,\ldots,\alpha_m) = \big[\sum_{j=1}^m(-1)^{j-1}j\sigma_j\big]/\big[1+\sum_{j=2}^m(-1)^{j-1}(j-1)\sigma_j\big]$. Note also that (2.8) implies that

$$\phi(\alpha_1,\ldots,\alpha_m) > \cfrac{1}{1+\cfrac{1}{\max\limits_{1\leqslant i\leqslant m}\dfrac{\alpha_i}{1-\alpha_i}}} = \max_{1\leqslant i\leqslant m}\alpha_i. \tag{2.12}$$

Remark 2.7 Let us compare the averagedness constant of Proposition 2.5 with alternative ones. Set

$$\widetilde{\phi}\colon\ ]0,1[^m \to]0,1[\,:(\alpha_1,\ldots,\alpha_m)\mapsto \frac{m\,\max\{\alpha_1,\ldots,\alpha_m\}}{(m-1)\max\{\alpha_1,\ldots,\alpha_m\}+1}, \tag{2.13}$$

and let $(\alpha_i)_{1\leqslant i\leqslant m}\in\ ]0,1[^m$.

(i) The averagedness constant of Proposition 2.5 is sharper than that of [11, Lemma 2.2(iii)], namely

$$\phi(\alpha_1,\ldots,\alpha_m)\leqslant\widetilde{\phi}(\alpha_1,\ldots,\alpha_m). \tag{2.14}$$

(ii) $\phi(\alpha_1,\ldots,\alpha_m) = \widetilde{\phi}(\alpha_1,\ldots,\alpha_m)$ if $\alpha_1=\cdots=\alpha_m$ and, in particular, if all the operators are firmly nonexpansive, i.e., $\alpha_1=\cdots=\alpha_m=1/2$.

(iii) If $m=2$, the averagedness constant of Proposition 2.5 is strictly sharper than that of [19, Lemma 3.2], namely (see also [8, Remark 2.2.38])

$$\phi(\alpha_1,\alpha_2)<\widehat{\phi}(\alpha_1,\alpha_2),\quad\text{where}\quad \widehat{\phi}(\alpha_1,\alpha_2)=\alpha_1+\alpha_2-\alpha_1\alpha_2. \tag{2.15}$$

In addition, $\phi(\alpha_1,\alpha_1)=\widetilde{\phi}(\alpha_1,\alpha_1)<\widehat{\phi}(\alpha_1,\alpha_1)$ while, for $\alpha_1=3/4$ and $\alpha_2=1/8$, $\widehat{\phi}(\alpha_1,\alpha_2)=25/32<6/7=\widetilde{\phi}(\alpha_1,\alpha_2)$, which shows that $\widetilde{\phi}$ and $\widehat{\phi}$ cannot be compared in general.

Proof. (i): Combine [8, Theorem 2.2.42], and [8, Corollary 2.2.17].

(ii): Set $\beta_1=\delta_1=\alpha_1$ and

$$(\forall k\in\{2,\ldots,m\})\quad \begin{cases}\beta_k=\cfrac{1}{1+\cfrac{1}{\sum\limits_{i=1}^{k}\dfrac{\alpha_i}{1-\alpha_i}}},\\[2ex] \delta_k=\dfrac{k\,\max\{\alpha_1,\ldots,\alpha_k\}}{(k-1)\max\{\alpha_1,\ldots,\alpha_k\}+1}.\end{cases} \tag{2.16}$$

Then (2.16) yields

$$(\forall k\in\{1,\ldots,m\})\quad \delta_k=\frac{k\alpha_1}{(k-1)\alpha_1+1}. \tag{2.17}$$

Let us show by induction that

$$(\forall k\in\{1,\ldots,m\})\quad \beta_k=\delta_k. \tag{2.18}$$

We have $\beta_1 = \delta_1 = \alpha_1$. Next, suppose that, for some $k \in \{1, \ldots, m-1\}$, $\beta_k = \delta_k$. Then $\alpha_{k+1} = \alpha_1$, while (2.10) and (2.17) yield

$$\beta_{k+1} = \cfrac{1}{1 + \cfrac{1}{\cfrac{1}{\beta_k^{-1} - 1} + \cfrac{\alpha_1}{1 - \alpha_1}}} = \cfrac{1}{1 + \cfrac{1}{\cfrac{1}{\delta_k^{-1} - 1} + \cfrac{\alpha_1}{1 - \alpha_1}}} = \frac{(k+1)\alpha_1}{k\alpha_1 + 1} = \delta_{k+1}. \tag{2.19}$$

This establishes (2.18).

(iii): This inequality was already obtained in [8, Remark 2.2.38]. It follows from the fact that

$$\widehat{\phi}(\alpha_1, \alpha_2) - \phi(\alpha_1, \alpha_2) = \frac{\alpha_1 \alpha_2 (1 - \alpha_1)(1 - \alpha_2)}{1 - \alpha_1 \alpha_2} > 0. \tag{2.20}$$

The remaining assertions are easily verified. ☐

3 Algorithms

We present applications of the bounds discussed in Section 2 to fixed point algorithms. Henceforth, we denote the set of fixed points of an operator $T\colon \mathcal{H} \to \mathcal{H}$ by $\operatorname{Fix} T$.

As a direct application of Proposition 2.2 and Proposition 2.5, we first consider so-called "string-averaging" iterations, which involve a mix of compositions and convex combinations of operators. In the case of projection operators, such iterations go back to [9].

Proposition 3.1 *Let $(T_i)_{i \in I}$ be a finite family of nonexpansive operators from $\mathcal{H}$ to $\mathcal{H}$ such that $\bigcap_{i \in I} \operatorname{Fix} T_i \neq \varnothing$, and let $(\alpha_i)_{i \in I}$ be real numbers in $]0,1[$ such that, for every $i \in I$, T_i is α_i-averaged. Let p be a strictly positive integer, for every $k \in \{1, \ldots, p\}$ let m_k be a strictly positive integer and let $\omega_k \in]0,1]$, and suppose that $\mathrm{i}\colon \big\{(k,l) \bigm| k \in \{1, \ldots, p\}, l \in \{1, \ldots, m_k\}\big\} \to I$ is surjective and that $\sum_{k=1}^{p} \omega_k = 1$. Define*

$$T = \sum_{k=1}^{p} \omega_k T_{\mathrm{i}(k,1)} \cdots T_{\mathrm{i}(k,m_k)}. \tag{3.1}$$

Then the following hold:

(i) *Set*

$$\alpha = \sum_{k=1}^{p} \cfrac{\omega_k}{1 + \cfrac{1}{\displaystyle\sum_{i=1}^{m_k} \frac{\alpha_{\mathrm{i}(k,i)}}{1 - \alpha_{\mathrm{i}(k,i)}}}}. \tag{3.2}$$

Then T is α-averaged and $\operatorname{Fix} T = \bigcap_{i \in I} \operatorname{Fix} T_i$.

(ii) *Let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1/\alpha[$ such that $\sum_{n \in \mathbb{N}} \lambda_n (1/\alpha - \lambda_n) = +\infty$. Furthermore, let $x_0 \in \mathcal{H}$ and set*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n \big(T x_n - x_n\big). \tag{3.3}$$

Then $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\bigcap_{i \in I} \operatorname{Fix} T_i$.

Proof. (i): The α-averagedness of T follows from Propositions 2.2 and 2.5. The remaining assertions follow from [6, Proposition 4.34 and Corollary 4.37].

(ii): This follows from (i) and [6, Proposition 5.15(iii)]. ☐

Remark 3.2 Proposition 3.1 improves upon [6, Corollary 5.18], where the averagedness constant α of (3.2) was replaced by

$$\alpha' = \max_{1\leqslant k\leqslant p} \rho_k, \quad \text{with} \quad (\forall k \in \{1,\ldots,p\}) \quad \rho_k = \cfrac{m_k}{m_k - 1 + \cfrac{1}{\max\{\alpha_{\mathrm{i}(k,1)},\ldots,\alpha_{\mathrm{i}(k,m_k)}\}}}. \tag{3.4}$$

In view of Remarks 2.3 and 2.7(i), $\alpha' \geqslant \alpha$ and therefore α provides a larger range for the relaxation parameters $(\lambda_n)_{n\in\mathbb{N}}$.

The subsequent applications require the following technical fact.

Lemma 3.3 [17, Lemma 2.2.2] *Let $(\alpha_n)_{n\in\mathbb{N}}$, $(\beta_n)_{n\in\mathbb{N}}$, and $(\varepsilon_n)_{n\in\mathbb{N}}$ be sequences in $[0,+\infty[$ such that $\sum_{n\in\mathbb{N}} \varepsilon_n < +\infty$ and $(\forall n \in \mathbb{N})\ \alpha_{n+1} \leqslant \alpha_n - \beta_n + \varepsilon_n$. Then $(\alpha_n)_{n\in\mathbb{N}}$ converges and $\sum_{n\in\mathbb{N}} \beta_n < +\infty$.*

Next, we introduce a general iteration process for finding a common fixed point of a countable family of averaged operators which allows for approximate computations of the operator values.

Proposition 3.4 *For every $n \in \mathbb{N}$, let $\alpha_n \in\,]0,1[$, let $\lambda_n \in\,]0,1/\alpha_n[$, let $e_n \in \mathcal{H}$, and let $T_n \colon \mathcal{H} \to \mathcal{H}$ be an α_n-averaged operator. Suppose that $S = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n \neq \varnothing$ and that $\sum_{n\in\mathbb{N}} \lambda_n \|e_n\| < +\infty$. Let $x_0 \in \mathcal{H}$ and set, for every $n \in \mathbb{N}$,*

$$x_{n+1} = x_n + \lambda_n\big(T_n x_n + e_n - x_n\big). \tag{3.5}$$

Then the following hold:

(i) *Let $n \in \mathbb{N}$, let $x \in S$, and set $\nu = \sum_{k\in\mathbb{N}} \lambda_k \|e_k\| + 2\sup_{k\in\mathbb{N}} \|x_k - x\|$. Then $\nu < +\infty$ and*

$$\|x_{n+1} - x\|^2 \leqslant \|x_n + \lambda_n(T_n x_n - x_n) - x\|^2 + \nu\lambda_n\|e_n\| \tag{3.6}$$
$$\leqslant \|x_n - x\|^2 - \lambda_n(1/\alpha_n - \lambda_n)\|T_n x_n - x_n\|^2 + \nu\lambda_n\|e_n\|. \tag{3.7}$$

(ii) $\sum_{n\in\mathbb{N}} \lambda_n(1/\alpha_n - \lambda_n)\|T_n x_n - x_n\|^2 < +\infty$.

(iii) *$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in S if and only if every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in S. In this case, the convergence is strong if $\operatorname{int} S \neq \varnothing$.*

(iv) *$(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in S if and only if $\underline{\lim}\, d_S(x_n) = 0$.*

Proof. (i): Set

$$R_n = (1 - 1/\alpha_n)\mathrm{Id} + (1/\alpha_n)T_n \quad \text{and} \quad \mu_n = \alpha_n\lambda_n. \tag{3.8}$$

Then $\operatorname{Fix} R_n = \operatorname{Fix} T_n$ and, by Proposition 2.1, R_n is nonexpansive. Furthermore, (3.5) can be written as

$$x_{n+1} = x_n + \mu_n\big(R_n x_n - x_n\big) + \lambda_n e_n, \quad \text{where} \quad \mu_n \in\,]0,1[\,. \tag{3.9}$$

Now set $z_n = x_n + \mu_n(R_n x_n - x_n)$. Since $x \in \operatorname{Fix} R_n$ and R_n is nonexpansive, we have

$$\begin{aligned}\|z_n - x\| &= \|(1-\mu_n)(x_n - x) + \mu_n(R_n x_n - R_n x)\| \\ &\leqslant (1-\mu_n)\|x_n - x\| + \mu_n\|R_n x_n - R_n x\| \\ &\leqslant \|x_n - x\|. \end{aligned} \tag{3.10}$$

Hence, (3.9) yields

$$\begin{aligned}\|x_{n+1} - x\| &\leqslant \|z_n - x\| + \lambda_n\|e_n\| \qquad (3.11)\\ &\leqslant \|x_n - x\| + \lambda_n\|e_n\| \qquad (3.12)\end{aligned}$$

and, since $\sum_{k\in\mathbb{N}} \lambda_k\|e_k\| < +\infty$, it follows from Lemma 3.3 that

$$\nu = \sum_{k\in\mathbb{N}} \lambda_k\|e_k\| + 2\sup_{k\in\mathbb{N}}\|x_k - x\| < +\infty. \tag{3.13}$$

Moreover, using (3.11), (3.10), and [6, Corollary 2.14], we can write

$$\begin{aligned}\|x_{n+1} - x\|^2 &\leqslant \|z_n - x\|^2 + (2\|z_n - x\| + \lambda_n\|e_n\|)\lambda_n\|e_n\| \\ &\leqslant \|z_n - x\|^2 + (2\|x_n - x\| + \lambda_n\|e_n\|)\lambda_n\|e_n\| \\ &\leqslant \|(1-\mu_n)(x_n - x) + \mu_n(R_n x_n - x)\|^2 + \nu\lambda_n\|e_n\| \qquad (3.14)\\ &= (1-\mu_n)\|x_n - x\|^2 + \mu_n\|R_n x_n - x\|^2 \\ &\quad - \mu_n(1-\mu_n)\|R_n x_n - x_n\|^2 + \nu\lambda_n\|e_n\| \\ &= (1-\mu_n)\|x_n - x\|^2 + \mu_n\|R_n x_n - R_n x\|^2 \\ &\quad - \mu_n(1-\mu_n)\|R_n x_n - x_n\|^2 + \nu\lambda_n\|e_n\| \\ &\leqslant \|x_n - x\|^2 - \mu_n(1-\mu_n)\|R_n x_n - x_n\|^2 + \nu\lambda_n\|e_n\| \\ &= \|x_n - x\|^2 - \lambda_n(1/\alpha_n - \lambda_n)\|T_n x_n - x_n\|^2 + \nu\lambda_n\|e_n\| \qquad (3.15)\\ &\leqslant \|x_n - x\|^2 + \nu\lambda_n\|e_n\|. \qquad (3.16)\end{aligned}$$

Thus, (3.6) follows from (3.8) and (3.14), and (3.15) provides (3.7).

(ii): This follows from (3.7), (3.13), and Lemma 3.3.

(iii): The weak convergence statement follows from (3.13), (3.16), and [10, Theorem 3.8], while the strong convergence statement follows from [10, Proposition 3.10].

(iv): By [6, Corollary 4.15], the sets $(\operatorname{Fix} T_n)_{n\in\mathbb{N}}$ are closed, and so is therefore their intersection S. Hence, the result follows from (3.13), (3.16), (ii), and [10, Theorem 3.11]. □

The main result of this section is the following.

Theorem 3.5 *Let $\varepsilon \in \,]0, 1/2[$, let $m \geqslant 2$ be an integer, let $x_0 \in \mathcal{H}$, and define ϕ as in (2.8). For every $i \in \{1, \ldots, m\}$ and every $n \in \mathbb{N}$, let $\alpha_{i,n} \in \,]0, 1[$, let $T_{i,n} \colon \mathcal{H} \to \mathcal{H}$ be $\alpha_{i,n}$-averaged, and let $e_{i,n} \in \mathcal{H}$. For every $n \in \mathbb{N}$, let $\lambda_n \in \,]0, (1-\varepsilon)(1 + \varepsilon\phi(\alpha_{1,n}, \ldots, \alpha_{m,n}))/\phi(\alpha_{1,n}, \ldots, \alpha_{m,n})]$ and set*

$$x_{n+1} = x_n + \lambda_n\Big(T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\Big) + e_{1,n} - x_n\Big). \tag{3.17}$$

Suppose that

$$S = \bigcap_{n\in\mathbb{N}} \operatorname{Fix}(T_{1,n}\cdots T_{m,n}) \neq \varnothing \quad \textit{and} \quad (\forall i \in \{1,\ldots,m\}) \quad \sum_{n\in\mathbb{N}} \lambda_n\|e_{i,n}\| < +\infty, \tag{3.18}$$

and define

$$(\forall i \in \{1,\ldots,m\})(\forall n \in \mathbb{N}) \quad T_{i+,n} = \begin{cases} T_{i+1,n}\cdots T_{m,n}, & \text{if } i \neq m;\\ \mathrm{Id}\,, & \text{if } i = m. \end{cases} \tag{3.19}$$

Then the following hold:

(i) $\sum_{n\in\mathbb{N}} \lambda_n(1/\phi(\alpha_{1,n},\ldots,\alpha_{m,n}) - \lambda_n)\|T_{1,n}\cdots T_{m,n}x_n - x_n\|^2 < +\infty$.

(ii) $(\forall x \in S)\ \max\limits_{1\leqslant i\leqslant m} \sum\limits_{n\in\mathbb{N}} \dfrac{\lambda_n(1-\alpha_{i,n})}{\alpha_{i,n}} \|(\mathrm{Id} - T_{i,n})T_{i+,n}x_n - (\mathrm{Id} - T_{i,n})T_{i+,n}x\|^2 < +\infty$.

(iii) $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* S *if and only if every weak sequential cluster point of* $(x_n)_{n\in\mathbb{N}}$ *is in* S. *In this case, the convergence is strong if* $\operatorname{int} S \neq \varnothing$.

(iv) $(x_n)_{n\in\mathbb{N}}$ *converges strongly to a point in* S *if and only if* $\varliminf d_S(x_n) = 0$.

Proof. Let $n \in \mathbb{N}$ and let $x \in S$. We can rewrite (3.17) as an instance of (3.5), namely

$$x_{n+1} = x_n + \lambda_n\big(T_n x_n + e_n - x_n\big), \tag{3.20}$$

where

$$T_n = T_{1,n}\cdots T_{m,n} \tag{3.21}$$

and

$$e_n = T_{1,n}\bigg(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n + e_{m,n}) + e_{m-1,n}\cdots\big) + e_{2,n}\bigg) + e_{1,n} - T_{1,n}\cdots T_{m,n}x_n. \tag{3.22}$$

It follows from Proposition 2.5 that

$$T_n \text{ is } \alpha_n\text{-averaged, where } \alpha_n = \phi(\alpha_{1,n},\ldots,\alpha_{m,n}). \tag{3.23}$$

Since $\alpha_n \in \left]0,1\right[$,

$$\frac{(1-\varepsilon)(1+\varepsilon\alpha_n)}{\alpha_n} < \frac{(1-\varepsilon)(1+\varepsilon)}{\alpha_n} = \frac{1-\varepsilon^2}{\alpha_n} < \frac{1}{\alpha_n} \tag{3.24}$$

and therefore $\lambda_n \in \left]0,1/\alpha_n\right[$, as required in Proposition 3.4.

(i): Using the nonexpansiveness of the operators $(T_{i,n})_{1\leqslant i\leqslant m}$, we derive from (3.22) that

$$\begin{aligned}
\|e_n\| &\leqslant \|e_{1,n}\|+\\
&\quad \left\|T_{1,n}\Big(T_{2,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n+e_{m,n})+e_{m-1,n}\cdots\big)+e_{2,n}\Big)-T_{1,n}\cdots T_{m,n}x_n\right\|\\
&\leqslant \|e_{1,n}\|+\\
&\quad \left\|T_{2,n}\Big(T_{3,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n+e_{m,n})+e_{m-1,n}\cdots\big)+e_{3,n}\Big)+e_{2,n}-T_{2,n}\cdots T_{m,n}x_n\right\|\\
&\leqslant \|e_{1,n}\|+\|e_{2,n}\|+\\
&\quad \left\|T_{3,n}\Big(T_{4,n}\big(\cdots T_{m-1,n}(T_{m,n}x_n+e_{m,n})+e_{m-1,n}\cdots\big)+e_{4,n}\Big)+e_{3,n}-T_{3,n}\cdots T_{m,n}x_n\right\|\\
&\ \ \vdots\\
&\leqslant \sum_{i=1}^{m}\|e_{i,n}\|. \qquad (3.25)
\end{aligned}$$

Accordingly, (3.18) yields

$$\sum_{k\in\mathbb{N}}\lambda_k\|e_k\|<+\infty. \qquad (3.26)$$

Hence, we deduce from Proposition 3.4(i) that

$$\nu=\sum_{k\in\mathbb{N}}\lambda_k\|e_k\|+2\sup_{k\in\mathbb{N}}\|x_k-x\|<+\infty \qquad (3.27)$$

and from Proposition 3.4(ii) that

$$\sum_{k\in\mathbb{N}}\lambda_k\Big(\frac{1}{\alpha_k}-\lambda_k\Big)\|T_kx_k-x_k\|^2<+\infty. \qquad (3.28)$$

(ii): We derive from Proposition 2.1 that

$$(\forall i\in\{1,\ldots,m\})(\forall(u,v)\in\mathcal{H}^2)$$
$$\|T_{i,n}u-T_{i,n}v\|^2\leqslant\|u-v\|^2-\frac{1-\alpha_{i,n}}{\alpha_{i,n}}\|(\mathrm{Id}\,-T_{i,n})u-(\mathrm{Id}\,-T_{i,n})v\|^2. \qquad (3.29)$$

Using this inequality m times leads to

$$\begin{aligned}
\|T_nx_n-x\|^2&=\|T_{1,n}\cdots T_{m,n}x_n-T_{1,n}\cdots T_{m,n}x\|^2\\
&\leqslant\|x_n-x\|^2-\sum_{i=1}^{m}\frac{1-\alpha_{i,n}}{\alpha_{i,n}}\|(\mathrm{Id}\,-T_{i,n})T_{i+,n}x_n-(\mathrm{Id}\,-T_{i,n})T_{i+,n}x\|^2\\
&\leqslant\|x_n-x\|^2-\frac{\beta_n}{\lambda_n}, \qquad (3.30)
\end{aligned}$$

where

$$\beta_n = \lambda_n \max_{1\leqslant i\leqslant m}\left(\frac{1-\alpha_{i,n}}{\alpha_{i,n}}\|(\mathrm{Id}\,-T_{i,n})T_{i+,n}x_n-(\mathrm{Id}\,-T_{i,n})T_{i+,n}x\|^2\right). \tag{3.31}$$

Note also that

$$\begin{aligned}\lambda_n\leqslant\frac{(1-\varepsilon)(1+\varepsilon\alpha_n)}{\alpha_n}\quad&\Rightarrow\quad\lambda_n\leqslant\frac{1+\varepsilon\alpha_n}{(1+\varepsilon)\alpha_n}\\&\Leftrightarrow\quad\left(1+\frac{1}{\varepsilon}\right)\lambda_n\leqslant\frac{1}{\varepsilon\alpha_n}+1\\&\Leftrightarrow\quad\lambda_n-1\leqslant\frac{1}{\varepsilon}\left(\frac{1}{\alpha_n}-\lambda_n\right).\end{aligned} \tag{3.32}$$

Thus, Proposition 3.4(i), (3.20), and [6, Corollary 2.14] yield

$$\begin{aligned}\|x_{n+1}-x\|^2&\leqslant\|(1-\lambda_n)(x_n-x)+\lambda_n(T_nx_n-x)\|^2+\nu\lambda_n\|e_n\|\\&=(1-\lambda_n)\|x_n-x\|^2+\lambda_n\|T_nx_n-x\|^2+\lambda_n(\lambda_n-1)\|T_nx_n-x_n\|^2+\nu\lambda_n\|e_n\|\\&\leqslant(1-\lambda_n)\|x_n-x\|^2+\lambda_n\|T_nx_n-x\|^2+\varepsilon_n,\end{aligned} \tag{3.33}$$

where

$$\varepsilon_n=\frac{\lambda_n}{\varepsilon}\left(\frac{1}{\alpha_n}-\lambda_n\right)\|T_nx_n-x_n\|^2+\nu\lambda_n\|e_n\|. \tag{3.34}$$

On the one hand, it follows from (3.26), (3.27), and (3.28) that

$$\sum_{k\in\mathbb{N}}\varepsilon_k<+\infty. \tag{3.35}$$

On the other hand, combining (3.30) and (3.33), we obtain

$$\|x_{n+1}-x\|^2\leqslant\|x_n-x\|^2-\beta_n+\varepsilon_n. \tag{3.36}$$

Consequently, Lemma 3.3 implies that $\sum_{k\in\mathbb{N}}\beta_k<+\infty$.

(iii)–(iv): These follow from their counterparts in Proposition 3.4. □

Remark 3.6 Theorem 3.5 extends the results of [11, Section 3], where the relaxations parameters $(\lambda_n)_{n\in\mathbb{N}}$ cannot exceed 1. Since these parameters control the step-lengths of the algorithm, the proposed extension can result in significant accelerations.

4 Application to forward-backward splitting

The forward-backward algorithm is one of the most versatile and powerful algorithms for finding a zero of the sum of two maximally monotone operators (see [12, 13] and the references therein for historical background and recent developments). In [11], the first author showed that the theory of averaged nonexpansive operators provided a convenient setting for analyzing this algorithm. In this section, we exploit the results of Sections 2 and 3 to further extend this analysis and obtain a new version of the forward-backward algorithm with an extended relaxation range.

Let us recall a few facts about monotone set-valued operators and convex analysis [6]. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator. The domain, the graph, and the set of zeros of A are respectively defined by $\operatorname{dom} A = \big\{x \in \mathcal{H} \;\big|\; Ax \neq \varnothing\big\}$, $\operatorname{gra} A = \big\{(x,u) \in \mathcal{H} \times \mathcal{H} \;\big|\; u \in Ax\big\}$, and $\operatorname{zer} A = \big\{x \in \mathcal{H} \;\big|\; 0 \in Ax\big\}$. The inverse of A is $A^{-1}\colon \mathcal{H} \mapsto 2^{\mathcal{H}}\colon u \mapsto \big\{x \in \mathcal{H} \;\big|\; u \in Ax\big\}$, and the resolvent of A is

$$J_A = (\mathrm{Id} + A)^{-1}. \tag{4.1}$$

This operator is firmly nonexpansive if A is monotone, i.e.,

$$(\forall (x,y) \in \mathcal{H} \times \mathcal{H})(\forall (u,v) \in Ax \times Ay) \quad \langle x - y \mid u - v\rangle \geqslant 0, \tag{4.2}$$

and $\operatorname{dom} J_A = \mathcal{H}$ if, furthermore, A is maximally monotone, i.e., there exists no monotone operator $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ such that $\operatorname{gra} A \subset \operatorname{gra} B$ and $A \neq B$. We denote by $\Gamma_0(\mathcal{H})$ the class of proper lower semicontinuous convex functions $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$. Let $f \in \Gamma_0(\mathcal{H})$. For every $x \in \mathcal{H}$, $f + \|x - \cdot\|^2/2$ possesses a unique minimizer, which is denoted by $\operatorname{prox}_f x$. We have

$$\operatorname{prox}_f = J_{\partial f}, \quad \text{where} \quad \partial f\colon \mathcal{H} \to 2^{\mathcal{H}}\colon x \mapsto \big\{u \in \mathcal{H} \;\big|\; (\forall y \in \mathcal{H})\;\; \langle y - x \mid u\rangle + f(x) \leqslant f(y)\big\} \tag{4.3}$$

is the Moreau subdifferential of f.

We start with a specialization of Theorem 3.5 to $m = 2$.

Corollary 4.1 *Let* $\varepsilon \in \left]0, 1/2\right[$ *and let* $x_0 \in \mathcal{H}$. *For every* $n \in \mathbb{N}$, *let* $\alpha_{1,n} \in \left]0, 1/(1+\varepsilon)\right]$, *let* $\alpha_{2,n} \in \left]0, 1/(1+\varepsilon)\right]$, *let* $T_{1,n}\colon \mathcal{H} \to \mathcal{H}$ *be* $\alpha_{1,n}$*-averaged, let* $T_{2,n}\colon \mathcal{H} \to \mathcal{H}$ *be* $\alpha_{2,n}$*-averaged, let* $e_{1,n} \in \mathcal{H}$, *and let* $e_{2,n} \in \mathcal{H}$. *In addition, for every* $n \in \mathbb{N}$, *let*

$$\lambda_n \in \left[\varepsilon, \frac{(1-\varepsilon)(1+\varepsilon\phi_n)}{\phi_n}\right], \quad \textit{where} \quad \phi_n = \frac{\alpha_{1,n} + \alpha_{2,n} - 2\alpha_{1,n}\alpha_{2,n}}{1 - \alpha_{1,n}\alpha_{2,n}}, \tag{4.4}$$

and set

$$x_{n+1} = x_n + \lambda_n\Big(T_{1,n}\big(T_{2,n}x_n + e_{2,n}\big) + e_{1,n} - x_n\Big). \tag{4.5}$$

Suppose that

$$S = \bigcap_{n\in\mathbb{N}} \operatorname{Fix}(T_{1,n}T_{2,n}) \neq \varnothing, \quad \sum_{n\in\mathbb{N}} \lambda_n\|e_{1,n}\| < +\infty, \quad \textit{and} \quad \sum_{n\in\mathbb{N}} \lambda_n\|e_{2,n}\| < +\infty. \tag{4.6}$$

Then the following hold:

(i) $(\forall x \in S)\; \sum_{n\in\mathbb{N}} \|T_{1,n}T_{2,n}x_n - T_{2,n}x_n + T_{2,n}x - x\|^2 < +\infty$.

(ii) $(\forall x \in S)\; \sum_{n\in\mathbb{N}} \|T_{2,n}x_n - x_n - T_{2,n}x + x\|^2 < +\infty$.

(iii) $\sum_{n\in\mathbb{N}} \|T_{1,n}T_{2,n}x_n - x_n\|^2 < +\infty$.

(iv) *Suppose that every weak sequential cluster point of* $(x_n)_{n\in\mathbb{N}}$ *is in* S. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* S, *and the convergence is strong if* $\operatorname{int} S \neq \varnothing$.

(v) *Suppose that* $\underline{\lim}\, d_S(x_n) = 0$. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly to a point in* S.

Proof. For every $n \in \mathbb{N}$, since $\phi_n \in]0,1[$, $\varepsilon < 1-\varepsilon < (1-\varepsilon)(1/\phi_n+\varepsilon)$ and λ_n is therefore well defined in (4.4). Overall, the present setting is encompassed by that of Theorem 3.5 with $m=2$.

(i)–(ii): Let $x \in S$. We derive from Theorem 3.5(ii) with $m=2$ that

$$\begin{cases} \displaystyle\sum_{n\in\mathbb{N}} \frac{\lambda_n(1-\alpha_{1,n})}{\alpha_{1,n}} \|(\mathrm{Id}\,-T_{1,n})T_{2,n}x_n - (\mathrm{Id}\,-T_{1,n})T_{2,n}x\|^2 < +\infty \\ \displaystyle\sum_{n\in\mathbb{N}} \frac{\lambda_n(1-\alpha_{2,n})}{\alpha_{2,n}} \|(\mathrm{Id}\,-T_{2,n})x_n - (\mathrm{Id}\,-T_{2,n})x\|^2 < +\infty. \end{cases} \tag{4.7}$$

However, it follows from the assumptions that

$$(\forall n \in \mathbb{N}) \quad T_{1,n}T_{2,n}x = x, \quad \frac{\lambda_n(1-\alpha_{1,n})}{\alpha_{1,n}} \geqslant \varepsilon^2, \quad \text{and} \quad \frac{\lambda_n(1-\alpha_{2,n})}{\alpha_{2,n}} \geqslant \varepsilon^2. \tag{4.8}$$

Combining (4.7) and (4.8) yields the claims.

(iii): Let $x \in S$. Then, for every $n \in \mathbb{N}$,

$$\begin{aligned} \|T_{1,n}T_{2,n}x_n - x_n\|^2 &= \|(T_{1,n}T_{2,n}x_n - T_{2,n}x_n + T_{2,n}x - x) + (T_{2,n}x_n - x_n - T_{2,n}x + x)\|^2 \\ &\leqslant 2\|T_{1,n}T_{2,n}x_n - T_{2,n}x_n + T_{2,n}x - x\|^2 + 2\|T_{2,n}x_n - x_n - T_{2,n}x + x\|^2. \end{aligned} \tag{4.9}$$

Hence the claim follows from (i)–(ii).

(iv)–(v): These follow from Theorem 3.5(iii)–(iv). ☐

Definition 4.2 [1, Definition 2.3] An operator $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is *demiregular* at $x \in \operatorname{dom} A$ if, for every sequence $((x_n,u_n))_{n\in\mathbb{N}}$ in $\operatorname{gra} A$ and every $u \in Ax$ such that $x_n \rightharpoonup x$ and $u_n \to u$, we have $x_n \to x$.

Here are some examples of demiregular monotone operators.

Lemma 4.3 [1, Proposition 2.4] *Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be monotone and suppose that $x \in \operatorname{dom} A$. Then A is demiregular at x in each of the following cases:*

(i) *A is uniformly monotone at x, i.e., there exists an increasing function $\theta\colon [0,+\infty[\to [0,+\infty]$ that vanishes only at 0 such that $(\forall u \in Ax)(\forall (y,v) \in \operatorname{gra} A)$ $\langle x-y \mid u-v\rangle \geqslant \theta(\|x-y\|)$.*

(ii) *A is strongly monotone, i.e., there exists $\alpha \in]0,+\infty[$ such that $A - \alpha\mathrm{Id}$ is monotone.*

(iii) *J_A is compact, i.e., for every bounded set $C \subset \mathcal{H}$, the closure of $J_A(C)$ is compact. In particular, $\operatorname{dom} A$ is boundedly relatively compact, i.e., the intersection of its closure with every closed ball is compact.*

(iv) *$A\colon \mathcal{H} \to \mathcal{H}$ is single-valued with a single-valued continuous inverse.*

(v) *A is single-valued on $\operatorname{dom} A$ and $\mathrm{Id} - A$ is demicompact, i.e., for every bounded sequence $(x_n)_{n\in\mathbb{N}}$ in $\operatorname{dom} A$ such that $(Ax_n)_{n\in\mathbb{N}}$ converges strongly, $(x_n)_{n\in\mathbb{N}}$ admits a strong cluster point.*

(vi) *$A = \partial f$, where $f \in \Gamma_0(\mathcal{H})$ is uniformly convex at x, i.e., there exists an increasing function $\theta\colon [0,+\infty[\to [0,+\infty]$ that vanishes only at 0 such that*

$$(\forall \alpha \in]0,1[)(\forall y \in \operatorname{dom} f) \quad f\big(\alpha x+(1-\alpha)y\big)+\alpha(1-\alpha)\theta(\|x-y\|) \leqslant \alpha f(x)+(1-\alpha)f(y). \tag{4.10}$$

(vii) $A = \partial f$, *where* $f \in \Gamma_0(\mathcal{H})$ *and, for every* $\xi \in \mathbb{R}$, $\big\{x \in \mathcal{H} \;\big|\; f(x) \leqslant \xi\big\}$ *is boundedly compact.*

Our extended forward-backward splitting scheme can now be presented.

Proposition 4.4 *Let* $\beta \in \left]0,+\infty\right[$, *let* $\varepsilon \in \left]0, \min\{1/2, \beta\}\right[$, *let* $x_0 \in \mathcal{H}$, *let* $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ *be maximally monotone, and let* $B\colon \mathcal{H} \to \mathcal{H}$ *be* β*-cocoercive, i.e.,*

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \langle x - y \mid Bx - By \rangle \geqslant \beta\|Bx - By\|^2. \tag{4.11}$$

Furthermore, let $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon, 2\beta/(1+\varepsilon)]$, *and let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be sequences in* $\mathcal{H}$ *such that* $\sum_{n\in\mathbb{N}}\|a_n\| < +\infty$ *and* $\sum_{n\in\mathbb{N}}\|b_n\| < +\infty$. *Suppose that* $\operatorname{zer}(A+B) \neq \varnothing$ *and, for every* $n \in \mathbb{N}$, *let*

$$\lambda_n \in \left[\varepsilon, (1-\varepsilon)\left(2 + \varepsilon - \frac{\gamma_n}{2\beta}\right)\right] \tag{4.12}$$

and set

$$x_{n+1} = x_n + \lambda_n\Big(J_{\gamma_n A}\big(x_n - \gamma_n(Bx_n + b_n)\big) + a_n - x_n\Big). \tag{4.13}$$

Then the following hold:

(i) $\sum_{n\in\mathbb{N}}\|J_{\gamma_n A}(x_n - \gamma_n Bx_n) - x_n\|^2 < +\infty$.

(ii) *Let* $x \in \operatorname{zer}(A+B)$. *Then* $\sum_{n\in\mathbb{N}}\|Bx_n - Bx\|^2 < +\infty$.

(iii) $(x_n)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\operatorname{zer}(A+B)$.

(iv) *Suppose that one of the following is satisfied:*

  (a) A *is demiregular at every point in* $\operatorname{zer}(A+B)$.

  (b) B *is demiregular at every point in* $\operatorname{zer}(A+B)$.

  (c) $\operatorname{int} S \neq \varnothing$.

  Then $(x_n)_{n\in\mathbb{N}}$ *converges strongly to a point in* $\operatorname{zer}(A+B)$.

Proof. We are going to establish the results as an application of Corollary 4.1. Set

$$(\forall n \in \mathbb{N}) \;\; T_{1,n} = J_{\gamma_n A}, \quad T_{2,n} = \mathrm{Id} - \gamma_n B, \quad e_{1,n} = a_n, \quad \text{and} \quad e_{2,n} = -\gamma_n b_n. \tag{4.14}$$

Then, for every $n \in \mathbb{N}$, $T_{1,n}$ is $\alpha_{1,n}$-averaged with $\alpha_{1,n} = 1/2$ [6, Remark 4.24(iii) and Corollary 23.8] and $T_{2,n}$ is $\alpha_{2,n}$-averaged with $\alpha_{2,n} = \gamma_n/(2\beta)$ [6, Proposition 4.33]. Moreover, for every $n \in \mathbb{N}$,

$$\phi_n = \frac{\alpha_{1,n} + \alpha_{2,n} - 2\alpha_{1,n}\alpha_{2,n}}{1 - \alpha_{1,n}\alpha_{2,n}} = \frac{2\beta}{4\beta - \gamma_n} \tag{4.15}$$

and, therefore,

$$\lambda_n \in [\varepsilon, (1-\varepsilon)(1+\varepsilon\phi_n)/\phi_n]\,, \tag{4.16}$$

in conformity with (4.4). In turn, (2.12) yields

$$(\forall n \in \mathbb{N}) \quad \lambda_n \leqslant \frac{1}{\phi_n} + \varepsilon < \frac{1}{\alpha_{1,n}} + \varepsilon = 2 + \varepsilon. \tag{4.17}$$

Consequently,

$$\sum_{n\in\mathbb{N}} \lambda_n \|e_{1,n}\| = (2+\varepsilon) \sum_{n\in\mathbb{N}} \|a_n\| < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \lambda_n \|e_{2,n}\| \leqslant 2(2+\varepsilon)\beta \sum_{n\in\mathbb{N}} \|b_n\| < +\infty. \tag{4.18}$$

On the other hand, [6, Proposition 25.1(iv)] yields

$$(\forall n \in \mathbb{N}) \quad \text{zer}\,(A+B) = \text{Fix}\,(T_{1,n}T_{2,n}). \tag{4.19}$$

Altogether, $S = \text{zer}\,(A+B) \neq \varnothing$, (4.6) is satisfied, and (4.13) is an instance of (4.5).

(i): This is a consequence of Corollary 4.1(iii) and (4.14).

(ii): Corollary 4.1(ii) and (4.14) yield

$$\begin{aligned}\sum_{n\in\mathbb{N}} \|Bx_n - Bx\|^2 &= \sum_{n\in\mathbb{N}} \gamma_n^{-2} \|T_{2,n}x_n - x_n - T_{2,n}x + x\|^2 \\ &\leqslant \varepsilon^{-2} \sum_{n\in\mathbb{N}} \|T_{2,n}x_n - x_n - T_{2,n}x + x\|^2 \\ &< +\infty. \end{aligned} \tag{4.20}$$

(iii): Let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ and let $y \in \mathcal{H}$ be such that $x_{k_n} \rightharpoonup y$. In view of Corollary 4.1(iv), it remains to show that $y \in \text{zer}\,(A+B)$. We set

$$(\forall n \in \mathbb{N}) \quad y_n = J_{\gamma_n A}(x_n - \gamma_n Bx_n) \quad \text{and} \quad u_n = \frac{x_n - y_n}{\gamma_n} - Bx_n, \tag{4.21}$$

and note that

$$(\forall n \in \mathbb{N}) \quad u_n \in Ay_n. \tag{4.22}$$

We derive from (i) that $y_n - x_n \to 0$, hence $y_{k_n} \rightharpoonup y$. Now let $x \in \text{zer}\,(A+B)$. Then (ii) implies that $Bx_n \to Bx$, hence $u_n \to -Bx$. However, since (4.11) implies that B is maximally monotone [6, Example 20.28], it follows from the properties $x_{k_n} \rightharpoonup y$ and $Bx_{k_n} \to Bx$ that $By = Bx$ [6, Proposition 20.33(ii)]. Thus, $y_{k_n} \rightharpoonup y$ and $u_{k_n} \to -By$, and it therefore follows from (4.22) and [6, Proposition 20.33(ii)] that $-By \in Ay$, i.e., $y \in \text{zer}\,(A+B)$.

(iv): By (iii), there exists $x \in \text{zer}\,(A+B)$ such that $x_n \rightharpoonup x$. In addition, we derive from (4.21), (i), and (ii) that $y_n \rightharpoonup x$ and $u_n \to -Bx \in Ax$.

(iv)(a): Suppose that A is demiregular at x. Then (4.22) yields $y_n \to x$ and (i) implies that $x_n \to x$.

(iv)(b): Suppose that B is demiregular at x. Since $x_n \rightharpoonup x$ and $Bx_n \to Bx$ by (ii), we have $x_n \to x$.

(iv)(c): This follows from (iii) and Corollary 4.1(iv). □

Remark 4.5 Proposition 4.4 extends [11, Corollary 6.5] and [1, Theorem 2.8], which impose the additional assumption that the relaxation parameters $(\lambda_n)_{n\in\mathbb{N}}$ satisfy $(\forall n \in \mathbb{N})\ \lambda_n \leqslant 1$. By contrast, the relaxation range allowed in (4.12) can be an arbitrarily large interval in $]0,2[$ and the maximum relaxation is always strictly greater than 1.

Remark 4.6 In Proposition 4.4, the parameters $(\gamma_n)_{n\in\mathbb{N}}$ are allowed to vary at each iteration. Now suppose that they are restricted to a fixed value $\gamma \in]0, 2\beta[$. Then, as in (3.20), (4.13) reduces to $x_{n+1} = x_n + \lambda_n(Tx_n + e_n - x_n)$, where $T = J_{\gamma A}(\mathrm{Id} - \gamma B)$ is α-averaged and e_n is given by (3.22). In this special case, the weak convergence of $(x_n)_{n\in\mathbb{N}}$ to a zero of $A + B$ can be derived from Proposition 3.4(iii) applied with $T_n \equiv T$, $\alpha_n \equiv \alpha$, and $(\lambda_n)_{n\in\mathbb{N}}$ in $]0, 1/\alpha[$ satisfying $\sum_{n\in\mathbb{N}} \lambda_n(1/\alpha - \lambda_n) = +\infty$ (see also [6, Proposition 5.15(iii)]). This approach was proposed in [6, Theorem 25.8(i)] with the constant $\alpha = \widetilde{\phi}(1/2, \gamma/(2\beta)) = 1/(1/2 + \min\{1, \beta/\gamma\})$ of (2.13), and revisited in [14, Lemma 4.4] in the case of subdifferentials of convex functions with the sharper constant $\alpha = \phi(1/2, \gamma/(2\beta)) = 2\beta/(4\beta - \gamma)$ of [16, Theorem 3(b)] (see Remark 2.7).

Proposition 4.7 *Let $\beta \in]0, +\infty[$, let $\varepsilon \in]0, \min\{1/2, \beta\}[$, let $x_0 \in \mathcal{H}$, let $f \in \Gamma_0(\mathcal{H})$, let $g\colon \mathcal{H} \to \mathbb{R}$ be convex and differentiable with a $1/\beta$-Lipschitz gradient, and suppose that the set S of solutions to the problem*

$$\underset{x\in\mathcal{H}}{\text{minimize}} \quad f(x) + g(x) \tag{4.23}$$

is nonempty. Furthermore, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, 2\beta/(1+\varepsilon)]$, and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$. For every $n \in \mathbb{N}$, let $\lambda_n \in [\varepsilon, (1-\varepsilon)(2+\varepsilon-\gamma_n/(2\beta))]$ and set

$$x_{n+1} = x_n + \lambda_n\Big(\mathrm{prox}_{\gamma_n f}\big(x_n - \gamma_n(\nabla g(x_n) + b_n)\big) + a_n - x_n\Big). \tag{4.24}$$

Then the following hold:

(i) $\sum_{n\in\mathbb{N}} \|\mathrm{prox}_{\gamma_n f}(x_n - \gamma_n \nabla g(x_n)) - x_n\|^2 < +\infty$.

(ii) *Let $x \in S$. Then $\sum_{n\in\mathbb{N}} \|\nabla g(x_n) - \nabla g(x)\|^2 < +\infty$.*

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

(iv) *Suppose that ∂f or ∇g is demiregular at every point in S, or that $\mathrm{int}\, S \neq \varnothing$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to a point in S.*

Proof. Using the same arguments as in [6, Section 27.3], one shows that this is the specialization of Proposition 4.4 to the case when $A = \partial f$ and $B = \nabla g$. ☐

References

[1] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, *SIAM J. Control Optim.*, vol. 48, pp. 3246–3270, 2010.

[2] J.-B. Baillon and R. E. Bruck, Optimal rates of asymptotic regularity for averaged nonexpansive mappings, in: *Proc. of the Second International Conference on Fixed Point Theory and Applications,* (K. K. Tan, ed.), World Scientific Press, London, 1992, pp. 27–66.

[3] J.-B. Baillon, R. E. Bruck, and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups, *Houston J. Math.*, vol. 4, pp. 1–9, 1978.

[4] J.-B. Baillon, P. L. Combettes, and R. Cominetti, Asymptotic behavior of compositions of under-relaxed nonexpansive operators, *J. Dynam. Games,* vol. 1, pp. 331–346, 2014.

[5] H. H. Bauschke and J. M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Rev.,* vol. 38, pp. 367–426, 1996.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer, New York, 2011.

[7] C. L. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Problems,* vol. 20, pp. 103–120, 2004.

[8] A. Cegielski, *Iterative Methods for Fixed Point Problems in Hilbert Spaces,* Lecture Notes in Mathematics, vol. 2057. Springer, Heidelberg, 2012.

[9] Y. Censor, T. Elfving, and G. T. Herman, Averaging strings of sequential iterations for convex feasibility problems, in: *Inherently Parallel Algorithms for Feasibility and Optimization*, (D. Butnariu, Y. Censor, and S. Reich, eds.), pp. 101–113. Elsevier, New York, 2001.

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

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

[12] P. L. Combettes, Dinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued Var. Anal.,* vol. 18, pp. 373–404, 2010.

[13] P. L. Combettes and B. C. Vũ, Variable metric forward-backward splitting with applications to monotone inclusions in duality, *Optimization,* vol. 63, pp. 1289–1318, 2014.

[14] L. Condat, A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, *J. Optim. Theory Appl.,* vol. 158, pp. 460–479, 2013.

[15] G. Crombez, On averaged, firmly nonexpansive and quasi-nonexpansive operators, and more of that, *J. Nonlinear Convex Anal.,* vol. 8, pp. 277–288, 2007.

[16] N. Ogura and I. Yamada, Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping, *Numer. Funct. Anal. Optim.,* vol. 23, pp. 113–137, 2002.

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

[18] H. Raguet, J. Fadili, and G. Peyré, A generalized forward-backward splitting, *SIAM J. Imaging Sci.,* vol. 6, pp. 1199–1226, 2013.

[19] Y.-W. Wen, M. K. Ng, and W.-K. Ching, Iterative algorithms based on decoupling of deblurring and denoising for image restoration, *SIAM J. Sci. Comput.,* vol. 30, pp. 2655–2674, 2008.

[20] I. Yamada, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings, in: *Inherently Parallel Algorithms for Feasibility and Optimization*, (D. Butnariu, Y. Censor, and S. Reich, eds.), pp. 473–504. Elsevier, New York, 2001.

[21] I. Yamada and N. Ogura, Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings, *Numer. Funct. Anal. Optim.,* vol. 25, pp. 619–655, 2004.