A Geometric Framework for Stochastic Iterations*

Patrick L. Combettes[1] and Javier I. Madariaga[2]

[1]North Carolina State University
Department of Mathematics
Raleigh, NC 27695, USA
plc@math.ncsu.edu

[2]North Carolina State University
Department of Mathematics
Raleigh, NC 27695, USA
jimadari@ncsu.edu

Abstract This paper concerns models and convergence principles for dealing with stochasticity in a wide range of algorithms arising in nonlinear analysis and optimization in Hilbert spaces. It proposes a flexible geometric framework within which existing solution methods can be recast and improved, and new ones can be designed. Almost sure weak, strong, and linear convergence results are established in particular for stochastic fixed point iterations, the stochastic gradient descent method, and stochastic extrapolated parallel algorithms for feasibility problems. In these areas, the proposed algorithms exceed the features and convergence guarantees of the state of the art. Numerical applications to signal and image recovery are provided.

Keywords. Convex feasibility, convex optimization, monotone inclusion, random fixed point algorithm, stochastic iterations.

MSC2020. 47N10, 90C15, 90C48, 47J26, 62L20.

*Contact author: P. L. Combettes. Email: plc@math.ncsu.edu. This work was supported by the National Science Foundation under grant DMS-2513409.

§1. Introduction

The objective of this paper is to propose a general algorithmic framework and convergence principles for dealing with stochasticity in a broad class of algorithms arising in optimization and numerical nonlinear analysis. Throughout, H is a separable real Hilbert space and the underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$ is complete. We denote by $\mathsf{Z} \subset \mathsf{H}$ the set of solutions to the problem of interest and assume that it is nonempty and closed.

In the recent paper [18], we showed that a simple geometry underlies most deterministic monotone operator splitting algorithms and that, by exploiting this geometry, the convergence analysis of existing methods could be simplified and improved. It was also argued that this geometric framework provides a flexible template to create new algorithms. The basic idea is to construct the update at iteration n of a deterministic algorithm for finding a point in the solution set Z as a relaxed projection $\mathsf{x}_{n+1} = \mathsf{x}_n + \lambda_n(\mathrm{proj}_{\mathsf{H}_n} \mathsf{x}_n - \mathsf{x}_n)$ onto a half-space $\mathsf{H}_n = \{\mathsf{z} \in \mathsf{H} \mid \langle \mathsf{z} \mid \mathsf{t}_n^* \rangle_{\mathsf{H}} \leqslant \eta_n\}$ containing Z as follows (see Fig. 1(a)).

Algorithm 1.1. Let $\mathsf{x}_0 \in \mathsf{H}$ and iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\text{take } \mathsf{t}_n^* \in \mathsf{H} \text{ and } \eta_n \in \mathbb{R} \text{ such that } (\forall \mathsf{z} \in \mathsf{Z})\ \ \langle \mathsf{z} \mid \mathsf{t}_n^* \rangle_{\mathsf{H}} \leqslant \eta_n \\
\alpha_n = \begin{cases} \dfrac{\langle \mathsf{x}_n \mid \mathsf{t}_n^* \rangle_{\mathsf{H}} - \eta_n}{\|\mathsf{t}_n^*\|_{\mathsf{H}}^2} & \text{if } \ \langle \mathsf{x}_n \mid \mathsf{t}_n^* \rangle_{\mathsf{H}} > \eta_n; \\ 0, & \text{otherwise} \end{cases} \\
\mathsf{d}_n = \alpha_n \mathsf{t}_n^* \\
\text{take } \lambda_n \in \left]0, 2\right[\\
\mathsf{x}_{n+1} = \mathsf{x}_n - \lambda_n \mathsf{d}_n.
\end{array}
\right.
\end{array}
\tag{1.1}
$$

Our approach consists in extending the above geometric construction to a general stochastic environment by making the following changes at iteration n:

- The deterministic quantities t_n^* and η_n are replaced by random ones.
- A stochastic tolerance is added in the construction of the outer approximation.
- The relaxation parameter λ_n is now random and no longer restricted to the interval $]0, 2[$.

This leads to the following algorithmic scheme (see Section 2.1 for notation).

Algorithm 1.2. Let $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\mathcal{X}_n = \sigma(x_0, \ldots, x_n) \\
\text{take } t_n^* \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \text{ and } \eta_n \in L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R}) \text{ such that} \\
\begin{cases}
\dfrac{1_{[t_n^* \neq 0]} 1_{[\langle x_n | t_n^* \rangle_{\mathsf{H}} > \eta_n]} \eta_n}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^* = 0]}} \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R}); \\
\alpha_n = \dfrac{1_{[t_n^* \neq 0]} 1_{[\langle x_n | t_n^* \rangle_{\mathsf{H}} > \eta_n]} \big(\langle x_n \mid t_n^* \rangle_{\mathsf{H}} - \eta_n\big)}{\|t_n^*\|_{\mathsf{H}}^2 + 1_{[t_n^* = 0]}}; \\
(\forall \mathsf{z} \in \mathsf{Z})\ \ \langle \mathsf{z} \mid \mathsf{E}(\alpha_n t_n^* \mid \mathcal{X}_n) \rangle_{\mathsf{H}} \leqslant \mathsf{E}(\alpha_n \eta_n \mid \mathcal{X}_n) + \varepsilon_n(\cdot, \mathsf{z}) \ \mathsf{P}\text{-a.s.} \\
\qquad\qquad \text{where } \varepsilon_n(\cdot, \mathsf{z}) \in \left[0, +\infty\right[\ \mathsf{P}\text{-a.s.}
\end{cases} \\
d_n = \alpha_n t_n^* \\
\text{take } \lambda_n \in L^\infty(\Omega, \mathcal{F}, \mathsf{P}; \left]0, +\infty\right[) \\
x_{n+1} = x_n - \lambda_n d_n.
\end{array}
\right.
\end{array}
\tag{1.2}
$$

Implicitly, Algorithm 1.2 constructs a random outer approximation S_n to Z, namely

$$Z \subset S_n = \big\{ z \in H \mid \langle z \mid E(\alpha_n t_n^* \mid \mathcal{X}_n)\rangle_H \leqslant E(\alpha_n \eta_n \mid \mathcal{X}_n) + \varepsilon_n(\cdot, z) \big\} \ \text{P-a.s.} \tag{1.3}$$

and the update x_{n+1} is obtained by performing a relaxed projection of the current iterate x_n onto the simpler set

$$H_n = \big\{ z \in H \mid \langle z \mid t_n^* \rangle_H \leqslant \eta_n \big\}, \tag{1.4}$$

which is a random affine half-space. It should be noted that, while $Z \subset S_n$, the more restrictive inclusion $Z \subset H_n$ does not hold in general (see Fig. 1(b)). In terms of modeling, choosing t_n^* and η_n such that $Z \subset H_n$ would restrict the scope of the processes we intend to model, whereas the more general inclusion $Z \subset S_n$ offers more flexibility. Let us observe that, if $\varepsilon_n = 0$, S_n is also a random half-space. However, as the following example shows, projecting onto it is not judicious.

Example 1.3. For every $k \in \{1, \dots, p\}$, let C_k be a closed convex subset of H. Suppose that $Z = \bigcap_{k=1}^{p} C_k \neq \varnothing$ and implement Algorithm 1.2 with $\lambda_n = 1$, $\varepsilon_n = 0$, $t_n^* = x_n - \text{proj}_{C_{k}} x_n$, and $\eta_n = \langle \text{proj}_{C_{k}} x_n \mid t_n^* \rangle_H$, where the random variable k is uniformly distributed in $\{1, \dots, p\}$. Then $E(t_n^* \mid \mathcal{X}_n) = x_n - p^{-1} \sum_{k=1}^{p} \text{proj}_{C_k} x_n$ and therefore

$$\begin{aligned} Z &\subset \left\{ z \in H \;\middle|\; \sum_{k=1}^{p} \big\langle z - \text{proj}_{C_k} x_n \;\big|\; x_n - \text{proj}_{C_k} x_n \big\rangle_H \leqslant 0 \right\} \\ &= \big\{ z \in H \;\big|\; \big\langle z \,\big|\, E(t_n^* \mid \mathcal{X}_n) \big\rangle_H \leqslant E\big(\eta_n \,\big|\, \mathcal{X}_n\big) \big\} \\ &= S_n \ \ \text{P-a.s.} \end{aligned} \tag{1.5}$$

Thus, Algorithm 1.2 yields the random iteration process $x_{n+1} = \text{proj}_{C_{k}} x_n$ in which a single, randomly selected set is projected onto at iteration n. By contrast, projecting onto S_n would yield the barycentric projection method $x_{n+1} = p^{-1} \sum_{k=1}^{p} \text{proj}_{C_k} x_n$, which is deterministic and imposes the computation of all p projections at each iteration.

Another new feature of Algorithm 1.2 is that the relaxation parameters $(\lambda_n)_{n\in\mathbb{N}}$ are random. In addition, they need not be confined to the range $]0, 2[$ imposed in deterministic algorithms [5, 11, 14, 18, 25]. We call such relaxations *super relaxations*.

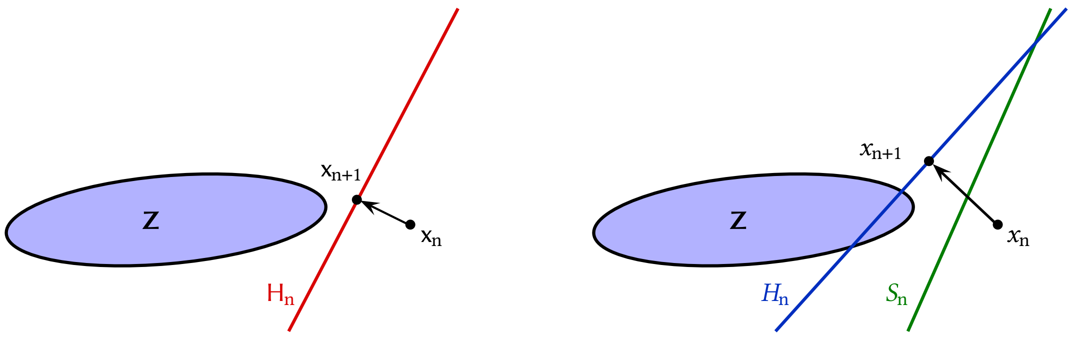

Figure 1: Geometry of algorithms for finding a point in Z with $\lambda_n = 1$. (a) Left: Iteration n of the deterministic Algorithm 1.1. (b) Right: Iteration n of the stochastic Algorithm 1.2 with $\varepsilon_n = 0$.

The deterministic setting of Algorithm 1.1 is known to capture a vast array of iterative methods in nonlinear analysis and optimization [18]. Our premise is that Algorithm 1.2 can serve the same purpose for their stochastic counterparts. Weak, strong, and linear convergence results will be established for Algorithm 1.2. In turn, these results will be applied to fixed point and feasibility problems, where they will be shown to provide new stochastic algorithms that go beyond the state of the art not only in terms of convergence guarantees but also of flexibility of implementation and scope.

The remainder of the paper is organized as follows. Notation and preliminary results are introduced in Section 2. The main theorems are presented in Section 3, where the asymptotic properties of Algorithm 1.2 are established. Section 4 is devoted to an application of the proposed theory to a randomly relaxed Krasnosel'skiĭ–Mann iteration process and includes new results on the convergence of the stochastic gradient method. Section 5 concerns an application to randomly activated and relaxed extrapolated fixed point methods for common fixed point problems in the presence of possibly uncountably many operators. These results significantly improve existing ones. Section 6 concludes the paper with applications to signal and image recovery. Applications of the results of Section 3 to the design and the analysis of stochastic splitting algorithms for monotone inclusion problems are addressed in the companion paper [19].

§2. Notation and background

2.1. Notation

We use sans-serif letters to denote deterministic variables and italicized serif letters to denote random variables.

The Hilbert space H has identity operator Id, scalar product $\langle\cdot\mid\cdot\rangle_{\mathsf{H}}$, and associated norm $\|\cdot\|_{\mathsf{H}}$. The symbols $\rightharpoonup$ and $\to$ denote weak and strong convergence in H, respectively. The sets of strong and weak sequential cluster points of a sequence $(\mathsf{x}_n)_{n\in\mathbb{N}}$ in H are denoted by $\mathfrak{S}(\mathsf{x}_n)_{n\in\mathbb{N}}$ and $\mathfrak{W}(\mathsf{x}_n)_{n\in\mathbb{N}}$, respectively. The distance function of a set $\mathsf{C}\subset\mathsf{H}$ is denoted by $\mathsf{d}_{\mathsf{C}}\colon\mathsf{x}\mapsto\inf_{\mathsf{y}\in\mathsf{C}}\|\mathsf{y}-\mathsf{x}\|_{\mathsf{H}}$ and the projection onto a nonempty closed convex set $\mathsf{C}\subset\mathsf{H}$ is denoted by $\mathrm{proj}_{\mathsf{C}}$. The fixed point set of an operator $\mathsf{T}\colon\mathsf{H}\to\mathsf{H}$ is $\operatorname{Fix}\mathsf{T}=\{\mathsf{x}\in\mathsf{H}\mid\mathsf{T}\mathsf{x}=\mathsf{x}\}$. The following notion will play an important role in Sections 4 and 5; see [2, Proposition 2.4] for examples of demiregular operators.

Definition 2.1. [2] $\mathsf{T}\colon\mathsf{H}\to\mathsf{H}$ is demiregular at $\mathsf{x}\in\mathsf{H}$ if, for every sequence $(\mathsf{x}_n)_{n\in\mathbb{N}}$ in H such that $\mathsf{x}_n\rightharpoonup\mathsf{x}$ and $\mathsf{T}\mathsf{x}_n\to\mathsf{T}\mathsf{x}$, we have $\mathsf{x}_n\to\mathsf{x}$.

Let $(\Xi,\mathcal{G})$ be a measurable space. A Ξ-valued random variable is a measurable mapping $x\colon(\Omega,\mathcal{F},\mathsf{P})\to(\Xi,\mathcal{G})$. Given $x\colon\Omega\to\Xi$ and $\mathsf{S}\in\mathcal{G}$, we set $[x\in\mathsf{S}]=\{\omega\in\Omega\mid x(\omega)\in\mathsf{S}\}$. Let x and y be random variables from $(\Omega,\mathcal{F},\mathsf{P})$ to $(\Xi,\mathcal{G})$. Then y is a copy of x if, for every $\mathsf{S}\in\mathcal{G}$, $\mathsf{P}([x\in\mathsf{S}])=\mathsf{P}([y\in\mathsf{S}])$. The Borel σ-algebra of H is denoted by $\mathcal{B}_{\mathsf{H}}$. An H-valued random variable is a measurable mapping $x\colon(\Omega,\mathcal{F})\to(\mathsf{H},\mathcal{B}_{\mathsf{H}})$. Let $\mathsf{p}\in[1,+\infty[$ and let $\mathcal{X}$ be a sub σ-algebra of $\mathcal{F}$. Then $L^{\mathsf{p}}(\Omega,\mathcal{X},\mathsf{P};\mathsf{H})$ denotes the space of equivalence classes of P-a.s. equal H-valued random variables $x\colon(\Omega,\mathcal{X},\mathsf{P})\to(\mathsf{H},\mathcal{B}_{\mathsf{H}})$ such that $\mathsf{E}\|x\|_{\mathsf{H}}^{\mathsf{p}}<+\infty$. Endowed with the norm

$$\|\cdot\|_{L^{\mathsf{p}}(\Omega,\mathcal{X},\mathsf{P};\mathsf{H})}\colon x\mapsto\mathsf{E}^{1/\mathsf{p}}\|x\|_{\mathsf{H}}^{\mathsf{p}}=\left(\int_\Omega\|x(\omega)\|_{\mathsf{H}}^{\mathsf{p}}\mathsf{P}(d\omega)\right)^{1/\mathsf{p}},\tag{2.1}$$

$L^{\mathsf{p}}(\Omega,\mathcal{X},\mathsf{P};\mathsf{H})$ is a real Banach space and $L^2(\Omega,\mathcal{X},\mathsf{P};\mathsf{H})$ is a real Hilbert space with scalar product

$$\langle\cdot\mid\cdot\rangle_{L^2(\Omega,\mathcal{X},\mathsf{P};\mathsf{H})}\colon(x,y)\mapsto\mathsf{E}\langle x\mid y\rangle_{\mathsf{H}}=\int_\Omega\big\langle x(\omega)\,\big|\,y(\omega)\big\rangle_{\mathsf{H}}\mathsf{P}(d\omega).\tag{2.2}$$

Further,

$$(\forall\mathsf{S}\in\mathcal{B}_{\mathsf{H}})\quad L^{\mathsf{p}}(\Omega,\mathcal{X},\mathsf{P};\mathsf{S})=\big\{x\in L^{\mathsf{p}}(\Omega,\mathcal{X},\mathsf{P};\mathsf{H})\;\big|\;x\in\mathsf{S}\ \mathsf{P}\text{-a.s.}\big\}.\tag{2.3}$$

The σ-algebra generated by a family Φ of random variables is denoted by $\sigma(\Phi)$. Let $\mathfrak{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ be a sequence of sub σ-algebras of $\mathcal{F}$ such that $(\forall n \in \mathbb{N})\ \mathcal{F}_n \subset \mathcal{F}_{n+1}$. Then $\ell_+(\mathfrak{F})$ is the set of sequences of $[0,+\infty[$-valued random variables $(\xi_n)_{n\in\mathbb{N}}$ such that, for every $n \in \mathbb{N}$, ξ_n is $\mathcal{F}_n$-measurable. We set

$$(\forall p \in \,]0,+\infty[) \quad \ell_+^p(\mathfrak{F}) = \left\{ (\xi_n)_{n\in\mathbb{N}} \in \ell_+(\mathfrak{F}) \;\middle|\; \sum_{n\in\mathbb{N}} \xi_n^p < +\infty \text{ P-a.s.} \right\}. \tag{2.4}$$

We say that $\varphi\colon \Omega \times H \to \mathbb{R}$ is a Carathéodory integrand if

$$\begin{cases} \text{for P-almost every } \omega \in \Omega, & \varphi(\omega,\cdot) \text{ is continuous;} \\ \text{for every } x \in H, & \varphi(\cdot,x) \text{ is } \mathcal{F}\text{-measurable.} \end{cases} \tag{2.5}$$

We denote by $\mathfrak{C}(\Omega,\mathcal{F},P;H)$ the class of Carathéodory integrands $\varphi\colon \Omega\times H \to [0,+\infty[$.

The reader is referred to [5] for background on convex analysis and fixed point theory, and to [31, 36] for background on probability in Hilbert spaces.

2.2. Preliminary results

Definition 2.2. Let $\mathcal{X}$ be a sub σ-algebra of $\mathcal{F}$, $C \in \mathcal{B}_H$, and x be an H-valued random variable. Then x is a C-valued $\mathcal{X}$-simple mapping if there exist a finite family of disjoint sets $(F_i)_{1\leqslant i\leqslant N}$ in $\mathcal{X}$ and a family of vectors $(z_i)_{1\leqslant i\leqslant N}$ in C such that

$$\bigcup_{i=1}^{N} F_i = \Omega \quad\text{and}\quad x = \sum_{i=1}^{N} 1_{F_i} z_i \quad \text{P-a.s.} \tag{2.6}$$

Remark 2.3. Let $p \in [1,+\infty[$. Then every C-valued $\mathcal{X}$-simple mapping is in $L^p(\Omega,\mathcal{X},P;C)$.

The following proposition is an adaptation of [31, Corollary 1.1.7].

Proposition 2.4. *Let* C *be a nonempty closed subset of* H, $\mathcal{X}$ *be a sub* σ*-algebra of* $\mathcal{F}$, $p \in [1,+\infty[$, *and* $x \in L^p(\Omega,\mathcal{X},P;C)$. *Then there exists a sequence* $(x_n)_{n\in\mathbb{N}}$ *of* C*-valued* $\mathcal{X}$*-simple mappings that converges strongly* P-a.s. *to* x *with* $\sup_{n\in\mathbb{N}} \|x_n\|_H^p \leqslant \|x\|_H^p + 1$ P-a.s.

Proof. Let $z \in C$ be such that $\|z\|_H^p \leqslant \inf_{y\in C} \|y\|_H^p + 1$ and let $\{z_n\}_{n\in\mathbb{N}}$ be a countable dense subset of C with $z_0 = z$. For every $n \in \mathbb{N}$ and every $y \in C$, define $I_{n,y} = \{ i \in \{0,\ldots,n\} \mid \|z_i\|_H^p \leqslant \|y\|_H^p + 1 \}$ and let $i_{n,y}$ be the smallest integer $i \in I_{n,y}$ such that $\|y - z_i\|_H = \min_{j\in I_{n,y}} \|y - z_j\|_H$. Define, for every $n \in \mathbb{N}$, $T_n\colon C \to C\colon y \mapsto z_{i_{n,y}}$. It follows from the density of $\{z_n\}_{n\in\mathbb{N}}$ in C that, for every $y \in C$, $T_n y \to y$ and

$$(\forall n \in \mathbb{N}) \quad \|T_n y\|_H^p \leqslant \|y\|_H^p + 1. \tag{2.7}$$

Set, for every $n \in \mathbb{N}$, $x_n = T_n x$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to x and

$$(\forall n \in \mathbb{N}) \quad \|x_n\|_H^p \leqslant \|x\|_H^p + 1 \text{ P-a.s.} \tag{2.8}$$

It remains to show that $(x_n)_{n\in\mathbb{N}}$ is a sequence of C-valued $\mathcal{X}$-simple mappings. Fix $n \in \mathbb{N}$. Then

$$[x_n = z_0] = \left\{ \omega \in \Omega \;\middle|\; \|x(\omega) - z_0\|_H = \min_{j\in I_{n,x(\omega)}} \|x(\omega) - z_j\|_H \right\} \tag{2.9}$$

and, for every $i \in \{1,\ldots,n\}$,

$$[x_n = z_i] = \left\{ \omega \in \Omega \;\middle|\; i \in I_{n,x(\omega)} \text{ and } \|x(\omega) - z_i\|_H = \min_{j\in I_{n,x(\omega)}} \|x(\omega) - z_j\|_H < \min_{j\in I_{i-1,x(\omega)}} \|x(\omega) - z_j\|_H \right\}. \tag{2.10}$$

By construction, (2.9) and (2.10) define sets in $\mathcal{X}$. Further,

$$\bigcup_{i=0}^{n}[x_n = z_i] = \Omega \quad \text{and} \quad x_n = \sum_{i=0}^{n} 1_{[x_n=z_i]} z_i, \tag{2.11}$$

which confirms that x_n is a C-valued $\mathcal{X}$-simple mapping. ☐

Lemma 2.5. *Let $\mathfrak{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ be a sequence of sub σ-algebras of $\mathcal{F}$ such that $(\forall n \in \mathbb{N})\ \mathcal{F}_n \subset \mathcal{F}_{n+1}$. Let $(\alpha_n)_{n\in\mathbb{N}} \in \ell_+(\mathfrak{F})$, $(\theta_n)_{n\in\mathbb{N}} \in \ell_+(\mathfrak{F})$, and $(\eta_n)_{n\in\mathbb{N}} \in \ell_+(\mathfrak{F})$. Then the following hold:*

(i) *Suppose that $(\eta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathfrak{F})$ and there exists a sequence $(\chi_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathfrak{F})$ satisfying*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}(\alpha_{n+1}\,|\,\mathcal{F}_n) + \theta_n \leqslant (1+\chi_n)\alpha_n + \eta_n \ \mathsf{P}\text{-a.s.} \tag{2.12}$$

Then $(\theta_n)_{n\in\mathbb{N}} \in \ell_+^1(\mathfrak{F})$ and $(\alpha_n)_{n\in\mathbb{N}}$ converges P-a.s. *to a $[0,+\infty[$-valued random variable.*

(ii) *Suppose that $\mathsf{E}\alpha_0 < +\infty$, $\sum_{n\in\mathbb{N}} \mathsf{E}\eta_n < +\infty$, and there exists a sequence $(\chi_n)_{n\in\mathbb{N}}$ in $[0,+\infty[$ satisfying $\varlimsup \chi_n < 1$ and*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}(\alpha_{n+1}\,|\,\mathcal{F}_n) + \theta_n \leqslant \chi_n\alpha_n + \eta_n \ \mathsf{P}\text{-a.s.} \tag{2.13}$$

Then $\sum_{n\in\mathbb{N}} \mathsf{E}\theta_n < +\infty$ and $\sum_{n\in\mathbb{N}} \mathsf{E}\alpha_n < +\infty$.

Proof. (i): This follows from [51, Theorem 1].
(ii): This follows from [22, Lemma 2.1(ii)]. ☐

Corollary 2.6. *Let $(\alpha_n)_{n\in\mathbb{N}}$, $(\theta_n)_{n\in\mathbb{N}}$, $(\eta_n)_{n\in\mathbb{N}}$, and $(\chi_n)_{n\in\mathbb{N}}$ be sequences in $[0,+\infty[$. Then the following hold:*

(i) *Suppose that $\sum_{n\in\mathbb{N}} \eta_n < +\infty$, $\sum_{n\in\mathbb{N}} \chi_n < +\infty$, and*

$$(\forall n \in \mathbb{N}) \quad \alpha_{n+1} + \theta_n \leqslant (1+\chi_n)\alpha_n + \eta_n. \tag{2.14}$$

Then $\sum_{n\in\mathbb{N}} \theta_n < +\infty$ and $(\alpha_n)_{n\in\mathbb{N}}$ converges to a positive real number.

(ii) *Suppose that $\sum_{n\in\mathbb{N}} \eta_n < +\infty$, $\varlimsup \chi_n < 1$, and*

$$(\forall n \in \mathbb{N}) \quad \alpha_{n+1} + \theta_n \leqslant \chi_n\alpha_n + \eta_n \ \mathsf{P}\text{-a.s.} \tag{2.15}$$

Then $\sum_{n\in\mathbb{N}} \theta_n < +\infty$ and $\sum_{n\in\mathbb{N}} \alpha_n < +\infty$.

Proof. An application of Lemma 2.5 with $(\forall n \in \mathbb{N})\ \mathcal{F}_n = \{\varnothing, \Omega\}$. ☐

Lemma 2.7. *Let $\xi \in L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$, let Φ be a family of random variables, set $\mathcal{X} = \sigma(\Phi)$, and let $\eta \in L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$ be independent of $\sigma(\{\xi\} \cup \Phi)$. Then $\mathsf{E}(\eta\xi\,|\,\mathcal{X}) = \mathsf{E}\eta\mathsf{E}(\xi\,|\,\mathcal{X})$.*

Proof. Note that $\mathcal{X} \subset \sigma(\{\xi\} \cup \Phi)$ and that ξ is $\sigma(\{\xi\} \cup \Phi)$-measurable. Hence, it follows from [54, Properties H*, K*, and J* in Section 2.7.4] that

$$\mathsf{E}(\eta\xi\,|\,\mathcal{X}) = \mathsf{E}\Big(\mathsf{E}\big(\eta\xi\,\big|\,\sigma(\{\xi\}\cup\Phi)\big)\,\Big|\,\mathcal{X}\Big) = \mathsf{E}\Big(\xi\mathsf{E}\big(\eta\,\big|\,\sigma(\{\xi\}\cup\Phi)\big)\,\Big|\,\mathcal{X}\Big) = \mathsf{E}\big(\xi\mathsf{E}\eta\,\big|\,\mathcal{X}\big) = \mathsf{E}\eta\mathsf{E}(\xi\,|\,\mathcal{X}), \tag{2.16}$$

which proves the identity. ☐

Lemma 2.8. *Let* $\boldsymbol{x} = (x_1, \ldots, x_{\mathsf{N}})$ *be an* H^{N}*-valued random variable, let* $(\mathsf{K}, \mathcal{K})$ *be a measurable space, and suppose that the random variable* $k\colon (\Omega, \mathcal{F}, \mathsf{P}) \to (\mathsf{K}, \mathcal{K})$ *is independent of* $\sigma(\boldsymbol{x})$*. Let* $\mathsf{f}\colon (\mathsf{K}\times\mathsf{H}, \mathcal{K}\otimes\mathcal{B}_{\mathsf{H}}) \to \mathbb{R}$ *be measurable and such that* $\mathsf{E}|\mathsf{f}(k, x_1)| < +\infty$*, and define* $\mathsf{g}\colon \mathsf{H} \to \mathbb{R}\colon \mathsf{x} \mapsto \mathsf{E}\mathsf{f}(k, \mathsf{x})$*. Then, for* P*-almost every* $\omega' \in \Omega$*,*

$$\mathsf{E}\big(\mathsf{f}(k, x_1)\,\big|\,\sigma(\boldsymbol{x})\big)(\omega') = \int_{\Omega} \mathsf{f}\big(k(\omega), x_1(\omega')\big)\mathsf{P}(d\omega) = \mathsf{g}\big(x_1(\omega')\big). \tag{2.17}$$

Proof. Define $\mathbf{f}\colon \mathsf{K}\times\mathsf{H}^{\mathsf{N}} \to \mathbb{R}\colon (\mathsf{k}, \mathbf{x}) \mapsto \mathsf{f}(\mathsf{k}, \mathsf{x}_1)$. Then $\mathbf{f}$ is an $\mathbb{R}$-valued measurable function. Let $\mathsf{S} \in \sigma(\boldsymbol{x})$. Then there exists $\mathsf{A} \in \bigotimes_{1\leqslant i\leqslant \mathsf{N}} \mathcal{B}_{\mathsf{H}}$ such that $\mathsf{S} = [\boldsymbol{x} \in \mathsf{A}]$. Thus, it follows from the image measure theorem [54, Theorem 2.6.7], the independence of k and $\sigma(\boldsymbol{x})$, and Fubini's theorem [54, Theorem 2.6.8] that

$$\begin{aligned}
\int_{\mathsf{S}} \mathsf{f}\big(k(\omega), x_1(\omega)\big)\mathsf{P}(d\omega) &= \int_{\Omega} \mathbf{f}\big(k(\omega), \boldsymbol{x}(\omega)\big)1_{\mathsf{A}}\big(\boldsymbol{x}(\omega)\big)\mathsf{P}(d\omega)\\
&= \int_{\mathsf{K}\times\mathsf{H}^{\mathsf{N}}} \mathbf{f}(\mathsf{k}, \mathbf{x})1_{\mathsf{A}}(\mathbf{x})\big(\mathsf{P}\circ(k, \boldsymbol{x})^{-1}\big)(d\mathsf{k}, d\mathbf{x})\\
&= \int_{\mathsf{K}\times\mathsf{H}^{\mathsf{N}}} \mathbf{f}(\mathsf{k}, \mathbf{x})1_{\mathsf{A}}(\mathbf{x})\big((\mathsf{P}\circ k^{-1}) \otimes (\mathsf{P}\circ \boldsymbol{x}^{-1})\big)(d\mathsf{k}, d\mathbf{x})\\
&= \int_{\mathsf{H}^{\mathsf{N}}} 1_{\mathsf{A}}(\mathbf{x})\left(\int_{\mathsf{K}} \mathbf{f}(\mathsf{k}, \mathbf{x})(\mathsf{P}\circ k^{-1})(d\mathsf{k})\right)(\mathsf{P}\circ \boldsymbol{x}^{-1})(d\mathbf{x})\\
&= \int_{\mathsf{H}^{\mathsf{N}}} 1_{\mathsf{A}}(\mathbf{x})\left(\int_{\mathsf{K}} \mathsf{f}(\mathsf{k}, \mathsf{x}_1)(\mathsf{P}\circ k^{-1})(d\mathsf{k})\right)(\mathsf{P}\circ \boldsymbol{x}^{-1})(d\mathbf{x})\\
&= \int_{\mathsf{H}^{\mathsf{N}}} 1_{\mathsf{A}}(\mathbf{x})\mathsf{g}(\mathsf{x}_1)(\mathsf{P}\circ \boldsymbol{x}^{-1})(d\mathbf{x})\\
&= \int_{\Omega} 1_{\mathsf{A}}\big(\boldsymbol{x}(\omega)\big)\mathsf{g}\big(x_1(\omega)\big)\mathsf{P}(d\omega)\\
&= \int_{\mathsf{S}} \mathsf{g}\big(x_1(\omega)\big)\mathsf{P}(d\omega).
\end{aligned} \tag{2.18}$$

Therefore $\mathsf{g}(x_1) = \mathsf{E}(\mathsf{f}(k, x_1)\,|\,\sigma(\boldsymbol{x}))$ P-a.s. ☐

Lemma 2.9. *Let* $\mathsf{p} \in\,]1, +\infty[$*, let* $(\xi_{\mathsf{n}})_{\mathsf{n}\in\mathbb{N}}$ *be a sequence in* $L^{\mathsf{p}}(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$ *such that* $\sup_{\mathsf{n}\in\mathbb{N}} \mathsf{E}|\xi_{\mathsf{n}}|^{\mathsf{p}} < +\infty$*, and let* $\xi\colon \Omega \to \mathbb{R}$ *be measurable. Suppose that* $\xi_{\mathsf{n}} \to \xi$ P*-a.s. Then* $\mathsf{E}|\xi| < +\infty$*,* $\xi_{\mathsf{n}} \to \xi$ *in* $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$*, and* $\mathsf{E}\xi_{\mathsf{n}} \to \mathsf{E}\xi$*.*

Proof. Set $\mathsf{q} = (\mathsf{p}-1)/\mathsf{p}$. It follows from the Hölder and Markov inequalities that

$$\begin{aligned}
0 &\leqslant \lim_{\beta\to+\infty} \sup_{\mathsf{n}\in\mathbb{N}} \int_{[|\xi_{\mathsf{n}}|\geqslant\beta]} |\xi_{\mathsf{n}}|d\mathsf{P}\\
&\leqslant \lim_{\beta\to+\infty} \sup_{\mathsf{n}\in\mathbb{N}} \Big(\mathsf{E}^{1/\mathsf{p}}|\xi_{\mathsf{n}}|^{\mathsf{p}}\mathsf{E}^{1/\mathsf{q}}1^{\mathsf{q}}_{[|\xi_{\mathsf{n}}|\geqslant\beta]}\Big)\\
&\leqslant \sup_{\mathsf{n}\in\mathbb{N}} \mathsf{E}^{1/\mathsf{p}}|\xi_{\mathsf{n}}|^{\mathsf{p}} \lim_{\beta\to+\infty} \sup_{\mathsf{n}\in\mathbb{N}} \Big(\mathsf{P}\big([|\xi_{\mathsf{n}}| \geqslant \beta]\big)\Big)^{1/\mathsf{q}}\\
&\leqslant \sup_{\mathsf{n}\in\mathbb{N}} \mathsf{E}^{1/\mathsf{p}}|\xi_{\mathsf{n}}|^{\mathsf{p}} \lim_{\beta\to+\infty} \sup_{\mathsf{n}\in\mathbb{N}} \frac{\mathsf{E}^{1/\mathsf{q}}|\xi_{\mathsf{n}}|^{\mathsf{p}}}{\beta^{\mathsf{p}/\mathsf{q}}}\\
&= 0,
\end{aligned} \tag{2.19}$$

which shows that $(\xi_{\mathsf{n}})_{\mathsf{n}\in\mathbb{N}}$ is uniformly integrable. We can therefore invoke [54, Theorem 2.6.4(b)], which asserts that $\xi \in L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$, $\mathsf{E}\xi_{\mathsf{n}} \to \mathsf{E}\xi$, and $\xi_{\mathsf{n}} \to \xi$ in $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$. ☐

Lemma 2.10. [31, Proposition 2.6.31] *Let* $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$, *let* $\mathcal{X}$ *be a sub* σ*-algebra of* $\mathcal{F}$, *and let* $y \in L^2(\Omega, \mathcal{X}, \mathsf{P}; \mathsf{H})$. *Then* $\mathsf{E}(\langle x \mid y\rangle_{\mathsf{H}} \,|\, \mathcal{X}) = \langle \mathsf{E}(x \mid \mathcal{X}) \,|\, y\rangle_{\mathsf{H}}$.

Lemma 2.11. *Let* C *be a nonempty closed subset of* H *and let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence of* H*-valued random variables. Define*

$$\mathfrak{X} = (\mathcal{X}_n)_{n\in\mathbb{N}}, \quad \textit{where} \quad (\forall n \in \mathbb{N}) \;\; \mathcal{X}_n = \sigma(x_0, \ldots, x_n). \tag{2.20}$$

Let $p \in [1, +\infty[$ *and suppose that, for every* $z \in \mathsf{C}$, *there exist* $(\mu_n(z))_{n\in\mathbb{N}} \in \ell_+^1(\mathfrak{X})$, $(\theta_n(z))_{n\in\mathbb{N}} \in \ell_+(\mathfrak{X})$, *and* $(\nu_n(z))_{n\in\mathbb{N}} \in \ell_+^1(\mathfrak{X})$ *such that*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}(\|x_{n+1} - z\|_{\mathsf{H}}^p \,|\, \mathcal{X}_n) + \theta_n(z) \leqslant (1 + \mu_n(z))\|x_n - z\|_{\mathsf{H}}^p + \nu_n(z) \;\; \mathsf{P}\text{-a.s.} \tag{2.21}$$

Then the following hold:

(i) *Let* $z \in \mathsf{C}$. *Then* $\sum_{n\in\mathbb{N}} \theta_n(z) < +\infty$ P-a.s.

(ii) $(\|x_n\|_{\mathsf{H}})_{n\in\mathbb{N}}$ *is bounded* P-a.s.

(iii) $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \neq \varnothing$ P-a.s.

(iv) *There exists* $\Omega' \in \mathcal{F}$ *such that* $\mathsf{P}(\Omega') = 1$ *and, for every* $\omega \in \Omega'$ *and every* $z \in \mathsf{C}$, $(\|x_n(\omega) - z\|_{\mathsf{H}})_{n\in\mathbb{N}}$ *converges.*

(v) *Suppose that* $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{C}$ P-a.s. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to a* C*-valued random variable.*

(vi) *Suppose that* $\mathfrak{S}(x_n)_{n\in\mathbb{N}} \cap \mathsf{C} \neq \varnothing$ P-a.s. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to a* C*-valued random variable.*

(vii) *Suppose that* $\mathfrak{S}(x_n)_{n\in\mathbb{N}} \neq \varnothing$ P-a.s. *and that* $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{C}$ P-a.s. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to an* C*-valued random variable.*

(viii) *Suppose that* $z \in \mathsf{C}$ *and* $(\chi_n)_{n\in\mathbb{N}}$ *in* $[0, +\infty[$ *satisfy*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}(\|x_{n+1} - z\|_{\mathsf{H}}^p \,|\, \mathcal{X}_n) \leqslant \chi_n \|x_n - z\|_{\mathsf{H}}^p \;\; \mathsf{P}\text{-a.s.} \;\; \textit{and} \;\; \overline{\lim}\, \chi_n < 1. \tag{2.22}$$

Then the following hold:

(a) *Let* $n \in \mathbb{N}$. *Then* $\mathsf{E}(\|x_{n+1} - z\|_{\mathsf{H}}^p \,|\, \mathcal{X}_0) \leqslant (\prod_{j=0}^{n} \chi_j)\|x_0 - z\|_{\mathsf{H}}^p$ P-a.s.

(b) *Suppose that* $x_0 \in L^p(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^p(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and* P-a.s. *to* z.

Proof. (i)-(vii): Apply [20, Proposition 2.3] with $\phi = |\cdot|^p$. The measurability of the weak limit in (v) relies on [20, Proposition 2.3(iv)] which invokes [47, Corollary 1.13]. The applicability of the latter follows from the separability of H and the completeness of $(\Omega, \mathcal{F}, \mathsf{P})$; see [31, Sections 1.1a–b] for details.
(viii): Apply [22, Lemma 2.2] with $\phi = |\cdot|^p$. ☐

§3. Main results

3.1. An abstract stochastic algorithm

The analysis of the asymptotic behavior of the following algorithm will serve as the backbone of subsequent convergence results. We recall from Section 1 that Z is the solution set of the problem under consideration and that it is assumed to be nonempty and closed.

Algorithm 3.1. Let $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. Iterate

$$
\begin{array}{l}
\text{for } \mathsf{n} = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\mathcal{X}_\mathsf{n} = \sigma(x_0, \ldots, x_\mathsf{n}) \\
\text{take } \lambda_\mathsf{n} \in L^\infty(\Omega, \mathcal{F}, \mathsf{P};]0, +\infty[),\; d_\mathsf{n} \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}),\; \text{and } \delta_\mathsf{n} \in \mathfrak{C}(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \text{ such that} \\
\qquad \begin{cases}
\mathsf{E}\big(\lambda_\mathsf{n}(2 - \lambda_\mathsf{n})\|d_\mathsf{n}\|_\mathsf{H}^2 \,\big|\, \mathcal{X}_\mathsf{n}\big) \geqslant 0 \;\; \mathsf{P}\text{-a.s.}; \\
(\forall \mathsf{z} \in \mathsf{Z}) \;\; \mathsf{E}\big(\lambda_\mathsf{n}\langle \mathsf{z} + d_\mathsf{n} - x_\mathsf{n} \mid d_\mathsf{n}\rangle_\mathsf{H} \,\big|\, \mathcal{X}_\mathsf{n}\big) \leqslant \delta_\mathsf{n}(\cdot, \mathsf{z})/2 \;\; \mathsf{P}\text{-a.s.}
\end{cases} \\
x_{\mathsf{n}+1} = x_\mathsf{n} - \lambda_\mathsf{n} d_\mathsf{n}.
\end{array}
\right.
\end{array}
\tag{3.1}
$$

Let us outline the weak and strong convergence properties of Algorithm 3.1.

Theorem 3.2. *Let $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ be the sequence generated by Algorithm 3.1. Then the following hold:*

(i) *$(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ is a well-defined sequence in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$.*

(ii) *Let $\mathsf{n} \in \mathbb{N}$ and $\mathsf{z} \in \mathsf{Z}$. Then*

$$
\mathsf{E}\big(\|x_{\mathsf{n}+1} - \mathsf{z}\|_\mathsf{H}^2 \,\big|\, \mathcal{X}_\mathsf{n}\big) \leqslant \|x_\mathsf{n} - \mathsf{z}\|_\mathsf{H}^2 - \mathsf{E}\big(\lambda_\mathsf{n}(2 - \lambda_\mathsf{n})\|d_\mathsf{n}\|_\mathsf{H}^2 \,\big|\, \mathcal{X}_\mathsf{n}\big) + \delta_\mathsf{n}(\cdot, \mathsf{z}) \;\; \mathsf{P}\text{-a.s.}
$$

(iii) *Let $\mathsf{n} \in \mathbb{N}$ and $z \in L^2(\Omega, \mathcal{X}_\mathsf{n}, \mathsf{P}; \mathsf{Z})$. Then*

$$
\mathsf{E}\big(\|x_{\mathsf{n}+1} - z\|_\mathsf{H}^2 \,\big|\, \mathcal{X}_\mathsf{n}\big) \leqslant \|x_\mathsf{n} - z\|_\mathsf{H}^2 - \mathsf{E}\big(\lambda_\mathsf{n}(2 - \lambda_\mathsf{n})\|d_\mathsf{n}\|_\mathsf{H}^2 \,\big|\, \mathcal{X}_\mathsf{n}\big) + \delta_\mathsf{n}(\cdot, z) \;\; \mathsf{P}\text{-a.s.}
$$

(iv) *Let $\mathsf{n} \in \mathbb{N}$ and $z \in L^2(\Omega, \mathcal{X}_\mathsf{n}, \mathsf{P}; \mathsf{Z})$. Then*

$$
\|x_{\mathsf{n}+1} - z\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})}^2 \leqslant \|x_\mathsf{n} - z\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})}^2 - \mathsf{E}\big(\lambda_\mathsf{n}(2 - \lambda_\mathsf{n})\|d_\mathsf{n}\|_\mathsf{H}^2\big) + \mathsf{E}\delta_\mathsf{n}(\cdot, z).
$$

(v) *Suppose that, for every $\mathsf{z} \in \mathsf{Z}$, $\sum_{\mathsf{n}\in\mathbb{N}} \delta_\mathsf{n}(\cdot, \mathsf{z}) < +\infty$ P-a.s. Then the following hold:*

  (a) *$(\|x_\mathsf{n}\|_\mathsf{H})_{\mathsf{n}\in\mathbb{N}}$ is bounded P-a.s.*

  (b) *Let z be a Z-valued random variable. Then $(\|x_\mathsf{n} - z\|_\mathsf{H})_{\mathsf{n}\in\mathbb{N}}$ converges P-a.s.*

  (c) *$\sum_{\mathsf{n}\in\mathbb{N}} \mathsf{E}(\lambda_\mathsf{n}(2 - \lambda_\mathsf{n})\|d_\mathsf{n}\|_\mathsf{H}^2 \mid \mathcal{X}_\mathsf{n}) < +\infty$ P-a.s.*

  (d) *Suppose that $\mathfrak{W}(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. Then $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges weakly P-a.s. to a Z-valued random variable.*

  (e) *Suppose that $\mathfrak{S}(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}} \cap \mathsf{Z} \neq \varnothing$ P-a.s. Then $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges strongly P-a.s. to a Z-valued random variable.*

  (f) *Suppose that $\mathfrak{S}(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}} \neq \varnothing$ P-a.s. and that $\mathfrak{W}(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. Then $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges strongly P-a.s. to a Z-valued random variable.*

(vi) *Suppose that, for every $\mathsf{z} \in \mathsf{Z}$, $\sum_{\mathsf{n}\in\mathbb{N}} \mathsf{E}\delta_\mathsf{n}(\cdot, \mathsf{z}) < +\infty$. Then the following hold:*

  (a) *$(\|x_\mathsf{n}\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{\mathsf{n}\in\mathbb{N}}$ is bounded.*

  (b) *Let $z \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$. Then $(\|x_\mathsf{n} - z\|_{L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{\mathsf{n}\in\mathbb{N}}$ converges.*

  (c) *$\sum_{\mathsf{n}\in\mathbb{N}} \mathsf{E}(\lambda_\mathsf{n}\,(2 - \lambda_\mathsf{n})\,\|d_\mathsf{n}\|_\mathsf{H}^2) < +\infty$.*

  (d) *Suppose that $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges weakly P-a.s. to an H-valued random variable x. Then $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges weakly in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ to x.*

  (e) *Let x be a Z-valued random variable. Then $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges strongly P-a.s. to x if and only if $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges strongly in $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ to x. In this case, $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$ and $(x_\mathsf{n})_{\mathsf{n}\in\mathbb{N}}$ converges weakly in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ to x.*

Proof. (i): By assumption, $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. Now suppose that $x_n \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. Then, since $d_n \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and $\lambda_n \in L^\infty(\Omega, \mathcal{F}, \mathsf{P};]0, +\infty[)$, $x_{n+1} = x_n - \lambda_n d_n \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. This establishes the claim by induction.

(ii): We derive from (3.1) that

$$\begin{aligned}
&\mathsf{E}\big(\|x_{n+1} - \mathsf{z}\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) \\
&\quad= \mathsf{E}\big(\|x_n - \mathsf{z}\|_{\mathsf{H}}^2 - 2\lambda_n \langle x_n - \mathsf{z} \mid d_n\rangle_{\mathsf{H}} + \lambda_n^2 \|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) \\
&\quad= \|x_n - \mathsf{z}\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + 2\mathsf{E}\big(\lambda_n \langle \mathsf{z} + d_n - x_n \mid d_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) \\
&\quad\leqslant \|x_n - \mathsf{z}\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \delta_n(\cdot, \mathsf{z}) \quad \mathsf{P}\text{-a.s.}
\end{aligned} \tag{3.2}$$

(iii): First, let s be a Z-valued $\mathcal{X}_n$-simple mapping. Then there exists a finite family of disjoint sets $(\mathsf{F}_i)_{i\in I}$ in $\mathcal{X}_n$ such that $\bigcup_{i\in I} \mathsf{F}_i = \Omega$, and a family $(\mathsf{z}_i)_{i\in I}$ in Z such that $s = \sum_{i\in I} 1_{\mathsf{F}_i} \mathsf{z}_i$. Then, by (ii),

$$\begin{aligned}
\mathsf{E}\big(\|x_{n+1} - s\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) &= \mathsf{E}\left(\left\|\sum_{i\in I} 1_{\mathsf{F}_i}(x_{n+1} - \mathsf{z}_i)\right\|_{\mathsf{H}}^2 \,\middle|\, \mathcal{X}_n\right) \\
&= \mathsf{E}\left(\sum_{i\in I} \|x_{n+1} - \mathsf{z}_i\|_{\mathsf{H}}^2 1_{\mathsf{F}_i} \,\middle|\, \mathcal{X}_n\right) \\
&= \sum_{i\in I} \mathsf{E}\big(\|x_{n+1} - \mathsf{z}_i\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) 1_{\mathsf{F}_i} \\
&\leqslant \sum_{i\in I} \|x_n - \mathsf{z}_i\|_{\mathsf{H}}^2 1_{\mathsf{F}_i} + \sum_{i\in I}\Big(-\mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \delta_n(\cdot, \mathsf{z}_i)\Big) 1_{\mathsf{F}_i} \\
&= \left\|\sum_{i\in I} 1_{\mathsf{F}_i}(x_n - \mathsf{z}_i)\right\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \sum_{i\in I} \delta_n(\cdot, \mathsf{z}_i) 1_{\mathsf{F}_i} \\
&= \|x_n - s\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \delta_n(\cdot, s) \quad \mathsf{P}\text{-a.s.}
\end{aligned} \tag{3.3}$$

Next, Proposition 2.4 guarantees the existence of a sequence of Z-valued $\mathcal{X}_n$-simple mappings $(s_j)_{j\in\mathbb{N}}$ such that $s_j \to z$ P-a.s. and $\sup_{j\in\mathbb{N}} \|s_j\|_{\mathsf{H}}^2 \leqslant \|z\|_{\mathsf{H}}^2 + 1$ P-a.s. Thus, we derive from (3.3) that

$$(\forall j \in \mathbb{N}) \quad \mathsf{E}\big(\|x_{n+1} - s_j\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) \leqslant \|x_n - s_j\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \delta_n(\cdot, s_j) \quad \mathsf{P}\text{-a.s.} \tag{3.4}$$

Additionally,

$$(\forall j \in \mathbb{N}) \quad \|x_{n+1} - s_j\|_{\mathsf{H}}^2 \leqslant 2\|x_{n+1}\|_{\mathsf{H}}^2 + 2\|s_j\|_{\mathsf{H}}^2 \leqslant 2\|x_{n+1}\|_{\mathsf{H}}^2 + 2\|z\|_{\mathsf{H}}^2 + 2 \quad \mathsf{P}\text{-a.s.} \tag{3.5}$$

Note that the right-hand term in (3.5) is integrable and that $\|x_{n+1} - s_j\|_{\mathsf{H}}^2 \to \|x_{n+1} - z\|_{\mathsf{H}}^2$ P-a.s. as $j \to +\infty$. Therefore, by the conditional dominated convergence theorem [54, Theorem 2.7.2(a)],

$$\mathsf{E}(\|x_{n+1} - s_j\|_{\mathsf{H}}^2 \mid \mathcal{X}_n) \to \mathsf{E}(\|x_{n+1} - z\|_{\mathsf{H}}^2 \mid \mathcal{X}_n) \ \mathsf{P}\text{-a.s.} \ \text{ as } \ j \to +\infty. \tag{3.6}$$

On the other hand, the continuity of δ_n with respect to the H-variable implies that $\delta_n(\cdot, s_j) \to \delta_n(\cdot, z)$ P-a.s. as $j \to +\infty$. Altogether, taking the limit as $j \to +\infty$ in (3.4) yields

$$\mathsf{E}\big(\|x_{n+1} - z\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) \leqslant \|x_n - z\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \delta_n(\cdot, z) \quad \mathsf{P}\text{-a.s.} \tag{3.7}$$

(iv): Take the expected value in (iii).

(v)(a): This follows from (ii) and Lemma 2.11(ii).

(v)(b): Let $\Omega'' \in \mathcal{F}$ be such that $\mathsf{P}(\Omega'') = 1$ and, for every $\omega \in \Omega''$, $z(\omega) \in \mathsf{Z}$. Further, let $\Omega' \in \mathcal{F}$ be given as in Lemma 2.11(iv), which holds as a consequence of (ii). Then

$$(\forall \omega \in \Omega' \cap \Omega'') \quad \big(\|x_n(\omega) - z(\omega)\|_{\mathsf{H}}\big)_{n\in\mathbb{N}} \text{ converges,} \tag{3.8}$$

which confirms that $(\|x_n - z\|_H)_{n\in\mathbb{N}}$ converges P-a.s. since $P(\Omega' \cap \Omega'') = 1$.

(v)(c): Let $z \in Z$. In view of (ii) and Lemma 2.11(i),

$$\sum_{n\in\mathbb{N}} E\big(\lambda_n(2-\lambda_n)\|d_n\|_H^2 \,\big|\, \mathcal{X}_n\big) < +\infty \text{ P-a.s.} \tag{3.9}$$

(v)(d)–(v)(f): These follow from (ii) and Lemma 2.11(v)–(vii).

(vi)(a): Note that $\{\varnothing, \Omega\} \subset \bigcap_{n\in\mathbb{N}} \mathcal{X}_n$. It follows from (iv) and the assumption that, for every $z \in Z$, $\sum_{n\in\mathbb{N}} E\delta_n(\cdot, z) < +\infty$, that $(x_n)_{n\in\mathbb{N}}$ is quasi-Fejér of Type III in $L^2(\Omega, \mathcal{F}, P; H)$ relative to $L^2(\Omega, \{\varnothing, \Omega\}, P; Z)$ [16, Definition 1.1]. Hence, [16, Proposition 3.3(i)] implies that $(x_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega, \mathcal{F}, P; H)$.

(vi)(b): It follows from (vi)(a) that $\sup_{n\in\mathbb{N}} E\|x_n - z\|_H^2 < +\infty$ and from (v)(b) that $(\|x_n - z\|_H)_{n\in\mathbb{N}}$ converges P-a.s. We then invoke Lemma 2.9 to deduce that $E\|x_n - z\|_H \to E\big(\lim \|x_n - z\|_H\big) < +\infty$.

(vi)(c): Let $z \in Z$. Then, in view of (iv) and Corollary 2.6(i),

$$\sum_{n\in\mathbb{N}} E\big(\lambda_n(2-\lambda_n)\|d_n\|_H^2\big) < +\infty. \tag{3.10}$$

(vi)(d): In view of (vi)(a), $(x_n)_{n\in\mathbb{N}}$ possesses a weak sequential cluster point $w \in L^2(\Omega, \mathcal{F}, P; H)$, i.e., there exists a strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $x_{k_n} \rightharpoonup w$ in $L^2(\Omega, \mathcal{F}, P; H)$. However, since H is separable, it contains a countable dense set $\{y_j\}_{j\in\mathbb{N}}$. Let us fix temporarily $j \in \mathbb{N}$ and identify y_j with a constant random variable in $L^2(\Omega, \mathcal{F}, P; H)$. Then $E\langle x_{k_n} - w \mid y_j\rangle_H \to 0$ and we can therefore extract a further subsequence $(x_{l_{k_n}})_{n\in\mathbb{N}}$ such that $\langle x_{l_{k_n}} - w \mid y_j\rangle_H \to 0$ P-a.s. On the other hand, the assumption yields $\langle x_{l_{k_n}} - x \mid y_j\rangle_H \to 0$ P-a.s. We deduce from the P-almost sure uniqueness of the limit that there exists $\Omega_j \in \mathcal{F}$ such that $P(\Omega_j) = 1$ and

$$(\forall \omega \in \Omega_j) \quad \langle x(\omega) \mid y_j\rangle_H = \langle w(\omega) \mid y_j\rangle_H. \tag{3.11}$$

Let us set $\Omega'' = \bigcap_{j\in\mathbb{N}} \Omega_j$ and note that $P(\Omega'') = 1$. Then (3.11) yields

$$(\forall j \in \mathbb{N})(\forall \omega \in \Omega'') \quad \langle x(\omega) - w(\omega) \mid y_j\rangle_H = 0. \tag{3.12}$$

By density, for every $\omega \in \Omega''$, there exists a strictly increasing sequence $(i_j)_{j\in\mathbb{N}}$ such that $y_{i_j} \to x(\omega) - w(\omega)$ and it results from (3.12) that

$$\|x(\omega) - w(\omega)\|_H^2 = \langle x(\omega) - w(\omega) \mid x(\omega) - w(\omega)\rangle_H = 0, \tag{3.13}$$

which shows that $x(\omega) = w(\omega)$. Thus, $x = w$ P-a.s. and it follows from [5, Lemma 2.46] that $x_n \rightharpoonup x$ in $L^2(\Omega, \mathcal{F}, P; H)$.

(vi)(e): Suppose that $x_n \to x$ P-a.s. Then it follows from (vi)(a) and Lemma 2.9 that $x \in L^1(\Omega, \mathcal{F}, P; Z)$ and $x_n \to x$ in $L^1(\Omega, \mathcal{F}, P; H)$. Conversely, suppose that $x \in L^1(\Omega, \mathcal{F}, P; Z)$ and $x_n \to x$ in $L^1(\Omega, \mathcal{F}, P; H)$, i.e., $E\|x_n - x\|_H \to 0$. Then there exists a strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $\|x_{k_n} - x\|_H \to 0$ P-a.s. On the other hand, (v)(b) guarantees that $(\|x_n - x\|_H)_{n\in\mathbb{N}}$ converges P-a.s. Since the P-almost sure limit of any subsequence coincides with the P-almost sure limit of the sequence, we conclude that $\|x_n - x\|_H \to 0$ P-a.s. Additionally, as P-almost sure strong convergence implies P-almost sure weak convergence, we deduce from (vi)(d) that $x \in L^2(\Omega, \mathcal{F}, P; H)$ and $x_n \rightharpoonup x$ in $L^2(\Omega, \mathcal{F}, P; H)$. ☐

We now assume that the tolerance variables $(\delta_n)_{n\in\mathbb{N}}$ are constant with respect to the H-variable and depend only on the Ω-variable.

Theorem 3.3. *Let $(x_n)_{n\in\mathbb{N}}$ be the sequence generated by Algorithm 3.1. For every $n \in \mathbb{N}$, assume that δ_n is constant with respect to the* H*-variable and set, for every $\omega \in \Omega$, $\vartheta_n(\omega) = \delta_n(\omega, 0)$. Then the following hold:*

(i) *Let* $\mathsf{n}\in\mathbb{N}$. *Then* $\mathsf{E}(\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}+1})\,|\,\mathcal{X}_{\mathsf{n}}) \leqslant \mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}}) + \vartheta_{\mathsf{n}}$ P-a.s.

(ii) *Let* $\mathsf{n}\in\mathbb{N}$. *Then* $\mathsf{E}\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}+1}) \leqslant \mathsf{E}\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}}) + \mathsf{E}\vartheta_{\mathsf{n}}$.

(iii) *Suppose that* $\sum_{\mathsf{n}\in\mathbb{N}}\vartheta_{\mathsf{n}} < +\infty$ P-a.s. *Then* $(\mathsf{d}_{\mathsf{Z}}(x_{\mathsf{n}}))_{\mathsf{n}\in\mathbb{N}}$ *converges* P-a.s.

(iv) *Suppose that* $\sum_{\mathsf{n}\in\mathbb{N}}\mathsf{E}\vartheta_{\mathsf{n}} < +\infty$. *Then the following hold:*

(a) $(\mathsf{E}\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}}))_{\mathsf{n}\in\mathbb{N}}$ *converges.*

(b) *Suppose that* Z *is convex and that* $\underline{\lim}\,\mathsf{E}\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}}) = 0$. *Then* $(x_{\mathsf{n}})_{\mathsf{n}\in\mathbb{N}}$ *converges strongly in* $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *and* P-a.s. *to a* Z*-valued random variable.*

(c) *Suppose that* Z *is convex and that there exists* $\chi\in\,]0,1[$ *such that*

$$(\forall \mathsf{n}\in\mathbb{N})\quad \mathsf{E}\big(\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}+1})\,\big|\,\mathcal{X}_{\mathsf{n}}\big) \leqslant \chi \mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}}) + \vartheta_{\mathsf{n}}\quad \text{P-a.s.} \tag{3.14}$$

Then the following are satisfied:

[A] *Let* $\mathsf{n}\in\mathbb{N}$. *Then* $\mathsf{E}\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}+1}) \leqslant \chi^{\mathsf{n}+1}\mathsf{E}\mathsf{d}_{\mathsf{Z}}^2(x_0) + \sum_{\mathsf{j}=0}^{\mathsf{n}}\chi^{\mathsf{n}-\mathsf{j}}\mathsf{E}\vartheta_{\mathsf{j}}$.

[B] *There exists* $x\in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{Z})$ *such that* $(x_{\mathsf{n}})_{\mathsf{n}\in\mathbb{N}}$ *converges strongly in* $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *and* P-a.s. *to* x, *and*

$$(\forall \mathsf{n}\in\mathbb{N})\quad \mathsf{E}\|x_{\mathsf{n}}-x\|_{\mathsf{H}}^2 \leqslant 4\chi^{\mathsf{n}}\mathsf{E}\mathsf{d}_{\mathsf{Z}}^2(x_0) + 4\sum_{\mathsf{j}=0}^{\mathsf{n}-1}\chi^{\mathsf{n}-\mathsf{j}-1}\mathsf{E}\vartheta_{\mathsf{j}} + 2\sum_{\mathsf{j}\geqslant\mathsf{n}}\mathsf{E}\vartheta_{\mathsf{j}}. \tag{3.15}$$

Proof. (i): Let $\mathsf{z}\in\mathsf{Z}$. Then Theorem 3.2(ii) yields $\mathsf{E}(\|x_{\mathsf{n}+1}-\mathsf{z}\|_{\mathsf{H}}^2\,|\,\mathcal{X}_{\mathsf{n}}) \leqslant \|x_{\mathsf{n}}-\mathsf{z}\|_{\mathsf{H}}^2 + \vartheta_{\mathsf{n}}$ P-a.s. On the other hand, $\mathsf{d}_{\mathsf{Z}}(x_{\mathsf{n}+1}) \leqslant \|x_{\mathsf{n}+1}-\mathsf{z}\|_{\mathsf{H}}$ P-a.s. Thus,

$$\mathsf{E}\big(\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}+1})\,\big|\,\mathcal{X}_{\mathsf{n}}\big) \leqslant \mathsf{E}\big(\|x_{\mathsf{n}+1}-\mathsf{z}\|_{\mathsf{H}}^2\,\big|\,\mathcal{X}_{\mathsf{n}}\big) \leqslant \|x_{\mathsf{n}}-\mathsf{z}\|_{\mathsf{H}}^2 + \vartheta_{\mathsf{n}}\ \text{P-a.s.} \tag{3.16}$$

Taking the infimum over $\mathsf{z}\in\mathsf{Z}$ yields the claim.

(ii): Take the expected value in (i).

(iii): This follows from (i) and Lemma 2.5(i).

(iv)(a): This follows from (ii) and Corollary 2.6(i).

(iv)(b): Let $\mathsf{n}\in\mathbb{N}$, $\mathsf{m}\in\mathbb{N}\smallsetminus\{0\}$, and $z\in L^2(\Omega,\mathcal{X}_{\mathsf{n}},\mathsf{P};\mathsf{Z})$. Then $z\in\bigcap_{1\leqslant\mathsf{j}\leqslant\mathsf{m}}L^2(\Omega,\mathcal{X}_{\mathsf{n}+\mathsf{j}},\mathsf{P};\mathsf{H})$ and we derive inductively from (3.1) and Theorem 3.2(iii) that

$$\begin{aligned}\mathsf{E}\big(\|x_{\mathsf{n}}-x_{\mathsf{n}+\mathsf{m}}\|_{\mathsf{H}}^2\,\big|\,\mathcal{X}_{\mathsf{n}}\big) &\leqslant 2\mathsf{E}\Big(\|x_{\mathsf{n}}-z\|_{\mathsf{H}}^2 + \|x_{\mathsf{n}+\mathsf{m}}-z\|_{\mathsf{H}}^2\,\Big|\,\mathcal{X}_{\mathsf{n}}\Big)\\ &\leqslant 2\|x_{\mathsf{n}}-z\|_{\mathsf{H}}^2 + 2\mathsf{E}\Big(\mathsf{E}\big(\|x_{\mathsf{n}+\mathsf{m}}-z\|_{\mathsf{H}}^2\,\big|\,\mathcal{X}_{\mathsf{n}+\mathsf{m}-1}\big)\,\Big|\,\mathcal{X}_{\mathsf{n}}\Big)\\ &\leqslant 4\|x_{\mathsf{n}}-z\|_{\mathsf{H}}^2 + 2\sum_{\mathsf{j}=\mathsf{n}}^{\mathsf{n}+\mathsf{m}-1}\vartheta_{\mathsf{j}}\quad\text{P-a.s.}\end{aligned} \tag{3.17}$$

Now assume that $z=\mathrm{proj}_{\mathsf{Z}}\,x_{\mathsf{n}}$ and recall that $\mathrm{proj}_{\mathsf{Z}}$ is nonexpansive [5, Proposition 4.16] while x_{n} is $(\mathcal{X}_{\mathsf{n}},\mathcal{B}_{\mathsf{H}})$-measurable. Consequently, z is $(\mathcal{X}_{\mathsf{n}},\mathcal{B}_{\mathsf{H}})$-measurable. Given $y\in L^2(\Omega,\mathcal{X}_{\mathsf{n}},\mathsf{P};\mathsf{Z})$,

$$\begin{aligned}\frac{1}{2}\mathsf{E}\|z\|_{\mathsf{H}}^2 = \frac{1}{2}\mathsf{E}\|z-y+y\|_{\mathsf{H}}^2 &\leqslant \mathsf{E}\|\mathrm{proj}_{\mathsf{Z}}\,x_{\mathsf{n}}-\mathrm{proj}_{\mathsf{Z}}\,y\|_{\mathsf{H}}^2 + \mathsf{E}\|y\|_{\mathsf{H}}^2\\ &\leqslant \|x_{\mathsf{n}}-y\|_{L^2(\Omega,\mathcal{X}_{\mathsf{n}},\mathsf{P};\mathsf{Z})}^2 + \|y\|_{L^2(\Omega,\mathcal{X}_{\mathsf{n}},\mathsf{P};\mathsf{Z})}^2,\end{aligned} \tag{3.18}$$

which shows that $z\in L^2(\Omega,\mathcal{X}_{\mathsf{n}},\mathsf{P};\mathsf{Z})$. Further, (3.17) yields

$$\mathsf{E}\big(\|x_{\mathsf{n}}-x_{\mathsf{n}+\mathsf{m}}\|_{\mathsf{H}}^2\,\big|\,\mathcal{X}_{\mathsf{n}}\big) \leqslant 4\mathsf{d}_{\mathsf{Z}}^2(x_{\mathsf{n}}) + 2\sum_{\mathsf{j}=\mathsf{n}}^{\mathsf{n}+\mathsf{m}-1}\vartheta_{\mathsf{j}}\quad\text{P-a.s.} \tag{3.19}$$

Therefore, upon taking expectations, we get

$$\mathsf{E}\|x_n - x_{n+m}\|_{\mathsf{H}}^2 \leqslant 4\mathsf{E}d_{\mathsf{Z}}^2(x_n) + 2\sum_{j=n}^{n+m-1}\mathsf{E}\vartheta_j. \tag{3.20}$$

The assumption $\underline{\lim}\,\mathsf{E}d_{\mathsf{Z}}^2(x_n) = 0$ and (iv)(a) yield $\lim \mathsf{E}d_{\mathsf{Z}}^2(x_n) = 0$. In addition,

$$(\forall m \in \mathbb{N} \smallsetminus \{0\}) \quad 0 \leqslant \sum_{j=n}^{n+m-1}\mathsf{E}\vartheta_j \leqslant \sum_{j\geqslant n}\mathsf{E}\vartheta_j \to 0 \text{ as } n \to +\infty. \tag{3.21}$$

We thus infer from (3.20) that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$, which implies that there exists $x \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ such that $x_n \to x$ in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Further, since $d_{\mathsf{Z}}^2 \colon \mathsf{H} \to [0,+\infty[$ is continuous, $d_{\mathsf{Z}}^2(x_n) \to d_{\mathsf{Z}}^2(x)$ P-a.s. In addition, it follows from Fatou's lemma that

$$0 \leqslant \mathsf{E}d_{\mathsf{Z}}^2(x) \leqslant \underline{\lim}\,\mathsf{E}d_{\mathsf{Z}}^2(x_n) = 0. \tag{3.22}$$

Hence $\mathsf{E}d_{\mathsf{Z}}^2(x) = 0$, $d_{\mathsf{Z}}^2(x) = 0$ P-a.s., and $x \in \mathsf{Z}$ P-a.s. Finally, Theorem 3.2(vi)(e) yields $x_n \to x$ P-a.s.

(iv)(c):

[A]: Taking expectations in (3.14) yields $\mathsf{E}d_{\mathsf{Z}}^2(x_{n+1}) \leqslant \chi\mathsf{E}d_{\mathsf{Z}}^2(x_n) + \mathsf{E}\vartheta_n$. The claim follows by induction.

[B]: It follows from Corollary 2.6(ii) that $\lim \mathsf{E}d_{\mathsf{Z}}^2(x_n) = 0$. Therefore, (iv)(b) implies that $(x_n)_{n\in\mathbb{N}}$ converges strongly in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and P-a.s. to a Z-valued random variable. Finally, arguing as in [16, Theorem 3.13(ii)], we obtain (3.15). ☐

3.2. A stochastic algorithm with super relaxations

We study an implementation of Algorithm 1.2 in which the standard condition that the relaxations are deterministic and bounded above by 2 is not imposed. In Section 1 we called such relaxations super relaxations.

Algorithm 3.4. In Algorithm 1.2 assume that, for every $n \in \mathbb{N}$, $\varepsilon_n \in \mathfrak{C}(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$, λ_n is independent of $\sigma(\{x_0,\ldots,x_n,d_n\})$, and $\mathsf{E}(\lambda_n(2-\lambda_n)) \geqslant 0$.

Proposition 3.5. *Algorithm 3.4 is a special case of Algorithm 3.1 where, for every* $n \in \mathbb{N}$, $\delta_n = 2\varepsilon_n\mathsf{E}\lambda_n$.

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be the sequence generated by Algorithm 3.4. Let us first show by induction that it is a well-defined sequence in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. By assumption, $x_0 \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Fix $n \in \mathbb{N}$ and note that d_n is measurable as a combination of measurable functions. Additionally, (1.2) yields

$$\begin{aligned}
\frac{1}{2}\mathsf{E}\|d_n\|_{\mathsf{H}}^2 &= \frac{1}{2}\mathsf{E}\|\alpha_n t_n^*\|_{\mathsf{H}}^2\\
&\leqslant \frac{1}{2}\mathsf{E}\left\|\frac{1_{[t_n^*\neq 0]}1_{[\langle x_n \mid t_n^*\rangle_{\mathsf{H}}>\eta_n]}\big(\langle x_n \mid t_n^*\rangle_{\mathsf{H}}-\eta_n\big)}{\|t_n^*\|_{\mathsf{H}}^2 + 1_{[t_n^*=0]}}\,t_n^*\right\|_{\mathsf{H}}^2\\
&= \frac{1}{2}\mathsf{E}\left|\frac{1_{[t_n^*\neq 0]}1_{[\langle x_n \mid t_n^*\rangle_{\mathsf{H}}>\eta_n]}\big(\langle x_n \mid t_n^*\rangle_{\mathsf{H}}-\eta_n\big)}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2\\
&\leqslant \mathsf{E}\left|\frac{1_{[t_n^*\neq 0]}1_{[\langle x_n \mid t_n^*\rangle_{\mathsf{H}}>\eta_n]}\langle x_n \mid t_n^*\rangle_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 + \mathsf{E}\left|\frac{1_{[t_n^*\neq 0]}1_{[\langle x_n \mid t_n^*\rangle_{\mathsf{H}}>\eta_n]}\eta_n}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2\\
&\leqslant \mathsf{E}\left|\frac{\|x_n\|_{\mathsf{H}}\,\|t_n^*\|_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 + \mathsf{E}\left|\frac{1_{[t_n^*\neq 0]}1_{[\langle x_n \mid t_n^*\rangle_{\mathsf{H}}>\eta_n]}\eta_n}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \mathsf{E}\|x_n\|_{\mathsf{H}}^2 + \mathsf{E}\left|\frac{1_{[t_n^*\neq 0]}1_{[\langle x_n \mid t_n^*\rangle_{\mathsf{H}}>\eta_n]}\eta_n}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2\\
&< +\infty. \qquad (3.23)
\end{aligned}
$$

Thus, $d_n \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and, since $\lambda_n \in L^\infty(\Omega, \mathcal{F}, \mathsf{P};]0, +\infty[)$, $x_{n+1} = x_n - \lambda_n d_n \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$, which completes the induction argument. The fact that $\lambda_n \in L^\infty(\Omega, \mathcal{F}, \mathsf{P};]0, +\infty[)$ also guarantees the integrability of λ_n and $\lambda_n(2-\lambda_n)$. Further, since λ_n is independent of $\sigma(\{x_0, \ldots, x_n, d_n\})$ and $\mathsf{E}(\lambda_n(2-\lambda_n)) \geqslant 0$, it follows from Lemma 2.7 that

$$
\mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) = \mathsf{E}\big(\lambda_n(2-\lambda_n)\big)\mathsf{E}\big(\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) \geqslant 0 \ \ \mathsf{P}\text{-a.s.} \qquad (3.24)
$$

Next, we infer from (1.2) that

$$
(\forall n \in \mathbb{N}) \quad \alpha_n\eta_n = \langle x_n \mid \alpha_n t_n^*\rangle_{\mathsf{H}} - \alpha_n^2\|t_n^*\|_{\mathsf{H}}^2 = \langle x_n \mid d_n\rangle_{\mathsf{H}} - \|d_n\|_{\mathsf{H}}^2, \qquad (3.25)
$$

which shows that $\alpha_n\eta_n \in L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$. Now set $\delta_n = 2\varepsilon_n\mathsf{E}\lambda_n \in \mathfrak{C}(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and let $\mathsf{z} \in \mathsf{Z}$. Then we deduce from (1.2), Lemma 2.10, and (3.25) that

$$
\begin{aligned}
(\forall n \in \mathbb{N}) \quad \mathsf{E}\big(\langle \mathsf{z} \mid d_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) &= \langle \mathsf{z} \mid \mathsf{E}(\alpha_n t_n^* \mid \mathcal{X}_n)\rangle_{\mathsf{H}}\\
&\leqslant \mathsf{E}(\alpha_n\eta_n \mid \mathcal{X}_n) + \varepsilon_n(\cdot, \mathsf{z})\\
&= \mathsf{E}\big(\langle x_n \mid d_n\rangle_{\mathsf{H}} - \|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \varepsilon_n(\cdot, \mathsf{z}) \ \ \mathsf{P}\text{-a.s.} \qquad (3.26)
\end{aligned}
$$

Finally, we derive from (3.26) and Lemma 2.7 that

$$
\begin{aligned}
(\forall n \in \mathbb{N}) \quad \mathsf{E}\big(\lambda_n\langle \mathsf{z} + d_n - x_n \mid d_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) &= \mathsf{E}\big(\langle \mathsf{z} + d_n - x_n \mid d_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big)\mathsf{E}\lambda_n\\
&= \mathsf{E}\big(\langle \mathsf{z} \mid d_n\rangle_{\mathsf{H}} + \|d_n\|_{\mathsf{H}}^2 - \langle x_n \mid d_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big)\mathsf{E}\lambda_n\\
&\leqslant \varepsilon_n(\cdot, \mathsf{z})\mathsf{E}\lambda_n\\
&= \frac{\delta_n(\cdot, \mathsf{z})}{2} \ \ \mathsf{P}\text{-a.s.}, \qquad (3.27)
\end{aligned}
$$

which yields the claim. ☐

The asymptotic behavior of Algorithm 3.4 is our next topic. We leverage Proposition 3.5 and Theorems 3.2 and 3.3 to obtain the following properties.

Theorem 3.6. *Let $(x_n)_{n\in\mathbb{N}}$ be the sequence generated by Algorithm 3.4.*

(i) *Suppose that, for every $\mathsf{z} \in \mathsf{Z}$, $\sum_{n\in\mathbb{N}} \varepsilon_n(\cdot, \mathsf{z})\mathsf{E}\lambda_n < +\infty$ P-a.s. Then the following hold:*
   - (a) $\sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}(\|d_n\|_{\mathsf{H}}^2 \mid \mathcal{X}_n) < +\infty$ P-a.s.
   - (b) *Suppose that $\inf_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n)) > 0$ and there exists $\rho \in [1, +\infty[$ such that $\sup_{n\in\mathbb{N}} \lambda_n < \rho$ P-a.s. Then $\sum_{n\in\mathbb{N}} \mathsf{E}(\|x_{n+1} - x_n\|_{\mathsf{H}}^2 \mid \mathcal{X}_n) < +\infty$ P-a.s. and $\sum_{n\in\mathbb{N}} \|x_{n+1} - x_n\|_{\mathsf{H}}^2 < +\infty$ P-a.s.*
   - (c) *Suppose that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to a Z-valued random variable.*
   - (d) *Suppose that $\mathfrak{S}(x_n)_{n\in\mathbb{N}} \cap \mathsf{Z} \neq \varnothing$ P-a.s. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to a Z-valued random variable.*
   - (e) *Suppose that $\mathfrak{S}(x_n)_{n\in\mathbb{N}} \neq \varnothing$ P-a.s. and that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to a Z-valued random variable.*

(ii) *Suppose that, for every $\mathsf{z} \in \mathsf{Z}$, $\sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n(\cdot, \mathsf{z})\mathsf{E}\lambda_n < +\infty$. Then the following hold:*

(a) $\sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\|d_n\|_{\mathsf{H}}^2 < +\infty$.

(b) *Suppose that* $\inf_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n)) > 0$ *and there exists* $\rho \in [1,+\infty[$ *such that* $\sup_{n\in\mathbb{N}} \lambda_n < \rho$ P-a.s. *Then* $\sum_{n\in\mathbb{N}} \mathsf{E}\|x_{n+1}-x_n\|_{\mathsf{H}}^2 < +\infty$.

(c) *Suppose that* $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *and weakly in* $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *to a random variable* $x \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{Z})$.

(d) *Suppose that* $\mathfrak{S}(x_n)_{n\in\mathbb{N}} \cap \mathsf{Z} \neq \varnothing$ P-a.s. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *and strongly in* $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *to a random variable* $x \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{Z})$. *Additionally,* $(x_n)_{n\in\mathbb{N}}$ *converges weakly in* $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *to* x.

(e) *Suppose that* Z *is convex, that, for every* $n \in \mathbb{N}$, ε_n *is constant with respect to the* H*-variable, and that* $\underline{\lim}\, \mathsf{E}d_{\mathsf{Z}}^2(x_n) = 0$. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *and* P-a.s. *to a* Z*-valued random variable.*

(f) *Suppose that* Z *is convex, that, for every* $n \in \mathbb{N}$, ε_n *is constant with respect to the* H*-variable, and that there exists* $\chi \in]0,1[$ *such that*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}\big(d_{\mathsf{Z}}^2(x_{n+1}) \,\big|\, \mathcal{X}_n\big) \leqslant \chi d_{\mathsf{Z}}^2(x_n) + 2\varepsilon_n \mathsf{E}\lambda_n \ \text{ P-a.s.} \tag{3.28}$$

Set, for every $n \in \mathbb{N}$ *and for every* $\omega \in \Omega$, $\vartheta_n(\omega) = \varepsilon_n(\omega, 0)$. *Then the following are satisfied:*

[A] *Let* $n \in \mathbb{N}$. *Then* $\mathsf{E}d_{\mathsf{Z}}^2(x_{n+1}) \leqslant \chi^{n+1}\mathsf{E}d_{\mathsf{Z}}^2(x_0) + 2\sum_{j=0}^{n} \chi^{n-j}\mathsf{E}\vartheta_j\mathsf{E}\lambda_j$.

[B] *There exists* $x \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{Z})$ *such that* $(x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *and* P-a.s. *to* x, *and*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}\|x_n - x\|_{\mathsf{H}}^2 \leqslant 4\chi^n \mathsf{E}d_{\mathsf{Z}}^2(x_0) + 8\sum_{j=0}^{n-1} \chi^{n-j-1}\mathsf{E}\vartheta_j\mathsf{E}\lambda_j + 4\sum_{j\geqslant n} \mathsf{E}\vartheta_j\mathsf{E}\lambda_j. \tag{3.29}$$

Proof. In view of Proposition 3.5, we appeal to Theorems 3.2 and 3.3 to establish the claims.

(i)(a): It follows from Theorem 3.2(v)(c) and Lemma 2.7 that

$$\sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\big(\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) = \sum_{n\in\mathbb{N}} \mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) < +\infty \ \text{P-a.s.} \tag{3.30}$$

(i)(a)⇒(i)(b): It follows from (1.2) that

$$\sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\left(\frac{1}{\lambda_n^2}\|x_{n+1}-x_n\|_{\mathsf{H}}^2 \,\middle|\, \mathcal{X}_n\right) = \sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\big(\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) < +\infty \ \text{P-a.s.} \tag{3.31}$$

Hence, the assumption $\inf_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n)) > 0$ yields $\sum_{n\in\mathbb{N}} \mathsf{E}\big(\|x_{n+1}-x_n\|_{\mathsf{H}}^2/\lambda_n^2 \,\big|\, \mathcal{X}_n\big) < +\infty$ P-a.s. Further,

$$(\forall n \in \mathbb{N}) \quad 0 < \frac{1}{\rho^2} \leqslant \frac{1}{\lambda_n^2} \ \text{ P-a.s.} \tag{3.32}$$

Thus,

$$\sum_{n\in\mathbb{N}} \mathsf{E}\big(\|x_{n+1}-x_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) < +\infty \ \text{P-a.s.} \tag{3.33}$$

In addition,

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}\left(\sum_{k=0}^{n+1}\|x_{k+1}-x_k\|_{\mathsf{H}}^2 \,\middle|\, \mathcal{X}_{n+1}\right) = \sum_{k=0}^{n}\|x_{k+1}-x_k\|_{\mathsf{H}}^2 + \mathsf{E}\big(\|x_{n+2}-x_{n+1}\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_{n+1}\big) \ \text{ P-a.s.} \tag{3.34}$$

It then follows from (3.33) and Lemma 2.5(i) that $(\sum_{k=0}^{n}\|x_{k+1}-x_k\|_H^2)_{n\in\mathbb{N}}$ converges P-a.s. to a $[0,+\infty[$-valued random variable, hence $\sum_{n\in\mathbb{N}}\|x_{n+1}-x_n\|_H^2<+\infty$ P-a.s.

(i)(c)–(i)(e): These follow from Theorem 3.2(v)(d)–(v)(f).

(ii)(a): It follows from Theorem 3.2(vi)(c) and Lemma 2.7 that

$$
\begin{aligned}
\sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\|d_n\|_H^2 &= \sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\Big(\mathsf{E}(\|d_n\|_H^2\,|\,\mathcal{X}_n)\Big)\\
&= \sum_{n\in\mathbb{N}} \mathsf{E}\Big(\mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}(\|d_n\|_H^2\,|\,\mathcal{X}_n)\Big)\\
&= \sum_{n\in\mathbb{N}} \mathsf{E}\Big(\mathsf{E}\Big(\lambda_n(2-\lambda_n)\|d_n\|_H^2\,\Big|\,\mathcal{X}_n\Big)\Big)\\
&= \sum_{n\in\mathbb{N}} \mathsf{E}\Big(\lambda_n(2-\lambda_n)\|d_n\|_H^2\Big)\\
&< +\infty. \qquad (3.35)
\end{aligned}
$$

Hence $\sum_{n\in\mathbb{N}}\mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\|d_n\|_H^2<+\infty$.

(ii)(a)⇒(ii)(b): It follows from (3.1) that

$$
\sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\bigg(\frac{1}{\lambda_n^2}\|x_{n+1}-x_n\|_H^2\bigg) = \sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n))\mathsf{E}\|d_n\|_H^2 < +\infty. \qquad (3.36)
$$

Thus, as in (i)(b), $\sum_{n\in\mathbb{N}}\mathsf{E}\|x_{n+1}-x_n\|_H^2<+\infty$.

(ii)(c)–(ii)(d): These follow from (i)(c)–(i)(d) and Theorem 3.2(vi)(d)–(vi)(e).

(ii)(e)–(ii)(f): These follow from Theorem 3.3(iv)(b)–(iv)(c). ☐

3.3. A stochastic algorithm with random relaxations bounded by 2

We present an implementation of Algorithm 1.2 with an alternative relaxation strategy.

Algorithm 3.7. In Algorithm 1.2, for every $n\in\mathbb{N}$, $\varepsilon_n\in\mathfrak{C}(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and $\lambda_n\in L^\infty(\Omega,\mathcal{X}_n,\mathsf{P};\,]0,2[)$.

Proposition 3.8. *Algorithm 3.7 is a special case of Algorithm 3.1 where, for every* $n\in\mathbb{N}$, $\delta_n=2\lambda_n\varepsilon_n$.

Proof. Set $(\forall n\in\mathbb{N})\ \delta_n=2\lambda_n\varepsilon_n$. Following the proof of Proposition 3.5, it is enough to show that

$$
(\forall n\in\mathbb{N})\quad
\begin{cases}
\delta_n\in\mathfrak{C}(\Omega,\mathcal{F},\mathsf{P};\mathsf{H});\\
\mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_H^2\,\big|\,\mathcal{X}_n\big)\geqslant 0\quad \text{P-a.s.};\\
(\forall z\in Z)\quad \mathsf{E}\big(\lambda_n\langle z+d_n-x_n\,|\,d_n\rangle_H\,\big|\,\mathcal{X}_n\big)\leqslant \delta_n(\cdot,z)/2\quad \text{P-a.s.}
\end{cases}
\qquad (3.37)
$$

Let $n\in\mathbb{N}$. It follows from the positivity and measurability of λ_n, as well the fact that $\varepsilon_n\in\mathfrak{C}(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$, that $\delta_n\in\mathfrak{C}(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Next, since $\lambda_n\in\,]0,2[$ P-a.s., we have $\lambda_n(2-\lambda_n)>0$ P-a.s. and hence

$$
\mathsf{E}\big(\lambda_n(2-\lambda_n)\|d_n\|_H^2\,\big|\,\mathcal{X}_n\big)\geqslant 0\quad\text{P-a.s.} \qquad (3.38)
$$

Finally, let $z\in Z$. It then follows from (3.26) and the fact that λ_n is positive and $\mathcal{X}_n$-measurable that

$$
\begin{aligned}
&\mathsf{E}\big(\lambda_n\langle z+d_n-x_n\,|\,d_n\rangle_H\,\big|\,\mathcal{X}_n\big)\\
&\qquad= \lambda_n\mathsf{E}\big(\langle z\,|\,d_n\rangle_H+\|d_n\|_H^2-\langle x_n\,|\,d_n\rangle_H\,\big|\,\mathcal{X}_n\big)\leqslant \lambda_n\varepsilon_n(\cdot,z)=\frac{\delta_n(\cdot,z)}{2}\quad\text{P-a.s.,} \qquad (3.39)
\end{aligned}
$$

which completes the proof. ☐

As in Section 3.2, we can derive weak, strong, and linear convergence results from Theorems 3.2 and 3.3. For brevity, we provide below only the weak convergence results but, as in Theorem 3.6, strong and linear convergence results can also be obtained.

Theorem 3.9. *Let $(x_n)_{n\in\mathbb{N}}$ be the sequence generated by Algorithm 3.7. Suppose that, for every* $z \in Z$, $\sum_{n\in\mathbb{N}} \lambda_n \varepsilon_n(\cdot, z) < +\infty$ P-a.s. *and that* $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset Z$ P-a.s. *Then $(x_n)_{n\in\mathbb{N}}$ converges weakly* P-a.s. *to a* Z-*valued random variable x. If, in addition, for every* $z \in Z$, $\sum_{n\in\mathbb{N}} \mathsf{E}(\lambda_n \varepsilon_n(\cdot, z)) < +\infty$, *then* $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and $(x_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ to x.*

Proof. In view of Proposition 3.8, the claim follows Theorem 3.2(v)(d) and 3.2(vi)(d). ☐

§4. Randomly relaxed Krasnosel'skiĭ–Mann iterations

Let us first recall some definitions about an operator $\mathsf{T} \colon \mathsf{H} \to \mathsf{H}$ [5, Chapter 4]. First, $\mathsf{T} \colon \mathsf{H} \to \mathsf{H}$ is nonexpansive if it is 1-Lipschitzian and α-averaged for some $\alpha \in \left]0, 1\right[$ if $\mathrm{Id} + \alpha^{-1}(\mathsf{T} - \mathrm{Id})$ is nonexpansive [3]. On the other hand, T is β-cocoercive for some $\beta \in \left]0, +\infty\right[$ if

$$(\forall x \in \mathsf{H})(\forall y \in \mathsf{H}) \quad \langle x - y \mid \mathsf{T}x - \mathsf{T}y\rangle_{\mathsf{H}} \geqslant \beta \|\mathsf{T}x - \mathsf{T}y\|_{\mathsf{H}}^2 \tag{4.1}$$

and it is firmly nonexpansive if it is 1-cocoercive.

The Krasnosel'skiĭ–Mann iterative process is a basic algorithm to construct fixed points of nonexpansive operators [5, 25, 29, 35, 38, 49]. We propose a study of its asymptotic behavior in a novel environment featuring random relaxations and stochastic errors.

Theorem 4.1. *Let* $\mathsf{T} \colon \mathsf{H} \to \mathsf{H}$ *be a nonexpansive operator such that* $\mathrm{Fix}\,\mathsf{T} \neq \varnothing$ *and let* $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. *Iterate*

$$\begin{array}{l} \text{for } n = 0, 1, \ldots \\ \left\lfloor \begin{array}{l} \text{take } e_n \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \text{ and } \mu_n \in L^\infty(\Omega, \mathcal{F}, \mathsf{P}; \left]0, 1\right[) \\ x_{n+1} = x_n + \mu_n\big(\mathsf{T}x_n + e_n - x_n\big). \end{array} \right. \end{array} \tag{4.2}$$

Set $(\forall n \in \mathbb{N})$ $\Phi_n = \{x_0, \ldots, x_n\}$ *and* $\mathcal{X}_n = \sigma(\Phi_n)$. *Suppose that* $\sum_{n\in\mathbb{N}} \mathsf{E}(\mu_n(1 - \mu_n)) = +\infty$ *and, for every* $n \in \mathbb{N}$, *μ_n is independent of* $\sigma(\{e_n\} \cup \Phi_n)$. *Then the following hold for some* Fix T*-valued random variable x:*

(i) *Suppose that* $\mathsf{E}(\|e_n\|_{\mathsf{H}}^2 \mid \mathcal{X}_n) \to 0$ P-a.s. *and* $\sum_{n\in\mathbb{N}} \mathsf{E}\mu_n \sqrt{\mathsf{E}(\|e_n\|_{\mathsf{H}}^2 \mid \mathcal{X}_n)} < +\infty$ P-a.s. *Then the following hold:*

  (a) $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to* 0.

  (b) $(x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to x.*

  (c) *Suppose that* $\mathsf{T} - \mathrm{Id}$ *is demiregular at every point in* Fix T. *Then $(x_n)_{n\in\mathbb{N}}$ converges strongly* P-a.s. *to x.*

(ii) *Suppose that* $\mathsf{E}\|e_n\|_{\mathsf{H}}^2 \to 0$ *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}\mu_n^2 \mathsf{E}\|e_n\|_{\mathsf{H}}^2} < +\infty$. *Then the following hold:*

  (a) $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and strongly* P-a.s. *to* 0.

  (b) $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathrm{Fix}\,\mathsf{T})$ *and $(x_n)_{n\in\mathbb{N}}$ converges weakly in* $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and weakly* P-a.s. *to x.*

  (c) *Suppose that* $\mathsf{T} - \mathrm{Id}$ *is demiregular at every point in* Fix T. *Then $(x_n)_{n\in\mathbb{N}}$ converges strongly in* $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and strongly* P-a.s. *to x.*

Proof. Let us show that the sequence $(x_n)_{n\in\mathbb{N}}$ constructed by (4.2) corresponds to a sequence generated by Algorithm 3.4. To see this, set $\mathsf{Z} = \operatorname{Fix}\mathsf{T}$ and observe that, since T is nonexpansive,

$$(\forall n \in \mathbb{N})\big(\forall z \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{Z})\big) \quad \mathsf{E}\|\mathsf{T}x_n - z\|_{\mathsf{H}}^2 \leqslant \mathsf{E}\|x_n - z\|_{\mathsf{H}}^2. \tag{4.3}$$

Thus if, for some $n \in \mathbb{N}$, $x_n \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$, then $\mathsf{T}x_n \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and (4.2) yields $x_{n+1} \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. This shows by induction that $(x_n)_{n\in\mathbb{N}}$ and $(\mathsf{T}x_n)_{n\in\mathbb{N}}$ lie in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Let us define

$$(\forall n \in \mathbb{N}) \quad \begin{cases} t_n^* = \dfrac{x_n - \mathsf{T}x_n - e_n}{2} \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H}); \\ \eta_n = \dfrac{\|x_n\|_{\mathsf{H}}^2 - \|\mathsf{T}x_n + e_n\|_{\mathsf{H}}^2}{4} \in L^1(\Omega,\mathcal{F},\mathsf{P};\mathbb{R}); \\ \alpha_n = 1_{[t_n^* \neq 0]}; \\ (\forall \mathsf{z} \in \mathsf{Z})\ \varepsilon_n(\cdot,\mathsf{z}) = \dfrac{1}{4}\mathsf{E}\big(\|e_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \dfrac{1}{2}\|x_n - \mathsf{z}\|_{\mathsf{H}}\sqrt{\mathsf{E}\big(\|e_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big)}; \\ d_n = t_n^*; \\ \lambda_n = 2\mu_n \in \left]0,2\right[\ \ \mathsf{P}\text{-a.s.} \end{cases} \tag{4.4}$$

Now set $\mathsf{F} = (\mathsf{T} + \mathrm{Id})/2$. Since T is nonexpansive, F is firmly nonexpansive (see [5, Proposition 4.4] or [28, Proposition 1.11.2]. Hence, we deduce from Lemma 2.10 and (4.4) that, for every $\mathsf{z} \in \mathsf{Z}$ and every $n \in \mathbb{N}$,

$$\begin{aligned} \big\langle \mathsf{z} \,\big|\, \mathsf{E}(\alpha_n t_n^* \,|\, \mathcal{X}_n)\big\rangle_{\mathsf{H}} &= \mathsf{E}\bigg(\Big\langle \mathsf{z} \,\Big|\, x_n - \mathsf{F}x_n - \frac{1}{2}e_n\Big\rangle_{\mathsf{H}} \,\bigg|\, \mathcal{X}_n\bigg) \\ &= \big\langle \mathsf{z} \,\big|\, x_n - \mathsf{F}x_n\big\rangle_{\mathsf{H}} - \frac{1}{2}\mathsf{E}\big(\langle \mathsf{z} \,|\, e_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) \\ &\leqslant \big\langle \mathsf{F}x_n \,\big|\, x_n - \mathsf{F}x_n\big\rangle_{\mathsf{H}} - \frac{1}{2}\mathsf{E}\big(\langle \mathsf{z} \,|\, e_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) \\ &= \mathsf{E}(\alpha_n\eta_n \,|\, \mathcal{X}_n) + \frac{1}{4}\mathsf{E}\big(\|e_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \frac{1}{2}\mathsf{E}\big(\langle \mathsf{T}x_n - \mathsf{z} \,|\, e_n\rangle_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) \\ &\leqslant \mathsf{E}(\alpha_n\eta_n \,|\, \mathcal{X}_n) + \frac{1}{4}\mathsf{E}\big(\|e_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \frac{1}{2}\|\mathsf{T}x_n - \mathsf{z}\|_{\mathsf{H}}\mathsf{E}\big(\|e_n\|_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) \\ &\leqslant \mathsf{E}(\alpha_n\eta_n \,|\, \mathcal{X}_n) + \frac{1}{4}\mathsf{E}\big(\|e_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) + \frac{1}{2}\|x_n - \mathsf{z}\|_{\mathsf{H}}\mathsf{E}\big(\|e_n\|_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) \\ &\leqslant \mathsf{E}(\alpha_n\eta_n \,|\, \mathcal{X}_n) + \varepsilon_n(\cdot,\mathsf{z}) \ \ \mathsf{P}\text{-a.s.} \end{aligned} \tag{4.5}$$

Next, it is clear from (4.4) that $\varepsilon_n \in \mathfrak{C}(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Furthermore, in view of (4.4), (4.2) can be written as

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n - \lambda_n d_n. \tag{4.6}$$

On the other hand, since $(\mu_n)_{n\in\mathbb{N}}$ lies almost surely in $\left]0,1\right[$, we have $\mathsf{E}(\lambda_n(2-\lambda_n)) \geqslant 0$. Additionally, for every $n \in \mathbb{N}$, μ_n is independent of $\sigma(\{e_n\} \cup \Phi_n)$ and, by (4.4), d_n is $\sigma(\{e_n\} \cup \Phi_n)$-measurable. Hence, $\sigma(\{d_n\} \cup \Phi_n) \subset \sigma(\{e_n\} \cup \Phi_n)$ and λ_n is independent of $\sigma(\{d_n\} \cup \Phi_n)$. Altogether, $(x_n)_{n\in\mathbb{N}}$ is a sequence generated by Algorithm 3.4. Finally, it follows from (4.2) and Lemma 2.7 that

$$\begin{aligned} (\forall n \in \mathbb{N})(\forall \mathsf{z} \in \mathsf{Z}) \quad & \mathsf{E}(\|x_{n+1} - \mathsf{z}\|_{\mathsf{H}} \,|\, \mathcal{X}_n) \\ &\leqslant \mathsf{E}\big((1-\mu_n)\|x_n - \mathsf{z}\|_{\mathsf{H}} + \mu_n\|\mathsf{T}x_n - \mathsf{z}\|_{\mathsf{H}} + \mu_n\|e_n\|_{\mathsf{H}} \,\big|\, \mathcal{X}_n\big) \\ &= (1 - \mathsf{E}\mu_n)\|x_n - \mathsf{z}\|_{\mathsf{H}} + \mathsf{E}\mu_n\|\mathsf{T}x_n - \mathsf{z}\|_{\mathsf{H}} + \mathsf{E}\mu_n\mathsf{E}(\|e_n\|_{\mathsf{H}} \,|\, \mathcal{X}_n) \\ &\leqslant \|x_n - \mathsf{z}\|_{\mathsf{H}} + \mathsf{E}\mu_n\mathsf{E}(\|e_n\|_{\mathsf{H}} \,|\, \mathcal{X}_n) \\ &\leqslant \|x_n - \mathsf{z}\|_{\mathsf{H}} + \mathsf{E}\mu_n\sqrt{\mathsf{E}(\|e_n\|_{\mathsf{H}}^2 \,|\, \mathcal{X}_n)} \ \ \mathsf{P}\text{-a.s.} \end{aligned} \tag{4.7}$$

Further, by invoking the nonexpansiveness of T and the Cauchy–Schwarz inequality, we obtain

$$
\begin{aligned}
(\forall n \in \mathbb{N})(\forall z \in Z) \quad & \|x_{n+1} - z\|_{L^2(\Omega,\mathcal{F},P;H)} \\
&= \sqrt{E\|x_{n+1} - z\|_H^2} \\
&= \sqrt{E\|(1-\mu_n)(x_n - z) + \mu_n(Tx_n - Tz) + \mu_n e_n\|_H^2} \\
&\leqslant \sqrt{E\big|\|x_n - z\|_H + \mu_n\|e_n\|_H\big|^2} \\
&= \sqrt{E\|x_n - z\|_H^2 + 2E\mu_n E(\|x_n - z\|_H\|e_n\|_H) + E\mu_n^2 E\|e_n\|_H^2} \\
&\leqslant \sqrt{E\|x_n - z\|_H^2 + 2\sqrt{E\mu_n^2}\sqrt{E\|x_n - z\|_H^2}\sqrt{\|e_n\|_H^2} + E\mu_n^2 E\|e_n\|_H^2} \\
&= \sqrt{\left|\sqrt{E\|x_n - z\|_H^2} + \sqrt{E\mu_n^2 E\|e_n\|_H^2}\right|^2} \\
&= \|x_n - z\|_{L^2(\Omega,\mathcal{F},P;H)} + \sqrt{E\mu_n^2 E\|e_n\|_H^2}.
\end{aligned} \tag{4.8}
$$

(i)(a): We derive from (4.7) and Lemma 2.11(ii) that $(\|x_n\|_H)_{n\in\mathbb{N}}$ is bounded P-a.s. Hence, for every $z \in Z$, $\sum_{n\in\mathbb{N}}\|x_n - z\|_H E\mu_n\sqrt{E(\|e_n\|_H^2 \,|\, \mathcal{X}_n)} < +\infty$ P-a.s. On the other hand, the assumptions $\lim E(\|e_n\|_H^2 \,|\, \mathcal{X}_n) = 0$ and $\sum_{n\in\mathbb{N}} E\mu_n\sqrt{E(\|e_n\|_H^2 \,|\, \mathcal{X}_n)} < +\infty$ P-a.s. yield

$$
\sum_{n\in\mathbb{N}} E\mu_n E(\|e_n\|_H^2 \,|\, \mathcal{X}_n) < +\infty \quad \text{P-a.s.} \tag{4.9}
$$

Therefore, for every $z \in Z$, $\sum_{n\in\mathbb{N}} \varepsilon_n(\cdot, z)E\lambda_n = 2\sum_{n\in\mathbb{N}} \varepsilon_n(\cdot, z)E\mu_n < +\infty$ P-a.s. It then follows from Theorem 3.6(i)(a) and the assumption $\sum_{n\in\mathbb{N}} E(\mu_n(1-\mu_n)) = +\infty$ that $\underline{\lim}\, E(\|d_n\|_H^2 \,|\, \mathcal{X}_n) = 0$ P-a.s. Hence,

$$
\begin{aligned}
0 &\leqslant \frac{1}{2}\,\underline{\lim}\|Tx_n - x_n\|_H^2 \\
&\leqslant \underline{\lim}\, E\big(\|Tx_n + e_n - x_n\|_H^2 + \|e_n\|_H^2 \,\big|\, \mathcal{X}_n\big) \\
&= \underline{\lim}\, E\big(\|Tx_n + e_n - x_n\|_H^2 \,\big|\, \mathcal{X}_n\big) + \lim E(\|e_n\|_H^2 \,|\, \mathcal{X}_n) \\
&= 4\,\underline{\lim}\, E\big(\|d_n\|_H^2 \,\big|\, \mathcal{X}_n\big) + \lim E(\|e_n\|_H^2 \,|\, \mathcal{X}_n) \\
&= 0 \quad \text{P-a.s.}
\end{aligned} \tag{4.10}
$$

Thus, Lemma 2.7 implies that, for every $n \in \mathbb{N}$,

$$
\begin{aligned}
&E\big(\|Tx_{n+1} - x_{n+1}\|_H \,\big|\, \mathcal{X}_n\big) \\
&\quad= E\big(\|Tx_{n+1} - Tx_n + (1-\mu_n)(Tx_n - x_n) - \mu_n e_n\|_H \,\big|\, \mathcal{X}_n\big) \\
&\quad\leqslant E\big(\|Tx_{n+1} - Tx_n\|_H \,\big|\, \mathcal{X}_n\big) + E\big((1-\mu_n)\|Tx_n - x_n\|_H \,\big|\, \mathcal{X}_n\big) + E\big(\mu_n\|e_n\|_H \,\big|\, \mathcal{X}_n\big) \\
&\quad\leqslant E\big(\|x_{n+1} - x_n\|_H \,\big|\, \mathcal{X}_n\big) + (1-E\mu_n)\|Tx_n - x_n\|_H + E\big(\mu_n\|e_n\|_H \,\big|\, \mathcal{X}_n\big) \\
&\quad= E\big(\mu_n\|Tx_n + e_n - x_n\|_H \,\big|\, \mathcal{X}_n\big) + (1-E\mu_n)\|Tx_n - x_n\|_H + E\big(\mu_n\|e_n\|_H \,\big|\, \mathcal{X}_n\big) \\
&\quad= E\big(\mu_n\|Tx_n - x_n\|_H \,\big|\, \mathcal{X}_n\big) + (1-E\mu_n)\|Tx_n - x_n\|_H + 2E\big(\mu_n\|e_n\|_H \,\big|\, \mathcal{X}_n\big) \\
&\quad= (E\mu_n)\|Tx_n - x_n\|_H + (1-E\mu_n)\|Tx_n - x_n\|_H + 2E\mu_n E(\|e_n\|_H \,|\, \mathcal{X}_n) \\
&\quad\leqslant \|Tx_n - x_n\|_H + 2E\mu_n\sqrt{E(\|e_n\|_H^2 \,|\, \mathcal{X}_n)} \quad \text{P-a.s.}
\end{aligned} \tag{4.11}
$$

Consequently, Lemma 2.5(i) secures the convergence P-a.s. of the sequence $(\|\mathsf{T}x_n - x_n\|_{\mathsf{H}})_{n\in\mathbb{N}}$, which, in view of (4.10), forces

$$\lim\|\mathsf{T}x_n - x_n\|_{\mathsf{H}} = 0 \text{ P-a.s.} \tag{4.12}$$

(i)(b): Let us show that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. Let $\omega \in \Omega$ be such that $\mathfrak{W}(x_n(\omega))_{n\in\mathbb{N}} \neq \varnothing$ and $\lim\|\mathsf{T}x_n(\omega) - x_n(\omega)\| = 0$. Let $\mathsf{x} \in \mathfrak{W}(x_n(\omega))_{n\in\mathbb{N}}$, say $x_{k_n}(\omega) \rightharpoonup \mathsf{x}$. The nonexpansiveness of T implies that $\mathrm{Id}-\mathsf{T}$ is demiclosed at 0 [5, Theorem 4.27]. In turn, $\mathsf{T}\mathsf{x} = \mathsf{x}$ and $\mathfrak{W}(x_n(\omega)) \subset \mathsf{Z}$. Since $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \neq \varnothing$ P-a.s. and $\lim\|\mathsf{T}x_n - x_n\| = 0$ P-a.s., we conclude that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. Thus, the claim follows from Theorem 3.6(i)(c).

(i)(c): By (i)(a) and (i)(b), there exists $\Omega' \in \mathcal{F}$ such that $\mathsf{P}(\Omega') = 1$ and

$$(\forall \omega \in \Omega') \quad \mathsf{T}x_n(\omega) - x_n(\omega) \to 0 \ \text{ and } \ x_n(\omega) \rightharpoonup x(\omega). \tag{4.13}$$

It then follows from the demiregularity of $\mathsf{T}-\mathrm{Id}$ that, for every $\omega \in \Omega'$, $x_n(\omega) \to x(\omega)$. Hence, $(x_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to x.

(ii)(a): We derive from (4.8) and Corollary 2.6(i) that $(\|x_n\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{n\in\mathbb{N}}$ is bounded. Therefore,

$$(\forall z \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})) \quad \sup_{n\in\mathbb{N}} \mathsf{E}\|x_n - z\|_{\mathsf{H}}^2 < +\infty. \tag{4.14}$$

In turn, for every $z \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$,

$$\sum_{n\in\mathbb{N}} \mathsf{E}\mu_n\sqrt{\mathsf{E}\|x_n - z\|_{\mathsf{H}}^2}\sqrt{\mathsf{E}\|e_n\|_{\mathsf{H}}^2} \leqslant \sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}\|x_n - z\|_{\mathsf{H}}^2}\sqrt{\mathsf{E}\mu_n^2\mathsf{E}\|e_n\|_{\mathsf{H}}^2} < +\infty. \tag{4.15}$$

On the other hand, since $\lim \mathsf{E}\|e_n\|_{\mathsf{H}}^2 = 0$ and $\sum_{n\in\mathbb{N}} \mathsf{E}\mu_n\sqrt{\mathsf{E}\|e_n\|_{\mathsf{H}}^2} < +\infty$, we have

$$\sum_{n\in\mathbb{N}} \mathsf{E}\mu_n\mathsf{E}\|e_n\|_{\mathsf{H}}^2 < +\infty. \tag{4.16}$$

Altogether, we deduce that

$$(\forall \mathsf{z} \in \mathsf{Z}) \quad \sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n(\cdot,\mathsf{z})\mathsf{E}\lambda_n = 2\sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n(\cdot,\mathsf{z})\mathsf{E}\mu_n < +\infty, \tag{4.17}$$

which shows in particular that, for every $\mathsf{z} \in \mathsf{Z}$, $\sum_{n\in\mathbb{N}} \varepsilon_n(\cdot,\mathsf{z})\mathsf{E}\lambda_n < +\infty$ P-a.s. Thus $\lim\|\mathsf{T}x_n - x_n\|_{\mathsf{H}} = 0$ P-a.s. On the other hand, it follows from Theorem 3.6(ii)(a) and the assumptions that $\underline{\lim}\, \mathsf{E}\|d_n\|_{\mathsf{H}}^2 = 0$. Hence, proceeding as in (4.10), we obtain $\underline{\lim}\, \mathsf{E}\|\mathsf{T}x_n - x_n\|_{\mathsf{H}}^2 = 0$. Moreover, taking expectations in (4.11) yields

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \mathsf{E}\|\mathsf{T}x_{n+1} - x_{n+1}\|_{\mathsf{H}} &\leqslant \mathsf{E}\|\mathsf{T}x_n - x_n\|_{\mathsf{H}} + 2\mathsf{E}\mu_n\sqrt{\mathsf{E}\|e_n\|_{\mathsf{H}}^2}\\ &\leqslant \mathsf{E}\|\mathsf{T}x_n - x_n\|_{\mathsf{H}} + 2\sqrt{\mathsf{E}\mu_n^2\mathsf{E}\|e_n\|_{\mathsf{H}}^2}.\end{aligned} \tag{4.18}$$

It then follows from Corollary 2.6(i) that $(\mathsf{E}\|\mathsf{T}x_n - x_n\|_{\mathsf{H}})_{n\in\mathbb{N}}$ converges. Since $\underline{\lim}\, \mathsf{E}\|\mathsf{T}x_n - x_n\|_{\mathsf{H}}^2 = 0$, this implies that $\lim \mathsf{E}\|\mathsf{T}x_n - x_n\|_{\mathsf{H}}^2 = 0$. Hence, appealing to (i)(a), $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$ converges strongly in $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and strongly P-a.s. to 0.

(ii)(b): We deduce weak convergence P-a.s. by arguing as in the proof of (i)(b), while weak convergence in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ follows from Theorem 3.6(ii)(c).

(ii)(c): As in the proof of (i)(c), it follows from (ii)(b) that $(x_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to x. Further, strong convergence in $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ follows from Theorem 3.6(ii)(d). ☐

Remark 4.2. Theorem 4.1(ii) extends [20, Corollary 2.7], where the relaxations are only deterministic and the weak limit is not shown to be in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. Another connected result is [8, Theorem 2.8], which focuses on the finite-dimensional setting (which implies that demiregularity holds [2, Proposition 2.4]) with deterministic relaxations and the weaker summability condition $\sum_{n\in\mathbb{N}} \mu_n \mathsf{E}(\|e_n\|_{\mathsf{H}} \,|\, \mathcal{X}_n) < +\infty$ P-a.s. The case of deterministic relaxations and deterministic errors was considered in [16, Theorem 5.5(i)], as an extension of the classical result error-free result of [29, Corollary 3].

The following application of Theorem 4.1 concerns averaged operators.

Corollary 4.3. *Let* $\alpha \in \left]0,1\right[$*, let* $\mathsf{T}\colon \mathsf{H} \to \mathsf{H}$ *be an* α*-averaged operator such that* $\operatorname{Fix} \mathsf{T} \neq \varnothing$*, and let* $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$*. Iterate*

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\text{take } e_n \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \text{ and } \mu_n \in L^\infty(\Omega, \mathcal{F}, \mathsf{P}; \left]0, 1/\alpha\right[) \\
x_{n+1} = x_n + \mu_n\big(\mathsf{T}x_n + e_n - x_n\big).
\end{array}
\right.
\end{array}
\tag{4.19}
$$

Set $(\forall n \in \mathbb{N})\ \Phi_n = \{x_0, \ldots, x_n\}$ *and* $\mathcal{X}_n = \sigma(\Phi_n)$*. Suppose that* $\sum_{n\in\mathbb{N}} \mathsf{E}(\mu_n(1-\alpha\mu_n)) = +\infty$ *and, for every* $n \in \mathbb{N}$*, that* μ_n *is independent of* $\sigma(\{e_n\} \cup \Phi_n)$*. Then the following hold for some* $\operatorname{Fix}\mathsf{T}$*-valued random variable* x*:*

(i) *Suppose that* $\mathsf{E}(\|e_n\|^2_{\mathsf{H}} \,|\, \mathcal{X}_n) \to 0$ P-a.s. *and* $\sum_{n\in\mathbb{N}} \mathsf{E}\mu_n \sqrt{\mathsf{E}(\|e_n\|^2_{\mathsf{H}} \,|\, \mathcal{X}_n)} < +\infty$ P-a.s. *Then the following hold:*

  (a) $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to* 0.

  (b) $(x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to* x.

  (c) *Suppose that* $\mathsf{T} - \mathrm{Id}$ *is demiregular at every point in* $\operatorname{Fix}\mathsf{T}$*. Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to* x.

(ii) *Suppose that* $\mathsf{E}\|e_n\|^2_{\mathsf{H}} \to 0$ *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}\mu_n^2 \mathsf{E}\|e_n\|^2_{\mathsf{H}}} < +\infty$*. Then the following hold:*

  (a) $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and strongly* P-a.s. *to* 0.

  (b) $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \operatorname{Fix}\mathsf{T})$ *and* $(x_n)_{n\in\mathbb{N}}$ *converges weakly in* $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and weakly* P-a.s. *to* x.

  (c) *Suppose that* $\mathsf{T} - \mathrm{Id}$ *is demiregular at every point in* $\operatorname{Fix}\mathsf{T}$*. Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and strongly* P-a.s. *to* x.

Proof. Apply Theorem 4.1 to the nonexpansive operator $\mathrm{Id} + \alpha^{-1}(\mathsf{T} - \mathrm{Id})$ and observe that it has the same fixed points as T. ☐

Remark 4.4. As discussed in [17, 18], the Krasnosel'skiĭ–Mann iterative process for averaged operators is at the core of monotone operator splitting strategies such as the three operator splitting scheme of [24], the Douglas–Rachford algorithm [37], and the constant proximal parameter version of the forward-backward algorithm [39]. Stochastically relaxed and perturbed extensions of these algorithms can be derived from Corollary 4.3 with weaker assumptions than those of [21, Theorem 4.1].

We now consider a stochastic version of the (forward) Euler method to find a zero of a cocoercive operator. For simplicity, we adopt deterministic step-sizes $(\gamma_n)_{n\in\mathbb{N}}$. This result extends those of [20, 21, 52] by establishing, under weaker assumptions, weak convergence P-almost surely and, in addition, proving for the first time weak convergence in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$.

Corollary 4.5. *Let* $\beta \in \left]0,+\infty\right[$ *and let* $\mathsf{B}\colon \mathsf{H} \to \mathsf{H}$ *be* β*-cocoercive, with* $\operatorname{zer}\mathsf{B} = \big\{\mathsf{z} \in \mathsf{H} \;\big|\; \mathsf{B}\mathsf{z} = 0\big\} \neq \varnothing$*. Let* $(\mathsf{K}, \mathcal{K})$ *be a measurable space, let* $k\colon (\Omega, \mathcal{F}, \mathsf{P}) \to (\mathsf{K}, \mathcal{K})$ *be a random variable, let* $\xi \in \left]0,+\infty\right[$*, and let*

$(\mathsf{B}_k)_{k\in\mathsf{K}}$ *be operators from* H *to* H *such that* $\mathbf{B}\colon(\mathsf{K}\times\mathsf{H},\mathcal{K}\otimes\mathcal{B}_{\mathsf{H}})\to(\mathsf{H},\mathcal{B}_{\mathsf{H}})\colon(k,x)\mapsto\mathsf{B}_k x$ *is measurable and*

$$(\forall x\in\mathsf{H})\quad \mathsf{E}(\mathsf{B}_k x)=\mathsf{B}x \quad\textit{and}\quad \mathsf{E}\|\mathsf{B}_k x-\mathsf{B}x\|_{\mathsf{H}}^2\leqslant\xi. \tag{4.20}$$

Let $\nu\in\,]2/3,1]$ *and* $x_0\in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. *Iterate*

$$\begin{array}{l}\text{for } n=0,1,\ldots\\ \left\lfloor\begin{array}{l}\mathcal{X}_n=\sigma(x_0,\ldots,x_n)\\ k_n \text{ is a copy of } k \textit{ and is independent of } \mathcal{X}_n\\ \gamma_n=\dfrac{2\beta}{(n+1)^{\nu}}\\ x_{n+1}=x_n-\gamma_n\mathsf{B}_{k_n}x_n.\end{array}\right.\end{array} \tag{4.21}$$

Then the following hold for some $x\in L^2(\Omega,\mathcal{F},\mathsf{P};\operatorname{zer}\mathsf{B})$*:*

(i) $(\mathsf{B}x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *and strongly* P-a.s. *to* 0.

(ii) $(x_n)_{n\in\mathbb{N}}$ *converges weakly in* $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *and weakly* P-a.s. *to* x.

(iii) *Suppose that* B *is demiregular at every point in* $\operatorname{zer}\mathsf{B}$. *Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ *and strongly* P-a.s. *to* x.

Proof. We apply Corollary 4.3 to the reduction technique of [8]. Fix $\theta\in\,]0,2\beta[$ and set

$$(\forall n\in\mathbb{N})\quad e_n=\mathsf{B}x_n-\mathsf{B}_{k_n}x_n=\mathsf{B}x_n-\mathbf{B}\circ(k_n,x_n)\ \text{ and }\ \mu_n=\frac{\gamma_n}{\theta}\in\left]0,\frac{2\beta}{\theta}\right[. \tag{4.22}$$

Then, for every $n\in\mathbb{N}$, e_n is measurable with $\mathsf{E}(e_n\,|\,\mathcal{X}_n)=0$ and $\mathsf{E}\mu_n=\gamma_n/\theta$. Let us also define $e_0'=\gamma_0e_0$ and

$$(\forall n\in\mathbb{N})\quad e_{n+1}'=(1-\gamma_{n+1})e_n'+\gamma_{n+1}e_{n+1}. \tag{4.23}$$

Set $\mathsf{T}=\mathrm{Id}-\theta\mathsf{B}$. Then $\operatorname{Fix}\mathsf{T}=\operatorname{zer}\mathsf{B}$ and T is $\theta/(2\beta)$-averaged [5, Proposition 4.39]. Finally, define, for every $n\in\mathbb{N}$, $y_n=x_n-e_n'$ and $\mathcal{Y}_n=\sigma(y_0,\ldots,y_n)$. Then, we infer from (4.21) that

$$(\forall n\in\mathbb{N})\quad y_{n+1}=y_n+\mu_n\big(\mathsf{T}y_n+e_n''-y_n\big), \tag{4.24}$$

where

$$(\forall n\in\mathbb{N})\quad \begin{cases}e_n''=\mathsf{T}x_n-\mathsf{T}y_n\in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H});\\ \|e_n''\|_{\mathsf{H}}\leqslant\|x_n-y_n\|_{\mathsf{H}}=\|e_n'\|_{\mathsf{H}}\ \ \mathsf{P}\text{-a.s.}\end{cases} \tag{4.25}$$

It follows from the choice of $(\gamma_n)_{n\in\mathbb{N}}$, the uniformly bounded variance in (4.20), and [8, Example 2.7 and Theorem 2.5] that

$$\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}\mu_n^2\mathsf{E}\|e_n'\|_{\mathsf{H}}^2}=\sum_{n\in\mathbb{N}}\frac{\gamma_n}{\theta}\sqrt{\mathsf{E}\|e_n'\|_{\mathsf{H}}^2}<+\infty, \tag{4.26}$$

and

$$\sum_{n\in\mathbb{N}}\mathsf{E}\bigg(\mu_n\Big(1-\frac{\theta}{2\beta}\mu_n\Big)\bigg)=\sum_{n\in\mathbb{N}}\frac{\gamma_n}{\theta}\Big(1-\frac{\gamma_n}{2\beta}\Big)=+\infty. \tag{4.27}$$

We also deduce from the proofs of [8, Lemma 2.4 and Theorem 2.5] that $\mathsf{E}\|e'_n\|_{\mathsf{H}}^2 \to 0$ and $\|e'_n\|_{\mathsf{H}} \to 0$ P-a.s. Consequently, by taking (4.25) into account, we obtain

$$\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}\mu_n^2 \mathsf{E}\|e''_n\|_{\mathsf{H}}^2} < +\infty \quad \text{and} \quad \mathsf{E}\|e''_n\|_{\mathsf{H}}^2 \to 0. \tag{4.28}$$

(i): It follows from Theorem 4.1(ii)(a) that $(\mathsf{B}y_n)_{n\in\mathbb{N}}$ converges strongly in $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and strongly P-a.s. to 0. On the other hand, (4.25) implies that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \theta\|\mathsf{B}x_n - \mathsf{B}y_n\|_{\mathsf{H}} &= \|\theta\mathsf{B}x_n - \theta\mathsf{B}y_n\|_{\mathsf{H}} \\ &\leqslant \|x_n - \theta\mathsf{B}x_n - (y_n - \theta\mathsf{B}y_n)\|_{\mathsf{H}} + \|x_n - y_n\|_{\mathsf{H}} \\ &= \|e''_n\|_{\mathsf{H}} + \|e'_n\|_{\mathsf{H}} \\ &\leqslant 2\|e'_n\|_{\mathsf{H}} \ \ \text{P-a.s.}\end{aligned} \tag{4.29}$$

Since $\mathsf{E}\|e'_n\|_{\mathsf{H}}^2 \to 0$ and $\|e'_n\|_{\mathsf{H}} \to 0$ P-a.s., we deduce that $(\mathsf{B}x_n - \mathsf{B}y_n)_{n\in\mathbb{N}}$ converges strongly in $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and strongly P-a.s. to 0 and, therefore, we obtain the convergence results for $(\mathsf{B}x_n)_{n\in\mathbb{N}}$.

(ii): We infer from Theorem 4.1(ii)(b) that $(y_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and weakly P-a.s. to some $x \in L^2(\Omega,\mathcal{F},\mathsf{P};\operatorname{zer}\mathsf{B})$. However, for every $n \in \mathbb{N}$, $x_n = y_n + e'_n$. Since $(e'_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. and strongly in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ to 0, we conclude that $(x_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and weakly P-a.s. to x.

(iii): This follows from Theorem 4.1(ii)(c) using the same arguments as in the proofs of (i) and (ii). ☐

The following special case of Corollary 4.5 concerns stochastic optimization and establishes new results on the convergence of the iterates generated by the standard stochastic gradient method, a method that goes back to the classical work of [7, 26, 50].

Corollary 4.6. *Let $\beta \in \,]0,+\infty[$ and let $\mathsf{f}\colon \mathsf{H} \to \mathbb{R}$ be convex, differentiable, and such that $\nabla\mathsf{f}$ is $1/\beta$-Lipschitzian, with $\operatorname{Argmin}\mathsf{f} \neq \varnothing$. Let $(\mathsf{K},\mathcal{K})$ be a measurable space, let $k\colon (\Omega,\mathcal{F},\mathsf{P}) \to (\mathsf{K},\mathcal{K})$ be a random variable, let $\xi \in \,]0,+\infty[$, and, for every $\mathsf{k} \in \mathsf{K}$, let $\mathsf{g}_{\mathsf{k}}\colon \mathsf{H} \to \mathbb{R}$ be differentiable and such that $\boldsymbol{\mathsf{B}}\colon (\mathsf{K}\times\mathsf{H},\mathcal{K}\otimes\mathcal{B}_{\mathsf{H}}) \to (\mathsf{H},\mathcal{B}_{\mathsf{H}})\colon (\mathsf{k},\mathsf{x}) \mapsto \nabla\mathsf{g}_{\mathsf{k}}(\mathsf{x})$ is measurable and*

$$(\forall \mathsf{x} \in \mathsf{H}) \quad \mathsf{E}\nabla\mathsf{g}_k(\mathsf{x}) = \nabla\mathsf{f}(\mathsf{x}) \quad \text{and} \quad \mathsf{E}\|\nabla\mathsf{g}_k(\mathsf{x}) - \nabla\mathsf{f}(\mathsf{x})\|_{\mathsf{H}}^2 \leqslant \xi. \tag{4.30}$$

Let $\nu \in \,]2/3,1]$ and $x_0 \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Iterate

$$\begin{array}{l}\text{for } n = 0,1,\ldots \\ \left\lfloor \begin{array}{l} \mathcal{X}_n = \sigma(x_0,\ldots,x_n) \\ k_n \text{ is a copy of } k \text{ and is independent of } \mathcal{X}_n \\ \gamma_n = \dfrac{2\beta}{(n+1)^\nu} \\ x_{n+1} = x_n - \gamma_n\nabla\mathsf{g}_{k_n}(x_n). \end{array}\right.\end{array} \tag{4.31}$$

Then the following hold for some $x \in L^2(\Omega,\mathcal{F},\mathsf{P};\operatorname{Argmin}\mathsf{f})$:

(i) *$(\nabla\mathsf{f}(x_n))_{n\in\mathbb{N}}$ converges strongly in $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and strongly P-a.s. to 0.*

(ii) *$(x_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and weakly P-a.s. to x.*

(iii) *Suppose that $\nabla\mathsf{f}$ is demiregular at every point in $\operatorname{Argmin}\mathsf{f}$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly in $L^1(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and strongly P-a.s. to x.*

Proof. Apply Corollary 4.5 to $\mathsf{B} = \nabla\mathsf{f}$, which is β-cocoercive [5, Corollary 18.17], and, for every $\mathsf{k} \in \mathsf{K}$, $\mathsf{B}_{\mathsf{k}} = \nabla\mathsf{g}_{\mathsf{k}}$. ☐

Remark 4.7. In Corollary 4.6(ii), the weak convergence P-a.s. and in L^2 results are new. In a finite-dimensional setting, we recover the P-a.s. convergence of [8, Corollary 4.5] with the novelty of the L^1 convergence. In the infinite-dimensional setting, we extend the result of [53] where the P-a.s. weak convergence is stated only for a subsequence of the iterates.

Remark 4.8. Variants of Corollary 4.5 can be explored by modifying the probabilistic assumptions in (4.20). In the context of Corollary 4.6, see for instance [13, 32, 46] and their bibliographies for possible candidates.

§5. Application to common fixed point problems

The problem under consideration is a common fixed point problem involving an arbitrary family of firmly quasinonexpansive operators. Recall that $T\colon H \to H$ is firmly quasinonexpansive [5, Definition 4.1(iv)] if

$$(\forall x \in H)(\forall y \in \operatorname{Fix} T) \quad \|Tx - y\|_H^2 + \|Tx - x\|_H^2 \leqslant \|x - y\|_H^2. \tag{5.1}$$

Example 5.1 ([4, Proposition 2.3]). Let $T\colon H \to H$. Then T is firmly quasinonexpansive if one of the following holds:

(i) C is a nonempty closed convex subset of H and $T = \operatorname{proj}_C$ is the projector onto C. Here, $\operatorname{Fix} T = C$.

(ii) $f\colon H \to]-\infty, +\infty]$ is a proper lower semicontinuous convex function and

$$T = \operatorname{prox}_f\colon H \to H\colon x \mapsto \underset{y \in H}{\operatorname{argmin}}\Big(f(y) + \frac{1}{2}\|x - y\|_H^2\Big). \tag{5.2}$$

Here, $\operatorname{Fix} T = \operatorname{Argmin} f$.

(iii) $A\colon H \to 2^H$ is maximally monotone and $T = J_A = (\operatorname{Id} + A)^{-1}$. Here, $\operatorname{Fix} T = \big\{z \in H \mid 0 \in Az\big\}$.

(iv) $f\colon H \to \mathbb{R}$ is a continuous convex function, $s\colon H \to H\colon x \mapsto s(x) \in \partial f(x)$ is a selection of ∂f, and

$$T = G_f\colon H \to H\colon x \mapsto \begin{cases} x - \dfrac{f(x)}{\|s(x)\|_H^2} s(x), & \text{if } f(x) > 0;\\ x, & \text{if } f(x) \leqslant 0, \end{cases} \tag{5.3}$$

is the subgradient projector onto $\operatorname{Fix} T = \big\{x \in H \mid f(x) \leqslant 0\big\}$.

The following formulation covers a wide range of problems in mathematics and its applications [11, 14, 16].

Problem 5.2. Let $(K, \mathcal{K})$ be a measurable space and $(T_k)_{k \in K}$ a family of firmly quasinonexpansive operators such that $\mathbf{T}\colon (K \times H, \mathcal{K} \otimes \mathcal{B}_H) \to (H, \mathcal{B}_H)\colon (k, x) \mapsto T_k x$ is measurable and, for every $k \in K$, $\operatorname{Id} - T_k$ is demiclosed at 0. Let $k\colon (\Omega, \mathcal{F}, P) \to (K, \mathcal{K})$ be a random variable. The task is to

$$\text{find} \;\; x \in Z = \big\{z \in H \mid z \in \operatorname{Fix} T_k \text{ P-a.s.}\big\}, \tag{5.4}$$

under the assumption that $Z \neq \varnothing$.

Remark 5.3. Z is a closed convex subset of H. Indeed, let $(z_n)_{n \in \mathbb{N}}$ be a sequence in Z that converges to $z \in H$. For every $n \in \mathbb{N}$, let $\Omega_n \in \mathcal{F}$ be such that $P(\Omega_n) = 1$ and, for every $\omega \in \Omega_n$, let $z_n \in \operatorname{Fix} T_{k(\omega)}$. Set $\Omega' = \bigcap_{n \in \mathbb{N}} \Omega_n$. Then $P(\Omega') = 1$ and

$$(\forall \omega \in \Omega')(\forall n \in \mathbb{N}) \quad z_n \in \operatorname{Fix} T_{k(\omega)}. \tag{5.5}$$

Since each set of fixed points is closed [16, Proposition 2.3(v)], we deduce that, for every $\omega \in \Omega'$, $z \in \operatorname{Fix} \mathsf{T}_{k(\omega)}$, i.e., $z \in Z$. So Z is closed. Likewise, let $z_1 \in Z$, $z_2 \in Z$, and $\alpha \in \left]0,1\right[$. Define almost sure events $\Omega_1 \in \mathcal{F}$ and $\Omega_2 \in \mathcal{F}$ as above. Then, it follows from the convexity of each set of fixed points [16, Proposition 2.3(v)] that

$$(\forall \omega \in \Omega_1 \cap \Omega_2) \quad \alpha z_1 + (1-\alpha) z_2 \in \operatorname{Fix} \mathsf{T}_{k(\omega)}. \tag{5.6}$$

Since $\mathsf{P}(\Omega_1 \cap \Omega_2) = 1$, we get $\alpha z_1 + (1-\alpha) z_2 \in Z$, which shows that Z is convex.

We propose the following stochastic variant of the extrapolated parallel block-iterative fixed point algorithm of [16]. It introduces stochasticity at four levels:

- The operators are indexed on a general measurable space rather than a countable set.
- The block of activated operators is randomly selected at each iteration.
- The evaluations of the operators at iteration n are averaged and extrapolated with random weights $(\beta_{i,n})_{1 \leqslant i \leqslant M}$.
- The relaxation parameter λ_n at iteration n is random and not confined to the interval $\left]0,2\right[$ as in traditional fixed point methods [5, 16, 25].

Theorem 5.4. *In the setting of Problem 5.2, let* $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$, $0 < M \in \mathbb{N}$, $\delta \in \left]0, 1/M\right[$, *and* $\rho \in \left[2, +\infty\right[$. *Iterate*

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\mathcal{X}_n = \sigma(x_0, \ldots, x_n) \\
\text{for } i = 1, \ldots, M \\
\left\lfloor
\begin{array}{l}
k_{i,n} \text{ is a copy of } k \text{ and is independent of } \mathcal{X}_n \\
p_{i,n} = \mathsf{T}_{k_{i,n}} x_n
\end{array}
\right. \\
(\beta_{i,n})_{1 \leqslant i \leqslant M} \textit{ are } [0,1]\text{-valued random variables such that} \\
\qquad \sum_{i=1}^{M} \beta_{i,n} = 1 \ \mathsf{P}\text{-a.s. and } (\forall i \in \{1, \ldots, M\})\ \beta_{i,n} \geqslant \delta 1_{[\|p_{i,n} - x_n\|_{\mathsf{H}} = \max_{1 \leqslant j \leqslant M} \|p_{j,n} - x_n\|_{\mathsf{H}}]} \\
p_n = \sum_{i=1}^{M} \beta_{i,n} p_{i,n} \\
L_n = \dfrac{\sum_{i=1}^{M} \beta_{i,n} \|p_{i,n} - x_n\|_{\mathsf{H}}^2 + 1_{[p_n = x_n]}}{\|p_n - x_n\|_{\mathsf{H}}^2 + 1_{[p_n = x_n]}} \\
a_n = x_n + L_n (p_n - x_n) \\
\text{take } \lambda_n \in L^\infty(\Omega, \mathcal{F}, \mathsf{P}; \left]0, \rho\right]) \\
x_{n+1} = x_n + \lambda_n (a_n - x_n).
\end{array}
\right.
\end{array}
\tag{5.7}
$$

Suppose that there exists $\mu \in \left]0,1\right[$ *such that* $\inf_{n \in \mathbb{N}} \mathsf{E}(\lambda_n(2 - \lambda_n)) \geqslant \mu$ *and that, for every* $n \in \mathbb{N}$, λ_n *is independent of* $\sigma(p_{1,n}, \ldots, p_{M,n}, \beta_{1,n}, \ldots, \beta_{M,n}, x_0, \ldots, x_n)$. *Then the following hold for some* $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; Z)$:

(i) $(x_n)_{n \in \mathbb{N}}$ *converges weakly in* $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and weakly* P-a.s. *to* x.

(ii) *Suppose that there exists* $S \in \mathcal{F}$ *such that*

$$S \subset \big\{ \omega \in \Omega \mid \mathsf{T}_{k(\omega)} - \mathrm{Id} \textit{ is demiregular at every point in } \operatorname{Fix} \mathsf{T}_{k(\omega)} \big\} \quad \textit{and} \quad \mathsf{P}(S) > 0. \tag{5.8}$$

Then $(x_n)_{n \in \mathbb{N}}$ *converges strongly in* $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and strongly* P-a.s. *to* x.

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

[A] *There exists* $\chi \in \left]0,1\right[$ *such that*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}\big(d_Z^2(x_{n+1}) \,\big|\, \mathcal{X}_n\big) \leqslant \chi d_Z^2(x_n) \ \ \mathsf{P}\text{-a.s.} \tag{5.9}$$

[B] T *is linearly regular in the sense that there exists* $\nu \in [1,+\infty[$ *such that*

$$(\forall x \in H) \quad d_Z^2(x) \leqslant \nu \mathsf{E}\|T_k x - x\|_H^2 = \nu \int_\Omega \big\|T_{k(\omega)} x - x\big\|_H^2 \mathsf{P}(d\omega), \tag{5.10}$$

in which case we set $\zeta = \inf_{j\in\mathbb{N}} \mathsf{E}\lambda_j^2$ *and* $\chi = 1 - \mu\delta\zeta/(\rho^2\nu)$.

Then $(x_n)_{n\in\mathbb{N}}$ *converges strongly in* $L^2(\Omega, \mathcal{F}, \mathsf{P}; H)$ *and strongly* P-a.s. *to* x, *and*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}\|x_n - x\|_H^2 \leqslant 4\chi^n \mathsf{E} d_Z^2(x_0). \tag{5.11}$$

Proof. We define

$$(\forall n \in \mathbb{N}) \quad \begin{cases} t_n^* = x_n - p_n; \\ \eta_n = \displaystyle\sum_{i=1}^M \beta_{i,n} \langle p_{i,n} \mid x_n - p_{i,n}\rangle_H; \\ \alpha_n = \dfrac{1_{[t_n^*\neq 0]} 1_{[\langle x_n \mid t_n^*\rangle_H > \eta_n]} \big(\langle x_n \mid t_n^*\rangle_H - \eta_n\big)}{\|t_n^*\|_H^2 + 1_{[t_n^*=0]}}; \\ \varepsilon_n = 0 \ \text{ P-a.s.}; \\ d_n = x_n - a_n \end{cases} \tag{5.12}$$

and shall show that, in this setting, the sequence $(x_n)_{n\in\mathbb{N}}$ constructed by (5.7) corresponds to one generated by Algorithm 3.4. Let $n \in \mathbb{N}$. We first infer from (5.12) and (5.7) that

$$\begin{aligned} d_n &= x_n - a_n \\ &= L_n\big(x_n - p_n\big) \\ &= \frac{\sum_{i=1}^M \beta_{i,n}\|p_{i,n} - x_n\|_H^2 + 1_{[p_n = x_n]}}{\|p_n - x_n\|_H^2 + 1_{[p_n=x_n]}}\big(x_n - p_n\big) \\ &= \frac{\sum_{i=1}^M \beta_{i,n}\|x_n - p_{i,n}\|_H^2 + 1_{[t_n^*=0]}}{\|t_n^*\|_H^2 + 1_{[t_n^*=0]}} t_n^* \\ &= \frac{\sum_{i=1}^M \beta_{i,n}\big(\langle x_n \mid x_n - p_{i,n}\rangle_H - \langle p_{i,n} \mid x_n - p_{i,n}\rangle_H\big)}{\|t_n^*\|_H^2 + 1_{[t_n^*=0]}} t_n^* \\ &= \frac{\langle x_n \mid t_n^*\rangle_H - \sum_{i=1}^M \beta_{i,n}\langle p_{i,n} \mid x_n - p_{i,n}\rangle_H}{\|t_n^*\|_H^2 + 1_{[t_n^*=0]}} t_n^* \\ &= \alpha_n t_n^* \ \text{ P-a.s.} \end{aligned} \tag{5.13}$$

Next, let us show that

$$L_n \geqslant 1 \ \text{ P-a.s.} \tag{5.14}$$

Fix $z \in Z$ and, for every $i \in \{1, \ldots, M\}$, let $\Omega_{i,n} \in \mathcal{F}$ be such that

$$\mathsf{P}(\Omega_{i,n}) = 1 \quad \text{and} \quad (\forall \omega \in \Omega_{i,n}) \quad z \in \operatorname{Fix} T_{k_{i,n}(\omega)}. \tag{5.15}$$

Thanks to (5.7), we then choose $\Omega_n \in \mathcal{F}$ such that

$$\mathsf{P}(\Omega_n) = 1 \quad \text{and} \quad (\forall \omega \in \Omega_n) \quad \bigcap_{1\leqslant i\leqslant M} \operatorname{Fix} T_{k_{i,n}(\omega)} \neq \varnothing \quad \text{and} \quad \sum_{i=1}^M \beta_{i,n}(\omega) = 1. \tag{5.16}$$

Given $\omega \in \Omega_n$, we consider the following two cases:

- Suppose that $p_n(\omega) = x_n(\omega)$. Then [16, Proposition 2.4] yields $x_n(\omega) \in \operatorname{Fix}(\sum_{i=1}^M \beta_{i,n}(\omega)\mathsf{T}_{k_{i,n}(\omega)}) = \bigcap_{1\leqslant i\leqslant M} \operatorname{Fix} \mathsf{T}_{k_{i,n}(\omega)}$, hence, $(\forall i \in \{1,\ldots,M\})$ $x_n(\omega) = p_{i,n}(\omega)$. Thus,

$$L_n(\omega) = \frac{\sum_{i=1}^M \beta_{i,n}(\omega)\|p_{i,n}(\omega) - x_n(\omega)\|_{\mathsf{H}}^2 + 1_{[p_n=x_n]}(\omega)}{\|p_n(\omega) - x_n(\omega)\|_{\mathsf{H}}^2 + 1_{[p_n=x_n]}(\omega)} = \frac{1_{[p_n=x_n]}(\omega)}{1_{[p_n=x_n]}(\omega)} = 1. \tag{5.17}$$

- Suppose that $p_n(\omega) \neq x_n(\omega)$. Then it follows from the convexity of $\|\cdot\|_{\mathsf{H}}^2$ that

$$0 < \|p_n(\omega) - x_n(\omega)\|_{\mathsf{H}}^2 = \left\|\sum_{i=1}^M \beta_{i,n}(\omega)\big(p_{i,n}(\omega) - x_n(\omega)\big)\right\|_{\mathsf{H}}^2 \leqslant \sum_{i=1}^M \beta_{i,n}(\omega)\big\|p_{i,n}(\omega) - x_n(\omega)\big\|_{\mathsf{H}}^2, \tag{5.18}$$

  which implies that $L_n(\omega) \geqslant 1$.

In view of (1.2), our next task is to show by induction that $(x_n)_{n\in\mathbb{N}}$ and $(t_n^*)_{n\in\mathbb{N}}$ are in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$, and that $(\eta_n)_{n\in\mathbb{N}}$ is in $L^1(\Omega,\mathcal{F},\mathsf{P};\mathbb{R})$. To this end, let $n \in \mathbb{N}$ and $i \in \{1,\ldots,M\}$, and suppose that $x_n \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Then $\mathsf{T}_{k_{i,n}}x_n = \mathbf{T}\circ(k_{i,n}, x_n)$ is measurable. On the other hand, for every $\omega \in \Omega_{i,n}$, $2\mathsf{T}_{k_{i,n}(\omega)} - \mathrm{Id}$ is quasinonexpansive with $\operatorname{Fix}(2\mathsf{T}_{k_{i,n}(\omega)} - \mathrm{Id}) = \operatorname{Fix}\mathsf{T}_{k_{i,n}(\omega)}$ [16, Proposition 2.2(v)] and hence

$$\begin{aligned}
2\big\|p_{i,n}(\omega)\big\|_{\mathsf{H}}^2 &= \frac{1}{2}\big\|2\mathsf{T}_{k_{i,n}(\omega)}x_n(\omega)\big\|_{\mathsf{H}}^2 \\
&\leqslant \big\|(2\mathsf{T}_{k_{i,n}(\omega)} - \mathrm{Id})x_n(\omega) - \mathsf{z}\big\|_{\mathsf{H}}^2 + \big\|x_n(\omega) + \mathsf{z}\big\|_{\mathsf{H}}^2 \\
&\leqslant \big\|x_n(\omega) - \mathsf{z}\big\|_{\mathsf{H}}^2 + \big\|x_n(\omega) + \mathsf{z}\big\|_{\mathsf{H}}^2.
\end{aligned} \tag{5.19}$$

Consequently, since $x_n \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and $\mathsf{z} \in \mathsf{H}$, we have $p_{i,n} \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ and (5.7) therefore yields $p_n \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. Thus, $t_n^* = x_n - p_n \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. On the other hand, it follows from the Cauchy–Schwarz inequalities in H as well in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathbb{R})$ that

$$\mathsf{E}\Big|\langle p_{i,n} \mid x_n - p_{i,n}\rangle_{\mathsf{H}}\Big| \leqslant \mathsf{E}\Big(\big\|p_{i,n}\big\|_{\mathsf{H}}\big\|x_n - p_{i,n}\big\|_{\mathsf{H}}\Big) \leqslant \sqrt{\mathsf{E}\big\|p_{i,n}\big\|_{\mathsf{H}}^2\mathsf{E}\big\|x_n - p_{i,n}\big\|_{\mathsf{H}}^2} < +\infty, \tag{5.20}$$

which shows that $\big\langle p_{i,n} \,\big|\, x_n - p_{i,n}\big\rangle_{\mathsf{H}} \in L^1(\Omega,\mathcal{F},\mathsf{P};\mathbb{R})$. Since this is true for every $i \in \{1,\ldots,M\}$, we obtain $\eta_n \in L^1(\Omega,\mathcal{F},\mathsf{P};\mathbb{R})$. Further, it follows from [5, Proposition 4.2(iv)] that

$$(\forall i \in \{1,\ldots,M\}) \quad \big\langle p_{i,n} - \mathsf{z} \,\big|\, x_n - p_{i,n}\big\rangle_{\mathsf{H}} = \big\langle \mathsf{T}_{k_{i,n}}x_n - \mathsf{z} \,\big|\, x_n - \mathsf{T}_{k_{i,n}}x_n\big\rangle_{\mathsf{H}} \geqslant 0 \quad \mathsf{P}\text{-a.s.} \tag{5.21}$$

In turn, the concavity of $\mathsf{y} \mapsto \langle \mathsf{y} - \mathsf{z} \mid x_n(\omega) - \mathsf{y}\rangle_{\mathsf{H}}$ yields

$$0 \leqslant \sum_{i=1}^M \beta_{i,n}\big\langle p_{i,n} - \mathsf{z} \,\big|\, x_n - p_{i,n}\big\rangle_{\mathsf{H}} \leqslant \big\langle p_n - \mathsf{z} \,\big|\, x_n - p_n\big\rangle_{\mathsf{H}} = \big\langle x_n - t_n^* - \mathsf{z} \,\big|\, t_n^*\big\rangle_{\mathsf{H}} \quad \mathsf{P}\text{-a.s.} \tag{5.22}$$

and therefore

$$\begin{aligned}
\frac{1}{2}\mathsf{E}\left|\frac{1_{[t_n^*\neq 0]}1_{[\langle x_n \mid t_n^*\rangle_{\mathsf{H}} > \eta_n]}\eta_n}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 &\leqslant \mathsf{E}\left|\frac{\eta_n - \sum_{i=1}^M \beta_{i,n}\langle \mathsf{z} \mid x_n - p_{i,n}\rangle_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 + \mathsf{E}\left|\frac{\sum_{i=1}^M \beta_{i,n}\langle \mathsf{z} \mid x_n - p_{i,n}\rangle_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 \\
&= \mathsf{E}\left|\frac{\sum_{i=1}^M \beta_{i,n}\langle p_{i,n} - \mathsf{z} \mid x_n - p_{i,n}\rangle_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 + \mathsf{E}\left|\frac{\langle \mathsf{z} \mid t_n^*\rangle_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 \\
&\leqslant \mathsf{E}\left|\frac{\big\langle x_n - t_n^* - \mathsf{z} \,\big|\, t_n^*\big\rangle_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 + \mathsf{E}\left|\frac{\|\mathsf{z}\|_{\mathsf{H}}\|t_n^*\|_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 \\
&\leqslant \mathsf{E}\left|\frac{\|x_n - t_n^* - \mathsf{z}\|_{\mathsf{H}}\|t_n^*\|_{\mathsf{H}}}{\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^*=0]}}\right|^2 + \mathsf{E}\|\mathsf{z}\|_{\mathsf{H}}^2 \\
&\leqslant \mathsf{E}\|x_n - t_n^* - \mathsf{z}\|_{\mathsf{H}}^2 + \|\mathsf{z}\|_{\mathsf{H}}^2.
\end{aligned} \tag{5.23}$$

Since $\{x_n, t_n^*\} \subset L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and $\mathsf{z} \in \mathsf{H}$, we thus obtain $1_{[t_n^* \neq 0]} \eta_n / (\|t_n^*\|_{\mathsf{H}} + 1_{[t_n^* = 0]}) \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R})$. Hence, arguing as in (3.23), we deduce from (5.13) that

$$x_n - a_n = \alpha_n t_n^* \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}). \tag{5.24}$$

It therefore results from (5.7) that $x_{n+1} \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$, which completes the induction argument. On the other hand, since $\alpha_n \in [0, +\infty[$ P-a.s., (5.22) yields

$$\langle \mathsf{z} \mid \alpha_n t_n^* \rangle_{\mathsf{H}} = \alpha_n \sum_{i=1}^{M} \beta_{i,n} \langle \mathsf{z} \mid x_n - p_{i,n} \rangle_{\mathsf{H}} \leqslant \alpha_n \sum_{i=1}^{M} \beta_{i,n} \langle p_{i,n} \mid x_n - p_{i,n} \rangle_{\mathsf{H}} = \alpha_n \eta_n \;\; \text{P-a.s.} \tag{5.25}$$

Thus, appealing to Lemma 2.10 and (5.12), we obtain

$$\Big\langle \mathsf{z} \,\Big|\, \mathsf{E}(\alpha_n t_n^* \,|\, \mathcal{X}_n) \Big\rangle_{\mathsf{H}} = \mathsf{E}\Big(\langle \mathsf{z} \mid \alpha_n t_n^* \rangle_{\mathsf{H}} \,\Big|\, \mathcal{X}_n \Big) \leqslant \mathsf{E}(\alpha_n \eta_n \,|\, \mathcal{X}_n) + \varepsilon_n \;\; \text{P-a.s.} \tag{5.26}$$

Altogether, the sequence $(x_n)_{n\in\mathbb{N}}$ constructed by (5.7) corresponds to one generated by Algorithm 3.4. Now set $\zeta = \inf_{j\in\mathbb{N}} \mathsf{E}\lambda_n^2$ and note that $\zeta \geqslant \inf_{j\in\mathbb{N}} \mathsf{E}^2\lambda_j \geqslant \mu^2/4 > 0$. Hence, we infer from (5.7) and Lemma 2.7 that

$$\begin{aligned}
\mathsf{E}(\|x_{n+1} - x_n\|_{\mathsf{H}}^2 \,|\, \mathcal{X}_n) &= \mathsf{E}\Big(\|\lambda_n(a_n - x_n)\|_{\mathsf{H}}^2 \,\Big|\, \mathcal{X}_n \Big) \\
&= \mathsf{E}\Big(\|\lambda_n L_n (p_n - x_n)\|_{\mathsf{H}}^2 \,\Big|\, \mathcal{X}_n \Big) \\
&= \mathsf{E}\Big(|\lambda_n L_n|^2 \|p_n - x_n\|_{\mathsf{H}}^2 \,\Big|\, \mathcal{X}_n \Big) \\
&= \mathsf{E}\bigg(\lambda_n^2 L_n \sum_{i=1}^{M} \beta_{i,n} \|p_{i,n} - x_n\|_{\mathsf{H}}^2 \,\bigg|\, \mathcal{X}_n \bigg) \\
&\geqslant \mathsf{E}\bigg(\lambda_n^2 \sum_{i=1}^{M} \beta_{i,n} \|p_{i,n} - x_n\|_{\mathsf{H}}^2 \,\bigg|\, \mathcal{X}_n \bigg) \\
&= (\mathsf{E}\lambda_n^2)\, \mathsf{E}\bigg(\sum_{i=1}^{M} \beta_{i,n} \|p_{i,n} - x_n\|_{\mathsf{H}}^2 \,\bigg|\, \mathcal{X}_n \bigg) \\
&\geqslant \zeta \mathsf{E}\bigg(\sum_{i=1}^{M} \beta_{i,n} \|p_{i,n} - x_n\|_{\mathsf{H}}^2 \,\bigg|\, \mathcal{X}_n \bigg) \\
&\geqslant \zeta \mathsf{E}\Big(\delta \max_{1\leqslant j\leqslant M} \|p_{j,n} - x_n\|_{\mathsf{H}}^2 \,\Big|\, \mathcal{X}_n \Big) \\
&\geqslant \delta\zeta \mathsf{E}\Big(\|p_{1,n} - x_n\|_{\mathsf{H}}^2 \,\Big|\, \mathcal{X}_n \Big) \\
&= \delta\zeta \mathsf{E}\Big(\|\mathsf{T}_{k_{1,n}} x_n - x_n\|_{\mathsf{H}}^2 \,\Big|\, \mathcal{X}_n \Big).
\end{aligned} \tag{5.27}$$

However, since $k_{1,n}$ is independent of $\mathcal{X}_n$, Lemma 2.8 implies that, for P-almost every $\omega' \in \Omega$,

$$\begin{aligned}
\mathsf{E}\Big(\|\mathsf{T}_{k_{1,n}} x_n - x_n\|_{\mathsf{H}}^2 \,\Big|\, \mathcal{X}_n \Big)(\omega') &= \int_\Omega \|\mathsf{T}_{k_{1,n}(\omega)} x_n(\omega') - x_n(\omega')\|_{\mathsf{H}}^2 \mathsf{P}(d\omega) \\
&= \int_\Omega \|\mathsf{T}_{k(\omega)} x_n(\omega') - x_n(\omega')\|_{\mathsf{H}}^2 \mathsf{P}(d\omega).
\end{aligned} \tag{5.28}$$

Therefore, for P-almost every $\omega' \in \Omega$, (5.27) implies that

$$\mathsf{E}(\|x_{n+1} - x_n\|_{\mathsf{H}}^2 \,|\, \mathcal{X}_n)(\omega') \geqslant \delta\zeta \int_\Omega \|\mathsf{T}_{k(\omega)} x_n(\omega') - x_n(\omega')\|_{\mathsf{H}}^2 \mathsf{P}(d\omega) \;\; \text{P-a.s.} \tag{5.29}$$

Upon taking the expected value in (5.27), summing over $n \in \mathbb{N}$, and invoking Theorem 3.6(ii)(b), we obtain

$$\mathsf{E}\left(\sum_{n\in\mathbb{N}} \int_\Omega \|\mathsf{T}_{k(\omega)}x_n - x_n\|_{\mathsf{H}}^2 \mathsf{P}(d\omega)\right) = \sum_{n\in\mathbb{N}} \mathsf{E}\left(\int_\Omega \|\mathsf{T}_{k(\omega)}x_n - x_n\|_{\mathsf{H}}^2 \mathsf{P}(d\omega)\right) < +\infty. \tag{5.30}$$

Hence,

$$\sum_{n\in\mathbb{N}} \int_\Omega \|\mathsf{T}_{k(\omega)}x_n - x_n\|_{\mathsf{H}}^2 \mathsf{P}(d\omega) < +\infty \quad \mathsf{P}\text{-a.s.} \tag{5.31}$$

Let $\Omega' \in \mathcal{F}$ such that $\mathsf{P}(\Omega') = 1$ and

$$(\forall \omega' \in \Omega') \quad \sum_{n\in\mathbb{N}} \int_\Omega \|\mathsf{T}_{k(\omega)}x_n(\omega') - x_n(\omega')\|_{\mathsf{H}}^2 \mathsf{P}(d\omega) < +\infty \quad \text{and} \quad \mathfrak{W}\big(x_n(\omega')\big)_{n\in\mathbb{N}} \neq \varnothing. \tag{5.32}$$

The existence of such a set Ω' follows from (5.31) as well as Theorem 3.2(v)(a). Fix $\omega' \in \Omega'$ and let $x(\omega') \in \mathfrak{W}(x_n(\omega'))_{n\in\mathbb{N}}$, say $x_{j_n}(\omega') \rightharpoonup x(\omega')$. On the other hand, it follows from the monotone convergence theorem that

$$\int_\Omega \sum_{n\in\mathbb{N}} \|\mathsf{T}_{k(\omega)}x_n(\omega') - x_n(\omega')\|_{\mathsf{H}}^2 \mathsf{P}(d\omega) = \sum_{n\in\mathbb{N}} \int_\Omega \|\mathsf{T}_{k(\omega)}x_n(\omega') - x_n(\omega')\|_{\mathsf{H}}^2 \mathsf{P}(d\omega) < +\infty. \tag{5.33}$$

Hence, for P-almost every $\omega \in \Omega$, $\sum_{n\in\mathbb{N}} \|\mathsf{T}_{k(\omega)}x_n(\omega') - x_n(\omega')\|_{\mathsf{H}}^2 < +\infty$. Therefore, there exists $\Omega'' \in \mathcal{F}$ such that $\mathsf{P}(\Omega'') = 1$ and

$$(\forall \omega \in \Omega'') \quad \mathsf{T}_{k(\omega)}x_n(\omega') - x_n(\omega') \to 0. \tag{5.34}$$

It then follows from the demiclosedness of the operators $(\mathrm{Id} - \mathsf{T}_k)_{k\in\mathsf{K}}$ at 0 that

$$(\forall \omega \in \Omega'') \quad \mathsf{T}_{k(\omega)}x(\omega') = x(\omega'). \tag{5.35}$$

Therefore $x(\omega') \in \big\{\mathsf{z} \in \mathsf{H} \mid \mathsf{z} \in \operatorname{Fix} \mathsf{T}_k \ \mathsf{P}\text{-a.s.}\big\} = \mathsf{Z}$. Since ω' is arbitrarily taken in Ω', we conclude that

$$\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{Z} \ \mathsf{P}\text{-a.s.} \tag{5.36}$$

(i): This follows from (5.36) and Theorems 3.6(i)(c) and 3.6(ii)(c).

(ii): Let $\omega' \in \Omega'$. In view of (5.8), (5.35), and (5.34), there exists $\mathcal{F} \ni \Omega''' \subset \Omega''$ such that $\mathsf{P}(\Omega''') > 0$ and

$$(\forall \omega \in \Omega''') \quad \begin{cases} \mathsf{T}_{k(\omega)} - \mathrm{Id} \text{ is demiregular at } x(\omega'); \\ \mathsf{T}_{k(\omega)}x_n(\omega') - x_n(\omega') \to 0. \end{cases} \tag{5.37}$$

However, (i) implies that, for P-almost every $\omega' \in \Omega$, $x_n(\omega') \rightharpoonup x(\omega')$. Therefore, by demiregularity, for P-almost every $\omega' \in \Omega$, we deduce from (5.37) that $x_n(\omega') \to x(\omega')$. Thus, $(x_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to x. Finally, the strong convergence in $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ follows from Theorem 3.6(ii)(d).

(iii): This follows from Theorem 3.6(ii)(f) when [A] holds. It remains to show that [B] implies [A]. Let us first show that

$$\chi \in \,]0, 1[. \tag{5.38}$$

First, the concavity of $\xi \mapsto \xi(2 - \xi)$ and Jensen's inequality yield

$$0 < \mu \leqslant \inf_{n\in\mathbb{N}} \mathsf{E}\big(\lambda_n(2 - \lambda_n)\big) \leqslant \inf_{n\in\mathbb{N}} \mathsf{E}\lambda_n(2 - \mathsf{E}\lambda_n). \tag{5.39}$$

This quadratic inequality forces

$$0 < 1 - \sqrt{1-\mu} \leqslant \inf_{n\in\mathbb{N}} \mathsf{E}\lambda_n, \tag{5.40}$$

and Jensen's inequality guarantees that $0 < \inf_{n\in\mathbb{N}} \mathsf{E}\lambda_n^2 = \zeta$. Next, since $\mu \in \left]0,1\right[$, $\delta \in \left]0,1\right[$, $\nu \in \left[1,+\infty\right[$, $\rho \in \left[2,+\infty\right[$, and $\lambda_n \in \left]0,\rho\right]$ P-a.s., we have $\lambda_n^2/\rho^2 \in \left]0,1\right]$ P-a.s. and

$$\frac{\zeta}{\rho^2} = \frac{\inf_{n\in\mathbb{N}} \mathsf{E}\lambda_n^2}{\rho^2} \in \left]0,1\right]. \tag{5.41}$$

It follows then that $\mu\delta\zeta/(\rho^2\nu) \in \left]0,1\right[$ and therefore that $\chi = 1-\mu\delta\zeta/(\rho^2\nu) \in \left]0,1\right[$. Next, let $n \in \mathbb{N}$ and $z \in L^2(\Omega,\mathcal{X}_n,\mathsf{P};\mathsf{Z})$. Theorem 3.2(iii), the independence assumption for λ_n, and (1.2) imply that

$$\begin{aligned}\mathsf{E}\big(\|x_{n+1}-z\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big) &\leqslant \|x_n-z\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\big)\mathsf{E}\big(\|d_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big)\\ &= \|x_n-z\|_{\mathsf{H}}^2 - \mathsf{E}\big(\lambda_n(2-\lambda_n)\big)\mathsf{E}\bigg(\frac{1}{\lambda_n^2}\|x_{n+1}-x_n\|_{\mathsf{H}}^2 \,\bigg|\, \mathcal{X}_n\bigg)\;\; \text{P-a.s.}\end{aligned} \tag{5.42}$$

Upon taking $z = \mathrm{proj}_{\mathsf{Z}}\, x_n$ in (5.42),

$$\begin{aligned}\mathsf{E}\big(d_{\mathsf{Z}}^2(x_{n+1}) \,\big|\, \mathcal{X}_n\big) &\leqslant d_{\mathsf{Z}}^2(x_n) - \mathsf{E}\big(\lambda_n(2-\lambda_n)\big)\mathsf{E}\bigg(\frac{1}{\lambda_n^2}\|x_{n+1}-x_n\|_{\mathsf{H}}^2 \,\bigg|\, \mathcal{X}_n\bigg)\\ &\leqslant d_{\mathsf{Z}}^2(x_n) - \mu\mathsf{E}\bigg(\frac{1}{\lambda_n^2}\|x_{n+1}-x_n\|_{\mathsf{H}}^2 \,\bigg|\, \mathcal{X}_n\bigg)\\ &\leqslant d_{\mathsf{Z}}^2(x_n) - \frac{\mu}{\rho^2}\mathsf{E}\big(\|x_{n+1}-x_n\|_{\mathsf{H}}^2 \,\big|\, \mathcal{X}_n\big).\end{aligned} \tag{5.43}$$

Thus, for P-almost every $\omega' \in \Omega$, we derive from (5.27) that

$$\begin{aligned}\mathsf{E}\big(d_{\mathsf{Z}}^2(x_{n+1}) \,\big|\, \mathcal{X}_n\big)(\omega') &\leqslant d_{\mathsf{Z}}^2(x_n)(\omega') - \frac{\mu\delta\zeta}{\rho^2}\int_\Omega \big\|\mathsf{T}_{k(\omega)}x_n(\omega') - x_n(\omega')\big\|_{\mathsf{H}}^2 \mathsf{P}(d\omega)\\ &\leqslant \chi d_{\mathsf{Z}}^2(x_n)(\omega').\end{aligned} \tag{5.44}$$

Hence, $\mathsf{E}(d_{\mathsf{Z}}^2(x_{n+1}) \,|\, \mathcal{X}_n) \leqslant \chi d_{\mathsf{Z}}^2(x_n)$ P-a.s. and, in view of (5.38), [A] holds. The conclusion follows from Theorem 3.6(ii)(f). ☐

Remark 5.5.

(i) In Algorithm (5.7), M is the batch size, i.e., the number of activated sets, p_n is the standard average of the selected operators, $L_n \geqslant 1$ is the extrapolation parameter, a_n is the extrapolated average, and λ_n is the relaxation parameter, which can exceed the standard bound 2 imposed by deterministic methods [16].

(ii) Problem 5.2 is studied in [27] for firmly nonexpansive operators with errors. A deterministic algorithm which activates all the operators at each iteration via a Bochner integral average is proposed. The weak convergence to a solution is established; see also [10] for a version in the context of projectors of Example 5.1(i). This result contrasts with Theorem 5.4 in which the convergence is guaranteed even when a finite number of operators are activated at each iteration.

(iii) In (5.7), we need not impose a lower bound on the weights $(\beta_{i,n})_{1\leqslant i\leqslant \mathsf{M}}$ if we assume that, for every $i \in \{1,\ldots,\mathsf{M}\}$, $\beta_{i,n}$ is independent of $\sigma(p_{i,n},x_0,\ldots,x_n)$. Indeed, in such a case, Lemma 2.7 asserts that

$$
\begin{aligned}
\mathsf{E}\bigg(\sum_{i=1}^{M}\beta_{i,n}\|p_{i,n}-x_n\|_{\mathsf{H}}^2\,\bigg|\,\mathcal{X}_n\bigg) &= \sum_{i=1}^{M}\mathsf{E}\big(\beta_{i,n}\|p_{i,n}-x_n\|_{\mathsf{H}}^2\,\big|\,\mathcal{X}_n\big)\\
&= \sum_{i=1}^{M}(\mathsf{E}\beta_{i,n})\,\mathsf{E}\big(\|p_{1,n}-x_n\|_{\mathsf{H}}^2\,\big|\,\mathcal{X}_n\big)\\
&= \mathsf{E}\big(\|p_{1,n}-x_n\|_{\mathsf{H}}^2\,\big|\,\mathcal{X}_n\big). \qquad (5.45)
\end{aligned}
$$

(iv) Suppose that, for every $k \in K$, $T_k\colon H \to H$ is continuous. Then, to obtain the joint measurability of $\mathbf{T}$, it is enough to suppose that, for every $x \in H$, $\mathbf{T}(\cdot, x)\colon k \mapsto T_k x$ is measurable [1, Lemma 4.51].

Remark 5.6. In the literature, convergence to solutions has been established in specific instances of Problem 5.2 and algorithm (5.7).

(i) Several works have focused on the sequential unrelaxed case, that is, the scenario in which

$$
M = 1,\ \ \lambda_n = 1,\ \text{ and therefore } x_{n+1} = a_n = p_n = p_{1,n} = T_{k_{1,n}} x_n. \qquad (5.46)
$$

In the context of the projectors of Example 5.1(i), [43] guarantees almost sure convergence to a solution when $H = \mathbb{R}^N$ and K is finite. This result is also found in [9] and in [34]. The setting of [34] involves a Euclidean space H and a general measurable space $(K, \mathcal{K})$, and it also shows convergence in $L^2(\Omega, \mathcal{F}, P; H)$. When the subsets are half-spaces or when the interior of Z is nonempty, [43] provides a rate for convergence in $L^2(\Omega, \mathcal{F}, P; H)$. For general separable Hilbert spaces and under the assumption that the operators are averaged mappings, [30] shows weak almost sure convergence. In addition, a convergence rate is established in $L^1(\Omega, \mathcal{F}, P; H)$ when (5.10) is satisfied. The paper [44] involves deterministic relaxations $\lambda_n \in\ ]0, 2[$ in the context of subgradient projectors of Example 5.1(iv) in $H = \mathbb{R}^N$. Assuming that (5.10) holds and, additionally, that the subgradients are uniform bounded, almost sure convergence to a solution is established.

(ii) We now discuss works that have studied algorithms for $M > 1$. In [33], K is countable, extrapolations are not allowed (hence $a_n = p_n$), λ_n is a deterministic parameter in $]0, 2]$, and the condition $\operatorname{int} Z \neq \varnothing$ is imposed. Finite convergence is established. In the context of projectors in $H = \mathbb{R}^N$, a similar approach to Algorithm 1.1 is studied in [40] and [42] with the following restrictions: deterministic relaxations $(\lambda_n)_{n\in\mathbb{N}}$ in $]0, 2[$ and iteration-independent fixed deterministic weights $\beta_{i,n} \equiv 1/M$. Mean-square rates of convergence are established by assuming that (5.10) holds, as well as ergodic convergence results. However, almost sure convergence is not proved. Similarly, [41] and [45] use a deterministic relaxation sequence $(\lambda_n)_{n\in\mathbb{N}}$ in $]0, 2[$ and iteration-independent fixed deterministic weights $\beta_{i,n} \equiv 1/M$ to solve Problem 5.2 in the context of subgradient projectors in $H = \mathbb{R}^N$. Under linear regularity assumptions and, additionally, uniform boundedness of the subgradients, rates of convergence in mean-square are provided. Nevertheless, almost sure convergence of the sequence of iterates is not guaranteed.

Remark 5.7. By combining the proofs to Theorem 4.1 and Theorem 5.4, it is possible to establish convergence results for the following error-tolerant algorithm for solving Problem 5.2: Let $x_0 \in L^2(\Omega, \mathcal{F}, P; H)$,

$0 < \mathsf{M} \in \mathbb{N}$, and $\delta \in]0, 1/\mathsf{M}[$. Iterate

$$
\begin{array}{l}
\text{for } \mathsf{n} = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\mathcal{X}_{\mathsf{n}} = \sigma(x_0, \ldots, x_{\mathsf{n}}) \\
\text{for } \mathsf{i} = 1, \ldots, \mathsf{M} \\
\left\lfloor
\begin{array}{l}
k_{\mathsf{i},\mathsf{n}} \text{ is a copy of } k \text{ and is independent of } \mathcal{X}_{\mathsf{n}} \\
\text{take } e_{\mathsf{i},\mathsf{n}} \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \\
p_{\mathsf{i},\mathsf{n}} = \mathsf{T}_{k_{\mathsf{i},\mathsf{n}}} x_{\mathsf{n}} + e_{\mathsf{i},\mathsf{n}}
\end{array}
\right. \\
(\beta_{\mathsf{i},\mathsf{n}})_{1 \leqslant \mathsf{i} \leqslant \mathsf{M}} \text{ are } [0,1] \text{-valued random variables such that} \\
\quad \sum_{\mathsf{i}=1}^{\mathsf{M}} \beta_{\mathsf{i},\mathsf{n}} = 1 \ \mathsf{P}\text{-a.s. and } (\forall \mathsf{i} \in \{1, \ldots, \mathsf{M}\}) \ \beta_{\mathsf{i},\mathsf{n}} \geqslant \delta 1_{[\|p_{\mathsf{i},\mathsf{n}} - x_{\mathsf{n}}\|_{\mathsf{H}} = \max_{1 \leqslant \mathsf{j} \leqslant \mathsf{M}} \|p_{\mathsf{j},\mathsf{n}} - x_{\mathsf{n}}\|_{\mathsf{H}}]} \\
p_{\mathsf{n}} = \sum_{\mathsf{i}=1}^{\mathsf{M}} \beta_{\mathsf{i},\mathsf{n}} p_{\mathsf{i},\mathsf{n}} \\
\text{take } \lambda_{\mathsf{n}} \in L^\infty(\Omega, \mathcal{F}, \mathsf{P};]0, 2[) \\
x_{\mathsf{n}+1} = x_{\mathsf{n}} + \lambda_{\mathsf{n}}(p_{\mathsf{n}} - x_{\mathsf{n}}).
\end{array}
\right.
\end{array}
\tag{5.47}
$$

Suppose that $\inf_{\mathsf{n}\in\mathbb{N}} \mathsf{E}(\lambda_{\mathsf{n}}(2 - \lambda_{\mathsf{n}})) > 0$, $\max_{1\leqslant \mathsf{i} \leqslant \mathsf{M}} \sum_{\mathsf{n}\in\mathbb{N}} \sqrt{\mathsf{E}\|e_{\mathsf{i},\mathsf{n}}\|_{\mathsf{H}}^2} < +\infty$, and that, for every $\mathsf{n} \in \mathbb{N}$, λ_{n} is independent of $\sigma(k_{1,\mathsf{n}}, \ldots, k_{\mathsf{M},\mathsf{n}}, e_{1,\mathsf{n}}, \ldots, e_{\mathsf{M},\mathsf{n}}, \beta_{1,\mathsf{n}}, \ldots, \beta_{\mathsf{M},\mathsf{n}}, x_0, \ldots, x_{\mathsf{n}})$. Then there exists $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$ such that $(x_{\mathsf{n}})_{\mathsf{n}\in\mathbb{N}}$ converges weakly in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and weakly P-a.s. to x.

§6. Numerical experiments

We illustrate numerically our results in the context of Problem 5.2 with applications of the stochastic algorithm (5.7) with the deterministic relaxation strategies

$$(\forall \mathsf{n} \in \mathbb{N}) \quad \lambda_{\mathsf{n}} = 1.0 \tag{6.1}$$

and

$$(\forall \mathsf{n} \in \mathbb{N}) \quad \lambda_{\mathsf{n}} = 1.9. \tag{6.2}$$

We also consider the random relaxation strategies

$$(\forall \mathsf{n} \in \mathbb{N}) \quad \mathsf{P}([\lambda_{\mathsf{n}} = 2.3]) = \frac{1}{2} \text{ and } \mathsf{P}([\lambda_{\mathsf{n}} = 1.5]) = \frac{1}{2} \tag{6.3}$$

and

$$(\forall \mathsf{n} \in \mathbb{N}) \quad \lambda_{\mathsf{n}} \sim \text{uniform}([1.5, 2.3]). \tag{6.4}$$

Note that (6.3) and (6.4) are super relaxation strategies that satisfy, for every $\mathsf{n} \in \mathbb{N}$, $\mathsf{E}(\lambda_{\mathsf{n}}(2 - \lambda_{\mathsf{n}})) > 0$, $\mathsf{P}([\lambda_{\mathsf{n}} > 2]) > 0$, and $\mathsf{E}\lambda_{\mathsf{n}} = 1.9$. Problem 5.2 is specialized to the standard Euclidean space $\mathsf{H} = \mathbb{R}^{\mathsf{N}}$ with $\|\cdot\|_{\mathsf{H}} = \|\cdot\|$, $\mathsf{K} = \{1, \ldots, \mathsf{p}\}$, and $k \sim \text{uniform}(\mathsf{K})$.

Problem 6.1. For every $\mathsf{k} \in \{1, \ldots, \mathsf{p}\}$, $\mathsf{f}_{\mathsf{k}} \colon \mathbb{R}^{\mathsf{N}} \to \mathbb{R}$ is a convex function and $\mathsf{C}_{\mathsf{k}} = \{\mathsf{x} \in \mathbb{R}^{\mathsf{N}} \mid \mathsf{f}_{\mathsf{k}}(\mathsf{x}) \leqslant 0\}$. It is assumed that $\mathsf{Z} = \bigcap_{1\leqslant \mathsf{k} \leqslant \mathsf{p}} \mathsf{C}_{\mathsf{k}} \neq \varnothing$. The task is to

$$\text{find } \mathsf{x} \in \mathbb{R}^{\mathsf{N}} \text{ such that } \mathsf{x} \in \mathsf{Z}. \tag{6.5}$$

Consider the setting of Problem 6.1. For every $\mathsf{k} \in \{1, \ldots, \mathsf{p}\}$, let $\mathsf{T}_{\mathsf{k}} \colon \mathbb{R}^{\mathsf{N}} \to \mathbb{R}^{\mathsf{N}}$ be the subgradient projector onto C_{k} of Example 5.1(iv), so that, by [5, Propositions 16.20 and 29.41]

$$\mathsf{T}_{\mathsf{k}} \text{ is firmly quasinonexpansive, } \operatorname{Fix} \mathsf{T}_{\mathsf{k}} = \mathsf{C}_{\mathsf{k}}, \text{ and } \mathrm{Id} - \mathsf{T}_{\mathsf{k}} \text{ is demiclosed at } 0. \tag{6.6}$$

Subgradient projectors extend the classical projection operators in the following sense. Let C be a nonempty closed and convex subset of $\mathbb{R}^{\mathsf{N}}$ and suppose that $\mathsf{f}_{\mathsf{k}} = d_{\mathsf{C}}$. Then $\mathsf{C}_{\mathsf{k}} = \mathsf{C}$ and $\mathsf{G}_{\mathsf{k}} = \operatorname{proj}_{\mathsf{C}}$ [5, Example 29.44]. Their importance in solving Problem 6.1 stems from the fact that they are generally much easier to implement than exact ones.

6.1. Signal restoration

The goal is to recover the original signal $\overline{x} \in \mathbb{R}^N$ ($N = 1024$) shown in Fig. 2(a) from 20 noisy observations $(r_k)_{1 \leqslant k \leqslant 20}$ given by

$$(\forall k \in \{1, \ldots, 20\}) \quad r_k = L_k \overline{x} + w_k \tag{6.7}$$

where $L_k : \mathbb{R}^N \to \mathbb{R}^N$ is a known linear operator, $\eta_k \in]0, +\infty[$, and $w_k \in [-\eta_k, \eta_k]^N$ is a bounded random noise vector. The parameters $(\eta_k)_{1 \leqslant k \leqslant 20} \in]0, +\infty[^{20}$ are known. The operators $(L_k)_{1 \leqslant k \leqslant 20}$ are Gaussian convolution filters with zero mean and standard deviation taken uniformly in $[10, 30]$, $\eta_k = 0.15$, and w_k is taken uniformly in $[-\eta_k, \eta_k]^N$. Set, for every $k \in \{1, \ldots, 20\}$ and every $j \in \{1, \ldots, N\}$,

$$C_{k,j} = \big\{ x \in \mathbb{R}^N \mid -\eta_n \leqslant \langle L_k x - r_k \mid e_j \rangle \leqslant \eta_n \big\}. \tag{6.8}$$

Since the intersection of these sets is nonempty and their projectors are computable explicitly [5, Example 29.21], we solve the feasibility problem

$$\text{find } x \in \mathbb{R}^N \text{ such that} \quad (\forall k \in \{1, \ldots, 20\})(\forall j \in \{1, \ldots, N\}) \;\; x \in C_{k,j} \tag{6.9}$$

by algorithm (5.7) implemented with exact projectors. We run two instances with $x_0 = 0$. In the first one, $M = 1$. Note that the relaxation scheme of (6.1) leads to the almost sure convergence result of [43] (see

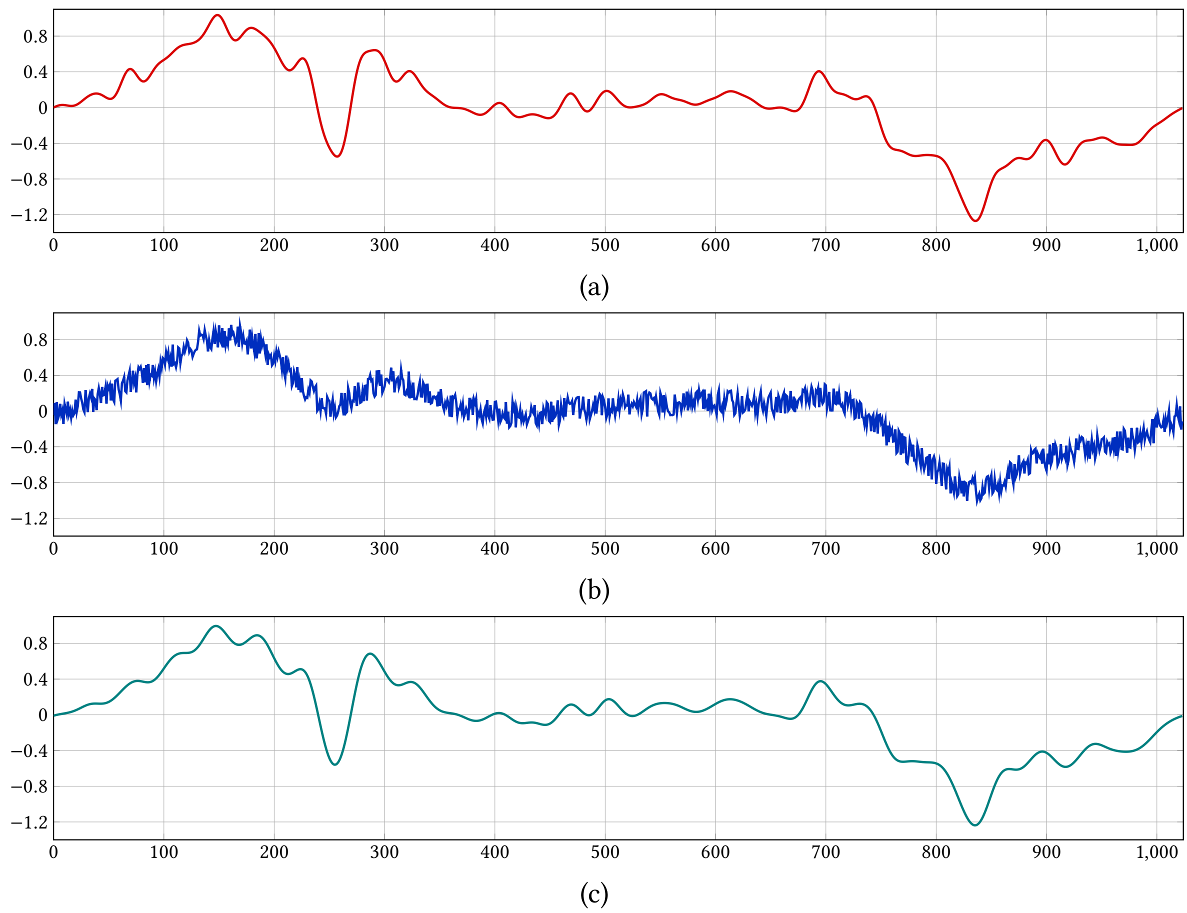

Figure 2: Experiment of Section 6.1. (a): Original signal $\overline{x}$. (b): Noisy observation r_1. (c): Solution produced by algorithm (5.7).

also [34]), while the relaxation schemes (6.2)–(6.4) are new even in this specialized context of randomly activated projection method. In the second instance $M = 16$. Fig. 3 displays the normalized error versus execution time.

Fig. 3 (top) shows the benefits of large relaxations when $M = 1$. Fig. 3 (bottom) shows the advantage of using $M > 1$ random blocks, in which case the extrapolation parameter L_n is not equal to 1 and can attain large values. This behavior was previously observed for deterministic algorithms [6, 12, 15, 48]. Fig. 3 also suggests that, on a single run, the use of the proposed random super relaxation scheme can further improve the speed of convergence. It is worth noting that the execution time can naturally be reduced if Algorithm 1.1 is implemented on a multi-core architecture where, at each iteration, each (subgradient) projector is assigned to a dedicated core and all the cores work in parallel.

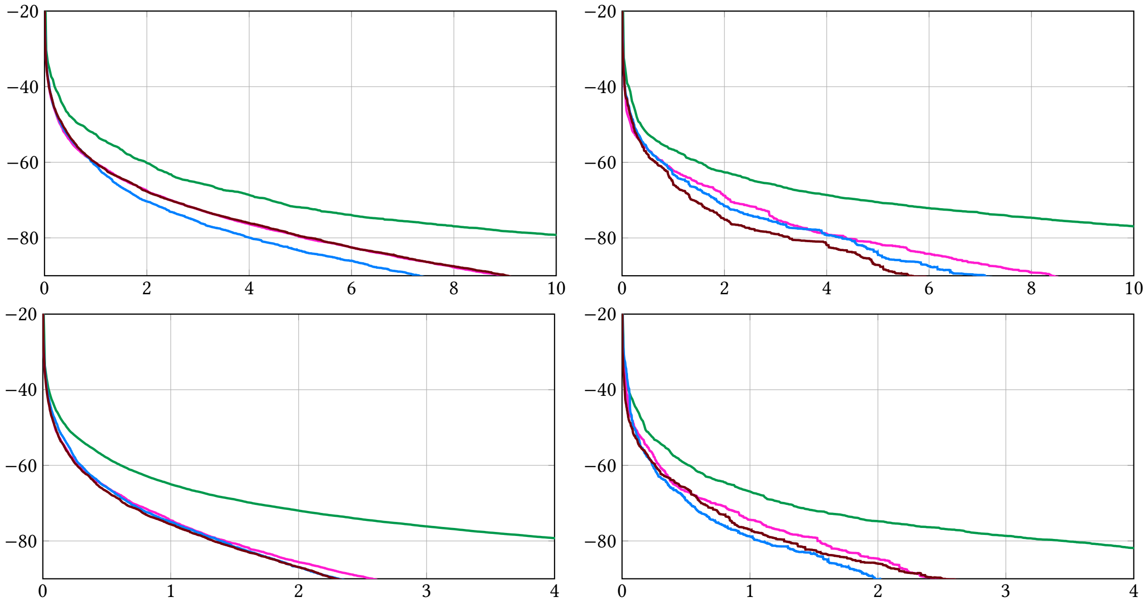

Figure 3: Experiment of Section 6.1. Normalized error $20\log(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus execution time (s) on single processor machine for various relaxation strategies. Green: (6.1). Magenta: (6.2). Blue: (6.3). Brown: (6.4). Left: Average over ten runs. Right: A single run. Top: $M = 1$. Bottom: $M = 16$.

6.2. Image restoration

The goal is to recover the original image $\bar{x} \in \mathbb{R}^{N\times N}$ ($N = 256$) shown in Fig. 4(a) from four observations $\{r_1, \ldots, r_4\}$ which are given by the degradation of $\bar{x}$ via a Gaussian kernel with a standard deviation of 8 and the addition of random noise. The noise distribution is uniform$([0,5]^{N\times N})$. Let L be the block-Toeplitz matrix associated with the convolutional blur. Then

$$(\forall k \in \{1,2,3,4\}) \quad r_k = L\bar{x} + w_k, \quad \text{where} \quad w_k \sim \text{uniform}\big([0,5]^{N\times N}\big). \tag{6.10}$$

The entries of the random vectors $(w_k)_{1\leqslant k\leqslant 4}$ are i.i.d. Therefore, as shown in [23], for every $k \in \{1,2,3,4\}$, with a 95% confidence coefficient

$$\bar{x} \in C_k = \big\{x \in \mathbb{R}^{N\times N} \mid \|r_k - Lx\|^2 \leqslant \xi\big\}, \tag{6.11}$$

where $\xi = N^2 E|u|^2 + 1.96N\sqrt{E|u|^4 - E^2|u|^2}$ with $u \sim$ uniform$([0,5])$. For every $k \in \{1,2,3,4\}$, we compute the subgradient projector onto C_k in (5.3) via the function $f_k\colon x \mapsto \|r_k - Lx\|^2 - \xi$. In addition, the

(a) (b) (c)

Figure 4: Experiment of Section 6.2. (a) Original image $\bar{x}$. (b) Noisy observation r_1. (c) Solution produced by algorithm (5.7).

boundedness on pixel values is incorporated as the property set $C_5 = [0, 255]^{N\times N}$. Finally, it is assumed that the discrete Fourier transform $\mathfrak{F}(\bar{x})$ of $\bar{x}$ is known on a portion of its support for low frequencies in both directions. That is, let S be the set of frequency pairs $\{0, \ldots, N/8 - 1\}^2$ as well as those resulting from the symmetry properties of the 2D discrete Fourier transform of real images. The associated set is $C_6 = \big\{x \in \mathbb{R}^{N\times N} \mid \mathfrak{F}(x)1_S = \mathfrak{F}(\bar{x})1_S\big\}$ and its projection is given by $\operatorname{proj}_{C_6} : x \mapsto \mathfrak{F}^{-1}(\mathfrak{F}(\bar{x})1_S + \mathfrak{F}(x)1_{\complement S})$. We run algorithm (5.7) with $x_0 = 0$ and $M = 2$. Fig. 5 displays the normalized error versus execution time. These results confirm the conclusions of Section 6.1.

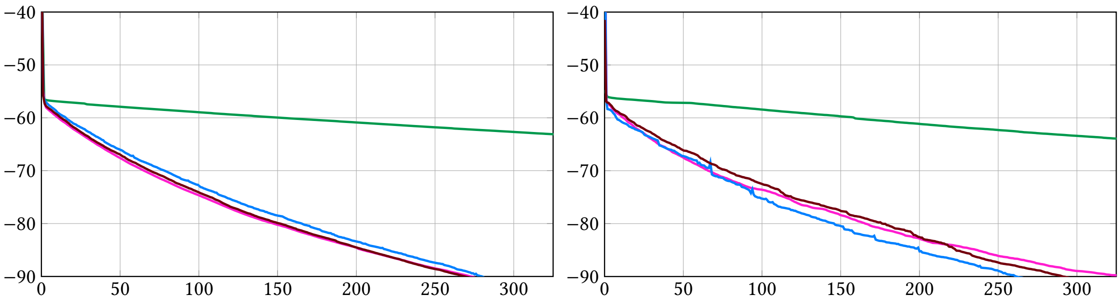

Figure 5: Experiment of Section 6.2 using $M = 2$. Normalized error $20\log(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus execution time (s) on a single processor machine for various relaxation strategies. Green: (6.1). Magenta: (6.2). Blue: (6.3). Brown: (6.4). Left: Average over ten runs. Right: A single run.

Acknowledgment. The authors thank Dr. Minh N. Bùi for bringing to their attention a gap in the proof of an earlier version of Theorem 5.4.

References

[1] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis*, 3rd ed. Springer, New York, 2006.

[2] 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.

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

[4] H. H. Bauschke and P. L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Math. Oper. Res.*, vol. 26, pp. 248–264, 2001.

[5] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Springer, New York, 2017.

[6] H. H. Bauschke, P. L. Combettes, and S. G. Kruk, Extrapolation algorithm for affine-convex feasibility problems, *Numer. Algorithms*, vol. 41, pp. 239–274, 2006.

[7] J. R. Blum, Multidimensional stochastic approximation methods, *Ann. Math. Stat.*, vol. 25, pp. 737–744, 1954.

[8] M. Bravo and R. Cominetti, Stochastic fixed-point iterations for nonexpansive maps: Convergence and error bounds, *SIAM J. Control Optim.*, vol. 62, pp. 191–219, 2024.

[9] L. Briceño-Arias, J. Deride, and C. Vega, Random activations in primal-dual splittings for monotone inclusions with a priori information, *J. Optim. Theory Appl.*, vol. 192, pp. 56–81, 2022.

[10] D. Butnariu and S. D. Flåm, Strong convergence of expected-projection methods in Hilbert spaces, *Numer. Funct. Anal. Optim.*, vol. 16, pp. 601–636, 1995.

[11] C. L. Byrne, *Iterative Optimization in Inverse Problems.* CRC Press, Boca Raton, FL, 2014.

[12] Y. Censor, W. Chen, P. L. Combettes, R. Davidi, and G. T. Herman, On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints, *Comput. Optim. Appl.*, vol. 51, pp. 1065–1088, 2012.

[13] Y.-L. Chen, S. Na, and M. Kolar, Convergence analysis of accelerated stochastic gradient descent under the growth condition, *Math. Oper. Res.*, vol. 49, pp. 2492–2526, 2024.

[14] P. L. Combettes, The convex feasibility problem in image recovery, in: *Advances in Imaging and Electron Physics* (P. Hawkes, Ed.), vol. 95, pp. 155–270. Academic Press, New York, 1996.

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

[16] 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.

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

[18] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numer.*, vol. 33, pp. 487–632, 2024.

[19] P. L. Combettes and J. I. Madariaga, Asymptotic analysis of an abstract stochastic scheme for solving monotone inclusions, arxiv, 2025. https://arxiv.org/pdf/2512.03023

[20] P. L. Combettes and J.-C. Pesquet, Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping, *SIAM J. Optim.*, vol. 25, pp. 1221–1248, 2015.

[21] P. L. Combettes and J.-C. Pesquet, Stochastic approximations and perturbations in forward-backward splitting for monotone operators, *Pure Appl. Funct. Anal.*, vol. 1, pp. 13–37, 2016.

[22] P. L. Combettes and J.-C. Pesquet, Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping II: Mean-square and linear convergence, *Math. Program.*, vol. B174, pp. 433–451, 2019.

[23] P. L. Combettes and H. J. Trussell, The use of noise properties in set theoretic estimation, *IEEE Trans. Signal Process.*, vol. 39, pp. 1630–1641, 1991.

[24] D. Davis and W. Yin, A three-operator splitting scheme and its optimization applications, *Set-Valued Var. Anal.*, vol. 25, pp. 829–858, 2017.

[25] Q.-L. Dong, Y. J. Cho, S. He, P. M. Pardalos, and T. M. Rassias, *The Krasnosel'skiĭ–Mann Iterative Method – Recent Progress and Applications.* Springer, New York, 2022.

[26] Yu. M. Ermol'ev and Z. V. Nekrylova, Some methods of stochastic optimization, *Kibernetika (Kiev)*, vol. 1966, pp. 96–98, 1966.

[27] S. D. Flåm, Successive averages of firmly nonexpansive mappings, *Math. Oper. Res.*, vol. 20, pp. 497–512, 1995.

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

[29] C. W. Groetsch, A note on segmenting Mann iterates, *J. Math. Anal. Appl.*, vol. 40, pp. 369–372, 1972.

[30] N. Hermer, D. R. Luke, and A. Sturm, Random function iterations for consistent stochastic feasibility, *Numer. Funct. Anal. Optim.*, vol. 40, pp. 386–420, 2019.

[31] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach Spaces. Volume I: Martingales and Littlewood–Paley Theory.* Springer, New York, 2016.

[32] A. Jofré and P. Thompson, On variance reduction for stochastic smooth convex optimization with multiplicative noise, *Math. Program.*, vol. A174, pp. 253–292, 2019.

[33] V. I. Kolobov, S. Reich, and R. Zalas, Finitely convergent deterministic and stochastic iterative methods for solving convex feasibility problems, *Math. Program.*, vol. A194, pp. 1163–1183, 2022.

[34] V. R. Kostić and S. Salzo, The method of randomized Bregman projections for stochastic feasibility problems, *Numer. Algorithms*, vol. 93, pp. 1269–1307, 2023.

[35] M. A. Krasnosel'skiĭ, Two remarks on the method of successive approximations, *Uspekhi Mat. Nauk*, vol. 10, pp. 123–127, 1955.

[36] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: Isoperimetry and Processes.* Springer, New York, 1991.

[37] P.-L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.*, vol. 16, pp. 964–979, 1979.

[38] W. R. Mann, Mean value methods in iteration, *Proc. Amer. Math. Soc.*, vol. 4, pp. 506–510, 1953.

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

[40] I. Necoara, Stochastic block projection algorithms with extrapolation for convex feasibility problems, *Optim. Methods Softw.*, vol. 37, pp. 1845–1875, 2022.

[41] I. Necoara and A. Nedić, Minibatch stochastic subgradient-based projection algorithms for feasibility problems with convex inequalities, *Comput. Optim. Appl.*, vol. 80, pp. 121–152, 2021.

[42] I. Necoara, P. Richtárik, and A. Patrascu, Randomized projection methods for convex feasibility: Conditioning and convergence rates, *SIAM J. Optim.*, vol. 29, pp. 2814–2852, 2019.

[43] A. Nedić, Random projection algorithms for convex set intersection problems, *Proc. 49th IEEE Conf. Decis. Control*, pp. 7655–7660, 2010.

[44] A. Nedić, Random algorithms for convex minimization problems, *Math. Program.*, vol. B129, pp. 225–253, 2011.

[45] A. Nedić and I. Necoara, Random minibatch subgradient algorithms for convex problems with functional constraints, *Appl. Math. Optim.*, vol. 80, pp. 801–833, 2019.

[46] L. M. Nguyen, P. H. Nguyen, P. Richtárik, K. Scheinberg, M. Takáč, and M. van Dijk, New convergence aspects of stochastic gradient algorithms, *J. Mach. Learn. Res.*, vol. 20, pp. 1–49, 2019.

[47] B. J. Pettis, On integration in vector spaces, *Trans. Amer. Math. Soc.*, vol. 44, pp. 277–304, 1938.

[48] G. Pierra, Decomposition through formalization in a product space, *Math. Program.*, vol. 28, pp. 96–115, 1984.

[49] S. Reich, Weak convergence theorems for nonexpansive mappings in Banach spaces, *J. Math. Anal. Appl.*, vol. 67, pp. 274–276, 1979.

[50] H. Robbins and S. Monro, A stochastic approximation method, *Ann. Math. Statist.*, vol. 22, pp. 400–407, 1951.

[51] H. Robbins and D. Siegmund, A convergence theorem for non negative almost supermartingales and some applications, in: *Optimizing Methods in Statistics*, (J. S. Rustagi, Ed.), pp. 233–257. Academic Press, New York, 1971.

[52] L. Rosasco, S. Villa, and B. C. Vũ, Stochastic forward-backward splitting for monotone inclusions, *J. Optim. Theory Appl.*, vol. 169, pp. 388–406, 2016.

[53] L. Rosasco, S. Villa, and B. C. Vũ, Convergence of stochastic proximal gradient algorithm, *Appl. Math. Optim.*, vol. 82, pp. 891–917, 2020.

[54] A. N. Shiryaev, *Probability–1*, 3rd ed. Springer, New York, 2016.