Optim Lett (2016) 10:781–803
DOI 10.1007/s11590-015-0904-5

ORIGINAL PAPER

Almost sure convergence of the forward–backward–forward splitting algorithm

Bằng Công Vũ[1,2]

Received: 13 October 2014 / Accepted: 9 May 2015 / Published online: 23 May 2015

Abstract In this paper, we propose a stochastic forward–backward–forward splitting algorithm and prove its almost sure weak convergence in real separable Hilbert spaces. Applications to composite monotone inclusion and minimization problems are demonstrated.

Keywords Monotone inclusion · Monotone operator · Operator splitting · Lipschitzian operators · Forward–backward–forward algorithm · Composite operator · Duality · Primal-dual algorithm

1 Introduction

Forward–backward–forward splitting algorithm was firstly proposed in [24] for solving the problem of finding a zero point of the sum of a maximally monotone operator $A: \mathcal{H} \to 2^{\mathcal{H}}$ and a monotone Lipschitzian operator $C: \mathcal{H} \to \mathcal{H}$, where $\mathcal{H}$ is a real Hilbert space. This splitting algorithm plays a role in solving a large class of composite monotone inclusions [3] and monotone inclusions involving the parallel sums [2,10,11,15] as well as applications to conposite convex optimization problem involving the infimal-convolutions [2–4,11,15]. However, these works are limitted to deterministic setting.

Very recently, we have found out in the literature that there appears the study of some splitting algorithms for solving monotone inclusions in the stochastic setting as in

✉ Bằng Công Vũ
cong.bang@iit.it

1 LCSL, Istituto Italiano di Tecnologia, Genova, Italy

2 Massachusetts Institute of Technology, Bldg. 46-5155, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

[12,19,21], and primal-dual splitting algorithm for composite monotone inclusions in [12,19]. Some iterations in [12,19,21] are designed for monotone inclusions involving cocoercive operators. For solving monotone inclusions involving Lipschitzian monotone operators, one can often use the iterations which has the structure of the forward–backward–forward splitting methods as cited above, but the convergence of their proposed methods is no longer available, in the literature, in the stochastic setting.

The objective of this note is to study the convergence of the forward–backward–forward splitting in the stochastic setting for monotone inclusions involving Lipschitzian monotone operators as well as for composite monotone inclusions involving parallel sums.

In Sect. 2, we recall some notations, background and preliminary results. We prove the almost sure convergence of the stochastic forward–backward–forward splitting algorithm in Sect. 3. In the last section, we provide applications to composite monotone inclusions involving the parallel sums as well as minimization problems involving infimal convolutions.

2 Notation–background and premilary results

Throughout, $\mathcal{H}$, $\mathcal{G}$, and $(\mathcal{G}_i)_{1\leq i\leq m}$ are real separable Hilbert spaces. Their scalar products and associated norms are respectively denoted by $\langle\cdot\mid\cdot\rangle$ and $\|\cdot\|$. We denote by $\mathcal{B}(\mathcal{H},\mathcal{G})$ the space of bounded linear operators from $\mathcal{H}$ to $\mathcal{G}$. The adjoint of $L\in\mathcal{B}(\mathcal{H},\mathcal{G})$ is denoted by L^*. We set $\mathcal{B}(\mathcal{H})=\mathcal{B}(\mathcal{H},\mathcal{H})$. Id denotes the identity operator. The symbols $\rightharpoonup$ and $\to$ denote weak and strong convergence, respectively. We denote by $\ell_+^1(\mathbb{N})$ the set of summable sequences in $[0,+\infty[$. The class of all proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty,+\infty]$ is denoted by $\Gamma_0(\mathcal{H})$. Let M_1 and M_2 be self-adjoint operators in $\mathcal{B}(\mathcal{H})$, we write $M_1\succcurlyeq M_2$ if and only if $(\forall x\in\mathcal{H})\ \langle M_1x\mid x\rangle\geq\langle M_2x\mid x\rangle$. Let $\alpha\in]0,+\infty[$. We set

$$\mathcal{P}_\alpha(\mathcal{H})=\big\{M\in\mathcal{B}(\mathcal{H})\mid M^*=M\quad\text{and}\quad M\succcurlyeq\alpha\,\mathrm{Id}\big\}. \tag{2.1}$$

Let $A\colon\mathcal{H}\to 2^{\mathcal{H}}$ be a set-valued operator. The domain of A is $\operatorname{dom}A=\big\{x\in\mathcal{H}\mid Ax\neq\varnothing\big\}$, and the graph of A is $\operatorname{gra}A=\big\{(x,u)\in\mathcal{H}\times\mathcal{H}\mid u\in Ax\big\}$. The set of zeros of A is $\operatorname{zer}A=\big\{x\in\mathcal{H}\mid 0\in Ax\big\}$, and the range of A is $\operatorname{ran}A=\big\{u\in\mathcal{H}\mid(\exists\,x\in\mathcal{H})\ u\in Ax\big\}$. The inverse of A is $A^{-1}\colon\mathcal{H}\mapsto 2^{\mathcal{H}}\colon u\mapsto\big\{x\in\mathcal{H}\mid u\in Ax\big\}$, and the resolvent of A is

$$J_A=(\mathrm{Id}+A)^{-1}. \tag{2.2}$$

Moreover, A is monotone if

$$(\forall(x,y)\in\mathcal{H}\times\mathcal{H})(\forall(u,v)\in Ax\times Ay)\quad\langle x-y\mid u-v\rangle\geq 0, \tag{2.3}$$

and maximally monotone if it is monotone and there exists no monotone operator $B\colon\mathcal{H}\to 2^{\mathcal{H}}$ such that $\operatorname{gra}A\subset\operatorname{gra}B$ and $A\neq B$. We say that A is uniformly monotone at $x\in\operatorname{dom}A$ if there exists an increasing function $\phi_A\colon[0,+\infty[\to[0,+\infty]$ vanishing only at 0 such that

$$\big(\forall u \in Ax\big)\big(\forall (y,v) \in \operatorname{gra} A\big) \quad \langle x-y \mid u-v\rangle \geq \phi_A(\|x-y\|). \tag{2.4}$$

Given a probability space $(\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}}, \mathsf{P})$, we denote by $\sigma(x)$ the σ-field generated by a random vector $x\colon \boldsymbol{\Omega} \to \mathcal{H}$, where $\mathcal{H}$ is endowed with the Borel σ-algebra. The expectation of a random variable x is denoted by $\mathbf{E}[x]$. The conditional expectation of x given a sub-sigma algebra $\mathcal{F} \subset \boldsymbol{\mathcal{F}}$ is denoted by $\mathbf{E}[x|\mathcal{F}]$. The conditional expectation of x given y is denoted by $\mathbf{E}[x|y]$.

Lemma 2.1 [23, Theorem 1] *Let $(\boldsymbol{\mathcal{F}}_n)_{n\in\mathbb{N}}$ be an increasing sequence of sub-sigma algebras of $\boldsymbol{\mathcal{F}}$. For every $n \in \mathbb{N}$, let z_n, ξ_n, ζ_n and t_n be non-negative, $\boldsymbol{\mathcal{F}}_n$-measurable random variable such that $(\zeta_n)_{n\in\mathbb{N}}$ and $(t_n)_{n\in\mathbb{N}}$ are summable and*

$$(\forall n \in \mathbb{N}) \quad \mathbf{E}[z_{n+1}|\boldsymbol{\mathcal{F}}_n] \leq (1+t_n)z_n + \zeta_n - \xi_n \quad \mathsf{P}-a.s. \tag{2.5}$$

Then $(z_n)_{n\in\mathbb{N}}$ converges and $(\xi_n)_{n\in\mathbb{N}}$ is summable P-a.s.

Lemma 2.2 [12, Proposition 2.3] *Let $\mathcal{H}$ be a real separable Hilbert space, let C be a non-empty closed subset of $\mathcal{H}$, let $\phi\colon [0,\infty[\to [0,\infty[$, let $(x_n)_{n\in\mathbb{N}}$ be a sequence of random vectors in $\mathcal{H}$. Suppose that, for every $x \in C$, there exist non-negative summable sequences of random variables $(\zeta_n(x))_{n\in\mathbb{N}}$ and $(t_n(x))_{n\in\mathbb{N}}$ such that, for every $n \in \mathbb{N}$, $\zeta_n(x)$ and $t_n(x)$ are $\boldsymbol{\mathcal{F}}_n = \sigma(x_0,\ldots,x_n)$-measurable, and*

$$(\forall n \in \mathbb{N}) \quad \mathbf{E}[\phi(\|x_{n+1}-x\|)|\boldsymbol{\mathcal{F}}_n] \leq (1+t_n(x))\phi(\|x_n-x\|) + \zeta_n(x) \quad \mathsf{P}\text{-}a.s. \tag{2.6}$$

Suppose that ϕ is strictly increasing and $\lim_{\xi\to\infty}\phi(\xi) = +\infty$. Then the following hold.

(i) *$(\|x_n - x\|)_{n\in\mathbb{N}}$ is bounded and converges P-a.s.*
(ii) *There exists a subset $\boldsymbol{\Omega}^*$ with $\mathsf{P}(\boldsymbol{\Omega}^*) = 1$ such that for every $x \in C$ and every $\omega \in \boldsymbol{\Omega}^*$, $(\|x_n(\omega) - x\|)_{n\in\mathbb{N}}$ converges.*
(iii) *$(x_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to a C-valued random vector if and only if every its weak cluster point is in C P-a.s.*

Remark 2.3 A sequence $(x_n)_{n\in\mathbb{N}}$ satisfying (2.6) is called a stochastic ϕ-quasi-Fejér monotone with respect to the target set C. The connections of Lemma 2.2 to existing work can be found in [12, Remark 2.4].

In view of the work in [13, Theorem 3.3], we also have a variable metric extension of Lemma 2.2.

Proposition 2.4 *Let $\mathcal{H}$ be a real separable Hilbert space, let C be a non-empty closed subset of $\mathcal{H}$, let $\phi\colon [0,\infty[\to [0,\infty[$, let $\alpha \in]0,\infty[$, let $W \in \mathcal{P}_\alpha(\mathcal{H})$ and $(W_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_\alpha(\mathcal{H})$ such that $W_n \to W$ pointwise, let $(x_n)_{n\in\mathbb{N}}$ be a sequence of random vectors in $\mathcal{H}$. Suppose that, for every $x \in C$, there exist non-negative summable sequences of random variables $(\zeta_n(x))_{n\in\mathbb{N}}$ and $(t_n(x))_{n\in\mathbb{N}}$ such that, for every $n \in \mathbb{N}$, $\zeta_n(x)$ and $t_n(x)$ are $\boldsymbol{\mathcal{F}}_n = \sigma(x_0,\ldots,x_n)$-measurable, and*

$$(\forall n \in \mathbb{N}) \quad \mathbf{E}[\phi(\|x_{n+1}-x\|_{W_{n+1}})|\boldsymbol{\mathcal{F}}_n] \leq (1+t_n(x))\phi(\|x_n-x\|_{W_n}) + \zeta_n(x) \quad \mathsf{P}\text{-}a.s. \tag{2.7}$$

Suppose that ϕ *is strictly increasing and* $\lim_{\xi\to\infty}\phi(\xi) = +\infty$. *Then the following hold.*

(i) $(\|x_n - x\|_{W_n})_{n\in\mathbb{N}}$ *is bounded and converges* P*-a.s.*
(ii) *There exists a subset* $\mathbf{\Omega}^*$ *with* $\mathsf{P}(\mathbf{\Omega}^*) = 1$ *such that for every* $x \in C$ *and every* $\omega \in \mathbf{\Omega}^*$, $(\|x_n(\omega) - x\|_{W_n})_{n\in\mathbb{N}}$ *converges.*
(iii) $(x_n)_{n\in\mathbb{N}}$ *converges weakly* P*-a.s. to a* C*-valued random vector if and only if every its weak cluster point is in* C P*-a.s.*

Proof (i): Set $(\forall n \in \mathbb{N})\ \xi_n = \|x_n - z\|_{W_n}$. It follows from (2.7) and Lemma 2.1 that $(\phi(\xi_n))_{n\in\mathbb{N}}$ converges P-a.s., say $\phi(\xi_n) \to \lambda$. In turn, since $\lim_{t\to+\infty}\phi(t) = +\infty$, $(\xi_n)_{n\in\mathbb{N}}$ is bounded P-a.s. Let $\omega \in \mathbf{\Omega}$ such that $(\xi_n(\omega))_{n\in\mathbb{N}}$ is bounded and, to show that it converges, it suffices to show that it cannot have two distinct cluster points. Suppose to the contrary that we can extract two subsequences $(\xi_{k_n}(\omega))_{n\in\mathbb{N}}$ and $(\xi_{l_n})_{n\in\mathbb{N}}(\omega)$ such that $\xi_{k_n}(\omega) \to \eta(\omega)$ and $\xi_{l_n}(\omega) \to \zeta(\omega) > \eta(\omega)$, and fix $\varepsilon \in]0, (\zeta - \eta)/2[$. Then, for n sufficiently large, $\xi_{k_n}(\omega) \le \eta(\omega) + \varepsilon < \zeta(\omega) - \varepsilon \le \xi_{l_n}(\omega)$ and, since ϕ is strictly increasing, $\phi(\xi_{k_n}(\omega)) \le \phi(\eta(\omega)+\varepsilon) < \phi(\zeta(\omega)-\varepsilon) \le \phi(\xi_{l_n}(\omega))$. Taking the limit as $n \to +\infty$ yields $\lambda(\omega) \le \phi(\eta(\omega) + \varepsilon) < \phi(\zeta(\omega) - \varepsilon) \le \lambda(\omega)$, which is impossible.

(ii): Since $\mathcal{H}$ is separable, so is C and hence there exists a countable subset X of C such that $\overline{X} = C$. In view of (i), for each $x \in X$, there exists a subset $\mathbf{\Omega}_x$ with probability 1 such that $(\|x_n(\omega) - x\|_{W_n})_{n\in\mathbb{N}}$ converges for every $\omega \in \mathbf{\Omega}_x$. Define $\mathbf{\Omega}^* = \bigcap_{x\in X}\mathbf{\Omega}_x$. Since X is countable, $\mathsf{P}(\mathbf{\Omega}^*) = 1$. Now, let $x_0 \in C$ and $\omega_0 \in \mathbf{\Omega}^*$. Then, there exists a sequence $(c_k)_{k\in\mathbb{N}}$ in X such that $c_k \to x_0$. By (i), we have

$$(\forall k \in \mathbb{N})(\exists \tau_k\colon \mathbf{\Omega} \to [0, +\infty[)(\forall \omega \in \mathbf{\Omega}_{c_k})\quad \|x_n(\omega) - c_k\|_{W_n} \to \tau_k(\omega). \tag{2.8}$$

Moreover, set $\mu = \sup_{n\in\mathbb{N}}\|W_n\|$. Then $\mu < +\infty$ by Banach–Steinhaus Theorem. Then, for every $n \in \mathbb{N}$ and $k \in \mathbb{N}$, we have

$$\begin{aligned}-\sqrt{\mu}\|c_k - x_0\| \le -\|c_k - x_0\|_{W_n} &\le \|x_n(\omega_0) - x_0\|_{W_n} - \|x_n(\omega_0) - c_k\|_{W_n}\\ &\le \|c_k - x_0\|_{W_n}\\ &\le \sqrt{\mu}\|c_k - x_0\|.\end{aligned} \tag{2.9}$$

Therefore,

$$\begin{aligned}(\forall k \in \mathbb{N})\ -\sqrt{\mu}\|c_k - x_0\| &\le \varliminf_{n\to\infty}\|x_n(\omega_0) - x_0\|_{W_n} - \lim_{n\to\infty}\|x_n(\omega_0) - c_k\|_{W_n}\\ &= \varliminf_{n\to\infty}\|x_n(\omega_0) - x_0\|_{W_n} - \tau_k(\omega_0)\\ &\le \varlimsup_{n\to\infty}\|x_n(\omega_0) - x_0\|_{W_n} - \tau_k(\omega_0)\\ &\le \sqrt{\mu}\|c_k - x_0\|.\end{aligned} \tag{2.10}$$

Now, let $k \to \infty$, we get $\lim_{n\to\infty}\|x_n(\omega_0) - x_0\|_{W_n} = \lim_{k\to\infty}\tau_k(\omega_0)$ which proves (ii).

(iii): Necessity is clear. To show sufficiency, let Ω be the set of all ω such that every weak sequential cluster point of $(x_n(\omega))_{n\in\mathbb{N}}$ is in C. Then Ω has probability 1, so is $\Omega_* = \Omega \cap \mathbf{\Omega}^*$. Let $\omega \in \Omega_*$ and $x(\omega)$ and $y(\omega)$ be two weak cluster points of $(x_n(\omega))_{n\in\mathbb{N}}$, say $x_{k_n}(\omega) \rightharpoonup x(\omega)$ and $x_{l_n}(\omega) \rightharpoonup y(\omega)$. Then it follows from (ii) that $(\|x_n(\omega) - x(\omega)\|_{W_n})_{n\in\mathbb{N}}$ and $(\|x_n(\omega) - y(\omega)\|_{W_n})_{n\in\mathbb{N}}$ converge. Moreover, $\|x(\omega)\|^2_{W_n} = \langle W_n x(\omega) \mid x(\omega)\rangle \to \langle W x(\omega) \mid x(\omega)\rangle$ and, likewise, $\|y(\omega)\|^2_{W_n} \to \langle W y(\omega) \mid y(\omega)\rangle$. Therefore, since

$$(\forall n \in \mathbb{N}) \quad \langle W_n x_n(\omega) \mid x(\omega) - y(\omega)\rangle = \frac{1}{2}\big(\|x_n(\omega) - y(\omega)\|^2_{W_n} - \|x_n(\omega) - x(\omega)\|^2_{W_n} + \|x(\omega)\|^2_{W_n} - \|y(\omega)\|^2_{W_n}\big), \tag{2.11}$$

the sequence $(\langle W_n x_n(\omega) \mid x(\omega) - y(\omega)\rangle)_{n\in\mathbb{N}}$ converges, say $\langle W_n x_n(\omega) \mid x(\omega) - y(\omega)\rangle \to \lambda(\omega) \in \mathbb{R}$, which implies that

$$\langle x_n(\omega) \mid W_n(x(\omega) - y(\omega))\rangle \to \lambda(\omega) \in \mathbb{R}. \tag{2.12}$$

However, since $x_{k_n}(\omega) \rightharpoonup x(\omega)$ and $W_{k_n}(x(\omega) - y(\omega)) \to W(x(\omega) - y(\omega))$, it follows from (2.12) and [5, Lemma 2.41(iii)] that $\langle x(\omega) \mid W(x(\omega) - y(\omega))\rangle = \lambda(\omega)$. Likewise, passing to the limit along the subsequence $(x_{l_n}(\omega))_{n\in\mathbb{N}}$ in (2.12) yields $\langle y(\omega) \mid W(x(\omega) - y(\omega))\rangle = \lambda$. Thus,

$$\begin{aligned} 0 &= \langle x(\omega) \mid W(x(\omega) - y(\omega))\rangle - \langle y(\omega) \mid W(x(\omega) - y(\omega))\rangle \\ &= \langle x(\omega) - y(\omega) \mid W(x(\omega) - y(\omega))\rangle \\ &\geq \alpha\|x(\omega) - y(\omega)\|^2. \end{aligned} \tag{2.13}$$

This shows that $x(\omega) = y(\omega)$. Upon invoking (ii) and [5, Lemma 2.38], we conclude that $x_n(\omega) \rightharpoonup x(\omega)$ and hence we obtain the conclusion. □

3 A stochastic forward–backward–forward splitting algorithm

The forward–backward–forward splitting algorithm was firstly proposed in [24] to solve inclusion involving the sum of a maximally monotone operator and a Lipschitzian monotone operator. In [3], it was revisited to include computational errors. Below, we extend it to a stochastic setting. The following theorem is a stochastic version of [25, Theorem 3.1].

Theorem 3.1 *Let $\mathcal{K}$ be a real separable Hilbert space with the scalar product $\langle\langle\cdot \mid \cdot\rangle\rangle$ and the associated norm $|||\cdot|||$. Let α and β be in $]0, +\infty[$, let $(\eta_n)_{n\in\mathbb{N}}$ be a sequence in $\ell^1_+(\mathbb{N})$, and let $(\boldsymbol{U}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{B}(\mathcal{K})$ such that*

$$\mu = \sup_{n\in\mathbb{N}} \|\boldsymbol{U}_n\| < +\infty \quad \text{and} \quad (\forall \boldsymbol{x} \in \mathcal{K})\ (1 + \eta_n)\langle\langle \boldsymbol{x} \mid \boldsymbol{U}_{n+1}\boldsymbol{x}\rangle\rangle \geq \langle\langle \boldsymbol{x} \mid \boldsymbol{U}_n\boldsymbol{x}\rangle\rangle \geq \alpha|||\boldsymbol{x}|||^2. \tag{3.1}$$

Let $\boldsymbol{A}\colon \mathcal{K} \to 2^{\mathcal{K}}$ be maximally monotone, let $\boldsymbol{B}\colon \mathcal{K} \to \mathcal{K}$ be a monotone and β-Lipschitzian operator on $\mathcal{K}$ such that $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \neq \varnothing$. *Let $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ be sequences of square integrable $\mathcal{K}$-valued random vectors. Let $\boldsymbol{x}_0$ be a* *square integrable $\mathcal{K}$-valued random vector, let* $\varepsilon \in]0, 1/(\beta\mu + 1)[$*, let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $[\varepsilon, (1-\varepsilon)/(\beta\mu)]$*, and set*

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{a}_n) \\ \boldsymbol{p}_n = J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}} \boldsymbol{y}_n + \boldsymbol{b}_n \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n. \end{array} \right. \tag{3.2}$$

Suppose that $\big(\sqrt{\mathbf{E}[|||\boldsymbol{a}_n|||^2|\mathcal{F}_n]}\big)_{n\in\mathbb{N}}$, $\big(\sqrt{\mathbf{E}[|||\boldsymbol{b}_n|||^2|\mathcal{F}_n]}\big)_{n\in\mathbb{N}}$, *and* $\big(\sqrt{\mathbf{E}[|||\boldsymbol{c}_n|||^2|\mathcal{F}_n]}\big)_{n\in\mathbb{N}}$ *are summable* P*-a.s., the following hold for some* $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$*-valued random vector* $\overline{\boldsymbol{x}}$*.*

(i) $\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{x}_n - \boldsymbol{p}_n|||^2|\mathcal{F}_n] < +\infty$ *and* $\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{y}_n - \boldsymbol{q}_n|||^2|\mathcal{F}_n] < +\infty$ P*-a.s.*

(ii) $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ *and* $J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}}(\boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\boldsymbol{x}_n) \rightharpoonup \overline{\boldsymbol{x}}$ P*-a.s.*

(iii) *Suppose that one of the following is satisfied for some subset* $\widetilde{\boldsymbol{\Omega}} \subset \boldsymbol{\Omega}$ *with* $\mathsf{P}(\widetilde{\boldsymbol{\Omega}}) = 1$.
   (a) $\boldsymbol{A}+\boldsymbol{B}$ *is demiregular (see* [1, Definition 2.3]) *at* $\overline{\boldsymbol{x}}(\omega)$ *for every* $\omega \in \widetilde{\boldsymbol{\Omega}}$.
   (b) $\boldsymbol{A}$ *or* $\boldsymbol{B}$ *is uniformly monotone at* $\overline{\boldsymbol{x}}(\omega)$ *for every* $\omega \in \widetilde{\boldsymbol{\Omega}}$.
   Then $\boldsymbol{x}_n \to \overline{\boldsymbol{x}}$ *and* $J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}}(\boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\boldsymbol{x}_n) \to \overline{\boldsymbol{x}}$ P*-a.s.*

Proof It follows from [14, Lemma 3.7] that the sequences $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{y}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{p}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{q}_n)_{n\in\mathbb{N}}$ are well defined. Moreover, using [13, Lemma 2.1(i)(ii)] and (3.1), for every sequence of random vectors $\mathcal{K}$-valued $(z_n)_{n\in\mathbb{N}}$, we have

$$\sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||z_n|||^2|\mathcal{F}_n]} < +\infty \quad \mathsf{P}\text{-a.s.} \quad \Leftrightarrow \quad \sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||z_n|||^2_{\boldsymbol{U}_n^{-1}}|\mathcal{F}_n]} < +\infty \quad \mathsf{P}\text{-a.s.} \tag{3.3}$$

and

$$\sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||z_n|||^2|\mathcal{F}_n]} < +\infty \quad \mathsf{P}\text{-a.s.} \quad \Leftrightarrow \quad \sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||z_n|||^2_{\boldsymbol{U}_n}|\mathcal{F}_n]} < +\infty \quad \mathsf{P}\text{-a.s.} \tag{3.4}$$

Let us set, for every $n \in \mathbb{N}$,

$$\begin{cases} \widetilde{\boldsymbol{y}}_n = \boldsymbol{x}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\boldsymbol{x}_n \\ \widetilde{\boldsymbol{p}}_n = J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}} \widetilde{\boldsymbol{y}}_n \\ \widetilde{\boldsymbol{q}}_n = \widetilde{\boldsymbol{p}}_n - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \\ \widetilde{\boldsymbol{x}}_{n+1} = \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n, \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{u}_n = \gamma_n^{-1} \boldsymbol{U}_n^{-1}(\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n) + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n - \boldsymbol{B}\boldsymbol{x}_n \\ \boldsymbol{e}_n = \widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1} \\ \boldsymbol{d}_n = \boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n + \widetilde{\boldsymbol{y}}_n - \boldsymbol{y}_n. \end{cases} \tag{3.5}$$

Then (3.5) yields

$$(\forall n \in \mathbb{N}) \quad \boldsymbol{u}_n = \gamma_n^{-1} \boldsymbol{U}_n^{-1}(\widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n) + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n \in \boldsymbol{A}\widetilde{\boldsymbol{p}}_n + \boldsymbol{B}\widetilde{\boldsymbol{p}}_n, \tag{3.6}$$

and (3.5), Lemma [14, Lemma 3.7(ii)], and the Lipschitzianity of $\boldsymbol{B}$ on $\mathcal{K}$ yield

$$(\forall n \in \mathbb{N}) \quad \begin{cases} |||\boldsymbol{y}_n - \widetilde{\boldsymbol{y}}_n|||_{\boldsymbol{U}_n^{-1}} \leq (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n} \\ |||\boldsymbol{p}_n - \widetilde{\boldsymbol{p}}_n|||_{\boldsymbol{U}_n^{-1}} \leq |||\boldsymbol{b}_n|||_{\boldsymbol{U}_n^{-1}} + (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n} \\ |||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n|||_{\boldsymbol{U}_n^{-1}} \leq 2\big(|||\boldsymbol{b}_n|||_{\boldsymbol{U}_n^{-1}} + (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n}\big) \\ \qquad\qquad + (\beta\mu)^{-1}|||\boldsymbol{c}_n|||_{\boldsymbol{U}_n}. \end{cases} \tag{3.7}$$

Since $\big(\sqrt{\mathbf{E}[|||\boldsymbol{a}_n|||^2|\boldsymbol{\mathcal{F}}_n]}\big)_{n\in\mathbb{N}}$, $\big(\sqrt{\mathbf{E}[|||\boldsymbol{b}_n|||^2|\boldsymbol{\mathcal{F}}_n]}\big)_{n\in\mathbb{N}}$, and $\big(\sqrt{\mathbf{E}[|||\boldsymbol{c}_n|||^2|\boldsymbol{\mathcal{F}}_n]}\big)_{n\in\mathbb{N}}$ are summable P-a.s., using Jensen's inequality, we derive from (3.3), (3.4), (3.5), and (3.7) that

$$\begin{cases} \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{p}_n - \widetilde{\boldsymbol{p}}_n||||\boldsymbol{\mathcal{F}}_n] < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{p}_n - \widetilde{\boldsymbol{p}}_n|||_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] \\ \quad < +\infty \quad \mathsf{P}\text{-a.s.} \\ \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n||||\boldsymbol{\mathcal{F}}_n] < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n|||_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] \\ \quad < +\infty \quad \mathsf{P}\text{-a.s.} \\ \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{d}_n||||\boldsymbol{\mathcal{F}}_n] < +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] < +\infty \quad \mathsf{P}\text{-a.s.} \end{cases} \tag{3.8}$$

Noting that

$$2\mathbf{E}[|||\boldsymbol{y}_n - \widetilde{\boldsymbol{y}}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] \leq 2(\beta\mu)^{-2}\mathbf{E}[|||\boldsymbol{a}_n|||^2_{\boldsymbol{U}_n}|\boldsymbol{\mathcal{F}}_n], \tag{3.9}$$

and

$$2\mathbf{E}[|||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] \leq 24\big(\mathbf{E}[|||\boldsymbol{b}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] + (\beta\mu)^{-2}\mathbf{E}[|||\boldsymbol{a}_n|||^2_{\boldsymbol{U}_n} + |||\boldsymbol{c}_n|||^2_{\boldsymbol{U}_n}|\boldsymbol{\mathcal{F}}_n]\big) \tag{3.10}$$

Therefore, upon setting $c = \max\{26(\beta\mu)^{-2}, 24\}$, and adding (3.9) and (3.10), we get

$$\begin{aligned} &2\mathbf{E}[|||\boldsymbol{y}_n - \widetilde{\boldsymbol{y}}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] + 2\mathbf{E}[|||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] \\ &\leq c\big(\mathbf{E}[|||\boldsymbol{a}_n|||^2_{\boldsymbol{U}_n}|\boldsymbol{\mathcal{F}}_n] + \mathbf{E}[|||\boldsymbol{b}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] + \mathbf{E}[|||\boldsymbol{c}_n|||^2_{\boldsymbol{U}_n}|\boldsymbol{\mathcal{F}}_n]\big). \end{aligned} \tag{3.11}$$

Now, using (3.11), (3.33), (3.34), (3.3), (3.4) and (3.5), we have

$$\begin{aligned} \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{d}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] &\leq 2\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{y}_n - \widetilde{\boldsymbol{y}}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] + 2\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{q}_n - \widetilde{\boldsymbol{q}}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] \\ &\leq c\Bigg(\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{a}_n|||^2_{\boldsymbol{U}_n}|\boldsymbol{\mathcal{F}}_n] + \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{b}_n|||^2_{\boldsymbol{U}_n^{-1}}|\boldsymbol{\mathcal{F}}_n] \\ &\qquad + \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{c}_n|||^2_{\boldsymbol{U}_n}|\boldsymbol{\mathcal{F}}_n]\Bigg) \end{aligned}$$

$$\leq c\tau_0\Bigg(\sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||\boldsymbol{a}_n|||^2_{\boldsymbol{U}_n}|\mathcal{F}_n]}+\sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||\boldsymbol{b}_n|||^2_{\boldsymbol{U}_n^{-1}}|\mathcal{F}_n]}$$
$$+\sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||\boldsymbol{c}_n|||^2_{\boldsymbol{U}_n}|\mathcal{F}_n]}\Bigg)$$
$$< +\infty \quad \mathsf{P}\text{-a.s.,} \tag{3.12}$$

where we define

$$\tau_0 = \sup_{n\in\mathbb{N}}\left\{\sqrt{\mathbf{E}[|||\boldsymbol{a}_n|||^2_{\boldsymbol{U}_n}|\mathcal{F}_n]}, \sqrt{\mathbf{E}[|||\boldsymbol{b}_n|||^2_{\boldsymbol{U}_n^{-1}}|\mathcal{F}_n]}, \sqrt{\mathbf{E}[|||\boldsymbol{c}_n|||^2_{\boldsymbol{U}_n}|\mathcal{F}_n]}\right\}$$
$$< +\infty \quad \mathsf{P}\text{-a.s.} \tag{3.13}$$

Now, let $\boldsymbol{x} \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$. Then, for every $n \in \mathbb{N}$, $(\boldsymbol{x}, -\gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}) \in \operatorname{gra}(\gamma_n\boldsymbol{U}_n\boldsymbol{A})$ and (3.5) yields $(\widetilde{\boldsymbol{p}}_n, \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n) \in \operatorname{gra}(\gamma_n\boldsymbol{U}_n\boldsymbol{A})$. Hence, by monotonicity of $\boldsymbol{U}_n\boldsymbol{A}$ with respect to the scalar product $\langle\langle\cdot \mid \cdot\rangle\rangle_{\boldsymbol{U}_n^{-1}}$, we have $\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}\rangle\rangle_{\boldsymbol{U}_n^{-1}} \leq 0$. Moreover, by monotonicity of $\boldsymbol{U}_n\boldsymbol{B}$ with respect to the scalar product $\langle\langle\cdot \mid \cdot\rangle\rangle_{\boldsymbol{U}_n^{-1}}$, we also have $\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x} - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle_{\boldsymbol{U}_n^{-1}} \leq 0$. By adding the last two inequalities, we obtain

$$(\forall n \in \mathbb{N}) \quad \langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle_{\boldsymbol{U}_n^{-1}} \leq 0. \tag{3.14}$$

In turn, we derive from (3.5) that

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad & 2\gamma_n\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle_{\boldsymbol{U}_n^{-1}}\\
&= 2\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \widetilde{\boldsymbol{p}}_n - \widetilde{\boldsymbol{y}}_n - \gamma_n\boldsymbol{U}_n\boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle_{\boldsymbol{U}_n^{-1}}\\
&\quad + 2\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}_n + \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n\rangle\rangle_{\boldsymbol{U}_n^{-1}}\\
&\leq 2\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \gamma_n\boldsymbol{U}_n\boldsymbol{B}\boldsymbol{x}_n + \widetilde{\boldsymbol{y}}_n - \widetilde{\boldsymbol{p}}_n\rangle\rangle_{\boldsymbol{U}_n^{-1}}\\
&= 2\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n\rangle\rangle_{\boldsymbol{U}_n^{-1}}\\
&= |||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - |||\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - |||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2_{\boldsymbol{U}_n^{-1}}.
\end{aligned} \tag{3.15}$$

Hence, using (3.5), (3.15), the β-Lipschitz continuity of $\boldsymbol{B}$, and [13, Lemma 2.1(ii)], for every $n \in \mathbb{N}$, we obtain

$$\begin{aligned}
|||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} &= |||\widetilde{\boldsymbol{q}}_n + \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}}\\
&= |||(\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}) + \gamma_n\boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n)|||^2_{\boldsymbol{U}_n^{-1}}\\
&= |||\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} + 2\gamma_n\langle\langle\widetilde{\boldsymbol{p}}_n - \boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n\rangle\rangle\\
&\quad + \gamma_n^2|||\boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n)|||^2_{\boldsymbol{U}_n^{-1}}\\
&\leq |||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - |||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2_{\boldsymbol{U}_n^{-1}}\\
&\quad + \gamma_n^2\mu\beta^2|||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2
\end{aligned}$$

$$\leq |||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - \mu^{-1}|||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2 + \gamma_n^2\mu\beta^2|||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2. \tag{3.16}$$

Hence, it follows from (3.1) and [13, Lemma 2.1(i)] that

$$(\forall n \in \mathbb{N}) \quad |||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}|||^2_{\boldsymbol{U}_{n+1}^{-1}} \leq (1+\eta_n)|||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - \mu^{-1}(1-\gamma_n^2\beta^2\mu^2)|||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2. \tag{3.17}$$

Consequently,

$$(\forall n \in \mathbb{N}) \quad |||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}|||_{\boldsymbol{U}_{n+1}^{-1}} \leq (1+\eta_n)|||\boldsymbol{x}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}}. \tag{3.18}$$

For every $n \in \mathbb{N}$, set

$$\varepsilon_n = \sqrt{\mu\alpha^{-1}}\Big(2\big(|||\boldsymbol{b}_n|||_{\boldsymbol{U}_n^{-1}} + (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n}\big) + (\beta\mu)^{-1}|||\boldsymbol{c}_n|||_{\boldsymbol{U}_n} + (\beta\mu)^{-1}|||\boldsymbol{a}_n|||_{\boldsymbol{U}_n}\Big). \tag{3.19}$$

Then $(\mathbf{E}[\varepsilon_n|\boldsymbol{\mathcal{F}}_n])_{n\in\mathbb{N}}$ is summable P-a.s. by (3.3) and we derive from [13, Lemma 2.1(ii)(iii)], and (3.8) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad |||\boldsymbol{e}_n|||_{\boldsymbol{U}_{n+1}^{-1}} &= |||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}|||_{\boldsymbol{U}_{n+1}^{-1}} \\ &\leq \sqrt{\alpha^{-1}}|||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}||| \\ &\leq \sqrt{\mu\alpha^{-1}}|||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}|||_{\boldsymbol{U}_n^{-1}} \\ &\leq \sqrt{\mu\alpha^{-1}}(|||\widetilde{\boldsymbol{y}}_n - \boldsymbol{y}_n|||_{\boldsymbol{U}_n^{-1}} + |||\widetilde{\boldsymbol{q}}_n - \boldsymbol{q}_n|||_{\boldsymbol{U}_n^{-1}}) \\ &\leq \varepsilon_n.\end{aligned} \tag{3.20}$$

In turn, we derive from (3.18) that

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad |||\boldsymbol{x}_{n+1} - \boldsymbol{x}|||_{\boldsymbol{U}_{n+1}^{-1}} &\leq |||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}|||_{\boldsymbol{U}_{n+1}^{-1}} + |||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}_{n+1}|||_{\boldsymbol{U}_{n+1}^{-1}} \\ &\leq |||\widetilde{\boldsymbol{x}}_{n+1} - \boldsymbol{x}|||_{\boldsymbol{U}_{n+1}^{-1}} + \varepsilon_n \\ &\leq (1+\eta_n)|||\boldsymbol{x}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}} + \varepsilon_n.\end{aligned} \tag{3.21}$$

By assumption, since $\mathbf{E}[\|\boldsymbol{x}_0\|^2]$ is finite, by induction, for every $n \in \mathbb{N}$, $\mathbf{E}[\|\boldsymbol{x}_n\|^2]$ is finite and hence $\mathbf{E}[\|\boldsymbol{x}_n\|]$ and $\mathbf{E}[\|\boldsymbol{x}_n\|_{\boldsymbol{U}_n^{-1}}]$ are finite too. By taking the conditional expectation with respect to $\boldsymbol{\mathcal{F}}_n$ and note that $|||\boldsymbol{x}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}}$ is $\boldsymbol{\mathcal{F}}_n$-measurable, we obtain

$$(\forall n \in \mathbb{N}) \quad \mathbf{E}[|||\boldsymbol{x}_{n+1} - \boldsymbol{x}|||_{\boldsymbol{U}_{n+1}^{-1}}|\boldsymbol{\mathcal{F}}_n] \leq (1+\eta_n)|||\boldsymbol{x}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}} + \mathbf{E}[\varepsilon_n|\boldsymbol{\mathcal{F}}_n]. \tag{3.22}$$

This shows that $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ is $|\cdot|$–quasi-Fejér monotone with respect to the target set $\mathrm{zer}(\boldsymbol{A}+\boldsymbol{B})$ relative to $(\boldsymbol{U}_n^{-1})_{n\in\mathbb{N}}$. Moreover, $(|||\boldsymbol{x}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}})_{n\in\mathbb{N}}$ is bounded. In turn, since $\boldsymbol{B}$ and $(J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}})_{n\in\mathbb{N}}$ are Lipschitzian, and $(\forall n \in \mathbb{N})$ $\boldsymbol{x} = J_{\gamma_n \boldsymbol{U}_n \boldsymbol{A}}(\boldsymbol{x} - \gamma_n \boldsymbol{U}_n \boldsymbol{B}\boldsymbol{x})$, we deduce from (3.5) that $(\widetilde{\boldsymbol{y}}_n)_{n\in\mathbb{N}}$, $(\widetilde{\boldsymbol{p}}_n)_{n\in\mathbb{N}}$, and $(\widetilde{\boldsymbol{q}}_n)_{n\in\mathbb{N}}$ are bounded. Therefore,

$$\tau = \sup_{n\in\mathbb{N}}\{|||\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}}, |||\boldsymbol{x}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}}\} < +\infty \quad \mathsf{P}\text{-a.s.} \tag{3.23}$$

Hence, using (3.5), Cauchy–Schwarz for the norms $(|||\cdot|||_{\boldsymbol{U}_n^{-1}})_{n\in\mathbb{N}}$, and (3.16), we get, for every $n \in \mathbb{N}$,

$$\begin{aligned} |||\boldsymbol{x}_{n+1} - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} &= |||\boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} \\ &= |||\widetilde{\boldsymbol{q}}_n + \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n - \boldsymbol{x} + \boldsymbol{d}_n|||^2_{\boldsymbol{U}_n^{-1}} \\ &\leq |||\widetilde{\boldsymbol{q}}_n + \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} + 2|||\widetilde{\boldsymbol{q}}_n + \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}}|||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}} \\ &\quad + |||\boldsymbol{d}_n|||^2_{\boldsymbol{U}_n^{-1}} \\ &\leq |||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - \mu^{-1}(1 - \gamma_n^2\beta^2\mu^2)|||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2 + \varepsilon_{1,n}, \end{aligned} \tag{3.24}$$

where $(\forall n \in \mathbb{N})$ $\varepsilon_{1,n} = 2|||\widetilde{\boldsymbol{q}}_n + \boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}}|||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}} + |||\boldsymbol{d}_n|||^2_{\boldsymbol{U}_n^{-1}}$. In turn, for every $n \in \mathbb{N}$, by (3.1) and [13, Lemma 2.1(i)],

$$\begin{aligned} |||\boldsymbol{x}_{n+1} - \boldsymbol{x}|||^2_{\boldsymbol{U}_{n+1}^{-1}} &\leq (1 + \eta_n)|||\boldsymbol{x}_{n+1} - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} \\ &\leq (1 + \eta_n)|||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - \mu^{-1}(1 - \gamma_n^2\beta^2\mu^2)|||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2 \\ &\quad + (1 + \eta_n)\varepsilon_{1,n}. \end{aligned} \tag{3.25}$$

Since, $J_{\gamma_n \boldsymbol{A}} \circ (\mathrm{Id} - \gamma_n \boldsymbol{B})$ is continuous, $\widetilde{\boldsymbol{p}}_n$ is $\boldsymbol{\mathcal{F}}_n$-measurable. In turn, for every $n \in \mathbb{N}$,

$$\begin{aligned} \mathbf{E}[|||\boldsymbol{x}_{n+1} - \boldsymbol{x}|||^2_{\boldsymbol{U}_{n+1}^{-1}} | \boldsymbol{\mathcal{F}}_n] \leq (1 + \eta_n)|||\boldsymbol{x}_n - \boldsymbol{x}|||^2_{\boldsymbol{U}_n^{-1}} - \mu^{-1}(1 - \gamma_n^2\beta^2\mu^2)|||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2 \\ + \mathbf{E}[(1 + \eta_n)\varepsilon_{1,n} | \boldsymbol{\mathcal{F}}_n]. \end{aligned} \tag{3.26}$$

Let us prove that

$$\sum_{n\in\mathbb{N}} \mathbf{E}[\varepsilon_{1,n} | \boldsymbol{\mathcal{F}}_n] < +\infty \quad \mathsf{P}\text{-a.s.} \tag{3.27}$$

Indeed, since $\mathrm{Id} - \gamma_n \boldsymbol{B}$ is continuous, $\widetilde{\boldsymbol{y}}_n$ and $\widetilde{\boldsymbol{q}}_n$ are $\boldsymbol{\mathcal{F}}_n$-measurable. Therefore, $|||\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}}$ is $\boldsymbol{\mathcal{F}}_n$-measurable and hence by (3.8) and (3.23), we obtain

$$\begin{aligned}
&\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}} |||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}} | \boldsymbol{\mathcal{F}}_n] \\
&\quad = \sum_{n\in\mathbb{N}} |||\boldsymbol{x}_n - \widetilde{\boldsymbol{y}}_n + \widetilde{\boldsymbol{q}}_n - \boldsymbol{x}|||_{\boldsymbol{U}_n^{-1}} \mathbf{E}[|||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}} | \boldsymbol{\mathcal{F}}_n] \\
&\quad \leq \tau \sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{d}_n|||_{\boldsymbol{U}_n^{-1}} | \boldsymbol{\mathcal{F}}_n] \\
&\quad < +\infty \quad \mathsf{P}\text{-a.s.},
\end{aligned} \tag{3.28}$$

which and (3.12) prove (3.27). It follows from Lemma 2.1 that

$$\sum_{n\in\mathbb{N}} |||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2 < +\infty \quad \mathsf{P}\text{-a.s.} \tag{3.29}$$

(i): It follows from (3.29) and (3.8) that

$$\begin{aligned}
&\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{x}_n - \boldsymbol{p}_n|||^2 | \boldsymbol{\mathcal{F}}_n] \leq 2\sum_{n\in\mathbb{N}} |||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2 + 2\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{p}_n - \widetilde{\boldsymbol{p}}_n|||^2 | \boldsymbol{\mathcal{F}}_n] \\
&\quad < +\infty \quad \mathsf{P}\text{-a.s.}
\end{aligned} \tag{3.30}$$

Furthermore, we derive from (3.8), (3.3) and (3.12) that

$$\begin{aligned}
\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{y}_n - \boldsymbol{q}_n|||^2] &= \sum_{n\in\mathbb{N}} \mathbf{E}[|||\widetilde{\boldsymbol{q}}_n - \widetilde{\boldsymbol{y}}_n + \boldsymbol{d}_n|||^2 | \boldsymbol{\mathcal{F}}_n] \\
&= \sum_{n\in\mathbb{N}} \mathbf{E}[|||\widetilde{\boldsymbol{p}}_n - \boldsymbol{x}_n + \gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n) + \boldsymbol{d}_n|||^2 | \boldsymbol{\mathcal{F}}_n] \\
&\leq 3\Big(\sum_{n\in\mathbb{N}} |||\boldsymbol{x}_n - \widetilde{\boldsymbol{p}}_n|||^2 + \mathbf{E}[|||\gamma_n \boldsymbol{U}_n(\boldsymbol{B}\boldsymbol{x}_n - \boldsymbol{B}\widetilde{\boldsymbol{p}}_n)|||^2 \\
&\qquad + |||\boldsymbol{d}_n|||^2 | \boldsymbol{\mathcal{F}}_n]\Big) \\
&< +\infty \quad \mathsf{P}\text{-a.s.}
\end{aligned} \tag{3.31}$$

(ii): Let $\boldsymbol{\Omega}_0$ be the set of all $\omega \in \boldsymbol{\Omega}$ such that $(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}}$ is bounded and (3.29) is satisfied. We have $\mathsf{P}(\boldsymbol{\Omega}_0) = 1$. Fix $\omega \in \boldsymbol{\Omega}_0$. Let $\boldsymbol{x}(\omega)$ be a weak cluster point of $(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}}$. Then there exists a subsequence $(\boldsymbol{x}_{k_n}(\omega))_{n\in\mathbb{N}}$ that converges weakly to $\boldsymbol{x}(\omega)$. Therefore $\widetilde{\boldsymbol{p}}_{k_n}(\omega) \rightharpoonup \boldsymbol{x}(\omega)$ by (3.29) and by the definition of $\boldsymbol{\Omega}_0$. Furthermore, it follows from (3.5) that $\boldsymbol{u}_{k_n}(\omega) \to 0$. Hence, since $(\forall n \in \mathbb{N})$ $(\widetilde{\boldsymbol{p}}_{k_n}(\omega), \boldsymbol{u}_{k_n}(\omega)) \in \operatorname{gra}(\boldsymbol{A}+\boldsymbol{B})$, we obtain, $\boldsymbol{x}(\omega) \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ [5, Proposition 20.33 (ii)]. Altogether, it follows Proposition 2.4 that $\boldsymbol{x}_n \rightharpoonup \overline{\boldsymbol{x}}$ and hence that $\widetilde{\boldsymbol{p}}_n \rightharpoonup \overline{\boldsymbol{x}}$.

Now, let $\boldsymbol{\Omega}_1$ be the set of all $\omega \in \boldsymbol{\Omega}$ such that $\boldsymbol{x}_n(\omega) \rightharpoonup \overline{\boldsymbol{x}}(\omega)$ and $\widetilde{\boldsymbol{p}}_n(\omega) \rightharpoonup \overline{\boldsymbol{x}}(\omega)$, and $\widetilde{\boldsymbol{p}}_n(\omega) - \boldsymbol{x}_n(\omega) \to 0$. Then $\mathsf{P}(\boldsymbol{\Omega}_1) = 1$ and hence $\mathsf{P}(\boldsymbol{\Omega}_1 \cap \widetilde{\boldsymbol{\Omega}}) = 1$.

(iii)(a): Fix $\omega \in \boldsymbol{\Omega}_1 \cap \widetilde{\boldsymbol{\Omega}}$. Then $\boldsymbol{x}_n(\omega) \rightharpoonup \overline{\boldsymbol{x}}(\omega)$ and $\widetilde{\boldsymbol{p}}_n(\omega) \rightharpoonup \overline{\boldsymbol{x}}(\omega)$. Furthermore, it follows from (3.5) that $\boldsymbol{u}_n(\omega) \to 0$. Hence, since $(\forall n \in \mathbb{N})$ $(\widetilde{\boldsymbol{p}}_n(\omega), \boldsymbol{u}_n(\omega)) \in \operatorname{gra}(\boldsymbol{A}+\boldsymbol{B})$ and since $\boldsymbol{A}+\boldsymbol{B}$ is demiregular at $\overline{\boldsymbol{x}}(\omega)$ by our assumption, by [1, Definition 2.3], $\widetilde{\boldsymbol{p}}_n(\omega) \to \overline{\boldsymbol{x}}(\omega)$, and therefore $\boldsymbol{x}_n(\omega) \to \overline{\boldsymbol{x}}(\omega)$.

(iii)(b): Fix $\omega \in \mathbf{\Omega}_1 \cap \widetilde{\mathbf{\Omega}}$. If $\boldsymbol{A}$ or $\boldsymbol{B}$ is uniformly monotone at $\overline{\boldsymbol{x}}(\omega)$, then $\boldsymbol{A}+\boldsymbol{B}$ is uniformly monotone at $\overline{\boldsymbol{x}}(\omega)$. Therefore, the result follows from [1, Proposition 2.4(i)]. □

Corollary 3.2 *Let $\mathcal{K}$ be a real separable Hilbert space with the scalar product $\langle\langle\cdot \mid \cdot\rangle\rangle$ and the associated norm $|||\cdot|||$. Let β be in $]0,+\infty[$, let $\boldsymbol{A}\colon \mathcal{K}\to 2^{\mathcal{K}}$ be maximally monotone, let $\boldsymbol{B}\colon \mathcal{K}\to\mathcal{K}$ be a monotone and β-Lipschitzian operator on $\mathcal{K}$ such that $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})\neq\varnothing$. Let $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ be sequences of square integrable $\mathcal{K}$-valued random vectors. Let $\boldsymbol{x}_0$ be a square integrable $\mathcal{K}$-valued random vector , let $\varepsilon\in]0,1/(\beta+1)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,(1-\varepsilon)/\beta]$, and set*

$$(\forall n\in\mathbb{N})\quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n=\boldsymbol{x}_n-\gamma_n(\boldsymbol{B}\boldsymbol{x}_n+\boldsymbol{a}_n)\\ \boldsymbol{p}_n=J_{\gamma_n\boldsymbol{A}}\boldsymbol{y}_n+\boldsymbol{b}_n\\ \boldsymbol{q}_n=\boldsymbol{p}_n-\gamma_n(\boldsymbol{B}\boldsymbol{p}_n+\boldsymbol{c}_n)\\ \boldsymbol{x}_{n+1}=\boldsymbol{x}_n-\boldsymbol{y}_n+\boldsymbol{q}_n. \end{array}\right. \tag{3.32}$$

Suppose that the following conditions are satisfied with $\boldsymbol{\mathcal{F}}_n=\sigma(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_n)$,

$$\sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||\boldsymbol{a}_n|||^2|\boldsymbol{\mathcal{F}}_n]}<\infty,\quad \sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||\boldsymbol{b}_n|||^2|\boldsymbol{\mathcal{F}}_n]}<\infty \tag{3.33}$$

and

$$\sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||\boldsymbol{c}_n|||^2|\boldsymbol{\mathcal{F}}_n]}<\infty\quad \mathsf{P}\text{-a.s.} \tag{3.34}$$

Then the following hold for some $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$-valued random vector $\overline{\boldsymbol{x}}$.

(i) $\sum_{n\in\mathbb{N}}\mathbf{E}[|||\boldsymbol{x}_n-\boldsymbol{p}_n|||^2|\boldsymbol{\mathcal{F}}_n]<+\infty$ *and* $\sum_{n\in\mathbb{N}}\mathbf{E}[|||\boldsymbol{y}_n-\boldsymbol{q}_n|||^2|\boldsymbol{\mathcal{F}}_n]<+\infty$ P*-a.s.*
(ii) $\boldsymbol{x}_n\rightharpoonup\overline{\boldsymbol{x}}$ *and* $J_{\gamma_n\boldsymbol{A}}(\boldsymbol{x}_n-\gamma_n\boldsymbol{B}\boldsymbol{x}_n)\rightharpoonup\overline{\boldsymbol{x}}$ P*-a.s.*
(iii) *Suppose that one of the following is satisfied for some subset $\widetilde{\mathbf{\Omega}}\subset\mathbf{\Omega}$ with $\mathsf{P}(\widetilde{\mathbf{\Omega}})=1$.*
  (a) *$\boldsymbol{A}+\boldsymbol{B}$ is demiregular (see [1, Definition 2.3]) at $\overline{\boldsymbol{x}}(\omega)$ for every $\omega\in\widetilde{\mathbf{\Omega}}$.*
  (b) *$\boldsymbol{A}$ or $\boldsymbol{B}$ is uniformly monotone at $\overline{\boldsymbol{x}}(\omega)$ for every $\omega\in\widetilde{\mathbf{\Omega}}$.*
  Then $\boldsymbol{x}_n\to\overline{\boldsymbol{x}}$ and $J_{\gamma_n\boldsymbol{A}}(\boldsymbol{x}_n-\gamma_n\boldsymbol{B}\boldsymbol{x}_n)\to\overline{\boldsymbol{x}}$ P-a.s.

Remark 3.3 Here are some remarks. In the case when $\boldsymbol{B}$ is a general multi-valued maximally monotone operator or a cocoercive operator, the almost sure convergence of the Douglas–Rachford or forward–backward are proved in [12] under the same type of condition on the stochastic errors. Furthermore, in the case when $\boldsymbol{B}$ is cocoercive and uniformly monotone, the almost sure convergence of the forward–backward splitting is also proved in [21] under different conditions on stepsize and stochastic errors. One of the early work concerns with Lipschitzian monotone operator was in [18].

Example 3.4 Let $\boldsymbol{f}\colon\mathcal{K}\to[-\infty,+\infty]$ be a proper lower semicontinuous convex function, let $\alpha\in]0,+\infty[$, let $\beta\in]0,+\infty[$, let $\boldsymbol{B}\colon\mathcal{K}\to\mathcal{K}$ be a monotone and β-Lipschitzian operator. Let $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ be sequences of square integrable $\mathcal{K}$-valued random vectors such that (3.33) and (3.34) are satisfied. Furthermore, let $\boldsymbol{x}_0$ be a square integrable $\mathcal{K}$-valued random vector, let

$\varepsilon \in]0, \min\{1, 1/(\beta+1)\}[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$. Suppose that the variational inequality

$$\text{find} \quad \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{K}} \quad \text{such that} \quad (\forall \boldsymbol{y} \in \boldsymbol{\mathcal{K}}) \quad \langle \overline{\boldsymbol{x}} - \boldsymbol{y} \mid \boldsymbol{B}\overline{\boldsymbol{x}} \rangle + \boldsymbol{f}(\overline{\boldsymbol{x}}) \le \boldsymbol{f}(\boldsymbol{y}) \tag{3.35}$$

admits at least one solution and set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{a}_n) \\ \boldsymbol{p}_n = \underset{\boldsymbol{x}\in\boldsymbol{\mathcal{K}}}{\arg\min} \left(\boldsymbol{f}(\boldsymbol{x}) + \frac{1}{2\gamma_n} |||\boldsymbol{x} - \boldsymbol{y}_n|||^2 \right) + \boldsymbol{b}_n \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n. \end{array} \right. \tag{3.36}$$

Then, for almost all $\omega \in \boldsymbol{\Omega}$, $(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}}$ converges weakly to a solution $\overline{\boldsymbol{x}}(\omega)$ to (3.35).

Proof Set $\boldsymbol{A} = \partial \boldsymbol{f}$ in Corollary 3.2(ii). □

Remark 3.5 Since $(\gamma_n)_{n\in\mathbb{N}}$ is bounded away from 0, we have

$$\sum_{n\in\mathbb{N}} \gamma_n = +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty. \tag{3.37}$$

While, in the standard stochastic gradient method [20], we often require

$$\sum_{n\in\mathbb{N}} \gamma_n = +\infty \quad \text{and} \quad \sum_{n\in\mathbb{N}} \gamma_n^2 < +\infty. \tag{3.38}$$

Under the condition (3.38), the conditions on the stochastic errors in the stochastic gradient method are weaker than (3.33)–(3.34) (see also [26, Assumption 2 and Eq (4)] for the case of the projected stochastic gradient method).

We end this section by noting that, in the case when $\boldsymbol{U}_n = \boldsymbol{U}$, we obtain a preconditioned version of (3.32). Some other preconditioned algorithms can be found in [17,19].

4 Monotone inclusions involving Lipschitzian operators

The applications of the forward–backward-forward splitting algorithm considered in [3,11,24] can be extended to a stochastic setting using Theorem 3.1. As an illustration, we present a stochastic version of the algorithm proposed in [11, Eq. (3.1)]. Recall that the parallel sum of $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ is [5]

$$A \,\square\, B = (A^{-1} + B^{-1})^{-1}. \tag{4.1}$$

Problem 4.1 Let $\mathcal{H}$ be a real separable Hilbert space, let m be a strictly positive integer, let $z \in \mathcal{H}$, let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone operator, let $C\colon \mathcal{H} \to \mathcal{H}$ be monotone and ν_0-Lipschitzian for some $\nu_0 \in]0, +\infty[$. For every $i \in \{1, \ldots, m\}$, let $\mathcal{G}_i$

be a real separable Hilbert space, let $r_i \in \mathcal{G}_i$, let $B_i : \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be maximally monotone operator, let $D_i : \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be monotone and such that D_i^{-1} is ν_i-Lipschitzian for some $\nu_i \in]0, +\infty[$, and let $L_i : \mathcal{H} \to \mathcal{G}_i$ be a nonzero bounded linear operator. Suppose that

$$z \in \operatorname{ran}\left(A + \sum_{i=1}^{m} L_i^*\big((B_i \,\square\, D_i)(L_i \cdot -r_i)\big) + C\right). \tag{4.2}$$

The problem is to solve the primal inclusion

$$z \in A\overline{x} + \sum_{i=1}^{m} L_i^*\big((B_i \,\square\, D_i)(L_i\overline{x} - r_i)\big) + C\overline{x}, \tag{4.3}$$

and the dual inclusion

$$(\forall i \in \{1, \ldots, m\}) \quad r_i \in -L_i(A + C)^{-1}\left(z - \sum_{i=1}^{m} L_i^*\overline{v}_i\right) + B_i^{-1}\overline{v}_i + D_i^{-1}\overline{v}_i. \tag{4.4}$$

We denote by $\mathcal{P}$ and $\mathcal{D}$ be the set of solutions to (4.3) and (4.4), respectively.

As shown in [11], Problem 4.1 covers a wide class of problems in nonlinear analysis and convex optimization problems. However, the algorithm in [11, Theorem 3.1] is studied in the deterministic. The following result extends this result to a stochastic setting.

Let us define $\mathcal{K} = \mathcal{H} \oplus \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_m$ the Hilbert direct sum of the Hilbert spaces $\mathcal{H}$ and $(\mathcal{G}_i)_{1 \le i \le m}$, the scalar product and the associated norm of $\mathcal{K}$ respectively defined by

$$\langle\langle \cdot \mid \cdot \rangle\rangle : ((x, \boldsymbol{v}), (y, \boldsymbol{w})) \mapsto \langle x \mid y \rangle + \sum_{i=1}^{m} \langle v_i \mid w_i \rangle \text{ and } |||\cdot||| : (x, \boldsymbol{v}) \mapsto \sqrt{\|x\|^2 + \sum_{i=1}^{m} \|v_i\|^2}, \tag{4.5}$$

where $\boldsymbol{v} = (v_1, \ldots, v_m)$ and $\boldsymbol{w} = (w_1, \ldots, w_m)$ are generic elements in $\mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_m$.

Corollary 4.2 *Let $(a_{1,n})_{n\in\mathbb{N}}$, $(b_{1,n})_{n\in\mathbb{N}}$, and $(c_{1,n})_{n\in\mathbb{N}}$ be sequences of square integrable $\mathcal{H}$-valued random vectors, and for every $i \in \{1, \ldots, m\}$, let $(a_{2,i,n})_{n\in\mathbb{N}}$, $(b_{2,i,n})_{n\in\mathbb{N}}$, and $(c_{2,i,n})_{n\in\mathbb{N}}$ be sequences of square integrable $\mathcal{G}_i$-valued random vectors. Furthermore, set*

$$\beta = \max\{\nu_0, \nu_1, \ldots, \nu_m\} + \sqrt{\sum_{i=1}^{m} \|L_i\|^2}, \tag{4.6}$$

let x_0 be a square integrable $\mathcal{H}$-valued random vector, and, for every $i \in \{1, \ldots, m\}$, let $v_{i,0}$ be a square integrable $\mathcal{G}_i$-valued random vector, let $\varepsilon \in]0, 1/(1+\beta)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon, (1-\varepsilon)/\beta]$. Set

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} y_{1,n} = x_n - \gamma_n\big(Cx_n + \sum_{i=1}^m L_i^* v_{i,n} + a_{1,n}\big) \\ p_{1,n} = J_{\gamma_n A}(y_{1,n} + \gamma_n z) + b_{1,n} \\ \text{for } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} y_{2,i,n} = v_{i,n} + \gamma_n\big(L_i x_n - D_i^{-1} v_{i,n} + a_{2,i,n}\big) \\ p_{2,i,n} = J_{\gamma_n B_i^{-1}}(y_{2,i,n} - \gamma_n r_i) + b_{2,i,n} \\ q_{2,i,n} = p_{2,i,n} + \gamma_n\big(L_i p_{1,n} - D_i^{-1} p_{2,i,n} + c_{2,i,n}\big) \\ v_{i,n+1} = v_{i,n} - y_{2,i,n} + q_{2,i,n} \end{array} \right. \\ q_{1,n} = p_{1,n} - \gamma_n\big(Cp_{1,n} + \sum_{i=1}^m L_i^* p_{2,i,n} + c_{1,n}\big) \\ x_{n+1} = x_n - y_{1,n} + q_{1,n}. \end{array} \right. \tag{4.7}$$

Suppose that the following conditions hold for $\boldsymbol{\mathcal{F}}_n = \sigma((x_k, (v_{i,k})_{1\le i\le m})_{0\le k\le n}$,

$$\begin{cases} \sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||(a_{1,n}, (a_{2,i,n})_{1\le i\le m})|||^2|\boldsymbol{\mathcal{F}}_n]} < +\infty \\ \sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[\|(b_{1,n}, (b_{2,i,n})_{1\le i\le m})|||^2|\boldsymbol{\mathcal{F}}_n]} < +\infty \\ \sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||(c_{1,n}, (c_{2,i,n})_{1\le i\le m})|||^2|\boldsymbol{\mathcal{F}}_n]} < +\infty. \end{cases} \tag{4.8}$$

Then the following hold.

(i) $\sum_{n\in\mathbb{N}} \mathbf{E}[\|x_n - p_{1,n}\|^2|\boldsymbol{\mathcal{F}}_n] < +\infty$ *and* $(\forall i \in \{1, \ldots, m\})$ $\sum_{n\in\mathbb{N}} \mathbf{E}[\|v_{i,n} - p_{2,i,n}\|^2|\boldsymbol{\mathcal{F}}_n] < +\infty$ P*-a.s.*

(ii) *There exist a $\mathcal{P}$-valued random vector $\overline{x}$ and a $\mathcal{D}$-valued random vector $(\overline{v}_1, \ldots, \overline{v}_m)$ such that the following hold.*
   (a) $x_n \rightharpoonup \overline{x}$ *and* $J_{\gamma_n A}(x_n - \gamma_n(Cx_n + \sum_{i=1}^m L_i^* v_{i,n}) + \gamma_n z) \rightharpoonup \overline{x}$ P*-a.s.*
   (b) $(\forall i \in \{1, \ldots, m\})$ $v_{i,n} \rightharpoonup \overline{v}_i$ *and* $J_{\gamma_n B_i^{-1}}(v_{i,n} + \gamma_n\big(L_i x_n - D_i^{-1} v_{i,n}) - \gamma_n r_i) \rightharpoonup \overline{v}_i$ P*-a.s.*
   (c) *Suppose that A or C is uniformly monotone at $\overline{x}(\omega)$ for every $\omega \in \widetilde{\boldsymbol{\Omega}} \subset \boldsymbol{\Omega}$ with $\mathsf{P}(\widetilde{\boldsymbol{\Omega}}) = 1$, then $x_n \to \overline{x}$ and $J_{\gamma_n A}(x_n - \gamma_n(Cx_n + \sum_{i=1}^m L_i^* v_{i,n}) + \gamma_n z) \to \overline{x}$ P-a.s.*
   (d) *Suppose that B_j^{-1} or D_j^{-1} is uniformly monotone at $\overline{v}_j(\omega)$ for every $\omega \in \widetilde{\boldsymbol{\Omega}} \subset \boldsymbol{\Omega}$ with $\mathsf{P}(\widetilde{\boldsymbol{\Omega}}) = 1$, for some $j \in \{1, \ldots, m\}$, then $v_{j,n} \to \overline{v}_j$ and $J_{\gamma_n B_j^{-1}}(v_{j,n} + \gamma_n\big(L_j x_n - D_j^{-1} v_{j,n}) - \gamma_n r_j) \to \overline{v}_j$ P-a.s.*

Proof Set

$$\begin{cases} \boldsymbol{A}\colon \boldsymbol{\mathcal{K}} \to 2^{\boldsymbol{\mathcal{K}}}\colon (x, v_1, \ldots, v_m) \mapsto (-z + Ax) \times (r_1 + B_1^{-1} v_1) \\ \qquad\qquad \times \cdots \times (r_m + B_m^{-1} v_m) \\ \boldsymbol{B}\colon \boldsymbol{\mathcal{K}} \to \boldsymbol{\mathcal{K}}\colon (x, v_1, \ldots, v_m) \mapsto \Big(Cx + \sum_{i=1}^m L_i^* v_i, D_1^{-1} v_1 - L_1 x, \ldots, \\ \qquad\qquad D_m^{-1} v_m - L_m x\Big). \end{cases} \tag{4.9}$$

Since $\boldsymbol{A}$ is maximally monotone [5, Propositions 20.22 and 20.23], $\boldsymbol{B}$ is monotone and β-Lipschitzian [11, Eq. (3.10)] with dom $\boldsymbol{B} = \boldsymbol{\mathcal{K}}$, $\boldsymbol{A} + \boldsymbol{B}$ is maximally monotone [5, Corollary 24.24(i)]. In addition, [5, Propositions 23.15(ii) and 23.16] yield $(\forall \gamma \in]0, +\infty[)(\forall n \in \mathbb{N})(\forall (x, v_1, \ldots, v_m) \in \boldsymbol{\mathcal{K}})$

$$J_{\gamma \boldsymbol{A}}(x, v_1, \ldots, v_m) = \Big(J_{\gamma A}(x + \gamma z), \big(J_{\gamma B_i^{-1}}(v_i - \gamma r_i) \big)_{1 \le i \le m} \Big). \tag{4.10}$$

It is shown in [11, Eq. (3.12)] and [11, Eq. (3.13)] that under the condition (4.2), $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \neq \varnothing$. Moreover, [11, Eq. (3.21)] and [11, Eq. (3.22)] yield

$$(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \Rightarrow \overline{x} \text{ solves (4.3) and } (\overline{v}_1, \ldots, \overline{v}_m) \text{ solves (4.4).} \tag{4.11}$$

Let us next set, for every $n \in \mathbb{N}$,

$$\begin{cases} \boldsymbol{x}_n = (x_n, v_{1,n}, \ldots, v_{m,n}) \\ \boldsymbol{y}_n = (y_{1,n}, y_{2,1,n}, \ldots, y_{2,m,n}) \\ \boldsymbol{p}_n = (p_{1,n}, p_{2,1,n}, \ldots, p_{2,m,n}) \\ \boldsymbol{q}_n = (q_{1,n}, q_{2,1,n}, \ldots, q_{2,m,n}) \end{cases} \text{and} \quad \begin{cases} \boldsymbol{a}_n = (a_{1,n}, a_{2,1,n}, \ldots, a_{2,m,n}) \\ \boldsymbol{b}_n = (b_{1,n}, b_{2,1,n}, \ldots, b_{2,m,n}) \\ \boldsymbol{c}_n = (c_{1,n}, c_{2,1,n}, \ldots, c_{2,m,n}). \end{cases} \tag{4.12}$$

Then our assumptions imply that

$$\sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||\boldsymbol{a}_n|||^2 | \boldsymbol{\mathcal{F}}_n]} < \infty, \quad \sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||\boldsymbol{b}_n|||^2 | \boldsymbol{\mathcal{F}}_n]} < \infty, \quad \text{and} \quad \sum_{n\in\mathbb{N}} \sqrt{\mathbf{E}[|||\boldsymbol{c}_n|||^2 | \boldsymbol{\mathcal{F}}_n]} < \infty. \tag{4.13}$$

Furthermore, it follows from the definition of $\boldsymbol{B}$, (4.10), and (4.12) that (4.7) can be rewritten in $\boldsymbol{\mathcal{K}}$ as

$$(\forall n \in \mathbb{N}) \quad \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \boldsymbol{x}_n - \gamma_n(\boldsymbol{B}\boldsymbol{x}_n + \boldsymbol{a}_n) \\ \boldsymbol{p}_n = J_{\gamma_n \boldsymbol{A}} \boldsymbol{y}_n + \boldsymbol{b}_n \\ \boldsymbol{q}_n = \boldsymbol{p}_n - \gamma_n(\boldsymbol{B}\boldsymbol{p}_n + \boldsymbol{c}_n) \\ \boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n, \end{array} \right. \tag{4.14}$$

which is (3.32). Moreover, every specific conditions in Corollary 3.2 are satisfied.

(i): By Corollary 3.2(i), $\sum_{n\in\mathbb{N}} \mathbf{E}[|||\boldsymbol{x}_n - \boldsymbol{p}_n|||^2 | \boldsymbol{\mathcal{F}}_n] < \infty$.

(ii)(a)&(ii)(b): It follows from Corollary 3.2(ii) that

$$x_n \rightharpoonup \overline{x} \quad \text{and} \quad (\forall i \in \{1, \ldots, m\}) \quad v_{i,n} \rightharpoonup \overline{v}_i \quad \mathsf{P} - a.s. \tag{4.15}$$

Corollary 3.2(ii) shows that $(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m) \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$. Hence, it follows from [11, Eq (3.19)] that $(\overline{x}, \overline{v}_1, \ldots, \overline{v}_m)$ satisfies the inclusions

$$\begin{cases} -\sum_{i=1}^m L_i^* \overline{v}_i - C\overline{x} \in -z + A\overline{x} \\ (\forall i \in \{1, \ldots, m\}) \; L_i \overline{x} - D_i^{-1} \overline{v}_i \in r_i + B_i^{-1} \overline{v}_i. \end{cases} \tag{4.16}$$

For every $n \in \mathbb{N}$ and every $i \in \{1, \ldots, m\}$, set

$$\begin{cases} \widetilde{y}_{1,n} = x_n - \gamma_n\big(Cx_n + \sum_{i=1}^m L_i^* v_{i,n}\big) \\ \widetilde{p}_{1,n} = J_{\gamma_n A}(\widetilde{y}_{1,n} + \gamma_n z) \end{cases} \quad \text{and} \quad \begin{cases} \widetilde{y}_{2,i,n} = v_{i,n} + \gamma_n\big(L_i x_n - D_i^{-1} v_{i,n}\big) \\ \widetilde{p}_{2,i,n} = J_{\gamma_n B_i^{-1}}(\widetilde{y}_{2,i,n} - \gamma_n r_i). \end{cases} \tag{4.17}$$

We note that (4.13) implies that

$$\mathbf{E}[|||\boldsymbol{a}_n|||^2|\boldsymbol{\mathcal{F}}_n] \to 0, \quad \mathbf{E}[|||\boldsymbol{b}_n|||^2|\boldsymbol{\mathcal{F}}_n] \to 0 \quad \text{and} \quad \mathbf{E}[|||\boldsymbol{c}_n|||^2|\boldsymbol{\mathcal{F}}_n] \to 0 \quad \mathsf{P}\text{-a.s.} \tag{4.18}$$

Then, using [5, Corollary 23.10], we get

$$\begin{cases} \|\widetilde{p}_{1,n} - p_{1,n}\| \le \|b_{1,n}\| + \beta^{-1}\|a_{1,n}\|, \\ (\forall i \in \{1, \ldots, m\}) \quad \|\widetilde{p}_{2,i,n} - p_{2,i,n}\| \le \|b_{2,i,n}\| + \beta^{-1}\|a_{2,i,n}\|, \end{cases} \tag{4.19}$$

which and (4.18) imply that

$$\begin{cases} \mathbf{E}[\|\widetilde{p}_{1,n} - p_{1,n}\|^2|\boldsymbol{\mathcal{F}}_n] \le 2\mathbf{E}[\|b_{1,n}\|^2 + \beta^{-2}\|a_{1,n}\|^2|\boldsymbol{\mathcal{F}}_n] \to 0 \ \mathsf{P}\text{-a.s.} \\ (\forall i \in \{1, \ldots, m\}) \ \mathbf{E}[\|\widetilde{p}_{2,i,n} - p_{2,i,n}\|^2|\boldsymbol{\mathcal{F}}_n] \le 2\mathbf{E}[\|b_{2,i,n}\|^2 \\ \quad + \beta^{-2}\|a_{2,i,n}\|^2|\boldsymbol{\mathcal{F}}_n] \to 0 \ \mathsf{P}\text{-a.s.} \end{cases} \tag{4.20}$$

Since $(x, v_1 \ldots, v_m) \mapsto J_{\gamma_n A}(x - \gamma_n\big(Cx + \sum_{i=1}^m L_i^* v_i) + \gamma_n z)$ is continuous from $\mathcal{K} \to \mathcal{H}$, $\widetilde{p}_{1,n}$ is $\boldsymbol{\mathcal{F}}_n$-measurable. By the same way, for every $i \in \{1, \ldots, m\}$, $\widetilde{p}_{2,i,n}$ is $\boldsymbol{\mathcal{F}}_n$-measurable. In turn, by (i),(ii)(a), and (ii)(b), we obtain

$$\begin{cases} \|\widetilde{p}_{1,n} - x_n\|^2 = \mathbf{E}[\|\widetilde{p}_{1,n} - x_n\|^2|\boldsymbol{\mathcal{F}}_n] \le 2\mathbf{E}[\|p_{1,n} - x_n\|^2 + \|\widetilde{p}_{1,n} \\ \qquad - p_{1,n}\|^2|\boldsymbol{\mathcal{F}}_n] \to 0 \ \mathsf{P}\text{-a.s.} \\ (\forall i \in \{1, \ldots, m\}) \ \|\widetilde{p}_{2,i,n} - v_{i,n}\|^2 = \mathbf{E}[\|\widetilde{p}_{2,i,n} - v_{i,n}\|^2|\boldsymbol{\mathcal{F}}_n] \\ \qquad\qquad \le 2\mathbf{E}[\|\widetilde{p}_{2,i,n} - p_{2,i,n}\|^2 + \|p_{2,i,n} \\ \qquad\qquad - v_{i,n}\|^2|\boldsymbol{\mathcal{F}}_n] \to 0 \ \mathsf{P}\text{-a.s.} \\ \widetilde{p}_{1,n} \rightharpoonup \overline{x} \quad \mathsf{P}\text{-a.s.} \quad \text{and} \quad (\forall i \in \{1, \ldots, m\}) \quad \widetilde{p}_{2,i,n} \rightharpoonup \overline{v}_i \quad \mathsf{P}\text{-a.s.} \end{cases} \tag{4.21}$$

(ii)(c): We derive from (4.17) that

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \gamma_n^{-1}(x_n - \widetilde{p}_{1,n}) - \sum_{i=1}^m L_i^* v_{i,n} - Cx_n \in -z + A\widetilde{p}_{1,n} \\ (\forall i \in \{1, \ldots, m\}) \ \gamma_n^{-1}(v_{i,n} - \widetilde{p}_{2,i,n}) + L_i x_n - D_i^{-1} v_{i,n} \in r_i \\ \qquad + B_i^{-1}\widetilde{p}_{2,i,n}. \end{cases} \tag{4.22}$$

Let $\boldsymbol{\Omega}_3$ be the set of all $\omega \in \boldsymbol{\Omega}$ such that $(x_n(\omega) - \overline{x}(\omega))_{n\in\mathbb{N}}$, $(\widetilde{p}_{1,n}(\omega) - \overline{x}(\omega))_{n\in\mathbb{N}}$ and $(\forall i \in \{1, \ldots, m\})$ $(v_{i,n}(\omega) - \overline{v}_i(\omega))_{n\in\mathbb{N}}$, $(\widetilde{p}_{2,i,n}(\omega) - \overline{v}_i(\omega))_{n\in\mathbb{N}}$ are bounded, and

$(\forall i \in \{1,\dots,m\})$ $\widetilde{p}_{2,i,n}(\omega)-\overline{v}_{i,n}(\omega) \to 0$, $\widetilde{p}_{1,n}(\omega)-x_n(\omega) \to 0$. Set $\mathbf{\Omega_4} = \mathbf{\Omega_3} \cap \widetilde{\mathbf{\Omega}}$. Then $\mathbf{\Omega_4}$ has probability 1. Now fix $\omega \in \mathbf{\Omega_4}$. Since A is uniformly monotone at $\overline{x}(\omega)$, using (4.16) and (4.22), there exists an increasing function $\phi_A\colon [0,+\infty[\to [0,+\infty]$ vanishing only at 0 such that, for every $n \in \mathbb{N}$,

$$\begin{aligned}\phi_A(\|\widetilde{p}_{1,n}(\omega)-\overline{x}(\omega)\|) &\leqslant \Big\langle \widetilde{p}_{1,n}(\omega)-\overline{x}(\omega) \mid \gamma_n(x_n(\omega) \\ &\quad -\widetilde{p}_{1,n}(\omega)) - \sum\nolimits_{i=1}^m (L_i^* v_{i,n}(\omega) - L_i^*\overline{v}_i(\omega))\Big\rangle - \chi_n(\omega) \\ &= \Big\langle \widetilde{p}_{1,n}(\omega)-\overline{x}(\omega) \mid \gamma_n^{-1}(x_n(\omega)-\widetilde{p}_{1,n}(\omega))\Big\rangle - \chi_n(\omega) \\ &\quad -\sum_{i=1}^m \big\langle \widetilde{p}_{1,n}(\omega)-\overline{x}(\omega) \mid L_i^* v_{i,n}(\omega) - L_i^*\overline{v}_i(\omega)\big\rangle, \end{aligned} \tag{4.23}$$

where we denote $\big(\forall n \in \mathbb{N}\big)$ $\chi_n(\omega) = \big\langle \widetilde{p}_{1,n}(\omega)-\bar{x}(\omega) \mid Cx_n(\omega) - C\bar{x}(\omega)\big\rangle$. Since $(B_i^{-1})_{1\leq i\leq m}$ are monotone, for every $i \in \{1,\dots,m\}$, we obtain

$$\begin{aligned}(\forall n \in \mathbb{N})\quad 0 &\leqslant \Big\langle \widetilde{p}_{2,i,n}(\omega)-\overline{v}_i(\omega) \mid L_i x_n(\omega) + \gamma_n^{-1}(v_{i,n}(\omega)-\widetilde{p}_{2,i,n}(\omega)) - L_i\overline{x}(\omega)\Big\rangle \\ &\quad - \beta_{i,n}(\omega) \\ &= \Big\langle \widetilde{p}_{2,i,n}(\omega)-\overline{v}_i(\omega) \mid L_i(x_n(\omega)-\overline{x}(\omega)) + \gamma_n^{-1}(v_{i,n}(\omega)-\widetilde{p}_{2,i,n}(\omega))\Big\rangle \\ &\quad - \beta_{i,n}(\omega), \end{aligned} \tag{4.24}$$

where $\big(\forall n \in \mathbb{N}\big)$ $\beta_{i,n}(\omega) = \Big\langle \widetilde{p}_{2,i,n}(\omega)-\overline{v}_i(\omega) \mid D_i^{-1}v_{i,n}(\omega) - D_i^{-1}\bar{v}_i(\omega)\Big\rangle$. Now, adding (4.24) from $i=1$ to $i=m$ and (4.23), we obtain, for every $n \in \mathbb{N}$,

$$\begin{aligned}\phi_A(\|\widetilde{p}_{1,n}(\omega)-\overline{x}(\omega)\|) &\leq \Big\langle \widetilde{p}_{1,n}(\omega)-\overline{x}(\omega) \mid \gamma_n^{-1}(x_n(\omega)-\widetilde{p}_{1,n}(\omega))\Big\rangle \\ &\quad + \Big\langle \widetilde{p}_{1,n}(\omega)-\overline{x}(\omega) \mid \sum_{i=1}^m L_i^*(\widetilde{p}_{2,i,n}(\omega)-v_{i,n}(\omega))\Big\rangle \\ &\quad + \sum_{i=1}^m \Big\langle \widetilde{p}_{2,i,n}(\omega)-\overline{v}_i(\omega) \mid L_i(x_n(\omega)-\widetilde{p}_{1,n}(\omega)) \\ &\quad +\gamma_n^{-1}(v_{i,n}(\omega)-\widetilde{p}_{2,i,n}(\omega))\Big\rangle \\ &\quad -\chi_n(\omega) - \sum_{i=1}^m \beta_{i,n}(\omega). \end{aligned} \tag{4.25}$$

For every $n \in \mathbb{N}$ and every $i \in \{1, \dots, m\}$, we expand $\chi_n(\omega)$ and $\beta_{i,n}(\omega)$ as

$$\begin{cases} \chi_n(\omega) = \langle x_n(\omega) - \overline{x}(\omega) \mid Cx_n(\omega) - C\overline{x}(\omega)\rangle \\ \qquad\qquad + \left\langle \widetilde{p}_{1,n}(\omega) - x_n(\omega) \mid Cx_n(\omega) - C\overline{x}(\omega)\right\rangle, \\ \beta_{i,n}(\omega) = \left\langle v_{i,n}(\omega) - \overline{v}_i(\omega) \mid D_i^{-1} v_{i,n}(\omega) - D_i^{-1}\overline{v}_i(\omega)\right\rangle \\ \qquad\qquad + \left\langle \widetilde{p}_{2,i,n}(\omega) - v_{i,n}(\omega) \mid D_i^{-1} v_{i,n}(\omega) - D_i^{-1}\overline{v}_i(\omega)\right\rangle. \end{cases} \tag{4.26}$$

By monotonicity of C and $(D_i^{-1})_{1\le i\le m}$,

$$(\forall n \in \mathbb{N}) \quad \begin{cases} \langle x_n(\omega) - \overline{x}(\omega) \mid Cx_n(\omega) - C\overline{x}(\omega)\rangle \geq 0, \\ (\forall i \in \{1, \dots, m\}) \ \left\langle v_{i,n}(\omega) - \overline{v}_i(\omega) \mid D_i^{-1} v_{i,n}(\omega) - D_i^{-1}\overline{v}_i(\omega)\right\rangle \geq 0. \end{cases} \tag{4.27}$$

Therefore, for every $n \in \mathbb{N}$, we derive from (4.26) and (4.25) that

$$\begin{aligned} \phi_A(\|\widetilde{p}_{1,n}(\omega) - \overline{x}(\omega)\|) &\leq \phi_A(\|\widetilde{p}_{1,n}(\omega) - \overline{x}(\omega)\|) + \langle x_n(\omega) - \overline{x}(\omega) \mid Cx_n(\omega) - C\overline{x}(\omega)\rangle \\ &\quad + \sum_{i=1}^m \left\langle v_{i,n}(\omega) - \overline{v}_i(\omega) \mid D_i^{-1} v_{i,n}(\omega) - D_i^{-1}\overline{v}_i(\omega)\right\rangle \\ &\leq \left\langle \widetilde{p}_{1,n}(\omega) - \overline{x}(\omega) \mid \gamma_n^{-1}(x_n(\omega) - \widetilde{p}_{1,n}(\omega))\right\rangle \\ &\quad + \left\langle \widetilde{p}_{1,n}(\omega) - \overline{x}(\omega) \mid \sum_{i=1}^m L_i^*(\widetilde{p}_{2,i,n}(\omega) - v_{i,n}(\omega))\right\rangle \\ &\quad + \sum_{i=1}^m \Big\langle \widetilde{p}_{2,i,n}(\omega) - \overline{v}_i(\omega) \mid L_i(x_n(\omega) - \widetilde{p}_{1,n}(\omega)) \\ &\quad + \gamma_n^{-1}(v_{i,n}(\omega) - \widetilde{p}_{2,i,n}(\omega))\Big\rangle \\ &\quad - \left\langle \widetilde{p}_{1,n}(\omega) - x_n(\omega) \mid Cx_n(\omega) - C\overline{x}(\omega)\right\rangle \\ &\quad - \sum_{i=1}^m \left\langle \widetilde{p}_{2,i,n}(\omega) - v_{i,n}(\omega) \mid D_i^{-1} v_{i,n}(\omega) - D_i^{-1}\overline{v}_i(\omega)\right\rangle. \end{aligned} \tag{4.28}$$

We set

$$\zeta = \max_{1\le i\le m} \sup_{n\in\mathbb{N}} \Big\{ \|x_n(\omega) - \overline{x}(\omega)\|, \|\widetilde{p}_{1,n}(\omega) - \overline{x}(\omega)\|, \|v_{i,n}(\omega) - \overline{v}_i(\omega)\|, \|\widetilde{p}_{2,i,n}(\omega) - \overline{v}_i(\omega)\| \Big\}. \tag{4.29}$$

Then it follows from the definition of $\mathbf{\Omega_4}$ that $\zeta < \infty$, and from our assumption that $(\forall n \in \mathbb{N})\ \gamma_n^{-1} \leq \varepsilon^{-1}$. Therefore, using the Cauchy–Schwarz inequality, and the Lipschitzianity of C and $(D_i^{-1})_{1\le i\le m}$, we derive from (4.28) that

$$\begin{aligned}\phi_A(\|\widetilde{p}_{1,n}(\omega)-\overline{x}(\omega)\|)\le\varepsilon^{-1}\zeta\,\|x_n(\omega)-\widetilde{p}_{1,n}(\omega)\|+\zeta\sum_{i=1}^m\big(\|L_i\|\ \|x_n(\omega)-\widetilde{p}_{1,n}(\omega)\|\\ +\,\varepsilon^{-1}\|v_{i,n}-\widetilde{p}_{2,i,n}\|\big)+\zeta\left(\sum_{i=1}^m\|L_i^*\|\,\|\widetilde{p}_{2,i,n}(\omega)-v_{i,n}(\omega)\|\right.\\ \left.+\nu_0\|\widetilde{p}_{1,n}(\omega)-x_n(\omega)\|+\sum_{i=1}^m\nu_i\,\|\widetilde{p}_{2,i,n}(\omega)-v_{i,n}(\omega)\|\right)\\ \to 0.\end{aligned}\tag{4.30}$$

We deduce from (4.30) and (4.21) that $\phi_A(\|\widetilde{p}_{1,n}(\omega)-\overline{x}(\omega)\|)\to 0$, which implies that $\widetilde{p}_{1,n}(\omega)\to\overline{x}(\omega)$. In turn, $x_n(\omega)\to\overline{x}(\omega)$. Likewise, if C is uniformly monotone at $\overline{x}(\omega)$, there exists an increasing function $\phi_C\colon[0,+\infty[\to[0,+\infty]$ that vanishes only at 0 such that

$$\begin{aligned}\phi_C(\|x_n(\omega)-\overline{x}(\omega)\|)\le\varepsilon^{-1}\zeta\,\|x_n(\omega)-\widetilde{p}_{1,n}(\omega)\|+\zeta\sum_{i=1}^m\big(\|L_i\|\ \|x_n(\omega)-\widetilde{p}_{1,n}(\omega)\|\\ +\,\varepsilon^{-1}\|v_{i,n}(\omega)-\widetilde{p}_{2,i,n}(\omega)\|\big)+\zeta\left(\sum_{i=1}^m\|L_i^*\|\,\|\widetilde{p}_{2,i,n}(\omega)\right.\\ \left.-\,v_{i,n}(\omega)\|+\nu_0\|\widetilde{p}_{1,n}(\omega)-x_n(\omega)\|+\sum_{i=1}^m\nu_i\,\|\widetilde{p}_{2,i,n}(\omega)-v_{i,n}(\omega)\|\right)\\ \to 0,\end{aligned}\tag{4.31}$$

in turn, $x_n(\omega)\to\overline{x}(\omega)$.

(ii)(d): Proceeding as in the proof of (ii)(c), we obtain the conclusions. □

We provide an application to minimization problems in [11, Section 4] which cover a wide class of convex optimization problems in the literature. We recall that the infimal convolution of the two functions f and g from $\mathcal{H}$ to $]-\infty,+\infty]$ is

$$f\,\square\,g\colon x\mapsto\inf_{y\in\mathcal{H}}(f(y)+g(x-y)).\tag{4.32}$$

The proximity operator of $f\in\Gamma_0(\mathcal{H})$, denoted by prox_f, which maps each point $x\in\mathcal{H}$ to the unique minimizer of the function $f+\frac{1}{2}\|x-\cdot\|^2$.

Example 4.3 Let m be a strictly positive integer. Let $\mathcal{H}$ be a real separable Hilbert space, let $z\in\mathcal{H}$, let $f\in\Gamma_0(\mathcal{H})$, let $h\colon\mathcal{H}\to\mathbb{R}$ be convex differentiable function with ν_0-Lipschitz continuous gradient, for some $\nu_0\in]0,+\infty[$. For every $k\in\{1,\ldots,m\}$, let $(\mathcal{G}_k,\langle\cdot\mid\cdot\rangle)$ be a real separable Hilbert space, let $r_k\in\mathcal{G}_k$, let $g_k\in\Gamma_0(\mathcal{G}_k)$, let $\ell_k\in\Gamma_0(\mathcal{G}_k)$ be $1/\nu_k$-strongly convex, for some $\nu_k\in]0,+\infty[$. For every $k\in\{1,\ldots,m\}$, let $L_k\colon\mathcal{H}\to\mathcal{G}_k$ be a bounded linear operator. The primal problems is to

$$\underset{x\in\mathcal{H}}{\text{minimize}}\big(f(x)-\langle x\mid z\rangle\big)+\sum_{k=1}^{m}(\ell_k\,\Box\,g_k)\Big(L_kx-r_k\Big)+h(x), \tag{4.33}$$

and the dual problem is to

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_m\in\mathcal{G}_m}{\text{minimize}}\;(f^*\,\Box\,h^*)\left(z-\sum_{i=1}^{m}L_i^*v_i\right)+\sum_{i=1}^{m}\big(g_i^*(v_i)+\ell_i^*(v_i)+\langle v_i\mid r_i\rangle\big). \tag{4.34}$$

We denote by $\mathcal{P}_1$ and $\mathcal{D}_1$ be the set of solutions to (4.33) and (4.34), respectively.

Corollary 4.4 *In Example 4.3, suppose that*

$$z\in\operatorname{ran}\left(\partial f+\sum_{i=1}^{m}L_i^*\big((\partial g_i\,\Box\,\partial\ell_i)(L_i\cdot-r_i)\big)+\nabla h\right). \tag{4.35}$$

Let $(a_{1,n})_{n\in\mathbb{N}}$, $(b_{1,n})_{n\in\mathbb{N}}$, and $(c_{1,n})_{n\in\mathbb{N}}$ be sequences of square integrable $\mathcal{H}$-valued random vectors, and for every $i\in\{1,\ldots,m\}$, let $(a_{2,i,n})_{n\in\mathbb{N}}$, $(b_{2,i,n})_{n\in\mathbb{N}}$, and $(c_{2,i,n})_{n\in\mathbb{N}}$ be sequences of square integrable $\mathcal{G}_i$-valued random vectors. Furthermore, set

$$\beta=\max\{\nu_0,\nu_1,\ldots,\nu_m\}+\sqrt{\sum_{i=1}^{m}\|L_i\|^2}, \tag{4.36}$$

let x_0 be a square integrable $\mathcal{H}$-valued random vector, and, for every $i\in\{1,\ldots,m\}$, let $v_{i,0}$ be a square integrable $\mathcal{G}_i$-valued random vector, let $\varepsilon\in]0,1/(1+\beta)[$, let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\varepsilon,(1-\varepsilon)/\beta]$. Set

$$(\forall n\in\mathbb{N})\quad\left\lfloor\begin{array}{l}y_{1,n}=x_n-\gamma_n\left(\nabla h(x_n)+\sum_{i=1}^{m}L_i^*v_{i,n}+a_{1,n}\right)\\ p_{1,n}=\operatorname{prox}_{\gamma_n f}(y_{1,n}+\gamma_n z)+b_{1,n}\\ \text{for } i=1,\ldots,m\\ \left\lfloor\begin{array}{l}y_{2,i,n}=v_{i,n}+\gamma_n\big(L_ix_n-\nabla\ell_i^*(v_{i,n})+a_{2,i,n}\big)\\ p_{2,i,n}=\operatorname{prox}_{\gamma_n g_i^*}(y_{2,i,n}-\gamma_n r_i)+b_{2,i,n}\\ q_{2,i,n}=p_{2,i,n}+\gamma_n\big(L_ip_{1,n}-\nabla\ell_i^*(p_{2,i,n})+c_{2,i,n}\big)\\ v_{i,n+1}=v_{i,n}-y_{2,i,n}+q_{2,i,n}\end{array}\right.\\ q_{1,n}=p_{1,n}-\gamma_n\left(\nabla h(p_{1,n})+\sum_{i=1}^{m}L_i^*p_{2,i,n}+c_{1,n}\right)\\ x_{n+1}=x_n-y_{1,n}+q_{1,n}.\end{array}\right. \tag{4.37}$$

Suppose that the following conditions hold for $\boldsymbol{\mathcal{F}}_n=\sigma((x_k,(v_{i,k})_{1\le i\le m})_{0\le k\le n}$,

$$\begin{cases}\sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||(a_{1,n},(a_{2,i,n})_{1\le i\le m})|||^2|\boldsymbol{\mathcal{F}}_n]}<+\infty\\ \sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[\|(b_{1,n},(b_{2,i,n})_{1\le i\le m})|||^2|\boldsymbol{\mathcal{F}}_n]}<+\infty\\ \sum_{n\in\mathbb{N}}\sqrt{\mathbf{E}[|||(c_{1,n},(c_{2,i,n})_{1\le i\le m})|||^2|\boldsymbol{\mathcal{F}}_n]}<+\infty.\end{cases} \tag{4.38}$$

Then the following hold.

(i) $\sum_{n\in\mathbb{N}} \mathbf{E}[\|x_n - p_{1,n}\|^2|\mathcal{F}_n] < +\infty$ *and* $(\forall i \in \{1,\ldots,m\})$ $\sum_{n\in\mathbb{N}} \mathbf{E}[\|v_{i,n} - p_{2,i,n}\|^2|\mathcal{F}_n] < +\infty$.

(ii) *There exist a* $\mathcal{P}_1$*-valued random vector* $\overline{x}$ *and a* $\mathcal{D}_1$*-valued random vector* $(\overline{v}_1,\ldots,\overline{v}_m)$ *such that the following hold.*

  (a) $x_n \rightharpoonup \overline{x}$ *and* $\operatorname{prox}_{\gamma_n f}(x_n - \gamma_n(\nabla h(x_n) + \sum_{i=1}^m L_i^* v_{i,n}) + \gamma_n z) \rightharpoonup \overline{x}$ P*-a.s.*

  (b) $(\forall i \in \{1,\ldots,m\})$ $v_{i,n} \rightharpoonup \overline{v}_i$ *and* $\operatorname{prox}_{\gamma g_i^*}(v_{i,n} + \gamma_n\big(L_i x_n - \nabla\ell_i^*(v_{i,n}) - \gamma_n r_i) \rightharpoonup \overline{v}_i$ P*-a.s.*

  (c) *Suppose that* f *or* ∇h *is uniformly convex at* $\overline{x}(\omega)$ *for every* $\omega \in \widetilde{\boldsymbol{\Omega}} \subset \boldsymbol{\Omega}$ *with* $\mathsf{P}(\widetilde{\boldsymbol{\Omega}}) = 1$*, then* $x_n \to \overline{x}$ *and* $\operatorname{prox}_{\gamma_n f}(x_n - \gamma_n(\nabla h(x_n) + \sum_{i=1}^m L_i^* v_{i,n}) + \gamma_n z) \to \overline{x}$ P*-a.s.*

  (d) *Suppose that* g_j^* *or* ℓ_j^* *is uniformly convex at* $\overline{v}_j(\omega)$ *for every* $\omega \in \widetilde{\boldsymbol{\Omega}} \subset \boldsymbol{\Omega}$ *with* $\mathsf{P}(\widetilde{\boldsymbol{\Omega}}) = 1$*, for some* $j \in \{1,\ldots,m\}$*, then* $v_{j,n} \to \overline{v}_j$ *and* $\operatorname{prox}_{\gamma_n g_j^*}(v_{j,n} + \gamma_n\big(L_j x_n - \nabla\ell_j^*(v_{j,n}) - \gamma_n r_j) \to \overline{v}_j$ P*-a.s.*

Proof Using the same argument as in the proof [11, Theorem 4.2], the conclusions follows from Corollary 4.2. □

Remark 4.5 Here are some comments.

(i) By using Remark 3.1, an extension of Corollary 4.2 to the variable metric setting is straightforward.

(ii) Almost sure convergence for some primal-dual splitting methods solving composite monotone inclusions and composite minimization problems are also presented in [12,19].

(iii) In the deterministic setting and in the case when each ℓ_k is the indicator function of $\{0\}$, and $(\forall k \in \{1,\ldots,m\}) r_k = 0$, and $z = 0$, a preconditioned algorithm for solving (4.33) can be found in [22].

Acknowledgments I thank Professor Patrick L. Combettes for helpful discussions. I thank the referees for their suggestions and correction which helped to improve the first version of the manuscript. This work is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant No. 102.01-2014.02.

References

1. Attouch, H., Briceño-Arias, L.M., Combettes, P.L.: A parallel splitting method for coupled monotone inclusions. SIAM J. Control Optim. **48**, 3246–3270 (2010)
2. Boţ, R.I., Hendrich, C.: An algorithm for solving monotone inclusions involving parallel sums of linearly composed maximally monotone operators (2013). arXiv:1306.3191
3. Briceño-Arias, L.M., Combettes, P.L.: A monotone+skew splitting model for composite monotone inclusions in duality. SIAM J. Optim. **21**, 1230–1250 (2011)
4. Briceño-Arias, L.M., Combettes, P.L.: Monotone operator methods for Nash equilibria in non-potential games. In: Bailey, D., Bauschke, H.H., Borwein, P., Garvan, F., Théra, M., Vanderwerff, J., Wolkowicz, H. (eds.) Computational and Analytical Mathematics. Springer, New York (2013)
5. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York (2011)
6. Barty, K., Roy, J.-S., Strugarek, C.: Hilbert-valued perturbed subgradient algorithms. Math. Oper. Res. **32**, 551–562 (2007)
7. Bennar, A., Monnez, J.-M.: Almost sure convergence of a stochastic approximation process in a convex set. Int. J. Appl. Math. **20**, 713–722 (2007)

8. Combettes, P.L.: Quasi-Fejérian analysis of some optimization algorithms. In: Butnariu, D., Censor, Y., Reich, S. (eds.) Inherently Parallel Algorithms in Feasibility and Optimization and their Applications, pp. 115–152. Elsevier, New York (2001)
9. Combettes, P.L.: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization **53**, 475–504 (2004)
10. Combettes, P.L.: Systems of structured monotone inclusions: duality, algorithms, and applications. SIAM J. Optim. **23**, 2420–2447 (2013)
11. Combettes, P.L., Pesquet, J.-C.: Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators. Set-Valued Var. Anal. **20**, 307–330 (2012)
12. Combettes, P.L., Pesquet, J.-C.: Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping, SIAM J. Optim (to appear)
13. Combettes, P.L., Vũ, B.C.: Variable metric quasi-Fejér monotonicity. Nonlinear Anal. **78**, 17–31 (2013)
14. Combettes, P.L., Vũ, B.C.: Variable metric forward-backward splitting with applications to monotone inclusions in duality. Optimization **63**, 1289–1318 (2014)
15. Dũng, D., Vũ, B.C.: A splitting algorithm for system of composite monotone inclusions, Vietnam J. Maths. (2014, to appear)
16. Ermol'ev, Y.M.: On the method of generalized stochastic gradients and quasi-Fejér sequences. Cybernetics **5**, 208–220 (1969)
17. Jiang, H., Xu, H.: Stochastic approximation approaches to the stochastic variational inequality problem. IEEE Trans. Automat. Control **53**, 1462–1475 (2008)
18. Juditsky, A., Nemirovski, A., Tauvel, C.: Solving variational inequalities with stochastic mirror-prox algorithm. Stoch. Syst. **1**, 17–58 (2011)
19. Pesquet, J.-C., Repetti, A.: A class of randomized primal-dual algorithms for distributed optimization (2014). arXiv:1406.6404
20. Polyak, B.: Introduction to Optimization. Optimization Software, Inc., New York (1987)
21. Rosasco, L., Villa, S., Vũ, B.C.: A stochastic forward-backward splitting method for solving monotone inclusions in Hilbert spaces (2014, preprint). arXiv:1403.7999
22. Repetti, A., Chouzenoux, E., Pesquet, J.-C.: A penalized weighted least squares approach for restoring data corrupted with signal-dependent noise. In: Proceedings of the 20th European Signal Processing (SIPCO 2012), Aug 27–31, pp. 1553–1557. Bucharest, Romania (2012)
23. Robbins, H., Siegmund, D.: A convergence theorem for non negative almost supermartingales and some applications. In: Rustagi, J.S. (ed.) Optimizing Methods in Statistic, pp. 233–257. Academic Press, New York (1971)
24. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**, 431–446 (2000)
25. Vũ, B.C.: A variable metric extension of the forward-backward-forward algorithm for monotone operators. Numer. Funct. Anal. Optim. **34**, 1050–1065 (2013)
26. Yousefian, F., Nedić, A., Shanbhag, U.V.: On stochastic gradient and subgradient methods with adaptive steplength sequences. Automatica **48**, 56–67 (2012)