SIAM J. OPTIM.
Vol. xx, pp. x x–x

Stochastic Quasi-Fejér Block-Coordinate Fixed Point Iterations with Random Sweeping*

Patrick L. Combettes† and Jean-Christophe Pesquet‡

Abstract. This work proposes block-coordinate fixed point algorithms with applications to nonlinear analysis and optimization in Hilbert spaces. The asymptotic analysis relies on a notion of stochastic quasi-Fejér monotonicity, which is thoroughly investigated. The iterative methods under consideration feature random sweeping rules to select arbitrarily the blocks of variables that are activated over the course of the iterations and they allow for stochastic errors in the evaluation of the operators. Algorithms using quasinonexpansive operators or compositions of averaged nonexpansive operators are constructed, and weak and strong convergence results are established for the sequences they generate. As a by-product, novel block-coordinate operator splitting methods are obtained for solving structured monotone inclusion and convex minimization problems. In particular, the proposed framework leads to random block-coordinate versions of the Douglas-Rachford and forward-backward algorithms and of some of their variants. In the standard case of $m = 1$ block, our results remain new as they incorporate stochastic perturbations.

Key words. Arbitrary sampling, block-coordinate algorithm, fixed-point algorithm, monotone operator splitting, primal-dual algorithm, stochastic quasi-Fejér sequence, stochastic algorithm, structured convex minimization problem.

AMS subject classifications. Primary 47H05; Secondary 65K05, 90C25, 94A08

1. Introduction.

The main advantage of block-coordinate algorithms is to result in implementations with reduced complexity and memory requirements per iteration. These benefits have long been recognized [3, 18, 50] and have become increasingly important in very large-scale problems. In addition, block-coordinate strategies may lead to faster [20] or distributed [41] implementations. In this paper, we propose a block-coordinate fixed point algorithmic framework to solve a variety of problems in Hilbertian nonlinear numerical analysis and optimization. Algorithmic fixed point theory in Hilbert spaces provides a unifying and powerful framework for the analysis and the construction of a wide array of solution methods in such problems [5, 7, 19, 22, 66]. Although several block-coordinate algorithms exist for solving specific optimization problems in Euclidean spaces, a framework for dealing with general fixed point methods in Hilbert spaces and which guarantees the convergence of the iterates does not seem to exist at present. In the proposed constructs, a random sweeping strategy is employed for selecting the blocks of coordinates which are activated over the iterations. The sweeping rule allows for an arbitrary sampling of the indices of the coordinates. Furthermore, the algorithms tolerate stochastic errors in the implementation of the operators. This paper provides

*Received by the editors XXX XX, 2014; accepted for publication XXX XX, 2015; published electronically DATE. http://www.siam.org/journals/mms/x-x/XXXX.html. This work was supported by the CNRS MASTODONS project under grant 2013MesureHD.

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

‡Université Paris-Est, Laboratoire d'Informatique Gaspard Monge – CNRS UMR 8049, F-77454, Marne la Vallée Cedex 2, France (jean-christophe.pesquet@univ-paris-est.fr)

SIAM J. OPTIM.
Vol. xx, pp. x x–x

the first general stochastic block-coordinate fixed point framework with guaranteed convergence of the iterates. It generates a wide range of new algorithms, which will be illustrated by numerical experiments elsewhere.

A main ingredient for proving the convergence of many fixed point algorithms is the fundamental concept of (quasi-)Fejér monotonicity [21, 23, 33, 57]. In Section 2, refining the seminal work of [34, 35, 36], we revisit this concept from a stochastic standpoint. By exploiting properties of almost super-martingales [59], we establish novel almost sure convergence results for an abstract stochastic iteration scheme. In Section 3, this scheme is applied to the design of block-coordinate algorithms for relaxed iterations of quasinonexpansive operators. A simple instance of such iterations is the Krasnosel'skiĭ–Mann method, which has found numerous applications [7, 17]. In Section 4, we design block-coordinate algorithms involving compositions of averaged nonexpansive operators. The results are used in Section 5 to construct block-coordinate algorithms for structured monotone inclusion and convex minimization problems. Splitting algorithms have recently become tools of choice in signal processing and machine learning; see, e.g., [17, 27, 29, 31, 56, 60]. Providing versatile block-coordinate versions of these algorithms is expected to benefit these emerging areas, as well as more traditional fields of applications of splitting methods, e.g., [39]. One of the offsprings of our work is an original block-coordinate primal-dual algorithm which can be employed to solve a large class of variational problems.

2. Stochastic quasi-Fejér monotonicity.

Fejér monotonicity has been exploited in various areas of nonlinear analysis and optimization to unify the convergence proofs of deterministic algorithms; see, e.g., [7, 23, 33, 57]. In the late 1960s, this notion was revisited in a stochastic setting in Euclidean spaces [34, 35, 36]. In this section, we investigate a notion of stochastic quasi-Fejér monotone sequence in Hilbert spaces and apply the results to a general stochastic iterative method. Throughout the paper, the following notation will be used.

Notation 2.1. H is a separable real Hilbert space with scalar product $\langle\cdot\mid\cdot\rangle$, associated norm $\|\cdot\|$, and Borel σ-algebra $\mathcal{B}$. Id denotes the identity operator on H and $\rightharpoonup$ and $\to$ denote, respectively, weak and strong convergence in H. 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 underlying probability space is $(\Omega,\mathcal{F},\mathsf{P})$. A H-valued random variable is a measurable map $x\colon(\Omega,\mathcal{F})\to(\mathsf{H},\mathcal{B})$. The σ-algebra generated by a family Φ of random variables is denoted by $\sigma(\Phi)$. Let $\mathscr{F}=(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a sequence of sub-sigma algebras of $\mathcal{F}$ such that $(\forall n\in\mathbb{N})$ $\mathcal{F}_n\subset\mathcal{F}_{n+1}$. We denote by $\ell_+(\mathscr{F})$ 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

$$(\forall p\in\,]0,+\infty[)\quad \ell_+^p(\mathscr{F})=\left\{(\xi_n)_{n\in\mathbb{N}}\in\ell_+(\mathscr{F})\;\middle|\;\sum_{n\in\mathbb{N}}\xi_n^p<+\infty\ \mathsf{P}\text{-a.s.}\right\}\tag{2.1}$$

and

$$\ell_+^\infty(\mathscr{F})=\left\{(\xi_n)_{n\in\mathbb{N}}\in\ell_+(\mathscr{F})\;\middle|\;\sup_{n\in\mathbb{N}}\xi_n<+\infty\ \mathsf{P}\text{-a.s.}\right\}.\tag{2.2}$$

Given a sequence $(x_n)_{n\in\mathbb{N}}$ of H-valued random variables, we define

$$\mathscr{X}=(\mathcal{X}_n)_{n\in\mathbb{N}},\quad\text{where}\quad(\forall n\in\mathbb{N})\quad\mathcal{X}_n=\sigma(x_0,\ldots,x_n).\tag{2.3}$$

Equalities and inequalities involving random variables will always be understood to hold P-almost surely, even if the expression "P-a.s." is not explicitly written. For background on probability in Hilbert spaces, see [37, 42].

Lemma 2.2.[59, Theorem 1] *Let $\mathscr{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ be a sequence of sub-sigma algebras of $\mathcal{F}$ such that $(\forall n \in \mathbb{N})\ \mathcal{F}_n \subset \mathcal{F}_{n+1}$. Let $(\alpha_n)_{n\in\mathbb{N}} \in \ell_+(\mathscr{F})$, $(\vartheta_n)_{n\in\mathbb{N}} \in \ell_+(\mathscr{F})$, $(\eta_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathscr{F})$, and $(\chi_n)_{n\in\mathbb{N}} \in \ell^1_+(\mathscr{F})$ be such that*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}(\alpha_{n+1}\,|\,\mathcal{F}_n) + \vartheta_n \leqslant (1+\chi_n)\alpha_n + \eta_n \quad \mathsf{P}\text{-a.s.} \tag{2.4}$$

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

Proposition 2.3. *Let F be a nonempty closed subset of H, let $\phi\colon [0,+\infty[\to [0,+\infty[$ be a strictly increasing function such that $\lim_{t\to+\infty}\phi(t) = +\infty$, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence of H-valued random variables. Suppose that, for every $\mathsf{z} \in \mathsf{F}$, there exist $(\chi_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell^1_+(\mathscr{X})$, $(\vartheta_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+(\mathscr{X})$, and $(\eta_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell^1_+(\mathscr{X})$ such that the following is satisfied* P-a.s.*:*

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}(\phi(\|x_{n+1}-\mathsf{z}\|)\,|\,\mathcal{X}_n) + \vartheta_n(\mathsf{z}) \leqslant (1+\chi_n(\mathsf{z}))\phi(\|x_n-\mathsf{z}\|) + \eta_n(\mathsf{z}). \tag{2.5}$$

Then the following hold:

(i) $(\forall \mathsf{z} \in \mathsf{F})\ \big[\sum_{n\in\mathbb{N}}\vartheta_n(\mathsf{z}) < +\infty$ P-a.s.$\big]$

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

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

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

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

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

Proof. (i): Fix $\mathsf{z} \in \mathsf{F}$. It follows from (2.5) and Lemma 2.2 that $\sum_{n\in\mathbb{N}}\vartheta_n(\mathsf{z}) < +\infty$ P-a.s.

(ii): Let $\mathsf{z} \in \mathsf{F}$ and set $(\forall n \in \mathbb{N})\ \xi_n = \|x_n - \mathsf{z}\|$. We derive from (2.5) and Lemma 2.2 that $(\phi(\xi_n))_{n\in\mathbb{N}}$ converges P-a.s., say $\phi(\xi_n) \to \alpha$ P-a.s., where α is a $[0,+\infty[$-valued random variable. In turn, since $\lim_{t\to+\infty}\phi(t) = +\infty$, $(\xi_n)_{n\in\mathbb{N}}$ is bounded P-a.s. and so is $(x_n)_{n\in\mathbb{N}}$. For subsequent use, let us also note that

$$(\|x_n - \mathsf{z}\|)_{n\in\mathbb{N}} \text{ converges to a } [0,+\infty[\text{-valued random variable } \mathsf{P}\text{-a.s.} \tag{2.6}$$

Indeed, take $\omega \in \Omega$ such that $(\xi_n(\omega))_{n\in\mathbb{N}}$ is bounded. Suppose that there exist $\tau(\omega) \in [0,+\infty[$, $\zeta(\omega) \in [0,+\infty[$, and subsequences $(\xi_{k_n}(\omega))_{n\in\mathbb{N}}$ and $(\xi_{l_n}(\omega))_{n\in\mathbb{N}}$ such that $\xi_{k_n}(\omega) \to \tau(\omega)$ and $\xi_{l_n}(\omega) \to \zeta(\omega) > \tau(\omega)$, and let $\delta(\omega) \in\,]0, (\zeta(\omega)-\tau(\omega))/2[$. Then, for n sufficiently large, $\xi_{k_n}(\omega) \leqslant \tau(\omega) + \delta(\omega) < \zeta(\omega) - \delta(\omega) \leqslant \xi_{l_n}(\omega)$ and, since ϕ is strictly increasing, $\phi(\xi_{k_n}(\omega)) \leqslant \phi(\tau(\omega)+\delta(\omega)) < \phi(\zeta(\omega)-\delta(\omega)) \leqslant \phi(\xi_{l_n}(\omega))$. Taking the limit as $n \to +\infty$ yields $\alpha(\omega) \leqslant \phi(\tau(\omega)+\delta(\omega)) < \phi(\zeta(\omega)-\delta(\omega)) \leqslant \alpha(\omega)$, which is impossible. It follows that $\tau(\omega) = \zeta(\omega)$ and, in turn, that $\xi_n(\omega) \to \tau(\omega)$. Thus, $\xi_n \to \tau$ P-a.s.

(iii): Since H is separable, there exists a countable set Z such that $\overline{\mathsf{Z}} = \mathsf{F}$. According to (2.6), for every $\mathsf{z} \in \mathsf{F}$, there exists a set $\Omega_{\mathsf{z}} \in \mathcal{F}$ such that $\mathsf{P}(\Omega_{\mathsf{z}}) = 1$ and, for every $\omega \in \Omega_{\mathsf{z}}$, the sequence $(\|x_n(\omega)-\mathsf{z}\|)_{n\in\mathbb{N}}$ converges. Now set $\widetilde{\Omega}=\bigcap_{\mathsf{z}\in\mathsf{Z}}\Omega_{\mathsf{z}}$ and let $\complement\widetilde{\Omega}$ be its complement. Then, since Z is countable, $\mathsf{P}(\widetilde{\Omega})=1-\mathsf{P}(\complement\widetilde{\Omega})=1-\mathsf{P}(\bigcup_{\mathsf{z}\in\mathsf{Z}}\complement\Omega_{\mathsf{z}})\geqslant 1-\sum_{\mathsf{z}\in\mathsf{Z}}\mathsf{P}(\complement\Omega_{\mathsf{z}})=1$, hence $\mathsf{P}(\widetilde{\Omega})=1$. We now fix $\mathsf{z}\in\mathsf{F}$. Since $\overline{\mathsf{Z}}=\mathsf{F}$, there exists a sequence $(\mathsf{z}_k)_{k\in\mathbb{N}}$ in Z such that $\mathsf{z}_k\to\mathsf{z}$. As just seen, (2.6) yields

$$(\forall k\in\mathbb{N})(\exists\,\tau_k\colon\Omega\to[0,+\infty[)(\forall\omega\in\Omega_{\mathsf{z}_k})\quad \|x_n(\omega)-\mathsf{z}_k\|\to\tau_k(\omega). \tag{2.7}$$

Now let $\omega\in\widetilde{\Omega}$. We have

$$(\forall k\in\mathbb{N})(\forall n\in\mathbb{N})\quad -\|\mathsf{z}_k-\mathsf{z}\|\leqslant\|x_n(\omega)-\mathsf{z}\|-\|x_n(\omega)-\mathsf{z}_k\|\leqslant\|\mathsf{z}_k-\mathsf{z}\|. \tag{2.8}$$

Therefore

$$\begin{aligned}(\forall k\in\mathbb{N})\quad -\|\mathsf{z}_k-\mathsf{z}\| &\leqslant \varliminf_{n\to+\infty}\|x_n(\omega)-\mathsf{z}\|-\lim_{n\to+\infty}\|x_n(\omega)-\mathsf{z}_k\|\\ &=\varliminf_{n\to+\infty}\|x_n(\omega)-\mathsf{z}\|-\tau_k(\omega)\\ &\leqslant\varlimsup_{n\to+\infty}\|x_n(\omega)-\mathsf{z}\|-\tau_k(\omega)\\ &=\varlimsup_{n\to+\infty}\|x_n(\omega)-\mathsf{z}\|-\lim_{n\to+\infty}\|x_n(\omega)-\mathsf{z}_k\|\\ &\leqslant\|\mathsf{z}_k-\mathsf{z}\|.\end{aligned} \tag{2.9}$$

Hence, taking the limit as $k\to+\infty$ in (2.9), we obtain that $(\|x_n(\omega)-\mathsf{z}\|)_{n\in\mathbb{N}}$ converges; more precisely, $\lim_{n\to+\infty}\|x_n(\omega)-\mathsf{z}\|=\lim_{k\to+\infty}\tau_k(\omega)$.

(iv): By assumption, there exists $\widehat{\Omega}\in\mathcal{F}$ such that $\mathsf{P}(\widehat{\Omega})=1$ and $(\forall\omega\in\widehat{\Omega})$ $\mathfrak{W}(x_n(\omega))_{n\in\mathbb{N}}\subset\mathsf{F}$. Now define $\widetilde{\Omega}$ as in the proof of (iii), let $\omega\in\widehat{\Omega}\cap\widetilde{\Omega}$, and let $x(\omega)$ and $y(\omega)$ be two points in $\mathfrak{W}(x_n(\omega))_{n\in\mathbb{N}}$, say $x_{k_n}(\omega)\rightharpoonup x(\omega)$ and $x_{l_n}(\omega)\rightharpoonup y(\omega)$. Then (iii) implies that $(\|x_n(\omega)-x(\omega)\|)_{n\in\mathbb{N}}$ and $(\|x_n(\omega)-y(\omega)\|)_{n\in\mathbb{N}}$ converge. In turn, since

$$\begin{aligned}(\forall n\in\mathbb{N})\quad &\langle x_n(\omega)\mid x(\omega)-y(\omega)\rangle\\ &=\frac{1}{2}\big(\|x_n(\omega)-y(\omega)\|^2-\|x_n(\omega)-x(\omega)\|^2+\|x(\omega)\|^2-\|y(\omega)\|^2\big),\end{aligned} \tag{2.10}$$

the sequence $(\langle x_n(\omega)\mid x(\omega)-y(\omega)\rangle)_{n\in\mathbb{N}}$ converges, say

$$\langle x_n(\omega)\mid x(\omega)-y(\omega)\rangle\to\varrho(\omega). \tag{2.11}$$

However, since $x_{k_n}(\omega)\rightharpoonup x(\omega)$, we have $\langle x(\omega)\mid x(\omega)-y(\omega)\rangle=\varrho(\omega)$. Likewise, passing to the limit along the subsequence $(x_{l_n}(\omega))_{n\in\mathbb{N}}$ in (2.11) yields

$$\langle y(\omega)\mid x(\omega)-y(\omega)\rangle=\varrho(\omega). \tag{2.12}$$

Thus,

$$0=\langle x(\omega)\mid x(\omega)-y(\omega)\rangle-\langle y(\omega)\mid x(\omega)-y(\omega)\rangle=\|x(\omega)-y(\omega)\|^2. \tag{2.13}$$

SIAM J. OPTIM.
Vol. xx, pp. x x–x

This shows that $x(\omega) = y(\omega)$. Since $\omega \in \widetilde{\Omega}$, $(x_n(\omega))_{n\in\mathbb{N}}$ is bounded and we invoke [7, Lemma 2.38] to conclude that $x_n(\omega) \rightharpoonup x(\omega) \in \mathsf{F}$. Altogether, since $\mathsf{P}(\widehat{\Omega}\cap\widetilde{\Omega}) = 1$, $x_n \rightharpoonup x$ P-a.s. and the measurability of x follows from [55, Corollary 1.13].

(v): Let $x \in \mathfrak{S}(x_n)_{n\in\mathbb{N}}\cap\mathsf{F}$ P-a.s. Then there exists $\widehat{\Omega}\in\mathcal{F}$ such that $\mathsf{P}(\widehat{\Omega}) = 1$ and $(\forall\omega\in\widehat{\Omega})$ $\underline{\lim}\,\|x_n(\omega) - x(\omega)\| = 0$. Now let $\widetilde{\Omega}$ be as in (iii) and let $\omega \in \widetilde{\Omega}\cap\widehat{\Omega}$. Then $\mathsf{P}(\widetilde{\Omega}\cap\widehat{\Omega}) = 1$, $x(\omega)\in\mathsf{F}$, and (iii) implies that $(\|x_n(\omega)-x(\omega)\|)_{n\in\mathbb{N}}$ converges. Thus, $\lim\|x_n(\omega)-x(\omega)\| = 0$. We conclude that $x_n \to x$ P-a.s.

(vi)⇒(v): Since $\varnothing \neq \mathfrak{S}(x_n)_{n\in\mathbb{N}} \subset \mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{F}$ P-a.s., we have $\mathfrak{S}(x_n)_{n\in\mathbb{N}}\cap\mathsf{F}\neq\varnothing$ P-a.s. ∎

Remark 2.4. Suppose that $\phi\colon t\mapsto t^2$ in (2.5). Then special cases of Proposition 2.3 are stated in several places in the literature. Thus, stochastic quasi-Fejér sequences were first discussed in [34] in the case when H is a Euclidean space and for every $n\in\mathbb{N}$, $\vartheta_n = \chi_n = 0$ and η_n is deterministic. A Hilbert space version of the results of [34] appears in [4] without proof. Finally, the case when all the processes are deterministic in (2.5) is discussed in [21].

The analysis of our main algorithms will rely on the following key illustration of Proposition 2.3. This result involves a general stochastic iterative process and it should also be of interest in the analysis of the asymptotic behavior of a broad class of stochastic algorithms, beyond those discussed in the present paper.

Theorem 2.5. *Let* F *be a nonempty closed subset of* H, *let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,1]$, *and let* $(t_n)_{n\in\mathbb{N}}$, $(x_n)_{n\in\mathbb{N}}$, *and* $(e_n)_{n\in\mathbb{N}}$ *be sequences of* H*-valued random variables. Suppose that the following hold:*

(i) $(\forall n\in\mathbb{N})\ x_{n+1} = x_n + \lambda_n(t_n + e_n - x_n)$.

(ii) $\sum_{n\in\mathbb{N}}\lambda_n\sqrt{\mathsf{E}(\|e_n\|^2\,|\,\mathfrak{X}_n)} < +\infty$ P-a.s.

(iii) *For every* $\mathsf{z}\in\mathsf{F}$, *there exist* $(\theta_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+(\mathscr{X})$, $(\mu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^\infty(\mathscr{X})$, *and* $(\nu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^\infty(\mathscr{X})$ *such that* $(\lambda_n\mu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^1(\mathscr{X})$, $(\lambda_n\nu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^{1/2}(\mathscr{X})$, *and the following is satisfied* P-a.s.*:*

$$(2.14)\qquad (\forall n\in\mathbb{N})\quad \mathsf{E}(\|t_n-\mathsf{z}\|^2\,|\,\mathfrak{X}_n) + \theta_n(\mathsf{z}) \leqslant (1+\mu_n(\mathsf{z}))\|x_n-\mathsf{z}\|^2 + \nu_n(\mathsf{z}).$$

Then

$$(2.15)\qquad (\forall\mathsf{z}\in\mathsf{F})\quad \left[\sum_{n\in\mathbb{N}}\lambda_n\theta_n(\mathsf{z}) < +\infty \text{ P-a.s.}\right]$$

and

$$(2.16)\qquad \sum_{n\in\mathbb{N}}\lambda_n(1-\lambda_n)\mathsf{E}(\|t_n-x_n\|^2\,|\,\mathfrak{X}_n) < +\infty \text{ P-a.s.}$$

Furthermore, suppose that:

(iv) $\mathfrak{W}(x_n)_{n\in\mathbb{N}}\subset\mathsf{F}$ P-a.s.

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* F*-valued random variable* x. *If, in addition,*

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

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

Proof. Let $\mathsf{z}\in\mathsf{F}$ and set

$$(2.17)\qquad (\forall n\in\mathbb{N})\qquad \varepsilon_n = \lambda_n\sqrt{\mathsf{E}(\|e_n\|^2\,|\,\mathfrak{X}_n)}.$$

It follows from Jensen's inequality and (iii) that

$$(\forall n\in\mathbb{N})\quad \mathsf{E}(\|t_n-\mathsf{z}\|\,|\,\mathcal{X}_n)\leqslant\sqrt{\mathsf{E}(\|t_n-\mathsf{z}\|^2\,|\,\mathcal{X}_n)}$$
$$\leqslant\sqrt{(1+\mu_n(\mathsf{z}))\|x_n-\mathsf{z}\|^2+\nu_n(\mathsf{z})}$$
$$\leqslant\sqrt{1+\mu_n(\mathsf{z})}\|x_n-\mathsf{z}\|+\sqrt{\nu_n(\mathsf{z})}$$
$$\leqslant(1+\mu_n(\mathsf{z})/2)\|x_n-\mathsf{z}\|+\sqrt{\nu_n(\mathsf{z})}. \tag{2.18}$$

On the other hand, (i) and the triangle inequality yield

$$(\forall n\in\mathbb{N})\quad \|x_{n+1}-\mathsf{z}\|\leqslant(1-\lambda_n)\|x_n-\mathsf{z}\|+\lambda_n\|t_n-\mathsf{z}\|+\lambda_n\|e_n\|. \tag{2.19}$$

Consequently,

$$(\forall n\in\mathbb{N})\quad \mathsf{E}(\|x_{n+1}-\mathsf{z}\|\,|\,\mathcal{X}_n)\leqslant(1-\lambda_n)\|x_n-\mathsf{z}\|+\lambda_n\mathsf{E}(\|t_n-\mathsf{z}\|\,|\,\mathcal{X}_n)$$
$$+\lambda_n\mathsf{E}(\|e_n\|\,|\,\mathcal{X}_n)$$
$$\leqslant\Big(1+\frac{\lambda_n\mu_n(\mathsf{z})}{2}\Big)\|x_n-\mathsf{z}\|$$
$$+\lambda_n\sqrt{\nu_n(\mathsf{z})}+\lambda_n\sqrt{\mathsf{E}(\|e_n\|^2\,|\,\mathcal{X}_n)}$$
$$=\Big(1+\frac{\lambda_n\mu_n(\mathsf{z})}{2}\Big)\|x_n-\mathsf{z}\|+\sqrt{\lambda_n\nu_n(\mathsf{z})}+\varepsilon_n. \tag{2.20}$$

Upon applying Proposition 2.3(ii) with $\phi\colon t\mapsto t$, we deduce from (2.20) that $(x_n)_{n\in\mathbb{N}}$ is almost surely bounded and, by virtue of assumption (iii), that $(\mathsf{E}(\|t_n-\mathsf{z}\|^2\,|\,\mathcal{X}_n))_{n\in\mathbb{N}}$ is likewise. Thus, there exist $]0,+\infty[$-valued random variables $\rho_1(\mathsf{z})$ and $\rho_2(\mathsf{z})$ such that, almost surely,

$$(\forall n\in\mathbb{N})\quad \|x_n-\mathsf{z}\|\leqslant\rho_1(\mathsf{z})\quad\text{and}\quad\sqrt{\mathsf{E}(\|t_n-\mathsf{z}\|^2\,|\,\mathcal{X}_n)}\leqslant\rho_2(\mathsf{z}). \tag{2.21}$$

Now set

$$(\forall n\in\mathbb{N})\quad \begin{cases}\chi_n(\mathsf{z})=\lambda_n\mu_n(\mathsf{z})\\ \xi_n(\mathsf{z})=2\lambda_n(1-\lambda_n)\|x_n-\mathsf{z}\|\,\|e_n\|+2\lambda_n^2\|t_n-\mathsf{z}\|\,\|e_n\|+\lambda_n^2\|e_n\|^2\\ \vartheta_n(\mathsf{z})=\lambda_n\theta_n(\mathsf{z})+\lambda_n(1-\lambda_n)\mathsf{E}(\|t_n-x_n\|^2\,|\,\mathcal{X}_n)\\ \eta_n(\mathsf{z})=\mathsf{E}(\xi_n(\mathsf{z})\,|\,\mathcal{X}_n)+\lambda_n\nu_n(\mathsf{z}).\end{cases} \tag{2.22}$$

On the one hand, it follows from (2.22), the Cauchy-Schwarz inequality, and (2.17) that

$$(\forall n\in\mathbb{N})\quad \mathsf{E}(\xi_n(\mathsf{z})\,|\,\mathcal{X}_n)=2\lambda_n(1-\lambda_n)\|x_n-\mathsf{z}\|\,\mathsf{E}(\|e_n\|\,|\,\mathcal{X}_n)$$
$$+2\lambda_n^2\mathsf{E}(\|t_n-\mathsf{z}\|\,\|e_n\|\,|\,\mathcal{X}_n)+\lambda_n^2\mathsf{E}(\|e_n\|^2\,|\,\mathcal{X}_n)$$
$$\leqslant 2\lambda_n\|x_n-\mathsf{z}\|\sqrt{\mathsf{E}(\|e_n\|^2\,|\,\mathcal{X}_n)}$$
$$+2\lambda_n\sqrt{\mathsf{E}(\|t_n-\mathsf{z}\|^2\,|\,\mathcal{X}_n)}\sqrt{\mathsf{E}(\|e_n\|^2\,|\,\mathcal{X}_n)}$$
$$+\lambda_n^2\mathsf{E}(\|e_n\|^2\,|\,\mathcal{X}_n)$$
$$\leqslant 2(\rho_1(\mathsf{z})+\rho_2(\mathsf{z}))\varepsilon_n+\varepsilon_n^2. \tag{2.23}$$

In turn, we deduce from (2.17), (2.22), (ii), and (iii) that

$$\big(\eta_n(\mathsf{z})\big)_{n\in\mathbb{N}} \in \ell_+^1(\mathscr{X}) \quad\text{and}\quad \big(\chi_n(\mathsf{z})\big)_{n\in\mathbb{N}} \in \ell_+^1(\mathscr{X}). \tag{2.24}$$

On the other hand, we derive from (i), [7, Corollary 2.14], and (2.22) that

$$\begin{aligned}(\forall n\in\mathbb{N})\quad \|x_{n+1}-\mathsf{z}\|^2 &= \|(1-\lambda_n)(x_n-\mathsf{z})+\lambda_n(t_n-\mathsf{z})\|^2\\ &\quad + 2\lambda_n\big\langle(1-\lambda_n)(x_n-\mathsf{z})+\lambda_n\big(t_n-\mathsf{z}\big)\;\big|\; e_n\big\rangle+\lambda_n^2\|e_n\|^2\\ &\leqslant (1-\lambda_n)\|x_n-\mathsf{z}\|^2+\lambda_n\|t_n-\mathsf{z}\|^2\\ &\quad -\lambda_n(1-\lambda_n)\|t_n-x_n\|^2+\xi_n(\mathsf{z}).\end{aligned} \tag{2.25}$$

Hence, (iii), (2.22), and (2.23) imply that

$$\begin{aligned}&(\forall n\in\mathbb{N})\quad \mathsf{E}(\|x_{n+1}-\mathsf{z}\|^2\,|\,\mathfrak{X}_n)\\ &\quad\leqslant (1-\lambda_n)\|x_n-\mathsf{z}\|^2+\lambda_n\mathsf{E}(\|t_n-\mathsf{z}\|^2\,|\,\mathfrak{X}_n)-\lambda_n(1-\lambda_n)\mathsf{E}(\|t_n-x_n\|^2\,|\,\mathfrak{X}_n)\\ &\qquad +\mathsf{E}(\xi_n(\mathsf{z})\,|\,\mathfrak{X}_n)\\ &\quad\leqslant (1-\lambda_n)\|x_n-\mathsf{z}\|^2+\lambda_n\big((1+\mu_n(\mathsf{z}))\|x_n-\mathsf{z}\|^2+\nu_n(\mathsf{z})-\theta_n(\mathsf{z})\big)\\ &\qquad -\lambda_n(1-\lambda_n)\mathsf{E}(\|t_n-x_n\|^2\,|\,\mathfrak{X}_n)+\mathsf{E}(\xi_n(\mathsf{z})\,|\,\mathfrak{X}_n)\\ &\quad\leqslant (1+\chi_n(\mathsf{z}))\|x_n-\mathsf{z}\|^2-\vartheta_n(\mathsf{z})+\eta_n(\mathsf{z}).\end{aligned} \tag{2.26}$$

Thus, in view of (2.24), applying Proposition 2.3(i) with $\phi\colon t\mapsto t^2$ yields $\sum_{n\in\mathbb{N}}\vartheta_n(\mathsf{z})<+\infty$ P-a.s. and it follows from (2.22) that (2.15) and (2.16) are established. Finally, the weak convergence assertion follows from (iv) and Proposition 2.3(iv) applied with $\phi\colon t\mapsto t^2$. Likewise, the strong convergence assertion follows from (iv)–(v) and Proposition 2.3(vi) applied with $\phi\colon t\mapsto t^2$. ■

Definition 2.6. *An operator* $\mathsf{T}\colon\mathsf{H}\to\mathsf{H}$ *is* nonexpansive *if it is* 1*-Lipschitz, and* demicompact *at* $\mathsf{y}\in\mathsf{H}$ *if for every bounded sequence* $(\mathsf{y}_n)_{n\in\mathbb{N}}$ *in* H *such that* $\mathsf{T}\mathsf{y}_n-\mathsf{y}_n\to\mathsf{y}$*, we have* $\mathfrak{S}(\mathsf{y}_n)_{n\in\mathbb{N}}\neq\varnothing$ *[54].*

Although our primary objective is to apply Theorem 2.5 to block-coordinate methods, it also yields new results for classical methods. As an illustration, the following application describes a Krasnosel'skiĭ–Mann iteration with stochastic errors.

Corollary 2.7. *Let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $[0,1]$ *such that* $\sum_{n\in\mathbb{N}}\lambda_n(1-\lambda_n)=+\infty$ *and let* $\mathsf{T}\colon\mathsf{H}\to\mathsf{H}$ *be a nonexpansive operator such that* $\operatorname{Fix}\mathsf{T}\neq\varnothing$*. Let* x_0 *and* $(e_n)_{n\in\mathbb{N}}$ *be* H*-valued random variables. Iterate*

$$\begin{array}{l}\text{for } n=0,1,\ldots\\ \big\lfloor\; x_{n+1}=x_n+\lambda_n\big(\mathsf{T}x_n+e_n-x_n\big).\end{array} \tag{2.27}$$

In addition, assume that $\sum_{n\in\mathbb{N}}\lambda_n\sqrt{\mathsf{E}(\|e_n\|^2\,|\,\mathfrak{X}_n)}<+\infty$ P-a.s. *Then the following hold:*

(i) $(x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to a* $(\operatorname{Fix}\mathsf{T})$*-valued random variable.*

(ii) *Suppose that* T *is demicompact at* 0 *(see Definition 2.6). Then* $(x_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to a* $(\operatorname{Fix}\mathsf{T})$*-valued random variable.*

Proof. Set $\mathsf{F} = \operatorname{Fix}\mathsf{T}$. Since T is continuous, T is measurable and F is closed. Now let $\mathsf{z} \in \mathsf{F}$ and set $(\forall n \in \mathbb{N})$ $t_n = \mathsf{T}x_n$. Then, using the nonexpansiveness of T, we obtain

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_{n+1} = x_n + \lambda_n(t_n + e_n - x_n) \\ \mathsf{E}(\|t_n - x_n\|^2 \,|\, \mathcal{X}_n) = \|\mathsf{T}x_n - x_n\|^2 \\ \mathsf{E}(\|t_n - \mathsf{z}\|^2 \,|\, \mathcal{X}_n) = \|\mathsf{T}x_n - \mathsf{T}\mathsf{z}\|^2 \leqslant \|x_n - \mathsf{z}\|^2. \end{cases} \tag{2.28}$$

It follows that properties (i)–(iii) in Theorem 2.5 are satisfied with $(\forall n \in \mathbb{N})$ $\theta_n = 0$, $\mu_n = 0$, and $\nu_n = 0$. Hence, (2.16) and (2.28) imply the existence of $\widetilde{\Omega} \in \mathcal{F}$ such that $\mathsf{P}(\widetilde{\Omega}) = 1$ and

$$(\forall \omega \in \widetilde{\Omega}) \quad \sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n)\|\mathsf{T}x_n(\omega) - x_n(\omega)\|^2 < +\infty. \tag{2.29}$$

Moreover,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \|\mathsf{T}x_{n+1} - x_{n+1}\| &= \|\mathsf{T}x_{n+1} - \mathsf{T}x_n + (1-\lambda_n)(\mathsf{T}x_n - x_n) - \lambda_n e_n\| \\ &\leqslant \|\mathsf{T}x_{n+1} - \mathsf{T}x_n\| + (1-\lambda_n)\|\mathsf{T}x_n - x_n\| + \lambda_n\|e_n\| \\ &\leqslant \|x_{n+1} - x_n\| + (1-\lambda_n)\|\mathsf{T}x_n - x_n\| + \lambda_n\|e_n\| \\ &\leqslant \lambda_n\|\mathsf{T}x_n - x_n\| + (1-\lambda_n)\|\mathsf{T}x_n - x_n\| + 2\lambda_n\|e_n\| \\ &= \|\mathsf{T}x_n - x_n\| + 2\lambda_n\|e_n\| \end{aligned} \tag{2.30}$$

and, therefore,

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad \mathsf{E}(\|\mathsf{T}x_{n+1} - x_{n+1}\| \,|\, \mathcal{X}_n) &\leqslant \|\mathsf{T}x_n - x_n\| + 2\lambda_n \mathsf{E}(\|e_n\| \,|\, \mathcal{X}_n) \\ &\leqslant \|\mathsf{T}x_n - x_n\| + 2\lambda_n\sqrt{\mathsf{E}(\|e_n\|^2 \,|\, \mathcal{X}_n)}. \end{aligned} \tag{2.31}$$

In turn, Lemma 2.2 implies that there exists $\widehat{\Omega} \subset \widetilde{\Omega}$ such that $\widehat{\Omega} \in \mathcal{F}$, $\mathsf{P}(\widehat{\Omega}) = 1$, and, for every $\omega \in \widehat{\Omega}$, $(\|\mathsf{T}x_n(\omega) - x_n(\omega)\|)_{n\in\mathbb{N}}$ converges.

(i): It is enough to establish property (iv) of Theorem 2.5. Let $\omega \in \widehat{\Omega}$ and let $\mathsf{x} \in \mathfrak{W}(x_n(\omega))_{n\in\mathbb{N}}$, say $x_{k_n}(\omega) \rightharpoonup \mathsf{x}$. In view of (2.29), since $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$, we have $\underline{\lim}\, \|\mathsf{T}x_n(\omega) - x_n(\omega)\| = 0$. Therefore,

$$\|\mathsf{T}x_n(\omega) - x_n(\omega)\| \to 0. \tag{2.32}$$

Altogether, $x_{k_n}(\omega) \rightharpoonup \mathsf{x}$ and $\mathsf{T}x_{k_n}(\omega) - x_{k_n}(\omega) \to 0$. Since T is nonexpansive, the demiclosedness principle [7, Corollary 4.18] asserts that $\mathsf{x} \in \mathsf{F}$.

(ii): It is enough to establish property (v) of Theorem 2.5. Let $\omega \in \widehat{\Omega}$. As shown above, $(x_n(\omega))_{n\in\mathbb{N}}$ converges weakly and it is therefore bounded [7, Lemma 2.38]. Hence, by demicompactness, (2.32) implies that $\mathfrak{S}(x_n(\omega))_{n\in\mathbb{N}} \neq \varnothing$. Thus, $\mathfrak{S}(x_n)_{n\in\mathbb{N}} \neq \varnothing$ P-a.s. ■

Remark 2.8. Corollary 2.7 extends [21, Theorem 5.5], which is restricted to deterministic processes and therefore less realistic error models. As shown in [7, 19, 21], the Krasnosel'skiĭ–Mann iteration process is at the core of many algorithms in variational problems and optimization. Corollary 2.7 therefore provides stochastically perturbed versions of these algorithms.

3. Single-layer random block-coordinate fixed point algorithms. In the remainder of the paper, the following notation will be used.

Notation 3.1. $\mathsf{H}_1,\ldots,\mathsf{H}_m$ are separable real Hilbert spaces and $\mathbf{H}=\mathsf{H}_1\oplus\cdots\oplus\mathsf{H}_m$ is their direct Hilbert sum. The scalar products and associated norms of these spaces are all denoted by $\langle\cdot\mid\cdot\rangle$ and $\|\cdot\|$, respectively, and $\mathbf{x}=(\mathsf{x}_1,\ldots,\mathsf{x}_m)$ denotes a generic vector in $\mathbf{H}$. Given a sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}=(x_{1,n},\ldots,x_{m,n})_{n\in\mathbb{N}}$ of $\mathbf{H}$-valued random variables, we set $(\forall n\in\mathbb{N})$ $\boldsymbol{\mathfrak{X}}_n=\sigma(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_n)$.

We recall that an operator $\mathbf{T}\colon\mathbf{H}\to\mathbf{H}$ with fixed point set $\operatorname{Fix}\mathbf{T}$ is quasinonexpansive if [7]

$$(\forall\mathbf{z}\in\operatorname{Fix}\mathbf{T})(\forall\mathbf{x}\in\mathbf{H})\quad\|\mathbf{T}\mathbf{x}-\mathbf{z}\|\leqslant\|\mathbf{x}-\mathbf{z}\|. \tag{3.1}$$

Theorem 3.2. *Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0,1[$ such that $\inf_{n\in\mathbb{N}}\lambda_n>0$ and $\sup_{n\in\mathbb{N}}\lambda_n<1$ and set $\mathsf{D}=\{0,1\}^m\smallsetminus\{\mathbf{0}\}$. For every $n\in\mathbb{N}$, let $\mathbf{T}_n\colon\mathbf{H}\to\mathbf{H}\colon\mathbf{x}\mapsto(\mathsf{T}_{i,n}\,\mathbf{x})_{1\leqslant i\leqslant m}$ be a quasinonexpansive operator where, for every $i\in\{1,\ldots,m\}$, $\mathsf{T}_{i,n}\colon\mathbf{H}\to\mathsf{H}_i$ is measurable. Let $\boldsymbol{x}_0$ and $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ be $\mathbf{H}$-valued random variables, and let $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ be identically distributed D-valued random variables. Iterate*

$$\begin{array}{l}\text{for } n=0,1,\ldots\\ \left\lfloor\begin{array}{l}\text{for } i=1,\ldots,m\\ \left\lfloor x_{i,n+1}=x_{i,n}+\varepsilon_{i,n}\lambda_n\big(\mathsf{T}_{i,n}\,(x_{1,n},\ldots,x_{m,n})+a_{i,n}-x_{i,n}\big),\right.\end{array}\right.\end{array} \tag{3.2}$$

and set $(\forall n\in\mathbb{N})$ $\boldsymbol{\mathcal{E}}_n=\sigma(\boldsymbol{\varepsilon}_n)$. In addition, assume that the following hold:

(i) $\mathbf{F}=\bigcap_{n\in\mathbb{N}}\operatorname{Fix}\mathbf{T}_n\neq\varnothing$.
(ii) $\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)}<+\infty$.
(iii) *For every $n\in\mathbb{N}$, $\boldsymbol{\mathcal{E}}_n$ and $\boldsymbol{\mathfrak{X}}_n$ are independent.*
(iv) $(\forall i\in\{1,\ldots,m\})$ $\mathsf{p}_i=\mathsf{P}[\varepsilon_{i,0}=1]>0$.

Then

$$\mathbf{T}_n\boldsymbol{x}_n-\boldsymbol{x}_n\to\mathbf{0}\ \ \mathsf{P}\text{-a.s.} \tag{3.3}$$

Furthermore, suppose that:

(v) $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}}\subset\mathbf{F}$ P-a.s.

Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to an $\mathbf{F}$-valued random variable $\boldsymbol{x}$. If, in addition,

(vi) $\mathfrak{S}(\boldsymbol{x}_n)_{n\in\mathbb{N}}\neq\varnothing$ P-a.s.,

then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to $\boldsymbol{x}$.

Proof. We define a norm $|||\cdot|||$ on $\mathbf{H}$ by

$$(\forall\mathbf{x}\in\mathbf{H})\quad|||\mathbf{x}|||^2=\sum_{i=1}^m\frac{1}{\mathsf{p}_i}\|\mathsf{x}_i\|^2. \tag{3.4}$$

We are going to apply Theorem 2.5 in $(\mathbf{H},|||\cdot|||)$. Let us set

$$(\forall n\in\mathbb{N})\quad\begin{cases}\boldsymbol{t}_n=(t_{i,n})_{1\leqslant i\leqslant m}\\ \boldsymbol{e}_n=(\varepsilon_{i,n}a_{i,n})_{1\leqslant i\leqslant m},\end{cases} \tag{3.5}$$

where $(\forall i\in\{1,\ldots,m\})$ $t_{i,n}=x_{i,n}+\varepsilon_{i,n}(\mathsf{T}_{i,n}\,\boldsymbol{x}_n-x_{i,n})$.

Then it follows from (3.2) that

$$(\forall n\in\mathbb{N})\quad \boldsymbol{x}_{n+1}=\boldsymbol{x}_n+\lambda_n\big(\boldsymbol{t}_n+\boldsymbol{e}_n-\boldsymbol{x}_n\big), \tag{3.6}$$

while (ii) implies that

$$\sum_{n\in\mathbb{N}}\lambda_n\mathsf{E}(|||\boldsymbol{e}_n|||^2\,|\,\boldsymbol{\mathfrak{X}}_n)\leqslant\sum_{n\in\mathbb{N}}\mathsf{E}(|||\boldsymbol{a}_n|||^2\,|\,\boldsymbol{\mathfrak{X}}_n)<+\infty. \tag{3.7}$$

Since the operators $(\mathbf{T}_n)_{n\in\mathbb{N}}$ are quasinonexpansive, $\mathbf{F}$ is closed [6, Section 2]. Now let $\mathbf{z}\in\mathbf{F}$ and set

$$(\forall n\in\mathbb{N})(\forall i\in\{1,\ldots,m\})\quad \mathsf{q}_{i,n}\colon \mathbf{H}\times\mathsf{D}\to\mathbb{R}\colon(\mathbf{x},\boldsymbol{\epsilon})\mapsto\|\mathsf{x}_i-\mathsf{z}_i+\epsilon_i(\mathsf{T}_{i,n}\,\mathbf{x}-\mathsf{x}_i)\|^2. \tag{3.8}$$

Note that, for every $n\in\mathbb{N}$ and every $i\in\{1,\ldots,m\}$, since $\mathsf{T}_{i,n}$ is measurable, so are the functions $(\mathsf{q}_{i,n}(\cdot,\boldsymbol{\epsilon}))_{\boldsymbol{\epsilon}\in\mathsf{D}}$. Consequently, since, for every $n\in\mathbb{N}$, (iii) asserts that the events $([\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}])_{\boldsymbol{\epsilon}\in\mathsf{D}}$ form an almost sure partition of Ω and are independent from $\boldsymbol{\mathfrak{X}}_n$, and since the random variables $(\mathsf{q}_{i,n}(\boldsymbol{x}_n,\boldsymbol{\epsilon}))_{\substack{1\leqslant i\leqslant m\\ \boldsymbol{\epsilon}\in\mathsf{D}}}$ are $\boldsymbol{\mathfrak{X}}_n$-measurable, we obtain [44, Section 28.2]

$$\begin{aligned}(\forall n\in\mathbb{N})(\forall i\in\{1,\ldots,m\})\quad \mathsf{E}(\|t_{i,n}-\mathsf{z}_i\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)&=\mathsf{E}\bigg(\mathsf{q}_{i,n}(\boldsymbol{x}_n,\boldsymbol{\varepsilon}_n)\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}1_{[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]}\,\bigg|\,\boldsymbol{\mathfrak{X}}_n\bigg)\\
&=\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}\mathsf{E}(\mathsf{q}_{i,n}(\boldsymbol{x}_n,\boldsymbol{\epsilon})1_{[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]}\,|\,\boldsymbol{\mathfrak{X}}_n)\\
&=\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}\mathsf{E}(1_{[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]}\,|\,\boldsymbol{\mathfrak{X}}_n)\mathsf{q}_{i,n}(\boldsymbol{x}_n,\boldsymbol{\epsilon})\\
&=\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\mathsf{q}_{i,n}(\boldsymbol{x}_n,\boldsymbol{\epsilon}).\end{aligned} \tag{3.9}$$

Thus, (3.4), (3.5), (3.9), (3.8), (iv), and (3.1) yield

$$\begin{aligned}(\forall n\in\mathbb{N})\quad &\mathsf{E}(|||\boldsymbol{t}_n-\mathbf{z}|||^2\,|\,\boldsymbol{\mathfrak{X}}_n)\\
&=\sum_{i=1}^m\frac{1}{\mathsf{p}_i}\mathsf{E}(\|t_{i,n}-\mathsf{z}_i\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)\\
&=\sum_{i=1}^m\frac{1}{\mathsf{p}_i}\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\|x_{i,n}-\mathsf{z}_i+\epsilon_i(\mathsf{T}_{i,n}\,\boldsymbol{x}_n-x_{i,n})\|^2\\
&=\sum_{i=1}^m\frac{1}{\mathsf{p}_i}\Bigg(\sum_{\boldsymbol{\epsilon}\in\mathsf{D},\epsilon_i=1}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\,\|\mathsf{T}_{i,n}\,\boldsymbol{x}_n-\mathsf{z}_i\|^2\\
&\quad+\sum_{\boldsymbol{\epsilon}\in\mathsf{D},\,\epsilon_i=0}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\,\|x_{i,n}-\mathsf{z}_i\|^2\Bigg)\\
&=\|\mathbf{T}_{\boldsymbol{n}}\boldsymbol{x}_n-\mathbf{z}\|^2+\sum_{i=1}^m\frac{1-\mathsf{p}_i}{\mathsf{p}_i}\|x_{i,n}-\mathsf{z}_i\|^2\\
&=|||\boldsymbol{x}_n-\mathbf{z}|||^2+\|\mathbf{T}_n\boldsymbol{x}_n-\mathbf{z}\|^2-\|\boldsymbol{x}_n-\mathbf{z}\|^2\\
&\leqslant|||\boldsymbol{x}_n-\mathbf{z}|||^2.\end{aligned} \tag{3.10}$$

Altogether, properties (i)–(iii) of Theorem 2.5 are satisfied with $(\forall n \in \mathbb{N})\ \theta_n = \mu_n = \nu_n = 0$. We therefore derive from (2.16) that $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n)\mathsf{E}(|||\boldsymbol{t}_n - \boldsymbol{x}_n|||^2 \,|\, \boldsymbol{\mathfrak{X}}_n) < +\infty$ P-a.s. In view of our conditions on $(\lambda_n)_{n\in\mathbb{N}}$, this yields

$$\mathsf{E}(|||\boldsymbol{t}_n - \boldsymbol{x}_n|||^2 \,|\, \boldsymbol{\mathfrak{X}}_n) \to 0 \ \ \mathsf{P}\text{-a.s.} \tag{3.11}$$

On the other hand, proceeding as in (3.9) leads to

$$(\forall n \in \mathbb{N})(\forall i \in \{1,\ldots,m\}) \quad \mathsf{E}(\|t_{i,n} - x_{i,n}\|^2 \,|\, \boldsymbol{\mathfrak{X}}_n) = \sum_{\boldsymbol{\epsilon}\in\mathsf{D}} \epsilon_i \mathsf{P}[\boldsymbol{\varepsilon}_n = \boldsymbol{\epsilon}] \|\mathsf{T}_{i,n}\, \boldsymbol{x}_n - x_{i,n}\|^2. \tag{3.12}$$

Hence, it follows from (3.4), (3.5), and (iv) that

$$\begin{aligned}
(\forall n \in \mathbb{N}) \quad \mathsf{E}(|||\boldsymbol{t}_n - \boldsymbol{x}_n|||^2 \,|\, \boldsymbol{\mathfrak{X}}_n) &= \sum_{i=1}^m \frac{1}{\mathsf{p}_i} \mathsf{E}(\|t_{i,n} - x_{i,n}\|^2 \,|\, \boldsymbol{\mathfrak{X}}_n) \\
&= \sum_{i=1}^m \frac{1}{\mathsf{p}_i} \sum_{\boldsymbol{\epsilon}\in\mathsf{D}} \epsilon_i \mathsf{P}[\boldsymbol{\varepsilon}_n = \boldsymbol{\epsilon}] \|\mathsf{T}_{i,n}\, \boldsymbol{x}_n - x_{i,n}\|^2 \\
&= \sum_{i=1}^m \frac{1}{\mathsf{p}_i} \sum_{\boldsymbol{\epsilon}\in\mathsf{D}, \epsilon_i=1} \mathsf{P}[\boldsymbol{\varepsilon}_n = \boldsymbol{\epsilon}] \|\mathsf{T}_{i,n}\, \boldsymbol{x}_n - x_{i,n}\|^2 \\
&= \|\mathbf{T}_n \boldsymbol{x}_n - \boldsymbol{x}_n\|^2.
\end{aligned} \tag{3.13}$$

Accordingly, (3.11) yields $\mathbf{T}_n\boldsymbol{x}_n - \boldsymbol{x}_n \to \mathbf{0}$ P-a.s. In turn, the weak and strong convergence assertions are consequences of Theorem 2.5. ■

Remark 3.3. Let us make a few comments about Theorem 3.2.

(i) The binary variable $\varepsilon_{i,n}$ signals whether the ith coordinate $\mathsf{T}_{i,n}$ of the operator $\mathbf{T}_n$ is activated or not at iteration n.

(ii) Assumption (iv) guarantees that each operator in $(\mathsf{T}_{i,n})_{1\leqslant i\leqslant m}$ is activated with a nonzero probability at each iteration n of algorithm (4.1). The simplest scenario corresponds to the case when the block sweeping process assigns nonzero probabilities to multivariate indices $\boldsymbol{\epsilon} \in \mathsf{D}$ having a single component equal to 1. Then only one of the operators in $(\mathsf{T}_{i,n})_{1\leqslant i\leqslant m}$ is activated randomly. In general, the sweeping rule allows for an arbitrary sampling of the indices $\{1,\ldots,m\}$.

(iii) In view of (3.3), (v) is satisfied if there exists $\widehat{\Omega} \in \mathcal{F}$ such that $\mathsf{P}(\widehat{\Omega}) = 1$ and

$$(\forall \omega \in \widehat{\Omega}) \quad \Big[\mathbf{T}_n\boldsymbol{x}_n(\omega) - \boldsymbol{x}_n(\omega) \to \mathbf{0} \quad \Rightarrow \quad \mathfrak{W}(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}} \subset \mathbf{F}\Big]. \tag{3.14}$$

In the deterministic case, this is akin to the focusing conditions of [5]; see [5, 6, 21] for examples of suitable sequences $(\mathbf{T}_n)_{n\in\mathbb{N}}$. Likewise, (vi) is satisfied if there exists $\widehat{\Omega} \in \mathcal{F}$ such that $\mathsf{P}(\widehat{\Omega}) = 1$ and

$$(\forall \omega \in \widehat{\Omega}) \quad \Big[\sup_{n\in\mathbb{N}} \|\boldsymbol{x}_n(\omega)\| < +\infty \quad \text{and} \quad \mathbf{T}_n\boldsymbol{x}_n(\omega) - \boldsymbol{x}_n(\omega) \to \mathbf{0}\Big] \quad \Rightarrow \quad \mathfrak{S}(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}} \neq \varnothing. \tag{3.15}$$

In the deterministic case, this is the demicompactness regularity condition of [21, Definition 6.5]. Examples of suitable sequences $(\mathbf{T}_n)_{n\in\mathbb{N}}$ are provided in [21].

Our first corollary is a random block-coordinate version of the Krasnosel'skiĭ–Mann iteration.

Corollary 3.4. *Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0,1[$ such that $\inf_{n\in\mathbb{N}}\lambda_n>0$ and $\sup_{n\in\mathbb{N}}\lambda_n<1$, set $\mathsf{D}=\{0,1\}^m\smallsetminus\{\mathbf{0}\}$, and let $\mathbf{T}\colon\mathbf{H}\to\mathbf{H}\colon\mathbf{x}\mapsto(\mathsf{T}_i\,\mathbf{x})_{1\leqslant i\leqslant m}$ be a nonexpansive operator such that $\operatorname{Fix}\mathbf{T}\neq\varnothing$ where, for every $i\in\{1,\ldots,m\}$, $\mathsf{T}_i\colon\mathbf{H}\to\mathsf{H}_i$. Let $\boldsymbol{x}_0$ and $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ be $\mathbf{H}$-valued random variables, and let $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ be identically distributed D-valued random variables. Iterate*

$$\text{(3.16)}\qquad \begin{array}{l}\text{for } n=0,1,\ldots\\ \left\lfloor\begin{array}{l}\text{for } i=1,\ldots,m\\ \left\lfloor\; x_{i,n+1}=x_{i,n}+\varepsilon_{i,n}\lambda_n\big(\mathsf{T}_i\,(x_{1,n},\ldots,x_{m,n})+a_{i,n}-x_{i,n}\big),\right.\end{array}\right.\end{array}$$

and set $(\forall n\in\mathbb{N})$ $\boldsymbol{\mathcal{E}}_n=\sigma(\boldsymbol{\varepsilon}_n)$. In addition, assume that properties (ii)–(iv) of Theorem 3.2 hold. Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. *to a* $(\operatorname{Fix}\mathbf{T})$*-valued random variable. The convergence is strong if $\mathbf{T}$ is demicompact at $\mathsf{0}$ (see Definition 2.6).*

Proof. This is an application of Theorem 3.2 with $\mathbf{F}=\operatorname{Fix}\mathbf{T}$ and $(\forall n\in\mathbb{N})$ $\mathbf{T}_n=\mathbf{T}$. Indeed, (3.14) follows from the demiclosedness principle [7, Corollary 4.18] and (3.15) follows from the demicompactness assumption. ■

Remark 3.5. A special case of Corollary 3.4 appears in [41]. It corresponds to the scenario in which $\mathbf{H}$ is finite-dimensional, $\mathbf{T}$ is firmly nonexpansive, and, for every $n\in\mathbb{N}$, $\lambda_n=1$, $\boldsymbol{a}_n=\mathbf{0}$, and only one block is activated as in Remark 3.3(ii). Let us also note that a renorming similar to that performed in (3.4) was employed in [49].

Next, we consider the construction of a fixed point of a family of averaged operators.

Definition 3.6. *Let $\mathsf{T}\colon\mathsf{H}\to\mathsf{H}$ be nonexpansive and let $\alpha\in\,]0,1[$. Then T is* averaged *with constant α, or α-averaged, if there exists a nonexpansive operator $\mathsf{R}\colon\mathsf{H}\to\mathsf{H}$ such that* $\mathsf{T}=(1-\alpha)\mathrm{Id}+\alpha\mathsf{R}$.

Proposition 3.7. [7, Proposition 4.25] *Let $\mathsf{T}\colon\mathsf{H}\to\mathsf{H}$ be nonexpansive and let $\alpha\in\,]0,1[$. Then T is α-averaged if and only if*

$$\text{(3.17)}\qquad (\forall\mathsf{x}\in\mathsf{H})(\forall\mathsf{y}\in\mathsf{H})\quad \|\mathsf{T}\mathsf{x}-\mathsf{T}\mathsf{y}\|^2\leqslant\|\mathsf{x}-\mathsf{y}\|^2-\frac{1-\alpha}{\alpha}\|(\mathrm{Id}\,-\mathsf{T})\mathsf{x}-(\mathrm{Id}\,-\mathsf{T})\mathsf{y}\|^2.$$

Corollary 3.8. *Let $\chi\in\,]0,1[$, let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence in $]0,1[$, and set $\mathsf{D}=\{0,1\}^m\smallsetminus\{\mathbf{0}\}$. For every $n\in\mathbb{N}$, let $\lambda_n\in[\chi/\alpha_n,(1-\chi)/\alpha_n]$ and let $\mathbf{T}_n\colon\mathbf{H}\to\mathbf{H}\colon\mathbf{x}\mapsto(\mathsf{T}_{i,n}\,\mathbf{x})_{1\leqslant i\leqslant m}$ be an α_n-averaged operator, where, for every $i\in\{1,\ldots,m\}$, $\mathsf{T}_{i,n}\colon\mathbf{H}\to\mathsf{H}_i$. Furthermore, let $\boldsymbol{x}_0$ and $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ be $\mathbf{H}$-valued random variables, and let $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ be identically distributed D-valued random variables. Iterate*

$$\text{(3.18)}\qquad \begin{array}{l}\text{for } n=0,1,\ldots\\ \left\lfloor\begin{array}{l}\text{for } i=1,\ldots,m\\ \left\lfloor\; x_{i,n+1}=x_{i,n}+\varepsilon_{i,n}\lambda_n\big(\mathsf{T}_{i,n}\,(x_{1,n},\ldots,x_{m,n})+a_{i,n}-x_{i,n}\big),\right.\end{array}\right.\end{array}$$

and set $(\forall n\in\mathbb{N})$ $\boldsymbol{\mathcal{E}}_n=\sigma(\boldsymbol{\varepsilon}_n)$. Furthermore, assume that there exists $\widehat{\Omega}\in\mathcal{F}$ such that $\mathsf{P}(\widehat{\Omega})=1$ and the following hold:

(i) $\mathbf{F}=\bigcap_{n\in\mathbb{N}}\operatorname{Fix}\mathbf{T}_n\neq\varnothing$.

(ii) $\sum_{n\in\mathbb{N}}\alpha_n^{-1}\sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2\,|\,\boldsymbol{\mathcal{X}}_n)}<+\infty$.

(iii) *For every* $n \in \mathbb{N}$, $\boldsymbol{\mathcal{E}}_n$ *and* $\boldsymbol{\mathfrak{X}}_n$ *are independent.*
(iv) $(\forall i \in \{1,\ldots,m\})$ $\mathsf{P}[\varepsilon_{i,0} = 1] > 0$.
(v) $(\forall \omega \in \widehat{\Omega})$ $\Big[\alpha_n^{-1}\big(\mathbf{T}_n\boldsymbol{x}_n(\omega) - \boldsymbol{x}_n(\omega)\big) \to \mathbf{0} \Rightarrow \mathfrak{W}(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}} \subset \mathbf{F}\Big]$.

Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* $\mathbf{F}$*-valued random variable* $\boldsymbol{x}$*. If, in addition,*
(vi) $(\forall \omega \in \widehat{\Omega})$ $\Big[\big[\sup_{n\in\mathbb{N}} \|\boldsymbol{x}_n(\omega)\| < +\infty$ *and* $\alpha_n^{-1}\big(\mathbf{T}_n\boldsymbol{x}_n(\omega) - \boldsymbol{x}_n(\omega)\big) \to \mathbf{0}\big] \Rightarrow \mathfrak{S}(\boldsymbol{x}_n(\omega))_{n\in\mathbb{N}} \neq \varnothing\Big]$,

then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to* $\boldsymbol{x}$.

Proof. Set $(\forall n \in \mathbb{N})$ $\mathbf{R}_n = (1-\alpha_n^{-1})\mathbf{Id} + \alpha_n^{-1}\mathbf{T}_n$ and $(\forall i \in \{1,\ldots,m\})$ $\mathsf{R}_{i,n} = (1-\alpha_n^{-1})\mathrm{Id} + \alpha_n^{-1}\mathsf{T}_{i,n}$. Moreover, set $(\forall n \in \mathbb{N})$ $\mu_n = \alpha_n\lambda_n$ and $\boldsymbol{b}_n = \alpha_n^{-1}\boldsymbol{a}_n$. Then $(\forall n \in \mathbb{N})$ $\operatorname{Fix}\mathbf{R}_n = \operatorname{Fix}\mathbf{T}_n$ and $\mathbf{R}_n$ is nonexpansive. In addition, we derive from (3.18) that

$$(\forall n \in \mathbb{N})(\forall i \in \{1,\ldots,m\}) \quad x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\mu_n\big(\mathsf{R}_{i,n}\,\boldsymbol{x}_n + b_{i,n} - x_{i,n}\big). \tag{3.19}$$

Since $(\mu_n)_{n\in\mathbb{N}}$ lies in $[\chi, 1-\chi]$ and

$$\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2 \,|\, \boldsymbol{\mathfrak{X}}_n)} = \sum_{n\in\mathbb{N}} \alpha_n^{-1}\sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2 \,|\, \boldsymbol{\mathfrak{X}}_n)} < +\infty, \tag{3.20}$$

the result follows from Theorem 3.2 and Remark 3.3(iii). ∎

Remark 3.9. In the special case of a single-block (i.e., $m = 1$) and of deterministic errors, Corollary 3.8 reduces to a scenario found in [22, Theorem 4.2].

4. Double-layer random block-coordinate fixed point algorithms.

The algorithm analyzed in this section comprises two successive applications of nonexpansive operators at each iteration. We recall that Notation 3.1 is in force.

Theorem 4.1. *Let* $(\alpha_n)_{n\in\mathbb{N}}$ *and* $(\beta_n)_{n\in\mathbb{N}}$ *be sequences in* $]0,1[$ *such that* $\sup_{n\in\mathbb{N}}\alpha_n < 1$ *and* $\sup_{n\in\mathbb{N}}\beta_n < 1$*, let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,1]$ *such that* $\inf_{n\in\mathbb{N}}\lambda_n > 0$*, and set* $\mathsf{D} = \{0,1\}^m \smallsetminus \{\mathbf{0}\}$*. Let* $\boldsymbol{x}_0$*,* $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$*, and* $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ *be* $\mathbf{H}$*-valued random variables, and let* $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ *be identically distributed* D*-valued random variables. For every* $n \in \mathbb{N}$*, let* $\mathbf{R}_n\colon \mathbf{H} \to \mathbf{H}$ *be* β_n*-averaged and let* $\mathbf{T}_n\colon \mathbf{H} \to \mathbf{H}\colon \mathbf{x} \mapsto (\mathsf{T}_{i,n}\mathbf{x})_{1\leqslant i\leqslant m}$ *be* α_n*-averaged, where* $(\forall i \in \{1,\ldots,m\})$ $\mathsf{T}_{i,n}\colon \mathbf{H} \to \mathsf{H}_i$*. Iterate*

$$\begin{array}{l} \text{for } n = 0,1,\ldots \\ \left\lfloor \begin{array}{l} \boldsymbol{y}_n = \mathbf{R}_n\boldsymbol{x}_n + \boldsymbol{b}_n \\ \text{for } i = 1,\ldots,m \\ \left\lfloor x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\lambda_n\big(\mathsf{T}_{i,n}\boldsymbol{y}_n + a_{i,n} - x_{i,n}\big), \right. \end{array} \right. \end{array} \tag{4.1}$$

and set $(\forall n \in \mathbb{N})$ $\boldsymbol{\mathcal{E}}_n = \sigma(\boldsymbol{\varepsilon}_n)$*. In addition, assume that the following hold:*
(i) $\mathbf{F} = \bigcap_{n\in\mathbb{N}} \operatorname{Fix}(\mathbf{T}_n \circ \mathbf{R}_n) \neq \varnothing$.
(ii) $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2 \,|\, \boldsymbol{\mathfrak{X}}_n)} < +\infty$ *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2 \,|\, \boldsymbol{\mathfrak{X}}_n)} < +\infty$.
(iii) *For every* $n \in \mathbb{N}$*,* $\boldsymbol{\mathcal{E}}_n$ *and* $\boldsymbol{\mathfrak{X}}_n$ *are independent.*
(iv) $(\forall i \in \{1,\ldots,m\})$ $\mathsf{p}_i = \mathsf{P}[\varepsilon_{i,0} = 1] > 0$.

Then

$$\big[\ (\forall \mathbf{z} \in \mathbf{F})\ \ \mathbf{T}_n(\mathbf{R}_n\boldsymbol{x}_n) - \mathbf{R}_n\boldsymbol{x}_n + \mathbf{R}_n\mathbf{z} \to \mathbf{z}\ \big] \text{ P-a.s.} \tag{4.2}$$

and

$$\Big[\,(\forall \mathsf{z} \in \mathsf{F})\;\; \boldsymbol{x}_n - \mathsf{R}_n\boldsymbol{x}_n + \mathsf{R}_n\mathsf{z} \to \mathsf{z}\,\Big]\ \mathsf{P}\text{-a.s.} \tag{4.3}$$

Furthermore, suppose that:
(v) $\mathfrak{W}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \subset \mathsf{F}$ P-a.s.
Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* F*-valued random variable* $\boldsymbol{x}$*. If, in addition,*
(vi) $\mathfrak{S}(\boldsymbol{x}_n)_{n\in\mathbb{N}} \neq \varnothing$ P-a.s.,
then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to* $\boldsymbol{x}$*.*

Proof. Let us prove that the result is an application of Theorem 2.5 in the renormed Hilbert space $(\mathsf{H}, |||\cdot|||)$, where $|||\cdot|||$ is defined in (3.4). Note that

$$(\forall \mathsf{x} \in \mathsf{H})\quad \|\mathsf{x}\|^2 \leqslant |||\mathsf{x}|||^2 \leqslant \frac{1}{\min\limits_{1\leqslant i\leqslant m} \mathsf{p}_i}\|\mathsf{x}\|^2 \tag{4.4}$$

and that, since the operators $(\mathsf{R}_n \circ \mathsf{T}_n)_{n\in\mathbb{N}}$ are nonexpansive, the sets $(\operatorname{Fix}(\mathsf{T}_n \circ \mathsf{R}_n))_{n\in\mathbb{N}}$ are closed [7, Corollary 4.15], and so is F. Next, for every $n \in \mathbb{N}$, set $\boldsymbol{r}_n = \mathsf{R}_n\boldsymbol{x}_n$, and define $\boldsymbol{t}_n$, $\boldsymbol{c}_n$, $\boldsymbol{d}_n$, and $\boldsymbol{e}_n$ coordinatewise by

$$(\forall i \in \{1,\ldots,m\})\quad \begin{cases} t_{i,n} = x_{i,n} + \varepsilon_{i,n}(\mathsf{T}_{i,n}\boldsymbol{r}_n - x_{i,n})\\ c_{i,n} = \varepsilon_{i,n}a_{i,n}\\ d_{i,n} = \varepsilon_{i,n}(\mathsf{T}_{i,n}\boldsymbol{y}_n - \mathsf{T}_{i,n}\boldsymbol{r}_n)\end{cases} \quad\text{and}\quad e_{i,n} = c_{i,n} + d_{i,n}. \tag{4.5}$$

Then (4.1) implies that

$$(\forall n \in \mathbb{N})\quad \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n\big(\boldsymbol{t}_n + \boldsymbol{e}_n - \boldsymbol{x}_n\big). \tag{4.6}$$

On the other hand, we derive from (4.5) that

$$\begin{aligned}(\forall n \in \mathbb{N})\quad \sqrt{\mathsf{E}(|||\boldsymbol{e}_n|||^2 \,|\, \mathfrak{X}_n)} &\leqslant \sqrt{\mathsf{E}(|||\boldsymbol{c}_n|||^2 \,|\, \mathfrak{X}_n)} + \sqrt{\mathsf{E}(|||\boldsymbol{d}_n|||^2 \,|\, \mathfrak{X}_n)}\\ &\leqslant \sqrt{\mathsf{E}(|||\boldsymbol{a}_n|||^2 \,|\, \mathfrak{X}_n)} + \sqrt{\mathsf{E}(|||\boldsymbol{d}_n|||^2 \,|\, \mathfrak{X}_n)}.\end{aligned} \tag{4.7}$$

However, it follows from (4.5), (4.4), and the nonexpansiveness of the operators $(\mathsf{T}_n)_{n\in\mathbb{N}}$ that

$$\begin{aligned}(\forall n \in \mathbb{N})\quad \mathsf{E}(|||\boldsymbol{d}_n|||^2 \,|\, \mathfrak{X}_n) &\leqslant \frac{1}{\min\limits_{1\leqslant i\leqslant m} \mathsf{p}_i}\mathsf{E}\bigg(\sum_{i=1}^m \|\varepsilon_{i,n}(\mathsf{T}_{i,n}\boldsymbol{y}_n - \mathsf{T}_{i,n}\boldsymbol{r}_n)\|^2 \,\bigg|\, \mathfrak{X}_n\bigg)\\ &\leqslant \frac{1}{\min\limits_{1\leqslant i\leqslant m} \mathsf{p}_i}\mathsf{E}(\|\mathsf{T}_n\boldsymbol{y}_n - \mathsf{T}_n\boldsymbol{r}_n\|^2 \,|\, \mathfrak{X}_n)\\ &\leqslant \frac{1}{\min\limits_{1\leqslant i\leqslant m} \mathsf{p}_i}\mathsf{E}(\|\boldsymbol{y}_n - \boldsymbol{r}_n\|^2 \,|\, \mathfrak{X}_n)\\ &= \frac{1}{\min\limits_{1\leqslant i\leqslant m} \mathsf{p}_i}\mathsf{E}(\|\boldsymbol{b}_n\|^2 \,|\, \mathfrak{X}_n).\end{aligned} \tag{4.8}$$

SIAM J. OPTIM.
Vol. xx, pp. x x–x

Consequently (4.4), (4.7), and (ii) yield

$$\sum_{n\in\mathbb{N}}\lambda_n\sqrt{\mathsf{E}(|||\boldsymbol{e}_n|||^2\,|\,\boldsymbol{\mathfrak{X}}_n)} \leqslant \frac{1}{\displaystyle\min_{1\leqslant i\leqslant m}\sqrt{\mathsf{p}_i}}\left(\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)}+\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)}\right) < +\infty. \tag{4.9}$$

Now let $\mathbf{z}\in\mathbf{F}$, and set

$$(\forall n\in\mathbb{N})(\forall i\in\{1,\ldots,m\})\quad \mathsf{q}_{i,n}\colon \mathbf{H}\times\mathsf{D}\to\mathbb{R}\colon(\mathbf{x},\boldsymbol{\epsilon})\mapsto\|\mathsf{x}_i-\mathsf{z}_i+\epsilon_i(\mathsf{T}_{i,n}(\mathbf{R}_n\mathbf{x})-\mathsf{x}_i)\|^2. \tag{4.10}$$

Observe that, for every $n\in\mathbb{N}$ and every $i\in\{1,\ldots,m\}$, by continuity of $\mathbf{R}_n$ and $\mathsf{T}_{i,n}$, $\mathsf{T}_{i,n}\circ\mathbf{R}_n$ is measurable, and the functions $(\mathsf{q}_{i,n}(\cdot,\boldsymbol{\epsilon}))_{\boldsymbol{\epsilon}\in\mathsf{D}}$ are therefore likewise. Consequently, using (iv) and arguing as in (3.9) leads to

$$\begin{aligned}(\forall n\in\mathbb{N})(\forall i\in\{1,\ldots,m\})\quad &\mathsf{E}(\|t_{i,n}-\mathsf{z}_i\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)\\ &=\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}\mathsf{E}(\mathsf{q}_{i,n}(\boldsymbol{x}_n,\boldsymbol{\epsilon})1_{[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]}\,|\,\boldsymbol{\mathfrak{X}}_n)\\ &=\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\|x_{i,n}-\mathsf{z}_i+\epsilon_i(\mathsf{T}_{i,n}\boldsymbol{r}_n-x_{i,n})\|^2.\end{aligned} \tag{4.11}$$

Hence, recalling (3.4) and (iv), we obtain

$$\begin{aligned}(\forall n\in\mathbb{N})\quad &\mathsf{E}(|||\boldsymbol{t}_n-\mathbf{z}|||^2\,|\,\boldsymbol{\mathfrak{X}}_n)\\ &=\sum_{i=1}^m\frac{1}{\mathsf{p}_i}\mathsf{E}(\|t_{i,n}-\mathsf{z}_i\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)\\ &=\sum_{i=1}^m\frac{1}{\mathsf{p}_i}\sum_{\boldsymbol{\epsilon}\in\mathsf{D}}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\|x_{i,n}-\mathsf{z}_i+\epsilon_i(\mathsf{T}_{i,n}\boldsymbol{r}_n-x_{i,n})\|^2\\ &=\sum_{i=1}^m\frac{1}{\mathsf{p}_i}\Bigg(\sum_{\boldsymbol{\epsilon}\in\mathsf{D},\epsilon_i=1}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\,\|\mathsf{T}_{i,n}\boldsymbol{r}_n-\mathsf{z}_i\|^2\\ &\qquad+\sum_{\boldsymbol{\epsilon}\in\mathsf{D},\epsilon_i=0}\mathsf{P}[\boldsymbol{\varepsilon}_n=\boldsymbol{\epsilon}]\,\|x_{i,n}-\mathsf{z}_i\|^2\Bigg)\\ &=\|\mathbf{T}_n\boldsymbol{r}_n-\mathbf{z}\|^2+\sum_{i=1}^m\frac{1-\mathsf{p}_i}{\mathsf{p}_i}\|x_{i,n}-\mathsf{z}_i\|^2\\ &=|||\boldsymbol{x}_n-\mathbf{z}|||^2+\|\mathbf{T}_n\boldsymbol{r}_n-\mathbf{z}\|^2-\|\boldsymbol{x}_n-\mathbf{z}\|^2.\end{aligned} \tag{4.12}$$

However, we deduce from (i) and Proposition 3.7 that

$$(\forall n\in\mathbb{N})\quad \|\mathbf{T}_n\boldsymbol{r}_n-\mathbf{z}\|^2+\frac{1-\alpha_n}{\alpha_n}\|\boldsymbol{r}_n-\mathbf{T}_n\boldsymbol{r}_n-\mathbf{R}_n\mathbf{z}+\mathbf{z}\|^2\leqslant\|\boldsymbol{r}_n-\mathbf{R}_n\mathbf{z}\|^2. \tag{4.13}$$

Combining (4.12) with (4.13) yields

$$\begin{aligned}(\forall n\in\mathbb{N})\quad &\mathsf{E}(|||\boldsymbol{t}_n-\mathbf{z}|||^2\,|\,\boldsymbol{\mathfrak{X}}_n)+\frac{1-\alpha_n}{\alpha_n}\|\boldsymbol{r}_n-\mathbf{T}_n\boldsymbol{r}_n-\mathbf{R}_n\mathbf{z}+\mathbf{z}\|^2\\ &\leqslant|||\boldsymbol{x}_n-\mathbf{z}|||^2+\|\mathbf{R}_n\boldsymbol{x}_n-\mathbf{R}_n\mathbf{z}\|^2-\|\boldsymbol{x}_n-\mathbf{z}\|^2.\end{aligned} \tag{4.14}$$

Now set $\chi = \min\{1/\sup_{k\in\mathbb{N}}\alpha_k, 1/\sup_{k\in\mathbb{N}}\beta_k\} - 1$. Then $\chi \in \,]0,+\infty[$ and since, for every $n \in \mathbb{N}$, $\mathbf{R}_n$ is β_n-averaged, Proposition 3.7 and (4.14) yield

$$(\forall n \in \mathbb{N}) \quad \mathsf{E}(|||\boldsymbol{t}_n - \mathbf{z}|||^2 \,|\, \mathcal{X}_n) + \theta_n(\mathbf{z}) \leqslant |||\boldsymbol{x}_n - \mathbf{z}|||^2, \tag{4.15}$$

where

$$(\forall n \in \mathbb{N}) \quad \theta_n(\mathbf{z}) = \chi\big(\|\boldsymbol{r}_n - \mathbf{T}_n\boldsymbol{r}_n - \mathbf{R}_n\mathbf{z} + \mathbf{z}\|^2 + \|\boldsymbol{x}_n - \boldsymbol{r}_n - \mathbf{z} + \mathbf{R}_n\mathbf{z}\|^2\big) \tag{4.16}$$

$$\leqslant \frac{1-\alpha_n}{\alpha_n}\|\boldsymbol{r}_n - \mathbf{T}_n\boldsymbol{r}_n - \mathbf{R}_n\mathbf{z} + \mathbf{z}\|^2 + \frac{1-\beta_n}{\beta_n}\|\boldsymbol{x}_n - \boldsymbol{r}_n - \mathbf{z} + \mathbf{R}_n\mathbf{z}\|^2. \tag{4.17}$$

We have thus shown that properties (i)–(iii) of Theorem 2.5 hold with $(\forall n \in \mathbb{N})$ $\mu_n = \nu_n = 0$. Next, let $\mathbf{Z}$ be a countable set which is dense in $\mathbf{F}$. Then (2.15) asserts that

$$(\forall \mathbf{z} \in \mathbf{Z})(\exists\, \Omega_{\mathbf{z}} \in \mathcal{F}) \quad \mathsf{P}(\Omega_{\mathbf{z}}) = 1 \quad \text{and} \quad (\forall \omega \in \Omega_{\mathbf{z}}) \quad \sum_{n\in\mathbb{N}} \lambda_n\theta_n(\mathbf{z},\omega) < +\infty. \tag{4.18}$$

Moreover, the event $\widetilde{\Omega} = \bigcap_{\mathbf{z}\in\mathbf{Z}} \Omega_{\mathbf{z}}$ is almost certain, i.e., $\mathsf{P}(\widetilde{\Omega}) = 1$. Now fix $\mathbf{z} \in \mathbf{F}$. By density, we can extract from $\mathbf{Z}$ a sequence $(\mathbf{z}_k)_{k\in\mathbb{N}}$ such that $\mathbf{z}_k \to \mathbf{z}$. In turn, since $\inf_{n\in\mathbb{N}} \lambda_n > 0$, we derive from (4.16) and (4.18) that

$$(\forall k \in \mathbb{N})(\forall \omega \in \widetilde{\Omega}) \quad \begin{cases} \boldsymbol{r}_n(\omega) - \mathbf{T}_n\boldsymbol{r}_n(\omega) - \mathbf{R}_n\mathbf{z}_k + \mathbf{z}_k \to \mathbf{0} \\ \boldsymbol{x}_n(\omega) - \boldsymbol{r}_n(\omega) - \mathbf{z}_k + \mathbf{R}_n\mathbf{z}_k \to \mathbf{0}. \end{cases} \tag{4.19}$$

Now set $\zeta = \sup_{n\in\mathbb{N}} \sqrt{\beta_n/(1-\beta_n)}$, and $(\forall n \in \mathbb{N})$ $\mathbf{S}_n = \mathbf{Id} - \mathbf{R}_n$ and $\boldsymbol{p}_n = \boldsymbol{r}_n - \mathbf{T}_n\boldsymbol{r}_n$. Then it follows from Proposition 3.7 that the operators $(\mathbf{S}_n)_{n\in\mathbb{N}}$ are ζ-Lipschitzian. Consequently

$$(\forall k \in \mathbb{N})(\forall n \in \mathbb{N})(\forall \omega \in \widetilde{\Omega}) \quad -\zeta\|\mathbf{z}_k - \mathbf{z}\| \leqslant -\|\mathbf{S}_n\mathbf{z}_k - \mathbf{S}_n\mathbf{z}\| \\ \leqslant \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}\| - \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}_k\| \leqslant \|\mathbf{S}_n\mathbf{z}_k - \mathbf{S}_n\mathbf{z}\| \leqslant \zeta\|\mathbf{z}_k - \mathbf{z}\| \tag{4.20}$$

and, therefore, (4.19) yields

$$\begin{aligned} (\forall k \in \mathbb{N}) \quad -\zeta\|\mathbf{z}_k - \mathbf{z}\| &\leqslant \varliminf_{n\to+\infty} \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}\| - \lim_{n\to+\infty} \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}_k\| \\ &= \varliminf_{n\to+\infty} \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}\| \\ &\leqslant \varlimsup_{n\to+\infty} \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}\| \\ &\leqslant \varlimsup_{n\to+\infty} \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}\| - \lim_{n\to+\infty} \|\boldsymbol{p}_n(\omega) + \mathbf{S}_n\mathbf{z}_k\| \\ &\leqslant \zeta\|\mathbf{z}_k - \mathbf{z}\|. \end{aligned} \tag{4.21}$$

Since $\|\mathbf{z}_k - \mathbf{z}\| \to 0$ and $\mathsf{P}(\widetilde{\Omega}) = 1$, we obtain $\boldsymbol{p}_n + \mathbf{S}_n\mathbf{z} \to \mathbf{0}$ P-a.s., which proves (4.2). Likewise, set $(\forall n \in \mathbb{N})$ $\boldsymbol{q}_n = \boldsymbol{x}_n - \boldsymbol{r}_n$. Then, proceeding as in (4.21), (4.19) yields $\boldsymbol{q}_n + \mathbf{S}_n\mathbf{z} \to \mathbf{0}$, which establishes (4.3). Finally, the weak and strong convergence claims follow from (v), (vi), and Theorem 2.5. ■

SIAM J. OPTIM.
Vol. xx, pp. x x–x

Remark 4.2.

(i) Consider the special case when only one block is present ($m = 1$) and when the error sequences $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, as well as $\boldsymbol{x}_0$, are deterministic. Then the setting of Theorem 4.1 is found in [22, Theorem 6.3]. Our framework therefore makes it possible to design block-coordinate versions of the algorithms which comply with the two-layer format of [22, Theorem 6.3], such as the forward-backward algorithm [22] or the algorithms of [14] and [56]. Theorem 4.1 will be applied to block-coordinate forward-backward splitting in Section 5.2.

(ii) Theorem 4.1(v) gives a condition for the P-a.s. weak convergence of a sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ produced by algorithm 4.1 to a solution $\boldsymbol{x}$. In infinite-dimensional spaces, examples have been constructed for which the convergence is only weak and not strong, i.e., $(\|\boldsymbol{x}_n - \boldsymbol{x}\|)_{n\in\mathbb{N}}$ does not converge to 0 P-a.s. [29, 40]. Even if, as in Theorem 4.1(vi), $(\|\boldsymbol{x}_n - \boldsymbol{x}\|)_{n\in\mathbb{N}}$ does converge to 0 P-a.s., there is in general no theoretical upper bound on the worst-case behavior of the rate of convergence, which can be arbitrarily slow [9]. The latter behavior is also possible in Euclidean spaces [8, 65].

5. Applications to operator splitting.

Let $\mathsf{A}\colon \mathsf{H} \to 2^{\mathsf{H}}$ be a set-valued operator and let A^{-1} be its inverse, i.e., $(\forall(\mathsf{x},\mathsf{u}) \in \mathsf{H}^2)\ \mathsf{x} \in \mathsf{A}^{-1}\mathsf{u} \Leftrightarrow \mathsf{u} \in \mathsf{A}\mathsf{x}$. The resolvent of A is $\mathsf{J}_{\mathsf{A}} = (\mathrm{Id} + \mathsf{A})^{-1}$. The domain of A is $\operatorname{dom}\mathsf{A} = \{\mathsf{x}\in\mathsf{H} \mid \mathsf{A}\mathsf{x}\neq\varnothing\}$ and the graph of A is $\operatorname{gra}\mathsf{A} = \{(\mathsf{x},\mathsf{u}) \in \mathsf{H}\times\mathsf{H} \mid \mathsf{u}\in\mathsf{A}\mathsf{x}\}$. If A is monotone, then J_{A} is single-valued and nonexpansive and, furthermore, if A is maximally monotone, then $\operatorname{dom}\mathsf{J}_{\mathsf{A}} = \mathsf{H}$. We denote by $\Gamma_0(\mathsf{H})$ the class of lower semicontinuous convex functions $\mathsf{f}\colon \mathsf{H} \to \left]-\infty,+\infty\right]$ such that $\mathsf{f} \not\equiv +\infty$. The Moreau subdifferential of $\mathsf{f}\in\Gamma_0(\mathsf{H})$ is the maximally monotone operator

$$\partial\mathsf{f}\colon \mathsf{H}\to 2^{\mathsf{H}}\colon \mathsf{x}\mapsto \big\{\mathsf{u}\in\mathsf{H} \;\big|\; (\forall\mathsf{y}\in\mathsf{H})\ \ \langle\mathsf{y}-\mathsf{x}\mid\mathsf{u}\rangle+\mathsf{f}(\mathsf{x})\leqslant\mathsf{f}(\mathsf{y})\big\}. \tag{5.1}$$

For every $\mathsf{x}\in\mathsf{H}$, $\mathsf{f} + \|\mathsf{x} - \cdot\|^2/2$ has a unique minimizer, which is denoted by $\operatorname{prox}_{\mathsf{f}}\mathsf{x}$ [47]. We have

$$\operatorname{prox}_{\mathsf{f}} = \mathsf{J}_{\partial\mathsf{f}}. \tag{5.2}$$

For background on convex analysis and monotone operator theory, see [7]. We continue to use the standing Notation 3.1.

5.1. Random block-coordinate Douglas-Rachford splitting.

We propose a random sweeping, block-coordinate version of the Douglas-Rachford algorithm with stochastic errors. The purpose of this algorithm is to construct iteratively a zero of the sum of two maximally monotone operators and it has found applications in numerous areas; see, e.g., [7, 11, 13, 24, 27, 32, 38, 43, 51, 52, 53].

Proposition 5.1. *Set $\mathsf{D} = \{0,1\}^m \smallsetminus \{\mathbf{0}\}$ and, for every $i \in \{1,\ldots,m\}$, let $\mathsf{A}_i\colon \mathsf{H}_i \to 2^{\mathsf{H}_i}$ be maximally monotone and let $\mathsf{B}_i\colon \mathbf{H} \to 2^{\mathsf{H}_i}$. Suppose that $\mathbf{B}\colon \mathbf{H} \to 2^{\mathbf{H}}\colon \mathbf{x} \mapsto \times_{i=1}^m \mathsf{B}_i\mathbf{x}$ is maximally monotone and that the set $\mathbf{F}$ of solutions to the problem*

$$\text{find}\ \ \mathsf{x}_1\in\mathsf{H}_1,\ldots,\mathsf{x}_m\in\mathsf{H}_m\ \ \text{such that}\ \ (\forall i\in\{1,\ldots,m\})\ \ 0\in\mathsf{A}_i\mathsf{x}_i+\mathsf{B}_i(\mathsf{x}_1,\ldots,\mathsf{x}_m) \tag{5.3}$$

is nonempty. Set $\mathbf{B}^{-1}\colon \mathbf{u}\mapsto \times_{i=1}^m \mathsf{C}_i\mathbf{u}$ *where, for every* $i\in\{1,\ldots,m\}$, $\mathsf{C}_i\colon \mathsf{H}\to 2^{\mathsf{H}_i}$. *We also consider the set* $\mathbf{F}^*$ *of solutions to the dual problem*

$$\text{(5.4)}\quad \text{find } \mathsf{u}_1\in\mathsf{H}_1,\ldots,\mathsf{u}_m\in\mathsf{H}_m \text{ such that } (\forall i\in\{1,\ldots,m\})\; 0\in -\mathsf{A}_i^{-1}(-\mathsf{u}_i)+\mathsf{C}_i(\mathsf{u}_1,\ldots,\mathsf{u}_m).$$

Let $\gamma\in\,]0,+\infty[$, *let* $(\mu_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,2[$ *such that* $\inf_{n\in\mathbb{N}}\mu_n>0$ *and* $\sup_{n\in\mathbb{N}}\mu_n<2$, *let* $\boldsymbol{x}_0$, $\boldsymbol{z}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, *and* $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ *be* $\mathbf{H}$*-valued random variables, and let* $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ *be identically distributed* D*-valued random variables. Set* $\mathbf{J}_{\gamma\mathbf{B}}\colon \mathbf{x}\mapsto(\mathsf{Q}_i\mathbf{x})_{1\leqslant i\leqslant m}$ *where, for every* $i\in\{1,\ldots,m\}$, $\mathsf{Q}_i\colon\mathbf{H}\to\mathsf{H}_i$, *iterate*

$$\text{(5.5)}\quad \begin{array}{l} \text{for } n=0,1,\ldots\\ \left\lfloor \begin{array}{l} \text{for } i=1,\ldots,m\\ \left\lfloor \begin{array}{l} z_{i,n+1}=z_{i,n}+\varepsilon_{i,n}\big(\mathsf{Q}_i(x_{1,n},\ldots,x_{m,n})+b_{i,n}-z_{i,n}\big)\\ x_{i,n+1}=x_{i,n}+\varepsilon_{i,n}\mu_n\big(\mathsf{J}_{\gamma\mathsf{A}_i}(2z_{i,n+1}-x_{i,n})+a_{i,n}-z_{i,n+1}\big), \end{array}\right. \end{array}\right. \end{array}$$

and set $(\forall n\in\mathbb{N})$ $\boldsymbol{\mathcal{E}}_n=\sigma(\boldsymbol{\varepsilon}_n)$. *Assume that the following hold:*

(i) $\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)}<+\infty$ *and* $\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2\,|\,\boldsymbol{\mathfrak{X}}_n)}<+\infty$.

(ii) *For every* $n\in\mathbb{N}$, $\boldsymbol{\mathcal{E}}_n$ *and* $\boldsymbol{\mathfrak{X}}_n$ *are independent.*

(iii) $(\forall i\in\{1,\ldots,m\})$ $\mathsf{p}_i=\mathsf{P}[\varepsilon_{i,0}=1]>0$.

Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to a* $\mathbf{H}$*-valued random variable* $\boldsymbol{x}$ *such that* $\boldsymbol{z}=\mathbf{J}_{\gamma\mathbf{B}}\boldsymbol{x}$ *is an* $\mathbf{F}$*-valued random variable and* $\boldsymbol{u}=\gamma^{-1}(\boldsymbol{x}-\boldsymbol{z})$ *is an* $\mathbf{F}^*$*-valued random variable. Furthermore, suppose that:*

(iv) $\mathsf{J}_{\gamma\mathbf{B}}$ *is weakly sequentially continuous and* $\boldsymbol{b}_n\rightharpoonup\mathbf{0}$ P-a.s.

Then $\boldsymbol{z}_n\rightharpoonup\boldsymbol{z}$ P-a.s. *and* $\gamma^{-1}(\boldsymbol{x}_n-\boldsymbol{z}_n)\rightharpoonup\boldsymbol{u}$ P-a.s.

Proof. Set $\mathbf{A}\colon\mathbf{H}\to 2^{\mathbf{H}}\colon \mathbf{x}\mapsto\times_{i=1}^m\mathsf{A}_i\mathsf{x}_i$ and $(\forall i\in\{1,\ldots,m\})$ $\mathsf{T}_i=(2\mathsf{J}_{\gamma\mathsf{A}_i}-\mathrm{Id}\,)\circ(2\mathsf{Q}_i-\mathrm{Id}\,)$. Then $\mathbf{T}=(2\mathbf{J}_{\gamma\mathbf{A}}-\mathbf{Id}\,)\circ(2\mathbf{J}_{\gamma\mathbf{B}}-\mathbf{Id}\,)$ is nonexpansive as the composition of two nonexpansive operators [7, Corollary 23.10(ii)]. Furthermore $\operatorname{Fix}\mathbf{T}\neq\varnothing$ since [22, Lemma 2.6(iii)]

$$\text{(5.6)}\qquad \mathbf{J}_{\gamma\mathbf{B}}(\operatorname{Fix}\mathbf{T})=\operatorname{zer}(\mathbf{A}+\mathbf{B})=\mathbf{F}\neq\varnothing.$$

Now set

$$\text{(5.7)}\qquad (\forall n\in\mathbb{N})\quad \begin{cases}\lambda_n=\mu_n/2\\ \boldsymbol{e}_n=2\big(\mathbf{J}_{\gamma\mathbf{A}}(2\mathbf{J}_{\gamma\mathbf{B}}\boldsymbol{x}_n+2\boldsymbol{b}_n-\boldsymbol{x}_n)-\mathbf{J}_{\gamma\mathbf{A}}(2\mathbf{J}_{\gamma\mathbf{B}}\boldsymbol{x}_n-\boldsymbol{x}_n)+\boldsymbol{a}_n-\boldsymbol{b}_n\big).\end{cases}$$

Then we derive from (5.5) that

$$\begin{aligned}
&(\forall n\in\mathbb{N})(\forall i\in\{1,\ldots,m\})\\
&\quad x_{i,n+1}=x_{i,n}+\varepsilon_{i,n}\mu_n\big(\mathsf{J}_{\gamma\mathsf{A}_i}\big(2\mathsf{Q}_i\boldsymbol{x}_n+2b_{i,n}-x_{i,n}\big)+a_{i,n}-z_{i,n+1}\big)\\
&\qquad\quad =x_{i,n}+\varepsilon_{i,n}\lambda_n\big(2\mathsf{J}_{\gamma\mathsf{A}_i}\big(2\mathsf{Q}_i\boldsymbol{x}_n-x_{i,n}\big)+e_{i,n}-2\mathsf{Q}_i\boldsymbol{x}_n\big)\\
\text{(5.8)}&\qquad\quad =x_{i,n}+\varepsilon_{i,n}\lambda_n\big(\mathsf{T}_i\boldsymbol{x}_n+e_{i,n}-x_{i,n}\big),
\end{aligned}$$

which is precisely the iteration process (3.16). Furthermore, we infer from (5.7) and the nonexpansiveness of $\mathsf{J}_{\gamma\mathsf{A}}$ [7, Corollary 23.10(i)] that

$$
\begin{aligned}
(\forall n \in \mathbb{N}) \quad \|\boldsymbol{e}_n\|^2 &\leqslant 4\|\mathsf{J}_{\gamma\mathsf{A}}(2\mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x}_n + 2\boldsymbol{b}_n - \boldsymbol{x}_n) - \mathsf{J}_{\gamma\mathsf{A}}(2\mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x}_n - \boldsymbol{x}_n) + \boldsymbol{a}_n - \boldsymbol{b}_n\|^2 \\
&\leqslant 12\big(\|\mathsf{J}_{\gamma\mathsf{A}}(2\mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x}_n + 2\boldsymbol{b}_n - \boldsymbol{x}_n) - \mathsf{J}_{\gamma\mathsf{A}}(2\mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x}_n - \boldsymbol{x}_n)\|^2 \\
&\quad + \|\boldsymbol{a}_n\|^2 + \|\boldsymbol{b}_n\|^2\big) \\
&\leqslant 12\big(\|\boldsymbol{a}_n\|^2 + 5\|\boldsymbol{b}_n\|^2\big) \qquad (5.9)
\end{aligned}
$$

and therefore that

$$
(\forall n \in \mathbb{N}) \quad \sqrt{\mathsf{E}(\|\boldsymbol{e}_n\|^2 \,|\, \mathcal{X}_n)} \leqslant 2\sqrt{3}\Big(\sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2 \,|\, \mathcal{X}_n)} + \sqrt{5}\sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2 \,|\, \mathcal{X}_n)}\Big). \tag{5.10}
$$

Thus, we deduce from (i) that $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{e}_n\|^2 \,|\, \mathcal{X}_n)} < +\infty$. Altogether, the almost sure weak convergence of $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ to a $(\operatorname{Fix}\mathsf{T})$-valued random variable $\boldsymbol{x}$ follows from Corollary 3.4. In turn, (5.6) asserts that $\boldsymbol{z} = \mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x} \in \mathsf{F}$ P-a.s. Now set $\boldsymbol{u} = \gamma^{-1}(\boldsymbol{x} - \boldsymbol{z})$. Then, P-a.s.,

$$
\boldsymbol{z} = \mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x} \;\Leftrightarrow\; \boldsymbol{x} - \boldsymbol{z} \in \gamma\mathsf{B}\boldsymbol{z} \;\Leftrightarrow\; \boldsymbol{z} \in \mathsf{B}^{-1}\boldsymbol{u} \tag{5.11}
$$

and

$$
\begin{aligned}
\boldsymbol{x} \in \operatorname{Fix}\mathsf{T} &\Leftrightarrow \boldsymbol{x} = (2\mathsf{J}_{\gamma\mathsf{A}} - \mathbf{Id}\,)(2\boldsymbol{z} - \boldsymbol{x}) \\
&\Leftrightarrow \boldsymbol{z} = \mathsf{J}_{\gamma\mathsf{A}}(2\boldsymbol{z} - \boldsymbol{x}) \\
&\Leftrightarrow \boldsymbol{z} - \boldsymbol{x} \in \gamma\mathsf{A}\boldsymbol{z} \\
&\Leftrightarrow -\boldsymbol{z} \in -\mathsf{A}^{-1}(-\boldsymbol{u}). \qquad (5.12)
\end{aligned}
$$

These imply that $\mathbf{0} \in -\mathsf{A}^{-1}(-\boldsymbol{u}) + \mathsf{B}^{-1}\boldsymbol{u}$ P-a.s., i.e., that $\boldsymbol{u} \in \mathsf{F}^*$ P-a.s. Finally, assume that (iv) holds. Then there exists $\widetilde{\Omega} \in \mathcal{F}$ such that $\mathsf{P}(\widetilde{\Omega}) = 1$ and $(\forall \omega \in \widetilde{\Omega})\ \mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x}_n(\omega) \rightharpoonup \mathsf{J}_{\gamma\mathsf{B}}\boldsymbol{x}(\omega) = \boldsymbol{z}(\omega)$. Now let $i \in \{1, \ldots, m\}$, $\omega \in \widetilde{\Omega}$, and $\mathsf{v} \in \mathsf{H}$. Then $\langle \mathsf{Q}_i\boldsymbol{x}_n(\omega) \mid \mathsf{v}_i\rangle \to \langle z_i(\omega) \mid \mathsf{v}_i\rangle$ and (5.5) yields

$$
\begin{aligned}
(\forall n \in \mathbb{N}) \quad \langle z_{i,n+1}(\omega) \mid \mathsf{v}_i\rangle &= \langle z_{i,n}(\omega) \mid \mathsf{v}_i\rangle \\
&\quad + \varepsilon_{i,n}(\omega)\big(\langle \mathsf{Q}_i\boldsymbol{x}_n(\omega) \mid \mathsf{v}_i\rangle + \langle b_{i,n}(\omega) \mid \mathsf{v}_i\rangle - \langle z_{i,n}(\omega) \mid \mathsf{v}_i\rangle\big).
\end{aligned} \tag{5.13}
$$

However, according to (iii), at the expense of possibly taking ω in a smaller almost sure event, $\varepsilon_{i,n}(\omega) = 1$ infinitely often. Hence, there exists a monotone sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $k_n \to +\infty$ and, for $n \in \mathbb{N}$ sufficiently large,

$$
\langle z_{i,n+1}(\omega) \mid \mathsf{v}_i\rangle = \langle \mathsf{Q}_i\boldsymbol{x}_{k_n}(\omega) \mid \mathsf{v}_i\rangle + \langle b_{i,k_n}(\omega) \mid \mathsf{v}_i\rangle. \tag{5.14}
$$

Thus, since $\langle \mathsf{Q}_i\boldsymbol{x}_{k_n}(\omega) \mid \mathsf{v}_i\rangle \to \langle z_i(\omega) \mid \mathsf{v}_i\rangle$ and $\langle b_{i,k_n}(\omega) \mid \mathsf{v}_i\rangle \to 0$, we have $\langle z_{i,n+1}(\omega) - z_i(\omega) \mid \mathsf{v}_i\rangle \to 0$. Hence,

$$
\langle \boldsymbol{z}_{n+1}(\omega) - \boldsymbol{z}(\omega) \mid \mathbf{v}\rangle = \sum_{i=1}^{m} \langle z_{i,n+1}(\omega) - z_i(\omega) \mid \mathsf{v}_i\rangle \to 0. \tag{5.15}
$$

This shows that $\boldsymbol{z}_n \rightharpoonup \boldsymbol{z}$ P-a.s., which allows us to conclude that $\gamma^{-1}(\boldsymbol{x}_n - \boldsymbol{z}_n) \rightharpoonup \boldsymbol{u}$ P-a.s. ■

Remark 5.2. Let us make some connections between Proposition 5.1 and existing results.

(i) In the standard case of a single block ($m = 1$) and when all the variables are deterministic, the above primal convergence result goes back to [32] and to [43] in the unrelaxed case.

(ii) In minimization problems, the alternating direction method of multipliers (ADMM) is strongly related to an application of the Douglas-Rachford algorithm to the dual problem [38]. This connection can be used to construct a random block-coordinate ADMM algorithm. Let us note that such an algorithm was recently proposed in [41] in a finite-dimensional setting, where single-block, unrelaxed, and error-free iterations were used.

Next, we apply Proposition 5.1 to devise a primal-dual block-coordinate algorithm for solving a class of structured inclusion problems investigated in [25].

Corollary 5.3. *Set* $\mathsf{D} = \{0,1\}^{m+p} \smallsetminus \{\mathbf{0}\}$*, let* $(\mathsf{G}_k)_{1\leqslant k\leqslant p}$ *be separable real Hilbert spaces, and set* $\mathbf{G} = \mathsf{G}_1 \oplus \cdots \oplus \mathsf{G}_p$*. For every* $i \in \{1,\dots,m\}$*, let* $\mathsf{A}_i \colon \mathsf{H}_i \to 2^{\mathsf{H}_i}$ *be maximally monotone and, for every* $k \in \{1,\dots,p\}$*, let* $\mathsf{B}_k \colon \mathsf{G}_k \to 2^{\mathsf{G}_k}$ *be maximally monotone, and let* $\mathsf{L}_{ki} \colon \mathsf{H}_i \to \mathsf{G}_k$ *be linear and bounded. It is assumed that the set* $\mathbf{F}$ *of solutions to the problem*

$$\text{find } \mathsf{x}_1 \in \mathsf{H}_1,\dots,\mathsf{x}_m \in \mathsf{H}_m \text{ such that } (\forall i \in \{1,\dots,m\})\; 0 \in \mathsf{A}_i\mathsf{x}_i + \sum_{k=1}^{p} \mathsf{L}_{ki}^*\mathsf{B}_k\Bigg(\sum_{j=1}^{m} \mathsf{L}_{kj}\mathsf{x}_j\Bigg) \tag{5.16}$$

is nonempty. We also consider the set $\mathbf{F}^*$ *of solutions to the dual problem*

$$\text{find } \mathsf{v}_1 \in \mathsf{G}_1,\dots,\mathsf{v}_p \in \mathsf{G}_p \text{ such that } (\forall k \in \{1,\dots,p\})\; 0 \in -\sum_{i=1}^{m} \mathsf{L}_{ki}\mathsf{A}_i^{-1}\Bigg(-\sum_{l=1}^{p} \mathsf{L}_{li}^*\mathsf{v}_l\Bigg) + \mathsf{B}_k^{-1}\mathsf{v}_k. \tag{5.17}$$

Let $\gamma \in \left]0,+\infty\right[$*, let* $(\mu_n)_{n\in\mathbb{N}}$ *be a sequence in* $\left]0,2\right[$ *such that* $\inf_{n\in\mathbb{N}} \mu_n > 0$ *and* $\sup_{n\in\mathbb{N}} \mu_n < 2$*, let* $\boldsymbol{x}_0$*,* $\boldsymbol{z}_0$*,* $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$*, and* $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ *be* $\mathbf{H}$*-valued random variables, let* $\boldsymbol{y}_0$*,* $\boldsymbol{w}_0$*,* $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$*, and* $(\boldsymbol{d}_n)_{n\in\mathbb{N}}$ *be* $\mathbf{G}$*-valued random variables, and let* $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ *be identically distributed* D*-valued random variables. Set*

$$\mathbf{V} = \Bigg\{(\mathsf{x}_1,\dots,\mathsf{x}_m,\mathsf{y}_1,\dots,\mathsf{y}_p) \in \mathbf{H}\oplus\mathbf{G} \;\Bigg|\; (\forall k \in \{1,\dots,p\})\; \mathsf{y}_k = \sum_{i=1}^{m} \mathsf{L}_{ki}\mathsf{x}_i\Bigg\}, \tag{5.18}$$

let $\mathbf{P}_{\mathbf{V}} \colon \mathbf{x} \mapsto (\mathsf{Q}_j\mathbf{x})_{1\leqslant j\leqslant m+p}$ *be its projection operator, where* $(\forall i \in \{1,\dots,m\})$ $\mathsf{Q}_i \colon \mathbf{H}\oplus\mathbf{G} \to \mathsf{H}_i$ *and* $(\forall k \in \{1,\dots,p\})$ $\mathsf{Q}_{m+k} \colon \mathbf{H}\oplus\mathbf{G} \to \mathsf{G}_k$*, iterate*

$$\begin{array}{l}
\text{for } n = 0,1,\dots\\
\left\lfloor\begin{array}{l}
\text{for } i = 1,\dots,m\\
\left\lfloor\begin{array}{l}
z_{i,n+1} = z_{i,n} + \varepsilon_{i,n}\big(\mathsf{Q}_i(x_{1,n},\dots,x_{m,n},y_{1,n},\dots,y_{p,n}) + c_{i,n} - z_{i,n}\big)\\
x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\mu_n\big(\mathsf{J}_{\gamma\mathsf{A}_i}(2z_{i,n+1} - x_{i,n}) + a_{i,n} - z_{i,n+1}\big)
\end{array}\right.\\
\text{for } k = 1,\dots,p\\
\left\lfloor\begin{array}{l}
w_{k,n+1} = w_{k,n} + \varepsilon_{m+k,n}\big(\mathsf{Q}_{m+k}(x_{1,n},\dots,x_{m,n},y_{1,n},\dots,y_{p,n}) + d_{k,n} - w_{k,n}\big)\\
y_{k,n+1} = y_{k,n} + \varepsilon_{m+k,n}\mu_n\big(\mathsf{J}_{\gamma\mathsf{B}_k}(2w_{k,n+1} - y_{k,n}) + b_{k,n} - w_{k,n+1}\big),
\end{array}\right.
\end{array}\right.
\end{array} \tag{5.19}$$

and set $(\forall n\in\mathbb{N})$ $\boldsymbol{\mathcal{Y}}_n=\sigma(\boldsymbol{x}_j,\boldsymbol{y}_j)_{0\leqslant j\leqslant n}$ *and* $\boldsymbol{\mathcal{E}}_n=\sigma(\boldsymbol{\varepsilon}_n)$*. In addition, assume that the following hold:*

(i) $\boldsymbol{c}_n\rightharpoonup \mathbf{0}$ P-a.s., $\boldsymbol{d}_n\rightharpoonup \mathbf{0}$ P-a.s., $\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2\,|\,\boldsymbol{\mathcal{Y}}_n)}<+\infty$, $\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2\,|\,\boldsymbol{\mathcal{Y}}_n)}<+\infty$, $\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{c}_n\|^2\,|\,\boldsymbol{\mathcal{Y}}_n)}<+\infty$, *and* $\sum_{n\in\mathbb{N}}\sqrt{\mathsf{E}(\|\boldsymbol{d}_n\|^2\,|\,\boldsymbol{\mathcal{Y}}_n)}<+\infty$.

(ii) *For every* $n\in\mathbb{N}$, $\boldsymbol{\mathcal{E}}_n$ *and* $\boldsymbol{\mathcal{Y}}_n$ *are independent.*

(iii) $(\forall j\in\{1,\ldots,m+p\})$ $\mathsf{P}[\varepsilon_{j,0}=1]>0$.

Then $(\boldsymbol{z}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* $\mathbf{F}$*-valued random variable, and* $(\gamma^{-1}(\boldsymbol{w}_n-\boldsymbol{y}_n))_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* $\mathbf{F}^*$*-valued random variable.*

Proof. Set $\mathbf{A}\colon\mathbf{H}\to 2^{\mathbf{H}}\colon\mathbf{x}\mapsto \times_{i=1}^m\mathsf{A}_i\mathsf{x}_i$, $\mathbf{B}\colon\mathbf{G}\to 2^{\mathbf{G}}\colon\mathbf{y}\mapsto \times_{k=1}^p\mathsf{B}_k\mathsf{y}_k$, and $\mathbf{L}\colon\mathbf{H}\to\mathbf{G}\colon\mathbf{x}\mapsto \big(\sum_{i=1}^m\mathsf{L}_{ki}\mathsf{x}_i\big)_{1\leqslant k\leqslant p}$. Furthermore, let us introduce

$$\mathbf{K}=\mathbf{H}\oplus\mathbf{G}\quad\text{and}\quad \mathbf{C}\colon\mathbf{K}\to 2^{\mathbf{K}}\colon(\mathbf{x},\mathbf{y})\mapsto\mathbf{A}\mathbf{x}\times\mathbf{B}\mathbf{y}. \tag{5.20}$$

Then the primal-dual problem (5.16)–(5.17) can be rewritten as

$$\text{find}\ \ (\mathbf{x},\mathbf{v})\in\mathbf{K}\ \ \text{such that}\ \begin{cases}\mathbf{0}\in\mathbf{A}\mathbf{x}+\mathbf{L}^*\mathbf{B}\mathbf{L}\mathbf{x}\\ \mathbf{0}\in-\mathbf{L}\mathbf{A}^{-1}(-\mathbf{L}^*\mathbf{v})+\mathbf{B}^{-1}\mathbf{v}.\end{cases} \tag{5.21}$$

The normal cone operator to $\mathbf{V}$ is [7, Example 6.42]

$$\mathbf{N}_{\mathbf{V}}\colon\mathbf{K}\to 2^{\mathbf{K}}\colon(\mathbf{x},\mathbf{y})\mapsto\begin{cases}\mathbf{V}^\perp, & \text{if}\ \ \mathbf{L}\mathbf{x}=\mathbf{y};\\ \varnothing, & \text{if}\ \ \mathbf{L}\mathbf{x}\neq\mathbf{y},\end{cases}\quad\text{where}\quad \mathbf{V}^\perp=\big\{(\mathbf{u},\mathbf{v})\in\mathbf{K}\;\big|\;\mathbf{u}=-\mathbf{L}^*\mathbf{v}\big\}. \tag{5.22}$$

Now let $(\mathbf{x},\mathbf{y})\in\mathbf{K}$. Then

$$\begin{aligned}(\mathbf{0},\mathbf{0})\in\mathbf{C}(\mathbf{x},\mathbf{y})+\mathbf{N}_{\mathbf{V}}(\mathbf{x},\mathbf{y}) &\Leftrightarrow \begin{cases}(\mathbf{x},\mathbf{y})\in\ \mathbf{V}\\ (\mathbf{0},\mathbf{0})\in(\mathbf{A}\mathbf{x}\times\mathbf{B}\mathbf{y})+\mathbf{V}^\perp\end{cases}\\ &\Leftrightarrow \begin{cases}\mathbf{L}\mathbf{x}=\mathbf{y}\\ (\exists\,\mathbf{u}\in\mathbf{A}\mathbf{x})(\exists\,\mathbf{v}\in\mathbf{B}\mathbf{y})\ \ \mathbf{u}=-\mathbf{L}^*\mathbf{v}\end{cases}\\ &\Rightarrow (\exists\,\mathbf{v}\in\mathbf{B}(\mathbf{L}\mathbf{x}))\ \ -\mathbf{L}^*\mathbf{v}\in\mathbf{A}\mathbf{x}\\ &\Rightarrow (\exists\,\mathbf{v}\in\mathbf{G})\ \ \mathbf{L}^*\mathbf{v}\in\mathbf{L}^*\mathbf{B}\mathbf{L}\mathbf{x}\ \ \text{and}\ \ -\mathbf{L}^*\mathbf{v}\in\mathbf{A}\mathbf{x}\\ &\Leftrightarrow \mathbf{x}\ \text{solves (5.16)}.\end{aligned} \tag{5.23}$$

Since $\mathbf{C}$ and $\mathbf{N}_{\mathbf{V}}$ are maximally monotone, it follows from [7, Proposition 23.16] that the iteration process (5.19) is an instance of (5.5) for finding a zero of $\mathbf{C}+\mathbf{N}_{\mathbf{V}}$ in $\mathbf{K}$. The associated dual problem consists of finding a zero of $-\mathbf{C}^{-1}(-\cdot)+\mathbf{N}_{\mathbf{V}}^{-1}$. Let $(\mathbf{u},\mathbf{v}) \in \mathbf{K}$. Then (5.22) yields

$$
\begin{aligned}
(\mathbf{0},\mathbf{0}) \in -\mathbf{C}^{-1}(-\mathbf{u},-\mathbf{v}) + \mathbf{N}_{\mathbf{V}}^{-1}(\mathbf{u},\mathbf{v})\\
&\Leftrightarrow (\mathbf{0},\mathbf{0}) \in -\mathbf{C}^{-1}(-\mathbf{u},-\mathbf{v}) + \mathbf{N}_{\mathbf{V}^\perp}(\mathbf{u},\mathbf{v})\\
&\Leftrightarrow \begin{cases}(\mathbf{u},\mathbf{v}) \in \ \mathbf{V}^\perp\\ (\mathbf{0},\mathbf{0}) \in \big(-\mathbf{A}^{-1}(-\mathbf{u}) \times -\mathbf{B}^{-1}(-\mathbf{v})\big) + \mathbf{V}\end{cases}\\
&\Leftrightarrow \begin{cases}\mathbf{u} = -\mathbf{L}^*\mathbf{v}\\ (\exists\, \mathbf{x} \in -\mathbf{A}^{-1}(-\mathbf{u}))(\exists\, \mathbf{y} \in -\mathbf{B}^{-1}(-\mathbf{v}))\ \mathbf{L}\mathbf{x} = \mathbf{y}\end{cases}\\
&\Rightarrow (\exists\, \mathbf{x} \in -\mathbf{A}^{-1}(\mathbf{L}^*\mathbf{v}))\ \mathbf{L}\mathbf{x} \in -\mathbf{B}^{-1}(-\mathbf{v})\\
&\Rightarrow (\exists\, \mathbf{x} \in \mathbf{H})\ \ \mathbf{L}\mathbf{x} \in -\mathbf{L}\mathbf{A}^{-1}(\mathbf{L}^*\mathbf{v})) \text{ and } -\mathbf{L}\mathbf{x} \in \mathbf{B}^{-1}(-\mathbf{v})\\
&\Leftrightarrow -\mathbf{v} \text{ solves (5.17).}
\end{aligned}
\tag{5.24}
$$

The convergence result therefore follows from Proposition 5.1 using (5.23), (5.24), and the weak continuity of $\mathbf{P}_{\mathbf{V}} = \mathbf{J}_{\gamma \mathbf{N}_{\mathbf{V}}}$ [7, Proposition 28.11(i)]. ■

Remark 5.4.The parameterization (5.20) made it possible to reduce the structured primal-dual problem (5.16)–(5.17) to a basic two-operator inclusion, to which the block-coordinate Douglas-Rachford algorithm (5.5) could be applied. A similar parameterization was used in [1] in a different context. We also note that, at each iteration of algorithm (5.19), components of the projector $\mathbf{P}_{\mathbf{V}}$ need to be activated. This operator is expressed as

$$
\big(\forall(\mathbf{x},\mathbf{y}) \in \mathbf{H}\oplus\mathbf{G}\big) \quad \mathbf{P}_{\mathbf{V}}\colon (\mathbf{x},\mathbf{y}) \mapsto (\mathbf{t},\mathbf{L}\mathbf{t}) = (\mathbf{x}-\mathbf{L}^*\mathbf{s}, \mathbf{y}+\mathbf{s})
\tag{5.25}
$$

where $\mathbf{t} = (\mathbf{Id}+\mathbf{L}^*\mathbf{L})^{-1}(\mathbf{x}+\mathbf{L}^*\mathbf{y})$ and $\mathbf{s} = (\mathbf{Id}+\mathbf{L}\mathbf{L}^*)^{-1}(\mathbf{L}\mathbf{x}-\mathbf{y})$ [1, Lemma 3.1]. This formula allows us to compute the components of $\mathbf{P}_{\mathbf{V}}$, which is especially simple when $\mathbf{Id}+\mathbf{L}^*\mathbf{L}$ or $\mathbf{Id}+\mathbf{L}\mathbf{L}^*$ is easily inverted.

The previous result leads to a random block-coordinate primal-dual proximal algorithm for solving a wide range of structured convex optimization problems.

Corollary 5.5.*Set* $\mathsf{D} = \{0,1\}^{m+p} \smallsetminus \{\mathbf{0}\}$, *let* $(\mathsf{G}_k)_{1\leqslant k\leqslant p}$ *be separable real Hilbert spaces, and set* $\mathbf{G} = \mathsf{G}_1\oplus\cdots\oplus\mathsf{G}_p$. *For every* $i \in \{1,\ldots,m\}$, *let* $\mathsf{f}_i \in \Gamma_0(\mathsf{H}_i)$ *and, for every* $k \in \{1,\ldots,p\}$, *let* $\mathsf{g}_k \in \Gamma_0(\mathsf{G}_k)$, *and let* $\mathsf{L}_{ki}\colon \mathsf{H}_i \to \mathsf{G}_k$ *be linear and bounded. It is assumed that there exists* $(\mathsf{x}_1,\ldots,\mathsf{x}_m) \in \mathbf{H}$ *such that*

$$
(\forall i \in \{1,\ldots,m\}) \quad 0 \in \partial\mathsf{f}_i(\mathsf{x}_i) + \sum_{k=1}^{p} \mathsf{L}_{ki}^*\partial\mathsf{g}_k\Bigg(\sum_{j=1}^{m}\mathsf{L}_{kj}\mathsf{x}_j\Bigg).
\tag{5.26}
$$

Let $\mathbf{F}$ *be the set of solutions to the problem*

$$
\underset{\mathsf{x}_1\in\mathsf{H}_1,\ldots,\mathsf{x}_m\in\mathsf{H}_m}{\text{minimize}} \ \ \sum_{i=1}^{m}\mathsf{f}_i(\mathsf{x}_i) + \sum_{k=1}^{p}\mathsf{g}_k\Bigg(\sum_{i=1}^{m}\mathsf{L}_{ki}\mathsf{x}_i\Bigg)
\tag{5.27}
$$

and let $\mathbf{F}^*$ *be the set of solutions to the dual problem*

$$
\underset{\mathsf{v}_1\in\mathsf{G}_1,\ldots,\mathsf{v}_p\in\mathsf{G}_p}{\text{minimize}} \ \ \sum_{i=1}^{m}\mathsf{f}_i^*\Bigg(-\sum_{k=1}^{p}\mathsf{L}_{ki}^*\mathsf{v}_k\Bigg) + \sum_{k=1}^{p}\mathsf{g}_k^*(\mathsf{v}_k).
\tag{5.28}
$$

Let $\gamma \in \,]0,+\infty[$, let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $]0,2[$ such that $\inf_{n\in\mathbb{N}}\mu_n > 0$ and $\sup_{n\in\mathbb{N}}\mu_n < 2$, let $\boldsymbol{x}_0$, $\boldsymbol{z}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ be $\mathbf{H}$-valued random variables, let $\boldsymbol{y}_0$, $\boldsymbol{w}_0$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{d}_n)_{n\in\mathbb{N}}$ be $\mathbf{G}$-valued random variables, and let $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ be identically distributed D-valued random variables. Define $\mathbf{V}$ as in (5.18) and set $\mathbf{P_V}\colon \mathbf{x}\mapsto(\mathsf{Q}_j\mathbf{x})_{1\leqslant j\leqslant m+p}$ where $(\forall i\in\{1,\ldots,m\})$ $\mathsf{Q}_i\colon \mathbf{H}\oplus\mathbf{G}\to\mathsf{H}_i$ and $(\forall k\in\{1,\ldots,p\})$ $\mathsf{Q}_{m+k}\colon \mathbf{H}\oplus\mathbf{G}\to\mathsf{G}_k$, and iterate

$$
\begin{array}{l}
\text{for } n=0,1,\ldots\\
\left\lfloor
\begin{array}{l}
\text{for } i=1,\ldots,m\\
\left\lfloor
\begin{array}{l}
z_{i,n+1}=z_{i,n}+\varepsilon_{i,n}\big(\mathsf{Q}_i(x_{1,n},\ldots,x_{m,n},y_{1,n},\ldots,y_{p,n})+c_{i,n}-z_{i,n}\big)\\
x_{i,n+1}=x_{i,n}+\varepsilon_{i,n}\mu_n\big(\mathrm{prox}_{\gamma\mathsf{f}_i}(2z_{i,n+1}-x_{i,n})+a_{i,n}-z_{i,n+1}\big)
\end{array}
\right.\\
\text{for } k=1,\ldots,p\\
\left\lfloor
\begin{array}{l}
w_{k,n+1}=w_{k,n}+\varepsilon_{m+k,n}\big(\mathsf{Q}_{m+k}(x_{1,n},\ldots,x_{m,n},y_{1,n},\ldots,y_{p,n})+d_{k,n}-w_{k,n}\big)\\
y_{k,n+1}=y_{k,n}+\varepsilon_{m+k,n}\mu_n\big(\mathrm{prox}_{\gamma\mathsf{g}_k}(2w_{k,n+1}-y_{k,n})+b_{k,n}-w_{k,n+1}\big).
\end{array}
\right.
\end{array}
\right.
\end{array}
\tag{5.29}
$$

In addition, assume that conditions (i)–(iii) of Corollary 5.3 are satisfied. Then $(\boldsymbol{z}_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. *to an $\mathbf{F}$-valued random variable, and $(\gamma^{-1}(\boldsymbol{w}_n-\boldsymbol{y}_n))_{n\in\mathbb{N}}$ converges weakly* P-a.s. *to an $\mathbf{F}^*$-valued random variable.*

Proof. Using the same arguments as in [25, Proposition 5.4] one sees that this is an application of Corollary 5.3 with, for every $i\in\{1,\ldots,m\}$, $\mathsf{A}_i=\partial\mathsf{f}_i$ and, for every $k\in\{1,\ldots,p\}$, $\mathsf{B}_k=\partial\mathsf{g}_k$. ∎

Remark 5.6. Sufficient conditions for (5.26) to hold are provided in [25, Proposition 5.3].

5.2. Random block-coordinate forward-backward splitting. The forward-backward algorithm addresses the problem of finding a zero of the sum of two maximally monotone operators, one of which has a strongly monotone inverse (see [2, 22] for historical background). It has been applied to a wide variety of problems among which are mechanics, partial differential equations, best approximation, evolution inclusions, signal and image processing, convex optimization, learning theory, inverse problems, statistics, and game theory [2, 7, 16, 17, 22, 27, 29, 31, 39, 46, 61, 62, 63]. In this section we design a block-coordinate version of this algorithm with random sweeping and stochastic errors.

Definition 5.7.[2, Definition 2.3] *An operator $\mathsf{A}\colon\mathsf{H}\to 2^{\mathsf{H}}$ is* demiregular *at $\mathsf{x}\in\operatorname{dom}\mathsf{A}$ if, for every sequence $((\mathsf{x}_n,\mathsf{u}_n))_{n\in\mathbb{N}}$ in $\operatorname{gra}\mathsf{A}$ and every $\mathsf{u}\in\mathsf{A}\mathsf{x}$ such that $\mathsf{x}_n\rightharpoonup\mathsf{x}$ and $\mathsf{u}_n\to\mathsf{u}$, we have $\mathsf{x}_n\to\mathsf{x}$.*

Lemma 5.8.[2, Proposition 2.4] *Let $\mathsf{A}\colon\mathsf{H}\to 2^{\mathsf{H}}$ be monotone and suppose that $\mathsf{x}\in\operatorname{dom}\mathsf{A}$. Then A is demiregular at x in each of the following cases:*

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

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

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

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

(v) *A is single-valued on $\operatorname{dom}\mathsf{A}$ and $\mathrm{Id}-\mathsf{A}$ is demicompact.*

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

$$(5.30)\quad (\forall \alpha \in\,]0,1[)(\forall \mathsf{y} \in \operatorname{dom} \mathsf{f}) \quad \mathsf{f}\big(\alpha\mathsf{x} + (1-\alpha)\mathsf{y}\big) + \alpha(1-\alpha)\theta(\|\mathsf{x}-\mathsf{y}\|) \leqslant \alpha\mathsf{f}(\mathsf{x}) + (1-\alpha)\mathsf{f}(\mathsf{y}).$$

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

Our block-coordinate forward-backward algorithm is the following.

Proposition 5.9. *Set* $\mathsf{D} = \{0,1\}^m \smallsetminus \{\mathbf{0}\}$ *and, for every* $i \in \{1,\ldots,m\}$, *let* $\mathsf{A}_i\colon \mathsf{H}_i \to 2^{\mathsf{H}_i}$ *be maximally monotone and let* $\mathsf{B}_i\colon \mathbf{H} \to \mathsf{H}_i$. *Suppose that*

$$(5.31)\qquad (\exists\, \vartheta \in\,]0,+\infty[)(\forall \mathbf{x} \in \mathbf{H})(\forall \mathbf{y} \in \mathbf{H}) \quad \sum_{i=1}^m \langle \mathsf{x}_i - \mathsf{y}_i \mid \mathsf{B}_i\mathbf{x} - \mathsf{B}_i\mathbf{y}\rangle \geqslant \vartheta \sum_{i=1}^m \big\|\mathsf{B}_i\mathbf{x} - \mathsf{B}_i\mathbf{y}\big\|^2$$

and that the set $\mathbf{F}$ *of solutions to the problem*

$$(5.32)\quad \text{find } \mathsf{x}_1 \in \mathsf{H}_1,\ldots,\mathsf{x}_m \in \mathsf{H}_m \text{ such that } (\forall i \in \{1,\ldots,m\}) \quad 0 \in \mathsf{A}_i\mathsf{x}_i + \mathsf{B}_i(\mathsf{x}_1,\ldots,\mathsf{x}_m)$$

is nonempty. Let $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,2\vartheta[$ *such that* $\inf_{n\in\mathbb{N}}\gamma_n > 0$ *and* $\sup_{n\in\mathbb{N}}\gamma_n < 2\vartheta$, *and let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,1]$ *such that* $\inf_{n\in\mathbb{N}}\lambda_n > 0$. *Let* $\boldsymbol{x}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, *and* $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ *be* $\mathbf{H}$*-valued random variables, and let* $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ *be identically distributed* D*-valued random variables. Iterate*

(5.33)

$$\begin{array}{l}\text{for } n = 0,1,\ldots\\ \left\lfloor \begin{array}{l}\text{for } i = 1,\ldots,m\\ \left\lfloor\; x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\lambda_n\big(\mathsf{J}_{\gamma_n\mathsf{A}_i}\big(x_{i,n} - \gamma_n(\mathsf{B}_i(x_{1,n},\ldots,x_{m,n}) + c_{i,n})\big) + a_{i,n} - x_{i,n}\big),\right.\end{array}\right.\end{array}$$

and set $(\forall n \in \mathbb{N})$ $\boldsymbol{\mathcal{E}}_n = \sigma(\boldsymbol{\varepsilon}_n)$. *Furthermore, assume that the following hold:*

(i) $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2 \,|\, \boldsymbol{\mathcal{X}}_n)} < +\infty$ *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{c}_n\|^2 \,|\, \boldsymbol{\mathcal{X}}_n)} < +\infty$.

(ii) *For every* $n \in \mathbb{N}$, $\boldsymbol{\mathcal{E}}_n$ *and* $\boldsymbol{\mathcal{X}}_n$ *are independent.*

(iii) $(\forall i \in \{1,\ldots,m\})$ $\mathsf{P}[\varepsilon_{i,0} = 1] > 0$.

Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* $\mathbf{F}$*-valued random variable* $\boldsymbol{x}$. *If, in addition, one of the following holds, then* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to* $\boldsymbol{x}$*:*

(iv) *For every* $\mathbf{x} \in \mathbf{F}$ *and every* $i \in \{1,\ldots,m\}$, A_i *is demiregular at* x_i.

(v) *The operator* $\mathbf{x} \mapsto (\mathsf{B}_i\mathbf{x})_{1\leqslant i\leqslant m}$ *is demiregular at every point in* $\mathbf{F}$.

Proof. We are going to apply Theorem 4.1. Set $\mathbf{A}\colon \mathbf{H} \to 2^{\mathbf{H}}\colon \mathbf{x} \mapsto \times_{i=1}^m \mathsf{A}_i\mathsf{x}_i$, $\mathbf{B}\colon \mathbf{H} \to \mathbf{H}\colon \mathbf{x} \mapsto (\mathsf{B}_i\mathbf{x})_{1\leqslant i\leqslant m}$, and, for every $n \in \mathbb{N}$, $\alpha_n = 1/2$, $\beta_n = \gamma_n/(2\vartheta)$, $\mathbf{T}_n = \mathbf{J}_{\gamma_n\mathbf{A}}$, $\mathbf{R}_n = \mathbf{Id} - \gamma_n\mathbf{B}$, and $\boldsymbol{b}_n = -\gamma_n\boldsymbol{c}_n$. Then, $\mathbf{F} = \operatorname{zer}(\mathbf{A}+\mathbf{B})$ and, for every $n \in \mathbb{N}$, $\mathbf{T}_n$ is α_n-averaged [7, Corollary 23.8], $\mathbf{T}_n\colon \mathbf{x} \mapsto (\mathsf{J}_{\gamma_n\mathsf{A}_i}\mathsf{x}_i)_{1\leqslant i\leqslant m}$ [7, Proposition 23.16], $\mathbf{R}_n$ is β_n-averaged [7, Proposition 4.33], and $\operatorname{Fix}(\mathbf{T}_n \circ \mathbf{R}_n) = \mathbf{F}$ [7, Proposition 25.1(iv)]. Moreover, $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2 \,|\, \boldsymbol{\mathcal{X}}_n)} \leqslant 2\vartheta \sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{c}_n\|^2 \,|\, \boldsymbol{\mathcal{X}}_n)} < +\infty$ and (5.33) is a special case of (4.1). Observe that (4.2) and (4.3) imply the existence of $\widetilde{\Omega} \in \mathcal{F}$ such that $\mathsf{P}(\widetilde{\Omega}) = 1$ and

$$(5.34)\qquad (\forall \omega \in \widetilde{\Omega})(\forall \mathbf{z} \in \mathbf{F}) \quad \begin{cases} \mathbf{T}_n(\mathbf{R}_n\boldsymbol{x}_n(\omega)) - \mathbf{R}_n\boldsymbol{x}_n(\omega) + \mathbf{R}_n\mathbf{z} \to \mathbf{z}\\ \mathbf{R}_n\boldsymbol{x}_n(\omega) - \boldsymbol{x}_n(\omega) - \mathbf{R}_n\mathbf{z} \to -\mathbf{z}. \end{cases}$$

Consequently, since $\inf_{n\in\mathbb{N}}\gamma_n > 0$,

$$(\forall\omega\in\widetilde{\Omega})(\forall \mathsf{z}\in\mathsf{F})\quad \begin{cases}\mathsf{J}_{\gamma_n\mathsf{A}}\big(\boldsymbol{x}_n(\omega)-\gamma_n\mathsf{B}\boldsymbol{x}_n(\omega)\big)-\boldsymbol{x}_n(\omega)=\mathsf{T}_n\big(\mathsf{R}_n\boldsymbol{x}_n(\omega)\big)-\boldsymbol{x}_n(\omega)\to\mathsf{0}\\ \mathsf{B}\boldsymbol{x}_n(\omega)\to\mathsf{Bz}.\end{cases} \tag{5.35}$$

Now set

$$(\forall n\in\mathbb{N})\quad \boldsymbol{y}_n=\mathsf{J}_{\gamma_n\mathsf{A}}(\boldsymbol{x}_n-\gamma_n\mathsf{B}\boldsymbol{x}_n)\quad\text{and}\quad \boldsymbol{u}_n=\gamma_n^{-1}(\boldsymbol{x}_n-\boldsymbol{y}_n)-\mathsf{B}\boldsymbol{x}_n. \tag{5.36}$$

Then (5.35) yields

$$(\forall\omega\in\widetilde{\Omega})(\forall \mathsf{z}\in\mathsf{F})\quad \boldsymbol{x}_n(\omega)-\boldsymbol{y}_n(\omega)\to\mathsf{0}\quad\text{and}\quad \boldsymbol{u}_n(\omega)\to-\mathsf{Bz}. \tag{5.37}$$

Now, let us establish condition (v) of Theorem 4.1. To this end, it is enough to fix $\mathsf{z}\in\mathsf{F}$, $\mathsf{x}\in\mathsf{H}$, a strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$, and $\omega\in\widetilde{\Omega}$ such that $\boldsymbol{x}_{k_n}(\omega)\rightharpoonup\mathsf{x}$ and to show that $\mathsf{x}\in\mathsf{F}$. It follows from (5.35) that $\mathsf{B}\boldsymbol{x}_{k_n}(\omega)\to\mathsf{Bz}$. Hence, since [7, Example 20.28] asserts that B is maximally monotone, we deduce from [7, Proposition 20.33(ii)] that $\mathsf{Bx}=\mathsf{Bz}$. We also derive from (5.37) that $\boldsymbol{y}_{k_n}(\omega)\rightharpoonup\mathsf{x}$ and $\boldsymbol{u}_{k_n}(\omega)\to-\mathsf{Bz}=-\mathsf{Bx}$. Since (5.36) implies that $(\boldsymbol{y}_{k_n}(\omega),\boldsymbol{u}_{k_n}(\omega))_{n\in\mathbb{N}}$ lies in the graph of A, it follows from [7, Proposition 20.33(ii)] that $-\mathsf{Bx}\in\mathsf{Ax}$, i.e., $\mathsf{x}\in\mathsf{F}$. This proves that $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to an F-valued random variable $\boldsymbol{x}$, say

$$\boldsymbol{x}_n(\omega)\rightharpoonup\boldsymbol{x}(\omega) \tag{5.38}$$

for every ω in some $\widehat{\Omega}\in\mathcal{F}$ such that $\widehat{\Omega}\subset\widetilde{\Omega}$ and $\mathsf{P}(\widehat{\Omega})=1$.

Finally take $\omega\in\widehat{\Omega}$. First, suppose that (iv) holds. Then A is demiregular at $\boldsymbol{x}(\omega)$. In view of (5.37) and (5.38), $\boldsymbol{y}_n(\omega)\rightharpoonup\boldsymbol{x}(\omega)$. Furthermore, $\boldsymbol{u}_n(\omega)\to-\mathsf{B}\boldsymbol{x}(\omega)$ and $(\boldsymbol{y}_n(\omega),\boldsymbol{u}_n(\omega))_{n\in\mathbb{N}}$ lies in the graph of A. Altogether $\boldsymbol{y}_n(\omega)\to\boldsymbol{x}(\omega)$ and, therefore $\boldsymbol{x}_n(\omega)\to\boldsymbol{x}(\omega)$. Now, suppose that (v) holds. Then, since (5.35) yields $\mathsf{B}\boldsymbol{x}_n(\omega)\to\mathsf{B}\boldsymbol{x}(\omega)$, (5.38) implies that $\boldsymbol{x}_n(\omega)\to\boldsymbol{x}(\omega)$. ∎

Remark 5.10. Here are a few remarks regarding Proposition 5.9.

(i) Proposition 5.9 generalizes [22, Corollary 6.5 and Remark 6.6], which does not allow for block-processing and uses deterministic variables.

(ii) Problem (5.32) was considered in [2], where it was shown to capture formulations encountered in areas such as evolution equations, game theory, optimization, best approximation, and network flows. It also models domain decomposition problems in partial differential equations [12].

(iii) Proposition 5.9 generalizes [2, Theorem 2.9], which uses a fully parallel deterministic algorithm in which all the blocks are used at each iteration, i.e., $(\forall n\in\mathbb{N})(\forall i\in\{1,\ldots,m\})\ \varepsilon_{i,n}=1$.

(iv) As shown in [26, 28], strongly monotone composite inclusion problems can be solved by applying the forward-backward algorithm to the dual problem. Using Proposition 5.9 we can obtain a block-coordinate version of this primal-dual framework. Likewise, it was shown in [28, 30, 64] that suitably renormed versions of the forward-backward algorithm applied in the primal-dual space yielded a variety of methods for solving composite inclusions in duality. Block-coordinate versions of these methods can be devised via Proposition 5.9.

Next, we present an application of Proposition 5.9 to block-coordinate convex minimization.

Corollary 5.11. *Set* $\mathsf{D} = \{0,1\}^m \smallsetminus \{\mathbf{0}\}$ *and let* $(\mathsf{G}_k)_{1\leqslant k\leqslant p}$ *be separable real Hilbert spaces. For every* $i \in \{1,\ldots,m\}$, *let* $\mathsf{f}_i \in \Gamma_0(\mathsf{H}_i)$ *and, for every* $k \in \{1,\ldots,p\}$, *let* $\tau_k \in \left]0,+\infty\right[$, *let* $\mathsf{g}_k\colon \mathsf{G}_k \to \mathbb{R}$ *be a differentiable convex function with a* τ_k*-Lipschitz-continuous gradient, and let* $\mathsf{L}_{ki}\colon \mathsf{H}_i \to \mathsf{G}_k$ *be linear and bounded. It is assumed that* $\min_{1\leqslant k\leqslant p}\sum_{i=1}^m \|\mathsf{L}_{ki}\|^2 > 0$ *and that the set* $\mathbf{F}$ *of solutions to the problem*

$$\underset{\mathsf{x}_1\in\mathsf{H}_1,\ldots,\mathsf{x}_m\in\mathsf{H}_m}{\text{minimize}} \quad \sum_{i=1}^m \mathsf{f}_i(\mathsf{x}_i) + \sum_{k=1}^p \mathsf{g}_k\Bigg(\sum_{i=1}^m \mathsf{L}_{ki}\mathsf{x}_i\Bigg) \tag{5.39}$$

is nonempty. Let

$$\vartheta \in \left]0, \Bigg(\sum_{k=1}^p \tau_k \Bigg\|\sum_{i=1}^m \mathsf{L}_{ki}\mathsf{L}_{ki}^*\Bigg\|\Bigg)^{-1}\right], \tag{5.40}$$

let $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $\left]0,2\vartheta\right[$ *such that* $\inf_{n\in\mathbb{N}}\gamma_n > 0$ *and* $\sup_{n\in\mathbb{N}}\gamma_n < 2\vartheta$, *and let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a sequence in* $\left]0,1\right]$ *such that* $\inf_{n\in\mathbb{N}}\lambda_n > 0$. *Let* $\boldsymbol{x}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, *and* $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ *be* $\mathbf{H}$*-valued random variables, and let* $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ *be identically distributed* D*-valued random variables. Iterate*

$$\begin{array}{l} \text{for } n = 0,1,\ldots \\ \left\lfloor \begin{array}{l} \text{for } i = 1,\ldots,m \\ \left\lfloor \begin{array}{l} r_{i,n} = \varepsilon_{i,n}\big(x_{i,n} - \gamma_n\big(\sum_{k=1}^p \mathsf{L}_{ki}^*\nabla\mathsf{g}_k\big(\sum_{j=1}^m \mathsf{L}_{kj}x_{j,n}\big) + c_{i,n}\big)\big) \\ x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\lambda_n\big(\operatorname{prox}_{\gamma_n\mathsf{f}_i} r_{i,n} + a_{i,n} - x_{i,n}\big). \end{array}\right. \end{array}\right. \end{array} \tag{5.41}$$

In addition, assume that conditions (i)–(iii) in Proposition 5.9 are satisfied. Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* $\mathbf{F}$*-valued random variable. If, furthermore, one of the following holds (see Lemma 5.8(vi)–(vii) for examples) then* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges strongly* P-a.s. *to* $\boldsymbol{x}$*:*

(i) *For every* $\mathbf{x} \in \mathbf{F}$ *and every* $i \in \{1,\ldots,m\}$, $\partial\mathsf{f}_i$ *is demiregular at* x_i.

(ii) *The operator* $\mathbf{x} \mapsto (\sum_{k=1}^p \mathsf{L}_{ki}^*\nabla\mathsf{g}_k(\sum_{j=1}^m \mathsf{L}_{kj}\mathsf{x}_j))_{1\leqslant i\leqslant m}$ *is demiregular at every point in* $\mathbf{F}$.

Proof. As shown in [2, Section 4], (5.39) is a special case of (5.32) with

$$\mathsf{A}_i = \partial\mathsf{f}_i \quad \text{and} \quad \mathsf{B}_i\colon (\mathsf{x}_j)_{1\leqslant j\leqslant m} \mapsto \sum_{k=1}^p \mathsf{L}_{ki}^*\nabla\mathsf{g}_k\Bigg(\sum_{j=1}^m \mathsf{L}_{kj}\mathsf{x}_j\Bigg). \tag{5.42}$$

Now set $\mathbf{h}\colon \mathbf{H} \to \mathbb{R}\colon \mathbf{x} \mapsto \sum_{k=1}^p \mathsf{g}_k\big(\sum_{i=1}^m \mathsf{L}_{ki}\mathsf{x}_i\big)$. Then $\mathbf{h}$ is a Fréchet-differentiable convex function and $\mathbf{B} = \nabla\mathbf{h}$ is Lipschitz-continuous with constant $1/\vartheta$, where ϑ is given in (5.40). It therefore follows from the Baillon-Haddad theorem [7, Theorem 18.15] that (5.31) holds with this constant. Since, in view of (5.2), (5.33) specializes to (5.41), the convergence claims follow from Proposition 5.9. ■

Remark 5.12. Here are a few observations about Corollary 5.11.

(i) If more assumptions are available about the problem, the Lipschitz constant ϑ of (5.40) can be improved. Some examples are given in [15].

(ii) Recently, some block-coordinate forward-backward methods have been proposed for not necessarily convex minimization problems in Euclidean spaces. Thus, when applied to convex functions satisfying the Kurdyka-Łojasiewicz inequality, the deterministic block-coordinate forward-backward algorithm proposed in [10, Section 3.6] corresponds to the special case of (5.39) in which

$$\mathsf{H} \text{ is a Euclidean space}, \quad p = 1, \quad \text{and} \quad (\forall \mathbf{x} \in \mathsf{H}) \quad \sum_{i=1}^{m} \mathsf{L}_{1i}\mathsf{x}_i = \mathbf{x}. \tag{5.43}$$

In that method, the sweeping proceeds by activating only one block at each iteration according to a periodic schedule. Moreover, errors and relaxations are not allowed. This approach was extended in [20] to an error-tolerant form with a cyclic sweeping rule whereby each block is used at least once within a preset number of consecutive iterations.

(iii) A block-coordinate forward-backward method with random seeping was proposed in [58] in the special case of (5.43). That method uses only one block at each iteration, no relaxation, and no error terms. The asymptotic analysis of [58] provides a lower bound on the probability that $(f + g_1)(x_n)$ be close to $\inf(f + g_1)(\mathsf{H})$, with no result on the convergence of the sequence $(x_n)_{n\in\mathbb{N}}$. Related work is presented in [45, 48].

REFERENCES

[1] M. A. Alghamdi, A. Alotaibi, P. L. Combettes, and N. Shahzad, *A primal-dual method of partial inverses for composite inclusions,* Optim. Lett., 8 (2014), pp. 2271–2284.

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

[3] A. Auslender, *Méthodes numériques pour la décomposition et la minimisation de fonctions non différentiables,* Numer. Math., 18 (1971/72), pp. 213–223.

[4] K. Barty, J.-S. Roy, and C. Strugarek, *Hilbert-valued perturbed subgradient algorithms,* Math. Oper. Res., 32 (2007), pp. 551–562.

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

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

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

[8] H. H. Bauschke, M. N. Dao, D. Noll, and H. M. Phan, *Proximal point algorithm, Douglas-Rachford algorithm and alternating projections: A case study,* J. Convex Anal., to appear.

[9] H. H. Bauschke, F. Deutsch, and H. Hundal, *Characterizing arbitrarily slow convergence in the method of alternating projections,* Int. Trans. Oper. Res., 16 (2009), pp. 413–425.

[10] J. Bolte, S. Sabach, and M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems,* Math. Program., 146 (2014), pp. 459–494.

[11] R. I. Boţ and C. Hendrich, *A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators,* SIAM J. Optim., 23 (2013), pp. 2541–2565.

[12] L. M. Briceño-Arias, *Problèmes d'Inclusions Couplées: Éclatement, Algorithmes et Applications,* Thèse de doctorat, Université Pierre et Marie Curie – Paris 6, Paris, 2011.

[13] L. M. Briceño-Arias, *A Douglas-Rachford splitting method for solving equilibrium problems,* Nonlinear Anal., 75 (2012), pp. 6053–6059.

[14] L. M. Briceño-Arias, *Forward-Douglas-Rachford splitting and forward-partial inverse method for solving monotone inclusions,* Optimization, 64 (2015), pp. 1239–1261.

[15] L. M. Briceño-Arias and P. L. Combettes, *Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery,* Numer. Math. Theory Methods Appl., 2 (2009), pp. 485–508.

[16] L. M. Briceño-Arias and P. L. Combettes, *Monotone operator methods for Nash equilibria in non-potential games,* in: Computational and Analytical Mathematics, (D. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. Vanderwerff, and H. Wolkowicz, eds.), pp. 143–159. Springer, New York, 2013.

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

[18] J. Céa, *Optimisation: Théorie et Algorithmes,* Dunod, Paris, 1971.

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

[20] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, *A block-coordinate variable metric forward-backward algorithm,*
`http://www.optimization-online.org/DB_HTML/2013/12/4178.html`

[21] 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.

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

[23] P. L. Combettes, *Fejér monotonicity in convex optimization,* in: Encyclopedia of Optimization, 2nd ed. (C. A. Floudas and P. M. Pardalos, eds.), pp. 1016–1024. Springer, New York, 2009.

[24] P. L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators,* J. Convex Anal., 16 (2009), pp. 727–748.

[25] P. L. Combettes, *Systems of structured monotone inclusions: Duality, algorithms, and applications,* SIAM J. Optim., 23 (2013), pp. 2420–2447.

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

[27] P. L. Combettes and J.-C. Pesquet, *Proximal splitting methods in signal processing,* in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering,* (H. H. Bauschke et al., eds), pp. 185–212. Springer, New York, 2011.

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

[29] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting,* Multiscale Model. Simul., 4 (2005), pp. 1168–1200.

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

[31] E. De Vito, V. Umanità, and S. Villa, *A consistent algorithm to solve Lasso, elastic-net and Tikhonov regularization,* J. Complexity, 27 (2011), pp. 188–200.

[32] J. Eckstein and D. P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators,* Math. Programming, 55 (1992), pp. 293–318.

[33] I. I. Eremin and L. D. Popov, *Fejér processes in theory and practice: Recent results,* Russian Math. (Iz. VUZ), 53 (2009), pp. 36–55.

[34] Yu. M. Ermol'ev, *On the method of generalized stochastic gradients and quasi-Fejér sequences,*

Cybernetics, 5 (1969), pp. 208–220.

[35] Yu. M. Ermol'ev, *On convergence of random quasi-Fejér sequences,* Cybernetics, 7 (1971), pp. 655–656.

[36] Yu. M. Ermol'ev and A. D. Tuniev, *Random Fejér and quasi-Fejér sequences,* Theory of Optimal Solutions – Akademiya Nauk Ukrainskoĭ SSR Kiev, 2 (1968), pp. 76–83; translated in: American Mathematical Society Selected Translations in Mathematical Statistics and Probability, 13 (1973), pp. 143–148.

[37] R. M. Fortet, *Vecteurs, Fonctions et Distributions Aléatoires dans les Espaces de Hilbert.* Hermès, Paris, 1995.

[38] D. Gabay, *Applications of the method of multipliers to variational inequalities,* in: M. Fortin and R. Glowinski (eds.), Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary Value Problems, pp. 299–331. North-Holland, Amsterdam, 1983.

[39] R. Glowinski and P. Le Tallec (eds.), *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics,* SIAM, Philadelphia, 1989.

[40] H. S. Hundal, *An alternating projection that does not converge in norm,* Nonlinear Anal., 57 (2004), pp. 35–61.

[41] F. Iutzeler, P. Bianchi, P. Ciblat, and W. Hachem, *Asynchronous distributed optimization using a randomized alternating direction method of multipliers,* Proc. 52nd Conf. Decision Control, pp. 3671–3676, Florence, Italy, Dec. 10–13, 2013.

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

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

[44] M. Loève, *Probability Theory II*, 4th ed. Springer, New York, 1978.

[45] Z. Lu and L. Xiao, *On the complexity analysis of randomized block-coordinate descent methods,* Math. Program., to appear.

[46] B. Mercier, *Topics in Finite Element Solution of Elliptic Problems* (Lectures on Mathematics, no. 63). Tata Institute of Fundamental Research, Bombay, 1979.

[47] J. J. Moreau, *Fonctions convexes duales et points proximaux dans un espace hilbertien,* C. R. Acad. Sci. Paris Sér. A, 255 (1962), pp. 2897–2899.

[48] I. Necoara and A. Patrascu, *A random coordinate descent algorithm for optimization problems with composite objective function and linear coupled constraints,* Comput. Optim. Appl., 57 (2014), pp. 307–337.

[49] Yu. Nesterov, *Efficiency of coordinate descent methods on huge-scale optimization problems,* SIAM J. Optim., 22 (2012), pp. 341–362.

[50] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables,* Academic Press, New York, 1970.

[51] N. Papadakis, G. Peyré, and E. Oudet, *Optimal transport with proximal splitting,* SIAM J. Imaging Sci., 7 (2014), pp. 212–238.

[52] T. Pennanen and M. Kallio, *A splitting method for stochastic programs,* Ann. Oper. Res., 142 (2006), pp. 259–268.

[53] J.-C. Pesquet and N. Pustelnik, *A parallel inertial proximal optimization method,* Pac. J. Optim., 8 (2012), pp. 273–305.

[54] W. V. Petryshyn, *Construction of fixed points of demicompact mappings in Hilbert space,* J. Math. Anal. Appl., 14 (1966), pp. 276–284.

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

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

[57] E. RAIK, *Fejér type methods in Hilbert space,* Eesti NSV Tead. Akad. Toimetised Füüs.-Mat., 16 (1967), pp. 286–293.

[58] P. RICHTÁRIK AND M. TAKÁČ, *Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function,* Math. Program., 144 (2014), pp. 1–38.

[59] 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.

[60] S. SRA, S. NOWOZIN, AND S. J. WRIGHT, (EDS.) *Optimization for Machine Learning.* MIT Press, Cambridge, MA, 2011.

[61] P. TSENG, *Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming,* Math. Programming, 48 (1990), pp. 249–263.

[62] P. TSENG, *Applications of a splitting algorithm to decomposition in convex programming and variational inequalities,* SIAM J. Control Optim., 29 (1991), pp. 119–138.

[63] S. VILLA, S. SALZO, L. BALDASSARRE, AND A. VERRI, *Accelerated and inexact forward-backward algorithms,* SIAM J. Optim., 23 (2013), pp. 1607–1633.

[64] B. C. VŨ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators,* Adv. Comput. Math., 38 (2013), pp. 667–681.

[65] D. C. YOULA, *Mathematical theory of image restoration by the method of convex projections,* in: H. Stark (ed.) Image Recovery: Theory and Application, pp. 29–77. Academic Press, San Diego, CA, 1987.

[66] E. ZEIDLER, *Nonlinear Functional Analysis and Its Applications I: Fixed-Point Theorems.* Springer-Verlag, New York, 1986.