

# A Class of Randomized Primal-Dual Algorithms for Distributed Optimization

Jean-Christophe Pesquet, Audrey Repetti

► **To cite this version:**

Jean-Christophe Pesquet, Audrey Repetti. A Class of Randomized Primal-Dual Algorithms for Distributed Optimization. 2014. hal-01077615

**HAL Id: hal-01077615**

**<https://hal.archives-ouvertes.fr/hal-01077615>**

Preprint submitted on 26 Oct 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Class of Randomized Primal-Dual Algorithms for Distributed Optimization

Jean-Christophe Pesquet and Audrey Repetti \*

Université Paris-Est  
Laboratoire d'Informatique Gaspard Monge – CNRS UMR 8049  
77454 Marne la Vallée Cedex 2, France  
[first.last@univ-paris-est.fr](mailto:first.last@univ-paris-est.fr)

## Abstract

Based on a preconditioned version of the randomized block-coordinate forward-backward algorithm recently proposed in [23], several variants of block-coordinate primal-dual algorithms are designed in order to solve a wide array of monotone inclusion problems. These methods rely on a sweep of blocks of variables which are activated at each iteration according to a random rule, and they allow stochastic errors in the evaluation of the involved operators. Then, this framework is employed to derive block-coordinate primal-dual proximal algorithms for solving composite convex variational problems. The resulting algorithm implementations may be useful for reducing computational complexity and memory requirements. Furthermore, we show that the proposed approach can be used to develop novel asynchronous distributed primal-dual algorithms in a multi-agent context.

**Keywords.** Block-coordinate algorithm, convex optimization, distributed algorithm, monotone operator, preconditioning, primal-dual algorithm, stochastic quasi-Fejér sequence.

**MSC.** 47H05, 49M29, 49M27, 65K10, 90C25.

---

\*This work was supported by the CNRS MASTODONS project (grant 2013 MesureHD).

# 1 Introduction

There has been recently a growing interest in primal-dual approaches for finding a zero of a sum of monotone operators or minimizing a sum of proper lower-semicontinuous convex functions (see [36] and the references therein). When various linear operators are involved in the formulation of the problem under investigation, solving jointly its primal and dual forms allows the design of strategies where none of the linear operators needs to be inverted. Avoiding such inversions may offer a significant advantage in terms of computational complexity when dealing with large-scale problems (see e.g. [6, 27, 31, 35, 46, 48, 52]).

Various classes of fixed-point primal-dual algorithms have been developed, in particular those based on the forward-backward iteration [13, 16, 18, 19, 24, 26, 28, 30, 32, 37, 45, 54], on the forward-backward-forward iteration [6, 9, 12, 17, 22], on the Douglas-Rachford iteration [8, 23], or those derived from other principles [1, 2, 15, 41]. This work is focused on the first class of primal-dual algorithms. When searching for a zero of a sum of monotone operators, the most recent versions of these methods can exploit the properties of each operator either in an implicit manner, through the use of its resolvent, or in a direct manner when the operator is cocoercive. When a sum of convex functions is minimized, this brings the ability either to make use of the proximity operator of each function or to employ its gradient if the function is Lipschitz differentiable. As discussed in [4, 21, 43], the proximity operator of a function is a versatile tool in convex optimization for tackling possibly nonsmooth problems, but it may be sometimes preferable, in particular for complexity reasons, to compute the gradient of the function when it enjoys some smoothness property.

Most of the aforementioned primal-dual methods make it possible to split the original problem in a sum of simpler terms whose associated operators can be addressed individually, in a parallel manner, at each iteration of the algorithm. Our objective in this paper is to add more flexibility to the existing primal-dual methods by allowing only a restricted number of these operators to be activated at each iteration. In the line of the work in [23], our approach will be grounded on the use of random sweeping techniques which are applicable to algorithms generating (quasi-)Fejér monotone sequences. One additional benefit of the proposed randomized approach is that it leads to algorithms which can be proved to be tolerant of stochastic errors satisfying some summability condition.

In the following, we will investigate two variants of forward-backward based primal-dual algorithms and we will design block-coordinate versions of both algorithms. These block-coordinate methods may be interesting for their own sake in order to reduce memory and computational loads, but it turns out that they are also instrumental in developing distributed strategies. More precisely, we will be interested in multi-agent problems where the performed updates can be limited to a neighborhood of a small number of agents in an asynchronous way. We will show that the proposed random distributed schemes apply not only to convex optimization problems, but also to general monotone inclusion ones. It is worth noting that, in the variational case, some distributed

primal-dual algorithms have already been proposed implementing subgradient steps [14, 55] (see also [53] for applications to data networks). As a general feature of (unaveraged) subgradient methods, their convergence requires the use of step-sizes converging to zero. Making use of proximity operators, which can be viewed as implicit subgradient descent steps, allows less restrictive step-size choices to be made. For example, convergence of the iterates can be established for constant step-size values.

The remainder of the paper is organized as follows. In Section 2 we provide some relevant background on monotone operator theory and convex analysis, and we introduce our notation. In Section 3, a preconditioned random block-coordinate version of the forward-backward iteration is presented. Based on this algorithm, in Section 4, we propose novel block-coordinate primal-dual methods for constructing iteratively a zero of a sum of monotone operators, and we study their convergence. In Section 5, similar block-coordinate primal-dual algorithms are developed for solving composite convex optimization problems. Finally, in Section 6, we show how the proposed random block-coordinate approaches are able to provide distributed iterative solutions to monotone inclusion and convex variational problems.

## 2 Notation

The reader is referred to [5] for background on monotone operator theory and convex analysis, and to [29] for background on probability in Hilbert spaces. Throughout this work,  $(\Omega, \mathcal{F}, P)$  is the underlying probability space. For simplicity, the same notation  $\langle \cdot | \cdot \rangle$  (resp.  $\| \cdot \|$ ) is used for the inner products (resp. norms) which equip all the Hilbert spaces considered in this paper. Let  $H$  be a separable real Hilbert space with Borel  $\sigma$ -algebra  $\mathcal{B}$ . A  $H$ -valued random variable is a measurable map  $x: (\Omega, \mathcal{F}) \rightarrow (H, \mathcal{B})$ . The smallest  $\sigma$ -algebra generated by a family  $\Phi$  of random variables is denoted by  $\sigma(\Phi)$ . The expectation is denoted by  $E(\cdot)$ .

Let  $G$  be a real Hilbert space. We denote by  $\mathcal{B}(H, G)$  the space of bounded linear operators from  $H$  to  $G$ , and we set  $\mathcal{B}(H) = \mathcal{B}(H, H)$ . Let  $L \in \mathcal{B}(H, G)$ , its adjoint is denoted by  $L^*$ .  $L \in \mathcal{B}(H)$  is a strongly positive self-adjoint operator if  $L^* = L$  and there exists  $\alpha \in ]0, +\infty[$  such that  $(\forall x \in H) \langle x | Lx \rangle \geq \alpha \|x\|^2$ . Then,  $L$  is an isomorphism and its inverse is a strongly positive self-adjoint operator in  $\mathcal{B}(H)$ . The square root of a strongly positive operator  $L$  is denoted by  $L^{1/2}$  and its inverse by  $L^{-1/2}$ .  $\text{Id}$  denotes the identity operator on  $H$ .

The power set of  $H$  is denoted by  $2^H$ . Let  $A: H \rightarrow 2^H$  be a set-valued operator. If, for every  $x \in H$ ,  $Ax$  is a singleton, then  $A$  will be identified with a mapping from  $H$  to  $H$ . We denote by  $\text{zer} A = \{x \in H \mid 0 \in Ax\}$  the set of zeros of  $A$  and by  $A^{-1}: H \mapsto 2^H: u \mapsto \{x \in H \mid u \in Ax\}$  the inverse of  $A$ . Operator  $A$  is monotone if  $(\forall (x, y) \in H^2) (\forall u \in Ax) (\forall v \in Ay) \langle x - y | u - v \rangle \geq 0$ . Such an operator is maximally monotone if there exists no other monotone operator whose graph includes the graph of  $A$ .  $A$  is  $\beta$ -strongly monotone for some  $\beta \in ]0, +\infty[$  if  $(\forall (x, y) \in H^2) (\forall u \in Ax) (\forall v \in Ay) \langle x - y | u - v \rangle \geq \beta \|x - y\|^2$ . Let  $B$  be a single-valued operator from  $H$  to  $H$ .  $B$  is  $\beta$ -

cocoercive for some  $\beta \in ]0, +\infty[$  if  $(\forall(x, y) \in H^2) \langle x - y \mid Bx - By \rangle \geq \beta \|Bx - By\|^2$ . Therefore,  $B$  is  $\beta$ -cocoercive if and only if  $B^{-1}: H \rightarrow 2^H$  is  $\beta$ -strongly monotone.  $B$  is  $\alpha$ -averaged with  $\alpha \in ]0, 1[$  if  $(\forall(x, y) \in H^2) \|Bx - By\|^2 \leq \|x - y\|^2 - \frac{1-\alpha}{\alpha} \|(\text{Id} - B)x - (\text{Id} - B)y\|^2$ . As a consequence of Minty's theorem, an operator  $A: H \rightarrow 2^H$  is maximally monotone if and only if its resolvent  $J_A = (\text{Id} + A)^{-1}$  is a firmly nonexpansive (i.e. 1-cocoercive) operator from  $H$  to  $H$ . As a generalization of Moreau's decomposition formula, if  $A: H \rightarrow 2^H$  is maximally monotone,  $U$  is a strongly positive self-adjoint operator in  $\mathcal{B}(H)$ , and  $\gamma \in ]0, +\infty[$ , then  $J_{\gamma U A}: H \rightarrow H$  is such that

$$J_{\gamma U A} = U^{1/2} J_{\gamma U^{1/2} A U^{1/2}} U^{-1/2} = \text{Id} - \gamma U J_{\gamma^{-1} U^{-1} A^{-1}} (\gamma^{-1} U^{-1} \cdot) \quad (2.1)$$

(see [24, Example 3.9]). The parallel sum of  $A: H \rightarrow 2^H$  and  $C: H \rightarrow 2^H$  is  $A \square C = (A^{-1} + C^{-1})^{-1}$ .

The domain of a function  $f: H \rightarrow ]-\infty, +\infty]$  is  $\text{dom } f = \{x \in H \mid f(x) < +\infty\}$ . A function with a nonempty domain is said to be proper. The class of proper, convex, lower-semicontinuous functions from  $H$  to  $]-\infty, +\infty]$  is denoted by  $\Gamma_0(H)$ . If  $f \in \Gamma_0(H)$ , then the Moreau subdifferential of  $f$  is the maximally monotone operator

$$\partial f: H \rightarrow 2^H: x \mapsto \{u \in H \mid (\forall y \in H) \langle y - x \mid u \rangle + f(x) \leq f(y)\}. \quad (2.2)$$

If  $f$  is proper and  $\beta$ -strongly convex for some  $\beta \in ]0, +\infty[$ , then  $\partial f$  is  $\beta$ -strongly monotone. If  $f \in \Gamma_0(H)$  is Gâteaux-differentiable at  $x \in H$ , then  $\partial f(x) = \{\nabla f(x)\}$  where  $\nabla f(x)$  is the gradient of  $f$  at  $x$ .  $f: H \rightarrow \mathbb{R}$  is  $\beta^{-1}$ -Lipschitz differentiable for some  $\beta \in ]0, +\infty[$  if it is Gâteaux-differentiable on  $H$  and  $(\forall(x, y) \in H^2) \beta \|\nabla f(x) - \nabla f(y)\| \leq \|x - y\|$ . The Baillon-Haddad theorem asserts that a differentiable convex function  $f$  defined on  $H$  is  $\beta^{-1}$ -Lipschitz differentiable if and only if its gradient  $\nabla f$  is  $\beta$ -cocoercive. If  $\Lambda$  is a nonempty subset of  $H$ , the indicator function of  $\Lambda$  is  $(\forall x \in H) \iota_\Lambda(x) = 0$  if  $x \in \Lambda$ , and  $+\infty$  otherwise. This function belongs to  $\Gamma_0(H)$  if and only if  $\Lambda$  is a nonempty closed convex set. Its subdifferential  $\partial \iota_\Lambda$  is the normal cone to  $\Lambda$ , denoted by  $N_\Lambda$ . The identity element of the parallel sum is  $N_{\{0\}}$ . The inf-convolution of two functions  $f: H \rightarrow ]-\infty, +\infty]$  and  $h: H \rightarrow ]-\infty, +\infty]$  is defined as  $f \square h: H \rightarrow ]-\infty, +\infty]: x \mapsto \inf_{y \in H} (f(y) + h(x - y))$ . The identity element of the inf-convolution is  $\iota_{\{0\}}$ . The conjugate of a function  $f \in \Gamma_0(H)$  is  $f^* \in \Gamma_0(H)$  such that  $(\forall v \in H) f^*(v) = \sup_{x \in H} (\langle x \mid v \rangle - f(x))$ . We have then  $\partial f^* = (\partial f)^{-1}$ . Let  $U$  be a strongly positive self-adjoint operator in  $\mathcal{B}(H)$ . The proximity operator of  $f \in \Gamma_0(H)$  relative to the metric induced by  $U$  is [33, Section XV.4]

$$\text{prox}_f^U: H \rightarrow H: x \rightarrow \underset{y \in H}{\text{argmin}} f(y) + \frac{1}{2} \langle x - y \mid U(x - y) \rangle. \quad (2.3)$$

We have thus  $\text{prox}_f^U = J_{U^{-1} \partial f}$ . When  $U = \text{Id}$ , we retrieve the standard definition of the proximity operator originally introduced in [39]. If  $\Lambda$  is a nonempty closed convex subset of  $H$ ,  $\Pi_\Lambda = \text{prox}_{\iota_\Lambda}^{\text{Id}}$  is the projector onto  $\Lambda$ . In the following, the relative interior of a subset  $\Lambda$  of  $H$  is denoted by  $\text{ri } \Lambda$ .

Let  $(G_i)_{1 \leq i \leq m}$  be real Hilbert spaces.  $\mathbf{G} = G_1 \oplus \cdots \oplus G_m$  is their Hilbert direct sum, i.e., their product space endowed with the scalar product  $(\mathbf{x}, \mathbf{y}) \mapsto \sum_{i=1}^m \langle x_i \mid y_i \rangle$ , where a generic element in  $\mathbf{G}$  is denoted by  $\mathbf{x} = (x_i)_{1 \leq i \leq m}$  with  $x_i \in G_i$ , for every  $i \in \{1, \dots, m\}$ . In addition,  $\mathbb{D}_m = \{0, 1\}^m \setminus \{\mathbf{0}\}$  denotes the set of nonzero binary strings of length  $m$ . We will keep on using this notation throughout the paper.

### 3 A preconditioned random block-coordinate forward-backward algorithm

In this section,  $m$  is a positive integer,  $K_1, \dots, K_m$  are separable real Hilbert spaces, and  $\mathbf{K} = K_1 \oplus \dots \oplus K_m$  is their Hilbert direct sum.

The algorithms in this paper are rooted in the forward-backward iteration [25] (see [3] for examples of problems which can be solved by this method). A block-coordinate version of the forward-backward method was recently proposed in [23, Section 5.2]. Stochastic versions of this algorithm were also presented in [40, 49] in a variational framework. Now, we show how a preconditioning operator can be included in the block-coordinate forward-backward algorithm through a metric change.

**Proposition 3.1** *Let  $\mathbf{Q}: \mathbf{K} \rightarrow 2^{\mathbf{K}}$  be a maximally monotone operator and let  $\mathbf{R}: \mathbf{K} \rightarrow \mathbf{K}$  be a cocoercive operator. Assume that  $\mathbf{Z} = \text{zer}(\mathbf{Q} + \mathbf{R})$  is nonempty. Let  $\mathbf{V}$  be a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{K})$  such that  $\mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$  is  $\vartheta$ -cocoercive with  $\vartheta \in ]0, +\infty[$ . Let  $(\gamma_n)_{n \in \mathbb{N}}$  be a sequence in  $\mathbb{R}$  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  $\mathbf{z}_0$ ,  $(\mathbf{s}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{t}_n)_{n \in \mathbb{N}}$  be  $\mathbf{K}$ -valued random variables, and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_m$ -valued random variables. For every  $n \in \mathbb{N}$ , set  $\mathbf{J}_{\gamma_n \mathbf{V} \mathbf{Q}}: \mathbf{z} \mapsto (\mathbb{T}_{i,n} \mathbf{z})_{1 \leq i \leq m}$  where  $(\forall i \in \{1, \dots, m\}) \mathbb{T}_{i,n}: \mathbf{K} \rightarrow K_i$ , iterate*

$$\begin{aligned} & \text{for } n = 0, 1, \dots \\ & \left[ \begin{array}{l} \mathbf{r}_n = \mathbf{V} \mathbf{R} \mathbf{z}_n \\ \text{for } i = 1, \dots, m \\ \left[ \begin{array}{l} z_{i,n+1} = z_{i,n} + \lambda_n \varepsilon_{i,n} (\mathbb{T}_{i,n}(\mathbf{z}_n - \gamma_n \mathbf{r}_n + \mathbf{s}_n) + \mathbf{t}_{i,n} - z_{i,n}), \end{array} \right. \end{array} \right. \end{aligned} \quad (3.1)$$

and set  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{Z}_n = \sigma(\mathbf{z}_0, \dots, \mathbf{z}_n)$ . In addition, assume that the following hold:

- (i)  $\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{s}_n\|^2 | \mathcal{Z}_n)} < +\infty$  and  $\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{t}_n\|^2 | \mathcal{Z}_n)} < +\infty$  P-a.s.
- (ii) For every  $n \in \mathbb{N}$ ,  $\mathcal{E}_n$  and  $\mathcal{Z}_n$  are independent and  $(\forall i \in \{1, \dots, m\}) \mathbb{P}[\varepsilon_{i,0} = 1] > 0$ .

Then  $(\mathbf{z}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\mathbf{Z}$ -valued random variable.

*Proof.* We have  $\mathbf{Z} = \text{zer}(\mathbf{V} \mathbf{Q} + \mathbf{V} \mathbf{R}) \neq \emptyset$ . Since  $\mathbf{V}$  is a strongly positive self-adjoint operator, we can renorm the space  $\mathbf{K}$  with the norm:

$$(\forall \mathbf{z} \in \mathbf{K}) \quad \|\mathbf{z}\|_{\mathbf{V}^{-1}} = \sqrt{\langle \mathbf{z} | \mathbf{V}^{-1} \mathbf{z} \rangle}. \quad (3.2)$$

Let  $\langle \cdot | \cdot \rangle_{\mathbf{V}^{-1}}$  denote the associated inner product. In this renormed space,  $\mathbf{V} \mathbf{Q}$  is maximally mono-

tone. In addition,

$$\begin{aligned}
(\forall (\mathbf{z}, \mathbf{z}') \in \mathbf{K}^2) \quad \|\mathbf{VRz} - \mathbf{VRz}'\|_{\mathbf{V}^{-1}}^2 &= \|\mathbf{V}^{1/2}\mathbf{Rz} - \mathbf{V}^{1/2}\mathbf{Rz}'\|^2 \\
&\leq \vartheta^{-1} \langle \mathbf{V}^{-1/2}\mathbf{z} - \mathbf{V}^{-1/2}\mathbf{z}' \mid \mathbf{V}^{1/2}\mathbf{Rz} - \mathbf{V}^{1/2}\mathbf{Rz}' \rangle \\
&= \vartheta^{-1} \langle \mathbf{z} - \mathbf{z}' \mid \mathbf{Rz} - \mathbf{Rz}' \rangle \\
&= \vartheta^{-1} \langle \mathbf{z} - \mathbf{z}' \mid \mathbf{VRz} - \mathbf{VRz}' \rangle_{\mathbf{V}^{-1}}, \tag{3.3}
\end{aligned}$$

which shows that  $\mathbf{VR}$  is  $\vartheta$ -cocoercive in  $(\mathbf{K}, \|\cdot\|_{\mathbf{V}^{-1}})$ . A forward-backward iteration can thus be employed to find an element of  $\mathbf{Z}$  by composing operators  $\mathbf{J}_{\gamma_n \mathbf{VQ}}$  and  $\mathbf{Id} - \gamma_n \mathbf{VR}$ . In  $(\mathbf{K}, \|\cdot\|_{\mathbf{V}^{-1}})$ , the first operator is firmly nonexpansive (hence,  $1/2$ -averaged) and the second one is  $\gamma_n/(2\vartheta)$ -averaged [5, Proposition 4.33]. The relaxed randomized algorithm given in [23, Section 4] then takes the form (3.1). The convergence result follows from [23, Theorem 4.1] by noticing that Assumption (i) leads to

$$\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{s}_n\|_{\mathbf{V}^{-1}}^2 \mid \mathfrak{Z}_n)} \leq \sqrt{\|\mathbf{V}^{-1}\|} \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{s}_n\|^2 \mid \mathfrak{Z}_n)} < +\infty \tag{3.4}$$

$$\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{t}_n\|_{\mathbf{V}^{-1}}^2 \mid \mathfrak{Z}_n)} \leq \sqrt{\|\mathbf{V}^{-1}\|} \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{t}_n\|^2 \mid \mathfrak{Z}_n)} < +\infty \tag{3.5}$$

and that weak convergences in the sense of  $\langle \cdot \mid \cdot \rangle$  and  $\langle \cdot \mid \cdot \rangle_{\mathbf{V}^{-1}}$  are equivalent.  $\square$

### Remark 3.2

- (i) If  $\mathbf{R} = \mathbf{L}^* \tilde{\mathbf{R}} \mathbf{L}$  where  $\mathbf{L} \in \mathcal{B}(\mathbf{K}, \tilde{\mathbf{K}})$ ,  $\tilde{\mathbf{R}}: \tilde{\mathbf{K}} \rightarrow \tilde{\mathbf{K}}$  is  $\tilde{\vartheta}$ -cocoercive with  $\tilde{\vartheta} \in ]0, +\infty[$ , and  $\tilde{\mathbf{K}}$  is a separable real Hilbert space, then  $\mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2}$  is  $\vartheta$ -cocoercive for every strongly positive self-adjoint operator  $\mathbf{V} \in \mathcal{B}(\mathbf{K})$  such that  $\vartheta \|\mathbf{L} \mathbf{V} \mathbf{L}^*\| = \tilde{\vartheta}$ .
- (ii) At iteration  $n \in \mathbb{N}$ ,  $\mathbf{s}_n$  and  $\mathbf{t}_n$  can be viewed as error terms when applying  $\mathbf{R}$  and  $\mathbf{J}_{\gamma_n \mathbf{VQ}}$ , respectively. The ability to consider summable stochastic errors offers more freedom than the assumption of summable deterministic errors which is often adopted in the literature. Note however that relative error models are considered in [38, 50, 51].
- (iii) Let  $n \in \mathbb{N}^*$ . In view of (3.1),  $\mathfrak{E}_n$  and  $\mathfrak{Z}_n$  are independent if  $\varepsilon_n$  is independent of  $(\mathbf{z}_0, (\varepsilon_{n'}, \mathbf{s}_{n'}, \mathbf{t}_{n'})_{0 \leq n' < n})$ .

## 4 Block-coordinate primal-dual algorithms for composite monotone inclusion problems

In the rest of this section,  $p$  and  $q$  are positive integers,  $(\mathbf{H}_j)_{1 \leq j \leq p}$  and  $(\mathbf{G}_k)_{1 \leq k \leq q}$  are separable real Hilbert spaces. In addition,  $\mathbf{H} = \mathbf{H}_1 \oplus \cdots \oplus \mathbf{H}_p$  and  $\mathbf{G} = \mathbf{G}_1 \oplus \cdots \oplus \mathbf{G}_q$  denote the Hilbert direct sums of  $(\mathbf{H}_j)_{1 \leq j \leq p}$  and  $(\mathbf{G}_k)_{1 \leq k \leq q}$ , respectively. We will also consider the product space  $\mathbf{K} = \mathbf{H} \oplus \mathbf{G}$ .

## 4.1 Problem

The following problem involving monotone operators which has drawn much attention in the last years (see e.g. [8, 11, 17, 22, 44, 47]) will play a prominent role throughout this work.

**Problem 4.1** For every  $j \in \{1, \dots, p\}$ , let  $A_j: H_j \rightarrow 2^{H_j}$  be maximally monotone, let  $C_j: H_j \rightarrow H_j$  be cocoercive and, for every  $k \in \{1, \dots, q\}$ , let  $B_k: G_k \rightarrow 2^{G_k}$  be maximally monotone, let  $D_k: G_k \rightarrow 2^{G_k}$  be maximally and strongly monotone, and let  $L_{k,j} \in \mathcal{B}(H_j, G_k)$ . It is assumed that

$$(\forall k \in \{1, \dots, q\}) \quad \mathbb{L}_k = \{j \in \{1, \dots, p\} \mid L_{k,j} \neq 0\} \neq \emptyset, \quad (4.1)$$

$$(\forall j \in \{1, \dots, p\}) \quad \mathbb{L}_j^* = \{k \in \{1, \dots, q\} \mid L_{k,j} \neq 0\} \neq \emptyset, \quad (4.2)$$

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

find  $x_1 \in H_1, \dots, x_p \in H_p$  such that

$$(\forall j \in \{1, \dots, p\}) \quad 0 \in A_j x_j + C_j x_j + \sum_{k=1}^q L_{k,j}^* (B_k \square D_k) \left( \sum_{j'=1}^p L_{k,j'} x_{j'} \right) \quad (4.3)$$

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

find  $v_1 \in G_1, \dots, v_q \in G_q$  such that

$$(\forall k \in \{1, \dots, q\}) \quad 0 \in - \sum_{j=1}^p L_{k,j} (A_j^{-1} \square C_j^{-1}) \left( - \sum_{k'=1}^q L_{k',j}^* v_{k'} \right) + B_k^{-1} v_k + D_k^{-1} v_k. \quad (4.4)$$

Our objective is to find a pair  $(\hat{x}, \hat{v})$  of random variables such that  $\hat{x}$  is  $\mathbf{F}$ -valued and  $\hat{v}$  is  $\mathbf{F}^*$ -valued.

The previous problem can be recast as a search for a zero of the sum of two maximally monotone operators in the product space  $\mathbf{K}$  as indicated below [26, 54].

**Proposition 4.2** Let us define  $\mathbf{A}: \mathbf{H} \rightarrow 2^{\mathbf{H}}: \mathbf{x} \mapsto \times_{j=1}^p A_j x_j$ ,  $\mathbf{B}: \mathbf{G} \rightarrow 2^{\mathbf{G}}: \mathbf{v} \mapsto \times_{k=1}^q B_k v_k$ ,  $\mathbf{C}: \mathbf{H} \rightarrow \mathbf{H}: \mathbf{x} \mapsto (C_j x_j)_{1 \leq j \leq p}$ ,  $\mathbf{D}: \mathbf{G} \rightarrow 2^{\mathbf{G}}: \mathbf{v} \mapsto \times_{k=1}^q D_k v_k$ , and  $\mathbf{L}: \mathbf{H} \rightarrow \mathbf{G}: \mathbf{x} \mapsto \left( \sum_{j=1}^p L_{k,j} x_j \right)_{1 \leq k \leq q}$ . Let us now introduce the operators

$$\mathbf{Q}: \quad \mathbf{K} \rightarrow 2^{\mathbf{K}} \\ (\mathbf{x}, \mathbf{v}) \mapsto (\mathbf{A}\mathbf{x} + \mathbf{L}^*\mathbf{v}) \times (-\mathbf{L}\mathbf{x} + \mathbf{B}^{-1}\mathbf{v}) \quad (4.5)$$

and

$$\mathbf{R}: \quad \mathbf{K} \rightarrow \mathbf{K} \\ (\mathbf{x}, \mathbf{v}) \mapsto (\mathbf{C}\mathbf{x}, \mathbf{D}^{-1}\mathbf{v}). \quad (4.6)$$



Then, the following hold:

- (i)  $\mathbf{Q}$  is maximally monotone and  $\mathbf{R}$  is cocoercive.
- (ii)  $\mathbf{Z} = \text{zer}(\mathbf{Q} + \mathbf{R})$  is nonempty.
- (iii) A pair  $(\hat{\mathbf{x}}, \hat{\mathbf{v}})$  of random variables is a solution to Problem 4.1 if and only if  $(\hat{\mathbf{x}}, \hat{\mathbf{v}})$  is  $\mathbf{Z}$ -valued.

The above properties suggest employing the block-coordinate forward-backward algorithm developed in Section 3 to solve numerically Problem 4.1. According to the choice of the involved preconditioning operator, several algorithms can be devised. Subsequently,  $\mathbf{L} \in \mathcal{B}(\mathbf{H}, \mathbf{G})$  is defined as in Proposition 4.2.

## 4.2 First algorithm subclass

We state two preliminary results which will be useful in the derivation of the algorithms proposed in this section.

**Lemma 4.3** *Let  $\mathbf{W} \in \mathcal{B}(\mathbf{H})$  and  $\mathbf{U} \in \mathcal{B}(\mathbf{G})$  be two strongly positive self-adjoint operators such that  $\|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\| < 1$ .*

- (i) *The operator defined by*

$$\begin{aligned} \mathbf{V}' : \quad & \mathbf{K} \rightarrow \mathbf{K} \\ (\mathbf{x}, \mathbf{v}) \mapsto & (\mathbf{W}^{-1}\mathbf{x} - \mathbf{L}^*\mathbf{v}, -\mathbf{L}\mathbf{x} + \mathbf{U}^{-1}\mathbf{v}) \end{aligned} \quad (4.7)$$

*is a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{K})$ . Its inverse given by*

$$\begin{aligned} \mathbf{V} : \quad & \mathbf{K} \rightarrow \mathbf{K} \\ (\mathbf{x}, \mathbf{v}) \mapsto & ((\mathbf{W}^{-1} - \mathbf{L}^*\mathbf{U}\mathbf{L})^{-1}\mathbf{x} + \mathbf{W}\mathbf{L}^*(\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)^{-1}\mathbf{v}, (\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)^{-1}(\mathbf{L}\mathbf{W}\mathbf{x} + \mathbf{v})) \end{aligned} \quad (4.8)$$

*is also a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{K})$ .*

- (ii) *Let  $\mathbf{C} : \mathbf{H} \rightarrow \mathbf{H}$ ,  $\mathbf{D} : \mathbf{G} \rightarrow 2^{\mathbf{G}}$ , and  $\mathbf{R} : \mathbf{K} \rightarrow \mathbf{K}$  be the operators defined in Proposition 4.2. If  $\mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}$  is  $\mu$ -cocoercive with  $\mu \in ]0, +\infty[$  and  $\mathbf{U}^{1/2}\mathbf{D}^{-1}\mathbf{U}^{1/2}$  is  $\nu$ -cocoercive with  $\nu \in ]0, +\infty[$ , then, for every  $\alpha \in ]0, +\infty[$ ,  $\mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$  is  $\vartheta_\alpha$ -cocoercive, where*

$$\vartheta_\alpha = (1 - \|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\|^2) \min \{ \mu(1 + \alpha\|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\|)^{-1}, \nu(1 + \alpha^{-1}\|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\|)^{-1} \}. \quad (4.9)$$

*Proof.* (i) The operators  $\mathbf{W}^{-1}$  and  $\mathbf{U}^{-1}$  being linear bounded and self-adjoint,  $\mathbf{V}'$  is linear bounded and self-adjoint. In addition, for every  $(\mathbf{x}, \mathbf{v}) \in \mathbf{K}$ ,

$$\begin{aligned} \langle \mathbf{x} \mid (\mathbf{W}^{-1} - \mathbf{L}^* \mathbf{U} \mathbf{L}) \mathbf{x} \rangle &= \langle \mathbf{W}^{-1/2} \mathbf{x} \mid (\mathbf{Id} - \mathbf{W}^{1/2} \mathbf{L}^* \mathbf{U} \mathbf{L} \mathbf{W}^{1/2}) \mathbf{W}^{-1/2} \mathbf{x} \rangle \\ &= \langle \mathbf{x} \mid \mathbf{W}^{-1} \mathbf{x} \rangle - \langle \mathbf{W}^{-1/2} \mathbf{x} \mid \mathbf{W}^{1/2} \mathbf{L}^* \mathbf{U} \mathbf{L} \mathbf{W}^{1/2} \mathbf{W}^{-1/2} \mathbf{x} \rangle \\ &\geq (1 - \|\mathbf{W}^{1/2} \mathbf{L}^* \mathbf{U} \mathbf{L} \mathbf{W}^{1/2}\|) \langle \mathbf{x} \mid \mathbf{W}^{-1} \mathbf{x} \rangle \\ &\geq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2) \|\mathbf{W}\|^{-1} \|\mathbf{x}\|^2 \end{aligned} \quad (4.10)$$

and

$$\begin{aligned} \langle \mathbf{v} \mid (\mathbf{U}^{-1} - \mathbf{L} \mathbf{W} \mathbf{L}^*) \mathbf{v} \rangle &\geq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W} \mathbf{L}^* \mathbf{U}^{1/2}\|) \langle \mathbf{v} \mid \mathbf{U}^{-1} \mathbf{v} \rangle \\ &\geq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2) \|\mathbf{U}\|^{-1} \|\mathbf{v}\|^2. \end{aligned} \quad (4.11)$$

We can deduce that

$$\begin{aligned} \langle (\mathbf{x}, \mathbf{v}) \mid \mathbf{V}'(\mathbf{x}, \mathbf{v}) \rangle &= \langle \mathbf{x} - \mathbf{W} \mathbf{L}^* \mathbf{v} \mid \mathbf{W}^{-1}(\mathbf{x} - \mathbf{W} \mathbf{L}^* \mathbf{v}) \rangle + \langle \mathbf{v} \mid (\mathbf{U}^{-1} - \mathbf{L} \mathbf{W} \mathbf{L}^*) \mathbf{v} \rangle \\ &\geq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2) \|\mathbf{U}\|^{-1} \|\mathbf{v}\|^2 \end{aligned} \quad (4.12)$$

and similarly,

$$\langle (\mathbf{x}, \mathbf{v}) \mid \mathbf{V}'(\mathbf{x}, \mathbf{v}) \rangle \geq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2) \|\mathbf{W}\|^{-1} \|\mathbf{x}\|^2. \quad (4.13)$$

The latter two inequalities yield

$$\begin{aligned} \langle (\mathbf{x}, \mathbf{v}) \mid \mathbf{V}'(\mathbf{x}, \mathbf{v}) \rangle &\geq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2) \min\{\|\mathbf{W}\|^{-1}, \|\mathbf{U}\|^{-1}\} \max\{\|\mathbf{x}\|^2, \|\mathbf{v}\|^2\} \\ &\geq \frac{1}{2} (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2) \min\{\|\mathbf{W}\|^{-1}, \|\mathbf{U}\|^{-1}\} (\|\mathbf{x}\|^2 + \|\mathbf{v}\|^2). \end{aligned} \quad (4.14)$$

This shows that  $\mathbf{V}'$  is a strongly positive operator. It is thus an isomorphism and its inverse is a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{K})$ .

Furthermore, (4.10) (resp. (4.11)) shows that  $\mathbf{W}^{-1} - \mathbf{L}^* \mathbf{U} \mathbf{L}$  (resp.  $\mathbf{U}^{-1} - \mathbf{L} \mathbf{W} \mathbf{L}^*$ ) is an isomorphism since it is a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{H})$  (resp.  $\mathcal{B}(\mathbf{G})$ ). The expression of the inverse of  $\mathbf{V}'$  can be checked by direct calculations.

(ii) Let  $\alpha \in ]0, +\infty[$ . Showing that  $\mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2}$  is  $\vartheta_\alpha$ -cocoercive is tantamount to establishing that

$$\begin{aligned} (\forall (\mathbf{z}, \mathbf{z}') \in \mathbf{K}^2) \quad \langle \mathbf{z} - \mathbf{z}' \mid \mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2} \mathbf{z} - \mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2} \mathbf{z}' \rangle &\geq \vartheta_\alpha \|\mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2} \mathbf{z} - \mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2} \mathbf{z}'\|^2 \\ \Leftrightarrow (\forall (\mathbf{z}, \mathbf{z}') \in \mathbf{K}^2) \quad \langle \mathbf{z} - \mathbf{z}' \mid \mathbf{R} \mathbf{z} - \mathbf{R} \mathbf{z}' \rangle &\geq \vartheta_\alpha \|\mathbf{R} \mathbf{z} - \mathbf{R} \mathbf{z}'\|_{\mathbf{V}}^2. \end{aligned} \quad (4.15)$$

Let  $\mathbf{z} = (\mathbf{x}, \mathbf{v}) \in \mathbf{K}$  and  $\mathbf{z}' = (\mathbf{x}', \mathbf{v}') \in \mathbf{K}$ . We have

$$\begin{aligned} \|\mathbf{R} \mathbf{z} - \mathbf{R} \mathbf{z}'\|_{\mathbf{V}}^2 &= \langle \mathbf{C} \mathbf{x} - \mathbf{C} \mathbf{x}' \mid (\mathbf{W}^{-1} - \mathbf{L}^* \mathbf{U} \mathbf{L})^{-1} (\mathbf{C} \mathbf{x} - \mathbf{C} \mathbf{x}') \rangle \\ &\quad + \langle \mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}' \mid (\mathbf{U}^{-1} - \mathbf{L} \mathbf{W} \mathbf{L}^*)^{-1} (\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}') \rangle \\ &\quad + 2 \langle \mathbf{C} \mathbf{x} - \mathbf{C} \mathbf{x}' \mid \mathbf{W} \mathbf{L}^* (\mathbf{U}^{-1} - \mathbf{L} \mathbf{W} \mathbf{L}^*)^{-1} (\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}') \rangle. \end{aligned} \quad (4.16)$$

On the other hand,

$$\begin{aligned}
& \langle \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}' \mid (\mathbf{W}^{-1} - \mathbf{L}^* \mathbf{U} \mathbf{L})^{-1} (\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}') \rangle \\
&= \langle \mathbf{W}^{1/2} (\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}') \mid (\mathbf{Id} - \mathbf{W}^{1/2} \mathbf{L}^* \mathbf{U} \mathbf{L} \mathbf{W}^{1/2})^{-1} \mathbf{W}^{1/2} (\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}') \rangle \\
&\leq \|(\mathbf{Id} - \mathbf{W}^{1/2} \mathbf{L}^* \mathbf{U} \mathbf{L} \mathbf{W}^{1/2})^{-1}\| \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}'\|_{\mathbf{W}}^2 \\
&= (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2)^{-1} \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}'\|_{\mathbf{W}}^2, \tag{4.17}
\end{aligned}$$

$$\begin{aligned}
& \langle \mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}' \mid (\mathbf{U}^{-1} - \mathbf{L} \mathbf{W} \mathbf{L}^*)^{-1} (\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}') \rangle \\
&\leq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2)^{-1} \|\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}'\|_{\mathbf{U}}^2, \tag{4.18}
\end{aligned}$$

$$\begin{aligned}
& \langle \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}' \mid \mathbf{W} \mathbf{L}^* (\mathbf{U}^{-1} - \mathbf{L} \mathbf{W} \mathbf{L}^*)^{-1} (\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}') \rangle \\
&\leq \|\mathbf{W}^{1/2} (\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}')\| \|\mathbf{W}^{1/2} \mathbf{L}^* \mathbf{U}^{1/2} (\mathbf{Id} - \mathbf{U}^{1/2} \mathbf{L} \mathbf{W} \mathbf{L}^* \mathbf{U}^{1/2})^{-1}\| \|\mathbf{U}^{1/2} (\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}')\| \\
&\leq \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\| (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2)^{-1} \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}'\|_{\mathbf{W}} \|\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}'\|_{\mathbf{U}} \\
&\leq \frac{1}{2} \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\| (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2)^{-1} (\alpha \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}'\|_{\mathbf{W}}^2 + \alpha^{-1} \|\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}'\|_{\mathbf{U}}^2). \tag{4.19}
\end{aligned}$$

Altogether, (4.16)-(4.19) and the cocoercivity assumptions on  $\mathbf{W}^{1/2} \mathbf{C} \mathbf{W}^{1/2}$  and  $\mathbf{U}^{1/2} \mathbf{D}^{-1} \mathbf{U}^{1/2}$  lead to the inequalities

$$\begin{aligned}
\|\mathbf{R}\mathbf{z} - \mathbf{R}\mathbf{z}'\|_{\mathbf{V}}^2 &\leq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2)^{-1} ((1 + \alpha \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|) \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}'\|_{\mathbf{W}}^2 \\
&\quad + (1 + \alpha^{-1} \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|) \|\mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}'\|_{\mathbf{U}}^2) \\
&\leq (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2)^{-1} (\mu^{-1} (1 + \alpha \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|) \langle \mathbf{x} - \mathbf{x}' \mid \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}' \rangle \\
&\quad + \nu^{-1} (1 + \alpha^{-1} \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|) \langle \mathbf{v} - \mathbf{v}' \mid \mathbf{D}^{-1} \mathbf{v} - \mathbf{D}^{-1} \mathbf{v}' \rangle) \\
&\leq \vartheta_{\alpha}^{-1} \langle \mathbf{z} - \mathbf{z}' \mid \mathbf{R}\mathbf{z} - \mathbf{R}\mathbf{z}' \rangle. \tag{4.20}
\end{aligned}$$

□

#### Remark 4.4

(i) In (4.9), we can simply choose  $\alpha = 1$ , yielding the cocoercivity constant

$$\vartheta_1 = (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|) \min\{\mu, \nu\}. \tag{4.21}$$

A tighter value of this constant is  $\vartheta_{\hat{\alpha}}$  where  $\hat{\alpha}$  is the maximizer of  $\alpha \mapsto \vartheta_{\alpha}$  on  $]0, +\infty[$ . It can be readily shown that

$$\hat{\alpha} = \frac{\mu - \nu + \sqrt{(\mu - \nu)^2 + 4\mu\nu \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2}}{2\nu \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|}. \tag{4.22}$$

(ii) When  $\mathbf{D}^{-1} = \mathbf{0}$ , the positive constant  $\nu$  can be chosen arbitrarily large. A cocoercivity constant of  $\mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2}$  is then equal to

$$\lim_{\substack{\alpha \rightarrow 0 \\ \alpha > 0}} \vartheta_{\alpha} = (1 - \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\|^2) \mu. \tag{4.23}$$

**Lemma 4.5** Let  $\mathbf{A}: \mathbf{H} \rightarrow 2^{\mathbf{H}}$ ,  $\mathbf{B}: \mathbf{G} \rightarrow 2^{\mathbf{G}}$ ,  $\mathbf{C}: \mathbf{H} \rightarrow \mathbf{H}$ ,  $\mathbf{D}: \mathbf{G} \rightarrow 2^{\mathbf{G}}$ ,  $\mathbf{Q}: \mathbf{K} \rightarrow 2^{\mathbf{K}}$ , and  $\mathbf{R}: \mathbf{K} \rightarrow \mathbf{K}$ . Let  $\mathbf{W} \in \mathcal{B}(\mathbf{H})$  and  $\mathbf{U} \in \mathcal{B}(\mathbf{G})$  be two strongly positive self-adjoint operators such that  $\|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\| < 1$ . Let  $\mathbf{V} \in \mathcal{B}(\mathbf{K})$  be defined by (4.8). For every  $\mathbf{z} = (\mathbf{x}, \mathbf{v}) \in \mathbf{K}$  and  $(\mathbf{c}, \mathbf{e}) \in \mathbf{K}$ , let

$$\begin{cases} \mathbf{y} = \mathbf{J}_{\mathbf{W}\mathbf{A}}(\mathbf{x} - \mathbf{W}(\mathbf{L}^*\mathbf{v} + \mathbf{C}\mathbf{x} + \mathbf{c})) \\ \mathbf{u} = \mathbf{J}_{\mathbf{U}\mathbf{B}^{-1}}(\mathbf{v} + \mathbf{U}(\mathbf{L}(2\mathbf{y} - \mathbf{x}) - \mathbf{D}^{-1}\mathbf{v} + \mathbf{e})). \end{cases} \quad (4.24)$$

Then,  $(\mathbf{y}, \mathbf{u}) = \mathbf{J}_{\mathbf{V}\mathbf{Q}}(\mathbf{z} - \mathbf{V}\mathbf{R}\mathbf{z} + \mathbf{s})$  where

$$\mathbf{s} = ((\mathbf{W}^{-1} - \mathbf{L}^*\mathbf{U}\mathbf{L})^{-1}(\mathbf{L}^*\mathbf{U}\mathbf{e} - \mathbf{c}), (\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)^{-1}(\mathbf{e} - \mathbf{L}\mathbf{W}\mathbf{c})). \quad (4.25)$$

*Proof.* Let  $\mathbf{z} = (\mathbf{x}, \mathbf{v}) \in \mathbf{K}$  and let  $\mathbf{s} = (\mathbf{c}', \mathbf{e}') \in \mathbf{K}$ . We have the following equivalences:

$$\begin{aligned} & (\mathbf{y}, \mathbf{u}) = \mathbf{J}_{\mathbf{V}\mathbf{Q}}(\mathbf{z} - \mathbf{V}\mathbf{R}\mathbf{z} + \mathbf{s}) \\ \Leftrightarrow & \mathbf{z} - \mathbf{V}\mathbf{R}\mathbf{z} + \mathbf{s} \in (\mathbf{Id} + \mathbf{V}\mathbf{Q})(\mathbf{y}, \mathbf{u}) \\ \Leftrightarrow & \mathbf{V}^{-1}(\mathbf{z} + \mathbf{s} - (\mathbf{y}, \mathbf{u})) - \mathbf{R}\mathbf{z} \in \mathbf{Q}(\mathbf{y}, \mathbf{u}) \\ \Leftrightarrow & \begin{cases} \mathbf{W}^{-1}(\mathbf{x} - \mathbf{y} + \mathbf{c}') - \mathbf{L}^*(\mathbf{v} + \mathbf{e}') - \mathbf{C}\mathbf{x} \in \mathbf{A}\mathbf{y} \\ \mathbf{U}^{-1}(\mathbf{v} - \mathbf{u} + \mathbf{e}') + \mathbf{L}(2\mathbf{y} - \mathbf{x} - \mathbf{c}') - \mathbf{D}^{-1}\mathbf{v} \in \mathbf{B}^{-1}\mathbf{u} \end{cases} \end{aligned} \quad (4.26)$$

$$\begin{aligned} \Leftrightarrow & \begin{cases} \mathbf{x} + \mathbf{c}' - \mathbf{W}(\mathbf{L}^*(\mathbf{v} + \mathbf{e}') + \mathbf{C}\mathbf{x}) \in (\mathbf{Id} + \mathbf{W}\mathbf{A})\mathbf{y} \\ \mathbf{v} + \mathbf{e}' + \mathbf{U}(\mathbf{L}(2\mathbf{y} - \mathbf{x} - \mathbf{c}') - \mathbf{D}^{-1}\mathbf{v}) \in (\mathbf{Id} + \mathbf{U}\mathbf{B}^{-1})\mathbf{u} \end{cases} \\ \Leftrightarrow & \begin{cases} \mathbf{y} = \mathbf{J}_{\mathbf{W}\mathbf{A}}(\mathbf{x} + \mathbf{c}' - \mathbf{W}(\mathbf{L}^*(\mathbf{v} + \mathbf{e}') + \mathbf{C}\mathbf{x})) \\ \mathbf{u} = \mathbf{J}_{\mathbf{U}\mathbf{B}^{-1}}(\mathbf{v} + \mathbf{e}' + \mathbf{U}(\mathbf{L}(2\mathbf{y} - \mathbf{x} - \mathbf{c}') - \mathbf{D}^{-1}\mathbf{v})), \end{cases} \end{aligned} \quad (4.27)$$

where, in (4.26), we have used the expression of  $\mathbf{Q}$  in (4.5) and the expression of the inverse of  $\mathbf{V}$  given by (4.7).

In order to conclude, let us note that, since  $\|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\| < 1$ , it has already been observed in the proof of Lemma 4.3(i) that  $\mathbf{W}^{-1} - \mathbf{L}^*\mathbf{U}\mathbf{L}$  and  $\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*$  are isomorphisms (as a result of (4.10) and (4.11)). Thus, for every  $(\mathbf{c}, \mathbf{e}) \in \mathbf{K}$ ,

$$\begin{cases} \mathbf{W}\mathbf{c} = \mathbf{W}\mathbf{L}^*\mathbf{e}' - \mathbf{c}' \\ \mathbf{U}\mathbf{e} = \mathbf{e}' - \mathbf{U}\mathbf{L}\mathbf{c}' \end{cases} \Leftrightarrow \begin{cases} \mathbf{c}' = (\mathbf{W}^{-1} - \mathbf{L}^*\mathbf{U}\mathbf{L})^{-1}(\mathbf{L}^*\mathbf{U}\mathbf{e} - \mathbf{c}) \\ \mathbf{e}' = (\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)^{-1}(\mathbf{e} - \mathbf{L}\mathbf{W}\mathbf{c}). \end{cases} \quad (4.28)$$

□

The above two lemmas allow us to obtain a first block-coordinate primal-dual algorithm to generate a solution to Problem 4.1.

**Proposition 4.6** Let

$$\mathbf{W}: \mathbf{H} \rightarrow \mathbf{H}: \mathbf{x} \mapsto (W_1\mathbf{x}_1, \dots, W_p\mathbf{x}_p) \quad \text{and} \quad \mathbf{U}: \mathbf{G} \rightarrow \mathbf{G}: \mathbf{v} \mapsto (U_1\mathbf{v}_1, \dots, U_q\mathbf{v}_q) \quad (4.29)$$

where, for every  $j \in \{1, \dots, p\}$ ,  $W_j$  is a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{H}_j)$  such that  $W_j^{1/2}\mathbf{C}_jW_j^{1/2}$  is  $\mu_j$ -cocoercive with  $\mu_j \in ]0, +\infty[$  and, for every  $k \in \{1, \dots, q\}$ ,  $U_k$  is a strongly positive

self-adjoint operator in  $\mathcal{B}(\mathbf{G}_k)$  such that  $U_k^{1/2}D_k^{-1}U_k^{1/2}$  is  $\nu_k$ -cocoercive with  $\nu_k \in ]0, +\infty[$ . Suppose that

$$(\exists \alpha \in ]0, +\infty[) \quad 2\vartheta_\alpha > 1 \quad (4.30)$$

where  $\vartheta_\alpha$  is defined by (4.9) with  $\mu = \min\{\mu_1, \dots, \mu_p\}$  and  $\nu = \min\{\nu_1, \dots, \nu_q\}$ . Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1[$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$ ,  $(\mathbf{a}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}$ -valued random variables, let  $\mathbf{v}_0$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_{p+q}$ -valued random variables. Iterate

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \left[ \begin{array}{l} \text{for } j = 1, \dots, p \\ \left[ \begin{array}{l} y_{j,n} = \varepsilon_{j,n} \left( \mathbf{J}_{W_j A_j} (x_{j,n} - W_j \left( \sum_{k \in \mathbb{L}_j^*} L_{k,j}^* v_{k,n} + C_j x_{j,n} + c_{j,n} \right)) + a_{j,n} \right) \\ x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} (y_{j,n} - x_{j,n}) \end{array} \right. \\ \text{for } k = 1, \dots, q \\ \left[ \begin{array}{l} u_{k,n} = \varepsilon_{p+k,n} \left( \mathbf{J}_{U_k B_k^{-1}} (v_{k,n} + U_k \left( \sum_{j \in \mathbb{L}_k} L_{k,j} (2y_{j,n} - x_{j,n}) - D_k^{-1} v_{k,n} + d_{k,n} \right)) + b_{k,n} \right) \\ v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n}), \end{array} \right. \end{array} \right. \quad (4.31) \end{array}$$

and set  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ . In addition, assume that the following hold:

- (i)  $\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{a}_n\|^2 | \mathcal{X}_n)} < +\infty$ ,  $\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{b}_n\|^2 | \mathcal{X}_n)} < +\infty$ ,  $\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{c}_n\|^2 | \mathcal{X}_n)} < +\infty$ , and  $\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{d}_n\|^2 | \mathcal{X}_n)} < +\infty$  P-a.s.
- (ii) For every  $n \in \mathbb{N}$ ,  $\mathcal{E}_n$  and  $\mathcal{X}_n$  are independent, and  $(\forall k \in \{1, \dots, q\}) \mathbb{P}[\varepsilon_{p+k,0} = 1] > 0$ .
- (iii) For every  $j \in \{1, \dots, p\}$  and  $n \in \mathbb{N}$ ,  $\bigcup_{k \in \mathbb{L}_j^*} \{\omega \in \Omega \mid \varepsilon_{p+k,n}(\omega) = 1\} \subset \{\omega \in \Omega \mid \varepsilon_{j,n}(\omega) = 1\}$ .

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

*Proof.* In view of (iii), for every  $j \in \{1, \dots, p\}$ ,  $\max\{\varepsilon_{j,n}, (\varepsilon_{p+k,n})_{k \in \mathbb{L}_j^*}\} = \varepsilon_{j,n}$ . Moreover, for every  $k \in \{1, \dots, q\}$ ,  $j \in \mathbb{L}_k \Leftrightarrow k \in \mathbb{L}_j^*$ . Iterations (4.31) are thus equivalent to

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \left[ \begin{array}{l} \text{for } j = 1, \dots, p \\ \left[ \begin{array}{l} \eta_{j,n} = \max\{\varepsilon_{j,n}, (\varepsilon_{p+k,n})_{k \in \mathbb{L}_j^*}\} \\ y_{j,n} = \eta_{j,n} \left( \mathbf{J}_{W_j A_j} (x_{j,n} - W_j \left( \sum_{k=1}^q L_{k,j}^* v_{k,n} + C_j x_{j,n} + c_{j,n} \right)) + a_{j,n} \right) \\ x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} (y_{j,n} - x_{j,n}) \end{array} \right. \\ \text{for } k = 1, \dots, q \\ \left[ \begin{array}{l} u_{k,n} = \varepsilon_{p+k,n} \left( \mathbf{J}_{U_k B_k^{-1}} (v_{k,n} + U_k \left( \sum_{j \in \mathbb{L}_k} L_{k,j} (2y_{j,n} - x_{j,n}) - D_k^{-1} v_{k,n} + d_{k,n} \right)) + b_{k,n} \right) \\ v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n}). \end{array} \right. \end{array} \right. \quad (4.32) \end{array}$$

On the other hand, according to Proposition 4.2(i)-(ii),  $\mathbf{Q}$  is maximally monotone,  $\mathbf{R}$  is cocoercive, and  $\mathbf{Z} = \text{zer}(\mathbf{Q} + \mathbf{R}) \neq \emptyset$ . It can be noticed that (4.9) and (4.30) imply that  $\|\mathbf{U}^{1/2}\mathbf{LW}^{1/2}\| < 1$ . Thus, by virtue of Lemma 4.5, Algorithm (4.32) can be rewritten under the form of Algorithm (3.1), where  $m = p + q$ ,  $\mathbf{V}$  is defined by (4.8) and, for every  $n \in \mathbb{N}$ ,

$$\mathbf{z}_n = (\mathbf{x}_n, \mathbf{v}_n), \quad (4.33)$$

$$\gamma_n = 1, \quad (4.34)$$

$$\mathbf{J}_{\mathbf{VQ}}: \mathbf{z} \mapsto (\mathbf{T}_{i,n}\mathbf{z})_{1 \leq i \leq m}, \quad (4.35)$$

$$(\forall j \in \{1, \dots, p\}) \quad \mathbf{T}_{j,n}: \mathbf{K} \rightarrow \mathbf{H}_j, \quad (4.36)$$

$$(\forall k \in \{1, \dots, q\}) \quad \mathbf{T}_{p+k,n}: \mathbf{K} \rightarrow \mathbf{G}_k, \quad (4.37)$$

$$\mathbf{t}_n = (\mathbf{a}_n, \mathbf{b}_n), \quad (4.38)$$

$$\mathbf{s}_n = ((\mathbf{W}^{-1} - \mathbf{L}^*\mathbf{U}\mathbf{L})^{-1}(\mathbf{L}^*\mathbf{U}\mathbf{e}_n - \mathbf{c}_n), (\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*)^{-1}(\mathbf{e}_n - \mathbf{LW}\mathbf{c}_n)), \quad (4.39)$$

$$\mathbf{e}_n = 2\mathbf{L}\mathbf{a}_n + \mathbf{d}_n. \quad (4.40)$$

Since  $\mathbf{W}$  and  $\mathbf{U}$  are two strongly positive self-adjoint operators such that  $\|\mathbf{U}^{1/2}\mathbf{LW}^{1/2}\| < 1$ , Lemma 4.3(i) allows us to claim that  $\mathbf{V}$  is a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{K})$ . In addition, for every  $(\mathbf{x}, \mathbf{x}') \in \mathbf{H}^2$ ,

$$\begin{aligned} \langle \mathbf{x} - \mathbf{x}' \mid \mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}\mathbf{x} - \mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}\mathbf{x}' \rangle &= \sum_{j=1}^p \langle \mathbf{x}_j - \mathbf{x}'_j \mid \mathbf{W}_j^{1/2}\mathbf{C}_j\mathbf{W}_j^{1/2}\mathbf{x}_j - \mathbf{W}_j^{1/2}\mathbf{C}_j\mathbf{W}_j^{1/2}\mathbf{x}'_j \rangle \\ &\geq \sum_{j=1}^p \mu_j \|\mathbf{W}_j^{1/2}\mathbf{C}_j\mathbf{W}_j^{1/2}\mathbf{x}_j - \mathbf{W}_j^{1/2}\mathbf{C}_j\mathbf{W}_j^{1/2}\mathbf{x}'_j\|^2 \\ &\geq \mu \|\mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}\mathbf{x} - \mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}\mathbf{x}'\|^2. \end{aligned} \quad (4.41)$$

Thus  $\mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}$  is  $\mu$ -cocoercive, and similarly,  $\mathbf{U}^{1/2}\mathbf{D}^{-1}\mathbf{U}^{1/2}$  is  $\nu$ -cocoercive. It follows from Lemma 4.3(ii) that  $\mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$  is  $\vartheta_\alpha$ -cocoercive, and our assumptions guarantee that  $1 = \sup_{n \in \mathbb{N}} \gamma_n < 2\vartheta_\alpha$ . Moreover, it can be deduced from Condition (i) and (4.38)-(4.40) that

$$\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{t}_n\|^2 \mid \mathcal{X}_n)} \leq \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{a}_n\|^2 \mid \mathcal{X}_n)} + \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{b}_n\|^2 \mid \mathcal{X}_n)} < +\infty, \quad (4.42)$$

$$\begin{aligned} &\sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{s}_n\|^2 \mid \mathcal{X}_n)} \\ &\leq \|(\mathbf{W}^{-1} - \mathbf{L}^*\mathbf{U}\mathbf{L})^{-1}\| \left( \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{c}_n\|^2 \mid \mathcal{X}_n)} + 2\|\mathbf{L}^*\mathbf{U}\mathbf{L}\| \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{a}_n\|^2 \mid \mathcal{X}_n)} \right) \\ &\quad + \|\mathbf{L}^*\mathbf{U}\| \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{d}_n\|^2 \mid \mathcal{X}_n)} + \|(\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*)^{-1}\| \left( 2\|\mathbf{L}\| \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{a}_n\|^2 \mid \mathcal{X}_n)} \right) \\ &\quad + \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{d}_n\|^2 \mid \mathcal{X}_n)} + \|\mathbf{LW}\| \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{c}_n\|^2 \mid \mathcal{X}_n)} < +\infty. \end{aligned} \quad (4.43)$$

In addition, since we have assumed that, for every  $j \in \{1, \dots, p\}$ ,  $\mathbb{L}_j^* \neq \emptyset$ , (ii) and (iii) guarantee that Condition (ii) in Proposition 3.1 is also fulfilled. All the assumptions of Proposition 3.1 are

then satisfied, which allows us to establish the almost sure convergence of  $(\mathbf{x}_n, \mathbf{v}_n)_{n \in \mathbb{N}}$  to a  $\mathbf{Z}$ -valued random variable. Finally, Proposition 4.2(iii) ensures that the limit is an  $\mathbf{F} \times \mathbf{F}^*$ -valued random variable.  $\square$

A number of observations can be made on Proposition 4.6.

**Remark 4.7**

- (i) The Boolean random variables  $(\varepsilon_{i,n})_{1 \leq i \leq p+q}$  signal the variables  $(x_{j,n})_{1 \leq j \leq p}$  and  $(v_{k,n})_{1 \leq k \leq q}$  that are activated at each iteration  $n$ . From a computational standpoint, when some of them are zero-valued, no update of the associated variables must be performed. Note that, in accordance with Condition (iii), for every  $j \in \{1, \dots, p\}$ ,  $y_{j,n}$  needs to be computed not only when  $x_{j,n}$  is activated, but also when there exists  $k \in \{1, \dots, q\}$  such that  $v_{k,n}$  is activated and  $L_{k,j} \neq 0$ .
- (ii) For every  $n \in \mathbb{N}$ ,  $j \in \{1, \dots, p\}$ , and  $k \in \{1, \dots, q\}$ ,  $a_{j,n}$ ,  $b_{k,n}$ ,  $c_{j,n}$ , and  $d_{k,n}$  model stochastic errors possibly arising at iteration  $n$ , when applying  $J_{W_j A_j}$ ,  $J_{U_k B_k^{-1}}$ ,  $C_j$ , and  $D_k^{-1}$ , respectively.
- (iii) Using the triangle and Cauchy-Schwarz inequalities yields

$$\begin{aligned}
(\mathbf{x} \in \mathbf{H}) \quad \|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2} \mathbf{x}\|^2 &= \sum_{k=1}^q \left\| \sum_{j=1}^p U_k^{1/2} L_{k,j} W_j^{1/2} x_j \right\|^2 \\
&\leq \sum_{k=1}^q \left( \sum_{j=1}^p \|U_k^{1/2} L_{k,j} W_j^{1/2}\| \|x_j\| \right)^2 \\
&\leq \sum_{k=1}^q \left( \sum_{j=1}^p \|U_k^{1/2} L_{k,j} W_j^{1/2}\|^2 \right) \left( \sum_{j=1}^p \|x_j\|^2 \right), \tag{4.44}
\end{aligned}$$

which shows that

$$\|\mathbf{U}^{1/2} \mathbf{L} \mathbf{W}^{1/2}\| \leq \left( \sum_{j=1}^p \sum_{k=1}^q \|U_k^{1/2} L_{k,j} W_j^{1/2}\|^2 \right)^{1/2}. \tag{4.45}$$

For every  $j \in \{1, \dots, p\}$ , let a cocoercivity constant of  $C_j$  be denoted by  $\tilde{\mu}_j \in ]0, +\infty[$  and, for every  $k \in \{1, \dots, q\}$ , let a strong monotonicity constant of  $D_k$  be denoted by  $\tilde{\nu}_k \in ]0, +\infty[$ . Then, one can choose

$$\mu = \min\{(\|W_j\|^{-1} \tilde{\mu}_j)_{1 \leq j \leq p}\}, \quad \nu = \min\{(\|U_k\|^{-1} \tilde{\nu}_k)_{1 \leq k \leq q}\}. \tag{4.46}$$

Therefore, by using Remark 4.4(i), a sufficient condition for (4.30) to be satisfied with  $\alpha = 1$  is

$$\left( 1 - \left( \sum_{j=1}^p \sum_{k=1}^q \|U_k^{1/2} L_{k,j} W_j^{1/2}\|^2 \right)^{1/2} \right) \min\{(\|W_j\|^{-1} \tilde{\mu}_j)_{1 \leq j \leq p}, (\|U_k\|^{-1} \tilde{\nu}_k)_{1 \leq k \leq q}\} > \frac{1}{2}. \tag{4.47}$$

When  $\mathbf{D}^{-1} = \mathbf{0}$ , in accordance with Remark 4.4(ii), this condition can be replaced by

$$\left(1 - \sum_{j=1}^p \sum_{k=1}^q \|\mathbf{U}_k^{1/2} \mathbf{L}_{k,j} \mathbf{W}_j^{1/2}\|^2\right) \min\{(\|\mathbf{W}_j\|^{-1} \tilde{\mu}_j)_{1 \leq j \leq p}\} > \frac{1}{2}. \quad (4.48)$$

(iv) The above algorithm extends a number of results existing in a deterministic setting, when  $p = 1$  and no random sweeping is performed. In most of these works,  $\mathbf{W}_1 = \tau \text{Id}$  and, for every  $k \in \{1, \dots, q\}$ ,  $\mathbf{U}_k = \rho_k \text{Id}$  where  $(\tau, \rho_1, \dots, \rho_q) \in ]0, +\infty[^{q+1}$ . In particular, in [54], a sufficient condition for (4.47) to be satisfied is employed, while in [26] it is assumed that  $\mathbf{D}^{-1} = \mathbf{0}$  and a condition similar to (4.48) is used. The proposed block-coordinate algorithm also extends the results in [24, Section 6] when a constant metric is considered.

Due to the symmetry existing between the primal and the dual problems, we can swap the roles of these two problems, so leading to a symmetric form of Algorithm (4.31):

**Proposition 4.8** *Let  $\mathbf{W}$ ,  $\mathbf{U}$ ,  $\mu$ , and  $\nu$  be defined as in Proposition 4.6. Suppose that (4.30) holds where  $\vartheta_\alpha$  is defined by (4.9). Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1]$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$ ,  $(\mathbf{a}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}$ -valued random variables, let  $\mathbf{v}_0$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_{p+q}$ -valued random variables. Iterate*

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \quad \left[ \begin{array}{l} \text{for } k = 1, \dots, q \\ \quad \left[ \begin{array}{l} u_{k,n} = \varepsilon_{p+k,n} \left( \mathbf{J}_{\mathbf{U}_k \mathbf{B}_k^{-1}}(v_{k,n} + \mathbf{U}_k \left( \sum_{j \in \mathbb{L}_k} \mathbf{L}_{k,j} x_{j,n} - \mathbf{D}_k^{-1} v_{k,n} + \mathbf{d}_{k,n} \right)) + \mathbf{b}_{k,n} \right) \\ v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n}) \end{array} \right. \\ \text{for } j = 1, \dots, p \\ \quad \left[ \begin{array}{l} y_{j,n} = \varepsilon_{j,n} \left( \mathbf{J}_{\mathbf{W}_j \mathbf{A}_j}(x_{j,n} - \mathbf{W}_j \left( \sum_{k \in \mathbb{L}_j^*} \mathbf{L}_{k,j}^* (2u_{k,n} - v_{k,n}) + \mathbf{C}_j x_{j,n} + \mathbf{c}_{j,n} \right)) + \mathbf{a}_{j,n} \right) \\ x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} (y_{j,n} - x_{j,n}). \end{array} \right. \end{array} \right. \end{array} \quad (4.49)$$

In addition, assume that Condition (i) in Proposition 4.6 is satisfied where  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ , and that the following hold:

- (ii) For every  $n \in \mathbb{N}$ ,  $\mathcal{E}_n$  and  $\mathcal{X}_n$  are independent, and  $(\forall j \in \{1, \dots, p\}) \mathbb{P}[\varepsilon_{j,0} = 1] > 0$ .
- (iii) For every  $k \in \{1, \dots, q\}$  and  $n \in \mathbb{N}$ ,  $\bigcup_{j \in \mathbb{L}_k} \{\omega \in \Omega \mid \varepsilon_{j,n}(\omega) = 1\} \subset \{\omega \in \Omega \mid \varepsilon_{p+k,n}(\omega) = 1\}$ .

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



### 4.3 Second algorithm subclass

We now consider a diagonal form of the operator  $\mathbf{V}$ , for which we proceed similarly to the approach followed in Section 4.2.

**Lemma 4.9** *Let  $\mathbf{W} \in \mathcal{B}(\mathbf{H})$  and  $\mathbf{U} \in \mathcal{B}(\mathbf{G})$  be two strongly positive self-adjoint operators such that  $\|\mathbf{U}^{1/2}\mathbf{LW}^{1/2}\| < 1$ .*

(i) *The operator defined by*

$$\begin{aligned} \mathbf{V}: \quad \mathbf{K} &\rightarrow \mathbf{K} \\ (\mathbf{x}, \mathbf{v}) &\mapsto (\mathbf{W}\mathbf{x}, (\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*)^{-1}\mathbf{v}) \end{aligned} \quad (4.50)$$

*is a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{K})$ .*

(ii) *Let  $\mathbf{C}: \mathbf{H} \rightarrow \mathbf{H}$ ,  $\mathbf{D}: \mathbf{G} \rightarrow 2^{\mathbf{G}}$ , and  $\mathbf{R}: \mathbf{K} \rightarrow \mathbf{K}$  be the operators defined in Proposition 4.2. If  $\mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}$  is  $\mu$ -cocoercive with  $\mu \in ]0, +\infty[$  and  $\mathbf{U}^{1/2}\mathbf{D}^{-1}\mathbf{U}^{1/2}$  is  $\nu$ -cocoercive with  $\nu \in ]0, +\infty[$ , then  $\mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$  is  $\vartheta$ -cocoercive, where*

$$\vartheta = \min \{ \mu, \nu(1 - \|\mathbf{U}^{1/2}\mathbf{LW}^{1/2}\|^2) \}. \quad (4.51)$$

*Proof.* (i) First note that, as we have already shown in (4.11), if  $\|\mathbf{U}^{1/2}\mathbf{LW}^{1/2}\| < 1$ , then  $\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*$  is a strongly positive operator in  $\mathcal{B}(\mathbf{G})$  and it is thus an isomorphism. We have then, for every  $(\mathbf{x}, \mathbf{v}) \in \mathbf{K}$ ,

$$\begin{aligned} \langle (\mathbf{x}, \mathbf{v}) \mid \mathbf{V}(\mathbf{x}, \mathbf{v}) \rangle &= \langle \mathbf{x} \mid \mathbf{W}\mathbf{x} \rangle + \langle \mathbf{v} \mid (\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*)^{-1}\mathbf{v} \rangle \\ &\geq \|\mathbf{W}^{-1}\|^{-1}\|\mathbf{x}\|^2 + \|\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*\|^{-1}\|\mathbf{v}\|^2 \\ &\geq \min \{ \|\mathbf{W}^{-1}\|^{-1}, \|\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*\|^{-1} \} (\|\mathbf{x}\|^2 + \|\mathbf{v}\|^2). \end{aligned} \quad (4.52)$$

Hence,  $\mathbf{V}$  is a strongly positive self-adjoint operator.

(ii) Let  $\mathbf{z} = (\mathbf{x}, \mathbf{v}) \in \mathbf{K}$  and  $\mathbf{z}' = (\mathbf{x}', \mathbf{v}') \in \mathbf{K}$ . We have

$$\begin{aligned} &\|\mathbf{R}\mathbf{z} - \mathbf{R}\mathbf{z}'\|_{\mathbf{V}}^2 \\ &= \langle \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}' \mid \mathbf{W}(\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}') \rangle + \langle \mathbf{D}^{-1}\mathbf{v} - \mathbf{D}^{-1}\mathbf{v}' \mid (\mathbf{U}^{-1} - \mathbf{LW}\mathbf{L}^*)^{-1}(\mathbf{D}^{-1}\mathbf{v} - \mathbf{D}^{-1}\mathbf{v}') \rangle \\ &\leq \|\mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}'\|_{\mathbf{W}}^2 + \|(\mathbf{Id} - \mathbf{U}^{1/2}\mathbf{LW}\mathbf{L}^*\mathbf{U}^{1/2})^{-1}\| \|\mathbf{D}^{-1}\mathbf{v} - \mathbf{D}^{-1}\mathbf{v}'\|_{\mathbf{U}}^2 \\ &\leq \mu^{-1}\langle \mathbf{x} - \mathbf{x}' \mid \mathbf{C}\mathbf{x} - \mathbf{C}\mathbf{x}' \rangle + \nu^{-1}(1 - \|\mathbf{U}^{1/2}\mathbf{LW}^{1/2}\|^2)^{-1}\langle \mathbf{v} - \mathbf{v}' \mid \mathbf{D}^{-1}\mathbf{v} - \mathbf{D}^{-1}\mathbf{v}' \rangle \\ &\leq \max \{ \mu^{-1}, \nu^{-1}(1 - \|\mathbf{U}^{1/2}\mathbf{LW}^{1/2}\|^2)^{-1} \} \langle \mathbf{z} - \mathbf{z}' \mid \mathbf{R}\mathbf{z} - \mathbf{R}\mathbf{z}' \rangle, \end{aligned} \quad (4.53)$$

which, in view of the remark made at the beginning of the proof of Lemma 4.3(ii), shows that  $\mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$  is  $\vartheta$ -cocoercive.  $\square$

**Lemma 4.10** *Let  $\mathbf{B}: \mathbf{G} \rightarrow 2^{\mathbf{G}}$ ,  $\mathbf{C}: \mathbf{H} \rightarrow \mathbf{H}$ ,  $\mathbf{D}: \mathbf{G} \rightarrow 2^{\mathbf{G}}$ ,  $\mathbf{Q}: \mathbf{K} \rightarrow 2^{\mathbf{K}}$ , and  $\mathbf{R}: \mathbf{K} \rightarrow \mathbf{K}$ . Assume that the operator  $\mathbf{A}$  defined in Proposition 4.2 is zero. Let  $\mathbf{W} \in \mathcal{B}(\mathbf{H})$  and  $\mathbf{U} \in \mathcal{B}(\mathbf{G})$  be two strongly*

positive self-adjoint operators such that  $\|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\| < 1$ . Let  $\mathbf{V} \in \mathcal{B}(\mathbf{K})$  be defined by (4.50). For every  $\mathbf{z} = (\mathbf{x}, \mathbf{v}) \in \mathbf{K}$  and  $(\mathbf{e}_1, \mathbf{e}_2) \in \mathbf{K}$ , let

$$\begin{cases} \mathbf{u} = \mathbf{J}_{\mathbf{U}\mathbf{B}^{-1}}(\mathbf{v} + \mathbf{U}(\mathbf{L}(\mathbf{x} - \mathbf{W}(\mathbf{C}\mathbf{x} + \mathbf{L}^*\mathbf{v})) - \mathbf{D}^{-1}\mathbf{v} + \mathbf{e}_2)) \\ \mathbf{y} = \mathbf{x} - \mathbf{W}(\mathbf{C}\mathbf{x} + \mathbf{L}^*\mathbf{u} + \mathbf{e}_1). \end{cases} \quad (4.54)$$

Then,  $(\mathbf{y}, \mathbf{u}) = \mathbf{J}_{\mathbf{V}\mathbf{Q}}(\mathbf{z} - \mathbf{V}\mathbf{R}\mathbf{z} + \mathbf{s})$  where

$$\mathbf{s} = (-\mathbf{W}\mathbf{e}_1, (\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)^{-1}(\mathbf{e}_2 + \mathbf{L}\mathbf{W}\mathbf{e}_1)). \quad (4.55)$$

*Proof.* Let  $\mathbf{z} = (\mathbf{x}, \mathbf{v}) \in \mathbf{K}$  and let  $\mathbf{s} = (\mathbf{e}'_1, \mathbf{e}'_2) \in \mathbf{K}$ . The following equivalences are obtained:

$$\begin{aligned} & (\mathbf{y}, \mathbf{u}) = \mathbf{J}_{\mathbf{V}\mathbf{Q}}(\mathbf{z} - \mathbf{V}\mathbf{R}\mathbf{z} + \mathbf{s}) \\ \Leftrightarrow & \mathbf{V}^{-1}(\mathbf{z} + \mathbf{s} - (\mathbf{y}, \mathbf{u})) - \mathbf{R}\mathbf{z} \in \mathbf{Q}(\mathbf{y}, \mathbf{u}) \\ \Leftrightarrow & \begin{cases} \mathbf{W}^{-1}(\mathbf{x} - \mathbf{y} + \mathbf{e}'_1) - \mathbf{L}^*\mathbf{u} - \mathbf{C}\mathbf{x} = \mathbf{0} \\ (\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)(\mathbf{v} - \mathbf{u} + \mathbf{e}'_2) + \mathbf{L}\mathbf{y} - \mathbf{D}^{-1}\mathbf{v} \in \mathbf{B}^{-1}\mathbf{u} \end{cases} \\ \Leftrightarrow & \begin{cases} \mathbf{y} = \mathbf{x} - \mathbf{W}(\mathbf{C}\mathbf{x} + \mathbf{L}^*\mathbf{u}) + \mathbf{e}'_1 \\ \mathbf{v} + \mathbf{e}'_2 + \mathbf{U}(\mathbf{L}(\mathbf{x} - \mathbf{W}(\mathbf{C}\mathbf{x} + \mathbf{L}^*\mathbf{v} + \mathbf{L}^*\mathbf{e}'_2) + \mathbf{e}'_1) - \mathbf{D}^{-1}\mathbf{v}) \in (\mathbf{Id} + \mathbf{U}\mathbf{B}^{-1})\mathbf{u} \end{cases} \\ \Leftrightarrow & \begin{cases} \mathbf{u} = \mathbf{J}_{\mathbf{U}\mathbf{B}^{-1}}(\mathbf{v} + \mathbf{e}'_2 + \mathbf{U}(\mathbf{L}(\mathbf{x} - \mathbf{W}(\mathbf{C}\mathbf{x} + \mathbf{L}^*\mathbf{v} + \mathbf{L}^*\mathbf{e}'_2) + \mathbf{e}'_1) - \mathbf{D}^{-1}\mathbf{v})) \\ \mathbf{y} = \mathbf{x} - \mathbf{W}(\mathbf{C}\mathbf{x} + \mathbf{L}^*\mathbf{u}) + \mathbf{e}'_1, \end{cases} \end{aligned} \quad (4.56)$$

which lead to (4.54) provided that

$$\begin{cases} -\mathbf{W}\mathbf{e}_1 = \mathbf{e}'_1 \\ \mathbf{U}\mathbf{e}_2 = \mathbf{U}\mathbf{L}(\mathbf{e}'_1 - \mathbf{W}\mathbf{L}^*\mathbf{e}'_2) + \mathbf{e}'_2. \end{cases} \quad (4.57)$$

Since  $\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*$  is an isomorphism, the latter equalities are equivalent to (4.55).  $\square$

From the above two lemmas, a second type of block-coordinate primal-dual algorithm can be deduced to solve Problem 4.1 in the case when  $\mathbf{A} = \mathbf{0}$ .

**Proposition 4.11** *Let  $\mathbf{W}$ ,  $\mathbf{U}$ ,  $\mu$ , and  $\nu$  be defined as in Proposition 4.6. Suppose that*

$$\min \{ \mu, \nu(1 - \|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\|^2) \} > \frac{1}{2}. \quad (4.58)$$

*Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1]$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$  and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}$ -valued random variables, let  $\mathbf{v}_0$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically*

distributed  $\mathbb{D}_{p+q}$ -valued random variables. Iterate

$$\begin{array}{l}
\text{for } n = 0, 1, \dots \\
\left[ \begin{array}{l}
\text{for } j = 1, \dots, p \\
\left[ \begin{array}{l}
\eta_{j,n} = \max \{ \varepsilon_{p+k,n} \mid k \in \mathbb{L}_j^* \} \\
w_{j,n} = \eta_{j,n} (x_{j,n} - \mathbf{W}_j (\mathbf{C}_j x_{j,n} + c_{j,n}))
\end{array} \right. \\
\text{for } k = 1, \dots, q \\
\left[ \begin{array}{l}
u_{k,n} = \varepsilon_{p+k,n} \left( \mathbf{J}_{\mathbf{U}_k \mathbf{B}_k^{-1}} (v_{k,n} + \mathbf{U}_k \left( \sum_{j \in \mathbb{L}_k} \mathbf{L}_{k,j} (w_{j,n} - \mathbf{W}_j \sum_{k' \in \mathbb{L}_j^*} \mathbf{L}_{k',j}^* v_{k',n}) - \mathbf{D}_k^{-1} v_{k,n} + d_{k,n} \right)) + b_{k,n} \right) \\
v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n})
\end{array} \right. \\
\text{for } j = 1, \dots, p \\
\left[ \begin{array}{l}
x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} \left( w_{j,n} - \mathbf{W}_j \sum_{k \in \mathbb{L}_j^*} \mathbf{L}_{k,j}^* u_{k,n} - x_{j,n} \right),
\end{array} \right.
\end{array} \right.
\end{array} \tag{4.59}$$

and set  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ . In addition, assume that

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

and Conditions (ii)-(iii) in Proposition 4.8 hold.

If, in Problem 4.1,  $(\forall j \in \{1, \dots, p\}) A_j = 0$ , then  $(\mathbf{x}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to an  $\mathbf{F}$ -valued random variable, and  $(\mathbf{v}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to an  $\mathbf{F}^*$ -valued random variable.

*Proof.* First note that, in view of Condition (iii) in Proposition 4.8 (since  $(\forall j \in \{1, \dots, p\}) \mathbb{L}_j^* \neq \emptyset$ ), Iterations (4.59) are equivalent to

$$\begin{array}{l}
\text{for } n = 0, 1, \dots \\
\left[ \begin{array}{l}
\text{for } k = 1, \dots, q \\
\left[ \begin{array}{l}
\zeta_{k,n} = \max \{ \varepsilon_{p+k,n}, (\varepsilon_{j,n})_{j \in \mathbb{L}_k} \} \\
\text{for } j = 1, \dots, p \\
\left[ \begin{array}{l}
\eta_{j,n} = \max \{ \varepsilon_{j,n}, (\zeta_{k,n})_{k \in \mathbb{L}_j^*} \} \\
w_{j,n} = \eta_{j,n} (x_{j,n} - \mathbf{W}_j (\mathbf{C}_j x_{j,n} + c_{j,n}))
\end{array} \right. \\
\text{for } k = 1, \dots, q \\
\left[ \begin{array}{l}
u_{k,n} = \zeta_{k,n} \left( \mathbf{J}_{\mathbf{U}_k \mathbf{B}_k^{-1}} (v_{k,n} + \mathbf{U}_k \left( \sum_{j \in \mathbb{L}_k} \mathbf{L}_{k,j} (w_{j,n} - \mathbf{W}_j \sum_{k' \in \mathbb{L}_j^*} \mathbf{L}_{k',j}^* v_{k',n}) - \mathbf{D}_k^{-1} v_{k,n} + d_{k,n} \right)) + b_{k,n} \right) \\
v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n})
\end{array} \right. \\
\text{for } j = 1, \dots, p \\
\left[ \begin{array}{l}
x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} \left( w_{j,n} - \mathbf{W}_j \sum_{k \in \mathbb{L}_j^*} \mathbf{L}_{k,j}^* u_{k,n} - x_{j,n} \right).
\end{array} \right.
\end{array} \right.
\end{array} \right.
\end{array} \tag{4.60}$$

Furthermore, Condition (4.58) implies that  $\|\mathbf{U}^{1/2}\mathbf{L}\mathbf{W}^{1/2}\| < 1$ . Hence, Lemma 4.10 allows us to show the equivalence between Algorithms (4.60) and (3.1) when  $\mathbf{V}$  is given by (4.50),  $\mathbf{Q}$  is given by (4.5) (with  $\mathbf{A} = \mathbf{0}$ ), and  $\mathbf{R}$  is given by (4.6), provided that, for every  $n \in \mathbb{N}$ , (4.33)-(4.37) hold and

$$\mathbf{t}_n = (\mathbf{0}, \mathbf{b}_n), \quad (4.61)$$

$$\mathbf{s}_n = \left( -\mathbf{W}\mathbf{e}_{1,n}, (\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)^{-1}(\mathbf{L}\mathbf{W}\mathbf{L}^*\mathbf{b}_n + \mathbf{d}_n) \right), \quad (4.62)$$

$$\mathbf{e}_{1,n} = \mathbf{L}^*\mathbf{b}_n + \mathbf{c}_n. \quad (4.63)$$

In the proof of Proposition 4.6, we have seen that  $\mathbf{W}^{1/2}\mathbf{C}\mathbf{W}^{1/2}$  is  $\mu$ -cocoercive and  $\mathbf{U}^{1/2}\mathbf{D}^{-1}\mathbf{U}^{1/2}$  is  $\nu$ -cocoercive. According to Lemma 4.9(ii),  $\mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$  is thus  $\vartheta$ -cocoercive where  $\vartheta$  is given by (4.51), and (4.58) means that  $1 = \sup_{n \in \mathbb{N}} \gamma_n < 2\vartheta$ . In addition,

$$\begin{aligned} \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{t}_n\|^2 | \mathcal{X}_n)} &= \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{b}_n\|^2 | \mathcal{X}_n)} < +\infty, \quad (4.64) \\ \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{s}_n\|^2 | \mathcal{X}_n)} &\leq \|\mathbf{W}\mathbf{L}^*\| \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{b}_n\|^2 | \mathcal{X}_n)} + \|\mathbf{W}\| \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{c}_n\|^2 | \mathcal{X}_n)} \\ &+ \|(\mathbf{U}^{-1} - \mathbf{L}\mathbf{W}\mathbf{L}^*)^{-1}\| \left( \|\mathbf{L}\mathbf{W}\mathbf{L}^*\| \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{b}_n\|^2 | \mathcal{X}_n)} + \sum_{n \in \mathbb{N}} \sqrt{\mathbb{E}(\|\mathbf{d}_n\|^2 | \mathcal{X}_n)} \right) < +\infty. \quad (4.65) \end{aligned}$$

Since we have assumed that, for every  $k \in \{1, \dots, q\}$ ,  $\mathbb{L}_k \neq \emptyset$ , Conditions (ii)-(iii) in Proposition 4.8 guarantee that Condition (ii) in Proposition 3.1 is satisfied. The convergence result then follows from this proposition.  $\square$

**Remark 4.12** For every  $j \in \{1, \dots, p\}$ , let a cocoercivity constant of  $\mathbf{C}_j$  be denoted by  $\tilde{\mu}_j \in ]0, +\infty[$  and, for every  $k \in \{1, \dots, q\}$ , let a strong monotonicity constant of  $\mathbf{D}_k$  be denoted by  $\tilde{\nu}_k \in ]0, +\infty[$ . Using (4.45)-(4.46), a necessary condition for (4.58) to be satisfied is

$$\min \left\{ (\|\mathbf{W}_j\|^{-1}\tilde{\mu}_j)_{1 \leq j \leq p}, \left( 1 - \sum_{j=1}^p \sum_{k=1}^q \|\mathbf{U}_k^{1/2}\mathbf{L}_{k,j}\mathbf{W}_j^{1/2}\|^2 \right) (\|\mathbf{U}_k\|^{-1}\tilde{\nu}_k)_{1 \leq k \leq q} \right\} > \frac{1}{2}. \quad (4.66)$$

In the case when, for every  $k \in \{1, \dots, q\}$ ,  $\mathbf{D}_k^{-1} = 0$ , the constants  $(\tilde{\nu}_k)_{1 \leq k \leq q}$  can be chosen arbitrarily large and the above condition reduces to

$$\begin{cases} \sum_{j=1}^p \sum_{k=1}^q \|\mathbf{U}_k^{1/2}\mathbf{L}_{k,j}\mathbf{W}_j^{1/2}\|^2 < 1 \\ \min \{ (\|\mathbf{W}_j\|^{-1}\tilde{\mu}_j)_{1 \leq j \leq p} \} > \frac{1}{2}. \end{cases} \quad (4.67)$$

This condition appears to be less restrictive than (4.48).

## 5 Block-coordinate primal-dual proximal algorithms for convex optimization problems

As we will show next, the results obtained in the previous section allow us to deduce a couple of novel primal-dual proximal splitting algorithms for solving a variety of (possibly nonsmooth) convex optimization problems. More precisely, we will turn our attention to the following class of optimization problems, the notation of the previous section being still in force:

**Problem 5.1** For every  $j \in \{1, \dots, p\}$ , let  $f_j \in \Gamma_0(H_j)$ , let  $h_j \in \Gamma_0(H_j)$  be Lipschitz-differentiable, and, for every  $k \in \{1, \dots, q\}$ , let  $g_k \in \Gamma_0(G_k)$ , let  $l_k \in \Gamma_0(G_k)$  be strongly convex, and let  $L_{k,j} \in \mathcal{B}(H_j, G_k)$ . Suppose that (4.1) and (4.2) hold, and that there exists  $(\bar{x}_1, \dots, \bar{x}_p) \in H_1 \oplus \dots \oplus H_p$  such that

$$(\forall j \in \{1, \dots, p\}) \quad 0 \in \partial f_j(\bar{x}_j) + \nabla h_j(\bar{x}_j) + \sum_{k=1}^q L_{k,j}^* (\partial g_k \square \partial l_k) \left( \sum_{j'=1}^p L_{k,j'} \bar{x}_{j'} \right). \quad (5.1)$$

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

$$\underset{x_1 \in H_1, \dots, x_p \in H_p}{\text{minimize}} \quad \sum_{j=1}^p (f_j(x_j) + h_j(x_j)) + \sum_{k=1}^q (g_k \square l_k) \left( \sum_{j=1}^p L_{k,j} x_j \right) \quad (5.2)$$

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

$$\underset{v_1 \in G_1, \dots, v_q \in G_q}{\text{minimize}} \quad \sum_{j=1}^p (f_j^* \square h_j^*) \left( - \sum_{k=1}^q L_{k,j}^* v_k \right) + \sum_{k=1}^q (g_k^*(v_k) + l_k^*(v_k)). \quad (5.3)$$

Our objective is to find a pair  $(\hat{x}, \hat{v})$  of random variables such that  $\hat{x}$  is  $\tilde{\mathbf{F}}$ -valued and  $\hat{v}$  is  $\tilde{\mathbf{F}}^*$ -valued.

Note that the inclusion condition in Problem 5.1 is satisfied under a number of relatively weak assumptions:

**Proposition 5.2** [17, Proposition 5.3] *Consider the setting of Problem 5.1. Suppose that (5.2) has a solution. Then, the existence of  $(\bar{x}_1, \dots, \bar{x}_p) \in H_1 \oplus \dots \oplus H_p$  satisfying (5.1) is guaranteed in each of the following cases:*

- (i) *For every  $j \in \{1, \dots, p\}$ ,  $f_j$  is real-valued and, for every  $k \in \{1, \dots, q\}$ ,  $(x_j)_{1 \leq j \leq p} \mapsto \sum_{j=1}^p L_{k,j} x_j$  is surjective.*
- (ii) *For every  $k \in \{1, \dots, q\}$ ,  $g_k$  or  $l_k$  is real-valued.*
- (iii)  *$(H_j)_{1 \leq j \leq p}$  and  $(G_k)_{1 \leq k \leq q}$  are finite-dimensional, and  $(\forall j \in \{1, \dots, p\}) (\exists x_j \in \text{ri dom } f_j)$  such that  $(\forall k \in \{1, \dots, q\}) \sum_{j=1}^p L_{k,j} x_j \in \text{ri dom } g_k + \text{ri dom } l_k$ .*

The following result can be deduced from Proposition 4.6:

**Proposition 5.3** *Let  $\mathbf{W}$  and  $\mathbf{U}$  be defined as in Proposition 4.6. For every  $j \in \{1, \dots, p\}$ , let  $\mu_j^{-1} \in ]0, +\infty[$  be a Lipschitz constant of the gradient of  $h_j \circ W_j^{1/2}$  and, for every  $k \in \{1, \dots, q\}$ , let  $\nu_k^{-1} \in ]0, +\infty[$  be a Lipschitz constant of the gradient of  $l_k^* \circ U_k^{1/2}$ . Suppose that (4.30) holds where  $\vartheta_\alpha$  is defined by (4.9),  $\mu = \min\{\mu_1, \dots, \mu_p\}$ , and  $\nu = \min\{\nu_1, \dots, \nu_q\}$ . Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1]$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$ ,  $(\mathbf{a}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}$ -valued random variables, let  $\mathbf{v}_0$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_{p+q}$ -valued random variables. Iterate*

$$\begin{array}{l}
 \text{for } n = 0, 1, \dots \\
 \left[ \begin{array}{l}
 \text{for } j = 1, \dots, p \\
 \left[ \begin{array}{l}
 y_{j,n} = \varepsilon_{j,n} \left( \text{prox}_{\mathbf{f}_j^j}^{W_j^{-1}} \left( x_{j,n} - W_j \left( \sum_{k \in \mathbb{L}_j^*} L_{k,j}^* v_{k,n} + \nabla h_j(x_{j,n}) + c_{j,n} \right) \right) + a_{j,n} \right) \\
 x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} (y_{j,n} - x_{j,n})
 \end{array} \right. \\
 \text{for } k = 1, \dots, q \\
 \left[ \begin{array}{l}
 u_{k,n} = \varepsilon_{p+k,n} \left( \text{prox}_{\mathbf{g}_k^k}^{U_k^{-1}} \left( v_{k,n} + U_k \left( \sum_{j \in \mathbb{L}_k} L_{k,j} (2y_{j,n} - x_{j,n}) - \nabla l_k^*(v_{k,n}) + d_{k,n} \right) \right) + b_{k,n} \right) \\
 v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n}).
 \end{array} \right.
 \end{array} \right.
 \end{array} \tag{5.4}$$

In addition, assume that Conditions (i)-(iii) in Proposition 4.6 hold, where  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ .

Then,  $(\mathbf{x}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\tilde{\mathbf{F}}$ -valued random variable, and  $(\mathbf{v}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\tilde{\mathbf{F}}^*$ -valued random variable.

*Proof.* Let us set, for every  $j \in \{1, \dots, p\}$ ,  $A_j = \partial f_j$ ,  $C_j = \nabla h_j$  and, for every  $k \in \{1, \dots, q\}$ ,  $B_k = \partial g_k$ , and  $D_k^{-1} = \nabla l_k^*$ . Then, it can be noticed that, for every  $j \in \{1, \dots, p\}$  and  $k \in \{1, \dots, q\}$ ,  $J_{W_j A_j} = \text{prox}_{\mathbf{f}_j^j}^{W_j^{-1}}$ ,  $J_{U_k B_k^{-1}} = \text{prox}_{\mathbf{g}_k^k}^{U_k^{-1}}$ , and that the Lipschitz-differentiability assumptions made on  $h_j$  and  $l_k^*$  are equivalent to the fact that  $W_j^{1/2} C_j W_j^{1/2}$  is  $\mu_j$ -cocoercive and  $U_k^{1/2} D_k^{-1} U_k^{1/2}$  is  $\nu_k$ -cocoercive [5, Corollaries 16.42 & 18.16]. Proposition 4.6 thus allows us to assert that  $(\mathbf{x}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to an  $\mathbf{F}$ -valued random variable, and  $(\mathbf{v}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to an  $\mathbf{F}^*$ -valued random variable, where  $\mathbf{F}$  and  $\mathbf{F}^*$  have been defined in Problem 4.1. Let us now show that the first limit is a  $\tilde{\mathbf{F}}$ -valued random variable, and the second one is a  $\tilde{\mathbf{F}}^*$ -valued random variable. Define the separable functions  $\mathbf{f} \in \Gamma_0(\mathbf{H})$ ,  $\mathbf{h} \in \Gamma_0(\mathbf{H})$ ,  $\mathbf{g} \in \Gamma_0(\mathbf{G})$ , and  $\mathbf{l} \in \Gamma_0(\mathbf{G})$  as

$$\mathbf{f}: \mathbf{x} \mapsto \sum_{j=1}^p f_j(x_j), \quad \mathbf{h}: \mathbf{x} \mapsto \sum_{j=1}^p h_j(x_j), \tag{5.5}$$

$$\mathbf{g}: \mathbf{v} \mapsto \sum_{k=1}^q g_k(v_k), \quad \mathbf{l}: \mathbf{v} \mapsto \sum_{k=1}^q l_k(v_k). \tag{5.6}$$

According to [5, Proposition 16.8], (5.1) can be reexpressed more concisely as

$$\mathbf{0} \in \partial \mathbf{f}(\bar{\mathbf{x}}) + \nabla \mathbf{h}(\bar{\mathbf{x}}) + \mathbf{L}^*(\partial \mathbf{g} \square \partial \mathbf{l})(\mathbf{L} \bar{\mathbf{x}}). \tag{5.7}$$

Since  $\text{dom } \mathbf{h} = \mathbf{H}$ ,  $\partial \mathbf{f} + \nabla \mathbf{h} = \partial(\mathbf{f} + \mathbf{h})$  [5, Propositions 16.38 & 17.26] and since  $\text{dom } \mathbf{l}^* = \mathbf{G}$ ,  $\partial \mathbf{g} \square \partial \mathbf{l} = \partial(\mathbf{g} \square \mathbf{l})$  [5, Proposition 24.27]. Equation (5.7) implies that  $\mathbf{L}(\text{dom}(\mathbf{f} + \mathbf{h})) \cap \text{dom}(\mathbf{g} \square \mathbf{l}) \neq \emptyset$  [5, Proposition 16.3(i)] and it follows from [5, Proposition 16.5] that

$$(\forall \mathbf{x} \in \mathbf{H}) \quad \partial \mathbf{f}(\mathbf{x}) + \nabla \mathbf{h}(\mathbf{x}) + \mathbf{L}^*(\partial \mathbf{g} \square \partial \mathbf{l})(\mathbf{L}\mathbf{x}) \subset \partial(\mathbf{f} + \mathbf{h} + (\mathbf{g} \square \mathbf{l}) \circ \mathbf{L})(\mathbf{x}). \quad (5.8)$$

As a consequence of (4.3) and Fermat's rule [5, Theorem 16.2], this allows us to conclude that

$$\mathbf{F} = \text{zer}(\partial \mathbf{f} + \nabla \mathbf{h} + \mathbf{L}^*(\partial \mathbf{g} \square \partial \mathbf{l})\mathbf{L}) \subset \text{zer}(\partial(\mathbf{f} + \mathbf{h} + (\mathbf{g} \square \mathbf{l}) \circ \mathbf{L})) = \tilde{\mathbf{F}}. \quad (5.9)$$

By a similar argument, the fact that  $\mathbf{F}^* = \text{zer}(-\mathbf{L}(\partial \mathbf{f}^* \square \partial \mathbf{h}^*)(-\mathbf{L}^*) + \partial \mathbf{g}^* + \nabla \mathbf{l}^*) \neq \emptyset$  allows us to show that  $\mathbf{F}^* \subset \tilde{\mathbf{F}}^*$ .  $\square$

In a quite similar way, Proposition 4.11 leads to the following result.

**Proposition 5.4** *Let  $\mathbf{W}$  and  $\mathbf{U}$  be defined as in Proposition 4.6. Let  $\mu$  and  $\nu$  be defined as in Proposition 5.3. Suppose that Condition (4.58) holds. Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1]$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$  and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}$ -valued random variables, let  $\mathbf{v}_0$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_{p+q}$ -valued random variables. Iterate*

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \left[ \begin{array}{l} \text{for } j = 1, \dots, p \\ \left[ \begin{array}{l} \eta_{j,n} = \max \{ \varepsilon_{p+k,n} \mid k \in \mathbb{L}_j^* \} \\ w_{j,n} = \eta_{j,n} (x_{j,n} - \mathbf{W}_j(\nabla \mathbf{h}_j(x_{j,n}) + \mathbf{c}_{j,n})) \end{array} \right. \\ \text{for } k = 1, \dots, q \\ \left[ \begin{array}{l} u_{k,n} = \varepsilon_{p+k,n} \left( \text{prox}_{\mathbf{g}_k^*}^{\mathbf{U}_k^{-1}} \left( v_{k,n} + \mathbf{U}_k \left( \sum_{j \in \mathbb{L}_k} \mathbf{L}_{k,j} (w_{j,n} - \mathbf{W}_j \sum_{k' \in \mathbb{L}_j^*} \mathbf{L}_{k',j}^* v_{k',n}) - \nabla \mathbf{l}_k^*(v_{k,n}) + \mathbf{d}_{k,n} \right) \right) + \mathbf{b}_{k,n} \right) \\ v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n}) \end{array} \right. \\ \text{for } j = 1, \dots, p \\ \left[ \begin{array}{l} x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} \left( w_{j,n} - \mathbf{W}_j \sum_{k \in \mathbb{L}_j^*} \mathbf{L}_{k,j}^* u_{k,n} - x_{j,n} \right). \end{array} \right. \end{array} \right. \end{array} \quad (5.10)$$

In addition, assume that Conditions (i) in Proposition 4.11, and (ii)-(iii) in Proposition 4.8 hold, where  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ .

If, in Problem 5.1,  $(\forall j \in \{1, \dots, p\}) f_j = 0$ , then  $(\mathbf{x}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\tilde{\mathbf{F}}$ -valued random variable, and  $(\mathbf{v}_n)_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\tilde{\mathbf{F}}^*$ -valued random variable.

At this point, it may appear interesting to examine the connections existing between the two proposed block-coordinate proximal algorithms and published works.

**Remark 5.5**

(i) In practice, one may be interested in problems of the form

$$\underset{x_1 \in H_1, \dots, x_p \in H_p}{\text{minimize}} \sum_{j=1}^p (f_j(x_j) + h_j(x_j)) + \sum_{k=1}^q g_k \left( \sum_{j=1}^p L_{k,j} x_j \right). \quad (5.11)$$

These are special cases of (5.2) where  $(\forall k \in \{1, \dots, q\}) l_k = \iota_{\{0\}}$ , i.e.  $l_k^* = 0$ .

- (ii) Algorithm (5.4) extends the deterministic approaches in [13, 26, 28, 32, 54], which deal with the case when  $p = 1$ , by introducing some random sweeping of the coordinates and by allowing the use of stochastic errors. Similarly, Algorithm (5.10) extends the algorithms in [16, 37] which were developed in a deterministic setting in the absence of errors, in the scenario where  $p = q = 1$ ,  $H_1$  and  $G_1$  are finite dimensional spaces,  $l_1 = \iota_{\{0\}}$ ,  $W_1 = \tau \text{Id}$  with  $\tau \in ]0, +\infty[$ ,  $U_1 = \rho \text{Id}$  with  $\rho \in ]0, +\infty[$ , and no relaxation ( $\lambda_n \equiv 1$ ) or a constant one ( $\lambda_n \equiv \lambda_0 < 1$ ) is performed. Recently, these works have been generalized to possibly infinite-dimensional Hilbert spaces when  $p = 1$  and  $q > 1$ , arbitrary preconditioning operators are employed, and deterministic summable errors are allowed [18]. The practical interest of introducing preconditioning operators for accelerating the convergence of primal-dual proximal methods was emphasized in [18, 45, 48].
- (iii) In [23, Corollary 5.5], another random block-coordinate primal-dual algorithm was proposed to solve an instance of Problem 5.1 obtained when  $(\forall j \in \{1, \dots, p\}) h_j = 0$  and  $(\forall k \in \{1, \dots, q\}) l_k = \iota_{\{0\}}$ . This algorithm is based on the Douglas-Rachford iteration which is also at the origin of the randomized Alternating Direction Method of Multipliers (ADMM) developed in finite dimensional spaces in [34]. Note however that the algorithm in [23, Corollary 5.5] requires to invert  $\text{Id} + \mathbf{L}\mathbf{L}^*$  or  $\text{Id} + \mathbf{L}^*\mathbf{L}$  (see [23, Remark 5.4]). By contrast, Algorithms (5.4) and (5.10) do not make it necessary to perform any linear operator inversion.

## 6 Asynchronous distributed algorithms

In this part,  $H, G_1, \dots, G_m$  are separable real Hilbert spaces,  $\mathbf{G} = G_1 \oplus \dots \oplus G_m$ , and the following problem is addressed:

**Problem 6.1** For every  $i \in \{1, \dots, m\}$ , let  $A_i: H \rightarrow 2^H$  be maximally monotone, let  $C_i: H \rightarrow H$  be cocoercive, let  $B_i: G_i \rightarrow 2^{G_i}$  be maximally monotone, let  $D_i: G_i \rightarrow 2^{G_i}$  be maximally monotone and strongly monotone, and let  $M_i$  be a nonzero operator in  $\mathcal{B}(H, G_i)$ . We assume that the set  $\widehat{F}$  of solutions to the problem:



$$\text{find } \mathbf{x} \in \mathbf{H} \text{ such that } 0 \in \sum_{i=1}^m \mathbf{A}_i \mathbf{x} + \mathbf{C}_i \mathbf{x} + \mathbf{M}_i^* (\mathbf{B}_i \square \mathbf{D}_i) (\mathbf{M}_i \mathbf{x}) \quad (6.1)$$

is nonempty. Our objective is to find a  $\widehat{\mathbf{F}}$ -valued random variable  $\widehat{\mathbf{x}}$ .

Problem (6.1) can be reformulated in the product space  $\mathbf{H}^m$  as

$$\text{find } (\mathbf{x}_1, \dots, \mathbf{x}_m) \in \Lambda_m \text{ such that } 0 \in \sum_{i=1}^m \mathbf{A}_i \mathbf{x}_i + \mathbf{C}_i \mathbf{x}_i + \mathbf{M}_i^* (\mathbf{B}_i \square \mathbf{D}_i) (\mathbf{M}_i \mathbf{x}_i) \quad (6.2)$$

where

$$\Lambda_m = \{(\mathbf{x}_1, \dots, \mathbf{x}_m) \in \mathbf{H}^m \mid \mathbf{x}_1 = \dots = \mathbf{x}_m\}. \quad (6.3)$$

This kind of reformulation was employed in [20, 44] to obtain parallel algorithms for finding a zero of a sum of maximal operators and it is also popular in consensus problems [10, 42]. To devise distributed algorithms, the involved linear constraint is further split in a set of similar constraints, each of them involving a reduced subset of variables. In this context indeed, each index  $i \in \{1, \dots, m\}$  corresponds to a given agent and a modeling of the topological relationships existing between the different agents is needed. To do so, we define nonempty subsets  $(\mathbb{V}_\ell)_{1 \leq \ell \leq r}$  of  $\{1, \dots, m\}$ , with cardinalities  $(\kappa_\ell)_{1 \leq \ell \leq r}$ , which are such that:

**Assumption 6.2** For every  $\mathbf{x} = (\mathbf{x}_i)_{1 \leq i \leq m} \in \mathbf{H}^m$ ,

$$\mathbf{x} \in \Lambda_m \quad \Leftrightarrow \quad (\forall \ell \in \{1, \dots, r\}) \quad (\mathbf{x}_i)_{i \in \mathbb{V}_\ell} \in \Lambda_{\kappa_\ell}. \quad (6.4)$$

This assumption is obviously satisfied if  $r = 1$  and  $\mathbb{V}_1 = \{1, \dots, m\}$ , or if  $r = m - 1$  and  $(\forall \ell \in \{1, \dots, m - 1\}) \mathbb{V}_\ell = \{\ell, \ell + 1\}$ . More generally if the sets  $(\mathbb{V}_\ell)_{1 \leq \ell \leq r}$  correspond to the hyperedges of a hypergraph with vertices  $\{1, \dots, m\}$ , then the assumption is equivalent to the fact that the hypergraph is connected.

In the following, we will need to introduce the notation:

$$\mathbf{H} = \mathbf{H}^{\kappa_1} \oplus \dots \oplus \mathbf{H}^{\kappa_r}, \quad \mathbf{\Lambda} = \Lambda_{\kappa_1} \oplus \dots \oplus \Lambda_{\kappa_r}, \quad (6.5)$$

$$\mathbf{A} = \times_{i=1}^m \mathbf{A}_i, \quad \mathbf{C} = \times_{i=1}^m \mathbf{C}_i, \quad (6.6)$$

$$\mathbf{B} = \times_{i=1}^m \mathbf{B}_i, \quad \mathbf{D} = \times_{i=1}^m \mathbf{D}_i, \quad (6.7)$$

$$\mathbf{S}: \mathbf{H}^m \rightarrow \mathbf{H}: \mathbf{x} \mapsto (\mathbf{S}_\ell \mathbf{x})_{1 \leq \ell \leq r}, \quad \mathbf{M}: \mathbf{H}^m \rightarrow \mathbf{G}: \mathbf{x} \mapsto (\mathbf{M}_i \mathbf{x}_i)_{1 \leq i \leq m}, \quad (6.8)$$

where, for every  $\ell \in \{1, \dots, r\}$ ,

$$\mathbf{S}_\ell: \mathbf{H}^m \rightarrow \mathbf{H}^{\kappa_\ell}: \mathbf{x} \mapsto (\mathbf{x}_i)_{i \in \mathbb{V}_\ell} = (\mathbf{x}_{i(\ell, j)})_{1 \leq j \leq \kappa_\ell} \quad (6.9)$$

and  $i(\ell, 1), \dots, i(\ell, \kappa_\ell)$  denote the elements of  $\mathbb{V}_\ell$  ordered in an increasing manner. Note that, for every  $\ell \in \{1, \dots, r\}$ , the adjoint of  $S_\ell$  is

$$S_\ell^*: \mathbb{H}^{\kappa_\ell} \rightarrow \mathbb{H}^m: \mathbf{z}_\ell = (z_{\ell,j})_{1 \leq j \leq \kappa_\ell} \mapsto (\mathbf{x}_i)_{1 \leq i \leq m} \quad (6.10)$$

where

$$(\forall i \in \{1, \dots, m\}) \quad \mathbf{x}_i = \begin{cases} z_{\ell,j} & \text{if } i = i(\ell, j) \text{ with } j \in \{1, \dots, \kappa_\ell\} \\ 0 & \text{otherwise.} \end{cases} \quad (6.11)$$

The adjoint of  $\mathbf{S}$  is thus given by

$$\mathbf{S}^*: \mathbb{H} \rightarrow \mathbb{H}^m: (\mathbf{z}_\ell)_{1 \leq \ell \leq r} \mapsto \sum_{\ell=1}^r S_\ell^* \mathbf{z}_\ell = (\mathbf{x}_i)_{1 \leq i \leq m} \quad (6.12)$$

where, for every  $i \in \{1, \dots, m\}$ ,

$$\mathbf{x}_i = \sum_{(\ell,j) \in \mathbb{V}_i^*} z_{\ell,j} \quad (6.13)$$

with

$$\mathbb{V}_i^* = \{(\ell, j) \mid \ell \in \{1, \dots, r\}, j \in \{1, \dots, \kappa_\ell\}, \text{ and } i(\ell, j) = i\}. \quad (6.14)$$

As a consequence of Assumption 6.2, the cardinality of  $\mathbb{V}_i^*$  (i.e. the number of sets  $(\mathbb{V}_\ell)_{1 \leq \ell \leq r}$  containing index  $i$ ) is nonzero.

The link between Problems 6.1 and 4.1 is now enlightened by the next result:

**Proposition 6.3** *Under Assumption 6.2, Problem (6.2) is equivalent to*

$$\text{find } \mathbf{x} \in \mathbb{H}^m \text{ such that } \mathbf{0} \in \mathbf{Ax} + \mathbf{Cx} + \mathbf{M}^*(\mathbf{B} \square \mathbf{D})(\mathbf{Mx}) + \mathbf{S}^* \mathbf{N}_\Lambda(\mathbf{Sx}). \quad (6.15)$$

*Proof.* For every  $\mathbf{x} \in \mathbb{H}^m$ , we have the following simple equivalences:

$$\begin{aligned} & \begin{cases} \mathbf{0} \in \sum_{i=1}^m A_i \mathbf{x}_i + C_i \mathbf{x}_i + M_i^*(B_i \square D_i)(M_i \mathbf{x}_i) \\ \mathbf{x} \in \Lambda_m \end{cases} \\ \Leftrightarrow & \begin{cases} (\forall i \in \{1, \dots, m\}) \quad \mathbf{0} \in A_i \mathbf{x}_i + C_i \mathbf{x}_i + M_i^*(B_i \square D_i)(M_i \mathbf{x}_i) + \mathbf{u}_i \\ \mathbf{x} \in \Lambda_m \\ \sum_{i=1}^m \mathbf{u}_i = \mathbf{0} \end{cases} \\ \Leftrightarrow & \begin{cases} \mathbf{0} \in \mathbf{Ax} + \mathbf{Cx} + \mathbf{M}^*(\mathbf{B} \square \mathbf{D})(\mathbf{Mx}) + \mathbf{u} \\ \mathbf{x} \in \Lambda_m \\ \mathbf{u} \in \Lambda_m^\perp \end{cases} \\ \Leftrightarrow & \mathbf{0} \in \mathbf{Ax} + \mathbf{Cx} + \mathbf{M}^*(\mathbf{B} \square \mathbf{D})(\mathbf{Mx}) + \mathbf{S}^* \mathbf{N}_\Lambda(\mathbf{Sx}), \quad (6.16) \end{aligned}$$

where we have used the fact that  $\mathbf{N}_{\Lambda_m} = \partial\iota_{\Lambda_m} = \partial(\iota_{\Lambda} \circ \mathbf{S}) = \mathbf{S}^* \partial\iota_{\Lambda} \mathbf{S} = \mathbf{S}^* \mathbf{N}_{\Lambda} \mathbf{S}$  since  $\Lambda + \text{ran}(\mathbf{S})$  is a closed subspace of  $\mathbf{H}$  [5, Propositions 6.19 & 16.42].  $\square$

**Remark 6.4** If we now reexpress (6.15) in terms of the notation used in Problem 4.1, we see that the equivalence of Problem 6.1 with Problem 4.1 is obtained by setting  $p = m$ ,  $q = m + r$ ,  $\mathbf{H}_1 = \dots = \mathbf{H}_m = \mathbf{H}$ ,  $(\forall \ell \in \{1, \dots, r\}) \mathbf{G}_{m+\ell} = \mathbf{H}^{\kappa_\ell}$ ,  $\mathbf{B}_{m+\ell} = \mathbf{N}_{\Lambda_{\kappa_\ell}}$ ,  $\mathbf{D}_{m+\ell} = \mathbf{N}_{\{0\}}$ , and

$$(\forall (k, i) \in \{1, \dots, m\}^2) \quad \mathbf{L}_{k,i} = \begin{cases} \mathbf{M}_i & \text{if } k = i \\ 0 & \text{otherwise,} \end{cases} \quad (6.17)$$

$$(\forall \ell \in \{1, \dots, r\})(\forall \mathbf{x} \in \mathbf{H}^m) \quad \sum_{i=1}^m \mathbf{L}_{m+\ell,i} \mathbf{x}_i = \mathbf{S}_\ell \mathbf{x} \quad (6.18)$$

(hence, (4.1) and (4.2) are satisfied).

Our goal now is to develop asynchronous distributed algorithms for solving Problem 6.1 in the sense that, at each iteration of these algorithms, a limited number of operators  $(\mathbf{A}_i)_{1 \leq i \leq m}$ ,  $(\mathbf{B}_i)_{1 \leq i \leq m}$ ,  $(\mathbf{C}_i)_{1 \leq i \leq m}$ , and  $(\mathbf{D}_i)_{1 \leq i \leq m}$  are activated in a random manner. Based on the above remark, the following convergence result can be deduced from Proposition 4.8.

**Proposition 6.5** *Let  $(\theta_\ell)_{1 \leq \ell \leq r} \in ]0, +\infty[^r$ . For every  $i \in \{1, \dots, m\}$ , let  $\mathbf{W}_i$  be a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{H})$  such that  $\mathbf{W}_i^{1/2} \mathbf{C}_i \mathbf{W}_i^{1/2}$  is  $\mu_i$ -cocoercive with  $\mu_i \in ]0, +\infty[$ , let  $\mathbf{U}_i$  be a strongly positive self-adjoint operator in  $\mathcal{B}(\mathbf{G}_i)$  such that  $\mathbf{U}_i^{1/2} \mathbf{D}_i^{-1} \mathbf{U}_i^{1/2}$  is  $\nu_i$ -cocoercive with  $\nu_i \in ]0, +\infty[$ , and let*

$$\bar{\theta}_i = \sum_{\ell \in \{\ell' \in \{1, \dots, r\} | i \in \mathbb{V}_{\ell'}\}} \theta_{\ell}. \quad (6.19)$$

Suppose that

$$(\exists \alpha \in ]0, +\infty[) \quad (1 - \chi) \min\{\mu(1 + \alpha\sqrt{\chi})^{-1}, \nu(1 + \alpha^{-1}\sqrt{\chi})^{-1}\} > \frac{1}{2} \quad (6.20)$$

where

$$\chi = \max_{i \in \{1, \dots, m\}} \|\mathbf{U}_i^{1/2} \mathbf{M}_i \mathbf{W}_i^{1/2}\|^2 + \bar{\theta}_i \|\mathbf{W}_i\|, \quad (6.21)$$

$\mu = \min\{\mu_1, \dots, \mu_m\}$ , and  $\nu = \min\{\nu_1, \dots, \nu_m\}$ . Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1]$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$ ,  $(\mathbf{a}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}^m$ -valued random variables, let  $(v_{i,0})_{1 \leq i \leq m}$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, for every  $\ell \in \{1, \dots, r\}$  let  $v_{m+\ell,0} = (v_{m+\ell,j,0})_{1 \leq j \leq \kappa_\ell}$  be a  $\mathbf{H}^{\kappa_\ell}$ -valued random variable, and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_{2m+r}$ -valued random

variables. Set  $(\forall \ell \in \{1, \dots, r\}) \bar{x}_{\ell,0} = \kappa_\ell^{-1} \sum_{i \in \mathbb{V}_\ell} x_{i,0}$ , iterate

$$\begin{array}{l}
\text{for } n = 0, 1, \dots \\
\left[ \begin{array}{l}
\text{for } \ell = 1, \dots, r \\
\left[ \begin{array}{l}
\bar{u}_{\ell,n} = \varepsilon_{2m+\ell,n} \left( \frac{1}{\kappa_\ell} \sum_{j=1}^{\kappa_\ell} v_{m+\ell,j,n} + \theta_\ell \bar{x}_{\ell,n} \right) \\
\text{for } j = 1, \dots, \kappa_\ell \\
\left[ \begin{array}{l}
w_{\ell,j,n} = \varepsilon_{2m+\ell,n} (2(\theta_\ell x_{i(\ell,j),n} - \bar{u}_{\ell,n}) + v_{m+\ell,j,n}) \\
\text{for } i = 1, \dots, m \\
\left[ \begin{array}{l}
u_{i,n} = \varepsilon_{m+i,n} \left( \mathbf{J}_{\mathbf{U}_i \mathbf{B}_i^{-1}} (v_{i,n} + \mathbf{U}_i (\mathbf{M}_i x_{i,n} - \mathbf{D}_i^{-1} v_{i,n} + d_{i,n})) + b_{i,n} \right) \\
y_{i,n} = \varepsilon_{i,n} \left( \mathbf{J}_{\mathbf{W}_i \mathbf{A}_i} (x_{i,n} - \mathbf{W}_i (\mathbf{M}_i^* (2u_{i,n} - v_{i,n}) + \sum_{(\ell,j) \in \mathbb{V}_i^*} w_{\ell,j,n} + \mathbf{C}_i x_{i,n} + c_{i,n})) + a_{i,n} \right) \\
v_{i,n+1} = v_{i,n} + \lambda_n \varepsilon_{m+i,n} (u_{i,n} - v_{i,n}) \\
x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (y_{i,n} - x_{i,n}) \\
\text{for } \ell = 1, \dots, r \\
\left[ \begin{array}{l}
v_{m+\ell,n+1} = v_{m+\ell,n} + \frac{\lambda_n}{2} \varepsilon_{2m+\ell,n} (w_{\ell,n} - v_{m+\ell,n}) \\
\eta_{\ell,n} = \max \{ \varepsilon_{i,n} \mid i \in \mathbb{V}_\ell \} \\
\bar{x}_{\ell,n+1} = \bar{x}_{\ell,n} + \eta_{\ell,n} \left( \frac{1}{\kappa_\ell} \sum_{i \in \mathbb{V}_\ell} x_{i,n+1} - \bar{x}_{\ell,n} \right),
\end{array} \right.
\end{array} \right.
\end{array} \right.
\end{array} \right.
\end{array}
\tag{6.22}$$

and set  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ . In addition, assume that the following hold:

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

(ii) For every  $n \in \mathbb{N}$ ,  $\mathcal{E}_n$  and  $\mathcal{X}_n$  are independent, and  $(\forall i \in \{1, \dots, m\}) \mathbb{P}[\varepsilon_{i,0} = 1] > 0$ .

(iii) For every  $n \in \mathbb{N}$ ,

$$(\forall i \in \{1, \dots, m\}) \quad \{\omega \in \Omega \mid \varepsilon_{i,n}(\omega) = 1\} \subset \{\omega \in \Omega \mid \varepsilon_{m+i,n}(\omega) = 1\}
\tag{6.23}$$

and

$$(\forall \ell \in \{1, \dots, r\}) \quad \bigcup_{i \in \mathbb{V}_\ell} \{\omega \in \Omega \mid \varepsilon_{i,n}(\omega) = 1\} \subset \{\omega \in \Omega \mid \varepsilon_{2m+\ell,n}(\omega) = 1\}.
\tag{6.24}$$

Then, under Assumption 6.2, for every  $i \in \{1, \dots, m\}$ ,  $(x_{i,n})_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\hat{\mathbb{F}}$ -valued random variable  $\hat{x}$  and, for every  $\ell \in \{1, \dots, r\}$ ,  $(\bar{x}_{\ell,n})_{n \in \mathbb{N}}$  converges weakly P-a.s. to  $\hat{x}$ .

*Proof.* By using Proposition 6.3, Remark 6.4, (6.9), (6.12)-(6.13), setting

$$(\forall \ell \in \{1, \dots, r\}) \quad \mathbf{U}_{m+\ell} = \theta_\ell \text{Id} \quad (6.25)$$

$$(\forall n \in \mathbb{N}) \quad b_{m+\ell, n} = d_{m+\ell, n} = 0, \quad (6.26)$$

and noticing that  $\mathbf{J}_{\mathbf{U}_{m+\ell} \mathbf{N}_{\Lambda_{\kappa_\ell}}^{-1}} = \text{Id} - \theta_\ell \Pi_{\Lambda_{\kappa_\ell}}(\cdot/\theta_\ell) = \text{Id} - \Pi_{\Lambda_{\kappa_\ell}}$  (see (2.1)), Algorithm (4.49) for solving Problem (6.2) reads

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \quad \left[ \begin{array}{l} \text{for } i = 1, \dots, m \\ \quad \left[ \begin{array}{l} u_{i,n} = \varepsilon_{m+i,n} \left( \mathbf{J}_{\mathbf{U}_i \mathbf{B}_i^{-1}}(v_{i,n} + \mathbf{U}_i(\mathbf{M}_i x_{i,n} - \mathbf{D}_i^{-1} v_{i,n} + d_{i,n})) + b_{i,n} \right) \\ v_{i,n+1} = v_{i,n} + \lambda_n \varepsilon_{m+i,n} (u_{i,n} - v_{i,n}) \end{array} \right. \\ \text{for } \ell = 1, \dots, r \\ \quad \left[ \begin{array}{l} u_{m+\ell,n} = \varepsilon_{2m+\ell,n} (v_{m+\ell,n} + \theta_\ell (x_{i,n})_{i \in \mathbb{V}_\ell} - \Pi_{\Lambda_{\kappa_\ell}}(v_{m+\ell,n} + \theta_\ell (x_{i,n})_{i \in \mathbb{V}_\ell})) \\ v_{m+\ell,n+1} = v_{m+\ell,n} + \lambda_n \varepsilon_{2m+\ell,n} (u_{m+\ell,n} - v_{m+\ell,n}) \end{array} \right. \\ \text{for } i = 1, \dots, m \\ \quad \left[ \begin{array}{l} y_{i,n} = \varepsilon_{i,n} \left( \mathbf{J}_{\mathbf{W}_i \mathbf{A}_i}(x_{i,n} - \mathbf{W}_i(\mathbf{M}_i^*(2u_{i,n} - v_{i,n}) + \sum_{(\ell,j) \in \mathbb{V}_i^*} (2u_{m+\ell,j,n} - v_{m+\ell,j,n}) \right. \\ \quad \left. + \mathbf{C}_i x_{i,n} + c_{i,n})) + a_{i,n} \right) \\ x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (y_{i,n} - x_{i,n}). \end{array} \right. \end{array} \right. \quad (6.27) \end{array}$$

Making explicit the form of the projections onto the vector spaces  $(\Lambda_{\kappa_\ell})_{1 \leq \ell \leq r}$  leads to

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \quad \left[ \begin{array}{l} \text{for } i = 1, \dots, m \\ \quad \left[ \begin{array}{l} u_{i,n} = \varepsilon_{m+i,n} \left( \mathbf{J}_{\mathbf{U}_i \mathbf{B}_i^{-1}}(v_{i,n} + \mathbf{U}_i(\mathbf{M}_i x_{i,n} - \mathbf{D}_i^{-1} v_{i,n} + d_{i,n})) + b_{i,n} \right) \\ v_{i,n+1} = v_{i,n} + \lambda_n \varepsilon_{m+i,n} (u_{i,n} - v_{i,n}) \end{array} \right. \\ \text{for } \ell = 1, \dots, r \\ \quad \left[ \begin{array}{l} \bar{u}_{\ell,n} = \varepsilon_{2m+\ell,n} \kappa_\ell^{-1} \left( \sum_{j=1}^{\kappa_\ell} v_{m+\ell,j,n} + \theta_\ell \sum_{i \in \mathbb{V}_\ell} x_{i,n} \right) \\ \text{for } j = 1, \dots, \kappa_\ell \\ \quad \left[ \begin{array}{l} u_{m+\ell,j,n} = \varepsilon_{2m+\ell,n} (v_{m+\ell,j,n} + \theta_\ell x_{i(\ell,j),n} - \bar{u}_{\ell,n}) \\ v_{m+\ell,n+1} = v_{m+\ell,n} + \lambda_n \varepsilon_{2m+\ell,n} (u_{m+\ell,n} - v_{m+\ell,n}) \end{array} \right. \\ \text{for } i = 1, \dots, m \\ \quad \left[ \begin{array}{l} y_{i,n} = \varepsilon_{i,n} \left( \mathbf{J}_{\mathbf{W}_i \mathbf{A}_i}(x_{i,n} - \mathbf{W}_i(\mathbf{M}_i^*(2u_{i,n} - v_{i,n}) + \sum_{(\ell,j) \in \mathbb{V}_i^*} (2u_{m+\ell,j,n} - v_{m+\ell,j,n}) \right. \\ \quad \left. + \mathbf{C}_i x_{i,n} + c_{i,n})) + a_{i,n} \right) \\ x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (y_{i,n} - x_{i,n}). \end{array} \right. \end{array} \right. \quad (6.28) \end{array} \right. \end{array}$$

By defining now, for every  $n \in \mathbb{N}$  and  $\ell \in \{1, \dots, r\}$ ,

$$\bar{x}_{\ell,n} = \frac{1}{\kappa_\ell} \sum_{i \in \mathbb{V}_\ell} x_{i,n}, \quad (6.29)$$

$$w_{\ell,n} = \varepsilon_{2m+\ell,n} (2u_{m+\ell,n} - v_{m+\ell,n}), \quad (6.30)$$

$$\eta_{\ell,n} = \max \{ \varepsilon_{i,n} \mid i \in \mathbb{V}_\ell \}, \quad (6.31)$$

and using (6.24) and the update equation

$$\bar{x}_{\ell,n+1} = \bar{x}_{\ell,n} + \eta_{\ell,n} \left( \frac{1}{\kappa_\ell} \sum_{i \in \mathbb{V}_\ell} x_{i,n+1} - \bar{x}_{\ell,n} \right), \quad (6.32)$$

(6.22) is obtained after reordering the computation steps in (6.28).

In order to apply Proposition 4.8, we shall now show that Condition (4.30) where  $\vartheta_\alpha$  is defined by (4.9) is fulfilled. Let

$$\mathbf{W}: \mathbb{H}^m \rightarrow \mathbb{H}^m: \mathbf{x} \mapsto (\mathbf{W}_i x_i)_{1 \leq i \leq m} \quad \text{and} \quad \mathbf{U}: \mathbf{G} \oplus \mathbf{H} \rightarrow \mathbf{G} \oplus \mathbf{H}: \mathbf{v} \mapsto (\mathbf{U}_k v_k)_{1 \leq k \leq m+r}. \quad (6.33)$$

According to Remark 6.4 and (6.25), we have

$$(\forall \mathbf{x} \in \mathbb{H}^m) \quad \mathbf{U}^{1/2} \mathbf{LW}^{1/2} \mathbf{x} = (\mathbf{U}_1^{1/2} \mathbf{M} \mathbf{W}^{1/2} \mathbf{x}, \mathbf{U}_2^{1/2} \mathbf{S} \mathbf{W}^{1/2} \mathbf{x}) \quad (6.34)$$

where  $\mathbf{U}_1: \mathbf{G} \rightarrow \mathbf{G}: (v_i)_{1 \leq i \leq m} \mapsto (\mathbf{U}_i v_i)_{1 \leq i \leq m}$  and  $\mathbf{U}_2: \mathbf{H} \rightarrow \mathbf{H}: (v_{m+\ell})_{1 \leq \ell \leq r} \mapsto (\theta_\ell v_{m+\ell})_{1 \leq \ell \leq r}$ . This allows us to deduce that

$$\begin{aligned} \|\mathbf{U}^{1/2} \mathbf{LW}^{1/2} \mathbf{x}\|^2 &= \|\mathbf{U}_1^{1/2} \mathbf{M} \mathbf{W}^{1/2} \mathbf{x}\|^2 + \|\mathbf{U}_2^{1/2} \mathbf{S} \mathbf{W}^{1/2} \mathbf{x}\|^2 \\ &= \sum_{i=1}^m \|\mathbf{U}_i^{1/2} \mathbf{M}_i \mathbf{W}_i^{1/2} x_i\|^2 + \left\langle \mathbf{W}^{1/2} \mathbf{x} \mid \mathbf{S}^* \mathbf{U}_2 \mathbf{S} \mathbf{W}^{1/2} \mathbf{x} \right\rangle. \end{aligned} \quad (6.35)$$

By using (6.9) and (6.12)-(6.13), it can be further noticed that

$$\mathbf{S}^* \mathbf{U}_2 \mathbf{S}: \mathbb{H}^m \rightarrow \mathbb{H}^m: (x_i)_{1 \leq i \leq m} \mapsto (\bar{\theta}_i x_i)_{1 \leq i \leq m} \quad (6.36)$$

which yields

$$\begin{aligned} \|\mathbf{U}^{1/2} \mathbf{LW}^{1/2} \mathbf{x}\|^2 &= \sum_{i=1}^m \|\mathbf{U}_i^{1/2} \mathbf{M}_i \mathbf{W}_i^{1/2} x_i\|^2 + \sum_{i=1}^m \bar{\theta}_i \|\mathbf{W}_i^{1/2} x_i\|^2 \\ &\leq \sum_{i=1}^m (\|\mathbf{U}_i^{1/2} \mathbf{M}_i \mathbf{W}_i^{1/2}\|^2 + \bar{\theta}_i \|\mathbf{W}_i\|) \|x_i\|^2 \leq \chi \|\mathbf{x}\|^2, \end{aligned} \quad (6.37)$$

so leading to  $\|\mathbf{U}^{1/2} \mathbf{LW}^{1/2}\|^2 \leq \chi$ . This shows that (6.20) implies (4.30).

In addition, Condition (iii) in Proposition 4.8 translates into Condition (iii) in the present proposition. It then follows from Propositions 4.8 and 6.3 that, for every  $i \in \{1, \dots, m\}$ ,  $(x_{i,n})_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\widehat{F}$ -valued random variable  $\widehat{x}$ . As a straightforward consequence of (6.29), for every  $\ell \in \{1, \dots, r\}$ ,  $(\bar{x}_{\ell,n})_{n \in \mathbb{N}}$  also converges weakly P-a.s. to  $\widehat{x}$ .  $\square$

**Remark 6.6**

- (i) The  $n$ -th iteration of Algorithm (6.22) basically consists of two kind of operations: the first ones update some of the variables  $(x_{i,n})_{1 \leq i \leq m}$  and  $(v_{i,n})_{1 \leq i \leq m}$  using the operators  $(J_{W_i A_i})_{1 \leq i \leq m}$ ,  $(J_{U_i B_i^{-1}})_{1 \leq i \leq m}$ ,  $(C_i)_{1 \leq i \leq m}$ , and  $(D_i^{-1})_{1 \leq i \leq m}$ , while the second ones can be viewed as merging steps performed on the sets  $(\mathbb{V}_\ell)_{1 \leq \ell \leq r}$ . In this context, a simple choice for the Boolean random variables  $(\varepsilon_{k,n})_{m+1 \leq k \leq 2m+r}$  to satisfy Condition (iii) is: for every  $n \in \mathbb{N}$ ,

$$(\forall i \in \{1, \dots, m\}) \quad \varepsilon_{m+i,n} = \varepsilon_{i,n}, \quad (6.38)$$

$$(\forall \ell \in \{1, \dots, r\}) \quad \varepsilon_{2m+\ell,n} = \eta_{\ell,n} = \max \{ \varepsilon_{i,n} \mid i \in \mathbb{V}_\ell \}. \quad (6.39)$$

- (ii) From (6.27), it can be noticed that, for every  $n \in \mathbb{N}$  and  $\ell \in \{1, \dots, r\}$ ,  $\Pi_{\Lambda_{\kappa_\ell}} u_{m+\ell,n} = 0$ , which implies that the following recursive relation holds:

$$\sum_{j=1}^{\kappa_\ell} v_{m+\ell,j,n+1} = (1 - \lambda_n \varepsilon_{2m+\ell,n}) \sum_{j=1}^{\kappa_\ell} v_{m+\ell,j,n}. \quad (6.40)$$

In particular, if the initial values  $(v_{m+\ell,0})_{1 \leq \ell \leq r}$  are chosen such that

$$(\forall \ell \in \{1, \dots, r\}) \quad \sum_{j=1}^{\kappa_\ell} v_{m+\ell,j,0} = 0, \quad (6.41)$$

then Algorithm (6.22) simplifies to

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \quad \left[ \begin{array}{l} \text{for } \ell = 1, \dots, r \\ \quad \left[ \begin{array}{l} \text{for } j = 1, \dots, \kappa_\ell \\ \quad \left[ w_{\ell,j,n} = \varepsilon_{2m+\ell,n} (2\theta_\ell (x_{i(\ell,j),n} - \bar{x}_{\ell,n}) + v_{m+\ell,j,n}) \right] \\ \text{for } i = 1, \dots, m \\ \quad \left[ \begin{array}{l} u_{i,n} = \varepsilon_{m+i,n} \left( J_{U_i B_i^{-1}} (v_{i,n} + U_i (M_i x_{i,n} - D_i^{-1} v_{i,n} + d_{i,n})) + b_{i,n} \right) \\ y_{i,n} = \varepsilon_{i,n} \left( J_{W_i A_i} (x_{i,n} - W_i (M_i^* (2u_{i,n} - v_{i,n}) + \sum_{(\ell,j) \in \mathbb{V}_i^*} w_{\ell,j,n} + C_i x_{i,n} + c_{i,n})) + a_{i,n} \right) \\ v_{i,n+1} = v_{i,n} + \lambda_n \varepsilon_{m+i,n} (u_{i,n} - v_{i,n}) \\ x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (y_{i,n} - x_{i,n}) \end{array} \right. \\ \text{for } \ell = 1, \dots, r \\ \quad \left[ \begin{array}{l} v_{m+\ell,n+1} = v_{m+\ell,n} + \frac{\lambda_n}{2} \varepsilon_{2m+\ell,n} (w_{\ell,n} - v_{m+\ell,n}) \\ \eta_{\ell,n} = \max \{ \varepsilon_{i,n} \mid i \in \mathbb{V}_\ell \} \\ \bar{x}_{\ell,n+1} = \bar{x}_{\ell,n} + \eta_{\ell,n} \left( \frac{1}{\kappa_\ell} \sum_{i \in \mathbb{V}_\ell} x_{i,n+1} - \bar{x}_{\ell,n} \right). \end{array} \right. \end{array} \right. \end{array} \end{array} \quad (6.42)$$

(iii) Similarly to Remark 4.7(iii), a sufficient condition for (6.20) to be satisfied is obtained by setting  $\alpha = 1$ :

$$(1 - \sqrt{\chi}) \min\{\mu, \nu\} > \frac{1}{2}. \quad (6.43)$$

(iv) When, for every  $i \in \{1, \dots, m\}$ ,  $D_i^{-1} = 0$ , a looser condition is

$$(1 - \chi)\mu > \frac{1}{2}. \quad (6.44)$$

In addition, if  $(\forall i \in \{1, \dots, m\}) \mathbf{B}_i = 0$ ,  $(\|\mathbf{M}_i\|)_{1 \leq i \leq m}$  can be chosen as small as desired, so that we can set  $\chi = \max_{i \in \{1, \dots, m\}} \bar{\theta}_i \|\mathbf{W}_i\|$ . In this case, Algorithm (6.42) can be simplified, by noting that, for every  $n \in \mathbb{N}$ , the computation of variables  $(u_{i,n})_{1 \leq i \leq m}$  and  $(v_{i,n})_{1 \leq i \leq m}$  becomes useless. By imposing (6.39) and (6.41), and by setting

$$(\forall n \in \mathbb{N})(\forall \ell \in \{1, \dots, m\}) \quad \tilde{v}_{\ell,n} = v_{m+\ell,n}, \quad (6.45)$$

we get

$$\left[ \begin{array}{l} \text{for } n = 0, 1, \dots \\ \left[ \begin{array}{l} \text{for } \ell = 1, \dots, r \\ \left[ \begin{array}{l} \eta_{\ell,n} = \max \{ \varepsilon_{i,n} \mid i \in \mathbb{V}_\ell \} \\ \text{for } j = 1, \dots, \kappa_\ell \\ \left[ \begin{array}{l} w_{\ell,j,n} = \eta_{\ell,n} (2\theta_\ell (x_{i(\ell,j),n} - \bar{x}_{\ell,n}) + \tilde{v}_{\ell,j,n}) \end{array} \right. \\ \text{for } i = 1, \dots, m \\ \left[ \begin{array}{l} y_{i,n} = \varepsilon_{i,n} \left( \mathbf{J}_{\mathbf{W}_i \mathbf{A}_i} (x_{i,n} - \mathbf{W}_i \left( \sum_{(\ell,j) \in \mathbb{V}_i^*} w_{\ell,j,n} + \mathbf{C}_i x_{i,n} + c_{i,n} \right)) + a_{i,n} \right) \\ x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (y_{i,n} - x_{i,n}) \end{array} \right. \\ \text{for } \ell = 1, \dots, r \\ \left[ \begin{array}{l} \tilde{v}_{\ell,n+1} = \tilde{v}_{\ell,n} + \frac{\lambda_n}{2} \eta_{\ell,n} (w_{\ell,n} - \tilde{v}_{\ell,n}) \\ \bar{x}_{\ell,n+1} = \bar{x}_{\ell,n} + \eta_{\ell,n} \left( \frac{1}{\kappa_\ell} \sum_{i \in \mathbb{V}_\ell} x_{i,n+1} - \bar{x}_{\ell,n} \right). \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \right. \quad (6.46)$$

(v) An alternative distributed algorithm can be deduced from Proposition 4.6, which however necessitates, at each iteration  $n \in \mathbb{N}$ , to update all the variables  $(x_{i,n})_{i \in \mathbb{V}_\ell}$  corresponding to the sets  $\mathbb{V}_\ell$  with  $\ell \in \{1, \dots, r\}$  which are randomly activated.



As an offspring of Proposition 4.11, another form of distributed algorithm is obtained:

**Proposition 6.7** Let  $(\theta_\ell)_{1 \leq \ell \leq r}$ ,  $(W_i)_{1 \leq i \leq m}$ ,  $(U_i)_{1 \leq i \leq m}$ ,  $\mu$ ,  $\nu$ , and  $\chi$  be defined as in Proposition 6.5. Suppose that

$$\min \{ \mu, \nu(1 - \chi) \} > \frac{1}{2}. \quad (6.47)$$

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1]$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$  and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}^m$ -valued random variables, let  $(v_{i,0})_{1 \leq i \leq m}$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, let  $(v_{m+\ell,0})_{1 \leq \ell \leq r}$  be a  $\mathbf{H}$ -valued random variable satisfying (6.41), and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_{2m+r}$ -valued random variables. Iterate

$$\begin{array}{l} \text{for } n = 0, 1, \dots \\ \quad \left[ \begin{array}{l} \text{for } i = 1, \dots, m \\ \quad \left[ \begin{array}{l} \eta_{i,n} = \max \{ \varepsilon_{m+i,n}, (\varepsilon_{2m+\ell,n})_{\ell \in \{1, \dots, r\} | i \in \mathbb{V}_{\ell'}} \} \\ w_{i,n} = \eta_{i,n} (x_{i,n} - W_i(\mathbf{C}_i x_{i,n} + \mathbf{c}_{i,n})) \\ \tilde{w}_{i,n} = \eta_{i,n} (w_{i,n} - W_i(M_i^* v_{i,n} + \sum_{(\ell,j) \in \mathbb{V}_i^*} v_{m+\ell,j,n})) \\ u_{i,n} = \varepsilon_{m+i,n} (\mathbf{J}_{U_i \mathbf{B}_i^{-1}}(v_{i,n} + U_i(M_i \tilde{w}_{i,n} - D_i^{-1} v_{i,n} + \mathbf{d}_{i,n})) + \mathbf{b}_{i,n}) \\ v_{i,n+1} = v_{i,n} + \lambda_n \varepsilon_{m+i,n} (u_{i,n} - v_{i,n}) \end{array} \right. \\ \text{for } \ell = 1, \dots, r \\ \quad \left[ \begin{array}{l} \bar{w}_{\ell,n} = \varepsilon_{2m+\ell,n} \frac{\theta_\ell}{\kappa_\ell} \sum_{i \in \mathbb{V}_\ell} \tilde{w}_{i,n} \\ \text{for } j = 1, \dots, \kappa_\ell \\ \quad \left[ \begin{array}{l} u_{m+\ell,j,n} = \varepsilon_{2m+\ell,n} (v_{m+\ell,j,n} + \theta_\ell \tilde{w}_{i(j,\ell),n} - \bar{w}_{\ell,n}) \\ v_{m+\ell,n+1} = v_{m+\ell,n} + \lambda_n \varepsilon_{2m+\ell,n} (u_{m+\ell,n} - v_{m+\ell,n}) \end{array} \right. \\ \text{for } i = 1, \dots, m \\ \quad \left[ \begin{array}{l} x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (w_{i,n} - W_i(M_i^* u_{i,n} + \sum_{(\ell,j) \in \mathbb{V}_i^*} u_{m+\ell,j,n}) - x_{i,n}), \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \quad (6.48)$$

and set  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ . In addition, assume that

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

and Conditions (ii)-(iii) in Proposition 6.5 hold.

If Assumption 6.2 holds and, in Problem 6.1,  $(\forall i \in \{1, \dots, m\}) A_i = 0$ , then, for every  $i \in \{1, \dots, m\}$ ,  $(x_{i,n})_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\hat{\mathbf{F}}$ -valued random variable  $\hat{x}$ .

*Proof.* By choosing  $(U_{m+l})_{1 \leq l \leq r}$  as in (6.25) and cancelling some error terms as in (6.26), Algorithm (4.59) for solving Problem (6.2) can be expressed as

$$\begin{array}{l}
\text{for } n = 0, 1, \dots \\
\left[ \begin{array}{l}
\text{for } i = 1, \dots, m \\
\left[ \begin{array}{l}
\eta_{i,n} = \max \{ \varepsilon_{m+i,n}, (\varepsilon_{2m+l,n})_{\ell \in \{ \ell' \in \{1, \dots, r\} | i \in \mathbb{V}_{\ell'} \}} \} \\
w_{i,n} = \eta_{i,n} (x_{i,n} - W_i(C_i x_{i,n} + c_{i,n})) \\
\tilde{w}_{i,n} = \eta_{i,n} (w_{i,n} - W_i(M_i^* v_{i,n} + \sum_{(\ell,j) \in \mathbb{V}_i^*} v_{m+l,j,n})) \\
u_{i,n} = \varepsilon_{m+i,n} (J_{U_i B_i^{-1}}(v_{i,n} + U_i(M_i \tilde{w}_{i,n} - D_i^{-1} v_{i,n} + d_{i,n})) + b_{i,n}) \\
v_{i,n+1} = v_{i,n} + \lambda_n \varepsilon_{m+i,n} (u_{i,n} - v_{i,n})
\end{array} \right. \\
\text{for } \ell = 1, \dots, r \\
\left[ \begin{array}{l}
u_{m+l,n} = \varepsilon_{2m+l,n} (v_{m+l,n} + \theta_\ell (\tilde{w}_{i,n})_{i \in \mathbb{V}_\ell} - \Pi_{\Lambda_{\kappa_\ell}}(v_{m+l,n} + \theta_\ell (\tilde{w}_{i,n})_{i \in \mathbb{V}_\ell})) \\
v_{m+l,n+1} = v_{m+l,n} + \lambda_n \varepsilon_{2m+l,n} (u_{m+l,n} - v_{m+l,n})
\end{array} \right. \\
\text{for } i = 1, \dots, m \\
\left[ \begin{array}{l}
x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (w_{i,n} - W_i(M_i^* u_{i,n} + \sum_{(\ell,j) \in \mathbb{V}_i^*} u_{m+l,j,n}) - x_{i,n}).
\end{array} \right.
\end{array} \right. \tag{6.49}
\end{array}$$

The rest of the proof is skipped due to its similarity with the proof of Proposition 6.5.  $\square$

**Remark 6.8** When  $(\forall i \in \{1, \dots, m\}) D_i^{-1} = 0$ , Condition (6.47) can be rewritten as

$$(\forall i \in \{1, \dots, m\}) \quad \|U_i^{1/2} M_i W_i^{1/2}\|^2 + \bar{\theta}_i \|W_i\| < 1 \quad \text{and} \quad \mu_i > 1/2. \tag{6.50}$$

As an illustration of the previous results in this section, let us consider variational problems which can be expressed as follows:

**Problem 6.9** For every  $i \in \{1, \dots, m\}$ , let  $f_i \in \Gamma_0(H)$ , let  $h_i \in \Gamma_0(H)$  be Lipschitz-differentiable, let  $g_i \in \Gamma_0(G_i)$ , let  $l_i \in \Gamma_0(G_i)$  be strongly convex, and let  $M_i$  be a nonzero operator in  $\mathcal{B}(H, G_i)$ . Suppose that there exists  $\bar{x} \in H$  such that

$$0 \in \sum_{i=1}^m \partial f_i(\bar{x}) + \nabla h_i(\bar{x}) + M_i^* (\partial g_i \square \partial l_i)(M_i \bar{x}). \tag{6.51}$$

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

$$\underset{x \in H}{\text{minimize}} \quad \sum_{i=1}^m f_i(x) + h_i(x) + (g_i \square l_i)(M_i x). \tag{6.52}$$

Our objective is to find a  $\check{F}$ -valued random variable  $\hat{x}$ .

A proximal algorithm for solving Problem 6.9 which results from Proposition 6.5 is described next:

**Proposition 6.10** Let  $(\theta_\ell)_{1 \leq \ell \leq r}$ ,  $(W_i)_{1 \leq i \leq m}$ ,  $(U_i)_{1 \leq i \leq m}$ , and  $\chi$  be defined as in Proposition 6.5. For every  $i \in \{1, \dots, m\}$ , let  $\mu_i^{-1} \in ]0, +\infty[$  be a Lipschitz constant of the gradient of  $h_i \circ W_i^{1/2}$  and let  $\nu_i^{-1} \in ]0, +\infty[$  be a Lipschitz constant of the gradient of  $l_i^* \circ U_i^{1/2}$ . Suppose that (6.20) holds where  $\mu = \min\{\mu_1, \dots, \mu_m\}$  and  $\nu = \min\{\nu_1, \dots, \nu_m\}$ . Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, 1]$  such that  $\inf_{n \in \mathbb{N}} \lambda_n > 0$ , let  $\mathbf{x}_0$ ,  $(\mathbf{a}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{c}_n)_{n \in \mathbb{N}}$  be  $\mathbf{H}^m$ -valued random variables, let  $(v_{i,0})_{1 \leq i \leq m}$ ,  $(\mathbf{b}_n)_{n \in \mathbb{N}}$ , and  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  be  $\mathbf{G}$ -valued random variables, let  $(v_{m+\ell,0})_{1 \leq \ell \leq r}$  be a  $\mathbf{H}$ -valued random variable satisfying (6.41), and let  $(\varepsilon_n)_{n \in \mathbb{N}}$  be identically distributed  $\mathbb{D}_{2m+r}$ -valued random variables. Iterate

$$\begin{array}{l}
\text{for } n = 0, 1, \dots \\
\left[ \begin{array}{l}
\text{for } \ell = 1, \dots, r \\
\left[ \begin{array}{l}
\text{for } j = 1, \dots, \kappa_\ell \\
\left[ w_{\ell,j,n} = \varepsilon_{2m+\ell,n} (2\theta_\ell (x_{i(\ell,j),n} - \bar{x}_{\ell,n}) + v_{m+\ell,j,n}) \right. \\
\text{for } i = 1, \dots, m \\
\left[ \begin{array}{l}
u_{i,n} = \varepsilon_{m+i,n} \left( \text{prox}_{\mathbf{g}_i^*}^{U_i^{-1}} (v_{i,n} + U_i (M_i x_{i,n} - \nabla l_i^*(v_{i,n}) + d_{i,n})) + b_{i,n} \right) \\
y_{i,n} = \varepsilon_{i,n} \left( \text{prox}_{\mathbf{f}_i}^{W_i^{-1}} (x_{i,n} - W_i (M_i^* (2u_{i,n} - v_{i,n}) + \sum_{(\ell,j) \in \mathbb{V}_i^*} w_{\ell,j,n} + \nabla h_i(x_{i,n}) + c_{i,n})) + a_{i,n} \right) \\
v_{i,n+1} = v_{i,n} + \lambda_n \varepsilon_{m+i,n} (u_{i,n} - v_{i,n}) \\
x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (y_{i,n} - x_{i,n})
\end{array} \right. \\
\text{for } \ell = 1, \dots, r \\
\left[ \begin{array}{l}
v_{m+\ell,n+1} = v_{m+\ell,n} + \frac{\lambda_n}{2} \varepsilon_{2m+\ell,n} (w_{\ell,n} - v_{m+\ell,n}) \\
\eta_{\ell,n} = \max \{ \varepsilon_{i,n} \mid i \in \mathbb{V}_\ell \} \\
\bar{x}_{\ell,n+1} = \bar{x}_{\ell,n} + \eta_{\ell,n} \left( \frac{1}{\kappa_\ell} \sum_{i \in \mathbb{V}_\ell} x_{i,n+1} - \bar{x}_{\ell,n} \right),
\end{array} \right.
\end{array} \right.
\end{array} \right.
\end{array} \tag{6.53}$$

where  $(\bar{x}_{\ell,0})_{1 \leq \ell \leq r}$  is initialized as in Proposition 6.5. In addition, assume that Conditions (i)-(iii) in Proposition 6.5 hold, where  $(\forall n \in \mathbb{N}) \mathcal{E}_n = \sigma(\varepsilon_n)$  and  $\mathcal{X}_n = \sigma(\mathbf{x}_{n'}, \mathbf{v}_{n'})_{0 \leq n' \leq n}$ .

Then, under Assumption 6.2, for every  $i \in \{1, \dots, m\}$ ,  $(x_{i,n})_{n \in \mathbb{N}}$  converges weakly P-a.s. to a  $\check{\mathbf{F}}$ -valued random variable  $\hat{x}$  and, for every  $\ell \in \{1, \dots, r\}$ ,  $(\bar{x}_{\ell,n})_{n \in \mathbb{N}}$  converges weakly P-a.s. to  $\hat{x}$ .

*Proof.* For every  $i \in \{1, \dots, m\}$ , set  $A_i = \partial f_i$ ,  $B_i = \partial g_i$ ,  $C_i = \nabla h_i$ , and  $D_i^{-1} = \nabla l_i^*$ . In view of (6.51) and [5, Proposition 16.5], we have

$$0 \in \sum_{i=1}^m A_i \bar{x} + C_i \bar{x} + M_i^* (B_i \square D_i) (M_i \bar{x}) \subset \partial \left( \sum_{i=1}^m f_i + h_i + (g_i \square l_i) \circ M_i \right) (\bar{x}), \tag{6.54}$$

which shows that  $\emptyset \neq \hat{\mathbf{F}} \subset \check{\mathbf{F}}$ . This allows us to conclude by applying Proposition 6.5 and using Remark 6.6(ii).  $\square$

**Remark 6.11**

- (i) Alternatively, a second distributed convex optimization algorithm can be deduced from Proposition 6.7.

- (ii) If  $(\forall i \in \{1, \dots, m\}) \mathbf{g}_i = \mathbf{0}$  and  $l_i = \iota_{\{0\}}$ ,  $(\forall \ell \in \{1, \dots, r\}) \kappa_\ell = 2$ , and (6.39) holds, then Algorithm (6.53) reduces to

$$\begin{array}{l}
 \text{for } n = 0, 1, \dots \\
 \left[ \begin{array}{l}
 \text{for } \ell = 1, \dots, r \\
 \left[ \begin{array}{l}
 \eta_{\ell,n} = \max \{ \varepsilon_{i,n} \mid i \in \mathbb{V}_\ell \} \\
 \tilde{v}_{\ell,1,n+1} = \tilde{v}_{\ell,1,n} + \frac{\lambda_n}{2} \eta_{\ell,n} \theta_\ell (x_{i(\ell,1),n} - x_{i(\ell,2),n}) \\
 \tilde{v}_{\ell,2,n+1} = -\tilde{v}_{\ell,1,n+1}
 \end{array} \right. \\
 \text{for } i = 1, \dots, m \\
 \left[ \begin{array}{l}
 y_{i,n} = \varepsilon_{i,n} \left( \text{prox}_{\mathbb{F}_i}^{\mathbf{W}_i^{-1}} \left( (1 - \mathbf{W}_i \bar{\theta}_i) x_{i,n} - \mathbf{W}_i \left( \sum_{(\ell,j) \in \mathbb{V}_i^*} (\tilde{v}_{\ell,j,n} - \theta_\ell x_{i(\ell,\bar{j}),n}) + \nabla \mathbf{h}_i(x_{i,n}) + c_{i,n} \right) \right) \right. \\
 \left. + a_{i,n} \right) \\
 x_{i,n+1} = x_{i,n} + \lambda_n \varepsilon_{i,n} (y_{i,n} - x_{i,n}),
 \end{array} \right.
 \end{array} \right.
 \end{array} \tag{6.55}$$

where we have set  $(\forall j \in \{1, 2\}) \bar{j} = 3 - j$ ,  $(\forall i \in \{1, \dots, m\}) v_{i,0} = 0$ ,  $(\forall n \in \mathbb{N}) \tilde{\mathbf{v}}_n = (v_{m+\ell,n})_{1 \leq \ell \leq r}$ , and  $\mathbf{b}_n = \mathbf{0}$ . The particular case when  $\mathbf{H}$  is an Euclidean space,  $(\forall n \in \mathbb{N}) \lambda_n = 1$ ,  $(\forall \ell \in \{1, \dots, r\}) \theta_\ell = \theta_1$ ,  $(\forall i \in \{1, \dots, m\}) \mathbf{W}_i = \tau_i \text{Id}$  with  $\tau_i \in ]0, +\infty[$ , and no error term is taken into account appears to be similar to the distributed iterative scheme developed in [7]. Then, the sets  $(\mathbb{V}_\ell)_{1 \leq \ell \leq r}$  can be viewed as the edges of a connected undirected graph, the nodes of which are indexed by  $i \in \{1, \dots, m\}$ .

## References

- [1] M. A. Alghamdi, A. Alotaibi, P. L. Combettes, and N. Shahzad. A primal-dual method of partial inverses for composite inclusions. *Optim. Letters*, 2014. Accepted, <http://arxiv.org/abs/1310.1152>.
- [2] A. Alotaibi, P. L. Combettes, and N. Shahzad. Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set. *SIAM J. Optim.*, 2014. Accepted, <http://arxiv.org/abs/1312.6696>.
- [3] H. Attouch, L. M. Briceño Arias, and P. L. Combettes. A parallel splitting method for coupled monotone inclusions. *SIAM J. Control Optim.*, 48:3246–3270, 2010.
- [4] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski. Optimization with sparsity-inducing penalties. *Found. Trends in Machine Learn.*, 4(1):1–106, 2012.
- [5] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.
- [6] S. R. Becker and P. L. Combettes. An algorithm for splitting parallel sums of linearly composed monotone operators, with applications to signal recovery. *J. Nonlinear Convex Anal.*, 15(1):137–159, Jan. 2014.
- [7] P. Bianchi, W. Hachem, and F. Iutzeler. A stochastic coordinate descent primal-dual algorithm and applications to large-scale composite optimization. 2014. <http://arxiv.org/abs/1407.0898>.

- [8] R. I. Boş 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(4):2541–2565, Dec. 2013.
- [9] R. I. Boş and C. Hendrich. Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization. *J. Math. Imaging Vision*, 49(3):551–568, Jul. 2014.
- [10] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Machine Learn.*, 8(1):1–122, 2011.
- [11] L. M. Briceño-Arias. A Douglas-Rachford splitting method for solving equilibrium problems. *Nonlinear Anal.*, 75(16):6053–6059, Nov. 2012.
- [12] L. M. Briceño-Arias and P. L. Combettes. A monotone + skew splitting model for composite monotone inclusions in duality. *SIAM J. Optim.*, 21(4):1230–1250, Oct. 2011.
- [13] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, 40(1):120–145, 2011.
- [14] T. Chang, A. Nedić, and A. Scaglione. Distributed constrained optimization by consensus-based primal-dual perturbation method. *IEEE Trans. Automat. Control*, 59(6):1524–1538, Jun. 2014.
- [15] G. Chen and M. Teboulle. A proximal-based decomposition method for convex minimization problems. *Math. Program.*, 64:81–101, 1994.
- [16] P. Chen, J. Huang, and X. Zhang. A primal-dual fixed point algorithm for convex separable minimization with applications to image restoration. *Inverse Problems*, 29(2):025011, 2013.
- [17] P. L. Combettes. Systems of structured monotone inclusions: duality, algorithms, and applications. *SIAM J. Optim.*, 23(4):2420–2447, Dec. 2013.
- [18] P. L. Combettes, L. Condat, J.-C. Pesquet, and B. C. Vũ. A forward-backward view of some primal-dual optimization methods in image recovery. In *Proc. Int. Conf. Image Process.*, Paris, France, 27-30 Oct. 2014. <http://arxiv.org/abs/1406.5439>.
- [19] P. L. Combettes, D. Dũng, and B. C. Vũ. Dualization of signal recovery problems. *Set-Valued Var. Anal.*, 18:373–404, Dec. 2010.
- [20] P. L. Combettes and J.-C. Pesquet. A proximal decomposition method for solving convex variational inverse problems. *Inverse Problems*, 24(6), Dec. 2008.
- [21] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer-Verlag, New York, 2011.
- [22] P. L. Combettes and J.-C. Pesquet. Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators. *Set-Valued Var. Anal.*, 20(2):307–330, June 2012.
- [23] P. L. Combettes and J.-C. Pesquet. Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping. 2014. [http://www.optimization-online.org/DB\\_HTML/2014/04/4333.html](http://www.optimization-online.org/DB_HTML/2014/04/4333.html).
- [24] P. L. Combettes and B. C. Vũ. Variable metric forward-backward splitting with applications to monotone inclusions in duality. *Optimization*, 63(9):1289–1318, Sept. 2014.

- [25] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. and Simul.*, 4(4):1168–1200, 2005.
- [26] L. Condat. A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.*, 158(2):460–479, Aug. 2013.
- [27] C. Couprie, L. Grady, L. Najman, J.-C. Pesquet, and H. Talbot. Dual constrained TV-based regularization on graphs. *SIAM J. Imaging Sci.*, 6:1246–1273, 2013.
- [28] E. Esser, X. Zhang, and T. Chan. A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *SIAM J. Imaging Sci.*, 3(4):1015–1046, 2010.
- [29] R. M. Fortet. *Fonctions et Distributions Aléatoires dans les Espaces de Hilbert*. Hermès, Paris, 1995.
- [30] T. Goldstein, E. Esser, and R. Baraniuk. Adaptive primal-dual hybrid gradient methods for saddle-point problems. 2013. <http://arxiv.org/abs/1305.0546>.
- [31] S. Harizanov, J.-C. Pesquet, and G. Steidl. Epigraphical projection for solving least squares Anscombe transformed constrained optimization problems. In A. Kuijper et al., editor, *4th International Conference on Scale-Space and Variational Methods in Computer Vision*, volume 7893 of *Lecture Notes in Computer Science*, pages 125–136, Schloss Seggau, Leibnitz, Austria, 2-6 June 2013. Springer-Verlag, Berlin.
- [32] B. He and X. Yuan. Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective. *SIAM J. Imaging Sci.*, 5(1):119–149, 2012.
- [33] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms, Part II : Advanced Theory and Bundle Methods*. Springer-Verlag, Berlin, Heidelberg, N.Y., 1993.
- [34] F. Iutzeler, P. Bianchi, P. Ciblat, and W. Hachem. Asynchronous distributed optimization using a randomized alternating direction method of multipliers. In *Proc. 52nd Conf. Decision Control*, pages 3671–3676, Florence, Italy, 10-13 Dec. 2013.
- [35] A. Jezierska, E. Chouzenoux, J.-C. Pesquet, and H. Talbot. A primal-dual proximal splitting approach for restoring data corrupted with Poisson-Gaussian noise. In *Proc. Int. Conf. Acoust., Speech Signal Process.*, pages 1085–1088, Kyoto, Japan, 25-30 Mar. 2012.
- [36] N. Komodakis and J.-C. Pesquet. Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems. 2014. [http://www.optimization-online.org/DB\\_HTML/2014/06/4398.html](http://www.optimization-online.org/DB_HTML/2014/06/4398.html).
- [37] I. Loris and C. Verhoeven. On a generalization of the iterative soft-thresholding algorithm for the case of non-separable penalty. *Inverse Problems*, 27(12):125007, 2011.
- [38] R. D. C. Monteiro and B. F. Svaiter. Convergence rate of inexact proximal point methods with relative error criteria for convex optimization. 2010. [http://www.optimization-online.org/DB\\_HTML/2010/08/2714.html](http://www.optimization-online.org/DB_HTML/2010/08/2714.html).
- [39] J. J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.
- [40] 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(2):307–337, Mar. 2014.
- [41] A. Nedić and A. Ozdaglar. Subgradient methods for saddle-point problems. *J. Optim. Theory Appl.*, 142(1):205–228, 2009.

- [42] A. Nedić and A. Ozdaglar. Cooperative distributed multi-agent optimization. In D. P. Palomar and Y. C. Eldar, editors, *Convex Optimization in Signal Processing and Communications*, chapter 10, pages 340–386. Cambridge University Press, Cambridge, UK, 2010.
- [43] N. Parikh and S. Boyd. Proximal algorithms. *Found. Trends Optim.*, 1(3):123–231, 2013.
- [44] J.-C. Pesquet and N. Pustelnik. A parallel inertial proximal optimization method. *Pac. J. Optim.*, 8(2):273–305, Apr. 2012.
- [45] T. Pock and A. Chambolle. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *Proc. IEEE Int. Conf. Comput. Vis.*, pages 1762–1769, Barcelona, Spain, Nov. 6-13 2011.
- [46] N. Pustelnik, P. Borgnat, and P. Flandrin. Empirical Mode Decomposition revisited by multicomponent nonsmooth convex optimization. *Signal Process.*, 102:313–331, Sept. 2014.
- [47] H. Raguet, J. Fadili, , and G. Peyré. A generalized forward-backward splitting. *SIAM J. Imaging Sci.*, 6(3):1199–1226, 2013.
- [48] A. Repetti, E. Chouzenoux, and J.-C. Pesquet. A penalized weighted least squares approach for restoring data corrupted with signal-dependent noise. In *Proc. Eur. Sig. and Image Proc. Conference*, pages 1553–1557, Bucharest, Romania, 27-31 Aug. 2012.
- [49] P. Richtárik and M. Takáč. Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function. *Math. Program.*, 144(1–2):1–38, Apr. 2014.
- [50] M. V. Solodov and B. F. Svaiter. A unified framework for some inexact proximal point algorithms. *Numer. Funct. Anal. Optim.*, 22:1013–1035, 2001.
- [51] B. F. Svaiter. A class of Fejér convergent algorithms, approximate resolvents and the hybrid proximal-extragradient method. *J. Optim. Theory Appl.*, 162:133–153, July 2014.
- [52] T. Teuber, G. Steidl, and R.-H. Chan. Minimization and parameter estimation for seminorm regularization models with  $I$ -divergence constraints. *Inverse Problems*, 29:035007, Mar. 2013.
- [53] Z. J. Towfic and A. H. Sayed. Stability and performance limits of adaptive primal-dual networks. 2014. <http://arxiv.org/pdf/1408.3693.pdf>.
- [54] B. C. Vũ. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comput. Math.*, 38(3):667–681, Apr. 2013.
- [55] D. Yuan, S. Xu, and H. Zhao. Distributed primal-dual subgradient method for multiagent optimization via consensus algorithms. *IEEE Trans. Systems, Man, and Cybernetics- Part B*, 41(6):1715–1724, Dec. 2011.