Stochastic Block-Coordinate Fixed Point Iterations with Applications to Splitting

Jean-Christophe Pesquet

Laboratoire d'Informatique Gaspard Monge - CNRS
Univ. Paris-Est, France

Sestri Levante – 11 September 2014

In collaboration with

P. Combettes

A. Repetti

Motivation

- In many areas (machine learning, inverse problems, computer vision,...), iterative solutions are needed.

- Sequence such that

$$(\forall n \in \mathbb{N}) \qquad \mathsf{x}_{n+1} = \mathsf{T}_n \mathsf{x}_n$$

where T_n denotes an operator from H to H, where H real Hilbert space with norm $\|\cdot\|$ and inner product $\langle \cdot \mid \cdot \rangle$.

- Examples: Forward-Backward, Douglas-Rachford,...
for finding a zero of a sum of maximally monotone operators
or for minimizing a sum of convex functions.

Extension to block-coordinate approaches

- Principle
  Optimized variable: $\mathbf{x} = (\mathsf{x}_1, \dots, \mathsf{x}_m)$
  At each iteration n, update only a subset of components
  ($\sim$ Gauss-Seidel).
- Advantage
  reduced complexity and memory requirements per iteration
  $\Rightarrow$ Useful for large-scale optimization
- Difficulty

$$\mathsf{T}\colon (\mathsf{x}_1, \mathsf{x}_2) \mapsto (-\mathsf{x}_2, \mathsf{x}_1) \text{ is nonexpansive}$$

but

$$\left.\begin{aligned} \mathsf{T}_1\colon (\mathsf{x}_1, \mathsf{x}_2) &\mapsto (\mathsf{x}_1, \mathsf{x}_1) \\ \mathsf{T}_2\colon (\mathsf{x}_1, \mathsf{x}_2) &\mapsto (-\mathsf{x}_2, \mathsf{x}_2) \\ &\mathsf{T}_1 \circ \mathsf{T}_2 \\ &\mathsf{T}_2 \circ \mathsf{T}_1 \end{aligned}\right\} \text{are not nonexpansive}$$

$\rightsquigarrow$ introduce stochasticity

Notation

- H separable real Hilbert space.
- $\mathfrak{W}(\mathsf{x}_n)_{n\in\mathbb{N}}$: set of weak sequential cluster points of a sequence $(\mathsf{x}_n)_{n\in\mathbb{N}}$ in H.
- $(\Omega, \mathcal{F}, \mathsf{P})$: underlying probability space
- $\sigma(\Phi)$: smallest σ-algebra generated by a family Φ of random variables
- Given a sequence $(x_n)_{n\in\mathbb{N}}$ of H-valued random variables,

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

- $\ell_+(\mathscr{X})$: set of sequences of $[0, +\infty[$-valued random variables $(\xi_n)_{n\in\mathbb{N}}$ such that, for every $n \in \mathbb{N}$, ξ_n is $\mathcal{X}_n$-measurable,

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

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

Stochastic quasi-Fejér sequences

F: a nonempty closed subset of H

$\phi\colon\ [0,+\infty[\ \to [0,+\infty[$: strictly increasing with $\lim_{t\to+\infty}\phi(t)=+\infty$.

- **Deterministic definition**: A sequence $(\mathsf{x}_n)_{n\in\mathbb{N}}$ of H is Fejér monotone if for every $\mathsf{z}\in$ F,

$$(\forall n\in\mathbb{N})\ \phi(\|\mathsf{x}_{n+1}-\mathsf{z}\|)\leqslant\phi(\|\mathsf{x}_n-\mathsf{z}\|).$$

Stochastic quasi-Fejér sequences

F: a nonempty closed subset of H
$\phi\colon\ [0,+\infty[\ \to [0,+\infty[$: strictly increasing with $\lim_{t\to+\infty}\phi(t)=+\infty$.

- **Stochastic definition**: A sequence $(x_n)_{n\in\mathbb{N}}$ of H-valued random variables is Fejér monotone if for every z $\in$ F, the following is satisfied P-a.s.:

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

Stochastic quasi-Fejér sequences

F: a nonempty closed subset of H

$\phi\colon [0,+\infty[\to [0,+\infty[$: strictly increasing with $\lim_{t\to+\infty}\phi(t) = +\infty$.

- **Stochastic definition**: A sequence $(x_n)_{n\in\mathbb{N}}$ of H-valued random variables is quasi-Fejér monotone if for every $\mathsf{z} \in$ F, there exist $(\chi_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^1(\mathscr{X})$, $(\vartheta_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+(\mathscr{X})$, and $(\eta_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^1(\mathscr{X})$ such that the following is satisfied P-a.s.:

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

Stochastic quasi-Fejér sequences

F: a nonempty closed subset of H
$\phi\colon [0,+\infty[\, \to [0,+\infty[$: strictly increasing with $\lim_{t\to+\infty}\phi(t)=+\infty$.

- **Stochastic definition**: A sequence $(x_n)_{n\in\mathbb{N}}$ of H-valued random variables is quasi-Fejér monotone if for every $\mathsf{z}\in\mathsf{F}$, there exist $(\chi_n(\mathsf{z}))_{n\in\mathbb{N}}\in\ell_+^1(\mathscr{X})$, $(\vartheta_n(\mathsf{z}))_{n\in\mathbb{N}}\in\ell_+(\mathscr{X})$, and $(\eta_n(\mathsf{z}))_{n\in\mathbb{N}}\in\ell_+^1(\mathscr{X})$ such that the following is satisfied P-a.s.:

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

Theorem

If $(x_n)_{n\in\mathbb{N}}$ is quasi-Fejér monotone as defined above, then

1. $(\forall \mathsf{z}\in\mathsf{F})\ \left[\ \sum_{n\in\mathbb{N}}\vartheta_n(\mathsf{z})<+\infty\ \text{P-a.s.}\right]$
2. $[\mathfrak{W}(x_n)_{n\in\mathbb{N}}\subset\mathsf{F}\ \ \text{P-a.s.}]\Rightarrow[(x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* F*-valued random variable*$]$.

An abstract iterative scheme

Theorem

Let F *be a nonempty closed subset of* H. *Suppose that* $(x_n)_{n\in\mathbb{N}}$, $(t_n)_{n\in\mathbb{N}}$, *and* $(e_n)_{n\in\mathbb{N}}$ *are sequences of* H*-valued random variables such that:*

1. $(\forall n \in \mathbb{N})\ x_{n+1} = x_n + \lambda_n(t_n + e_n - x_n), \quad \lambda_n \in]0,1]$.
2. $\sum_{n\in\mathbb{N}} \lambda_n \sqrt{\mathsf{E}(\|e_n\|^2 \,|\, \mathfrak{X}_n)} < +\infty$ P-a.s.
3. *For every* $\mathsf{z} \in \mathsf{F}$, *there exist* $(\theta_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+(\mathscr{X})$, $(\mu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^\infty(\mathscr{X})$, *and* $(\nu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^\infty(\mathscr{X})$ *such that* $(\lambda_n\mu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^1(\mathscr{X})$, $(\lambda_n\nu_n(\mathsf{z}))_{n\in\mathbb{N}} \in \ell_+^{1/2}(\mathscr{X})$, *and the following is satisfied* P-a.s.*:*

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

Then $(\forall \mathsf{z} \in \mathsf{F}) \quad \left[\sum_{n\in\mathbb{N}} \lambda_n\theta_n(\mathsf{z}) < +\infty \text{ P-a.s.} \right]$ *and* $[\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{F} \text{ P-a.s.}] \Rightarrow (x_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to an* F*-valued random variable.*

Random block-coordinate Krasnosel'skiĭ–Mann iteration

- Algorithm

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\quad \text{for } i = 1, \ldots, m \\
\quad\quad x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\lambda_n\big(\mathsf{T}_i\,(x_{1,n}, \ldots, x_{m,n}) + a_{i,n} - x_{i,n}\big)
\end{array}
$$

where

- $\boldsymbol{x}_0$ and $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ **H**-valued random variables where $\mathbf{H} = \mathsf{H}_1 \oplus \cdots \oplus \mathsf{H}_m$, $(\mathsf{H}_i)_{1\leqslant i\leqslant m}$ separable real Hilbert spaces, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$: error term
- $\mathbf{T}\colon \mathbf{H} \to \mathbf{H}\colon \mathbf{x} \mapsto (\mathsf{T}_i\,\mathbf{x})_{1\leqslant i\leqslant m}$ nonexpansive operator such that $\operatorname{Fix}\mathbf{T} \neq \varnothing$ where, for every $i \in \{1, \ldots, m\}$, $\mathsf{T}_i\colon \mathbf{H} \to \mathsf{H}_i$.
- $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ identically distributed D-valued random variables with $\mathsf{D} = \{0, 1\}^m \smallsetminus \{\mathbf{0}\}$
- $(\forall n \in \mathbb{N})$ $\lambda_n \in\,]0, 1[$ with $\inf_{n\in\mathbb{N}} \lambda_n > 0$ and $\sup_{n\in\mathbb{N}} \lambda_n < 1$.

Random block-coordinate Krasnosel'skiĭ–Mann iteration

- Algorithm

for $n = 0, 1, \ldots$
  for $i = 1, \ldots, m$
    $x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\lambda_n \big(\mathsf{T}_i\,(x_{1,n}, \ldots, x_{m,n}) + a_{i,n} - x_{i,n}\big)$

Theorem

Set $(\forall n \in \mathbb{N})\ \mathcal{X}_n = \sigma(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n)$ *and* $\mathcal{E}_n = \sigma(\boldsymbol{\varepsilon}_n)$. *Assume that*

1. $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2 \,|\, \mathcal{X}_n)} < +\infty$.
2. *For every* $n \in \mathbb{N}$, $\mathcal{E}_n$ *and* $\mathcal{X}_n$ *are independent.*
3. *For every* $i \in \{1, \ldots, m\}$, $\mathsf{P}[\varepsilon_{i,0} = 1] > 0$.

Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to a* $(\mathrm{Fix}\,\mathbf{T})$*-valued random variable.*

Random block-coordinate Douglas-Rachford splitting

- Variational problem

Find an element of the set **F** of solutions to the primal problem

$$\underset{\mathsf{x}_1\in\mathsf{H}_1,\ldots,\mathsf{x}_m\in\mathsf{H}_m}{\text{minimize}} \quad \sum_{i=1}^{m} \mathsf{f}_i(\mathsf{x}_i) + \mathbf{g}(\mathsf{x}_1,\ldots,\mathsf{x}_m),$$

where $(\forall i \in \{1,\ldots,m\})$ $\mathsf{f}_i \in \Gamma_0(\mathsf{H}_i)$ and $\mathbf{g} \in \Gamma_0(\mathbf{H})$, and an element of the set $\mathbf{F}^*$ of solutions to the dual problem

$$\underset{\mathsf{u}_1\in\mathsf{H}_1,\ldots,\mathsf{u}_m\in\mathsf{H}_m}{\text{minimize}} \quad \sum_{i=1}^{m} \mathsf{f}_i^*(-\mathsf{u}_i) + \mathbf{g}^*(\mathsf{u}_1,\ldots,\mathsf{u}_m).$$

We assume that there exists $(\mathsf{x}_1,\ldots,\mathsf{x}_m) \in \mathbf{H}$ such that

$$\mathbf{0} \in \times_{i=1}^{m} \partial \mathsf{f}_i(\mathsf{x}_i) + \partial \mathbf{g}(\mathsf{x}_1,\ldots,\mathsf{x}_m).$$

Random block-coordinate Douglas-Rachford splitting

- Algorithm

for $n = 0, 1, \ldots$
  for $i = 1, \ldots, m$
    $z_{i,n+1} = z_{i,n} + \varepsilon_{i,n}\big(\mathsf{Q}_i(x_{1,n}, \ldots, x_{m,n}) + b_{i,n} - z_{i,n}\big)$
    $x_{i,n+1} = x_{i,n} + \varepsilon_{i,n}\mu_n\big(\mathrm{prox}_{\gamma \mathsf{f}_i}(2z_{i,n+1} - x_{i,n}) + a_{i,n} - z_{i,n+1}\big)$

where

- $\boldsymbol{x}_0$, $\boldsymbol{z}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ **H**-valued random variables, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$: error terms
- $\mathrm{prox}_{\gamma \mathbf{g}} \colon \mathbf{x} \mapsto (\mathsf{Q}_i\mathbf{x})_{1\leqslant i\leqslant m}$ where $(\forall i \in \{1, \ldots, m\})$ $\mathsf{Q}_i \colon \mathbf{H} \to \mathsf{H}_i$
- $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ identically distributed D-valued random variables with $\mathsf{D} = \{0, 1\}^m \smallsetminus \{\mathbf{0}\}$
- $\gamma \in \left]0, +\infty\right[$
- $(\forall n \in \mathbb{N})$ $\mu_n \in \left]0, 2\right[$ such that $\inf_{n\in\mathbb{N}} \mu_n > 0$ and $\sup_{n\in\mathbb{N}} \mu_n < 2$.

Random block-coordinate Douglas-Rachford splitting

- Algorithm

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

Theorem

Assume that

1. $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2 \,|\, \boldsymbol{\mathcal{X}}_n)} < +\infty$ *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2 \,|\, \boldsymbol{\mathcal{X}}_n)} < +\infty$.
2. *For every* $n \in \mathbb{N}$, $\boldsymbol{\mathcal{E}}_n$ *and* $\boldsymbol{\mathcal{X}}_n$ *are independent.*
3. *For every* $i \in \{1, \ldots, m\}$, $\mathsf{P}[\varepsilon_{i,0} = 1] > 0$.

Then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges weakly* P-a.s. *to a random variable* $\boldsymbol{x}$ *such that* $\boldsymbol{z} = \operatorname{prox}_{\gamma \mathbf{g}} \boldsymbol{x}$ *is* **F**-*valued and* $\gamma^{-1}(\boldsymbol{x} - \boldsymbol{z})$ *is* $\mathbf{F}^*$-*valued.*

Double-layer random block-coordinate algorithms

- Algorithm

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

  where

  - $\boldsymbol{x}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ **H**-valued random variables
    $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$: errors terms
  - For every $n \in \mathbb{N}$, $\mathbf{R}_n : \mathbf{H} \to \mathbf{H}$ β_n-averaged and
    $\mathbf{T}_n : \mathbf{H} \to \mathbf{H} : \mathbf{x} \mapsto (\mathsf{T}_{i,n}\mathbf{x})_{1\leqslant i\leqslant m}$ α_n-averaged, where
    $(\forall i \in \{1, \ldots, m\})$ $\mathsf{T}_{i,n} : \mathbf{H} \to \mathsf{H}_i$, $\alpha_n \in]0,1[$, $\beta_n \in]0,1[$
  - $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ identically distributed D-valued random variables.

Double-layer random block-coordinate algorithms

- Algorithm

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

Theorem

Assume that

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

Then, $\big[\, (\forall \mathbf{z} \in \mathbf{F}) \;\; \mathbf{T}_n(\mathbf{R}_n \boldsymbol{x}_n) - \mathbf{R}_n \boldsymbol{x}_n + \mathbf{R}_n \mathbf{z} \to \mathbf{z} \,\big]$ P-a.s.
and $\big[\, (\forall \mathbf{z} \in \mathbf{F}) \;\; \boldsymbol{x}_n - \mathbf{R}_n \boldsymbol{x}_n + \mathbf{R}_n \mathbf{z} \to \mathbf{z} \,\big]$ P-a.s.

Double-layer random block-coordinate algorithms

- Algorithm

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

Theorem

Assume that

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

Furthermore, $[\mathfrak{W}(\boldsymbol{x}_n)_{n \in \mathbb{N}} \subset \mathbf{F}\ \ \mathsf{P}\text{-a.s.}] \Rightarrow (\boldsymbol{x}_n)_{n \in \mathbb{N}}$ *converges weakly* P-a.s. *to an* $\mathbf{F}$*-valued random variable.*

Random block-coordinate forward-backward splitting

- Variational problem

Find an element of the set **F** of solutions to

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

where $(\forall i \in \{1,\ldots,m\})$ $\mathsf{f}_i \in \Gamma_0(\mathsf{H}_i)$, $(\forall k \in \{1,\ldots,p\})$ $\mathsf{g}_k\colon \mathsf{G}_k \to \mathbb{R}$ τ_k-Lipschitz differentiable convex function with $\tau_k \in \left]0,+\infty\right[$, G_k separable real Hilbert space, $\mathsf{L}_{k,i}\colon \mathsf{H}_i \to \mathsf{G}_k$ linear and bounded.
We assume that **F** $\neq \varnothing$ and $\min_{1\leqslant k\leqslant p} \sum_{i=1}^{m} \|\mathsf{L}_{k,i}\|^2 > 0$.

Random block-coordinate forward-backward splitting

- Algorithm

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

where

- $\boldsymbol{x}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ **H**-valued random variables, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$: error terms
- $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ identically distributed D-valued random variables with $\mathsf{D} = \{0,1\}^m \smallsetminus \{\mathbf{0}\}$
- $(\gamma_n)_{n\in\mathbb{N}}$ sequence in $]0, 2\vartheta[$ such that $\inf_{n\in\mathbb{N}} \gamma_n > 0$ and $\sup_{n\in\mathbb{N}} \gamma_n < 2\vartheta$ with

$$
\vartheta = \left(\sum_{k=1}^{p} \tau_k \left\| \sum_{i=1}^{m} \mathsf{L}_{k,i}\mathsf{L}_{k,i}^{*} \right\| \right)^{-1}
$$

- $(\forall n \in \mathbb{N})\ \lambda_n \in\,]0,1]$ such that $\inf_{n\in\mathbb{N}} \lambda_n > 0$.

Random block-coordinate forward-backward splitting

- Algorithm

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

Remark: extension of [Attouch et al., 2010], similar stochastic algorithm in [Richtárik,Takáč, 2014]

Theorem

Assume that

1. $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{a}_n\|^2 \,|\, \mathcal{X}_n)} < +\infty$ *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{b}_n\|^2 \,|\, \mathcal{X}_n)} < +\infty$.
2. *For every* $n \in \mathbb{N}$ $\mathcal{E}_n$ *and* $\mathcal{X}_n$ *are independent.*
3. *For every* $i \in \{1, \ldots, m\}$, $\mathsf{P}[\varepsilon_{i,0} = 1] > 0$.

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

Extended form of random block-coordinate forward-backward splitting

- Monotone inclusion problem

Find an element of the set **F** of solutions to

find $(\mathsf{x}_1, \ldots, \mathsf{x}_m) \in \mathbf{K}$ such that $\mathbf{0} \in \mathbf{Q}(\mathsf{x}_1, \ldots, \mathsf{x}_m) + \mathbf{R}(\mathsf{x}_1, \ldots, \mathsf{x}_m)$

where $\mathbf{K} = \mathsf{K}_1 \oplus \cdots \oplus \mathsf{K}_m$, $(\mathsf{K}_i)_{1 \leqslant i \leqslant m}$ separable real Hilbert spaces, $\mathbf{Q} \colon \mathbf{K} \to 2^{\mathbf{K}}$ maximally monotone operator, and $\mathbf{R} \colon \mathbf{K} \to \mathbf{K}$ μ-cocoercive operator with $\mu \in \left]0, +\infty\right[$.
We assume that $\mathbf{F} \neq \varnothing$.

Extended form of random block-coordinate forward-backward splitting

- Algorithm

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\boldsymbol{r}_n = \mathbf{R}\boldsymbol{z}_n \\
\text{for } i = 1, \ldots, m \\
\left\lfloor z_{i,n+1} = z_{i,n} + \lambda_n \varepsilon_{i,n} \big(\mathsf{T}_{i,n}(\boldsymbol{z}_n - \gamma_n \boldsymbol{r}_n + \boldsymbol{s}_n) + t_{i,n} - z_{i,n}\big) \right.
\end{array}
\right.
\end{array}
$$

where

- $\boldsymbol{x}_0$, $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ $\mathbf{K}$-valued random variables, $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$: error terms
- $(\forall n \in \mathbb{N})$ $\mathbf{J}_{\gamma_n \mathbf{Q}} \colon \mathbf{z} \mapsto (\mathsf{T}_{i,n}\mathbf{z})_{1\leqslant i\leqslant m}$ with $(\forall i \in \{1,\ldots,m\})$ $\mathsf{T}_{i,n} \colon \mathbf{K} \to \mathsf{K}_i$
- $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ identically distributed D-valued random variables with $\mathsf{D} = \{0,1\}^m \smallsetminus \{\mathbf{0}\}$
- $(\gamma_n)_{n\in\mathbb{N}}$ sequence in $]0, 2\mu[$ such that $\inf_{n\in\mathbb{N}} \gamma_n > 0$ and $\sup_{n\in\mathbb{N}} \gamma_n < 2\mu$
- $(\forall n \in \mathbb{N})$ $\lambda_n \in]0,1]$ such that $\inf_{n\in\mathbb{N}} \lambda_n > 0$.

Extended form of random block-coordinate forward-backward splitting

- Algorithm

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\boldsymbol{r}_n = \mathbf{V}\mathbf{R}\boldsymbol{z}_n \\
\text{for } i = 1, \ldots, m \\
\left\lfloor\; z_{i,n+1} = z_{i,n} + \lambda_n \varepsilon_{i,n}\big(\mathsf{T}_{i,n}(\boldsymbol{z}_n - \gamma_n \boldsymbol{r}_n + \boldsymbol{s}_n) + t_{i,n} - z_{i,n}\big)\right.
\end{array}
\right.
\end{array}
$$

where

- $\boldsymbol{x}_0$, $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{t}_n)_{n\in\mathbb{N}}$ $\mathbf{K}$-valued random variables
- $\mathbf{V}\colon \mathbf{K} \to \mathbf{K}$ strongly positive self-adjoint bounded operator such that $\mathbf{V}^{1/2}\mathbf{R}\mathbf{V}^{1/2}$ is ϑ-cocoercive with $\vartheta \in\,]0, +\infty[$
- $(\forall n \in \mathbb{N})$ $\mathbf{J}_{\gamma_n \mathbf{V}\mathbf{Q}}\colon \mathbf{z} \mapsto (\mathsf{T}_{i,n}\mathbf{z})_{1\leqslant i\leqslant m}$ with $(\forall i \in \{1, \ldots, m\})$ $\mathsf{T}_{i,n}\colon \mathbf{K} \to \mathsf{K}_i$
- $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ identically distributed D-valued random variables with $\mathsf{D} = \{0, 1\}^m \smallsetminus \{\mathbf{0}\}$
- $(\gamma_n)_{n\in\mathbb{N}}$ sequence in $]0, 2\vartheta[$ such that $\inf_{n\in\mathbb{N}} \gamma_n > 0$ and $\sup_{n\in\mathbb{N}} \gamma_n < 2\vartheta$
- $(\forall n \in \mathbb{N})$ $\lambda_n \in\,]0, 1]$ such that $\inf_{n\in\mathbb{N}} \lambda_n > 0$.

Extended form of random block-coordinate forward-backward splitting

- Algorithm

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor \begin{array}{l}
\boldsymbol{r}_n = \mathbf{VR}\boldsymbol{z}_n \\
\text{for } i = 1, \ldots, m \\
\left\lfloor \; z_{i,n+1} = z_{i,n} + \lambda_n \varepsilon_{i,n} \big(\mathsf{T}_{i,n}(\boldsymbol{z}_n - \gamma_n \boldsymbol{r}_n + \boldsymbol{s}_n) + t_{i,n} - z_{i,n} \big) \right.
\end{array} \right.
\end{array}
$$

Theorem

Set $(\forall n \in \mathbb{N})\ \mathcal{Z}_n = \sigma(\boldsymbol{z}_0, \ldots, \boldsymbol{z}_n)$ *and* $\mathcal{E}_n = \sigma(\boldsymbol{\varepsilon}_n)$. *Assume that*

1. $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{s}_n\|^2 \,|\, \mathcal{Z}_n)} < +\infty$ *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}(\|\boldsymbol{t}_n\|^2 \,|\, \mathcal{Z}_n)} < +\infty$.
2. *For every* $n \in \mathbb{N}$, $\mathcal{E}_n$ *and* $\mathcal{Z}_n$ *are independent.*
3. *For every* $i \in \{1, \ldots, m\}$, $\mathsf{P}[\varepsilon_{i,0} = 1] > 0$.

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

Random block-coordinate primal-dual algorithm

- Variational problem

Find an element of the set **F** of solutions to

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

where

- $(\forall j\in\{1,\dots,p\})$ $f_j\in\Gamma_0(H_j)$, h_j convex μ_j-Lipschitz differentiable with $\mu_j\in\,]0,+\infty[$
- $(\forall k\in\{1,\dots,q\})$ $g_k\in\Gamma_0(G_k)$, $l_k\in\Gamma_0(G_k)$ ν_k-strongly convex with $\nu_k\in\,]0,+\infty[$ $\Leftrightarrow$ $l_k^*\in\Gamma_0(G_k)$ ν_k-Lipschitz differentiable
- $L_{k,j}\colon H_j\to G_k$ linear and bounded
- $g_k\,\square\,l_k$ inf-convolution of g_k and l_k:

$$(\forall v_k\in G_k)\qquad (g_k\,\square\,l_k)(v_k)=\inf_{v_k'\in G_k} g_k(v_k')+l_k(v_k-v_k')$$

$\Rightarrow$ $(g_k\,\square\,l_k)^*=g_k^*+l_k^*$

$g_k\,\square\,\iota_{\{0\}}=g_k$

Random block-coordinate primal-dual algorithm

- Variational problem

Find an element of the set **F** of solutions to

$$\underset{\mathsf{x}_1\in\mathsf{H}_1,\ldots,\mathsf{x}_p\in\mathsf{H}_p}{\text{minimize}} \quad \sum_{j=1}^{p} \big(\mathsf{f}_j(\mathsf{x}_j)+\mathsf{h}_j(\mathsf{x}_j)\big) + \sum_{k=1}^{q} (\mathsf{g}_k \,\square\, \mathsf{l}_k)\Big(\sum_{j=1}^{p} \mathsf{L}_{k,j}\mathsf{x}_j\Big)$$

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

$$\underset{\mathsf{v}_1\in\mathsf{G}_1,\ldots,\mathsf{v}_q\in\mathsf{G}_q}{\text{minimize}} \quad \sum_{j=1}^{p} (\mathsf{f}_j^* \,\square\, \mathsf{h}_j^*)\Big(-\sum_{k=1}^{q} \mathsf{L}_{k,j}^*\mathsf{v}_k\Big) + \sum_{k=1}^{q} \big(\mathsf{g}_k^*(\mathsf{v}_k)+\mathsf{l}_k^*(\mathsf{v}_k)\big).$$

We assume that

$$(\forall j\in\{1,\ldots,p\}) \quad 0\in\partial\mathsf{f}_j(\overline{\mathsf{x}}_j)+\nabla\mathsf{h}_j(\overline{\mathsf{x}}_j) + \sum_{k=1}^{q} \mathsf{L}_{k,j}^*(\partial\mathsf{g}_k \,\square\, \partial\mathsf{l}_k)\Big(\sum_{j'=1}^{p} \mathsf{L}_{k,j'}\overline{\mathsf{x}}_{j'}\Big).$$

Random block-coordinate primal-dual algorithm

Theorem (Characterization of Kuhn-Tucker points [Condat,2013][Vũ,2013])

Let $\mathbf{K} = \mathbf{H} \oplus \mathbf{G}$ *with* $\mathbf{H} = \mathsf{H}_1 \oplus \cdots \oplus \mathsf{H}_p$ *and* $\mathbf{G} = \mathsf{G}_1 \oplus \cdots \oplus \mathsf{G}_q$,
$\mathbf{A}\colon \mathbf{H} \to 2^{\mathbf{H}} \colon \mathbf{x} \mapsto \times_{j=1}^{p} \partial \mathsf{f}_j(\mathsf{x}_j)$, $\mathbf{B}\colon \mathbf{G} \to 2^{\mathbf{G}} \colon \mathbf{v} \mapsto \times_{k=1}^{q} \partial \mathsf{g}_k(\mathsf{v}_k)$,
$\mathbf{C}\colon \mathbf{H} \to \mathbf{H} \colon \mathbf{x} \mapsto \big(\nabla \mathsf{h}_j(\mathsf{x}_j)\big)_{1\leqslant j\leqslant p}$, $\mathbf{D}\colon \mathbf{G} \to 2^{\mathbf{G}} \colon \mathbf{v} \mapsto \times_{k=1}^{q} \partial \mathsf{l}_k(\mathsf{v}_k)$, *and*
$\mathbf{L}\colon \mathbf{H} \to \mathbf{G} \colon \mathbf{x} \mapsto \big(\sum_{j=1}^{p} \mathsf{L}_{k,j}\mathsf{x}_j\big)_{1\leqslant k\leqslant q}$,

$$\mathbf{Q}\colon \mathbf{K} \to 2^{\mathbf{K}} \colon (\mathbf{x}, \mathbf{v}) \mapsto (\mathbf{A}\mathbf{x} + \mathbf{L}^*\mathbf{v}) \times (-\mathbf{L}\mathbf{x} + \mathbf{B}^{-1}\mathbf{v})$$

and

$$\mathbf{R}\colon \mathbf{K} \to \mathbf{K} \colon (\mathbf{x}, \mathbf{v}) \mapsto \big(\mathbf{C}\mathbf{x}, \mathbf{D}^{-1}\mathbf{v}\big).$$

Then, the following hold:

1. $\mathbf{Q}$ *is maximally monotone and* $\mathbf{R}$ *is cocoercive.*
2. $\mathbf{Z} = \operatorname{zer}(\mathbf{Q} + \mathbf{R})$ *is nonempty.*
3. $(\widehat{\mathbf{x}}, \widehat{\mathbf{v}}) \in \mathbf{F} \times \mathbf{F}^*$ *if and only if* $(\widehat{\mathbf{x}}, \widehat{\mathbf{v}}) \in \mathbf{Z}$.

Random block-coordinate primal-dual algorithm

- Algorithm

$$
\begin{aligned}
&\text{for } n = 0, 1, \ldots\\
&\quad \text{for } j = 1, \ldots, p\\
&\quad\quad y_{j,n} = \varepsilon_{j,n}\Big(\mathrm{prox}_{\mathsf{f}_j}^{\mathsf{W}_j^{-1}}\big(x_{j,n} - \mathsf{W}_j\big(\sum_{k\in\mathbb{L}_j^*} \mathsf{L}_{k,j}^* v_{k,n} + \nabla \mathsf{h}_j(x_{j,n}) + c_{j,n}\big)\big)\\
&\quad\quad\quad + a_{j,n}\Big)\\
&\quad\quad x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n}(y_{j,n} - x_{j,n})\\
&\quad \text{for } k = 1, \ldots, q\\
&\quad\quad u_{k,n} = \varepsilon_{p+k,n}\Big(\mathrm{prox}_{\mathsf{g}_k^*}^{\mathsf{U}_k^{-1}}\big(v_{k,n} + \mathsf{U}_k\big(\sum_{j\in\mathbb{L}_k} \mathsf{L}_{k,j}(2y_{j,n} - x_{j,n}) - \nabla \mathsf{l}_k^*(v_{k,n})\\
&\quad\quad\quad + d_{k,n}\big)\big) + b_{k,n}\Big)\\
&\quad\quad v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n}(u_{k,n} - v_{k,n}),
\end{aligned}
$$

where

- $\boldsymbol{x}_0$, $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$ **H**-valued random variables, $\boldsymbol{v}_0$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{d}_n)_{n\in\mathbb{N}}$ **G**-valued random variables,
  $(\boldsymbol{a}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{c}_n)_{n\in\mathbb{N}}$, and $(\boldsymbol{d}_n)_{n\in\mathbb{N}}$: error terms

Random block-coordinate primal-dual algorithm

- Algorithm

$$
\begin{aligned}
&\text{for } n = 0,1,\ldots\\
&\quad \text{for } j = 1,\ldots,p\\
&\quad\quad y_{j,n} \simeq \varepsilon_{j,n} \mathrm{prox}_{\mathsf{f}_j}^{\mathsf{W}_j^{-1}} \Big(x_{j,n} - \mathsf{W}_j\big(\sum_{k\in\mathbb{L}_j^*} \mathsf{L}_{k,j}^* v_{k,n} + \nabla \mathsf{h}_j(x_{j,n})\big)\Big)\\
&\quad\quad x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n}(y_{j,n} - x_{j,n})\\
&\quad \text{for } k = 1,\ldots,q\\
&\quad\quad u_{k,n} \simeq \varepsilon_{p+k,n} \mathrm{prox}_{\mathsf{g}_k^*}^{\mathsf{U}_k^{-1}} \Big(v_{k,n} + \mathsf{U}_k\big(\sum_{j\in\mathbb{L}_k} \mathsf{L}_{k,j}(2y_{j,n} - x_{j,n}) - \nabla \mathsf{l}_k^*(v_{k,n})\big)\Big)\\
&\quad\quad v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n}(u_{k,n} - v_{k,n}),
\end{aligned}
$$

where

- $(\forall j \in \{1,\ldots,p\})$ $\mathsf{W}_j \colon \mathsf{H}_j \to \mathsf{H}_j$ strongly positive self-adjoint bounded operator and $(\forall k \in \{1,\ldots,q\})$ $\mathsf{U}_k \colon \mathsf{G}_k \to \mathsf{G}_k$ strongly positive self-adjoint bounded operator such that
$$\left(1 - \Big(\sum_{j=1}^p \sum_{k=1}^q \|\mathsf{U}_k^{1/2}\mathsf{L}_{k,j}\mathsf{W}_j^{1/2}\|^2\Big)^{1/2}\right) \times \min\{(\|\mathsf{W}_j\|^{-1}\widetilde{\mu}_j)_{1\leqslant j\leqslant p}, (\|\mathsf{U}_k\|^{-1}\widetilde{\nu}_k)_{1\leqslant k\leqslant q}\} > \frac{1}{2}.$$

Random block-coordinate primal-dual algorithm

- Algorithm

$$
\begin{aligned}
&\text{for } n = 0, 1, \ldots\\
&\quad \text{for } j = 1, \ldots, p\\
&\qquad y_{j,n} \simeq \varepsilon_{j,n} \mathrm{prox}_{\mathsf{f}_j}^{\mathsf{W}_j^{-1}} \Big(x_{j,n} - \mathsf{W}_j \big(\sum_{k \in \mathbb{L}_j^*} \mathsf{L}_{k,j}^* v_{k,n} + \nabla \mathsf{h}_j(x_{j,n})\big)\Big)\\
&\qquad x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} (y_{j,n} - x_{j,n})\\
&\quad \text{for } k = 1, \ldots, q\\
&\qquad u_{k,n} \simeq \varepsilon_{p+k,n} \mathrm{prox}_{\mathsf{g}_k^*}^{\mathsf{U}_k^{-1}} \Big(v_{k,n} + \mathsf{U}_k \big(\sum_{j \in \mathbb{L}_k} \mathsf{L}_{k,j} (2 y_{j,n} - x_{j,n}) - \nabla \mathsf{l}_k^*(v_{k,n})\big)\Big)\\
&\qquad v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n}),
\end{aligned}
$$

where

- $(\forall k \in \{1, \ldots, q\})\ \mathbb{L}_k = \big\{ j \in \{1, \ldots, p\} \;\big|\; \mathsf{L}_{k,j} \neq 0 \big\} \neq \varnothing$,
  $(\forall j \in \{1, \ldots, p\})\ \mathbb{L}_j^* = \big\{ k \in \{1, \ldots, q\} \;\big|\; \mathsf{L}_{k,j} \neq 0 \big\} \neq \varnothing$,
  $\mathrm{prox}_{\mathsf{f}_j}^{\mathsf{W}_j^{-1}}$ proximity operator of f_j in $(\mathsf{H}_j, \|\cdot\|_{\mathsf{W}_j^{-1}})$,
  $\mathrm{prox}_{\mathsf{g}_k^*}^{\mathsf{U}_k^{-1}}$ proximity operator of g_k^* in $(\mathsf{G}_k, \|\cdot\|_{\mathsf{U}_k^{-1}})$

Random block-coordinate primal-dual algorithm

- Algorithm

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\text{for } j = 1, \ldots, p \\
\left\lfloor
\begin{array}{l}
y_{j,n} \simeq \varepsilon_{j,n} \mathrm{prox}_{\mathsf{f}_j}^{\mathsf{W}_j^{-1}} \Big(x_{j,n} - \mathsf{W}_j \big(\sum_{k \in \mathbb{L}_j^*} \mathsf{L}_{k,j}^* v_{k,n} + \nabla \mathsf{h}_j(x_{j,n}) \big) \Big) \\
x_{j,n+1} = x_{j,n} + \lambda_n \varepsilon_{j,n} (y_{j,n} - x_{j,n})
\end{array}
\right. \\
\text{for } k = 1, \ldots, q \\
\left\lfloor
\begin{array}{l}
u_{k,n} \simeq \varepsilon_{p+k,n} \mathrm{prox}_{\mathsf{g}_k^*}^{\mathsf{U}_k^{-1}} \Big(v_{k,n} + \mathsf{U}_k \big(\sum_{j \in \mathbb{L}_k} \mathsf{L}_{k,j} (2 y_{j,n} - x_{j,n}) - \nabla \mathsf{l}_k^*(v_{k,n}) \big) \Big) \\
v_{k,n+1} = v_{k,n} + \lambda_n \varepsilon_{p+k,n} (u_{k,n} - v_{k,n}),
\end{array}
\right.
\end{array}
\right.
\end{array}
$$

where

- $(\boldsymbol{\varepsilon}_n)_{n\in\mathbb{N}}$ identically distributed $\mathbb{D}$-valued random variables with $\mathbb{D} = \{0,1\}^{p+q} \smallsetminus \{\mathbf{0}\}$
- $(\forall n \in \mathbb{N})\ \lambda_n \in \left]0,1\right]$ such that $\inf_{n\in\mathbb{N}} \lambda_n > 0$.

Random block-coordinate primal-dual algorithm

Theorem

Set $(\forall n \in \mathbb{N})\ \mathcal{X}_n = \sigma(\boldsymbol{x}_{n'}, \boldsymbol{v}_{n'})_{0 \leqslant n' \leqslant n}$ *and* $\mathcal{E}_n = \sigma(\varepsilon_n)$*. Assume that*

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

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

Proof: Use extended form of random block-coordinate forward-backward splitting for an appropriate metric operator **V**.

Random block-coordinate primal-dual algorithm

- Advantages
  - No linear operator inversion.
  - Use of proximable or/and differentiable functions.
  - Use of preconditioning linear operators.
  - Flexibility in the random activation of primal/dual components.
  - Similar results for other primal-dual methods.
  - Applications to asynchronous distributed optimization.

Open issues

- Can we activate the components according to a Markov model ?
- Can we quantify the convergence speed ?
- Is there a way to do all this in a deterministic manner ?

P. Combettes and J.-C Pesquet
Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping
2014, http://arxiv.org/abs/1404.7536.

N. Komodakis and J.-C. Pesquet
Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems
2014, http://arxiv.org/abs/1406.5429.

J.-C. Pesquet and A. Repetti
A class of randomized primal-dual algorithms for distributed optimization
2014, http://arxiv.org/abs/1406.6404.