Optimization of Energy Production and Transport $\lozenge$ Approaches by Decomposition under Stochasticity

P. Carpentier, J.-P. Chancelier, A. Lenoir, F. Pacaud [1]

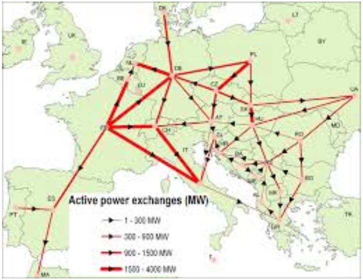
ENPC ParisTech — EDF Lab — Efficacity — ENSTA ParisTech

VAME Workshop, May 2017

[1] *Work supported by the FMJH Program Gaspard Monge for Optimization.*

Motivation

An energy production and transport optimization problem on a grid modeling energy exchange across countries.[2]



- Stochastic dynamical problem.
- Discrete time formulation (one-day time step).
- Large-scale problem (many countries).

[2] But the framework remains valid for smaller energy management problems.

Goal

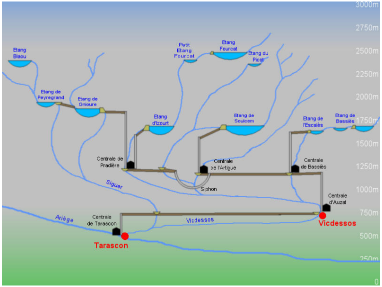
Obtain cost-to-go functions for such a large scale stochastic optimal control problem in discrete time.

- In order to obtain these functions ($\rightsquigarrow$ decision strategies), we have to use dynamic programming or related methods.
  - **Assumption**: Markovian case,
  - **Difficulty**: curse of dimensionality.
- To overcome the barrier of the dimension we want to use decomposition/coordination techniques (by country), which makes it difficult to take into account the information pattern induced by the stochasticity in the optimization problem.

This study is part of a broader project, aiming to develop decision analysis tools for long-term investment problems.

Previous work

We studied the application of stochastic decomposition to the optimization of large hydraulic valleys.



Valley: a tree structure with

- node: hydroelectric dam,
- arc: inter-dams connection.

We solved these problems using a price-decomposition approach (see [Carpentier et al, 2017]).

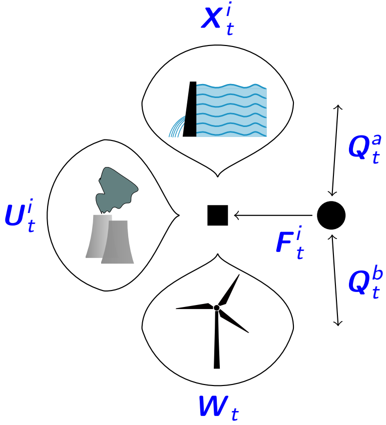
We want to extend this work in two directions:

- more complex topologies (graphs rather than trees)
- other decomposition algorithms (allocation, prediction).

Lecture outline

Production at each node of the grid

At each node i of the grid, we formulate a production problem on a discrete time horizon $[\![0, T]\!]$, involving the following variables at each time t:



- $\boldsymbol{X}^i_t$: state variable (dam volume)
- $\boldsymbol{U}^i_t$: control variable (energy production)
- $\boldsymbol{F}^i_t$: grid flow (import/export from the grid)
- $\boldsymbol{W}_t$: noise (consumption, renewable)

The noise $\boldsymbol{W}_t$ is supposed to be shared across the different nodes.

A stochastic optimization problem decoupled in space

At each node i of the grid, we have to solve a stochastic optimal control subproblem depending on the grid flow process $\boldsymbol{F}^i$:[3]

$$J^i_{\mathfrak{P}}[\boldsymbol{F}^i] = \min_{\boldsymbol{X}^i, \boldsymbol{U}^i} \mathbb{E}\Big(\sum_{t=0}^{T-1} L^i_t(\boldsymbol{X}^i_t, \boldsymbol{U}^i_t, \boldsymbol{F}^i_t, \boldsymbol{W}_{t+1}) + K^i(\boldsymbol{X}^i_T) \Big) ,$$

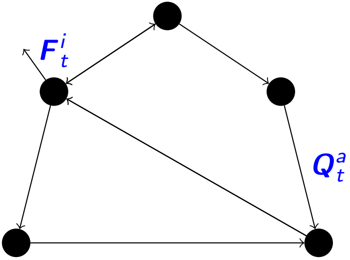
$$\begin{aligned} \text{s.t.} \quad & \boldsymbol{X}^i_{t+1} = f^i_t(\boldsymbol{X}^i_t, \boldsymbol{U}^i_t, \boldsymbol{F}^i_t, \boldsymbol{W}_{t+1}) , \\ & \boldsymbol{X}^i_t \in \mathcal{X}^{i,\mathrm{ad}}_t , \quad \boldsymbol{U}^i_t \in \mathcal{U}^{i,\mathrm{ad}}_t , \\ & \boldsymbol{U}^i_t \preceq \mathcal{F}_t , \end{aligned}$$

The last equation is the measurability constraint, where $\mathcal{F}_t$ is the σ-field generated by the noises $\{\boldsymbol{W}_\tau\}_{\tau=1\ldots t}$ up to time t.

[3] The notation $J^i_{\mathfrak{P}}[\cdot]$ means that the argument of $J^i_{\mathfrak{P}}$ is a *random variable*.

Modeling exchanges between countries

The grid is represented by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$. At each time $t \in [\![0, T-1]\!]$ we have:



- a flow $\boldsymbol{Q}_t^a$ through each arc a, inducing a cost $c_t^a(\boldsymbol{Q}_t^a)$, modeling the exchange between two countries
- a grid flow $\boldsymbol{F}_t^i$ at each node i, resulting from the balance equation

$$\boldsymbol{F}_t^i = \sum_{a \in input(i)} \boldsymbol{Q}_t^a - \sum_{b \in output(i)} \boldsymbol{Q}_t^b$$

A transport cost decoupled in time

At each time step $t \in [\![0, T-1]\!]$, we define the transport cost as the sum of the cost of the flows $\boldsymbol{Q}_t^a$ through the arcs a of the grid:

$$J_{\mathfrak{T},t}[\boldsymbol{Q}_t] = \mathbb{E}\Big(\sum_{a \in \mathcal{A}} c_t^a(\boldsymbol{Q}_t^a)\Big) ,$$

where the c_t^a's are easy to compute functions (say quadratic).

Kirchhof's law

The balance equation stating the conservation between $\boldsymbol{Q}_t$ and $\boldsymbol{F}_t$ rewrites in the following matrix form:

$$A\boldsymbol{Q}_t + \boldsymbol{F}_t = 0 ,$$

where A is the node-arc incidence matrix of the grid.

The overall production transport problem

The *production cost* $J_{\mathfrak{P}}$ aggregates the costs at all nodes i:

$$J_{\mathfrak{P}}[\boldsymbol{F}] = \sum_{i \in \mathcal{N}} J^i_{\mathfrak{P}}[\boldsymbol{F}^i] ,$$

and the *transport cost* $J_{\mathfrak{T}}$ aggregates the costs at all time t:

$$J_{\mathfrak{T}}[\boldsymbol{Q}] = \sum_{t=0}^{T-1} J_{\mathfrak{T},t}[\boldsymbol{Q}_t] .$$

The compact production-transport problem formulation writes:

$$\begin{aligned} \min_{\boldsymbol{Q},\boldsymbol{F}} \quad & J_{\mathfrak{P}}[\boldsymbol{F}] + J_{\mathfrak{T}}[\boldsymbol{Q}] \\ \text{s.t.} \quad & A\boldsymbol{Q} + \boldsymbol{F} = 0 \quad \leftrightsquigarrow \quad \text{coupling} . \end{aligned}$$

1. Introduction
   - The production and transport problem
   - Mixing decomposition and dynamic programming
2. Decomposition methods
   - Price decomposition
   - Resource allocation
   - Interaction prediction
3. Discussion

Introducing decomposition methods

The decomposition/coordination methods we want to deal with are iterative algorithms involving the following ingredients.

- Decompose the global problem in several subproblems of smaller size by dealing with the constraint $A\boldsymbol{Q} + \boldsymbol{F} = 0$,
- Coordinate at each iteration the subproblems using either a price or an allocation.

$$A\boldsymbol{Q} + \underbrace{\boldsymbol{F}}_{\textit{allocation}} = 0 \quad \rightsquigarrow \underbrace{\boldsymbol{\lambda}}_{\textit{price}}$$

- Solve the subproblems using Dynamic Programming (when a state is available in the subproblem), taking into account the price or the allocation transmitted by the coordination.

Production subproblems induced by decomposition

The i-th production subproblem at iteration k formulates as follows.

- Price transmission case

$$\begin{aligned} \min_{\boldsymbol{X}^i, \boldsymbol{U}^i, \boldsymbol{F}^i} \ & \mathbb{E}\Big(\sum_{t=0}^{T-1} L_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{F}_t^i, \boldsymbol{W}_{t+1}) + \langle \boldsymbol{\lambda}_t^{(k)}, \boldsymbol{F}_t^i \rangle + K^i(\boldsymbol{X}_T^i) \Big) , \\ \text{s.t.} \ & \boldsymbol{X}_{t+1}^i = f_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{F}_t^i, \boldsymbol{W}_{t+1}) , \\ & \boldsymbol{U}_t^i \preceq \mathcal{F}_t . \end{aligned}$$

- Allocation transmission case

$$\begin{aligned} \min_{\boldsymbol{X}^i, \boldsymbol{U}^i} \ & \mathbb{E}\Big(\sum_{t=0}^{T-1} L_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{F}_t^{i,(k)}, \boldsymbol{W}_{t+1}) + K^i(\boldsymbol{X}_T^i) \Big) , \\ \text{s.t.} \ & \boldsymbol{X}_{t+1}^i = f_t^i(\boldsymbol{X}_t^i, \boldsymbol{U}_t^i, \boldsymbol{F}_t^{i,(k)}, \boldsymbol{W}_{t+1}) , \\ & \boldsymbol{U}_t^i \preceq \mathcal{F}_t . \end{aligned}$$

Approximating the subproblems

In both cases, the subproblems encompass a new "noise", that is, either a price multiplier $\boldsymbol{\lambda}_t^{(k)}$ or a flow allocation $\boldsymbol{F}_t^{i,(k)}$, which may be correlated in time. The white noise assumption fails.

Dynamic Programming cannot be used for solving the subproblems.

In order to overcome this difficulty, we use a trick that involves approximating the new noise (either $\boldsymbol{\lambda}_t^k$ or $\boldsymbol{F}_t^k$) by its conditional expectation w.r.t. a chosen random variable $\boldsymbol{Y}_t$.

Assume that the process $\boldsymbol{Y}$ has a given dynamics:

$$\boldsymbol{Y}_{t+1} = h_t(\boldsymbol{Y}_t, \boldsymbol{W}_{t+1}) .$$

If noises $\boldsymbol{W}_t$'s are time independent, then $(\boldsymbol{X}_t^i, \boldsymbol{Y}_t)$ is a valid state for the i-th subproblem and Dynamic Programming applies.[4]

[4] See [Barty et al, 2010] for further details.

Approximating the subproblems

In both cases, the subproblems encompass a new "noise", that is, either a price multiplier $\boldsymbol{\lambda}_t^{(k)}$ or a flow allocation $\boldsymbol{F}_t^{i,(k)}$, which may be correlated in time. The white noise assumption fails.

Dynamic Programming cannot be used for solving the subproblems.

In order to overcome this difficulty, we use a trick that involves approximating the new noise (either $\boldsymbol{\lambda}_t^k$ or $\boldsymbol{F}_t^k$) by its conditional expectation w.r.t. a chosen random variable $\boldsymbol{Y}_t$.

Assume that the process $\boldsymbol{Y}$ has a given dynamics:

$$\boldsymbol{Y}_{t+1} = h_t(\boldsymbol{Y}_t, \boldsymbol{W}_{t+1}) .$$

If noises $\boldsymbol{W}_t$'s are time independent, then $(\boldsymbol{X}_t^i, \boldsymbol{Y}_t)$ is a valid state for the i-th subproblem and Dynamic Programming applies.[4]

[4] See [Barty et al, 2010] for further details.

Approximating the subproblems

In both cases, the subproblems encompass a new "noise", that is, either a price multiplier $\boldsymbol{\lambda}_t^{(k)}$ or a flow allocation $\boldsymbol{F}_t^{i,(k)}$, which may be correlated in time. The white noise assumption fails.

Dynamic Programming cannot be used for solving the subproblems.

In order to overcome this difficulty, we use a trick that involves approximating the new noise (either $\boldsymbol{\lambda}_t^k$ or $\boldsymbol{F}_t^k$) by its conditional expectation w.r.t. a chosen random variable $\boldsymbol{Y}_t$.

Assume that the process $\boldsymbol{Y}$ has a given dynamics:

$$\boldsymbol{Y}_{t+1} = h_t(\boldsymbol{Y}_t, \boldsymbol{W}_{t+1}) \ .$$

If noises $\boldsymbol{W}_t$'s are time independent, then $(\boldsymbol{X}_t^i, \boldsymbol{Y}_t)$ is a valid state for the i-th subproblem and Dynamic Programming applies.[4]

[4] See [Barty et al, 2010] for further details.

Price decomposition

The production and transport optimization problem writes

$$\min_{\boldsymbol{Q},\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] + J_{\mathfrak{T}}[\boldsymbol{Q}] \qquad \text{s.t.} \quad A\boldsymbol{Q} + \boldsymbol{F} = 0 \,. \qquad (\mathcal{P})$$

The decomposition scheme consists in dualizing the constraint, and then in approximating the multiplier $\boldsymbol{\lambda}$ by its conditional expectation w.r.t. $\boldsymbol{Y}$. This trick leads to the following problem

$$\max_{\boldsymbol{\lambda}} \min_{\boldsymbol{Q},\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] + J_{\mathfrak{T}}[\boldsymbol{Q}] + \left\langle \mathbb{E}(\boldsymbol{\lambda} \mid \boldsymbol{Y}) \,, A\boldsymbol{Q} + \boldsymbol{F} \right\rangle \,.$$

It is not difficult to prove that this dual problem is associated to the following relaxed primal problem:

$$\min_{\boldsymbol{Q},\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] + J_{\mathfrak{T}}[\boldsymbol{Q}] \qquad \text{s.t.} \quad \mathbb{E}\Big(A\boldsymbol{Q} + \boldsymbol{F} \;\Big|\; \boldsymbol{Y}\Big) = 0 \,,$$

and hence provides a lower bound of $(\mathcal{P})$.

A dual gradient-like algorithm

Applying the Uzawa algorithm to the dual problem

$$\max_{\boldsymbol{\lambda}} \min_{\boldsymbol{Q},\boldsymbol{F}} \; J_{\mathfrak{P}}[\boldsymbol{F}] + J_{\mathfrak{T}}[\boldsymbol{Q}] + \left\langle \mathbb{E}(\boldsymbol{\lambda} \mid \boldsymbol{Y}) , A\boldsymbol{Q} + \boldsymbol{F} \right\rangle ,$$

leads to a decomposition between production and transport:

$$\boldsymbol{F}^{(k+1)} \in \arg\min_{\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] + \left\langle \mathbb{E}\big(\boldsymbol{\lambda}^{(k)} \mid \boldsymbol{Y}\big) , \boldsymbol{F} \right\rangle , \qquad \text{Production}$$

$$\boldsymbol{Q}^{(k+1)} \in \arg\min_{\boldsymbol{Q}} J_{\mathfrak{T}}[\boldsymbol{Q}] + \left\langle \mathbb{E}\big(\boldsymbol{\lambda}^{(k)} \mid \boldsymbol{Y}\big) , A\boldsymbol{Q} \right\rangle , \qquad \text{Transport}$$

$$\mathbb{E}\big(\boldsymbol{\lambda}^{(k+1)} \mid \boldsymbol{Y}\big) = \mathbb{E}\big(\boldsymbol{\lambda}^{(k)} \mid \boldsymbol{Y}\big) + \rho\, \mathbb{E}\big(A\boldsymbol{Q}^{(k+1)} + \boldsymbol{F}^{(k+1)} \mid \boldsymbol{Y}\big) . \qquad \text{Update}$$

Note that the update step may implement a much more elaborated formula than the one corresponding to a fixed-step gradient...

Decomposing the transport subproblem

The transport subproblem

$$\min_{\boldsymbol{Q}} \; J_{\mathfrak{T}}[\boldsymbol{Q}] + \left\langle \mathbb{E}\left(\boldsymbol{\lambda}^{(k)} \;\middle|\; \boldsymbol{Y}\right), A\boldsymbol{Q} \right\rangle ,$$

writes in a detailled manner

$$\min_{\boldsymbol{Q}} \sum_{t=0}^{T-1} \mathbb{E}\Big(\sum_{a \in \mathcal{A}} c_t^a(\boldsymbol{Q}_t^a) + \left\langle A^\top \mathbb{E}\left(\boldsymbol{\lambda}_t^{(k)} \;\middle|\; \boldsymbol{Y}_t\right), \boldsymbol{Q}_t \right\rangle \Big) .$$

This minimization subproblem is evidently decomposable in time (t by t) and in space (arc by arc), leading to a collection of easy to solve subproblems.

Decomposing the production subproblem

The production subproblem

$$\min_{\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] + \left\langle \mathbb{E}\left(\boldsymbol{\lambda}^{(k)} \;\middle|\; \boldsymbol{Y}\right), \boldsymbol{F} \right\rangle ,$$

evidently decomposes node by node

$$\min_{\boldsymbol{F}^i} J^i_{\mathfrak{P}}[\boldsymbol{F}^i] + \left\langle \mathbb{E}\left(\boldsymbol{\lambda}^{i,(k)} \;\middle|\; \boldsymbol{Y}\right), \boldsymbol{F}^i \right\rangle ,$$

hence a stochastic optimal control subproblem for each node i:

$$\begin{aligned}
\min_{\boldsymbol{X}^i,\boldsymbol{U}^i,\boldsymbol{F}^i} \; & \mathbb{E}\left(\sum_{t=0}^{T-1} \left(L^i_t(\boldsymbol{X}^i_t, \boldsymbol{U}^i_t, \boldsymbol{F}^i_t, \boldsymbol{W}_{t+1}) + \left\langle \mathbb{E}\left(\boldsymbol{\lambda}^{i,(k)}_t \;\middle|\; \boldsymbol{Y}_t\right), \boldsymbol{F}^i_t \right\rangle \right) + K^i(\boldsymbol{X}^i_T) \right) \\
\text{s.t. } \; & \boldsymbol{X}^i_{t+1} = f^i_t(\boldsymbol{X}^i_t, \boldsymbol{U}^i_t, \boldsymbol{F}^i_t, \boldsymbol{W}_{t+1}) \\
& \boldsymbol{U}^i_t \preceq \mathcal{F}_t \; .
\end{aligned}$$

Solving the production subproblems by DP

Assuming that

- the process $\boldsymbol{W}$ is a white noise,
- the process $\boldsymbol{Y}$ follows a dynamics $\boldsymbol{Y}_{t+1} = h_t(\boldsymbol{Y}_t, \boldsymbol{W}_{t+1})$,

Dynamic Programming applies for production subproblems:

$$
\begin{aligned}
V_T^i(x,y) &= K^i(x) \\
V_t(x,y) &= \min_{u,f} \mathbb{E}\Big(L_t^i(x,u,f,\boldsymbol{W}_{t+1}) \\
&\qquad + \big\langle \mathbb{E}\big(\boldsymbol{\lambda}_t^{i,(k)} \mid \boldsymbol{Y}_t = y\big), f \big\rangle + V_{t+1}^i(\boldsymbol{X}_{t+1}^i, \boldsymbol{Y}_{t+1}) \Big) \\
&\quad \text{s.t.} \quad \boldsymbol{X}_{t+1}^i = f_t^i(x,u,f,\boldsymbol{W}_{t+1}) , \\
&\qquad\quad\;\; \boldsymbol{Y}_{t+1} = h_t(y, \boldsymbol{W}_{t+1}) .
\end{aligned}
$$

1. Introduction
   - The production and transport problem
   - Mixing decomposition and dynamic programming

2. Decomposition methods
   - Price decomposition
   - Resource allocation
   - Interaction prediction

3. Discussion

Resource allocation decomposition

Resource allocation decomposition applied to the problem

$$\min_{\boldsymbol{Q},\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] + J_{\mathfrak{T}}[\boldsymbol{Q}] \quad \text{s.t.} \quad A\boldsymbol{Q} + \boldsymbol{F} = 0 \, . \qquad (\mathcal{P})$$

consists in rewriting the constraint $A\boldsymbol{Q} + \boldsymbol{F} = 0$ by introducing a new variable $\boldsymbol{V}$ (the allocation), that is,

$$A\boldsymbol{Q} + \boldsymbol{V} = 0 \quad \text{and} \quad \boldsymbol{F} - \boldsymbol{V} = 0 \, .$$

Here the trick consists in limiting the measurability of variable $\boldsymbol{V}$, that is, $\boldsymbol{V} \preceq \boldsymbol{Y}$. This approximation leads to solve the following restricted primal problem (hence providing an upper bound of $(\mathcal{P})$)

$$\min_{\boldsymbol{V}\preceq\boldsymbol{Y}} \Bigg(\min_{\boldsymbol{F}} \Big(J_{\mathfrak{P}}[\boldsymbol{F}] \quad \text{s.t.} \quad \boldsymbol{F} - \boldsymbol{V} = 0 \Big) + \min_{\boldsymbol{Q}} \Big(J_{\mathfrak{T}}[\boldsymbol{Q}] \quad \text{s.t.} \quad A\boldsymbol{Q} + \boldsymbol{V} = 0 \Big) \Bigg) \, .$$

A primal gradient-like algorithm

Applying a gradient-like algorithm w.r.t. $\boldsymbol{V}$ to the problem

$$\min_{\boldsymbol{V} \preceq \boldsymbol{Y}} \left(\min_{\boldsymbol{F}} \Big(J_{\mathfrak{P}}[\boldsymbol{F}] \quad \text{s.t.} \quad \boldsymbol{F} - \boldsymbol{V} = 0 \Big) + \min_{\boldsymbol{Q}} \Big(J_{\mathfrak{T}}[\boldsymbol{Q}] \quad \text{s.t.} \quad A\boldsymbol{Q} + \boldsymbol{V} = 0 \Big) \right) ,$$

leads to a decomposition between production and transport:[5]

$$\min_{\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] \quad \text{s.t.} \quad \boldsymbol{F} - \boldsymbol{V}^{(k)} = 0 \quad \rightsquigarrow \quad \boldsymbol{\lambda}^{(k+1)} \qquad \text{Production}$$

$$\min_{\boldsymbol{Q}} J_{\mathfrak{T}}[\boldsymbol{Q}] \quad \text{s.t.} \quad A\boldsymbol{Q} + \boldsymbol{V}^{(k)} = 0 \quad \rightsquigarrow \quad \boldsymbol{\nu}^{(k+1)} \qquad \text{Transport}$$

$$\boldsymbol{V}^{(k+1)} = \mathrm{proj}_{\boldsymbol{V} \preceq \boldsymbol{Y}} \Big(\boldsymbol{V}^{(k)} + \rho \big(\boldsymbol{\lambda}^{(k+1)} - \boldsymbol{\nu}^{(k+1)} \big) \Big) \qquad \text{Update}$$

[5] Note that we must ensure at each iteration that $\boldsymbol{V}_t^{(k)} \in \mathrm{Im} A$.

Decomposing the transport subproblem

The transport subproblem

$$\min_{\boldsymbol{Q}} \; J_{\mathfrak{T}}[\boldsymbol{Q}] \quad \text{s.t.} \quad A\boldsymbol{Q} + \boldsymbol{V}^{(k)} = 0 \; ,$$

writes in a detailled manner

$$\min_{\boldsymbol{Q}} \sum_{t=0}^{T-1} \mathbb{E}\Big(\sum_{a \in \mathcal{A}} c_t^a(\boldsymbol{Q}_t^a) \Big) \quad \text{s.t.} \quad A\boldsymbol{Q}_t + \boldsymbol{V}_t^{(k)} = 0 \;\; \forall t \; .$$

This minimization subproblem is evidently decomposable in time (t by t), but not in space (coupling between the arcs). However, the resulting subproblems are still easy to solve.

Decomposing the production subproblem

The production subproblem

$$\min_{\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] \quad \text{s.t.} \quad \boldsymbol{F} - \boldsymbol{V}^{(k)} = 0 \ ,$$

evidently decomposes node by node

$$\min_{\boldsymbol{F}^i} J^i_{\mathfrak{P}}[\boldsymbol{F}^i] \quad \text{s.t.} \quad \boldsymbol{F}^i - \boldsymbol{V}^{i,(k)} = 0 \ ,$$

hence a stochastic optimal control subproblem for each node i:

$$\begin{aligned} \min_{\boldsymbol{X}^i, \boldsymbol{U}^i} \ & \mathbb{E}\bigg(\sum_{t=0}^{T-1} L^i_t(\boldsymbol{X}^i_t, \boldsymbol{U}^i_t, \boldsymbol{V}^{i,(k)}_t, \boldsymbol{W}_{t+1}) + K^i(\boldsymbol{X}^i_T) \bigg) \ , \\ \text{s.t.} \ & \boldsymbol{X}^i_{t+1} = f^i_t(\boldsymbol{X}^i_t, \boldsymbol{U}^i_t, \boldsymbol{V}^{i,(k)}_t, \boldsymbol{W}_{t+1}) \\ & \boldsymbol{U}^i_t \preceq \mathcal{F}_t \ . \end{aligned}$$

Solving the production subproblems by DP

Assuming that

- the process $\boldsymbol{W}$ is a white noise,
- the process $\boldsymbol{Y}$ follows a dynamics $\boldsymbol{Y}_{t+1} = h_t(\boldsymbol{Y}_t, \boldsymbol{W}_{t+1})$,

Dynamic Programming applies for production subproblems:[6]

$$V_T^i(x,y) = K^i(x)$$

$$V_t(x,y) = \min_u \mathbb{E}\Big(L_t^i(x,u,\psi_t^{i,(k)}(y), \boldsymbol{W}_{t+1}) + V_{t+1}^i(\boldsymbol{X}_{t+1}^i, \boldsymbol{Y}_{t+1}) \Big)$$

$$\text{s.t.} \quad \boldsymbol{X}_{t+1}^i = f_t^i(x,u,\psi_t^{i,(k)}(y), \boldsymbol{W}_{t+1}) \ ,$$

$$\boldsymbol{Y}_{t+1} = h_t(y, \boldsymbol{W}_{t+1}) \ .$$

[6] $\boldsymbol{V}_t^{i,(k)}$, being measurable w.r.t. $\boldsymbol{Y}_t$, writes as a functional $\psi_t^{i,(k)}$ of $\boldsymbol{Y}_t$.

Interaction Prediction Principle

As in resource allocation, we introduce a new variable $\boldsymbol{V}$ and rewrite the constraint $A\boldsymbol{Q} + \boldsymbol{F} = 0$ as

$$A\boldsymbol{Q} + \boldsymbol{V} = 0 \quad \text{and} \quad \boldsymbol{F} - \boldsymbol{V} = 0 \ .$$

We again limit the measurability of variable $\boldsymbol{V}$, that is, $\boldsymbol{V} \preceq \boldsymbol{Y}$. The interaction prediction is, in this specific case, a mix of price decomposition and resource allocation, aiming at solving

$$\min_{\boldsymbol{V} \preceq \boldsymbol{Y}} \max_{\boldsymbol{\mu}} \left(\min_{\boldsymbol{F}} \Big(J_{\mathfrak{P}}[\boldsymbol{F}] \quad \text{s.t.} \quad \boldsymbol{F} - \boldsymbol{V} = 0 \Big) + \min_{\boldsymbol{Q}} \Big(J_{\mathfrak{T}}[\boldsymbol{Q}] + \langle \boldsymbol{\mu} \, , A\boldsymbol{Q} + \boldsymbol{V} \rangle \Big) \right) ,$$

that is, a part of the constraint is handled as such (production), whereas the other part is treated by duality (transport).

A fixed-point algorithm

Applying a fixed-point algorithm w.r.t. $\boldsymbol{V}$ and $\boldsymbol{\mu}$ to the problem

$$\min_{\boldsymbol{V} \preceq \boldsymbol{Y}} \max_{\boldsymbol{\mu}} \Bigg(\min_{\boldsymbol{F}} \Big(J_{\mathfrak{P}}[\boldsymbol{F}] \quad \text{s.t.} \quad \boldsymbol{F} - \boldsymbol{V} = 0 \Big) + \min_{\boldsymbol{Q}} \Big(J_{\mathfrak{T}}[\boldsymbol{Q}] + \langle \boldsymbol{\mu} \, , A\boldsymbol{Q} + \boldsymbol{V} \rangle \Big) \Bigg) \, ,$$

leads to a decomposition between production and transport:

$$\min_{\boldsymbol{F}} J_{\mathfrak{P}}[\boldsymbol{F}] \quad \text{s.t.} \quad \boldsymbol{F} - \boldsymbol{V}^{(k)} = 0 \quad \rightsquigarrow \quad \boldsymbol{\lambda}^{(k+1)} \, , \qquad \text{Production}$$

$$\min_{\boldsymbol{Q}} J_{\mathfrak{T}}[\boldsymbol{Q}] + \langle \boldsymbol{\mu}^{(k)} \, , A\boldsymbol{Q} \rangle \quad \rightsquigarrow \quad \boldsymbol{Q}^{(k+1)} \, , \qquad \text{Transport}$$

$$\big(\boldsymbol{V}^{(k+1)}, \boldsymbol{\mu}^{(k+1)} \big) = \big(-\mathbb{E}(A\boldsymbol{Q}^{(k+1)} \mid \boldsymbol{Y}) \, , \, \boldsymbol{\lambda}^{(k+1)} \big) \, . \qquad \text{Update}$$

Decomposing the production and transport subproblems

In prediction decomposition, the production subproblem is solved in the same way as in ressource allocation, whereas the transport subproblem is solved in the same way as in price decomposition.

All that has been seen above therefore applies:

- the production subproblem decomposes node by node and Dynamic Programming applies;
- the transport subproblem decomposes in time and in space which leads to easy to solve subproblems.

Research agenda

We aim at benchmarking these three decomposition methods:

- numerical comparison on the (simplified) European grid,
- convergence and convergence rate of the method,
- proper choice of the information process $\boldsymbol{Y}$,
- gap between the lower and upper bounds:

$$\mathfrak{J}^{price} \leq \mathfrak{J}^{\sharp} \leq \mathfrak{J}^{resource} = \mathfrak{J}^{prediction} ,$$

- application to energy management in a urban district (dozens of houses equipped with solar panels, batteries and connected by a private network).

We also aim at comparing these methods with some augmented Lagrangian based methods such as ADMM (work in progress in cooperation with Ph. Mahey).

K. Barty, P. Carpentier, G. Cohen and P. Girardeau.
Price decomposition in large-scale stochastic optimal control.
arXiv, math.OC 1012.2092, 2010.

P. Carpentier, J.-P. Chancelier, V. Leclère and F. Pacaud.
Stochastic decomposition applied to large-scale hydro valleys management.
Working Paper, 2017.

G. Cohen.
Auxiliary Problem Principle and Decomposition of Optimization Problems.
Journal of Optimization Theory and Applications 32, 277-305, 1980.

P. Girardeau.
Résolution de grands problèmes en optimisation stochastique dynamique.
Thèse de doctorat, Université Paris-Est, 2010.

V. Leclère.
Contributions aux méthodes de décomposition en optimisation stochastique.
Thèse de doctorat, Université Paris-Est, 2014.

A. Lenoir and P. Mahey.
A survey of monotone operator splitting methods and decomposition of convex programs.
RAIRO Operations Research 51, 17-41, 2017.

Philippe Mahey, Jonas Koko, Arnaud Lenoir and Luc Marchand.
Coupling decomposition with dynamic programming for a stochastic spatial model for long-term energy pricing problem.
Working Paper, 2017.