# Space Decentralization and Primal-Dual Techniques for Stochastic Optimal Control

SESO'2014, ENSTA

Arnaud Lenoir Philippe Mahey

EDF R&D - OSIRIS

LIMOS - Université Blaise Pascal - UMR 6158

24 june 2014

# Introduction

- Short-term production planning (1 year) simulated for a far future (2025).
- Economic and political decisions for the European market.
- Facing offer and demand scenarios in order to analyze electricity pricing strategies .

# Example of an interconnected network

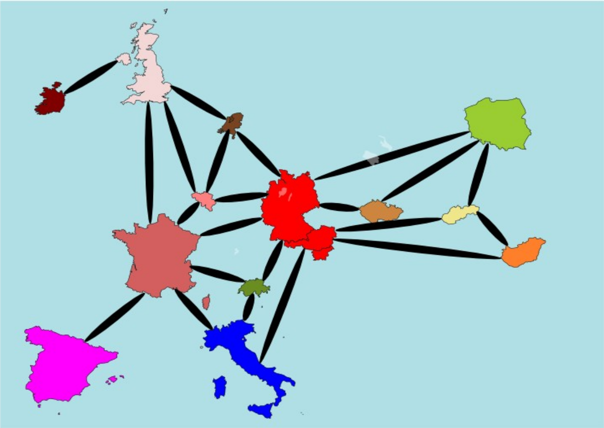


Figure : The European market network for electricity

# Example of an interconnected network

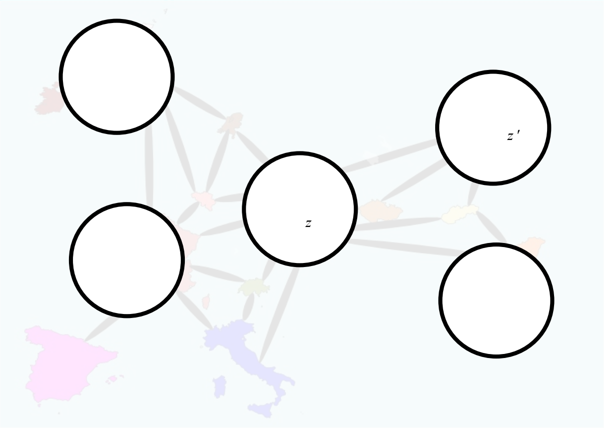



Figure : The European market network for electricity

# Example of an interconnected network

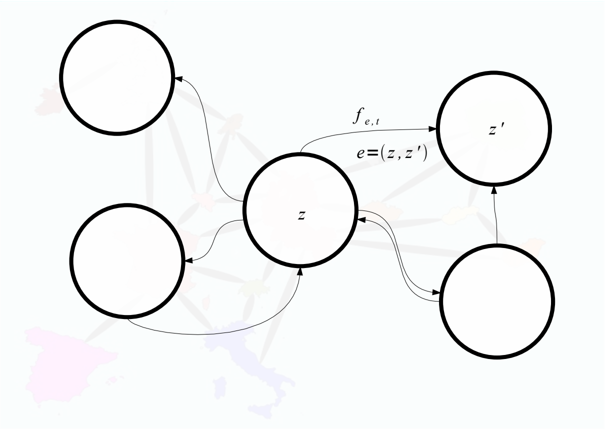



Figure : The European market network for electricity

# Example of an interconnected network

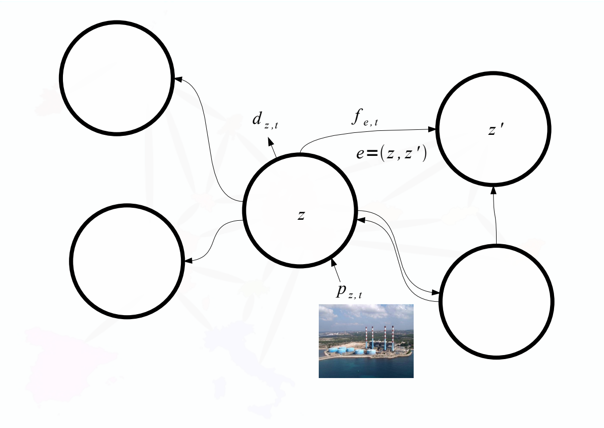



Figure : The European market network for electricity

# Example of an interconnected network

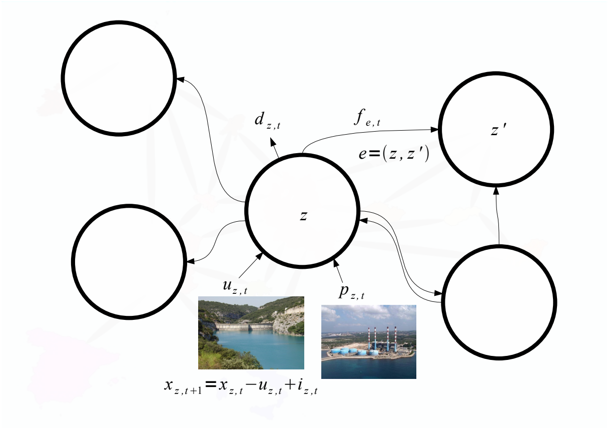



Figure : The European market network for electricity

# Electrical production

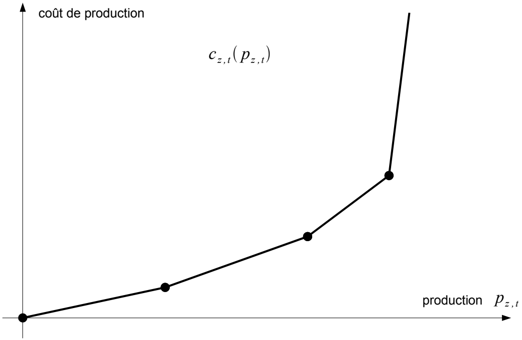
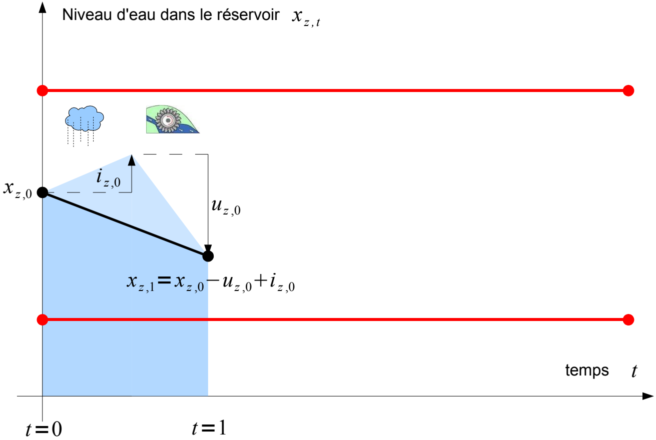
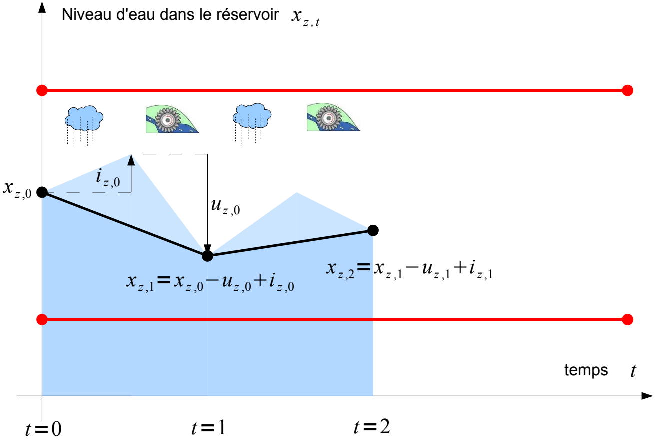



Figure : Thermic production cost

# Hydroelectric production



# Hydroelectric production



# Transportation costs

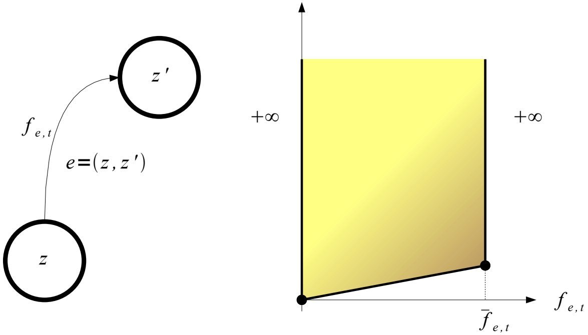



Figure : Exchange costs

# Model entries

- The network is simplified with a node for each zone and arcs between pairs of zones (graph $(Z, E \subset Z \times Z)$).
- Zonal demand is aggregated by $d_{z,t}$ for each period $t$.
- Thermic production is aggregated in each zone by $p_{z,t}$ with cost $c_z(p_{z,t})$,
- Hydro-electric production (or consume) is aggregated by $x_(z, t)$ and follows a storage dynamics $x_{z,t+1} = x_{z,t} - u_{z,t} + i_{z,t}$, where $u_(z, t)$ and $i(z, t)$ represents pumping and release of water respectively.
- Through interconnection arc $e = (z, z')$, zone $z$ can transfer some power flow $f_{e,t}$ to neighbour zone $z'$ at cost $l_e(f_{e,t})$.

# Production offer to meet demand

When the production of zone $z$ is offered at cost $c_z(\cdot)$, satisfying demand at minimal cost reduces to :

$$\tilde{Q}(d_t) = \min_{p_t, f_t} \quad \sum_{z \in Z} c_z(p_{z,t}) + \sum_{e \in E} l_e(f_{e,t})$$

$$s.t \quad p_{z,t} + \sum_{e \in z^+} f_{e,t} = d_{z,t} + \sum_{e \in z^-} f_{e,t} \quad )\lambda_{z,t}$$

$$\implies \lambda_t \in \partial \tilde{Q}(d_t)$$

# Value function of the hydro-electric production

- The cost of hydro-electric production is represented by the Bellman function associated with the value of the water storage level for the future periods.
- Assuming perfect and pure competition, producers coordinate the value of the water global storage to offer energy at the best price on the whole horizont. : $x_{t+1} \mapsto V_{t+1}(x_{z=1,t+1}, \ldots, x_{z=|Z|,t+1})$.

$$Q_t(d_t, x_t) = \min_{p_t, f_t, x_{t+1}, u_t} \quad \sum_{z \in Z} c_z(p_{z,t}) + \sum_{e \in E} l_e(f_{e,t}) + V_{t+1}(x_{t+1})$$

$$s.t \quad p_{z,t} + u_{z,t} + \sum_{e \in z^+} f_{e,t} = d_{z,t} + \sum_{e \in z^-} f_{e,t} \quad )\lambda_{z,t}$$

$$x_{z,t+1} = x_{z,t} - u_{z,t} + i_{z,t}$$

$$\implies \lambda_{z,t} \in \partial_{d_t} Q_t(d_t, x_t)$$

# Joint value of the storage

Function $V_t(\cdot)$ represents the future cost beginning with a given water level :

$$V_t(x_t) = \min_{(p_\tau, u_\tau, f_\tau, x_\tau))} \sum_{\tau=t,\dots,T-1} \left\{ \begin{array}{l} \sum_{z\in Z} c_z(p_{z,\tau}) \\ +\sum_{e\in E} l_e(f_{e,\tau}) \end{array} \right\}$$

$$s.c \quad p_{z,\tau} + u_{z,\tau} + \sum_{e\in z^+} f_{e,\tau} = d_{z,\tau} + \sum_{e\in z^-} f_{e,\tau}$$

$$x_{z,\tau+1} = x_{z,\tau} - u_{z,\tau} + i_{z,\tau}$$

# Modelling with uncertainties

- Let consider a probability space $(\Omega, \mathcal{F}, \mu)$ and random variables $\omega \mapsto \boldsymbol{d}_{z,t}(\omega)$ et $\omega \mapsto \boldsymbol{i}_{z,t}(\omega)$ associated with the demand and the water supply $d_{z,t}$ and $i_{z,t}$ .
- $\mathcal{F}_t$ is the filtration generated by the past history of the random events up to period $t$:

$$\mathcal{F}_t = \sigma((\boldsymbol{i}_{z,\tau}, \boldsymbol{d}_{z,\tau})_{z \in Z, \tau=0,\ldots,t})$$

$$\boldsymbol{v}_t \preceq \mathcal{F}_t \implies \exists \varphi(\cdot) / \boldsymbol{v}_t = \varphi((\boldsymbol{i}_{z,\tau}, \boldsymbol{d}_{z,\tau})_{z \in Z, \tau=0,\ldots,t})$$

# Joint storage value

Function $V_t(\cdot)$ is defined as the solution of a multistage stochastic optimization problem:

$$
\begin{aligned}
& V_t(x_t) = \\
& \min_{(\boldsymbol{p}_\tau, \boldsymbol{u}_\tau, \boldsymbol{f}_\tau, \boldsymbol{x}_\tau) \in L^2(\Omega)} \mathbb{E}\left( \sum_{\tau=t,\dots,T-1} \left\{ \begin{array}{l} \sum_{z \in Z} c_z(\boldsymbol{p}_{z,\tau}) \\ + \sum_{e \in E} l_e(\boldsymbol{f}_{e,\tau}) \end{array} \right\} \right) \\
& s.c \quad \boldsymbol{p}_{z,\tau} + \boldsymbol{u}_{z,\tau} + \sum_{e \in z^+} \boldsymbol{f}_{e,\tau} = \boldsymbol{d}_{z,\tau} + \sum_{e \in z^-} \boldsymbol{f}_{e,\tau} \\
& \quad \boldsymbol{x}_{z,\tau+1} = \boldsymbol{x}_{z,\tau} - \boldsymbol{u}_{z,\tau} + \boldsymbol{i}_{z,\tau} \\
& \quad \boldsymbol{x}_t \equiv x_t \\
& \quad (\boldsymbol{x}_{\tau+1}, \boldsymbol{u}_\tau, \boldsymbol{f}_\tau, \boldsymbol{p}_\tau) \preceq \mathcal{F}_t
\end{aligned}
$$

# Dynamic programming equations

When the random events are independent on time :$(\boldsymbol{d}_{z,t}, \boldsymbol{i}_{z,t})_{z \in Z} \perp\!\!\!\perp \mathcal{F}_{t-1}$, alors $V_t(\cdot)$ verifies the dynamic programming equations:

$$
\begin{aligned}
V_T(x_T) &\equiv 0 \\
V_t(x_t) &= \mathbb{E}\left[Q_t(\boldsymbol{d}_t, \boldsymbol{i}_t, x_t)\right] \\
Q_t(d_t, i_t, x_t) &= \min_{p_t, x_{t+1}, u_t} \left[\sum_{z \in Z} c_z(p_{z,t}) + \sum_{e \in E} l_e(f_{e,t}) + V_{t+1}(x_{t+1})\right] \\
&\quad s.c\ p_{z,t} + u_{z,t} + \sum_{e \in z^+} f_{e,t} = d_{z,t} + \sum_{e \in z^-} f_{e,t} \\
&\quad x_{z,t+1} = x_{z,t} - u_{z,t} + i_{z,t}
\end{aligned}
$$

# Facing the curse of dimensionality.

- To solve numerically the dynamic programming equations is too complex when the dimension of the state $x_t$ is too large.

$\Longrightarrow$ We should simplify the problem to turn the computation of the value functions decentralized : for each zone $z$, compute $V_{z,t}(x_{z,t})$.

# Incomplete information

When optimizing the storage, we suppose that the decision maker doesn't know the exact quantity he should produce but just an average estimation conditioned by an information $\sigma(\boldsymbol{y}_{z,t})$:

$$
\begin{aligned}
& V_t(x_t) = \\
& \min_{(\boldsymbol{p}_\tau,\boldsymbol{u}_\tau,\boldsymbol{f}_\tau,\boldsymbol{x}_\tau)\in L^2(\Omega)} \mathbb{E}\left(\sum_{\tau=t,\dots,T-1} \left\{ \begin{array}{l} \sum_{z\in Z} c_z(\boldsymbol{p}_{z,\tau}) \\ +\sum_{e\in E} l_e(\boldsymbol{f}_{e,\tau}) \end{array} \right\}\right) \\
& \qquad s.c \quad \boldsymbol{p}_{z,\tau} + \boldsymbol{u}_{z,\tau} = \mathbb{E}\left(\boldsymbol{d}_{z,\tau} + \sum_{e\in z^-} \boldsymbol{f}_{e,\tau} - \sum_{e\in z^+} \boldsymbol{f}_{e,\tau} | \boldsymbol{y}_{z,t}\right) \\
& \qquad\qquad \boldsymbol{x}_{z,\tau+1} = \boldsymbol{x}_{z,\tau} - \boldsymbol{u}_{z,\tau} + \boldsymbol{i}_{z,\tau} \\
& \qquad\qquad \boldsymbol{x}_t \equiv x_t \\
& \qquad\qquad (\boldsymbol{x}_{\tau+1}, \boldsymbol{u}_\tau, \boldsymbol{f}_\tau, \boldsymbol{p}_\tau) \preceq \mathcal{F}_t
\end{aligned}
$$

# Examples of information.

1. If $\boldsymbol{y}_{z,t} = \{\boldsymbol{d}_{z,\tau}, \boldsymbol{i}_{z,\tau}\}_{0 \leq \tau \leq t, z \in Z}$ then $\sigma(\boldsymbol{y}_{z,t}) = \mathcal{F}_t$ and we retrieve the original problem.
2. If $\boldsymbol{y}_{z,t} \equiv 0$ (or any constant) then the decisor of the zone should decide to satisfy an average demand.
3. If $\boldsymbol{y}_{z,t} = \boldsymbol{d}_{z,t}$ then the decisor of the zone should decide to satisfy the demand plus the expected export and import flows knowing the internal demand.

# Decentralization

The problem of determining $V_t$ can be rewritten as the minimization of the sum of two functions:

$$V_t(x_t) = \min_{\boldsymbol{q}_{t,T} \in L^2(\Omega)} \quad C_t(\boldsymbol{q}_{t,T}) + \underbrace{\sum_{z \in Z} F_{z,t}(\boldsymbol{q}_{z,t,T}, x_{z,t})}_{F(\boldsymbol{q}_{t,T}, x_t)}$$

where $F_{z,t}$ is the expected local production cost of the quantity $\boldsymbol{q}^z$:

$$\begin{aligned} F_{z,t}(\boldsymbol{q}_{z,t,T}, x_{z,t}) = \min_{\boldsymbol{p},\boldsymbol{x},\boldsymbol{u} \in L^2(\Omega)} \quad & \mathbb{E}\left[\sum_{\tau=t}^{T-1}\left(\sum_{z \in Z} c_{z,\tau}(\boldsymbol{p}_{z,\tau})\right)\right] \\ s.c \quad & \boldsymbol{p}_{z,\tau} + \boldsymbol{u}_{z,\tau} = \mathbb{E}\left(\boldsymbol{q}_{z,\tau} | \boldsymbol{y}_{z,\tau}\right) \\ & \boldsymbol{x}_{z,\tau+1} = \boldsymbol{x}_{z,\tau} - \boldsymbol{u}_{z,\tau} + \boldsymbol{i}_{z,\tau} \\ & \boldsymbol{p}_{z,\tau}, \boldsymbol{u}_{z,\tau}, \boldsymbol{x}_{z,\tau+1} \preceq \mathcal{F}_\tau \end{aligned}$$

# Decentralization

... and $C_t(\boldsymbol{q}_{t,T})$ is the expected transportation cost coupling the local production processes $\boldsymbol{q}_{t,T}$:

$$C_t(\boldsymbol{q}_{t,T}) = \min_{\boldsymbol{f} \in L^2(\Omega)} \quad \mathbb{E}\left[\sum_{\tau=t}^{T-1} \left(\sum_{e \in E} l_e(\boldsymbol{f}_{e,\tau})\right)\right]$$

$$s.c \quad \boldsymbol{d}_{z,\tau} + \sum_{e \in z^-} \boldsymbol{f}_{e,\tau} - \sum_{e \in z^+} \boldsymbol{f}_{e,\tau} = \boldsymbol{q}_{z,\tau}$$

$$\boldsymbol{f}_{e,\tau} \preceq \mathcal{F}_\tau$$

# Suppose the optimal flows are known ...

If we choose $y_{z,t}$ such that $y_{z,t} \preceq \mathcal{F}_t$ and $y_{z,t} \perp\!\!\!\perp \mathcal{F}_{t-1}$ then $F_{z,t}(\boldsymbol{q}_{z,t,T}, x_{z,t})$ can be computed using Dynamic Programming.

# Reformulation of the static model

Consider the multizonal model and relax first the dynamic features to focus on decoupling the zones to get the following model :

$$
\begin{aligned}
&\textbf{(P)} \\
\min_{p,f} \quad & \sum_{z \in Z} c_z(p_z) + \sum_{e \in E} l_e(f_e) \\
s.t \quad & p_z + \sum_{e \in z^+} f_e = d_z + \sum_{e \in z^-} f_e, \quad \forall z \in Z \\
& p_z \in P_z \quad \forall z, f_e \geq 0 \quad \forall e \in E
\end{aligned}
$$

Observe that we did not include the control variable $u_z$ as it is embedded in $p_z$

# Coupling constraints

Let rewrite the constraints in matrix form as :

$$p + Af = d$$

where $A$ is the $(n \times n(n-1))$ incidence node-arc matrix for graph $G$. We can reorganize the columns (arcs) of $A$ by ordering the zones and the outgoing arcs in each zone :

$$A = \left[\begin{array}{ccc|ccc|ccc|ccc}
-1 & -1 & -1 & 1 & & & 1 & & & 1 & & \\
\hline
1 & & & -1 & -1 & -1 & & 1 & & & 1 & \\
\hline
& 1 & & & 1 & & -1 & -1 & -1 & & & 1 \\
\hline
& & 1 & & & 1 & & & 1 & -1 & -1 & -1
\end{array}\right]$$

Each diagonal block is denoted by $A_z$. Let too $f_z$ denote the vector of outgoing flows $f_e, e \in z^+$.

# Variable splitting reformulation (1)

We introduce copies $\phi_{ez}$ of the ingoing flows for each zone $z$ associated with each 1 in row $z$ of the incidence matrix to obtain the following equivalent reformulation :

$$p_z - A_z f_z + \sum_{e \in z^-} \phi_{ez} = d_z$$

$$\phi_{ez} = f_{ez'}, \quad \text{for} \quad e = (z', z) \in {z'}^+ \cap z^-$$

Observe that the demand equation is now completely decentralized with additional variables $\phi_{ez}$. The second set of equations represents a coupling subspace, denoted hereafter by $\mathcal{A}$, between zones in $I\!R^{2n(n-1)}$.

# Variable splitting reformulation (2)

**(P')**

$$\min_{p,f} \quad \sum_{z\in Z}[c_z(p_z) + \sum_{e\in z^+} l_e(f_e)]$$

$$s.t \quad p_z - A_z f_z + \sum_{e\in z^-} \phi_{ez} = d_z \quad \forall z \in Z$$

$$p_z \in P_z \forall z, f_e \geq 0, \phi_e \geq 0, \forall e \in E$$

$$(f, \phi) \in \mathcal{A}$$

which can be condensed into :

$$\min_{f_z,\phi_z} \quad \sum_{z\in Z} C_z(p_z, f_z, \phi_z) \quad \text{s.t.} \quad (f, \phi) \in \mathcal{A}$$

# Proximal Decomposition

We describe now the *Proximal Decomposition* algorithm.
The key ingredient is that we need a pair of primal and dual variables lying respectively in subspace $\mathcal{A}$ and its orthogonal $\mathcal{A}^{\perp}$. We denote the dual variables by $(u, v)$ so that we have the following dual relations :

$$X = \begin{bmatrix} f \\ \phi \end{bmatrix} = \begin{bmatrix} I\!I \\ I\!I \end{bmatrix} \alpha, \text{for some}\, \alpha \in I\!R^{n(n-1)} \quad (1)$$

$$W = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} I\!I \\ -I\!I \end{bmatrix} \beta, \text{for some}\, \beta \in I\!R^{n(n-1)} \quad (2)$$

The method alternates proximal steps on the primal variables and projection steps on each subspace.

# Zonal subproblem

$$\begin{aligned} \min_{p_z, X_z} \quad & C_z(p_z, X_z) + \tfrac{1}{2\mu}\|X_z - X_z^t - \mu W_z^t\|^2 \\ \text{s.t.} \quad & p_z - A_z f_z + \textstyle\sum_{e\in z^-} \phi_{ez} = d_z \\ & p_z \in P_z, X_z \geq 0 \end{aligned}$$

which solution is denoted by $\tilde{X_z}^t$. The dual update is given by :

$$\tilde{W_z}^t = \mu^{-1}(X_z^t + \mu W_z^t - \tilde{X_z}^t)$$

# Projection step

$$\begin{aligned} f_z^{t+1} = \phi_z^{t+1} = & \quad \frac{\tilde{f_z}^t + \tilde{\phi_z}^t}{2} \\ u_z^{t+1} = -v_z^{t+1} = & \quad \tilde{u_z}^t - \frac{\tilde{u_z}^t + \tilde{v_z}^t}{2} \end{aligned}$$

Observe that $X^t \in \mathcal{A}$ and $W^t \in \mathcal{A}^\perp, \forall t$, so that $f^t$ is a feasible flow for graph $G$. On the other hand, $\tilde{f}^t$ is not feasible but it will converge asymptotically.
Numerical enhancements :

- Adjustable scaling parameter $\mu^t$
- Additional relaxation parameter with Generalized Douglas-Rachford

# “Ideal” prices in practice

Get a sample $\{\omega^\nu\}_{\nu=0,\dots,N-1} \subset \Omega$. In practice, we can solve the backward dynamic programming equation:

$$
\begin{aligned}
V_T(\boldsymbol{x}_t) \equiv& 0 \\
V_t(\boldsymbol{x}_t) =& \mathbb{E}_N \left[ \min_{\boldsymbol{p}_t, \boldsymbol{x}_t, \boldsymbol{u}_t} \sum_{z \in Z} c_z(p_{z,t}) + \sum_{e \in E} l_e(f_{e,t}) + V_{t+1}(\boldsymbol{x}_{t+1}) \right] \\
& s.c \quad p_{z,t} + u_{z,t} + \sum_{e \in z^+} f_{e,t} = d_{z,t} + \sum_{e \in z^-} \phi_{e,t} \quad \text{(supplying)} \\
& x_{z,t+1} = x_{z,t} - u_{z,t} + i_{z,t} \quad \text{(stock balance)}
\end{aligned}
$$

Then we proceed in the forward phase to simulate strategy and get dual variables

# Conclusions

- Alternative reformulation to obtain completely decentralized zonal subsystems
- Many splitting algorithms can be compared on the static model
- First interpretation for the long-term pricing strategy
- Ongoing PGMO project with EDF R & D

# Perspectives

- Study the numerical behaviour of the splitting methods
- Analyze the loss of information with different scenarios
- Further decompose the zonal subproblems with different production units
- Reformulate as a competitive equilibrium problem
- Compare with Mixed Decomposition

# More on Mixed Decomposition (1)

Back to the static model, introducing a single auxiliary variable for each zone :

$$\begin{aligned} p_z - A_z f_z &= \quad d_z - \Phi_z \\ \sum_{e \in z^-} f_{ez'} - \Phi_z &= \quad 0, e = (z', z) \end{aligned}$$

By dualizing the second (coupling) constraint, we get the following subproblem :

$$\begin{aligned} \min_{p_z, f_z} \quad & C_z(p_z, f_z) + \textstyle\sum_{e=(z,z')} \lambda_z^t f_{ez} \\ \text{s.t.} \quad & p_z - A_z f_z = d_z - \Phi_z^t \\ & p_z \in P_z, f_z \geq 0 \end{aligned}$$

# More on Mixed Decomposition (2)

Direct coordination between zones :
Subproblem $z$ computes $f_z^t$ and $\lambda_z^t$.

$$\Phi_z^{t+1} = \sum_{e=(z',z)} f_{ez'}^t$$

Each subproblem updates primal allocations $\Phi_z$ and transfers dual prices to its neighbours.
*Prediction of interaction* (see G. Cohen 1982)