Summer School CEA-EDF-INRIA 2012
Stochastic Optimization
$\Diamond$
Information Constraints in Stochastic Control

P. Carpentier — *ENSTA ParisTech*
J.-P. Chancelier — *École des Ponts ParisTech*
M. De Lara — *École des Ponts ParisTech*

June 25, 2012 — July 6, 2012

Decomposition-coordination methods (P. Carpentier)

Wednesday, July 4, 2012

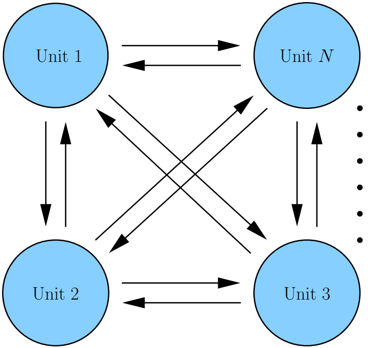
Aim of the lecture

Obtain reasonable strategies on large scale stochastic optimal control problems, for example the optimal management over a given time horizon of a system involving a large amount of production units.

- Use Dynamic Programming (or related methods) in order to obtain the decision strategies:
  - Markovian case,
  - curse of dimensionality.
- Specify the information patterns so that the optimization problems can be solved by decomposition/coordination.

Lecture outline

Decomposition and coordination



Interconnected units

- The system to be optimized consists of interconnected subsystems: we want to use this structure in order to formulate optimization subproblems of reasonable complexity.
- But the presence of interactions requires a level of coordination.
- Coordination must provide a local model of the interactions to each subproblem: it is an iterative process.
- The ultimate goal is to obtain the solution of the overall problem by concatenation of the solutions of subproblems.

Decomposition and decentralization

In the standard stochastic optimization framework, one considers that the entire information is in the hand of a single decision maker. This situation is that of centralized information.

In team problems, several decision makers have their private information and try to optimize the system in a collaborative manner. This situation is termed decentralization.

The last situation includes the case of non classical information patterns. It can be a source of huge difficulties in stochastic optimal control (dual effect, Witsenhausen counterexample).

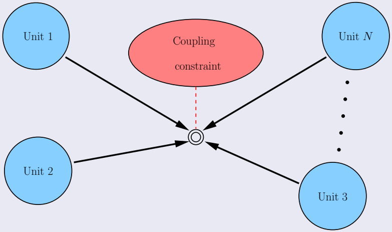
In this lecture, we only deal with the centralized case: only the computations are distributed in the decomposition/coordination approach.

Two prototype models

"Flower model"

$$\min_{u} \sum_{i=1}^{N} J_i(u_i) \ ,$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \Theta_i(u_i) = \theta \ .$$

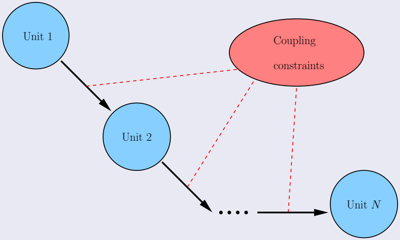


"Cascade model"

$$\min_{u,v} \sum_{i=1}^{N} J_i(u_i, v_i) \ ,$$

$$\text{s.t.} \quad H_i(u_i, v_i) = v_{i+1} \ .$$

Link with the flower model:
$\Theta_i \rightsquigarrow \big(0, \ldots, -v_i, H_i(u_i, v_i), \ldots, 0\big)^\top$.



General model

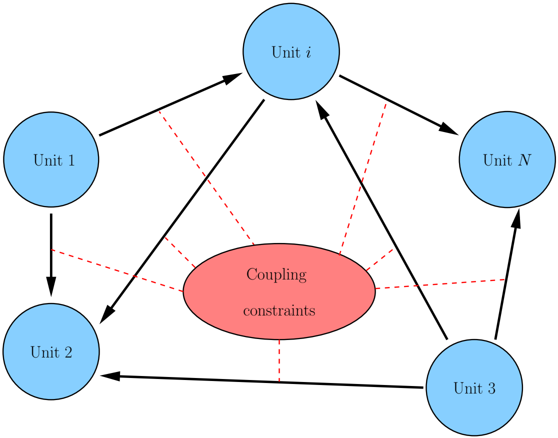


Figure: Units interacting through a network (smartgrid?)

Price decomposition (flower model)

$$\min_{u} \sum_{i=1}^{N} J_i(u_i) \quad \text{subject to} \quad \sum_{i=1}^{N} \Theta_i(u_i) - \theta = 0 \; .$$

1. Form the Lagrangian of the problem and assume that a saddle point exists. The initial problem is equivalent to:

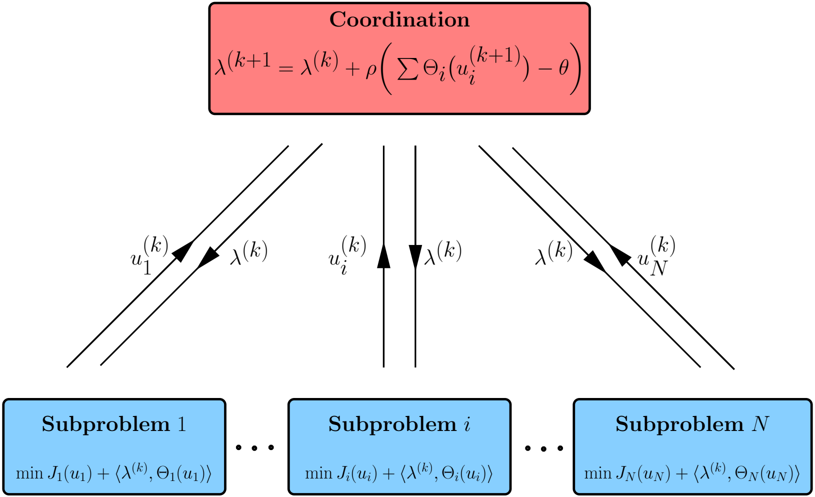
$$\max_{\lambda} \min_{u} \sum_{i=1}^{N} \Big(J_i(u_i) + \langle \lambda \, , \Theta_i(u_i) \rangle \Big) - \langle \lambda \, , \theta \rangle \; ,$$

2. Solve this problem by the Uzawa algorithm:

$$u_i^{(k+1)} \in \arg\min_{u_i} J_i(u_i) + \langle \lambda^{(k)} \, , \Theta_i(u_i) \rangle \; , \quad \forall i \; ,$$

$$\lambda^{(k+1)} = \lambda^{(k)} + \rho \bigg(\sum_{i=1}^{N} \Theta_i \Big(u_i^{(k+1)} \Big) - \theta \bigg) \; .$$

Price decomposition



Resource allocation (flower model)

$$\min_u \sum_{i=1}^N J_i(u_i) \quad \text{subject to} \quad \sum_{i=1}^N \Theta_i(u_i) - \theta = 0 \; .$$

1. Form an equivalent problem obtained by introducing new variables $(v_1, \dots, v_N)$ ("allocation") and new constraints:

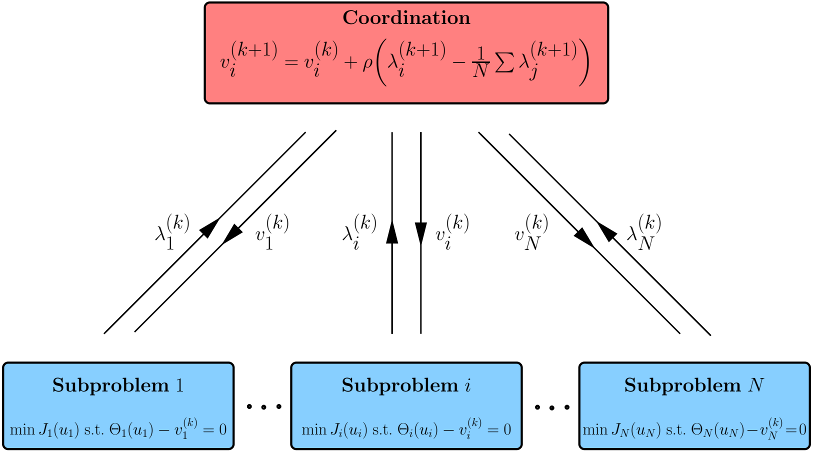
$$\min_v \sum_{i=1}^N \Big(\min_{u_i} J_i(u_i) \ \text{ s.t. } \ \Theta_i(u_i) - v_i = 0 \Big) \ \text{ s.t. } \ \sum_{i=1}^N v_i = \theta \; ,$$

2. Solve this problem by a projected gradient method w.r.t. v:

$$\min_{u_i} J_i(u_i) \ \text{ s.t. } \ \Theta_i(u_i) - v_i^{(k)} = 0 \quad \rightsquigarrow \quad \lambda_i^{(k+1)} \; , \ \ \forall i \; ,$$

$$v_i^{(k+1)} = v_i^{(k)} + \rho \Big(\lambda_i^{(k+1)} - \frac{1}{N} \sum_{j=1}^N \lambda_j^{(k+1)} \Big) \; .$$

Resource allocation



Prediction (cascade model)

$$\min_{u,v} \sum_{i=1}^{N} J_i(u_i, v_i) \quad \text{subject to} \quad v_{i+1} - H_i(u_i, v_i) = 0 \; .$$

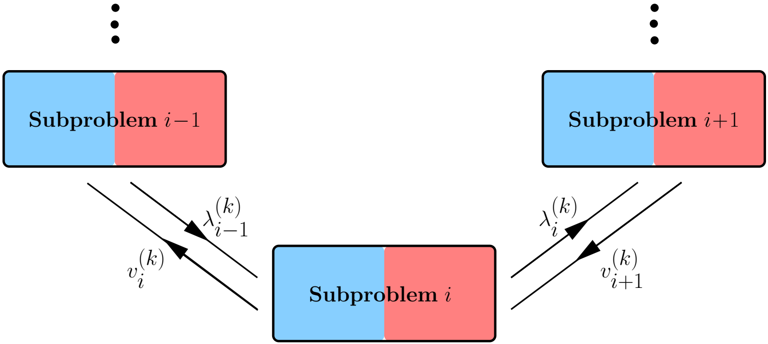
- At iteration k, the interaction variable v_{i+1} is predicted to subproblem i, hence the constraint: $v_{i+1}^{(k)} - H_i(u_i, v_i) = 0$,
- and the optimal multiplier $\lambda_i^{(k)}$ of this constraint is used to add the term $\langle \lambda_i^{(k)} \,, v_{i+1} \rangle$ in the cost of subproblem $i+1$.

The i-th subproblem at iteration k reads

$$\min_{u_i, v_i} J_i(u_i, v_i) + \langle \lambda_{i-1}^{(k)} \,, v_i \rangle \quad \text{s.t.} \quad v_{i+1}^{(k)} - H_i(u_i, v_i) = 0 \; ,$$

and the set of primal-dual solutions $\left(u_i^{(k+1)}, v_i^{(k+1)}, \lambda_i^{(k+1)}\right)$ is used to formulate the subproblems at the next iteration.

Prediction


Subproblem $i-1$
Subproblem $i+1$
$\lambda_{i-1}^{(k)}$
$v_i^{(k)}$
$\lambda_i^{(k)}$
$v_{i+1}^{(k)}$
Subproblem i

Auxiliary Problem Principle

The three decomposition schemes we have presented, which seem to crucially depend on the additive structure of the problems under consideration, can actually be generalized to problems of the form:

$$\min_{u} J(u_1, \ldots, u_N) \quad \text{subject to} \quad \Theta(u_1, \ldots, u_N) = 0 \ .$$

This generalization is achieved by the Auxiliary Problem Principle (APP), which consists in linearizing J and Θ around a point $u^{(k)}$:

$$J(u) \rightsquigarrow \sum_{i=1}^{N} \left\langle \nabla_{u_i} J(u^{(k)}) \, , u_i \right\rangle \quad , \quad \Theta(u) \rightsquigarrow \sum_{i=1}^{N} \left(\Theta'_{u_i}(u^{(k)})\right)^{\top} \cdot u_i \ ,$$

in order to obtain again an additive structure.

For further readings on decomposition/coordination and APP, see:

- G. Cohen, "Auxiliary Problem Principle and Decomposition of Optimization Problems", *Journal of Optimization Theory and Applications*, **32**, 1980.
- G. Cohen, "Optimisation des grands systèmes", *Cours du DEA Modélisation et Méthodes Mathématiques en Économie*, 2004.

General remark about decomposition

Whatever the decomposition/coordination scheme used (price, allocation, prediction, APP), we have observed that new variables, namely primal variables $u^{(k)}$ and/or dual variables $\lambda^{(k)}$, appear in the subproblems arising at iteration k of the optimization process.

All these new variables are considered as fixed when solving the subproblems (they only depend on the iteration index k). They are nothing but constants, and therefore do not cause any difficulty in the deterministic case.

Stochastic optimal control problems

We consider a SOC problem in the following form:

$$\min_{\mathbf{U},\mathbf{X}} \mathbb{E}\Big(\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) + K(\mathbf{X}_T) \Big) , \tag{1a}$$

subject to the constraints:

$$\mathbf{X}_0 = f_{\text{-}1}(\mathbf{W}_0) , \tag{1b}$$

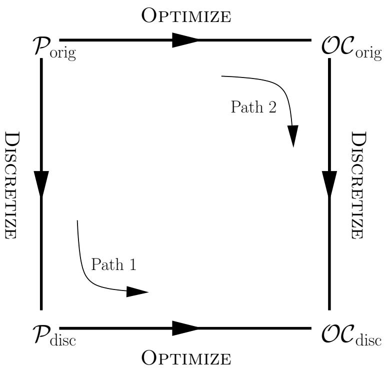
$$\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) , \quad \forall t = 0, \ldots, T-1 , \tag{1c}$$

$$\mathbf{U}_t \preceq \mathcal{F}_t := \sigma(\mathbf{W}_0, \ldots, \mathbf{W}_t) , \quad \forall t = 0, \ldots, T-1 , \tag{1d}$$

$$\mathbf{U}_t \in C_t , \quad \forall t = 0, \ldots, T-1 , \tag{1e}$$

with $\mathbf{U}_t = (\mathbf{U}_t^1, \ldots, \mathbf{U}_t^N)$ and $\mathbf{X}_t = (\mathbf{X}_t^1, \ldots, \mathbf{X}_t^N)$, with N large.

Approximation schemes for SOC problems



1. Following Path 1 (discretize, then optimize), we solve a deterministic approximation of Problem (1). All decomposition methods are thus available, and some of them are more specifically well-suited for an implementation in this context (progressive hedging).
2. Following Path 2 (optimize, then discretize) we directly make use of a decomposition/coordination method on Problem (1) and then discretize the subproblems.
   $\rightsquigarrow$ **Today's lecture**.

Dynamic Programming yields centralized controls

Remember that we want to solve Problem (1) either by variational approaches or by Dynamic Programming (and related methods, e.g. SDDP). For the sake of simplicity, we assume that we are in the Markovian setting and restrict ourselves to the DP approach.

The system is made of N interconnected subsystems. We denote by $\mathbf{U}_t^i$ and $\mathbf{X}_t^i$ the command and state of subsystem i at time t. We know that the optimal control of subsystem i is a function of the whole system state, that is, a centralized feedback:

$$\mathbf{U}_t^i = \gamma_t^i(\mathbf{X}_t^1, \ldots, \mathbf{X}_t^N) \ ,$$

whereas any decentralized feedback, that is, a control such as

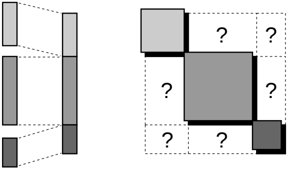
$$\mathbf{U}_t^i = \widehat{\gamma}_t^i(\mathbf{X}_t^i)$$

will be, in most cases, far from being optimal. . .

Straightforward decomposition of Dynamic Programming?

The crucial point is that the optimal feedback of a subsystem a priori depends on the state of all other subsystems, so that using a decomposition scheme by subsystems is far from being obvious...

As far as we have to deal with Dynamic Programming, the central concern for decomposition/coordination purpose is resumed as:



- how to decompose a feedback γ_t w.r.t. its domain $\mathbb{X}_t$ rather than its range $\mathbb{U}_t$?

  And the answer is:
- impossible in the general case!

In the deterministic case, there is no reason to search for feedbacks and the optimal controls correspond to values rather than functions.

Decomposition/coordination in the stochastic case

- Apply a decomposition/coordination scheme, for example the price decomposition, to a SOC problem subject to coupling constraints (theory available in general Hilbert spaces).
- As pointed out in the deterministic case, new variables, that is, dual multipliers $\boldsymbol{\Lambda}_t^{(k)}$, appear in the subproblems arising at iteration k: these variables, given at this stage of calculation, corresponds to random variables.
- The process $\boldsymbol{\Lambda}^{(k)}$ acts as an additional input (data) in the subproblems, but the structure of this process is a priori unknown: it may be correlated in time, so that the white noise assumption, crucial for Dynamic Programming, has no reason to be fulfilled in that context!

Summary

- On the one hand, it seems that Dynamic Programming cannot be decomposed in a straightforward manner.
- On the other hand, applying a decomposition scheme to a SOC problem introduces coordination instruments in the subproblems, that is, multipliers $\boldsymbol{\Lambda}_t^{(k)}$ in the case of price decomposition, which correspond to additional fixed random variables whose structure is unknown.

Question: how to handle these coordination instruments in order to obtain (an approximation of) the overall optimum of the problem?

Simplified energy management problem

$$\min_{\mathbf{U},\mathbf{X}} \mathbb{E}\left(\sum_{i=1}^{N} \left(\sum_{t=0}^{T-1} \frac{c^i}{2} (\mathbf{U}_t^i)^2 + \frac{d^i}{2} (\mathbf{X}_T^i - x_{\mathrm{f}}^i)^2 \right) \right) ,$$

subject to:

$$\mathbf{X}_{t+1}^i = \mathbf{X}_t^i - \mathbf{U}_t^i + \mathbf{A}_{t+1}^i ,$$

$$\mathbf{D}_t - \sum_{i=1}^{N} \mathbf{U}_t^i = 0 ,$$

$$\mathbf{D}_t - \sum_{i=1}^{N} \mathbf{U}_t^i = 0 ,$$

$$\mathbf{U}_t^i \preceq \sigma\big(\mathbf{A}_0, \ldots, \mathbf{A}_t, \mathbf{D}_0, \ldots, \mathbf{D}_t\big) ,$$

with $\mathbf{A}_t = (\mathbf{A}_t^1, \ldots, \mathbf{A}_t^N)$.

Price decomposition

Consider the subproblem i obtained by price decomposition:

$$\min_{\mathbf{U}^i,\mathbf{X}^i} \mathbb{E}\bigg(\sum_{t=0}^{T-1} \Big(\frac{c^i}{2} (\mathbf{U}_t^i)^2 - \mathbf{\Lambda}_t \cdot \mathbf{U}_t^i \Big) + \frac{d^i}{2} (\mathbf{X}_T^i - x_{\mathrm{f}}^i)^2 \bigg) ,$$

subject to:

$$\mathbf{X}_{t+1}^i = \mathbf{X}_t^i - \mathbf{U}_t^i + \mathbf{A}_{t+1}^i ,$$
$$\mathbf{U}_t^i \preceq \sigma\big(\mathbf{A}_0, \ldots, \mathbf{A}_t, \mathbf{D}_0, \ldots, \mathbf{D}_t\big) ,$$

It incorporates the stochastic process $(\mathbf{\Lambda}_0, \ldots, \mathbf{\Lambda}_{T-1})$, so that Dynamic Programming cannot be applied considering only the "state" variable $\mathbf{X}_t^i$.

The notation $\mathbf{\Lambda}_t \cdot \mathbf{U}_t^i$ is used to represent the standard product of the values taken by the two random variables, namely $\mathbf{\Lambda}_t(\omega)\mathbf{U}_t^i(\omega)$. The scalar product (in L^2) of these variables is: $\langle \mathbf{\Lambda}_t \, , \mathbf{U}_t^i \rangle = \mathbb{E}\big(\mathbf{\Lambda}_t \cdot \mathbf{U}_t^i\big) = \int_\Omega \mathbf{\Lambda}_t(\omega)\mathbf{U}_t^i(\omega) \, \mathrm{d}\mathbb{P}(\omega)$.

Dynamics of the multipliers

Proposition 1 (Strugarek)

Assume that the random variables $\mathbf{W}_t = (\mathbf{A}_t, \mathbf{D}_t)$ *are independent over time (white noise). Let* $\beta = \sum_{i=1}^{N} 1/c^i$ *and assume that*

$$\exists \alpha > 0 \ , \quad c^i = \alpha d^i \ , \quad \forall i = 1, \ldots, N \ ,$$

Then the optimal multiplier process $\mathbf{\Lambda}^\sharp$ *is such that*

$$\mathbf{\Lambda}_0^\sharp \quad = \frac{1}{\beta}\left((1-\alpha)\mathbf{D}_0 - \alpha \sum_{\tau=1}^{T-1} \mathbb{E}\big(\mathbf{D}_\tau\big) - \alpha \sum_{\tau=1}^{T} \mathbb{E}\big(\mathbf{A}_\tau\big)\right) \ ,$$

$$\mathbf{\Lambda}_{t+1}^\sharp = \mathbf{\Lambda}_t^\sharp + \frac{1}{\beta}\Bigg((1+\alpha)\mathbf{D}_{t+1} - \mathbf{D}_t - \alpha\mathbb{E}\big(\mathbf{D}_{t+1}\big) - \alpha\Big(\mathbf{A}_{t+1} - \mathbb{E}\big(\mathbf{A}_{t+1}\big)\Big)\Bigg) \ .$$

Price decomposition and Dynamic Programming

Thanks to Proposition 1, each subproblem i can be solved using Dynamic Programming provided we take into account both the dynamics of $\boldsymbol{\Lambda}_t^\sharp$ and the dynamical effect induced by $\mathbf{D}_t$.

This leads to use the 3-dimensional state variable $(\mathbf{X}_t^i, \boldsymbol{\Lambda}_t, \boldsymbol{\Xi}_t)$:

$$\mathbf{X}_{t+1}^i = \mathbf{X}_t^i - \mathbf{U}_t^i + \mathbf{A}_{t+1}^i \ ,$$

$$\boldsymbol{\Lambda}_{t+1} = \boldsymbol{\Lambda}_t + \frac{1}{\beta}\Big((1+\alpha)\mathbf{D}_{t+1} - \boldsymbol{\Xi}_t - \alpha\mathbb{E}\big(\mathbf{D}_{t+1}\big) - \alpha\Big(\mathbf{A}_{t+1} - \mathbb{E}\big(\mathbf{A}_{t+1}\big)\Big)\Big) \ ,$$

$$\boldsymbol{\Xi}_{t+1} = \mathbf{D}_{t+1} \ .$$

The optimal solution of the N-dimensional state initial problem is obtained by solving N subproblems with state dimension equal to 3. The optimal control is of the form: $\mathbf{U}_t^{i\,\sharp} = \gamma_t^{i\,\sharp}\big(\mathbf{X}_t^i, \boldsymbol{\Lambda}_t, \mathbf{D}_t\big)$.

What can be drawn from this example

Although the multiplier $\boldsymbol{\Lambda}_t$ a priori depends on $\left(\mathbf{X}_t^1, \ldots, \mathbf{X}_t^N, \mathbf{D}_t\right)$ in the Markovian case,[1] we saw in our example that $\boldsymbol{\Lambda}_t$ may depend only on a reduced subset of variables. So we were able to break the curse of dimensionality and to implement Dynamic Programming.

In the sequel, we will try to identify a reduced subset of variables which "explains" the multiplier $\boldsymbol{\Lambda}_t$ in order to efficiently solve the subproblems which arise in price decomposition. Usually, such an identification will be done only approximately: indeed, the very specific situation of Strugarek's example (which led to an exact multiplier dynamics) can not be extended to the general case.

Identifying directly a dynamics for the multiplier (e.g. using ARMA models) is not satisfactory: thus, the subset of AR(1) processes is not convex!

[1] and on the past noises $\left(\mathbf{A}_0, \ldots, \mathbf{A}_t, \mathbf{D}_0, \ldots, \mathbf{D}_t\right)$ in the general case

Optimization problem

The problem is a "flower model" stochastic optimization problem:

$$\min_{\mathbf{U},\mathbf{X}} \mathbb{E}\left(\sum_{i=1}^{N} \left(\sum_{t=0}^{T-1} L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) + K^i(\mathbf{X}_T^i) \right) \right) ,$$

subject to dynamics constraints:

$$\mathbf{X}_0^i = f_{-1}^i(\mathbf{W}_0) ,$$
$$\mathbf{X}_{t+1}^i = f_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) ,$$

to measurability constraints:

$$\mathbf{U}_t^i \preceq \sigma(\mathbf{W}_0, \dots, \mathbf{W}_t, \mathbf{W}_{t+1}) , \quad \textbf{Hazard-Decision setting}$$

and to instantaneous coupling constraints

$$\sum_{i=1}^{N} g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) = 0 . \qquad \textbf{Feasible constraints}$$

Optimization problem

The problem is a "flower model" stochastic optimization problem:

$$\min_{\mathbf{U},\mathbf{X}} \mathbb{E}\bigg(\sum_{i=1}^{N} \Big(\sum_{t=0}^{T-1} L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) + K^i(\mathbf{X}_T^i) \Big) \bigg) , \tag{2a}$$

subject to dynamics constraints:

$$\mathbf{X}_0^i = f_{-1}^i(\mathbf{W}_{-1}) , \tag{2b}$$

$$\mathbf{X}_{t+1}^i = f_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) , \tag{2c}$$

to measurability constraints:

$$\mathbf{U}_t^i \preceq \sigma(\mathbf{W}_{-1}, \ldots, \mathbf{W}_t) , \tag{2d}$$

and to instantaneous coupling constraints

$$\sum_{i=1}^{N} g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) = 0 . \tag{2e}$$

Decomposable problem setting

The considered system consists of N subsystems, whose dynamics and costs functions are independent one from another. The state and the control of the global system write

$$\mathbf{X}_t = \left(\mathbf{X}_t^1, \ldots, \mathbf{X}_t^N\right) ,$$
$$\mathbf{U}_t = \left(\mathbf{U}_t^1, \ldots, \mathbf{U}_t^N\right) ,$$

and all the random variables belong to Hilbert spaces:
$\mathbf{X}_t^i \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^{n_i})$, $\mathbf{U}_t^i \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^{m_i})$, $\mathbf{W}_t \in L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^{p})$.

Note that there are three types of coupling in the problem:

1. temporal coupling induced by the state dynamics,
2. informational coupling induced by nonanticipativity,
3. spatial coupling induced by the instantaneous constraint.

Assumptions

Assumption 1 (White noise)

Noises $\mathbf{W}_{-1}, \dots, \mathbf{W}_{T-1}$ are independent over time.

Remember that we have also assumed that we are in the case of full noise observation:

$$\mathbf{U}_t^i \preceq \sigma(\mathbf{W}_{-1}, \dots, \mathbf{W}_t) \ .$$

We thus are in the Markovian case , so that DP applies.

Notice that, in our Hazard–Decision framework, the decision maker at time t observes the current noise value $\mathbf{W}_t$ before choosing the control $\mathbf{U}_t$. In such a setting, the optimal decision at time t depends on both the state variable $\mathbf{X}_t$ and the noise variable $\mathbf{W}_t$ (whereas the Bellman function only depends on the state variable $\mathbf{X}_t$).

Lagrangian formulation

We dualize the coupling constraint and obtain the Lagrangian:

$$\mathcal{L}(\mathbf{X}, \mathbf{U}, \mathbf{\Lambda}) = \mathbb{E}\Bigg(\sum_{i=1}^{N} \bigg(\sum_{t=0}^{T-1} L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) + K^i(\mathbf{X}_T^i) + \sum_{t=0}^{T-1} \mathbf{\Lambda}_t \cdot g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) \bigg) \Bigg) ,$$

where $\mathbf{\Lambda}_t$ is a $\sigma(\mathbf{W}_{\text{-}1}, \dots, \mathbf{W}_t)$-measurable random variable.

Under standard assumptions, a saddle point of $\mathcal{L}$ exists, and

$$\min_{\mathbf{U},\mathbf{X}} \max_{\mathbf{\Lambda}} \mathcal{L}(\mathbf{X}, \mathbf{U}, \mathbf{\Lambda}) = \max_{\mathbf{\Lambda}} \min_{\mathbf{U},\mathbf{X}} \mathcal{L}(\mathbf{X}, \mathbf{U}, \mathbf{\Lambda}) .$$

Uzawa algorithm

At iteration k of the algorithm,

1. Solve subproblem i, $i = 1, \dots, N$, with $\mathbf{\Lambda}^{(k)}$ fixed:

$$\min_{\mathbf{U}^i, \mathbf{X}^i} \mathbb{E}\left(\sum_{t=0}^{T-1} \left(L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) + \mathbf{\Lambda}_t^{(k)} \cdot g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) \right) + K^i(\mathbf{X}_T^i) \right) , \tag{3a}$$

subject to

$$\mathbf{X}_{t+1}^i = f_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) , \tag{3b}$$

$$\mathbf{U}_t^i \preceq \sigma(\mathbf{W}_1, \dots, \mathbf{W}_t) , \tag{3c}$$

whose solution is denoted $\left(\mathbf{U}^{i,(k)}, \mathbf{X}^{i,(k)}\right)$.

2. Update the multipliers $\mathbf{\Lambda}_t$:

$$\mathbf{\Lambda}_t^{(k+1)} = \mathbf{\Lambda}_t^{(k)} + \rho_t \left(\sum_{i=1}^{N} g_t^i\left(\mathbf{X}_t^{i,(k)}, \mathbf{U}_t^{i,(k)}, \mathbf{W}_t\right) \right) . \tag{4}$$

Main idea of DADP

As already pointed out, $\boldsymbol{\Lambda}_t^{(k)}$ depends on $(\mathbf{W}_{-1}, \ldots, \mathbf{W}_t)$, so that solving a subproblem is as complex as solving the original problem.

In order to overcome the difficulty, let us choose at each time t a random variable $\mathbf{Y}_t^i$ which is measurable w.r.t. the past noises $(\mathbf{W}_{-1}, \ldots, \mathbf{W}_t)$. We call $\mathbf{Y}^i = (\mathbf{Y}_0^i, \ldots, \mathbf{Y}_{T-1}^i)$ the information process for subsystem i.

The core idea is to replace the multiplier $\boldsymbol{\Lambda}_t^{(k)}$ at iteration k by its conditional expectation w.r.t. $\mathbf{Y}_t^i$, namely $\mathbb{E}(\boldsymbol{\Lambda}_t^{(k)} \mid \mathbf{Y}_t^i)$. Making this change is of interest if the following two conditions are met:

1. $\mathbf{Y}_t^i$ is (strongly) correlated to the random variable $\boldsymbol{\Lambda}_t$,
2. $\mathbf{Y}^i$ is a short memory process.

Note that we require that the information process is not influenced by controls.

Subproblem approximation

Following this idea, we replace Subproblem (3) by:

$$\min_{\mathbf{U}^i,\mathbf{X}^i} \mathbb{E}\bigg(\sum_{t=0}^{T-1} \Big(L_t^i(\mathbf{X}_t^i,\mathbf{U}_t^i,\mathbf{W}_t) + \mathbb{E}(\mathbf{\Lambda}_t^{(k)} \mid \mathbf{Y}_t^i) \cdot g_t^i(\mathbf{X}_t^i,\mathbf{U}_t^i,\mathbf{W}_t) \Big) + K^i(\mathbf{X}_T^i) \bigg) ,$$

subject to

$$\begin{aligned} \mathbf{X}_{t+1}^i &= f_t^i(\mathbf{X}_t^i,\mathbf{U}_t^i,\mathbf{W}_t) , \\ \mathbf{U}_t^i &\preceq \sigma(\mathbf{W}_{\text{-}1},\ldots,\mathbf{W}_t) . \end{aligned}$$

The conditional expectation $\mathbb{E}(\mathbf{\Lambda}_t^{(k)} \mid \mathbf{Y}_t^i)$ corresponds to a given function of the variable $\mathbf{Y}_t^i$, so that the Subproblem (3) involves the two exogenous random processes $\mathbf{W}$ and $\mathbf{Y}^i$. If $\mathbf{Y}^i$ exhibits a dynamics of small size, the subproblem can be solved by DP.

Possible choices for the information process

1. **Perfect memory**: $\mathbf{Y}_t^i = \big(\mathbf{W}_{-1}, \dots, \mathbf{W}_t\big)$.
   - $\mathbb{E}(\boldsymbol{\Lambda}_t^{(k)} \mid \mathbf{Y}_t^i) = \boldsymbol{\Lambda}_t^{(k)}$: no approximation!
   - The state size of the subproblem increases with time...
2. **Minimal information**: $\mathbf{Y}_t^i \equiv \mathrm{cste}$.
   - $\boldsymbol{\Lambda}_t^{(k)}$ is approximated by its expectation $\mathbb{E}(\boldsymbol{\Lambda}_t^{(k)})$.
   - The information variable is summarized by a constant value.
3. **Static information**: $\mathbf{Y}_t^i = h_t^i\big(\mathbf{W}_t\big)$.
   - Such a choice is guided by the intuition that a part of $\mathbf{W}_t$ mostly "explains" the optimal multiplier.
4. **Dynamic information**: $\mathbf{Y}_t^i = h_{t-1}^i\big(\mathbf{Y}_{t-1}^i, \mathbf{W}_t\big)$.
   - In the Dynamic Programming equation, the state vector is augmented by embedding $\mathbf{Y}_t^i$, that is, the necessary memory to compute the information variable at the next time step.

Dynamic Programming equation

In the last case (dynamic information), the DP equation writes:

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

$$V_t^i(x,y) = \mathbb{E}\Bigg(\min_u \Big(L_t^i(x,u,\mathbf{W}_t) + \mathbb{E}\big(\boldsymbol{\Lambda}_t^{(k)} \mid \mathbf{Y}_t^i\big) \cdot g_t^i(x,u,\mathbf{W}_t) + V_{t+1}^i\big(\mathbf{X}_{t+1}^i, \mathbf{Y}_t^i\big)\Big)\Bigg) \ ,$$

with $\mathbf{X}_{t+1}^i = f_t^i(x,u,\mathbf{W}_t)$ and $\mathbf{Y}_t^i = h_{t-1}^i\big(y,\mathbf{W}_t\big)$.

The index gap between the information variable and and the stock variable comes from our (bad) notations in the hazard-decision setting: the information used at time t to take decisions is the conjunction of the information kept in memory (that has index $t-1$) and of the noise $\mathbf{W}_t$.

About the coordination

The task of coordination is performed in a scenario-wise manner.

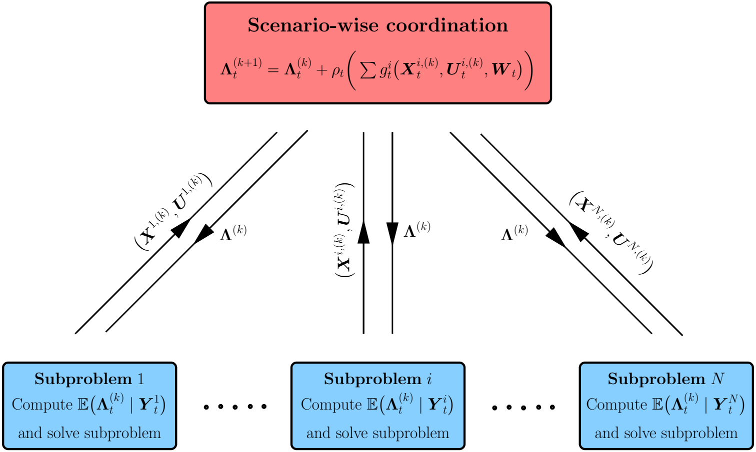
- A bunch of noise scenarios is given once for all, and the trajectories of the information process $\mathbf{Y}^i$ are simulated along these scenarios.
- At iteration k, the optimal trajectories of both the state process $\mathbf{X}^{i,(k)}$ and the control process $\mathbf{U}^{i,(k)}$ are simulated along the noise scenarios for all the subsystems.
- The dual multipliers are updated along the noise scenarios according to the formula:

$$\mathbf{\Lambda}_t^{(k+1)} = \mathbf{\Lambda}_t^{(k)} + \rho_t \left(\sum_{i=1}^{N} g_t^i(\mathbf{X}_t^{i,(k)}, \mathbf{U}_t^{i,(k)}, \mathbf{W}_t) \right) .$$

- The conditional expectations $\mathbb{E}(\mathbf{\Lambda}_t^{(k+1)} \mid \mathbf{Y}_t^i)$ are obtained by regression of the trajectories of $\mathbf{\Lambda}_t^{(k+1)}$ on those of $\mathbf{Y}_t^i$.

DADP algorithm

The algorithm, called Dual Approximate Dynamic Programming, (DADP) is summarized as follows.



Interpretation of DADP (1)

The approximation made on the dual process gives us a tractable way of computing strategies for the subsystems. Let us examine precisely the consequences in terms of constraints.

From now on, assume that the information variable $\mathbf{Y}_t$ is the same for all subsystems. We consider a new problem derived from (2):

$$\min_{\mathbf{U},\mathbf{X}} \mathbb{E}\left(\sum_{i=1}^{N} \left(\sum_{t=0}^{T-1} L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) + K^i(\mathbf{X}_T^i) \right) \right) , \tag{5a}$$

subject to the modified coupling constraints:

$$\mathbb{E}\left(\sum_{i=1}^{N} g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) \;\middle|\; \mathbf{Y}_t \right) = 0 . \tag{5b}$$

Interpretation of DADP (2)

Proposition 2

Suppose the Lagrangian associated with Problem (5) has a saddle point. Then the DADP algorithm can be interpreted as the Uzawa algorithm applied to Problem (5).

Proof. Since the term $\left\langle \mathbb{E}(\boldsymbol{\Lambda}_t^{(k)} \mid \mathbf{Y}_t) \, , g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) \right\rangle$ which appears in the cost function of subproblem i in DADP can be written:

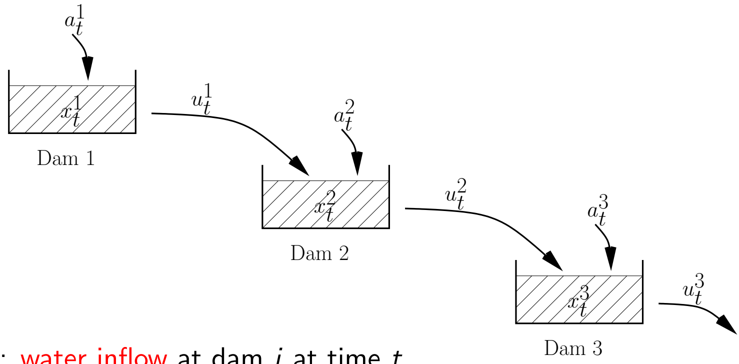
$$\left\langle \mathbb{E}(\boldsymbol{\Lambda}_t^{(k)} \mid \mathbf{Y}_t) \, , g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) \right\rangle = \left\langle \boldsymbol{\Lambda}_t^{(k)} \, , \mathbb{E}(g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) \mid \mathbf{Y}_t) \right\rangle \, ,$$

the global constraint really handled by DADP is:

$$\mathbb{E}\left(\sum_{i=1}^{N} g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t) \;\middle|\; \mathbf{Y}_t \right) = 0 \, . \qquad \square$$

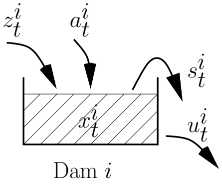
DADP thus consists in replacing an almost-sure constraint by its conditional expectation w.r.t. the information variable $\mathbf{Y}_t$.

Operating scheme



a_t^i : water inflow at dam i at time t,
x_t^i : water volume of dam i at time t,
u_t^i : water turbinated by dam i at time t,
p_t^i : water price at dam i at time t,
$w_t^i = (a_t^i, p_t^i)$ and $w_t = (w_t^1, \ldots, w_t^N)$.

Dynamics and costs



Dam dynamics:

$$\begin{aligned} x_{t+1}^i &= f_t^i(x_t^i, u_t^i, w_t^i, z_t^i) \; , \\ &= x_t^i - u_t^i + a_t^i + z_t^i - s_t^i \; . \end{aligned}$$

with

$$\begin{aligned} z_t^i &= g_t^{i-1}(x_t^{i-1}, u_t^{i-1}, w_t^{i-1}, z_t^{i-1}) \; , \\ &= u_t^{i-1} + \underbrace{\max\left\{0, x_t^{i-1} - u_t^{i-1} + a_t^{i-1} + z_t^{i-1} - \overline{x}^{i-1}\right\}}_{s_t^{i-1}} \; . \end{aligned}$$

We assume the Hazard-Decision information structure (u_t^i is chosen once w_t^i is observed), so that $\underline{u}^i \leq u_t^i \leq \min\left\{\overline{u}^i, x_t^i + a_t^i + z_t^i - \underline{x}^i\right\}$.

Cost at time $t < T$: $\quad L_t^i(x_t^i, u_t^i, w_t^i, z_t^i) = -p_t^i u_t^i + \epsilon (u_t^i)^2$.

Final cost at time T: $\quad K^i\left(x_T^i\right) = \kappa^i \min\{0, x_T^i - \widehat{x}^i\}^2$.

Stochastic optimization problem

The global optimization problem reads:

$$\min_{(\mathbf{X},\mathbf{U},\mathbf{Z})} \mathbb{E}\left(\sum_{i=1}^{N} \left(\sum_{t=0}^{T-1} L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t^i, \mathbf{Z}_t^i) + K^i(\mathbf{X}_T^i) \right) \right), \tag{6a}$$

subject to:

$$\mathbf{X}_{t+1}^i = f_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t^i, \mathbf{Z}_t^i) \ , \ \ \forall i \ , \ \ \forall t \ , \tag{6b}$$

$$\mathbf{Z}_t^{i+1} = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t^i, \mathbf{Z}_t^i) \ , \ \ \forall i \ , \ \ \forall t \ , \tag{6c}$$

as well as measurability constraints:

$$\mathbf{U}_t^i \preceq \sigma\big(\mathbf{W}_0, \ldots, \mathbf{W}_t\big) \ , \ \ \forall i \ , \ \ \forall t \ . \tag{6d}$$

$\rightsquigarrow$ "Cascade model" stochastic optimisation problem.

Price decomposition and DADP algorithm

Our aim is to dualize Constraint (6c) and to solve Problem (6) by using DADP algorithm: at iteration k, the multiplier associated to (6c) is a fixed random variable $\boldsymbol{\Lambda}_t^{i+1,(k)}$, and the term (under the expectation) induced by duality in the cost function is

$$\mathbb{E}\big(\boldsymbol{\Lambda}_t^{i+1,(k)} \mid \mathbf{Y}_t^i\big) \cdot \big(\mathbf{Z}_t^{i+1} - g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t^i, \mathbf{Z}_t^i)\big) .$$

It can be decomposed as the sum of two terms

- $\mathbb{E}\big(\boldsymbol{\Lambda}_t^{i+1,(k)} \mid \mathbf{Y}_t^i\big) \cdot \mathbf{Z}_t^{i+1}$, pertaining to dam $i+1$, and
- $-\mathbb{E}\big(\boldsymbol{\Lambda}_t^{i+1,(k)} \mid \mathbf{Y}_t^i\big) \cdot g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t^i, \mathbf{Z}_t^i)$, pertaining to dam i.

Subproblems and information variables

The expression of Subproblem i is:

$$\min_{\mathbf{U}^i,\mathbf{Z}^i,\mathbf{X}^i} \mathbb{E}\Bigg(\sum_{t=0}^{T-1} \Big(L_t^i(\mathbf{X}_t^i,\mathbf{U}_t^i,\mathbf{W}_t^i,\mathbf{Z}_t^i) + \mathbb{E}\big(\boldsymbol{\Lambda}_t^{i,(k)} \bigm| \mathbf{Y}_t^{i-1}\big)\cdot \mathbf{Z}_t^i - \mathbb{E}\big(\boldsymbol{\Lambda}_t^{i+1,(k)} \bigm| \mathbf{Y}_t^{i}\big)\cdot g_t^i(\mathbf{X}_t^i,\mathbf{U}_t^i,\mathbf{W}_t^i,\mathbf{Z}_t^i)\Big) + K^i(\mathbf{X}_T^i)\Bigg) .$$

Possible choices for $\mathbf{Y}_t^i$ are:[2]

1. $\mathbf{Y}_t^i \equiv \mathrm{cste}$: we deal with the constraint in expectation,
2. $\mathbf{Y}_t^i = \mathbf{W}_t^{i-1}$: we incorporate the noise $\mathbf{W}_t^{i-1}$ in Subproblem i,
3. $\mathbf{Y}_t^i = \widetilde{f}_t^{i-1}(\mathbf{Y}_{t-1}^i,\mathbf{W}_t^{i-1})$: we mimic the dynamics of $\mathbf{X}_t^{i-1}$.

[2] Remember that $\mathbf{Y}_t^i$ is related to a constraint involving both $\mathbf{W}_t^{i-1}$ and $\mathbf{X}_t^{i-1}$.

The particular case $\mathbf{Y}_t^i \equiv \mathrm{cste}$

1. The multipliers $\mathbf{\Lambda}_t^{i,(k)}$ appear only in the subproblems by means of their expectations $\mathbb{E}\big(\mathbf{\Lambda}_t^{i,(k)}\big)$, so that each subproblem involves a 1-dimensional state variable.
2. The coordination task reduces to:

$$\mathbb{E}\big(\mathbf{\Lambda}_t^{i,(k+1)}\big) = \mathbb{E}\big(\mathbf{\Lambda}_t^{i,(k)}\big) + \rho_t \mathbb{E}\Big(\mathbf{Z}_t^{i+1,(k)} - g_t^i\big(\mathbf{X}_t^{i,(k)}, \mathbf{U}_t^{i,(k)}, \mathbf{W}_t^i, \mathbf{Z}_t^{i,(k)}\big)\Big) .$$

3. The constraints taken into account by DADP are in fact the expected ones:

$$\mathbb{E}\Big(\mathbf{Z}_t^{i+1} - g_t^i\big(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t^i, \mathbf{Z}_t^i\big)\Big) = 0 ,$$

so that the solutions given by DADP do not satisfy the inter-dam constraints.

To be continued
during
the practical works.

G. Cohen.
Optimisation des grands systèmes. Notes de cours Université Paris I.
http://cermics.enpc.fr/ cohen-g/documents/ParisIcours-A4-NB.pdf, 2004.

C. Strugarek.
Approches variationnelles et contributions en optimisation stochastique.
Thèse de doctorat, Ecole Nationale des Ponts et Chausées, 2006.

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

K. Barty, P. Carpentier and P. Girardeau,
Decomposition of large-scale stochastic optimal control problems.
RAIRO Operations Research, 44, 167-183, 2010.

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