Variational approach to SOC problems

Between Stochastic Programming and Dynamic Programming

Pierre Carpentier and SOWG [1]

ENSTA ParisTech

July 12, 2013

[1] **Systems and Optimization Working Group**
Joint work with G. Cohen and A. Dallagi,
with the support of Électricité de France R&D.

Lecture outline

1. Problem formulation and optimality conditions
   - Stochastic optimal control problem
   - General optimality conditions
2. Several possible implementations
   - Standard optimality conditions
   - Adapted optimality conditions
   - Markovian case
3. Numerical algorithm and example
   - The particle method
   - A simple benchmark problem
   - Results and comments

Introduction

Stochastic Optimal Control (SOC) problems.

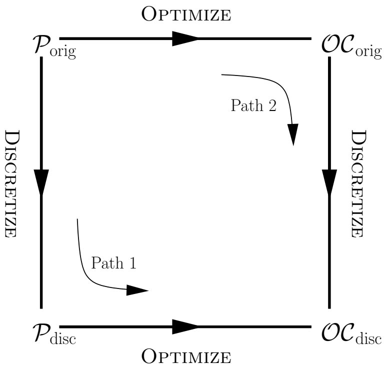
- Stochastic discrete time formulation:
  noise, state, control variables, cost function, constraints.
- Algebraic point of view:
  measurability constraints between random variables.
- Variational approach:
  necessary optimality conditions "à la Kuhn-Tucker".
- Numerical resolution methods.

$\rightsquigarrow$ **Standard way to solve the problem**: $\min_{\mathbf{U} \in \mathcal{U}^{\mathrm{ad}}} J(\mathbf{U})$

Another approach for such problems: Dynamic Programming (functional point of view, sufficient conditions).

Introduction

Two main paths when solving infinite dimensional problems:



Noncommutative diagram!

1. either obtain a finite dimensional approximation of the problem (discretize) and then solve the associated optimality conditions (optimize),
2. or obtain optimality conditions of the problem (optimize) and solve a finite dimensional approximation of these conditions (discretize).

This lecture: **Path 2** (*Path 1* $\rightsquigarrow$ *Scenario tree*)

SOC problem formulation

Consider a fixed discrete time horizon T.

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

subject to the constraints:

$$\begin{aligned} \mathbf{X}_0 &= f_{\text{-}1}(\mathbf{W}_0) , \\ \mathbf{X}_{t+1} &= f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) , \quad \forall t = 0, \ldots, T-1 , \end{aligned}$$

$$\begin{aligned} &\mathbf{U}_t \preceq \mathcal{F}_t , && \forall t = 0, \ldots, T-1 , \\ &\mathbf{U}_t \in \Gamma_t \ \ \mathbb{P}\text{-a.s.} , && \forall t = 0, \ldots, T-1 . \end{aligned}$$

All variables $\mathbf{W}_t$, $\mathbf{U}_t$ and $\mathbf{X}_t$ are assumed to be square integrable random variables defined on $(\Omega, \mathcal{A}, \mathbb{P})$ and valued on appropriate finite dimensional spaces $\mathbb{W}_t$, $\mathbb{U}_t$ and $\mathbb{X}_t$.

Compact formulation

We denote by $\mathbf{W} = (\mathbf{W}_0, \dots, \mathbf{W}_T) \in \mathcal{W}$, $\mathbf{U} = (\mathbf{U}_0, \dots, \mathbf{U}_{T-1}) \in \mathcal{U}$ and $\mathbf{X} = (\mathbf{X}_0, \dots, \mathbf{X}_T) \in \mathcal{X}$ the noise, control and the "state" processes.

- $\mathbf{X}$ being an intermediate process depending on $\mathbf{U}$ and $\mathbf{W}$, the cost function may be written in the following form:

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

- $\mathbf{U}_t$ has to be measurable w.r.t. the σ-field $\mathcal{F}_t$ generated by $(\mathbf{W}_0, \dots, \mathbf{W}_t)$. This constraint defines a linear subspace:

$$\mathbf{U}_t \in \mathcal{U}_t^{\mathrm{me}} = L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{U}_t) .$$

- $\mathbf{U}_t$ is subject to the almost sure constraint: $\mathbf{U}_t(\omega) \in \Gamma_t$ $\mathbb{P}$-a.s. which defines a closed convex subset of random variables:

$$\mathbf{U}_t \in \mathcal{U}_t^{\mathrm{as}} .$$

General optimality conditions

Using all previous notations, the SOC problem boils down to

$$\min_{\mathbf{U} \in \mathcal{U}} J(\mathbf{U}) \quad \text{s.t.} \quad \mathbf{U}_t \in \mathcal{U}_t^{\mathrm{as}} \cap \mathcal{U}_t^{\mathrm{me}} \quad \forall t = 0, \ldots, T-1 \, ,$$

and the associated optimality conditions are as follows:

$$\mathbb{E}\left(\nabla_{\mathbf{U}_t} J(\mathbf{U}^\sharp) \;\middle|\; \mathcal{F}_t\right) \in -\partial \chi_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{U}_t^\sharp) \quad \forall t = 0, \ldots, T-1 \, .$$

Sketch of proof

Write the standard optimality conditions: $\nabla_{\mathbf{U}_t} J(\mathbf{U}^\sharp) \in -\partial \chi_{\mathcal{U}_t^{\mathrm{as}} \cap \mathcal{U}_t^{\mathrm{me}}}(\mathbf{U}_t^\sharp)$,

and use the specific structure of the feasible set:

- $\mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}} \cap \mathcal{U}_t^{\mathrm{me}}} = \mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}} \circ \mathrm{proj}_{\mathcal{U}_t^{\mathrm{me}}}$,
- $\mathrm{proj}_{\mathcal{U}_t^{\mathrm{me}}}(\mathbf{U}_t)$ is the conditional expectation $\mathbb{E}(\mathbf{U}_t \mid \mathcal{F}_t)$,

Initial formulation of the optimality conditions

Computing the gradient of the cost function J using the adjoint (co-state) method, we obtain a first set of detailed optimality conditions for the SOC problem.

If $\mathbf{U}^\sharp$ is a solution of the problem, then

$$\mathbb{E}\Big(\nabla_u L_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1}) + \nabla_u f_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1})\boldsymbol{\lambda}_{t+1}^\sharp \;\Big|\; \mathcal{F}_t\Big) \in -\partial \chi_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{U}_t^\sharp) ,$$

where $\mathbf{X}^\sharp$ and $\boldsymbol{\lambda}^\sharp$ are given by

$$\begin{aligned}
\mathbf{X}_0^\sharp &= f_{\text{-}1}(\mathbf{W}_0) , \\
\mathbf{X}_{t+1}^\sharp &= f_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1}) , \\
\boldsymbol{\lambda}_T^\sharp &= \nabla K(\mathbf{X}_T^\sharp) , \\
\boldsymbol{\lambda}_t^\sharp &= \nabla_x L_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1}) + \nabla_x f_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1}^\sharp)\boldsymbol{\lambda}_{t+1}^\sharp .
\end{aligned}$$

Optimality conditions with adapted co-states

Starting from the previous set of optimality conditions, taking the conditional expectation w.r.t. $\mathcal{F}_t$, we obtain a new set of optimality conditions that only depends on $\mathbf{\Lambda}_t = \mathbb{E}(\boldsymbol{\lambda}_t \mid \mathcal{F}_t)$.

If $\mathbf{U}^\sharp$ is a solution of the problem, then

$$\mathbb{E}\Big(\nabla_u L_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1}) + \nabla_u f_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1})\mathbf{\Lambda}_{t+1}^\sharp \;\Big|\; \mathcal{F}_t\Big) \in -\partial\chi_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{U}_t^\sharp) \ ,$$

where $\mathbf{X}^\sharp$ and $\mathbf{\Lambda}^\sharp$ are given by

$$\mathbf{X}_0^\sharp = f_{-1}(\mathbf{W}_0) \ ,$$

$$\mathbf{X}_{t+1}^\sharp = f_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1}) \ ,$$

$$\mathbf{\Lambda}_T^\sharp = \nabla K(\mathbf{X}_T^\sharp) \ ,$$

$$\mathbf{\Lambda}_t^\sharp = \mathbb{E}\Big(\nabla_x L_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1}) + \nabla_x f_t(\mathbf{X}_t^\sharp, \mathbf{U}_t^\sharp, \mathbf{W}_{t+1})\mathbf{\Lambda}_{t+1}^\sharp \;\Big|\; \mathcal{F}_t\Big) \ .$$

Optimality conditions in the Markovian case

Assuming that the random variables $\mathbf{W}_0, \dots, \mathbf{W}_T$ are independent over time (white noise), one can prove that the optimal control $\mathbf{U}^\sharp_t$ is $\mathbf{X}^\sharp_t$-measurable, hence a third set of optimality conditions.

If $\mathbf{U}^\sharp$ is a solution of the problem, then

$$\mathbb{E}\Big(\nabla_u L_t(\mathbf{X}^\sharp_t, \mathbf{U}^\sharp_t, \mathbf{W}_{t+1}) + \nabla_u f_t(\mathbf{X}^\sharp_t, \mathbf{U}^\sharp_t, \mathbf{W}_{t+1}) \mathbf{\Lambda}^\sharp_{t+1} \;\Big|\; \mathbf{X}^\sharp_t\Big) \in -\partial\chi_{\mathcal{U}^{\mathrm{as}}_t}(\mathbf{U}^\sharp_t) \ ,$$

where $\mathbf{X}^\sharp$ and $\mathbf{\Lambda}^\sharp$ are given by

$$\begin{aligned}
\mathbf{X}^\sharp_0 &= f_{\text{-}1}(\mathbf{W}_0) \ , \\
\mathbf{X}^\sharp_{t+1} &= f_t(\mathbf{X}^\sharp_t, \mathbf{U}^\sharp_t, \mathbf{W}_{t+1}) \ , \\
\mathbf{\Lambda}^\sharp_T &= \nabla K(\mathbf{X}^\sharp_T) \ , \\
\mathbf{\Lambda}^\sharp_t &= \mathbb{E}\Big(\nabla_x L_t(\mathbf{X}^\sharp_t, \mathbf{U}^\sharp_t, \mathbf{W}_{t+1}) + \nabla_x f_t(\mathbf{X}^\sharp_t, \mathbf{U}^\sharp_t, \mathbf{W}_{t+1}) \mathbf{\Lambda}^\sharp_{t+1} \;\Big|\; \mathbf{X}^\sharp_t\Big) \ .
\end{aligned}$$

Optimality conditions: functional point of view

$\mathbf{U}_t^\sharp$ and $\mathbf{\Lambda}_t^\sharp$ being $\mathbf{X}_t^\sharp$-measurable, there exist measurable mappings $U_t^\sharp$ and $\Lambda_t^\sharp$ such that $\mathbf{U}_t^\sharp = U_t^\sharp(\mathbf{X}_t^\sharp)$ and $\mathbf{\Lambda}_t^\sharp = \Lambda_t^\sharp(\mathbf{X}_t^\sharp)$. Using them in the expression of the co-state equation, we obtain: [2]

$$\Lambda_t^\sharp(\mathbf{X}_t^\sharp) = \mathbb{E}\Big(\nabla_x L_t\big(\mathbf{X}_t^\sharp, U_t^\sharp(\mathbf{X}_t^\sharp), \mathbf{W}_{t+1}\big) + \nabla_x f_t\big(\mathbf{X}_t^\sharp, U_t^\sharp(\mathbf{X}_t^\sharp), \mathbf{W}_{t+1}\big) \Lambda_{t+1}^\sharp\big(f_t(\mathbf{X}_t^\sharp, U_t^\sharp(\mathbf{X}_t^\sharp), \mathbf{W}_{t+1})\big) \;\Big|\; \mathbf{X}_t^\sharp\Big) ,$$

which only involves the two independent r.v. $\mathbf{X}_t^\sharp$ and $\mathbf{W}_{t+1}$.

The conditional expectation reduces to an expectation over the distribution of $\mathbf{W}_{t+1}$, and hence a functional condition:

$$\Lambda_t^\sharp(\cdot) = \mathbb{E}\Big(\nabla_x L_t\big(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) + \nabla_x f_t\big(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) \Lambda_{t+1}^\sharp\big(f_t(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1})\big)\Big) .$$

[2] The same reasoning holds true for the condition involving the gradients.

Numerical implementation

We now consider the numerical implementation of the functional optimality conditions obtained in the Markovian case.

Since we have expressions of the gradient of function J, we aim at implementing methods akin to the projected gradient algorithm.

We face two concerns:

- expectations must be evaluated:
  $\rightsquigarrow$ Monte Carlo,
- discrete representation of functions must be obtained:
  $\rightsquigarrow$ interpolation-regression.

Particle method: noise scenarios

The noise process $\mathbf{W}$ is represented using a Monte Carlo approximation.

We thus obtain a set of N noise scenarios $\left\{(w_0^i, \ldots, w_T^i)\right\}_{i=1,\ldots,N}$ associated to a N-sample of the noise process $\mathbf{W}$.[3]

Unlike the scenario tree technique, there is no need to derive a tree structure: the noise scenarios are used as they are!

[3] And remember that the noises are independent over time. . .

Particle method: state dynamics

Optimality conditions

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

At iteration (k), control scenarios $\left\{\left(u_0^{i,(k)}, \ldots, u_{T-1}^{i,(k)}\right)\right\}_{i=1,\ldots,N}$ are available, that is, control values along each scenario.

Obtain the state values $x_t^{i,(k)}$ along the scenarios by integrating the state dynamics in forward time:

$$x_0^{i,(k)} = f_{\text{-}1}(w_0^i) ,$$
$$x_{t+1}^{i,(k)} = f_t\left(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^i\right) ,$$

and thus obtain state scenarios $\left\{\left(x_0^{i,(k)}, \ldots, x_T^{i,(k)}\right)\right\}_{i=1,\ldots,N}$

Particle method: co-state dynamics

Optimality conditions

$$\Lambda_T^\sharp(\cdot) = \nabla K(\cdot) \, ,$$

$$\Lambda_t^\sharp(\cdot) = \mathbb{E}\Big(\nabla_x L_t\big(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) + \nabla_x f_t\big(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1}\big)\Lambda_{t+1}^\sharp\big(f_t(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1})\big)\Big) \, .$$

Obtain the co-state values $\ell_t^{i,(k)}$ by integrating in backward time:

$$\ell_T^{i,(k)} = \nabla K(x_T^{i,(k)}) \, ,$$

$$\ell_t^{i,(k)} = \frac{1}{N}\sum_{j=1}^{N}\Big(\nabla_x L_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big) + \nabla_x f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big)\times \underbrace{\Lambda_{t+1}^{(k)}\Big(f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big)\Big)}_{\neq \ell_{t+1}^{j,(k)} \ \forall j}\Big) \, .$$

$\rightsquigarrow$ *use an interpolation operator*: $\Lambda_{t+1}^{(k)} = \mathfrak{R}_{\mathbb{X}_{t+1}}\big(\mathbf{x}_{t+1}^{(k)}, \boldsymbol{\ell}_{t+1}^{(k)}\big)$.

Particle method: projected gradient

Optimality conditions

$$\mathbb{E}\Big(\nabla_u L_t\big(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) + \nabla_u f_t\big(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1}\big)\Lambda_{t+1}^\sharp\big(f_t(\cdot, U_t^\sharp(\cdot), \mathbf{W}_{t+1})\big)\Big) \in -\partial\chi_{\mathcal{U}_t^{\mathrm{as}}}\big(U_t^\sharp(\cdot)\big) \;.$$

Compute the gradient values $g_t^{i,(k)}$ along the scenarios:

$$g_t^{i,(k)} = \frac{1}{N}\sum_{j=1}^{N}\Big(\nabla_u L_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big) + \nabla_u f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big)\times \Lambda_{t+1}^{(k)}\Big(f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big)\Big)\Big) \;,$$

update the control values using a projected gradient step:

$$u_t^{i,(k+1)} = \mathrm{proj}_{\Gamma_t}\Big(u_t^{i,(k)} - \varepsilon^{(k)} g_t^{i,(k)}\Big) \;,$$

and obtain new control scenarios $\big\{\big(u_0^{i,(k+1)}, \ldots, u_{T-1}^{i,(k+1)}\big)\big\}_{i=1,\ldots,N}$.

A (very) simple benchmark problem

Production management of an hydro-electric dam.

- **Horizon**: $T = 24$ (one day with one hour time steps).
- **Dynamics**:

$$\begin{aligned} \mathbf{X}_0 &= \mathbf{W}_0 \ , \\ \mathbf{X}_{t+1} &= \min\big(\max(\mathbf{X}_t - \mathbf{U}_t + \mathbf{A}_{t+1}, \underline{x}), \overline{x}\big) \ . \end{aligned}$$

- **Cost function**:

$$\sum_t c_t(\mathbf{D}_{t+1} - \mathbf{P}_{t+1}) + K(\mathbf{X}_T) \ ,$$

where $\mathbf{P}_{t+1} = g(\mathbf{U}_t, \mathbf{X}_t, \mathbf{A}_{t+1})$ is the electricity production

- **Constraints**:
  - *measurability*: $\mathbf{U}_t \preceq (\mathbf{W}_0, \ldots, \mathbf{W}_t)$, with $\mathbf{W}_t = (\mathbf{A}_t, \mathbf{D}_t)$.
  - *bounds*: $\mathbf{U}_t \in [\, \underline{u} \,, \overline{u} \,]$.

A simple benchmark problem: data

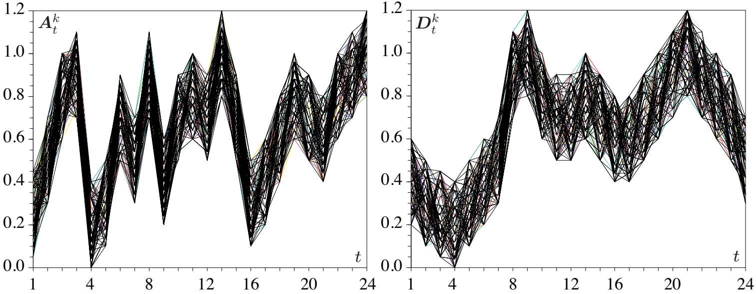


Figure: Water inflow and electricity demand trajectories

Results: Dynamic Programming

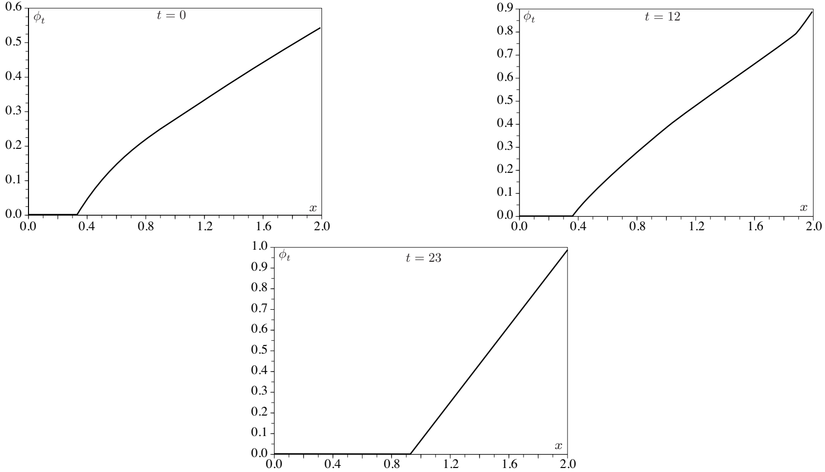
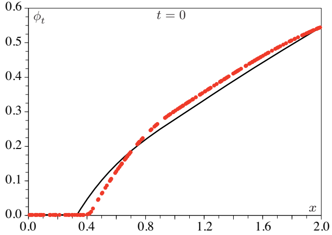
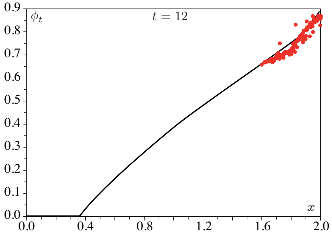


Figure: Dynamic Programming: optimal feedback for three time instants

Results: particle method





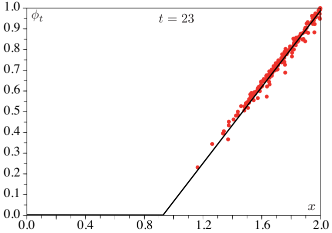


Figure: Particle method: optimal pairs (x, u) at three time instants

Final comments

- The sampling is done once and for all, and that there is no need to derive a tree structure from these noise trajectories.
- The state space discretization is “self-constructive” and adapted to the optimal solution of the problem: the state grids are not designed a priori by the user, as in the case of the DP resolution, but they are automatically produced by the algorithm itself. In fact, the state grids reflect the optimal state distribution of the problem under consideration.
- The fact that the particle method is able to construct a grid in the state space which is adapted to the optimal state distribution, as illustrated by our benchmark problem, should be considered as an advantage (but of course not a definitive answer) to alleviate the curse of dimensionality.

I. Ekeland and R. Temam.
Convex Analysis and Variational Problems.
SIAM, 1999

M. Broadie and P. Glasserman.
A stochastic mesh method for pricing high dimensional american options.
The journal of computational finance, 7(4), 2004.

P. Carpentier, G. Cohen and A. Dallagi.
Particle methods for stochastic optimal control problems.
arXiv, math.OC 0907.4663, 2009. Accepted for publication in *Computational Optimization and Applications*.

A. Dallagi.
Méthodes particulaires en commande optimale stochastique.
Thèse de doctorat, Université Paris I Panthéon-Sorbonne, 2007.

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