Summer School CEA-EDF-INRIA 2012
Stochastic Optimization
◊
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

Variational approach to SOC problems (P. Carpentier)

Tuesday, June 26, 2012

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.

$\rightsquigarrow$ *Classical way to solve the optimization 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:

$$\begin{tikzcd}
\mathcal{P}_{\text{orig}} \arrow[r, "\textsc{Optimize}"] \arrow[d, "\textsc{Discretize}"'] \arrow[dr, phantom, "\text{Path 1} \quad \text{Path 2}"] & \mathcal{OC}_{\text{orig}} \arrow[d, "\textsc{Discretize}"] \\
\mathcal{P}_{\text{disc}} \arrow[r, "\textsc{Optimize}"'] & \mathcal{OC}_{\text{disc}}
\end{tikzcd}$$

Noncommutative diagram!

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

De Lara's lecture: **Path 1** — Today's lecture: **Path 2**

Introduction

On the agenda.

- Obtain tractable optimality conditions for SOC problems:
  - express the gradient of the criterion w.r.t. control variables (co-state variables),
  - Express both the pointwise and the measurability constraints, the latter corresponding to projections on linear subspaces (conditional expectation),

  $\rightsquigarrow$ adequate combination of conditions involving gradients and projections.
- Devise algorithms to numerically solve these conditions.

Lecture outline

1. Formulation of the problem
   - Stochastic optimal control problem
   - Compact formulation
2. Optimality conditions
3. Application to SOC problems
4. Numerical algorithm and example

Overview

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:

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

$$\begin{aligned} &\mathbf{U}_t \preceq \mathcal{G}_t , && \forall t = 0, \dots, T-1 , \\ &\mathbf{U}_t \in \mathbf{C}_t , && \forall t = 0, \dots, T-1 . \end{aligned}$$

All variables $\mathbf{U}_t$, $\mathbf{X}_t$ and $\mathbf{W}_t$ are random variables over $(\Omega, \mathcal{A}, \mathbb{P})$.

Dynamics

The system dynamics follows the equations ($\mathbb{P}$-a.s.):

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

- $\mathbf{W}_t \in \mathcal{W}_t = L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{W}_t)$ is the noise variable at time t, a random variable with values in $\mathbb{W}_t := \mathbb{R}^{d_w}$;
- $\mathbf{U}_t \in \mathcal{U}_t = L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{U}_t)$ is the control variable at time t, a random variable with values in $\mathbb{U}_t := \mathbb{R}^{d_u}$;
- $\mathbf{X}_t \in \mathcal{X}_t = L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{X}_t)$ is the "state" variable at time t, a random variable with values in $\mathbb{X}_t := \mathbb{R}^{d_x}$.

Each function f_t is assumed to be continuously differentiable w.r.t. its first two arguments and to be a normal integrand.

Cost function

We refer to $\mathbf{W} = (\mathbf{W}_0, \ldots, \mathbf{W}_T)$ as the noise stochastic process, which is an element of the product space $\mathcal{W} := \mathcal{W}_0 \times \cdots \times \mathcal{W}_T$.
Similar notations apply (among others) to

- the control process $\mathbf{U} = (\mathbf{U}_0, \ldots, \mathbf{U}_{T-1})$ and the associated space $\mathcal{U}$,
- the "state" process $\mathbf{X} = (\mathbf{X}_0, \ldots, \mathbf{X}_T)$ and the associated space $\mathcal{X}$.

The cost function $\widetilde{j}$ involves an integral term L_t and a final term K:

$$\widetilde{j}(\mathbf{U}, \mathbf{X}, \mathbf{W}) := \sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) + K(\mathbf{X}_T) \ ,$$

and the criterion to be minimized is the expectation:

$$\mathbb{E}\Big(\widetilde{j}(\mathbf{U}, \mathbf{X}, \mathbf{W})\Big) \ .$$

L_t is continuously differentiable w.r.t. its first two arguments and is a normal integrand. K is continuously differentiable.

Measurability constraints (1)

Let $\mathcal{F}_t$ be the σ-field generated by $(\mathbf{W}_0, \dots, \mathbf{W}_t)$. $\mathcal{F}_t$ represents the information available at time t when the decision maker has a complete observation and a full memory of past noises.

Let $\mathcal{G}_t$ be a subfield of $\mathcal{F}_t$. We require that the decision variable $\mathbf{U}_t$ is measurable w.r.t. $\mathcal{G}_t$:

$$\mathbf{U}_t \preceq \mathcal{G}_t \ ,$$

that is, $\sigma(\mathbf{U}_t) \subset \mathcal{G}_t \subset \mathcal{F}_t = \sigma(\mathbf{W}_0, \dots, \mathbf{W}_t)$.

This constraint defines a linear subspace of $\mathcal{U}_t$ denoted $\mathcal{U}_t^{\mathrm{me}}$:

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

Note that the projection onto $\mathcal{U}_t^{\mathrm{me}}$ is a linear operator.

Measurability constraints (2)

The subfield $\mathcal{G}_t$ may be generated by an *observation* variable $\mathbf{Y}_t$:

1. $\mathbf{Y}_t = (\mathbf{W}_0, \dots, \mathbf{W}_t)$ (full noise observation),
2. $\mathbf{Y}_t = (\mathbf{X}_0, \dots, \mathbf{X}_t)$ (full state observation),
3. $\mathbf{Y}_t = \mathbf{X}_t$ (last state observation).

What is important to stress in the framework considered in this lecture is that $\mathbf{Y}_t$ or $\mathcal{G}_t$ should not depend upon past decisions. This restriction is essential to be able to use differential calculus. This assumption holds true only for the first example here above (in the last two examples, $\mathbf{X}_t$ depends on $(\mathbf{U}_0, \dots, \mathbf{U}_{t-1})$).

We shall model measurability constraints using conditional expectations $\mathbb{E}(\cdot \mid \mathcal{G}_t)$*. Removing this assumption would imply that one knows how to differentiate w.r.t.* $\mathbf{Y}_t \dots$

Pointwise constraints

Consider, for every t, a set-valued mapping $\mathbf{C}_t : \Omega \rightrightarrows \mathbb{U}_t$. We require that $\mathbf{U}_t$ is subject to the almost sure constraint:

$$\mathbf{U}_t(\omega) \in \mathbf{C}_t(\omega) \ \ \mathbb{P}\text{-a.s.}\ .$$

This constraint defines a subset of $\mathcal{U}_t$ denoted $\mathcal{U}_t^{\mathrm{as}}$.

The set-valued random mapping $\mathbf{C}_t$ assumes nonempty, closed, convex values $\mathbb{P}$-a.s.. Then, $\mathcal{U}_t^{\mathrm{as}}$ is a closed convex subset of $\mathcal{U}_t$.

We moreover assume that $\mathbf{C}_t$ is $\mathcal{G}_t$-measurable for all t.

Assume that $\mathcal{G}_t$ is generated by $\mathbf{Y}_t$. For any pair (ω, ω') s.t. $\mathbf{Y}_t(\omega) = \mathbf{Y}_t(\omega')$, we have $\mathbf{U}_t(\omega) = \mathbf{U}_t(\omega')$ (measurability constraint). But at the same time, if $\mathbf{C}_t(\omega) \cap \mathbf{C}_t(\omega') = \emptyset$, then the pointwise and the measurability constraints are incompatible. This is prevented by the measurability assumption on $\mathbf{C}_t$:

$$\mathbf{Y}_t(\omega) = \mathbf{Y}_t(\omega') \Rightarrow \mathbf{C}_t(\omega) = \mathbf{C}_t(\omega')\ .$$

To sum up

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

subject to the constraints:

$$\mathbf{X}{=}F(\mathbf{U},\mathbf{W}) : \begin{cases} \mathbf{X}_0 = f_{\text{-}1}(\mathbf{W}_0) \;, \\ \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) \;, \quad \forall t = 0, \dots, T-1 \;, \end{cases} \tag{1b}$$

$$\mathbf{U}{\in}\mathcal{U}^{\mathrm{me}} \quad : \quad \mathbf{U}_t \preceq \mathcal{G}_t \;, \qquad \forall t = 0, \dots, T-1 \;, \tag{1c}$$

$$\mathbf{U}{\in}\mathcal{U}^{\mathrm{as}} \quad : \quad \mathbf{U}_t(\omega) \in \mathbf{C}_t(\omega) \;\; \mathbb{P}\text{-a.s.} \;, \qquad \forall t = 0, \dots, T-1 \;. \tag{1d}$$

A compact formulation

The state $\mathbf{X}$ is just an intermediate stochastic process completely determined by $\mathbf{U}$ and $\mathbf{W}$ through (1b). Let $j(\mathbf{U},\mathbf{W})$ be the value of the cost $\widetilde{j}(\mathbf{U},\mathbf{X},\mathbf{W})$ when replacing $\mathbf{X}$ by $F(\mathbf{U},\mathbf{W})$, and

$$J(\mathbf{U}) := \mathbb{E}\big(j(\mathbf{U},\mathbf{W})\big) \ .$$

Using (1c)—(1d), the SOC problem (1) boils down to

$$\min_{\mathbf{U}\in\mathcal{U}} J(\mathbf{U}) \quad \text{subject to} \quad \mathbf{U} \in \mathcal{U}^{\mathrm{me}} \cap \mathcal{U}^{\mathrm{as}} \ . \tag{2}$$

Optimality conditions for Problem (2) express that the gradient of J at a solution $\mathbf{U}^\sharp$ belongs to the cone orthogonal to the constraints.

- How to operate a projection onto the intersection $\mathcal{U}^{\mathrm{me}} \cap \mathcal{U}^{\mathrm{as}}$?
- How to compute the gradient $\nabla J(\mathbf{U})$ of J (defined implicitly)?

About the notation

Forget for a while the variable $\mathbf{W}$.

- The notation $j(\mathbf{U})$ is used to represent the cost function under the integral sign. In this situation, $j(\mathbf{U})$ is a random variable obtained by the function composition $j \circ \mathbf{U}$:

$$\begin{array}{rcccccc} j(\mathbf{U}) & : & \Omega & \to & \mathbb{U} & \to & \mathbb{R} \\ & & \omega & \mapsto & \mathbf{U}(\omega) & \mapsto & j\big(\mathbf{U}(\omega)\big) \ . \end{array}$$

- We use the notation $J(\mathbf{U})$ to represent the expected cost:

$$\begin{array}{rcccc} J & : & \mathcal{U} & \to & \mathbb{R} \\ & & \mathbf{U} & \mapsto & \mathbb{E}\big(j(\mathbf{U},\mathbf{W})\big) \ . \end{array}$$

  In this situation, $J(\mathbf{U})$ is the value taken by J at point $\mathbf{U}$, and by no means the composite function $J \circ \mathbf{U}$.

$\rightsquigarrow$ Notational ambiguity, lifted by the context.

Projection on $\mathcal{U}^{\mathrm{me}}$ and $\mathcal{U}^{\mathrm{as}}$

Recall that $\mathcal{U}_t$ is the Hilbert space $L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{U}_t)$, and that

$$\mathcal{U}_t^{\mathrm{me}} = \{\mathbf{U}_t \in \mathcal{U}_t \mid \mathbf{U}_t \preceq \mathcal{G}_t\} ,$$
$$\mathcal{U}_t^{\mathrm{as}} = \{\mathbf{U}_t \in \mathcal{U}_t \mid \mathbf{U}_t(\omega) \in \mathbf{C}_t(\omega) \ \ \mathbb{P}\text{-a.s.}\} .$$

1. $\mathrm{proj}_{\mathcal{U}_t^{\mathrm{me}}}(\mathbf{U}_t)$ is the conditional expectation $\mathbb{E}(\mathbf{U}_t \mid \mathcal{G}_t)$.
   $\mathbb{E}(\cdot \mid \mathcal{G}_t)$ is by definition the orthogonal projection on $L^2(\Omega, \mathcal{G}_t, \mathbb{P}; \mathbb{U}_t)$.
2. $\mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{U}_t)$ is the random variable: $\omega \mapsto \mathrm{proj}_{\mathbf{C}_t(\omega)}(\mathbf{U}_t(\omega))$.

$$\mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{U}_t) = \operatorname*{arg\,min}_{\mathbf{V} \in \mathcal{U}_t} \Big(\|\mathbf{V} - \mathbf{U}_t\|_{\mathcal{U}_t}^2 + \chi_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{V}) \Big)$$
$$= \operatorname*{arg\,min}_{\mathbf{V} \in \mathcal{U}_t} \int_\Omega \Big(\|\mathbf{V}(\omega) - \mathbf{U}_t(\omega)\|_{\mathbb{U}_t}^2 + \chi_{\mathbf{C}_t(\omega)}(\mathbf{V}(\omega)) \Big) \mathrm{d}\mathbb{P}(\omega) ,$$

since $\chi_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{V}) = \int_\Omega \chi_{\mathbf{C}_t(\omega)}(\mathbf{V}(\omega)) \, \mathrm{d}\mathbb{P}(\omega)$.

Otherwise stated, $\mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}}$ operates "ω per ω" (pointwise).

Projection on $\mathcal{U}^{\mathrm{me}} \cap \mathcal{U}^{\mathrm{as}}$

Proposition 1

Assume that $\mathbf{C}_t$ *is* $\mathcal{G}_t$*-measurable, closed convex valued. Then*

$$\mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}} \cap \mathcal{U}_t^{\mathrm{me}}} = \mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}} \circ \mathrm{proj}_{\mathcal{U}_t^{\mathrm{me}}} \ .$$

Note first that the pointwise projection of a $\mathcal{G}_t$-measurable function is also a $\mathcal{G}_t$-measurable function:

$$\mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}}\left(\mathcal{U}_t^{\mathrm{me}}\right) \subset \mathcal{U}_t^{\mathrm{me}} \ .$$

Then, the projection $\mathbf{U}_t^\natural$ of $\mathbf{U}_t \in \mathcal{U}_t$ on $\mathcal{U}^{\mathrm{me}} \cap \mathcal{U}^{\mathrm{as}}$ is characterized by

$$\langle \mathbf{U}_t - \mathbf{U}_t^\natural \, , \mathbf{V} - \mathbf{U}_t^\natural \rangle \leq 0 \ , \ \ \forall \, \mathbf{V} \in \mathcal{U}^{\mathrm{me}} \cap \mathcal{U}^{\mathrm{as}} \ .$$

But $\mathbf{U}_t^\flat := \mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}} \circ \mathrm{proj}_{\mathcal{U}_t^{\mathrm{me}}}(\mathbf{U}_t)$ is such that

$$\langle \mathbf{U}_t - \mathbf{U}_t^\flat \, , \mathbf{V} - \mathbf{U}_t^\flat \rangle \leq \langle \mathbf{U}_t - \mathrm{proj}_{\mathcal{U}_t^{\mathrm{me}}}\left(\mathbf{U}_t\right) \, , \mathbf{V} - \mathbf{U}_t^\flat \rangle = 0 \ .$$

Stationary conditions in the general case

Let $J : \mathcal{U} \to \mathbb{R}$ be a differentiable function over a closed convex subset $\mathcal{U}^{\mathrm{ad}}$ of a Hilbert space $\mathcal{U}$. Consider the problem:

$$\min_{\mathbf{U} \in \mathcal{U}^{\mathrm{ad}}} J(\mathbf{U}) \ .$$

The following statements are three equivalent necessary conditions for $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{ad}}$ to be optimal:

$$\forall \mathbf{U} \in \mathcal{U}^{\mathrm{ad}} \ , \ \ \left\langle \nabla J(\mathbf{U}^\sharp) \, , \mathbf{U} - \mathbf{U}^\sharp \right\rangle \geq 0 \ , \tag{3a}$$

$$\nabla J(\mathbf{U}^\sharp) \in -\partial \chi_{\mathcal{U}^{\mathrm{ad}}}(\mathbf{U}^\sharp) \ , \tag{3b}$$

$$\forall \varepsilon > 0 \ , \ \ \mathbf{U}^\sharp = \mathrm{proj}_{\mathcal{U}^{\mathrm{ad}}} \Big(\mathbf{U}^\sharp - \varepsilon \nabla J(\mathbf{U}^\sharp) \Big) \ . \tag{3c}$$

Equivalence (3a)—(3b) stems from the fact that the subdifferential $\partial \chi_{\mathcal{U}^{\mathrm{ad}}}(\mathbf{U})$ of the **characteristic function** $\chi_{\mathcal{U}^{\mathrm{ad}}}$ is the normal cone to $\mathcal{U}^{\mathrm{ad}}$ at point $\mathbf{U}$. The equivalence (3a)—(3c) is immediate.

Stationary conditions: case $\mathcal{U}^{\mathrm{ad}} = \mathcal{U}^{\mathrm{me}} \cap \mathcal{U}^{\mathrm{as}}$

Proposition 2

Assume that $\mathcal{U}^{\mathrm{ad}} = \mathcal{U}^{\mathrm{me}} \cap \mathcal{U}^{\mathrm{as}}$ and that Proposition 1 applies. Then a necessary condition for $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{ad}}$ to be optimal is:

$$\mathrm{proj}_{\mathcal{U}^{\mathrm{me}}}\Big(\nabla J(\mathbf{U}^\sharp)\Big) \in -\partial\chi_{\mathcal{U}^{\mathrm{as}}}(\mathbf{U}^\sharp) . \tag{4}$$

Condition (3c) writes:

$$\begin{aligned}
\forall\varepsilon > 0 , \ \ \mathbf{U}^\sharp &= \mathrm{proj}_{\mathcal{U}^{\mathrm{as}} \cap \mathcal{U}^{\mathrm{me}}}\Big(\mathbf{U}^\sharp - \varepsilon\nabla J(\mathbf{U}^\sharp)\Big) \ , \\
&= \mathrm{proj}_{\mathcal{U}^{\mathrm{as}}} \circ \mathrm{proj}_{\mathcal{U}^{\mathrm{me}}}\Big(\mathbf{U}^\sharp - \varepsilon\nabla J(\mathbf{U}^\sharp)\Big) \ , \\
&= \mathrm{proj}_{\mathcal{U}^{\mathrm{as}}}\Big(\mathbf{U}^\sharp - \varepsilon\,\mathrm{proj}_{\mathcal{U}^{\mathrm{me}}}\Big(\nabla J(\mathbf{U}^\sharp)\Big)\Big) \ ,
\end{aligned}$$

($\mathrm{proj}_{\mathcal{U}^{\mathrm{me}}}$ is a linear operator), hence the result thanks to Condition (3b).

Stationarity conditions: application to Problem (1)

Condition (4) has to be written at each time t. As already seen, $\mathrm{proj}_{\mathcal{U}_t^{\mathrm{me}}}(\cdot) = \mathbb{E}(\cdot \mid \mathcal{G}_t)$, so that the stationarity conditions can be more explicitly written, that is, for $t = 0, \ldots, T-1$:

$$\mathbb{E}\big(\nabla_{\mathbf{U}_t} J(\mathbf{U}^\sharp) \bigm| \mathcal{G}_t\big) \in -\partial\chi_{\mathcal{U}_t^{\mathrm{as}}}(\mathbf{U}_t^\sharp) . \tag{5}$$

Note that the expression of the gradient $\mathbb{E}\big(\nabla_{\mathbf{U}_t} J(\mathbf{U}) \bigm| \mathcal{G}_t\big)$ can be used inside a gradient-like algorithm in order to obtain the optimal solution $\mathbf{U}_t^\sharp$:

$$\mathbf{U}_t^{(k+1)} = \mathrm{proj}_{\mathcal{U}_t^{\mathrm{as}}}\Big(\mathbf{U}_t^{(k)} - \varepsilon\, \mathbb{E}\big(\nabla_{u_t} J(\mathbf{U}^{(k)}) \bigm| \mathcal{G}_t\big)\Big) .$$

Computation of the cost gradient (1)

A classical way to obtain the gradient of $J(\mathbf{U}) = \mathbb{E}\big(j(\mathbf{U},\mathbf{W})\big)$ (obtained by replacing $\mathbf{X}$ with $F(\mathbf{U},\mathbf{W})$) is to introduce the so-called co-state variables $\boldsymbol{\lambda}_t$. The method is the following.

- Form the "pseudo-Lagrangian" $\mathfrak{L}$:

$$\mathfrak{L}(\mathbf{X},\mathbf{U},\boldsymbol{\lambda}) = \mathbb{E}\Bigg(\boldsymbol{\lambda}_0^\top\big(f_{-1}(\mathbf{W}_0) - \mathbf{X}_0\big) + \sum_{t=0}^{T-1} \boldsymbol{\lambda}_{t+1}^\top\big(f_t(\mathbf{X}_t,\mathbf{U}_t,\mathbf{W}_{t+1}) - \mathbf{X}_{t+1}\big) + \sum_{t=0}^{T-1} L_t(\mathbf{X}_t,\mathbf{U}_t,\mathbf{W}_{t+1}) + K(\mathbf{X}_T)\Bigg),$$

($\boldsymbol{\lambda}_t$ is a random variable which belongs to $L^2(\Omega,\mathcal{A},\mathbb{P};\mathbb{X}_t)$).

Computation of the cost gradient (2)

- Draw $\mathbf{X}$ from $\nabla_{\boldsymbol{\lambda}}\mathfrak{L}(\mathbf{X}, \mathbf{U}, \boldsymbol{\lambda}) = 0$ (forward dynamics):

$$\begin{aligned} \mathbf{X}_0 &= f_{\text{-}1}(\mathbf{W}_0) \ , \\ \mathbf{X}_{t+1} &= f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) \ , \end{aligned} \qquad .$$

- Draw $\boldsymbol{\lambda}$ from $\nabla_{\mathbf{X}}\mathfrak{L}(\mathbf{X}, \mathbf{U}, \boldsymbol{\lambda}) = 0$ (backward dynamics):

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

- Obtain the **gradient** $\nabla J(\mathbf{U})$ from $\nabla_{\mathbf{U}}\mathfrak{L}(\mathbf{X}, \mathbf{U}, \boldsymbol{\lambda})$:

$$\nabla_{\mathbf{U}_t} J(\mathbf{U}) = \nabla_u L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) + \nabla_u f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1})\boldsymbol{\lambda}_{t+1} \ .$$

Initial formulation of the optimality conditions (1)

Gathering the results obtained at the previous section, that is, the computation of the cost gradient and the stationarity condition (5), we deduce a first set of optimality conditions for Problem (1).

If $\mathbf{U}^\sharp$ is a solution of Problem (1), there exist $\mathbf{X}^\sharp$ and $\boldsymbol{\lambda}^\sharp$ such that

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

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

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

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

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

Initial formulation of the optimality conditions (2)

Note that the co-state random variable $\boldsymbol{\lambda}_t$ is not $\mathcal{F}_t$-adapted (the dynamics of $\boldsymbol{\lambda}_t$ propagates in backward time). It would however be normal to be because $\boldsymbol{\lambda}_t$ corresponds to the multiplier associated to the $\mathcal{F}_t$-adapted constraint $f_{t-1}(\mathbf{X}_{t-1}, \mathbf{U}_{t-1}, \mathbf{W}_t) - \mathbf{X}_t = 0$.

Decompose $\boldsymbol{\lambda}_t$ into its $\mathcal{F}_t$-measurable component, namely

$$\boldsymbol{\Lambda}_t = \mathbb{E}(\boldsymbol{\lambda}_t \mid \mathcal{F}_t) \; ,$$

on the one hand, and its orthogonal complement $\boldsymbol{\lambda}_t - \boldsymbol{\Lambda}_t$ on the other hand. Only the former component contributes to the duality product $\mathbb{E}\big(\boldsymbol{\lambda}_t \cdot \big(f_{t-1}(\mathbf{X}_{t-1}, \mathbf{U}_{t-1}, \mathbf{W}_t) - \mathbf{X}_t\big)\big)$. Hence it should be possible to get optimality conditions involving only $\boldsymbol{\Lambda}_t$, that is, an adapted co-state process...

Optimality conditions with adapted co-states (1)

Starting from the previous set of optimality conditions involving the non $\mathcal{F}_t$-adapted co-state variables $\boldsymbol{\lambda}_t$, and taking the conditional expectation w.r.t. $\mathcal{F}_t$ of the co-state equations, we obtain:

If $\mathbf{U}^\sharp$ is a solution of Problem (1), there exist $\mathbf{X}^\sharp$ and $\boldsymbol{\lambda}^\sharp$ such that

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

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

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

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

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

Optimality conditions with adapted co-states (2)

From the general property $\mathcal{G} \subset \mathcal{F} \Rightarrow \mathbb{E}(\cdot \mid \mathcal{G}) = \mathbb{E}\big(\mathbb{E}(\cdot \mid \mathcal{F}) \bigm| \mathcal{G}\big)$, we deduce that this set of optimality conditions only depends on $\mathbf{\Lambda}_t = \mathbb{E}(\boldsymbol{\lambda}_t \mid \mathcal{F}_t)$, hence a second set of optimality conditions:

If $\mathbf{U}^\sharp$ is a solution of Problem (1), there exist $\mathbf{X}^\sharp$ and $\mathbf{\Lambda}^\sharp$ such that

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

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

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

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

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

Markovian setting and assumptions

Two kinds of assumptions:

- on the one hand, the noise process $\mathbf{W}$ should forget the past,
- on the other hand, constraints on the decision process $\mathbf{U}$ should not reintroduce past observations.

Assumption 1 (White noise)

The random variables $\mathbf{W}_0, \ldots, \mathbf{W}_T$ are independent over time.

Assumption 2 (Decision constraints)

1. $\mathcal{G}_t = \mathcal{F}_t$.
2. The set-valued mappings $\mathbf{C}_t$ involved in the pointwise constraints are constant (deterministic) and denoted C_t.

Optimality conditions (1)

We start from the second set of optimality conditions, that is, the conditions involving adapted co-state variables $\boldsymbol{\Lambda}_t = \mathbb{E}(\boldsymbol{\lambda}_t \mid \mathcal{F}_t)$.

If $\mathbf{U}^\sharp$ is a solution of Problem (1), there exist $\mathbf{X}^\sharp$ and $\boldsymbol{\Lambda}^\sharp$ such that

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

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

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

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

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

Optimality conditions (2)

Using implicit measurable selection theorems, we can prove the following property by induction (under Assumptions 1 and 2).

If $\mathbf{U}^\sharp$ is a solution of Problem (1), there exist $\mathbf{X}^\sharp$ and $\mathbf{\Lambda}^\sharp$ such that

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

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

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

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

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

Moreover, we have that $\mathbf{U}_t^\sharp$ is $\mathbf{X}_t^\sharp$-measurable: $\mathbf{U}_t^\sharp \preceq \mathbf{X}_t^\sharp$.

Remarks about the Markovian case

- The new optimality conditions involve conditional expectations w.r.t. the state $\mathbf{X}_t^\sharp$, that is, a conditioning term of fixed size.
- The random variables $\mathbf{U}_t^\sharp$ and $\mathbf{\Lambda}_t^\sharp$ involved in the optimality conditions are $\mathbf{X}_t^\sharp$-measurable. Note that the property $\mathbf{U}_t^\sharp \preceq \mathbf{X}_t^\sharp$ instead of $\mathbf{U}_t^\sharp \preceq \mathcal{G}_t$ is a valuable result, strongly related to Dynamic Programming.
- These new optimality conditions have been derived from the adapted optimality conditions, in the Markovian case. We could have started from the constraint $\mathbf{U}_t \preceq \mathbf{X}_t$ and have attempted to directly obtain optimality conditions. This would have been a difficult challenge because the feasible set $\mathcal{U}_t^{\mathrm{me}}$ would have depend on the past decision variables through the state $\mathbf{X}_t$. The indirect path we followed has been a way to circumvent this difficulty.

Optimality conditions from a functional point of view (1)

Since $\mathbf{U}_t^\sharp$ and $\mathbf{\Lambda}_t^\sharp$ are both $\mathbf{X}_t^\sharp$-measurable, there exist measurable mappings $\gamma_t^\sharp : \mathbb{X}_t \to \mathbb{U}_t$ and $\Lambda_t^\sharp : \mathbb{X}_t \to \mathbb{X}_t$ such that $\mathbf{U}_t^\sharp = \gamma_t^\sharp(\mathbf{X}_t^\sharp)$ and $\mathbf{\Lambda}_t^\sharp = \Lambda_t^\sharp(\mathbf{X}_t^\sharp)$. Then, the last two optimality conditions write:

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

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

Apart from $\mathbf{X}_t^\sharp$, these expressions only involve $\mathbf{W}_{t+1}$, which is independent of $\mathbf{X}_t^\sharp$. Therefore, the conditional expectation w.r.t. $\mathbf{X}_t^\sharp$ reduces to a simple expectation over the distribution of $\mathbf{W}_{t+1}$.

Optimality conditions from a functional point of view (2)

We obtain optimality conditions which involve only expectations.

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

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

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

In the present Markovian case, the optimal solution can also be obtained by the Dynamic Programming equation:

$$V_T(\cdot) = K(\cdot) \ , \ \ V_t(\cdot) = \min_{u \in C_t} \mathbb{E}\Big(L_t\big(\cdot, u, \mathbf{W}_{t+1}\big) + V_{t+1}\big(f_t(\cdot, u, \mathbf{W}_{t+1})\big)\Big) \ ,$$

and it can be proved (by induction) that: $\Lambda_t^\sharp = \nabla V_t$.

Implementation

We now consider the numerical implementation of the (functional) optimality conditions obtained in the Markovian case:

$$\Lambda_t^\sharp(\cdot) = \mathbb{E}\Big(\nabla_x L_t\big(\cdot, \gamma_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) + \nabla_x f_t\big(\cdot, \gamma_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) \Lambda_{t+1}^\sharp\big(f_t(\cdot, \gamma_t^\sharp(\cdot), \mathbf{W}_{t+1})\big)\Big) , \quad (6)$$

$$\mathbb{E}\Big(\nabla_u L_t\big(\cdot, \gamma_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) + \nabla_u f_t\big(\cdot, \gamma_t^\sharp(\cdot), \mathbf{W}_{t+1}\big) \Lambda_{t+1}^\sharp\big(f_t(\cdot, \gamma_t^\sharp(\cdot), \mathbf{W}_{t+1})\big)\Big) \in -\partial\chi_{\mathcal{U}_t^{\mathrm{as}}}\big(\gamma_t^\sharp(\cdot)\big) . \quad (7)$$

We face two concerns:

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

Interpolation-regression

Consider an unknown function $\gamma_t : \mathbb{X}_t \to \mathbb{U}_t$ and two known grids:

- $\mathbf{x}_t = \left(x_t^1, \dots, x_t^N\right)$: grid of N points of $\mathbb{X}_t$,
- $\mathbf{u}_t = \left(u_t^1, \dots, u_t^N\right)$: grid of N points of $\mathbb{U}_t$,

such that

$$u_t^i = \gamma_t(x_t^i) \ .$$

In order to evaluate the function at points outside of the grid, we introduce an interpolation-regression operator:

$$\mathfrak{R}_{\mathbb{U}_t} : \mathbb{X}_t^N \times \mathbb{U}_t^N \to \mathbb{U}_t^{\mathbb{X}_t} \ ,$$

so that $\widetilde{\gamma}_t = \mathfrak{R}_{\mathbb{U}_t}(\mathbf{x}_t, \mathbf{u}_t)$ is an approximation of γ_t.

There are various ways to define this operator (polynomial and spline interpolation, kernel regression, closest neighbor...).

Algorithm (1)

Initialization

- Obtain a set of N realizations $\{(w_0^i, \dots, w_T^i)\}_{i=1,\dots,N}$ of the noise process $\mathbf{W}$.
- Obtain grids $\mathbf{u}_t^{(0)} = \big(u_t^{1,(0)}, \dots, u_t^{N,(0)}\big)$, $t = 0, \dots, T-1$, for the control process $\mathbf{U}$.

Iteration (k) (beginning)

- Using the control grids $\mathbf{u}_t^{(k)}$, compute the state values $x_t^{i,(k)}$:

$$x_0^{i,(k)} = f_{-1}(w_0^i) \; ,$$
$$x_{t+1}^{i,(k)} = f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^i\big) \; .$$

which yields the state grids $\mathbf{x}_t^{(k)}$, $t = 0, \dots, T$.

Algorithm (2)

Iteration (k) (continued)

- The approximations $\Lambda_t^{(k)}$ of the co-state functions $\Lambda_t^\sharp$ defined by Equation (6) are constructed by backward recursion:
  - $$\ell_t^{i,(k)} = \frac{1}{N}\sum_{j=1}^{N}\Big(\nabla_x L_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^{j}\big) + \nabla_x f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^{j}\big)\Lambda_{t+1}^{(k)}\Big(f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^{j}\big)\Big)\Big) ,$$
  - $\boldsymbol{\ell}_t^{(k)} = \big\{\ell_t^{i,(k)}\big\}_{i=1,\dots,N}$,
  - $\Lambda_t^{(k)} = \mathfrak{R}_{\mathbb{X}_t}\big(\mathbf{x}_t^{(k)}, \boldsymbol{\ell}_t^{(k)}\big)$.

Note that the computation of $l_t^{i,(k)}$ is such that the function $\Lambda_{t+1}^{(k)}$ has to be evaluated at points which lie outside of the grid $\mathbf{x}_{t+1}^{(k)}$.

Algorithm (3)

Iteration (k) (end)

- Finally, Equation (7) is used to update the control particles, using a gradient step (with stepsize $\epsilon^{(k)}$) projected on the feasible set C_t:

$$u_t^{i,(k+1)} = \mathrm{proj}_{C_t}\Big(u_t^{i,(k)} - \frac{\varepsilon^{(k)}}{N} \sum_{j=1}^{N} \Big(\nabla_u L_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big) + \nabla_u f_t\big(x_t^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big) \Lambda_{t+1}^{(k)}\Big(f_t\big(x^{i,(k)}, u_t^{i,(k)}, w_{t+1}^j\big)\Big)\Big)\Big) .$$

Terminaison

Assuming the convergence of state and control particle values at $\mathbf{x}_t^{(\infty)} = \{x_t^{i,(\infty)}\}_{i=1,\dots,N}$ and $\mathbf{u}_t^{(\infty)} = \{u_t^{i,(\infty)}\}_{i=1,\dots,N}$, we build up an approximation of the solution by interpolation-regression:

$$\gamma_t^{(\infty)} = \mathfrak{R}_{\mathbb{U}_t}\big(\mathbf{x}_t^{(\infty)}, \mathbf{u}_t^{(\infty)}\big) .$$

Algorithm (4)

Algorithm summary.

Initialization: Set $k = 0$ and guess initial control grids $\mathbf{u}_t^{(0)}$ for $t = 0, \ldots, T-1$ and $i = 1, \ldots, N$.

Iteration k:

1. Compute the state grids $\mathbf{x}_t^{(k)}$ using a forward recursion.
2. Compute the co-state functions $\Lambda_t^{(k)}$ using a backward recursion.
3. Update the control grids to $u_t^{(k+1)}$ using a gradient step.
4. Iterate with $k+1 \leftarrow k$ or stop if stationarity.

Termination: With the limit values $\mathbf{x}_t^{(\infty)}$ and $\mathbf{u}_t^{(\infty)}$, build up feedback functions $\gamma_t^{(\infty)}$ for $t = 0, \ldots, T-1$.

A 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

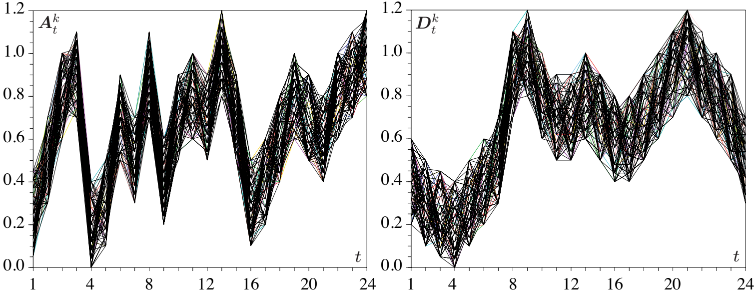

Figure: Water inflow and electricity demand trajectories

Results: Dynamic Programming

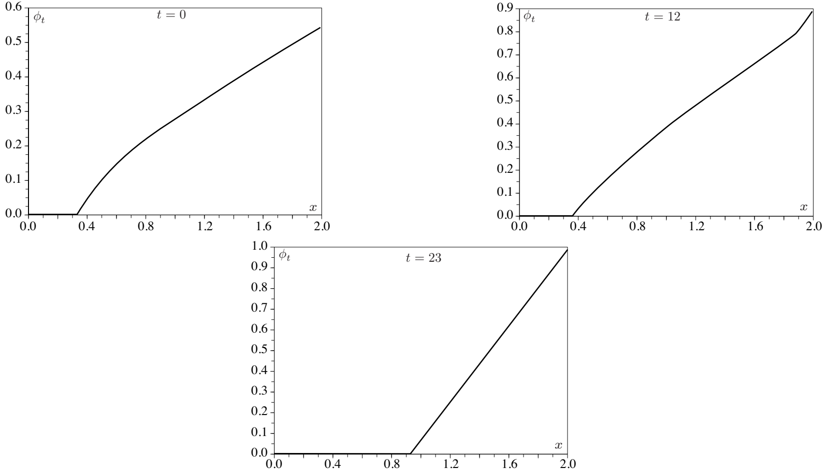

Figure: Dynamic Programming: optimal feedback for three time instants

Results: particle method

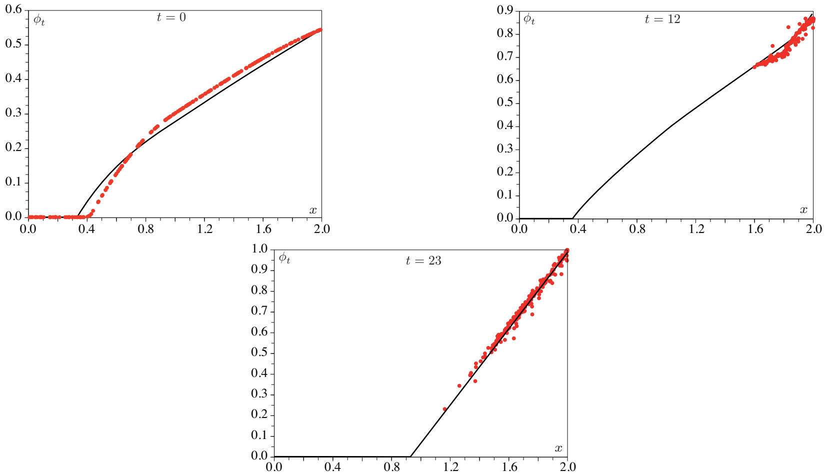

Figure: Scenario tree: optimal pairs (x, u) at three time instants

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 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.

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

P. Carpentier, G. Cohen and A. Dallagi.
Particle methods for stochastic optimal control problems.
arXiv, math.OC 0907.4663, 2009.