Multi-Stage Stochastic Optimization for Clean Energy Transition

◇

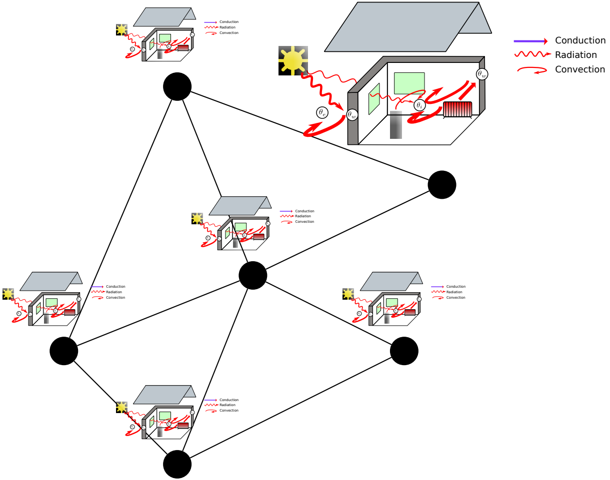
An Overview of Decomposition/Coordination Methods for Multistage Stochastic Optimization Problems

P. Carpentier — J.-P. Chancelier — M. De Lara

ENSTA Paris — ENPC ParisTech

BIRS-CMO Workshop 19w5091, 22–27 September 2019

Motivation



Conduction
Radiation
Convection
θ_e
θ_w
θ_i
θ_w

Lecture outline

Decomposition and coordination

- The three dimensions of stochastic optimization problems
- A bird's eye view of decomposition methods: the cube

A brief insight into three decomposition methods

- Scenario decomposition methods
- Spatial (price/resource) decomposition methods
- Time decomposition methods

Summary and research agenda

Outline of the presentation

Decomposition and coordination

A brief insight into three decomposition methods

Summary and research agenda

Outline of the presentation

Decomposition and coordination

- The three dimensions of stochastic optimization problems
- A bird's eye view of decomposition methods: the cube

A brief insight into three decomposition methods

- Scenario decomposition methods
- Spatial (price/resource) decomposition methods
- Time decomposition methods

Summary and research agenda

Temporal, scenario and spatial structures in multistage stochastic optimization problems

In multistage stochastic optimization problems,
we consider that the control variable

$$\mathbf{U}_t^i(\omega)$$

is indexed by

- Time/stages $t \in \mathbb{T}$ $(= \{0, \ldots, T-1\})$
- Scenarios $\omega \in \Omega$
- Space/units/agents $i \in \mathbb{I}$ $(= \{1, \ldots, N\})$

The letter U comes from the Russian word *upravlenie* for control

Let us fix problem and notations

$$\min_{\mathbf{U},\mathbf{X}} \overbrace{\mathbb{E}\Big(\sum_{i \in \mathbb{I}} \sum_{t \in \mathbb{T}} L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)}^{\text{additive costs}} \qquad \text{subject to}$$

dynamics constraints

$$\underbrace{\mathbf{X}_{t+1}^i}_{\text{state}} = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \underbrace{\mathbf{W}_{t+1}}_{\text{uncertainty}}) \ , \ \mathbf{X}_0^i = g_{-1}^i(\mathbf{W}_0)$$

measurability constraints (nonanticipativity of the control $\mathbf{U}_t^i$)

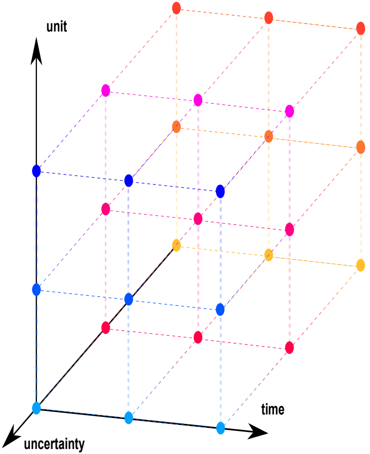
$$\sigma\big(\mathbf{U}_t^i\big) \subset \sigma(\mathbf{W}_0, \ldots, \mathbf{W}_t) \iff \mathbf{U}_t^i = \mathbb{E}\big(\mathbf{U}_t^i \mid \mathbf{W}_0, \ldots, \mathbf{W}_t\big)$$

spatially coupling constraints

$$\sum_{i \in \mathbb{I}} \Theta_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i) = 0$$

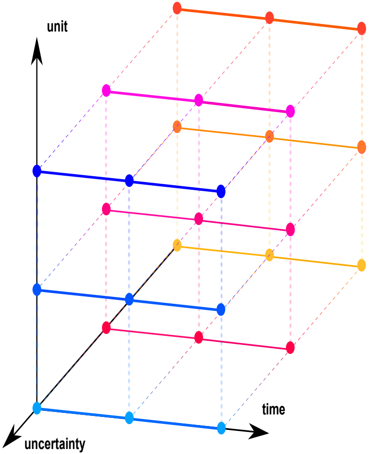
Outline of the presentation

Couplings for stochastic problems



$$\min \quad \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

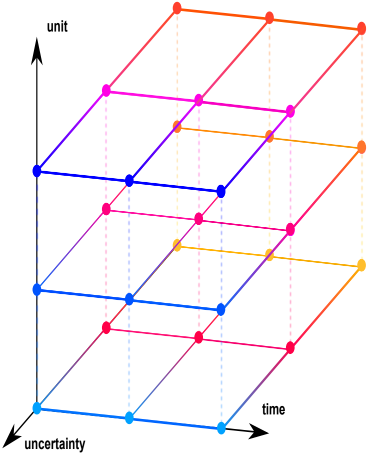
Couplings for stochastic problems: in time



$$\min \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

$$\text{s.t.} \quad \mathbf{X}_{t+1}^i = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1})$$

Couplings for stochastic problems: in uncertainty

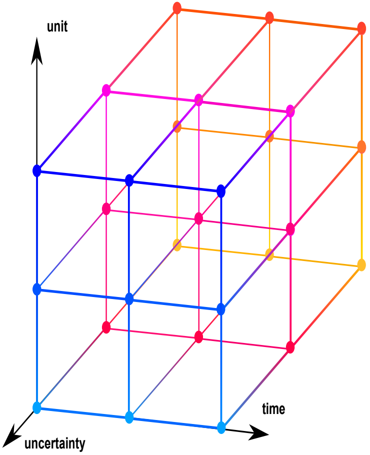


$$\min \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

$$\text{s.t. } \mathbf{X}_{t+1}^i = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1})$$

$$\mathbf{U}_t^i = \mathbb{E}\big(\mathbf{U}_t^i \mid \mathbf{W}_0, \dots, \mathbf{W}_t\big)$$

Couplings for stochastic problems: in space



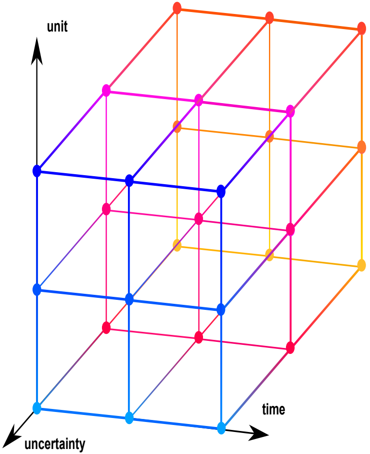
$$\min \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

s.t. $\mathbf{X}_{t+1}^i = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1})$

$$\mathbf{U}_t^i = \mathbb{E}\big(\mathbf{U}_t^i \bigm| \mathbf{W}_0, \dots, \mathbf{W}_t\big)$$

$$\sum_i \Theta_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i) = 0$$

Can we decouple stochastic optimization problems?



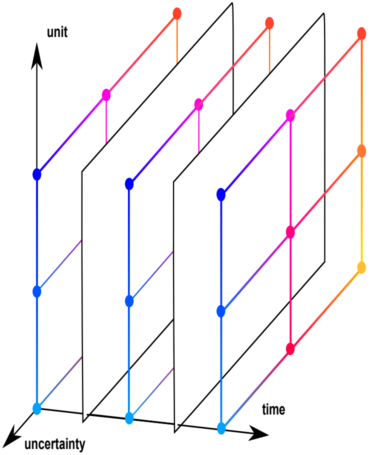
$$\min \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

s.t. $$\mathbf{X}_{t+1}^i = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1})$$

$$\mathbf{U}_t^i = \mathbb{E}\big(\mathbf{U}_t^i \mid \mathbf{W}_0, \dots, \mathbf{W}_t\big)$$

$$\sum_i \Theta_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i) = 0$$

Sequential decomposition in time



$$\min \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

s.t. $\mathbf{X}_{t+1}^i = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1})$

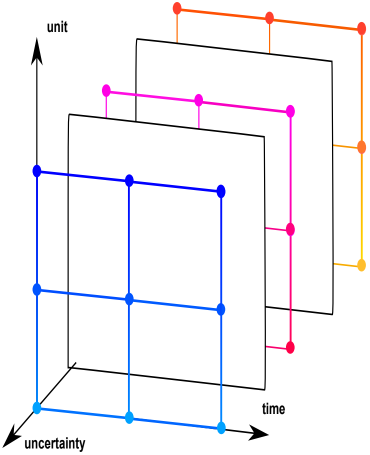
$$\mathbf{U}_t^i = \mathbb{E}\big(\mathbf{U}_t^i \mid \mathbf{W}_0, \dots, \mathbf{W}_t\big)$$

$$\sum_i \Theta_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i) = 0$$

Dynamic Programming (DP)

Bellman (56)

Parallel decomposition in uncertainty/scenarios



$$\min \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

s.t. $\mathbf{X}_{t+1}^i = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1})$

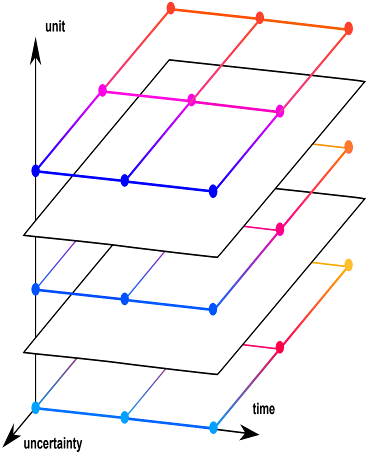
$$\mathbf{U}_t^i = \mathbb{E}\big(\mathbf{U}_t^i \mid \mathbf{W}_0, \dots, \mathbf{W}_t\big)$$

$$\sum_i \Theta_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i) = 0$$

Progressive Hedging

Rockafellar-Wets (91)

Parallel decomposition in space/units



$$\min \mathbb{E}\Big(\sum_i \sum_t L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) \Big)$$

$$\text{s.t.} \quad \mathbf{X}_{t+1}^i = g_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1})$$

$$\mathbf{U}_t^i = \mathbb{E}\big(\mathbf{U}_t^i \mid \mathbf{W}_0, \ldots, \mathbf{W}_t\big)$$

$$\sum_i \Theta_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i) = 0$$

Price and Resource Decompositions

Decomposition-coordination: divide and conquer

- Temporal decomposition
  - A state is an information summary
  - Time coordination realized through Dynamic Programming, by value functions (of the state)
  - Hard nonanticipativity constraints
- Scenario decomposition
  - Along each scenario, subproblems are deterministic (powerful algorithms)
  - Scenario coordination realized through Progressive Hedging, by updating nonanticipativity multipliers
  - Soft nonanticipativity constraints
- Spatial decomposition
  - By prices (multipliers of the spatial coupling constraint)
  - By resources (splitting the spatial coupling constraint)

Outline of the presentation

Decomposition and coordination

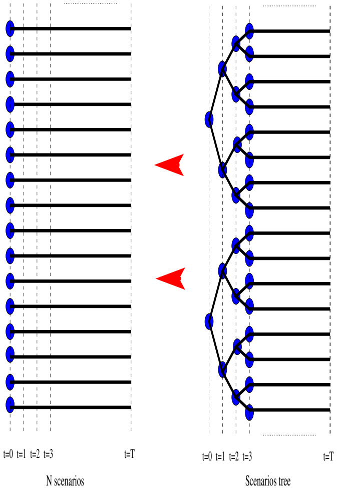
A brief insight into three decomposition methods

Summary and research agenda

Outline of the presentation

Moving from tree to fan (and scenarios)

Equivalent formulations of the nonanticipativity constraints



- On a (scenario) tree,
  the nonanticipativity constraints

  $$\sigma(\mathbf{U}_t) \subset \sigma(\mathbf{W}_0, \dots, \mathbf{W}_t)$$

  are "hardwired"
- On a fan,
  the nonanticipativity constraints
  write as linear equality constraints

  $$\mathbf{U}_t = \mathbb{E}\big(\mathbf{U}_t \mid \mathbf{W}_0, \dots, \mathbf{W}_t\big)$$

Progressive Hedging stands as a scenario decomposition method

Rockafellar-Wets (91) dualize the nonanticipativity constraints

$$\mathbf{U}_t = \mathbb{E}\big(\mathbf{U}_t \mid \mathbf{W}_0, \ldots, \mathbf{W}_t\big)$$

- When the criterion is strongly convex, one uses a Lagrangian relaxation (algorithm "à la Uzawa") to obtain a scenario decomposition
- When the criterion is linear, Rockafellar-Wets (91) propose to use an augmented Lagrangian, and obtain the Progressive Hedging algorithm

Data: step $\rho > 0$, initial multipliers $\left\{\lambda_s^{(0)}\right\}_{s\in\mathbb{S}}$ and mean first decision $\overline{\mathbf{u}}^{(0)}$;

Result: optimal first decision $\mathbf{u}$;

repeat

forall *scenarios* $s \in \mathbb{S}$ **do**

Solve the deterministic minimization problem for scenario s,

with a penalization $+\lambda_s^{(k)}\left(\mathbf{u}_s^{(k+1)} - \overline{\mathbf{u}}^{(k)}\right)$,

and obtain optimal first decision $\mathbf{u}_s^{(k+1)}$;

Update the mean first decisions

$$\overline{\mathbf{u}}^{(k+1)} = \sum_{s\in\mathbb{S}} \pi_s \mathbf{u}_s^{(k+1)} \ ;$$

Update the multiplier by

$$\lambda_s^{(k+1)} = \lambda_s^{(k)} + \rho\left(\mathbf{u}_s^{(k+1)} - \overline{\mathbf{u}}^{(k+1)}\right) \ , \ \ \forall s \in \mathbb{S} \ ;$$

until $\mathbf{u}_s^{(k+1)} - \sum_{s'\in\mathbb{S}} \pi_{s'} \mathbf{u}_{s'}^{(k+1)} = 0 \ , \ \ \forall s \in \mathbb{S}$;

Outline of the presentation

We consider an additive model

Consider the following minimization problem

$$\min_{u \in \mathcal{U}_{\mathrm{ad}} \subset \mathcal{U}} J(u) \quad \text{subject to} \quad \Theta(u) - \theta = 0 \in \mathcal{V}$$

for which exists a decomposition of the space $\mathcal{U} = \mathcal{U}^1 \times \ldots \times \mathcal{U}^N$, so that $u \in \mathcal{U}$ writes $u = \big(u^1, \ldots, u^N\big)$ with $u^i \in \mathcal{U}^i$, and also

- $\mathcal{U}_{\mathrm{ad}} = \mathcal{U}_{\mathrm{ad}}^1 \times \cdots \times \mathcal{U}_{\mathrm{ad}}^N$ $\qquad \mathcal{U}_{\mathrm{ad}}^i \subset \mathcal{U}^i$
- $J(u) = J^1(u^1) + \cdots + J^N(u^N)$ $\qquad u^i \in \mathcal{U}^i$
- $\Theta(u) = \Theta^1(u^1) + \cdots + \Theta^N(u^N)$ $\qquad u^i \in \mathcal{U}^i$

Then the problem displays the following additive structure

$$\min_{\substack{u^1 \in \mathcal{U}_{\mathrm{ad}}^1 \\ \vdots \\ u^N \in \mathcal{U}_{\mathrm{ad}}^N}} \sum_{i=1}^{N} J^i(u^i) \quad \text{subject to} \quad \sum_{i=1}^{N} \Theta^i(u^i) - \theta = 0$$

Additive model — Price decomposition I

$$\min_{u \in \mathcal{U}_{\mathrm{ad}}} \sum_{i=1}^{N} J^i(u^i) \quad \text{subject to} \quad \sum_{i=1}^{N} \Theta^i(u^i) - \theta = 0$$

1. Form the Lagrangian of the problem
   We assume that a saddle point exists,
   so that solving the initial problem is equivalent to

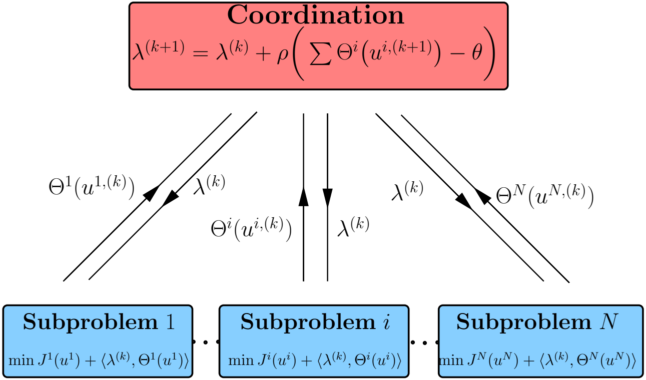
$$\max_{\lambda \in \mathcal{V}} \min_{u \in \mathcal{U}_{\mathrm{ad}}} \sum_{i=1}^{N} \Big(J^i(u^i) + \langle \lambda , \Theta^i(u^i) \rangle \Big) - \langle \lambda , \theta \rangle$$

2. Solve this problem by the Uzawa algorithm

$$u^{i,(k+1)} \in \operatorname*{arg\,min}_{u^i \in \mathcal{U}^i_{\mathrm{ad}}} J^i(u^i) + \langle \lambda^{(k)} , \Theta^i(u^i) \rangle , \quad i = 1 \ldots , N$$

$$\lambda^{(k+1)} = \lambda^{(k)} + \rho \Big(\sum_{i=1}^{N} \Theta^i \Big(u^{i,(k+1)} \Big) - \theta \Big)$$

Additive model — Price decomposition II



Additive model — Resource allocation I

$$\min_{u \in \mathcal{U}_{\mathrm{ad}}} \sum_{i=1}^{N} J^i(u^i) \quad \text{subject to} \quad \sum_{i=1}^{N} \Theta^i(u^i) - \theta = 0$$

1. Write the constraint in a equivalent manner by introducing new variables $v = (v^1, \ldots, v^N)$ (the so-called "allocation")

$$\sum_{i=1}^{N} \Theta^i(u^i) - \theta = 0 \quad \Leftrightarrow \quad \Theta^i(u^i) - v^i = 0 \ \text{and} \ \sum_{i=1}^{N} v^i = \theta$$

and minimize the criterion w.r.t. u and v

$$\min_{v \in \mathcal{V}^N} \sum_{i=1}^{N} \Big(\min_{u^i \in \mathcal{U}^i_{\mathrm{ad}}} J^i(u^i) \ \text{s.t.} \ \Theta^i(u^i) - v^i = 0 \Big) \ \text{s.t.} \ \sum_{i=1}^{N} v^i = \theta$$

Additive model — Resource allocation II

$$\min_{v\in\mathcal{V}^N}\sum_{i=1}^{N}\underbrace{\Big(\min_{u^i\in\mathcal{U}^i_{\mathrm{ad}}} J^i(u^i)\ \text{ s.t. }\ \Theta^i(u^i)-v^i=0\Big)}_{G^i(v^i)}\ \text{ s.t. }\ \sum_{i=1}^{N}v^i=\theta$$

$$\Updownarrow$$

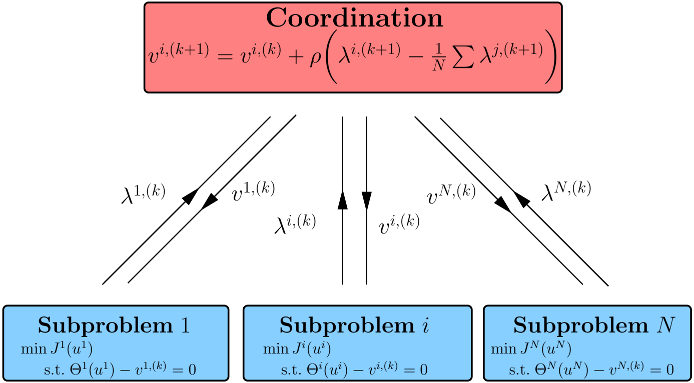
$$\min_{v\in\mathcal{V}^N}\sum_{i=1}^{N}G^i(v^i)\quad \text{s.t.}\quad \sum_{i=1}^{N}v^i=\theta$$

2. Solve the last problem using a projected gradient method

$$G^i(v^{i,(k)})=\min_{u^i\in\mathcal{U}^i_{\mathrm{ad}}} J^i(u^i)\ \text{ s.t. }\ \Theta^i(u^i)-v^{i,(k)}=0 \ \rightsquigarrow\ \lambda^{i,(k+1)}$$

$$v^{i,(k+1)}=v^{i,(k)}+\rho\Big(\lambda^{i,(k+1)}-\frac{1}{N}\sum_{j=1}^{N}\lambda^{j,(k+1)}\Big)$$

Additive model — Resource allocation III



Preparing Pierre Carpentier's talk

We can also use price/resource decomposition to bound a minimization problem

$$
\begin{aligned}
V_0^\star = \inf_{u^1 \in \mathbb{U}^1_{\mathrm{ad}}, \cdots, u^N \in \mathbb{U}^N_{\mathrm{ad}}} & \sum_{i=1}^N J^i(u^i) \\
\text{s.t.} \quad & \underbrace{\big(\Theta^1(u^1), \cdots, \Theta^N(u^N)\big) \in S}_{\text{coupling constraint}}
\end{aligned}
$$

- $u^i \in \mathbb{U}^i$ be a local decision variable
- $J^i : \mathbb{U}^i \to \mathbb{R},\ i \in [\![1, N]\!]$ be a local objective function
- $\mathbb{U}^i_{\mathrm{ad}}$ be a subset of the local decision set $\mathbb{U}^i$
- $\Theta^i : \mathbb{U}^i \to \mathcal{C}^i$ be a local constraint mapping
- S be a subset of $\mathcal{C} = \mathcal{C}^1 \times \cdots \times \mathcal{C}^N$

We denote by S^o the polar cone of S

$$
S^o = \big\{ p \in \mathcal{C}^\star \mid \langle p \,, r \rangle \leq 0 \,, \ \forall r \in S \big\}
$$

Price and resource local value functions

For each $i \in [\![1, N]\!]$,

- for any price $p^i \in (\mathcal{C}^i)^\star$, we define the local price value

$$\underline{V}_0^i[p^i] = \inf_{u^i \in \mathbb{U}^i_{\mathrm{ad}}} \; J^i(u^i) + \left\langle p^i \,, \Theta^i(u^i) \right\rangle$$

- for any resource $r^i \in \mathcal{C}^i$, we define the local resource value

$$\overline{V}_0^i[r^i] = \inf_{u^i \in \mathbb{U}^i_{\mathrm{ad}}} \; J^i(u^i) \quad \text{s.t.} \quad \Theta^i(u^i) = r^i$$

Proposition (upper and lower bounds for optimal value)

- *For any admissible price* $p = (p^1, \cdots, p^N) \in S^o$
- *For any admissible resource* $r = (r^1, \cdots, r^N) \in S$

$$\sum_{i=1}^N \underline{V}_0^i[p^i] \;\; \leq \;\; V_0^\star \;\; \leq \;\; \sum_{i=1}^N \overline{V}_0^i[r^i]$$

Outline of the presentation

Brief literature review on dynamic programming

	Bellman	Puterman	Bertsekas Schreve	Evstignev	Witsenhausen (standard form)
	1957	1994	1996	1976	1973
State	X	X	X	–	$(\omega, U_{1:t-1})$
Dynamics	$f(X, U, W)$	$P^u_{x,x'}$	$f(X, U, W)$	–	$X_t = (X_{t-1}, U_t)$
Uncertainties	Indep.	–	ρ	$(\Omega, \mathcal{F})$	$(\Omega, \mathcal{F})$
Cost	$\sum_t$	$\sum_t$	$\sum_t$	$j(\omega, U)$	$j(\omega, U)$
Controls	$\gamma(X)$	$\gamma(X)$ $\gamma(H)$	$\gamma(X)$ $\gamma(H)$	$\mathcal{F}_t$-meas.	$\gamma(x_t)$ $\mathcal{I}_t$-meas.
History	–	$(X, U, \ldots)_t$	$(W, U, \ldots)_t$	–	X_t

We introduce the history

- The timeline is

$$w_0 \rightsquigarrow u_0 \rightsquigarrow w_1 \rightsquigarrow u_1 \rightsquigarrow \quad \ldots \quad \rightsquigarrow w_{T-1} \rightsquigarrow u_{T-1} \rightsquigarrow w_T$$

- and the history is

$$\begin{aligned}\overbrace{h_t}^{\text{history}} &= (\overbrace{w_0}^{\text{uncertainty}}, \overbrace{u_0}^{\text{control}}, \overbrace{w_1}^{\text{uncertainty}}, u_1, \ldots, u_{t-1}, w_t) \\ &\in \mathbb{H}_t = \mathbb{W}_0 \times \prod_{s=0}^{t-1} (\underbrace{\mathbb{U}_s}_{\text{control space}} \times \underbrace{\mathbb{W}_{s+1}}_{\text{uncertainty space}})\end{aligned}$$

History is the largest state

The history follows the dynamics

$$\begin{aligned} h_{t+1} &= (\overbrace{w_0, u_0, w_1, u_1, \ldots, u_{t-1}, w_t}^{\text{history } h_t}, u_t, w_{t+1}) \\ &= \Big(h_t, \underbrace{u_t}_{\text{control}}, \underbrace{w_{t+1}}_{\text{uncertainty}}\Big) \end{aligned}$$

We formulate a sequence of minimization problems over increasing history spaces

- Once given
  - a criterion $j : \mathbb{H}_T \to \mathbb{R}$
  - a sequence of stochastic kernels $\rho_{t:t+1} : \mathbb{H}_t \to \Delta(\mathbb{W}_{t+1})$
- we define, for any history h_t, a minimization problem

$$V_t(h_t) = \underbrace{\inf_{\gamma_{t:T-1} \in \Gamma_{t:T-1}}}_{\text{history feedbacks}} \int_{\mathbb{H}_T} \overbrace{j(h'_T)}^{\text{criterion}} \underbrace{\rho^{\gamma}_{t:T}(h_t, \mathrm{d}h'_T)}_{\text{controlled stochastic kernel}}$$

There is a Bellman equation involving value functions over increasing history spaces without white noise assumption

$$
\begin{aligned}
V_T &= j \\
V_t &= \mathcal{B}_{t+1:t} V_{t+1}
\end{aligned}
$$

with

$$
\big(\mathcal{B}_{t+1:t}\varphi\big)(h_t) = \inf_{u_t \in \mathbb{U}_t} \int_{\mathbb{W}_{t+1}} \varphi(h_t, u_t, w_{t+1}) \rho_{t:t+1}(h_t, dw_{t+1})
$$

Preparing Jean-Philippe Chancelier's talk

Towards state reduction by time blocks

- History h_t is itself a canonical state variable, which lives in the history space $\mathbb{H}_t = \mathbb{W}_0 \times \prod_{s=0}^{t-1} (\mathbb{U}_s \times \mathbb{W}_{s+1})$
- However the size of this canonical state increases with t, which is a nasty feature for dynamic programming
- We will now
  - introduce "state" spaces $\mathbb{X}_t$
  - and then reduce the history with a mapping $\theta_r : \mathbb{H}_r \to \mathbb{X}_r$
  - to obtain a compressed "state" variable $\theta_t(h_t) = x_t \in \mathbb{X}_t$
  - but only at some specified times $0 = t_0 < t_1 < \cdots < t_N = T$
- As an application, we will handle stochastic independence between time blocks but possible dependence *within* time blocks

State reduction graphically

The triplet $(\theta_r, \theta_t, f_{r:t})$ is a state reduction across $(r:t)$ if

- the following diagram, for the dynamics, commutes

$$
\begin{tikzcd}
\mathbb{H}_r \times \mathbb{H}_{r+1:t} \arrow[r, "I_d"] \arrow[d, "\theta_r"] \arrow[d, "I_d"] & \mathbb{H}_t \arrow[d, "\theta_t"] \\
\mathbb{X}_r \times \mathbb{H}_{r+1:t} \arrow[r, "f_{r:t}"] & \mathbb{X}_t
\end{tikzcd}
$$

- the following diagrams, for the stochastic kernels, commute

$$
\begin{tikzcd}
\mathbb{H}_r \times \mathbb{H}_{r+1:s-1} \arrow[r, "\rho_{s-1:s}"] \arrow[d, "\theta_r"] \arrow[d, "I_d"] & \Delta(\mathbb{W}_s) \\
\mathbb{X}_r \times \mathbb{H}_{r+1:s-1} \arrow[ur, "\widetilde{\rho}_{s-1:s}"'] &
\end{tikzcd}
$$

Bellman operator across $(r:t)$

$\mathcal{B}_{r:t} : \mathbb{L}^0_+(\mathbb{H}_r, \mathcal{H}_r) \to \mathbb{L}^0_+(\mathbb{H}_t, \mathcal{H}_t)$ is defined by

$$\mathcal{B}_{r:t} = \mathcal{B}_{t+1:t} \circ \cdots \circ \mathcal{B}_{r:r-1} \ ,$$

where the one time step operators $\mathcal{B}_{s:s-1}$ are

$$\big(\mathcal{B}_{s:s-1}\varphi\big)(h_{s-1}) = \inf_{u_{s-1} \in \mathbb{U}_{s-1}} \int_{\mathbb{W}_s} \varphi(h_{s-1}, u_{s-1}, w_s) \rho_{s-1:s}(h_{s-1}, dw_s)$$

State reduction and Dynamic Programming

Denoting by $\theta_r^\star : \mathbb{L}_+^0(\mathbb{X}_r, \mathcal{X}_r) \to \mathbb{L}_+^0(\mathbb{H}_r, \mathcal{H}_r)$
the operator defined by

$$\theta_r^\star(\widetilde{\varphi}_r) = \widetilde{\varphi}_r \circ \theta_r \ , \ \ \forall \widetilde{\varphi}_r \in \mathbb{L}_+^0(\mathbb{X}_r, \mathcal{X}_r) \ ,$$

there exists a reduced Bellman operator across $(r\!:\!t)$ such that

$$\theta_t^\star \circ \widetilde{\mathcal{B}}_{r:t} = \mathcal{B}_{r:t} \circ \theta_r^\star \ ,$$

that is, the following diagram is commutative

$$\begin{tikzcd}
\mathbb{L}_+^0(\mathbb{H}_t, \mathcal{H}_t) \arrow[r, "\mathcal{B}_{t:r}"] & \mathbb{L}_+^0(\mathbb{H}_r, \mathcal{H}_r) \\
\mathbb{L}_+^0(\mathbb{X}_t, \mathcal{X}_t) \arrow[u, "\theta_t^\star"] \arrow[r, "\widetilde{\mathcal{B}}_{t:r}"] & \mathbb{L}_+^0(\mathbb{X}_r, \mathcal{X}_r) \arrow[u, "\theta_r^\star"']
\end{tikzcd}$$

Outline of the presentation

We have sketched three main decomposition methods in multistage stochastic optimization

- time: Dynamic Programming
- scenario: Progressive Hedging
- space: decomposition by prices or by resources

Numerical walls are well-known

- in dynamic programming,
  the bottleneck is the dimension of the state
- in stochastic programming,
  the bottleneck is the number of stages

Here is our research agenda for stochastic decomposition

- Designing risk criteria compatible with decomposition
- Combining different decomposition methods
  - time: Dynamic Programming
  - scenario: Progressive Hedging
  - space: decomposition by prices or by resources
- to produce blends and tackle large scale energy applications
  - time blocks + prices/resources
    (talk of Jean-Philippe Chancelier)
    - dynamic programming across time blocks
      + prices/resources decomposition by time block
    - application to two time scales battery management
  - time + space
    (talk of Pierre Carpentier)
    - nodal decomposition by prices or by resources
      + dynamic programming within node
    - application to large scale microgrid management