Introduction to Decomposition Methods in Stochastic Optimization

V. Leclère

2016, December 13

École des Ponts ParisTech

UNIVERSITÉ PARIS-EST

Presentation Outline

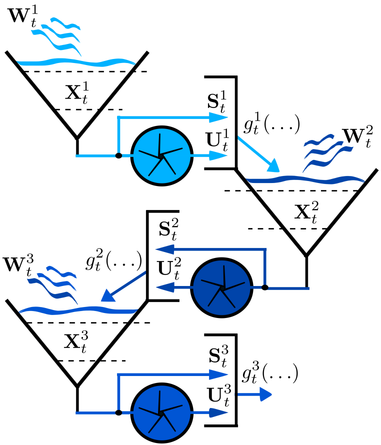
Mulstistage Stochastic Optimization: an Example

Objective function:

$$\mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=0}^{T-1} L_t^i(\underbrace{\boldsymbol{x}_t^i}_{\text{state}}, \underbrace{\boldsymbol{u}_t^i}_{\text{control}}, \underbrace{\boldsymbol{w}_{t+1}}_{\text{noise}})\right]$$

Constraints:

- dynamics:
  $\boldsymbol{x}_{t+1} = f_t(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{w}_{t+1})$,
- nonanticipativity:
  $\boldsymbol{u}_t \preceq \mathcal{F}_t$,
- spatial coupling:
  $\boldsymbol{z}_t^{i+1} = g_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}^i)$.



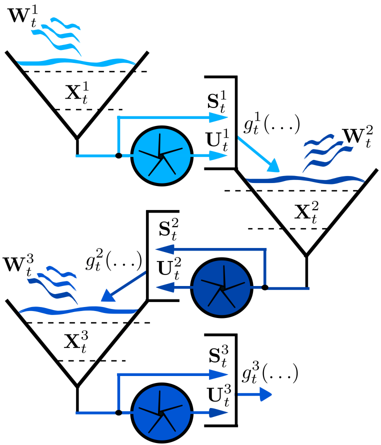
Mulstistage Stochastic Optimization: an Example

Objective function:

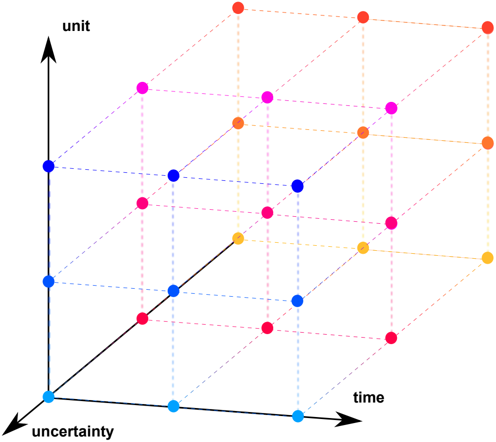
$$\mathbb{E}\Big[\sum_{i=1}^{N}\sum_{t=0}^{T-1} L_t^i(\underbrace{\boldsymbol{x}_t^i}_{\text{state}}, \underbrace{\boldsymbol{u}_t^i}_{\text{control}}, \underbrace{\boldsymbol{w}_{t+1}}_{\text{noise}})\Big]$$

Constraints:

- dynamics:
  $\boldsymbol{x}_{t+1} = f_t\big(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{w}_{t+1}\big)$,
- nonanticipativity:
  $\boldsymbol{u}_t \preceq \mathcal{F}_t$,
- spatial coupling:
  $\boldsymbol{z}_t^{i+1} = g_t^i\big(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}^i\big)$.

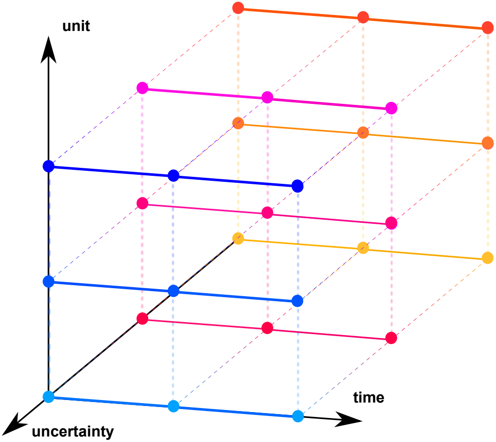


Couplings for Stochastic Problems



$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

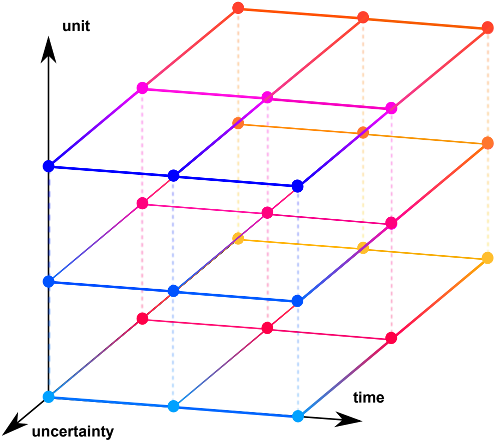
Couplings for Stochastic Problems: in Time



$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

s.t. $\boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$

Couplings for Stochastic Problems: in Uncertainty

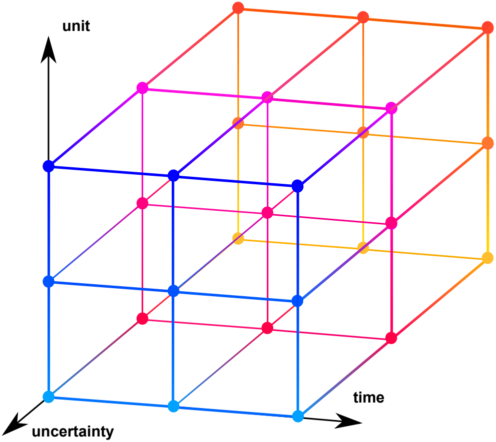


$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

s.t. $\boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$

$$\boldsymbol{u}_t^i \preceq \mathcal{F}_t = \sigma\big(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t\big)$$

Couplings for Stochastic Problems: in Space



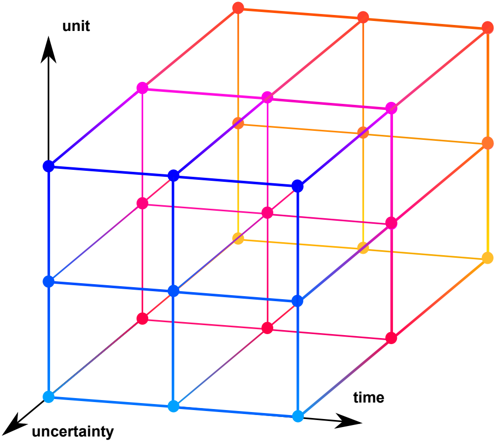
$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\text{s.t.} \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\boldsymbol{u}_t^i \preceq \mathcal{F}_t = \sigma\big(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t\big)$$

$$\sum_i \Theta_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i) = 0$$

Couplings for Stochastic Problems: a Complex Problem



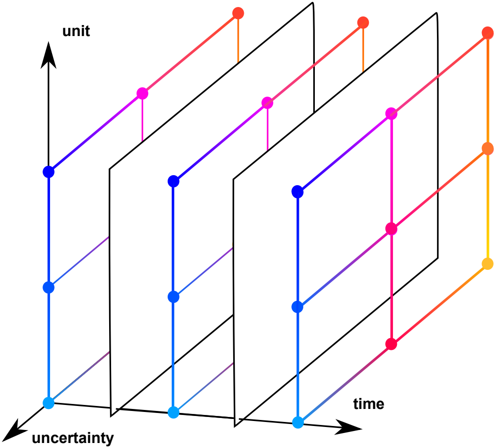
$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\text{s.t.} \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\boldsymbol{u}_t^i \preceq \mathcal{F}_t = \sigma\big(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t\big)$$

$$\sum_i \Theta_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i) = 0$$

Decompositions for Stochastic Problems: in Time



$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\text{s.t.} \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

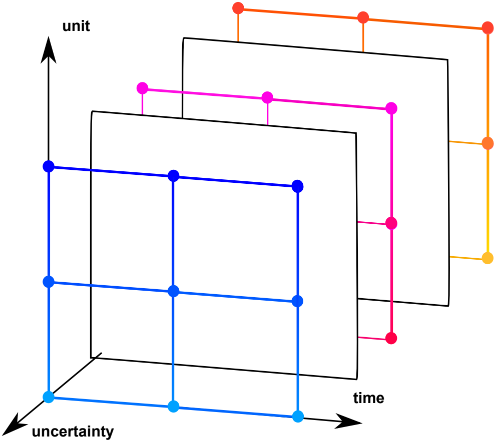
$$\boldsymbol{u}_t^i \preceq \mathcal{F}_t = \sigma\big(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t\big)$$

$$\sum_i \Theta_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i) = 0$$

Dynamic Programming

Bellman (56)

Decompositions for Stochastic Problems: in Uncertainty



$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\text{s.t.} \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

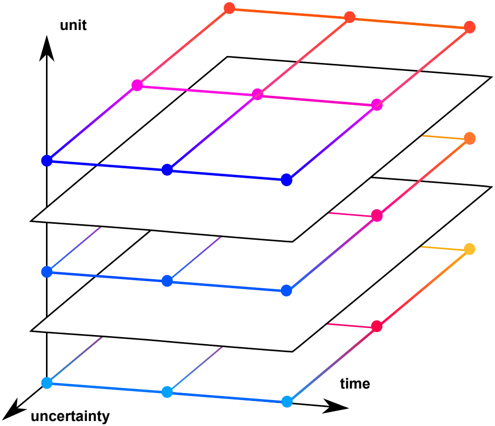
$$\boldsymbol{u}_t^i \preceq \mathcal{F}_t = \sigma\big(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t\big)$$

$$\sum_i \Theta_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i) = 0$$

Progressive Hedging

Rockafellar - Wets (91)

Decompositions for Stochastic Problems: in Space



$$\min \sum_{\omega} \sum_{i} \sum_{t} \pi_\omega L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\text{s.t.} \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1})$$

$$\boldsymbol{u}_t^i \preceq \mathcal{F}_t = \sigma\big(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t\big)$$

$$\sum_i \Theta_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i) = 0$$

Dual Approximate
Dynamic Programming

Presentation Outline

1. Decompositions of Mulstistage Stochastic Optimization
2. Dynamic Programming
3. Spatial Decomposition

Optimization Problem

We want to solve the following optimization problem

$$\min \quad \mathbb{E}\Big[\sum_{t=0}^{T-1} L_t\big(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{w}_{t+1}\big) + K\big(\boldsymbol{x}_T\big)\Big] \tag{1a}$$

$$s.t. \quad \boldsymbol{x}_{t+1} = f_t(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{w}_{t+1}), \qquad \boldsymbol{x}_0 = x_0 \tag{1b}$$

$$\boldsymbol{u}_t \in U_t(\boldsymbol{x}_t) \tag{1c}$$

$$\sigma(\boldsymbol{u}_t) \subset \sigma\big(\boldsymbol{w}_0, \cdots, \boldsymbol{w}_t\big) \tag{1d}$$

Dynamic Programming Principle

Assume that the noises $\boldsymbol{w}_t$ are independent and exogeneous.

Then, there exists an optimal solution, called a strategy, of the form $\boldsymbol{u}_t = \pi_t(\boldsymbol{x}_t)$, given by

$$\pi_t(x) = \underset{u \in U_t(x)}{\arg\min}\, \mathbb{E}\Big[\underbrace{L_t(x, u, \boldsymbol{w}_{t+1})}_{\text{current cost}} + \underbrace{V_{t+1} \circ f_t(x, u, \boldsymbol{w}_{t+1})}_{\text{future costs}} \Big] ,$$

where (Dynamic Programming Equation)

$$\begin{cases} V_T(x) &= K(x) \\ V_t(x) &= \min_{u \in U_t(x)} \mathbb{E}\Big[L_t(x, u, \boldsymbol{w}_{t+1}) + V_{t+1} \circ \underbrace{f_t(x, u, \boldsymbol{w}_{t+1})}_{"\boldsymbol{X}_{t+1}"} \Big] \end{cases}$$

Dynamic Programming Principle

Assume that the noises $\boldsymbol{w}_t$ are independent and exogeneous.

Then, there exists an optimal solution, called a strategy, of the form $\boldsymbol{u}_t = \pi_t(\boldsymbol{x}_t)$, given by

$$\pi_t(x) = \arg\min_{u \in U_t(x)} \mathbb{E}\Big[\underbrace{L_t(x, u, \boldsymbol{w}_{t+1})}_{\text{current cost}} + \underbrace{V_{t+1} \circ f_t\big(x, u, \boldsymbol{w}_{t+1}\big)}_{\text{future costs}} \Big] ,$$

where (Dynamic Programming Equation)

$$\begin{cases} V_T(x) & = K(x) \\ V_t(x) & = \min_{u \in U_t(x)} \mathbb{E}\Big[L_t(x, u, \boldsymbol{w}_{t+1}) + V_{t+1} \circ \underbrace{f_t\big(x, u, \boldsymbol{w}_{t+1}\big)}_{\text{"}\boldsymbol{X}_{t+1}\text{"}} \Big] \end{cases}$$

Interpretation of Bellman Value

The Bellman's value function $V_{t_0}(x)$ can be interpreted as the value of the problem starting at time t_0 from the state x. More precisely we have

$$\begin{aligned} V_{t_0}(x) = \min \quad & \mathbb{E}\Big[\sum_{t=t_0}^{T-1} L_t\big(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{w}_{t+1}\big) + K\big(\boldsymbol{x}_T\big)\Big] && \text{(2a)} \\ s.t. \quad & \boldsymbol{x}_{t+1} = f_t(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{w}_{t+1}), \qquad \boldsymbol{x}_{t_0} = x && \text{(2b)} \\ & \boldsymbol{u}_t \in U_t(\boldsymbol{x}_t) && \text{(2c)} \\ & \sigma(\boldsymbol{u}_t) \subset \sigma\big(\boldsymbol{w}_0, \cdots, \boldsymbol{w}_t\big) && \text{(2d)} \end{aligned}$$

Dynamic Programming Algorithm : Discrete Case

Data: Problem parameters
Result: optimal control and value;
$V_T \equiv K$;
for $t : T \to 0$ **do**
    for $x \in \mathbb{X}_t$ **do**
        $V_t(x) = \infty$;
        for $u \in U_t(x)$ **do**
            $v_u = \mathbb{E}\Big[L_t(x, u, \boldsymbol{w}_{t+1}) + V_{t+1} \circ f_t\big(x, u, \boldsymbol{w}_{t+1}\big)\Big]$;
            if $v_u < V_t(x)$ **then**
                $V_t(x) = v_u$;
                $\pi_t(x) = u$;
            end
        end
    end
end

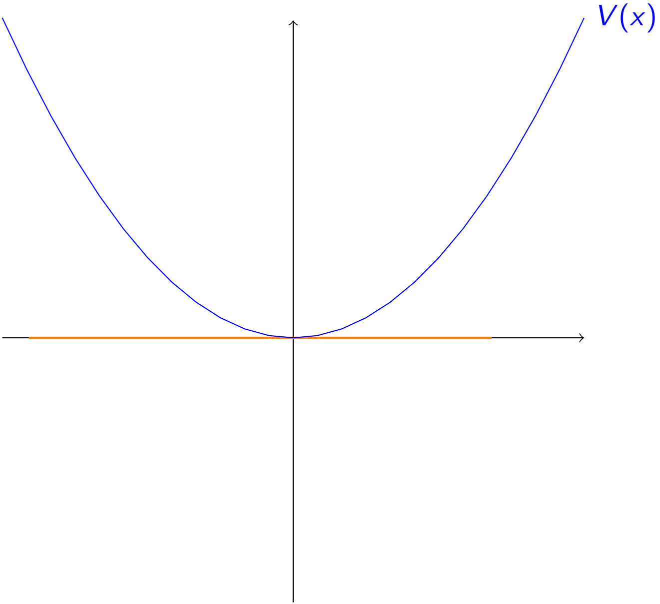
Number of flops: $O(T \times |\mathbb{X}_t| \times |\mathbb{U}_t| \times |\mathbb{W}_t|)$.

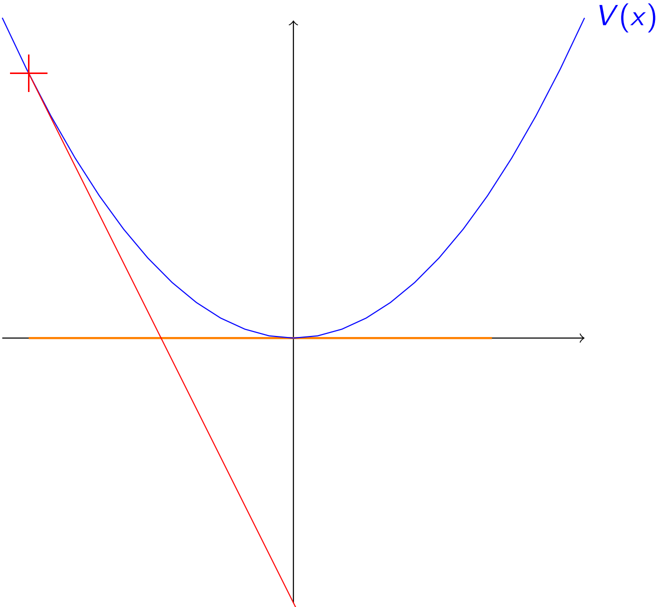
3 curses of dimensionality

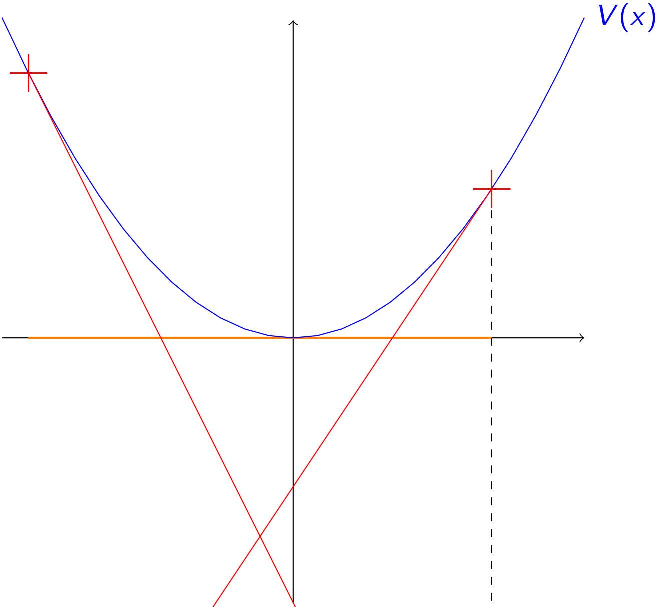
1. State. If we consider 3 independent states each taking 10 values, then $|\mathbb{X}_t| = 10^3 = 1000$. In practice DP is not applicable for states of dimension more than 5.
2. Decision. The decision are often vector decisions, that is a number of independent decision, hence leading to huge $|U_t(x)|$.
3. Expectation. In practice random information came from large data set. Without a proper statistical treatment computing an expectation is costly. Monte-Carlo approach are costly too, and unprecise.

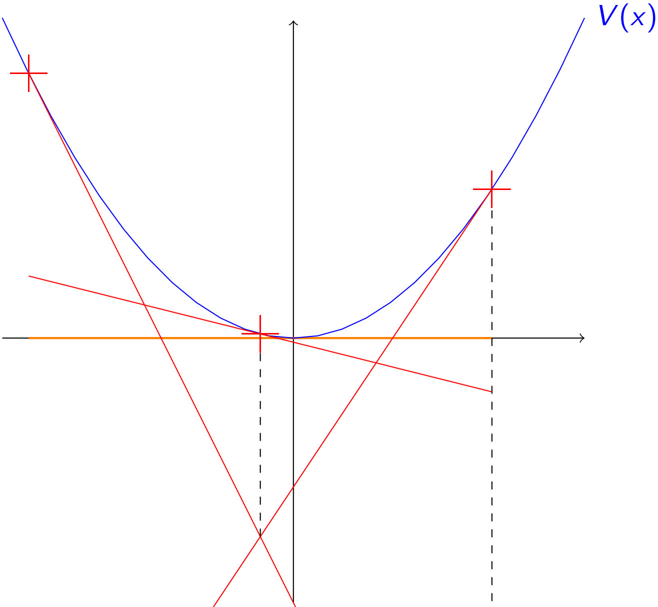
Dynamic Programming : continuous and convex case

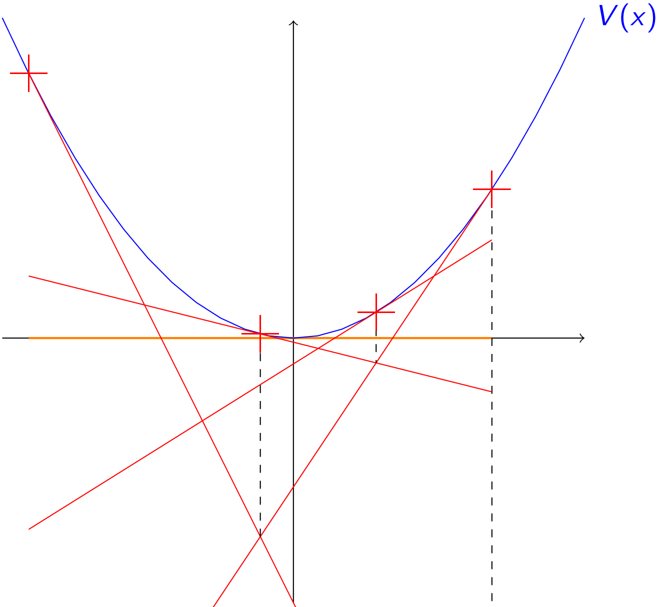
- If the problem has continuous states and control the classical approach consists in discretizing.
- With further assumption on the problem (convexity, linearity) we can look at a dual approach:
  - Instead of discretizing and interpolating the Bellman function we choose to do a polyhedral approximation.
  - Indeed we choose a "smart state" in which we compute the value of the function and its marginal value (tangeant).
  - Knowing that the problem is convex and using the power of linear solver we can efficiently approximate the Bellman function.
- This approach is known as SDDP in the electricity community and widely used in practice.


$V(x)$


$V(x)$


$V(x)$

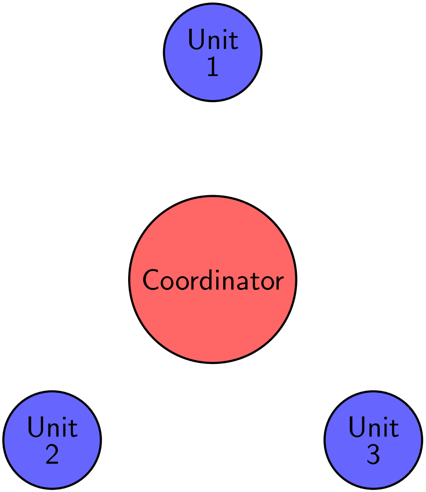

$V(x)$


$V(x)$

Presentation Outline

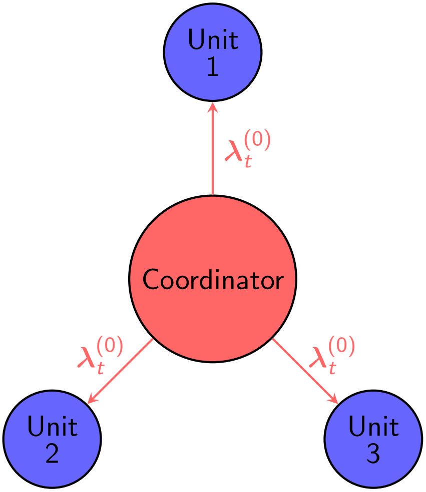
Intuition of Spatial Decomposition

- Satisfy a demand (over T time step) with N units of production at minimal cost.
- Price decomposition:
  - the coordinator sets a sequence of price λ_t,
  - the units send their production planning $\boldsymbol{u}_t^{(i)}$,
  - the coordinator compares total production and demand and updates the price,
  - and so on...



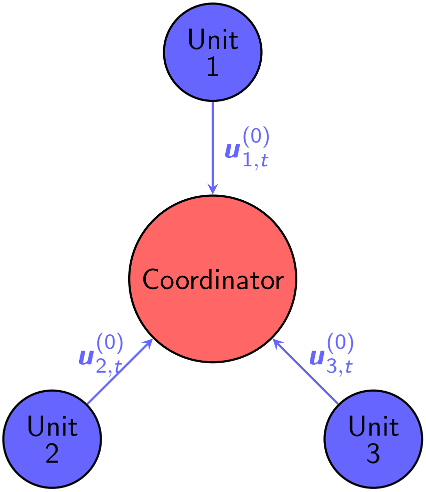
Intuition of Spatial Decomposition

- Satisfy a demand (over T time step) with N units of production at minimal cost.
- Price decomposition:
  - the coordinator sets a sequence of price λ_t,
  - the units send their production planning $\boldsymbol{u}_t^{(i)}$,
  - the coordinator compares total production and demand and updates the price,
  - and so on...



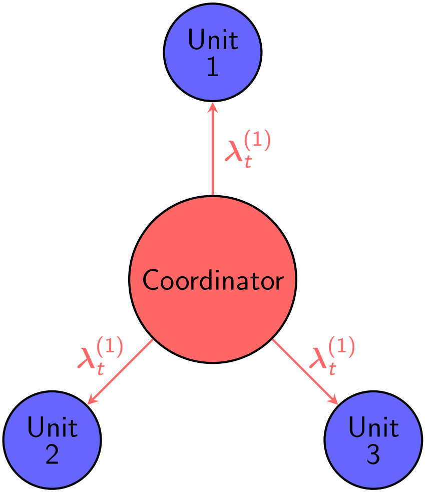
Intuition of Spatial Decomposition

- Satisfy a demand (over T time step) with N units of production at minimal cost.
- Price decomposition:
  - the coordinator sets a sequence of price λ_t,
  - the units send their production planning $\boldsymbol{u}_t^{(i)}$,
  - the coordinator compares total production and demand and updates the price,
  - and so on...



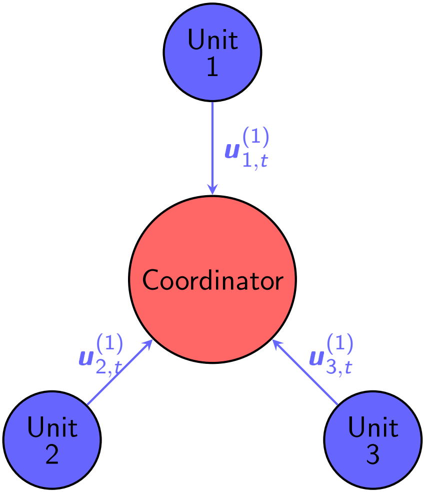
Intuition of Spatial Decomposition

- Satisfy a demand (over T time step) with N units of production at minimal cost.
- Price decomposition:
  - the coordinator sets a sequence of price λ_t,
  - the units send their production planning $\boldsymbol{u}_t^{(i)}$,
  - the coordinator compares total production and demand and updates the price,
  - and so on...



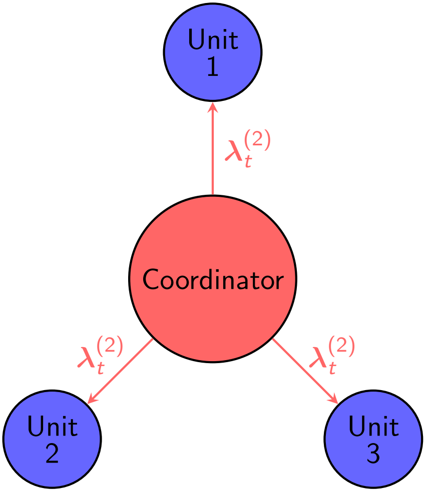
Intuition of Spatial Decomposition

- Satisfy a demand (over T time step) with N units of production at minimal cost.
- Price decomposition:
  - the coordinator sets a sequence of price λ_t,
  - the units send their production planning $\boldsymbol{u}_t^{(i)}$,
  - the coordinator compares total production and demand and updates the price,
  - and so on...



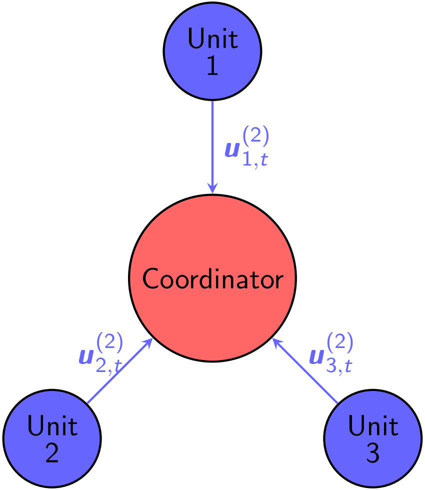
Intuition of Spatial Decomposition

- Satisfy a demand (over T time step) with N units of production at minimal cost.
- Price decomposition:
  - the coordinator sets a sequence of price λ_t,
  - the units send their production planning $\boldsymbol{u}_t^{(i)}$,
  - the coordinator compares total production and demand and updates the price,
  - and so on...

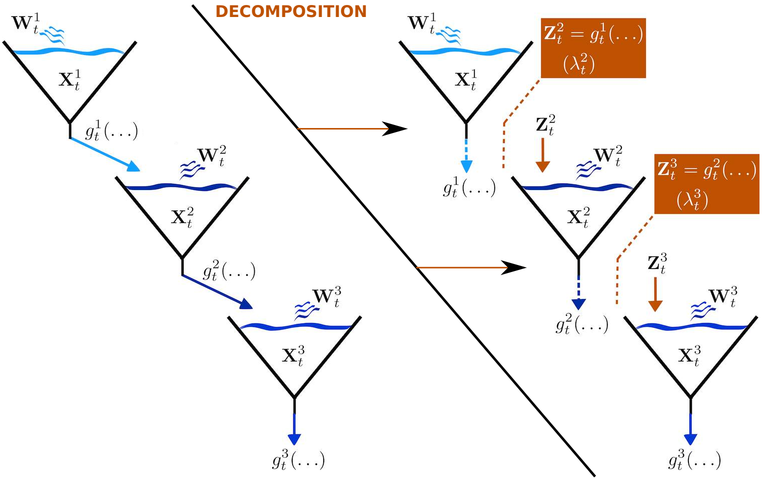


Intuition of Spatial Decomposition

- Satisfy a demand (over T time step) with N units of production at minimal cost.
- Price decomposition:
  - the coordinator sets a sequence of price λ_t,
  - the units send their production planning $\boldsymbol{u}_t^{(i)}$,
  - the coordinator compares total production and demand and updates the price,
  - and so on...



Application to dam management



Primal Problem

$$\min_{\boldsymbol{x},\boldsymbol{u}} \sum_{i=1}^{N} \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}) + K^i(\boldsymbol{x}_T^i)\Big]$$

$$\forall\ i, \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}), \quad \boldsymbol{x}_0^i = x_0^i,$$

$$\forall\ i, \quad \boldsymbol{u}_t^i \in \mathcal{U}_{t,i}^{ad}, \quad \boldsymbol{u}_t^i \preceq \mathcal{F}_t,$$

$$\sum_{i=1}^{N} \theta_t^i(\boldsymbol{u}_t^i) = 0$$

Solvable by DP with state $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N)$

Primal Problem

$$\min_{\boldsymbol{x},\boldsymbol{u}} \sum_{i=1}^{N} \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}) + K^i(\boldsymbol{x}_T^i)\Big]$$

$$\forall\ i, \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}), \quad \boldsymbol{x}_0^i = x_0^i,$$

$$\forall\ i, \quad \boldsymbol{u}_t^i \in \mathcal{U}_{t,i}^{ad}, \quad \boldsymbol{u}_t^i \preceq \mathcal{F}_t,$$

$$\sum_{i=1}^{N} \theta_t^i(\boldsymbol{u}_t^i) = 0 \qquad \rightsquigarrow \boldsymbol{\lambda}_t \quad \text{multiplier}$$

Solvable by DP with state $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N)$

Primal Problem with Dualized Constraint

$$\min_{\boldsymbol{x},\boldsymbol{u}} \max_{\boldsymbol{\lambda}} \sum_{i=1}^{N} \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}) + \langle \boldsymbol{\lambda}_t, \theta_t^i(\boldsymbol{u}_t^i)\rangle + K^i(\boldsymbol{x}_T^i)\Big]$$

$$\forall\, i, \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}), \quad \boldsymbol{x}_0^i = x_0^i,$$

$$\forall\, i, \quad \boldsymbol{u}_t^i \in \mathcal{U}_{t,i}^{ad}, \quad \boldsymbol{u}_t^i \preceq \mathcal{F}_t,$$

Coupling constraint dualized $\Longrightarrow$ all constraints are unit by unit

Dual Problem

$$\max_{\boldsymbol{\lambda}} \min_{\boldsymbol{x},\boldsymbol{u}} \sum_{i=1}^{N} \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}) + \langle \boldsymbol{\lambda}_t, \theta_t^i(\boldsymbol{u}_t^i)\rangle + K^i(\boldsymbol{x}_T^i)\Big]$$

$$\forall\, i, \quad \boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}), \quad \boldsymbol{x}_0^i = x_0^i,$$

$$\forall\, i, \quad \boldsymbol{u}_t^i \in \mathcal{U}_{t,i}^{ad}, \quad \boldsymbol{u}_t^i \preceq \mathcal{F}_t,$$

Exchange operator min and max to obtain a new problem

Decomposed Dual Problem

$$\max_{\boldsymbol{\lambda}} \sum_{i=1}^{N} \min_{\boldsymbol{x}^i,\boldsymbol{u}^i} \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}) + \langle \boldsymbol{\lambda}_t, \theta_t^i(\boldsymbol{u}_t^i)\rangle + K^i(\boldsymbol{x}_T^i)\Big]$$

$$\boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}), \quad \boldsymbol{x}_0^i = x_0^i,$$

$$\boldsymbol{u}_t^i \in \mathcal{U}_{t,i}^{ad}, \quad \boldsymbol{u}_t^i \preceq \mathcal{F}_t,$$

For a given λ, minimum of sum is sum of minima

Inner Minimization Problem

$$\min_{\boldsymbol{x}^i, \boldsymbol{u}^i} \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}) + \langle \boldsymbol{\lambda}_t, \theta_t^i(\boldsymbol{u}_t^i) \rangle + K^i(\boldsymbol{x}_T^i)\Big]$$

$$\boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}), \quad \boldsymbol{x}_0^i = x_0^i,$$

$$\boldsymbol{u}_t^i \in \mathcal{U}_{t,i}^{ad}, \quad \boldsymbol{u}_t^i \preceq \mathcal{F}_t,$$

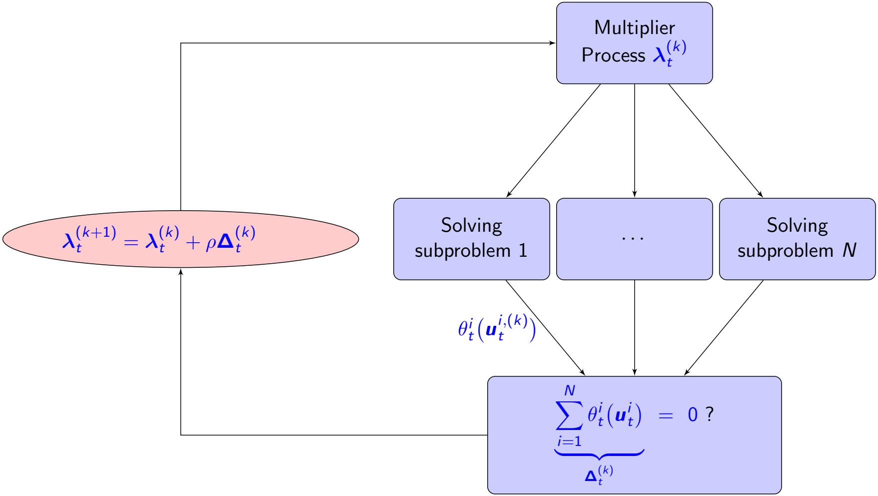
We have N smaller subproblems. Can they be solved by DP ?

Inner Minimization Problem

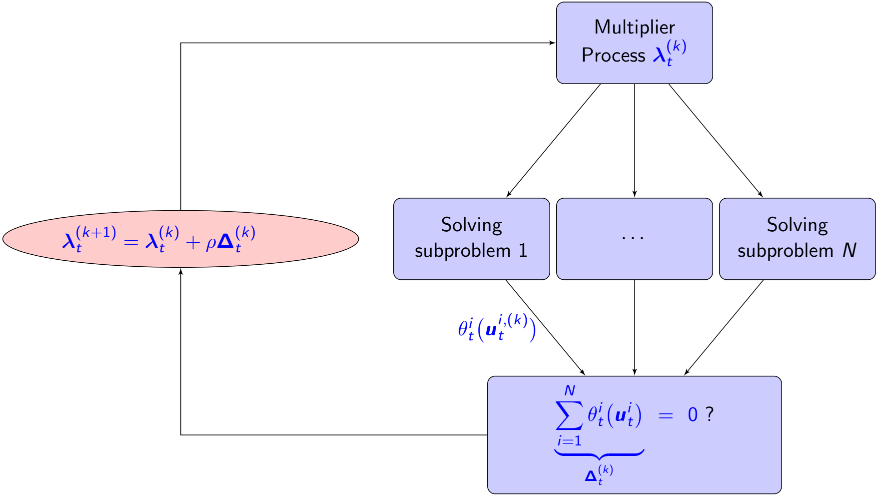
$$\min_{\boldsymbol{x}^i, \boldsymbol{u}^i} \mathbb{E}\Big[\sum_{t=0}^{T} L_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}) + \langle \boldsymbol{\lambda}_t, \theta_t^i(\boldsymbol{u}_t^i)\rangle + K^i(\boldsymbol{x}_T^i)\Big]$$
$$\boldsymbol{x}_{t+1}^i = f_t^i(\boldsymbol{x}_t^i, \boldsymbol{u}_t^i, \boldsymbol{w}_{t+1}), \quad \boldsymbol{x}_0^i = x_0^i,$$
$$\boldsymbol{u}_t^i \in \mathcal{U}_{t,i}^{ad}, \quad \boldsymbol{u}_t^i \preceq \mathcal{F}_t,$$

No : $\boldsymbol{\lambda}$ is a time-dependent noise $\rightsquigarrow$ state $\big(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t\big)$

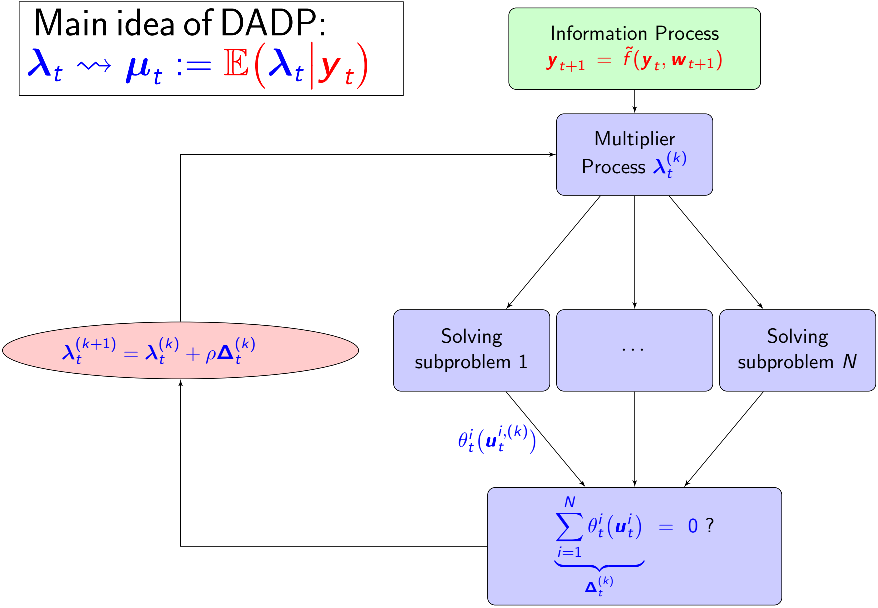
Stochastic spatial decomposition scheme



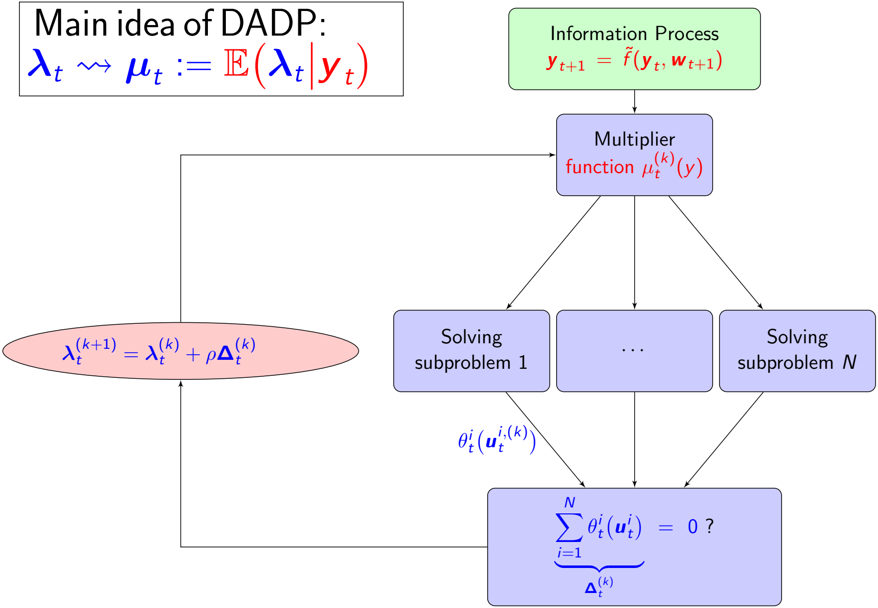
Main idea of DADP:
$\boldsymbol{\lambda}_t \rightsquigarrow \boldsymbol{\mu}_t := \mathbb{E}\big(\boldsymbol{\lambda}_t \big| \boldsymbol{y}_t\big)$



Main idea of DADP:
$\boldsymbol{\lambda}_t \rightsquigarrow \boldsymbol{\mu}_t := \mathbb{E}\big(\boldsymbol{\lambda}_t \big| \boldsymbol{y}_t\big)$

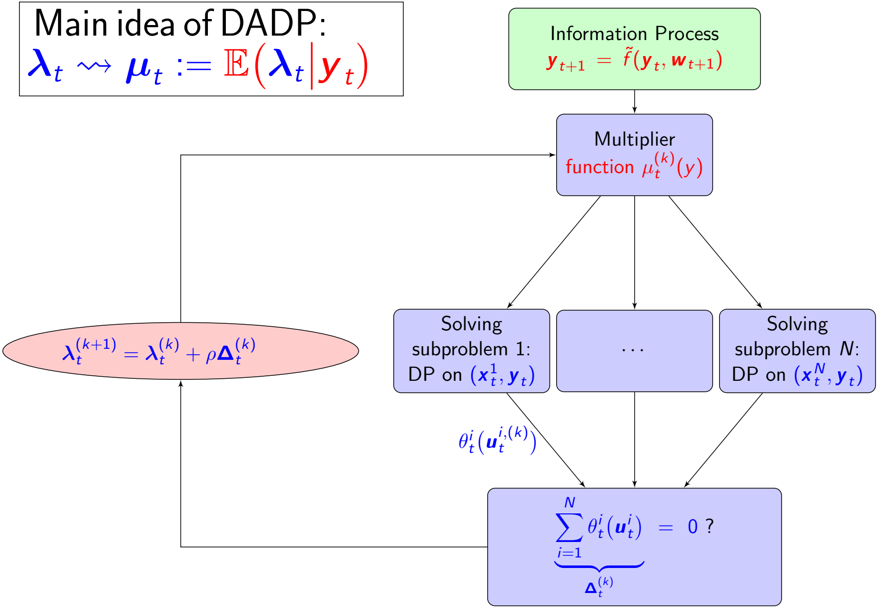


Main idea of DADP:
$$\boldsymbol{\lambda}_t \rightsquigarrow \boldsymbol{\mu}_t := \mathbb{E}\big(\boldsymbol{\lambda}_t \big| \boldsymbol{y}_t\big)$$



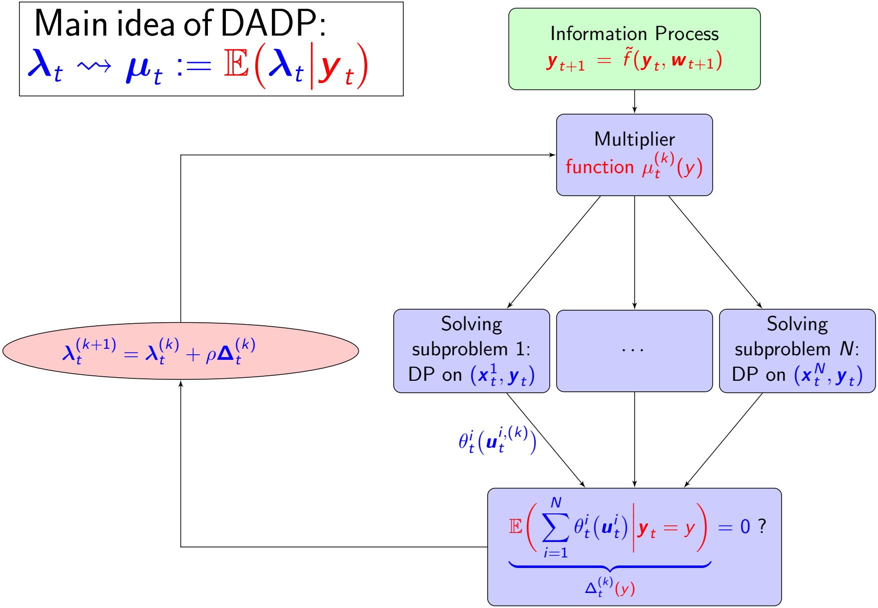
Main idea of DADP:

$$\boldsymbol{\lambda}_t \rightsquigarrow \boldsymbol{\mu}_t := \mathbb{E}\big(\boldsymbol{\lambda}_t \big| \boldsymbol{y}_t\big)$$

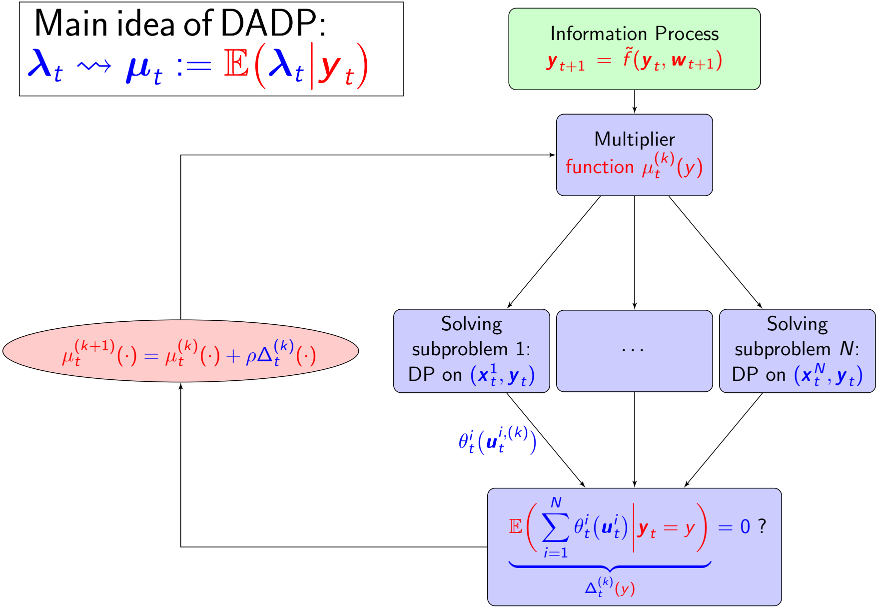


Main idea of DADP:

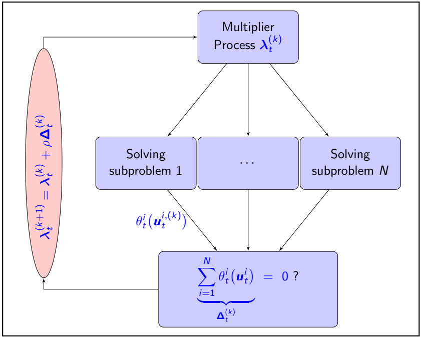
$$\boldsymbol{\lambda}_t \rightsquigarrow \boldsymbol{\mu}_t := \mathbb{E}\big(\boldsymbol{\lambda}_t \big| \boldsymbol{y}_t\big)$$

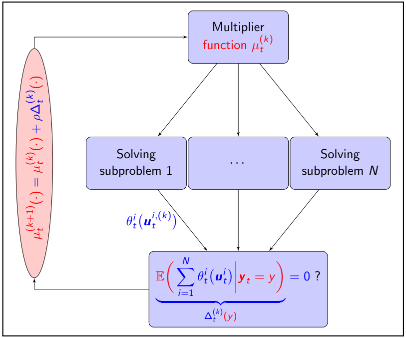


Main idea of DADP:
$$\boldsymbol{\lambda}_t \rightsquigarrow \boldsymbol{\mu}_t := \mathbb{E}\big(\boldsymbol{\lambda}_t \big| \boldsymbol{y}_t\big)$$



Main idea of DADP: $\boldsymbol{\lambda}_t \rightsquigarrow \boldsymbol{\mu}_t := \mathbb{E}\left(\boldsymbol{\lambda}_t \middle| \boldsymbol{y}_t\right)$





Main problems:

- Subproblems not easily solvable by DP
- $\boldsymbol{\lambda}^{(k)}$ live in a huge space

Advantages:

- Subproblems solvable by DP with state $\left(\boldsymbol{x}_t^i, \boldsymbol{y}_t\right)$
- $\mu^{(k)}$ live in a smaller space

Three Interpretations of DADP

- DADP as an approximation of the optimal multiplier

$$\boldsymbol{\lambda}_t \quad \rightsquigarrow \quad \mathbb{E}\big(\boldsymbol{\lambda}_t \big| \boldsymbol{y}_t\big) .$$

- DADP as a decision-rule approach in the dual

$$\max_{\boldsymbol{\lambda}} \min_{\boldsymbol{u}} L\big(\boldsymbol{\lambda}, \boldsymbol{u}\big) \quad \rightsquigarrow \quad \max_{\boldsymbol{\lambda}_t \preceq \boldsymbol{y}_t} \min_{\boldsymbol{u}} L\big(\boldsymbol{\lambda}, \boldsymbol{u}\big) .$$

- DADP as a constraint relaxation in the primal

$$\sum_{i=1}^{n} \theta_t^i\big(\boldsymbol{u}_t^i\big) = 0 \quad \rightsquigarrow \quad \mathbb{E}\Big(\sum_{i=1}^{n} \theta_t^i\big(\boldsymbol{u}_t^i\big) \Big| \boldsymbol{y}_t\Big) = 0 .$$

Conclusion

- Large multistage stochastic program are numerically difficult.
- To tackle such problems one can use decomposition methods.
- If the number of stages is small enough, decomposition per scenario (like Progressive-Hedging) is numerically efficient, and use special deterministic methods.
- If the noises are time-independent Dynamic Programming equations are available.
  - If the state dimension is small enough direct discretized dynamic programming is available.
  - If dynamics is linear and cost are convex SDDP approach allow for larger states
  - Finally we can also spatially decompose problems, and with an approximation recover Dynamic Programming equations for the subproblems.