

# Séminaire EDF

## Méthodes numériques pour la commande optimale



Méthodes classiques et alternatives  
en commande optimale stochastique.

Pierre Carpentier, Pierre Girardeau et SOWG<sup>1</sup>

EDF — ENPC — ENSTA

2 avril 2009

---

<sup>1</sup>**Systems and Optimization Working Group**

L. Andrieu, K. Barty, P. Carpentier, J.-P. Chancelier, G. Cohen, A. Dallagi, P. Girardeau & M. de Lara



# Plan de l'exposé

- 1 Introduction et motivations
- 2 Méthodes particulières
- 3 Décomposition

- 1 Introduction et motivations
  - Formulation du problème
  - Programmation dynamique : limitations
  - Arbres de scénarios : avantages et inconvénients
  - Objectifs pour la suite
- 2 Méthodes particulières
- 3 Décomposition

On pose un problème d'optimisation stochastique en **temps discret** (instant final  $T$ ), structure d'information causale (**boucle fermée**).

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace de probabilité. Les variables aléatoires intervenant dans le problème d'optimisation sont les suivantes.

- $\mathbf{W}_t : \Omega \rightarrow \mathbb{W}$  : **bruit** affectant le système à l'instant  $t$ .
- $\mathbf{U}_t : \Omega \rightarrow \mathbb{U}$  : **commande** appliquée à l'instant  $t$ .  
On note  $\mathbf{U} = (\mathbf{U}_0, \dots, \mathbf{U}_{T-1})$ , et  $\mathcal{U} = L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{U}^T)$ .
- $\mathbf{X}_t : \Omega \rightarrow \mathbb{X}$  : **état** physique du système à l'instant  $t$ .  
On note  $\mathbf{X} = (\mathbf{X}_0, \dots, \mathbf{X}_T)$ , et  $\mathcal{X} = L^2(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{X}^{T+1})$ .
- $\mathbf{Y}_t : \Omega \rightarrow \mathbb{Y}$  : **information** disponible à l'instant  $t$ .

**Exemples** :  $\mathbf{Y}_t = (\mathbf{W}_0, \dots, \mathbf{W}_t)$  ou  $\mathbf{Y}_t = \mathbf{X}_t$ .

## Problème d'optimisation stochastique en temps discret

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

sous les contraintes  $\mathbf{X}_0 = \mathbf{W}_0$ ,

(dynamique)  $\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}), \quad (1b)$

(mesurabilité)  $\mathbf{U}_t \preceq \mathbf{Y}_t, \quad (1c)$

(ponctuelle)  $\mathbf{U}_t \in \mathcal{U}_t^{\text{as}} \quad \mathbb{P}\text{-p.s.} \quad (1d)$

## Forme épurée du problème

$$\min_{\mathbf{U} \in \mathcal{U}} \mathbb{E}(j(\mathbf{U}, \mathbf{W})), \quad (2a)$$

sous les contraintes  $\mathbf{U} \in \mathcal{U}^{\text{me}}, \quad (2b)$

$\mathbf{U} \in \mathcal{U}^{\text{as}}. \quad (2c)$

## Cadre markovien

Sous les hypothèses  $\mathbf{Y}_t = \mathbf{X}_t$  et  $(\mathbf{W}_0, \dots, \mathbf{W}_T)$  est un bruit blanc, le contrôle optimal à  $t$  est en **feedback** sur l'état :  $\mathbf{U}_t = \gamma_t(\mathbf{X}_t)$ .

## Algorithme de la programmation dynamique

Pour tout  $x \in \mathbb{X}$ ,

$$V_T(x) = K(x),$$

$$V_t(x) = \min_{u \in \mathbb{U}_t^{\text{as}}} \mathbb{E} \left( L_t(x, u, \mathbf{W}_{t+1}) + V_{t+1}(f_t(x, u, \mathbf{W}_{t+1})) \right).$$

## Difficultés d'application

- Explosion combinatoire en  $x$  (malédiction de la dimension).
- Pas d'application directe des méthodes de décomposition.

## Principes de la méthode des arbres de scénarios

L'approche est basée sur un schéma **Discrétisation – Optimisation**.

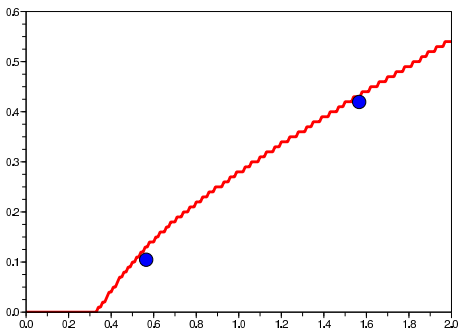
On suppose (pour simplifier) que  $\mathbf{Y}_t = (\mathbf{W}_0, \dots, \mathbf{W}_t)$

- Partant de  $N$  scénarios  $(\mathbf{W}_t^k)_{t=0, \dots, T}^{k=1, \dots, N}$ , générer un **arbre** : en tout nœud, passé unique et (idéalement) plusieurs futurs.
- Écrire le problème d'optimisation sur l'arbre (la contrainte  $\mathbf{U}_t \preceq (\mathbf{W}_0, \dots, \mathbf{W}_t)$  est automatiquement satisfaite...)
- Résoudre ( $\rightsquigarrow$  couple de particules  $(\mathbf{X}_i^\#, \mathbf{U}_i^\#)$  en tout nœud  $i$ ).

## Principales caractéristiques

- Résultats **aléatoires**.
- Explosion combinatoire en  $t$ .
- Pas de difficulté particulière de décomposition.
- Pour  $t \approx 0$ , petit nombre de particules de bonne qualité.

$t = 0$  ; number of points = 2



$t = 12$  ; number of points = 200

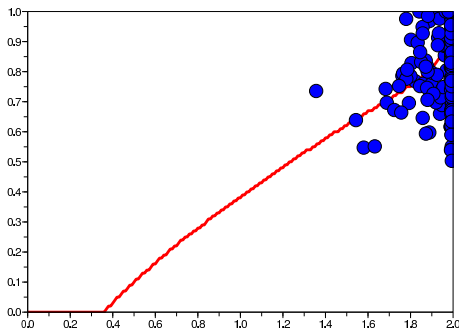



Figure: Programmation dynamique / Arbres de scénarios. 



On s'intéresse tout particulièrement à trouver des méthodes de résolution qui :

- permettent d'obtenir des **stratégies** de commandes (comme dans la programmation dynamique),
- permettent de s'attaquer à des problèmes de **grande taille** (comme dans la programmation stochastique).

- 1 Introduction et motivations
- 2 Méthodes particulières
  - Principes et outils
  - Conditions d'optimalité
  - Algorithme et résultats
  - Conclusions
- 3 Décomposition

L'approche est basée sur un schéma **Optimisation – Discrétisation**.

- On écrit les conditions d'optimalité du problème (1).
- On approxime les opérateurs d'espérance (conditionnelle) apparaissant dans ces conditions à l'aide de scénarios de bruit.

Les conditions d'optimalité approximées permettent de mettre en œuvre des algorithmes de type gradient pour résoudre le problème.

On cherche donc :

- à se placer dans la lignée des méthodes de type Monte Carlo (comme la méthodes des arbres de scénarios) . . .
- sans pour autant avoir à construire un tel arbre de scénarios (qui “appauvrit” les chroniques de bruits dont on dispose).

Référence : [Dallagi, 2007].

## Condition d'optimalité d'un problème sous contraintes

Soit  $\mathbf{U}^\#$  une solution du problème :

$$\min_{\mathbf{U} \in \mathcal{U}} \mathbb{E} (j(\mathbf{U}, \mathbf{W})) \quad \text{sous} \quad \mathbf{U} \in \mathcal{U}^{\text{me}} \cap \mathcal{U}^{\text{as}}.$$

- $\mathcal{U}^{\text{me}} = L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{U})$ ,  $\mathcal{F}$  sous-tribu de  $\mathcal{A}$  (mesurabilité),
- $\mathcal{U}^{\text{as}} = \{\mathbf{U} \in \mathcal{U}, \mathbf{U}(\omega) \in \mathbb{U}^{\text{as}} \text{ } \mathbb{P}\text{-p.s.}\}$  (contraintes ponctuelles).

Une condition nécessaire d'optimalité est :

$$\begin{aligned} \mathbf{U}^\# &= \text{proj}_{\mathcal{U}^{\text{as}} \cap \mathcal{U}^{\text{me}}} \left( \mathbf{U}^\# - \epsilon j_u'^{\top}(\mathbf{U}^\#, \mathbf{W}) \right), \\ &= \text{proj}_{\mathcal{U}^{\text{as}}} \left( \mathbf{U}^\# - \epsilon \mathbb{E} (j_u'^{\top}(\mathbf{U}^\#, \mathbf{W}) \mid \mathcal{F}) \right). \end{aligned}$$

La preuve s'appuie sur les deux résultats suivants :

- Notant  $J(\mathbf{U}) = \mathbb{E} (j(\mathbf{U}, \mathbf{W}))$ , on montre que :  $J'(\mathbf{U})(\omega) = j_u'(\mathbf{U}(\omega), \mathbf{W}(\omega))$ ,  $\mathbb{P}$ -p.s..
- Les deux ensembles  $\mathcal{U}^{\text{as}}$  et  $\mathcal{U}^{\text{me}}$  sont tels que l'on a :  $\text{proj}_{\mathcal{U}^{\text{as}} \cap \mathcal{U}^{\text{me}}} = \text{proj}_{\mathcal{U}^{\text{as}}} \circ \text{proj}_{\mathcal{U}^{\text{me}}}$ .

## Élimination des variables d'état

Incorporation récursive de la dynamique (1b) dans le coût (1a) pour éliminer les variables **intermédiaires** d'état  $\mathbf{X}_t$  :

$$j(u, w) = \sum_{t=0}^{T-1} \tilde{L}_t(u_0, \dots, u_t, w_0, \dots, w_{t+1}) + \tilde{K}(u_0, \dots, u_{T-1}, w_0, \dots, w_T),$$

## Dérivation de la fonction $j$ (technique de l'état-adjoint)

$$(j)'_{u_t}(u, w) = (L_t)'_u(x_t, u_t, w_{t+1}) + \lambda_{t+1}^\top (f_t)'_u(x_t, u_t, w_{t+1}).$$

- $(x_0, \dots, x_T)$  suit la dynamique "forward" :

$$x_0 = w_0, \quad x_{t+1} = f_t(x_t, u_t, w_{t+1}).$$

- $(\lambda_1, \dots, \lambda_T)$  est **choisi** pour suivre la dynamique "backward" :

$$\lambda_T = K'^\top(x_T), \quad \lambda_t = (L_t)'_x^\top(x_t, u_t, w_{t+1}) + (f_t)'_x^\top(x_t, u_t, w_{t+1})\lambda_{t+1}.$$

## Condition d'optimalité : cas non-adapté

On suppose que  $\mathbf{Y}_t = (\mathbf{W}_0, \dots, \mathbf{W}_t)$  et on note  $\mathcal{F}_t = \sigma(\mathbf{Y}_t)$ .  
Soit  $(\mathbf{U}^\#, \mathbf{X}^\#)$  une solution du problème (1). Alors, il existe un processus  $(\lambda_1^\#, \dots, \lambda_T^\#)$  vérifiant :

$$\mathbf{X}_0^\# = \mathbf{W}_0, \quad \mathbf{X}_{t+1}^\# = f_t(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) \quad (\text{pour mémoire}),$$

$$\lambda_T^\# = K'^\top(\mathbf{X}_T^\#), \quad \lambda_t^\# = (L_t)_x'^\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) + (f_t)_x'^\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1})\lambda_{t+1}^\#,$$

$$\mathbf{U}_t^\# = \text{proj}_{\mathcal{U}_t^{\text{as}}} \left( \mathbf{U}_t^\# - \epsilon \mathbb{E} \left( (L_t)_u'^\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) + (f_t)_u'^\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1})\lambda_{t+1}^\# \mid \mathcal{F}_t \right) \right).$$

La preuve est une conséquence des deux résultats vus précédemment :

- calcul du gradient par la technique de l'état-adjoint,
- conditions d'optimalité sur l'ensemble  $\mathcal{U}^{\text{me}} \cap \mathcal{U}^{\text{as}}$ .

## Condition d'optimalité : cas adapté

On obtient un résultat analogue au précédent en remplaçant le processus  $(\lambda_1^\#, \dots, \lambda_T^\#)$  par sa version adaptée :

$$\Lambda_t^\# = \mathbb{E}(\lambda_t^\# \mid \mathcal{F}_t).$$

Les conditions d'optimalité deviennent alors :

$$\Lambda_T^\# = K'^\top(\mathbf{X}_T^\#),$$

$$\Lambda_t^\# = \mathbb{E}\left(\left(L_t\right)'_\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) + \left(f_t\right)'_\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1})\Lambda_{t+1}^\# \mid \mathcal{F}_t\right),$$

$$\mathbf{U}_t^\# = \text{proj}_{\mathbb{U}_t^{\text{as}}}\left(\mathbf{U}_t^\# - \epsilon \mathbb{E}\left(\left(L_t\right)'_\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) + \left(f_t\right)'_\top(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1})\Lambda_{t+1}^\# \mid \mathcal{F}_t\right)\right).$$

La preuve est une conséquence directe de  $\mathbb{E}(\cdot \mid \mathcal{F}_t) = \mathbb{E}(\mathbb{E}(\cdot \mid \mathcal{F}_{t+1}) \mid \mathcal{F}_t)$ .

## Condition d'optimalité : cas markovien

Sous l'hypothèse  $(\mathbf{W}_0, \dots, \mathbf{W}_T)$  bruit blanc, on montre que la commande et l'état-adjoint sont mesurables par rapport à l'état :

$$\begin{aligned}\boldsymbol{\Lambda}_{t+1}^{\#} \preceq \mathbf{X}_{t+1}^{\#} &\quad \rightsquigarrow \quad \boldsymbol{\Lambda}_{t+1}^{\#} = \mu_{t+1}(\mathbf{X}_{t+1}^{\#}), \\ \mathbf{U}_t^{\#} \preceq \mathbf{X}_t^{\#} &\quad \rightsquigarrow \quad \mathbf{U}_t^{\#} = \gamma_t(\mathbf{X}_t^{\#}),\end{aligned}$$

Les conditions d'optimalité deviennent dans ce cas :

$$\boldsymbol{\Lambda}_T^{\#} = K'^T(\mathbf{X}_T^{\#}),$$

$$\boldsymbol{\Lambda}_t^{\#} = \mathbb{E} \left( (L_t)'_x(\mathbf{X}_t^{\#}, \mathbf{U}_t^{\#}, \mathbf{W}_{t+1}) + (f_t)'_x(\mathbf{X}_t^{\#}, \mathbf{U}_t^{\#}, \mathbf{W}_{t+1}) \boldsymbol{\Lambda}_{t+1}^{\#} \mid \mathbf{X}_t^{\#} \right),$$

$$\mathbf{U}_t^{\#} = \text{proj}_{\mathbb{U}_t^{\text{ad}}} \left( \mathbf{U}_t^{\#} - \epsilon \mathbb{E} \left( (L_t)'_u(\mathbf{X}_t^{\#}, \mathbf{U}_t^{\#}, \mathbf{W}_{t+1}) + (f_t)'_u(\mathbf{X}_t^{\#}, \mathbf{U}_t^{\#}, \mathbf{W}_{t+1}) \boldsymbol{\Lambda}_{t+1}^{\#} \mid \mathbf{X}_t^{\#} \right) \right).$$

La preuve se fait par récurrence. Le point techniquement délicat est de montrer que  $\mathbf{U}_t^{\#} \preceq \mathbf{X}_t^{\#}$ .



## Passage de l'espérance conditionnelle à l'espérance

On reprend le calcul de l'état-adjoint :

$$\Lambda_t^\# = \mathbb{E} \left( (L_t)'_{\mathbf{x}} (\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) + (f_t)'_{\mathbf{x}} (\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) \Lambda_{t+1}^\# \mid \mathbf{X}_t^\# \right).$$

- Comme  $\Lambda_{t+1}^\# = \mu_{t+1}(\mathbf{X}_{t+1}^\#) = \mu_{t+1}(f_t(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}))$ , le terme sous l'espérance conditionnelle est de la forme  $H_t(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1})$  :

$$\Lambda_t^\# = \mathbb{E} \left( H_t(\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) \mid \mathbf{X}_t^\# \right).$$

- Comme  $\mathbf{U}_t^\# \preceq \mathbf{X}_t^\#$  et comme  $\mathbf{W}_{t+1}$  et  $\mathbf{X}_t^\#$  sont indépendantes, le calcul de cette espérance conditionnelle revient à "figer" les variables aléatoires  $\mathbf{X}_t^\#$  et  $\mathbf{U}_t^\#$  et à intégrer par rapport à  $\mathbf{W}_{t+1}$ . C'est donc un calcul d'espérance.

On obtient une conclusion similaire pour le calcul du gradient :

$$\mathbb{E} \left( (L_t)'_{\mathbf{u}} (\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) + (f_t)'_{\mathbf{u}} (\mathbf{X}_t^\#, \mathbf{U}_t^\#, \mathbf{W}_{t+1}) \Lambda_{t+1}^\# \mid \mathbf{X}_t^\# \right).$$

## Initialisation.

- Tirer un  $N$ -échantillon  $(\mathbf{W}_t^k)_{t=0, \dots, T}^{k=1, \dots, N}$  des trajectoires du bruit.
- Initialiser les trajectoires  $(\mathbf{U}_t^k)_{t=0, \dots, T-1}^{k=1, \dots, N}$  de la commande.

## Itération $\ell$ .

- Intégrer l'état le long de chaque trajectoire du bruit :

$$\begin{aligned}\mathbf{x}_0^k &= \mathbf{W}_0^k, \\ \mathbf{x}_{t+1}^k &= f_t(\mathbf{x}_t^k, \mathbf{U}_t^k, \mathbf{W}_{t+1}^k).\end{aligned}$$

- Intégrer l'état-adjoint en temps rétrograde :

$$\boldsymbol{\Lambda}_T^k = K'^\top(\mathbf{x}_T^k),$$

$$\boldsymbol{\Lambda}_t^k = \frac{1}{N} \sum_{j=1}^N (L_t)'^\top(\mathbf{x}_t^k, \mathbf{U}_t^k, \mathbf{W}_{t+1}^j) + (f_t)'^\top(\mathbf{x}_t^k, \mathbf{U}_t^k, \mathbf{W}_{t+1}^j) \hat{\mu}_{t+1}(f_t(\mathbf{x}_t^k, \mathbf{U}_t^k, \mathbf{W}_{t+1}^j)).$$

L'**espérance** est obtenue par approximation de Monte Carlo et l'approximation  $\hat{\mu}_{t+1}$  de  $\mu_{t+1}$  est un **opérateur d'interpolation** basé sur les couples  $(\mathbf{x}_{t+1}^k, \boldsymbol{\Lambda}_{t+1}^k)_{k=1, \dots, N}$ .

## Itération $\ell$ (suite).

- Calculer les particules de gradient :

$$\mathbf{G}_t^k = \frac{1}{N} \sum_{j=1}^N (L_t)'_u (\mathbf{X}_t^k, \mathbf{U}_t^k, \mathbf{W}_{t+1}^j) + (f_t)'_u (\mathbf{X}_t^k, \mathbf{U}_t^k, \mathbf{W}_{t+1}^j) \hat{\mu}_{t+1} (f_t(\mathbf{X}_t^k, \mathbf{U}_t^k, \mathbf{W}_{t+1}^j)).$$

- Mettre à jour les particules de commande :

$$(\mathbf{U}_t^k)^+ = \text{proj}_{\mathcal{U}_t^{\text{as}}} (\mathbf{U}_t^k - \epsilon \mathbf{G}_t^k).$$

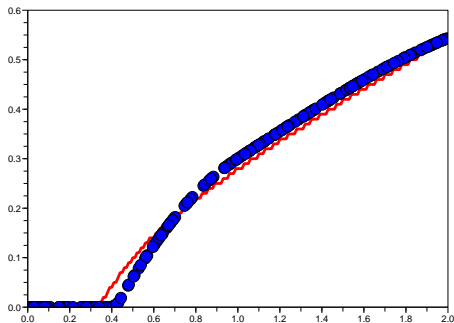
- Test de convergence.

## Étape finale.

- Obtenir une approximation  $\hat{\gamma}_t$  du feedback  $\gamma_t$  à l'aide d'un opérateur d'interpolation basé sur les couples  $(\mathbf{X}_t^k, \mathbf{U}_t^k)^{k=1, \dots, N}$ .

**Remarque.** On pourrait faire les itérations sur les lois de feedback approximées  $\hat{\gamma}_t$  plutôt que sur les particules  $(\mathbf{U}_t^k)^{k=1, \dots, N}$ .

t = 0 ; number of points = 200



t = 12 ; number of points = 200

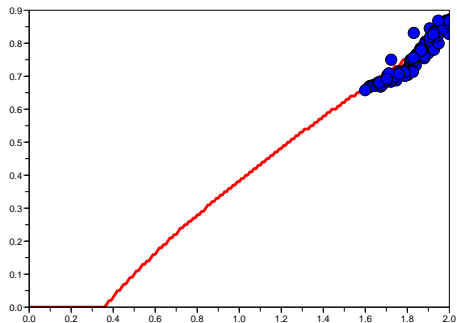


Figure: Programmation dynamique / Méthodes particulières.



On dispose d'une batterie d'algorithmes de nature variationnelle.

Ces algorithmes utilisent des chroniques "brutes" de scénarios.

L'algorithme se cantonne aux zones de l'espace où "vit" la solution.

Dans le cas Markovien, on se ramène à des calculs d'espérance.

Influence de l'opérateur d'interpolation dans la mise en œuvre ?

- 1 Introduction et motivations
- 2 Méthodes particulières
- 3 **Décomposition**
  - Un problème séparable
  - Décomposition par les prix "classique"
  - Dual Approximate Dynamic Programming

# Introduction

- Nous supposons maintenant que le problème de commande optimale a une *structure décomposable*.
- On est à la recherche de stratégies en *feedback* sur un état : le cadre naturel est celui de la programmation dynamique.

## But

L'objectif de la décomposition est de remplacer la résolution d'un "grand problème" par la résolution de plusieurs problèmes de plus petite taille.

# Notations

On considère trois types de variables aléatoires :

- un état  $\mathbf{X}_t^i$  à l'instant  $t$  pour l'unité  $i$  ;
- une commande  $\mathbf{U}_t^i$  à l'instant  $t$  pour l'unité  $i$  ;
- un aléa  $\mathbf{W}_t$  à l'instant  $t$ .

On supposera ici :

- que les bruits sont blancs (ou qu'ils ont été "blanchis") ;
- qu'on observe tout le passé des bruits.



# Le problème

## Problème considéré

$$\left\{ \begin{array}{l}
 \min_{\mathbf{X}, \mathbf{U}} \mathbb{E} \left( \sum_{t=1}^{T-1} \sum_{i=1}^n L_t^i(\mathbf{x}_t^i, \mathbf{u}_t^i, \mathbf{w}_{t+1}) + \sum_{i=1}^n K^i(\mathbf{x}_T^i) \right) \\
 \text{s.c. } \mathbf{x}_{t+1}^i = f_t^i(\mathbf{x}_t^i, \mathbf{u}_t^i, \mathbf{w}_{t+1}), \quad \boxed{\text{couplage temporel}} \\
 \sum_{i=1}^n \mathbf{u}_t^i = 0, \quad \boxed{\text{couplage spatial}} \\
 \mathbf{u}_t \text{ est } \sigma\{\mathbf{w}_1, \dots, \mathbf{w}_t\}\text{-mesurable,} \quad \boxed{\text{couplage probabiliste}} \\
 + \text{ contraintes de borne sur } \mathbf{X} \text{ et } \mathbf{U}.
 \end{array} \right.$$

## Remarques

- Ce type de problème est *courant en pratique* : gestion de production, gestion multi-portefeuilles.
- On pourrait de la même manière envisager un *couplage dans la fonction objectif*.

## Remarques

On sait qu'à l'optimum la commande est  $\mathbf{X}_t$ -mesurable : le cadre naturel est celui de la programmation dynamique.

### Malédiction de la dimension

L'équation de la programmation dynamique

- ne peut être résolue "frontalement" pour une dimension d'état  $> 4$  environ.
- ne se décompose pas :
  - *la stratégie d'une unité dépend en général de l'état de toutes les unités* :  $\mathbf{U}_t^i \preceq \mathbf{X}_t^1, \dots, \mathbf{X}_t^n$  ;
  - le choix de *feedbacks locaux* :  $\mathbf{U}_t^i \preceq \mathbf{X}_t^i$  est généralement très *sous-optimal*.

# Dualité (1)

On a toujours le droit d'écrire que le problème original est équivalent à :

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{U}} \max_{\lambda} \quad & \mathbb{E} \left( \sum_{t=0}^{T-1} \left( \sum_{i=1}^n L_t^i(\mathbf{x}_t^i, \mathbf{U}_t^i, \mathbf{w}_{t+1}) - \lambda_t \sum_{i=1}^n \mathbf{U}_t^i \right) + \sum_{i=1}^n K^i(\mathbf{x}_T^i) \right), \\ \text{s.c.} \quad & \mathbf{x}_{t+1}^i = f_t^i(\mathbf{x}_t^i, \mathbf{U}_t^i, \mathbf{w}_{t+1}), \\ & \mathbf{U}_t \text{ est } \mathcal{F}_t\text{-mesurable.} \end{aligned}$$

$\Leftrightarrow$  Le multiplicateur  $\lambda_t$  est une variable aléatoire  $\mathcal{F}_t$ -mesurable.

## Dualité (2)

Si le lagrangien a un point-selle, on peut *échanger minimisation et maximisation*:

$$\begin{aligned} \max_{\lambda} \min_{\mathbf{x}, \mathbf{U}} \quad & \mathbb{E} \left( \sum_{t=0}^{T-1} \left( \sum_{i=1}^n L_t^i(\mathbf{x}_t^i, \mathbf{u}_t^i, \mathbf{w}_{t+1}) - \lambda_t \sum_{i=1}^n \mathbf{u}_t^i \right) + \sum_{i=1}^n K^i(\mathbf{x}_T^i) \right), \\ \text{s.c.} \quad & \mathbf{x}_{t+1}^i = f_t^i(\mathbf{x}_t^i, \mathbf{u}_t^i, \mathbf{w}_{t+1}), \\ & \mathbf{U}_t \text{ est } \mathcal{F}_t\text{-mesurable.} \end{aligned}$$

## Dualité (3)

À  $\lambda$  fixé, le problème de minimisation est décomposable:

$$\begin{aligned} \min_{\mathbf{x}^i, \mathbf{u}^i} \quad & \mathbb{E} \left( \sum_{t=0}^{T-1} (L_t^i(\mathbf{x}_t^i, \mathbf{u}_t^i, \mathbf{w}_{t+1}) - \lambda_t \mathbf{u}_t^i) + K^i(\mathbf{x}_T^i) \right), \\ \text{s.c.} \quad & \mathbf{x}_{t+1}^i = f_t^i(\mathbf{x}_t^i, \mathbf{u}_t^i, \mathbf{w}_{t+1}), \\ & \mathbf{u}_t^i \text{ est } \mathcal{F}_t\text{-mesurable.} \end{aligned}$$

### Problème

On ne connaît rien sur les corrélations en temps de  $\lambda$ .

⇒ *On ne sait pas résoudre les sous-problèmes par programmation dynamique sur l'état local (ou sur un état "augmenté" de petite taille).*

# Dual Approximate Dynamic Programming (1)

- L'heuristique consiste à *restreindre l'espace dual afin d'être capable de résoudre les sous-problèmes* par programmation dynamique sur un état de petite taille.

## Hypothèse

Nous supposons que les multiplicateurs ont une certaine dynamique *indépendante de la commande*.

- Cette hypothèse a été justifiée sur un problème simple par [Strugarek, 2006].

## Dual Approximate Dynamic Programming (2)

Ainsi, on est amené à résoudre :

Le  $i$ -ème sous-problème

$$\begin{aligned} \min_{\mathbf{x}^i, \mathbf{U}^i} \quad & \mathbb{E} \left( \sum_{t=0}^{T-1} (L_t^i(\mathbf{x}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}) - \lambda_t \mathbf{U}_t^i) + K^i(\mathbf{x}_T^i) \right), \\ \text{s.c.} \quad & \mathbf{x}_{t+1}^i = f_t^i(\mathbf{x}_t^i, \mathbf{U}_t^i, \mathbf{W}_{t+1}), \\ & \lambda_{t+1} = \alpha_t \lambda_t + \beta_t \mathbf{W}_{t+1} + \gamma_t, \\ & \mathbf{U}_t^i \text{ est } \mathcal{F}_t\text{-mesurable.} \end{aligned}$$

La coordination revient à *mettre à jour les paramètres*  $\alpha$ ,  $\beta$ ,  $\gamma$  de la dynamique.



## Justification de la dynamique des multiplicateurs

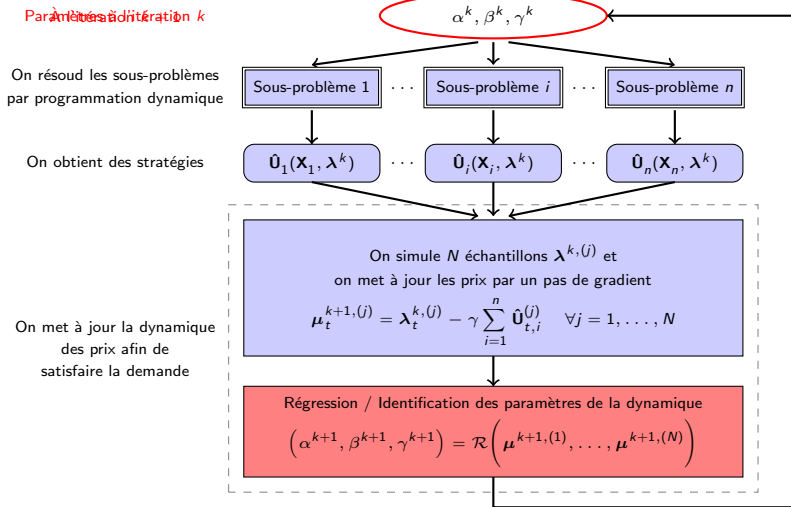
[Strugarek, 2006] montre, sur un problème très particulier, que les *prix suivent une dynamique en marche avant* :

$$\lambda_{t+1} = \lambda_t + \gamma \left[ (1 + \alpha) \mathbf{D}_{t+1} - \mathbf{D}_t - \alpha \mathbb{E}(\mathbf{D}_{t+1}) - \alpha (\mathbf{w}_{t+1} - \mathbb{E}(\mathbf{w}_{t+1})) \right],$$

avec  $\mathbf{D}_t$  une partie du bruit.

↔ Si on ajoute  $\mathbf{D}_t$  et  $\lambda_t$  dans l'état, les sous-problèmes peuvent alors être résolus par *programmation dynamique en dimension 3*.

# L'algorithme [Barty et al., 2009]



## Remarques

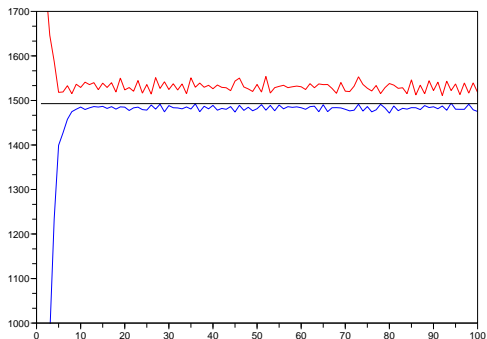
- Contrairement au cas déterministe, on doit ici ajouter cette *étape de régression*.

### Parallèle

- On se sert des variables duales (multiplicateurs) pour "*résumer*" l'état global du système.
  - Une approche classique dans l'industrie, la *décomposition par agrégation* [Turgeon, 1980], consiste à résumer l'état global du système à l'aide de variables primales.
- ↪ Il semble plus évident de mettre en œuvre DADP sur des systèmes hétérogènes.

# Convergence des coûts primaux et duaux

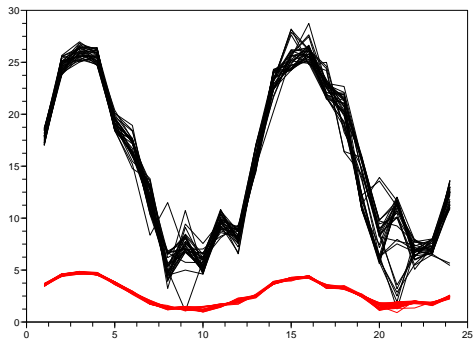
Sur un problème convexe avec une fonction objectif fortement convexe :



- Coûts primaux (en rouge) et duaux (en bleu) au cours des itérations.

# Convergence du multiplicateur

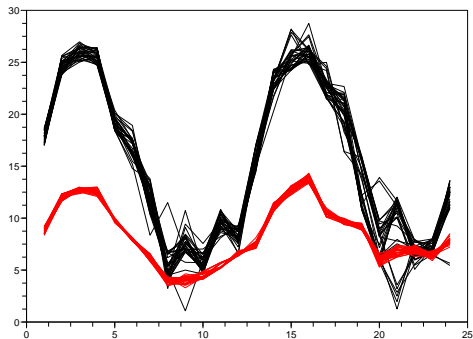
## Itération 1



- Les prix optimaux (en noir) sont obtenus par programmation dynamique.
- Les instabilités dans le bas de la courbe sont numériques.

# Convergence du multiplicateur

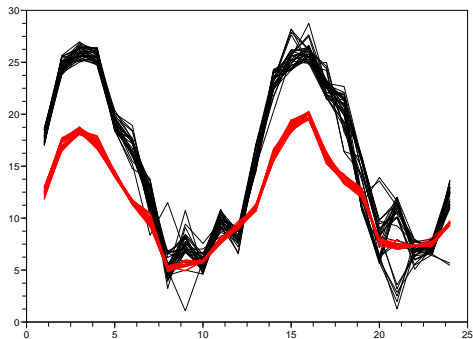
Itération 10



- Les prix optimaux (en noir) sont obtenus par programmation dynamique.
- Les instabilités dans le bas de la courbe sont numériques.

# Convergence du multiplicateur

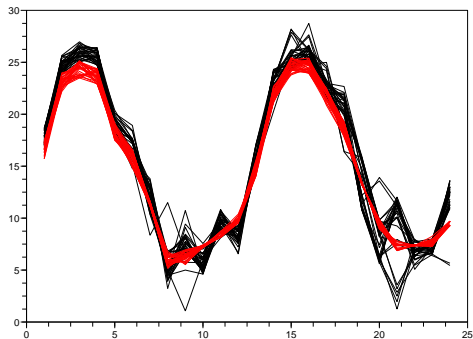
Itération 20



- Les prix optimaux (en noir) sont obtenus par programmation dynamique.
- Les instabilités dans le bas de la courbe sont numériques.

# Convergence du multiplicateur

Itération 50

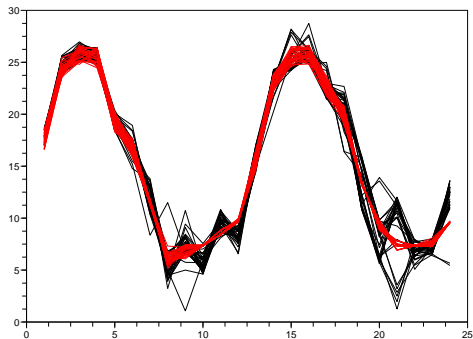


- Les prix optimaux (en noir) sont obtenus par programmation dynamique.
- Les instabilités dans le bas de la courbe sont numériques.



# Convergence du multiplicateur

Itération 90



- Les prix optimaux (en noir) sont obtenus par programmation dynamique.
- Les instabilités dans le bas de la courbe sont numériques.

## Conclusion

- Les méthodes classiques présentées atteignent rapidement leurs *limites lorsque la taille du problème augmente* (horizon de temps ou dimension d'état).
- L'alternative apportée par les *méthodes particulières* a l'avantage
  - de *calculer des stratégies raisonnables* ;
  - d'être une *technique variationnelle* (qui se mariera bien à une technique de décomposition lagrangienne) ;
  - de se concentrer dans les *régions "intéressantes" de l'espace d'état* (adaptée au cas où la loi de l'état optimal a un petit support comparé à l'ensemble de l'espace admissible).

mais ne semble pas échapper complètement à la malédiction de la dimension.

## Conclusion (suite)

- D'où l'intérêt des *méthodes de décomposition*, qui offrent des *résultats préliminaires prometteurs*. Des pistes d'améliorations comprennent :
  - S'affranchir de la restriction de l'espace dual (dynamique des multiplicateurs).
  - Marier décomposition et méthodes particulières.

# Références I



Barty, K., Carpentier, P., and Girardeau, P. (2009).  
Decomposition of large-scale stochastic optimal control problems.  
*arXiv:0903.1148v1*.



Dallagi, A. (2007).  
*Méthodes particulières en commande optimale stochastique*.  
PhD dissertation, Université Paris I Panthéon-Sorbonne.



Strugarek, C. (2006).  
*Approches variationnelles et autres contributions en optimisation stochastique*.  
PhD dissertation, École Nationale des Ponts et Chaussées.



Turgeon, A. (1980).  
Optimal operation of multi-reservoir power systems with stochastic inflows.  
*Water Resources Research*, 16(2):275–283.

## Construction d'un arbre de scénarios

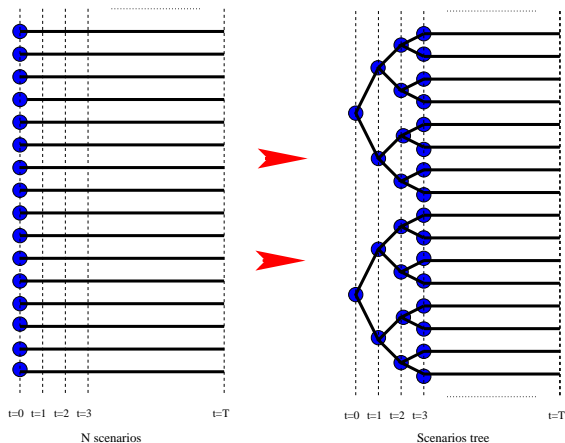


Figure: Construction d'un arbre de scénarios. ◀

# Conditions d'optimalité équivalentes

On considère le problème :

$$\min_{\mathbf{U} \in \mathcal{U}^{\text{ad}}} J(\mathbf{U}),$$

où  $\mathcal{U}^{\text{ad}}$  est un convexe fermé d'un espace de Hilbert  $\mathcal{U}$ .

Soit  $\mathbf{U}^\#$  une solution de ce problème.

Les **conditions nécessaires d'optimalité** suivantes :

$$\langle J'^{\top}(\mathbf{U}^\#), \mathbf{U} - \mathbf{U}^\# \rangle \geq 0 \quad \forall \mathbf{U} \in \mathcal{U}^{\text{ad}},$$

$$\mathbf{U}^\# = \text{proj}_{\mathcal{U}^{\text{ad}}} \left( \mathbf{U}^\# - \epsilon J'^{\top}(\mathbf{U}^\#) \right) \quad \forall \epsilon > 0.$$

sont **équivalentes**. 

## Espérance conditionnelle et variables indépendantes

Soit  $\mathbf{X}$  et  $\mathbf{Y}$  deux variables aléatoires **indépendantes**.

Le calcul de l'**espérance conditionnelle** suivante :

$$\mathbb{E}(f(\mathbf{X}, \mathbf{Y}) \mid \mathbf{X})$$

se réduit en fait à un simple **calcul d'espérance**. Posant :

$$F(x) = \mathbb{E}(f(x, \mathbf{Y})),$$

on a l'égalité :

$$\mathbb{E}(f(\mathbf{X}, \mathbf{Y}) \mid \mathbf{X}) = F(\mathbf{X}).$$

