

# Méthode des états intermédiaires pour le contrôle optimal.

Julien Salomon

CEREMADE, UMR CNRS 7534, Université Paris-Dauphine, France.

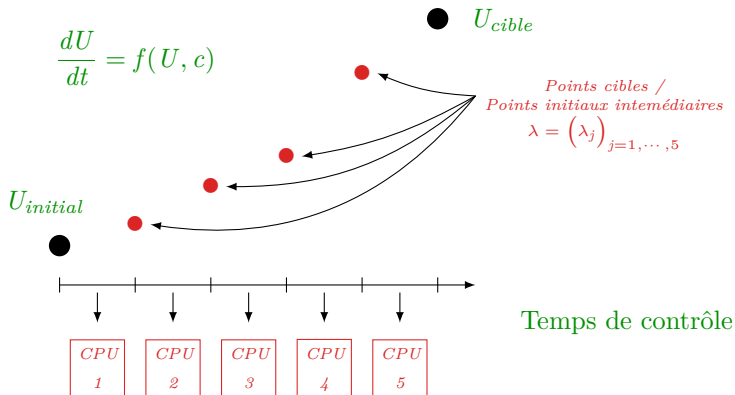
*Coll. Y. Maday, K. Riahi, D. Sugny, G. Turinici*

- Enjeu : accélérer le calcul en utilisant une décomposition en temps dans un problème de contrôle optimal.
- Problématique liée au Multishooting.

Méthode	Solver interne	Solver global
Multishooting	ordre 2	-
Nous	ordre 1	ordre 1

*Ordre de convergence des méthodes.*

# Points intermédiaires



Accélération : parallélisation du calcul / problèmes plus simples car plus petits.

- Comment choisir les états intermédiaires de telle sorte que les solutions concaténées donne la solution ?
- Comment éviter des "propagations" sur  $[0, T]$  lors du calcul des états intermédiaires ?

# Plan de l'exposé

- ① Cas de la RMN et du contrôle quantique.
- ② Cas du contrôle de l'équation de la chaleur.
- ③ Conclusion et perspectives.

Équation représentative : équation de Bloch

$$i\partial_t U(t) = [H_0 + \sum_{\ell}^L \Omega_{\ell}(t)H_{\ell}]U(t)$$
$$U(t=0) = U_0.$$

Problème de contrôle optimal :

trouver  $\Omega^*(t) = (\Omega_1^*(t), \dots, \Omega_L^*(t))$ , qui résoud

$$\Omega^* = \operatorname{argmax} J(\Omega)$$

avec

$$J(\Omega) = \operatorname{Re}\langle U(T), U_{\text{target}} \rangle.$$

Quelques caractéristiques de notre approche :

- Alternance de la mise à jour des contrôles partiels  $(\Omega_{I_j})_{j=1, \dots, J}$  et celle des états intermédiaires  $(\lambda_j)_{j=0, \dots, J}$ .
- On suppose disposer d'une méthode de résolution des problèmes partiels.

La formule de mise à jour des états est :

$$\lambda^\Omega(t) = \frac{T-t}{T}U(t) + \frac{t}{T}P(t),$$



On pose

$$J_j^\lambda(\Omega_j) = \operatorname{Re}\langle U_j(T_{j+1}), \lambda_{j+1} \rangle.$$

**Algorithme :** Étant donnés  $\Omega^k, \lambda^k$  à l'étape  $k$  :

① résoudre en parallèle sur  $[T_j, T_{j+1}]$

$$\max_{\Omega_j} J_j^\lambda(\Omega_j) \rightarrow \Omega_j^{k+1},$$

② définir  $\Omega^{k+1}$  comme la concaténation de  $\Omega_j^{k+1}$ ,

③ définir  $\lambda^{k+1}$  par la formule d'interpolation.

Propriétés : en posant

$$J_{\parallel}(\Omega, \lambda) = \sum_{j=0}^{N-1} \beta_j J_j^{\lambda}(\Omega_j).$$

**Théorème 1.** Étant donné  $\Omega$

$$J_{\parallel}(\Omega, \lambda^{\Omega}) = \min_{\lambda} J_{\parallel}(\Omega, \lambda) = J(\Omega).$$

**Théorème 2.** Supposons :

- 1 la méthode d'optimisation converge,
- 2  $\lambda^k = \lambda^{\Omega^k}$ , then  $\lim \Omega^k = \Omega^{\infty}$ ,  $\nabla J(\Omega^{\infty}) = 0$ .

# Nouvelle interprétation de la méthode

Avec la définition des états intermédiaires, on a en fait :

$$\nabla J(\Omega)_{|[T_j, T_{j+1}]} = \frac{T_{j+1} - T_j}{T} \nabla J_j^\lambda(\Omega_{|[T_j, T_{j+1}]}).$$

POUR TOUT  $\Omega$ !

d'où la nouvelle interprétation :

la méthode donne une décomposition du calcul du gradient.

# Efficacité de la méthode ?

Le gradient peut être calculé en parallèle, mais pour calculer  $\lambda(t)$ ,  $U(t)$  et  $P(t)$  semblent nécessiter une résolution sur  $[0, T]$  ?

NON!!!  $\rightarrow$  seulement  $U(t_j)$  et  $P(t_j)$  sont nécessaires.

Pour les systèmes de petite dimension, les propagateurs  $t_j \rightarrow t_{j+1}$  peuvent être calculés en parallèle, au moment de la phase d'optimisation des contrôles partiels !

# Efficacité de la méthode ?

Le gradient peut être calculé en parallèle, mais pour calculer  $\lambda(t)$ ,  $U(t)$  et  $P(t)$  semblent nécessiter une résolution sur  $[0, T]$  ?

NON!!! → seulement  $U(t_j)$  et  $P(t_j)$  sont nécessaires.

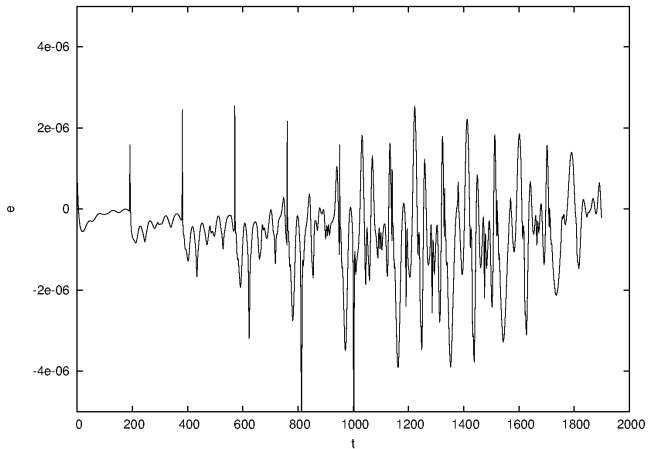
Pour les systèmes de petite dimension, les propagateurs  $t_j \rightarrow t_{j+1}$  peuvent être calculés en parallèle, au moment de la phase d'optimisation des contrôles partiels !

Tests sur le contrôle de l'équation de Schrödinger :

$$i\partial_t\psi = [H_0 - \mu(x)\varepsilon(t)]\psi$$

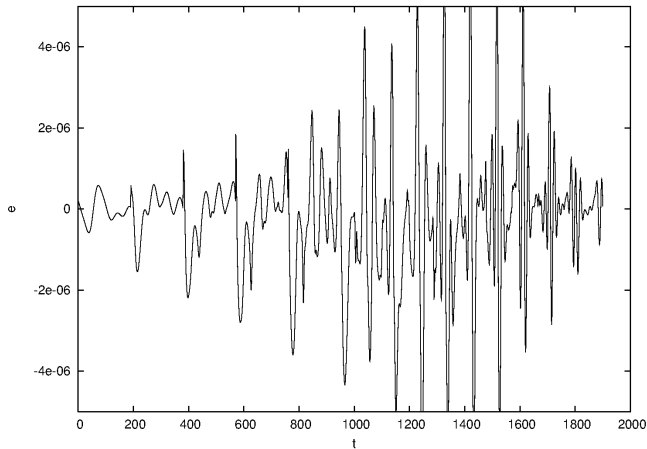
# Résultats numériques : Schrödinger

Après une itération :



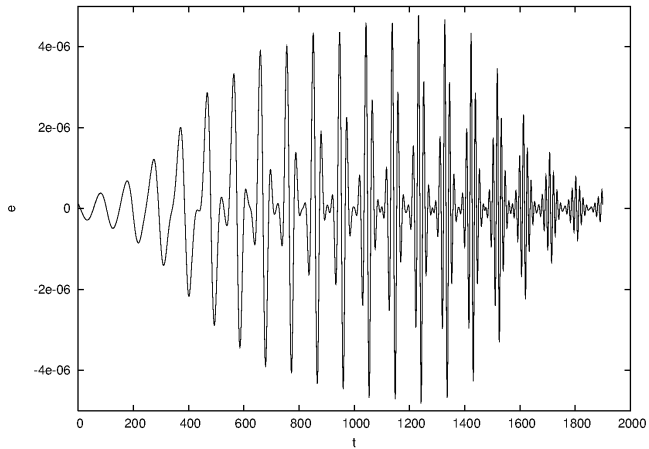
# Résultats numériques : Schrödinger

Après 10 :

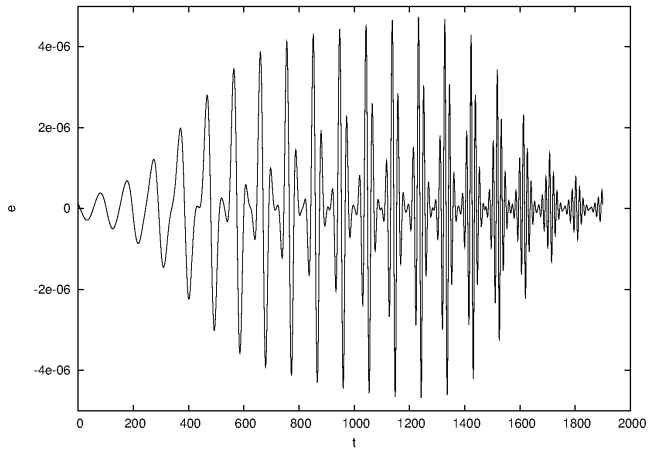




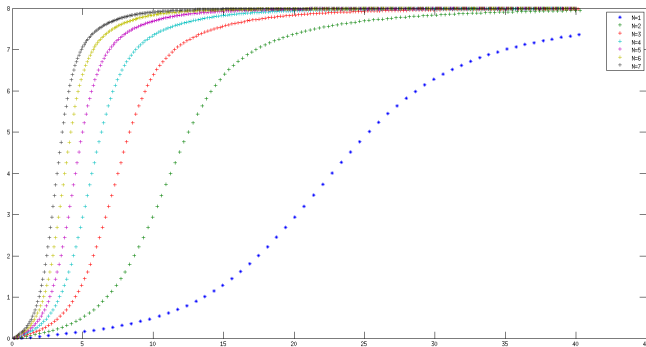
# Résultats numériques : Schrödinger



# Résultats numériques : Schrödinger



# Résultats numériques : RMN



# Cas du contrôle de l'équation de la chaleur

On considère maintenant la fonctionnelle :

$$J(c) = \frac{1}{2} \|u(T) - u_{target}\|_2^2 + \frac{\alpha}{2} \int_0^T \|c(t)\|^2 dt,$$

associée à  $u$ , contrôlé par  $c$  :

$$\partial_t u = \Delta u + B(c).$$

Système d'optimalité :

$$\begin{cases} \partial_t p & = & -\Delta p \\ p(T) & = & u(T) - u_{target}. \end{cases}$$

$$\nabla J(c) = \alpha c + \langle p, \partial_c B(c) \rangle = 0.$$

On définit cette fois-ci deux trajectoires :

$$\Lambda = (\underline{\lambda}(t), \bar{\lambda}(t)),$$

et les sous-problèmes :

$$J_j^\Lambda(c_j) = \frac{1}{2} \|u_j(T_{j+1}) - \bar{\lambda}(T_{j+1})\|_2^2 + \frac{\alpha}{2} \int_{T_j}^{T_{j+1}} \|c_j(t)\|^2 dt,$$

où l'état  $u_j$  défini sur  $[T_j, T_{j+1}]$  est contrôlé par  $c$  selon :

$$\partial_t u_j = \Delta u_j + B(c_j),$$

avec :

$$u_j(T_j) = \underline{\lambda}(T_j).$$

Les deux trajectoires sont définies par

$$\bar{\lambda}(t) = u(t) - p(t),$$

de telle sorte que  $\bar{\lambda}(T) = u_{target}$ , et

$$\underline{\lambda}(t) = u(t).$$

**Théorème 1 :** Soit  $c^*$  et  $c_j^*$  les solutions des problèmes locaux et globaux.  
Alors :

$$c_j^* = c_{|[T_j, T_{j+1}]}$$

**Théorème 2 :**

- 1 Si on effectue la mise à jour  $c_j^k \rightarrow c_j^{k+1}$  par un pas de la méthode du gradient,
- 2  $\bar{\lambda}^k = u^k - p^k$ ,  $\underline{\lambda}^k = u^k$ .

Alors  $\lim c^k = c^\infty$ ,  $\nabla J(c^\infty) = 0$ .



# Cas du contrôle de l'équation de la chaleur

Revenons au système d'optimalité :

$$\begin{cases} \partial_t u &= \Delta u + Bc \\ u(0) &= u_0. \end{cases}$$

$$\begin{cases} \partial_t p &= -\Delta p \\ p(T) &= u(T) - u_{target}. \end{cases}$$

$$\alpha c + B^* p = 0.$$

# Cas du contrôle de l'équation de la chaleur

Après discrétisation totale :

$$\underbrace{\begin{bmatrix} \mathbf{E} & \mathbf{N} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} & \mathbf{N}^T \\ \mathbf{K} & \mathbf{0} & \mathbf{E}^T \end{bmatrix}}_M \underbrace{\begin{bmatrix} \mathbf{U} \\ \mathbf{C} \\ \mathbf{P} \end{bmatrix}}_X = \underbrace{\begin{bmatrix} \mathbf{D}_0 \\ \mathbf{0} \\ \mathbf{K}\Lambda \end{bmatrix}}_b .$$

$$\rightarrow MX = b$$

# Cas du contrôle de l'équation de la chaleur

En réordonnant :

$$\begin{bmatrix} \mathbf{E}_1 & \mathbf{N}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_1 & \mathbf{N}_1^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{E}_1^T & \mathbf{0} & \mathbf{0} & \mathbf{E}_{21}^T \\ \mathbf{E}_{21} & \mathbf{0} & \mathbf{0} & \mathbf{E}_2 & \mathbf{N}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_2 & \mathbf{N}_2^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_2 & \mathbf{0} & \mathbf{E}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{C}_1 \\ \mathbf{P}_1 \\ \mathbf{U}_2 \\ \mathbf{C}_2 \\ \mathbf{P}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{K}_2 \Lambda \end{bmatrix}$$

# Cas du contrôle de l'équation de la chaleur

$$\left[ \begin{array}{ccc|ccc} \mathbf{E}_1 & \mathbf{N}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_1 & \mathbf{N}_1^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_1 & \mathbf{0} & \mathbf{E}_1^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_2 & \mathbf{N}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_2 & \mathbf{N}_2^T \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_2 & \mathbf{0} & \mathbf{E}_2^T \end{array} \right] \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{C}_1 \\ \mathbf{P}_1 \\ \mathbf{U}_2 \\ \mathbf{C}_2 \\ \mathbf{P}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{0} \\ \mathbf{K}_1 \mathbf{U}_1 - \mathbf{E}_{21}^T \mathbf{P}_2 \\ -\mathbf{E}_{21} \mathbf{U}_1 \\ \mathbf{0} \\ \mathbf{K}_2 \Lambda_{2,2} \end{bmatrix}$$

$$MX = b \rightarrow DX = -(L + U)X + b.$$

# Cas du contrôle de l'équation de la chaleur

$$\left[ \begin{array}{ccc|ccc} \mathbf{E}_1 & \mathbf{N}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{G}_1 & \mathbf{N}_1^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{K}_1 & \mathbf{0} & \mathbf{E}_1^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_2 & \mathbf{N}_2 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_2 & \mathbf{N}_2^T \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_2 & \mathbf{0} & \mathbf{E}_2^T \end{array} \right] \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{C}_1 \\ \mathbf{P}_1 \\ \mathbf{U}_2 \\ \mathbf{C}_2 \\ \mathbf{P}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{0} \\ \mathbf{K}_1 \Lambda_{1,1} \\ \mathbf{d}_1 \\ \mathbf{0} \\ \mathbf{K}_2 \Lambda_{2,2} \end{bmatrix}$$

# Cas du contrôle de l'équation de la chaleur

L'algorithme se ré-écrit :  $X^k$

- 1  $\tilde{X}^{k+1} = D^{-1}(L + U)X^k$ , (en parallèle)
- 2 Extraire  $C^{k+1}$  de  $\tilde{X}^{k+1}$ ,
- 3 Calculer  $X^{k+1}$  qui correspond à  $C^{k+1}$ . (séquentielle)

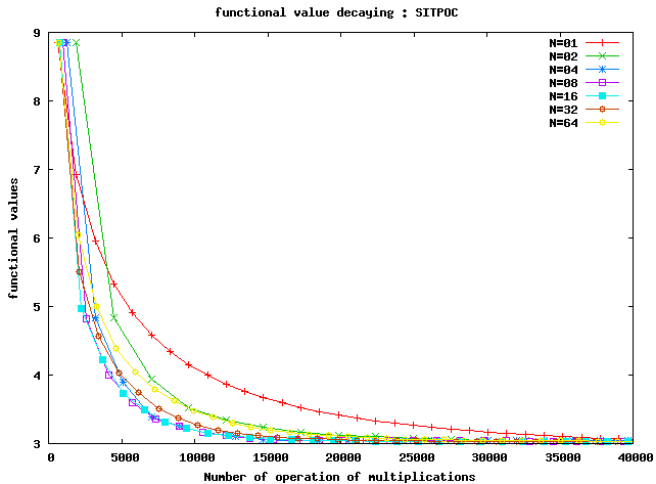
Il reste une résolution sur  $[0, T]$ .

# Cas du contrôle de l'équation de la chaleur

Problème : le problème est trop gros pour transmettre les propagateurs partiels.

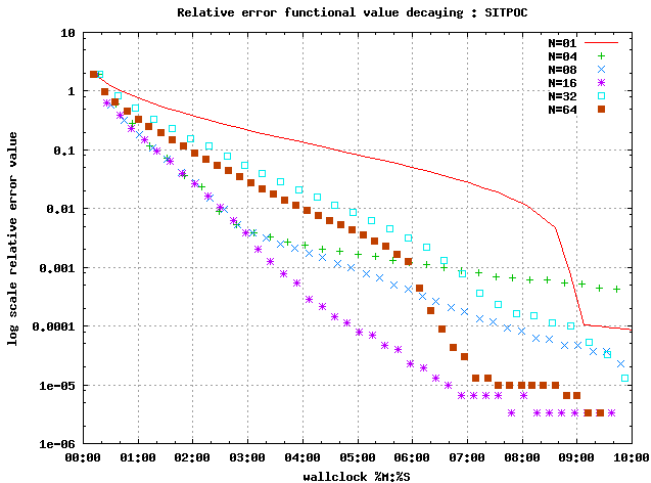
Solution : on utilise le schéma "pararéel", qui gagne du temps en n'effectuant les propagations sur  $[0, T]$  de manière grossière.

# Résultats numériques

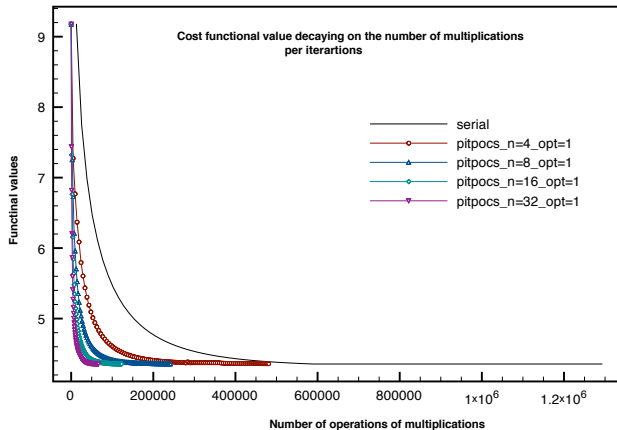




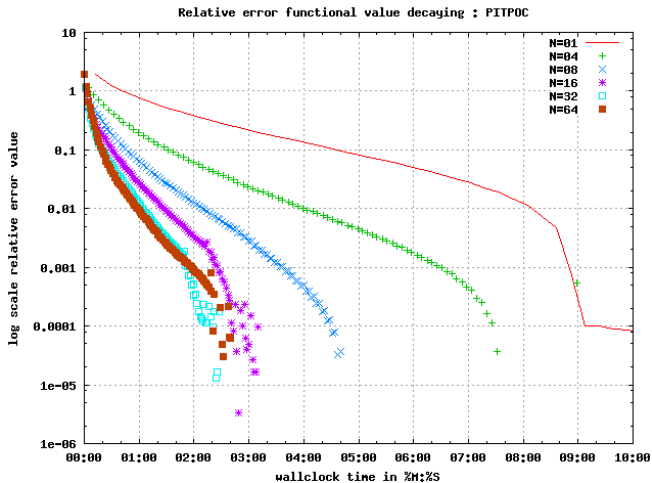
# Résultats numériques



Combinaison avec l'algorithme parallèle.



# Résultats numériques



## Ce qu'on a obtenu :

- Cadre général de parallélisation en temps pour les algorithmes classiques.
- Efficacité "totale" sur des exemples en petites dimensions.

## Ce qu'on projette :

- Obtenir une version adaptée à la convergence quadratique.
- Utiliser des techniques d'apprentissage et de réduction pour accélérer la partie séquentielle dans le cas de grand systèmes.
- Traiter le cas de dynamiques non-linéaires.

Merci !