

Une méthode de réduction de variance pour les processus de diffusion

Antoine Lejay

en collaboration avec

Madalina Deaconu

Projet OMEGA

(INRIA Lorraine, Institut Élie Cartan, Nancy)

Journées thématique du GdR MOMAS

« Problèmes inverses et analyse d'incertitudes »

Université de Nice — octobre 2006

Méthodes de Monte Carlo et processus de diffusion

Simuler le déplacement d'un grand nombre de **particules** associées à un **opérateur de diffusion** de type

$$L = \underbrace{\frac{1}{2} \sum_{i,j=1}^d a_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}}_{\text{diffusion}} + \underbrace{\sum_{i=1}^d b_i \frac{\partial}{\partial x_i}}_{\text{transport}}$$

et en calculer la moyenne de certaines quantités (fonctions de la position, du temps d'atteinte, du temps passé dans certaines zones, ...)

Quelques exemples

- Résolution d'EDP elliptiques ou paraboliques
- Calcul du coefficient effectif de perméabilité
- Calcul du coefficient d'échange pour le modèle « double porosité »
- Évaluation d'événements rares (=événement de petite probabilité)
- Calcul de la première valeur propre
- ...

Avantages et inconvénients

- ☺ Le déplacement des particules est spécifié par les valeurs locales des coefficients de l'opérateur différentiel : utile pour les problèmes de grande taille (réservoir pétrolier, ...)
- ☺ Algorithmes simples à mettre en œuvre
- ☺ Pas de problème avec la dimension (augmentation linéaire, ne nécessite pas forcément de maillage, ...)
- ☺ Parallélisme implicite (les particules sont indépendantes)
- ☹ Il faut simuler beaucoup de particules
- ☹ Cela prend du temps
- ☹ Incertitude sur le résultat final

Méthodes de Monte Carlo : vitesse de convergence ?

Dans beaucoup de cas, il s'agit d'évaluer des intégrales de type

$$J = \int f(x)K(x) dx = \mathbb{E} [f(X)], \quad K(x) \text{ densité de la variable aléatoire (v.a.) } X$$

Soient X^1, \dots, X^N des **réalisations indépendantes** de X

Si $J_N = \frac{1}{N} \sum_{i=1}^N f(X^i)$, alors J_N approche J

Si X^1, \dots, X^N sont des **copies indépendantes** de X ,

$$J_N = \frac{1}{N} \sum_{i=1}^N f(X^i) \xrightarrow[N \rightarrow \infty]{} J \text{ presque sûrement (Loi des grands nombres)}$$

Peut on donner la **vitesse de convergence** de la v.a. J_N vers J ?

$$\sqrt{N}(J_N - J) = \frac{1}{\sqrt{N}} \sum_{i=1}^N (f(X^i) - \mathbb{E} [f(X)]) \xrightarrow[N \rightarrow \infty]{} \mathcal{N}(0, \text{Var}(f(X))) \text{ (TCL)}$$

Asymptotiquement, si $Z \sim \mathcal{N}(0, \text{Var}(f(X)))$, alors

$$J_N \approx J + \frac{1}{\sqrt{N}}Z \text{ et } \mathbb{P} [Z \in [-3 \text{Var}(f(X)), 3 \text{Var}(f(X))]] = 99,9 \%$$

Méthodes de Monte Carlo : vitesse de convergence ?

On a donc deux informations sur la vitesse : la **précision** est contrôlée par

$$\text{Var}(f(X)) = \mathbb{E} [f(X)^2] - \mathbb{E} [f(X)]^2 \text{ et } N$$

Exemple : On cherche à évaluer la probabilité p que les particules atteignent une certaine zone A sur le bord du domaine

$$X \text{ position où la particule atteint le bord du domaine ; } f(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

$\text{Var}(f(X)) = p(1 - p) \implies$ pour une précision de 10^{-k} (à 99,9 % de chance),

$$N \sim \frac{9 \cdot 10^{2k}}{p^2(1 - p)^2}$$

Le coût devient exorbitant si p est proche de 0 (ou proche de 1)

Nil desperandum !

Il existe des techniques dites de réduction de variance pour diminuer la variance (réduction de N ou augmentation de la précision) et/ou estimer des événements rares

Beaucoup de ces techniques reposent sur le **tirage d'importance**

Au lieu de simuler la v.a. X , on simule une v.a. Y de densité L

$$J = \mathbb{E} [f(X)] = \int f(x)K(x) dx = \int \frac{K(x)}{L(x)} f(x)L(x) dx = \mathbb{E} \left[\frac{K(Y)}{L(Y)} f(Y) \right]$$

$$\text{Var}(K(Y)f(Y)/L(Y)) = \mathbb{E} \left[\frac{K^2(Y)}{L^2(Y)} f^2(Y) \right] - J^2 = \mathbb{E} \left[\frac{K(X)}{L(X)} f^2(X) \right] - J^2$$

Le but est de **bien choisir** la densité L

Exemple (suite)

$$\text{Si } f = \begin{cases} 1 & \text{si } X \in A \subset [0, 1], \\ 0 & \text{si } X \in \bar{A} \end{cases}$$

K, L densité par rapport à la mesure de Lebesgue

$$K(x) = \begin{cases} p/|A| & \text{si } x \in A \\ (1-p)/|\bar{A}| & \text{si } x \in \bar{A} \end{cases} \implies \mathbb{P}[X \in A] = \int_A K(x) dx = p$$

$$L(y) = \begin{cases} q/|A| & \text{si } x \in A \\ (1-q)/|\bar{A}| & \text{si } x \in \bar{A} \end{cases} \implies \mathbb{P}[Y \in A] = \int_A L(x) dx = q$$

$$\mathbb{E} \left[\frac{K(X)}{L(X)} f^2(X) \right] = \mathbb{E} \left[\frac{p}{q}; X \in A \right] = \frac{p^2}{q}$$

$$\implies \text{Var}(K(Y)f(Y)/L(Y)) = \frac{p^2}{q} - p^2 = p^2 \left(\frac{1}{q} - 1 \right)$$

alors que

$$\text{Var}(f(X)) = p(1-p)$$

Il y a **réduction de variance** dès que $q \geq p$

Mais ce choix dépend de f, K, L, \dots

Résolution d'EDP

On s'intéresse à un processus de diffusion dans un domaine $[0, T] \times D \subset \mathbb{R}_+ \times \mathbb{R}^d$ et plus précisément à la loi de (τ, X_τ) , où τ est le **premier temps de sortie** des particules de $[0, T] \times D$ (ou de D)

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + Lu(t, x) = 0 \text{ sur } [0, T] \times \mathbb{R}^d \\ u(T, x) = f(x) \end{cases}$$

$$\implies u(t, x) = \mathbb{E}_x [f(X_{T-t})] = \int_{\mathbb{R}^d} \underbrace{p(T-t, x, y)}_{\text{densité} \equiv \text{solution fondamentale}} f(y) dy$$

$$\begin{cases} Lu(x) = 0, \\ u(x) = \varphi(x) \text{ sur } \partial D \end{cases} \implies u(x) = \mathbb{E}_x [\varphi(X_\tau)] = \int_{\partial D} \underbrace{K(x, y)}_{\text{noyau de Poisson}} f(y) d\sigma_y$$

Simuler X_t ou (τ, X_τ) revient donc à calculer numériquement la **solution fondamentale** ou le **noyau de poisson** de L

Une méthode de simulation par Monte Carlo : marche aléatoire sur les carrés/rectangles

☺ On sait très bien simuler un processus de diffusion dans tout l'espace (schéma d'Euler, ...)

☹ La prise en compte de condition aux bords (Dirichlet, Neumann) est un peu plus délicate

On s'intéresse ici au cas simple de $L = \frac{1}{2}\Delta$
(mais on peut considérer L plus général)

Le processus de diffusion X associé est le **mouvement brownien**

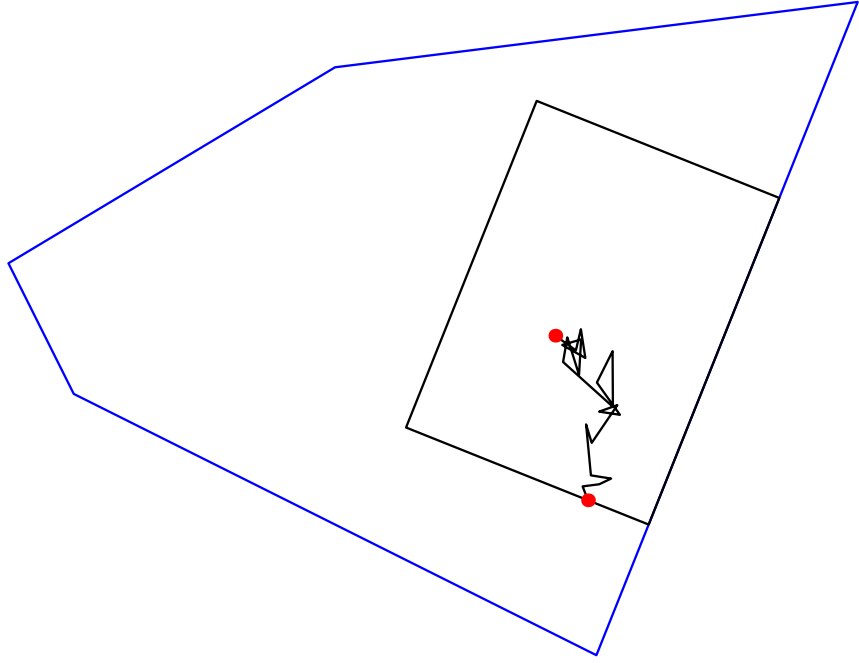
Dans tout l'espace, la loi de X_t à l'instant t est $\mathcal{N}(0, t)$

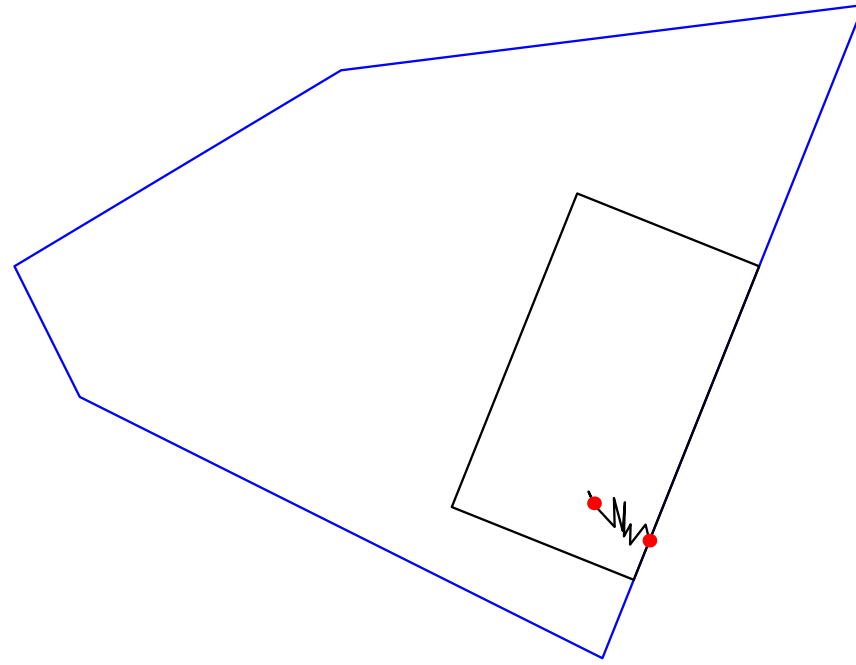
On sait par ailleurs simuler (τ, X_τ) pour une sphère (Müller), un carré centré (Milstein & Tretyakov), un rectangle (D&L, 2006)

L'idée est donc de simuler successivement les temps et positions de sortie pour la particule de ces formes simples

☺ l'on sait ou non si l'on a atteint le bord du domaine

☺ dans certain cas, on peut faire des simulations exactes





Réduction de variance

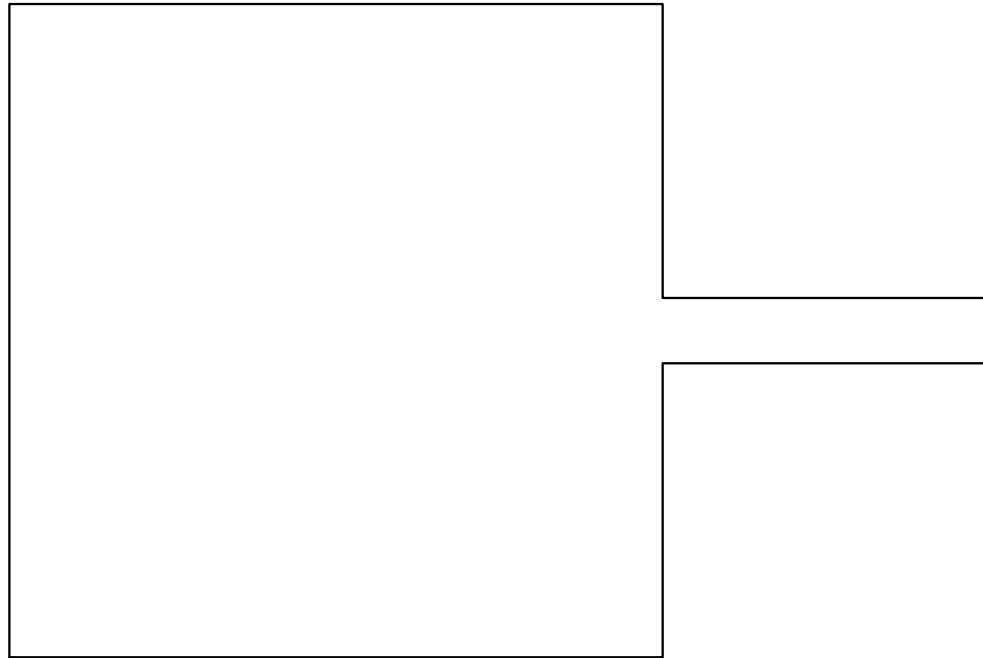
Au lieu de simuler exactement (θ, X_θ) (θ temps de sortie d'une forme simple), l'idée est de tirer (θ, X_θ) selon une **loi arbitraire**

On change la probabilité en fonction d'un **poids w** , et l'on multiplie les poids à chaque tirage

Dans le cas de la marche aléatoire sur les carrés, rectangles, les poids se calculent facilement (numériquement parlant)

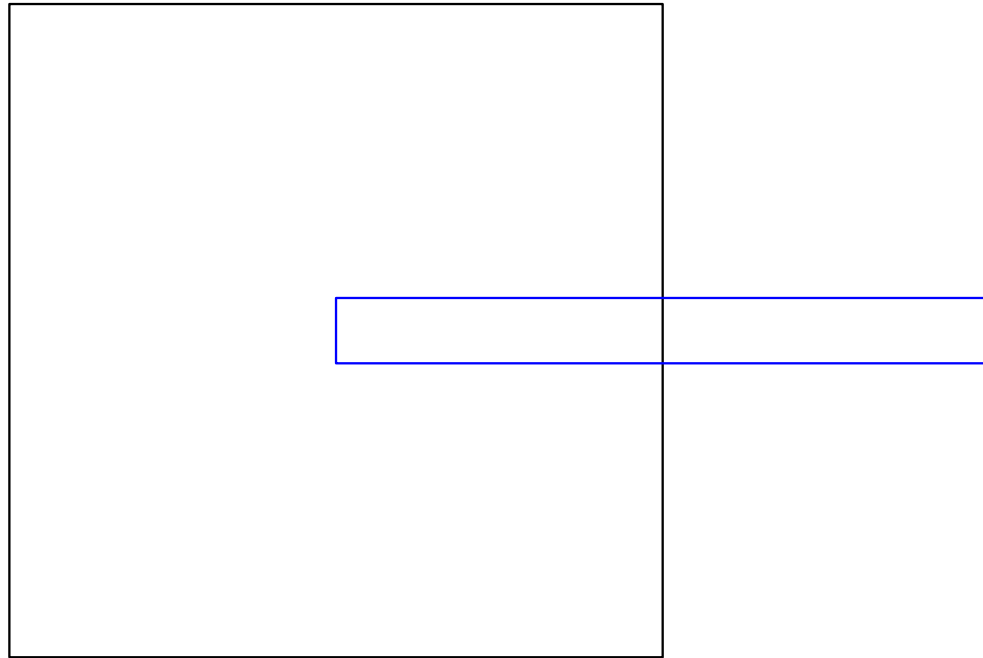
- ☺ On peut « forcer » le processus à aller là où il ne va pas en général
- ☹ Le produit des poids a tendance à exploser
- ☹ Il faut trouver les bonnes « stratégies »

Une application numérique : simulation d'événements rares



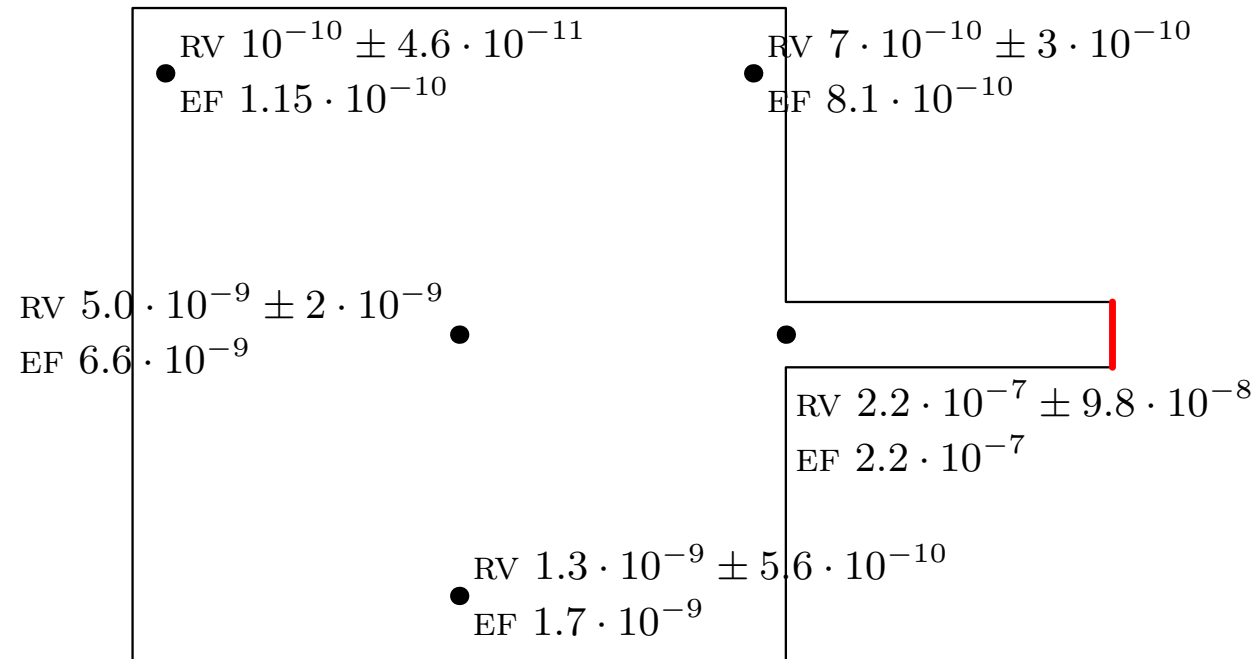
Qu'elle est la probabilité de sortir par le côté rouge ?

Une application numérique : simulation d'événements rares



On utilise deux rectangles et on change les proba. de sortie par chacun des côtés

Une application numérique : simulation d'événements rares



Le nombre de tirages est de 100 000 !

Réduction de variance : futurs travaux

Méthodes **automatiques** de réduction de variance : il s'agit d'un problème d'optimisation, parfois en grande dimension

Pour les diffusions, c'est une problématique en pleine effervescence, avec notamment l'utilisation d'**algorithmes stochastiques**

Qu'en est-il pour notre méthode ?