

Analyse de l'Algorithme "Adaptive Multilevel Splitting" pour la simulation d'événements rares

Charles-Edouard Bréhier

CERMICS (ENPC) & Inria

Collaboration avec T.Lelièvre, M.Rousset, M.Gazeau, L.Goudenège

Journées MAS Toulouse - 27/08/2014

Session Probabilités Numériques

Problématique: simulation de la **métastabilité**.

Particule dans un double-puits:

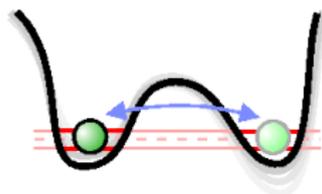


Figure: Potentiel: $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$.

Dynamique:

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon}dB_t.$$

Objectif: simuler les transitions (rares) quand $\epsilon \rightarrow 0$.

Introduction

Double-puits:

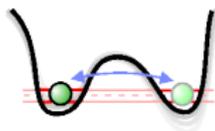


Figure: Potentiel: $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$.

Modèle généralisé: chaînes de N particules.

Limite $N \rightarrow +\infty$: équation d'Allen-Cahn

$$\frac{\partial u^\epsilon(t, x)}{\partial t} = \gamma \Delta u^\epsilon(t, x) - \nabla V(u^\epsilon(t, x)) + \sqrt{2\epsilon} \dot{W}(t, x),$$

avec conditions au bord de type Neumann sur $(-1, 1)$.

Etats stables ($\epsilon = 0$): $u(t, x) = -1$ et $u(t, x) = +1$.

Simulations Allen-Cahn (CEMRACS 2013)

$\updownarrow: u$ $\nwarrow: t$ $\nearrow: x$

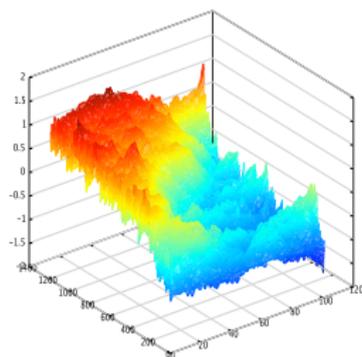
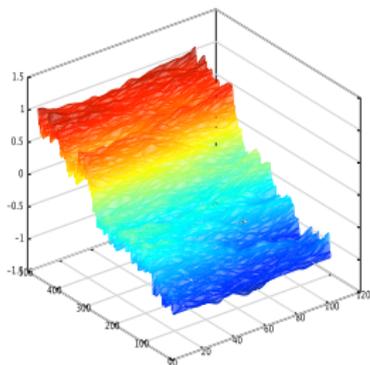


Figure: Transitions pour l'équation d'Allen-Cahn, $\gamma = 1$ et $\gamma = 0.1$

Analysis and simulation of rare events for SPDEs,
avec M. Gazeau, L. Goudenège, M. Rousset; Proceeding CEMRACS; 2013.

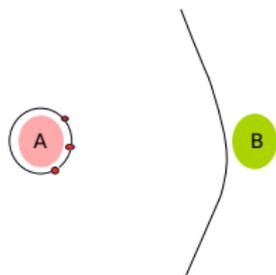
Cadre général

Cadre: $Y = (Y_t)_{t \geq 0}$ un processus de Markov, à trajectoires continues.

$A, B \subset E$ disjoints;

Hypothèse: temps d'atteinte $\tau_A, \tau_B < +\infty$ p.s.

Condition initiale: $y_0 \notin A, \notin B$.



Objectif: estimer $\mathbb{P}^{y_0}(\tau_B < \tau_A)$.

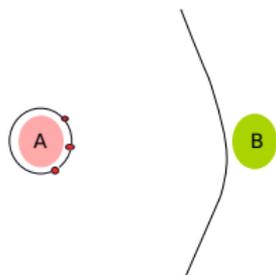
Cadre général

Cadre: $Y = (Y_t)_{t \geq 0}$ un processus de Markov, à trajectoires continues.

$A, B \subset E$ disjoints;

Hypothèse: temps d'atteinte $\tau_A, \tau_B < +\infty$ p.s.

Condition initiale: $y_0 \notin A, \notin B$.



Objectif: estimer $\mathbb{P}^{y_0}(\tau_B < \tau_A)$.

Calcul des **niveaux**: $\xi : E \rightarrow \mathbb{R}$ continue, **coordonnée de réaction**.

Hypothèse: $\xi(B) \subset (a, +\infty)$.

Exemple (Allen-Cahn): $\xi(u) = \int_0^1 u(x) dx, u \in L^1(0, 1)$.

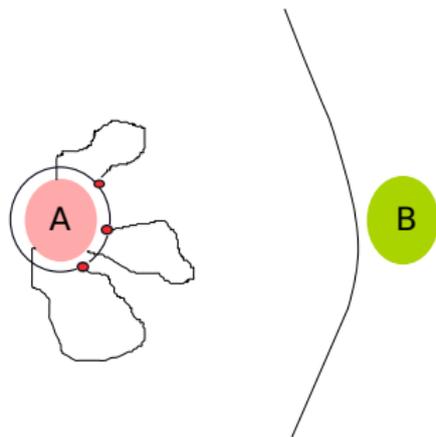
L'algorithme Adaptive Multilevel Splitting.

Paramètres:

n répliques

$k \in \{1, \dots, n-1\}$

niveau d'arrêt a .



Initialisation: Y_1, \dots, Y_n iid.

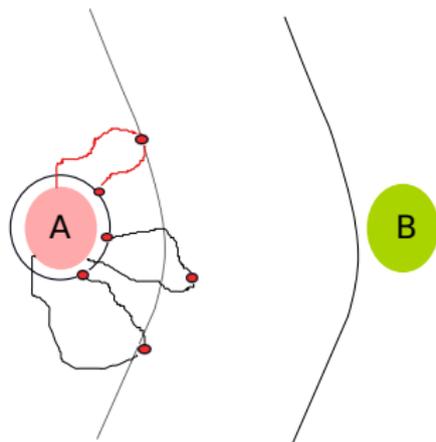
L'algorithme Adaptive Multilevel Splitting.

Paramètres:

n répliques

$k \in \{1, \dots, n-1\}$

niveau d'arrêt a .

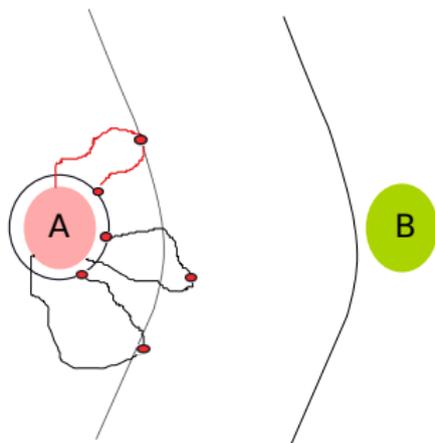


Initialisation: Y_1, \dots, Y_n iid.

Niveaux individuels: $X_i = \sup_{0 \leq t \leq \tau_A \wedge \tau_B} \xi(Y_i(t))$.

Niveau initial: $Z = X_{(k)}$, où $X_{(1)} \leq \dots \leq X_{(k)} \leq \dots \leq X_{(n)}$.

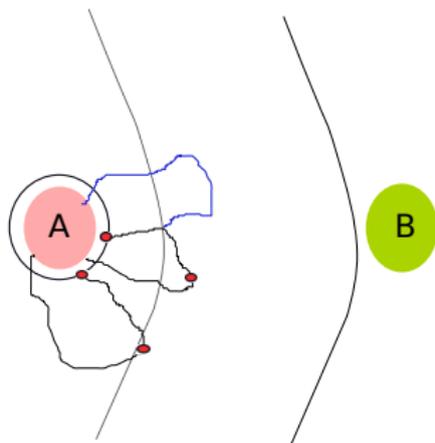
L'algorithme Adaptive Multilevel Splitting.



Itérations: tant que $Z < a$,

- ▶ on supprime $Y_{(1)}, \dots, Y_{(k)}$.
- ▶ pour les remplacer: on tire $n_{(1)}, \dots, n_{(k)}$ iid dans $\{(k+1), \dots, (n)\}$.
- ▶ $Y_{(i)}(t) = Y_{n_{(i)}}(t)$ si $t \leq T^{(i)} := \inf \left\{ s \geq 0; \xi(Y_{n_{(i)}}(s)) > Z \right\}$.
- ▶ $Y_{(i)}(t)$ pour $t \geq T^{(i)}$: trajectoire Markovienne issue de $Y_{n_{(i)}}(T^{(i)})$.
- ▶ on met à jour Z .

L'algorithme Adaptive Multilevel Splitting.



Itérations: tant que $Z < a$,

- ▶ on supprime $Y_{(1)}, \dots, Y_{(k)}$.
- ▶ pour les remplacer: on tire $n_{(1)}, \dots, n_{(k)}$ iid dans $\{(k+1), \dots, (n)\}$.
- ▶ $Y_{(i)}(t) = Y_{n_{(i)}}(t)$ si $t \leq T^{(i)} := \inf \left\{ s \geq 0; \xi(Y_{n_{(i)}}(s)) > Z \right\}$.
- ▶ $Y_{(i)}(t)$ pour $t \geq T^{(i)}$: trajectoire Markovienne issue de $Y_{n_{(i)}}(T^{(i)})$.
- ▶ on met à jour Z .

L'estimateur de la probabilité

Calcul de l'estimateur: $\hat{p}^{n,k} = C^{n,k} \left(1 - \frac{k}{n}\right)^{Q^{n,k}}.$

avec $Q^{n,k}$: nombre d'itérations

$$C^{n,k} = \frac{\#\{i; \tau_B^i < \tau_A^i\}}{n}.$$

Théorème

L'algorithme fournit un estimateur non biaisé de la probabilité p .

Dans 2 situations:

1. cas "idéal"
2. dynamique à temps discret: avec modifications de l'algorithme

Résultats: cas idéal

Cas plus simple: $E = \mathbb{R}$, $\xi(x) = x$, $p = \mathbb{P}(X \geq a)$.

Hypothèses: $F = F_X$ **continue**, $F(0) = 0$.

A chaque itération: simulations selon $\mathcal{L}(X|X > Z)$.

Théorème (B.-Lelièvre-Rousset 2014)

L'estimateur $\hat{p}^{n,k} = C^{n,k} \left(1 - \frac{k}{n}\right)^{Q^{n,k}}$ est **non biaisé**:

$$\mathbb{E}\hat{p}^{n,k} = p.$$

Résultats: cas idéal

Cas plus simple: $E = \mathbb{R}$, $\xi(x) = x$, $p = \mathbb{P}(X \geq a)$.

Hypothèses: $F = F_X$ **continue**, $F(0) = 0$.

A chaque itération: simulations selon $\mathcal{L}(X|X > Z)$.

Théorème (B.-Lelièvre-Rousset 2014)

L'estimateur $\hat{p}^{n,k} = C^{n,k} \left(1 - \frac{k}{n}\right)^{Q^{n,k}}$ est **non biaisé**:

$$\mathbb{E}\hat{p}^{n,k} = p.$$

Théorème (B.-Lelièvre-Rousset 2014, B-Goudenège 2014)

Propriétés asymptotiques $n \rightarrow +\infty$ (k, p fixés)

$$\text{Var}(\hat{p}^{n,k}) \sim \frac{-p^2 \log(p)}{n} \quad \text{et} \quad \mathbb{E}[Q^{n,k}] \sim \frac{-n \log(p)}{k}$$
$$\sqrt{n}(\hat{p}^{n,k} - p) \rightarrow \mathcal{N}(0, -p^2 \log(p)).$$

Littérature:

$k = 1$ connu ($C^{n,1} = 1$, $Q^{n,1} \sim \mathcal{P}(-n \log(p))$).

$k > 1$: étudié avec un autre choix de Z et asymptotiques k/n fixé.

Références: [Cérou, Guyader 2007], [Guyader, Hengartner, Matzner-Løber 2011].

Technique de preuve

Hypothèse: X à densité f .

Technique de preuve

Hypothèse: X à densité f .

Probabilité estimée $p = P(0)$ avec $P(x) = \mathbb{P}(X > a | X > x)$, pour $0 \leq x \leq a$.

Estimateurs associés: $\hat{p}^{n,k}(x)$.

Technique de preuve

Hypothèse: X à densité f .

Probabilité estimée $p = P(0)$ avec $P(x) = \mathbb{P}(X > a | X > x)$, pour $0 \leq x \leq a$.

Estimateurs associés: $\hat{p}^{n,k}(x)$.

Equation fonctionnelle sur $x \mapsto p^{n,k}(x) := \mathbb{E}[\hat{p}^{n,k}(x)]$:

$$p^{n,k}(x) = \int_x^a \left(1 - \frac{k}{n}\right) p^{n,k}(y) f_{n,k}(y; x) dy + \sum_{\ell=0}^{k-1} \frac{n-\ell}{n} \mathbb{P}(X_{(\ell)}(x) \leq a < X_{(\ell+1)}(x))$$

Technique de preuve

Hypothèse: X à densité f .

Probabilité estimée $p = P(0)$ avec $P(x) = \mathbb{P}(X > a | X > x)$, pour $0 \leq x \leq a$.

Estimateurs associés: $\hat{p}^{n,k}(x)$.

Equation fonctionnelle sur $x \mapsto p^{n,k}(x) := \mathbb{E}[\hat{p}^{n,k}(x)]$:

$$p^{n,k}(x) = \int_x^a \left(1 - \frac{k}{n}\right) p^{n,k}(y) f_{n,k}(y; x) dy + \sum_{\ell=0}^{k-1} \frac{n-\ell}{n} \mathbb{P}(X_{(\ell)}(x) \leq a < X_{(\ell+1)}(x))$$

P est l'unique solution, donc $p^{n,k} = P$.

Technique de preuve

Hypothèse: X à densité f .

Probabilité estimée $p = P(0)$ avec $P(x) = \mathbb{P}(X > a | X > x)$, pour $0 \leq x \leq a$.

Estimateurs associés: $\hat{p}^{n,k}(x)$.

Equation fonctionnelle sur $x \mapsto p^{n,k}(x) := \mathbb{E}[\hat{p}^{n,k}(x)]$:

$$p^{n,k}(x) = \int_x^a \left(1 - \frac{k}{n}\right) p^{n,k}(y) f_{n,k}(y; x) dy + \sum_{\ell=0}^{k-1} \frac{n-\ell}{n} \mathbb{P}(X_{(\ell)}(x) \leq a < X_{(\ell+1)}(x))$$

P est l'unique solution, donc $p^{n,k} = P$.

Autres résultats:

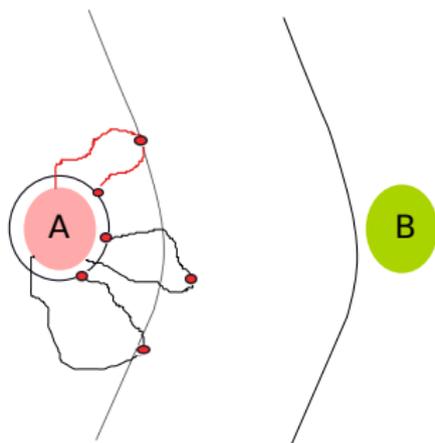
- ▶ une équation fonctionnelle;
- ▶ utiliser $-\log(1 - F_X(X)) \sim \mathcal{E}(1)$;
- ▶ trouver une EDO linéaire d'ordre k à coefficients constants;
- ▶ prendre la limite $n \rightarrow +\infty$.

L'intérêt de $k > 1$

Analyse du coût: optimal pour $k = 1$.
Mais utile si dynamique en temps discret

L'intérêt de $k > 1$

Analyse du coût: optimal pour $k = 1$.
Mais utile si dynamique en temps discret



Possibilité d'égalité de niveaux.

Generalized AMS: modifications

Principe: k_{level} (fixé) et k_{samp} (adaptatif).

Generalized AMS: modifications

Principe: k_{level} (fixé) et k_{samp} (adaptatif).

Une itération:

- ▶ calcul du niveau: $Z = X_{(k_{level})}$
- ▶ nombre de trajectoires à modifier: $k_{samp} \geq k_{level}$ tel que $X_{(k_{samp})} = Z < X_{(k_{samp}+1)}$
- ▶ si $k_{samp} = n$: "extinction", estimateur = 0
- ▶ choix du point de départ: premier instant i tel que $X_i > Z$
- ▶ mise à jour de l'estimateur: $\hat{p} \leftarrow \hat{p} \left(1 - \frac{k_{samp}}{n}\right)$.

Même critère d'arrêt (ou extinction) + même mise à jour finale.

Generalized AMS: modifications

Principe: k_{level} (fixé) et k_{samp} (adaptatif).

Une itération:

- ▶ calcul du niveau: $Z = X_{(k_{level})}$
- ▶ nombre de trajectoires à modifier: $k_{samp} \geq k_{level}$ tel que $X_{(k_{samp})} = Z < X_{(k_{samp}+1)}$
- ▶ si $k_{samp} = n$: "extinction", estimateur = 0
- ▶ choix du point de départ: premier instant i tel que $X_i > Z$
- ▶ mise à jour de l'estimateur: $\hat{p} \leftarrow \hat{p} \left(1 - \frac{k_{samp}}{n}\right)$.

Même critère d'arrêt (ou extinction) + même mise à jour finale.

Théorème

On a un estimateur non biaisé de la probabilité: $\mathbb{E}\hat{p} = p$ avec

$$\hat{p} = C \left(1 - \frac{k_{samp}^1}{n}\right) \dots \left(1 - \frac{k_{samp}^Q}{n}\right).$$

Important: ne dépend pas du choix de ξ .

Exemple numérique

Exemple 1D:

$$dX_t = -dt + \sqrt{2}dB_t$$

Discrétisation:

$$X_{i+1} = X_i - \Delta t + \sqrt{2\Delta t}G_i$$

Condition initiale: $X_0 = 1$. Pas de temps: $\Delta t = 0.1$.

$$p = \mathbb{P}(\tau_{1.9} < \tau_{0.1})$$

où $\tau_{1.9} = \inf\{i \geq 1, X_i > 1.9\}$ et $\tau_{0.1} = \inf\{i \geq 1, X_i < 0.1\}$.

Pour l'algorithme: $a = 1.9$.

Résultats numériques

Référence: simulation Monte-Carlo, $N_{exp} = 16.10^6$ réalisations.
 $\hat{p} = 0.2386 \pm 0.0002$

Résultats numériques

Référence: simulation Monte-Carlo, $N_{exp} = 16.10^6$ réalisations.

$$\hat{p} = 0.2386 \pm 0.0002$$

Algorithme Generalized AMS:

N_{exp}	10000	160000	160000
n	10	20	20
k	1	1	5
$\mathbb{E}(\hat{p})$	0.238 ± 0.002	$0.2388 \pm 0.0.0004$	0.2387 ± 0.0007
N_{exp}	40000	40000	
n	50	50	
k	1	10	
$\mathbb{E}(\hat{p})$	0.2386 ± 0.004	0.2386 ± 0.0005	

Table: GAMS

Résultats numériques: variantes biaisées ($k = 1$)

Si on repart de i tel que $X_i \geq Z$:

N_{exp}	10000	1000	100
n	10	100	1000
$\mathbb{E}(\hat{p})$	0.214 ± 0.002	0.212 ± 0.002	0.211 ± 0.002

Table: Variante 1

Résultats numériques: variantes biaisées ($k = 1$)

Si on repart de i tel que $X_i \geq Z$:

N_{exp}	10000	1000	100
n	10	100	1000
$\mathbb{E}(\hat{p})$	0.214 ± 0.002	0.212 ± 0.002	0.211 ± 0.002

Table: Variante 1

Si on rejette les i tels que $X_i \geq Z$:

N_{exp}	1000	1000	1000	100
n	10	50	100	1000
(i), $\mathbb{E}(\hat{p})$	0.227 ± 0.006	0.242 ± 0.003	0.238 ± 0.002	0.240 ± 0.002
(ii), $\mathbb{E}(\hat{p})$	0.230 ± 0.007	0.243 ± 0.003	0.238 ± 0.002	0.240 ± 0.002
(iii), $\mathbb{E}(\hat{p})$	0.256 ± 0.007	0.266 ± 0.003	0.261 ± 0.002	0.262 ± 0.002

Table: Variante 2

Conclusion

Analyse de l'algorithme AMS:

- ▶ cas idéal: pas de biais, analyse asymptotique
- ▶ intérêt de la variante $k > 1$: temps discret \rightarrow algorithme modifié \rightarrow pas de biais.

Résultats numériques: quantitatif 1D et qualitatif ∞ D (Allen-Cahn)

Conclusion

Analyse de l'algorithme AMS:

- ▶ cas idéal: pas de biais, analyse asymptotique
- ▶ intérêt de la variante $k > 1$: temps discret \rightarrow algorithme modifié \rightarrow pas de biais.

Résultats numériques: quantitatif 1D et qualitatif ∞ D (Allen-Cahn)

Une question: rôle de la coordonnée de réaction (cas test 2D)

Conclusion

Analyse de l'algorithme AMS:

- ▶ cas idéal: pas de biais, analyse asymptotique
- ▶ intérêt de la variante $k > 1$: temps discret \rightarrow algorithme modifié \rightarrow pas de biais.

Résultats numériques: quantitatif 1D et qualitatif ∞ D (Allen-Cahn)

Une question: rôle de la coordonnée de réaction (cas test 2D)

Références:

Analysis and simulation of rare events for SPDEs, B.-Gazeau-Goudenège-Rousset; Proceeding CEMRACS; 2013.

Analysis of adaptive multilevel splitting algorithms in an idealized setting, 2014, B.-Lelièvre-Rousset.

A Central Limit Theorem for adaptive multilevel splitting algorithms in an idealized setting, en préparation, B.-Goudenège.

Unbiased property of generalized adaptive multilevel splitting, en préparation, B.-Gazeau-Goudenège-Lelièvre-Rousset.

Merci pour votre attention.