

Document de synthèse présenté pour obtenir
L'HABILITATION À DIRIGER DES RECHERCHES
par

Emmanuel GOBET

Centre de Mathématiques Appliquées
École Polytechnique
91128 PALAISEAU Cedex

Contributions à la simulation et à l'analyse de discrétisation de processus, et applications.

Présenté le 24 novembre 2003, devant le jury composé de:

Patrick CATTIAUX

Nicole EL KAROUI

Laure ELIE (rapporteur)

Valentine GENON-CATALOT

Jean JACOD (rapporteur)

Damien LAMBERTON (rapporteur)

David NUALART (rapporteur)

Wolfgang RUNGALDIER

Introduction

Nous présentons quelques contributions à la simulation et à l'analyse de discrétisation de processus, avec leurs applications notamment en finance. Nous avons regroupé nos travaux selon 4 thèmes :

1. statistique des processus avec observations discrètes ;
2. couverture en temps discret en finance ;
3. sensibilités d'espérances ;
4. analyses d'erreurs de discrétisation.

Ce choix ne respecte absolument pas l'ordre chronologique de nos recherches. Le premier chapitre sur la statistique des processus est assez indépendant du reste. En revanche, les trois autres chapitres correspondent à une cohérence et une progression dans les questions soulevées. Néanmoins au fil de la lecture, on remarquera des liens entre les quatre parties : différentiation par rapport à des domaines et amélioration de simulation de temps de sortie, sensibilités d'espérances et statistique asymptotique avec le calcul de Malliavin, sensibilités d'espérances et analyse d'erreur etc... Les preuves des résultats s'appuient notamment sur les outils du calcul de Malliavin, des martingales, des Équations aux Dérivées Partielles (EDP) et de leurs liens avec les Équations Différentielles Stochastiques (EDS).

Pour présenter nos résultats, nous avons choisi de commencer par rappeler leurs motivations venues des applications ou des questions mathématiques. Les théorèmes sont formulés sous des jeux d'hypothèses un peu imprécis parfois, juste par soucis de ne pas alourdir la rédaction. Lorsque c'est possible, des tests numériques agrémentent et illustrent l'analyse du problème ou la pertinence de la méthode développée. Les travaux présentés correspondent à 8 articles et 5 prépublications, tous numérotés par ordre chronologique de 1 à 13, puis de 14 à 16 pour les actes de congrès et notes.

Articles parus ou à paraître

- [1] E. Gobet. Euler schemes for the weak approximation of killed diffusion. *Stochastic Processes and their Applications*, 87 :167–197, 2000.
- [2] E. Gobet and E. Temam. Discrete time hedging errors for options with irregular pay-offs. *Finance and Stochastics*, 5(3) :357–367, 2001.
- [3] E. Gobet. Local asymptotic mixed normality property for elliptic diffusion : a Malliavin calculus approach. *Bernoulli*, 7(6) :899–912, 2001.

- [4] E. Gobet. Euler schemes and half-space approximation for the simulation of diffusions in a domain. *ESAIM : Probability and Statistics*, 5 :261–297, 2001.
- [5] E. Gobet. LAN property for ergodic diffusion with discrete observations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(5) :711–737, 2002.
- [6] G. Bernis, E. Gobet, and A. Kohatsu-Higa. Monte Carlo evaluation of Greeks for multi-dimensional barrier and lookback options. *Mathematical Finance*, 13(1) :99–113, 2003.
- [7] E. Gobet and A. Kohatsu-Higa. Computation of Greeks for barrier and lookback options using Malliavin calculus. *Electronic Communications in Probability*, 8 :51–62, 2003.
- [8] E. Gobet, M. Hoffmann, and M. Reiß. Nonparametric estimation of scalar diffusions based on low frequency data. A paraître dans *The Annals of Statistics*. 30 pages.

Articles soumis

- [9] L. Carassus, E. Gobet, and E. Temam. Effective computations of super-replication prices with discrete rebalancing. 16 pages, 2001. Soumis à *Decisions in Economics and Finance*.
- [10] E. Gobet and R. Munos. Sensitivity analysis using Itô-Malliavin calculus and martingales. Application to stochastic control problem. 39 pages, 2002. Soumis à *SIAM Journal on Control and Optimization*.
- [11] E. Gobet and S. Menozzi. Discrete approximations of killed Itô processes. 33 pages, 2003. Soumis à *Stochastic Processes and their Applications*.
- [12] M. Bossy, E. Gobet, and D. Talay. Symmetrized euler scheme for an efficient approximation of reflected diffusions. 13 pages, 2003. Soumis à *Journal of Applied Probability*.
- [13] C. Costantini, N. El Karoui and E. Gobet. Représentation de Feynman-Kac dans des domaines temps-espace et sensibilité par rapport au domaine. 4 pages, 2003. Note soumise au *C. R. Acad. Sci. Paris Sér. I Math.*.

Actes de congrès avec comité de lecture

- [14] E. Gobet. Efficient schemes for the weak approximation of reflected diffusions. *Monte Carlo Methods Appl.*, 7(1-2) :193–202, 2001. Monte Carlo and probabilistic methods for partial differential equations (Monte Carlo, 2000).

Notes aux Compte-Rendus de l’Académie des Sciences

- [15] E. Gobet. Schéma d’Euler continu pour des diffusions tuées et options barrière. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(12) :1411–1414, 1998.
- [16] E. Gobet. Schéma d’Euler discret pour diffusion multidimensionnelle tuée. *C. R. Acad. Sci. Paris Sér. I Math.*, 328(6) :515–520, 1999.

Thèse

- [T] E. Gobet. *Schémas d’Euler pour diffusion tuée. Application aux options barrière*. Thèse de doctorat de l’Université Denis Diderot Paris 7, 1998, dont sont tirés les notes [15][16] et l’article [1].

Chapitre 1

Statistique des processus avec observations discrètes

Étant donnée une observation $(X_{t_k})_{0 \leq k \leq N}$ d'une solution d'une EDS à des instants déterministes équidistants $(t_k = k\Delta_N)_{0 \leq k \leq N}$, l'objectif est d'inférer sur les coefficients définissant la dynamique du processus dans l'asymptotique d'un nombre d'observations N tendant vers l'infini.

Nous donnons dans [3, 5, 8] différentes contributions à l'estimation et l'efficacité dans des situations de données haute fréquence ($\Delta_N \rightarrow 0$ avec horizon fixe $N\Delta_N = 1$ ou long $N\Delta_N \rightarrow \infty$) ou basse fréquence ($\Delta_N = \Delta > 0$). Il faut bien avouer que la frontière entre données basse et haute fréquence est assez floue du point de vue de l'expérimentateur. Se situer dans une situation $\Delta_N \rightarrow 0$, c'est croire d'une certaine manière que la fréquence d'échantillonnage est suffisamment élevée pour assurer la validité de l'approximation diffusive pour les chaînes de Markov.

1.1 Propriété LAMN dans le cas haute fréquence avec $N\Delta_N = 1$

Dans [3], nous considérons $X = (X_t)_{0 \leq t \leq 1}$, solution de l'EDS multidimensionnelle

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(\beta_0, s, X_s) dW_s,$$

où β_0 est un paramètre réel inconnu (rappelons qu'avec des données sur $[0, 1]$, il n'est pas possible d'inférer sur b). Il s'agit d'estimer β_0 à partir de l'observation de l'échantillon $(X_{i/N})_{0 \leq i \leq N}$. Une méthode heuristique possible consiste à construire à partir des données la matrice de variation/covariation quadratique discrète, et obtenir asymptotiquement le processus $(\sigma\sigma^*(\beta_0, s, X_s))_{0 \leq t \leq 1}$: sous une hypothèse d'identifiabilité du modèle, on en déduit β_0 . Ce n'est pas la seule manière de procéder et ces questions ont été résolues par Genon-Catalot et Jacod [GCJ94] qui construisent des estimateurs $\hat{\beta}_N$ de type contraste : ils convergent à la vitesse \sqrt{N} et l'erreur renormalisée $\sqrt{N}(\hat{\beta}_N - \beta_0)$ converge en loi vers une variable gaussienne conditionnellement à X (ou mixture de gaussiennes). Le problème que nous avons résolu est celui de l'efficacité asymptotique de ces estimateurs, en terme de variance optimale. On sait depuis les années 60-70 que ces questions sont en fait étroitement liées à la convergence des rapports de vraisemblance

locale (voir [Cam53, Cam60, Haj70, Haj71, Cam71], [Jeg82, Jeg83] dans ce cadre, ou encore [CY00]). Pour préciser les notations, si $Z_{N,\beta_0} = \log \left(\frac{d\mathbf{P}^{\beta_0+uN^{-1/2}}}{d\mathbf{P}^{\beta_0}} \right) (X_0, \dots, X_{i/N}, \dots, X_1)$ désigne le logarithme du rapport de vraisemblance de l'échantillon sous deux lois proches $\mathbf{P}^{\beta_0+uN^{-1/2}}$ et \mathbf{P}^{β_0} , la propriété LAMN (Local Asymptotic Mixed Normality) est satisfaite en β_0 lorsque Z_{N,β_0} converge en loi sous \mathbf{P}^{β_0} vers $u V_{\beta_0} \Gamma - \frac{1}{2} u^2 V_{\beta_0}^2$, où Γ est de loi gaussienne centrée réduite indépendante de la variable aléatoire $V_{\beta_0} : N^{1/2}$ et $[V_{\beta_0}]^{-1}$ sont alors les vitesses et variances asymptotiques optimales.

La propriété LAMN dans ce cadre a été prouvée par Donhal [Doh87] en dimension 1, sous hypothèse d'ellipticité. Notant $p^\beta(t_i, t_{i+1}, x_i, x_{i+1})$ la densité de transition de X sous la loi \mathbf{P}^β , la propriété de Markov fournit

$$Z_{N,\beta_0} = \sum_{i=0}^{N-1} \log \left(\frac{p^{\beta_0+uN^{-1/2}}(t_i, t_{i+1}, X_{i/N}, X_{(i+1)/N})}{p^{\beta_0}(t_i, t_{i+1}, X_{i/N}, X_{(i+1)/N})} \right); \quad (1.1)$$

en dimension 1, $p^\beta(t_i, t_{i+1}, x_i, x_{i+1})$ a une représentation quasi-explicite en termes de ponts browniens [Rog85] et l'approche dans [Doh87] consiste alors à en calculer un développement limité pour (t_i, x_i) et (t_{i+1}, x_{i+1}) proches, d'où la convergence de Z_{N,β_0} se déduit. Cette technique ne peut pas s'étendre systématiquement aux dimensions supérieures, même en utilisant les résultats de [Aze83]. Pour passer au cadre multidimensionnel, notre approche dans [3] a consisté à réécrire

$$\log \left(\frac{p^{\beta_0+uN^{-1/2}}(t_i, t_{i+1}, x_i, x_{i+1})}{p^{\beta_0}(t_i, t_{i+1}, x_i, x_{i+1})} \right) = \int_{\beta_0}^{\beta_0+uN^{-1/2}} \frac{\partial_\beta p^\beta(t_i, t_{i+1}, x_i, x_{i+1})}{p^\beta(t_i, t_{i+1}, x_i, x_{i+1})} d\beta,$$

puis exprimer les rapports $\frac{\partial_\beta p^\beta}{p^\beta}$ sous forme d'espérances conditionnelles à l'aide du calcul de Malliavin [Nua95] :

$$\frac{\partial_\beta p^\beta}{p^\beta}(t_i, t_{i+1}, x_i, x_{i+1}) = \mathbf{E}^\beta(H_i | X_{t_i} = x_i, X_{t_{i+1}} = x_{i+1}) \quad (1.2)$$

(on remarquera l'analogie avec les formules de Bismut [Bis84]). Ensuite, il suffit de procéder à quelques approximations stochastiques des variables H_i et surtout par des résultats de martingales, de bien contrôler la convergence de somme d'espérances conditionnelles de type $\sum_{i=0}^{N-1} \mathbf{E}^\beta(H_i | X_{t_i}, X_{t_{i+1}})$. La vraie difficulté technique est due au fait que les espérances conditionnelles soient calculées sous \mathbf{P}^β alors que la convergence doit avoir lieu sous $\mathbf{P}^{\beta_0} \neq \mathbf{P}^\beta$. Nous finissons par prouver la propriété LAMN en dimension quelconque sous condition d'ellipticité.

Théorème 1.1.1 ([3]) *Sous quelques hypothèses de régularité sur b et σ et sous une condition d'ellipticité ($\sigma(\beta_0, t, x)$ est symétrique et inversible pour tout $(t, x) \in [0, 1] \times \mathbf{R}^d$), alors la propriété LAMN est satisfaite avec $V_{\beta_0} = 2 \int_0^1 \text{Tr}([\partial_\beta \sigma \sigma^{-1}]^2(\beta_0, t, X_t)) dt$.*

Ce résultat n'est pas vraiment une surprise, mais il est remarquable de constater que cette approche par calcul de Malliavin fait apparaître assez vite la limite cherchée, ce qui n'est pas aussi clair avec l'approche précédente en dimension 1. Notons pour conclure que le Théorème 1.1.1 justifie l'optimalité des estimateurs proposés dans [GCJ94].

1.2 Propriété LAN dans le cas haute fréquence avec $N\Delta_N \rightarrow \infty$

Ce travail s'inscrit dans la lignée du précédent, mais cette fois, avec un modèle d'EDS dépendant de deux paramètres (multidimensionnels) dans les coefficients de dérive et de diffusion

$$X_t = x + \int_0^t b(\alpha, X_s) ds + \int_0^t \sigma(\beta, X_s) dW_s.$$

et des données aux temps $(i\Delta_N)_{0 \leq i \leq N}$ avec un horizon d'observations tendant vers l'infini. C'est une asymptotique qui permet de retrouver à la fois les effets haute fréquence pour l'estimation de β et l'effet temps grand pour l'estimation de α . Le processus est supposé pour cela satisfaire une propriété d'ergodicité.

Les procédures d'estimation ont été étudiées dans [Rao83, FZ89, Yos92, Kes97] mais à chaque fois sous des restrictions techniques sur la vitesse de décroissance de Δ_N ; dans [Kes97], il suffit de $N\Delta_N^p \rightarrow 0$ pour un $p > 1$. Les estimations de α et β ont lieu aux vitesses $\sqrt{N\Delta_N}$ et $\sqrt{\Delta_N}$.

La question de l'optimalité est restée ouverte jusqu'à notre travail [5] où nous avons établi la propriété LAN (Local Asymptotic Normality), analogue de LAMN avec une variance¹ inverse limite *déterministe* dans la convergence de la log-vraisemblance locale²

$$Z_{N,\alpha_0,\beta_0} = \sum_{i=0}^{N-1} \log \left(\frac{p^{\alpha_0+u[N\Delta_N]^{-1/2}, \beta_0+v[\Delta_N]^{-1/2}}(\Delta_N, X_{i\Delta_N}, X_{(i+1)\Delta_N})}{p^{\alpha_0,\beta_0}(\Delta_N, X_{i\Delta_N}, X_{(i+1)\Delta_N})} \right)$$

sous $\mathbf{P}^{\alpha_0,\beta_0}$. Le jeu suivant d'hypothèses sur b et σ assure l'ergodicité de X sous $\mathbf{P}^{\alpha,\beta}$ [Has80].

- (H1) *Régularité.* Les coefficients b et σ sont localement assez réguliers en α , β et x .
- (H2) *Croissances des coefficients.* b est au plus à croissance linéaire en x , alors que σ et les différentes dérivées de b et σ sont uniformément bornées.
- (H3) *Ellipticité.* σ est symétrique et satisfait une condition d'ellipticité uniforme.
- (H4) *Ergodicité.* Une condition de dérive fortement rentrante est vérifiée : $b(\alpha, x) \cdot x \leq -C|x|^2$ pour x assez grand ($C > 0$).

Ne pas supposer b bornée est assez essentiel pour les applications, car il s'agit d'englober les processus de type Ornstein-Uhlenbeck.

Théorème 1.2.1 ([5]) *Sous les hypothèses (H1) à (H4), la propriété LAN est satisfaite sans condition supplémentaire sur la décroissance de Δ_N , avec pour matrice d'information de*

$$\text{Fisher limite } V^{\alpha_0,\beta_0} = \left(\begin{array}{c|c} V_b^{\alpha_0,\beta_0} & 0 \\ \hline 0 & V_\sigma^{\alpha_0,\beta_0} \end{array} \right), \text{ déterministe et explicite.}$$

Ainsi, les estimateurs donnés dans [Kes97] sont optimaux, et on retrouve leur indépendance asymptotique, au vue de la structure par bloc de V^{α_0,β_0} .

Notre approche utilise comme dans le cas $N\Delta_N = 1$ la représentation de Z_{N,α_0,β_0} à l'aide du calcul de Malliavin. Néanmoins, au niveau des convergences, des différences importantes

¹ici, les paramètres étant multidimensionnels, il s'agit plutôt d'une matrice de variance/covariance.

² X étant homogène en temps, nous notons maintenant $p^{\alpha,\beta}(\Delta, x, y)$ à la place de $p^{\alpha,\beta}(t, t + \Delta, x, y)$.

apparaissent. D'abord, les approximations stochastiques des variables H_i apparaissant dans (1.2) doivent être faites en gardant des *termes d'erreur centrés*, sans quoi des conditions $N\Delta_N^p \rightarrow 0$ apparaissent comme dans [Kes97]. Ensuite, pour bien contrôler ces erreurs au niveau de Z_{N,α_0,β_0} , des encadrements de type Aronson [Aro67] de $p^{\alpha,\beta}$ sont nécessaires, ce qui n'est pas standard car σ est borné et b est sous-linéaire. En introduisant le noyau gaussien $g(t, x, y) = [2\pi t]^{-d/2} \exp\left(-\frac{|x-y|^2}{2t}\right)$, nous montrons les estimations suivantes.

Proposition 1.2.2 ([5]) *Sous les hypothèses (H1) à (H3), il existe des constantes $c > 1$ and $K > 1$ telles que pour tout $(t, x, y) \in]0, 1] \times \mathbf{R}^d \times \mathbf{R}^d$, on a*

$$\begin{aligned} \frac{1}{C} g(c^{-1}t, x, y) \exp(-ct|x|^2) &\leq p^{\alpha,\beta}(t, x, y) \leq C g(ct, x, y) \exp(ct|x|^2), \\ t^{-1/2} |\partial_\alpha p^{\alpha,\beta}(t, x, y)| + |\partial_\beta p^{\alpha,\beta}(t, x, y)| &\leq C g(ct, x, y) \exp(ct|x|^2). \end{aligned}$$

Il est facile de voir que les termes $\exp(\pm ct|x|^2)$ peuvent être remplacés par $\exp(\pm ct|y|^2)$, quitte à prendre t assez petit. Ces nouveaux termes proviennent de l'hypothèse de croissance linéaire sur b , comme il est facile de le vérifier dans le cas explicite du processus d'Ornstein Uhlenbeck. La Proposition 1.2.2 traitent des situations intermédiaires entre le cas de coefficients bornés (où l'encadrement par noyaux gaussiens est valide [Fri75]) et le cas de coefficients à croissance linéaire (où des majorations à décroissance plus rapide que tout polynôme en $|x-y|$ sont valables [KS87]).

1.3 Estimation non-paramétrique à basse fréquence $\Delta_N = \Delta > 0$

Les deux approches précédentes s'inscrivent dans une situation où une forme paramétrique de la dynamique du processus est connue, mais ce choix n'est pas toujours réaliste. On peut lui préférer un point de vue plus général, où la recherche des coefficients b et σ s'effectue dans un certain espace fonctionnel de régularité.

Des résultats traitent le cadre haute fréquence [Hof99, Jac00] mais le cas de données basse fréquence est resté longtemps hors de portée. En effet, les phénomènes sont très différents du cas $\Delta_N \rightarrow 0$ où asymptotiquement, le comportement infinitésimal de l'EDS est retrouvé : dans le cas $\Delta_N = \Delta > 0$, nous sommes seulement en présence d'une chaîne de Markov dont on observe l'opérateur de transition $P_\Delta^{b,\sigma}$.

Avec Marc Hoffmann et Markus Reiß, nous avons considéré ce cadre en se plaçant sous l'hypothèse technique supplémentaire que X est une EDS unidimensionnelle uniformément elliptique, vivant dans un intervalle disons $I = [0, 1]$, avec réflexion aux bords (et donc ergodique avec une unique probabilité invariante de densité $\mu^{b,\sigma}$) :

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s + \int_0^t n(X_s) dk_s$$

avec $(k_t)_{t \geq 0}$ croissant seulement en 0 ou 1, et $n(0) = 1 = -n(1)$. Nous considérons dans [8] un critère de perte \mathbf{L}_2 et déterminons alors les vitesses minimax associées.

Théorème 1.3.1 ([8]) *Supposons des coefficients bornés b et σ dans une boule \mathcal{B}_s de régularité \mathbf{L}_2 -Sobolev d'indice $s - 1$ et s ($s > 1$). Alors, pour tout $0 < a < b < 1$, il existe une constante $C > 0$ telle que pour tout estimateur $(\hat{b}, \hat{\sigma})$, on a*

$$\sup_{(b,\sigma) \in \mathcal{B}_s} [\mathbf{E}^{b,\sigma} \|\hat{b} - b\|_{\mathbf{L}_2([a,b])}^2]^{1/2} \geq CN^{-\frac{(s-1)}{2s+3}} \quad \text{et} \quad \sup_{(b,\sigma) \in \mathcal{B}_s} [\mathbf{E}^{b,\sigma} \|\hat{\sigma} - \sigma\|_{\mathbf{L}_2([a,b])}^2]^{1/2} \geq CN^{-\frac{s}{2s+3}},$$

avec un estimateur atteignant ces bornes.

Ces résultats sont assez inattendus. En effet, d'une part, les deux vitesses (ramenées aux mêmes régularités en b et σ) sont différentes de celles du cadre haute fréquence, et d'autre part elles ne sont pas égales entre elles *a contrario* du cadre $\Delta_N \rightarrow 0$: en fait, elles s'interprètent comme celles de problèmes inverses d'ordre 2 pour b et 1 pour σ .

La construction d'estimateurs s'appuie sur la possibilité asymptotiquement d'avoir accès à la densité invariance $\mu^{b,\sigma}$ et à un couple fonction propre/valeur propre pour $P_\Delta^{b,\sigma}$ (et donc pour le générateur infinitésimal $L^{b,\sigma}$). Les deux équations $[L^{b,\sigma}]^* \mu^{b,\sigma} = 0$ et $L^{b,\sigma} \phi = \nu \phi$ permettent alors d'identifier b et σ . L'optimalité des vitesses se prouve de manière standard en contrôlant la distance de Kullback pour des modèles proches : les plus difficiles à séparer statistiquement correspondent à ceux donnant la même densité invariante $\mu^{b,\sigma}$.

Chapitre 2

Couverture en temps discret en finance

Marquée par les idées de Black-Scholes [BS73] et Merton [Mer73], la valorisation moderne des options repose sur la capacité à répliquer un flux aléatoire dans le futur par un portefeuille dynamique (dit de couverture) investi en 1 actif non risqué et d actifs risqués.

Notons $V_t^{x,\delta}$ la valeur à l'instant t du portefeuille autofinçant, de richesse initiale x , contenant à l'instant t un nombre δ_t d'actifs risqués, le reste étant investi à la banque au taux sans risque (supposé ici constant et égal à r). Supposons un marché financier parfait, à savoir sans arbitrage, sans coût de transaction, avec des actifs liquides et des stratégies en temps continu... Le marché est dit complet (et incomplet sinon) si en particulier, toute option européenne de flux raisonnable Ψ en T est répliquable par un portefeuille de couverture : $\mathbf{P}(V_T^{x,\delta} = \Psi) = 1$. La richesse initiale x est alors le prix de l'option et vaut généralement $x = \mathbf{E}^{\mathbf{Q}}(e^{-rT}\Psi)$, l'espérance du flux actualisé sous une nouvelle probabilité $\mathbf{Q} \sim \mathbf{P}$ dite *probabilité martingale*¹, rendant martingales les actifs actualisés [KS98]. En marché incomplet, d'autres critères de prix doivent être utilisés, comme par exemple le prix de sur-réplication [KQ95], défini comme la plus petite richesse initiale x permettant de sur-répliquer Ψ en T : $\mathbf{P}(V_T^{x,\delta} \geq \Psi) = 1$.

Nos contributions [2, 9] se concentrent sur l'impact de stratégie en temps discret dans ces problèmes de réplication en marché complet, ou de sur-réplication en marché incomplet.

2.1 Couverture approchée d'options européennes

Considérons un modèle d'actif risqué en dimension $d = 1$, de dynamique sous la probabilité historique \mathbf{P}

$$\frac{dS_t}{S_t} = \mu(S_t)dt + \sigma(S_t)dW_t,$$

sur lequel est écrite une option de maturité T et de payoff $\Psi = f(S_T)$. Les exemples pertinents sont notamment ceux du Call ($f(x) = (x - K)_+$) ou de l'option digitale ($f(x) = \mathbf{1}_{x > K}$). Sous l'hypothèse que σ ne s'annule pas, le marché est complet pour des payoffs dans \mathbf{L}_2 : le prix en $t = 0$ est $u(0, S_0)$, où u solution d'une EDP de type Cauchy, alors que la couverture vaut $\delta_t = \partial_x u(t, S_t)$.

¹ou probabilité *neutre au risque*

En pratique, les stratégies de couverture en temps continu ne sont pas réalistes et il est naturel de se limiter à des dates discrètes de transactions, disons $(t_i = iT/N)_{0 \leq i \leq N}$: il convient alors d'analyser l'impact d'une telle discrétisation, en considérant par exemple la nouvelle stratégie (constante par morceaux) $\tilde{\delta}_t = \partial_x u(\phi(t), S_{\phi(t)})$ où $\phi(t) = t_i$ pour $t_i \leq t < t_{i+1}$. L'erreur de couverture $\Delta_N(f)$ est définie alors par

$$\Delta_N(f) := f(S_T) - V_T^{u(0, S_0), \tilde{\delta}} = e^{rT} \int_0^T [\partial_x u(t, S_t) - \partial_x u(\phi(t), S_{\phi(t)})] d(e^{-rt} S_t). \quad (2.1)$$

Clairement, $\Delta_N(f)$ tend vers 0 en probabilité (sous **P** et **Q**) et il convient d'examiner sa vitesse de convergence, pour obtenir ainsi des renseignements sur la distribution des pertes et profits en T due cette couverture discrète. Dans la publication [2] réalisée avec Emmanuel Temam, nous avons retenu un critère quadratique pour mesurer cette erreur.

En fait, la convergence en loi de l'intégrale stochastique $\Delta_N(f)$ définie par (2.1) a lieu à la vitesse $N^{1/2}$ (voir [Roo80]), mais la variable limite peut ne pas être de carré intégrable : ces résultats ne renseignent donc pas sur la convergence dans \mathbf{L}_2 .

Dans le cas du Call, la vitesse de convergence dans \mathbf{L}_2 est bien égale à $N^{1/2}$ comme établi dans [Zha99]. La situation se détériore pour des payoffs f moins réguliers : c'est un phénomène bien connu des praticiens.

Théorème 2.1.1 ([2]) *Sous des hypothèses de régularité de μ et σ , pour l'option digitale $f(x) = \mathbf{1}_{x > K}$, nous avons*

$$\lim_{N \rightarrow \infty} N^{1/4} \|\Delta_N(f)\|_{\mathbf{L}_2} = C > 0,$$

la norme \mathbf{L}_2 pouvant être calculée sous **P** ou sous **Q**.

Nous montrons aussi que d'autres vitesses intermédiaires entre $N^{1/4}$ et $N^{1/2}$ apparaissent pour des régularités intermédiaires de f (de type Höldérienne). L'obtention de ces résultats nécessitent des estimations précises de $u(t, x)$ et de ses dérivées lorsque t tend vers T , notamment pour x au voisinage de K : la représentation de la densité de transition de $(S_t)_{t \geq 0}$ en termes de ponts browniens [Rog85] nous permet de calculer les équivalents recherchés. Des généralisations du Théorème 2.1.1 à des normes BMO ont été par la suite étudiées dans [Gei02], ainsi que des extensions multidimensionnelles dans [Tem03].

2.2 Prix de sur-réplication explicites en temps discret

Au lieu d'une couverture approchée, nous considérons dans [9] avec Laurence Carassus et Emmanuel Temam les prix de sur-réplication avec des stratégies à temps discret, appliquées aux dates $(0 = t_0 < \dots < t_i < \dots < t_N = T)$ déterministes (pas nécessairement équidistantes). Quitte à changer de numéraire, supposons que le taux sans risque r est nul. Les processus de prix des $d(\geq 1)$ actifs risqués peuvent être définis en temps continu, mais on suppose qu'en fait ils sont seulement observés aux dates de transaction : la filtration \mathcal{F}_{t_j} est donc celle générée par les variables $(S_{t_i})_{0 \leq i \leq j}$. Le prix de sur-réplication de l'option européenne de payoff $\Psi = \Psi(S_{t_0}, \dots, S_{t_N})$ est défini par

$$x_0 = \inf\{x : \exists \text{ une stratégie à temps discret } \delta \text{ telle que } V_T^{x, \delta} \geq \Psi \text{ avec probabilité } 1\}. \quad (2.2)$$

A cette formulation primale nous faisons correspondre le point de vue dual : dans notre cas, c'est une adaptation sans changement essentiel de [Pha00] (voir [KQ95, FK97] dans le cas continu) et cela donne

$$x_0 = \sup_{Q \in \mathcal{P}} \mathbf{E}^Q[\Psi], \quad (2.3)$$

pourvu que le supremum soit fini et sous des hypothèses usuelles, dont l'absence d'opportunité d'arbitrage. Dans (2.3), \mathcal{P} désigne l'ensemble des probabilités martingales $\mathbf{Q} \sim \mathbf{P}$ avec $\frac{d\mathbf{Q}}{d\mathbf{P}}|_{\mathcal{F}_T}$ borné. Malheureusement, cette représentation duale (2.3), tout comme (2.2), ne permet qu'assez rarement de calculer explicitement le prix de surréplication : voir [EJ97, BCS98, CPT99] dans le cas de payoff du type $\Psi = f(S_T)$.

Dans ce contexte à temps discret, en mixant les arguments des deux formulations, nous parvenons à obtenir des formules fermées, pour tout type d'option, toute dynamique de sous-jacents (sans hypothèse de Markov) et toute dimension d'actifs, à la différence des précédentes résolutions (formulées en continu, il est vrai). L'énoncé du résultat demande quelques notations :

- la loi conditionnelle de S_j sachant $\mathcal{F}_{t_{j-1}}$ est notée $\mu_j(S_{t_0}, \dots, S_{t_{j-1}})$;
- la loi du vecteur $(S_{t_0}, \dots, S_{t_j})$ est notée \mathbf{P}_j ;
- pour Ψ mesurable de $(\mathbf{R}^d)^{N+1}$ dans \mathbf{R} , nous définissons une suite d'opérateurs $(\Gamma_j^e \Psi)_{0 \leq j \leq N}$, définie par une équation de programmation dynamique à base d'enveloppes concaves :

$$\begin{aligned} \Gamma_N^e \Psi(x_0, \dots, x_N) &= \Psi(x_0, \dots, x_N), \\ \Gamma_j^e \Psi(x_0, \dots, x_j) &= \operatorname{ess\,inf}_{(\alpha, \beta) \in \mathbf{R} \times \mathbf{R}^d} \{f_{\alpha, \beta}^{\Gamma_{j+1}^e \Psi}\}(x_0, \dots, x_j) \quad \mathbf{P}_j - p.p., \quad 0 \leq j \leq N-1, \\ f_{\alpha, \beta}^u(x_0, \dots, x_j) &= \begin{cases} \alpha + \beta \cdot x_j & \text{si } \mu_{j+1}(x_0, \dots, x_j) \{z : \alpha + \beta \cdot z < u(x_0, \dots, x_j, z)\} = 0 \\ +\infty & \text{sinon.} \end{cases} \end{aligned}$$

Théorème 2.2.1 ([9]) *Sous les hypothèses d'obtention de (2.3), nous avons*

$$x_0 = [\Gamma_0^e \Psi](S_0).$$

Dans [9], des versions plus générales de (2.3) et du Théorème 2.2.1 englobent des contraintes coniques sur les stratégies et permettent de traiter le cas d'options américaines. À la vue des opérateurs $(\Gamma_j^e)_{1 \leq j \leq N}$, le calcul explicite des prix ne dépend que du support des lois des actifs sous-jacents, ce qui est intuitivement satisfaisant. Notamment, pour bien des situations (modèles d'EDS browniennes, modèle de Merton incluant des sauts poissonniens, modèles à volatilité stochastique), les lois μ_j sont équivalentes à λ_+ , la mesure de Lebesgue sur $]0, \infty[^d$ (et similairement pour \mathbf{P}_j) : pour calculer effectivement $(\Gamma_j^e)_{1 \leq j \leq N}$, il reste alors à déterminer des enveloppes concaves à des ensembles λ_+ -négligeables près. Pour les fonctions usuelles de payoff (souvent linéaires par morceaux), ces enveloppes sont immédiates à obtenir et la détermination de x_0 en résulte très facilement, et ceci pour toute fonction de payoff avec exercices européen et américain : cela donne ainsi un traitement assez complet du problème de sur-réplication en temps discret.

Comme on pouvait s'y attendre [EJ97], les prix obtenus sont en pratique élevés et correspondent souvent à des stratégies triviales. L'applicabilité de ces résultats à des situations réelles est donc finalement assez limitée.

Chapitre 3

Sensibilités d'espérances

Dans de nombreuses applications, il est important de pouvoir évaluer numériquement la sensibilité d'espérance de variables aléatoires $X(\alpha)$ par rapport à un paramètre α qui les définit : α peut être de dimension finie, mais aussi infinie (une fonction).

Les exemples auxquels nous nous sommes intéressés sont les suivants.

1. En finance, comme rappelé au chapitre 2, il est essentiel de déterminer le gradient du prix d'option à la date courante (donné sous forme d'espérance) par rapport aux cours des actifs sous-jacents à la même date, car en général cela fournit la stratégie de couverture de l'option. Plus généralement, les sensibilités (appelés aussi *Greeks*) du prix par rapport aux paramètres du modèle servent d'indicateurs de risque. Nous avons étudié cette question de calcul des Greeks dans le cas d'option barrière ou lookback, où le payoff est une fonction des minima/maxima des actifs sous-jacents, d'abord pour un modèle de diffusion assez général en dimension 1 [7], puis en dimension supérieure avec un modèle log-normal [6].
2. Dans [10], nous étudions la sensibilité de $\mathbf{E}(f(X_T))$ par rapport aux coefficients définissant la dynamique de l'EDS satisfaite par $X = (X_t)_{0 \leq t \leq T}$. Cette question se pose assez naturellement lorsqu'on analyse l'impact d'une erreur de modèles (en finance par exemple), ou lorsqu'on cherche à résoudre un problème de contrôle stochastique de manière approchée, en paramétrant d'abord le contrôle pour résoudre ensuite un problème d'optimisation paramétrique à l'aide de sensibilités par rapport à la paramétrisation. Nous déterminons dans [10] des estimateurs de ces sensibilités.
3. Enfin, la résolution de problème d'arrêt optimal [Kar81] (options américaines en finance [Kar88][Lam98]) peut être menée en cherchant à optimiser la *région d'exercice* \mathcal{E} [Gar01]. Pour ce problème, la sensibilité de $\mathbf{E}(g(\tau, X_\tau))$ (avec $\tau = \inf\{T \geq 0 : (t, X_t) \in \mathcal{E}\}$) par rapport à \mathcal{E} est une quantité pertinente, que nous déterminons dans [13].

Quelle forme pour des estimateurs de sensibilité? Au moins dans le cas où α est un paramètre fini-dimensionnel, le calcul de $\partial_\alpha \mathbf{E}(X(\alpha))$ par simulation peut être faite par la *méthode de resimulation* [LP94], consistant à calculer $\mathbf{E}(X(\alpha'))$ par méthodes de Monte Carlo pour différents α' , pour en déduire $\partial_\alpha \mathbf{E}(X(\alpha))$ par des formules de dérivées discrètes : cette approche est coûteuse lorsque le paramètre est de grande dimension et fournit des estimateurs biaisés. La seconde approche (*sensibilité trajectorielle*) se base sur la formule de dérivation sous l'espérance

$\partial_\alpha \mathbf{E}(X(\alpha)) = \mathbf{E}(\partial_\alpha X(\alpha))$, valable si $\alpha \mapsto X(\alpha)$ est dérivable ω par ω : cette condition est une vraie restriction dans beaucoup de situations, elle n'est pas satisfaite en général dans les exemples évoqués précédemment. La troisième méthode consiste à procéder à une dérivation au niveau des lois, en écrivant

$$\partial_\alpha \mathbf{E}(X(\alpha)) = \mathbf{E}(Y(\alpha)) \quad (3.1)$$

pour une certaine variable aléatoire $Y(\alpha)$, propice à la simulation : les outils de type Calcul de Malliavin peuvent être particulièrement adaptés dans ces questions, comme nous l'illustrons dans [7, 6]. Mais l'EDP sous-jacente à un cadre markovien d'EDS, les martingales et le calcul stochastique peuvent se révéler tout aussi efficaces (voir [10, 13]). Dans la suite, pour les calculs et formules relatives au calcul de Malliavin, nous utilisons les notations de Nualart [Nua95].

Mentionnons avant de rentrer dans le détail, que trouver des variables $Y(\alpha)$ dans (3.1) est une réponse au problème théorique. Ensuite, pour la pertinence numérique, la simulation de ces variables doit être aussi efficace que possible (en termes de temps de calcul), pour une variance compétitive : les expérimentations numériques valident alors les formules déterminées.

3.1 Calcul des Greeks pour les options barrière et lookback

Dans [7] réalisé avec Arturo Kohatsu-Higa, nous nous intéressons à déterminer les dérivées¹ de $\mathbf{E}(\Phi(\max_{s \in I} S_s, \min_{s \in I} S_s, S_T))$ par rapport à S_0 pour $S = (S_t)_{t \geq 0}$ une EDS unidimensionnelle, un ensemble $I \subset [0, T]$ de dates de constatation ($I = \{t_i : 0 \leq i \leq N\}$ ou $I = [0, T]$ par exemple) et une fonction de payoff Φ sans régularité supposée. Au facteur d'actualisation près, ces sensibilités déterminent la couverture d'option barrière ou lookback (selon la forme de la fonction Φ) en temps discret ou continu.

L'approche standard pour ce type de problème a été donnée dans [FLL⁺99] : cela consiste d'abord à faire passer la différentiation dans l'espérance (après avoir éventuellement régularisé Φ), puis d'intégrer par parties au sens du calcul de Malliavin. La difficulté technique supplémentaire qu'il faut surmonter dans notre situation provient de la faible régularité des variables $\max_{s \in I} S_s$ et $\min_{s \in I} S_s$, seulement dans $\mathbf{D}^{1,p}$ [NV88], ce qui empêche une formule d'intégration par parties directe.

Pour y parvenir, nous utilisons une intégration par parties locale, dans l'esprit de celle effectuée pour le maximum du drap brownien [Nua95], en introduisant un processus Y dominant (adapté, croissant et continu à droite) :

$$\forall t \in I \quad |\log(S_t) - \log(S_0)| \leq Y_t. \quad (3.2)$$

Le processus Y doit aussi satisfaire certaines estimations, notamment de régularité au sens du calcul de Malliavin : lorsque I est discret, des candidats pour Y sont faciles à trouver, alors qu'en temps continu, on en obtient via le lemme de Garsia-Rodemich-Rumsey [GRR70].

Une subtilité subsiste encore : il apparaît que la matrice de covariance de Malliavin du vecteur $(\max_{s \in I} S_s, \min_{s \in I} S_s, S_T)$ calculée avec localisation est cette fois plus régulier, mais reste dans tous les cas non inversible. Toutefois, la sensibilité spécifique des Greeks permet de contourner

¹en fait les deux premières

cet écueil. Sous une restriction sur la forme de la fonction de payoff, nous montrons des formules du type suivant.

Théorème 3.1.1 ([7]) *Considérons un modèle d'actif de dynamique $(S_t)_{t \geq 0}$ donnée par un brownien géométrique. Sous la condition que la fonction $\Phi(M, m, S)$ ne dépend pas de M et de m dès lors que $M \leq S_0 + a$ ou $m \geq S_0 - a$ (pour $a > 0$), il existe une variable aléatoire H (ne dépendant de Φ que par l'intermédiaire de a) telle que*

$$\partial_{S_0} \mathbf{E}(\Phi(\max_{s \in I} S_s, \min_{s \in I} S_s, S_T)) = \mathbf{E}(\Phi(\max_{s \in I} S_s, \min_{s \in I} S_s, S_T) H). \quad (3.3)$$

Dans [7], nous montrons une formule analogue pour la seconde dérivée, pour un modèle de diffusion elliptique générale en dimension 1 : en fait, via un changement de variable et un changement de probabilité, il est standard de pouvoir se ramener au cas brownien. Des tests numériques dans le cas continu $I = [0, T]$ illustrent que dans certaines situations, la variance des simulations est plus faible que par les autres méthodes : toutefois, c'est à comparer au surcoût de calcul lié à la simulation de H (qui s'exprime à l'aide d'intégrales de Skorohod). Ce travail a été préparé lors de notre visite à Arturo Kohatsu-Higa à Barcelone au printemps 2000 : sa publication a été très retardée car le comité de rédaction du journal où le travail a été soumis initialement a avoué au bout d'un an et demi avoir perdu nos manuscrits, après quoi une nouvelle soumission a été faite ailleurs.



FIG. 3.1 – Évolution de 3 estimateurs Monte-Carlo en fonction du nombre de trajectoires M : pour $M = 200000$, les intervalles de confiance à 95% ont pour largeur 7.08×10^{-4} (différences finies), 1.82×10^{-3} (estimateur de [FLL⁺99]) et 1.7×10^{-4} (notre estimateur).

Plus tard, Guillaume Bernis nous a rejoint, pour étendre ces sensibilités au cas de d actifs sous-jacents :

$$\Phi \left(\max_{s \in I} S_s^1, \dots, \max_{s \in I} S_s^d, \min_{s \in I} S_s^1, \dots, \min_{s \in I} S_s^d, S_T^1, \dots, S_T^d \right).$$

Cette extension [6] n'a été possible que dans le cas de mouvements browniens géométriques (avec corrélation) sous des conditions sur Φ du même type qu'avant, mais un peu affaiblies. Les résultats numériques pour des dates de constatations discrètes sont concluants en faveur de cette approche, notamment lorsque la fréquence de constatations est grande. Le graphe 3.1 tiré de [6] en donne une illustration : nous traçons l'estimateur des différences finies, celui de [FLL⁺99] qui exploite une intégration par parties sur le premier intervalle de temps (d'où une variance très grande) et notre estimateur pour un certain processus dominant Y .

La condition sur Φ n'est pas très restrictive car elle est automatique dans le cas d'options barrières, mais elle permet de traiter seulement certaines options lookback : néanmoins, dans [6], nous montrons qu'elle est aussi d'une certaine manière nécessaire pour pouvoir obtenir des représentations du type (3.3).

3.2 Sensibilités par rapport aux coefficients d'une EDS

Dans [10] réalisé avec Rémi Munos, nous considérons un modèle général d'EDS inhomogène, avec coefficients réguliers dépendant d'un paramètre α (disons uni-dimensionnel pour simplifier)

$$X_t = x + \int_0^t b(s, X_s, \alpha) ds + \sum_{j=1}^q \int_0^t \sigma_j(s, X_s, \alpha) dW_s^j; \quad (3.4)$$

nous cherchons des estimateurs de $\partial_\alpha \mathbf{E}(f(X_T))$. Des résultats de Kushner et Yang [KY91] permettent de traiter le cas où f est dérivable, en différentiant dans l'espérance et introduisant la dérivée trajectorielle de X_t par rapport à α , solution de

$$\partial_\alpha X_t = \int_0^t (\partial_\alpha b_s + b'_s \partial_\alpha X_s) ds + \sum_{j=1}^q \int_0^t (\partial_\alpha \sigma_{j,s} + \sigma'_{j,s} \partial_\alpha X_s) dW_s^j, \quad (3.5)$$

en notant pour simplifier $\partial_\alpha b_s, b'_s$ etc. à la place de $\partial_\alpha b(\alpha, s, X_s), \nabla_x b(\alpha, s, X_s)$ etc.. Il en découle

Proposition 3.2.1 (Dérivée trajectorielle [KY91]) *Si f est dérivable (avec certaines conditions d'intégrabilité sur f'), nous avons*

$$\partial_\alpha \mathbf{E}(f(X_T)) = \mathbf{E}(f'(X_T) \partial_\alpha X_T). \quad (3.6)$$

Pour f sans régularité et une matrice σ carrée inversible indépendante de α , on peut faire porter la différentiation juste sur la vraisemblance par rapport à une probabilité \mathbf{P}^{α_0} fixe (obtenu par le théorème Girsanov) : c'est la méthode dite du *score* ou du *rapport de vraisemblance* [Gly86, Gly87, RW86].

Proposition 3.2.2 (Méthode du score [KY91]) *Pour $q = d$ et σ inversible indépendante de α , nous avons*

$$\partial_\alpha \mathbf{E}(f(X_T)) = \mathbf{E}\left(f(X_T) \int_0^T [\sigma_t^{-1} \partial_\alpha b_t]^* dW_t\right). \quad (3.7)$$

Notre travail a consisté à déterminer des estimateurs couvrant les situations où f n'est pas régulière et σ dépend de α , en affaiblissant éventuellement l'hypothèse d'ellipticité.

Si l'on part de (3.6), il est assez naturel d'intégrer par parties au sens du calcul de Malliavin : cela a été développé dans un cadre elliptique dans [FLL⁺99] avec à la clé, des formules assez simples, mais cela se généralise facilement sous des conditions plus faibles que l'ellipticité.

Proposition 3.2.3 (Approche calcul de Malliavin [10]) *Supposons que γ_T , la matrice de covariance de Malliavin de X_T , soit inversible avec des inverses dans tous les $\mathbf{L}_p, p \geq 1$. Alors, nous avons*

$$\partial_\alpha \mathbf{E}(f(X_T)) = \mathbf{E} \left(f(X_T) \delta(\partial_\alpha X_T^* \gamma_T^{-1} \mathcal{D}.X_T) \right). \quad (3.8)$$

La condition sur γ_T est satisfaite sous des hypothèses usuelles de type Hörmander : voir [Nua95] ou [CM02] pour de récents développements. Si X_T est remplacé par quelques-unes de ses composantes, la condition se ramène à de l'ellipticité (partielle) sur ces composantes, ce qui a de l'intérêt en finance pour l'optimisation de portefeuilles. Même si cette représentation est assez générale, elle est assez complexe à implémenter numériquement (voir les tests et discussions dans [10]). Dans les situations $q = d$ et σ inversible (ce que nous supposons dans la suite), nous proposons des représentations plus commodes.

Une première approche s'inspire de formulation des sensibilités à l'aide d'états adjoints (voir [Ben88, Pen90] en contrôle stochastique). Introduisons $u(\alpha, t, x) = \mathbf{E}(f(X_T) | X_t = x)$, solution de l'EDP parabolique $\partial_t u(\alpha, \cdot) + L^\alpha u(\alpha, \cdot) = 0$ avec condition de Cauchy $u(\alpha, T, \cdot) = f(\cdot)$. Au moins formellement, on obtient par différentiation directe, que $v(\alpha, \cdot) = \partial_\alpha u(\alpha, \cdot)$ est solution de $\partial_t v(\alpha, \cdot) + L^\alpha v(\alpha, \cdot) = -\partial_\alpha [L^\alpha] u(\alpha, \cdot)$ avec $v(\alpha, T, \cdot) = 0$. Cette équation s'interprète par la formule de Feynman-Kac sous la forme

$$\partial_\alpha \mathbf{E}(f(X_T)) = \int_0^T \mathbf{E} \left(\sum_{i=1}^d \partial_\alpha b_{i,t} \partial_{x_i} u(\alpha, t, X_t) + \frac{1}{2} \sum_{i,j=1}^d \partial_\alpha [\sigma \sigma^*]_{i,j,t} \partial_{x_i, x_j}^2 u(\alpha, t, X_t) \right) dt.$$

Pour obtenir un estimateur explicite, il suffit de représenter les dérivées de u au point (t, X_t) en termes d'espérances conditionnelles impliquant $f(X_T)$ et des quantités connues. Pour cela et pour garder des formules simples (évitant des intégrales de Skorohod), nous utilisons des techniques de calcul stochastique de Bismut [Bis84] (voir aussi [Tha97, EJM99, Pic02]). Remarquant que $(u(\alpha, t, X_t))_{0 \leq t < T}$ est une martingale, il en est de même pour $(M_t = \nabla_x u(\alpha, t, X_t) Y_t)_{0 \leq t < T}$ en différentiant par rapport à la condition initiale x (avec les notations $Y_t = \nabla_x X_t$ et $Z_t = [Y_t]^{-1}$). Cette propriété combinée avec la décomposition explicite du théorème de représentation prévisible $f(X_T) = u(t, X_t) + \int_t^T \nabla_x u(\alpha, s, X_s) \sigma_s dW_s$ donne la formule $\partial_x u(\alpha, t, X_t) Y_t = \frac{1}{T-t} \mathbf{E}(f(X_T) [\int_t^T [\sigma_s^{-1} Y_s]^* dW_s]^* | \mathcal{F}_t)$. Nous généralisons ce type de formule pour la dérivée seconde et obtenons finalement

Théorème 3.2.4 (Approche par les adjoints [10]) *Supposons $q = d$ et σ inversible, et notons e^i le i -ème vecteur de la base canonique de \mathbf{R}^d . Alors, nous avons $\partial_\alpha \mathbf{E}(f(X_T)) =$*

$\mathbf{E}(H_T^{b,Adj.} + \frac{1}{2}H_T^{\sigma,Adj.})$ avec

$$\begin{aligned}
H_T^{b,Adj.} &= f(X_T) \int_0^T dt \partial_\alpha b_t \cdot \frac{Z_t^*}{T-t} \int_t^T [\sigma_s^{-1} Y_s]^* dW_s \\
H_T^{\sigma,Adj.} &= \int_0^T dt \sum_{i,j=1}^d \partial_\alpha [\sigma \sigma^*]_{i,j,t} [f(X_T) - f(X_t)] \left(\frac{2e^j}{T-t} \cdot [Z_t^* \int_{\frac{T+t}{2}}^T [\sigma_s^{-1} Y_s]^* dW_s] \right. \\
&\quad \left. \times \frac{2e^i}{T-t} \cdot [Z_t^* \int_t^{\frac{T+t}{2}} [\sigma_s^{-1} Y_s]^* dW_s] + \frac{2e^i}{T-t} \cdot \{ \nabla_x [Z_t^* \int_t^{\frac{T+t}{2}} [\sigma_s^{-1} Y_s]^* dW_s] Z_t e^j \} \right), \quad (3.9)
\end{aligned}$$

sous l'hypothèse que $\int_0^T \frac{\|f(X_T) - f(X_t)\|_{L^{p_0}}}{T-t} dt < +\infty$ pour un certain $p_0 > 1$. L'estimateur de sensibilité $H_T^{b,Adj.} + \frac{1}{2}H_T^{\sigma,Adj.}$ appartient alors à \mathbf{L}_p pour $p < p_0$.

Si l'on réécrit la condition sur f à l'aide des densités de transition du processus et qu'on utilise les estimations d'Aronson [Fri75] pour ces dernières, on voit apparaître que cette hypothèse sur f se lit comme une condition de régularité faible de type Sobolev W^{s,p_0} à poids avec $s > 0$. En particulier, elle est satisfaite pour f hölderienne, ou f égale à l'indicatrice d'un domaine régulier ou d'un polyèdre, ce qui englobe beaucoup de situations intéressantes pour les applications.

La représentation (3.9) est en fait plus simple d'utilisation pour des fins numériques que (3.8), même si la formule est plus longue à écrire. Elle s'avère utile si le nombre de paramètres est grand car alors, il y a une factorisation des calculs grâce aux pseudo²-adjoints (en facteur de $\partial_\alpha b_t$ et $\partial_\alpha [\sigma \sigma^*]_t$).

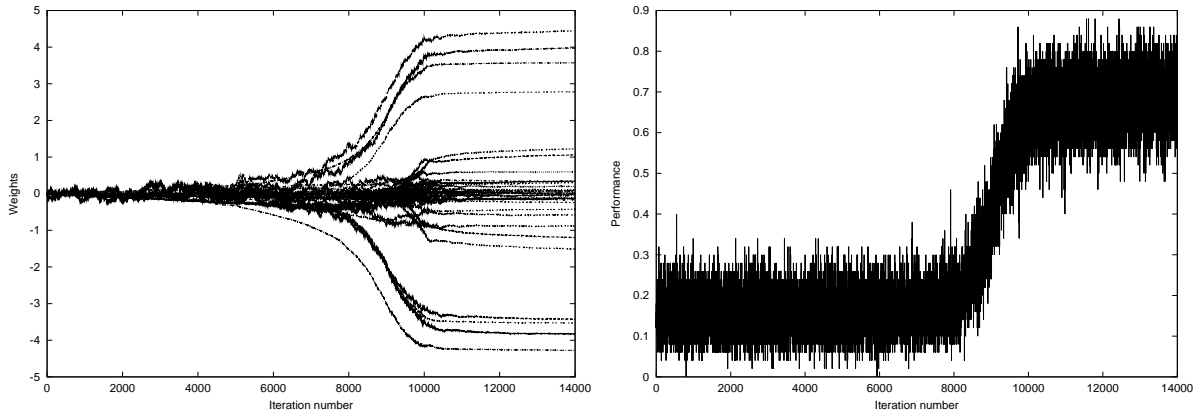


FIG. 3.2 – Évolution des paramètres et de la performance au cours des itérations de la méthode de gradient stochastique, en utilisant pour le calcul des gradients des estimateurs de sensibilités.

L'autre approche que nous développons se base essentiellement sur des arguments de martingales. Partant de la martingale $u(\alpha, t, X_t)$, nous déduisons par différentiation par rapport à α que $(N_t = \partial_\alpha u(\alpha, t, X_t) + \nabla_x u(\alpha, t, X_t) \partial_\alpha X_t)_{0 \leq t < T}$ est aussi une martingale : cette propriété conduit à une équation intégrale pour $\mathbf{E}(\partial_\alpha u(\alpha, t, X_t))$ dont la solution en $t = 0$ est la quantité cherchée. Nous obtenons

²l'adjectif *pseudo* renvoie au fait qu'ils ne sont pas adaptés.

Théorème 3.2.5 (Approche martingale [10]) *Sous les hypothèses du théorème 3.2.4, nous avons $\partial_\alpha \mathbf{E}(f(X_T)) = \mathbf{E}(H_T^{Mart.})$ avec*

$$H_T^{Mart.} = \frac{f(X_T)}{T} \int_0^T [\sigma_s^{-1} \partial_\alpha X_s]^* dW_s + \int_0^T dr \frac{[f(X_T) - f(X_r)]}{(T-r)^2} \int_r^T [\sigma_s^{-1} (\partial_\alpha X_s - Y_s Z_r \partial_\alpha X_r)]^* dW_s. \quad (3.10)$$

Cette représentation est très intéressante et assez surprenante.

- Cette formule n’engage que les dérivées premières des coefficients, ce qui est inhabituel dans des formules d’intégration par parties du calcul de Malliavin.
- D’autre part, elle est très simple et sa mise en oeuvre numérique conduit à un temps calcul comparable à celui de l’estimateur de la dérivée trajectorielle (3.6) (et avec une variance compétitive en générale).

Nous analysons aussi dans [10] les problèmes liés à la simulation et à la discrétisation de tous ces estimateurs de sensibilité : nous renvoyons au chapitre 4 pour ces aspects.

Le graphe 3.2 tiré de [10] donne une application de ces calculs de sensibilités dans un problème de contrôle stochastique avec paramétrisation du contrôle (ici par un réseau de neurones) : l’objectif est de maximiser la probabilité d’atteindre une certaine cible stochastique. Sont représentés l’évolution des paramètres (*weights*) et de la probabilité correspondante (*performance*), en fonction du nombre d’itérations d’un algorithme de gradient stochastique pour l’optimisation.

3.3 Sensibilités par rapport au domaine

Dans ce travail [13] réalisé avec Cristina Costantini et Nicole El Karoui, des questions de sensibilité moins standards sont abordées. Nous considérons une EDS inhomogène $X = (X_s^{t,x})_{s \geq t}$ à valeurs dans \mathbf{R}^d , partant de x à l’instant t , et $\tau^{t,x} = \inf\{s \geq t : (s, X_s^{t,x}) \notin \mathcal{D}\}$ le premier temps de sortie du processus $(s, X_s^{t,x})_{s \geq t}$ du domaine temps-espace \mathcal{D} , ouvert borné de $]-\infty, T[\times \mathbf{R}^d$. Nous considérons des espérances de fonctionnelle stoppée en $\tau^{t,x}$ de la forme

$$u(t, x) = \mathbf{E} \left(g(\tau^{t,x}, X_{\tau^{t,x}}^{t,x}) e^{-\int_t^{\tau^{t,x}} c(r, X_r^{t,x}) dr} - \int_t^{\tau^{t,x}} e^{-\int_t^s c(r, X_r^{t,x}) dr} f(s, X_s^{t,x}) ds \right). \quad (3.11)$$

Sous des hypothèses de régularité de \mathcal{D} et de continuité (höldérienne) des fonctions g, f et c , il est standard de vérifier que u défini ainsi est la représentation de Feynman-Kac de l’unique solution de classe $C^0(\bar{\mathcal{D}}) \cap C^{1,2}(\mathcal{D}, \mathbf{R})$ de l’EDP parabolique avec condition de Cauchy-Dirichlet

$$\begin{cases} \partial_t u + Lu - cu = f & \text{dans } \mathcal{D} \\ u = g & \text{sur } \mathcal{PD}. \end{cases} \quad (3.12)$$

Dans (3.12), \mathcal{PD} est la frontière parabolique de \mathcal{D} : nous renvoyons à [Lie96] pour des définitions et résultats précis pour les EDP paraboliques dans des domaines temps-espaces.

Motivés par l’évaluation des options américaines via une optimisation de la frontière d’exercice, nous cherchons à déterminer la sensibilité de u par rapport au domaine \mathcal{D} . Nous montrons

Théorème 3.3.1 ([13]) *Pour $\Theta : [0, T] \times \mathbf{R}^d \mapsto \mathbf{R}^d$ régulier, considérons le domaine perturbé $\mathcal{D}_\epsilon = \{(t, x) : (t, x + \epsilon\Theta(t, x)) \in \mathcal{D}\}$ et le temps de sortie correspondant $\tau_\epsilon^{t,x} := \inf\{s \geq t : (s, X_s^{t,x}) \notin \mathcal{D}_\epsilon\}$. Sous des conditions de régularité sur f, g, c et d'ellipticité uniforme de σ , l'application $J : \epsilon \mapsto J(\epsilon)(t, x) = \mathbf{E}(g(\tau_\epsilon^{t,x}, X_{\tau_\epsilon^{t,x}}^{t,x})e^{-\int_t^{\tau_\epsilon^{t,x}} c(r, X_r^{t,x})dr} - \int_t^{\tau_\epsilon^{t,x}} e^{-\int_t^s c(r, X_r^{t,x})dr} f(s, X_s^{t,x})ds)$ est différentiable en $\epsilon = 0$ pour $(t, x) \in \mathcal{D} \cap \{0 \leq t \leq T\}$ avec*

$$\partial_\epsilon J(\epsilon)(t, x)|_{\epsilon=0} = \mathbf{E} \left[e^{-\int_t^{\tau^{t,x}} c(r, X_r^{t,x})dr} [(\nabla u - \nabla g)\Theta](\tau^{t,x}, X_{\tau^{t,x}}^{t,x}) \right].$$

La sensibilité d'EDP par rapport aux domaines est une question classique en optimisation de formes, avec de nombreuses applications aux calculs de structures mécaniques optimales : les travaux de Murat et Simon [MS76] posent un formalisme complet pour ce problème étudié il y a longtemps par Hadamard. En probabilité, c'est beaucoup moins standard et moins naturel, car évidemment l'application qui à \mathcal{D} associe le temps de sortie correspondant n'est pas régulier ω par ω : néanmoins, comme nous le montrons avec le théorème précédent, les lois de fonctionnelles arrêtées sont régulières.

Quelle est notre apport probabiliste vis-à-vis des analystes ?

1. Le cas stationnaire (EDP elliptique) est complètement traité dans [MS76, Sim80, Pir84], pour des EDS générales (mais non dégénérées).
2. Pour le cas parabolique, [SZ92] analyse le cas du laplacien dans des domaines cylindriques $\mathcal{D} =]-\infty, T[\times D$. Nous étendons ces résultats au cas de diffusions plus générales, mais aussi au domaine temps-espace (pertinents dans le cadre de l'arrêt optimal).
3. Notre contribution se situe aussi au niveau de la preuve, qui à l'aide des processus et du calcul stochastique, est assez simple.

L'application de cette analyse de sensibilité à l'arrêt optimal est en cours de développement. Mentionnons enfin que cette formule a de l'intérêt aussi pour l'amélioration des procédures de simulation de temps de sortie. Dans le chapitre 4, nous montrons que la correction d'erreur de discrétisation peut être réalisée en décalant la frontière de manière adéquate (voir Théorème 4.2.4) : le Théorème 3.3.1 permet d'analyser l'impact d'une telle modification de la frontière.

Chapitre 4

Analyses d'erreurs de discrétisation

Cette partie est relative plus particulièrement aux articles [1, 4, 12, 10, 11].

De manière générale, nous nous intéressons dans ce chapitre à l'évaluation d'espérance de fonctionnelles Ψ de la trajectoire d'un processus $X = (X_t)_{0 \leq t \leq T}$ par méthodes de Monte-Carlo. Sauf cas exceptionnel, la simulation exacte de Ψ n'est pas envisageable, et seule celle d'une approximation Ψ^N sera possible, via par exemple l'utilisation d'un schéma de discrétisation en temps de X avec pas $h = T/N$ (N étant le nombre de pas). Ainsi, on se retrouve en mesure de simuler des copies indépendantes de Ψ^N : l'estimateur Monte-Carlo de $\mathbf{E}(X)$ qui en résulte cumule deux erreurs de nature différente. D'une part, l'erreur statistique, qui s'estime par le Théorème de la Limite Centrale à la vitesse \sqrt{M} (avec des variances en général stables en N). D'autre part, l'erreur de discrétisation $\mathbf{E}(\Psi^N) - \mathbf{E}(\Psi)$, qui ne dépend que de la procédure de discrétisation choisie : c'est celle qui retient notre attention dans ce chapitre, l'objectif étant de déterminer précisément la vitesse de convergence en fonction de N .

Cette question d'analyse de discrétisation est importante en théorie car elle permet évidemment de valider la qualité d'une procédure d'approximation, ou au contraire met à jour ses insuffisances (comme dans le cas de simulation discrète de temps de sortie [11]). Cette problématique est aussi cruciale en pratique, et ce pour deux raisons.

1. Si l'on souhaite augmenter la précision finale de l'algorithme, on dispose d'une règle concrète permettant de répartir optimalement l'effort de calcul en équilibrant les deux sources d'erreur [DG95].
2. Si au delà de la vitesse de convergence de l'erreur $\mathbf{E}(\Psi^N) - \mathbf{E}(\Psi)$, on obtient aussi son développement limité, alors la convergence numérique peut être accélérée (sans pour autant beaucoup modifier la complexité globale) : il s'agit des techniques d'extrapolation de Romberg [TL90], consistant à calculer des combinaisons linéaires astucieuses de résultats numériques obtenus pour différents nombres de pas N .

Les fonctionnelles que nous avons considérées dans nos travaux sont du type suivant.

- Elles dépendent d'une marginale du processus dans les cas les plus simples [4, 12] : $\Psi = f(X_T)$.
- Elles sont parfois davantage compliquées en se mettant sous la forme $\Psi = f(X_T)H$ où H

est une variable aléatoire définie à partir de certaines intégrales stochastiques d'Itô ou de Skorohod [10]. Cette configuration est très fréquente dans les problèmes de sensibilités du chapitre 3.

- Elles peuvent correspondre à la fonctionnelle de meurtre du processus tué à sa sortie d'un domaine D : $\Psi = f(X_T)\mathbf{1}_{T < \tau}$ où $\tau = \inf\{t \geq 0 : X_t \notin D\}$ [1, 4, 11]. Cette situation trouve des applications notamment en finance avec le calcul de prix d'options barrière.

Les processus sous-jacents avec lesquels nous avons travaillé sont continus, soit solutions d'EDS définies dans tout l'espace [1, 4, 10, 11], soit plus généralement des processus d'Itô [11] ou encore des EDS avec réflexion au bord d'un domaine [4, 12].

4.1 Approximations d'EDS dans tout l'espace

Considérons X solution de l'EDS vectorielle

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \quad (4.1)$$

W étant un mouvement brownien multidimensionnel. Dans toute la suite, nous supposons pour simplifier b and σ de classe C^∞ avec des dérivées premières bornées.

L'approximation de processus de diffusion a maintenant une quarantaine d'années d'histoire : à la fin des années 60, des résultats généraux de Stroock et Varadhan [SV69a, SV69b, SV71] justifient des approximations par chaînes de Markov pour la convergence en loi. Mais pour établir des vitesses de convergence plus facilement, les travaux de recherche dans les années 70 s'orientent davantage vers des discrétisations trajectorielles de X à l'aide de développement de type Itô-Taylor [Mil74, Mil76, PW82]. Précisément, considérons N pour le nombre de pas de temps dans l'intervalle $[0, T]$ et définissons des instants de discrétisation $(t_k)_{0 \leq k \leq N}$, uniformément répartis (c'est-à-dire $t_k = kh$ avec $h = T/N$ le pas de temps). La procédure d'approximation la plus simple est l'approximation d'Euler définie par

$$X_0^N = X_0, \quad X_{t_{k+1}}^N - X_{t_k}^N = b(X_{t_k}^N)(t_{k+1} - t_k) + \sigma(X_{t_k}^N)(W_{t_{k+1}} - W_{t_k}), \quad (4.2)$$

qui est très commode à simuler numériquement : voir [KP95] pour un panorama des autres schémas a priori plus précis théoriquement, mais dont la simulation est en général plus complexe. On peut aussi considérer l'extension continue de (4.2) en posant pour $t_k \leq t < t_{k+1}$

$$X_t^N - X_{t_k}^N = b(X_{t_k}^N)(t - t_k) + \sigma(X_{t_k}^N)(W_t - W_{t_k}). \quad (4.3)$$

Les techniques usuelles de calcul stochastique permettent de montrer que l'erreur en norme \mathbf{L}_p est d'ordre $N^{-1/2}$, mais dans beaucoup de situations (valorisation d'options ou calcul de risque en finance, résolution probabiliste d'équations aux dérivées partielles), le critère de convergence faible (consistant notamment à estimer l'erreur $\mathbf{E}(f(X_T^N)) - \mathbf{E}(f(X_T))$) est le plus pertinent.

4.1.1 Analyse de l'erreur faible via la propriété de Markov et l'EDP

C'est au milieu des années 80 que Milshtein [Mil85] et Talay [Tal86] prouvent que le schéma d'Euler converge faiblement à la vitesse N pour des fonctions f de classe C^4 . Plus tard, Talay et

Tubaro [TL90] justifient que l'erreur admet en fait un développement asymptotique en puissance de $\frac{1}{N}$ pour des fonctions indéfiniment dérivables : comme rappelé précédemment, ce type de résultat permet d'utiliser les techniques d'extrapolation de Romberg et d'augmenter très simplement l'ordre de convergence. Du coup, les autres schémas d'ordre directement plus élevé (avec hélas, en général, des complexités bien plus grandes) perdent de l'intérêt en pratique. Cette analyse d'erreur a été étendue par Bally et Talay [BT96b] au cas où f est irrégulière sous conditions d'hypoellipticité uniforme pour X . Les résultats de majoration d'erreur pour l'approximation $\mathbf{E}(f(X_T^N))$ se résument ainsi :

Théorème 4.1.1 ([TL90, BT96b]) *On a*

$$|\mathbf{E}(f(X_T^N)) - \mathbf{E}(f(X_T))| \leq \frac{C}{N} \quad (4.4)$$

sous l'une des 2 hypothèses suivantes :

1. f est de classe C^4 à croissance polynomiale ainsi que ses dérivées.
2. f est mesurable bornée et les coefficients b, σ satisfont à une condition de Hörmander faible uniforme en $x \in \mathbf{R}^d$.

La preuve, sous l'une ou l'autre des hypothèses, consiste à décomposer l'erreur globale en une somme d'erreurs locales analysées à l'aide du calcul stochastique, en introduisant la fonction $u(t, x) = \mathbf{E}(f(X_T) | X_t = x)$ solution d'une EDP avec condition de Cauchy. Lorsque f est régulière, u l'est aussi, simplement par la régularité du flot $x \mapsto X_t^{t,x}$; pour f mesurable, la régularité de u provient de celle des densités de la loi de X_T conditionnellement à $X_t = x$, ce qui est assuré par les conditions d'hypoellipticité du théorème (voir [KS87]). Remarquons que la structure markovienne de la fonctionnelle d'intérêt $f(X_T)$ apparaît cruciale : cela écarte tout espoir de pouvoir traiter par cette méthode des fonctionnelles de type $f(X_T) \int_0^T h(X_s) dW_s$, comme on a pu en rencontrer dans le chapitre 3 sur les sensibilités.

4.1.2 Analyse directe de l'erreur faible via le calcul de Malliavin

L'approche que nous avons développée avec Rémi Munos dans [10] a été précisément motivée par les estimateurs de sensibilités introduits au chapitre 3, correspondant par exemple à des fonctionnelles de la forme $f(X_T)H$ où H s'exprime à l'aide d'intégrales d'Itô ou de Skorohod. L'utilisation du calcul de Malliavin pour montrer des estimations d'erreur faible n'est pas nouvelle : elle a été utilisée pour la première fois par Kohatsu-Higa [KH01] dans le cas d'EDS anticipatives. Nous adaptons ces techniques pour traiter des fonctionnelles compliquées de type $f(X_T)H$, avec des techniques de localisation spécifiques, différentes de celles de [BT96b] ou [KHP02]. Le résultat que nous montrons, puis appliquons à différentes situations, prend la forme suivante.

Théorème 4.1.2 ([10]) *Considérons une fonction f mesurable et une variable aléatoire H_T dans \mathbf{D}^∞ , avec les conditions $\mathbf{E}(f^2(X_T)) + \mathbf{E}(f^2(X_T^N)) < \infty$. Alors, on a*

$$|\mathbf{E}(f(X_T^N)H_T^N) - \mathbf{E}(f(X_T)H_T)| \leq \frac{C}{N} \quad (4.5)$$

sous les 2 conditions suivantes :

- (A) la différence $H_T^N - H_T$ satisfait un développement de type Itô-Taylor de la forme (en notant $dW_t^0 = dt$ et $\phi(t) = t_i$ si $t_i \leq t < t_{i+1}$)

$$H_T^N - H_T = \sum_{0 \leq i, j \leq q} c_{i,j}^0(T) \int_0^T c_{i,j}^1(t) \left(\int_{\phi(t)}^t c_{i,j}^2(s) dW_s^i \right) dW_t^j$$

pour des processus adaptés $(c_{i,j}^k(t))_{0 \leq t \leq T}$, tels que les variables aléatoires $c_{i,j}^k(t)$ appartiennent à \mathbf{D}^∞ avec des estimations uniformes en N et en temps des normes Sobolev : $\sup_N \sup_{t \in [0, T]} \|c_{i,j}^k(t)\|_{k', p'} < \infty$ pour tout $k', p' \geq 1$.

- (B) La matrice γ^{X_T} de covariance de Malliavin de X_T est inversible avec des inverses dans tous les \mathbf{L}_p , $p \geq 1$.

La condition (A) est en pratique simple à vérifier. Par exemple, si H et H_T^N sont respectivement les valeurs en T d'une EDS à coefficients réguliers (ou d'une de ses composantes) et de son approximation d'Euler, la décomposition dans (A) ainsi que les estimations rattachées sont automatiquement satisfaites : en effet, on reconnaît une des premières étapes utilisées dans l'étude de la convergence en loi du processus d'erreur renormalisée $[\sqrt{N}(X_t^N - X_t)]_{t \geq 0}$ (voir [KP91] ou [JP98] plus généralement). Quant à la condition (B), elle est satisfaite si X est solution d'une EDS sous une hypothèse de Hörmander faible seulement en X_0 . Cela permet de revisiter les résultats du théorème 4.1.1 en prenant $H_T = H_T^N = 1$, pour obtenir

Corollaire 4.1.3 *Supposons des coefficients b et σ de classe C^∞ à dérivées bornées. Alors, on a $|\mathbf{E}(f(X_T^N)) - \mathbf{E}(f(X_T))| \leq \frac{C}{N}$ sous la condition que f est mesurable bornée et que les coefficients b, σ satisfont à une condition de Hörmander faible en X_0 (au lieu de uniformément sur tout l'espace dans [BT96b]).*

Précisons que le théorème 4.1.2 est aussi valable si X est la solution d'une EDS inhomogène en temps ou même seulement égal à une ou plusieurs composantes de celle-ci : nous retrouvons ainsi les résultats de [TZ02] sous des hypothèses d'ellipticité partielle plus faibles.

Mentionnons par ailleurs que cette approche directe par calcul de Malliavin est intéressante aussi lorsque f est régulière car elle montre que l'estimation (4.4) reste valable si f n'est que 3 fois dérivables au lieu de 4 fois dans [TL90].

Avec des variantes du théorème 4.1.2 étendu aux EDS inhomogènes en temps, nous montrons par exemple que la discrétisation de l'estimateur de sensibilité par la méthode de l'adjoint (voir Théorème 3.2.4) est d'ordre $\frac{1}{N}$: cela illustre la portée de cette approche qui est à même d'analyser le cas de fonctionnelles sans structure markovienne.

Si elle est facile d'implémentation, cette procédure n'est pas satisfaisante au niveau de la précision : il a été largement observé *numériquement* que l'erreur $\mathbf{E}(\mathbf{1}_{T < \tau_d^N} f(X_T^N)) - \mathbf{E}(\mathbf{1}_{T < \tau} f(X_T))$ est positive (lorsque f est positive) et d'ordre $\frac{1}{\sqrt{N}}$ [RR91, BL94]. L'existence d'un biais est facile à comprendre intuitivement car s'il n'y a pas d'approximation du processus aux temps de discrétisation, on a $\tau \geq \tau_d^N$ avec probabilité 1. Approcher les valeurs de X aux instants $(t_k)_{0 \leq k \leq N}$ ne change pas *en espérance* cette surestimation : c'est ce que nous avons démontré.

Théorème 4.2.2 ([1, 11]) *Supposons la condition de frontière non caractéristique (C) et une condition de Hörmander fort uniforme dans D . Pour un domaine de classe C^3 et une fonction f mesurable bornée positive non nulle s'annulant¹ au voisinage de ∂D , on a*

$$\frac{C_1}{\sqrt{N}} \leq \mathbf{E}(\mathbf{1}_{T < \tau_d^N} f(X_T^N)) - \mathbf{E}(\mathbf{1}_{T < \tau} f(X_T)) \leq \frac{C_2}{\sqrt{N}} \quad (4.6)$$

pour deux constantes $C_2 \geq C_1 > 0$.

Seule la majoration de droite dans le cas elliptique est prouvée dans [1], et ce n'est que plus récemment dans [11] que nous justifions le biais systématique avec la constante $C_1 > 0$, sous des conditions hypoelliptiques fortes.

Par analogie au cas dans tout l'espace [BT96b], nous basons notre preuve sur la propriété de Markov pour faire rentrer en jeu la fonction $v(t, x) = \mathbf{E}(\mathbf{1}_{\forall s \in [t, T], X_s \in D} f(X_T) | X_t = x)$, solution dans D d'une EDP parabolique avec conditions de Cauchy et Dirichlet. La nouvelle difficulté tient au fait que la fonction v a une discontinuité au niveau de sa dérivée normale en ∂D . Pour dépasser ce problème, nous établissons dans [1] une généralisation dans \mathbf{R}^d de la formule d'Itô-Tanaka pour de telles fonctions : elle s'écrit à l'aide du temps local en 0 de $(F(X_t^N))_{0 \leq t \leq T}$, où F est la distance signée à la frontière (c'est-à-dire $F > 0$ dans D , $F \leq 0$ en dehors de D , F régulière et $|F(x)| = d(x, \partial D)$ au voisinage de ∂D). L'erreur dans (4.6) se réécrit alors plus précisément.

Théorème 4.2.3 ([11]) *Sous les conditions du Théorème 4.2.2, nous avons*

$$\mathbf{E}(\mathbf{1}_{T < \tau_d^N} f(X_T^N)) - \mathbf{E}(\mathbf{1}_{T < \tau} f(X_T)) = \mathbf{E} \left(\int_0^T \frac{\partial v}{\partial n}(t, X_t^N) d[F(X_{t \wedge \tau_d^N}^N)]^- \right) + O\left(\frac{1}{N}\right). \quad (4.7)$$

Cette représentation, qui sert à dériver en fait l'encadrement (4.6), fait apparaître $[F(X_{t \wedge \tau_d^N}^N)]_{0 \leq t \leq T}^-$, le processus d'*overshoot* arrêté au temps de sortie discret. Cela justifie heuristiquement pourquoi la convergence est d'ordre $\frac{1}{\sqrt{N}}$, puisque ses accroissements sont aussi de cet ordre.

Il est naturel d'exploiter l'égalité (4.7) pour déduire un développement limité de l'erreur, améliorant ainsi le Théorème 4.2.2 qui donne seulement un encadrement : cela passe notamment, au vue de (4.7), par l'étude de la loi asymptotique du processus d'*overshoot*. C'est une question délicate que nous n'avons pas su résoudre en toute généralité : les techniques habituelles s'appuient sur des théorèmes de renouvellement pour chaînes de Markov [Als94, FL01], qui ne sont valables que sous certaines hypothèses d'ergodicité qui ne peuvent pas être satisfaites dans notre contexte. Dans le cas spécifique du demi-espace pour D et du mouvement brownien avec dérive et corrélation pour X (pour lequel $X = X^N$), le problème de l'*overshoot* se ramène à

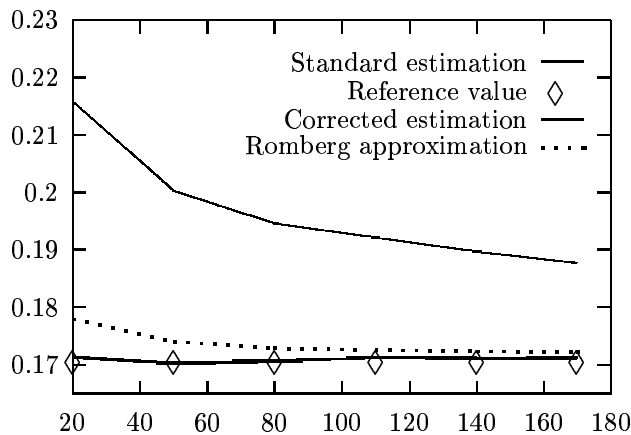


FIG. 4.1 – Résultats de convergence en fonction du nombre de pas de temps N .

celui de la marche aléatoire en dimension 1, étudié dans [SY82] par exemple : dans ce cas-là, le développement limité de l'erreur est valable [11], avec un reste en $o(N^{-1/2})$.

Pour compenser le biais du temps de sortie discret et éliminer la contribution d'ordre $1/\sqrt{N}$, on peut modifier le domaine en décalant la frontière : cette méthode proposée dans [BGK99] en dimension 1 pour le brownien se généralise en dimension supérieure comme nous le montrons.

Théorème 4.2.4 ([11]) *Considérons $X_t = x + \mu t + \sigma W_t$ (avec $\sigma \sigma^*$ inversible et $\mu \in \mathbf{R}^d$) et $D := \{x \in \mathbf{R}^d : x.a < b\}$, $b \in \mathbf{R}$, $a \neq 0$. Supposons aussi que f satisfait aux conditions du Théorème 4.2.2. Alors, en posant $D_N = \{x \in \mathbf{R}^d : x.a < b - c\|\sigma^* a\|\sqrt{\frac{T}{N}}\}$ avec $c = 0.5823(\dots)$ une constante universelle (voir [Sie79]) et $\tau_{D_N}^N = \inf\{t_i : X_{t_i} \notin D_N\}$, nous avons*

$$\mathbf{E}(\mathbf{1}_{T < \tau_{D_N}^N} f(X_T)) - \mathbf{E}(\mathbf{1}_{T < \tau} f(X_T)) = o\left(\frac{1}{\sqrt{N}}\right).$$

Cette méthode est très intéressante en pratique, car elle est simple d'implémentation et rapide ; de plus, elle semble même plus compétitive que l'extrapolation de Romberg comme le montre les résultats numériques sur le graphe 4.1 (tiré de [11]). Il serait tout à fait intéressant pour les applications de généraliser ces résultats à des domaines moins particuliers, notamment des polyèdres.

4.2.2 Processus d'Itô et temps de sortie discret

Dans le cas d'EDS, l'outil de base pour analyser l'impact de discrétisation de temps de sortie est l'EDP satisfaite par l'espérance étudiée : le cadre markovien semble donc important pour l'estimation d'erreur. Néanmoins, on a vu que la vitesse \sqrt{N} provient pour bonne partie de l'estimation de l'overshoot, qui reste en fait d'ordre $1/\sqrt{N}$ même si le processus sous-jacent est un processus d'Itô général. La question qui vient à l'esprit alors est la suivante.

¹L'hypothèse sur le support de f peut être affaiblie en conditions d'annulation si f est régulière.

Dans quelle mesure la propriété markovienne du processus est importante pour assurer que l'approximation discrète de temps de sortie a lieu à la vitesse \sqrt{N} ?

En d'autres termes, pour quelle classe de processus a-t'on

$$\mathbf{E}(\mathbf{1}_{T < \tau_d^N} f(X_T)) - \mathbf{E}(\mathbf{1}_{T < \tau} f(X_T)) = O\left(\frac{1}{\sqrt{N}}\right), \quad (4.8)$$

en notant $\tau_d^N = \inf\{t_i : X_{t_i} \notin D\}$?

Remarquons que (4.8) ne concerne que l'approximation du temps de sortie, à la différence de (4.6) qui prenait en compte aussi l'approximation d'Euler.

Dans [11], nous montrons que les processus d'Itô, vérifiant seulement une hypothèse de frontière caractéristique analogue à (C), satisfont à l'estimation (4.8).

Théorème 4.2.5 ([11]) *Considérons un processus d'Itô sous la forme $X_t = x + \int_0^t b_s ds + \int_0^t \sigma_s dW_s$, avec des coefficients progressivement mesurables et bornés. Supposons*

- Condition de frontière non caractéristique. *Il existe des constantes $a_0 > 0$ et $R > 0$ telles que p.s. pour $s \in [0, T]$ et $d(X_s, \partial D) \leq R$, on a $n(\pi(X_s)) \cdot \sigma_s \sigma_s^* n(\pi(X_s)) \geq a_0$ avec $\pi(x)$ la projection orthogonale de x sur ∂D .*
- Continuité en probabilité de σ . *Pour tout $\delta > 0$, on a p.s. pour $s \in]t_i, t_{i+1}[$ avec $X_s \in \partial D$, $\mathbf{P}(|\int_s^{t_{i+1}} (\sigma_u - \sigma_s) dW_u| \geq \delta \sqrt{t_{i+1} - s} \mid \mathcal{F}_s) \xrightarrow{h \rightarrow 0} 0$ uniformément en ω .*
- *La fonction f est mesurable bornée et s'annule au voisinage de ∂D .*

Alors, l'inégalité (4.8) est satisfaite.

La preuve de ce résultat utilise de manière astucieuse le processus $V_t = \mathbf{E}(\mathbf{1}_{T < \tau_t} f(X_T) \mid \mathcal{F}_t)$ avec $\tau_t = \inf\{s \geq t : X_s \notin D\}$. Nous remarquons que $(V_{t \wedge \tau})_{0 \leq t \leq T}$ est une martingale, avec $V_T = f(X_T)$ et $V_t = 0$ pour $X_t \notin D$, qui sont les homologues pour $v(t, x)$ de l'EDP et des conditions de Cauchy-Dirichlet. La description quantitative de $(V_t)_{0 \leq t \leq T}$ et de $(X_t)_{0 \leq t \leq T}$ au voisinage de ∂D à l'aide de browniens changés de temps conduit au résultat.

4.2.3 EDS et temps de sortie continu

En dehors des méthodes de correction de domaine (Théorème 4.2.4), les approximations avec temps de sortie discret conduisent à la vitesse \sqrt{N} , insatisfaisante dans les applications.

Pour remédier à cela, on peut considérer le schéma d'Euler en temps continu (4.3) avec comme temps de sortie $\tau_c^N = \inf\{t \geq 0 : X_t^N \notin D\}$. Ainsi, on prend en compte les excursions éventuelles de X^N en dehors de D : ce principe, introduit dans [LS89, Bal95], est connu pour donner numériquement des méthodes plus efficaces [ABR96, BDZ97, BCI99]. Nous justifions dans [1] sa convergence théorique, en montrant que nous retrouvons ainsi la vitesse de convergence N valable dans tout l'espace.

Théorème 4.2.6 ([1]) *Pour un domaine de classe C^5 avec une fonction f mesurable bornée s'annulant au voisinage de ∂D , nous avons*

$$\mathbf{E}(\mathbf{1}_{T < \tau_c^N} f(X_T^N)) - \mathbf{E}(\mathbf{1}_{T < \tau} f(X_T)) = \frac{C}{N} + o\left(\frac{1}{N}\right). \quad (4.9)$$

Suivant les idées de [BT96a], il n'est pas très difficile d'étendre le résultat pour des fonctions f quelconques à des distributions, et d'avoir ainsi une estimation analogue à (4.9) sur les densités de transition des processus tués : c'est ce que nous prouvons dans [15].

La simulation effective de l'évènement $\{T < \tau_c^N\}$ demande, une fois obtenues les valeurs $(X_{t_k}^N)_{0 \leq k \leq N}$, de calculer les quantités $p(t_{k+1} - t_k, x_k, x_{k+1}) = \mathbf{P}(\exists t \in [t_k, t_{k+1}] X_t^N \notin D | X_{t_k}^N = x_k, X_{t_{k+1}}^N = x_{k+1})$ qui sont les probabilités de sortie de certains ponts browniens. En dehors de la dimension 1 et du cas du demi-espace en dimension supérieure, $p(t_{k+1} - t_k, x_k, x_{k+1})$ n'est pas explicite.

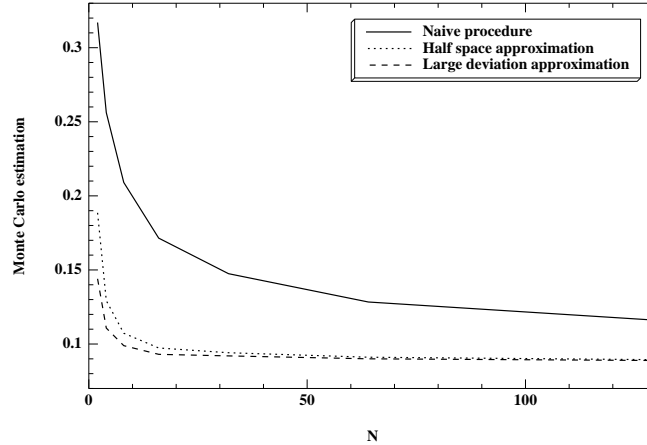


FIG. 4.2 – Comparaison entre temps de sortie discret, estimation de type grande déviation et approximation par demi-espace local, en fonction du nombre de pas de temps.

Baldi [Bal95] donne un développement asymptotique complet de ces quantités lorsque $t_{k+1} - t_k$ tend vers 0, avec des arguments de grande déviation. Le calcul numérique de ces termes est complexe en réalité et nous avons développé dans [4] une autre approche, appelée *approximation par demi-espace local*, a priori un peu moins précise, mais beaucoup plus rapide : sur l'exemple tiré de [4] et considéré sur le graphe 4.2, la précision est comparable et il y a un gain de 10 sur le facteur temps. Pour x_k et x_{k+1} dans D et proches de ∂D , le domaine D dans le calcul de $p(t_{k+1} - t_k, x_k, x_{k+1})$ est remplacé par le demi-espace délimité par l'hyperplan tangent à ∂D au point $\pi(x_k)$, avec à la clé le calcul élémentaire d'une approximation de $p(t_{k+1} - t_k, x_k, x_{k+1})$. Nous avons montré la convergence à la vitesse N de cet algorithme.

Théorème 4.2.7 ([4]) *Sous les hypothèses et notations du Théorème 4.2.6, le schéma d'Euler continu avec approximation par demi-espace local donne une erreur faible d'ordre $\frac{1}{N}$.*

4.2.4 Commentaires sur l'hypothèse sur le support de la fonction f

L'essentiel de nos résultats de vitesse de convergence sont établis sans hypothèse de régularité sur f , mais sous la condition assez restrictive que f s'annule au voisinage de ∂D . En fait, dans ces preuves d'analyse d'erreur, l'usage de l'intégration par parties du calcul de Malliavin pour le schéma d'Euler tué apparaît techniquement nécessaire. Le calcul des variations stochastiques loin

du bord ∂D est possible, au prix de localisations éventuelles, en adaptant des idées de Cattiaux [Cat91]; en revanche, près de ∂D , il est déjà particulièrement astucieux dans le cas d'EDS et nous n'avons pas su l'adapter au cas du schéma d'Euler. L'hypothèse sur f permet d'éviter les intégrations par parties près de ∂D .

4.3 Approximation d'EDS réfléchie dans un domaine

Nous examinons ici le problème d'approximation d'EDS avec réflexion normale ou oblique à la frontière ∂D , en considérant pour le processus $X = (X_t)_{t \geq 0}$ la solution de

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s + \int_0^t \gamma(X_s) dk_s,$$

où k_t est le processus de temps local croissant seulement sur ∂D et γ est la direction de réflexion uniformément non tangentielle à ∂D . Dans la suite, D est régulier à frontière compacte et les coefficients b , σ et γ sont suffisamment réguliers. Les fonctionnelles que nous avons étudiées sont de la forme $f(X_T) - \int_0^T h(X_t) dk_t$.

Dans la littérature, il existe trois principales approches pour l'approximation de X au sens trajectorien ou dans \mathbf{L}_p : schéma d'Euler avec projection introduit dans [Sai87], méthode avec pénalisation [Men83, LS84], schéma d'Euler réfléchi [Slo94]. Peu de travaux traitent en fait de l'approximation faible : une contribution essentielle sur cette question est due à Costantini, Pacchiarotti and Sartoretto [CPS98], qui ont montré que le schéma d'Euler projeté converge faiblement à la vitesse $N^{1/2-\epsilon}$ pour tout $\epsilon > 0$, sous hypothèse de régularité et compatibilité des fonctions f et h .

Nous avons analysé dans [4] et [12] (et annoncé dans [14]) deux nouveaux algorithmes, avec pour objectif d'atteindre la vitesse N de convergence faible (comme dans tout l'espace).

4.3.1 Schéma d'Euler réfléchi avec approximation par demi-espace local

Le principe d'approcher sur chaque intervalle de temps le domaine par un demi-espace est analogue au cas du temps de sortie (Théorème 4.2.7).

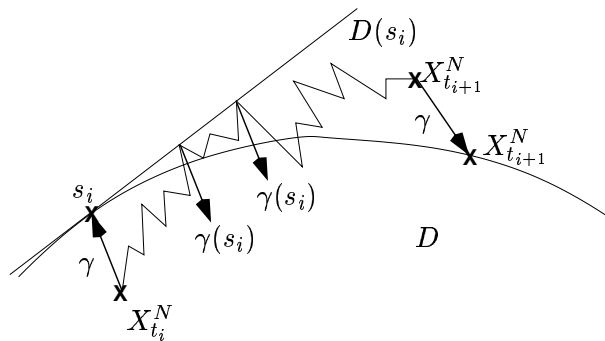


FIG. 4.3 – Schéma d'Euler avec approximation par demi-espace local.

L'algorithme de simulation de $(X_{t_i}^N)_{0 \leq i \leq N}$ peut être décrit ainsi, partant de $X_0^N = X_0$ et $k_0^N = 0$. Supposons $X_{t_i}^N \in D$ et $k_{t_i}^N$ obtenus.

1. Si $X_{t_i}^N$ est loin de ∂D , $X_{t_{i+1}}^N$ est obtenu par (4.2) en ignorant toute réflexion possible.
2. Sinon, on remplace D par le demi-espace $D(s_i)$ délimité par l'hyperplan tangent à ∂D passant par $s_i = \pi^\gamma(X_{t_i}^N)$ (le projeté de $X_{t_i}^N$ parallèlement à γ sur ∂D) et $X_{t_{i+1}}^N, k_{t_{i+1}}^N - k_{t_i}^N$ s'obtiennent comme solution d'un problème de réflexion avec coefficients gelés $b(X_{t_i}^N)$ et $\sigma(X_{t_i}^N)$, dans la direction $\gamma(s_i)$ et dans le demi-espace $D(s_i)$. La solution est explicite comme l'a remarqué Lépingle [Lép93, Lép95]. Si après cela $X_{t_{i+1}}^N \notin D$, il convient en plus de le projeter parallèlement à γ sur ∂D , en modifiant également le temps local.

Une approximation de $f(X_T) - \int_0^T h(X_t) dk_t$ s'en déduit facilement. Nous avons prouvé que cette procédure converge à la vitesse N dans le cas conormal ($\gamma \parallel \sigma\sigma^*n$).

Théorème 4.3.1 ([4]) *Supposons un domaine de classe C^5 et une condition d'uniforme ellipticité pour $\sigma\sigma^*$. Considérons des fonctions f et h satisfaisant l'une des 2 hypothèses suivantes :*

1. *Les fonctions f et h sont régulières et vérifient des conditions de compatibilité de Neumann à la frontière.*
2. *La fonction f est mesurable bornée et s'annule sur un voisinage de ∂D , alors que h est nulle.*

Alors, le schéma d'Euler réfléchi avec approximation par demi-espace local converge faiblement à la vitesse N dans le cas conormal et à la vitesse \sqrt{N} sinon.

4.3.2 Schéma d'Euler symétrisé

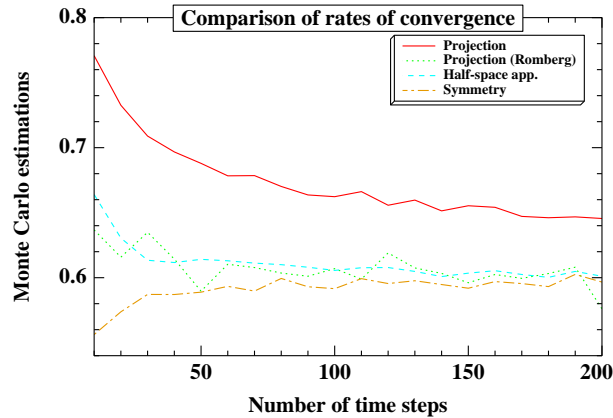


FIG. 4.4 – Le schéma d'Euler avec approximation par demi-espace local et le schéma d'Euler symétrisé améliorent la précision du schéma projeté (même avec extrapolation) .

Les procédés de symétrisation sont assez naturels pour les processus réfléchis. Par exemple, une identité de Lévy exprime que la valeur absolue du mouvement brownien réel a même loi que celle d'un mouvement brownien réfléchi sur le demi-axe positif. Par ailleurs, Freidlin [Fre85] construit la solution d'EDS réfléchies par un procédé de symétrie à la frontière.

Avec Mireille Bossy et Denis Talay, nous nous inspirons de ces anciennes idées pour construire un schéma d'Euler symétrisé, convergeant à la vitesse N . L'étape de construction de $X_{t_{i+1}}^N$ à partir $X_{t_i}^N \in D$ est la suivante :

1. Si $X_{t_i}^N$ est loin de ∂D , la réflexion est négligée et $X_{t_{i+1}}^N$ s'obtient par l'équation (4.2).
2. Sinon, $X_{t_{i+1}}^N$ s'obtient en prenant le symétrique de $X_{t_i}^N$ par rapport à ∂D dans la direction γ : si $s_i = \pi^\gamma(X_{t_i}^N)$, alors $X_{t_{i+1}}^N = s_i + |X_{t_{i+1}}^N - s_i|\gamma(s_i)$.

Le procédé est donc très simple et rapide en pratique, mais il ne permet d'approcher de manière efficace que des fonctionnelles de type $f(X_T)$.

Théorème 4.3.2 ([12]) *Sous des hypothèses analogues au Théorème 4.3.1, nous avons*

$$\mathbf{E}(f(X_T^N)) - \mathbf{E}(f(X_T)) = O\left(\frac{1}{N}\right).$$

Le graphe 4.4 tiré [12] illustre la précision de divers procédés de discrétisation en fonction de N .

Bibliographie

- [ABR96] L. Andersen and R. Brotherton-Ratcliffe. Exact exotics. *Risk*, 9 :85–89, 1996.
- [Als94] G. Alsmeyer. On the Markov renewal theorem. *Stochastic Process. Appl.*, 50(1) :37–56, 1994.
- [Aro67] D.G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bulletin of the American Mathematical society*, 73 :890–903, 1967.
- [Aze83] R. Azencott. Densité des diffusions en temps petit : développements asymptotiques. partie i. *Séminaire de Probabilités XVIII - Lecture Notes in Mathematics 1059*, 1982–83.
- [Bal95] P. Baldi. Exact asymptotics for the probability of exit from a domain and applications to simulation. *The Annals of Probability*, 23(4) :1644–1670, 1995.
- [BCI99] P. Baldi, L. Caramellino, and M.G. Iovino. Pricing general barrier options : a numerical approach using sharp large deviations. *Math. Finance*, 9(4) :293–322, 1999.
- [BCS98] M. Broadie, J. Cvitanic, and M. Soner. Optimal replication of contingent claims under portfolio constraints. *Rev. of Financial Studies*, 11 :59–79, 1998.
- [BDZ97] D.R. Beaglehole, P.H. Dybvig, and G. Zhou. Going to extremes : Correcting simulation bias in exotic option valuation. *Financial Analysts Journal*, 53 :62–68, 1997.
- [Ben88] A. Bensoussan. *Perturbation methods in optimal control*. Wiley/Gauthier-Villars Series in Modern Applied Mathematics. John Wiley & Sons Ltd., Chichester, 1988. Translated from the French by C. Tomson.
- [BGK99] M. Broadie, P. Glasserman, and S. Kou. Connecting discrete and continuous path-dependent options. *Finance and Stochastics*, 3 :55–82, 1999.
- [Bis84] J.M. Bismut. *Large deviations and the Malliavin calculus*. Birkhäuser Boston Inc., Boston, MA, 1984.
- [BL94] P.P. Boyle and S.H. Lau. Bumping up against the barrier with the binomial method. *Journal of Derivatives*, 1 :6–14, 1994.
- [BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.*, 81 :637–654, 1973.
- [BT96a] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations : II. Convergence rate of the density. *Monte-Carlo Methods and Appl.*, 2(2) :93–128, 1996.
- [BT96b] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations : I. Convergence rate of the distribution function. *Probab. Theory Related Fields*, 104-1 :43–60, 1996.
- [Cam53] L. Le Cam. On some asymptotic properties of maximum likelihood estimates and related bayes' estimates. *Univ. California, Publ. Statist.*, 1 :277–330, 1953.
- [Cam60] L. Le Cam. Locally asymptotic normal families of distributions. *Univ. California, Publ. Statist.*, 3 :37–98, 1960.

- [Cam71] L. Le Cam. Limits of experiments. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press*, pages 245–261, 1971.
- [Cat91] P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre - II. Problème de Dirichlet. *Bull. Sc. Math., 2ème série*, 115 :81–122, 1991.
- [CM02] P. Cattiaux and L. Mesnager. Hypocoelliptic non-homogeneous diffusions. *Probab. Theory Related Fields*, 123(4) :453–483, 2002.
- [CPS98] C. Costantini, B. Pacchiarotti, and F. Sartoretto. Numerical approximation for functionals of reflecting diffusion processes. *SIAM J. Appl. Math.*, 58(1) :73–102, 1998.
- [CPT99] J. Cvitanić, H. Pham, and N. Touzi. Super-replication in stochastic volatility models under portfolio constraints. *J. Appl. Probab.*, 36(2) :523–545, 1999.
- [CY00] L. Le Cam and G. Lo Yang. *Asymptotics in Statistics*. Springer Verlag, second edition, 2000.
- [DG95] D. Duffie and P. Glynn. Efficient Monte Carlo simulation of security prices. *Ann. Appl. Probab.*, 5(4) :897–905, 1995.
- [Doh87] G. Dohnal. On estimating the diffusion coefficient. *J. Appl. Prob.*, 24 :105–114, 1987.
- [EJ97] E. Eberlein and J. Jacod. On the range of options prices. *Finance Stoch.*, 1(2) :131–140, 1997.
- [EJL99] K.D. Elworthy, Y. Le Jan, and Xue-Mei Li. *On the geometry of diffusion operators and stochastic flows*. Springer-Verlag, Berlin, 1999.
- [FK97] H. Föllmer and D. Kramkov. Optional decompositions under constraints. *Probab. Theory Related Fields*, 109(1) :1–25, 1997.
- [FL01] C.D. Fuh and T.L. Lai. Asymptotic expansions in multidimensional Markov renewal theory and first passage times for Markov random walks. *Adv. in Appl. Probab.*, 33(3) :652–673, 2001.
- [FLL⁺99] E. Fournié, J.M. Lasry, J. Lebuchoux, P.L. Lions, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3(4) :391–412, 1999.
- [Fre85] M. Freidlin. *Functional integration and partial differential equations*. Annals of Mathematics Studies - Princeton University Press, 1985.
- [Fri75] A. Friedman. *Stochastic differential equations and applications. Vol. 1*. New York - San Francisco - London : Academic Press, a subsidiary of Harcourt Brace Jovanovich, Publishers. XIII, 1975.
- [FZ89] D. Florens-Zmirou. Approximate discrete time schemes for statistics of diffusion processes. *Statistics*, 20 :547–557, 1989.
- [Gar01] D. Garcia. Convergence and biases of monte carlo estimates of american option prices using a parametric exercise rule. Technical report, 2001.
- [GCJ94] V. Genon-Catalot and J. Jacod. Estimation of the diffusion coefficient for diffusion processes : random sampling. *Scandinavian Journal of Statistics*, 21 :193–221, 1994.
- [Gei02] S. Geiss. Quantitative approximation of certain stochastic integrals. *Stoch. Stoch. Rep.*, 73(3-4) :241–270, 2002.
- [Gly86] P.W. Glynn. Stochastic approximation for Monte Carlo optimization. In J. Wilson, J. Henriksen, and S. Roberts, editors, *Proceedings of the 1986 Winter Simulation Conference*, pages 356–365, 1986.
- [Gly87] P.W. Glynn. Likelihood ratio gradient estimation : an overview. In A. Thesen, H. Grant, and W.D. Kelton, editors, *Proceedings of the 1987 Winter Simulation Conference*, pages 366–375, 1987.

- [GRR70] A.M. Garsia, E. Rodemich, and H.Jr. Rumsey. A real variable lemma and the continuity of paths of some gaussian processes. *Indiana University Mathematics Journal*, 20(6) :565–578, 1970.
- [Haj70] J. Hajek. A characterization of limiting distributions of regular estimates. *Z. Wahrsch. Verw. Gebiete*, 14 :323–330, 1970.
- [Haj71] J. Hajek. Local asymptotic minimax and admissibility in estimation. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press*, pages 175–194, 1971.
- [Has80] R.Z. Has'minskii. *Stochastic stability of differential equations*. Sijthoff & Noordhoff, 1980.
- [Hof99] M. Hoffmann. Adaptive estimation in diffusion processes. *Stochastic Process. Appl.*, 79(1) :135–163, 1999.
- [Jac00] J. Jacod. Non-parametric kernel estimation of the coefficient of a diffusion. *Scand. J. Statist.*, 27(1) :83–96, 2000.
- [Jeg82] P. Jeganathan. On the asymptotic theory of estimation when the limit of the loglikelihood is mixed normal. *Sankhya A*, 44 :173–212, 1982.
- [Jeg83] P. Jeganathan. Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. *Sankhya A*, 45 :66–86, 1983.
- [JP98] J. Jacod and P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26(1) :267–307, 1998.
- [Kar81] N. El Karoui. Les aspects probabilistes du contrôle stochastique. In *Ninth Saint Flour Probability Summer School—1979 (Saint Flour, 1979)*, volume 876 of *Lecture Notes in Math.*, pages 73–238. Springer, Berlin, 1981.
- [Kar88] I. Karatzas. On the pricing of American options. *Appl. Math. Optim.*, 17(1) :37–60, 1988.
- [Kes97] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24 :211–229, 1997.
- [KH01] A. Kohatsu-Higa. Weak approximations. A Malliavin calculus approach. *Math. Comp.*, 70(233) :135–172, 2001.
- [KHP02] A. Kohatsu-Higa and R. Pettersson. Variance reduction methods for simulation of densities on Wiener space. *SIAM J. Numer. Anal.*, 40(2) :431–450, 2002.
- [KP91] T.G. Kurtz and P. Protter. Wong-Zakai corrections, random evolutions and numerical schemes for SDEs. in *Stochastic Analysis, Academic Press*, pages 331–346, 1991.
- [KP95] P.E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer Verlag, 1995.
- [KQ95] N. El Karoui and M.C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.*, 33(1) :29–66, 1995.
- [KS87] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus III. *J. Fac. Sci. Univ. Tokyo*, 34 :391–442, 1987.
- [KS98] I. Karatzas and S.E. Shreve. *Methods of mathematical finance*. Springer-Verlag, New York, 1998.
- [KY91] H.J. Kushner and J. Yang. A Monte Carlo method for sensitivity analysis and parametric optimization of nonlinear stochastic systems. *SIAM J. Control Optim.*, 29(5) :1216–1249, 1991.

- [Lam98] D. Lamberton. American options. In D.J. Hand and S.D. Saul, editors, *Statistics in finance*, chapter 10. Arnold Applications of Statistics Series. London : Arnold., 1998.
- [Lép93] D. Lépingle. Un schéma d'Euler pour équations différentielles stochastiques réfléchies. *C. R. Acad. Sci. Paris Sér. I Math.*, 316(6) :601–605, 1993.
- [Lép95] D. Lépingle. Euler scheme for reflected stochastic differential equations. *Math. Comput. Simulation*, 38 :119–126, 1995.
- [Lie96] G.M. Lieberman. *Second order parabolic differential equations*. World Scientific Publishing Co. Inc., River Edge, NJ, 1996.
- [LP94] P. L'Ecuyer and G. Perron. On the convergence rates of IPA and FDC derivative estimators. *Oper. Res.*, 42(4) :643–656, 1994.
- [LS84] P.L. Lions and A.S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, 37 :511–537, 1984.
- [LS89] H.R. Lerche and D. Siegmund. Approximate exit probabilities for a Brownian bridge on a short time interval, and applications. *Adv. in Appl. Probab.*, 21(1) :1–19, 1989.
- [LS01] F. Longstaff and E.S. Schwartz. Valuing american options by simulation : A simple least squares approach. *The Review of Financial Studies*, 14 :113–147, 2001.
- [Men83] J.L. Menaldi. Stochastic variational inequality for reflected diffusion. *Indiana Univ. Math. J.*, 32(5) :733–744, 1983.
- [Mer73] R.C. Merton. Theory of rational option pricing. *Bell J. Econom. and Management Sci.*, 4 :141–183, 1973.
- [Mil74] G.N. Milshtein. Approximate integration of stochastic differential equations. *Theory Probab. Appl.*, 19 :557–562, 1974.
- [Mil76] G.N. Milshtein. A method of second order accuracy integration of SDE. *Theory Probab. Appl.*, 23 :396–401, 1976.
- [Mil85] G.N. Milshtein. Weak approximation of solutions of systems of stochastic differential equations. *Theory Probab. Appl.*, 30 :750–766, 1985.
- [MS76] F. Murat and J. Simon. Etude de problèmes d'optimal design. In *Optim. Tech., Part 2, Proc. 7th IFIP Conf., Nice 1975*, Lect. Notes Comput. Sci. 41, pages 54–62, 1976.
- [Nua95] D. Nualart. *Malliavin calculus and related topics*. Springer Verlag, 1995.
- [NV88] D. Nualart and J. Vives. Absolute continuity of the law of the maximum of a continuous process. *C. R. Acad. Sci., Paris*, 307(7) :349–354, 1988.
- [Pen90] S.G. Peng. A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.*, 28(4) :966–979, 1990.
- [Pha00] H. Pham. Dynamic l^p -hedging in discrete time under cone constraints. *SIAM J. Control Optimization*, 38(3) :665–682, 2000.
- [Pic02] J. Picard. Gradient estimates for some diffusion semigroups. *Probab. Theory Related Fields*, 122 :593–612, 2002.
- [Pir84] P. Pironneau. *Optimal shape design for elliptic systems*. Springer Series in Computational Physics. Springer-Verlag, New York, 1984.
- [PW82] E. Platen and W. Wagner. On a Taylor formula for a class of Itô processes. *Probab. Math. Statist.*, 3(1) :37–51, 1982.
- [Rao83] B.L.S. Prakasa Rao. Asymptotic theory for non-linear least square estimator for diffusion processes. *Math. Operationsforsch. Statist. Ser. Stat*, 14 :195–209, 1983.

- [Rog85] L.C.G. Rogers. Smooth transition densities for one-dimensional diffusions. *Bull. Lond. Math Soc.*, 17 :157–161, 1985.
- [Roo80] H. Rootzen. Limit distributions for the error in approximations of stochastic integrals. *The Annals of Probability*, 8(2) :241–251, 1980.
- [RR91] M. Rubinstein and E. Reiner. Breaking down the barriers. *Risk*, 4-8 :28–35, 1991.
- [RW86] M.I. Reiman and A. Weiss. Sensitivity analysis via likelihood ratios. In J. Wilson, J. Henriksen, and S. Roberts, editors, *Proceedings of the 1986 Winter Simulation Conference*, pages 285–289, 1986.
- [Sai87] Y. Saisho. Stochastic differential equations for multi-dimensional domain with reflecting boundary. *Probab. Theory Related Fields*, 74 :455–477, 1987.
- [Sie79] D. Siegmund. Corrected diffusion approximations in certain random walk problems. *Adv. in Appl. Probab.*, 11(4) :701–719, 1979.
- [Sim80] J. Simon. Differentiation with respect to the domain in boundary value problems. *Numer. Funct. Anal. Optim.*, 2(7-8) :649–687, 1980.
- [Slo94] L. Słomiński. On approximation of solutions of multidimensional SDEs with reflecting boundary conditions. *Stochastic Process. Appl.*, 50(2) :197–219, 1994.
- [SV69a] D.W. Stroock and S.R.S. Varadhan. Diffusion processes with continuous coefficients. I. *Comm. Pure Appl. Math.*, 22 :345–400, 1969.
- [SV69b] D.W. Stroock and S.R.S. Varadhan. Diffusion processes with continuous coefficients. II. *Comm. Pure Appl. Math.*, 22 :479–530, 1969.
- [SV71] D.W. Stroock and S.R.S. Varadhan. Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.*, 24 :147–225, 1971.
- [SY82] D. Siegmund and Y.S. Yuh. Brownian approximations for first passage probabilities. *Z. Wahrsch. verw. Gebiete*, 59 :239–248, 1982.
- [SZ92] J. Sokolowski and J.P. Zolésio. *Introduction to shape optimization*. Springer-Verlag, Berlin, 1992. Shape sensitivity analysis.
- [Tal86] D. Talay. Discrétisation d’une EDS et calcul approché d’espérances de fonctionnelles de la solution. *Math. Modelling Numer. Anal.*, 20 :141–179, 1986.
- [Tem03] E. Temam. Analysis of error with Malliavin calculus : application to hedging. *Mathematical Finance*, 13(1) :201–214, 2003.
- [Tha97] A. Thalmaier. On the differentiation of heat semigroups and Poisson integrals. *Stochastics Stochastics Rep.*, 61(3-4) :297–321, 1997.
- [TL90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8-4 :94–120, 1990.
- [TZ02] D. Talay and Z. Zheng. Approximation of quantiles of components of diffusion processes. Technical report, <http://www.uwm.edu/~ziyu/publication.html>, 2002.
- [Yos92] N. Yoshida. Estimation for diffusion processes for discrete observations. *J. Multivariate Anal.*, 41 :220–242, 1992.
- [Zha99] R. Zhang. *Couverture approchée des options Européennes*. PhD thesis, Ecole Nationale des Ponts et Chaussées, <http://cermics.enpc.fr/theses/99/zhang-ruotao.ps.gz>, 1999.