HAL
open science

Discrétisation de processus stochastiques, estimées de densités et applications

Stephane Menozzi

To cite this version:

Stephane Menozzi. Discrétisation de processus stochastiques, estimées de densités et applications. Mathématiques [math]. Université Paris-Diderot - Paris VII, 2010. tel-00533333

HAL Id: tel-00533333

https://theses.hal.science/tel-00533333v1

Submitted on 5 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Habilitation à Diriger des Recherches en Sciences

Spécialité : **Mathématiques Appliquées**

présentée par

Stéphane Menozzi

à l'Université Paris 7-Denis Diderot

Discrétisation de processus stochastiques, estimées de densité et applications.

Soutenue le 10 Novembre 2010 devant le jury composé de :

Vlad BALLY	Professeur à l'Université de Marne la Vallée-Paris Est
Patrick CATTIAUX	Professeur à l'Université Toulouse 3-Paul Sabatier (Rapporteur)
Josselin GARNIER	Professeur à l'Université Paris 7-Denis Diderot (Rapporteur)
Emmanuel GOBET	Professeur à l'Ecole Polytechnique
Sylvie MÉLÉARD	Professeur à l'Ecole Polytechnique
Gilles PAGÈS	Professeur à l'Université Paris 6-Pierre et Marie Curie
Alexander VERETENNIKOV	Professeur à l'Université de Leeds (Rapporteur)

Présentation du manuscrit

Ce mémoire présente et met en perspective l'ensemble des travaux de recherche effectués pendant ma thèse, de 2001 à 2004, à l'Université Pierre et Marie Curie (Paris VI) et à l'Ecole Polytechnique, sous la direction de Vlad Bally et Emmanuel Gobet, et par la suite depuis mon recrutement en 2005 à l'Université Denis Diderot (Paris VII).

Ces travaux se sont développés autour de deux axes principaux d'intersection non vide : le premier relatif à la discrétisation de processus stochastiques, le second à l'estimation de densité de processus dégénérés et à certaines applications associées.

Les quatre premiers chapitres relèvent majoritairement du premier axe. Le Chapitre 1 est introductif. Nous y présentons les principaux types de problèmes associés à la discrétisation de processus, les résultats existants ainsi que les outils utilisés pour les obtenir. Nous détaillons plus particulièrement les techniques employées pour analyser l'approximation faible.

Le Chapitre 2 regroupe les résultats des articles [**1,3,5,7**]. Nous y étudions principalement l'approximation faible d'un processus de diffusion tué/stoppé à la sortie d'un domaine régulier [**1,7**]. Sont également abordés le cas du cône pour le mouvement Brownien [**3**], ainsi qu'une extension de l'analyse de l'erreur à des processus non Markoviens [**5**].

Au Chapitre 3, nous rappelons les principes de la méthode Parametrix qui permet d'obtenir une représentation de la densité de transition d'une équation différentielle stochastique (EDS) comme somme d'une série où chacun des termes est la convolution de la densité d'une EDS à coefficients constants et des convolutions itérées d'un noyau régularisant. En appliquant cette approche à temps discret nous en déduisons un théorème limite local pour l'approximation par chaîne de Markov d'un processus de diffusion dégénéré [**8**] et un développement d'erreur sur les densités pour une EDS dirigée par un processus de Lévy stable symétrique et approchée par son schéma d'Euler [**9**].

Après une brève présentation des Equations Différentielles Progressives Rétrogrades (EDSR) et de leurs connexions avec certaines équations aux dérivées partielles non linéaires, nous donnerons au Chapitre 4 un aperçu des méthodes numériques développées pour leur approximation. Nous détaillerons plus particulièrement les articles [**2,4**] qui traitent du cas quasi-linéaire (EDSR fortement couplées) qui constitue le cadre le plus général pour les applications en physique ou en finance. Enfin, nous exposerons certaines propriétés de schémas de semi-discrétisation pour des EDSR tuées à la sortie d'un domaine, article [**6**].

Le Chapitre 5 relève lui majoritairement du deuxième axe, i.e. l'estimation de densité. Il débute par une partie introductive destinée à présenter une synthèse des techniques intervenant dans l'estimation quantitative de densités de processus stochastiques ou en termes analytique de solutions fondamentales d'équations aux dérivées partielles. Nous y présentons ensuite les résultats de l'article [**10**]. Nous obtenons des estimées de type Aronson pour la densité de transition d'une équation différentielle stochastique dont la dynamique peut se voir comme celle d'une équation différentielle ordinaire (EDO) dont la première composante serait perturbée par un bruit Brownien. Bien sûr, pour que la densité existe, les coefficients de l'EDO doivent vérifier une condition de non dégénérescence de type Hörmander faible dans la mesure où seule la première composante est non dégénérée. Ces résultats font intervenir contrôle stochastique et techniques analytiques, méthode parametrix présentée au Chapitre 3. Nous concluons le chapitre par une application des estimées de densité : les contrôles non asymptotiques d'intervalles de confiance dans les méthodes de Monte Carlo. Nous montrons pour des schémas d'Euler, éventuellement dégénérés, des estimées d'Aronson et nous utilisons la domination Gaussienne ainsi que des techniques de concentration de la mesure : inégalités de Sobolev logarithmique, argument de Herbst pour arriver au résultat, cf. [**11**].

La note [**12**] établit une condition suffisante pour avoir l'unicité en loi pour les systèmes dégénérés considérés au Chapitre 5 et abordés dans les articles [**10**],[**11**]. Les idées de ce travail sont présentées à l'Appendice A. En particulier, ces résultats apportent des éléments de réponse à une question soulevée par le rapport d'Alexander Veretennikov.

Travaux présentés pour l'habilitation

[**1**] E. Gobet et S. Menozzi."Exact Approximation rate of killed hypoelliptic diffusions using the discrete Euler scheme". *Stochastic Processes and their Applications*, 112-2 (2004), 201-223.

[**2**] F. Delarue et S. Menozzi. "A forward backward algorithm for quasi-linear PDEs". *Annals of Applied Probability*, 16-1 (2006), 140-184.

[**3**] S. Menozzi. "Improved simulation for the killed Brownian motion in a cone". *SIAM Journal on Numerical Analysis*, 44-6 (2006), 2610-2632.

[**4**] E. Gobet et S. Menozzi. "Discrete Sampling of functionals of Itô processes". *Séminaire de Probabilités*, XL (2007), 355-375.

[**5**] F. Delarue et S. Menozzi. "An interpolated Stochastic Algorithm for Quasi-Linear PDEs" *Mathematics of Computation*, 261-77 (2008), 125-158.

[**6**] B. Bouchard and S. Menozzi. "Strong Approximations of BSDEs in a domain". *Bernoulli*, 15-4 (2009), 1117-1147.

[**7**] E. Gobet et S. Menozzi. "Stopped diffusion processes : overshoots and boundary correction". *Stochastic Processes and their Applications*, 120-2 (2010), 130-162.

[**8**] V. Konakov, S. Menozzi et S. Molchanov. "Explicit parametrix and local limit theorems for some de--generate diffusion processes". *Annales de l'Institut Henri Poincaré, Série B*, Vol. 46-4 (2010), 908-923.

[**9**] V. Konakov et S. Menozzi."Weak Error for stable driven SDEs : expansion of the densities". A paraître dans *Journal of Theoretical Probability*, 25 pages.

[**10**] F. Delarue et S. Menozzi. "Density Estimates for a Random Noise Propagating through a Chain of Differential Equations". *Journal of Functional Analysis*, 259-6 (2010), 1577-1630.

[**11**] V. Lemaire et S. Menozzi. "On some Non Asymptotic Bounds for the Euler Scheme". *Electronic Journal of Probability*, Volume 15 (2010), 1645-1681.

Travail non présenté

[**12**] S. Menozzi. "Parametrix techniques and martingale problem for some degenerate Kolmogorov's equations".

Chapitre 1

Discrétisation de processus : un bref tour d'horizon

Dans ce chapitre nous présentons brièvement différents problèmes associés à la discrétisation d'un processus stochastique de "diffusion-sauts" en distiguant tout d'abord approximation forte et faible. Ces approximations sont de nature très différentes, tant en terme des résultats de convergence auxquels elles permettent d'aboutir, que dans les techniques mathématiques utilisées pour les établir.

Nous abordons premièrement le cas de l'*approximation forte* qui est celle associée *trajectoriellement* à la différence entre le processus initial et une approximation donnée. Les contrôles que l'on souhaite obtenir pour cette erreur sont soit presque sûrs soit en norme L^p. Pour les démontrer, on a recours aux outils de l'analyse stochastique : développements d'Itô-Taylor, inégalités de martingales de type Burkholder Davies Gundy (BDG), associés à des arguments de stabilité semblables à ceux utilisés pour l'approximation d'équations différentielles ordinaires.

Nous poursuivons ensuite par l'*approximation faible*. Dans ce cadre, on ne souhaite plus avoir un contrôle trajectoriel, mais proposer un schéma d'approximation pour l'espérance d'une fonctionnelle d'un processus. Cette fonctionnelle peut être par exemple une fonction du processus observé à une date donnée fixée ou à un temps d'arrêt, une fonction du maximum de l'une des composantes ou de l'intégrale du processus sur la période. Le travail fondateur de Talay et Tubaro [TT90] a mis en évidence l'importance fondamentale pour l'analyse de ce type d'erreur dans le cas Markovien, du lien entre la quantité à approcher et l'équation aux dérivées partielles dont cette quantité n'est autre que la représentation probabiliste de la solution, formule de Feynman-Kac.

Nous avons essayé de rassembler quelques résultats majeurs de ces vingt dernières années qui seront utiles à la compréhension des sujets abordés dans les chapitres suivants. Nous insistons particulièrement sur l'approximation faible et les techniques d'étude de l'erreur associée en distinguant les deux types d'hypothèses requises : régularité de la fonctionnelle et des coefficients ou non dégénerescence. En effet, cette thématique sera l'objet d'étude des Chapitres 2 (cas des processus stoppés) et 3 (processus dégénérés et de certains processus à sauts).

Nous conclurons par une section dédiée à quelques autres problèmes de discrétisation : l'*approximation en loi*, on entend par là la caractérisation en terme de dynamique d'une renormalisation convenable de la différence entre le processus initial et le schéma de discrétisation, l'*approximation de mesures invariantes* et de *dynamiques non linéaires* où les contrôles sur l'erreur forte interviennent de nouveau.

1.1 Modèle initial et discrétisation

Soit une équation différentielle stochastique (EDS) à valeurs dans $\mathbb{R}^d$ de dynamique

$$X_t = x + \int_{0^+}^t b(s, X_{s-})ds + \int_{0^+}^t \sigma(s, X_{s-})dZ_s,\ t \geq 0, \tag{1.1}$$

où les coefficients $b : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^r$ sont mesurables bornés en temps, Lipschitziens en espace et où $(Z_t)_{t\geq 0}$ est un processus de Lévy r-dimensionnel de décomposition de Lévy-Khintchine

$$\mathbb{E}[\exp(i\langle u, Z_t\rangle)] = \exp(t\psi(u)),\ \psi(u) = \exp\left(-\frac{|u|^2}{2} + \int_{\mathbb{R}^r} \left(\exp(i\langle u, z\rangle) - 1 - i\langle u, z\rangle \mathbb{I}_{|z|\leq 1}\right) \nu(dz)\right). \tag{1.2}$$

Nous supposons sans perte de généralité que la partie de dérive est nulle dans (1.2) dans la mesure où nous l'avons introduite explicitement dans la dynamique de (1.1). De même, la partie diffusive non normalisée et/ou

anisotrope de Z peut être prise en compte dans le coefficient σ de (1.1). Enfin, la mesure de Lévy ν de Z vérifie la condition d'intégrabilité usuelle $\int_{\mathbb{R}^r}(1 \wedge |z|^2)\nu(dz) < +\infty$.

Ainsi l'équation (1.1) est dirigée par un processus Z qui s'exprime comme la somme d'un mouvement Brownien standard W et d'un processus de saut pur N de mesure de Lévy ν.

L'EDS (1.1) intervient dans de nombreux domaines d'applications. En particulier, des liens profonds existent entre équations différentielles stochastiques et équations aux dérivées partielles (EDP) via les formules de représentation de Feynman Kac, voir Freidlin [Fre85] ou Friedman [Fri75]. Ainsi, dans de nombreux cas détaillés ci-après, cf. Section **1.3**, il est possible d'exprimer la solution d'une EDP linéaire, parabolique ou elliptique, comme espérance d'une fonctionnelle d'un processus de dynamique (1.1).

Dans le cas où le processus directeur $Z = W$, le processus X solution de (1.1) interviendra dans les représentations probabilistes des équations de diffusion de la chaleur en milieu inhomogène ou des équations de la neutronique. Le cas où Z est un processus de saut pur N intervient quant à lui de façon naturelle dans la modélisation du trafic dans des réseaux informatiques ou routiers. Enfin, en mathématiques financières, l'équation (1.1) décrit la dynamique d'un actif dont les variations peuvent faire intervenir composantes browniennes et de sauts, ces dernières intervenant de façon significative dans la modélisation en cette période de crise. Dans de nombreux cas, le prix d'une option sur l'actif X peut là aussi s'exprimer comme l'espérance d'une fonctionnelle de X.

Il est donc naturel d'introduire une procédure d'approximation de (1.1) qui de façon générale ne peut être simulée de façon exacte. Mentionnons à cet égard dans le cas monodimensionnel continu non dégénére, i.e. $d = r = 1$, $Z = W$ et $\sigma \geq \sigma_0 > 0$, le travail de Beskos *et al.* [BPR06] qui permet une simulation exacte.

En toute généralité l'approximation la plus naturelle de (1.1) est celle du schéma d'Euler, qui par ailleurs fournit dans certains cas un procédé constructif de résolution de (1.1), cf. Revuz et Yor [RY99] pour le cas continu. Pour un pas de temps fixé $h > 0$ on pose pour $i \in \mathbb{N}, t_i := ih$, et l'on définit

$$X_t^h = x + \int_0^t b(\phi(s), X_{\phi(s)}^h)ds + \int_0^t \sigma(\phi(s), X_{\phi(s)}^h)dZ_s,\ t \geq 0, \tag{1.3}$$

où $\phi(s) = \sup\{t_i : t_i \leq s\}$. En d'autres termes, pour une grille temporelle donnée $(t_i)_{i\in\mathbb{N}^*}$, le schéma d'Euler est obtenu en gelant sur chaque pas de temps la dynamique spatiale et temporelle aux points de discrétisation en temps. Précisons que dans le cas où $Z = W + N$ possède une composante de sauts, le schéma d'Euler n'est pas toujours exactement simulable, cf. Section **1.3.2**. Il le reste en revanche pour de nombreux cas importants pour les applications : si N est un processus stable, Γ, ...

Pour la suite nous allons distinguer deux grands types de résultats d'approximation qui font intervenir $(X_t)_{t\geq 0}$ solution de (1.1) et son approximation $(X_t^h)_{t\geq 0}$ solution de (1.3). Historiquement, l'approximation de processus remonte à Stroock et Varadhan qui à la suite de l'étude des problèmes de martingales s'étaient intéréssés à des approximations en loi des processus de diffusions avec des contrôles "à la Donsker", i.e. convergence sans vitesse de l'approximation. Pour un aperçu de cette approche, voir le Chapitre 11 de [SV79] pour les approximations dans tout l'espace ou l'article [SV71] pour les problèmes de Dirichlet.

Il a fallu attendre les travaux de Milstein [Mil74, Mil76] pour que des approximations de type Itô-Taylor permettent de quantifier plus précisément l'erreur forte trajectorielle. C'est ensuite dans les années 80 que Milstein [Mil85] et Talay [Tal86] ont montré que l'erreur faible était d'ordre h sous de bonnes hypothèses.

1.2 Erreur forte

On désigne par erreur forte associée au schéma d'Euler la quantité $\sup_{t\in[0,T]} |X_t - X_t^h|$ pour $T > 0$ fixé.

Dans le cas d'un processus de diffusion, si l'on suppose que les coefficients b, σ dans (1.1) sont uniformément Lipshitz continus en espace et 1/2-Hölder en temps, des développements d'Itô-Taylor permettent de déduire que pour tout $p \geq 1$, $\exists C := C(p, T, b, \sigma)$,

$$\mathbb{E}[\sup_{t\in[0,T]} |X_t - X_t^h|^p]^{1/p} \leq C\sqrt{h}, \tag{1.4}$$

où la dépendance en b, σ dans la constante C ci-dessus intervient via les modules de continuité en temps et espace des coefficients. Par rapport à l'analyse de la discrétisation d'une EDO, ce sont les mêmes arguments de stabilité qui sont utilisés avec une étape supplémentaire pour contrôler les intégrales stochastiques à l'aide des inégalités BDG. C'est lors de cette étape qu'apparaît l'ordre de l'erreur en $h^{1/2}$ qui correspond à la norme L^p de l'accroissement Brownien sur un pas de temps. L'utilisation du Lemme de Gronwall pour conclure nous indique

ensuite qu'en toute généralité ces estimations ne seront pertinentes que pour un horizon de temps $T > 0$ fini, du fait de la dépendance exponentielle de C en ce paramètre. Un corollaire immédiat de (1.4) et du lemme de Borel Cantelli est la convergence presque sûre de $h^{-\alpha} \sup_{t\in[0,T]} |X_t - X_t^h|$ vers 0 avec h pour tout $\alpha < 1/2$.

C'est pour compenser cette vitesse que Milstein a introduit son schéma [Mil76] qui permet de retrouver la vitesse usuelle en h. En revanche, ce schéma est difficilement simulable en dimension $d > 1$, sans condition supplémentaire de commutation sur les colonnes de la matrice de diffusion identifiées à des champs de vecteurs. Il existe de nombreux schémas d'ordre élévés pour l'erreur forte. Nous renvoyons à l'ouvrage de Kloeden et Platten [KP92] pour un panorama de ces schémas. Mentionnons toutefois que leur simulation reste beaucoup plus complexe que celle du schéma d'Euler et que ces schémas ne sont pas forcément les plus appropriés dans de nombreux cas, voir Section **1.3**.

Des arguments similaires peuvent être utilisés pour caractériser l'erreur forte, avec la même vitesse, lorsque Z est un processus de Lévy de carré intégrable, voir Kohatsu-Higa et Protter [KHP94]. Mentionnons enfin les travaux de Bruti Liberati et Platten [BLP07] relatifs à la construction de schémas d'ordre plus élévé pour l'erreur forte dans le cas mixte $Z = W + N$ ou de saut pur $Z = N$ lorsque l'intensité du processus de saut est finie, i.e. $\nu(\mathbb{R}^r) < +\infty$ dans (1.2). Ces auteurs utilisent en particulier l'idée d'une grille de discrétisation temporelle adaptée aux instants de saut.

1.3 Erreur faible

Dans de nombreuses applications : approximation probabiliste de la solution d'une EDP linéaire, valorisation ou couverture d'option en finance, il n'est toutefois pas nécéssaire d'avoir un contrôle trajectoriel sur l'approximation du processus. En effet, dans ces cas là les quantités à estimer sont de la forme $\mathbb{E}[\Psi\left((X_t)_{t\geq 0}\right)]$, où $\Psi : C(\mathbb{R}^+, \mathbb{R}^d) \to \mathbb{R}$ est une fonctionnelle de la trajectoire du processus. La représentation en terme d'espérance de la quantité à estimer tend naturellement à faire envisager une méthode de Monte Carlo pour l'approximation sous réserve d'intégrabilité. Par ailleurs, outre le problème déjà évoqué de l'impossibilité de simuler de façon exacte la diffusion, s'ajoute ici celui de l'approximation de la fonctionnelle Ψ. Si l'on désigne par Ψ^h une approximation de Ψ et par $\left((X_t^{h,i})_{t\geq 0}\right)_{i\in\mathbb{N}^*}$ une suite i.i.d. de trajectoires du schéma d'Euler, l'estimateur de Monte Carlo naturel s'écrit pour tout $M \geq 1$, $\Theta_M^h := \frac{1}{M}\sum_{i=1}^M \Psi^h\left((X_t^{h,i})_{t\geq 0}\right)$. Ainsi, l'erreur globale entre la quantité explicitement simulable sur ordinateur et la quantité exacte se décompose comme :

$$\begin{aligned} \mathcal{E}_M^h &= \left(\Theta_M^h - \mathbb{E}[\Psi^h\left((X_t^h)_{t\geq 0}\right)]\right) + \left(\mathbb{E}[\Psi^h\left((X_t^h)_{t\geq 0}\right)] - \mathbb{E}[\Psi\left((X_t)_{t\geq 0}\right)]\right) \\ &:= \mathcal{E}_{M,h}^S + \mathcal{E}_h^D. \end{aligned} \tag{1.5}$$

Le terme $\mathcal{E}_{M,h}^S$ est un terme d'*erreur statistique*. Si $\sup_{h>0} \Psi^h\left((X_t^h)_{t\geq 0}\right) \in L^2(\mathbb{P})$, ce terme sera contrôlé par le théorème central limite à l'ordre $\sigma^h/\sqrt{M}$ avec forte probabilité, où l'écart-type $\sigma^h := \mathrm{Var}(\Psi^h\left((X_t^h)_{t\geq 0}\right))^{1/2}$ est en général assez stable, ou tout du moins uniformément dominé, en h. Nous reviendrons au Chapitre 5.3 sur les contrôles non asymptotiques d'intervalles de confiance.

L'étude de l'*erreur faible* consiste à analyser le comportement du terme $\mathcal{E}_h^D$. Ce terme correpond à l'*erreur de discrétisation* : du processus, et de la fonctionnelle. Pour évaluer la pertinence de l'approximation choisie, ou mettre en évidence ses limites comme pour le problème de Dirichlet, cf. Section 1.3.3 et Chapitre 2, la première étape consiste à quantifier en fonction de h la convergence de $\mathcal{E}_h^D$ vers 0. Le découpage (1.5) permet ensuite de déterminer la précision globale de l'approximation numérique. Il est alors possible d'équilibrer de façon optimale les paramètres M de Monte Carlo et h de discrétisation de sorte à atteindre un seuil fixé.

Si l'on arrive par ailleurs à obtenir un développement asympotique de $\mathcal{E}_h^D$ en fonction de h, il est alors possible, à complexité équivalente, d'utiliser des techniques usuelles et purement algébriques d'extrapolation pour accélérer de façon significative la convergence de $\mathcal{E}_h^D$. Cette démarche, connue sous le nom d'extrapolation de Richardson-Romberg est présentée en détails dans l'article [TT90]. Nous la rappelons en Section 1.3.1.

Dans la suite de ce chapitre nous allons rassembler des résultats relatifs à trois grands types de fonctionnelles :

1. $\psi\left((X_t)_{t\geq 0}\right) = g(X_T)$ pour $T > 0$ fixé et une fonction g donnée.
2. $\psi\left((X_t)_{t\geq 0}\right) = g(\tau \wedge T, X_{\tau\wedge T})$, $\tau := \inf\{t \geq 0 : X_t \notin D\}$, pour D domaine de $\mathbb{R}^d$.
3. $\psi\left((X_t)_{t\geq 0}\right) = g(T^{-1}\int_0^T X_t dt, X_T)$.

La technique eprouvée dans le cas Markovien pour l'analyse de l'erreur $\mathcal{E}_h^D$, consiste à identifier la quantité à estimer $\mathbb{E}[\Psi\left((X_t)_{t\geq 0}\right)]$ à la solution d'une EDP linéaire avec conditions aux limites appropriées en fonction du problème considéré. Dans le cas 1. ci-dessus nous serons dans le cadre du problème de Cauchy. Dans le cas 2. du problème de Cauchy-Dirichlet et dans le cas 3. toujours dans le cadre du problème de Cauchy mais pour

un opérateur dégénéré. L'intégrale est vue comme une variable supplémentaire à laquelle ne sera pas associé de terme d'ordre 2 dans le générateur infinitésimal du couple $(X_t, Y_t)_{t\geq 0}, Y_t := \int_0^t X_s ds$.

1.3.1 Problème de Cauchy : cas continu

Nous considérons ici d'abord le cas où X dans (1.1) est un processus de diffusion à coefficients réguliers, i.e. b, σ sont supposés C^∞ et à dérivées bornées et $\psi\left((X_t)_{t\geq 0}\right) = g(X_T)$ où g est une fonction C^∞ à croissance polynomiale. L'idée originale des articles [Tal86] et [TT90], consiste tout d'abord à observer que sous les hypothèses précédentes, la fonction $u(t,x) := \mathbb{E}[g(X_T^{t,x})] := \mathbb{E}[g(X_T)|X_t = x]$ définie pour $(t,x) \in [0,T]\times\mathbb{R}^d$ satisfait l'EDP parabolique

$$\begin{cases} (\partial_t + L_t)u(t,x) = 0, \ (t,x) \in [0,T)\times\mathbb{R}^d, \\ u(T,x) = g(x), \end{cases} \tag{1.6}$$

où $L_t\varphi(x)$ désigne le générateur infinitésimal de (1.1). Pour $\varphi \in C_0^\infty(\mathbb{R}^d)$, $L_t\varphi(x) = \langle b(t,x), \nabla_x\varphi(x)\rangle + \frac{1}{2}\mathrm{tr}(a(t,x)D_x^2\varphi(x)),\ a(t,x) := \sigma\sigma^*(t,x)$. Ces hypothèses permettent ensuite de déduire des contrôles sur les dérivées de u à l'aide d'estimées L^p sur les dérivées du flot de (1.1). Précisément, pour tout multi-indice de dérivation α, il existe $k_\alpha(T), C_\alpha(T) > 0$ tels que pour tout $(t,x) \in [0,T]\times\mathbb{R}^d$, $|\partial_\alpha u(t,x)| \leq C_{\alpha(T)}(1+|x|^{k_\alpha(T)})$. C'est à dire que les dérivées gardent une croissance polynomiale en espace, ce qui permet en particulier de déduire des contrôles usuels sur le schéma d'Euler que pour tout $p \geq 1, i \in [\![0,N]\!], T := Nh,\ |\partial_\alpha u(t_i, X_{t_i}^h)| \in L^p(\mathbb{P})$. L'étape suivante consiste à utiliser (1.6) pour réécrire l'erreur de discrétisation comme :

$$\mathcal{E}_h^D = \mathbb{E}_x[g(X_T^h)] - \mathbb{E}_x[g(X_T)] = \mathbb{E}_x[u(T,X_T^h)] - u(0,x) = \sum_{i=0}^{N-1}\mathbb{E}[u(t_{i+1}, X_{t_{i+1}}^h) - u(t_i, X_{t_i}^h)].$$

A partir de cette dernière formulation, l'idée est ensuite d'appliquer un développement de Taylor en temps-espace autour de $(t_i, X_{t_i}^h)$ pour chaque $i \in [\![0, N-1]\!]$ grâce à la régularité *a priori* de u. En exploitant que u satisfait (1.6), on obtient que les termes d'ordre h s'annulent dans $\mathbb{E}[u(t_{i+1}, X_{t_{i+1}}^h) - u(t_i, X_{t_i}^h)] := \mathcal{E}_h^{D,i}$. Pour établir ensuite que $\mathcal{E}_h^{D,i} = O(h^2)$ il est nécessaire de poursuivre le développement à l'ordre deux en temps et quatre en espace, en exploitant que l'on peut éliminer les moments d'ordre 3 par centrage. Ainsi, si l'on a *a priori* que $u \in C^{2,4}([0,T]\times\mathbb{R}^d)$, on déduit que $\mathcal{E}_D^h = O(h)$ par simple sommation.

Notons que dans la procédure décrite ci-avant, il n'est pas nécessaire d'avoir un aléa Gaussien dans le schéma d'Euler (1.3). En effet, le développement ne fait intervenir que les propriétés de centrage des moments d'ordre 1 et 3 et que la covariance de l'aléa est la matrice identité. Cette dernière observation permet de gagner en complexité en utilisant par exemple des lois de Bernoulli convenablement normalisées en lieu et place des accroissements Browniens.

Si l'on poursuit maintenant les développements à l'ordre 3 en temps et 6 en espace, en tuant toujours par centrage les moments impairs des accroissements, on aboutit au développement d'erreur :

$$\mathcal{E}_h^D = Ch + O(h^2), \tag{1.7}$$

où la constante $C := C(T,a,b,u,x)$, cf. Théorème 1 de [TT90] pour une expression explicite. Du développement précédent il est aisé de voir que si l'on considère le schéma d'Euler de pas $h/2$, il vient : $\mathcal{E}_{h/2}^D = C\frac{h}{2} + O(h^2)$. Une simple manipulation algébrique donne ensuite

$$2\mathcal{E}_{h/2}^D - \mathcal{E}_h^D = \left\{2\mathbb{E}_x[g(X_T^{h/2})] - \mathbb{E}_x[g(X_T^h)]\right\} - \mathbb{E}_x[g(X_T)] = O(h^2). \tag{1.8}$$

Cette identité suggère d'utiliser l'estimateur de Monte Carlo $\hat{\Theta}_M^h = \frac{1}{M}\left\{\sum_{i=1}^M \left(2g(X_T^{h/2,i}) - g(X_T^{h,i})\right)\right\}$, où les $(X_T^{h/2,i}, X_T^{h,i})_{i\in[\![1,M]\!]}$ sont des réalisations i.i.d. et pour tout $i \in [\![1,M]\!]$ les $X_T^{h/2,i}, X_T^{h,i}$ sont simulés à partir des mêmes accroissements Browniens. Ce procédé est connu sous le nom d'extrapolation de Richardson-Romberg. Notons que sous les hypothèses précédentes, la régularité de u est suffisante pour obtenir un développement asymptotique à l'ordre n, i.e. $\mathcal{E}_h^d = \sum_{i=1}^n C_i h^i + O(h^{n+1})$ à partir duquel il serait possible d'itérer le procédé d'extrapolation et d'obtenir une vitesse d'ordre h^{n+1} en multipliant la complexité par $n+1$. Mentionnons toutefois que cette procédure peut accroître la variance de l'estimateur $\hat{\Theta}_M^h$. On considère en effet $n+1$ fois plus de variables aléatoires.

L'analyse de l'erreur faible s'effectuera toujours suivant le schéma présenté ci-dessus, avec les spécificités associées à chaque problème, voir Sections 1.3.2 et 1.3.3. Le point clé est d'avoir de la régularité sur la fonction u et des contrôles sur ses dérivées, afin d'effectuer des développements d'Itô-Taylor. Sous les hypothèses

précédentes, ces contrôles s'obtiennent par des techniques usuelles de calcul stochastique. On a cependant besoin pour cela de beaucoup de régularité, sur b, σ et g. Or, dans de nombreux cas cette régularité n'est pas nécessaire. C'est en particulier vrai si l'on peut bénéficier de l'effet régularisant des noyaux de la chaleur. En d'autres termes, il est possible de relaxer la régularité de la fonction g si par exemple X_T admet une densité de transition régulière dont on contrôle bien les dérivées. En effet dans ce cas $u(t,x) = \int_{\mathbb{R}^d} g(y)p(t,T,x,y)dy$ où $p(t,T,x,.)$ désigne la densité de $X_T^{t,x}$. Dans le cas de coefficients homogènes en temps nous noterons pour simplifier l'écriture $p(t,T,x,.) = p(T-t,x,.)$.

Néanmoins, prouver que $X_T^{t,x}$ possède une telle densité nécéssite une hypothèse de non-dégénérescence de l'équation (1.1). On considère deux jeux d'hypothèses qui garantissent que l'on obtient les propriétés souhaitées : l'hypoellipticité à la Hörmander (hypothèse **(H)**) et l'uniforme ellipticité de la matrice de diffusion a (hypothèse **(UE)**).

Hypothèse (H) : Supposons que les coefficients b, σ sont C^∞, bornés, homogènes en temps et réécrivons $dX_t = A_0(X_t)dt + \sum_{i=1}^r A_i(X_t) \circ dW_t^i$ (i.e. les champs de vecteurs $(A_i)_{i\in[\![1,r]\!]}$ correspondent aux colonnes de la matrice de diffusion, $A_0 = b - \frac{1}{2}\sigma.\nabla\sigma$ et $\circ dW$ désigne l'intégrale de Stratonovitch). Si l'espace vectoriel engendré par

$$A_1, \cdots, A_r,\ [A_i, A_j]_{0\le i,j\le r},\ [A_i,[A_j,A_k]]_{0\le i,j,k\le r}, \cdots$$

au point initial x est $\mathbb{R}^d$ (où $[.,.]$ désigne le crochet de Lie entre champs de vecteurs), alors pour tout $t > 0$, $X_t^{0,x}$ admet une densité, voir par exemple le Chapitre 2.3 de Nualart [Nua95].

Si l'on suppose de plus qu'il existe $L \in \mathbb{N}^*$ tel que que pour tout $x \in \mathbb{R}^d$ l'espace $\mathbb{R}^d$ est engendré en ne considérant que les crochets de longueur inférieure ou égale à L, alors les travaux de Kusuoka et Stroock [KS85], donnent les contrôles suivants sur la densité. Soient m, k des entiers et des multi-indices α, β tels que $2m + |\alpha| \le k$. Il existe des constantes $C(T), c, q, Q$ dépendant de L, m, k, α, a, b, telles que

$$|\partial_t^m \partial_x^\alpha p(t,x,y)| \le \frac{C(T)(1+|x|^Q)}{t^q(1+|y-x|)^k} \exp\left(-c\frac{|x-y|^2}{t(1+|x|)^2}\right). \tag{1.9}$$

Par ailleurs, l'inverse de la matrice de Malliavin que nous noterons $\Gamma_t(x)$ est contrôlée dans tous les L^p. Pour tout $p \ge 1$ il existe $C(T), \mu$ t.q.

$$\mathbb{E}[|\Gamma_t(x)|^p]^{1/p} \le C(T)\frac{1+|x|^\mu}{t^{dL}}. \tag{1.10}$$

Le propos du travail de Bally et Talay [BT96a] est d'exploiter les contrôles (1.9), (1.10) pour étendre le développement d'erreur (1.7) au cas d'une diffusion hypoelliptique sans régularité *a priori* sur g, que l'on peut ici supposer mesurable et à croissance exponentielle.

Rappelons que lors des développements d'Itô-Taylor précédents apparaissaient pour tout $i \in [\![0, N-1]\!]$, des quantités du type $Q_i^h := \mathbb{E}[\partial_x^\alpha u(t_i, X_{t_i}^h) G(t_i, X_{t_i}^h)]$, où G fait intervenir les coefficients b, σ et leurs dérivées. Le contrôle sur la densité (1.9) permet de contrôler les Q_i^h pour $i \le N/2$. Pour les $i > N/2$, l'idée naturelle est d'utiliser une intégration par parties de Malliavin. Notons $Q_i = \mathbb{E}[\partial_x^\alpha u(t_i, X_{t_i}) G(t_i, X_{t_i})]$, i.e. on remplace $X_{t_i}^h$ par X_{t_i}. L'intégration par parties pour Q_i s'écrit $Q_i = \mathbb{E}[u(t_i, X_{t_i}) H_\alpha(X_{t_i}, G(t_i, X_{t_i}))]$. Les inégalités de Meyer permettent alors d'obtenir qu'il existe $p \ge 1$ tel que $|Q_i| \le C\mathbb{E}[|\Gamma_{t_i}(x)|^p]^{1/p}$, où C dépend de normes des dérivées de Malliavin de X_{t_i}.

On pourra faire l'intégration par parties dans Q_i^h sur les événements de forte probabilité pour lesquels l'inverse de la matrice du schéma d'Euler $X_{t_i}^h$ est "proche" de l'inverse de la matrice de Malliavin $\Gamma_{t_i}(x)$ de la diffusion (1.1) et contrôler la probabilité des événements complémentaires, cf. [BT96a].

L'idée sous jacente globale est que l'hypothèse d'hypoellipticité permet de contrôler les dérivées en utilisant en temps petit les estimées ponctuelles sur la densité qui sont pertinentes dans la mesure où $u(t,x) = \int_{\mathbb{R}} g(y)p(T-t,x,y)dy$, puis d'exploiter qu'en temps plus long c'est la matrice inverse $(\Gamma_t)_{t\ge T/2}$ qui est non-dégénérée. Quant à la croissance de g, c'est le contrôle (1.9) qui rend cette condition suffisante.

Mentionnons que si la non-dégénérescence de l'inverse de la matrice de diffusion est bien contrôlée dans le cas de la diffusion X, il n'en est pas de même pour son schéma d'Euler X^h. C'est ce qui a motivé l'introduction des Q_i ci-dessus pour présenter l'intégration par parties. Il n'y a à ma connaissance pas de preuve d'existence de la densité pour le schéma d'Euler hypoelliptique. C'est d'ailleurs ce dernier point qui a amené Bally et Talay à considérer un schéma d'Euler légèrement perturbé pour obtenir un développement d'erreur sur les densités, cf. [BT96b].

Hypothèse (UE) : On suppose la matrice de diffusion a uniformément elliptique et bornée, i.e. $\exists \Lambda \ge 1,\ \forall (t,x) \in [0,T] \times \mathbb{R}^d,\ \forall \xi \in \mathbb{R}^d,\ \Lambda^{-1}|\xi|^2 \le \langle a(t,x)\xi, \xi\rangle \le \Lambda|\xi|^2$. Pour b et σ Lipschitziens en espace et b borné, il est bien connu que l'unique solution forte de (1.1) admet une densité vérifiant une estimée supérieure Gaussienne, i.e.

$\forall T > 0,\ \exists (C,c) := (C,c)(T,|b|_\infty,\Lambda),\ \forall (t,x,y) \in (0,T] \times \mathbb{R}^d \times \mathbb{R}^d, p(t,x,y) \leq \frac{C}{t^{d/2}} \exp(-c\frac{|x-y|^2}{t})$. Ce contrôle s'obtient aisément à partir de la représentation en série parametrix de la densité, voir Friedman [Fri64] ou McKean et Singer [MS67]. L'idée de ce type de méthode est d'écrire :

$$p(t,x,y) = \tilde{p}(t,x,y) + \sum_{i=1}^{+\infty} \tilde{p} \otimes H^{(i)}(t,x,y). \tag{1.11}$$

où $\tilde{p}(t,x,y)$ est une densité Gaussienne, dans laquelle on a gelé les cofficients de (1.1) au point d'arrivée y et où les termes $\tilde{p} \otimes H^{(i)}$ sont des convolutions temps-espace de densités Gaussiennes gelées et de noyaux $H^{(i)}$ toujours plus régularisants en temps avec i. En d'autres termes, en temps petit, la contribution des $\tilde{p} \otimes H^{(i)}$ est de plus en plus négligeable avec i. Nous reviendrons en détail sur la méthode parametrix au Chapitre 3.

L'approche développée dans [MS67] s'est en fait révelée très adaptée aux discrétisations. Considérons le schéma d'Euler généralisé de (1.1) où l'on remplace les accroissements Browniens $(W_{t_i} - W_{t_{i-1}})_{i \in [\![1,N]\!]}$ dans (1.3) par les réalisations i.i.d. $(\xi_i^h)_{i \in [\![1,N]\!]}$ d'une loi à densité plus générale ayant mêmes moments que l'accroissement Brownien jusqu'à un certain ordre. Si ce dernier admet une densité p^N, il est possible de la représenter aux instants de discrétisation sous la forme

$$p^N(t,x,y) = \tilde{p}^N(t,x,y) + \sum_{i=1}^{+\infty} \tilde{p}^N \otimes_N H^{N,(i)}(t,x,y). \tag{1.12}$$

Dans l'équation ci-dessus $\tilde{p}^N$ designe la densité de la somme convenablement normalisée des ξ_i^h, $\otimes_N$ désigne une convolution temps-espace discrète en temps et $H^{N,(i)}$ une version discrète des noyaux régularisants de (1.11). Cette écriture et une analyse de la différence entre les représentations (1.11) et (1.12) a permis à Konakov et Mammen [KM00] d'obtenir un théorème limite local à l'ordre $h^{1/2}$ pour la différence $(p - p^N)(T,x,y)$. Cette vitesse correspond à la vitesse d'approximation classique du théorème limite local en régime Gaussien, cf. Bhattacharya et Rao [BR76]. C'est celle qui contrôle la convergence de $\tilde{p}^N$ vers $\tilde{p}$. Le résultat entre p et p^N découle ensuite d'une longue analyse de type stabilité. Nous présenterons au Chapitre 3 une extension de ce résultat pour certains processus dégénérés, article **[8]**.

Dans le cas particulier du schéma d'Euler (1.3), i.e. lorsque les $(\xi_i^h)_{i \in [\![1,N]\!]}$ sont déjà Gaussiens, on retrouve la vitesse d'ordre h comme dans le cas de l'erreur faible mais sur les densités [KM02] et leur dérivées, cf. [Guy06] :

$$\begin{aligned} (p - p^N)(1,x,y) &= \pi(1,x,y)h + r(1,x,y)h^2, \\ \exists (c,C) := (c,C)(1,|b|_\infty,\Lambda),\ \forall (x,y),\ (|\pi| + |r|)(1,x,y) &\leq C \exp\left(-c|x-y|^2\right). \end{aligned} \tag{1.13}$$

Sous des hypothèses plus fortes que celles courantes de régularité sur les coefficients, Gobet et Labart [GL08] ont précisément quantifié l'ordre du premier terme pour T éventuellement petit. Ils ont établi $|p - p^N|(T,x,y) \leq ChT^{-(d+1)/2} \exp(-c|x-y|^2/T)$, où les constantes C, c sont uniformes sur un compact en temps. C'est l'utilisation du calcul de Malliavin qui impose une certaine régularité sur les coefficients (b, σ supposés $C_b^{1,3}$ et $\partial_t \sigma \in C_b^{0,1}$). L'ordre de la singularité en temps du terme de reste est néanmoins intuitivement clair : les développements parametrix (1.11) et (1.12) font intervenir dans les noyaux H et H^N des termes d'ordre deux en espace. Lorsque l'on fait la différence de ses deux développements on a une sensibilité qui fait donc intervenir des termes d'ordre trois. Les contrôles usuels sur les dérivées des noyaux de la chaleurs elliptiques, cf. Friedman [Fri64] donnent que ces termes auront une singularité d'ordre 3/2 en temps. Celle-ci doit néanmoins être intégrée en temps dans la convolution ce qui, après intégrations par parties, donne bien une singularité globale additionnelle en $T^{-1/2}$ correspondant au résultat de Gobet et Labart. Le même résultat peut se retrouver à partir des travaux de Konakov et Mammen [KM09] qui précisent les résultats de [KM00] pour T éventuellement petit toujours pour une chaîne de Markov générale approchant (1.1).

L'avantage des développements de type (1.13) pour l'étude de l'erreur faible, par rapport aux approches précédentes, est de pouvoir précisément relâcher les hypothèses sur la fonction g. Guyon montre en particulier que le développement (1.7) reste valable si g est une distribution tempérée, incluant en particulier les masses de Dirac et les dérivées de masse de Dirac. Ce dernier point est en particulier important pour les applications au calcul de sensibilité, ou d'estimation des dérivées de l'EDP (1.6), qui sont cruciales en finance.

Enfin, l'approche parametrix de [KM00], [KM09] pour l'approximation permettrait d'étendre les résultats mentionnés à des équations à coefficient de diffusion a uniformément Hölder en espace, borné, mesurable en temps et dérive b bornée dans le cadre du problème de martingale classique. On retrouverait de la sorte les hypothèses naturelles en analyse pour obtenir les contrôles sur les noyaux de la chaleur associés à des opérateurs sous forme non divergence elliptiques, cf. Friedman [Fri64] ou Sheu [She91] pour une approche probabiliste de ces résultats.

En conclusion de cette présentation, notons que l'approche originelle de Talay et Tubaro ne requiert pas de condition de non-dégénérescence. Elle est robuste dès lors que l'on a de la régularité. Pour se passer de régularité sur g, une hypothèse de non dégénérescence (hypoellipticité ou uniforme ellipticité) est nécessaire. Cela permet de faire porter les dérivations spatiales intervenant lors de l'analyse de l'erreur, sur la densité du processus que l'on sait alors contrôler.

Mentionnons également la recherche active ces dernières années sur les schémas d'erreur faible élevée. Ces derniers sont issus des travaux de Kusuoka [Kus01] dont le schéma a ensuite été développé par Lyons et Victoir [LV04] et Ninomiya *et al.* [NV08], [NN09]. L'idée est toujours de procéder à des développements d'Itô-Taylor dans les intégrales stochastiques et de remplacer ensuite les intégrales stochastiques itérées apparaissant par des variables aléatoires ayant les mêmes moments jusqu'à un certain ordre en utilisant des formules de quadrature pour ces intégrales Browniennes multiples. Le gros du travail est ensuite de déterminer des bonnes fonctions de quadrature, ce qui requiert un travail algébrique conséquent.

Ces schémas imposent néanmoins beaucoup de régularité sur les coefficients b, σ ce qui n'est pas *a priori* nécessaire pour avoir un développement (1.7) (cf. [KM02]) à partir duquel l'extrapolation de Richardson Romberg est toujours possible.

1.3.2 Problème de Cauchy : cas discontinu

L'approche initiale de [TT90] est suffisamment robuste pour s'étendre naturellement au cas où le processus Z dans (1.1) comprend une partie à saut. L'idée sous jacente est la même et consiste à contrôler les dérivées de u qui peut dans ce cadre se voir comme la solution de l'EDP intégro différentielle :

$$\begin{cases} \partial_t u(t,x) + \langle b(t,x), D_x u(t,x)\rangle + \frac{1}{2}\mathrm{tr}(a(t,x)D_x^2 u(t,x)) \\ + \int_{\mathbb{R}^r} \left(u(t,x+\sigma(t,x)z) - u(t,x) - \langle \nabla u(t,x), \sigma(t,x)z\rangle \mathbb{I}_{|z|\leq 1}\right)\nu(dz) = 0,\ (t,x)\in[0,T)\times\mathbb{R}^d, \\ u(T,x) = g(x),\ x\in\mathbb{R}^d. \end{cases} \tag{1.14}$$

Sous des hypothèses similaires à celles du cas continu, i.e. régularité des coefficients (supposés homogènes) et de la condition terminale g, en imposant par ailleurs des conditions d'intégrabilité assez fortes sur la mesure de Lévy ν, Protter et Talay [PT97] ont réussi à obtenir un développement de l'erreur faible similaire à (1.7) lorsque (1.1) est approchée par son schéma d'Euler (1.3), cf. leur Théorème 2.2. L'idée reste celle de dérivation du flot. Les hypothèses d'intégrabilité sur ν sont nécessaires pour contrôler en norme L^p, où p est relié à l'indice de dérivation, les dérivées du flot.

Dans le cas d'EDS dirigées par des processus de Lévy généraux Z se pose également la question de simulation du schéma d'Euler. En effet, bien que les accroissements de processus de Lévy soient explictement simulables dans certains cas : processus stables, Gamma, Poisson composé, ..., ils ne le sont pas pas en toute généralité. L'analyse de l'erreur faible associée à un schéma "simulable" dérivant du schéma d'Euler est effectuée dans Jacod *et al* [JKMP05]. Les auteurs introduisent le schéma

$$\bar{X}^h_{t_{i+1}} = \bar{X}^h_{t_i} + b(t_i, \bar{X}^h_{t_i})h + \sigma(t_i, \bar{X}^h_{t_i})\zeta^h_{i+1},\ i\in[\![0, N-1]\!], \bar{X}^h_0 = x,$$

où les $(\zeta_i)_{i\geq 1}$ sont des variables i.i.d. approchant en loi les accroissements $(Z_{t_{i+1}} - Z_{t_i})_{i\geq 1}$. Toujours sous des hypothèses fortes concernant l'intégrabilité de la mesure de Lévy ν, ils obtiennent une majoration d'erreur de la forme $|\mathbb{E}[g(\bar{X}^h_T) - g(X^h_T)]| \leq C(\gamma_h \vee h)$. Dans l'inégalité précédente γ_h est tel que pour $H\in C^4_0(\mathbb{R}^d)$, $|\mathbb{E}[H(\zeta_h) - H(Z_h)]| \leq Ch\gamma_h$. Sous des hypothèses toujours plus fortes de régularité et d'intégrabilité sur b, σ, g, ν, un développement d'erreur est également obtenu.

L'hypothèse d'intégrabilité de la mesure ν requise dans les travaux précédents, et cruciale du point de vue des techniques utilisées, est en revanche génante pour de nombreuses applications qui font intervenir des processus à queues lourdes. Le cas des processus stables d'indice $\alpha\in(0,2)$, i.e. on exclut le cas Brownien, illustre bien cet aspect. Ils interviennent dans de nombreux champs d'applications : physique mathématique [IP06], télécommunications [SK74], finance [JMW96] et ne vérifient pas du tout les conditions d'intégrabilité précédentes. Par ailleurs, le schéma d'Euler est dans ce cas parfaitement simulable, cf. [PT97] ou [ST94]. Dans ce cadre, Hausenblas [Hau02] a obtenu dans le cas scalaire une majoration de l'erreur faible à l'ordre 1 en h pour $g\in L^\infty$ en utilisant le calcul de Malliavin.

Au chapitre 3, nous développerons comment à partir d'une représentation en série parametrix pour la densité de l'EDS (1.1) dirigée par un processus stable, ou en termes analytiques de la solution fondamentale de l'EDP (1.14), on peut obtenir un développement d'erreur entre la densité de la diffusion et de son schéma d'Euler, article **[9]**. En utilisant que les queues de la densité ont un comportement asymptotique en espace de l'ordre

de $(1+|x|^{d+\alpha})^{-1}$, qui est optimal au sens où une minoration du même ordre existe, nous pourrons également relaxer les hypothèses de bornitude de g pour permettre des croissances en $|g(x)| \leq C(1+|x|^{\beta}),\ \beta < \alpha$.

Mentionnons dans le cas général pour Z le travail de Kohatsu Higa et Tankov [KHT10] qui étudie l'erreur faible pour des schémas de pas adaptés à l'amplitude des sauts avec une amélioration de l'erreur de convergence usuelle en h du schéma d'Euler. L'approche reste néanmoins celle de dérivation du flot et requiert donc de la régularité.

1.3.3 Autres fonctionnelles

Le problème de Cauchy Dirichlet : cas continu

Supposons que l'équation (1.1) soit dirigée par un mouvement Brownien. Considérons pour un domaine $D \subset \mathbb{R}^d$ (i.e. un ouvert connexe) $\tau := \inf\{t \geq 0 : X_t \notin D\}$. Si l'on considère la fonctionnelle $\Psi\left((X_t)_{t\in[0,T]}\right) = g(T\wedge\tau, X_{T\wedge\tau})$, la question de l'approximation de $\mathbb{E}_x[\Psi\left((X_t)_{t\in[0,T]}\right)]$ est liée à des problématiques de valorisation d'options à barrières en finance et de façon plus générale à la résolution numérique de problèmes de Cauchy-Dirichlet pour l'équation de la chaleur.

Dans la pratique, une question supplémentaire se pose par rapport à l'étude du problème sans condition de bord, celle de l'approximation de Ψ^h. Un choix naturel et peu coûteux numériquement est de prendre $\Psi^h\left((X^h_t)_{t\in[0,T]}\right) = g(T\wedge\tau^h, \Pi_{\bar{D}}(X^h_{T\wedge\tau^h}))$, où $\tau^h := \inf\{t_i := ih, X^h_{t_i} \notin D\}$ et $\Pi_{\bar{D}}$ désigne la projection sur l'adhérence du domaine. En effet, ce choix permet de ne vérifier qu'aux instants de discrétisation si le schéma d'Euler se trouve ou non dans le domaine. L'analyse de l'erreur de discrétisation

$$\mathcal{E}^D_h = \mathbb{E}_x[g(T\wedge\tau^h, \Pi_{\bar{D}}(X^h_{T\wedge\tau^h}))] - \mathbb{E}_x[g(T\wedge\tau, X_{T\wedge\tau})] = \mathbb{E}[u(T\wedge\tau^h, \Pi_{\bar{D}}(X^h_{T\wedge\tau^h}))] - u(0,x),$$

s'effectue ensuite suivant le même schéma que dans le cas sans bord en utilisant les propriétés de régularité de u solution de

$$\begin{cases} (\partial_t + L_t)u(t,x) = 0,\ (t,x) \in [0,T)\times D, \\ u(t,x) = g(t,x),\ (t,x) \in [0,T)\times\partial D \cup \{T\}\times\bar{D}. \end{cases} \tag{1.15}$$

Des hypothèses de non dégénérescence similaires à celles présentées à la section précédente (hypoellipticité ou uniforme ellipticité) permettent d'avoir des estimées si le domaine D est suffisamment régulier. Dans ce cadre, Gobet [Gob00] a montré que $\mathcal{E}^D_h = O(h^{1/2})$ avec certaines conditions de support et/ou régularité sur g. Il est à noter que ce résultat n'est pas dû à l'approximation du processus de diffusion mais est inhérent au choix de l'approximation discrète du temps d'arrêt. En effet, dans le cas monodimensionnel, pour $X_t = W_t$ et qu'il n'y a donc pas d'erreur de discrétisation du processus, on peut déduire des travaux de Siegmund et Yuh [SY82] que pour l'approximation discrète de la fonction de répartition du processus stoppé, i.e. $g(t,x) = \mathbb{I}_{t=T}\mathbb{I}_{x>A},\ A$ fixé, $\mathcal{E}^D_h = Ch^{1/2} + o(h^{1/2})$. Ce résultat a motivé les travaux **[1]** et **[7]** qui ont permis d'établir respectivement une borne inférieure et un développement d'erreur à l'ordre 1/2 pour $\mathcal{E}^h_D$.

Par ailleurs, cette mauvaise vitesse en comparaison du cas sans bord a donné lieu à une recherche active pour l'amélioration de la convergence des méthodes numériques : approximation plus fine du temps de sortie par ponts Browniens [Gob00, Gob01], techniques de grandes déviations pour les asymptotiques de probabilités de sortie [Bal95, BCI99], correction de domaine [BGK97]. Détaillons ici cette dernière. Dans le cas du modèle de Black et Scholes monodimensionnel et pour un demi-espace, Broadie *et al.* [BGK97], ont introduit l'idée de baisser convenablement en fonction de h la barrière de sorte à compenser la surestimation des trajectoires prises en compte du fait de l'estimation discrète. Cette méthode a été généralisée à une large classe de domaines non cylindriques réguliers dans **[7]** pour des processus de diffusions.

Nous reviendrons en détail sur ces aspects au Chapitre 2 dédié aux approximations de processus stoppés.

Les modèles cinétiques et options asiatiques : cas continu

Nous considérons ici le cas où $\psi\left((X_t)_{t\in[0,T]}\right) = g(T^{-1}\int_0^T X_t dt, X_T)$. Typiquement, si l'on considère le cas d'une option asiatique en finance $g(T^{-1}\int_0^T X_t dt, X_T) = (T^{-1}\int_0^T X_t dt - X_T)^+$ pour un *call*. Le facteur de normalisation en T^{-1} devant l'intégrale permet aux deux quantités d'être à la même échelle globale et la comparaison entre valeur d'un actif à l'instant T et sa moyenne sur la période. Pour l'analyse de l'erreur faible $\mathcal{E}^D_h := \mathbb{E}_x[g(T^{-1}\displaystyle\int_0^T X^h_{\phi(t)}dt, X^h_T)] - \mathbb{E}_x[g(T^{-1}\displaystyle\int_0^T X_t dt, X_T)]$, le découpage précédent peut s'effectuer en considérant l'EDP :

$$\begin{cases} (\partial_t + L_t)u(t,\mathbf{x}) = 0,\ (t,\mathbf{x}) \in [0,T)\times\mathbb{R}^{2d}, \\ u(T,\mathbf{x}) = g(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^{2d}, \end{cases} \tag{1.16}$$

où l'on note $\mathbf{x} = (x, y)$ et $L_t\varphi(\mathbf{x}) = \langle b(t,x), \nabla_x\varphi(\mathbf{x})\rangle + \frac{1}{2}\mathrm{tr}\left(a(t,x)D_x^2\varphi(\mathbf{x})\right) + \langle x, \partial_y\varphi(\mathbf{x})\rangle$. Cette équation est dégénérée et le générateur L_t est celui du couple $(X_t, Y_t)_{t\geq 0}, Y_t = y + \int_0^y X_s ds$. Lorsque les coefficients b, σ dépendent également de y alors l'équation (1.16) peut être associée à un système Hamiltonien stochastique. Pour un Hamiltonien donné de la forme $H(x,y) = V(y) + \frac{|x|^2}{2}$ où V est un potentiel et $\frac{|x|^2}{2}$ l'énergie cinétique d'une particule de masse unité, le système Hamiltonien stochastique associé correspondrait à $b(X_s, Y_s) = -(\partial_y V(Y_s) + F(X_s, Y_s)X_s)$, où $F \geq 0$ est un terme de friction.

Dans le cas d'un mouvement Brownien géométrique à coefficient de diffusion aléatoire, l'analyse de $\mathcal{E}_h^D$ a été effectuée par Temam [Tem01]. L'approximation par chaîne de Markov de la solution fondamentale de (1.16), avec b, σ dépendant de $\mathbf{x} = (x, y)$, est l'objet du travail **[8]** et sera présentée au Chapitre 3. A l'aide de développements parametrix nous obtiendrons un théorème limite local en régime Gaussien pour la différence entre les densités du processus et de la chaîne de Markov. L'étude de systèmes dégénérés de type (1.16) sera développée au Chapitre 5.2. Nous y présenterons les résultats de **[10]** qui donnent en particulier des estimées "*à la Aronson*" pour la solution fondamentale de (1.16).

1.4 Quelques autres problèmes usuels d'approximation de processus

Erreur en loi normalisée

Lorsque l'on s'intéresse à l'approximation de fonctionnelles de la trajectoire, une autre question naturelle est celle de la limite en loi (au sens des processus) de l'erreur normalisée, i.e. la limite de $U^h := \rho_h(X - X^h)$ pour une renormalisation $\rho_h \underset{h\to 0}{\to} +\infty$ à préciser. La dynamique de cette limite a été étudiée initialement par Kurtz et Protter [KP91] pour des processus de diffusion puis dans un cadre plus général par Jacod et Protter [JP98]. Ce travail inclut en particulier le cas déterministe et de certains processus de Lévy.

Pour les processus de diffusion, la bonne renormalisation est $\rho_h = h^{-1/2}$, ce qui pouvait se prévoir à partir des contrôles sur l'erreur forte. Si $b = 0, Z = W$ dans (1.1), la dynamique limite de l'erreur s'écrit

$$U_t = \int_0^t \sigma'(X_s)U_s dZ_s - \frac{1}{\sqrt{2}}\int_0^t (\sigma\sigma')(X_s)d\tilde{W}_s, \tag{1.17}$$

où $\tilde{W}$ est un mouvement Brownien indépendant de W. Cette dynamique peut intuitivement se déduire du développement au premier ordre de la dynamique de la différence :

$$U_t^h = \int_0^t \sigma'(X_s)U_s^h dW_s + \int_0^t \sigma\sigma'(X_{\phi(s)}^h)h^{-1/2}(W_s - W_{\phi(s)})dW_s + R_t^h,$$

où R_t^h est un terme de reste.

Le mouvement Brownien indépendant du deuxième terme dans (1.17) apparaît grâce à l'indépendance asymptotique des deux intégrales stochastiques dans la dynamique de U^h. Le terme de normalisation en $1/\sqrt{2}$ provient quant à lui de la caractérisation à l'aide de la variation quadratique.

Le cas des processus de Lévy est plus délicat. Mentionnons le travail de Jacod [Jac04] qui relie le paramètre de normalisation ρ_h à la concentration de la mesure de Lévy ν de Z autour de 0 (cf. (1.2)). La dynamique limite de U est similaire à celle de (1.17) mais l'intégration pour le deuxième terme ne fait plus intervenir un Brownien mais un autre processus de Lévy V, lié au processus directeur Z mais toujours indépendant de ce dernier. En particulier, le cas stable est entièrement caractérisé.

Le problème de caractérisation de la dynamique limite de l'erreur a été abordé par Rubenthaler [Rub03] dans le cas d'une version simulable du schéma de discrétisation de l'EDS dirigée par un procéssus de Lévy général.

Mesures invariantes

Lorsque le processus X vérifie certaines propriétés de stabilité de type Lyapunov, la question de l'approximation de mesures invariantes se pose de façon naturelle.

Des schémas à pas décroissants ont été introduits dans le cas continu par Lamberton et Pagès [LP02, LP03]. Le raffinement du pas de temps permet de converger vers la mesure invariante du processus initial. Cette approche présente par ailleurs l'avantage de pouvoir s'appliquer au cas de processus dégénérés. Elle a été généralisée par Panloup pour des d'EDS dirigées par des processus de Lévy [Pan08] et dans des cas continus à dérive singulière par Lemaire [Lem07].

Dans le cadre des systèmes Hamiltoniens, i.e. lorsque le terme de friction $F > 0$ dans l'exemple ci-dessus, Talay [Tal02], Mattingly *et al.* [MSH02] ont étudié différents schémas d'approximation de la mesure invariante

par chaîne de Markov à pas constants. Ces derniers présupposent une meilleure connaissance *a priori* du semi-groupe de la diffusion. En effet dans ces travaux c'est la mesure invariante de la chaîne qui est approchée et il est ensuite nécessaire d'avoir des contrôles précis sur l'écart avec la mesure invariante du processus initial.

Transferts de propriétés du modèle continu au modèle discret

Nous mentionnerons ici les travaux de Veretennikov et ses coauteurs qui se sont intéressés au transfert de propriétés de mixing [KV06] et de principes de grandes déviations du modèle continu de diffusion dirigée par un Brownien à un schéma numérique discret [Ver03].

Vers des modèles non linéaires

L'essor de la théorie des équations différentielles stochastiques rétrogrades (EDSR) initiée par Pardoux et Peng [PP90], a permis d'étendre les représentations de Feynman Kac à une certaine classe de problèmes non linéaires : équations semi/quasi-linéaires, problèmes de frontière libre.

Ce lien a donné lieu à une série de travaux sur l'approximation de telles équations. Citons Zhang [Zha04] pour l'analyse de l'erreur d'un schéma de semi-discrétisation d'une EDSR, puis Gobet et Labart [GL07] pour un développement de l'erreur. Il ressort de ces travaux que l'erreur globale en norme L^2 commise sur le triplet (X, Y, Z) est guidée par l'erreur forte associée à la composante progressive. Si la composante rétrograde est un peu régulière, on propage simplement en temps rétrograde l'erreur usuelle forte associée à l'approximation de la composante progressive. On retrouve ce même phénomène pour les équations couplées, cf. **[2]** et **[5]**, les approximations d'EDSR dans un domaine **[6]**, ou les problèmes de frontière libre, voir Bally, Pagès *et al.* [BP03].

Nous reviendrons en détail sur ces aspects au Chapitre 4.

Chapitre 2

Approximation de processus tués/stoppés

Ce chapitre vise à faire une synthèse des travaux [**1**], [**3**], [**4**], [**7**] écrits en partie avec Emmanuel Gobet de l'ENSIMAG et relatifs à différents problèmes d'approximation pour des processus tués et/ou stoppés. L'idée est tout d'abord d'établir une vitesse de convergence pour l'erreur faible lorsque l'on fait le choix de stopper le schéma d'approximation à temps discret. Par opposition au cas de l'espace tout entier, cette vitesse est d'ordre $1/2$ en h si h désigne le pas de temps uniforme de discrétisation. Cet ordre correpond à celui de l'*overshoot* du schéma d'Euler, i.e. la distance au bord du domaine lorsque le schéma sort à temps discret. Cet ordre est encore celui de l'accroissement Brownien.

Le travail initial [**1**] a permis d'isoler l'overshoot comme terme dominant dans l'erreur et d'obtenir un encadrement de l'erreur à l'ordre $1/2$ pour des diffusions hypoelliptiques et des domaines cylindriques réguliers vérifiant une condition de frontière non caractéristique.

Des théorèmes limites sur l'overshoot existent depuis longtemps dans le cadre du Brownien monodimensionnel tué à temps discret. En particulier, Siegmund et Yuh [SY82] ont établi l'indépendance asymptotique du temps de sortie discret et de l'overshoot renormalisé. C'est précisément ce type de résultat qui est à la base du développement de l'erreur faible. Une grande partie de [**7**] a consisté à caractériser la loi limite du triplet temps de sortie discret, position de sortie et overshoot renormalisé pour le schéma d'Euler d'une diffusion uniformément elliptique stoppée à la sortie d'un domaine régulier non cylindrique. Il s'agit là de l'un des résultats principaux de ce chapitre. Dans le cas particulier de cônes formés par l'intersection de demi-espaces un résultat similaire avait été obtenu pour le Brownien dans [**3**]. Ce théorème limite a dans les deux cas mentionnés permis d'aboutir au développement de l'erreur faible (avec une extension au problème stationnaire pour un domaine régulier).

Comme nous l'avons indiqué au Chapitre 1.3.1, un développement d'erreur permet d'utiliser des méthodes d'accélération de la convergence pour l'approximation numérique. En particulier, la méthode d'extrapolation de Romberg est toujours possible, puisque purement algébrique, sans connaissance *a priori* de la constante dans le développement limité. Dans le cas de l'approximation faible de processus tués, une caractérisation "explicite" de cette constante permet d'utiliser une technique alternative d'accélération de convergence. Celle-ci consiste à tuer le processus à temps discret à la sortie d'un domaine plus petit que celui initial que l'on peut choisir de sorte à tuer le terme dominant de l'erreur. La construction de ce domaine dépend du pas de temps h, et explicitement de l'expression de la constante dans le développement limité de l'erreur. L'idée est d'utiliser la notion de sensibilité du problème de Dirichlet par rapport au domaine, qui avait d'abord été introduite en optimisation de formes, cf. Sokoloski et Zolesio [SZ92], puis développée dans un cadre probabiliste par Costantini *et al.* [CKG06].

Nous conclurons ce chapitre en présentant les résulats de [**4**] où l'on montre comment étendre l'analyse de l'erreur de discrétisation au cas non Markovien des processus d'Itô pour lesquels il n'est plus possible d'utiliser une représentation de l'erreur via une EDP.

2.1 Modèle général et lien avec le problème de Dirichlet

Soit $(X_t)_{t\geq 0}$ un processus de diffusion d-dimensionnel de dynamique

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s \tag{2.1}$$

où W est un mouvement Brownien de $\mathbb{R}^d$, et les coefficients b, σ sont Lipschitz en espace et localement bornés en temps.

On se donne une famille $(D_t)_{t\geq 0}$, dépendant du temps de domaines réguliers bornés de $\mathbb{R}^d$. On suppose également cette famille régulière en temps. Pour un instant $T > 0$ déterministe fixé, on associe à la famille précédente le domaine temps-espace (voir Figure 2.1)

$$\mathcal{D} = \bigcup_{0<t<T} \{t\} \times D_t = \{(t,x) : 0 < t < T, x \in D_t\} \subset]0,T[\times\mathbb{R}^d.$$

Les domaines cylindriques sont des cas particuliers des domaines temps-espace précédents. On a dans ce cas

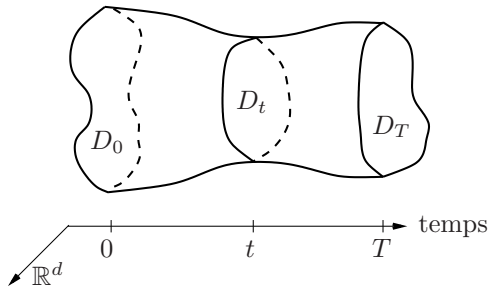


FIGURE 2.1 – Domaine temps espace et ses sections en temps.

$\mathcal{D} =]0,T[\times D$ où D est un domaine de $\mathbb{R}^d$ (i.e. $D_t = D$ pour tout t).

Définissons $\tau = \inf\{t > 0 : X_t \notin D_t\}$. Ainsi, $\tau \wedge T$ est le premier temps de sortie de $(t, X_t)_{t\geq 0}$ du domaine temps-espace $\mathcal{D}$. Pour des fonctions continues $g, f, k : \bar{\mathcal{D}} \to \mathbb{R}$, nous présenterons des résultats relatifs à l'approximation numérique de quantités de la forme

$$Q(T,g,f,k,x) := \mathbb{E}_x[g(\tau\wedge T, X_{\tau\wedge T})Z_{\tau\wedge T} + \int_0^{\tau\wedge T} Z_s f(s,X_s)ds], Z_s = \exp(-\int_0^s k(r,X_r)dr). \tag{2.2}$$

Sous de bonnes hypothèses de régularité des fonctions précédentes et de non-dégénerescence du coefficient de diffusion dans un voisinage du bord de $\mathcal{D}$, l'expression ci-dessus n'est autre que la représentation Feynman-Kac de la solution de l'EDP

$$\begin{cases} (\partial_t u + L_t u - ku + f)(t,x) = 0, \ (t,x) \in \mathcal{D}, \\ u(t,x) = g(t,x), \ (t,x) \in \mathcal{PD}, \end{cases} \tag{2.3}$$

où $L_t\varphi(x) = \langle b(t,x), \nabla\varphi(x)\rangle + \frac{1}{2}\mathrm{tr}(a(t,x)D_x^2\varphi(x))$ est le générateur infinitésimal de X_t, $\mathcal{PD} := \partial\mathcal{D}\backslash[\{0\}\times D_0]$ désigne la frontière parabolique du domaine $\mathcal{D}$ et les fonctions k et f correspondent respectivement à des termes potentiels et source dans l'équation de la chaleur ci-dessus.

L'approximation numérique la plus simple de la quantité (2.2) consiste à utiliser le schéma d'Euler stoppé à temps discret. Pour un pas de temps $h = T/m > 0$, $m \in \mathbb{N}^*$ fixé, on définit les instants de discrétisation $t_i = ih = iT/m)_{i\geq 0}$ et pour $t \geq 0$, $\phi(t) := t_i$ si $t_i \leq t < t_{i+1}$. Le schéma d'Euler s'écrit alors :

$$X_t^h = x + \int_0^t b(\phi(s), X_{\phi(s)}^h)ds + \int_0^t \sigma(\phi(s), X_{\phi(s)}^h)dW_s. \tag{2.4}$$

Nous associons au schéma d'Euler son temps de sortie discret $\tau^h := \inf\{t_i > 0 : X_{t_i}^h \notin D_{t_i}\}$, ce qui conduit à approcher (2.2) par :

$$Q^h(T,g,f,k,x) := \mathbb{E}_x[g(\tau^h\wedge T, X_{\tau^h\wedge T}^h)Z_{\tau^h\wedge T}^h + \int_0^{\tau^h\wedge T} Z_{\phi(s)}^h f(\phi(s), X_{\phi(s)}^h)ds], \ Z_t^h = e^{-\int_0^t k(\phi(r),X_{\phi(r)}^h)dr}. \tag{2.5}$$

L'erreur faible associée à cette procédure de discrétisation est la différence des équations (2.2) et (2.5), $\mathcal{E}_h^D := (Q^h - Q)(T,g,f,k,x)$. Précisons que dans l'approximation (2.5) ci-dessus la fonction g est presque sûrement évaluée hors de l'adhérence du domaine. Cette approximation peut à première vue sembler grossière. Elle n'affecte pas néanmoins la vitesse de convergence et évite numériquement l'étape potentiellement coûteuse de projection sur $\bar{\mathcal{D}}$.

Une première hypothèse de travail est celle de frontière non caractéristique du domaine. Précisément, si $F(t,x)$ désigne la distance signée de $x \in \mathbb{R}^d$ au bord du domaine D_t [1], on suppose que

1. Dans le cas d'un domaine cylindrique $F(t,x) = F(x)$, distance signée au bord du domaine D.

(FNC) Pour un $r_0 > 0$ il existe $a_0 > 0$ tel que $\nabla F(t,x)[\sigma\sigma^*](t,x)\nabla F(t,x)^* \geq a_0$ pour tout $(t,x) \in \bigcup_{0\leq t\leq T}\{t\} \times V_{\partial D_t}(r_0),\ V_{\partial D_t}(r_0) := \{x \in D_t : d(x, \partial D_t) \leq r_0\}$.

Cette condition est suffisante pour garantir la convergence de $\mathcal{E}_h^D$ vers 0, cf. Proposition 3 de [**1**]. Elle apparaît comme minimale dans la mesure où des contre-exemples existent lorsqu'elle n'est pas vérifiée, cf. [Gob00]. Elle exprime que dans un voisinage du bord le processus de distance au bord est non dégénéré, ou dit autrement que la matrice de diffusion est non dégénérée le long des normales au bord des sections D_t du domaine $\mathcal{D}$.

Comme le processus de distance au bord est mono-dimensionnel, il est en quelque sorte similaire à un mouvement Brownien changé de temps. C'est la formalisation de cette observation heuristique qui a permis de généraliser au schéma d'Euler les résultats asymptotiques sur l'overshoot à partir des résultats Browniens, voir Section 2.2.1.

Comme nous l'avons indiqué en Section 1.3.1, un outil essentiel pour l'analyse de l'erreur faible est le contrôle des dérivées de la solution u de (2.3). Nous avons considéré deux types d'hypothèses :

(A) Le domaine est cylindrique, i.e. $\mathcal{D} :=]0,T[\times D$. Les coefficients b, σ sont homogènes en temps et de classe C_b^∞ (bornés à dérivées bornées). Le coefficient de diffusion a vérifie **(FNC)** et une hypothèse de type Hörmander forte, i.e. seuls les crochets associés à la partie diffusive engendrent l'espace.

Dans ce cadre le noyau de la chaleur tué possède de bonnes propriétés, cf. Cattiaux [Cat91]. Ainsi, lorsque $g|_{\mathcal{PD}\setminus[\{T\}\times D_T]} = 0$ et que le support de $g(T,.)$ est à une distance positive du bord, il est possible de contrôler les dérivées de u en adaptant au cas tué les intégrations par parties de Malliavin utilisées par Bally et Talay [BT96a] (Article [**1**]).

(B) Le domaine $\mathcal{D}$, les coefficients b, σ, f, g sont suffisamment réguliers et a est uniformément elliptique.

Dans ce cas il est possible d'utiliser des estimées de type Schauder pour les équations dans des domaines non cylindriques, cf. Lieberman [Lie96] (Article [**7**]).

Précisons que sous les hypothèses **(A)** nous avons étudié l'erreur faible sans terme potentiel ni source, i.e. $f = k = 0$. En revanche, sous **(B)** nous avons considéré le cas général de l'équation (2.3). Dans les deux cas mentionnés, les formules d'Itô-Taylor nécessaires pour contrôler $\mathcal{E}_h^D$ font néanmoins intervenir un développement en semi-martingale de $\Pi_{\bar{\mathcal{D}}}(X_t^h)$. En particulier, le contrôle des restes, voir équation (2.6), requiert des estimées assez fines sur la mesure d'occupation d'un voisinage du bord pour le schéma et sur son temps local au bord, i.e. temps local de $F(t, X_t^h)_{t\geq 0}$ en 0. Nous renvoyons à l'annexe A de [**7**] et à la Section 3 de [**1**] pour des détails.

Sous les hypothèses **(A)**, pour lesquelles $g(s,x) = \mathbb{I}_{s=T}g_0(x),\ (s,x) \in \mathcal{PD}$, nous avons établi dans [**1**] que :

$$\mathcal{E}_h^D := \mathbb{E}_x[g_0(X_T^h)\mathbb{I}_{\tau^h>T}] - \mathbb{E}[g_0(X_T)\mathbb{I}_{\tau>T}] = \mathbb{E}_x\left[\int_0^T \partial_n u(t, X_t^h) d[F^-(X_{t\wedge\tau^h}^h)]\right] + O(h). \tag{2.6}$$

Ce contrôle illustre l'importance de l'overshoot pour le développement d'erreur. Pour g_0 positive, nous avons déduit de (2.6) dans [**1**] un encadrement de l'erreur pour h suffisamment petit, i.e. $\exists C \geq 1,\ C^{-1}h^{1/2} \leq \mathcal{E}_h^D \leq Ch^{1/2}$, en utilisant la stricte positivité de la dérivée normale. C'est la caractérisation de la loi limite de l'overshoot qui manquait pour obtenir le développement. C'est précisément cette loi que nous avons pu établir dans [**7**] pour des domaines temps-espace vérifiant la condition de frontière non caractéristique **(FNC)**.

2.2 Contrôles sur l'erreur faible

2.2.1 Théorème limite sur l'overshoot

Le résultat principal de cette section est le théorème limite sur la loi jointe du triplet formé par le temps de sortie discret, la position de sortie et l'overshoot renormalisé du schéma de discrétisation :

$$(\tau^h, X_{\tau^h}^h, h^{-1/2}F^-(\tau^h, X_{\tau^h}^h)).$$

Les lois limites d'overshoot ont été étudiées dans de nombreux cadres. Mentionnons les travaux de Siegmund [Sie79] et Siegmund et Hu [SY82] dans le cas de la marche simple scalaire, ceux d'Alsmeyer [Als94] ou Fuh et Lai [FL01] pour les chaînes de Markov ergodiques, et enfin Doney et Kyprianou [DK06] pour les processus de Lévy. Le point commun à toutes ces approches est l'utilisation d'arguments de renouvellement. Une théorie du renouvellement non linéaire a même été développée afin d'étudier les overshoot de marches aléatoires par rapport à des frontières courbes, cf. [Woo82] et [Zha88]. Nous étendons ce type de résultats au schéma d'Euler.

Théorème 2.2.1 (Lois jointes limites associées à l'overshoot) *Supposons que* **(FNC)** *est vérifiée, que le domaine $\mathcal{D}$ est de classe C^2 et les coefficients $b, \sigma \in C^{(1+\theta)/2,1+\theta}(\bar{\mathcal{D}})$ où $\theta \in]0,1[$ avec les notations usuelles pour les espaces de Hölder*[2]*. Soit φ une fonction continue à support compact. Pour tout $t \in [0,T],\ x \in D_0,\ y \geq 0$,*

$$\mathbb{E}_x[\mathbb{I}_{\tau^h \leq t} Z^h_{\tau^h} \varphi(X^h_{\tau^h}) \mathbb{I}_{F^-(\tau^h, X^h_{\tau^h}) \geq y\sqrt{h}}] \underset{h\to 0}{\longrightarrow} \mathbb{E}_x\big[\mathbb{I}_{\tau \leq t} Z_\tau \varphi(X_\tau)\big(1 - H(y/|\nabla F\sigma(\tau, X_\tau)|)\big)\big],$$

où $H(y) := (\mathbb{E}_0[s_{\tau^+}])^{-1} \int_0^y \mathbb{P}_0[s_{\tau^+} > z]dz$ et $s_0 := 0, \forall n \geq 1, s_n := \sum_{i=1}^n G^i$, les G^i étant des lois normales centrées standards i.i.d., $\tau^+ := \inf\{n \geq 0 : s_n > 0\}$.

En d'autres termes, le triplet $(\tau^h, X^h_{\tau^h}, h^{-1/2}F^-(\tau^h, X^h_{\tau^h}))$ converge donc faiblement vers $(\tau, X_\tau, |\nabla F\sigma(\tau, X_\tau)|Y)$ où Yest une variable aléatoire indépendante de (τ, X_τ) et de fonction de répartition H. En fait, Y est la loi asymptotique de l'overshoot Brownien, voir [Sie79]. La moyenne de l'overshoot est une quantité importante qui interviendra dans la procédure de correction de domaine. Notons que $\mathbb{E}(Y) = \frac{\mathbb{E}_0[s^2_{\tau^+}]}{2\mathbb{E}_0[s_{\tau^+}]} := c_0$, et d'après Siegmund [Sie79]

$$c_0 = -\frac{\zeta(1/2)}{\sqrt{2\pi}} = 0.5826... \tag{2.7}$$

où ζ désigne la fonction de Riemann.

L'idée clé pour prouver le Théorème 2.2.1 est de pratiquer une approximation de la frontière du domaine par un demi-espace et du schéma d'Euler par un processus à coefficients gelés dans un voisinage du bord. Une double localisation est toutefois nécessaire, voir Figure 2.2. En effet, l'approximation à domaine et coefficients constants sera pertinente lorsque le schéma d'Euler se trouve dans un voisinage de taille h^α du bord, $\alpha > 0$. Mais pour pouvoir appliquer les arguments de renouvellement, il faut également que le schéma d'Euler se trouve hors d'un voisinage de taille $h^{1/2-\varepsilon} < h^\alpha$. Par rapport à l'échelle Brownienne typique de $h^{1/2}$, le schéma se trouve suffisamment loin du bord. Le théorème résulte ensuite de contrôles assez précis sur l'approximation évoquée,

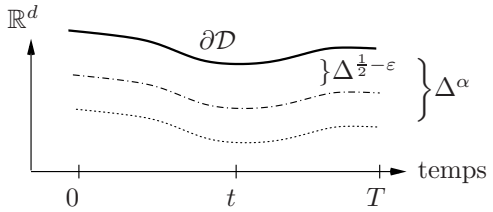


FIGURE 2.2 – Les deux voisinages de localisation avec $\alpha < \frac{1}{2} - \varepsilon$.

avec en particulier des estimées sur les échelles typiques pour les temps de sortie en fonction de l'échelle du voisinage, voir Section 3 de [**7**].

2.2.2 Application au développement d'erreur

Le Théorème 2.2.1 était l'argument manquant pour obtenir le développement d'erreur dans [**1**]. Nous avons utilisé ce résultat dans [**7**] pour établir le :

Théorème 2.2.2 (Développement au premier ordre) *Supposons que les hypothèses* **(B)** *soient vérifiées. Alors, pour h suffisamment petit*

$$\mathcal{E}^D_h = c_0\sqrt{h}\mathbb{E}_x[\mathbb{I}_{\tau \leq T} Z_\tau(\nabla u - \nabla g)(\tau, X_\tau) \cdot \nabla F(\tau, X_\tau)|\nabla F\sigma(\tau, X_\tau)|] + o(\sqrt{h}), \tag{2.8}$$

où c_0 est défini en (2.7).

Nous avons indiqué dans **(B)** qu'une régularité *suffisante* était requise. La régularité minimale considérée dans [**7**] présuppose que le domaine est C^2, les fonctions g, k, f sont $C^{(1+\theta)/2,1+\theta}(\bar{\mathcal{D}})$ de même que les coefficients b, σ.

Dans le cas particulier du jeu d'hypothèses **(A)**, abordé dans [**1**], i.e. cas hypoelliptique tué sans second membre ni potentiel, on déduirait sans trop de difficultés le développement à partir du Théorème 2.2.1 et de l'équation (2.6), après une intégration par parties pour faire explicitement apparaître l'overshoot $F^-(X^h_{T\wedge\tau^h})$ et non plus ses variations.

2. On renvoie pour cela au livre de Friedman [Fri64] ou Lieberman [Lie96]

2.2.3 Correction de domaine

Le Théorème 2.2.2 caractérise précisément la "mauvaise" convergence de l'approximation discrète Q^h. De nombreuses techniques avaient été utilisées numériquement pour y remédier. Mentionnons les approches par grandes déviations de Baldi *et al.*, [Bal95, BCI99], ou celles par ponts Browniens, pour affiner l'estimation du temps de sortie entre deux instants de discrétisation, cf. [Gob00, Gob01]. Ces dernières sont néanmoins coûteuses pour des géométries complexes dans la mesure où il faut procéder à des approximations par plan tangent.

D'autres méthodes numériques ont été développées. Indiquons celles par marches aléatoires sur des sphères, cf. Milstein [Mil98], ou sur des rectangles, cf. Milstein et Tretyakov [MT99] et Deaconu et Lejay [DL06], qui sont d'autant plus efficaces que la géométrie du domaine est proche de leur structure naturelle. Nous pensons néanmoins que notre approche reste une des plus directes.

Par ailleurs, le développement d'erreur du Théorème 2.2.2 ouvre de nouvelles perspectives. Il permet déjà d'effectuer l'extrapolation de Romberg. Néanmoins, la connaissance explicite de la constante c_0 ainsi que la forme du développement, qui fait intervenir les dérivées normales de la solution de l'EDP (2.3) et le gradient sur la frontière parabolique de la condition de bord, nous a amenés à proposer une technique alternative d'accélération de convergence. Celle-ci a par ailleurs un moindre impact sur la variance que l'extrapolation de Romberg.

Il s'agit simplement de stopper à temps discret le processus lorsque ce dernier sort d'un domaine plus petit que celui initial que l'on déforme suivant une amplitude et une direction qui permettent de supprimer le premier terme du développement dans (2.8). Précisément, introduisons $\mathcal{D}^h \subset \mathcal{D}$ dont les sections en temps sont définies par $D_t^h = \{x \in D_t : d(x, \partial D_t) > c_0 h^{1/2} |\nabla F \sigma(t,x)|\}$, voir Figure 2.3. Notons $\hat{\tau}^h = \inf\{t_i > 0 : X_{t_i}^h \in D_{t_i}^h\}$ le

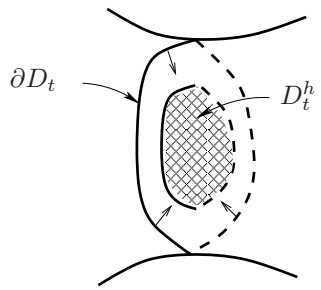


FIGURE 2.3 – Frontière ∂D_t et domaine restreint D_t^h.

temps de sortie discret de $\mathcal{D}^h$. Cette procédure de *correction de frontière* avait initialement été introduite dans le cas mono-dimensionnel Brownien par Broadie *et al.* [BGK99]. Elle amène à simuler la quantité

$$g(X_{\hat{\tau}^h}^h) Z_{\hat{\tau}^h}^h + \int_0^{\hat{\tau}^h} Z_{\phi(s)}^h f(X_{\phi(s)}^h) ds. \tag{2.9}$$

Notons $\hat{Q}^h(T, g, f, k, x) := \mathbb{E}[g(X_{\hat{\tau}^h}^h) Z_{\hat{\tau}^h}^h + \int_0^{\hat{\tau}^h} Z_{\phi(s)}^h f(X_{\phi(s)}^h) ds]$. Nous avons alors le résultat suivant :

Théorème 2.2.3 (Correction de frontière) *Sous les hypothèses du Théorème 2.2.2 et si $\nabla F(.)|\nabla F\sigma(.)|$ est de classe $C^{1,2}$, alors pour h suffisamment petit :*

$$\hat{\mathcal{E}}_h^D := (\hat{Q}^h - Q)(T, g, f, k, x) = o(\sqrt{h}).$$

L'hypothèse supplémentaire sur $\nabla F(.)|\nabla F\sigma(.)|$ permet d'assurer que le domaine modifié reste de classe C^2. L'idée de la preuve repose sur le développement :

$$\begin{aligned}\hat{\mathcal{E}}_h^D := &\left(\mathbb{E}[g(X_{\hat{\tau}^h}^h) Z_{\hat{\tau}^h}^h + \int_0^{\hat{\tau}^h} Z_{\phi(s)}^h f(X_{\phi(s)}^h) ds] - \mathbb{E}_x[g(\tau_{\mathcal{D}^h} \wedge T, X_{\tau_{\mathcal{D}^h} \wedge T}) Z_{\tau_{\mathcal{D}^h} \wedge T} + \int_0^{\tau_{\mathcal{D}^h} \wedge T} Z_s f(s, X_s) ds]\right) \\ &+ \left(\mathbb{E}_x[g(\tau_{\mathcal{D}^h} \wedge T, X_{\tau_{\mathcal{D}^h} \wedge T}) Z_{\tau_{\mathcal{D}^h} \wedge T} + \int_0^{\tau_{\mathcal{D}^h} \wedge T} Z_s f(s, X_s) ds] - \mathbb{E}_x[g(\tau \wedge T, X_{\tau \wedge T}) Z_{\tau \wedge T} + \int_0^{\tau \wedge T} Z_s f(s, X_s) ds]\right) \\ &:= \hat{\mathcal{E}}_h^{D,1} + \hat{\mathcal{E}}_h^{D,2},\end{aligned}$$

où $\tau_{\mathcal{D}^h} := \inf\{t > 0 : X_t \notin D_t^h\}$. Le terme $\hat{\mathcal{E}}_h^{D,1}$ est de même nature que $\mathcal{E}_h^D$. En particulier, les estimées de type Schauder sur l'EDP associée au domaine $\mathcal{D}^h$ sont uniformes pour h petit et la solution de l'EDP dans le

domaine perturbé et son gradient vont converger respectivement vers u et ∇u. Le développement du Théorème 2.2.2 va donc rester valable pour $\hat{\mathcal{E}}_h^{D,1}$.

Le terme $\hat{\mathcal{E}}_h^{D,2}$ peut être relié à la sensibilité du problème de Dirichlet, cf. Theorème 2.2 dans [CKG06]. Assez intuitivement, cette sensibilité fait intervenir les gradients de u et g sur la frontière parabolique, l'intensité de la pertubation, $-c_0 h^{1/2}|\nabla F\sigma|$ dans la mesure où l'on rétrécit le domaine, et la direction normale ∇F qui fait rentrer "le plus vite" dans le domaine. On obtient :

$$\hat{\mathcal{E}}_h^{D,2} = -c_0 h^{1/2}\mathbb{E}_x[\mathbb{I}_{\tau\leq T} Z_\tau(\nabla u - \nabla g)(X_\tau)\cdot\nabla F(X_\tau)|\nabla F\sigma(X_\tau)|] + o(h^{1/2}),$$

ce qui va exactement compenser le terme du développement du Théorème 2.2.2 et donner l'estimée du Théorème 2.2.3.

Insistons sur le fait qu'il était crucial pour pouvoir effectuer cette correction de connaître explicitement la constante c_0 dans le développement (2.8). C'est la forme de ce développement qui permet le parallèle avec la sensibilité du problème de Dirichlet. La connaissance précise de la constante est nécessaire pour prendre la bonne intensité de perturbation dans le schéma corrigé. Cette technique reste donc très liée au type de problème considéré.

Nous conclurons en indiquant que le développement d'erreur et la correction associée ont pu être démontrés sous des hypothèses de type **(B)**, i.e. domaine et coefficients réguliers, uniforme ellipticité, pour un problème stationnaire, voir Section 4 de [**7**] et les résultats numériques ci-dessous (Section 2.3). Un autre cadre est celui du mouvement Brownien tué à la sortie d'un cône, cf. [**3**]. Ce dernier cas est important en finance où les domaines de validité des options à barrières sont souvent définis en terme d'espaces produits. Un théorème limite similaire au Théorème 2.2.1 a été obtenu pour une intersection de demi-espaces. Le développement et la correction n'ont en revanche été établis en toute généralité que pour des cônes bidimensionnels lorsque g est nulle sur un voisinage de la singularité du cône. La difficulté est ici de contrôler le gradient de la solution de l'EDP jusqu'au bord, y compris au voisinage de la singularité. Pour ce dernier point, nous obtenons la régularité gâce à l'hypothèse sur le support de g et à partir d'un contrôle radial sur les dérivées de la densité du processus tué. Celle-ci se décompose en une série dont chaque terme fait intervenir une partie sphérique et radiale : la partie sphérique provient de la décomposition spectrale du Laplacien pour le problème de Dirichlet associé à la trace du cône sur la sphère, la partie radiale fait intervenir des fonctions de Bessel d'indices associés à la dimension et aux valeurs propres du Laplacien. Dans le cas bidimensionnel, l'expression explicite de cette densité sous forme de série est bien connue et s'obtient comme extension de la méthode des images, cf. Carslaw et Jaeger [CJ59]. Dans le cas $d \geq 2$, la formulation générale a été obtenue par Bañuelos et Smits [BS97].

2.3 Résultats numériques

Pour des exemples numériques dans le cas du Brownien à la sortie d'un cône nous renvoyons à la Section 3 de [**3**]. Nous choisissons de présenter ici un exemple pour un processus tridimensionnel stoppé à la sortie d'une boule, problème elliptique.

Soit $(X_t)_{t\geq 0}$ de dynamique :

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dW_t,\ \forall x\in\mathbb{R}^3,\ b(x) = (x_2\ x_3\ x_1)^*, \\ \sigma(x) &= \begin{pmatrix} (1+|x_3|)^{1/2} & 0 & 0 \\ \frac{1}{2}(1+|x_1|)^{1/2} & \left(\frac{3}{4}\right)^{1/2}(1+|x_1|)^{1/2} & 0 \\ 0 & \frac{1}{2}(1+|x_2|)^{1/2} & \left(\frac{3}{4}\right)^{1/2}(1+|x_2|)^{1/2} \end{pmatrix}, \end{aligned} \tag{2.10}$$

et X_0 appartenant à $D = B(0,2)$. On se donne une fonction $u(x) = x_1x_2x_3$ définie sur $\bar{D}$. On obtient l'EDP de type (2.3), en version stationnaire, associée à (2.10) satisfaite par u en prenant $g = u|_{\partial D}$, $f = -Lu$, où L désigne le générateur infinitésimal de X dans (2.10) et $k = 0$. Ce procédé permet donc d'avoir une expression explicite de la solution à approcher.

Pour x_0 tel que $(x_0^i)_{1\leq i\leq 3} \in \{-.7, -.3, .3, .7\}$, nous avons pris $N_{MC} = 10^6$ trajectoires pour la simulation de Monte Carlo et et nous avons fait varier h dans $\{.01, .05, .1\}$. Pour tous les calculs effectués, la taille de l'intervalle de confiance à 95% était comprise dans $[1.5\times 10^{-3}, 2\times 10^{-3}]$. Nous présentons au tableau 2.1 les résultats obtenus pour le *supremum* de la valeur absolue des erreurs absolues et relatives pour les $3\times 4^3 = 192$ points de la grille spatiale. Ces résultats semblent indiquer que pour la correction le terme de reste en $o(h^{1/2})$ dans le Théorème 2.2.3 est en fait un $O(h)$.

Nous concluons cette section en illustrant numériquement le résultat du Théorème 2.2.2. Pour $x_0 = (-.7, .3, .7)$ et les précédentes valeurs de h, nous représentons en Figure 2.4 la quantité $-\log(\mathrm{Err}_{MC})$ en fonction de $-\log(h)$,

h	Sans correction	Correction de domaine
.1	0.169 (199%)	0.0220 (24.4%)
.05	0.114 (133%)	0.0115 (13.1%)
.01	0.0471 (54.7%)	0.0026 (2.98%)

TABLE 2.1 – *Supremum* de la valeur absolue du schéma d'Euler (erreur relative en % entre parenthèses)

où $\mathrm{Err}_{MC} := \left\{ \frac{1}{MC} \sum_{i=1}^{MC} \left(g(X^{h,i}_{\tau^{h,i}}) + \int_0^{\tau^{h,i}} f(X^{h,i}_{\phi(s)}) ds \right) \right\} - u(x_0)$. La courbe ainsi obtenue est assez proche d'une droite de pente 1/2 telle que l'indique le théorème.

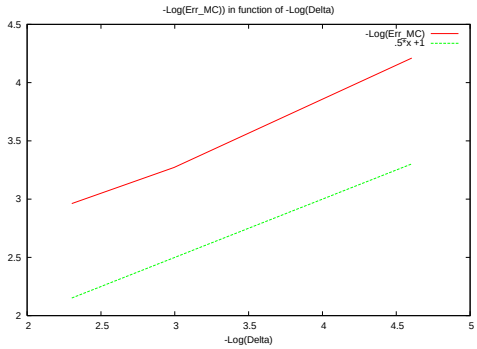


FIGURE 2.4 – Erreur de la méthode de Monte Carlo (sans correction) en fonction de h (échelle logarithmique). Evaluation en $x_0 = (-.7, .3, .7)$.

2.4 Echantillonage discret de processus d'Itô

En ce qui concerne l'étude de l'erreur faible pour des processus stoppés/tués, le dernier axe abordé est celui de l'observation discrète de processus d'Itô.

Le point principal dans ce cadre non Markovien est que l'on ne dispose plus d'une EDP de référence (2.3) pour analyser l'erreur. L'idée développée dans **[4]** pour un domaine cylindrique régulier, $\mathcal{D} :=]0, T[\times D$, consiste à exploiter que le processus $U_t := \mathbb{E}[g(T \wedge \tau_t, \Pi_{\bar{D}}(X_{T \wedge \tau_t}))|\mathcal{F}_t]$, $\tau_t := \inf\{s \geq t : X_s \notin D\}$, $t \geq 0$ est une martingale. Ce processus est un objet naturel, en effet dans le cas Markovien, pour $f = k = 0$, on a bien $U_t := u(t, X_t)$ où u est solution de (2.3). La projection sur $\bar{D}$ est introduite pour la cohérence avec les observations discrètes. En écrivant les accroissements de cette martingale entre les instants de discrétisation, on arrive à se ramener à une discussion sur les temps de sortie proche de celles de **[1]**,**[7]**.

Sous des hypothèses de type **(A)**, de frontière non caractéristique, bornitude des coefficients b, σ, de continuité en probabilité de l'intégrale stochastique à l'ordre usuel 1/2 en temps, pour un domaine D de classe C^2 est une fonction $g \in C^{1,2}([0, T] \times \mathbb{R}^d)$, nous retrouvons le contrôle

$$|\mathbb{E}_x[g(T \wedge \tau^h, \Pi_{\bar{D}}(X_{T \wedge \tau^h}))] - \mathbb{E}_x[g(T \wedge \tau, \Pi_{\bar{D}}(X_{T \wedge \tau}))]| \leq C\sqrt{h},$$

où $\tau^h := \inf\{t_i := ih : X_{t_i} \notin D\}$. Dans le cas tué, i.e. lorsque $g(s, x) = \mathbb{I}_{s=T} g_0(x)$, $(s, x) \in [0, T] \times \bar{D}$, et que g_0 bornée vérifie une condition de support, $d(\mathrm{supp}(g_0), \partial D) \geq \varepsilon > 0$, le résultat précédent s'étend à une intersection de domaines réguliers.

Extensions et perspectives

Le bilan des travaux **[1]**,**[3]**,**[4]**,**[7]** est que nous avons pu obtenir dans le cas des processus de diffusion un développement de l'erreur et une procédure d'accélération de convergence associée en contrôlant précisément l'overshoot du schéma d'Euler.

Une question naturelle, tant pour la modélisation financière que pour l'approximation numérique d'équations intégro-différentielles, est celle du comportement de l'erreur faible lorsque le processus dirigeant l'EDS (2.1)

comporte une partie de saut. Intuitivement il semble naturel de penser que dans le cas où l'on ajoute dans la dynamique différentielle les variations d'un processus de Lévy d'intensité finie, il sera possible de distinguer dans l'analyse de l'erreur les sorties dues aux *grands sauts* de celles dues à l 'observation discrète du schéma. Par ailleurs, les propriétés de la solution de l'EDP associée ont été étudiées par Garroni et Menaldi [GM02] dans le cas elliptique ou Bensoussan et Lions dans le cas parabolique [BL82], cf. Chapitre 3 Section 3. La difficulté reste de bien contrôler le gradient ponctuellement jusqu'au bord du domaine. Dans ce cadre, l'erreur faible devrait néanmoins rester du même ordre.

Les résultats numériques de la Section 2.3 suggèrent que sous de bonnes hypothèses de régularité le terme de reste du Théorème 2.2.3 est un $O(h)$. La question de la poursuite du développement de l'erreur et de la correction associée aux ordres supérieurs reste ouverte.

Un dernier développement possible de cette thématique est associé à la caractérisation de la régularité des noyaux de la chaleur dans des domaines non-réguliers. Des techniques variationnelles ont été développées, voir par exemple Grisvard [Gri92] ou Kozlov *et al.* [KMR97]. L'approche par décomposition spectrale proposée dans [BS97] et les contrôles sur le problème elliptique en dimension 2, que l'on peut penser être de même nature que ceux associés à la trace d'un cône tridimensionnel sur la sphère, devraient permettre de passer à la dimension 3 et de procéder par suite à une sorte de récurrence sur la dimension. Indiquons également que pour des cônes convexes des techniques de fonctions barrières permettent d'obtenir assez directement la Lipschitz continuité jusqu'au bord, cf. Chapitre 4, Section 4.4.

Chapitre 3

La méthode parametrix : théorèmes limites et développement d'erreur

Nous présentons dans ce chapitre les résultats des articles [**8**] (en collaboration avec V. Konakov de l'Université de Moscou et S. Molchanov de l'Université de Charlotte) et [**9**] (en collaboration avec V. Konakov). Le point commun à ces deux travaux est l'utilisation de la représentation de la densité de certains processus en terme de série *parametrix*. L'idée de cette représentation est ancienne et remonte pour les opérateurs paraboliques d'ordre deux non dégénérés sous forme non divergence à E. Levi, cf. Friedman [Fri64] pour une présentation. Elle se base sur une idée de continuité ou de méthode perturbative. En temps petit, la densité d'une EDS à coefficients variables est *a priori* proche de la densité d'une EDS à coefficients constants, pour laquelle on a de bons contrôles de densité. L'idée de la méthode est ensuite d'utiliser les équations de Kolmogorov satisfaites par ces deux densités pour précisément estimer la différence.

Outre l'approche de Levi, d'autres versions de la méthode *parametrix* ont été développées. En particulier, c'est l'approche proposée par McKean et Singer [MS67] pour obtenir des développements asymptotiques du spectre du Laplacien sur une variété en fonction de sa courbure, qui dans la mesure où elle admet un pendant discrétisé direct, est la plus propice pour analyser les erreurs associées aux schémas d'approximation. C'est cette approche qui a permis à Konakov et Mammen d'obtenir dans le cadre des processus de diffusion non dégénérés inhomogènes un théorème limite local sur les densités pour l'approximation par chaîne de Markov [KM00] ainsi qu'un développement d'erreur pour le schéma d'Euler [KM02].

Nous présenterons d'abord dans le cadre abstrait des processus de "diffusion-saut" la méthode parametrix de Mc Kean et Singer, cf. Section 3.1. Nous illustrerons ensuite la robustesse de cette approche pour les problèmes de discrétisation à travers deux exemples :

- l'approximation par chaîne de Markov de processus dégénérés, qui interviennent entre autres dans des modèles cinétiques et en finance (options asiatiques), cf. Section 3.2 et article [**8**].
- le développement d'erreur pour des EDS dirigées par des processus stables. cf. Section 3.3 et article [**9**].

3.1 Parametrix à la Mc Kean et Singer pour les processus de "diffusion-saut"

3.1.1 Présentation de la méthode

Soit $(X_t)_{t\geq 0}$ un processus de Markov à valeurs dans $\mathbb{R}^d$ de générateur de la forme :

$$L_t\varphi(x) = \langle b(t,x), \nabla\varphi(x)\rangle + \frac{1}{2}\mathrm{tr}\left(a(t,x)D_x^2\varphi(x)\right) + \int_{\mathbb{R}^d}\left(\varphi(x+y)-\varphi(x)-\frac{\langle y,\nabla\varphi(x)\rangle}{1+|y|^2}\right)\nu(t,x,dy), \quad (3.1)$$

pour $t\geq 0$, $a(t,.)$ matrice de diffusion semi-définie positive, $b(t,.)$ coefficient de dérive borné, ν mesure de Lévy et $\varphi\in C_0^\infty(\mathbb{R}^d)$ fonction test. Les opérateurs de la forme 3.1 vérifient le principe du maximum faible et sont quasi locaux, i.e. si $\mathrm{supp}(\varphi)\in B(x,\varepsilon)$, $|L_t\varphi(x)|\leq C_\varepsilon|\varphi|_\infty$, cf. Hille and Philips [HP57].

Intuitivement, pour $h>0$ petit, l'incrément $X_{t+h}-X_t$ est "proche" de l'incrément du processus à accroissements indépendants et stationnaires (PAIS) de caractéristiques gelées en $b(t,X_t), a(t,X_t), \nu(t,X_t,.)$. C'est là l'idée qui est à la base de l'approche de Levi. Supposons désormais que le processus X admet pour tout $t>0$

une densité p par rapport à la mesure de Lebesgue et que l'on s'intéresse à l'estimation de $p(t, t+h, X_t, x')$, i.e. densité en $t+h$ au point x'. Il est alors également naturel de dire que l'accroissement $X_{t+h} - X_t$ est proche de celui du PAIS de caractéristiques gelées en espace au point final où l'on considère la densité, i.e. $b(t, x'), a(t, x'), \nu(t, x', .)$. Cette autre approximation est à la base du travail de Mc Kean et Singer.

Pour obtenir une représentation de la densité p à partir de ces approximations heuristiques, il faut ensuite quantifier précisément l'erreur d'approximation. Comme l'on a supposé que X admettait une densité pour tout $t > 0$ il en est de même pour le processus à caractéristiques gelées. Notons $\tilde{p}^{x'}$ cette densité, où l'indice x' indique explicitement le point de congélation, et $\tilde{L}^{x'}$ son générateur :

$$\tilde{L}_t^{x'}\varphi(x) = \langle b(t, x'), \nabla\varphi(x)\rangle + \frac{1}{2}\mathrm{tr}\left(a(t, x')D_x^2\varphi(x)\right) + \int_{\mathbb{R}^d}\left(\varphi(x+y) - \varphi(x) - \frac{\langle y, \nabla\varphi(x)\rangle}{1+|y|^2}\right)\nu(t, x', dy). \tag{3.2}$$

Sous de bonnes hypothèses, la densité p vérifie les équations de Kolmogorov progressives et rétrogrades, i.e

$$\begin{aligned}\partial_t p(s, t, x, y) &= L_t^* p(s, t, x, y),\ t > s, (x, y) \in (\mathbb{R}^d)^2,\ p(s, t, x, .) \underset{t\downarrow s}{\rightarrow} \delta_x(.),\\ \partial_s p(s, t, x, y) &= -L_s p(s, t, x, y),\ t > s, (x, y) \in (\mathbb{R}^d)^2,\ p(s, t, ., y) \underset{s\uparrow t}{\rightarrow} \delta_y(.),\end{aligned} \tag{3.3}$$

où L_t^* désigne l'adjoint formel de L_t (nous n'avons en effet pas précisé d'hypothèses de régularité sur les coefficients dans (3.1)). La densité $\tilde{p}^{x'}$ vérifie également l'équation rétrograde :

$$\partial_s\tilde{p}^{x'}(s, t, x, y) = -\tilde{L}_s^{x'}\tilde{p}^{x'}(s, t, x, y),\ t > s, (x, y) \in (\mathbb{R}^d)^2,\ \tilde{p}^{x'}(s, t, ., y) \underset{s\uparrow t}{\rightarrow} \delta_y(.). \tag{3.4}$$

Ainsi,

$$\begin{aligned}(p - \tilde{p}^{x'})(s, t, x, x') &= \int_s^t du\partial_u\int_{\mathbb{R}^d} dz p(s, u, x, z)\tilde{p}^{x'}(u, t, z, x')\\ &= \int_s^t du\int_{\mathbb{R}^d} dz\left(\partial_u p(s, u, x, z)\tilde{p}^{x'}(u, t, z, x') + p(s, u, x, z)\partial_u\tilde{p}^{x'}(u, t, z, x')\right)\\ &= \int_s^t du\int_{\mathbb{R}^d} dz\left(L_u^* p(s, u, x, z)\tilde{p}^{x'}(u, t, z, x') - p(s, u, x, z)\tilde{L}_u^{x'}\tilde{p}^{x'}(u, t, z, x')\right)\\ &= \int_s^t du\int_{\mathbb{R}^d} dz p(s, u, x, z)(L_u - \tilde{L}_u^{x'})\tilde{p}^{x'}(u, t, z, x'),\end{aligned} \tag{3.5}$$

où l'on utilise la convergence vers la masse de Dirac pour la première égalité, les équations de Kolmogorov (3.3) et (3.4) pour la troisième et le passage à l'adjoint pour la dernière. Nous avons aussi formellement dérivé sous l'intégrale en espace pour la deuxième égalité.

Introduisons la notation $f \otimes g(s, t, x, x') = \int_s^t du\int_{\mathbb{R}^d} dz f(s, u, x, z)g(u, t, z, x')$ pour la convolution temps-espace et posons $\tilde{p}(s, t, x, x') := \tilde{p}^{x'}(s, t, x, x'),\ H(s, t, x, x') := (L_s - \tilde{L}_s^{x'})\tilde{p}(s, t, x, x')$. L'équation (3.5) se réécrit alors :

$$(p - \tilde{p})(s, t, x, x') = p \otimes H(s, t, x, x').$$

A partir de cette expression, l'idée est ensuite d'itérer cette procédure pour $p(s, u, x, z)$ dans (3.5) en utilisant cette fois ci la densité de transition d'un processus de caractéristiques gelées en z variable d'intégration. On fait de la sorte apparaître les convolutions itérées du noyau H pour obtenir le développement formel :

$$p(s, t, x, x') = \sum_{r=0}^{\infty}\tilde{p} \otimes H^{(r)}(s, t, x, x'), \tag{3.6}$$

où $H^{(0)} = I, H^{(r)} = H \otimes H^{(r-1)},\ r \geq 1$. Déduire des estimées sur p à partir de l'expression formelle (3.6) suppose d'avoir de bons contrôles sur le membre de droite. De façon générale, on utilisera les contrôles *a priori* connus sur la densité gelée et des propriétés de régularisation du noyau de convolution H. Nous renvoyons à la Section 3.1.2 pour une illustration dans le cas des diffusions.

Nous avons jusqu'à présent supposé l'existence de la densité p. Lorsque les coefficients dans (3.1) sont réguliers (en particulier Lipschitz continus en espace), il est facile de voir X_t comme la solution forte d'une EDS. Sous des hypothèses de non dégénérescence des coefficients, le calcul de Malliavin permet d'obtenir l'existence et la

régularité de la densité, voir Norris [Nor86] pour le cadre de diffusions hypoelliptiques homogènes, Cattiaux et Mesnager [CM02] pour le cas inhomogène, et Bichteller *et al.* [BGJ87] pour le cas avec sauts.

Réciproquement, il est toujours possible de partir du terme de droite de l'équation (3.6) résultant du précédent développement formel. Si la série ainsi que ses dérivées par rapport aux variables de temps et d'espace sont bien définies, il vient bien par construction que p est une solution des équations de Kolmogorov. Le principe du maximum donne l'unicité. Si cette procédure aboutit, on construit bien ainsi l'unique solution fondamentale de l'EDP parabolique intégro-différentielle associée à (3.1). L'identification à la densité du processus de Markov de générateur (3.1) se fait ensuite par un argument *à la Dynkin*, cf. Théorème 2.3 de [Dyn63] .

C'est ce que nous allons faire dans la section suivante dans le cadre de processus de diffusion inhomogènes à coefficient de diffusion uniformément Hölder continu en espace, mesurable en temps et dérive bornée.

3.1.2 L'exemple des processus de diffusion non dégénérés

Considérons le cas $\nu = 0$ dans (3.1) et supposons également que la matrice de diffusion est uniformément elliptique bornée de constante $\Lambda \geq 1$ (i.e. $\Lambda^{-1} I_d \leq a(t,x) \leq \Lambda I_d,\ \forall (t,x) \in \mathbb{R}^+ \times \mathbb{R}^d$) et uniformément η-Hölder continue en espace, $\eta \in (0,1]$. Le coefficient b est supposé mesurable borné.

Le problème de martingale est alors bien posé, cf. Théorème 7.2.1 de Stroock et Varadhan [SV79]. On peut *définir* p comme en (3.6). La densité gelée $\tilde{p}(s,t,x,x')$ n'est autre que celle du processus Gaussien $X_t^{x'} = x + \int_s^t b(u,x')du + \int_s^t \sigma(u,x')dW_u$, où $\sigma\sigma^*(u,x') = a(u,x')$, au point de congélation x'. Les hypothèses précédentes sur les coefficients a, b donnent en particulier que pour $T > 0$ fixé, il existe $C := C(T, \Lambda, |b|_\infty),\ c := c(\Lambda)$ telles que pour tout multi-indice de dérivation $\alpha, |\alpha| \leq 4$, et pour tout $0 < s < t < T$, $(x,x') \in (\mathbb{R}^d)^2$,

$$|\partial_x^\alpha \tilde{p}(s,t,x,x')| \leq \frac{C}{(t-s)^{|\alpha|/2}} p_c(t-s,x,x'),\ p_c(t-s,x,x') := \left(\frac{c}{2\pi(t-s)}\right)^{d/2} \exp\left(-c\frac{|x-x'|^2}{t-s}\right). \tag{3.7}$$

Le noyau de convolution s'écrit :

$$\begin{aligned} H(s,t,x,x') &= (L_s - \tilde{L}_s^{x'})\tilde{p}(s,t,x,x') \\ &= \langle b(s,x) - b(s,x'), \nabla_x \tilde{p}(s,t,x,x')\rangle + \frac{1}{2}\mathrm{tr}\left((a(s,x) - a(s,x'))D_x^2 \tilde{p}(s,t,x,x')\right). \end{aligned} \tag{3.8}$$

En particulier, le contrôle (3.7) donne :

$$\begin{aligned} |H(s,t,x,x')| &\leq C\left(\frac{2|b|_\infty}{(t-s)^{1/2}} + \frac{|a(s,x) - a(s,x')|}{(t-s)}\right) p_c(t-s,x,x') \\ &\leq C((t^{-1/2} + (t-s)^{-1+\frac{\eta}{2}})p_c(t-s,x,x'), \end{aligned}$$

où l'on utilise la η-Hölder continuité de a pour compenser la singularité en temps non intégrable des termes d'ordre 2 en modifiant le c pour la dernière inégalité. Le noyau H est donc régularisant, au sens où lorsqu'on le convole on obtient :
$\exists C := C(T, \Lambda, |b|_\infty, \eta),\ c := c(\Lambda, \eta),\ \forall r \geq 1$,

$$|\tilde{p} \otimes H^{(r)}(s,t,x,x')| \leq C^{r+1}(t-s)^{r\eta/2} \prod_{i=1}^{r+1} B\left(1 + \frac{(i-1)\eta}{2}, \frac{\eta}{2}\right) p_c(t-s,x,x'), \tag{3.9}$$

où $B(m,n) := \int_0^1 s^{m-1}(1-s)^{n-1}ds$ désigne la fonction β. Pour $t-s$ petit, ce terme est donc de plus en plus négligeable lorsque r croît.

De (3.9) et (3.6) on déduit directement des asymptotiques de la fonction β une borne supérieure de type Aronson pour p ainsi définie, i.e. $\exists C := C(T, \Lambda, |b|_\infty, \eta),\ c := c(\Lambda, \eta),\ \forall 0 \leq s < t \leq T, (x,x') \in (\mathbb{R}^d)^2,\ p(s,t,x,x') \leq Cp_c(t-s,x,x')$.

Par ailleurs, l'équation (3.6) permet de justifier que p vérifie l'équation de Kolmogorov rétrograde $(\partial_s + L_s)p(s,t,x,x') = 0,\ 0 \leq s < t \leq T$. La propriété régularisante en temps de H, cf. (3.9), permet ensuite de déduire $p(s,t,.,x') \underset{s\uparrow t}{\rightarrow} \delta_{x'}(.)$. On a donc bien construit une solution fondamentale de l'EDP parabolique. L'identification avec la densité de transition du processus X solution du problème de martingale se fait de façon standard, cf. Chapitre 3 de [SV79] ou Dynkin [Dyn63].

3.1.3 Représentation parametrix pour un schéma de discrétisation

Considérons un processus de diffusion-saut $(X_t)_{t\geq 0}$ à valeurs dans $\mathbb{R}^d$ de dynamique :

$$X_t = x + \int_0^t b(s, X_{s-})ds + \int_0^t \sigma(s, X_{s-})dZ_s, \tag{3.10}$$

pour $(Z_t)_{t\geq 0}$ processus de Lévy dont le générateur peut se mettre sous la forme (3.1) et admet une densité. Supposons également que X_t possède pour tout $t \in (0, T]$, $T > 0$ fixé, une densité que l'on peut exprimer sous la forme (3.6).

Soit $h > 0$ un pas de discrétisation tel que $T/h = N \in \mathbb{N}^*$ et $\{(t_i = ih)_{i\in[\![0,N]\!]}\}$ la grille temporelle associée. Une approximation naturelle de (3.10) par chaîne de Markov est donnée par le schéma

$$X^h_{t_{i+1}} = X^h_{t_i} + b(t_i, X^h_{t_i})h + \sigma(t_i, X^h_{t_i})\xi^h_{i+1},\ i \in [\![0, N-1]\!],\ X^h_0 = x, \tag{3.11}$$

où les $(\xi^h_i)_{i\in\mathbb{N}^*}$ sont des variables à densité i.i.d. qui approchent l'incrément $Z_{t_{i+1}} - Z_{t_i} \overset{\text{(loi)}}{=} Z_h$ (i.e. $Z_0 = 0$). Si le coefficient σ vérifie une hypothèse de non-dégénérescence, alors pour tout $i \in [\![1, N]\!]$, $X^h_{t_i}$ admet une densité. Notons p^h cette densité.

La méthode parametrix de Mc Kean et Singer présentée en Section 3.1 présente le grand avantage de pouvoir assez directement se transposer au cas discret et d'établir une représentation de p^h similaire à (3.6). Si l'on s'intéresse à la représentation de la densité $p^h(t_j, t_{j'}, x, x'), 0 \leq j < j' \leq N$, l'idée analogue au cas continu est d'introduire une chaîne de Markov $\tilde{X}^{h,x'}$ à coefficients gelés au point d'arrivée x' où l'on considère la densité. Sa dynamique s'écrit,

$$\tilde{X}^{h,x'}_{t_{i+1}} = \tilde{X}^{h,x'}_{t_i} + b(t_i, x')h + \sigma(t_i, x')\xi^h_{i+1},\ i \in [\![j, j'-1]\!],\ \tilde{X}^{h,x'}_{t_j} = x. \tag{3.12}$$

On notera $\tilde{p}^{h,x'}$ sa densité. Introduisons désormais les analogues discrets des générateurs définis en (3.1) et (3.2). Soit pour $\varphi \in C^\infty_0(\mathbb{R}^d)$ et $j \in [\![0, j')$,

$$L^h_{t_j}\varphi(x) = \frac{\mathbb{E}[\varphi(X^h_{t_j+h})|X^h_{t_j} = x] - \varphi(x)}{h}, \quad \text{et} \quad \widetilde{L}^{h,x'}_{t_j}\varphi(x) = \frac{\mathbb{E}[\varphi(\widetilde{X}^{h,x'}_{t_j+h})|\widetilde{X}^{h,x'}_{t_j} = x] - \varphi(x)}{h}.$$

En utilisant la notation $\widetilde{p}^h(t_j, t_{j'}, x, x') = \widetilde{p}^{h,x'}(t_j, t_{j'}, x, x')$, on définit de façon similaire à (3.8) le noyau discret H^h par

$$\begin{aligned} H^h(t_j, t_{j'}, x, x') &= \left(L^h_{t_j} - \widetilde{L}^{h,x'}_{t_j}\right)\widetilde{p}^h(t_j + h, t_{j'}, x, x') \\ &= h^{-1}\int_{\mathbb{R}^d}\left[p^h - \widetilde{p}^{h,x'}\right](t_j, t_{j+1}, x, z)\widetilde{p}^{h,x'}(t_{j+1}, t_{j'}, z, x')dz, \quad 0 \leq j < j' \leq N. \end{aligned}$$

Une simple manipulation algébrique, cf. Lemme 3.6 de [KM00], permet ensuite d'établir que pour $0 \leq t_j < t_{j'} \leq T$,

$$p^h(t_j, t_{j'}, x, x') = \sum_{r=0}^{j'-j}\left(\widetilde{p}^h \otimes_h H^{h,(r)}\right)(t_j, t_{j'}, x, x'), \tag{3.13}$$

où $f \otimes_h g(t_j, t_{j'}, x, x') = \sum_{k=0}^{j'-j-1} h\int_{\mathbb{R}^d} g(t_j, t_{j+k}, x, z)f(t_{j+k}, t_{j'}, z, x')dz$ est l'opérateur de convolution discret et $H^{h,(0)} = I$ et $H^{h,(r)} = H^h \otimes_h H^{h,(r-1)},\ r \geq 1$.

D'après (3.6) et (3.13), les densités respectives du processus X et de son schéma d'Euler X^h peuvent se représenter sous des formes similaires. Les étapes pour contrôler l'écart $(p - p^h)(0, T, x, x')$ sont les suivantes :

1. Contrôler les queues de la série dans (3.6), i.e. les termes $r \geq N$.
2. Quantifier l'impact de la discrétisation de l'opérateur de convolution dans (3.6), i.e. lorsque l'on remplace $\otimes$ par $\otimes_h$ dans le développement tronqué en $r = N$.
3. Contrôler la convergence de $\tilde{p}^h$ vers $\tilde{p}$.
4. Appliquer des arguments fins de stabilité pour quantifier les différences entre les noyaux de convolution continus H et discrets H^h.

Dans [KM00], Konakov et Mammen ont pu appliquer ce programme pour des processus de diffusion inhomogènes à coefficients bornés Lipschitziens et non dégénérés, i.e. la matrice de diffusion a est uniformément elliptique. Dans ce cadre, c'est l'étape 3. qui donne la vitesse de l'erreur d'approximation qui est de $h^{1/2}$ et qui est celle du théorème limite local sur les densités en régime Gaussien, cf. Bhattacharya et Rao [BR76]. Les auteurs établissent que $|(p-p^h)(0,T,x,x')| \leq h^{1/2}\pi(T,x,x'),\ \pi(T,x,x') \leq CT^{-d/2}(1+|\frac{x'-x}{T^{1/2}}|^S)^{-1}$ où l'exposant S est fonction des propriétés d'intégrabilité des variables $(\xi_i^h)_{i\geq 1}$ et C est une constante dépendant des coefficients, uniforme sur les compacts en temps. Ces mêmes techniques leur ont permis d'établir des asymptotiques précises de l'erreur en temps T petit [KM09].

Lorsque les ξ_{i+1}^h ont même loi que l'accroissement $Z_{t_{i+1}} - Z_{t_i}$, i.e. X^h est le schéma d'Euler, l'étape 3. ci-dessus n'est pas nécessaire. On retrouve alors que l'erreur dominante est celle associée à la discrétisation de l'opérateur de convolution qui est intuitivement de l'ordre de h. C'est ce qu'ont montré Konakov et Mammen [KM02] pour des processus de diffusion homogènes non dégénérés à coefficients bornés et Lipschitziens, i.e. pour $T=1$, $(p-p^h)(1,x,x') = h\pi(1,x,x') + h^2 r(1,x,x')$, $(|\pi|+|r|)(1,x,x') \leq C\exp(-c|x-x'|^2)$, $C := C(1,\Lambda,|b|_\infty)$, $c := c(\Lambda)$. L'extension au cas inhomogène ne présente pas de difficultés et se ferait suivant la même procédure.

Un développement au même ordre a été obtenu dans le cadre hypoelliptique par Bally et Talay [BT96b] pour un schéma d'Euler perturbé. L'une des difficultés dans ce cadre est en effet de prouver l'existence d'une densité pour le schéma d'approximation ce que permet la perturbation. Nous avons retrouvé cette difficulté dans **[8]** que nous présentons dans la section suivante.

3.2 Théorème limite local en régime Gaussien dans un cas dégénéré

Nous présentons dans cette section les résultats de **[8]**. Nous avons pu établir un théorème limite local pour l'approximation par chaîne de Markov d'un processus de diffusion dégénéré $(X_t^1, X_t^2)_{t\geq 0}$ à valeurs dans $\mathbb{R}^d \times \mathbb{R}^d$ de dynamique :

$$\begin{aligned} X_t^1 &= x_1 + \int_0^t b(X_s^1, X_s^2)ds + \int_0^t \sigma(X_s^1, X_s^2)dW_s, \\ X_t^2 &= x_2 + \int_0^t X_s^1 ds, \end{aligned} \tag{3.14}$$

où W est un mouvement Brownien de $\mathbb{R}^d$ et où l'on suppose

(S) b, σ sont bornés, Lipschitz continus et de classe C^1.

(UE) La matrice de diffusion $a : \mathbb{R}^{2d} \to \mathbb{R}^d \otimes \mathbb{R}^d$ est uniformémement elliptique, ce qui associé à l'hypothèse **(S)** donne : $\exists \Lambda \geq 1,\ \forall x = (x_1,x_2) \in \mathbb{R}^{2d},\ \forall \xi \in \mathbb{R}^d,\ \Lambda^{-1}|\xi|^2 \leq \langle a(x)\xi, \xi\rangle \leq \Lambda|\xi|^2$.

Comme nous l'avons mentionné au Chapitre 1, ce type d'équations intervient dans de nombreuses applications, en finance pour les problèmes de valorisation/couverture d'options Asiatiques, en physique pour les modèles cinétiques ou les systèmes Hamiltoniens.

3.2.1 Contrôles sur l'objet continu

Sous **(S)**, **(UE)**, le théorème de Hörmander donne directement l'existence de la densité pour $X_t, t > 0$. Pour comprendre plus précisément la structure de cette densité, considérons d'abord le cas de coefficients b, σ constants pour lequel X_t est un processus Gaussien. Il suffit alors pour exprimer sa densité de calculer sa moyenne et sa matrice de covariance. L'expression de cette densité remonte au travail de Kolmogorov [Kol34]. Dans ce cas dégénéré très simple on remarque déjà l'échelle intrinsèque non diffusive due à la dégénérescence, i.e. la première composante vit à l'échelle de temps caractéristique du Brownien en $t^{1/2}$, la deuxième composante qui est l'intégrale de la première vit elle à l'échelle $t^{3/2}$. Un autre aspect remarquable est dû à la non bornitude du coefficient de la deuxième composante, qui induit un transport de la condition initiale x_1 dans X_t^2, i.e. $\mathbb{E}[X_t^2] = x_2 + tx_1 + \int_0^t \mathbb{E}[\int_0^s b(X_u^1, X_u^2)du]ds$.

La solution fondamentale de l'EDP parabolique associée au générateur de (3.14) a, sous des hypothèses supplémentaires de régularité des coefficients, été étudiée par Weber [Web51]. L'approche est celle du parametrix *à la Levi*. L'auteur obtient une borne supérieure sur la densité p de (3.14) similaire à celle de l'exemple de

Kolmogorov, soit pour $T > 0$ fixé,

$$\exists C := C(T, \Lambda, |b|_\infty),\ c := c(\Lambda),\ \forall (t,x,x') \in (0,T] \times \mathbb{R}^d \times \mathbb{R}^d,\ p(t,x,x') \leq C p_c(t,x,x'),$$
$$p_c(t,x,x') := \left(\frac{\sqrt{3}c}{2\pi(t-s)^2}\right)^{d/2} \exp\left(-c\left\{\frac{|x_1'-x_1|^2}{4t} + 3\frac{|x_2'-x_2-\frac{x_1+x_1'}{2}t|^2}{t^3}\right\}\right). \tag{3.15}$$

A la Section 3.1.2 nous avons gelé les coefficients au point d'arrivée où l'on souhaitait estimer la densité. Au vu des contrôles (3.7) et de la définition de l'opérateur de convolution (3.8), cela permettait en effet bien d'équilibrer la singularité associée au termes d'ordre deux. Le point clé est qu'il y avait dans ce cas "compatibilité" entre le terme apparaissant dans l'exponentielle $\exp(-c|x'-x|^2/t)$ et celui résultant de $|a(t,x)-a(t,x')|t^{-1} \leq C\frac{|x'-x|^\eta}{t^{\eta/2}}t^{-1+\eta/2}$, i.e. cette singularité était compensée par l'exponentielle. Dans le cas présent ce choix de processus gelé n'est pas satisfaisant dans la mesure où il ne permet pas de prendre en compte le transport de la condition initiale pour la deuxième composante qui apparaît dans l'expression explicite (3.15). Pour adapter la méthode de Mc Kean et Singer au cas de processus à coefficient de dérive non borné, l'idée est de construire un processus Gaussien à coefficients gelés qui permette d'avoir lors de la différence des générateurs un terme homogène à la composante dans l'exponentielle. Le candidat naturel est :

$$\begin{aligned}
\widetilde{X}_s^{1,t,x'} &= x_1 + \int_0^s \sigma(x_1', x_2' - x_1'(t-u))dW_u + b(x_1', x_2')s, \\
\widetilde{X}_s^{2,t,x'} &= x_2 + \int_0^s \widetilde{X}_u^{t,x'}du,\ s \in [0,t].
\end{aligned} \tag{3.16}$$

En effet, les parties dégénérées des deux générateurs sont les mêmes et se simplifient. Par ailleurs, la différence entre les termes d'ordre 2 des générateurs de (3.14) et (3.16) est dominée par $|a(x_1', x_2' - x_1't) - a(x)|t^{-1}\hat{p}_c(t,x,x') \leq C(|x_1'-x_1|(1+\frac{t}{2}) + |x_2'-x_2-\frac{x_1'+x_1}{2}t|)t^{-1}\hat{p}_c(t,x,x') \leq Ct^{-1/2}p_c(t,x,x')$ grâce à l'hypothèse de Lipschitz continuité de a et avec une modification de c pour la dernière inégalité. Ce qui permet bien de rattrapper la singularité. Nous pouvons ainsi retrouver sous nos hypothèses courantes, plus faibles, le contrôle de type Aronson obtenu dans [Web51].

Proposition 3.2.1 *Soit $T > 0$ fixé. Sous* **(UE)**, **(S)**, *il existe des constantes $C := C(T, \Lambda, |b|_\infty)$, $c := c(\Lambda)$ telles que pour tout $(t,x,x') \in (0,T] \times (\mathbb{R}^d)^2$,*

$$p(t,x,x') \leq C p_c(t,x,x'),$$

où p_c est défini en (3.15).

En fait, de façon similaire à la Section 3.1.2, il suffit pour obtenir cette borne de considérer le coefficient a, η-Hölder continu en espace et b borné. De même une dépendance en temps des coefficients n'induirait aucune difficulté additionnelle.

Pour conclure cette partie, mentionnons que le choix des paramètres de congélation qui permet de contrôler le développement (3.6) est fortement lié à la métrique dérivant du problème de contrôllabilité déterministe associé à l'EDS de départ. Le terme dans l'exponentielle apparaît alors comme l'énergie du contrôle optimal qui permet de joindre x à x' en un temps t. Nous reviendrons en détails sur ce point au Chapitre 5.2.

3.2.2 Approximation par chaîne de Markov

On introduit tout d'abord un schéma de discrétisation du processus gelé (3.16). Soit $h_0 > 0$ un pas de temps "microscopique", on définit pour $(x,x') \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}, t > 0$ fixés, $X_0^{h_0,t,x'} = \left(\widetilde{X}_0^{h_0,1,t,x'}, X_0^{h_0,2,t,x'}\right) = x = (x_1, x_2)$ et pour $i \geq 0$,

$$\begin{aligned}
\widetilde{X}_{t_{i+1}}^{h_0,1,t,x'} &= \widetilde{X}_{t_i}^{h_0,1,t,x'} + b(x')h_0 + \sigma(x_1', x_2' - tx_1')\sqrt{h_0}\widetilde{\xi}_{i+1}, \\
\widetilde{X}_{t_{i+1}}^{h_0,2,t,x'} &= \widetilde{X}_{t_i}^{h_0,2,t,x'} + \widetilde{X}_{t_{i+1}}^{h_0,1,t,x'}h_0 \\
&= \widetilde{X}_{t_i}^{h_0,2,t,x'} + h_0\widetilde{X}_{t_i}^{h_0,1,t,x'} + h_0^2 b(x') + h_0^{3/2}\sigma(x_1', x_2' - tx_1')\widetilde{\xi}_{i+1},
\end{aligned} \tag{3.17}$$

où les $(\widetilde{\xi}_i)_{i\in\mathbb{N}^*}$ sont des variables centrées i.i.d. de densité q, de covariance l'identité. Si l'on itère n fois le schéma il vient

$$\begin{aligned}
\widetilde{X}_{t_n}^{h_0,1,t,x'} &= x_1 + (nh_0)b(x') + \sigma(x_1', x_2' - x_1't)\sqrt{nh_0}\widetilde{\xi}_n^{(1)}, \\
\widetilde{X}_{t_n}^{h_0,2,t,x'} &= x_2 + (nh_0)x_1 + \frac{\gamma_n}{2}(nh_0)^2 b(x') + (nh_0)^{3/2}\sigma(x_1', x_2' - x_1't)\widetilde{\xi}_n^{(2)},
\end{aligned} \tag{3.18}$$

où $\gamma_n = (1+\frac{1}{n})$ et $\widetilde{\xi}_n^{(1)} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \widetilde{\xi}_i$, $\widetilde{\xi}_n^{(2)} = \frac{1}{\sqrt{n}}\sum_{i=1}^n (\frac{n-(i-1)}{n})\widetilde{\xi}_i$.

Sous des hypothèses d'intégrabilité et régularité de la densité q des $(\widetilde{\xi}_i)_{i\in\mathbb{N}^*}$, on peut montrer en calculant explicitement la fonction caractéristique que pour n suffisamment grand $(\widetilde{\xi}_n^{(1)}, \widetilde{\xi}_n^{(2)})$ admet une densité q_n possédant des propriétés de régularité et intégrabilité similaires à celles de q, cf. Proposition 4.2 de [**8**]. Par ailleurs $\mathrm{Cov}(\widetilde{\xi}_n^{(1)}, \widetilde{\xi}_n^{(2)}) = \begin{pmatrix} 1 & \gamma_n/2 \\ \gamma_n/2 & 1/3\gamma_n(1+1/(2n)) \end{pmatrix}$ qui est "proche" pour n grand de $\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} = \mathrm{Cov}(W_1, \int_0^1 W_s ds)$.

Si l'on pose maintenant $h := nh_0$, pas de temps "macroscopique", et que l'on s'intéresse à l'approximation de la densité en $t_j := jh,\ j \in (0, N]\!],\ Nh = T$, on introduit le schéma gelé agrégé, dérivant de (3.18) et résultant du regroupement de n itérations de (3.17) où l'on prend $t = t_j - t_i$ au $i^{\text{ème}}$ instant de discrétisation de l'échelle macroscopique. Précisément, pour $(x, x') \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$, $j \in (0, N]\!]$ on définit $\widetilde{X}_0^{h,t_j,x'} = x$ et pour $i \in [\![0, j-1]\!]$,

$$\begin{aligned}
\widetilde{X}_{t_{i+1}}^{h,1,t_j,x'} &= \widetilde{X}_{t_i}^{h,1,t_j,x'} + b(x')h + \sigma(x_1', x_2' - x_1'(t_j - t_i))\sqrt{h}\widetilde{\eta}_{i+1}^1, \\
\widetilde{X}_{t_{i+1}}^{h,2,t_j,x'} &= \widetilde{X}_{t_i}^{h,2,t_j,x'} + \left\{\widetilde{X}_{t_i}^{h,1,t_j,x'} + \frac{\gamma_n}{2}b(x')h + \sigma(x_1', x_2' - x_1'(t_j - t_i))\sqrt{h}\widetilde{\eta}_{i+1}^2\right\} h,
\end{aligned} \tag{3.19}$$

où les variables i.i.d. $(\tilde{\eta}_i^1, \tilde{\eta}_i^2)_{i\in(0,j]\!]}$ ont pour densité $q_n(.)$.

Le schéma précédent guide le choix de la dynamique de la chaîne de Markov qui va approcher X solution de (3.14). Soit pour $(x, x') \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}$, $X_0^h = x$ et pour $i \in [\![0, N-1]\!]$,

$$\begin{aligned}
X_{t_{i+1}}^{h,1} &= X_{t_i}^{h,1} + b(X_{t_i}^h)h + \sigma(X_{t_i}^h)\sqrt{h}\eta_{i+1}^1, \\
X_{t_{i+1}}^{h,2} &= X_{t_i}^{h,2} + (X_{t_i}^{h,1} + \frac{\gamma_n}{2}b(X_{t_i}^h)h + \sigma(X_{t_i}^h)\sqrt{h}\eta_{i+1}^2)h,
\end{aligned} \tag{3.20}$$

où les $(\eta_i^1, \eta_i^2)_{i\in(0,N]\!]}$ ont également pour densité q_n.

En résumé, l'itération du schéma "naturel" gelé (3.17) permet d'obtenir une densité après suffisamment de pas de temps "microscopiques". Ceci donne la forme du schéma gelé à l'échelle "macroscopique", cf. (3.19). La structure de ce schéma donne ensuite celle de la chaîne de Markov associée à l'approximation du processus initial. A partir des équations (3.20), (3.19), on peut obtenir pour la densité p^h de la chaîne de Markov X^h une représentation de type (3.13) où $\tilde{p}^h$ est la densité associée à (3.19). Cette représentation sera comparable au développement (3.6) obtenu pour le processus continu à la Section 3.2.1 à partir de (3.16). En suivant ensuite le programme détaillé en fin de Section 3.1.3 on obtient :

Théorème 3.2.1 (Théorème limite local pour les densités) .
Supposons **(UE)** *et* **(B)**. *Si pour tout multi-indice de dérivation* ν, $|\nu| \le 4$, $\int_{\mathbb{R}^d} |u|^S |D_u^\nu q_n(u)| du < +\infty$, $S > 4(d+1) + 2d^2$, *alors* $\exists c > 0$,

$$\sup_{x,x'\in\mathbb{R}^{2d}} \left[(1+|x_1'|+|x_1|) \sup_{\delta\in[0,1]} p_c(T(1+\delta), x, x') + \chi_{T^{1/2}}\left(x_1' - x_1, x_2' - x_2 - T\left(\frac{x_1'+x_1}{2}\right)\right)\right]^{-1}$$
$$\times |p_h(T, x, x') - p(T, x, x')| = O(h^{1/2}),$$

où p_c *est défini en* (3.15) *et* $\forall u = (u_1, u_2) \in \mathbb{R}^{2d}$,

$$\chi_{T^{1/2}}(u) = T^{-2d}\left(1 + \left(\left|\frac{u_1}{T^{1/2}}\right|^2 + \left|\frac{u_2}{T^{3/2}}\right|^2\right)^{(S-4)/4d-1}\right)^{-1}.$$

La vitesse de convergence reste celle du théorème limite local Gaussien dans la mesure où la densité $\tilde{p}$ du processus gelé continu (3.16) reste Gaussienne. L'ordre de $h^{1/2}$ est celui de la convergence de $\tilde{p}^h$ vers $\tilde{p}$. Le terme en p_c provient d'une correction technique nécessaire pour traiter le coefficient non borné dans l'analyse de stabilité.

Pour illustrer ce résultat on peut considérer l'exemple de l'approximation numérique de $\mathbb{P}_x[(T^{-1}X_T^2 - X_T^1)^+ > K]$ pour $K > 0$ donné. Cette quantité correspond en finance au prix d'une option *Call digital* de maturité T lorsque la dynamique de l'actif est donnée par la première composante de (3.14). Si l'on associe maitenant les composantes du système (3.14) à la vitesse et à la position d'une particule, $\mathbb{P}_x[(T^{-1}X_T^2 - X_T^1)^+ > K]$ est un quantile de la partie positive de l'écart entre vitesse moyenne et vitesse instantanée en T. Le théorème limite local 3.2.1 permet de directement relier les densités des objets discrets et continus de dynamiques respectives données en (3.14) et (3.20). Cette approche est préférable à une discrétisation de la composante non-dégénérée associée une estimation numérique de l'intégrale dans la mesure où dans ce cas là le schéma d'approximation peut ne pas avoir de densité.

3.3 Approximation d'EDS dirigées par un stable : développement d'erreur des densités

On présente ici les résultats de l'article [**9**]. Nous considérons le cas où X à valeurs dans $\mathbb{R}^d$ suit une dynamique de type (3.10). Les coefficients b, σ sont homogènes en temps et Z est un processus α-stable symétrique, i.e.

$$X_t = x + \int_0^t b(X_{s-})ds + \int_0^t \sigma(X_{s-})dZ_s,$$

et pour $t \geq 0$, $u \in \mathbb{R}^d$,

$$\mathbb{E}[\exp(i\langle u, Z_t\rangle)] \quad = \quad \exp(it\langle \gamma, u\rangle - t\int_{S^{d-1}} |\langle s, u\rangle|^\alpha \lambda(ds)), \tag{3.21}$$

où λ est une mesure symétrique sur la sphère unité S^{d-1}, $\alpha \in (0, 2)$. On exclut le cas Brownien pour lequel on peut se reporter à [KM02] pour des résultats similaires.

Dans ce cadre, le schéma d'Euler associé à (3.11) est bien défini et explicitement simulable, cf. [PT97] ou [ST94]. Pour un pas $h > 0$ et la grille de discrétisation $\{(t_i := ih)_{i\in\mathbb{N}^*}\}$ associée, on pose $X_0^h = x$ et pour $i \geq 0$,

$$X_{t_{i+1}}^h = X_{t_i}^h + b(X_{t_i}^h)h + \sigma(X_{t_i}^h)(Z_{t_{i+1}} - Z_{t_i}).$$

En revanche, les hypothèses d'intégrabilité sur la mesure de Lévy requises dans Protter et Talay [PT97] pour l'analyse de l'erreur faible ne sont pas satisfaites par les processus stables. Ces derniers ont en effet des queues lourdes. Dans le cas monodimensionnel et pour des fonctions tests bornées, le contrôle de l'erreur faible avait été obtenu par Hausenblas [Hau02] à l'aide du calcul de Malliavin. L'approche parametrix permet de contrôler l'erreur entre les densités.

Sous de bonnes hypothèses spécifiées ci-après, Kolokoltsov a donné une représentation en série (3.6) de la densité p de X. Il a également établi des contrôles *à la Aronson* avec une échelle de temps caractéristique en $t^{1/\alpha}$ généralisant le cas Gaussien, cf. Théorème 3.2 de [Kol00]. Nous avons de façon similaire au procédé présenté en Section 3.1.3 obtenu le développement (3.13) pour la densité du schéma d'Euler. En suivant le programme détaillé en fin de Section 3.1.3 nous avons pu établir un développement d'erreur sur les densités à l'ordre h. Pour ce faire il a fallu dans l'analyse de stabilité raffiner les arguments de [Kol00] et utiliser des découpages minutieux dans les parties d'analyse de Fourier utilisées pour représenter densités gelées et noyau de convolution.

Soit $q \in \mathbb{N}^*$. Introduisons les hypothèses

(A1-q) Pour $d \geq 2$, la mesure sphérique λ a une densité de surface $C^q(S^{d-1})$ et pour $d \geq 1$, il existe $0 < C_1 \leq C_2 < +\infty$, $\forall p \in \mathbb{R}^d$,

$$C_1|p|^\alpha \leq \int_{S^{d-1}} |\langle p, s\rangle|^\alpha \lambda(ds) \leq C_2|p|^\alpha.$$

(A2-q) Les coefficients b, σ et leurs dérivées jusqu' à l'ordre q sont uniformément bornées. Ainsi, pour $1 < \alpha < 2$, $B(x) := b(x) + \sigma(x)\gamma$ est uniformément borné. On suppose pour $0 < \alpha \leq 1$, que $B(x) = 0$ pour tout $x \in \mathbb{R}^d$.

(A3) Il existe $\Lambda \geq 1$ tel que pour tout $x \in \mathbb{R}^d, \xi \in \mathbb{R}^d$, $\Lambda^{-1}|\xi|^2 \leq \langle \sigma(x)\xi, \xi\rangle \leq \Lambda|\xi|^2$.

Les hypothèses **(A1-q)** et **(A3)** sont des hypothèses classiques de non-dégénérescence, de la mesure sphérique λ et du coefficient σ.

La condition de dérive nulle dans **(A2-q)** provient du fait que pour $\alpha \in (0, 1]$ l'ajout d'une dérive d'ordre t ne correspond pas en temps petit à un terme négligeable par rapport à l'échelle caractéristique en $t^{1/\alpha}$. Nous renvoyons à l'Appendice B de [**9**] pour des détails.

Théorème 3.3.1 *Supposons $q > d+4$ et que* **(A1-q)**, **(A2-q)** *et* **(A3)** *sont vérifiées. Soit $0 < M \leq q-(d+4)$. Il existe une fonction $R_M(T, x, y)$, où $|R_M(T, x, y)| \leq C_M(T)\left(\frac{1}{1+|y-x|^{d+\alpha}}\right) := \rho_{\alpha,M}(T, y-x)$ pour une constante positive $C_M(T)$, telle que*

$$(p - p^h)(T, x, y) = \sum_{l=1}^{M-1} \frac{h^l}{(l+1)!}\pi_l^h(T, x, y) + h^M R_M(T, x, y),$$

où $\sum_{l=1}^{M-1} \left|\pi_l^h(T, x, y)\right| \leq \rho_{\alpha,M}(T, y - x)$, i.e. les contrôles sur les $(\pi_l^h)_{l\in[\![1,M-1]\!]}$ sont uniformes en h.

Les termes dans le développement précédent dépendent de h. Il est néanmoins possible de faire des développements supplémentaires des quantités π_l^h de sorte à obtenir des termes indépendants de h dans l'expression ci-dessus, ce que l'uniformité en h du contrôle laisse d'ailleurs à penser.

Le théorème précédent permet également de relâcher l'hypothèse de bornitude de la fonction test de [Hau02] et d'obtenir un développement de l'erreur faible $\mathcal{E}_h^D := \mathbb{E}_x[g(X_T^h)] - \mathbb{E}_x[g(X_T)]$ pour des fonctions g mesurables vérifiant la condition de croissance $\exists C > 0,\ \forall x \in \mathbb{R}^d,\ |g(x)| \leq C(1 + |x|^\beta), \beta < \alpha$, qui est optimale pour le cas considéré. En particulier aucune hypothèse de régularité sur g n'est ici requise, contrairement aux approches utilisées dans [TT90], [PT97]. Rappelons que d'un point de vue numérique, le développement de l'erreur faible permet d'utiliser la méthode d'extrapolation de Romberg Richardson pour accélérer la convergence, cf. Chapitre 1.3.1.

Enfin, le développement du Théorème 3.3.1 est utile pour étudier la sensibilité de l'erreur faible $\mathcal{E}_h^D$ par rapport à x sans conditions supplémentaires sur g. Ce point est fondamental pour les applications en finance (couverture). Nous renvoyons à Guyon [Guy06] pour des détails das le cas diffusif.

Extensions et perspectives

Nous avons dans les articles **[8]**,**[9]** utilisé la méthode parametrix de McKean et Singer qui dans des cas non dégénérés ou faiblement dégénérés est très robuste et se prête naturellement à la discrétisation.

Pour l'instant les théorèmes limites locaux sur les densités ont été obtenus à partir des développements parametrix (3.6), (3.13) et du théorème limite local en régime Gaussien. Il serait intéressant d'étendre ce type de résultats à des processus dont la densité n'a pas de comportement de type Gaussien. A titre d'exemple, si l'on remplace en Section 3.3 les accroissements stables par des variables i.i.d. renormalisées qui sont dans le bassin d'attraction de la loi stable initiale, quel type de théorème limite local peut-on espérer établir ? Quelle serait la vitesse en fonction de l'indice α du stable ?

Nous avons également jusqu'à présent toujours énoncé des résultats à horizon $T > 0$ fini. Dans le cas où le processus admet une mesure invariante la question de l'existence du théorème limite local et de la vitesse associée pour une estimation par chaîne de Markov, éventuellement à pas décroissant, reste ouverte.

Chapitre 4

Approximation d'Equations Différentielles Rétrogrades (EDSR)

Nous présentons dans ce chapitre les travaux **[2]**, **[5]** (en collaboration avec F. Delarue de l'Université de Nice), **[6]** (en collaboration avec B. Bouchard de l'Université Paris Dauphine) relatifs à l'approximation d'équations différentielles stochastiques (progressives) rétrogrades (EDS(P)R). Introduites par Pardoux et Peng [PP90], les EDSR ont permis d'étendre les réprésentations de Feynman-Kac à des EDP paraboliques ou elliptiques semi ou quasi-linéaires. Ce sont par ailleurs des objets qui apparaissent de façon naturelle en finance, cf. El Karoui *et al.* [KPQ97], dans la mesure où les problèmes de réplication d'options se formulent directement en terme d'EDSR.

C'est un sujet qui a fait l'objet d'une active recherche tant du point de vue théorique que numérique. Nous présenterons en Section 4.1 un bref aperçu de la théorie des EDSR et de leurs liens avec les EDP non linéaires. En terme numérique l'approximation d'EDSR fournit un procédé alternatif aux méthodes déterministes de différences/éléments finis pour les problèmes non linéaires. Nous verrons ci-après que la complexité des algorithmes proposés dépend de façon hautement non linéaire de la dimension du problème sous-jacent. On ne retrouve donc pas la complexité linéaire en la dimension spécifique des méthodes de Monte Carlo pour l'approximation probabiliste de solutions d'EDP linéaires et l'on se heurte à la "fléau de la dimension" bien connue dans les approximations numériques déterministes d'EDP.

Cette compléxité provient de l'estimation d'espérances conditionnelles, qui est une étape inévitable pour l'approximation d'EDSR. Trois grands types de techniques ont été introduites à cet effet :

- Les méthodes de Monte Carlo associées à des estimateurs dérivant du calcul de Malliavin. Cette approche a été développée par Bouchard et Touzi [BT04] pour des EDSR faiblement couplées.
- Les méthodes de régression linéaire, utilisées par Gobet *et al.* [GLW05] toujours dans le cadre des EDSR faiblement couplées.
- Les méthodes de quantification, introduites par Bally et Pagès dans le cadre des EDSR réfléchies, problème d'arrêt optimal pour les options américaines [BP03]. Le premier algorithme de discrétisation d'EDSR proposé par Chevance [Che97] dans un cadre plus restrictif utilise pour l'estimation des espérances conditionnelles une méthode qui s'apparente à cette approche. C'est par ailleurs celle que nous avons utilisée dans **[2]** et **[5]**.

Nous présenterons brièvement en Section 4.2 les algorithmes de [BT04] et [GLW05]. En retenant l'approche par quantification pour l'approximation des espérances conditionnelles, nous avons pu dans **[2]**, **[5]** proposer des schémas de discrétisation d'EDSR fortement couplées associées à des EDP quasi-linéaires. Nous détaillerons ces algorithmes en Section 4.3. Enfin, nous développerons en Section 4.4, les résultats de **[6]** où nous avons analysé l'erreur forte d'un schéma de semi-discrétisation, i.e. on ne propose pas de schéma d'approximation des espérances conditionnelles, pour une EDSR dans un domaine.

4.1 Quelques résultats sur les EDSR et lien avec les EDP

4.1.1 Introduction à la solvabilité des EDSR et couplage faible

La dynamique d'une EDSR s'écrit de façon générique

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \ t \in [0, T], \tag{4.1}$$

où $(W_t)_{t\geq 0}$ est un mouvement Brownien standard d-dimensionnel sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ de filtration naturelle $(\mathcal{F}_t)_{t\geq 0}$ et $T > 0$ est un instant arbitraire. La condition terminale ξ est $\mathcal{F}_T$ mesurable et généralement supposée de carré intégrable. Le coefficient f est appelé *générateur* de l'EDSR. Ce dernier peut être aléatoire à condition de supposer que le processus $(f(t, y, z))_{t\geq 0}$ est pour tout (y, z) progressivement mesurable.

Un exemple typique de dépendance aléatoire du générateur est fourni par le couplage faible avec une EDS progressive, i.e. f s'écrit $f(s, X_s, Y_s, Z_s)$, $\xi = g(X_T)$ où X est la solution d'une EDS progressive. L'EDSR faiblement couplée s'écrit alors :

$$\begin{aligned} Y_t &= g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \\ X_t &= x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \ t \in [0, T]. \end{aligned} \tag{4.2}$$

Ce sont les EDSR de type (4.2) qui permettent d'établir une connexion avec les EDP semi-linéaires.

Principes de résolution

Résoudre une EDSR de type (4.1) consiste à trouver un **couple** $(Y_t, Z_t)_{t\in[0,T]}$ progressivement mesurable tel que (4.1) soit vérifiée pour tout $t \in [0, T]$, $\mathbb{P}$-presque sûrement. L'hypothèse de progressive mesurabilité est essentielle. Le terme de droite dans (4.1) est *a priori* $\mathcal{F}_T$ mesurable là où l'on cherche à ce que celui de gauche soit $\mathcal{F}_t$ mesurable. C'est le rôle du processus $(Z_t)_{t\in[0,T]}$ que de ramener condition terminale et générateur dans la bonne tribu.

Ce point s'illustre directement dans le cas d'un générateur nul. En effet, si $f = 0$ et que l'on considère une condition terminale ξ de carré intégrable, le théorème de représentation des martingales donne l'existence d'un unique processus $(Z_t)_{t\in[0,T]}$ progressivement mesurable et de carré intégrable tel que $\xi = \mathbb{E}[\xi] + \int_0^T Z_s dW_s$. Ainsi $Y_t := \mathbb{E}[\xi|\mathcal{F}_t] = \mathbb{E}[\xi] + \int_0^t Z_s dW_s$ est bien solution de (4.1). L'unicité découle de celle du processus Z de représentation.

En bref, c'est la condition de mesurabilité sur le couple $(Y_t, Z_t)_{t\in[0,T]}$ qui permet de "fermer" l'équation (4.1) à deux inconnues : à la relation différentielle on ajoute la condition de mesurabilité.

Sous des conditions de structure du générateur f, l'équation (4.1) est alors bien posée. Lorsque f est Lipschitzien en (y, z), on peut construire une suite $(Y^n, Z^n)_{n\in\mathbb{N}}$ définie par $(Y^0, Z^0) = 0, Y_t^{n+1} = \xi + \int_t^T f(s, Y_s^n, Z_s^n)ds - \int_t^T Z_s^{n+1} dW_s$, $t \in [0, T]$, $n \geq 0$. La condition de Lipschitz continuité de f permet d'identifier le processus Z^{n+1} au processus de représentation associé à la variable aléatoire $\xi + \int_t^T f(s, Y_s^n, Z_s^n)ds$ qui est bien de carré intégrable. Les arguments usuels d'analyse différentielle, lemme de Gronwall, et stochastique, inégalités BDG, permettent de déduire que cette suite admet un unique point fixe solution forte de (4.1). La même démarche reste possible si l'on remplace l'hypothèse de Lipschitz continuité du générateur en y par une condition de monotonie associée à une croissance au plus linéaire. Ces résultats de solvabilité sont initialement dûs à Pardoux et Peng [PP90]. Par la suite de nombreux auteurs ont affaibli ces hypothèses pour la solvabilité de (4.1), indiquons Briand *et al.* [BDH+03] pour le cas d'une condition terminale $L^p, p \in (1, 2)$ ou Kobylanski [Kob00] pour le cas d'un générateur à croissance quadratique en Z, hypothèse par ailleurs naturelle du point de vue des EDP, cf. Ladyzhenskaya *et al.* [LSU68].

Lien avec les EDP semi-linéaires

Nous considérerons pour l'illustration du lien entre EDSR et EDP le cas parabolique, le cas elliptique a été abordé par Darling et Pardoux [DP97].

De l'EDP vers l'EDSR. Comme pour la formule de Feynman-Kac dans le cadre linéaire c'est l'application de la formule d'Itô qui permet de relier EDP et EDSR. Considérons l'EDP parabolique semi-linéaire

$$\begin{cases} (\partial_t + L_t)u(t, x) + f(t, x, u(t, x), ((\sigma\nabla u)(t, x)) = 0, \ (t, x) \in [0, T) \times \mathbb{R}^d, \\ u(T, x) = g(x), \ x \in \mathbb{R}^d, \end{cases} \tag{4.3}$$

où $L_t\varphi(x) = \langle b(t, x), \nabla\varphi(x)\rangle + \frac{1}{2}\mathrm{tr}\left(a(t, x)D_x^2\varphi\right)$, $a(t, x) = \sigma\sigma^*(t, x)$ est le générateur infinitésimal d'une équation de diffusion $X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s$ soluble au moins au sens faible. S'il existe une solution "classique" à (4.3), i.e. bornée et à dérivées d'ordres un en temps et deux en espace bornées, il est possible d'appliquer sur l'espace de probabilité sous-jacent la formule d'Itô au processus $Y_t := u(t, X_t)$. On identifie alors $Z_t = (\nabla u\sigma)(t, X_t)$ qui permet de retrouver la formulation couplée de l'équation (4.2).

Ainsi, la donnée d'une solution régulière à l'EDP a permis de construire par un argument de vérification une solution à l'EDSR. Cette méthode peut s'étendre à des solutions de (4.3) ayant des dérivées temps espace au sens de Sobolev à l'aide de la formule d'Itô-Krylov, cf. [Kry96].

De l'EDSR vers l'EDP. Réciproquement, une question naturelle consiste à essayer de retrouver l'EDP à partir de la solution de l'EDSR. Supposons que l'EDSR (4.2) admette une solution unique et définissons u : $(t,x) \in [0,T] \times \mathbb{R}^d \mapsto u(t,x) := Y_t^{t,x}$ qui désigne la valeur de l'EDSR (4.2) en t pour la condition initiale $X_t = x$. De cette définition de u il est aisé de voir que $u(0,x)$ s'écrit comme la solution en 0 d'une EDSR sur $[0,t]$ avec condition terminale $u(t, X_t^{0,x})$. Ce principe de programmation dynamique satisfait par u permet d'établir qu'elle est *solution de viscosité* de (4.3). Nous renvoyons à l'article de Crandall *et al* [CIL92] pour la définition.

Lorsque les coefficients de l'EDSR sont réguliers il est possible de la dériver par rapport à la condition initiale à l'aide de techniques usuelles de flots stochastiques sur les composantes progressives et rétrogrades, voir Kunita [Kun90] ou Ikeda et Watanabe [IW89]. Cette opération permet d'obtenir la différentiabilité de u à l'ordre un en temps et deux en espace. Le calcul de Malliavin permet ensuite l'identification de $(Z_t)_{t\in[0,T]}$ avec $((\sigma\nabla u)(t,X_t))_{t\in[0,T]}$ dont on déduit ensuite grâce à (4.2) que u est solution classique de (4.3). Cette démarche est celle de l'article de Pardoux et Peng [PP92]. Notons que dans cette approche aucune condition de non dégénérescence n'est a priori requise, seule la régularité des coefficients intervient. Si l'équation de départ est en revanche non dégénérée, cette approche n'exploite pas l'effet régularisant du noyau de la chaleur.

4.1.2 EDSR fortement couplées

De nombreuses applications, équations des milieux poreux, problème de gros investisseur en finance, font intervenir des équations aux dérivées partielles pour lesquelles la solution et/ou son gradient apparaissent dans l'opérateur différentiel. On parle alors d'EDP quasi-linéaires. De façon générale celles-ci s'écrivent :

$$\begin{cases} \partial_t u(t,x) + \langle b\left(t,x,u(t,x),(\sigma\nabla u)(t,x)\right), \nabla u(t,x)\rangle + \frac{1}{2}\mathrm{tr}\left(a\left(t,u(t,x)\right) D_x^2 u(t,x)\right) \\ \qquad + f\left(t,x,u(t,x),(\sigma\nabla u)(t,x)\right) = 0,\ (t,x) \in [0,T) \times \mathbb{R}^d, \\ u(T,x) = g(x),\ x \in \mathbb{R}^d. \end{cases} \tag{4.4}$$

Ce sont des équations de ce type pour lesquelles nous avons proposé dans **[2]**, **[5]** des algorithmes de discrétisation que nous détaillerons en Section 4.3. En gardant à l'esprit les relations $Y_t = u(t,X_t), Z_t = (\sigma\nabla u)(t,X_t)$ qui lient l'EDSR à l'EDP il apparaît naturel que l'objet probabiliste naturellement associé à (4.4) sera une EDSR "fortement couplée" de dynamique

$$\begin{aligned} X_t &= x + \int_0^t b(s,X_s,Y_s,Z_s)ds + \int_0^t \sigma(s,X_s,Y_s)dW_s \\ Y_t &= g(X_T) + \int_t^T f(s,X_s,Y_s,Z_s)ds - \int_t^T Z_s dW_s. \end{aligned} \tag{4.5}$$

Le couplage est dit fort dans la mesure où les coefficients de la composante progressive dépendent de l'équation rétrograde. Il n'est ainsi plus possible de résoudre l'équation progressive de façon autonome ce qui induit une difficulté supplémentaire par rapport à la Section précédente.

Solvabilité en temps court. Le cadre naturel de résolution de (4.5) est celui de coefficients Lipschitziens. On cherche alors à construire un triplet $(X_t,Y_t,Z_t)_{t\in[0,T]}$ adapté tel que $\mathbb{E}[\sup_{t\in[0,T]}|X_t|^2] + \mathbb{E}[\sup_{t\in[0,T]}|Y_t|^2] < +\infty$, et $\mathbb{E}[\int_0^T |Z_s|^2 ds] < +\infty$ avec Z progressivement mesurable. L'idée naturelle pour obtenir une telle solution consiste à construire, en suivant la démarche du travail de [PP90], une contraction. Les espaces fonctionnels sont ceux naturellement associés aux conditions ci-avant. Si l'on se donne X_0 on définit pour $n \geq 0$, $Y_t^{n+1} = g(X_T^n) + \int_t^T f(s,X_s^n,Y_s^{n+1},Z_s^{n+1})ds - \int_t^T Z_s^{n+1}dW_s, X_t^{n+1} = x + \int_0^t b(s,X_s^{n+1},Y_s^{n+1},Z_s^{n+1})ds + \int_0^t \sigma(s,X_s^{n+1},Y_s^{n+1})dW_s$. Sous les hypothèses de Lipschitz continuité uniforme des coefficients b,σ,f,g, Antonelli [Ant93] a montré que le procédé précédent définit une contraction et que son unique point fixe est l'unique solution forte de (4.5) si $T \leq c_K$ où K désigne la constante de Lipschitz des coefficients. Sans autres hypothèses ce résultat est optimal dans la mesure où il existe des contre-exemples. Considérons le cas particulier déterministe $X_0 = x, Y_T = -X_T, dX_t = Y_t dt, dY_t = -X_t dt,\ t \in [0,T]$, pour lequel $Z = 0$. La forme générale des solutions s'écrit $X_t = A\cos(t) + B\sin(t), Y_t = -A\sin(t) + B\cos(t),\ t \in [0,T]$ ce qui impose $A = x$ et $A\cos(T) + B\sin(T) = A\sin(T) - B\cos(T)$ pour vérifier les conditions de bord. Pour $T = 3\pi/4$ cette condition conduit à $A = x$ et $A = 0$, ce qui met la solvabilité en défaut.

Schéma en quatre temps. Cette approche est due à Ma, Protter et Yong [MPY94]. L'idée sous-jacente est de "décorréler" les composantes progressives et rétrogrades de (4.5) en utilisant des propriétés de la solution supposée connue de l'EDP (4.4) sous-jacente. Précisément, si u solution de (4.4) est une solution "classique", l'identification formelle $Y_t = u(t, X_t), Z_t = (\sigma\nabla u)(t, X_t)$ précédente conduit à considérer l'équation progressive autonome

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

Si les coefficients, b, σ ainsi que u et son gradient sont Lipschitziens, alors l'équation ci-dessus est fortement soluble. Sous des hypothèses usuelles de croissance du générateur f, l'application de la formule d'Itô donne que le triplet $(X_t, Y_t, Z_t)_{t\in[0,T]} = (X_t, u(t, X_t), (\sigma\nabla u)(t, X_t))_{t\in[0,T]}$ est bien solution de (4.5).

C'est encore là une méthode de type *vérification.* On construit l'objet probabiliste à partir des propriétés de l'EDP. Pour l'unicité, partant d'une solution arbitraire $(\bar{X}, \bar{Y}, \bar{Z})$ de (4.5) il faut l'identifier avec (X, Y, Z) construite précédemment. L'idée est ici d'utiliser de nouveau la fonction u solution de l'EDP (supposée bornée ainsi que ses dérivées d'ordre un et deux en espace) pour identifier $\bar{Y}_t = u(t, \bar{X}_t), \bar{Z}_t = (\sigma\nabla u)(t, \bar{X}_t)$. Pour cela, on applique la formule d'Itô à $(u(t, \bar{X}_t) - \bar{Y}_t)^2,\ t \in [0, T]$. Les outils usuels en équations rétrogrades : inégalités BDG, inégalités de Young et Lemme de Gronwall, permettent de conclure.

Une fois effectuée l'identification pour $(\bar{Y}, \bar{Z})$, les hypothèses précédentes de Lipschitzianité sur les coefficients donnent que $X = \bar{X}$ car les deux composantes progressives vérifient la même EDS. Dans le cadre d'un générateur Lipschitzien l'identification $(\bar{Y}, \bar{Z}) = (Y, Z)$ provient ensuite de l'unique solubilité de l'équation rétrograde.

Pour pouvoir appliquer ce programme il faut donc disposer de coefficients b, σ, f Lipschitziens et d'une solution classique de (4.4) bornée ainsi que ses dérivées d'ordre un et deux en espace. Un jeu typique d'hypothèses qui garantit ces propriétés est le suivant :

(a) Les coefficients b, f, σ et g sont réguliers et à dérivées d'ordre un en (x, y, z) bornées.

(b) $\exists\Lambda > 0,\ |(b, f)(t, x, y, z)| \leq \Lambda(1 + |y| + |z|),\ |g(x)| \leq \Lambda$.

(c) $\exists\lambda > 0,\ \lambda I_d \leq a(t, x, y) \leq \Lambda I_d$.

(d) $g \in C^2(\mathbb{R}^d)$ a des dérivées d'ordre un et deux bornées et uniformément η-Hölderiennes, $\eta \in (0, 1]$.

Théorème 4.1.1 (cf. Ladyzhenskaya *et al.* [LSU68]) *Supposons que (a),(b),(c),(d) soient satisfaites. Alors, l'équation admet une unique solution classique bornée ainsi que ses dérivées d'ordre un et deux en espace.*

C'est sous des hypothèses de ce type que nous nous placerons en Section 4.3. Nous reprendrons pour étudier la convergence de schémas d'approximation d'EDP/EDSR la démarche présentée ci-dessus pour l'unicité en exploitant les propriétés de u.

Solvabilité en temps arbitraire dans le cadre d'hypothèses "usuelles".

Sous l'hypothèse d'uniforme ellipticité (c) et de coefficients Lipschitziens, l'unique solvabilité de (4.5) pour $T > 0$ arbitraire a été établie par Delarue [Del02]. Pour passer du temps court au temps long, l'idée consiste à itérer le résultat en temps court de [Ant93] en contrôlant que la solution de viscosité de (4.4) associée à l'EDSR ainsi construite, $u(t, x) := Y_t^{t,x}$, reste Lipschitzienne en espace (au sens où sa constante de Lipschitz n'explose pas) sur chacun des intervalles de temps "courts". Ceci permet en effet d'itérer la procédure en considérant comme nouvelle condition terminale la solution précédemment obtenue et Lipschitzienne en espace. Celle-ci sert alors d'entrée à une nouvelle EDSR fortement couplée fortement uniquement soluble en temps court.

L'hypothèse de non-dégénérescence est ici cruciale pour pouvoir construire une solution pour T arbitraire, comme l'on pouvait l'envisager à partir du contre-exemple ci-avant.

Pour la solvabilité au sens faible d'EDSR fortement couplées dans un cadre non dégénéré nous renvoyons au travail de Delarue et Guatteri [DG06].

4.2 Algorithmes de discrétisation d'EDSR faiblement couplées

Au regard des parties précédentes le cas faiblement couplé apparaît plus favorable à la discrétisation dans la mesure où l'approximation de la composante progressive peut être effectuée de façon autonome. Par ailleurs les méthodes numériques pour celle-ci sont bien comprises, voir Chapitre 1. Nous présentons tout d'abord un premier schéma de semi-discrétisation pour l'EDSR (4.2). On entend par là un schéma théorique qui ne propose pas encore d'approximation des espérances conditionnelles en présence. Nous présenterons ensuite deux techniques numériques pour l'approximation des espérances conditionnelles. Nous analyserons précisément leur complexité algorithmique.

4.2.1 Un premier schéma de semi-discrétisation

Soit $\{(t_i := ih)_{i\in[\![0,N]\!]}\}, Nh = T$, une grille en temps de pas h donnée. Pour discrétiser (4.2), l'idée naturelle consiste à partir de l'écriture "locale" en temps :

$$Y_{t_i} = \mathbb{E}[Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(s, X_s, Y_s, Z_s)|\mathcal{F}_{t_i}],\ i \in [\![0, N-1]\!]. \tag{4.6}$$

Pour $i = N-1$, $Y_{t_{i+1}} = Y_T = g(X_T)$, condition terminale connue. Pour déduire un schéma d'approximation de l'écriture (4.6), si l'on pense par exemple à une méthode des rectangles pour la discrétisation de l'intégrale, il est fondamental d'avoir une approximation du processus de représentation. A ce propos, partant de la formulation (4.2) entre t_i et t_{i+1}, $i \in [\![0, N-1]\!]$, si l'on multiplie les deux membres de la composante rétrograde par l'accroissement Brownien $W_{t_{i+1}} - W_{t_i}$ et que l'on passe à l'espérance conditionnelle en $\mathcal{F}_{t_i}$ il vient :

$$\begin{aligned}\mathbb{E}[\int_{t_i}^{t_{i+1}} Z_s ds|\mathcal{F}_{t_i}] &= \mathbb{E}[(Y_{t_{i+1}} - Y_{t_i})(W_{t_{i+1}} - W_{t_i})|\mathcal{F}_{t_i}] + \mathbb{E}[\int_{t_i}^{t_{i+1}} f(s, X_s, Y_s, Z_s)ds(W_{t_{i+1}} - W_{t_i})|\mathcal{F}_{t_i}] \\ &= \mathbb{E}[Y_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})|\mathcal{F}_{t_i}] + O(h^{3/2}),\end{aligned} \tag{4.7}$$

sous les hypothèses typiques de solvabilité de EDSR. Par ailleurs il est naturel de chercher une approximation de $(Z_s)_{s\in[t_i,t_{i+1}]}$ qui soit $\mathcal{F}_{t_i}$ mesurable. Si l'on néglige la contribution associée au générateur dans (4.7), la meilleure approximation $\mathcal{F}_{t_i}$ mesurable de $(Z_s)_{s\in[t_i,t_{i+1}]}$ dans $L^2([t_i, t_{i+1}] \times \Omega, dt \otimes d\mathbb{P})$ est donnée par :

$$\hat{Z}_{t_i} = h^{-1}\mathbb{E}[Y_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})|\mathcal{F}_{t_i}].$$

Cette technique d'"incrément de martingale" pour l'estimation du Z apparaît initialement chez Bally et Pagès [BP03]. Supposons que l'on dispose désormais d'une approximation $(X^h_{t_i})_{i\in[\![0,N]\!]}$ de la composante progressive de (4.2), par exemple le schéma d'Euler. L'itération dérivant des écritures précédentes s'écrit :

$$\begin{aligned}Y^h_{t_i} &= \mathbb{E}[Y^h_{t_{i+1}}|\mathcal{F}_{t_i}] + h\mathbb{E}[f(t_i, X^h_{t_i}, Y^h_{t_{i+1}}, \hat{Z}^h_{t_i})|\mathcal{F}_{t_i}], \\ \hat{Z}^h_{t_i} &= h^{-1}\mathbb{E}[Y^h_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})|\mathcal{F}_{t_i}],\ i \in [\![0, N-1]\!].\end{aligned} \tag{4.8}$$

La vitesse de convergence de (4.8) a été obtenue pas Zhang [Zha04] sous des hypothèses usuelles de solvabilité. Précisément, lorsque X^h est le schéma d'Euler, il établit qu'il existe C dépendant de T et des constantes de Lipschitz des coefficients dans (4.2) tel que :

$$\max_{i\in[\![0,N]\!]} \mathbb{E}[|Y^h_{t_i} - Y_{t_i}|^2] + \sum_{i=0}^{N-1} \mathbb{E}[\int_{t_i}^{t_{i+1}} |\hat{Z}^h_{t_i} - Z_s|^2 ds] \leq Ch.$$

On retrouve que l'erreur obtenue sur la composante rétrograde est du même ordre que celle de l'équation progressive, voir également Gobet et Labart [GL07] pour un développement d'erreur qui met en évidence cette propriété.

Pour passer à partir de (4.8) à un schéma "implémentable" il est nécessaire de préciser une méthode d'approximation des espérances conditionnelles.

4.2.2 Méthodes de Monte Carlo et calcul de Malliavin

Cette approche a été utilisée par Bouchard et Touzi [BT04] à partir d'un schéma légèrement différent de (4.8), i.e. l'itération sur la composante rétrograde est implicite $Y^h_{t_i} = \mathbb{E}[Y^h_{t_{i+1}}|\mathcal{F}_{t_i}] + hf(t_i, X^h_{t_i}, Y^h_{t_i}, \hat{Z}^h_{t_i})$. Sous les hypothèses usuelles et pour h suffisamment petit, ce schéma est également bien défini.

Lorsque le coefficient de diffusion σ est uniformément elliptique, et que les coefficients b, σ sont réguliers, le schéma d'Euler X^h possède une densité. L'estimation des espérances conditionnelles se réduit ainsi à un calcul de lois conditionnelles. Les auteurs utilisent ensuite des formules d'intégration par partie pour représenter la densité et des fonctions du schéma d'Euler.

Soit une variable aléatoire $R := \rho(X^h_{t_{i+1}}, \xi)A(W_{t_{i+1}} - W_{t_i})$ où ρ et une fonction à valeurs réelles, A une application affine et ξ une variable aléatoire *indépendante* de $\sigma\{(X^h_{t_i}), i \in [\![1, N]\!]\}$. Les auteurs établissent l'expression

$$\mathbb{E}[R|X^h_{t_i} = x] := \frac{\mathbb{E}[H_x(X^h_{t_i})\rho(X^h_{t_{i+1}}, \xi)\mathcal{S}(A(W_{t_{i+1}} - W_{t_i}), x, X^h_{t_i})]}{\mathbb{E}[H_x(X^h_{t_i})\mathcal{S}(1, x, X^h_{t_i})]}, \tag{4.9}$$

où $H_x(y) = \prod_{j=1}^d \mathbb{I}_{x_j \leq y_j}$ est un produit de fonctions de Heaviside résultant de l'intégration par parties. L'expression de $\mathcal{S}$ fait intervenir des intégrales de Skorohod et un terme de localisation qui permet de contrôler la variance pour l'estimation de Monte Carlo associée.

Ainsi, pour une trajectoire du schéma d'Euler, l'estimation par Monte Carlo de la quantité ci-avant va nécessiter la simulation de $M+1$ trajectoires i.i.d. du schéma d'Euler (en incluant la trajectoire initiale notée $X^{h,0} := X^h$). Dans la pratique $M = KN$ et à chaque instant $(t_i)_{i\in[\![0,N-1]\!]}$ de discrétisation K trajectoires des $\left((X_{t_i}^{h,k})_{i\in[\![0,N]\!]}\right)_{k\in[\![0,M]\!]}$ sont utilisées pour l'approximation de l'espérance conditionnelle. Ces K variables indépendantes interviennent implicitement dans l'expression de ξ de l'équation (4.9) lorsque R est associée à l'approximation de l'EDSR. Précisément, si l'on note $\mathcal{N}_i := \{(i-1)K+1, ..., iK\}, i \in [\![1,N]\!]$ le schéma s'écrit : $\forall l \in \mathcal{N}_N,\ Y_T^{h,l} = g(X_T^{h,l}),\ \forall j \in \{0\} \cup \mathcal{N}_i,\ i \in [\![0,N-1]\!]$,

$$\begin{aligned}
\hat{Z}_{t_i}^{h,j} &= \frac{\frac{1}{K}\sum_{l\in\mathcal{N}_{i+1}} H_{X_{t_i}^{h,j}}(X_{t_i}^{h,l})\hat{Y}_{t_{i+1}}^{h,l}\mathcal{S}_i(W_{t_{i+1}}^l - W_{t_i}^l, X_{t_i}^{h,j}, X_{t_i}^{h,l})}{\frac{1}{K}\sum_{l\in\mathcal{N}_{i+1}} H_{X_{t_i}^{h,j}}(X_{t_i}^{h,l})\mathcal{S}_i(1, X_{t_i}^{h,j}, X_{t_i}^{h,l})}, \\
\hat{Y}_{t_i}^{h,j} &= \frac{\frac{1}{K}\sum_{l\in\mathcal{N}_{i+1}} H_{X_{t_i}^{h,j}}(X_{t_i}^{h,l})\hat{Y}_{t_{i+1}}^{h,l}\mathcal{S}_i(1, X_{t_i}^{h,j}, X_{t_i}^{h,l})}{\frac{1}{K}\sum_{l\in\mathcal{N}_{i+1}} H_{X_{t_i}^{h,j}}(X_{t_i}^{h,l})\mathcal{S}_i(1, X_{t_i}^{h,j}, X_{t_i}^{h,l})} + hf(t_i, X_{t_i}^{h,j}, Y_{t_i}^{h,j}, \hat{Z}_{t_i}^{h,j}).
\end{aligned}$$

La fonction A dans (4.9) est soit l'accroissement Brownien pour la mise à jour du processus de représentation approché, soit 1 pour la composante rétrograde. L'analyse de l'erreur associée, cf. Théorème 6.2 de [BT04] donne que pour retrouver l'erreur attendue d'ordre $h^{1/2}$ en norme L^2, qui était celle du schéma de semi-discrétisation, il faut prendre $K \geq N^{6+d/2}$, ce qui conduit à considérer $N^{7+d/2}$ trajectoires du schéma d'Euler.

La complexité apparente d'une telle procédure est de l'ordre de K^2N. Néanmoins les fonctions de Heaviside intervenant dans l'estimateur permettent de passer à $K\log(K)^{(d-1)\vee 1}N$ en triant les réalisations. Ce type d'estimateur reste délicat à calibrer dans la pratique, en particulier le paramètre de troncation implicite dans $\mathcal{S}$. Enfin, il faut également approcher les intégrales de Skorohod dans ce même terme.

4.2.3 Techniques de régression

Une autre approche est basée sur des méthodes de régression. Elle a été développée par Gobet, Lemor et Warin [GLW05, GLW06] selon une démarche voisine de celle proposée par Longstaff et Schwarz [LS01] dans le cadre des options américaines. L'idée sous-jacente est d'utiliser M trajectoires $\left((X_{t_i}^{h,j})_{i\in[\![0,N]\!]}\right)_{j\in[\![1,M]\!]}$ indépendantes du schéma d'Euler et de "régresser" les approximations naturelles associées à l'EDSR, voir (4.8), sur une base de fonctions de dimension finie. La phase d'initialisation de l'algorithme reste la même et consiste à poser $Y_T^{h,j} = g(X_T^{h,j}),\ j \in [\![1,M]\!]$. Pour X à valeurs dans $\mathbb{R}^d$, si l'on se donne $(p_l)_{l\in[\![1,d]\!]}$ bases de fonctions de cardinaux respectifs $(K_l)_{l\in[\![1,d]\!]}$, la première étape pour la mise à jour de l'approximation du Z à l'instant $(t_i)_{i\in[\![0,N-1]\!]}$ consiste à résoudre le problème de moindre carré :

$$\inf_{\alpha\in\mathbb{R}^{K_l}} \frac{1}{M}\sum_{j=1}^M \left| h^{-1}Y_{t_{i+1}}^{h,j}(W_{t_{i+1}} - W_{t_i})_l - \langle \alpha, p_l(X_{t_i}^{h,j})\rangle \right|^2,\ l \in [\![1,d]\!].$$

Si $\hat{\alpha}_{t_i}^l$ désigne le α optimal, on pose ensuite $(\hat{Z}_{t_i}^{h,j})_l = \langle \hat{\alpha}_{t_i}^l, p_l(X_{t_i}^{h,j})\rangle,\ l \in [\![1,d]\!]$. L'estimation de $Y_{t_i}^{h,j}$ se fait suivant le même schéma. Si

$$\hat{\alpha}_{t_i}^0 = \mathrm{arginf}_{\alpha\in\mathbb{R}^{K_0}} \left\{ \frac{1}{M}\sum_{j=1}^M \left| Y_{t_{i+1}}^{h,j} + hf(t_i, X_{t_i}^{h,j}, Y_{t_{i+1}}^{h,j}, \hat{Z}_{t_i}^h) - \langle \alpha, p_0(X_{t_i}^{h,j})\rangle \right|^2 \right\},$$

où p_0 désigne une base de fonctions de cardinal K_0, on pose $Y_{t_i}^{h,j} = \langle \hat{\alpha}_{t_i}^0, p_0(X_{t_i}^{h,j})\rangle$. L'approximation est ici comme dans (4.8) explicite.

Nous renvoyons aux articles sus-mentionnés pour le détail de l'analyse de la convergence. La complexité globale de l'algorithme est de l'ordre de $NMd\log(K)$ si l'on choisit $K = K_l,\ l \in [\![0,d]\!]$ et des fonctions de base qui sont des indicatrices de cellules. Pour retrouver une erreur de l'ordre de $h^{1/2}$ en norme L^2 sous de bonnes hypothèses de Lipschitz continuité des coefficients, l'équilibre des paramètres N, M, K conduit à une complexité algorithmique l'ordre de $N^{2d+4}\log(N)$ à mettre en regard avec l'ordre de $N^{7+d/2}\log(N)$ de l'algorithme précédent pour obtenir cette vitesse. En revanche, l'approche par régression ne requiert aucune non dégénérescence des coefficients et peut s'étendre à des EDSR faiblement couplées qui font intervenir dans

leur dynamique rétrograde les variations d'une martingale orthogonale au Brownien ainsi que d'éventuelles composantes de sauts pour la partie progressive [GLW06].

Nous voyons sur ces exemples que même dans le cas faiblement couplé les complexités algorithmiques associées à ces modèles non-linéaires sont hautement explosives avec la dimension.

4.3 Discrétisation d'EDSR fortement couplées

La discrétisation d'une EDSR fortement couplée présente les mêmes difficultés que celles rencontrées pour la construction des objets continus, cf. Section 4.1.2. Il n'est pas possible de séparer estimations des composantes progressives et rétrogrades à moins de les "décorréler" complètement, c'est à dire de reproduire numériquement la procédure de point fixe qui a permis de construire l'objet continu en résolvant une suite d'EDSR faiblement couplées. Dans ce cadre, il devient possible numériquement d'utiliser les algorithmes de la section précédente. Nous renvoyons pour cette approche aux travaux de Rivière [Riv05] et Bender et Zhang [BZ08].

Nous avons choisi pour les articles **[2]**, **[5]** d'utiliser les liens forts entre l'EDSR et l'EDP quasi-linéaire et d'exploiter que, sous les hypothèses du Théorème 4.1.1, la solution u de (4.4) est régulière. Cette approche est également celle retenue par Milstein et Tretyakov [MT04]. L'idée commune à ces travaux est de procéder à une itération rétrograde et à chaque étape en temps d'injecter dans la dynamique de l'équation progressive les estimées de la solution et de son gradient obtenues à l'étape précédente. La régularité de la vraie solution permet de contrôler la propagation de l'erreur due à cette approximation. De fait, cette procédure fournit un algorithme probabiliste d'approximation de la solution de l'EDP en exploitant la dynamique locale en temps de l'EDSR (4.5). L'approximation de l'EDSR elle même se fait dans un second temps après la résolution approchée de l'EDP de façon similaire à Douglas *et al.* [DMP96] qui proposaient eux un algorithme déterministe pour (4.4) avant d'injecter les résultats de l'approximation dans l'EDSR.

4.3.1 Etapes de la discrétisation

Posons pour la suite pour tout $(t,x) \in [0,T] \times \mathbb{R}^d, v(t,x) := \big(\sigma \nabla u\big)(t,x)$. Pour $t \in [0, T-h]$, l'équation (4.5) et le lien avec u solution de (4.4) donnent le principe de programmation dynamique :

$$u(t, X_t) := \mathbb{E}[u(t+h, X_{t+h})|\mathcal{F}_t] + \mathbb{E}[\int_t^{t+h} f(s, X_s, Y_s, Z_s)ds|\mathcal{F}_t].$$

Cette écriture ne peut pas directement s'exploiter numériquement dans la mesure où la dynamique de X entre t et $t+h$ fait intervenir $(Y_s, Z_s)_{s\in[t,t+h]}$ soit la solution u de (4.4) et son gradient v. Si l'on dispose d'approximations $\big(\tilde{u}(t+h,.), \tilde{v}(t+h,.)\big)$ de $\big(u(t+h,.), v(t+h,.)\big)$ sur tout l'espace, on définit une transition "approchée" entre t et $t+h$:

$$\forall x \in \mathbb{R}^d,\ \mathcal{T}(t,x) := b(t,x,\tilde{u}(t+h,x), \tilde{v}(t+h,x))h + \sigma(t,x,\tilde{u}(t+h,x))(W_{t+h} - W_t).$$

On utilise ensuite cette transition dans la dynamique du processus progressif pour mettre à jour l'estimation du gradient par incrément de martingales, i.e. pour tout $x \in \mathbb{R}^d$,

$$\tilde{v}(t,x) := h^{-1}\mathbb{E}[\tilde{u}(t, x + \mathcal{T}(t,x))(W_{t+h} - W_t)]. \tag{4.10}$$

On définit enfin la nouvelle approximation $\tilde{u}(t,.)$ par

$$\tilde{u}(t,x) \quad := \quad \mathbb{E}[\tilde{u}(t+h, x+\mathcal{T}(t,x))] + hf(t,x,\tilde{u}(t+h,x), \tilde{v}(t,x)). \tag{4.11}$$

Partant de la condition initiale $\tilde{u}(T,.) = g(.), \tilde{v}(T,.) = (\sigma\nabla g)(T,.)$, ce schéma est bien défini le long d'une grille de discrétisation en temps $\{(t_i := ih)_{i\in[\![0,N]\!]}\}$.

Deux obstacles subsistent toutefois pour sa mise en oeuvre :

(a) Le schéma ne peut être défini en pratique sur tout l'espace.

(b) Les espérances doivent être estimées numériquement.

Grilles spatiales et prolongement. Pour le premier point nous nous sommes restreints à une grille spatiale ce qui nous rapproche des algorithmes déterministes. Dans le cas découplé, des simulations autonomes permettent d'avoir une idée *a priori* de la "géométrie" associée à la composante progressive, i.e. zones plus visitées que d'autres. Le cas couplé n'autorise pas une telle procédure et sans autre connaissance *a priori* nous avons retenu des grilles Cartésiennes tronquées. C'est également le choix de Milstein et Tretyakov.

La restriction à une grille spatiale amène à préciser le sens des relations (4.10),(4.11) pour $t = t_i,\ i \in [\![0, N-1]\!]$. En effet, si l'on se restreint en t_{i+1} à l'estimation de u, v sur une grille spatiale $\mathcal{C}_{i+1}$, la transition $x + \mathcal{T}(t_i, x),\ x \in \mathcal{C}_i$ n'appartiendra pas à $\mathcal{C}_{i+1}$. Il est alors nécessaire de comprendre $\tilde{u}(t_{i+1}, .)$ dans (4.10), (4.11) comme un prolongement à tout l'espace de la fonction définie aux points d'une grille spatiale. Dans **[2]** nous avons considéré un prolongement au plus proche voisin, i.e. $\forall (x, i) \in \mathbb{R}^d \times [\![0, N]\!],\ \tilde{u}(t_i, x) = \tilde{u}(t_i, \Pi_{\mathcal{C}_i}(x))$ où $\Pi_{\mathcal{C}_i}$ désigne la projection sur la grille Cartésienne $\mathcal{C}_i$ (avec une convention pour le centre des mailles).

Dans **[5]** nous avons en revanche utilisé une interpolation multilinéaire qui peut par ailleurs s'interpréter de façon naturelle en termes probabilistes. Cette approche a permis de tirer parti de la régularité *a priori* de u sous les hypothèses du Théorème 4.1.1 et d'améliorer la vitesse de convergence de notre approximation.

Quantification. Pour le point (b), nous avons choisi d'utiliser une méthode de quantification plutôt que l'approximation usuelle de Monte Carlo. L'idée de la quantification de variables aléatoires est d'approcher une variable donnée par une loi discrète à support fini. Cette procédure permet de ramener le calcul d'espérances à celui d'une somme finie. Nous renvoyons à l'ouvrage de Graf et Luschgy [GL00] pour une référence théorique à cette technique.

Des équations ci-dessus nous devons quantifier l'accroissement Brownien soit à un changement d'échelle près la loi normale. Une première approximation consisterait à l'approcher par un produit de lois de Bernoulli symétriques correctement normalisées, c'est le choix effectué par Milstein et Tretyakov [MT04]. Nous avons choisi d'utiliser une quantification "optimale" de la loi gaussienne. Si nous notons G la loi gaussienne standard de $\mathbb{R}^d$ et $\hat{G}$ une quantifiée optimale $\hat{G}_M$ à M points (non nécessairement unique) vérifie (pour M suffisamment grand) l'inégalité de "distortion" :

$$\exists C := C(G, d),\ \|G - \hat{G}_M\|_{L^2} \leq C M^{-1/d}. \tag{4.12}$$

Les quantifieurs optimaux vérifient également la propriété projective (ou de stationnarité) $\mathbb{E}[G|\hat{G}_M] = \hat{G}_M$ dont on déduit que $\mathbb{E}[\hat{G}_M] = 0$ et que pour une fonction f de classe C^2 à dérivées bornées,

$$\begin{aligned}\mathbb{E}[f(G) - f(\hat{G}_M)] &= \mathbb{E}[\langle \nabla f(\hat{G}_M), \underbrace{\mathbb{E}[G - \hat{G}_M|\hat{G}_M]}_{=0}\rangle] \\ &\quad + \textstyle\int_0^1 d\lambda (1-\lambda) \mathbb{E}[\langle D^2 f(\hat{G}_M + \lambda(G - \hat{G}_M))(G - \hat{G}_M), G - \hat{G}_M\rangle] \\ &:= R_M,\ |R_M| \leq \tfrac{|D^2 f|_\infty}{2} \mathbb{E}[|G - \hat{G}_M|^2] \leq C M^{-2/d},\end{aligned} \tag{4.13}$$

en exploitant la propriété projective pour la première égalité et (4.12) pour la dernière inégalité.

Ainsi, contrairement à la méthode de Monte Carlo, la quantification permet d'exploiter la régularité des fonctions dont on veut estimer l'espérance. Le comportement de cette méthode se rapproche des techniques de quadrature ce qui se retrouve sur la vitesse de convergence associée qui dépend fortement de la dimension d de l'espace. Intuitivement, pour M points, la vitesse de $M^{-1/d}$ provient d'une répartition du même ordre dans chaque direction. Nous avons tiré tout l'avantage de cette approche dans **[5]**.

4.3.2 Algorithmes

Les deux algorithmes proposés dans **[2]** et **[5]** dérivent de (4.10), (4.11) avec les procédures d'extension précédentes dues aux restrictions à des grilles spatiales. Dans **[2]**, on remplace simplement l'accroissement Brownien par la quantifiée optimale gaussienne renormalisée en considérant la projection au plus proche voisin. Il vient

$$\begin{aligned}\bar{u}(T, x) &= g(x),\ x \in \mathcal{C}_N, \\ \mathcal{T}(t_i, x) &:= b(t_i, x, \bar{u}(t_{i+1}, x), \bar{v}(t_{i+1}, x))h + \sigma(t_i, x, \bar{u}(t_{i+1}, x))h^{1/2}\hat{G}_M, \\ \bar{v}(t_i, x) &= h^{-1}\mathbb{E}[\bar{u}(t_{i+1}, x + \mathcal{T}(t_i, x))h^{1/2}\hat{G}_M], \\ \bar{u}(t_i, x) &= \mathbb{E}[\bar{u}(t_i, x + \mathcal{T}(t_i, x))] + hf(t_i, x, \bar{u}(t_{i+1}, x), \bar{v}(t_i, x)),\ x \in \mathcal{C}_i,\end{aligned} \tag{4.14}$$

pour tout $i \in [\![0, N-1]\!]$.

Dans **[5]**, pour l'analyse de l'erreur, il est nécessaire que la matrice de covariance de la variable aléatoire qui approche la loi normale soit l'identité. Or ce n'est pas forcément le cas pour une quantification $\hat{G}_M$ optimale de G. Pour M assez grand, $\Gamma_M := \mathbb{E}[\hat{G}_M \hat{G}_M^*]$ est inversible et la technique naturelle consiste à préfixer l'incrément $h^{1/2}\hat{G}_M$ par γ_M^* où γ_M est la matrice triangulaire inférieure de la factorisation de Cholesky de $\Gamma_M^{-1} = \gamma_M \gamma_M^*$.

On a ainsi l'algorithme

$$\begin{aligned}
\bar{u}(T,x) &= g(x),\ x \in \mathcal{C}_N, \\
\mathcal{T}(t_i,x) &:= hb(t_i,x,\bar{u}(t_{i+1},x),\bar{v}(t_{i+1},x))h + \sigma(t_i,x,\bar{u}(t_{i+1},x))\gamma_M^* h^{1/2}\hat{G}_M, \\
\bar{v}(t_i,x) &= h^{-1}\mathbb{E}[\bar{u}(t_{i+1},x+\mathcal{T}(t_i,x))h^{1/2}\gamma_M^*\hat{G}_M], \\
\bar{u}(t_i,x) &= \mathbb{E}[\bar{u}(t_i,x+\mathcal{T}(t_i,x))] + hf(t_i,x,\bar{u}(t_{i+1},x),\bar{v}(t_i,x)),\ x \in \mathcal{C}_i, i \in [\![0,N-1]\!],
\end{aligned} \tag{4.15}$$

en considérant l'interpolation multilinéaire comme procédure de prolongement à tout l'espace.

Indiquons que plus que l'aspect *optimal* c'est l'aspect *stationnaire* qui est crucial pour l'analyse de convergence dans [**5**]. Or, il est possible de construire de telles grilles à partir de lois de Bernoulli. Ce point de vue est celui adopté par Milstein et Tretyakov dans le livre [MT04] et les articles [MT06b, MT07]. Nous verrons par ailleurs au paragraphe suivant que lorsque l'on exploite la stationnarité, pour une solution régulière de (4.4) l'erreur de quantification est négligeable devant les autres composantes de l'erreur.

4.3.3 Théorèmes de convergence

Nous présentons les résultats de convergence lorsque le coefficient b dans les algorithmes précédents n'ont pas de dépendance en gradient. Des solutions alternatives pour traiter ce cas ont été proposées dans [**2**] et [**5**]. Dans les deux cas elles font intervenir une transition intermédiaire additionnelle par rapport à $\mathcal{T}$ et induisent une dépendance déraisonnable entre M et h pour pouvoir analyser la convergence de cette méthode. Par ailleurs, cette contrainte n'est nullement justifiée par les observations numériques effectuées dans ces cas là. Nous renvoyons aux sections 7 de [**2**] et 7 de [**5**] pour une description détaillée.

Algorithme avec projection "au plus proche voisin".

Théorème 4.3.1 (Delarue et Menozzi [2]) *Supposons les coefficients b, σ, f dans* (4.4) *homogènes et uniformément Lipschitziens en x, y, z, uniformément bornés en espace et à croissance au plus linéaire en (y, z), b indépendant de z, σ uniformément elliptique et la fonction terminale $g \in C^{2+\alpha}, \alpha \in (0,1]$ bornée et à dérivées bornées, sur le modèle de l'énoncé du Théorème 4.1.1. Si l'on prend des grilles spatiales Cartésiennes de pas d'espace $\delta > 0$ avec $(\mathcal{C}_i)_{i\in[\![1,N]\!]}$ de rayon $R+\rho$, $R, \rho > 0$ et $\mathcal{C}_0$ de rayon R on a le contrôle suivant pour l'erreur associée à l'algorithme* (4.14) *:*

$$\sup_{x\in\mathcal{C}_0} |u(0,x) - \bar{u}(0,x)| \leq C\left(\mathcal{E}(\text{temps}) + \mathcal{E}(\text{espace}) + \mathcal{E}(\text{quantification}) + \mathcal{E}(\text{troncature})\right),$$

où C dépend des hypothèses sur les coefficients et est uniforme par rapport aux paramètres de discrétisation. Les ordres des erreurs sont : $\mathcal{E}(\text{temps}) = h^{1/2}$, $\mathcal{E}(\text{espace}) = h^{-1}\delta$, $\mathcal{E}(\text{quantification}) = h^{-1/2}M^{-1/d}$ et $\mathcal{E}(\text{troncature}) = \rho/(\rho+R)$.

Ces ordres s'expliquent de la façon suivante :

- $\mathcal{E}$(temps). L'ordre 1/2 provient de la régularité en temps de l'accroissement Brownien et de la propriété de 1/2-Hölder continuité en temps de $\nabla_x u$ sous les hypothèses du théorème.
- $\mathcal{E}$(espace). Pour une fonction régulière l'erreur associée à la projection au plus proche voisin sur une grille spatiale Cartésienne de pas δ est d'ordre δ. Comme ce type d'erreur apparaît à chaque itération en temps on obtient bien $h^{-1}\delta$.
- $\mathcal{E}$(quantification). L'erreur de quantification de l'accroissement Brownien est en norme L^1 d'ordre $h^{1/2}M^{-1/2}$ d'après (4.12). La vitesse indiquée correspond de nouveau à la sommation de $N = T/h$ erreurs le long de la grille en temps.
- $\mathcal{E}$(troncature). Ce terme est obtenu par des contrôles de type Tchebychev. Nous n'avons pour l'instant pas su tirer profit de l'échelle caractéristique de temps en $t^{1/2}$ de la diffusion pour mieux calibrer nos grilles spatiales.

Précisons que le résultat est présenté pour des coefficients homogènes en temps par souci de cohérence avec les énoncés de [**2**]. Ce choix avait été motivé par le souci d'alléger les notations. Le résultat ci-dessus resterait vrai sans difficultés techniques additionnelles pour des coefficients uniformément 1/2-Hölder en temps, cf. Théorème 4.1.1.

Algorithme interpolé.

Théorème 4.3.2 (Delarue et Menozzi [5]) *Supposons les coefficients b, σ, f dans* (4.4) *homogènes et uniformément α-Hölderiens, $\alpha \in (0,1]$, en x, uniformément Lipschitziens en y, z, uniformément bornés en espace et à croissance au plus linéaire en (y,z), b indépendant de z, σ uniformément elliptique et la fonction terminale $g \in C^{2+\alpha}$, bornée et à dérivées bornées. Si l'on prend des grilles spatiales Cartésiennes infinies de pas d'espace $\delta > 0$ on a le contrôle suivant pour l'erreur associée à l'algorithme* (4.15) *:*

$$\sup_{x \in \mathcal{C}_0} |u(0,x) - \bar{u}(0,x)| \leq C \left(\mathcal{E}(\text{temps}) + \mathcal{E}(\text{espace}) + \mathcal{E}(\text{quantification}) \right),$$

où C dépend des hypothèses sur les coefficients et est uniforme par rapport aux paramètres de discrétisation. Les ordres des erreurs sont : $\mathcal{E}(\text{temps}) = h^{\alpha/2}$, $\mathcal{E}(\text{espace}) = h^{-1}\delta^2$, $\mathcal{E}(\text{quantification}) = h^{\alpha/2} M^{-2/d}$.

Le propos du travail **[5]** était d'estimer le gain découlant de l'interpolation et de l'utilisation intensive de la stationnarité du quantifieur. Nous n'avons donc pas pris en compte les erreurs associées à la troncation des grilles Cartésiennes qui pourraient être analysées de façon similaire au travail précédent. Sous les hypothèses du théorème nous avons toujours que u est une solution classique de (4.4) [1], cf. Théorème 7.1. du Chapitre 7 de [LSU68] et [MPY94], à une procédure de régularisation près des coefficients. L'EDSR fortement couplée associée est alors à comprendre au sens faible, cf. Delarue et Guatteri [DG06].

Les ordres précédents se justifient comme suit :

- $\mathcal{E}(\text{temps})$. L'ordre $\alpha/2$ provient de la composition de la régularité α-Hölder des coefficients en x avec la régularité en temps de l'accroissement Brownien.
- $\mathcal{E}(\text{espace})$. Pour une fonction régulière l'erreur associée l'interpolation linéaire sur une grille spatiale Cartésienne de pas δ est d'ordre δ^2. L'itération le long de la grille en temps conduit à l'ordre indiqué.
- $\mathcal{E}(\text{quantification})$. L'espérance d'une fonction "régulière" de la quantification Gaussienne optimale renormalisée par $h^{1/2}$ est d'après (4.13) d'ordre $hM^{-2/d}$. La fonction régulière n'est autre dans l'analyse de convergence que la solution u de (4.4). Ainsi l'itération de la procédure devrait conduire à un contrôle de $M^{-2/d}$. Nous obtenons une vitesse de $h^{\alpha/2}M^{-2/d}$ en utilisant la α-Hölder continuité des dérivées secondes en espace.

Comme nous l'avons indiqué en Section 4.1.2, la démarche pour l'analyse de l'erreur est similaire à celle utilisée dans [MPY94] pour établir l'unicité de la solution à l'EDSR fortement couplée. Cela nous permet d'étendre les résultats de Milstein et Tretyakov qui supposent que la solution u de (4.4) est quatre fois dérivable et utilisent un produit tensoriel de lois de Bernoulli symétriques pour approcher l'incrément gaussien. Par ailleurs si l'on prend formellement $\alpha = 2, M = 2^d, \delta = h$ dans le théorème précédent on retrouve bien l'ordre h obtenu dans [MT04, MT06b, MT07].

Nous conclurons cette section en insistant sur le faible impact de la quantification sur l'erreur globale, ce dont atteste la borne sur $\mathcal{E}(\text{quantification})$.

4.3.4 Résultats numériques

Plusieurs exemples ont été développés dans **[2]** et **[5]** : équations de Burgers et KPZ déterministe (voir ci-après) et des milieux poreux. Par souci de concision nous ne présentons que deux exemples issus de **[5]** dans le cadre multidimensionnel.

Equation de Burgers multi-dimensionnelle. Les équations de Burgers sont une forme simplifiée des équations de Navier-Stokes. Les parties convectives et dissipatives sont les mêmes mais elles ne font intervenir ni le gradient de pression, ni la contrainte d'incompressibilité. Les équations s'écrivent :

$$\begin{aligned} &\partial_t u - (u.\nabla)u + \frac{\varepsilon^2}{2}\Delta u = 0, \ (t,x) \in [0,T) \times \mathbb{R}^d, \ \varepsilon > 0, \\ &u(T,x) = g(x), \ x \in \mathbb{R}^d, \end{aligned} \tag{4.16}$$

où $\forall i \in [\![1,d]\!]$, $((u.\nabla)u)_i = \nabla u_i \times u$. Bien que le Théorème 4.3.2 soit établi pour des fonctions à valeurs réelles, l'analyse précédente s'étendrait sans difficultés à des systèmes d'équations.

En dimension un il est bien connu que l'équation de Burgers "visqueuse" possède une solution explicite donnée par une factorisation de type Cole-Hopf, voir e.g. Whitham [Whi73]. Cette approche s'étend au cas

1. i.e. bornée et à dérivées spatiales d'ordre un et deux bornées.

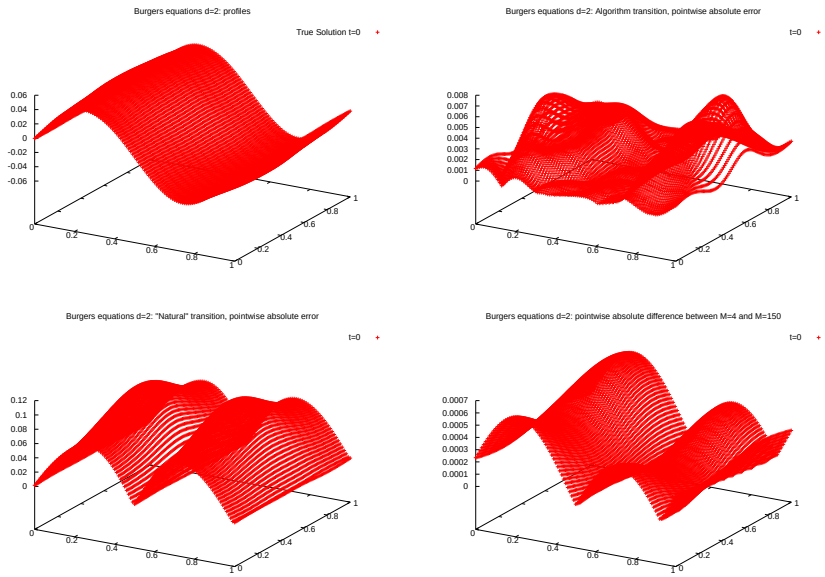
multi-dimensionnel lorsque la condition terminale dérive d'un potentiel, i.e. $g = \nabla g_0$ où g_0 est une fonction à valeurs réelles. On a dans ce cas : $\forall (t,x) \in [0,T] \times \mathbb{R}^d$,

$$u(t,x) = \frac{\mathbb{E}[\nabla g_0(x + \varepsilon W_{T-t}) \exp(-\varepsilon^{-2} g_0(x + \varepsilon W_{T-t}))]}{\mathbb{E}[\exp(-\varepsilon^{-2} g_0(x + \varepsilon W_{T-t}))]}.$$

Nous avons pour les résultats ci-après considéré l'interprétation "couplée" de l'équation (4.16), i.e. la non linéarité apparaît en injectant l'approximation de la solution dans la dérive de l'équation progressive. Ce choix semble plus robuste numériquement et dans le cas particulier de (4.16) permet même d'éviter l'étape d'estimation du gradient.

Nous considérons le cas $d = 2$ pour la fonction $H_0(x) = \prod_{i=1,2} \sin^2(\pi x_i)$ qui est symétrique et périodique permettant ainsi d'éviter les problèmes de troncation. Une des difficultés est de construire une grille spatiale qui permette de "voir" les transitions de la diffusion approchée qui sont de l'ordre de $\varepsilon h^{1/2}$. Ceci impose que δ soit plus petit que $\varepsilon h^{1/2}$. Nous prenons le jeu suivant de paramètres : $T = 3/8, h = 2.5 \times 10^{-2}, \delta = .01, \varepsilon^2 = .4$. Comme la solution est symétrique nous ne présentons que les résultats associés à la première composante u_1 de $u = (u_1, u_2)^*$.

Nous présentons les profils de la solution, ainsi que la différence entre la valeur de référence, obtenue à partir de la représentation explicite pour 600 points de quantification, et l'approximation de l'algorithme (4.15). Un autre graphique représente ensuite les résultats obtenus sans facteur de normalisation γ_M^* pour la quantification de l'accroissement Brownien et illustre l'importance de cette procédure. Nous étudions enfin l'influence du nombre de points M utilisés pour la quantification en prenant successivement $M = 4$ et $M = 150$. Les résultats corroborent la conclusion de la section précédente, à savoir que le nombre de points de quantification n'a qu'un faible impact.



Equation de KPZ déterministe.

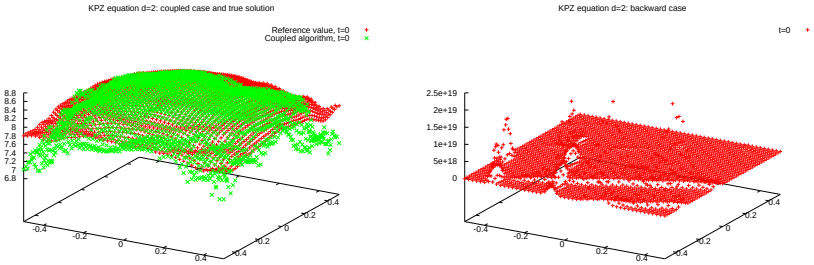
L'équation KPZ [2]peut se voir comme la primitive de l'équation de Burgers. Elle s'écrit :

$$\begin{aligned} &\partial_t u + \frac{1}{2}|\nabla u|^2 + \frac{1}{2}\Delta u = 0, \ (t,x) \in [0,T) \times \mathbb{R}^d, \\ &u(T,x) = g(x), \ x \in \mathbb{R}^d. \end{aligned} \tag{4.17}$$

Cette équation possède également une représentation explicite obtenue par factorisation Cole-Hopf : $u(t,x) = \log(\mathbb{E}[\exp(g(x + W_{T-t}))])$. C'est une équation quadratique en gradient pour laquelle les approches proposées ci-avant ne sont pas tout à fait satisfaisantes quant à l'analyse de l'erreur. En revanche, dans l'exemple suivant, le comportement numérique des algorithmes "couplés", où l'on voit le gradient comme la dérive dans la composante progressive, et découplés, où le carré du gradient apparaît au générateur, est très différent.

Soit $d = 2$, $g(x) = 10\cos(5|x|^2), T = .1, h = .02, \delta = .02, M = 4$. On tronque la grille pour $|x| \geq 1$. La valeur de référence est obtenue par 200 points de quantification. Il vient :

2. Cette équation apparaît en mécanique statistique où elle est de plus forcée par un bruit blanc, cf. [KPZ86]. Elle intervient dans des modèles de croissance ou d'interface.



Dans la pratique l'intégration numérique de grands gradients provoque des *overflow*, et incite dans ce cas de figure à préférer l'interprétation couplée plus "robuste". Indiquons qu'un exemple tridimensionnel pour (4.17) est présenté dans [**5**]. Pour la grille Cartésienne de pas δ, le nombre de noeuds pour lesquels il faut mettre à jour la solution et son gradient est de l'ordre de δ^{-d} donnant une complexité globale de $h^{-1}M\delta^{-d}$ ce qui interdit dans la pratique les grandes dimensions.

4.3.5 Extensions et perspectives

1. L'analyse précédente se base sur les propriétés de régularité des solutions d'EDP quasi-linéaires à coefficients réguliers et terme de diffusion elliptique. Ce contexte permet de mener à bien une analyse de stabilité "usuelle". Un enjeu serait de proposer une extension de la méthode précédente pour l'approximation de solutions faibles ou variationnelles de l'équation (4.4). Un point de départ pourrait être le travail de Barles et Lesigne [BL97] qui établit dans le cadre de problèmes semi-linéaires sous forme divergence une équivalence entre formulation variationnelle usuelle et formulation variationnelle associant l'EDSR et des fonctions tests aléatoires, cf. également Bally et Matoussi pour un cadre infini-dimensionnel [BM01].
2. Le problème stationnaire pour les EDP quasi-linéaires non dégénérées est bien compris, cf. Ladyzenskaya et Uralts'eva [LU68] ou Gilbarg et Trudinger [GT77]. En revanche l'approximation numérique reste ouverte. En effet comment procéder autrement que par la résolution approchée d'une succession de problèmes "découplés" avec temps d'atteinte ? Dans ce cas là, que dire du nombre d'itérations nécessaire pour obtenir une convergence ? Un autre problème lié à ce contexte consiste d'un point de vue théorique à s'interroger sur l'extension des résultats d'existence et d'unicité à l'EDSR fortement couplée avec temps d'arrêt sous les hypothèses naturelles de Lipschitz continuité et non dégénérescence des coefficients. Là aussi l'approche itérative par succession de problèmes découplés semble la plus naturelle.
3. Nous avons abordé dans la partie numérique le cas de l'équation de Burgers. Belopolskaya et Milstein [BM03] ont proposé pour les équations de Navier-Stokes dans le cas périodique tridimensionnel en temps court un algorithme d'approximation basé sur une itération rétrograde. L'idée est d'exploiter la périodicité pour faire un développement en série de Fourier de la solution approchée et d'utiliser la décomposition d'Helmoltz pour conserver pour le schéma la contrainte d'incompressibilité. Le gradient de pression est mis à jour en estimant l'équation $\Delta p = \sum_{(i,j)\in[\![1,d]\!]^2} \partial_{x_i} u^j \partial_{x_j} u^i$ qui relie le Laplacien de la pression aux gradients de la solution supposés réguliers, i.e. au moins continus. La difficulté de cette approche est de passer des contrôles L^2 naturels à ce contexte aux contrôles ponctuels nécessaires pour avoir la convergence.

   Dans l'optique de mieux appréhender ce problème, il serait intéressant d'envisager une représentation de la solution en terme d'EDSR. Une idée serait de procéder par pénalisation pour avoir l'incompressibilité. La gestion du gradient de pression est plus délicate dans la mesure où ce terme est formellement lié à des dérivées secondes que l'on retrouveraient en terme source. Notons que les formulations vortex semblent mal adaptées aux interprétations rétrogrades. La structure de convolution rend naturelles l'approche de Mc Kean et Vlasov et les méthodes numériques particulaires associées, cf. Méléard [Mél01].

   Nous discutons avec Bruno Bouchard de Dauphine à propos de problèmes d'existence et d'unicité pour des EDSR avec contraintes sur le Z.
4. Le cas quadratique est mal compris du point de vue numérique alors qu'il est assez naturel et bien compris tant du point de vue des EDSR [Kob00] que des EDP [LSU68]. Pour l'approximation ce sont bien sûr les contrôles sur le gradient qui sont à raffiner.

4.4 Semi-discrétisation d'EDSR dans un domaine

Nous présentons dans cette Section le travail [**6**], en commun avec B. Bouchard de l'Université Paris Dauphine, relatif à l'approximation forte d'EDSR dans un domaine. Il s'agit là de l'analyse d'un schéma de semi-

discrétisation dont la spécificité est de prendre en compte des domaines non-réguliers et des coefficients de diffusion dégénérés dans un contexte parabolique, i.e. un instant d'observation final $T > 0$ est fixé.

4.4.1 Cadre de travail

Nous considérons l'EDS faiblement couplée

$$\begin{aligned} X_t &= x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \\ Y_t &= g(\tau, X_\tau) + \int_t^T \mathbb{I}_{s<\tau} f(X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \end{aligned} \tag{4.18}$$

où les coefficients b, σ, f, g de (4.18) sont uniformément Lipschitziens en espace et τ est le temps de sortie de (t, X_t) d'un domaine cylindrique $\mathcal{D} = [0, T) \times D$ où $D \subset \mathbb{R}^d$ est un domaine régulier par morceaux.

Sous de bonnes hypothèses on peut montrer que $u(t, x) := Y_t^{t,x}$ est l'unique solution de viscosité de

$$\begin{cases} (\partial_t + L)u(t, x) + f(t, x, u(t, x), (\sigma\nabla u)(t, x)) = 0, \ (t, x) \in [0, T) \times D, \\ u(t, x) = g(t, x), \ (t, x) \in [0, T) \times \partial D \cup \{T\} \times \bar{D}, \end{cases}$$

où L désigne le générateur infinitésimal de X, cf. Darling et Pardoux [DP97] pour le cas elliptique qui peut s'adapter à (4.18).

Considérons maintenant le schéma de semi-discrétisation suivant le long de la grille en temps $(t_i := ih)_{i\in[\![1,N]\!]}$ où $Nh = T$:

$$\begin{aligned} Y_T^h &= g(\tau^h, X_{\tau^h}^h), \\ Y_{t_i}^h &= \mathbb{E}[Y_{t_{i+1}}^h|\mathcal{F}_{t_i}] + \mathbb{I}_{t_i<\tau^h} h f(X_{t_i}^h, Y_{t_i}^h, Z_{t_i}^h), \ Z_{t_i}^h = h^{-1}\mathbb{E}[Y_{t_{i+1}}^h(W_{t_{i+1}} - W_{t_i})|\mathcal{F}_{t_i}], \ i \in [\![0, N-1]\!], \end{aligned} \tag{4.19}$$

où X^h est le schéma d'Euler de l'équation progressive de (4.18), $\tau^h := \inf\{t_i := ih : (t_i, X_{t_i}^h) \notin \mathcal{D}\}$ et g désigne une extension de la condition de bord à tout l'espace. Remarquons que $Y_{t_i}^h \mathbb{I}_{t_i \geq \tau^h} = g(\tau^h, X_{\tau^h}^h)$ et par conséquent $Z_{t_i}^h \mathbb{I}_{t_i \geq \tau^h} = 0$.

Le propos est de fournir en fonction du pas de discrétisation un contrôle sur l'erreur :

$$\mathrm{Err}(h, T)^2 := \max_{i\leq N} \mathbb{E}[\sup_{t\in[t_i,t_{i+1}]} |Y_t - Y_{t_i}^h|^2] + \mathbb{E}[\int_0^T |Z_t - Z_{\phi(t)}^h|^2 dt],$$

où $\phi(t) := \sup\{t_i := ih : t_i \leq t\}$.

4.4.2 Résultats

Supposons que la condition terminale g vérifie :

(Hg) $g \in C^{1,2}([0, T] \times \mathbb{R}^d)$ et il existe $C_0 > 0$ telle que $|\partial_t g|_\infty + |\nabla g|_\infty + |D^2 g|_\infty \leq C_0$.

Nous avons alors le résultat suivant :

Proposition 4.4.1 *Il existe C dépendant uniquement des coefficients de* (4.18)*, de la borne dans* **(Hg)** *et de T telle que*

$$\mathrm{Err}(h, T)^2 \leq C \left(h + \mathcal{R}(Y)^h + \mathcal{R}(Z)^h + \mathbb{E}\left[\xi|\tau - \tau^h| + \mathbb{I}_{\tau^h<\tau} \int_{\tau^h}^{\tau} |Z_s|^2\right]\right),$$

où ξ est une variable aléatoire positive telle que $\mathbb{E}[\xi^p] \leq C^p$, $p \geq 2$, et $\mathcal{R}(Y)^h := \max_{i\leq N} \mathbb{E}[\sup_{t\in[t_i,t_{i+1}]} |Y_t - Y_{t_i}|^2]$, $\mathcal{R}(Z)^h := \mathbb{E}[\int_0^T |Z_t - \hat{Z}_{\phi(t)}|^2 dt]$, $\hat{Z}_{t_i} := h^{-1}\mathbb{E}[\int_{t_i}^{t_{i+1}} Z_s ds|\mathcal{F}_{t_i}]$, $i \in [\![0, N-1]\!]$.

Les termes $\mathcal{R}(Y)^h$ et $\mathcal{R}(Z)^h$ sont usuels dans l'analyse des schémas de semi-discrétisation d'EDSR et apparaissent déjà dans [Zha04]. A l'aune du lien entre EDSR et EDP, il apparaît néanmoins clairement qu'obtenir un contrôle à l'ordre h pour $\mathcal{R}(Y)^h, \mathcal{R}(Z)^h$ nécessite d'avoir des propriétés de Lipschitz continuité en espace de la solution de viscosité de l'EDP (4.19). Pour obtenir ces estimées et contrôler l'écart entre les temps d'arrêts il est nécessaire de rajouter quelques hypothèses de "structure" sur le domaine, cf. Chapitre 2.

Introduisons :

(D) Le domaine $D := \cap_{l=1}^m D_l,\ m \in \mathbb{N}^*$ et les $(D_l)_{l\in[\![1,m]\!]}$ sont de classe C^2 à frontière compacte. Nous supposerons également que D vérifie une condition de sphère extérieure uniforme et de cône intérieur uniforme (cf. Section 2.2 de [**6**] pour une définition précise de ces deux dernières conditions).

Désignons par $\mathcal{C} := \bigcap_{l\neq k=1}^m \partial D_l \cap \partial D_k$ l'ensemble des coins et par $B(\mathcal{C},\eta)$ un voisinage de rayon $\eta > 0$ de cet ensemble. Nous supposerons également :

(C) Le coefficient de diffusion a vérifie une condition de frontière non caractéristique uniforme hors d'un voisinage $B(\mathcal{C},\eta), \eta > 0$ des coins, i.e $\exists C > 0, \forall x \in \partial D \backslash B(\mathcal{C},\eta),\ \langle a(x)n(x), n(x)\rangle \geq C$ où n désigne la normale rentrante au domaine qui est bien définie hors d'un voisinage des coins, cf. Chapitre 2. Le coefficient a est uniformément elliptique dans un voisinage des coins, i.e. $\exists C > 0, \forall x \in \bar{D} \cap B(\mathcal{C},\eta),\ \forall \xi \in S^{d-1},\ \langle a(x), \xi, \xi\rangle \geq C$.

Ces conditions permettent d'utiliser des arguments d'EDP de type fonctions barrières pour contrôler le module de continuité au bord du domaine. A l'intérieur du domaine, où le coefficient de diffusion peut dégénérer, ce même contrôle est obtenu par des techniques probabilistes plus standards de flots stochastiques. Nous obtenons ainsi le résultat suivant :

Théorème 4.4.1 *Si les coefficients b, σ, f, g sont uniformément Lipschitziens de constante K et que les hypothèses* **(Hg)**, **(D)**, **(C)** *sont vérifiées*

$$\exists C := C(K, T, (\mathbf{Hg}), (\mathbf{D}), (\mathbf{C})),\ \mathcal{R}(Y)^h + \mathcal{R}(Z)^h \leq Ch,$$

et l'unique solution de viscosité u de (4.19) *dans la classe des solutions continues à croissance polynomiale est uniformément* 1/2 *Hölder continue en temps et Lipschitz en espace, i.e. pour le même C :*

$$|u(t,x) - u(t',x')| \leq C(|t-t'|^{1/2} + |x-x'|),\ \forall (t,x),(t',x') \in \mathcal{D}.$$

Nous avons par ailleurs le contrôle suivant pour le temps d'arrêts.

Théorème 4.4.2 *Sous les hypothèses du Théorème 4.4.1, pour tout $\varepsilon \in (0, \frac{1}{2})$, $\exists C := C(\varepsilon, K, T, (\mathbf{Hg}), (\mathbf{D}), (\mathbf{C}))$,*

$$\mathbb{E}[|\tau - \tau^h|] \leq Ch^{1/2-\varepsilon}.$$

Cette estimée s'obtient par des techniques de martingales et des arguments de changement de temps assez similaires à ceux introduits au Chapitre 2. Par ailleurs, à ε près ce contrôle est optimal, cf. Théorème 2.2.2.

C'est ainsi l'approximation du temps d'arrêt qui domine dans l'erreur et l'on déduit de la Proposition 4.4.1 et des Théorèmes 4.4.1, 4.4.2 la borne finale :

$$\forall \varepsilon \in (0, \frac{1}{2}),\ \exists C := C(\varepsilon, K, T, (\mathbf{Hg}), (\mathbf{D}), (\mathbf{C})),\ \mathrm{Err}(h,T)^2 \leq Ch^{1/2-\varepsilon}.$$

4.4.3 Extensions et perspectives

Nous avons retenu pour la clarté de la présentation dans [**6**] un domaine cylindrique. Le résultat précédent s'étendrait sans difficultés supplémentaires à domaine temps-espace vérifiant des hypothèses similaires à **(C)**,**(D)**. De même, pour un coefficient de diffusion non dégénéré il semble naturel de retrouver les même estimées pour le problème elliptique, dans la mesure où les moments des temps de sorties sont bien contrôlés pour le schéma d'Euler, voir Gobet et Maire [GM05] ou [**7**]. En revanche les situations elliptiques dégénérées restent à analyser, en particulier le transfert des propriétés des temps d'atteinte associés à l'objet continu au schéma de discrétisation.

Chapitre 5

Estimation de densité : cas dégénérés et applications

Nous présentons ici les résultats des articles [**10**], écrit en collaboration avec François Delarue de l'Université de Nice, et [**11**], en collaboration avec Vincent Lemaire de l'Université Paris 6.

Nous commençons par une synthèse des principales techniques intervenant dans l'estimation quantitative de densités de processus stochastiques. On soulignera en particulier les hypothèses de nature différente en fonction des outils utilisés : hypoellipticité et régularité des coefficients pour le calcul de Malliavin, et du côté analytique, uniforme ellipticité et mesurabilité pour des opérateurs sous forme divergence et uniforme ellipticité et Hölder continuité pour des opérateurs sous forme non divergence, voir aussi le Chapitre 3. Dans ce dernier cas, la notion de problème de martingale de Stroock et Varadhan permet d'unifier l'approche probabiliste et analytique.

Nous détaillons ensuite l'article [**10**] où nous considérons une EDS vérifiant une hypothèse d'Hörmander faible, i.e. c'est le coefficient de dérive qui propage le bruit dans le système. Sous de bonnes hypothèses, proches de celles naturelles du point de vue analytique, nous retrouvons des estimées Gaussiennes de type Aronson en régime non diffusif. Pour obtenir la minoration, une première approche consiste à utiliser la transformée de Fleming qui permet d'exprimer le logarithme de la densité de la diffusion comme solution d'un problème de contrôle stochastique. Nous donnerons également une seconde preuve plus géométrique. La majoration découle quant à elle de l'approche *parametrix*.

Les techniques développées dans [**10**] peuvent s'adapter à des schémas de discrétisation de certains processus dégénérés ou non. Dans [**11**] nous avons obtenu des estimées Gaussiennes pour les densités de certains schémas d'Euler. L'utilisation de techniques de concentration de la mesure : inégalités de Sobolev logarithmique, argument de Herbst, nous a permis d'en déduire des intervalles de confiance non asymptotiques pour l'approximation de Monte Carlo.

5.1 Estimation de densité : quelques techniques et résultats

5.1.1 Opérateurs non dégénérés sous forme divergence

Considérons un opérateur sous forme divergence

$$L\varphi(x) = \frac{1}{2}\mathrm{div}(a(x)\nabla\varphi(x)) + \langle b(x), \nabla\varphi(x)\rangle,\ \varphi \in C_0^\infty(\mathbb{R}^d),$$

où les coefficients a, b sont mesurables, b est borné, a est uniformément elliptique borné, i.e $\exists\Lambda \geq 1$, $\forall\xi \in \mathbb{R}^d$, $\Lambda^{-1}|\xi|^2 \leq \langle a(x)\xi, \xi\rangle \leq \Lambda|\xi|^2$. Dans ce cadre, des techniques analytiques désormais classiques donnent que pour tout $T > 0$, il existe des constantes $C := C(T, \Lambda, |b|_\infty), c := c(\Lambda)$, $(c, C) \in (0, 1] \times [1, +\infty)$ telles que si p désigne la solution fondamentale associée à L,

$$\frac{C^{-1}}{t^{d/2}} \exp\left(-c^{-1}\frac{|x-y|^2}{t}\right) \leq p(t, x, y) \leq \frac{C}{t^{d/2}} \exp\left(-c\frac{|x-y|^2}{t}\right), \forall(t, x, y) \in (0, T] \times \mathbb{R}^d \times \mathbb{R}^d. \tag{5.1}$$

Si par ailleurs $b = 0$ les estimées sont uniformes en temps, i.e. les constantes $(c, C) := (c, C)(\Lambda)$ et (5.1) est valable pour tout $t > 0$. Nous appellerons pour la suite de ce chapitre inégalités de Aronson [1] des inégalités de la forme (5.1), avec éventuellement une échelle diagonale non diffusive.

1. qui dans son travail fondateur [Aro67] a établi les inégalités (5.1)

La borne supérieure s'obtient typiquement à l'aide des inégalités de Nash et d'opérations sur la forme de Dirichlet associée à L. La borne inférieure s'obtient via la représentation spectrale du noyau de la chaleur tué de L et un argument de *chaînage*, voir Chapitre 7 de Bass [Bas97] ou Stroock [Str88].

Dans cette approche seule la non dégénérescence de a compte. Pour $b = 0$, l'aspect auto-adjoint est également très important dans les preuves et donne l'uniformité en temps.

5.1.2 Opérateurs sous forme non divergence

Soit maintenant

$$L_t\varphi(x) = \frac{1}{2}\mathrm{tr}(a(t,x)D^2\varphi(x)) + \langle b(t,x), \nabla\varphi(x)\rangle,\ t \geq 0,\ \varphi \in C_0^\infty(\mathbb{R}^d).$$

Nous avons vu au Chapitre 3.1.2 que pour a uniformément elliptique borné, η-Hölder continu en espace, $\eta > 0$ et b mesurable borné, on obtenait la borne supérieure de (5.1) à l'aide de la représentation en série parametrix, voir aussi Friedman [Fri64]. Cette représentation fournit également la borne inférieure de (5.1) pour $|x-y|/t^{1/2} \leq C_0, t \leq T_0$, ce qui avait été initialement remarqué par Il'in *et al.* [IKO62]. L'extension à $|x-y|/t^{1/2} > C_0, t \leq T_0$ se fait par un argument de chaînage tout à fait similaire à celui effectué pour les opérateurs sous forme divergence. Nous reviendrons sur la structure du chaînage que l'on peut en fait relier à un problème de contrôle déterministe en Section 5.2. L'extension à un temps arbitraire s'obtient par convolution ou par des arguments de changement d'échelle (voir Section 2.3 de [**10**]).

Par cette approche la constante C de (5.1) dépend du temps même pour $b = 0$. Nous obtenons toutefois les bornes Gaussiennes de façon assez directe pour la densité de l'unique solution du problème de martingale associé à (a, b). Cette analyse s'étend sous les mêmes hypothèses au schéma d'Euler, voir Section 5.3. Même si l'on ne peut plus ignorer toute régularité spatiale sur a, la condition est assez peu contraignante.

5.1.3 Opérateurs hypoelliptiques et calcul de Malliavin

Considérons l'opérateur $L = \frac{1}{2}\sum_{i=1}^r A_i^2 + A_0$ sous forme de Hörmander, où les $(A_i)_{i\in[\![1,r]\!]}$ sont des champs de vecteurs réguliers sur $\mathbb{R}^d$, et qui est le générateur de la diffusion $X_t = x + \int_0^t A_0(X_s)ds + \int_0^t A_i(X_s) \circ dW_s^i$ où $\circ dW$ désigne l'intégrale de Stratonovitch. Si la dimension de l'espace vectoriel engendré par

$$A_1, \cdots, A_r,\ [A_i, A_j]_{0\leq i,j\leq r},\ [A_i, [A_j, A_k]]_{0\leq i,j,k\leq r}, \cdots$$

au point initial x est $\mathbb{R}^d$ (où $[,]$ désigne le crochet de Lie entre champs de vecteurs) [2], alors pour tout $t > 0$, X_t admet une densité p_t, cf. Norris [Nor86] ou Nualart [Nua98]. La question de l'estimation de cette densité est dans le cadre hypoelliptique général très difficile. Il est en effet très délicat de mettre en évidence une échelle globale intrinsèque au système.

Plusieurs résultats ont néanmoins été obtenus l'aide du calcul de Malliavin. Cet outil (voir Malliavin [Mal97], Stroock [Str83] et Nualart [Nua95]) permet de quantifier la sensibilité du système par rapport au bruit et cette sensibilité se lit sur sur les dérivées de Malliavin et la matrice de covariance associée. La non-dégénérescence de cette matrice donne l'existence de la densité. L'utilisation de cette technique pour l'estimation de densité est due à Kusuoka et Stroock [KS84, KS85, KS87]. Dans le dernier travail, les auteurs établissent des estimées de type Aronson sous des hypothèses d'Hörmander fortes, i.e. le bruit ne se propage dans le système qu'à travers la composante diffusive. La normalisation diagonale provient alors du volume des boules hypoelliptiques de rayon $t^{1/2}$ centrées au point initial. C'est en effet une normalisation diffusive mais les boules ne sont plus euclidiennes et le caractère "local" de leur volume apparaît clairement. La décroissance hors diagonale est donnée par la métrique de Carnot-Carathéodory, ou métrique de contrôle, associée aux champs $(A_i)_{i\in[\![1,r]\!]}$. Ces résultats ne sont valables que pour une dérive nulle ou engendrée par les champs de vecteurs de la partie diffusive.

Pour d'autres exemples d'application du calcul de Malliavin nous renvoyons à Bally [Bal90], pour une discussion sur les liens entre matrice de Malliavin et hypothèses de Hörmander, à Cattiaux [Cat90], pour les propriétés de résolvantes de diffusions hypoelliptiques, à Cattiaux et Mesnager [CM02] pour l'existence de la densité dans le cadre non-homogène. Pour des résultats plus spécifiques sur l'estimation de densité nous renvoyons à Ben Arous et Léandre [BL91], qui ont mis en évidence plusieurs régimes pour le comportement en temps petit de la densité de certaines diffusions hypoelliptiques en fonction de la forme de leur dérive.

En résumé, dans un cadre dégénéré l'utilisation du calcul de Malliavin pour obtenir des estimées de densités demande de la régularité sur les coefficients. Cette régularité n'est pas toujours nécessaire. Dans la section suivante nous arrivons à obtenir pour une diffusion faiblement hypoelliptique des estimées de densités sous

2. Dans l'article initial [Hör67], Hörmander supposait cette condition vérifiée en tout point de l'espace.

des hypothèses proches de celles naturelles pour les opérateurs non dégénérés sous forme non divergence. En particulier peu de régularité des coefficients est nécessaire.

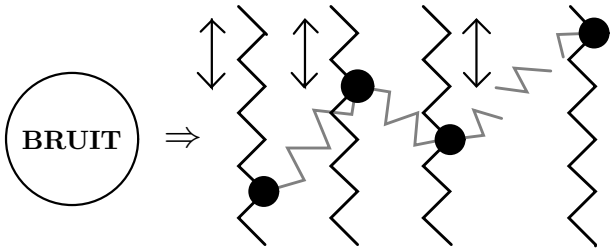
5.2 Estimations de densités pour un bruit propagé dans une chaîne d'Equations différentielles

Le modèle que nous avons considéré pour l'article **[10]** est le suivant :

$$\begin{aligned} dX_t^1 &= F_1(t, X_t^1, \dots, X_t^n)dt + \sigma(t, X_t^1, \dots, X_t^n)dW_t, \\ dX_t^2 &= F_2(t, X_t^1, \dots, X_t^n)dt, \\ dX_t^3 &= F_3(t, X_t^2, \dots, X_t^n)dt, \qquad t \geq 0, \\ &\dots \\ dX_t^n &= F_n(t, X_t^{n-1}, X_t^n)dt, \end{aligned} \tag{5.2}$$

où $(W_t)_{t\geq 0}$ désigne un mouvement Brownien d-dimensionnel et pour tout $i \in [\![1, n]\!]$, $(X_t^i)_{t\geq 0}$, est aussi à valeurs dans $\mathbb{R}^d$.

Un exemple typique pour (5.2) est donné par un système de n oscillateurs couplés, chacun d'entre eux bougeant verticalement et étant directement relié à son plus proche voisin, lorsque le premier oscillateur est excité par un bruit aléatoire.



Précisons que les F_i dans (5.2) ne peuvent pas dépendre des positions $X_t^1, \dots, X_t^{i-2}$: d'un point de vue physique le bruit doit traverser les $(i-1)$ premiers oscillateurs avant d'atteindre le $i^{\text{ème}}$. Dans la dynamique différentielle n'intervient que la composante qui transmet le bruit. Intuitivement cette condition permet la *séparation des échelles*. Supposons que le bruit soit en effet transmis par la composante X^{i-1} à la composante X^i, ce qui sera assuré par la condition de non dégénérescence des $D_{x_{i-1}}F_i$ (hypothèse de type Hörmander faible). Alors, l'échelle caractéristique de X^i est $t^{i-1/2}$ pour tout $i \in [\![1, n]\!]$.

Des systèmes de ce type apparaissent dans différents champs d'applications. Pour $n = 2$, (5.2) décrit la dynamique de certains systèmes hamiltoniens, voir Soize [Soi94] pour une synthèse ou Talay [Tal02] et Hérau et Nier [HN04] pour l'aspect convergence à l'équilibre. Toujours pour $n = 2$, (5.2) correspond à la dynamique utilisée en finance pour les options asiatiques, cf. [BPV01] pour des questions de régularité du prix associé. Pour $n \geq 2$, cette équation apparaît dans des modèles de conduction de la chaleur, cf. Eckmann et al. [EPRB99] et Rey-Bellet et Thomas [RBT00] où la chaîne d'oscillateurs est chauffée aux deux extrémités par des bains de chaleur. Mentionnons également le travail récent de Bodineau et Lefevere [BL08]).

Pour le reste de la section notre jeu d'hypothèses (noté **(A)**) est le suivant :

(a) Le spectre de la fonction matricielle $(t, \mathbf{x}) := (t, x_1, \cdots, x_n) \in \mathbb{R}_+ \times \mathbb{R}^{nd} \mapsto a(t, \mathbf{x}) = [\sigma\sigma^*](t, \mathbf{x})$ est inclus dans $[\Lambda^{-1}, \Lambda]$pour $\Lambda \geq 1$.

(b) Les fonctions $F_1, \dots, F_n$ et σ sont respectivement uniformément Lipschitz et η-Hölder continues ($\eta \in (0, 1]$) par rapport aux variables d'état, pour une constante positive κ.

(c) Pour tout $i \in [\![2, n]\!]$, $(t, (x_i, \dots, x_n)) \in \mathbb{R}_+ \times \mathbb{R}^{(n-i+1)d}$, la fonction $x_{i-1} \in \mathbb{R}^d \mapsto F_i(t, x_{i-1}, \dots, x_n)$ est continûment différentiable, la dérivée notée $(t, x_{i-1}, \dots, x_n) \in \mathbb{R}_+ \times \mathbb{R}^{(n-i+2)d} \mapsto D_{x_{i-1}}F_i(t, x_{i-1}, \dots, x_n)$, est supposée η-Hölder continue de constante κ en la première variable spatiale x_{i-1}.

(d) Il existe un ensemble convexe fermé $\mathcal{E}_{i-1} \subset GL_d(\mathbb{R})$ (ensemble des matrices inversible $d \times d$) tel que pour tout $t \geq 0$ et $(x_{i-1}, \dots, x_n) \in \mathbb{R}^{(n-i+2)d}$, la matrice $D_{x_{i-1}}F_i(t, x_{i-1}, \dots, x_n)$ appartienne à $\mathcal{E}_{i-1}$. Par exemple, $\mathcal{E}_i$, $1 \leq i \leq n-1$, peut-être une boule fermée incluse dans l'ouvert $GL_d(\mathbb{R})$.

Le point (a) est une condition classique d'uniforme ellipticité de la matrice de diffusion. La condition (b) est assez naturelle du point de vue des EDO/EDP. Elle garantit que l'EDO sous-jacente à (5.2), i.e. lorsque l'on

supprime le forçage stochastique, est bien posée. Par ailleurs on retrouve sur a l'hypothèse usuelle de régularité pour les opérateurs sous forme non divergence. La condition (c) est utilisée pour l'analyse de densité : on suppose un peu plus de régularité sur F_i par rapport à la composante x_{i-1} qui transmet le bruit.

Le point (d) traduit une condition de type Hörmander faible qui permet la propagation du bruit dans le système. L'hypothèse de convexité est technique. Nous renvoyons à la Section 3 de [**10**] pour plus de détails.

En particulier, on retrouve comme au Chapitre 3 que les coefficients, supposés mesurables, peuvent être irréguliers en temps ($F(t,\mathbf{0})$ borné).

Théorème 5.2.1 *Supposons* **(A)** *vérifiée est que l'EDS* (5.2) *est uniquement soluble en loi. Alors, la solution (faible) de* (5.2) *admet pour tout $t>0$ une densité $p(t,.,.)$ qui vérifie des estimées de type d'Aronson. Précisément, pour tout $T>0$, $\exists C_T \geq 1, C_T := C_T(\Lambda,\eta,\kappa,(\mathcal{E}_i)_{i\in[\![1,n-1]\!]},n,d)$, telle que pour tout $t\in(0,T]$, $(\mathbf{x},\mathbf{y})\in\mathbb{R}^{nd}\times\mathbb{R}^{nd}$:*

$$C_T^{-1}t^{-n^2d/2}\exp\big(-C_Tt\big|\mathbb{T}_t^{-1}\big(\boldsymbol{\theta}_t(\mathbf{x})-\mathbf{y}\big)\big|^2\big)\leq p(t,\mathbf{x},\mathbf{y})\leq C_Tt^{-n^2d/2}\exp\big(-C_T^{-1}t\big|\mathbb{T}_t^{-1}\big(\boldsymbol{\theta}_t(\mathbf{x})-\mathbf{y}\big)\big|^2\big).$$

Ci-dessus, $\mathbb{T}_t$ est la matrice d'"échelle" du système. C'est une matrice diagonale de taille $nd\times nd$ qui se décompose en n blocs diagonaux de taille $d\times d$ où le $i^{\text{ème}}$ bloc diagonal est donné par t^iI_d, où I_d est la matrice identité $d\times d$. De plus, $\boldsymbol{\theta}_t(\mathbf{x})=(\theta_t^1(\mathbf{x}),\theta_t^2(\mathbf{x}),\ldots,\theta_t^n(\mathbf{x}))$ désigne la valeur dans $\mathbb{R}^{nd}$ au temps t de la solution de l'EDO déterministe associée à (5.2) *:*

$$\begin{aligned}\dot{\theta_t}^1 &= F_1(t,\theta_t^1,\ldots,\theta_t^n),\\ \dot{\theta_t}^i &= F_i(t,\theta_t^{i-1},\ldots,\theta_t^n),\quad 2\leq i\leq n,\ \boldsymbol{\theta}_0(\mathbf{x})=(\theta_0^1(\mathbf{x}),\ldots,\theta_n^1(\mathbf{x}))=\mathbf{x}.\end{aligned}\tag{5.3}$$

Le transport de la condition initiale $\mathbf{x}$ par le flot déterministe $\boldsymbol{\theta}$ peut se voir comme une généralisation non linéaire de l'exemple de Kolmogorov abordé en Section 3.2.1. La matrice d'échelle traduit le fait que $X_t^i,\ i\in[\![1,n]\!]$ oscille autour de $\theta_t^i(\mathbf{x})$ avec des fluctuations d'ordre $t^{i-1/2}$. Pour $i=1$, l'action du flot (d'ordre t) est négligeable devant le bruit (d'ordre $t^{1/2}$) en temps petit. A l'opposé, pour $i\geq 2$, les fluctuations de X_t^i peuvent être beaucoup plus petites que la distance entre $\theta^i(\mathbf{x})$ et x_i : le terme de transport joue alors un rôle clé dans l'estimée de densité pour ces coordonnées.

L'effet multi-échelle dans l'estimation ci-dessus provient de la structure des coefficients $(F_i)_{i\in[\![1,n]\!]}$ et de la condition de Hörmander (d) qui nous donne que le bruit se propage suivant une seule direction dans le système.

Nous avons supposé ci-avant que F_1 était Lipschitz, en autorisant de fait les croissances linéaires. Sous réserve d'unicité en loi le résultat précédent resterait valable pour F_1 mesurable borné.

Pour la suite nous réécrirons la dynamique de (5.2) sous la forme condensée

$$d\mathbf{X}_t=\mathbf{F}(t,\mathbf{X}_t)+B\sigma(t,\mathbf{X}_t)dW_t,$$

où $\mathbf{F}=(F_1,\ldots,F_n)$ et B désigne la matrice d'injection de $\mathbb{R}^d$ dans $\mathbb{R}^{nd}$.

5.2.1 Transformée de Fleming

Pour prouver le Théorème précédent nous allons exploiter la formulation de $-\log p(t,\mathbf{x},\mathbf{y})$ en terme de problème de contrôle stochastique. Ce procédé, connu sous le nom de *transformée de Fleming*, a été utilisé par Sheu [She91] pour retrouver les estimées d'Aronson pour des opérateurs sous forme non-divergence dans le cadre du Paragraphe correspondant à la Section 5.1. La représentation Feynman-Kac associée à cette transformation amène à considérer un processus contrôlé où le paramètre de contrôle intervient comme dérive additionnelle sur la première composante, non dégénérée, de l'équation initiale. Un choix pertinent de contrôle permettra d'obtenir une minoration de la densité.

Nous supposerons par la suite, pour justifier l'existence de la densité que les coefficients $(F_i)_{i\in[\![1,n]\!]}$ et σ sont réguliers en espace et en temps de sorte que le théorème de Hörmander appliqué à l'opérateur parabolique ∂_t+L_t donne l'existence de la densité. Tous les résultats que nous obtenons sont uniformes par rapport aux paramètres de régularisation. Le théorème se déduit ensuite par passage à la limite, sous réserve d'unicité en loi, par des arguments de type Radon-Nikodym.

Fixons $(T,\mathbf{y}_0)\in(0,+\infty)\times\mathbb{R}^{nd}$ et posons pour $\varepsilon>0$ et $(t,\mathbf{x})\in[0,T-\varepsilon]\times\mathbb{R}^{nd}$, $u_\varepsilon(t,\mathbf{x})=\mathbb{E}[\eta_\varepsilon(\mathbf{X}_{T-\varepsilon}^{t,\mathbf{x}})]$ où $\mathbf{X}^{t,\mathbf{x}}$ désigne la solution de (5.2) pour $\mathbf{X}_t=\mathbf{x}$ et $(\eta_\varepsilon)_{\varepsilon>0}$ est une suite d'approximations régulières **strictement positives** de $\delta_{\mathbf{y}_0}$. La fonction u_ε est solution du problème de Cauchy

$$\begin{cases}\partial_tu_\varepsilon(t,\mathbf{x})+L_tu_\varepsilon(t,\mathbf{x})=0,\ (t,x)\in[0,T-\varepsilon)\times\mathbb{R}^{nd},\\ u_\varepsilon(T-\varepsilon,\mathbf{x})=\eta_\varepsilon(\mathbf{x}),\ \mathbf{x}\in\mathbb{R}^{nd}.\end{cases}$$

Par ailleurs $\lim_{\varepsilon\to 0} u_\varepsilon(0,\mathbf{x}) = \lim_{\varepsilon\to 0} \mathbb{E}[\eta_\varepsilon(\mathbf{X}^{0,\mathbf{x}}_{T-\varepsilon})] = p(T,\mathbf{x},\mathbf{y}_0)$.

Introduisons maintenant pour $(t,\mathbf{x}) \in [0,T-\varepsilon]\times\mathbb{R}^{nd}$, $J_\varepsilon(t,\mathbf{x}) = -\ln[u_\varepsilon(t,\mathbf{x})]$. Cette fonction vérifie l'EDP quadratique suivante

$$\begin{cases} \partial_t J_\varepsilon(t,\mathbf{x}) + L_t J_\varepsilon(t,\mathbf{x}) - \frac{1}{2}\langle a(t,\mathbf{x}) D_{x_1} J_\varepsilon(t,\mathbf{x}), D_{x_1} J_\varepsilon(t,\mathbf{x})\rangle = 0,\ (t,x)\in[0,T-\varepsilon)\times\mathbb{R}^{nd}, \\ J_\varepsilon(T-\varepsilon,\mathbf{x}) = -\ln(\eta_\varepsilon(\mathbf{x})),\ \mathbf{x}\in\mathbb{R}^{nd}. \end{cases} \tag{5.4}$$

Le point clé consiste à remarquer que la partie différentielle précédente peut se réécrire

$$\partial_t J_\varepsilon(t,\mathbf{x}) + L_t J_\varepsilon(t,\mathbf{x}) + \inf_{v\in\mathbb{R}^d}\big[\langle v, D_{x_1} J_\varepsilon(t,\mathbf{x})\rangle + \frac{1}{2}\langle a^{-1}(t,\mathbf{x})v, v\rangle\big] = 0, \tag{5.5}$$

où l'infimum est atteint en $v = v(t,\mathbf{x}) := -a(t,\mathbf{x}) D_{x_1} J_\varepsilon(t,\mathbf{x})$ et redonne (5.4). Notons $\mathcal{P}(T-\varepsilon)$ l'ensemble des processus progressivement mesurables $(v_t)_{t\in[0,T-\varepsilon]}$ à valeurs dans $\mathbb{R}^d$ tels que $\mathbb{E}[\int_0^{T-\varepsilon}|v_t|^2 dt] < +\infty$. A tout $(v_t)_{0\le t\le T-\varepsilon} \in \mathcal{P}(T-\varepsilon)$ on peut associer une version contrôlée de (5.2), i.e.

$$d\boldsymbol{\chi}_t = [\mathbf{F}(t,\boldsymbol{\chi}_t) + Bv_t]\,dt + B\sigma(t,\boldsymbol{\chi}_t)dW_t,\ \boldsymbol{\chi}_0 = \mathbf{x}. \tag{5.6}$$

En exploitant (5.5), nous pouvons écrire $J_\varepsilon(0,\mathbf{x})$ comme la fonction valeur d'un problème de contrôle stochastique :

$$J_\varepsilon(0,\mathbf{x}) = \inf_{(v_t)_t\in\mathcal{P}(T-\varepsilon)} \mathbb{E}\left[\frac{1}{2}\int_0^{T-\varepsilon}\langle a^{-1}(t,\boldsymbol{\chi}^{0,\mathbf{x}}_t)v_t, v_t\rangle dt - \ln[\eta_\varepsilon(\boldsymbol{\chi}^{0,\mathbf{x}}_{T-\varepsilon})]\right]. \tag{5.7}$$

L'équation (5.7) associée à l'identité $-\ln(p(T,\mathbf{x},\mathbf{y}_0)) = \lim_{\varepsilon\to 0} J_\varepsilon(0,\mathbf{x})$ constitue l'élément clé de notre preuve. Elle intervient pour les deux bornes du théorème.

5.2.2 Etapes pour la borne inférieure

L'idée principale est la suivante. Pour obtenir une borne inférieure sur $p(T,\mathbf{x}_0,\mathbf{y}_0)$, $\mathbf{x}_0\in\mathbb{R}^{nd}$, il est suffisant d'obtenir une borne supérieure pour $J_\varepsilon(0,\mathbf{x}_0)$, uniformément en $\varepsilon > 0$. D'après la formulation (5.7), tout le problème consiste à trouver un contrôle $(v_t)_{t\in[0,T-\varepsilon]}$ pertinent pour tout $\varepsilon > 0$.

Le choix de $(v_t)_{0\le t\le T-\varepsilon}$ s'effectue en trois étapes :

(i). Nous avons d'abord caractérisé la densité dans le cas linéaire, i.e. lorsque la dérive $\mathbf{F}$ est affine en espace, et la matrice σ ne dépend que du temps. La solution $(\mathbf{X}_t)_{t\ge 0}$ de (5.2) est alors un processus Gaussien. Sous $(\mathbf{A})$, la densité de transition existe et possède une forme explicite. De plus le terme hors diagonale de la densité peut s'identifier à l'énergie du contrôle optimal pour le problème de contrôllabilité linéaire déterministe, cf. Section 3 de **[10]**. En utilisant la convexité des ensembles du point (d) de **(A)**, on peut alors déduire que la densité est homogène aux bornes du théorème. Enfin, le contrôle optimal dans (5.6–5.7) s'écrit explicitement.

(ii). Pour se rapprocher le plus possible du cas précédent, l'idée naturelle est ensuite de linéariser l'équation contrôlée (5.6). Le point crucial est celui du choix de la courbe déterministe $(\boldsymbol{\phi}_t)_{t\in[0,T]}$ autour de qui opérer cette linéarisation. On écrit $\mathbf{F}(t,\boldsymbol{\chi}^{0,\mathbf{x}_0}_t) = \mathbf{F}(t,\boldsymbol{\phi}_t) + \mathbf{D_x F}(t,\boldsymbol{\phi}_t)(\boldsymbol{\chi}^{0,\mathbf{x}_0}_t - \boldsymbol{\phi}_t) + o(|\boldsymbol{\chi}^{0,\mathbf{x}_0}_t - \boldsymbol{\phi}_t|)$, où $\mathbf{D_x F}\in\mathcal{M}_{nd}(\mathbb{R})$ est la dérivée spatiale de $\mathbf{F}$. De même nous approchons le coefficient de diffusion $\sigma(t,\boldsymbol{\chi}^{0,\mathbf{x}_0}_t)$ par $\sigma(t,\boldsymbol{\phi}_t)$. Comme $(\boldsymbol{\phi}_t)_{0\le t\le T}$ est déterministe, l'application $\mathbf{z}\in\mathbb{R}^{nd}\mapsto \mathbf{F}(t,\boldsymbol{\phi}_t) + \mathbf{D_x F}(t,\boldsymbol{\phi}_t)(\mathbf{z}-\boldsymbol{\phi}_t)$ est bien affine, le coefficient de diffusion $(\sigma(t,\boldsymbol{\phi}_t))_{0\le t\le T}$ est déterministe, et l'on retrouve le cadre de (i). Un choix naturel consiste à prendre $(\boldsymbol{\phi}_t)_{0\le t\le T}$ comme la solution du problème de contrôle déterministe pour l'ODE naturellement associée à (5.2), i.e.

$$\dot{\boldsymbol{\phi}}_t = \mathbf{F}(t,\boldsymbol{\phi}_t) + B\varphi_t, t\in[0,T];\quad \boldsymbol{\phi}_0 = \mathbf{x}_0, \tag{5.8}$$

où $(\varphi_t)_{0\le t\le T}$ est un contrôle déterministe de $L^2([0,T],\mathbb{R}^d)$ solution du problème de minimisation :

$$I(T,\mathbf{x}_0,\mathbf{y}_0) := \inf\Big\{\int_0^T |\varphi_t|^2 dt : \boldsymbol{\phi}_0 = \mathbf{x}_0,\ \boldsymbol{\phi}_T = \mathbf{y}_0\Big\}. \tag{5.9}$$

Des arguments usuels de contrôle, cf. Coron [Cor07] nous ont permis d'établir que $I(T,\mathbf{x}_0,\mathbf{y}_0)$ est finie, i.e. l'équation (5.8) est contrôllable. Nous avons également établi que $I(T,\mathbf{x}_0,\mathbf{y}_0)$ est du même ordre que $T|\mathbb{T}_T^{-1}(\boldsymbol{\theta}_T(\mathbf{x}_0) - \mathbf{y}_0)|^2$. La comparaison avec l'énoncé du Théorème 5.2.1, nous donne donc que l'énergie du contrôle optimal déterministe va fournir la décroissance hors diagonale de la densité. De cette étape on décide de choisir le contrôle v_t dans (5.6) sous la forme $v_t := \varphi_t + v^0_t$.

(iii). Il reste donc à choisir $(v_t^0)_{t\in[0,T-\varepsilon]}$. Pour cela remarquons que d'après (5.6) et (5.8), la dynamique de $(\boldsymbol{\chi}_t^{0,\mathbf{x}} - \boldsymbol{\phi}_t)_{0\le t\le T-\varepsilon}$ s'écrit :

$$d\big[\boldsymbol{\chi}_t^{0,\mathbf{x}} - \boldsymbol{\phi}_t\big] = \big[\mathbf{D_x F}(t,\boldsymbol{\phi}_t)(\boldsymbol{\chi}_t^{0,\mathbf{x}} - \boldsymbol{\phi}_t) + B(v_t - \varphi_t)\big]dt + B\sigma(t,\boldsymbol{\phi}_t)dW_t + dR_t, \tag{5.10}$$

où $(R_t)_{0\le t\le T-\varepsilon}$ désigne un terme de reste qu'une analyse de stabilité technique, car non linéaire, permet de contrôler. Observons que $\boldsymbol{\chi}_0^{0,\mathbf{x}} - \boldsymbol{\phi}_0 = \mathbf{0}$ et que $\boldsymbol{\chi}_{T-\varepsilon}^{0,\mathbf{x}} - \boldsymbol{\phi}_{T-\varepsilon}$ est *a priori* proche de $\mathbf{y}_0 - \mathbf{y}_0 = \mathbf{0}$. Ces observations heuristiques font qu'un candidat naturel pour v^0 est le contrôle optimal Gaussien qui partant de $\mathbf{0}$ permet d'atteindre $\mathbf{0}$ en T. Par l'étape (i), nous savons également calculer son coût (au sens de la formulation (5.7)) : celui-ci est à une constante additive près de l'ordre de $-\ln(T^{-n^2d/2})$ et correspond dans le Théorème 5.2.1 au contrôle de la densité sur la "diagonale".

En résumé, de la formulation contrôle (5.7) on choisit un contrôle particulier $v_t = \varphi_t + v_t^0$, où $(\varphi_t)_{t\in[0,T-\varepsilon]}$ permet d'isoler les contributions hors diagonale, $(v_t^0)_{t\in[0,T]}$ donne la vitesse diagonale. C'est l'étape de contrôle des restes dans (5.10) qui est délicate. En effet, nous devons à ce propos utiliser des arguments de type Gronwall non linéaire qui font que l'on établira d'abord les bornes du théorème en temps court. Un argument de convolution ou de changement d'échelle permet ensuite de les étendre à temps arbitraire fixé.

La fonction $I(T,\mathbf{x},\mathbf{y})$ introduite dans (5.9) est la fonctionnelle d'action des grandes déviations. Elle établit un lien naturel entre le problème de contrôle déterministe (5.8–5.9) et la densité de transition de (5.2), cf. Freidlin and Wentzell [FW98] pour le cas non-dégénéré et Ben Arous et Léandre [BL91] pour des résultats sous condition de Hörmander forte. L'utilisation de la transformée de Fleming, introduite par Sheu [She91] et développée dans cette section permet une extension de ce lien lorsque le temps ne tend pas forcément vers 0. Comme indiqué ci-avant, nous conservons à cause de l'étude de stabilité une contrainte de temps court.

5.2.3 Parametrix et majoration

La formulation (5.7) n'est pas la plus adaptée pour obtenir une majoration. En effet, il faut minorer J_ε pour espérer majorer la densité. L'expression en terme d'infimum oblige alors à partir du contrôle optimal.

En revanche la structure du transport intervenant dans le terme exponentiel donne une bonne idée pour construire une famille de processus Gaussiens de densités homogènes à celle du Théorème 5.2.1 et qui serviront de *parametrix*.

Notons par $\tilde{p}^{T,\mathbf{y}_0}$ la densité Gaussienne qui en temps petit et $\mathbf{y}_0$ doit être "proche" de $p(T,\cdot,\mathbf{y}_0)$. L'exposant $(T,\mathbf{y}_0)$ de $\tilde{p}^{T,\mathbf{y}_0}$ indique que moyenne et covariance de $\tilde{p}$ dépendent du temps final et du point d'arrivée où l'on considère la densité. Comme pour la minoration l'idée est de linéariser la dynamique de la diffusion (5.2) autour d'une courbe déterministe. Comme c'est la variable spatiale d'arrivée qui est connue on va développer autour de la solution de l'EDO *rétrograde*

$$\frac{d}{dt}\boldsymbol{\theta}_{t,T}(\mathbf{y}_0) = \mathbf{F}\big(t,\boldsymbol{\theta}_{t,T}(\mathbf{y}_0)\big), \quad t\in[0,T]; \quad \boldsymbol{\theta}_{T,T}(\mathbf{y}_0) = \mathbf{y}_0.$$

Le modèle Gaussien associé s'écrit :

$$d\tilde{\mathbf{X}}_t = \big[\mathbf{F}(t,\boldsymbol{\theta}_{t,T}(\mathbf{y}_0)) + \mathbf{D_x F}(t,\boldsymbol{\theta}_{t,T}(\mathbf{y}_0))\big(\tilde{\mathbf{X}}_t - \boldsymbol{\theta}_{t,T}(\mathbf{y}_0)\big)\big]dt + B\sigma(t,\boldsymbol{\theta}_{t,T}(\mathbf{y}_0))dW_t, \quad t\in[0,T].$$

Désignons par $(\tilde{L}_t^{T,\mathbf{y}_0})_{0\le t\le T}$ le générateur de $\tilde{\mathbf{X}}$ gelé au temps T et au point $\mathbf{y}_0$. On rappelle que la quantité fondamentale pour le développement parametrix est le noyau

$$H(t,T,\mathbf{x},\mathbf{y}_0) = \big[L_{t,\mathbf{x}} - \tilde{L}_{t,\mathbf{x}}^{T,\mathbf{y}_0}\big]\big(\tilde{p}^{T,\mathbf{y}_0}(t,T,\mathbf{x},\mathbf{y}_0)\big),$$

cf. [MS67] ou Chapitre 3. Nous obtenons pour tout $N\ge 1$ le développement :

$$\begin{aligned} p(T,\mathbf{x}_0,\mathbf{y}_0) = \tilde{p}(0,T,\mathbf{x}_0,\mathbf{y}_0) &+ \sum_{k=1}^{N}\int_0^T\int_{\mathbb{R}^{nd}} \tilde{p}(0,t,\mathbf{x}_0,\mathbf{z})H^{\otimes k}(t,T,\mathbf{z},\mathbf{y}_0)d\mathbf{z}dt \\ &+ \int_0^T\int_{\mathbb{R}^{nd}} p(t,\mathbf{x}_0,\mathbf{z})H^{\otimes(N+1)}(t,T,\mathbf{z},\mathbf{y}_0)d\mathbf{z}dt, \end{aligned} \tag{5.11}$$

avec $\tilde{p}(0,t,\mathbf{x}_0,\mathbf{z}) := \tilde{p}^{t,\mathbf{z}}(0,t,\mathbf{x}_0,\mathbf{z})$ et où l'on rappelle

$$H^{\otimes(k+1)}(t,T,\mathbf{z},\mathbf{y}_0) := \int_t^T\int_{\mathbb{R}^{nd}} H(t,s,\mathbf{z},\mathbf{z}')H^{\otimes k}(s,T,\mathbf{z}',\mathbf{y}_0)d\mathbf{z}'ds.$$

Malheureusement nous n'arrivons pas à faire tendre N vers l'infini dans (5.11). A cause du transport *non linéaire* $\boldsymbol{\theta}$, nous n'avons pas réussi à contrôler les noyaux itérés $(H^{\otimes k})_{k\geq 1}$ uniformément en k. Nous renvoyons au Chapitre 3 pour le cas linéaire auquel la démarche usuelle s'applique.

Nous devons donc ici tronquer la série. Le point clé est que la formulation (5.7) permet d'établir que, pour N suffisamment grand, le terme de reste

$$\int_0^T \int_{\mathbb{R}^{nd}} p(t,\mathbf{x}_0,\mathbf{z}) H^{\otimes(N+1)}(t,T,\mathbf{z},\mathbf{y}_0) d\mathbf{z} dt$$

dans (5.11) admet une borne Gaussienne dont on déduit l'estimée énoncée. Cette technique est à notre connaissance nouvelle.

5.2.4 Une seconde approche pour la minoration : parametrix et chaînage

Nous avons développé dans [**10**] une deuxième méthode pour obtenir la minoration du Théorème 5.2.1. A partir du développement (5.11), en exploitant la propriété régularisante en temps du noyau de convolution H et les contrôles utilisés pour la borne supérieure, nous pouvons déduire dans l'esprit de [IKO62] la borne inférieure sur les "compacts" de la métrique en temps petit. Il existe $T_0 > 0$ tel que pour tout $(T,\mathbf{x},\mathbf{y}) \in (0,T_0] \times (\mathbb{R}^{nd})^2$ lorsque la métrique $d_T(\mathbf{x},\mathbf{y}) := T^{1/2}|\mathbb{T}_T^{-1}(\boldsymbol{\theta}_T(\mathbf{x}) - \mathbf{y})| \leq C_0,\ p(0,T,\mathbf{x},\mathbf{y}) \geq C_1 T^{-n^2 d/2}$.

Il reste ensuite à l'obtenir pour $d_T(\mathbf{x},\mathbf{y}) > C_0$. La propriété de Markov donne

$$p(T,\mathbf{x},\mathbf{y}) := \int_{(\mathbb{R}^{nd})^{N-1}} p(0,t_1,\mathbf{x},\mathbf{z}_1) \prod_{i=1}^{N-2} p(t_i,t_{i+1},\mathbf{z}_i,\mathbf{z}_{i+1}) p(t_{N-1},T,\mathbf{z}_{N-1},\mathbf{y}) d\mathbf{z}_1 ... d\mathbf{z}_{N-1}, \tag{5.12}$$

où $t_0 = 0 < t_1 < t_2 < ... < t_N = T,\ N \in \mathbb{N}^*$.

L'idée est ensuite de choisir N, les $(t_i)_{i\in[\![0,N]\!]}$ et des ensembles $(B_i)_{i\in[\![0,N]\!]}$ (avec les conventions $B_0 = \{\mathbf{x}\}, B_N = \{\mathbf{y}\}$) tels que pour $(\mathbf{z}_i,\mathbf{z}_{i+1}) \in B_i \times B_{i+1}, i \in [\![0,N-1]\!]$, l'on puisse appliquer la borne sur les compacts de la métrique pour $p(t_i,t_{i+1},\mathbf{z}_i,\mathbf{z}_{i+1})$. Si $(t_{i+1}-t_i)^{1/2}|\mathbb{T}_{t_{i+1}-t_i}^{-1}(\boldsymbol{\theta}_{t_{i+1},t_i}(\mathbf{z}_i) - \mathbf{z}_{i+1})| \leq C_0$, où $\dot{\boldsymbol{\theta}}_{s,t_i}(\mathbf{z}_i) = \mathbf{F}(s,\boldsymbol{\theta}_{s,t_i}(\mathbf{z}_i)),\ s \geq t_i,\ \boldsymbol{\theta}_{t_i,t_i}(\mathbf{z}_i) = \mathbf{z}_i$, alors $p(t_i,t_{i+1},\mathbf{z}_i,\mathbf{z}_{i+1}) \geq C_1/(t_{i+1}-t_i)^{n^2 d/2}$. De (5.12) on écrit

$$\begin{aligned} p(T,\mathbf{x},\mathbf{y}) &\geq \int_{\prod_{i=1}^{N-1} B_i} p(0,t_1,\mathbf{x},\mathbf{z}_1) \prod_{i=1}^{N-2} p(t_i,t_{i+1},\mathbf{z}_i,\mathbf{z}_{i+1}) p(t_{N-1},T,\mathbf{z}_{N-1},\mathbf{y}) d\mathbf{z}_1 ... d\mathbf{z}_{N-1} \\ &\geq \frac{C_1}{t_1^{n^2 d/2}} \prod_{i=1}^{N-1} \frac{C_1}{(t_{i+1}-t_i)^{n^2 d/2}} |B_i|. \end{aligned}$$

Si maintenant $t_{i+1} - t_i \leq T/N, |B_i| \geq C_2 (T^{1/2} d_T(\mathbf{x},\mathbf{y})/N)^{n^2 d}, \forall i \in [\![0,N-1]\!]$, il vient

$$p(T,\mathbf{x},\mathbf{y}) \geq C_1^N C_2^{N-1} \left(\frac{N}{T}\right)^{n^2 d/2} \left(\frac{d_T^2(\mathbf{x},\mathbf{y})}{N}\right)^{(N-1)n^2 d/2}.$$

Pour $N = \lceil K d_T^2(\mathbf{x},\mathbf{y}) \rceil \geq 1$, nous obtenons $p(t,\mathbf{x},\mathbf{y}) \geq \frac{C_3^N}{T^{n^2 d/2}}$, où $C_3 = C_1 C_2/(K+1) < 1$ pour K suffisamment grand. On retrouve donc la borne inférieure en temps petit à partir de l'écriture $p(T,\mathbf{x},\mathbf{y}) \geq \frac{\exp(N \ln(C_3))}{T^{-n^2 d/2}}$ en utilisant la définition de N. L'extension à un temps arbitraire se fait de nouveau par convolution ou changement d'échelle.

La construction des $(t_i)_{i\in[\![0,N]\!]}, (B_i)_{i\in[\![0,N]\!]}$ vérifiant les conditions précédentes, est présentée en appendice de [**10**]. L'idée est d'interpréter la construction usuelle dans le cas elliptique, cf. Chapitre 7 de Bass [Bas97], en terme de problème de contrôle déterministe de sorte à pouvoir le transposer à notre cadre.

Pour le cas non dégénéré, les ensembles B_i sont des boules dont les centres sont uniformément répartis sur la ligne droite reliant point de départ et d'arrivée. La ligne droite est ici la solution optimale du problème de contrôllabilité déterministe, $\boldsymbol{\phi}_t = \mathbf{x} + \frac{t}{T}(\mathbf{y}-\mathbf{x})$. Le déplacement $\frac{\mathbf{y}-\mathbf{x}}{T}$ est constant ce qui explique l'équirépartition. L'idée dans notre cas dégénéré consiste à disposer le centre des ensembles B_i le long de la trajectoire optimale $(\boldsymbol{\phi}_t)_{t\in[0,T]}$ solution de (5.8). A une localisation technique près[3] on définit $t_0 = 0,\ \forall i \geq 1,\ t_i := \inf\{t \geq t_{i-1} : \int_{t_{i-1}}^t |\varphi_s|^2 ds = I(T,\mathbf{x},\mathbf{y})/N\}$, i.e. on considère une équirépartition en terme de niveau d'énergie. Les centres des $(B_i)_{i\in[\![1,N-1]\!]}$ sont associés à $(\boldsymbol{\phi}_{t_i})_{i\in[\![1,N-1]\!]}$, la géométrie de ces ensembles est liée à la métrique naturelle.

3. qui consiste à prendre l'infimum avec T/N dans la définition ci-après en s'assurant que le nombre de pas de temps reste de l'ordre de N.

5.2.5 Extensions et perspectives

Ce travail ouvre de nombreuses perspectives. Tout d'abord le Théorème 5.2.1 est énoncé sous réserve d'unicité en loi. Si les coefficients de (5.2) vérifient **(A)** avec $\eta > 1/2$ nous obtenons cette identité en loi comme corollaire d'un principe de comparaison de Ishii et Lions [IL90] pour les solutions de viscosité d'équations non linéaires. Cet argument n'est pas très approprié dans notre cas dans la mesure où il n'utilise pas du tout l'aspect régularisant du noyau de la chaleur. Par ailleurs, dans les preuves aucune limite de ce type sur η n'est apparue. Dans un travail en cours[4], je m'intéresse au problème de martingale pour (5.2), dans le cas où a est uniformément elliptique et continu en espace, en gardant inchangées les autres hypothèses dans **(A)**. Comme dans le cas non dégénéré, cf. [SV79], l'idée est d'établir des inégalités de Calderón-Zygmund pour le modèle Gaussien. Une des difficultés est dans notre cadre d'introduire une bonne métrique parabolique qui tienne compte de l'aspect multi-échelle. De l'unicité du problème de martingale on déduirait immédiatement que le Théorème 5.2.1 est valable pour $\eta \in (0, 1]$. Les estimées de Calderón-Zygmund permettraient également d'aborder des EDP dégénérées semi-linéaires. J'essaie également de donner des contrôles ponctuels sur les dérivées de la densité à partir de l'analyse précédente.

En terme d'applications, on peut penser aux systèmes Hamiltoniens, deux problèmes associés à l'équation (5.2) considérée apparaissent naturellement : le développement asymptotique de la densité en temps petit, et l'estimation du régime stationnaire, ainsi que de la vitesse de convergence vers ce dernier, sous de bonnes conditions de stabilité. Pour l'estimée en temps petit, on s'attend à ce que le contrôle optimal dans la formulation (5.7) soit proche du contrôle optimal déterministe dans (5.8). Il faudrait raffiner l'analyse de stabilité pour l'établir rigoureusement. Adapter la démarche précédente à l'estimation en temps long sous conditions de type Lyapunov reste à faire. Une idée pourrait être de choisir comme contrôle dans (5.7) un contrôle optimal explicite associé à un processus "canonique" pour le régime limite considéré, i.e. Ornstein-Ulhenbeck si la force de rappel est à croissance linéaire.

Si l'on rajoute dans (5.2) un bruit sur la dernière composante on retrouve le cadre des modèles microscopiques de mécanique statistique pour la diffusion de chaleur. Dans un modèle linéaire d'interaction aux plus proches voisins comme celui de Bodineau et Lefevere [BL08], il serait intéressant de reprendre l'analyse précédente et d'essayer d'obtenir des bornes de déviations non asymptotiques en temps. Nous avons prévu avec Raphaël Lefevere de discuter à ce propos.

Nous avons considéré ici un régime Gaussien qui permettait de faire apparaître naturellement, et avec de bonnes estimées, le lien avec le contrôle déterministe. Une extension naturelle consiste à considérer d'autres régimes, i.e. lorsque des crochets de Lie d'ordre strictement supérieur à 1 interviennent pour engendrer l'espace. L'exemple le plus simple consiste à étudier la densité du couple $(x + W_t, \int_0^t (x + W_s)^k ds)_{t\geq 0}$ où $(W_t)_{t\geq 0}$ est le mouvement Brownien réel standard.

Pour $k = 2$ des formules "explicites" existent, cf. formule 1.9.8 de Borodin et Salminen [BS03] mais ne sont pas directement exploitables pour déduire des contrôles *à la Aronson* sur la densité. Pour $x = 0$ et k impair Chiara Cinti et Sergio Polidoro, des Universités de Bologne et Modena, ont obtenu dans [CP08] une borne inférieure. Les outils sont ceux de la théorie du potentiel pour les opérateurs sous-elliptiques, cf. Bonfiglioli *et al.* [BLU07]. Nous avons avec ces auteurs étendu ces résultats pour k quelconque. Ce travail est en cours de rédaction.

Par ailleurs, avec V. Konakov de l'Université de Moscou, nous avons montré par des techniques de calcul de Malliavin que ces estimées sont précises, i.e. qu'il existe une borne supérieure du même ordre. Pour $x_1 = 0$ et k quelconque nous obtenons $\exists c \in (0, 1], C := C(T) \geq 1$ telles que pour tout $(t, (0, x_2), (y_1, y_2)) \in (0, T] \times (\{0\} \times \mathbb{R}) \times \mathbb{R}^2$:

$$\frac{C^{-1}}{t^{(k+3)/2}} \exp\left(-c^{-1}\left\{\frac{|y_1|^2}{t} + \frac{|y_2 - x_2|^{2/k}}{t^{1+2/k}}\right\}\right) \leq p(t, (0, x_2), (y_1, y_2)) \leq \frac{C}{t^{(k+3)/2}} \exp\left(-c\left\{\frac{|y_1|^2}{t} + \frac{|y_2 - x_2|^{2/k}}{t^{1+2/k}}\right\}\right).$$

Ce travail est également en cours. L'idée est de bien comprendre la structure du transport de la condition initiale pour $x \neq 0$. Nous y arrivons lorsque $|y_2 - x_2| \geq t|x_1|^k$, i.e. lorsque l'on est en grand régime par rapport au transport déterministe. Le terme dans l'exponentielle précédente s'écrit alors $\frac{|y_1-x_1|^2}{t} + \frac{(|y_2-x_2|-t|x_1|^k)^{2/k}}{t^{1+2/k}}$.

Le transport en petits régimes reste à analyser. Une piste peut être d'utiliser des inégalités de Tchebychev exponentielles à l'aide des transformées de Laplace, voir Mansuy et Yor [MY09] pour le cas $k = 2$. La caractérisation complète, i.e. pour tous les régimes, est une étape nécessaire en vue d'appliquer une méthode paramétrix pour remplacer le Brownien par une diffusion.

4. S.Menozzi "Well posedness of the martingale problem for Kolmogorov's equations"

5.3 Estimation de densité et concentration

Nous présentons ici une application des estimées de type Aronson : l'obtention d'intervalles de confiance non asymptotiques pour la méthode de Monte Carlo. La domination Gaussienne permet en effet d'utiliser les outils usuels en concentration de la mesure, inégalité de Sobolev logarithmique et argument de Herbst (cf. Ledoux [Led99]), pour déduire des inégalités de déviation non asymptotiques.

Nous avons abordé le cas des diffusions :

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \tag{5.13}$$

où $(X_t)_{t\geq 0}$ est à valeurs dans $\mathbb{R}^d$ et $(W_t)_{t\geq 0}$ est un mouvement Brownien de $\mathbb{R}^d$. Par la suite nous ferons référence à l'équation précédente en parlant du cas (a). Nous avons également considéré :

$$\begin{aligned} X_t^1 &= x_1 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \\ X_t^2 &= x_2 + \int^t X_s^1 ds, \end{aligned} \tag{5.14}$$

où $X_t = (X_t^1, X_t^2)$ est à valeurs dans $\mathbb{R}^d \times \mathbb{R}^d$. L'équation (5.14) est un cas particulier de (5.2) pour $n = 2$ et $F(x_1, x_2) = (b(x_1, x_2), x_1)^*$. Nous en parlerons comme du cas (b).

Les coefficients b, σ des équations ci-dessus vérifient les hypothèses de la partie *opérateurs sous forme non divergence* de la Section 5.1 dans le cas non dégénéré (a), et **(H)** de la Section 5.2 en supposant $F_1 = b$ borné pour le cas dégénéré (b).

Dans le cas (a), nous avons considéré le schéma d'Euler de pas de temps $h := T/N$, $N \in \mathbb{N}^*$. Posons $\forall i \in \mathbb{N},\ t_i = ih$, et pour $t \geq 0$, $\phi(t) := \sup\{t_i : t_i \leq t\}$. La dynamique s'écrit :

$$X_t^h = x + \int_0^t b(\phi(s), X_{\phi(s)}^h)ds + \int_0^t \sigma(\phi(s), X_{\phi(s)}^h)dW_s. \tag{5.15}$$

Pour (b) nous avons défini :

$$\begin{aligned} X_t^{h,1} &= x_1 + \int_0^t b_1(\phi(s), X_{\phi(s)}^h)ds + \int_0^t \sigma(\phi(s), X_{\phi(s)}^h)dW_s, \\ X_t^{h,2} &= x_2 + \int_0^t X_s^{h,1} ds, \end{aligned} \tag{5.16}$$

où $X_t^h := (X_t^{h,1}, X_t^{h,2})$. L'équation (5.16) definit un schéma complètement simulable à incréments Gaussiens. Sur chaque pas de temps, les d dernières composantes sont l'intégrale d'un processus Gaussien.

L'erreur faible pour ces modèles est bien comprise, voir Chapitre 1 et 3. Il reste donc à contrôler, avec les notations de l'équation (1.5) :

$$\mathcal{E}_{M,h}^S := \frac{1}{M}\sum_{i=1}^M f(T, (X_T^h)^i) - \mathbb{E}_x[f(T, X_T^h)], \tag{5.17}$$

où $T > 0$ est fixé et $\left((X_T^h)^i\right)_{i\in\mathbb{N}^*}$ est une suite i.i.d de réalisations du schéma d'Euler (5.15) dans le cas (a) et (5.16) pour le cas (b).

Dans le cadre ergodique, le contrôle non asymptotique de l'erreur a été étudié par Malrieu et Talay [MT06a]. Les auteurs ont réussi à établir que le schéma d'Euler d'un processus de diffusion multi-dimensionnel à coefficient de diffusion constant vérifie une inégalité de Sobolev logarithmique, qui implique la concentration gaussienne (voir e.g. Ledoux [Led99]). La borne de concentration non asymptotique en découle aisément. Toujours dans le contexte ergodique, mentionnons le travail de Joulin et Ollivier pour les chaînes de Markov [JO09].

Notre but est ici différent. Nous souhaitons donner un contrôle pour $T > 0$ et ne pouvons appliquer directement la machinerie log-Sobolev, par ailleurs assez rigide pour le passage aux discrétisations, directement. Sous les hypothèses précédentes nous avons néanmoins pu établir des bornes Gaussiennes dans les cas (a) et (b). L'idée pour en dériver une borne non asymptotique est d'utiliser ensuite que la densité Gaussienne qui domine vérifie une inégalité de Sobolev logarithmique. On peut alors modifier l'argument de Herbst pour obtenir l'inégalité de déviation souhaitée. Il faut toutefois gérer la constante qui apparaît devant la densité Gaussienne dans l'estimée d'Aronson suivante.

Théorème 5.3.1 (Estimées d'"Aronson" Gaussiennes pour le schéma d'Euler) *Il existe* $c \in (0,1]$, $C \geq 1$, $C := C(T,d,\Lambda,|b|)$, $c := c(d,\Lambda)$, *tels que pour* $0 \leq j < j' \leq N$ *:*

$$C^{-1} p_{c^{-1}}(t_{j'} - t_j, x, x') \leq p^h(t_j, t_{j'}, x, x') \leq C p_c(t_{j'} - t_j, x, x'), \tag{5.18}$$

où p^h *désigne la densité du schéma d'Euler. Pour* $0 \leq s < t \leq T$, *dans le cas (a),* $p_c(t-s,x,x') := \left(\frac{c}{2\pi(t-s)}\right)^{d/2} \exp\left(-c\frac{|x'-x|^2}{2(t-s)}\right)$ *et dans le cas (b)*

$$p_c(t-s,x,x') := \left(\frac{\sqrt{3}c}{2\pi(t-s)^2}\right)^d \exp\left(-c\left\{\frac{|x_1'-x_1|^2}{4(t-s)} + 3\frac{|x_2'-x_2-\frac{x_1+x_1'}{2}(t-s)|^2}{(t-s)^3}\right\}\right).$$

Pour établir ce théorème, les techniques sont les mêmes qu'auparavant : i.e. représentation parametrix discrète de la densité des schémas pour obtenir la borne supérieure, et méthode de chaînage adaptée au cas semi-Markovien du schéma d'Euler pour la minoration. Néanmoins, ce résultat est à ma connaissance le premier sous les hypothèses usuelles précédentes. Il peut dans le cas non dégénéré se déduire de Gobet et Labart [GL08] sous des hypothèses plus fortes de régularité de b, σ.

Nous obtenons ensuite le résultat de concentration.

Théorème 5.3.2 (Concentration Gaussienne) *Considérons pour le cas (a) une fonction* $f : \mathbb{R}^d \to \mathbb{R}$ *Lipschitzienne en espace et mesurable bornée en temps, vérifiant* $|\nabla f(T,.)|_\infty \leq 1$. *Pour le cas (b) on suppose que* $f(T,x) = g(T, \mathbb{T}_T^{-1}x)$ *où* $\mathbb{T}_T^{-1} = \begin{pmatrix} I_d & 0 \\ 0 & T^{-1}I_d \end{pmatrix}$ *et* g *est une fonction Lipschitzienne en espace et mesurable bornée en temps, vérifiant* $|\nabla g(T,.)|_\infty \leq 1$.

Pour les constantes c, C *du Théorème 5.3.1, pour tout* $h = T/N$, $N \geq 1$,

$$\forall r > 0, \quad \forall M \geq 1, \quad \mathbb{P}_x\left[|\mathcal{E}_{M,h}^S| \geq r + \delta_{C,\alpha(T)}\right] \leq 2\exp\left(-\frac{M}{\alpha(T)}r^2\right),$$

où

$$\frac{1}{\alpha(T)} = \begin{cases} \frac{c}{2T} & \text{dans le cas (a)}, \\ (4-\sqrt{13})\frac{c}{2T} & \text{dans le cas (b)}, \end{cases}$$

et $\delta_{C,\alpha(T)} = 2\sqrt{\alpha(T)\log C}$.

Lorsque de plus f tend vers l'infini dans une direction, nous établissons une borne inférieure du même ordre. Nous renvoyons à **[10]** pour un énoncé complet du théorème.

Le terme correctif $\delta_{C,\alpha(T)}$ correspond à une borne supérieure pour la distance de Wasserstein entre les lois de densités $p^h(T,x,.)$ et $p_c(T,x,.)$ lorsque celles-ci vérifient la condition de domination de l'équation (5.18). Si $C = 1$ ce terme correctif disparaît. Dans ce cas là on est directement Gaussien. C'est en revanche le prix à payer dans le cas général. Ce terme est fixé une fois pour toutes. Pour M grand, l'ordre de déviation typique donné par le théorème central limite est $r = \frac{c_0}{\sqrt{M}}$ pour $c_0 > 0$ dépendant d'un seuil fixé et de la variance de $f(T, X_T^h)$. Du fait de la technique employée, i.e. domination par une densité qui vérifie une inégalité de log-Sobolev, notre borne n'est pas pertinente pour M tendant vers l'infini à T fixé. En revanche elle l'est tant que $r = \frac{c_0}{\sqrt{M}} \geq \left(\frac{2T}{c}\right)^{1/2}$ et en particulier lorsque $r = \frac{c_0}{\sqrt{M}}$ et $\left(\frac{2T}{c}\right)^{1/2}$ sont du même ordre. Ainsi, si le temps T est "petit", la borne non asymptotique restera valable jusqu'à M assez grand, relié à T par l'inégalité précédente. Un autre cas intéressant est celui où pour $T > 0$ fixé il est difficile de simuler un nombre important de trajectoires de Monte Carlo. On peut penser pour illustrer ce dernier cas à un schéma d'Euler dont l'évaluation des coefficients ferait appel à des codes de calculs lourds.

Annexe A

Unicité en loi pour certains systèmes dégénérés

Nous indiquons dans cette section les idées principales de la note **[12]** qui établit un lien entre la méthode parametrix *à la Mc Kean et Singer* présentée ci-avant et la technique employée par Bass et Perkins [BP09] pour prouver l'unicité du problème de martingale pour un opérateur sous forme non divergence non dégénéré.

Cette identification et les contrôles obtenus dans l'article **[10]** permettent ensuite de prouver l'unicité du problème de martingale associé à l'équation dégénérée (5.2) sous les hypothèses **(A)** de la Section 5.2 pour tout $\eta \in (0,1]$. Ceci étend la validité du Théorème 5.2.1, comme nous l'espérions (cf. Section 5.2.5).

A.1 Rappel sur le modèle et la construction du noyau H

Nous rappelons la dynamique de $\mathbf{X}_t$ qui sous forme compacte s'écrivait :

$$d\mathbf{X}_t = \mathbf{F}(t, \mathbf{X}_t) + B\sigma(t, \mathbf{X}_t)dW_t, \tag{A.1}$$

La linéarisation utilisée dans la méthode parametrix pour la majoration de la densité faisait intervenir pour $T > 0, \mathbf{y} \in \mathbb{R}^{nd}$ l'équation linéaire :

$$\begin{aligned} d\tilde{\mathbf{X}}_t^{T,\mathbf{y}} &= \big[\mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y})) + D\mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y}))\big(\tilde{\mathbf{X}}_t^{T,\mathbf{y}} - \boldsymbol{\theta}_{t,T}(\mathbf{y})\big)\big]dt \\ &\quad + B\sigma(t, \boldsymbol{\theta}_{t,T}(\mathbf{y}))dW_t, \quad 0 \le t \le T, \end{aligned} \tag{A.2}$$

où $(\boldsymbol{\theta}_{t,T}(\mathbf{y}))_{t\ge 0}$ est solution de l'EDO $[d/dt]\boldsymbol{\theta}_{t,T}(\mathbf{y}) = \mathbf{F}(t, \boldsymbol{\theta}_{t,T}(\mathbf{y}))$, $t \ge 0$, avec condition terminale $\boldsymbol{\theta}_{T,T}(\mathbf{y}) = \mathbf{y}$ et $\forall (t, \mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$, $D\mathbf{F}(t, \mathbf{x}) = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ D_{\mathbf{x}_1}\mathbf{F}_2(t,\mathbf{x}) & 0 & \cdots & \cdots & 0 \\ 0 & D_{\mathbf{x}_2}\mathbf{F}_3(t,\mathbf{x}) & 0 & 0 & \vdots \\ \vdots & 0 & \ddots & \vdots & \\ 0 & \cdots & 0 & D_{\mathbf{x}_{n-1}}\mathbf{F}_n(t,\mathbf{x}) & 0 \end{pmatrix}$ est la sous-diagonale de la matrice Jacobienne $\mathbf{D_x F}$. Désignons par $\tilde{p}^{T,\mathbf{y}}(t, T, \mathbf{x}, .)$ la densité de $\tilde{\mathbf{X}}_T^{T,\mathbf{y}}$ partant de $\mathbf{x}$ en t.

Pour $0 \le s < t \le T$, $\mathbf{z}, \mathbf{y} \in \mathbb{R}^{nd}$, définissons le noyau de convolution H comme :

$$\begin{aligned} H(s,t,\mathbf{z},\mathbf{y}) &= \{\langle \mathbf{F}(s,\mathbf{z}), \mathbf{D_z}\tilde{p}^{t,\mathbf{y}}(s,t,\mathbf{z},\mathbf{y})\rangle + \frac{1}{2}\mathrm{Tr}(a(s,\mathbf{z})D^2_{\mathbf{z}_1}\tilde{p}^{t,\mathbf{y}}(s,t,\mathbf{z},\mathbf{y}))\} - \\ &\quad \{\langle \mathbf{F}(s, \boldsymbol{\theta}_{s,t}(\mathbf{y})) + D\mathbf{F}(s, \boldsymbol{\theta}_{s,t}(\mathbf{y}))(\mathbf{z} - \boldsymbol{\theta}_{s,t}(\mathbf{y})), \mathbf{D_z}\tilde{p}^{t,\mathbf{y}}(s,t,\mathbf{z},\mathbf{y})\rangle \\ &\quad + \frac{1}{2}\mathrm{Tr}(a(s, \boldsymbol{\theta}_{s,t}(\mathbf{y}))D^2_{\mathbf{z}_1}\tilde{p}^{t,\mathbf{y}}(s,t,\mathbf{z},\mathbf{y}))\} = (L_{s,\mathbf{z}} - \tilde{L}^{t,\mathbf{y}}_{s,\mathbf{z}})\tilde{p}^{t,\mathbf{y}}(s,t,\mathbf{z},\mathbf{y}), \end{aligned} \tag{A.3}$$

où L est le générateur de $\mathbf{X}$ et $\tilde{L}^{t,\mathbf{y}}_{s,\mathbf{z}}$ et $\tilde{p}^{t,\mathbf{y}}(s,t,\mathbf{z},.)$ désignent respectivement le générateur à l'instant s et la densité à l'instant t de $\tilde{\mathbf{X}}^{t,\mathbf{y}}$, $\tilde{\mathbf{X}}^{t,\mathbf{y}}_s = \mathbf{z}$ avec coefficients "gelés" en $t, \mathbf{y}$. L'indice $\mathbf{z}$ précise le paramètre de différentiation.

Nous avons établi dans **[10]** (cf. Lemma 5.5)

Lemme A.1.1 *Soit $T_0 > 0$ fixé. Il existe une constante $C_{A.1.1} > 0$, dépendant de (**A**) et T_0 t.q. pour tout $t \in [0,T)$, $T \leq T_0$ et $\mathbf{z}, \mathbf{y} \in \mathbb{R}^{nd}$,*

$$|H(t,T,\mathbf{z},\mathbf{y})| \leq C_{A.1.1}(T-t)^{\frac{n}{2}-1} g_{C_{A.1.1},T-t}\big(\mathbf{z} - \boldsymbol{\theta}_{t,T}(\mathbf{y})\big),$$

où pour tout $a > 0, t > 0$, $g_{a,t}(\mathbf{y}) = t^{-n^2\frac{d}{2}} \exp(-a^{-1}t|\mathbb{T}_t^{-1}\mathbf{y}|^2)$, $\quad a, t > 0$, $\mathbf{y} \in \mathbb{R}^{nd}$.

Par le même type de techniques que celles utilisées pour prouver le lemme précédent, on peut également établir :

Lemme A.1.2 *Soit une fonction $h \in C^0(\mathbb{R}^{nd}, \mathbb{R})$. Définissons pour tout $(s,\mathbf{x}) \in [0,T] \times \mathbb{R}^{nd}$,*

$$\forall \varepsilon > 0,\ \boldsymbol{\Xi}^{\varepsilon}(s,\mathbf{x}) := \int_{\mathbb{R}^{nd}} d\mathbf{y} h(s,\mathbf{y}) \tilde{p}^{s+\varepsilon,\mathbf{y}}(s, s+\varepsilon, \mathbf{x}, \mathbf{y}).$$

Alors $\boldsymbol{\Xi}^{\varepsilon}(s,x)$ converge de façon bornée et ponctuellement vers $h(s,\mathbf{x})$ lorsque $\varepsilon \to 0$.

A.2 Unicité du problème de martingale sous (A)

A.2.1 Résultats principaux

Le résultat principal de la note est le suivant :

Théorème A.2.1 *Sous* **(A)** *et pour $\eta \in (0,1]$, le problème de martingale associé au générateur L dans* (A.3) *est bien posé. En particulier, on a l'unicité en loi pour l'EDS* (A.1).

Nous en déduisons le corollaire suivant.

Corollaire A.2.1 *Sous* **(A)** *et pour $\eta \in (0,1]$, l'unique solution faible de* (A.1) *admet pour tout $t > 0$ une densité qui vérifie les estimées d'Aronson du théorème 5.2.1.*

A.2.2 Preuve du théorème A.2.1

Supposons que l'on ait deux solutions $\mathbb{P}_1, \mathbb{P}_2$ du problème de martingale associé à $(L_t)_{t\in[s,T]}$ partant de $\mathbf{x}$ en s. Nous pouvons sans perte de généralité supposer $T \leq 1$. Pour une fonction $f : [0,T] \times \mathbb{R}^{nd} \to \mathbb{R}$ borélienne bornée, définissons :

$$S^i f := \mathbb{E}_i \left[\int_s^T dt f(t, \mathbf{X}_t) \right],\ i \in [\![1,2]\!],$$

où $(\mathbf{X}_t)_{t\in[s,T]}$ désigne ici le processus canonique associé à $(\mathbb{P}_i)_{i\in[\![1,2]\!]}$. Précisons (comme il est mentionné dans [BP09]) que Sf est seulement une fonctionnelle linéaire et non une fonction dans la mesure où $\mathbb{P}_i$ ne provient pas forcément d'un processus de Markov. Posons :

$$S^{\Delta} f := S^1 f - S^2 f,\ \Theta := \sup_{\|f\|_\infty \leq 1} |S^{\Delta} f|.$$

Remarquons que $\Theta \leq T - s$.

Si $f \in C_0^2([0,T) \times \mathbb{R}^{nd}, \mathbb{R})$, par définition du problème de martingale nous obtenons :

$$f(s,\mathbf{x}) + \mathbb{E}_i \left[\int_s^T dt (\partial_t + L_t) f(t, \mathbf{X}_t) \right] = 0,\ i \in [\![1,2]\!]. \tag{A.4}$$

Pour $\mathbf{y} \in \mathbb{R}^{nd}$ fixé et $\varepsilon \geq 0$, introduisons $\forall f \in C_0^{1,2}([0,T) \times \mathbb{R}^{nd}, \mathbb{R})$ la fonction de Green

$$\forall (s,\mathbf{x}) \in [0,T) \times \mathbb{R}^{nd},\ G^{\varepsilon,\mathbf{y}} f(s,\mathbf{x}) := \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{z} \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t,\mathbf{x},\mathbf{z}) f(t,\mathbf{z}). \tag{A.5}$$

On vérifie aisément que

$$(\partial_s + \tilde{L}_{s,\mathbf{x}}^{t+\varepsilon,\mathbf{y}}) \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t,\mathbf{x},\mathbf{z}) = 0, \forall (s,\mathbf{x},\mathbf{z}) \in [0,t) \times (\mathbb{R}^{nd})^2,\ \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t,\mathbf{x},.) \underset{t\downarrow s}{\longrightarrow} \delta_{\mathbf{x}}(.). \tag{A.6}$$

Si l'on définit pour toute fonction $f \in C_0^{1,2}([0,T) \times \mathbb{R}^{nd}, \mathbb{R}),\ (s,x) \in [0,T) \times \mathbb{R}^{nd}$,

$$M_{s,\mathbf{x}}^{\varepsilon,\mathbf{y}} f(s,\mathbf{x}) := \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{z} \tilde{L}_{s,\mathbf{x}}^{t+\varepsilon,\mathbf{y}} \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t,\mathbf{x},\mathbf{z}) f(t,\mathbf{z}),$$

nous obtenons des équations (A.5), (A.6) :

$$(\partial_s G^{\varepsilon,\mathbf{y}} f + M_{s,\mathbf{x}}^{\varepsilon,\mathbf{y}} f)(s,\mathbf{x}) = -f(s,\mathbf{x}), \forall (s,\mathbf{x}) \in [0,T) \times \mathbb{R}^{nd}. \tag{A.7}$$

Introduisons désormais pour une fonction régulière $h \in C_0^{1,2}([0,T) \times \mathbb{R}^{nd}, \mathbb{R})$ et pour tout $(s,\mathbf{x}) \in [0,T) \times \mathbb{R}^{nd}$:

$$\begin{aligned} \Phi^{\varepsilon,\mathbf{y}}(s,\mathbf{x}) &:= \tilde{p}^{s+\varepsilon,\mathbf{y}}(s,s+\varepsilon,\mathbf{x},\mathbf{y}) h(s,\mathbf{y}), \\ \Psi_\varepsilon(s,\mathbf{x}) &:= \int_{\mathbb{R}^{nd}} d\mathbf{y} G^{\varepsilon,\mathbf{y}}(\Phi^{\varepsilon,\mathbf{y}})(s,\mathbf{x}). \end{aligned}$$

L'identité (A.5) donne :

$$\begin{aligned} \Psi_\varepsilon(s,\mathbf{x}) &:= \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{z} \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t,\mathbf{x},\mathbf{z}) \Phi^{\varepsilon,\mathbf{y}}(t,\mathbf{z}) \\ &= \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{z} \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t,\mathbf{x},\mathbf{z}) \tilde{p}^{t+\varepsilon,\mathbf{y}}(t,t+\varepsilon,\mathbf{z},\mathbf{y}) h(t,\mathbf{y}) \\ &= \int_{\mathbb{R}^{nd}} d\mathbf{y} \int_s^T dt \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t+\varepsilon,\mathbf{x},\mathbf{y}) h(t,\mathbf{y}), \end{aligned} \tag{A.8}$$

en utilisant la propriété de semi-groupe de la densité gelée. Ecrivons pour tout $(s,x) \in [0,T) \times \mathbb{R}^{nd}$,

$$\begin{aligned} (\partial_s + L_s)\Psi_\varepsilon(s,x) &= \int_{\mathbb{R}^{nd}} d\mathbf{y} (\partial_s + L_s)(G^{\varepsilon,\mathbf{y}} \Phi^{\varepsilon,\mathbf{y}})(s,\mathbf{x}) \\ &= \int_{\mathbb{R}^{nd}} d\mathbf{y} (\partial_s G^{\varepsilon,\mathbf{y}} \Phi^{\varepsilon,\mathbf{y}} + M_{s,\mathbf{x}}^{\varepsilon,\mathbf{y}} \Phi^{\varepsilon,\mathbf{y}})(s,\mathbf{x}) + \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_s G^{\varepsilon,\mathbf{y}} \Phi^{\varepsilon,\mathbf{y}} - M_{s,\mathbf{x}}^{\varepsilon,\mathbf{y}} \Phi^{\varepsilon,\mathbf{y}})(s,\mathbf{x}) \\ &= -\int_{\mathbb{R}^{nd}} d\mathbf{y} \Phi^{\varepsilon,\mathbf{y}}(s,\mathbf{x}) + \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_s G^{\varepsilon,\mathbf{y}} \Phi^{\varepsilon,\mathbf{y}} - M_{s,\mathbf{x}}^{\varepsilon,\mathbf{y}} \Phi^{\varepsilon,\mathbf{y}})(s,\mathbf{x}) := I_1^\varepsilon + I_2^\varepsilon, \end{aligned}$$

en utilisant (A.7) pour l'avant dernière égalité. Le Lemme A.1.2 donne $I_1^\varepsilon \underset{\varepsilon \to 0}{\longrightarrow} -h(s,\mathbf{x})$. Par ailleurs,

$$\begin{aligned} I_2^\varepsilon &= \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} (L_s - \tilde{L}_{s,\mathbf{x}}^{t+\varepsilon,\mathbf{y}}) \tilde{p}^{t+\varepsilon,\mathbf{y}}(s,t+\varepsilon,\mathbf{x},\mathbf{y}) h(t,\mathbf{y}) \\ &= \int_s^T dt \int_{\mathbb{R}^{nd}} d\mathbf{y} H(s,t+\varepsilon,\mathbf{x},\mathbf{y}) h(t,\mathbf{y}). \end{aligned}$$

Du Lemme A.1.1,

$$\begin{aligned} |I_2^\varepsilon| &\leq C_{A.1.1} \|h\|_\infty \int_s^T dt (t+\varepsilon-s)^{-1+\eta/2} \int_{\mathbb{R}^{nd}} d\mathbf{y}\, g_{C_{A.1.1},t+\varepsilon-s}(\mathbf{x} - \boldsymbol{\theta}_{s,t+\varepsilon}(\mathbf{y})) \\ &\leq \|h\|_\infty C_{A.9} \int_s^T dt (t+\varepsilon-s)^{-1+\eta/2} \int_{\mathbb{R}^{nd}} d\mathbf{y}\, g_{C_{A.9},t+\varepsilon-s}(\boldsymbol{\theta}_{t+\varepsilon,s}(\mathbf{x}) - \mathbf{y}) \leq C_{A.9} \{(T-s) \vee \varepsilon\}^{\eta/2} \|h\|_\infty, \end{aligned} \tag{A.9}$$

en utilisant la propriété de bi-Lipschitz continuité du flot pour l'avant dernière inégalité, en modifiant $C_{A.9}$ dans la dernière (cette constante modifiée ne dépend toutefois que de **(A)**). Ainsi, pour T suffisamment petit nous avons par (A.9)

$$|I_2^\varepsilon| \leq \frac{1}{2} \|h\|_\infty. \tag{A.10}$$

L'équation (A.4) et la définition précédente de S^Δ donnent :

$$S^\Delta((\partial_s + L_s)\Psi_\varepsilon) = 0 \Rightarrow |S^\Delta I_1^\varepsilon| = |S^\Delta I_2^\varepsilon|.$$

Par la partie convergence bornée du Lemme A.1.2 et (A.10), nous obtenons :

$$|S^{\Delta}h| = \lim_{\varepsilon\to 0} |S^{\Delta} I_1^{\varepsilon}| = \lim_{\varepsilon\to 0} |S^{\Delta} I_2^{\varepsilon}| \leq \Theta \limsup_{\varepsilon\to 0} |I_2^{\varepsilon}| \leq \Theta \frac{\|h\|_{\infty}}{2}.$$

Par un argument de classe monotone, l'inégalité précédente reste vraie pour toute fonction h mesurable bornée, à support compact dans $[0,T)\times\mathbb{R}^{nd}$. En prenant le supremum sur $\|h\|_{\infty}\leq 1$, on arrive à $\Theta \leq \frac{1}{2}\Theta$ ce qui donne $\Theta = 0$ comme $\Theta < +\infty$. Ainsi, $\mathbb{E}_1\left[\int_s^T dth(t,\mathbf{X}_t)\right] = \mathbb{E}_2\left[\int_s^T dth(t,\mathbf{X}_t)\right]$ ce qui prouve le résultat sur l'intervalle $[0,T]$. Les probabilités conditionnelles régulières permettent d'étendre le résultat sur $\mathbb{R}^+$ (voir Chapitre 6.2 de Stroock et Varadhan [SV79]).

Bibliographie

[Als94] G. Alsmeyer. On the Markov renewal theorem. *Stoch. Proc. Appl.*, 50(1) :37–56, 1994.

[Ant93] F. Antonelli. Backward-Forward Stochastic Differential Equations. *Ann. Appl. Probab.*, 3-3 :777–793, 1993.

[Aro67] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73 :890–896, 1967.

[Bal90] V. Bally. On the connection between the Malliavin covariance matrix and Hörmander's condition. *J. Funct. Anal.*, 96–2 :219–255, 1990.

[Bal95] P. Baldi. Exact asymptotics for the probability of exit from a domain and applications to simulation. *Ann. Prob.*, 23(4) :1644–1670, 1995.

[Bas97] R. F. Bass. *Diffusions and Elliptic Operators*. Springer, 1997.

[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.

[BDH+03] P. Briand, B. Delyon, Y. Hu, E. Pardoux, and L. Stoica. L^p solutions of backward stochastic differential equations. *Stoch. Proc. Appl.*, 108 :109–129, 2003.

[BGJ87] K. Bichteler, J.B. Gravereaux, and J. Jacod. *Malliavin Calculus for processes with jumps*. Stochastics Monographs, volume 2, 1987.

[BGK97] M. Broadie, P. Glasserman, and S. Kou. A continuity correction for discrete barrier options. *Mathematical Finance*, 7 :325–349, 1997.

[BGK99] M. Broadie, P. Glasserman, and S. Kou. Connecting discrete and continuous path-dependent options. *Fin. and Stoch.*, 3 :55–82, 1999.

[BL82] A. Bensoussan and J. L. Lions. *Contrôle impulsionnel et inéquations quasi variationnelles*. Gauthiers Villars, 1982.

[BL91] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. *Prob. Th. Rel. Fields*, 90(3) :377–402, 1991.

[BL97] G. Barles and E. Lesigne. SDE, BSDE and PDE. Backward stochastic differential equations (Paris 95-96). *Stat. and Prob. Letters*, Pitman research, 364 :47–80, 1997.

[BL08] T. Bodineau and L. Lefevere. Large deviations of lattice Hamiltonian dynamics coupled to stochastic thermostats. *Journal of Statistical Physics*, 133 :1–27, 2008.

[BLP07] N. Bruti-Liberati and E. Platen. Strong approximations of stochastic differential equations with jumps. *J. Comput. Appl. Math.*, 205–2 :982–1001, 2007.

[BLU07] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni. *Stratified Lie groups and potential theory for their sub-Laplacians*. Springer, 2007.

[BM01] V. Bally and A. Matoussi. Weak solutions for SPDEs and backward doubly stochastic differential equations. *J. Theoret. Probab.*, 14–1 :125–164, 2001.

[BM03] Y. Belopolskaya and G. N. Milstein. An approximation method for Navier–Stokes equations based on probabilistic approach. *Stat. and Prob. Letters*, 64 :201–211, 2003.

[BP03] V. Bally and G. Pagès. A quantization algorithm for solving discrete time multi-dimensional optimal stopping problems. *Bernoulli*, 9–6 :1003–1049, 2003.

[BP09] R.F. Bass and E.A. Perkins. A new technique for proving uniqueness for martingale problems. *From Probability to Geometry (I) : Volume in Honor of the 60th Birthday of Jean-Michel Bismut*, pages 47–53, 2009.

[BPR06] A. Beskos, O. Papaspiliopoulos, and G. O. Roberts. Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, 12–6 :1077–1098, 2006.

[BPV01] E. Barucci, S. Polidoro, and V. Vespri. Some results on partial differential equations and asian options. *Math. Models Methods Appl. Sci*, 3 :475–497, 2001.

[BR76] R. Bhattacharya and R. Rao. *Normal approximations and asymptotic expansions.* Wiley and sons, 1976.

[BS97] R. Bañuelos and R.G. Smits. Brownian Motion in cones. *Prob. Th. Rel. Fields*, 108 :299–319, 1997.

[BS03] A. N. Borodin and P. Salminen. *Handbook of Brownian motion : facts and formulae.* Birkhauser, 2003.

[BT96a] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations : I. Convergence rate of the distribution function. *Prob. Th. Rel. Fields*, 104-1 :43–60, 1996.

[BT96b] 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 :93–128, 1996.

[BT04] B. Bouchard and N. Touzi. Discrete time approximation and Monte-Carlo simulation of Backward Stochastic Differential Equations. *Stoch. Proc. Appl.*, 111 :175–206, 2004.

[BZ08] C. Bender and J. Zhang. Time discretization and Markovian Iteration for Coupled FBSDEs. *Annals of Applied Probability*, 18 :143–177, 2008.

[Cat90] P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre - I. Résolvantes, théorème de Hörmander et applications. *Bull. Sc. Math., 2ème série*, 114 :421–462, 1990.

[Cat91] P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre - II. Problème de Dirichlet. *Bull. Sc. Math., 2ème série*, 115 :81–122, 1991.

[Che97] D. Chevance. Numerical methods for Backward Stochastic Differential Equations. *Publ. Newton Inst., Cambridge University Press*, pages 232–234, 1997.

[CIL92] M. G. Crandall, I. Ishii, and Lions P. L. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27 :1–67, 1992.

[CJ59] H. Carslaw and J. Jaeger. *Conduction of Heat in Solids.* Oxford Univ. Press, 1959.

[CKG06] C. Costantini, N. El Karoui, and E. Gobet. Boundary sensitivities for diffusion processes in time dependent domains. *Appl. Math. Optim.*, 54–2 :159–187, 2006.

[CM02] P. Cattiaux and L. Mesnager. Hypoelliptic non-homogeneous diffusions. *Prob. Th. Rel. Fields*, 123(4) :453–483, 2002.

[Cor07] J.-M. Coron. *Control and nonlinearity.* Mathematical Surveys and Monographs, 136, AMS, 2007.

[CP08] C. Cinti and S. Polidoro. Harnack inequalities and lifting procedure for evolution hypoelliptic equations. *Lect. Notes Semin. Interdisc. Mat.*, 7 :93–105, 2008.

[Del02] F. Delarue. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stoch. Proc. and App.*, 99 :209–286, 2002.

[DG06] F. Delarue and G. Guatteri. Weak Existence and Uniqueness for FBSDEs. *Stoc. Proc. and Appl.*, 116 :1712–1742, 2006.

[DK06] R. A. Doney and A. E. Kyprianou. Overshoots and undershoots of Lévy processes. *Ann. Appl. Probab.*, 16–1 :91–106, 2006.

[DL06] M. Deaconu and A. Lejay. A random walk on rectangles algorithm. *Methodology And Computing In Applied Probability*, 8(1) :135–151, 2006.

[DMP96] J. Douglas, J. Ma, and P. Protter. Numerical methods for Forward-Backward Stochastic Differential Equations. *Ann. Appl. Prob.*, 6 :940–968, 1996.

[DP97] R. Darling and E. Pardoux. Backward SDE with random terminal time and applications to semilinear elliptic PDE. *Ann. Probab*, 25 :1135–1159, 1997.

[Dyn63] E. B Dynkin. *Markov Processes.* Springer Verlag, 1963.

[EPRB99] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Comm. Math. Phys.*, 201–3 :657–697, 1999.

[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.

[Fre85] M. Freidlin. *Functional integration and Partial differential equations.* Annals of Mathematics studies, Princeton University Press, 1985.

[Fri64] A. Friedman. *Partial differential equations of parabolic type.* Prentice-Hall, 1964.

[Fri75] A. Friedman. *Stochastic differential equations and applications. Vol. 1.* Academic Press, 1975.

[FW98] M. Freidlin and A. Wentzell. *Random Perturbations of Dynamical Systems.* Springer, 1998.

[GL00] S. Graf and H. Lushgy. *Foundations of quantization for random vectors.* LNM-1730, Springer-Verlag, 2000.

[GL07] E. Gobet and C. Labart. Error expansion for the discretization of Backward Stochastic Differential Equations. *Stochastic Processes and Applications*, 117–7 :803–829, 2007.

[GL08] E. Gobet and C. Labart. Sharp estimates for the convergence of the density of the euler scheme in small time. *Electron. Commun. Probab.*, 13 :352–363, 2008.

[GLW05] E. Gobet, J.P. Lemor, and X. Warin. A regression-based Monte-Carlo method to solve backward stochastic differential equations. *Ann. Appl. Proba.*, 15 :2172–2202, 2005.

[GLW06] E. Gobet, J.P. Lemor, and X. Warin. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, 12 :889–916, 2006.

[GM02] M.G. Garroni and J. L. Menaldi. *Elliptic problems for integro-differential operators.* Chapman and Hall, 02.

[GM05] E. Gobet and S. Maire. Sequential Control Variates for Functionals of Markov Processes. *Siam. Journal of Num. Analysis*, 43-3 :1256–1275, 2005.

[Gob00] E. Gobet. Euler schemes for the weak approximation of killed diffusion. *Stoch. Proc. Appl.*, 87 :167–197, 2000.

[Gob01] E. Gobet. Euler schemes and half-space approximation for the simulation of diffusions in a domain. *ESAIM : Prob. Stat.*, 5 :261–297, 2001.

[Gri92] P. Grisvard. *Variational methods in the study of singularities.* LMN, Springer, 1992.

[GT77] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order.* Springer Verlag, 1977.

[Guy06] J. Guyon. Euler scheme and tempered distributions. *Stochastic Process. Appl.*, 116–6 :877–904, 2006.

[Hau02] E. Hausenblas. Error analysis for approximation of stochastic differential equations driven by Poisson random measures. *SIAM J. Num. Anal.*, 40–1 :87–113, 2002.

[HN04] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the fokker-planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, 171–2 :151–218, 2004.

[Hör67] L. Hörmander. Hypoelliptic second order differential operators. *Acta. Math.*, 119 :147–171, 1967.

[HP57] E. Hille and R. S. Philips. *Functional Analysis and semigroups.* AMS, 57.

[IKO62] A. M. Il'in, A. S. Kalashnikov, and O. A. Oleinik. Second-order linear equations of parabolic type. *Uspehi Mat. Nauk*, 17–3(105) :3–146, 1962.

[IL90] H. Ishii and P.-L. Lions. Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. *J. Diff. Equations*, 83 :26–78, 1990.

[IP06] P. Imkeller and I. Pavlyukevich. First exit times of sdes driven by stable lÈvy processes. *Stoc. Proc. Appl*, 116–4 :611–642, 2006.

[IW89] N. Ikeda and . Watanabe. *Stochastic Differential Equations and Diffusion processes.* North Holland, 1989.

[Jac04] J. Jacod. The Euler scheme for lévy driven stochastic differential equations : limit theorems. *Ann. Probab.*, 5(32) :1830–1872, 2004.

[JKMP05] J. Jacod, T. G. Kurtz, S. Méléard, and P. Protter. The approximate euler method for Lévy driven stochastic differential equations. *Ann. Inst. H. PoincarÈ Probab. Statist*, 41–3 :523–558, 2005.

[JMW96] A. Janicki, Z. Michna, and A. Weron. Approximation of stochastic differential equations driven by α-stable Lévy motion. *Appl. Math. (Warsaw)*, 24–2 :149–168, 1996.

[JO09] A. Joulin and Y. Ollivier. Curvature, concentration, and error estimates for Markov chain Monte Carlo. *http ://arxiv.org/abs/0904.1312*, 2009.

[JP98] J. Jacod and P. Protter. Asymptotic error distributions for the euler method for stochastic differential equations. *Ann. Probab*, 26–1 :267–307, 1998.

[KHP94] A. Kohatsu-Higa and P. Protter. The Euler scheme for SDEs driven by semimartingales. *in Stochastic analysis on infinite dimensional spaces, 1994. H. Kunita and H.Kuo (Eds.)*, 310 :141–151, 1994.

[KHT10] A. Kohatsu Higa and P. Tankov. Jump-adapted discretization schemes for Lévy-driven sdes. *Preprint*, 2010.

[KM00] V. Konakov and E. Mammen. Local limit theorems for transition densities of Markov chains converging to diffusions. *Prob. Th. Rel. Fields*, 117 :551–587, 2000.

[KM02] V. Konakov and E. Mammen. Edgeworth type expansions for euler schemes for stochastic differential equations. *Monte Carlo Methods Appl.*, 8–3 :271–285, 2002.

[KM09] V. Konakov and E. Mammen. Small time edgeworth-type expansions for weakly convergent nonhomogeneous markov chains. *Probab. Theory Related Fields*, 143–1-2 :137–176, 2009.

[KMR97] V.A. Kozlov, V.G. Maz'ya, and J. Rossmann. *Elliptic boundary values problem in domains with singularities.* Math. Surveys and Monographs, 52. AMS, 1997.

[Kob00] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Prob.*, 28 :558–602, 2000.

[Kol34] A. N. Kolmogorov. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math.*, 2-35 :116–117, 1934.

[Kol00] V. Kolokoltsov. Symmetric stable laws and stable-like jump diffusions. *Proc. London Math. Soc.*, 80 :725–768, 2000.

[KP91] T.G. Kurtz and P. Protter. *Wong-Zakai corrections, random evolutions and numerical schemes for SDEs.* Academic Press, 1991.

[KP92] P. Kloeden and E. Platen. *Numerical solution of stochastic differential equations.* Springer-Verlag Berlin Heidelberg New-York, 1992.

[KPQ97] N El Karoui, S. Peng, and M.C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7 :1–71, 1997.

[KPZ86] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.*, 56 :889–892, 1986.

[Kry96] N. V. Krylov. *Controlled diffusion processes.* Springer, 1996.

[KS84] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. I. *Stochastic analysis (Katata/Kyoto, 1982)*, North-Holland Math. Library, 32 :271–306, 1984.

[KS85] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math*, 32 :1–76, 1985.

[KS87] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math*, 34 :391–442, 1987.

[Kun90] H. Kunita. *Stochastic flows and stochastic differential equations.* Cambridge University Press, 1990.

[Kus01] S. Kusuoka. Approximation of expectation of diffusion process and mathematical finance. *Taniguchi Conference on Mathematics Nara '98. Adv. Stud. Pure Math., Math. Soc. Japan, Tokyo*, 31 :147–165, 2001.

[KV06] S. A. Klokov and A. Yu Veretennikov. On mixing and convergence rates for a family of markov processes approximating SDEs. *Random Oper. Stochastic Equations*, 14–2 :103–126, 2006.

[Led99] M. Ledoux. Concentration of measure and logarithmic sobolev inequalities. *Séminaire de Probabilités XXXIII. LNM, 1709*, page 120–216, 1999.

[Lem07] V. Lemaire. An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.*, 117–10 :1491–1518, 2007.

[Lie96] G.M. Lieberman. *Second Order parabolic differential equations, 1st edn.* World Scientfic, River Edge, NJ, 1996.

[LP02] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8–3 :367–405, 2002.

[LP03] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion : the case of a weakly mean reverting drift. *Stoch. Dyn.*, 3–4 :435–451, 2003.

[LS01] F. A. Longstaff and R. S. Schwartz. Valuing American options by simulation : a simple least-square approach. *Rev. Financ. Studies*, 14 :113–147, 2001.

[LSU68] O.A. Ladyzenskaja, V.A. Solonnikov, and N.N. Ural'ceva. *Linear and quasi-linear equations of parabolic type.* Vol.23 Trans. Math. Monog., AMS, Providence, 1968.

[LU68] O.A. Ladyzhenskaya and N.N. Ural'ceva. *Linear and quasilinear equations of elliptic type.* Academic Press, 1968.

[LV04] T. Lyons and N. Victoir. Cubature on wiener space, Stochastic analysis with applications to mathematical finance. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci*, 460 :169–198, 2004.

[Mal97] P. Malliavin. *Stochastic analysis.* Springer-Verlag, Berlin, 1997.

[Mél01] S. Méléard. Monte-Carlo approximations for 2d Navier-Stokes equations with measure initial data. *Prob. Th. Rel. Fields*, 121–3 :367–388, 2001.

[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.

[Mil98] G. N. Milstein. Weak approximation of a diffusion process in a bounded domain. *Stoch. Stoch. Reports*, 64 :211–233, 1998.

[MPY94] J. Ma, P. Protter, and J. Yong. Solving Forward-Backward Stochastic Differential Equations explicitly - a four step scheme. *Prob. Th. Rel. Fields*, 98 :339–359, 1994.

[MS67] H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1 :43–69, 1967.

[MSH02] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for sdes and approximations : locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101–2 :185–232, 2002.

[MT99] G. N. Milstein and M. V. Tretyakov. Simulation of a space-time bounded diffusion. *Ann. Appl. Prob.*, 9–3 :732–779, 1999.

[MT04] G.N. Milstein and M.V. Tretyakov. *Stochastic numerics for mathematical physics*. Springer-Verlag, Berlin, 2004.

[MT06a] F. Malrieu and D. Talay. Concentration inequalities for euler schemes. *Monte Carlo and Quasi-Monte Carlo Methods 2004, H. Niederreiter and D. Talay (Eds.)*, pages 355–37, 2006.

[MT06b] G. N. Milstein and M. V. Tretyakov. Numerical algorithms for forward-backward stochastic differential equations. *SIAM J. of Sci. Comput.*, 28 :561–582, 2006.

[MT07] G. N. Milstein and M. V. Tretyakov. Discretization of forward-backward stochastic differential equations and related quasi-linear parabolic equations. *SIAM J. of Sci. Comput.*, 1 :24–44, 2007.

[MY09] R. Mansuy and M. Yor. *Aspects of Brownian motion*. Springer, 2009.

[NN09] M. Ninomiya and S. Ninomiya. A new higher-order weak approximation scheme for stochastic differential equations and the runge-kutta method. *Finance and Stochastics*, 13–3 :415–443, 2009.

[Nor86] J. R. Norris. Simplified Malliavin Calculus. *Séminaire de Probabilités*, XX :101–130, 1986.

[Nua95] D. Nualart. *The Malliavin calculus and related topics. Probability and its Applications*. Springer-Verlag, New York, 1995.

[Nua98] D. Nualart. Analysis on Wiener space and anticipating stochastic calculus. In *Lectures on probability theory and statistics (Saint-Flour, 1995)*, pages 123–227. LNM 1690. Springer, Berlin, 1998.

[NV08] S. Ninomiya and N. Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance*, 15–1-2 :107–121, 2008.

[Pan08] F. Panloup. Computation of the invariant measure for a Lévy driven SDE : rate of convergence. *Stochastic Process. Appl.*, 118–8 :1351–1384, 2008.

[PP90] E. Pardoux and S.G. Peng. Adapted solution of a Backward Stochastic Differential Equation. *Systems Control Lett*, 14-1 :55–61, 1990.

[PP92] E. Pardoux and S.G. Peng. Backward Stochastic Differential Equations and quasilinear parabolic partial differential equations. *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, 176 :200–217, 1992.

[PT97] P. Protter and D. Talay. The Euler scheme for lévy driven stochastic differential equations. *Ann. Probab.*, 1(25) :393–423, 1997.

[RBT00] L. Rey-Bellet and L. Thomas. Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. *Comm. Math. Phys.*, 215–1 :1–24, 2000.

[Riv05] O. Rivière. *Equations Différentielles Stochastiques Rétrogrades couplées : EDPs et discrétisation*. Thèse de doctorat. Université Paris Descartes, 2005.

[Rub03] S. Rubenthaler. Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. *Stochastic Process. Appl.*, 103–2 :311–349, 2003.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion. 3rd ed.* Grundlehren der Mathematischen Wissenschaften. 293. Berlin : Springer, 1999.

[She91] S. J. Sheu. Some estimates of the transition density of a nondegenerate diffusion Markov process. *Ann. Probab.*, 19–2 :538–561, 1991.

[Sie79] D. Siegmund. Corrected diffusion approximations in certain random walk problems. *Adv. in Appl. Probab.*, 11(4) :701–719, 1979.

[SK74] B.W. Stuck and B. Z. Kleiner. A statistical analysis of telephone noise. *The Bell System Technical Journal*, 53 :1263–1320, 1974.

[Soi94] C. Soize. *The Fokker-Planck equation for stochastic dynamical systems and its explicit steady state solutions.* Series on Advances in Mathematics for Applied Sciences, 17. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[ST94] G. Samorodnittsky and M. Taqqu. *Stable non Gaussian random processes, stochastic models with infinite variance.* Chapman and Hall, New York, 1994.

[Str83] D.W. Stroock. Some applications of stochastic calculus to partial differential equations. *In : Eleventh Saint Flour probability summer school—1981 (Saint Flour, 1981), Lecture Notes in Math., 976, Springer, Berlin*, pages 267–382, 1983.

[Str88] D. W. Stroock. Diffusion semigroups corresponding to uniformly elliptic divergence form operators. *Séminaire de Probabilités*, XXII :316–347, 1988.

[SV71] D.W. Stroock and S.R.S. Varadhan. Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.*, 24 :147–225, 1971.

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes.* Springer-Verlag Berlin Heidelberg New-York, 1979.

[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. Shape sensitivity analysis*, volume 16. Springer Series in Computational Mathematics, 1992.

[Tal86] D. Talay. Discrétisation d'une eds et calcul approché d'espérances de fonctionnelles de la solution. *Math. Model ling Numer. Anal.*, 20 :141–179, 1986.

[Tal02] D. Talay. Stochastic Hamiltonian dissipative systems : exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8–2 :163–198, 2002.

[Tem01] E. Temam. Couverture Approchée d'Options Exotiques, Pricing des Options Asiatiques. *Ph. Dissertation. University of Paris VI*, 2001.

[TT90] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. and App.*, 8-4 :94–120, 1990.

[Ver03] A. Yu. Veretennikov. On large deviations for approximations of SDEs. *Probab. Theory Related Fields*, 125–1 :135–152, 2003.

[Web51] M. Weber. The fundamental solution of a degenerate partial differential equation of parabolic type. *Trans. Amer. Math. Soc.*, 71 :24–37, 1951.

[Whi73] N. Whitham. *Linear and Non linear waves.* Wiley, 1973.

[Woo82] M. Woodroofe. *Nonlinear renewal theory in sequential analysis.* SIAM, 1982.

[Zha88] C.H. Zhang. A nonlinear renewal theory. *Ann. Prob.*, 16-2 :793–824, 1988.

[Zha04] J. Zhang. A numerical scheme for BSDEs. *Annals of Applied Probability*, 14–1 :459–488, 2004.