Discretizations associated to a process in a domain and probabilistic numerical schemes for quasilinear parabolic PDEs

Stephane Menozzi

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS VI

Spécialité: **Mathématiques Appliquées**

présentée par

Stéphane MENOZZI

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PARIS VI

Sujet:

**Discrétisations associées à un processus dans un domaine
et
Schémas numériques probabilistes pour les EDP paraboliques quasi-linéaires.**

Soutenue le 15 décembre 2004 devant le jury composé de :

Vlad BALLY	Directeur de thèse
Nicole EL KAROUI	Examinateur
Emmanuel GOBET	Directeur de thèse
Arturo KOHATSU HIGA	Rapporteur
Gilles PAGÈS	Examinateur
Huyên PHAM	Examinateur
Denis TALAY	Rapporteur

Remerciements

Je tiens tout d'abord à exprimer ma sincère reconnaissance à Messieurs Vlad Bally et Emmanuel Gobet qui ont encadré cette thèse. Au cours de ces trois années, leur pédagogie, disponibilité et patience à mon égard furent en tout point exemplaires.

Au cours de cette thèse, j'ai eu l'occasion de rencontrer de nombreux chercheurs et de travailler avec certains d'entre eux.

Je souhaite en particulier remercier François Delarue pour une collaboration aussi instructive qu'agréable. Jacques Printems m'a donné de précieux conseils par rapport à certains aspects numériques développés dans cette thèse. Je remercie Marc Yor pour m'avoir à plusieurs reprises indiqué des références qui ont été capitales pour le développement de certaines parties de la thèse. Merci enfin à tous ceux qui m'ont consacré du temps.

Je suis très reconnaissant à Arturo Kohatsu Higa et Denis Talay d'avoir accepté de rapporter ce travail dans l'urgence. Je les remercie par ailleurs de leurs diverses suggestions.

Je suis très honoré que Nicole El Karoui ait accepté de faire partie du jury malgré un emploi du temps chargé. Je l'en remercie chaleureusement.

Enfin, les travaux de Gilles Pagès et Huyên Pham sur l'utilisation de méthodes de quantification en probabilités numériques ont eu une influence considérable sur la dernière partie de la thèse. Je suis donc heureux de les compter parmi les membres de ce jury.

En conclusion, je remercie ma famille, mes amis et mes camarades matheux pour leurs encouragements et leur soutien.

Contents

Introduction

Deux problématiques sont abordées dans cette thèse. La première, développée dans les trois premiers chapitres, concerne l'approximation à temps discret de la loi, à un instant $T > 0$ donné, d'un processus de diffusion multidimensionnel (ou plus généralement d'un processus d'Itô) tué/stoppé à la sortie d'un domaine. La deuxième, développée au Chapitre 4, est relative à l'étude d'un algorithme probabiliste pour les équations aux dérivées partielles (EDP en abrégé) quasi-linéaires. Le pendant probabiliste de cet algorithme est un procédé de discrétisation d'équations différentielles stochastiques progressives rétrogrades (EDSPR en abrégé).

1 Discrétisations associées à des processus dans un domaine

Dans cette partie, qui regroupe trois chapitres, nous nous sommes intéressés pour un instant $T > 0$ donné à l'approximation de quantités du type

$$\mathbb{E}[g(T \wedge \tau, X_{T\wedge\tau})] \tag{1}$$

où g est une fonction borélienne, $(X_t)_{t\geq 0}$ est un processus d'Itô d-dimensionnel et $\tau := \inf\{t \geq 0 : X_t \notin D\}$ où D est un domaine de $\mathbb{R}^d$.

Les espérances de fonctionnelles stoppées ou tuées de type (1) apparaissent dans de nombreux contextes. En mathématiques financières, elles peuvent représenter le prix d'options à barrière de pay-off g lorsque la dynamique de l'actif risqué sous-jacent est donnée par celle de X (cf. Karatzas et Shreve [KS98]). Dans un cadre Markovien et sous des hypothèses appropriées, l'expression (1) correspond à la représentation Feynman-Kac de la solution à l'instant initial et au point $x = X_0$ d'une EDP parabolique avec des conditions aux limites de type Cauchy-Dirichlet (cf. Freidlin [Fre85]).

Dans ces contextes, il paraît essentiel de pouvoir approcher (1) en ayant de bons contrôles des vitesses de convergence par rapport aux paramètres d'approximation.

Dans le premier chapitre de cette thèse, nous considérons le cas où X est un processus de diffusion, dont les coefficients vérifient une hypothèse de type hypoellipticité, lorsque le domaine D est suffisamment régulier. Pour approcher (1), nous discrétisons ensuite la diffusion par son schéma d'Euler tué à temps discret le long d'un maillage régulier en temps de pas $h = T/N$, $N \in \mathbb{N}^*$. Sous les hypothèses précisées ci-après, nous isolons d'abord le terme principal de l'erreur faible associée, i.e. celui qui reste à développer dans la perspective d'obtenir un équivalent. Jusqu'à présent, seule une majoration de l'erreur faible avait été prouvée, cf. Gobet [Gob00], et ce sous une hypothèse d'uniforme ellipticité du coefficient de diffusion. Nous obtenons ici un encadrement de l'erreur sous une hypothèse d'hypoellipticité restreinte exhibant ainsi la vitesse exacte de convergence de l'erreur (qui est $\sqrt{h}$) dans un cadre où le coefficient de diffusion peut dégénérer.

Dans le deuxième chapitre, nous étudions l'impact dans (1) de la discrétisation du temps de sortie τ le long d'un maillage en temps régulier de pas $h = T/N$, $N \in \mathbb{N}^*$, lorsque X est un processus d'Itô. Sous des hypothèses peu contraignantes nous retrouvons, pour un domaine D régulier (ou une intersection finie de domaines réguliers), la vitesse précédente de $\sqrt{h}$ pour l'erreur faible. Ceci met en évidence que l'ordre précédemment trouvé est en quelque sorte intrinsèque à l'approximation discrète du temps de sortie et n'est pas relié à l'utilisation du schéma d'Euler. Par ailleurs, dans ce cadre non Markovien, pour palier à l'absence d'EDP sous-jacente nous avons introduit des nouvelles techniques spécifiques au calcul stochastique pour développer et contrôler l'erreur faible.

Enfin, le troisième chapitre est relatif au cas particulier du mouvement Brownien lorsque le domaine D est une intersection de m demi-espaces ($m \in [\![1,d]\!]$). Notons que lorsque $m \geq 2$, le domaine D présente des singularités le long des intersections des demi-espaces. Dans ce contexte nous avons obtenu un développement de l'erreur faible associée à la discrétisation du temps de sortie. D'un point de vue numérique, les développements d'erreur sont utiles dans la mesure où ils légitiment l'emploi de techniques d'accélération de la convergence de type Romberg. Dans notre cas, nous nous appuyons sur le développement limité de l'erreur pour proposer une méthode d'accélération originale basée sur une correction appropriée du domaine qui permet de compenser le biais de discrétisation dans (1). Notre technique se révèle moins coûteuse que Romberg et les résultats numériques ainsi obtenus sont encourageants.

1.1 Vitesse de convergence pour des diffusions hypoelliptiques tuées approchées par le schéma d'Euler discret

Ce sujet est abordé dans le Chapitre 1 de la thèse. Soit $(X_t)_{t\geq 0}$ un processus de diffusion d-dimensionnel de dynamique

$$X_t := x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{2}$$

où $(W_t)_{t\geq 0}$ désigne un mouvement Brownien (MB en abrégé) standard de $\mathbb{R}^{d'}$.

Considérons un domaine D, i.e. un ouvert connexe, de $\mathbb{R}^d$ tel que $x \in D$ et définissons $\tau := \inf\{t \geq 0 : X_t \notin D\}$. Pour un temps déterministe $T > 0$ donné, nous nous sommes dans ce chapitre intéréssés à l'approximation de quantités du type

$$\mathbb{E}_x [f(X_T)\mathbb{I}_{\tau>T}] \tag{3}$$

où f est une fonction borélienne.

Pour estimer la quantité (3), la stratégie naturelle consiste à utiliser un schéma de discrétisation de la diffusion X, puis à approcher l'espérance de la fonctionnelle du processus discrétisé par une méthode de Monte-Carlo en précisant un procédé d'approximation du temps de sortie associé au schéma discrétisé.

Pour discrétiser la diffusion, notre choix s'est porté sur le schéma d'Euler. Pour une subdivision régulière de pas $h = T/N$, $N \in \mathbb{N}^*$ de l'intervalle $[0,T]$, notons pour tout $i \in [\![0,N]\!]$, $t_i = ih$. Le schéma d'Euler discret $(X^N_{t_i})_{i\in[\![0,N]\!]}$ associé à (2) sur l'intervalle $[0,T]$ est défini comme suit

$$\begin{aligned}
&X^N_0 = x,\\
&\forall i \in [\![1,N]\!],\ X^N_{t_i} = X^N_{t_{i-1}} + b(X^N_{t_{i-1}})h + \sigma(X^N_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}).
\end{aligned}$$

Indiquons que l'on peut prolonger ce processus constant par morceaux en un processus continu en posant sur $[t_{i-1}, t_i)$

$$X^N_t = X^N_{t_{i-1}} + b(X^N_{t_{i-1}})(t - t_{i-1}) + \sigma(X^N_{t_{i-1}})(W_t - W_{t_{i-1}}).$$

Au schéma d'Euler discret (resp. continu) on associe le temps de sortie discret $\tau^N := \inf\{t_i \geq 0 : X^N_{t_i} \notin D\}$ (resp. le temps de sortie continu $\tau^{N,c} := \inf\{t \geq 0 : X^N_t \notin D\}$).

Dans cette partie, nous avons restreint par choix l'étude au cas du schéma d'Euler tué à temps discret pour sa simplicité algorithmique. En effet, pour un nombre donné N_{MC} de trajectoires du schéma d'Euler, l'approximation par méthode de Monte-Carlo de $\mathbb{E}_x[f(X^N_T)\mathbb{I}_{\tau^N>T}]$ est donnée par la moyenne empirique

$$E_{N_{MC}}(f,N) = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} f(X^{N,j}_T)\mathbb{I}_{\tau^{N,j}>T}$$

où les $(X^j_{t_i})_{(i,j)\in[\![0,N]\!]\times[\![1,N_{MC}]\!]}$ désignent des réalisations indépendantes du schéma d'Euler. Rappelons que dans l'estimateur ci-dessus la simulation du schéma d'Euler ne requiert que la simulation de variables Gaussiennes standard quelle que soit la dimension d du processus de diffusion sous-jacent. Ensuite, en ce qui concerne les $\tau^{N,j}$, il suffit de vérifier à chaque instant de discrétisation si le schéma d'Euler est hors du domaine, auquel cas l'on tue la trajectoire et celle-ci n'apporte aucune contribution dans la moyenne empirique, ou bien s'il est dans le domaine et l'on poursuit alors la simulation jusqu'au temps final T ou jusqu'à ce que le schéma d'Euler sorte de D à un instant de discrétisation ultérieur.

L'erreur totale entre la quantité à estimer et ce que l'on simule dans la pratique se décompose de la façon suivante

$$\begin{aligned}\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] - E_{N_{MC}}(f,N) &= (\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] - \mathbb{E}_x[f(X_T^N)\mathbb{I}_{\tau^N>T}]) + (\mathbb{E}_x[f(X_T^N)\mathbb{I}_{\tau^N>T}] - E_{N_{MC}}(f,N)) \\ &:= \mathrm{Err}(T,h,f,x) + E_S.\end{aligned}$$

L'erreur $\mathrm{Err}(T,h,f,x)$ est associée à la discrétisation du processus et du temps de sortie. Elle est contrôlée par les relations (5) et (6) ci-dessous. Le terme E_S correspond à une erreur statistique qui, sous les hypothèses qui permettent d'établir (5), est avec grande probabilité contrôlée à l'ordre 1/2 par rapport à N_{MC}^{-1}, vitesse usuelle du théorème de limite centrale. Une vitesse de convergence presque sûre pour ce terme est donnée par la loi du logarithme itéré, i.e. $\frac{\log(\log(N_{MC}))}{\sqrt{N_{MC}}}$.

Indiquons par ailleurs que des techniques de simulations existent pour approcher le temps de sortie continu $\tau^{N,c}$ du schéma d'Euler. Il suffit par exemple sur l'événement $\{\tau^{N,c} > t_i\}$ et conditionnellement à la réalisation de $X_{t_i}^N, X_{t_{i+1}}^N$ où $X_{t_{i+1}}^N \in D$, de simuler une loi de Bernoulli de paramètre $p := \mathbb{P}[\exists s \in [t_i, t_{i+1}] : X_s \notin D | X_{t_i}^N, X_{t_{i+1}}^N]$. Si la réalisation de la simulation vaut 1, i.e. le processus est sorti entre les instants t_i et t_{i+1}, on annule la contribution, on continue sinon. Par définition du schéma d'Euler, p est liée à la probabilité de sortie du pont Brownien d'un domaine donné. Si l'expression de cette probabilité est explicite en dimension 1 pour un domaine de la forme $(-\infty, b)$, il n'en est pas de même en dimension supérieure et/ou pour des domaines plus généraux. A ce propos, mentionnons le travail de Baldi, [Bal95], qui a obtenu des asymptotiques de cette probabilité pour des domaines assez généraux lorsque le pas de temps $h \to 0$, à l'aide de techniques de grandes déviations. Ces procédés permettent d'améliorer la vitesse de convergence de $\mathrm{Err}(T,h,f,x)$, cf. Gobet [Gob00], ou dans un autre contexte Baldi [Bal95] ainsi que Baldi, Caramellino et Iovino [BCI99] pour des applications financières. Ils se révèlent néanmoins numériquement coûteux pour des domaines arbitraires en dimension d.

Principaux résultats

Dans un premier temps, nous montrons que lorsque les coefficients b, σ de (2) sont Lipschitziens et que σ vérifie une condition de frontière non caractéristique par rapport à ∂D de classe C^2 et bornée (hypothèse **(C)** du Chapitre 1), alors pour toute fonction continue f à croissance polynomiale

$$\boxed{\mathrm{Err}(T,h,f,x) \underset{h\longrightarrow 0}{\longrightarrow} 0.}$$

Mentionnons que cette hypothèse **(C)** de frontière non caractéristique est en quelque sorte minimale dans la mesure où il existe des exemples pour lesquels **(C)** n'est pas vérifiée et $\mathrm{Err}(T,h,f,x) \underset{h\longrightarrow 0}{\not\longrightarrow} 0$.

Dans la suite de cette section nous supposons $f \geq 0$. Si les coefficients b, σ sont réguliers et vérifient des propriétés de non dégénérescence de type hypoellipticité générale, **(H')**, alors, pour un domaine régulier à frontière bornée et une fonction f bornée s'annulant sur un voisinage de la frontière, nous pouvons isoler un terme dominant dans l'erreur $\mathrm{Err}(T,h,f,x)$. Précisément on obtient

$$\boxed{\mathrm{Err}(T,h,f,x) = \frac{1}{2}\mathbb{E}_x[\int_0^T \frac{\partial v}{\partial n}(s, X_s^N) dL^0_{s\wedge\tau^N}(F(X^N))] + o(h^{1/2})} \qquad (4)$$

où $v(t,x) := \mathbb{E}_x[f(X_{T-t})\mathbb{I}_{\tau>T-t}]$, F représente la distance signée à la frontière et $L^0(F(X^N))$ désigne le temps local du schéma d'Euler à la frontière.

Pour prouver ce résultat les outils utilisés proviennent essentiellement des travaux de Talay et Tubaro, [TT90], pour l'analyse de l'erreur à l'aide de l'EDP satisfaite par v. Rappelons néanmoins que dans notre cas v est solution d'un problème parabolique de type Cauchy-Dirichlet contrairement à la référence indiquée où seul le problème de Cauchy est considéré. Ceci est source de quelques difficultés techniques supplémentaires (i.e. développement en semi-martingale de la projection sur $\bar{D}$ du schéma d'Euler X^N).

L'utilisation du calcul de Malliavin conditionnel pour contrôler les dérivées de v lorsque t est proche de T sont présentes chez Cattiaux, [Cat91]. Toutefois, pour obtenir les estimées nécessaires à nos calculs, des contrôles trajectoriels assez fins doivent être en plus employés. Nous sommes notamment amenés à raffiner certaines estimées de Bally et Talay [BT96a].

Indiquons enfin que dans un cadre d'uniforme ellipticité l'on peut s'affranchir de l'hypothèse de support sur la fonction f lorsque celle-ci est suffisament régulière et vérifie des conditions de compatibilité sur la frontière. Il est également possible dans ce contexte d'affaiblir les hypothèses de régularité portant sur le domaine et les coefficients.

Le deuxième résultat de ce chapitre concerne la vitesse de convergence de $\mathrm{Err}(T,h,f,x)$ lorsque h tend vers 0. Sous les hypothèses précédentes de régularité et d'hypoellipticité générale sur les coefficients b, σ, de régularité sur le domaine et de support sur f, il vient

$$\boxed{\exists C > 0,\ h \le h_0,\ |\mathrm{Err}(T,h,f,x)| \le C\sqrt{h}.} \tag{5}$$

Ce résultat avait été obtenu par Gobet, [Gob00], dans un cadre d'uniforme ellipticité du coefficient de diffusion. Nous le généralisons à un cadre où le terme σ peut dégénérer, i.e. sous **(H')**.

En considérant ensuite une hypothèse d'hypoellipticité restreinte, cf. **(H)**, qui garantit la stricte positivité de la densité de la diffusion tuée, on a également

$$\boxed{\exists C' > 0,\ \forall h \le h_0(x),\ C'\sqrt{h} \le \mathrm{Err}(T,h,f,x).} \tag{6}$$

De ce dernier contrôle l'on déduit en particulier que $\mathrm{Err}(T,h,f,x) \ge 0$. Cette propriété apparaissait clairement pour le mouvement Brownien, pour lequel $X^N = X$, dans la mesure où $f \ge 0$ et $\tau^N \ge \tau$ p.s. En revanche, dans le cas général cela n'était pas a priori évident, bien que cela eût déjà été remarqué numériquement dans plusieurs cas de figures (voir Rubinstein et Reiner [RR91], Boyle et Lau [BL94] ou Baldi [Bal95]). De (6) nous déduisons donc que ce biais de surestimation est inhérent au choix de meurtre discret.

De façon générale, ces résultats mettent en évidence que la vitesse de convergence du schéma d'Euler discret est exactement d'ordre $1/2$. Dans le cadre Brownien, ceci avait été prouvé par Broadie, Glasserman et Kou [BGK97] qui ont obtenu un développement de l'erreur avec un terme de premier ordre en $\sqrt{h}$. Nous n'avons malheureusement obtenu dans le cas général qu'un encadrement, néanmoins les simulations numériques effectuées tendent à mettre en évidence l'existence d'un développement limité de l'erreur à l'ordre $1/2$. Rappelons que du point de vue numérique un développement de l'erreur permet d'utiliser des techniques d'amélioration de la vitesse de convergence de type extrapolation de Romberg, voir Talay et Tubaro [TT90] et Chapitre 3.

Un aspect essentiel pour l'analyse de l'erreur est que le terme principal mis en évidence dans (4) peut s'exprimer comme une moyenne de l'overshoot du schéma d'Euler discret (voir les preuves du Théorème 1.5.2 et du Lemme 1.4.5), où l'overshoot désigne la distance à la frontière du processus lorsqu'il quitte le domaine. Ceci donne une intuition de l'origine de l'erreur: les incréments du processus sont d'ordre $1/2$ et il en est de même pour l'overshoot. Les bornes (5) et (6) de l'erreur se déduisent ensuite des contrôles sur l'overshoot et respectivement de la bornitude et stricte positivité de la dérivée normale $\dfrac{\partial v}{\partial n}$.

Indiquons que dans une perspective de développement de l'erreur, il faudrait obtenir la loi asymptotique de l'overshoot. Ce type de problèmes est généralement analysé à l'aide de techniques de renouvellement pour les chaînes de Markov. Cependant, dans un contexte multidimensionnel général, les hypothèses d'ergodicité requises, voir Alsmeyer [Als94], Fuh et Lai [FL01], ne sont jamais satisfaites sur le schéma d'Euler changé d'échelle. Mentionnons que dans le cas Brownien, i.e. pour $X_t = x + \mu t + \sigma W_t$, il est possible d'obtenir la loi asymptotique de l'overshoot lorsque D est une intersection de demi-espaces, voir Section 1.3 et Chapitre 3. De cette loi asymptotique l'on déduira sous certaines hypothèses de régularité sur v un développement de l'erreur.

1.2 Echantillonage discret de fonctionnelles de processus d'Itô

La vitesse obtenue précédemment en (5) et (6) soulève la question suivante. Le terme dominant de l'ordre $1/2$ en h provient-t-il de la discrétisation d'Euler ou bien de l'approximation discrète du temps de sortie? Nous répondons à cette question au Chapitre 2. Dans le contexte des processus de diffusion uniformément elliptiques, et pour un domaine régulier, Gobet, cf. [Gob00], avait montré que l'ordre $1/2$ était intrinsèque au temps de sortie discret. Dans le Chapitre 2, nous étendons ce résultat au cas de processus d'Itô, donc a priori non Markoviens, dont le coefficient de diffusion peut en un certain sens dégénérer. Dans la mesure où l'on n'a plus forcément comme au chapitre précédent d'EDP sous jacente, d'autres types d'outils doivent être introduits pour analyser l'erreur.

En fait on a recours à des techniques de type martingales qui se révèlent particulièrement efficaces dans la mesure où l'on obtient avec des arguments assez standards de calcul stochastique (changements de temps de Dambis-Dubins-Schwarz, inégalités BDG,...) le contrôle à l'ordre $1/2$ pour une classe de fonctions plus grande qu'au chapitre précédent. Indiquons notamment que si la fonction est suffisamment régulière, il n'est pas nécessaire d'imposer qu'elle s'annule à la frontière pour obtenir la majoration attendue. Ce type d'outils permet également d'étendre aisément ce résultat à une intersection de domaines réguliers.

Principaux résultats

Soit $(X_t)_{t\geq 0}$ un processus d'Itô d-dimensionnel de dynamique

$$X_t = x + \int_0^t b_s ds + \int_0^t \sigma_s dW_s$$

où comme précédemment W désigne un MB standard de $\mathbb{R}^{d'}$. Par la suite on supposera que les coefficients b, σ sont bornés. Définissons de nouveau $\tau := \inf\{t \geq 0 : X_t \notin D\}$. Pour un domaine D donné et un temps final fixe $T > 0$, nous nous sommes intéréssés à l'impact de la discrétisation de τ dans des quantités du type

$$\mathbb{E}[g(T \wedge \tau, X_{T\wedge\tau})].$$

Précisément, pour un maillage régulier de pas $h = T/N,\ N \in \mathbb{N}^*$ de l'intervalle $[0, T]$, si $t_i = ih,\ i \in [\![0, N]\!]$ et $\tau^N := \inf\{t_i \geq 0, X_{t_i} \notin D\}$, on souhaite contrôler la convergence vers 0 en fonction de h de la quantité

$$\mathrm{Err}(T, h, g, x) = \mathbb{E}[g(T \wedge \tau^N \wedge \tau_R, \pi_{\bar{D}}(X_{T\wedge\tau^N\wedge\tau_R}))] - \mathbb{E}[g(T \wedge \tau, X_{T\wedge\tau})].$$

Ci-dessus $\tau_R := \inf\{t \geq 0 : X_t \notin D(R)\}$, où $D(R) := \{x \in \mathbb{R}^d : d(x, D) \leq R\}$ est un domaine pour lequel la projection $\pi_{\bar{D}}$ sur $\bar{D}$ est définie de façon unique.

Pour un domaine D de classe C^2 à frontière compacte, si le coefficient de diffusion $(\sigma_t)_{t\geq 0}$ satisfait p.s. une condition de frontière non caractéristique (hypothèse **(C)** du Chapitre 2) et si la fonction g est borélienne bornée et vérifie une condition de support, cf. **(G1)**, ou bien si $g \in C_b^{1,2}([0, T] \times \bar{D})$, cf. **(G2)**, i.e. g est bornée et continûment différentiable (1 fois en temps, 2 fois en espace) et que toutes ses dérivées partielles sont bornées, alors

$$\boxed{\mathrm{Err}(T, h, g, x) = O(\sqrt{h}).}$$

Rappelons que le résultat ci-dessus est relatif à l'impact de la discrétisation du temps de sortie *seulement* et non du processus X. Il ne peut donc être directement comparé avec les résultats énoncés à la Section 1.1, sauf dans le cas particulier du mouvement Brownien. Dans ce contexte, indiquons que l'on obtient beaucoup plus simplement la borne supérieure de l'équation (5).

Nous montrons ensuite que sous **(G1)** le résultat précédent reste valable pour une intersection de domaines de classe C^2 et à frontière compacte. Ce dernier point est intéressant même dans le cadre Markovien du mouvement Brownien. En effet pour des domaines non réguliers, il n'est pas facile d'utiliser les techniques usuelles d'analyse de l'erreur qui requièrent la régularité de l'EDP satisfaite par v introduite dans la section précédente jusqu'à la frontière, cf. Chapitre 3 à titre illustratif. Dans le cadre du Chapitre 3, nous souhaitons obtenir un développement de l'erreur faible lorsque D est un orthant et n'avons pas réussi à ce propos à nous passer de l'EDP sous-jacente. En revanche l'extension du résultat principal du Chapitre 2 à une intersection de domaines met en relief le fait que dans un contexte Brownien, l'erreur faible se majore simplement et que la principale difficulté ne réside pas dans l'absence de régularité du domaine.

Dans ce chapitre nous avons par ailleurs également étudié l'impact d'une discrétisation en temps pour d'autres fonctionnelles du processus d'Itô X. Des arguments simples de type Fubini permettent de montrer que pour une fonction ψ Lipschitzienne, si l'on définit pour $t \in [0,T]$, $\phi(t) := \inf\{t_i : t_i \leq t < t_{i+1}\}$ alors

$$\boxed{\psi\left(\int_0^T X_t dt\right) - \psi\left(\int_0^T X_{\phi(t)} dt\right) \underset{\mathrm{L}_p(\mathbb{P})}{=} O(h),\ p \geq 1.}$$

Enfin, dans le cadre unidimensionnel, i.e. $d = d' = 1$, nous nous sommes intéréssés à l'approximation discrète du maximum d'un processus d'Itô. Notons $M_T := \sup_{t\in[0,T]} X_t,\ M_T^N := \sup_{t\in[0,T]} X_{\phi(t)}$. Dans le cas particulier où $\sigma_s = \sigma(X_s)$ et $\sigma \geq \sigma_0 > 0$, i.e. le coefficient de diffusion est Markovien et uniformément elliptique, alors, en utilisant les techniques de martingales préalablement introduites pour les processus stoppés/tués, on arrive à montrer

$$\boxed{\exists C > 0,\ \mathbb{E}[M_T - M_T^N] \leq C\sqrt{h}.} \tag{7}$$

Dans le cadre Brownien, ce résultat est une conséquence d'un travail de Asmussen, Glynn et Pitman [AGP95], qui avaient obtenu, à l'aide de décompositions trajectorielles en ponts de Bessel du Brownien autour de son maximum, la loi limite de $\sqrt{N}(M_T - M_T^N)$. Nous retrouvons la majoration précédente avec des techniques élémentaires. L'hypothèse de coefficient de diffusion Markovien et uniformément elliptique permet ensuite de se ramener au cas Brownien à l'aide de la transformation de Rogers [Rog85]. Dans des travaux futurs nous chercherons à prouver (7) sans l'hypothèse de Markovianité sur σ.

1.3 Développement d'erreur et correction dans un orthant: le cas Brownien

Le dernier aspect de cette première partie est relatif à l'étude de l'erreur faible dans le cadre de domaines non réguliers lorsque X est un mouvement Brownien d-dimensionnel de dynamique $X_t = x + \mu t + \sigma W_t$ où $\sigma\sigma^*$ est définie positive. Nous nous sommes restreints à des cônes de la forme $D = \cap_{j=1}^m D^j,\ m \in [\![1, d]\!]$ où les $(D^j)_{j\in[\![1,m]\!]}$ sont des demi-espaces d'intersection non vide.

Développement d'erreur

Pour une fonction f s'annulant sur ∂D, si l'on a des bonnes propriétés jusqu'à la frontière de la fonction $v(t,x) = \mathbb{E}_x[f(X_{T-t})\mathbb{I}_{\tau > T-t}]$ alors on peut montrer que pour h assez petit et avec les notations des deux paragraphes précédents

$$\boxed{\mathrm{Err}(T,h,f,x)=\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau^N>T}]-\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]=C\sqrt{h}+o(\sqrt{h})} \tag{8}$$

où C est une constante dépendant de D, f, x. Comme indiqué dans (4), le terme dominant reste celui associé au temps local du processus tué à temps discret à la frontière, en l'occurrence ici le mouvement Brownien. Précisément, si $D=\{y\in\mathbb{R}^d : y_i > b_0^i,\ i\in[\![1,m]\!]\}$ et la fonction $v\in C^{1,2}([0,T]\times\bar{D})$ on obtient

$$\mathrm{Err}(T,h,f,x)=\frac{1}{2}\sum_{i=1}^{m}\mathbb{E}_x[\int_0^{T\wedge\tau^N}\partial_{x_i}v(s,\pi_{\bar{D}}(X_s))dL_s^{b_0^i}(X^i)]. \tag{9}$$

De l'expression précédente et de la propriété $v(s,.)|_{\partial D}=0$, on déduit qu'il n'y a pas dans l'erreur de contributions associées aux parties singulières du domaine. En effet, si pour $x\in\bar{D}$ il existe $j\in[\![1,d]\!]$, $x_j\in\partial D$, alors pour $i\neq j$, $\partial_{x_i}v(s,x)=0$.

Comme indiqué à la section 1.1, la partie dominante du temps local est en fait l'overshoot du processus au dessus de la frontière, où l'on rappelle que l'overshoot est la distance du processus à la frontière lorsque ce dernier quitte le domaine. Dans le cas particulier du mouvement Brownien et pour une intersection de demi-espaces, il est possible d'obtenir la loi limite de l'overshoot et donc un équivalent du temps local. A ce propos nous étendons les travaux de Siegmund [Sie79] qui avait identifié dans un cadre unidimensionnel la loi limite de l'overshoot et prouvé son indépendance asymptotique avec le temps de sortie discret du domaine.

On déduit ensuite le développement de l'erreur (8) à partir de (9) à l'aide d'intégrations par parties. C'est pour justifier cette dernière étape que de fortes hypothèses de régularité sont imposées sur v (dans la pratique on la supposera $C_b^{2,4}([0,T]\times\bar{D})$). Obtenir une telle régularité est un point délicat dans la mesure où D n'est pas régulier lorsque $m\geq 2$.

Concernant la régularité jusqu'au bord de la solution d'une EDP parabolique avec conditions aux limites de type Cauchy-Dirichlet pour une intersection de domaines réguliers, des résultats ont été publiés. Nous mentionnons à cet égard les travaux de Azzam [Azz85] et Azzam et Kreyszig [AK80], [AK81]. Néanmoins, à notre sens, les preuves sont en partie incorrectes et nous n'avons pas utilisé les résultats énoncés dans ces articles. Nous avons réussi à obtenir la régularité requise pour v dans le cas de la dimension 2 à l'aide de l'expression explicite de la densité du mouvement Brownien tué hors d'un cône, cf. Carslaw et Jaeger [CJ59] et Iyengar [Iye85]. Cette dernière s'obtient par extension de la méthode des images qui, lorsque l'angle α du cône est de la forme $\pi/m_0,\ m_0\in\mathbb{N}^*$, permet d'exprimer cette densité comme somme finie de noyaux gaussiens. La principale difficulté que nous avons rencontrée dans le cas d'un angle α général est relative au contrôle des dérivées de la densité qui s'exprime comme série de fonctions de Bessel modifiées. En dimension supérieure à 2, indiquons que la méthode des images reste valable pour des angles de la forme $(\alpha_i=\pi/m_0^i,\ m_0^i\in\mathbb{N}^*)_{i\in[\![1,m]\!]}$, $m\in[\![1,d]\!]$ et que dans le cas général, Bañuelos et Smits [BS97] ont obtenu l'expression de la densité de transition du MB tué hors d'un orthant. Cette dernière n'est toutefois pas aisée à manipuler en ce qui concerne le contrôle de ses dérivées et nous réservons ces calculs pour des recherches ultérieures.

Correction de domaine et accélération de la convergence

Rappelons que d'un point de vue numérique, le fait d'avoir un développement de la forme (8) permet d'appliquer des techniques d'accéleration de la convergence de type Romberg, cf. Talay et Tubaro [TT90]. Rappelons brièvement ce principe. Pour h assez petit on déduit de (8) que

$$\frac{1}{\sqrt{2}-1}\mathbb{E}[f(X_T)(\sqrt{2}\mathbb{I}_{\tau^{2N}>T}-\mathbb{I}_{\tau^N>T})]-\mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}]=o(\sqrt{h}).$$

Ceci suggère de remplacer l'estimateur de Monte-Carlo standard $E_{N_{MC}}(f,N)$ de $\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]$ introduit précédemment par

$$E_{N_{MC}}^{\mathrm{Romberg}}(f,N):=\frac{1}{\sqrt{2}-1}\left(\sqrt{2}(\sum_{i=1}^{N_{MC}}f(X_T^i)\mathbb{I}_{\tau^{2N,i}>T})-(\sum_{i=1}^{N_{MC}}f(X_T^i)\mathbb{I}_{\tau^{N,i}>T})\right).$$

On se débarasse ainsi, à l'erreur statistique près, du terme dominant de l'erreur. Notons malgré tout que cette méthode nécéssite le raffinement du pas de temps et engendre donc un coût supplémentaire de complexité.

Nous proposons une méthode alternative de simulation qui va permettre, de façon similaire à l'extrapolation de Romberg, d'annuler le terme principal de l'erreur tout en gardant la même complexité algorithmique que l'estimateur de Monte-Carlo usuel $E_{N_{MC}}(f,N)$. L'idée intuitive de cette correction consiste à tuer le mouvement Brownien non plus lorsqu'il quitte D à un instant de discrétisation mais lorsque qu'il quitte un domaine plus restreint D_h à préciser. En effet, nous avons indiqué à la section 1.1 que pour $f \geq 0$ le meurtre à temps discret entraîne une surestimation de la quantité $\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]$. Une idée naturelle est donc de chercher à contraindre le domaine de sorte à compenser cette surestimation. Indiquons qu'une démarche de même nature est au centre d'un travail de Broadie, Glasserman et Kou [BGK99] qui dans un cas unidimensionnel ont établi

$$\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau^N>T}] = \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{\tilde{D}_h}>T}]+o(\sqrt{h}), \tau_{\tilde{D}_h} := \inf\{s \geq 0 : X_s \notin \tilde{D}_h\},\ D \subset \tilde{D}_h := \{y \in \mathbb{R} : y-\tilde{C}\sqrt{h} \in D\}$$

pour $D = (-\infty, b),\ f(x) = e^{-rT}(\exp(x) - K)^+,\ K \in \mathbb{R}^+$. Leur propos était d'approcher l'espérance avec temps de sortie discret par l'espérance avec un temps de sortie continu d'un domaine modifié pour laquelle l'on possède, dans ce cas particulier, une expression explicite, i.e. il suffit d'exploiter la densité du mouvement Brownien unidimensionnel tué.

Cette relation indique également que pour compenser la surestimation due au meurtre discret, il est nécessaire de modifier le domaine à l'ordre $1/2$ par rapport à h.

Pour établir notre correction dans le cadre multidimensionnel d'une intersection de demi-espaces, nous utilisons, et étendons dans une moindre mesure au cas de certains domaines non réguliers, les travaux de Costantini, El Karoui et Gobet, [CKG03], relatifs à la sensibilité du problème de Dirichlet par rapport au domaine. Soit $D_h := \cap_{i=1}^m D_h^i,\ D_h^i := \{x \in \mathbb{R}^d : x - C_i\sqrt{h}n_i \in D^i\}$ où n_i désigne la normale rentrante au domaine D^i, et C_i est une constante positive dépendant de D_i, du coefficient σ et d'une constante universelle provenant des techniques de renouvellement employées pour obtenir l'équivalent du temps local. Définissons $\tau^N_{D^h} := \inf\{s_i \geq 0 : X_{s_i} \notin D_h\}$. On montre alors que

$$\boxed{\mathrm{Err}'(T,h,f,x) := \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau^N_{D^h}>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] = o(\sqrt{h}).}$$

L'estimateur de Monte-Carlo associé s'écrit

$$E_{N_{MC}}^{\text{Correction Domaine}} := \frac{1}{N_{MC}}\sum_{j=1}^{N_{MC}} f(X_T^{N,j})\mathbb{I}_{\tau^{N,j}_{D_h}>T}.$$

Du point de vue numérique, nous avons comparé les résultats fournis par les estimateurs $E_{N_{MC}}^{\text{Correction Domaine}}$ et $E_{N_{MC}}^{\text{Romberg}}$ sur plusieurs exemples (demi-espace et cône en dimension 2 pour différentes fonctions). En termes d'erreurs relatives par rapport à une quantité de référence donnée pour $\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]$, $E_{N_{MC}}^{\text{Correction Domaine}}$ s'est toujours révélé plus précis, à coût moindre, que $E_{N_{MC}}^{\text{Romberg}}$. De plus sa variance empirique était également toujours inférieure.

1.4 Applications des résultats précédents en mathématiques financières

Dans ce cadre, l'incertain est modélisé par un espace de probabilité filtré $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ où $\mathbb{P}$ désigne la probabilité historique et $\mathcal{F}_t$ représente l'information disponible à l'instant t. Nous considérons un ensemble de d actifs sous jacents $(S^1, ..., S^d)$ dont la dynamique sous $\mathbb{P}$ est donnée par

$$\forall i \in [\![1,d]\!],\ \frac{dS_t^i}{S_t^i} = B_t^i dt + \sum_{j=1}^d \Sigma_t^{ij} dW_t^j$$

où W est un $\mathbb{P}$-mouvement Brownien d-dimensionnel. Nous supposons que Σ_t est p.s. inversible, ce qui assure, au prix de quelques hypothèses d'intégrabilité, l'existence d'une unique probabilité équivalente $\mathbb{Q}$ (probabilité neutre au risque) sous laquelle les actifs actualisés sont des martingales.

La théorie moderne des options [KS98] montre qu'en l'absence d'opportunités d'arbitrage et dans un marché sans frictions le prix d'une option de maturité T et de flux ψ est donné par

$$\mathbb{E}^{\mathbb{Q}}[e^{-rT}\psi((S_t)_{t\in[0,T]})]$$

si le taux d'intérêt r est constant, et

$$B(0,T)\mathbb{E}^{\mathbb{Q}^T}[\psi((S_t)_{t\in[0,T]})]$$

si le taux est stochastique, où $B(0,T)$ désigne alors le zéro coupon de maturité T et $\mathbb{Q}^T$ la probabilité forward neutre.

Une option barrière de pay-off f, échéance $T>0$ et domaine D de type "*out*" sur S donne le flux $f(S_T)$ à son détenteur à l'horizon T si S est resté dans D sur l'intervalle $[0,T]$ et se trouve désactivée sinon. Avec les notations précédentes ceci correspond au choix de $\psi((S_t)_{t\in[0,T]}) = f(S_T)\mathbb{I}_{\tau>T}$, où $\tau := \inf\{t\geq 0 : S_t \notin D\}$. L'option barrière peut également être avec "*rebate*" c'est à dire que le flux est $R_t \neq 0$ lorsque la barrière est touchée à la date $t\in[0,T]$. Le pay-off prend alors la forme plus générale $\psi((S_t)_{t\in[0,T]}) = g(T\wedge\tau, S_{T\wedge\tau})$.

En conclusion, les étapes précédentes montrent qu'après le changement de variables $X_t^i = \log(S_t^i)$ (la dynamique de X^i est alors celle des rendements du $i^{\text{ème}}$ actif risqué), la valorisation des options barrière se ramène au calcul de

$$\mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}] \text{ ou } \mathbb{E}[g(T\wedge\tau, X_{T\wedge\tau})]$$

avec de nouvelles fonctions f, g et domaines D, sous une certaine probabilité ($\mathbb{Q}$ ou $\mathbb{Q}^T$).

Dans les paragraphes suivants nous donnons une interprétation de nos résultats en termes d'options barrière.

Chapitre 1

On a considéré dans ce chapitre le cas où X a une dynamique Markovienne de la forme

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t.$$

Les résultats donnés en (5) et (6) sont donc des indicateurs de la vitesse de convergence de l'espérance de la fonctionnelle du schéma d'Euler de X tué à temps discret vers le prix de l'option lorsque celle-ci est de type "*out*". Ces contrôles sont essentiels pour la simulation. A titre indicatif, la vitesse a priori donnée par (6) va permettre de calibrer le nombre N_{MC} de trajectoires dans $E_{N_{MC}}(f,N)$ pour atteindre en un certain sens une précision donnée. Indiquons également que l'hypothèse d'hypoelliticité permet à σ de dégénérer, ce qui autorise le cas des options Asiatiques à barrière de pay-off $f(S_T, \int_0^T S_t dt)\mathbb{I}_{\sup_{t\in[0,T]} S_t<U}$ où U est un seuil donné.

Chapitre 2

Dans ce chapitre, nous avons quantifié la différence entre les prix d'options lorsque l'on observe la position de l'actif par rapport à la barrière à temps continu et à temps discret. Pour une dynamique assez générale, i.e. X est ici un processus d'Itô, nous avons montré que cette différence est d'ordre $1/2$ par rapport à la période h d'observation dans le cas d'options de type "*out*" ou "*rebate*" lorsque le pay-off est assez régulier.

Le résultat concernant le maximum permet quant à lui d'établir le même type de contrôle pour les options "*look-back*", i.e. lorsque l'on a $\psi((S_t)_{t\in[0,T]}) = f(S_T, \max_{t\in[0,T]} S_t)$. En effet, si f est Lipschitz continue, et que l'on s'attache à la différence des prix entre les options faisant intervenir un maximum continu (resp. discret), on peut déduire du contrôle mentionné

$$|\mathbb{E}[f(S_T, \max_{t\in[0,T]} S_t)] - \mathbb{E}[f(S_T, \max_{t\in\{(ih)_{i\in[\![0,N]\!]}\}} S_t)]| \leq C\sqrt{h}.$$

Nous nous sommes ici restreints au cas où

$$X_t = x + \mu t + \sigma W_t$$

qui correspond au modèle de Black-Merton-Scholes (i.e. S est un MB géométrique standard) et au cas de l'option "*out*".

Dans la pratique il est assez naturel de choisir D (associé à l'actif S) sous la forme d'un domaine produit $\Pi_{i=1}^{m}(0, b_i)$, $m \in [\![1, d]\!]$. Ceci signifie que pour que l'option "*out*" reste active aucune des m premières composantes de l'actif S ne doit dépasser un seuil donné sur $[0, T]$. Dans ce contexte et pour une certaine classe de fonctions f, nos résultats précédents permettent donc d'avoir le développement de la différence des prix entre option à barrière désactivée à temps continu et à temps discret. On pourrait également retrouver, à l'aide des techniques précédentes de correction de domaine, un résultat similaire à celui de Broadie et *al.*, [BGK97], dans notre cadre multidimensionnel où, de plus, le domaine présente des singularités. Précisément, nous pouvons indiquer comment compenser, en modifiant le domaine de l'option, la surestimation associée à l'observation discrète.

2 Algorithme probabiliste pour les EDP paraboliques quasi-linéaires

Les liens entre processus de diffusion et EDP linéaires sont classiques. La formule de Feynman-Kac établit une correspondance claire entre un problème aux limites linéaire et l'espérance d'une fonctionnelle de processus de diffusion.

D'un point de vue algorithmique, cette représentation offre l'intérêt de pouvoir estimer la solution du problème de départ en simulant des trajectoires d'un schéma de discrétisation (schéma d'Euler, Milstein ou autre) associé à la diffusion sous-jacente. Cette approche est particulièrement adaptée à la simulation en grandes dimensions de par sa simplicité algorithmique, qui permet d'éviter l'inversion de grands systèmes linéaires associés aux méthodes analytiques de différences ou éléments finis.

Au début des années 90, Pardoux et Peng [PP90] ont introduit avec les équations différentielles stochastiques rétrogrades (EDSR en abrégé) le bon cadre probabiliste pour donner une représentation de type Feynman-Kac de certaines EDP non linéaires, dites semi-linéaires. Au Chapitre 4, nous nous intéressons à une classe plus générale d'EDP, dites quasi-linéaires, ayant précisément pour un $T > 0$ donné, sur $[0, T[\times \mathbb{R}^d$ la forme :

$$(\mathcal{E}) \begin{cases} \partial_t u + \langle b(x, u(t,x), \nabla_x u(t,x)\sigma(x, u(t,x))), \nabla_x u(t,x)\rangle + \dfrac{1}{2}\mathrm{tr}(\sigma\sigma^*(x, u(t,x))\nabla^2_{x,x} u(t,x)) \\ + f(x, u(t,x), \nabla_x u(t,x)\sigma(x, u(t,x))) = 0, \\ u(T,x) = H(x), \end{cases}$$

où u est à valeurs dans $\mathbb{R}$.

Sous des hypothèses garantissant une régularité suffisante de la solution u de $(\mathcal{E})$, d'un point de vue probabiliste, l'outil permettant de donner une interprétation de type Feynman-Kac pour u est fourni par la théorie des équations différentielles stochastiques progressives rétrogrades (EDSPR en abrégé).

Sur un espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ et pour un $\mathbb{P}$-mouvement Brownien d-dimensionnel B, résoudre une EDSPR c'est trouver un triplet (U, V, W), $(\mathcal{F}_t)_{t\in[0,T]}$ progressivement mesurable solution du couple d'équations

$$(E) \begin{cases} \forall t \in [0,T], \\ U_t = x_0 + \displaystyle\int_0^t b(U_s, V_s, W_s)ds + \int_0^t \sigma(U_s, V_s)dB_s, \\ V_t = H(U_T) + \displaystyle\int_t^T f(U_s, V_s, W_s)ds - \int_t^T W_s dB_s \end{cases}$$

et satisfaisant la condition d'intégrabilité

$$\mathbb{E}[\int_0^T (|U_s|^2 + |V_s|^2 + |W_s|^2)ds] < +\infty.$$

Ces équations ont été introduites par Antonelli [Ant93] qui a prouvé des résultats d'existence et unicité en temps court. Ensuite, Ma, Protter et Yong [MPY94] ont mis en évidence une méthode constructive de tels objets à partir de la régularité de la solution de $(\mathcal{E})$, c'est le *Four step scheme*. Néanmoins cette technique requiert une régularité assez forte des coefficients, qui est par ailleurs assez inhabituelle dans le cadre des EDS et des EDSR. Enfin, d'un point de vue plus général, Delarue [Del02] a montré l'existence et l'unicité sous des hypothèses usuelles de Lipschitz continuité pour les coefficients de (E) et d'uniforme ellipticité de σ. En termes d'applications, ces équations interviennent dans de nombreux domaines tels que les mathématiques financières (problème du gros investisseur) ou la formulation Hamiltonienne de problèmes de contrôles. Nous renvoyons au livre de Ma et Yong [MY99] pour un aperçu des différents champs d'applications.

Sous de bonnes hypothèses la relation de type Feynman-Kac entre $(\mathcal{E})$ et (E) se résume de la façon suivante

$$\forall t \in [0,T],\ V_t = u(t,U_t),\ W_t = \nabla_x u(t,U_t)\sigma(U_t,V_t),\ V_t = \mathbb{E}[V_T|\mathcal{F}_t] + \mathbb{E}\Big[\int_t^T f(U_s,V_s,W_s)ds|\mathcal{F}_t\Big].$$

Au Chapitre 4, nous déduisons de la théorie des EDSPR un algorithme complètement implémentable pour approcher la solution u de l'équation $(\mathcal{E})$. La contrepartie probabiliste de cet algorithme est de fournir une procédure de discrétisation du triplet (U, V, W).

La plupart des méthodes numériques proposées jusqu'à présent pour discrétiser des EDSPR reposent sur des techniques analytiques de type différences finies pour approcher la solution u de $(\mathcal{E})$. Cette approximation est ensuite utilisée pour obtenir une discrétisation d'Euler de l'équation progressive. Le schéma proposé par Douglas, Ma et Protter [DMP96] s'inscrit dans ce cadre.

A l'opposé, nous proposons de déduire de la représentation en terme d'EDSPR de la solution u de $(\mathcal{E})$ un schéma numérique pour aprocher u. Notre stratégie s'inspire du schéma de discrétisation temps-espace initialement introduit par Chevance [Che97] dans le cas découplé. Dans ce cadre, les coefficients b et σ ne dépendent pas de V et W de sorte que l'équation progressive se réduit à une EDS "classique". Le processus U apparaît donc comme une "diffusion objective". Indiquons que dans ce cas particulier, l'idée de discrétisation temps-espace permet d'utiliser une approche usuelle de "principe de programmation dynamique".

Du point de vue numérique, deux autres approches ont été développées dans le cas rétrograde. La première est basée sur des méthodes de Monte-Carlo, couplées à des intégrations par parties de type Malliavin chez Bouchard et Touzi [BT04] ou à des résolutions de problèmes de moindres carrés chez Gobet, Lemor et Warin [GLW04]. La deuxième repose sur des techniques de quantification d'une version discrétisée de l'équation progressive. La quantification consiste à approcher une variable aléatoire par une loi discrète appropriée. Cette technique fournit une alternative numériquement peu coûteuse et efficace aux méthodes de Monte-Carlo pour l'estimation d'espérances. Dans les travaux de Bally, Pagès [BP03] ou Bally, Pagès et Printems [BPP02] relatifs à l'estimation du prix et de la couverture d'options américaines, l'idée principale consiste à quantifier de façon "optimale" une version discrétisée du processus de diffusion, de sorte à approcher une fois pour toutes par une méthode de Monte-Carlo le semi-groupe associé. On effectue ensuite une descente de programmation dynamique standard. Pour d'autres applications de la quantification, nous renvoyons aux travaux de Pagès, Pham et Printems, [PPP04] ou Pagès et Printems, [PP04].

Stratégie de discrétisation

Dans le cas couplé, soit dans le cadre quasi-linéaire, la diffusion U n'est plus "objective". En effet, du fait de la non-linéarité de l'équation $(\mathcal{E})$, les coefficients de l'équation progressive sous-jacente dépendent de la solution ainsi que de son gradient.

En particulier, il n'est plus possible de quantifier un schéma de discrétisation du processus de diffusion comme indiqué ci-dessus dans la mesure où nous ne pouvons pas obtenir "a priori", i.e. sans approcher u, de grille optimale pour ce dernier. Nous quantifions donc les incréments Browniens qui apparaissent dans la discrétisation de l'EDS progressive et nous choisissons de définir la diffusion approchée sur une suite de grilles Cartésiennes tronquées d-dimensionnelles. Indiquons que la procédure de discrétisation de U est désormais fortement couplée à l'approximation de $(u, \nabla_x u)$ laquelle est également calculée le long de la suite de grilles Cartésiennes tronquées d-dimensionnelles. Notons-la de façon générique $(\overline{u}, \overline{v})$.

On peut définir $(\overline{u}, \overline{v})$ ainsi que les approximations des transitions de U de façon à retrouver une sorte de "principe de programmation dynamique". En effet, considérons un maillage régulier en temps $(t_i = ih)_{i\in\{0,\ldots,N\}}$ de $[0,T]$, où h est le pas de discrétisation. A chaque instant de discrétisation t_i associons une grille Cartésienne spatiale $\mathcal{C}_i \equiv \{(x_k^i)_{k\in\{1,\ldots,N_i\}}\}$ telle que $\forall i \in \{0,\ldots,N-1\}$, $\mathcal{C}_i \subset \mathcal{C}_{i+1}$. Partant de $t_N = T$ pour lequel la solution de $(\mathcal{E})$ et son gradient sont connus, on injecte les valeurs de $\overline{u}(t_{i+1},.)$ et $\overline{v}(t_{i+1},.)$ sur la grille $\mathcal{C}_{i+1}$ dans la transition de la diffusion approchée entre t_i to t_{i+1}, $i \in \{0,\ldots,N-1\}$. Cela permet d'exprimer $\overline{u}(t_i,.)$ à l'aide d'une version discrétisée de la formule de Feynman-Kac.

Il reste désormais à détailler la façon de mettre à jour l'approximation du gradient de la solution u. Précisément notre méthode permet d'approcher $\nabla_x u(t_k,.)\sigma(.,u(t_k,.))$ et non $\nabla_x u(t_k,.)$, d'où l'écriture spécifique de l'EDP $(\mathcal{E})$. Nous procédons en deux étapes. Nous utilisons tout d'abord une procédure de type incrément de martingale similaire à celle proposée dans le schéma de discrétisation de Bouchard et Touzi [BT04], ou encore employée dans Bally et *al.* [BPP02]. La deuxième étape consiste à quantifier les incréments Gaussiens qui apparaissent dans la représentation précédente.

Références complémentaires

Certains aspects préliminaires de notre approche sont utilisés par Milstein et Tretyakov [MT99] dans le cas particulier où $(b,f)(x,u(t,x),\nabla_x u(t,x)\sigma(t,x,u(t,x)))$ se réduit à $(b,f)(x,u(t,x))$. Mentionnons toutefois que la preuve de convergence du schéma numérique proposé dans cette référence est valable pour des équations à "petit paramètre" (i.e. à petite matrice de diffusion). De façon générale, les auteurs doivent contrôler les propriétés de régularité de la solution du problème de transport associé à l'équation $(\mathcal{E})$ (i.e. la même équation que $(\mathcal{E})$, sans terme d'ordre 2). Aucune condition de ce type n'apparaît dans ce chapitre: en particulier nous supposerons la matrice $\sigma\sigma^*$ uniformément elliptique. Aussi, nous pensons que le travail de Milstein et Tretyakov [MT99] s'insère dans un contexte différent du notre. Pour cette raison nous éviterons toute comparaison ultérieure entre les deux situations. Indiquons pour conclure que Makarov [Mak03] a adapté avec succès la stratégie de Milstein et Tretyakov [MT99] au cas $(b,f) \equiv (b,f)(x,u(t,x),\nabla_x u(t,x)\sigma(t,x,u(t,x)))$ sous des hypothèses de régularité suffisante des coefficients. Bien entendu, la condition associée au "petit paramètre" reste nécessaire.

Quelques caractéristiques de l'Algorithme

Convergence sous des hypothèses faibles

Dans Douglas, Ma and Protter [DMP96], les auteurs se placent sous des hypothèses de régularité qui permettent d'exprimer le gradient de u comme la solution de l'EDP dérivée.

Par notre stratégie d'incréments de martingales, nous évitons cette différentiation et pouvons donc affaiblir les hypothèses de régularité portant à la fois sur les coefficients de (E) et sur la régularité de la solution u de $(\mathcal{E})$. Dans [DMP96], les coefficients sont supposés différentiables et bornés. Nous les supposons juste Lipschitziens et bornés en x. Dans la référence précédente, la solution u de $(\mathcal{E})$ est au moins bornée dans $C^{2+\alpha/2,4+\alpha}([0,T]\times\mathbb{R}^d), \alpha \in]0,1[$. Nous imposons seulement que u appartienne à $C^{1,2}([0,T]\times\mathbb{R}^d)$ avec des dérivées bornées d'ordre 1 en temps et d'ordre 1 et 2 en x. Sous ces hypothèses nous obtenons des estimations de convergence de $\overline{u}$ vers u, et des processus discrets construits vers (U, V, W), en fonction des paramètres de discrétisation.

Un algorithme complètement implémentable

Rappelons que dans [DMP96], les auteurs considèrent toujours le cas de grilles infinies. Ceci est plus simple pour l'analyse de la convergence mais ne fournit pas en toute généralité un algorithme implémentable. Nous précisons de notre côté l'impact de la troncature des grilles et quantifions cette contribution dans l'erreur globale.

Finalement, indiquons qu'une procédure d'interpolation linéaire est nécessaire à Douglas et *al.* pour définir leur algorithme. Cela peut être lourd en grande dimension. Notre algorithme permet de définir les solutions approchées aux seuls noeuds de la grille spatiale. De la sorte nous pensons que notre méthode est plus simple à implémenter et numériquement moins coûteuse. Par ailleurs nous évitons également les inversions de "grands" systèmes linéaires inhérentes aux techniques usuelles d'analyse numériques.

Du point de vue numérique, les résultats obtenus par notre algorithme pour l'équation de Burgers unidimensionnelle avec coefficient de viscosité non nul, l'équation KPZ déterministe en dimension 2 et l'équation des milieux poreux en dimension 1 sont bons et encourageants pour des situations plus complexes. Notamment, les techniques employées pour prouver la convergence de l'algorithme devraient pouvoir s'adapter au cas d'un problème quasi-linéaire de type Cauchy-Dirichlet.

Soumissions et Publications

Chaque chapitre de cette thèse a fait ou fera l'objet d'une soumission à une revue scientifique avec comité de lecture. Précisément:

- Chapitre 1: E. Gobet et S. Menozzi. "Exact approximation rate of killed hypoelliptic diffusions using the discrete Euler scheme". *Stochastic Processes and their Applications*, vol. 112 (2004), p. 201-223.
- Chapitre 2: E. Gobet et S. Menozzi. "Discrete sampling of functionals of Itô processes". Soumis au *Séminaire de Probabilités.*
- Chapitre 3: S. Menozzi. "Improved Simulation of a discretely killed Brownian Motion in a cone". En préparation.
- Chapitre 4: F. Delarue et S. Menozzi. "A Forward-Backward Stochastic Algorithm for Quasi-Linear PDEs". Soumis à *Annals of Applied Probability.*

Chapter 1

Convergence rate of the Discrete Euler scheme for killed diffusions

Introduction

Let $(X_t)_{0\leq t\leq T}$ be a d-dimensional diffusion process, whose dynamics is given by

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{1.0.1}$$

with a fixed initial data x and a fixed terminal time T. Here, W is a d'-dimensional standard Brownian motion (BM in short) and the mappings b, σ are Lipschitz continuous.

Let D be a domain, i.e. an open connected subset, of $\mathbb{R}^d$. In this chapter, we are interested in approximating the law of this diffusion process, killed when it exits from D. Namely, for a measurable function f, we want to estimate the quantity

$$\mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}] \tag{1.0.2}$$

with $\tau := \inf\{t \geq 0 : X_t \notin D\}$.

Consider now a regular time mesh of the interval $[0,T]$ with N time steps, $(t_i = ih)_{i\in[\![0,N]\!]}$, $h = T/N$ being the step size. For a simulation procedure of (1.0.2), we define the Euler approximation of (1.0.1) by

$$X_t^N = x + \int_0^t b(X_{\phi(s)}^N)ds + \int_0^t \sigma(X_{\phi(s)}^N)dW_s \tag{1.0.3}$$

where $\phi(t) := \sup\{t_i : t_i \leq t < t_{i+1}\}$. Note that the values $(X_{t_i}^N)_{0\leq i\leq N}$ are straightforward to obtain using the simulations of the Brownian increments along the mesh. The so-called *discrete Euler scheme* corresponds to the killing time $\tau^N := \inf\{t_i \geq 0 : X_{t_i}^N \notin D\}$. Thus, in a Monte-Carlo procedure the random variable to simulate is simply given by $f(X_T^N)\mathbb{I}_{\tau^N>T}$ which can directly be derived from the realization of $(X_{t_i}^N)_{0\leq i\leq N}$.

We will focus on the discretization error

$$\mathrm{Err}(T,h,f,x) := \mathbb{E}_x[f(X_T^N)\mathbb{I}_{\tau^N>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}].$$

Main results

If b and σ are Lipschitz continuous and D is smooth with ∂D bounded, we first show that under a non-characteristic boundary condition **(C)**, for every continuous function f with polynomial growth

$$\boxed{\mathrm{Err}(T,h,f,x) \underset{h\longrightarrow 0}{\longrightarrow} 0.}$$

We also indicate that this condition is somehow minimal to ensure a convergent approximation.

For a smooth domain D with compact boundary, if σ is in some-sense non-degenerated one can obtain a precise error decomposition and estimate the speed of convergence.

In the following we always assume f to be non-negative. Under the "usual" hypoellipticity conditions **(H')** on the coefficients b, σ of (1.0.1) and for a bounded measurable function f vanishing on a neighbourhood of the boundary we obtain

$$\boxed{\mathrm{Err}(T,h,f,x) = \frac{1}{2}\mathbb{E}_x[\int_0^T \frac{\partial v}{\partial n}(s, X_s^N) dL^0_{s\wedge\tau^N}(F(X^N))] + o(h^{1/2})}$$

where $v(t,x) := \mathbb{E}_x[f(X_{T-t})\mathbb{I}_{\tau>T-t}]$, F is the signed distance to the boundary and $L^0(F(X^N))$ is the local time of the Euler scheme on the boundary.

Under uniform ellipticity conditions on $\sigma\sigma^*(.)$, the same holds true without support condition under some additional smoothness assumptions on f.

Concerning the speed of convergence, in the uniformly elliptic case, Gobet, cf. [Gob00] and [Gob98], proved with the above conditions on f

$$\boxed{\exists C > 0,\ h \le h_0,\ |\mathrm{Err}(T,h,f,x)| \le C\sqrt{h}.}$$

We first extend this result under **(H')** when $d(\mathrm{supp}(f), \partial D) \ge 2\varepsilon > 0$. Then, strengthening the hypoellipticity assumption, cf. **(H)**, we prove that

$$\boxed{\exists C' > 0,\ \forall h \le h_0(x),\ C'\sqrt{h} \le \mathrm{Err}(T,h,f,x).}$$

In particular, we derive that in this case $\mathrm{Err}(T,h,f,x) \ge 0$. The step $h_0(x)$ depends on the distance $d(x,\partial D)$ of the initial point x to the boundary. If the domain is bounded, the constant C' of the lower bound is uniform w.r.t. the starting point $x \in D$. For unbounded domains, this constant is only locally uniform and also depends on $d(x,\partial D)$.

These results show that anyway the order of convergence for the discrete Euler scheme is exactly $1/2$. In a Brownian setting, this had been observed by Broadie and *al.*, see [BGK97], who obtained an error expansion with first order term in $\sqrt{h}$, thus proving that it was in that case the true speed of convergence. Concerning the positivity of $\mathrm{Err}(T,h,f,x)$ for $f \ge 0$ it is straightforward if $X_t = W_t$ since in this case $X_t = X_t^N$ and $\tau^N \ge \tau$ *a.s.* It had been numerically observed in the general case, see Boyle and Lau [BL94], Baldi [Bal95], but not yet proved.

We also provide numerical examples that confirm the theoretical convergence results. Actually, they are somehow better since they tend to show, numerically speaking, that an expansion holds true even in the general case.

An essential feature which comes up is that the main part of the error $\mathrm{Err}(T,h,f,x)$ can be expressed as a suitable average with positive weights of the overshoot of the discretely killed process above the boundary (see the proof of Theorem 1.5.2 and Lemma 1.4.5) (the overshoot being defined as the distance to the boundary of the process when it exits the domain). Hence, it provides a clear explanation of the main origin of the error:

roughly speaking, the increments are of order 1/2, and hence (but this is not so straightforward as it will be seen), the same estimate holds for the overshoot.

Then, the lower and upper bounds for the error derive from the controls on the overshoot and the respectively positivity and boundedness of the normal derivatives of the solution of the underlying PDE satisfied by v introduced for the error analysis.

The derivation of an expansion for $\mathrm{Err}(T, h, f, x)$ at the order $\frac{1}{2}$ would require the computation of the asymptotic law of the overshoot: this is a classical issue which is usually analyzed with the renewal theory for Markov chains. Unfortunately, in a general multidimensional setting, the available results only hold under ergodicity type conditions (see Alsmeyer [Als94], Fuh and Lai [FL01] and references therein), which are never satisfied on the relevant process (i.e. the time-rescaled Euler scheme). However, in a Brownian framework, i.e. when $X_t = x + \mu t + \sigma W_t$, it is possible to obtain an error development when D is an intersection of half-spaces, see Chapter 3.

Plan of the chapter

In Section 1 we state our working assumptions and also recall some basic properties concerning smooth domains and the Euler scheme.

In Section 2 we introduce the PDE satisfied by $v(t, x) = \mathbb{E}_x[f(X_{T-t})\mathbb{I}_{\tau > T-t}]$ and give some smoothness properties of its Green kernel, i.e. the transition density of the underlying killed diffusion process.

Writing an Itô-like formula using the smoothness of the PDE we then decompose $\mathrm{Err}(T, h, f, x)$ in Section 3.

The various contributions of the former error decomposition are analyzed in Section 4. This is the technical core of the chapter. This Section is divided into three parts: the first one concerns the estimation of the boundary terms, overshoot and normal derivative of v. The second part is dedicated to the control of the expectations of derivatives of v. The last one concerns the analysis of a residual in the error. For the last two parts, in the hypoelliptic framework, we adapt ideas from Bally and Talay, see [BT96a], using some more or less standard Malliavin calculus computations.

We state our main results in Section 5 putting together the previous controls. We first isolate the leading term in the error decomposition. Then, we give the upper and lower bounds for the error in the hypoelliptic case when f satisfies a support condition, cf. **(F1)**. Finally, we detail the case of a smooth function, cf. **(F2-l)**, without support condition when the diffusion is uniformly elliptic.

In Section 6 we present some numerical results.

Eventually, we conclude in Section 7 giving some extensions and evoking some remaining open problems.

1 Notations and assumptions

1.1 About the process

For the coefficients of equation (1.0.1) we introduce the assumption

(S-k) $b \in C_b^k(\mathbb{R}^d, \mathbb{R}^d)$ and $\sigma \in C_b^k(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^{d'})$,

where $C_b^k(\mathbb{R}^d, \mathbb{R}^d)$ (resp. $C_b^k(\mathbb{R}^d, \mathbb{R}^d \otimes \mathbb{R}^{d'})$), $k \in [\![0, \infty]\!]$, denotes the space of bounded $\mathbb{R}^d$ (resp. $\mathbb{R}^d \otimes \mathbb{R}^{d'}$) valued functions having k bounded continuous derivatives on $\mathbb{R}^d$.

We also require the coefficients to satisfy some Hörmander like assumptions. Identifying the coefficients $(\sigma_i)_{1 \le i \le d'}$ with the vector fields which they define, we denote by $L_M(x)$ the set of these vector fields and of their Lie brackets of length lower or equal to M taken at point x. We define the strong Hörmander assumption by

(H) $\exists M \in \mathbb{N}, \exists C > 0, \forall x \in \mathbb{R}^d,\ \forall z \in \mathbb{R}^d,\ \sum_{Y \in L_M(x)} \langle Y, z\rangle^2 \ge C|z|^2.$

To introduce the second type of Hörmander-like assumption we first need to rewrite (1.0.1) in Stratonovitch form. It comes

$$X_t = x + \int_0^t A_0(X_s)ds + \sum_{i=1}^{d'} \int_0^t \sigma_i(X_s) \circ dW_s^i,$$

where the $\circ$ stands for the Stratonovitch integral and $A_0(.) = (b - \frac{1}{2}\sum_{i=1}^{d'} \sigma_i^{\nabla} \sigma_i)(.)$ with for all $j \in [\![1, d]\!]$, $(\sigma_i^{\nabla}\sigma_i)_j(.) = \sum_{l=1}^{d} \sigma_i^l(.)\partial_l \sigma_i^j(.)$. More generally, $\forall (j,k) \in [\![1, d']\!]^2$, $\sigma_j^{\nabla}\sigma_k := (\sum_{l=1}^{d} \sigma_j^l \partial_l \sigma_k^i)_{i \in [\![1,d]\!]}$ denotes the covariant derivative of σ_k in the direction σ_j. The notation $[\sigma_j, \sigma_k] := \sigma_j^{\nabla}\sigma_k - \sigma_k^{\nabla}\sigma_j$ stands for the Lie bracket between σ_j and σ_k.

Identifying once again the coefficient A_0 with the vector field it defines, we denote by $L'_M(x)$ the set of the fields $(\sigma_i)_{i \in [\![1,d']\!]}$ and of the Lie brackets of length lower or equal to M of the fields $(\sigma_i)_{i \in [\![1,d']\!]} \cup A_0$ taken at point x. We introduce:

(H') $\exists M \in \mathbb{N}, \exists C > 0, \forall x \in \mathbb{R}^d,\ \forall z \in \mathbb{R}^d,\ \sum_{Y \in L'_M(x)} \langle Y, z\rangle^2 \ge C|z|^2.$

Remark 1.1.1 *In* **(H)** *and* **(H')**, *we took $x \in \mathbb{R}^d$ for notational convenience. It is actually sufficient, for the convergence analysis, to have the previous statements for $x \in D$, where D denotes the domain at hand.*

Assumption **(H')** is often referred to as the general Hörmander assumption. **(H')** is sufficient to obtain the existence and smoothness of the density for the law of X_t for smooth coefficients b, σ, cf. Theorems 2.3.2 and 2.3.3 in [Nua95]. However, it does not guarantee the positivity of the density which is needed to obtain a lower bound for the error. This positivity property holds true under **(H)**, see Property **(P)** in Subsection 2.2.

In the following we say we are in a hypoelliptic framework if **(H)** or **(H')** is in force[1].

Remark 1.1.2 *For a $\mathbb{R}^d$ valued diffusion process, the uniform ellipticity condition*

(UE) $\exists C > 0,\ \forall x \in \mathbb{R}^d,\ \forall z \in \mathbb{R}^d,\ \langle \sigma\sigma^*(x)z, z\rangle \ge C|z|^2,$

is simply a special case of **(H)** *or* **(H')** *with $M = 0$.*

[1] A differential operator $\mathcal{A}$ with smooth (i.e. infinitely differentiable) coefficients is said to be hypoelliptic in an open set G if for $u \in \mathcal{D}'(G)$, $u \in C^{\infty}(G)$ as soon as $\mathcal{A}u \in C^{\infty}(G)$. For $\mathcal{A} = \sum_{i=1}^{d'} (A_i)^2 + A_0$, Hörmander's Theorem, cf. [Hör67], states that if $\mathrm{Lie}(A_0, A_1, ... A_{d'})$ has full rank at every point $x \in \mathbb{R}^d$ then $\mathcal{A}$ is hypoelliptic.

As usual, the index x in $\mathbb{E}_x$ and $\mathbb{P}_x$ refers to the initial value of a given process for which we compute the expectation or the probability: this will be clear from the context. When needed, we will use the usual notation $X^{t_0,x}$ for the solution of (1.0.1) starting from x at time t_0.

The notation L stands for the infinitesimal generator of the diffusion, $Lg(x) = b(x).\nabla g(x) + \frac{1}{2}\mathrm{tr}(\sigma\sigma^*(x)H_g(x))$, where ∇g (resp. H_g) denotes the gradient (resp. the Hessian matrix) of g. We additionally define for all $z \in \mathbb{R}^d$ the operator $\mathcal{L}_z$ by $\mathcal{L}_z g(x) = b(z).\nabla g(x) + \frac{1}{2}\mathrm{tr}(\sigma\sigma^*(z)H_g(x))$, which can locally be interpreted as the generator of the Euler approximation.

1.2 About the domain

In the following of the chapter, we consider a domain $D \subset \mathbb{R}^d$, i.e. an open connected set, which satisfies some smoothness assumptions.

Definition 1.1.1 *We recall, cf. Gilbarg and Trudinger [GT77] pp. 88-89, that a domain $D \subset \mathbb{R}^d, d \geq 2$ is said to be of class $C^k, k \geq 1$ if for all $s \in \partial D$ there exists a ball $B = B(s)$ and a C^k diffeomorphism ψ from B to $B' \subset \mathbb{R}^d$ s.t.*

$$\begin{cases} \psi(B \cap D) \subset \mathbb{R}^d_+ := \{y \in \mathbb{R}^d, y_1 > 0\}, \\ \psi(B \cap \partial D) \subset \partial\mathbb{R}^d_+ := \{y \in \mathbb{R}^d, y_1 = 0\}, \\ \psi \in C^k(B), \ \psi^{-1} \in C^k(B'). \end{cases}$$

We introduce the hypothesis

(D-k) The domain D is of class C^k with bounded boundary ∂D.

In practice, D will satisfy at least **(D-2)**. Such a smoothness ensures the existence of a tangent space at every point of ∂D and is also needed for some PDE results, see Section 2. Note also that under **(D-k)** the domain D can be unbounded.

For $x \in \partial D$, denote by $n(x)$ the unit inward normal vector at x. For $r \geq 0$, set $V_{\partial D}(r) := \{z \in \mathbb{R}^d : d(z, \partial D) \leq r\}$ and $D(r) := \{z \in \mathbb{R}^d : d(z, D) \leq r\}$. $B(z, r)$ stands for the closed ball with center z and radius r.

We now state standard facts on the distance to the boundary and the orthogonal projection on ∂D (see Lemma 1 and its proof from [GT77] p. 382).

Proposition 1.1.1 *Assume* **(D-k)**, *$k \geq 2$. There is a constant $R > 0$ such that:*

i) for any $x \in V_{\partial D}(R)$, there are unique $s = \pi_{\partial D}(x) \in \partial D$ and $F(x) \in \mathbb{R}$ such that $x = \pi_{\partial D}(x) + F(x)n(\pi_{\partial D}(x))$.

ii) The function $x \longmapsto \pi_{\partial D}(x)$ is the normal projection of x on ∂D: this is a C^{k-1}-function on $V_{\partial D}(R)$.

iii) The function $x \longmapsto F(x)$ is the signed normal distance of x to ∂D: this is a C^k-function on $V_{\partial D}(R)$, which can be extended to a C^k function on $\mathbb{R}^d$ with bounded derivatives. This extension satisfies $F(x) \geq d(x, \partial D) \wedge R$ on D, $F(x) \leq -[d(x, \partial D) \wedge R]$ on D^c and $F = 0$ on ∂D.

iv) For $x \in V_{\partial D}(R)$, one has $\nabla F(x) = n(\pi_{\partial D}(x))$.

We now define the non characteristic boundary condition by assuming

(C) $\exists a_0 > 0, \forall x \in V_{\partial D}(R),\ \alpha(x) := \nabla F(x).\sigma\sigma^*(x)\nabla F(x) \geq a_0$.

From property *iv)* of the previous proposition, assumption **(C)** means that σ is non-tangential to the boundary ∂D.

1.3 About the function

In this chapter we will consider two kinds of assumptions for f. The first one is given by

(F1) f is a non-negative bounded measurable function with support strictly included in D:

$$d(\mathrm{supp}(f), \partial D) \geq 2\varepsilon > 0.$$

We may assume that $2\varepsilon \leq R$ and that f is not identically equal to 0.

To define **(F2-l)** we need to introduce the spaces of Hölder continuous functions. We follow the notations of Ladyzenskaja and *al.*, [LSU68], pp. 7-8. Let $l > 0, l = m + \alpha,\ m \in \mathbb{N}, \alpha \in (0,1)$. We denote by $H^l(\bar{D})$ the Banach space of functions u of class $C_b^m(\bar{D})$ having an α-Hölder continuous m-th derivative endowed with the norm

$$\|u\|_D^l = \langle u\rangle_D^l + \sum_{j=0}^{m} \langle u\rangle_D^{(j)},$$

where

$$\langle u\rangle_D^{(0)} = |u|_D^{(0)} = \sup_{x\in D} |u(x)|,\ \langle u\rangle_D^{(j)} = \sum_{\chi, |\chi|=j} |\partial_x^\chi u|_D^{(0)},\ \langle u\rangle_D^{(l)} = \sum_{\chi,|\chi|=m} \langle \partial_x^\chi u\rangle_D^{(\alpha)},$$
$$\langle u\rangle_D^{(\alpha)} = \sup_{x,x'\in D} \frac{|u(x) - u(x')|}{|x - x'|^\alpha}.$$

For l satisfying the above conditions we define

(F2-l) f is a non-negative function and belongs to $H^l(\bar{D})$.

One can now extend the definition of **(S-l)** and **(D-l)** for a nonintegral number l. We simply mean by this that for **(S-l)**, $l \notin \mathbb{N}$, the coefficients b, σ belong to $H^l(\bar{D})$ and for **(D-l)** that the domain is of class H^l, i.e. the function ψ in Definition 1.1.1 belongs to $H^l(\bar{D})$.

1.4 Miscellaneous

For smooth functions $g(t,x)$, we denote by $\partial_x^\alpha g(t,x)$ the derivative of g w.r.t. x according to the multi-index α, whereas time derivatives of g are denoted by $\partial_t g(t,x), \partial_t^2 g(t,x), \ldots$. The notation $\frac{\partial g}{\partial n}(t,x) = \nabla g(t,x).n(x)$ is the normal derivative on the boundary.

The distribution function of the standard normal law is denoted by Φ.

We will keep the same notation C (or C') for all finite, non-negative constants which will appear in our computations: they may depend on D, T, b, σ or f, but they will not depend on the number of time steps N and the initial value x. We reserve the notation c and c' for constants also independent of x, T and f.

In the following $O_{pol}(h)$ (resp. $O(h)$) stands for every quantity $R(h)$ such that $\forall n \in \mathbb{N}$, for some $C > 0$, one has $|R(h)| \leq Ch^n$ (resp. $|R(h)| \leq Ch$) (uniformly in x).

1.5 Usual controls.

We now give some basic estimates for the Euler scheme X^N, which will be useful in the whole chapter. They only exploit the boundedness of the coefficients (and are thus also valid for X).

Lemma 1.1.2 (Bernstein's type inequality) *Assume* **(S-1)**. *Consider two stopping times S, S' upper bounded by T with $0 \leq S' - S \leq \Delta \leq T$. Then for any $p \geq 1$ and $c' > 0$, there are some constants $c > 0$ and C, such that for any $\eta \geq 0$, one has a.s:*

$$\mathbb{P}[\sup_{t\in[S,S']} \|X_t^N - X_S^N\| \geq \eta \mid \mathcal{F}_S] \leq C \exp\left(-c\frac{\eta^2}{\Delta}\right),$$
$$\mathbb{E}[\sup_{t\in[S,S']} \|X_t^N - X_S^N\|^p \mid \mathcal{F}_S] \leq C\Delta^{p/2},$$
$$\mathbb{E}[\exp\left(-c'\frac{d^2(X_{S'}^N, \partial D)}{\Delta}\right) \mid \mathcal{F}_S] \leq C \exp\left(-c\frac{d^2(X_S^N, \partial D)}{\Delta}\right).$$

Proof. To simplify the notations we prove the first inequality for $d = d' = 1$. The case $d \wedge d' > 1$ is similar. Two cases can be distinguished.

- If $\eta \leq 2|b|_\infty\Delta$, we upper bound the probability by 1. Namely, $\mathbb{P}[\sup_{t\in[S,S']}|X_t^N - X_S^N| \geq \eta|\mathcal{F}_S] \leq 1 \leq \exp(c4|b|_\infty^2 T)\exp(-c\frac{\eta^2}{\Delta})$.

- If $\eta > 2|b|_\infty\Delta$, we write $\mathbb{P}[\sup_{t\in[S,S']}|X_t^N - X_S^N| \geq \eta|\mathcal{F}_S] \leq \mathbb{P}\left[\sup_{t\in[S,S']}\left|\int_S^t \sigma(X_{\phi(s)}^N)dW_s\right| \geq \frac{\eta}{2}\middle|\mathcal{F}_S\right]$.

We then introduce the martingale: $(M_t := \int_0^t \mathbb{I}_{S<s\leq S'}\sigma(X_{\phi(s)}^N)dW_s)_{t\geq 0}$, whose bracket satisfies $\langle M\rangle_\infty = \langle M\rangle_T \leq |\sigma^2|_\infty\Delta$. Hence, as a consequence of Bernstein's inequality, cf. Revuz-Yor [RY99] p.154 we get

$$\mathbb{P}\left[\sup_{t\in[S,S']}|X_t^N - X_S^N| \geq \eta \middle| \mathcal{F}_S\right] \quad\leq\quad \mathbb{P}[\sup_{t\in[0,T]}|M_t| \geq \frac{\eta}{2}] \leq 2\exp\left(-\frac{\eta^2}{8|\sigma|_\infty^2\Delta}\right)$$

which proves the result.

The second inequality in Lemma 1.1.2 is a direct consequence of the first one noting that

$$\mathbb{E}[\sup_{t\in[S,S']}\|X_t^N - X_S^N\|^p|\mathcal{F}_S] = p\int_{\mathbb{R}+} d\eta\eta^{p-1}\mathbb{P}[\sup_{\in[S,S']}\|X_t^N - X_S^N\| \geq \eta|\mathcal{F}_S].$$

For the last inequality we consider the events $A := \{\|X_{S'}^N - X_S^N\| < \frac{1}{2}d(X_S^N, \partial D)\}$ and A^C. On A one has $d(X_{S'}^N, \partial D) \geq \frac{1}{2}d(X_S^N, \partial D)$. Hence,

$$\mathbb{E}[\mathbb{I}_A\exp\left(-c'\frac{d^2(X_{S'}^N, \partial D)}{\Delta}\right) \mid \mathcal{F}_S] \leq \exp\left(-c'\frac{d^2(X_S^N, \partial D)}{4\Delta}\right).$$

On A^C we simply write $\mathbb{E}[\mathbb{I}_{A^C}\exp\left(-c'\frac{d^2(X_{S'}^N, \partial D)}{\Delta}\right) \mid \mathcal{F}_S] \leq \mathbb{P}[\|X_{S'}^N - X_S^N\| \geq \frac{1}{2}d(X_S^N, \partial D)|\mathcal{F}_S]$ and derive the result from the first inequality.

□

2 Some smoothness and positivity properties for v

As explained before, we are going to analyze the error using the mixed Cauchy-Dirichlet problem satisfied by

$$v(t,x) := \mathbb{E}_x[f(X_{T-t})\mathbb{I}_{\tau>T-t}], (t,x) \in [0,T] \times \bar{D}.$$

In this section, we specify some smoothness and positivity properties for v under our two main working assumptions. Namely:

- Hypoellipticity conditions **(H)** or **(H')** associated to a non characteristic boundary condition **(C)**, and to the assumptions **(D-∞)** for the domain, **(S-∞)** for the coefficients, **(F1)** for the function.

- Uniform ellipticity conditions **(UE)** associated to the assumptions **(D-$l+2$)**, $l \in (0,1)$, for the domain, **(S-1)** for the coefficients and **(F2-$l+2$)** for the function.

2.1 Smoothness properties

Hypoelliptic case

Proposition 1.2.1 *Assume* **(C)**, **(D-∞)**, **(H')**, **(S-∞)**. *For every bounded Borel function f, v satisfies the following parabolic PDE*

$$\begin{cases} \partial_t v + Lv = 0 \text{ on } [0,T) \times \bar{D}, \\ v(t,x) = 0 \text{ on } [0,T] \times D^c, \\ \lim_{t\to T, y\to x} v(t,y) = v(T,x) = f(x), \text{ if } x \in D \text{ and } f \text{ is continuous in } x. \end{cases} \tag{1.2.1}$$

The function v belongs to $C^\infty([0,T)\times\bar{D},\mathbb{R})$ and for any multi-index α there exist constants ζ, C s.t.

$$\forall t\in[0,T),\ \sup_{x\in\bar{D}}|\partial_x^\alpha v(t,x)|\le C\|f\|_\infty/(T-t)^\zeta.$$

Proof. It is a direct consequence of the existence of a smooth kernel $q_s(.,.)$ s.t.

$$v(t,x)=\int_D q_{T-t}(x,y)f(y)dy.$$

From a probabilistic viewpoint, q_{T-t} represents the transition density of the killed process at time $T-t$.

Under our assumptions, Cattiaux, cf. [Cat91], proved that the function q is $C^\infty((0,T]\times\bar{D}\times\bar{D},\mathbb{R})$ and satisfies Kolmogorov's forward and backward equations. Hence, v satisfies (1.2.1).

From Proposition 3.44 in [Cat91] and following the arguments used in the proof of Proposition 1.12 in [Cat90] we also derive that for any multi-index α there exist constants $c>0,\zeta>0$ and C s.t.

$$\forall(s,x,y)\in(0,T]\times\bar{D}\times\bar{D},\ |\partial^\alpha q_s(x,y)|\le\frac{C}{s^\zeta}\exp(-c\frac{\|x-y\|^2}{s}). \tag{1.2.2}$$

This gives the second statement of the Proposition.

□

Note that if we additionally assume **(F)**, as a consequence of (1.2.2), v and its derivatives near ∂D are uniformly bounded and exponentially decreasing as $t\to T$: for all multi-index α there exist constants $c>0,\zeta>0$ and C such that

$$\forall(t,x)\in[0,T)\times V_{\partial D}(\varepsilon),\ |\partial_x^\alpha v(t,x)|\le C\frac{\|f\|_\infty}{1\wedge\varepsilon^\zeta}\exp(-c\frac{\varepsilon^2}{T-t}). \tag{1.2.3}$$

Uniformly elliptic case

Before giving the main result of this section we need to introduce some additional functional spaces. As in Section 1.3 we follow the notations of Ladyzenskaja and *al.*, [LSU68] pp. 7-8.

For a positive nonintegral number l, we denote by $H^{l/2,l}([0,T]\times\bar{D})$ the Banach space of functions $u(t,x)$ that are continuous in $[0,T]\times\bar{D}$ together with all derivatives of the form $\partial_t^r\partial_x^s u$ for $2r+s<l$, and have finite norm

$$\|u\|^{(l)}_{(0,T)\times D}=\langle u\rangle^{(l)}_{(0,T)\times D}+\sum_{j=0}^{\lfloor l\rfloor}\langle u\rangle^{(j)}_{(0,T)\times D},$$

where

$$\langle u\rangle^{(0)}_{(0,T)\times D}=|u|^{(0)}_{(0,T)\times D}=\sup_{(t,x)\in(0,T)\times D}|u(t,x)|,\ \langle u\rangle^{(j)}_{(0,T)\times D}=\sum_{2r+s=j}|\partial_t^r\partial_x^s u|^{(0)}_{(0,T)\times D},$$

$$\langle u\rangle^{(l)}_{(0,T)\times D}=\langle u\rangle^{(l)}_{x,(0,T)\times D}+\langle u\rangle^{(l/2)}_{t,(0,T)\times D},$$

$$\langle u\rangle^{(l)}_{x,(0,T)\times D}=\sum_{2r+s=\lfloor l\rfloor}\langle\partial_t^r\partial_x^s u\rangle^{(l-\lfloor l\rfloor)}_{x,(0,T)\times D},\ \langle u\rangle^{(l/2)}_{t,(0,T)\times D}=\sum_{0<l-2r-s<2}\langle\partial_t^r\partial_x^s u\rangle^{(\frac{l-2r-s}{2})}_{t,(0,T)\times D},$$

$$\langle u\rangle^{(\alpha)}_{x,(0,T)\times D}=\sup_{(t,x),(t',x')\in(0,T)\times D}\frac{|u(t,x)-u(t',x')|}{|x-x'|^\alpha},\ \langle u\rangle^{(\alpha)}_{t,(0,T)\times D}=\sup_{(t,x),(t',x')\in(0,T)\times D}\frac{|u(t,x)-u(t',x')|}{|t-t'|^\alpha},$$

$\alpha\in(0,1)$.

Proposition 1.2.2 *Let $l\in(0,1)$. Under* **(UE)**, **(D-**$l+2$**)**, **(S-**1**)**, **(F2-**$l+2$**)**, *if f satisfies the compatibility conditions $f|_{\partial D}=Lf|_{\partial D}=0$, v is of class $H^{l/2+1,l+2}([0,T]\times\bar{D})$ and satisfies*

$$\begin{cases}\partial_t v+Lv=0 \text{ on } [0,T]\times\bar{D},\\ v(t,x)=0 \text{ on } [0,T]\times D^c,\\ v(T,x)=f(x),x\in D.\end{cases} \tag{1.2.4}$$

Furthermore, one has

$$\|v\|_{(0,T)\times D}^{l/2+1,l+2} \le C\|f\|_D^{(l+2)}. \qquad (1.2.5)$$

Proof. It is a direct application of a well known result in PDE theory. Indeed, under the assumptions of the proposition, Theorem 5.2 p 320 from [LSU68] states that there exists a unique solution u satisfying (1.2.4) and (1.2.5). For $l > 0$ this solution u is at least $C^{1,2}([0,T]\times\bar{D})$. Hence, for all $(t,x)\in[0,T]\times\bar{D}$, we get that $u(t,x) = v(t,x)$. It suffices to write the semi-martingale decomposition of $M_s := u(s\wedge\tau, X_{s\wedge\tau}^{t,x})$ between t and T using Itô's formula and to take the expectation. This is the usual Feynman-Kac verification procedure.

□

2.2 Positivity properties in the domain

We first state a positivity property for v which is used for the lower bound and requires the stronger assumption **(H)**. Since we assume the function f is not identically equal to 0, we have:

(P) Under **(C)**, **(D-∞)**, **(F1)**, **(H)**, **(S-∞)**, (resp. **(UE)**, **(D-$l+2$)**, **(F2-$l+2$)**, $l\in(0,1)$, **(S-1)**) $v(t,x) > 0$ on $[0,T)\times D$.

Proof. We are reduced to check the positivity of $q_t(x,y)$ on $(0,T]\times D\times D$. For the first set of assumptions, this property follows from the arguments used for Lemma 5.37 in [Cat92], that can be adapted to our case. For the second set of assumptions, i.e. under a uniform ellipticity condition, we refer to Friedman, [Fri64] Theorem 11 p 44. For SDEs in the whole space under **(H)**, see Ben Arous and Léandre, [BL91], and for a Malliavin calculus approach, Theorem 3.3.6.1 in Michel and Pardoux [MP90] or Section 4.3 in Nualart [Nua98].

□

To derive a lower bound for the error we also need to have a positivity property for $\frac{\partial v}{\partial n}$. We first give a general result from which we derive in Lemma 1.2.4 the required one in our framework.

Actually, the type of result stated in Lemma 1.2.3 is known in the PDE theory as the *Hopf boundary point* lemma: in the *uniformly parabolic* case, see Friedman, [Fri64]; for *partially degenerate* elliptic operators see Lieberman, [Lie85]. We give here a variant of this result, using a probabilistic proof under the sole assumption **(C)** and without smoothness properties on u, which seems to be new.

Lemma 1.2.3 *Assume* **(C)**, **(D-2)** *and* **(S-1)**. *Consider* $(t_0,x_0)\in\mathbb{R}^+\times\partial D$ *and the time-space set* $\mathcal{D} = [t_0,t_0+\delta)\times(D\cap V_{\partial D}(R))$ *(with* $\delta>0$ *and* R *defined as in Proposition 1.1.1). If* u *is a bounded continuous function defined on* $\overline{\mathcal{D}}$ *such that* $U^x = (U_s^x = u(s\wedge\tau_{\mathcal{D}}, X_{s\wedge\tau_{\mathcal{D}}}^{t_0,x}))_{s\ge t_0}$ *(with* $\tau_{\mathcal{D}} = \inf\{s\ge t_0 : (s, X_s^{t_0,x})\notin\mathcal{D}\}\le t_0+\delta$*) defines a super-martingale and* $u(t,x) > u(t_0,x_0)$ *for* $(t,x)\in\mathcal{D}$*, then one has*

$$\liminf_{\lambda\downarrow 0}\frac{u(t_0,x_0+\lambda n(x_0)) - u(t_0,x_0)}{\lambda} > 0.$$

Proof. The main idea is to consider a closed subset $A\subset\overline{\mathcal{D}}$ containing the points $(t_0,x_0+\lambda n(x_0))_{0\le\lambda\le\lambda_0}$ ($\lambda_0>0$ small enough) and a $C_b^\infty(\overline{\mathcal{D}})$ function w with the four following requirements:

i) $w(t_0,x_0) = 0$.

ii) $\partial_t w + Lw \ge 0$ on A.

iii) $\frac{\partial w}{\partial n}(t_0,x_0) > 0$.

iv) $u \ge u(t_0,x_0) + \varepsilon_0 w$ on ∂A for some $\varepsilon_0 > 0$.

Then for such A and w, if we set τ_A for the exit time of $(s, X_s^{t_0,x})_{s\ge t_0}$ from A (for (t_0,x) in A), we easily deduce by **ii)** that $(Z_s := u(s\wedge\tau_A, X_{s\wedge\tau_A}^{t_0,x}) - u(t_0,x_0) - \varepsilon_0 w(s\wedge\tau_A, X_{s\wedge\tau_A}^{t_0,x}))_{s\ge 0}$ is a super-martingale, and thus using **iv)** and **i)** $0\le\mathbb{E}[Z_{\tau_A}]\le Z_{t_0} = u(t_0,x) - u(t_0,x_0) - \varepsilon_0(w(t_0,x) - w(t_0,x_0))$. Take $(t_0,x) = (t_0,x_0+\lambda n(x_0))\in A$ with $\lambda\downarrow 0$ to get the result considering **iii)**.

Now, we turn to the construction of A, w and ε_0. Assumption **(C)** is here crucial. Up to modifying u for

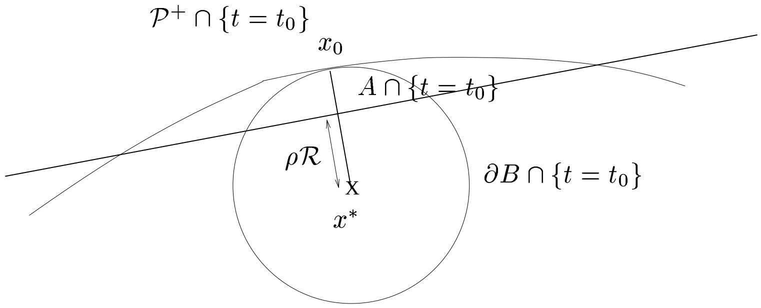


Figure 1.1: Space representation of A at $t = t_0$.

$t < t_0$, we can assume that $\mathcal{D}$ is of the form $\mathcal{D} = (t_0 - \delta, t_0 + \delta) \times (D \cap V_{\partial D}(R))$. Under **(D)**, x_0 satisfies an interior sphere condition in D that permits to construct a time-space ball $B := B(P^*, \mathcal{R}) \subset \overline{\mathcal{D}}$ (w.l.o.g. $\mathcal{R} < \delta \wedge R/2$), $P^* = (t_0, x^*)$ s.t. $x^* - x_0 = \mathcal{R} n(x_0)$ and $B \cap \partial\mathcal{D} = \{(t_0, x_0)\}$ (see Figure 1.1). Now, introduce the time cylindrical half-space $\mathcal{P}^+ := [t_0 - \delta, t_0 + \delta] \times \{z \in \mathbb{R}^d : (x^* - z).n(x_0) \geq \rho\mathcal{R}\}$ for $\rho \in (0,1)$ and denote $A := B \cap \mathcal{P}^+$ the expected set. For $\alpha > 0$, we define $w_\alpha(t,x) = \exp\left(-\alpha r^2\right) - \exp\left(-\alpha\mathcal{R}^2\right)$ where $r^2 := \|x - x^*\|^2 + (t - t_0)^2$: easily, we get $[\partial_t + L]w_\alpha(t,x) \geq \exp(-\alpha r^2)\left(2\langle\sigma\sigma^*(x)(x - x^*), x - x^*\rangle\alpha^2 - C\alpha\right)$ for $(t,x) \in A$. Since by continuity $\lim_{x\to x_0}\langle\sigma\sigma^*(x)(x - x^*), x - x^*\rangle \geq a_0\mathcal{R}^2$ under **(C)**, it is clear that we can choose the cutting-level ρ close enough to 1 to ensure $[\partial_t + L]w_\alpha \geq \exp(-\alpha r^2)(a_0\mathcal{R}^2\alpha^2 - C\alpha) \geq 0$ on A for α big enough: for such α, $w = w_\alpha$ satisfies **iii)**. Indeed, $\frac{\partial w}{\partial n}(t_0, x_0) = 2\alpha\exp(-\alpha\mathcal{R}^2) > 0$. Statements **i)** and **ii)** are straightforward to check. It remains to exhibit $\varepsilon_0 > 0$ in **iv)**: since $w = 0$ on ∂B, we may consider only $(t,x) \in \overline{\partial A \backslash \partial B}$. But on this compact set, $u > u(t_0, x_0)$ and thus, **iv)** holds true for ε_0 small enough.

□

As a direct consequence of the previous Lemma applied to the function v we have

Lemma 1.2.4 (Positivity of the inner normal derivative). *Under* **(C)**, **(D-∞)**, **(F1)**, **(H)**, **(S-∞)** *(resp.* **(UE)**, **(D-$l+2$)**, **(F2-$l+2$)**, $l \in (0,1)$, **(S-1)**, $f|_{\partial D} = Lf|_{\partial D} = 0$*), for any* $(t,x) \in [0,T[\times\partial D$*, we have* $\frac{\partial v}{\partial n}(t,x) > 0$.

Proof. The martingale property for the above U^x taking $u = v$ easily follows from the one for $(v(t\wedge\tau, X_{t\wedge\tau}) = \mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}|\mathcal{F}_t])_{0\leq t\leq T}$ (Markov property). Since $v(t_0, x_0) = 0$ for $(t_0, x_0) \in [0,T[\times\partial D$, Property **(P)**, introduced in Section 2.2, provides the required strict lower bound for v.

At last, since as a consequence of Proposition 1.2.1 (resp. Proposition 1.2.2) v is smooth up to the boundary, the lim inf in Lemma 1.2.3 equals the normal derivative.

□

3 Error decomposition

From now on, R denotes the constant introduced in Proposition 1.1.1 for domains satisfying **(D-k)**, $k \geq 2$. Recall that in this framework, $\forall x \in D_R$ the projection $\pi_{\bar{D}}(x)$ is uniquely defined. Put $\tau_R^N := \inf\{s \geq 0 : X_s^N \notin D_R\}$. Under the assumptions of Proposition 1.2.1 or Proposition 1.2.2 we decompose the error $\mathrm{Err}(T,h,f,x)$ as follows:

$$\begin{aligned}
\mathrm{Err}(T,h,f,x) &= \mathbb{E}_x\left[f(X_T^N)\mathbb{I}_{\tau^N>T}\right] - \mathbb{E}_x\left[f(X_T)\mathbb{I}_{\tau>T}\right] \\
&= \mathbb{E}_x\left[f(X_T^N)\mathbb{I}_{\tau^N\wedge\tau_R^N>T}\right] + \mathbb{E}_x\left[f(X_T^N)\mathbb{I}_{\tau^N>T\geq\tau_R^N}\right] - v(0,x) \\
&= \mathbb{E}_x\left[v(\tau^N\wedge\tau_R^N\wedge(T-h), X^N_{\tau^N\wedge\tau_R^N\wedge(T-h)})\right] - v(0,x) \\
&\quad +\mathbb{E}_x\left[v(\tau^N\wedge\tau_R^N\wedge T, X^N_{\tau^N\wedge\tau_R^N\wedge T})\right] - \mathbb{E}_x\left[v(\tau^N\wedge\tau_R^N\wedge(T-h), X^N_{\tau^N\wedge\tau_R^N\wedge(T-h)})\right] \\
&\quad +\mathbb{E}_x\left[\mathbb{I}_{\tau_R^N\leq T<\tau^N}f(X_T^N)\right] := E_1(h) + E_2(h) + E_3(h).
\end{aligned} \tag{1.3.1}$$

Lemma 1.1.2 gives $E_3 = O_{pol}(h)$. The term $E_2(h)$ is the contribution to the error of the last time step. It is shown in Section 4.4, Lemma 1.4.12 and Lemma 1.4.13, that it is a term of order 1 w.r.t. h.

For $E_1(h)$ we recall from Lemma 1.2.4 that ∇v is discontinuous on the boundary. Thus, we can not directly apply Itô's formula and we use the semi-martingale decomposition of $v(t\wedge\tau_R^N, X^N_{t\wedge\tau_R^N})$ as in [Gob00] Corollary 3.1. In order to be self-contained we briefly recall the main steps that yield this development. Note that since $\forall t \in [0,T]$, $\text{supp}(v(t,.)) \subset \bar{D}$, one has $E_1(h) = \mathbb{E}_x[v(\tau^N\wedge\tau_R^N\wedge(T-h), \pi_{\bar{D}}(X^N_{\tau^N\wedge\tau_R^N\wedge(T-h)}))] - v(0,x)$. Hence, to obtain the semi-martingale decomposition of $E_1(h)$ it suffices to show that the projection of the Euler scheme $(\pi_{\bar{D}}(X^N_{t\wedge\tau_R^N}))_{t\in[0,T]}$ is a semi-martingale and to give its decomposition. Indeed, the result then follows from the smoothness properties of the previous section regarding v and a direct application of Itô's formula.

3.1 Semi-martingale decomposition for $\pi_{\bar{D}}(X^N_{t\wedge\tau_R^N})$

Let us first consider the case of a half space $D = \{x\in\mathbb{R}^d : x_1 > a\}$. The algebraic distance of $x\in\mathbb{R}^d$ to ∂D writes $x_1 - a$. Thus, for the projection of the Euler scheme we have:

$$\pi_{\bar{D}}(X_s^N) = \left((X_s^N)_1 - a)^+ + a, (X_s^N)_2, \ldots, (X_s^N)_d\right)^*.$$

From Tanaka's formula, $\pi_{\bar{D}}(X_s^N)$ is a semi-martingale with decomposition

$$\pi_{\bar{D}}(X_s^N) = \left((x_1-a)^+ + \int_0^s \mathbb{I}_{X_u^N\in D} d(X_u^N)_1 + \frac{1}{2}L_s^a((X_.^N)_1), (X_s^N)_2, \ldots, (X_s^N)_d\right)^*.$$

Assumption **(D-k)**, $k\geq 2$ allows to locally map some neighborhood of the boundary into a portion of half-space using local chart changes. We now specify, without proof, some properties stated in Proposition 1.1.1. For details see Gilbarg and Trudinger [GT77] pp. 381-384, Ladyzenskaja and *al.*, [LSU68], Chapter 4 §4 and Gobet [Gob98] Chapter 3 Section 3.3.

Property 1.3.1 Normal Chart change
Assume **(D-k)**, $k\geq 2$. *There exist a constant* $R_0 > 0$ *s.t. for every* $s\in\partial D$ *there exists a* C^{k-1}*-diffeomorphism* J^s, $(K^s = (J^s)^{-1})$ *from a neighbourhood* U^s *of* s, *in* $[-2R_0, 2R_0]\times V^s$, *s.t.:*

$$\begin{cases} U^s\subset\mathbb{R}^d & \longrightarrow & [-2R_0,2R_0]\times V^s\subset\mathbb{R}\times\mathbb{R}^{d-1} & \\ x & \longmapsto & (z_1, z) & s.t.\ \ x = g^s(z) + z_1 n(g^s(z))\end{cases}$$

where g^s *is a local parameterization of* ∂D.

Remark 1.3.1 *Note that the compactness assumption in* **(D-k)** *is needed in order to choose a global* R_0 *for every* $s\in\partial D$.

Remark 1.3.2 *The component* $J_1^s(x) = z_1(x)$ *represents the algebraic distance from* x *to* ∂D. *Observing that* $V_{\partial D}(2R_0) = \bigcup_{s\in\partial D} U_s$, *we get that* J_1^s *does not depend on the point* $s\in\partial D$. *As in Proposition 1.1.1 we denote it by* F. *We recall that, under* **(D-k)**, F *is a* C^k *function on* $V_{\partial D}(R)$.

From Property 1.3.1, we derive that for every $s\in\partial D$, $x\in U_s$ the projections on ∂D and $\bar{D}$ are respectively given by the mappings

$$K^s\begin{pmatrix} 0 \\ J_2^s(x) \\ \vdots \\ J_d^s(x)\end{pmatrix},\ K^s\begin{pmatrix} [F(x)]^+ \\ J_2^s(x) \\ \vdots \\ J_d^s(x)\end{pmatrix}. \tag{1.3.2}$$

To go from this local description of the projection to a global one we introduce a partition of unity of $\bar{D}$.

Property 1.3.2 Partition of unity.
Assume **(D-**k**)**, $k \geq 2$. *By compactness there exist a finite number of points* $(s_i)_{i\in[\![1,k]\!]}$, $s_i \in \partial D$ *for which we consider the associated* U^i, V^i, J^i, K^i *from Property 1.3.1 and s.t.* $\bigcup_{1\leq i\leq k} U^i$ *covers* $V_{\partial D}(3/2R_0)$. *Consider an open set* U_0 *s.t.* $d(\partial D, \overline{U^0}) > 0$ *and* $\bigcup_{0\leq i\leq k} U^i$ *covers* $D_{3/2R_0}$. *One can build a partition of unity subordinated to* $\bigcup_{0\leq i\leq k} U^i$, *i.e. a set of non-negative* C_b^∞ *functions* $(\phi_i)_{i\in[0,k]}$ *s.t.:*

$$D \subset \bar{D}_{R_0} \subset \bigcup_{i\in[0,k]} U_i,\ \forall i \in [0,k],\ \mathrm{supp}(\phi_i) \subset U_i,\ \sum_{i=0}^{k} \phi^i(x) = 1 \text{ on } \bar{D}_{R_0}.$$

So far, the functions J^i (resp. K^i) have been well defined only on U^i (resp. $(-2R_0, 2R_0) \times V^i$). We extend them smoothly on $\mathbb{R}^d$. Anyhow this extension is only formal since we are going to localize using the above partition of unity. As a matter of fact, the constant R of Proposition 1.1.1 is actually equal to R_0.

From (1.3.2) and Property 1.3.2 we obtain

Lemma 1.3.1 Projection on $\bar{D}$.
Assume **(D-**k**)**, $k \geq 2$. *For* $x \in D_R$ *the projection on* $\bar{D}$ *writes*

$$\pi_{\bar{D}}(x) = \phi^0(x) + \sum_{i=1}^{k} \phi^i(x) K^i \begin{pmatrix} [F(x)]^+ \\ J_2^i(x) \\ \vdots \\ J_d^i(x) \end{pmatrix}. \tag{1.3.3}$$

Remark 1.3.3 *The restriction to* D_R *is needed to uniquely define the projection on* $\bar{D}$. *If the domain* D *is convex, one can take* $R = +\infty$.

We are now in position to give the semi-martingale decomposition of the projection of the stopped Euler-scheme.

Proposition 1.3.2 *Assume* **(D-**3**)**, **(S-**1**)**. *The process* $\pi_{\bar{D}}(X^N_{t\wedge\tau_R^N})_{t\geq 0}$ *is a semi-martingale with decomposition*

$$\pi_{\bar{D}}(X^N_{t\wedge\tau_R^N}) = x + \int_0^{t\wedge\tau_R^N} \mathbb{I}_{X_s^N\in D} dX_s^N + \int_0^{t\wedge\tau_R^N} \mathbb{I}_{X_s^N\notin D} dX_s^{N,\partial D} + \frac{1}{2} n(X^N_{t\wedge\tau_R^N}) L^0_{t\wedge\tau_R^N}(F(X^N))$$

where on the set $\{t < \tau_R^N\}$, $X^{N,\partial D}$ *is a continuous semi-martingale with decomposition:*

$$dX_t^{N,\partial D} = \sum_{i=1}^{k} \phi^i(X_t^N) d\left(K^i \begin{pmatrix} 0 \\ J_2^i(X_t^N) \\ \vdots \\ J_d^i(X_t^N) \end{pmatrix} \right).$$

Proof. Since we have assumed **(D-3)** for the domain, we get from Property 1.3.1 that the functions K^i, J^i, ϕ^i in (1.3.3) are at least C^2. Hence, the result follows identifying the terms in the Itô differential of $\pi_{\bar{D}}(X_{t\wedge\tau_R^N})$ using (1.3.3) and Tanaka's formula for the term $[F(X_s^N)]^+$. For details, see Proposition 3.1 in [Gob00].

□

3.2 Semi-martingale and error decomposition of $E_1(h)$

Propositions 1.3.2 and 1.2.1 (resp. 1.2.2) readily give

Proposition 1.3.3 *Assume* **(C)**, **(D-∞)**, **(H')**, **(S-∞)** *(resp.* **(UE)**, **(D-3)**, **(S-$l+1$)**, **(F2-$l+2$)**, $l \in (0,1)$, $f|_{\partial D} = Lf|_{\partial D} = 0$*). One has:*

$$
\begin{aligned}
E_1(h) = \mathbb{E}_x \Bigg[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} & \mathbb{I}_{X_t^N \in D} \partial_t v(t, X_t^N) dt + \frac{1}{2}\frac{\partial v}{\partial n}(t, X_t^N) dL_t^0(F(X^N)) \\
& + \mathbb{I}_{X_t^N \in D} \left(\nabla v(t, X_t^N) \cdot dX_t^N + \frac{1}{2}\mathrm{tr}(H_v(t, X_t^N) d\langle X^N \rangle_t) \right) \\
& + \mathbb{I}_{X_t^N \notin D} \left(\nabla v(t, \pi_{\partial D}(X_t^N)) \cdot dX_t^{N,\partial D} + \frac{1}{2}\mathrm{tr}(H_v(t, \pi_{\partial D}(X_t^N)) d\langle X^{N,\partial D} \rangle_t) \right) \Bigg].
\end{aligned}
$$

Let us now rewrite $E_1(h)$ exploiting the properties of Section 2. Using the PDE one gets

$$
\mathbb{E}_x \left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \mathbb{I}_{X_t^N \in D} \partial_t v(t, X_t^N) dt \right] = -\mathbb{E}_x \left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \mathbb{I}_{X_t^N \in D} L v(t, X_t^N) dt \right].
$$

Under the assumptions of Proposition 1.3.3, we derive from Proposition 1.2.1 (resp. 1.2.2) that the first order derivatives of v are bounded. Hence, the martingale terms cancel taking the expectation. It comes,

$$
\begin{aligned}
& \mathbb{E}_x \left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \mathbb{I}_{X_t^N \in D} \left(\nabla v(t, X_t^N) \cdot dX_t^N + \frac{1}{2}\mathrm{tr}(H_v(t, X_t^N) d\langle X^N \rangle_t) \right) \right] \\
= \ & \mathbb{E}_x \left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \mathbb{I}_{X_t^N \in D} \mathcal{L}_{X^N_{\phi(t)}} v(t, X_t^N) dt \right].
\end{aligned}
$$

Finally, we isolate three main components in $E_1(h)$. Namely,

$$
\begin{aligned}
E_1(h) \ = \ & \mathbb{E}_x \left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \frac{1}{2}\frac{\partial v}{\partial n}(t, X_t^N) dL_t^0(F(X^N)) \right. \\
& + \mathbb{E}_x \left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \mathbb{I}_{X_t^N \in D} \left((\mathcal{L}_{X^N_{\phi(t)}} - L) v(t, X_t^N) \right) dt \right] \\
& + \mathbb{E}_x \left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \mathbb{I}_{X_t^N \notin D} \left(\nabla v(t, \pi_{\partial D}(X_t^N)) \cdot dX_t^{N,\partial D} + \frac{1}{2}\mathrm{tr}(H_v(t, \pi_{\partial D}(X_t^N)) d\langle X^{N,\partial D} \rangle_t) \right) \right] \\
:= \ & E_1^{LT}(h) + E_1^I(h) + E_1^O(h). \qquad (1.3.4)
\end{aligned}
$$

3.3 Main decomposition of the error $\mathrm{Err}(T, h, f, x)$

From (1.3.1) and (1.3.4) we eventually write

$$
\mathrm{Err}(T, h, f, x) = E_1^{LT}(h) + E_1^I(h) + E_1^O(h) + E_2(h) + O_{pol}(h). \qquad (1.3.5)
$$

It therefore remains to quantify the contribution of each term in the above decomposition. This is the purpose of the next section.

4 Error Analysis

Let us first analyze expression (1.3.5).

- The term $E_1^{LT}(h)$ involves the local time of the Euler scheme on the boundary and the normal derivative of the function v. It is actually the leading term in the error development. We give bounds on the expectation of the local time as well as results concerning the positivity of $\frac{\partial v}{\partial n}$ in Section 4.1.

- For the term $E_1^I(h)$ we have $\mathbb{I}_{X_t\in D}$ in the integrand. Regarding its analysis, the main idea is to apply recursively Itô's formula controlling the derivatives of v. In the uniformly elliptic case we use the smoothness results given by Proposition 1.2.2. In the hypoelliptic case the key tool is the integration by parts formula from the Malliavin calculus. These facts are detailed in Section 4.3. The useful results from Malliavin calculus are given in Section 4.2.

- Controlling $E_1^O(h)$ amounts in both the uniformly and hypoelliptic case, thanks to the boundedness of $(\partial_x^\alpha v, |\alpha|\le 2)$ on the boundary, to estimate $\displaystyle\int_0^T dt\mathbb{P}_x[\tau^N > t, X_t^N \notin D]$. Since this kind of integral is also needed for the control of the local time, we upper-bound it in Section 4.1.

- The term $E_2(h)$ is associated to the last time step. As for $E_1^I(h)$ we have to distinguish the two cases. We analyze this contribution in Section 4.4

4.1 Boundary terms

We prove in this section various boundary estimates under mild assumptions, namely **(C)**, **(D-**2**)**, **(S-**1**)**, so that they are useful in both the uniformly elliptic and hypoelliptic cases. This significantly improves the previous results obtained by Gobet in the uniformly elliptic case [Gob00]. We mention that the controls obtained therein for the terms outside the domain, i.e. $E_1^O(h)$, were not sharp enough to derive our current main results.

Control of the local time

We first state a preliminary bound for the integral of the exit probability.

Lemma 1.4.1 *Under* **(C)**, **(D-**2**)**, **(S-**1**)**, *we have* $\int_0^T \mathbb{P}_x[X_t^N \notin D, \tau^N > t]dt \le C\sqrt{h}$.

Proof. Applying twice Lemma 1.1.2 first with $S'=t$, $S=\phi(t)$ and then with $S'=\phi(t)+h$, $S=t$, we easily get:

$$\begin{aligned}\int_0^T \mathbb{P}_x[t<\tau^N, X_t^N\notin D]dt \le& Ch\sum_{i=0}^{N-1}\mathbb{E}_x\left[\mathbb{I}_{\tau^N>t_i}\exp\left(-c\frac{d^2(X_{t_i}^N,\partial D)}{h}\right)\right]\\ \le& C\int_0^T\mathbb{E}_x\left[\mathbb{I}_{\tau^N>t}\exp\left(-c\frac{d^2(X_t^N,\partial D)}{h}\right)\right]dt+Ch.\end{aligned}$$

We now wish to apply the occupation times formula and use for this a localization argument. Namely, under **(C)**, **(D-**2**)**, it is clear that for some $r_0>0$ (w.l.o.g. $r_0\le R/2$):

$$\forall z\in V_{\partial D}(R/2), \forall y\in B(z,r_0), \langle\sigma\sigma^*(z)\nabla F(y),\nabla F(y)\rangle \ge a_0/2. \tag{1.4.1}$$

Thus, for $X_{\phi(t)}^N \in V_{\partial D}(R/2)$ and $X_t^N\in B(X_{\phi(t)}^N, r_0)$, we have $X_t^N\in V_{\partial D}(R)$, $d^2(X_t^N,\partial D)=F^2(X_t^N)$ and thus $d\langle F(X^N)\rangle_t = \|\sigma(X_{\phi(t)}^N)^*\nabla F(X_t^N)\|^2dt \ge a_0/2dt$. The occupation times formula gives:

$$\begin{aligned}&\int_0^T \mathbb{P}_x[t<\tau^N, X_t^N\notin D]dt\\ &\le C\int_0^T\mathbb{E}_x\left[\mathbb{I}_{\tau^N>t, X_{\phi(t)}^N\in V_{\partial D}(R/8), X_t^N\in B(X_{\phi(t)}^N, r_0/4)}\exp\left(-c\frac{d^2(X_t^N,\partial D)}{h}\right)\right]dt + Ch\\ &\le \frac{2C}{a_0}\int_{-R/4}^{R/4}dy\exp(-c\frac{y^2}{h})\mathbb{E}_x[L_{T\wedge\tau^N}^y(F(X^N))]+Ch\end{aligned}\tag{1.4.2}$$

where the discarded events in the second inequality are neglected using Lemma 1.1.2. Note that $\mathbb{E}_x[L_{T\wedge\tau^N}^y(F(X^N))]\le C$ uniformly in $y\in[-R/4,R/4]$: the proof is complete.

□

The preliminary bound from Lemma 1.4.1 helps now to prove the upper bound for the expectation of the local time.

Lemma 1.4.2 *Under* **(C)**, **(D-2)**, **(S-1)**, *we have* $\mathbb{E}_x[L^0_{T\wedge\tau^N}(F(X^N))] \le C\sqrt{h}$.

Proof. As a consequence of Tanaka's formula and after taking the expectation, we have:

$$\left|\mathbb{E}_x[\frac{1}{2}L^0_{T\wedge\tau^N}(F(X^N)) - F^-(X^N_{T\wedge\tau^N})]\right| \le C\int_0^T \mathbb{P}_x[X_t^N \notin D, \tau^N > t]dt. \tag{1.4.3}$$

Hence using Lemma 1.4.1 it remains to control the expectation of the overshoot:

$$\begin{aligned}\mathbb{E}_x[F^-(X^N_{T\wedge\tau^N})] &= \sum_{i=1}^N \mathbb{E}_x[F^-(X^N_{t_i})\mathbb{I}_{\tau^N=t_i}] = \sum_{i=1}^N \mathbb{E}_x[\mathbb{I}_{\tau^N>t_{i-1}}\mathbb{E}_{X^N_{t_{i-1}}}[F^-(X^N_{t_i})]] \\ &= \sum_{i=1}^N \mathbb{E}_x[\mathbb{I}_{\tau^N>t_{i-1}}\mathbb{I}_{X^N_{t_{i-1}}\in V_{\partial D}(R/2)}\mathbb{E}_{X^N_{t_{i-1}}}[F^-(X^N_{t_i})]] + O_{pol}(h).\end{aligned} \tag{1.4.4}$$

On the set $\{X^N_{t_{i-1}} \in V_{\partial D}(R/2)\}$ we have to upper-bound: $\mathbb{E}_{X^N_{t_{i-1}}}[F^-(X^N_{t_i})] = \mathbb{E}_{X^N_{t_{i-1}}}[\mathbb{I}_{\tau^N_{t_{i-1}}\le t_i}\mathbb{E}[F^-(X^N_{t_i}) \mid \mathcal{F}_{\tau^N_{t_{i-1}}}]]$ with $\tau^N_{t_{i-1}} := \inf\{t > t_{i-1} : X^N_t \notin D\}$. Remind that F^- is Lipschitz so $\mathbb{E}_{X^N_{t_{i-1}}}[F^-(X^N_{t_i})] \le C\sqrt{h}\mathbb{P}_{X^N_{t_{i-1}}}[\tau^N_{t_{i-1}} \le t_i]$. We conclude the proof using Lemma 1.4.3 and summing over i.

□

Lemma 1.4.3 *Under* **(C)**, **(D-2)**, **(S-1)**, *for h small enough, we have for $x \in V_{\partial D}(R/2)\cap D$:*

$$\mathbb{P}_x[\tau_0^N \le h] \le C\mathbb{P}_x[X_h^N \notin D] + O_{pol}(h).$$

Proof. We adapt ideas from [Gob00]: in the quoted paper, the uniform ellipticity condition enabled to use a Gaussian type lower bound for the transition density of X_h^N w.r.t. the Lebesgue measure, together with some computations related to a cone exterior to D. Here, under **(C)**, the law of X_h^N may degenerate and our proof relies on the scaling invariance of the cone and of the Brownian increments.
We restrict to the event $\mathcal{A} = \{X^N_{\tau_0^N} \in B(x, r_0)\}$, noting thanks to Lemma 1.1.2 that $\mathbb{P}_x[\mathcal{A}^c] = O_{pol}(h)$. Furthermore, on $\mathcal{A}$, equation (1.4.1) guarantees that

$$\alpha_{\tau_0^N} := \langle\sigma\sigma^*(x)n(X^N_{\tau_0^N}), n(X^N_{\tau_0^N})\rangle \ge a_0/2. \tag{1.4.5}$$

It is enough to prove that $a.s$ on $\mathcal{A}\cap\{\tau_0^N \le h\}$, one has

$$\mathbb{P}[X_h^N \notin D \mid \mathcal{F}_{\tau_0^N}] \ge \frac{1}{C}. \tag{1.4.6}$$

Indeed, it follows that $\mathbb{P}_x[X_h^N \notin D] \ge \mathbb{E}_x[\mathbb{I}_{\mathcal{A}}\mathbb{I}_{\tau_0^N\le h}\mathbb{P}[X_h^N \notin D \mid \mathcal{F}_{\tau_0^N}]] \ge \frac{\mathbb{P}_x[\tau_0^N\le h]}{C} + O_{pol}(h)$ and Lemma 1.4.3 is proved.
To get (1.4.6), write $X_h^N = X^N_{\tau_0^N} + \sigma(x)(W_h - W_{\tau_0^N}) + b(x)(h - \tau_0^N)$. The domain D is of class C^2, and thus satisfies a uniform exterior sphere condition with radius $R/2$ (R defined in Proposition 1.1.1): for any $z \in \partial D$, $B(z - \frac{R}{2}n(z), \frac{R}{2}) \subset D^c$. In particular, if we define for $\theta \in]0, \pi/2[$ the cone $\mathcal{K}(\theta, z) = \{y \in \mathbb{R}^d : (y - z).[-n(z)] \ge \|y - z\|\ \cos(\theta)\}$, then one has $\mathcal{K}(\theta, z)\cap B(z, R(\theta)) \subset B(z - \frac{R}{2}n(z), \frac{R}{2}) \subset D^c$ for some appropriate choice of

the *positive* function $R(.)$. Then, it follows that

$$
\begin{aligned}
&\mathbb{P}[X_h^N \notin D \mid \mathcal{F}_{\tau_0^N}] \geq \mathbb{P}[X_h^N \in \mathcal{K}(\theta, X_{\tau_0^N}^N) \cap B(X_{\tau_0^N}^N, R(\theta)) \mid \mathcal{F}_{\tau_0^N}] \\
&\geq \mathbb{P}[X_h^N \in \mathcal{K}(\theta, X_{\tau_0^N}^N) \mid \mathcal{F}_{\tau_0^N}] - \mathbb{P}[X_h^N \notin B(X_{\tau_0^N}^N, R(\theta)) \mid \mathcal{F}_{\tau_0^N}] \\
&\geq \mathbb{P}[(X_h^N - X_{\tau_0^N}^N).(-n(X_{\tau_0^N}^N)) \geq \sqrt{\alpha_{\tau_0^N}(h-\tau_0^N)} \geq \|X_h^N - X_{\tau_0^N}^N\| \cos(\theta) \mid \mathcal{F}_{\tau_0^N}] \\
&\quad - \mathbb{P}[X_h^N \notin B(X_{\tau_0^N}^N, R(\theta)) \mid \mathcal{F}_{\tau_0^N}] \geq A_1 - A_2(\theta) - A_3(\theta),
\end{aligned}
\tag{1.4.7}
$$

$$
\begin{aligned}
\text{where }\; A_1 &= \mathbb{P}[(X_h^N - X_{\tau_0^N}^N).(-n(X_{\tau_0^N}^N)) \geq \sqrt{\alpha_{\tau_0^N}(h-\tau_0^N)} \mid \mathcal{F}_{\tau_0^N}], \\
A_2(\theta) &= \mathbb{P}[\sqrt{\alpha_{\tau_0^N}(h-\tau_0^N)} < \|X_h^N - X_{\tau_0^N}^N\| \cos(\theta) \mid \mathcal{F}_{\tau_0^N}], \\
A_3(\theta) &= \mathbb{P}[X_h^N \notin B(X_{\tau_0^N}^N, R(\theta)) \mid \mathcal{F}_{\tau_0^N}].
\end{aligned}
$$

Term A_1. Clearly, one has $A_1 \geq \mathbb{P}[(-n(X_{\tau_0^N}^N)).\sigma(x)(W_h - W_{\tau_0^N}) \geq 2\sqrt{\alpha_{\tau_0^N}(h-\tau_0^N)} \mid \mathcal{F}_{\tau_0^N}] - \mathbb{P}[|n(X_{\tau_0^N}^N).b(x)(h-\tau_0^N)| \geq \sqrt{\alpha_{\tau_0^N}(h-\tau_0^N)} \mid \mathcal{F}_{\tau_0^N}] := A_{11} - A_{12}$. The random variable $(-n(X_{\tau_0^N}^N)).\sigma(x)(W_h - W_{\tau_0^N})$ is conditionally to $\mathcal{F}_{\tau_0^N}$ a centered Gaussian variable with positive variance $\alpha_{\tau_0^N}(h-\tau_0^N)$. Thus $A_{11} = \Phi(-2) > 0$. Since b is bounded, we have $A_{12} = 0$ uniformly, for h small enough.

Term $A_2(\theta)$. From Markov's inequality, $A_2(\theta) \leq \frac{\mathbb{E}[\|X_h^N - X_{\tau_0^N}^N\|^2 \cos^2(\theta) \mid \mathcal{F}_{\tau_0^N}]}{\alpha_{\tau_0^N}(h-\tau_0^N)} \leq C\cos^2(\theta)$ using (1.4.5) and estimates of Lemma 1.1.2. In particular, taking θ close to $\pi/2$ ensures that $A_2(\theta) \leq \frac{A_{11}}{4}$.

Term $A_3(\theta)$. Using Lemma 1.1.2, one readily gets $A_3(\theta) \leq C\exp\big(-c\frac{R^2(\theta)}{h}\big) \leq \frac{A_{11}}{4}$ for h small enough $(R(\theta) > 0)$.

Putting together estimates for $A_1, A_2(\theta)$ and $A_3(\theta)$ into (1.4.7) give $\mathbb{P}[X_h^N \notin D \mid \mathcal{F}_{\tau_0^N}] \geq \frac{A_{11}}{2}$. This proves (1.4.6).

□

Thanks to the previous Lemmas 1.4.1 and 1.4.2 , we now specify the behaviour of the expectation of the local time near the boundary. Namely,

Lemma 1.4.4 *Under* **(C)**, **(D-**2**)**, **(S-**1**)**, *we have for* $y \in [-R/4, R/4]$

$$\mathbb{E}_x[L^y_{T\wedge\tau^N}(F(X^N))] \leq C(|y| + h^{1/2}).$$

Proof. Tanaka's formula gives:

$$
\begin{aligned}
\mathbb{E}_x[L^y_{T\wedge\tau^N}(F(X^N))] =& 2\mathbb{E}_x[(F(X^N_{T\wedge\tau^N}) - y)^- - (F(x)-y)^-] \\
&+ 2\mathbb{E}_x[\int_0^T \mathbb{I}_{F(X_t^N)\leq y}\mathbb{I}_{\tau^N > t} d(F(X_t^N))].
\end{aligned}
\tag{1.4.8}
$$

Using Lemmas 1.4.2, 1.4.1 and estimates (1.4.3), we obtain that the first term of the r.h.s. above is upper bounded by $2(\mathbb{E}_x[F^-(X^N_{T\wedge\tau^N})] + |y|) \leq Ch^{1/2} + 2|y|$. For the other term it is enough to prove that $\omega(y) := \mathbb{E}_x[\int_0^T \mathbb{I}_{F(X_t^N)\leq y}\mathbb{I}_{\tau^N>t}dt] \leq C(\sqrt{h} + |y|)$. Since ω is increasing, for $y \leq 0$ one has $\omega(y) \leq \omega(0) \leq C\sqrt{h}$ by Lemma 1.4.1. For $y > 0$, it is enough to upper bound $\omega(y) - \omega(0)$ by $C(y + \sqrt{h})$: write $\omega(y) - \omega(0) = \mathbb{E}_x[\int_0^T \mathbb{I}_{0<F(X_t^N)\leq y}\mathbb{I}_{\tau^N>t}\mathbb{I}_{X^N_{\phi(t)}\in V_{\partial D}(R/2)}dt] + O_{pol}(h)$ using Lemma 1.1.2 (with $|y| \leq R/4$). The localization technique of Lemma 1.4.1 associated to the occupation times formula gives:

$$\mathbb{E}_x[\int_0^T \mathbb{I}_{0<F(X_t^N)\leq y}\mathbb{I}_{\tau^N>t}dt] \leq C\int_0^y du\,\mathbb{E}_x[L^u_{T\wedge\tau^N}(F(X^N))] + O_{pol}(h). \tag{1.4.9}$$

The expected local time in the above integral is uniformly bounded in $u \in [0, R/4]$, and this gives $\omega(y) - \omega(0) \leq Cy + O_{pol}(h)$.

$\square$

As a direct consequence of (1.4.2) and the above Lemma 1.4.4 one can refine the result of Lemma 1.4.1 to obtain

Lemma 1.4.5 (A sharp control for the integral of the exit probability). *Under* **(C)**, **(D-2)**, **(S-1)**, *we have* $\int_0^T \mathbb{P}_x[t<\tau^N, X_t^N \notin D]dt \le Ch$.

Let us now prove the lower bound for the expectation of the local time.

Lemma 1.4.6 *Under* **(C)**, **(D-2)**, **(H)**, **(S-1)**, *we have for h small enough (depending on $d(x,\partial D)>0$)*

$$\mathbb{E}_x[L^0_{T/2\wedge\tau^N}(F(X^N))] \ge C\sqrt{h} \tag{1.4.10}$$

with $C>0$. If the domain D is bounded the constant C is uniform w.r.t. $x\in D$. Otherwise it also depends on $d(x,\partial D)$ but is still locally uniform.

Proof. For notational convenience, we prove the result for the local time at time T instead of $T/2$. Set $L^N = \mathbb{E}_x[L^0_{T\wedge\tau^N}(F(X^N))]$; starting from (1.4.3), (1.4.4) and using Lemma 1.4.5, one has:

$$L^N \ge 2\sum_{i=1}^N \mathbb{E}_x[\mathbb{I}_{\tau^N>t_{i-1},X^N_{t_{i-1}}\in V_{\partial D}(c_0h^{1/2})}\mathbb{E}_{X^N_{t_{i-1}}}[F^-(X^N_{t_i})]] - Ch \tag{1.4.11}$$

where c_0 denotes a constant to be fixed later on. If we write $F(X^N_{t_i}) = F(X^N_{t_{i-1}}) + \nabla F(X^N_{t_{i-1}}).\sigma(X^N_{t_{i-1}})(W_{t_i}-W_{t_{i-1}})+R_i^N$, then $\mathbb{E}_{X^N_{t_{i-1}}}[|R_i^N|]\le Ch$ and thus $\mathbb{E}_{X^N_{t_{i-1}}}[F^-(X^N_{t_i})] \ge \mathbb{E}_{X^N_{t_{i-1}}}[(F(X^N_{t_{i-1}})+\nabla F(X^N_{t_{i-1}}).\sigma(X^N_{t_{i-1}})(W_{t_i}-W_{t_{i-1}}))^-] - Ch$. A direct computation gives:

$$\mathbb{E}_{X^N_{t_{i-1}}}[F^-(X^N_{t_i})] \ge \alpha(X^N_{t_{i-1}})h^{1/2} g\left(\frac{F(X^N_{t_{i-1}})}{\alpha(X^N_{t_{i-1}})h^{1/2}}\right) - Ch \tag{1.4.12}$$

where $g(z) := \dfrac{\exp(-\frac{z^2}{2})}{(2\pi)^{1/2}} - z\Phi(-z)$ is a positive decreasing function on $\mathbb{R}^+$. Note that for h small enough ($c_0h^{1/2}\le R$), one has $\alpha(x)\ge a_0$ for $x\in V_{\partial D}(c_0h^{1/2})$ (Assumption **(C)**); thus, plugging (1.4.12) into (1.4.11) it comes:

$$L^N \ge 2a_0h^{1/2}\sum_{i=1}^N \mathbb{E}_x\left[\mathbb{I}_{\tau^N>t_{i-1}}\mathbb{I}_{F(X^N_{t_{i-1}})\in(0,c_0h^{1/2}]} g\left(\frac{F(X^N_{t_{i-1}})}{a_0h^{1/2}}\right)\right] - S(h)$$

with $S(h) = Ch\sum_{i=1}^N \mathbb{P}_x\left[\tau^N>t_{i-1}, F(X^N_{t_{i-1}})\in(0,c_0h^{1/2}]\right]$. Assume for a while that $S(h)\le Ch^{3/4}$ and consider the other contribution. Use that $\forall i\in[\![2,N]\!]$, $t\in[t_{i-2},t_{i-1}]$, $\mathbb{I}_{t<\tau^N}\le \mathbb{I}_{t_{i-1}<\tau^N}$, and that g does not vanish on the compact sets of $\mathbb{R}^+$, to obtain:

$$\begin{aligned}
&\mathbb{E}_x\left[\mathbb{I}_{\tau^N>t_{i-1}}\mathbb{I}_{F(X^N_{t_{i-1}})\in(0,c_0h^{1/2}]} g\left(\frac{F(X^N_{t_{i-1}})}{a_0h^{1/2}}\right)\right]\\
&\ge C_1\mathbb{E}_x\left[\mathbb{I}_{\tau^N>t}\mathbb{I}_{F(X^N_t)\in[c_0h^{1/2}/4,3c_0h^{1/2}/4]} g\left(\frac{F(X^N_t)}{a_0h^{1/2}}\right)\mathbb{P}\left[F(X^N_{t_{i-1}})\in(0,c_0h^{1/2}]|\mathcal{F}_t\right]\right].
\end{aligned}$$

where $C_1>0$. On $\{F(X^N_t)\in[c_0h^{1/2}/4,3c_0h^{1/2}/4]\}$, we easily conclude by Lemma 1.1.2: $\mathbb{P}[F(X^N_{t_{i-1}})\notin(0,c_0h^{1/2}]|\mathcal{F}_t]\le C\exp(-cc_0^2/16)$, so that $\mathbb{P}[F(X^N_{t_{i-1}})\in(0,c_0h^{1/2}]|\mathcal{F}_t]\ge 1/2$ for c_0 large enough. We have obtained:

$$\begin{aligned}
L^N &\ge a_0C_1h^{-1/2}\int_0^{T-h}\mathbb{E}_x[\mathbb{I}_{\tau^N>t}\mathbb{I}_{F(X^N_t)\in[c_0h^{1/2}/4,3c_0h^{1/2}/4]}g(\frac{F(X^N_t)}{a_0h^{1/2}})]dt - Ch^{3/4}\\
&\ge \frac{a_0C_1h^{-1/2}}{\|\sigma^*\|^2_\infty\|\nabla F\|^2_\infty}\int_{c_0h^{1/2}/4}^{3c_0h^{1/2}/4} g\left(\frac{y}{a_0h^{1/2}}\right)\mathbb{E}_x[L^y_{(T-h)\wedge\tau^N}(F(X^N))]dy - Ch^{3/4}\\
&\ge \frac{a_0C_1h^{-1/2}}{\|\sigma^*\|^2_\infty\|\nabla F\|^2_\infty}\int_{c_0h^{1/2}/4}^{3c_0h^{1/2}/4} g\left(\frac{y}{a_0h^{1/2}}\right)\mathbb{E}_x[L^y_{T/2\wedge\tau^N}(F(X^N))]dy - Ch^{3/4}
\end{aligned}$$

where the occupation times formula is once again the key tool for the last but one inequality (we do not need here the rather tedious localization procedure of Lemmas 1.4.1 and 1.4.5, and only use σ's boundedness). The last inequality is a direct consequence of the increasing property in time of the local time.

Lemma 1.4.7 below and equation (1.4.8), written with $T/2$ instead of T, then yield: $\mathbb{E}_x[L^y_{T/2\wedge\tau^N}(F(X^N))] = 2(\mathbb{E}_x[(F(X^N_{T/2\wedge\tau^N})-y)^-] - (F(x)-y)^-) - Ch = 2\mathbb{E}_x[(F(X^N_{T/2\wedge\tau^N})-y)^-] - Ch$ for y small enough (namely $y \le F(x)$). This explains the dependence in $d(x,\partial D)$ for the step size h.

If we put $C_2 = \frac{2a_0C_1}{\|\sigma^*\|^2_\infty\|\nabla F\|^2_\infty} > 0$, it follows that

$$\begin{aligned}
L^N \ge& C_2 h^{-1/2}\int_{c_0h^{1/2}/4}^{3c_0h^{1/2}/4} g\left(\frac{y}{a_0h^{1/2}}\right)\mathbb{E}_x[(F(X^N_{T/2\wedge\tau^N})-y)^-]dy - Ch^{3/4}\\
\ge& C_2\int_{c_0/4}^{3c_0/4} g\left(\frac{z}{a_0}\right)\mathbb{E}_x[(zh^{1/2}-F(X^N_{T/2\wedge\tau^N}))\mathbb{I}_{zh^{1/2}>F(X^N_{T/2\wedge\tau^N})}]dz - Ch^{3/4}\\
\ge& C_2h^{1/2}(\int_{c_0/4}^{3c_0/4} zg(z/a_0)dz)\mathbb{P}_x[\tau^N \le T/2] - Ch^{3/4}
\end{aligned}$$

noting that $(zh^{1/2}-F(X^N_{T/2\wedge\tau^N}))\mathbb{I}_{zh^{1/2}>F(X^N_{T/2\wedge\tau^N})} \ge zh^{1/2}\mathbb{I}_{0\ge F(X^N_{T/2\wedge\tau^N})} = zh^{1/2}\mathbb{I}_{\tau^N\le T/2}$. To conclude the proof, note that $\mathbb{P}_x[\tau^N \le T/2] \ge \mathbb{P}_x[X^N_{T/2} \notin D]$ which converges pointwise (see [BT96a]) in $x \in D$ to $\mathbb{P}_x[X_T \notin D]$. Under **(H)**, this last quantity is positive (see [BL91]).

If the domain D is bounded the previous convergence is actually uniform (see Theorem 3.1 in [BT96a]) and $\mathbb{P}_x[X_{T/2} \notin D] \ge \inf_{y\in\bar{D}}\mathbb{P}_y[X_{T/2}\notin D] = C > 0$. Indeed, under **(H)**, $f_D(x) := \mathbb{P}_x[X_{T/2}\notin D]$ defines a continuous function. We take its infimum over a compact set. This infimum is positive.

If the domain D is not bounded the constant C in (1.4.10) depends on $d(x,\partial D)$. Anyhow, from Theorem 3.1 in [BT96a] we get that this constant can be made uniform for every given neighbourhood $V(x)\subset D$ of x.

It remains to estimate $S(h)$. For this, remark that for $\forall i \in [\![2,N]\!]$, $t_{i-2}\le t < t_{i-1}$

$$\mathbb{P}_x\left[\tau^N > t_{i-1}, F(X^N_{t_{i-1}}) \in (0, c_0h^{1/2}]\right] \le \mathbb{P}_x\left[\tau^N > t, F(X^N_t) \in (-c_0h^{3/8}, c_0h^{3/8}]\right] + O_{pol}(h).$$

This provides the way to transform the sum over i in an integral over t. We conclude using Lemma 1.4.7.

□

Lemma 1.4.7 *Under* **(C)**, **(D-2)**, **(S-1)**, *we have for* $y \le R/4$

$$\int_0^T \mathbb{P}_x[F(X^N_t)\le y, \tau^N > t]dt \le C(h+y^2).$$

Proof. The contribution associated to $y \le 0$ is already controlled by Lemma 1.4.5. For $y \in (0, R/4]$, by (1.4.9), write $\int_0^T \mathbb{P}_x[F(X^N_t)\le y, \tau^N > t]dt \le C\int_0^y du\mathbb{E}_x[L^u_{T\wedge\tau^N}(F(X^N))] + Ch \le C(h + y\sqrt{h} + y^2)$ using Lemma 1.4.4.

□

4.2 Reminders on Malliavin Calculus

We briefly recall in this section elementary aspects of Malliavin calculus and present the main tools that will be needed in our analysis. Namely

- The integration by parts formula, intensively used in the next two sections for the hypoelliptic case.

- The local criterion of absolute continuity for the law of a one-dimensional random variable. This will be used to show a convergence result for $\mathrm{Err}(T,h,f,x)$ under weak assumptions, cf. Section 5.

The various results are stated without proof. We refer to the monograph of Nualart, [Nua95], for details.

Notations and definitions

Let W be a standard d'-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Consider the Hilbert space $H = \mathrm{L}_2([0,T], \mathbb{R}^{d'})$. For $h \in H$ we denote $W(h)$ the Wiener-integral $\int_0^T h(t) \cdot dW_t$.

The space $\mathcal{S}$ of simple functionals is the space of random variables of the form $F = f(W(h_1), ..., W(h_n))$ where $f \in C_p^\infty(\mathbb{R}^n), (h_1, ..., h_n) \in H^n, n \geq 1$. For a given simple functional F we define the derivative operator D as the H-valued random variable that writes

$$DF = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(h_1), ..., W(h_n))h_i.$$

As an operator from $\mathrm{L}_p(\Omega)$ into $\mathrm{L}_p(\Omega, H)$, D is closeable for $p \geq 1$. Let us denote $\mathbb{D}^{1,p}$ its domain with the norm

$$\|F\|_{1,p} = (\mathbb{E}[|F|^p] + \mathbb{E}[|DF|_H^p])^{1/p}.$$

For $F \in \mathcal{S}$ and $k \in \mathbb{N}, k > 1$ one defines the k^{th} derivative, denoted $D^k F$, iterating k times the operator D. The variable $D^k F$ is a $H^{\otimes k}$ valued random variable. As for $k = 1$, the operator D^k is closeable from $\mathrm{L}_p(\Omega)$ into $\mathrm{L}_p(\Omega, H^{\otimes k})$. Let us denote $\mathbb{D}^{k,p}$ its domain with the norm

$$\|F\|_{k,p} = \left(\mathbb{E}[|F|^p] + \sum_{i=1}^k \mathbb{E}[|D^i F|_{H^{\otimes i}}^p]\right)^{1/p}.$$

We denote $\mathbb{D}^\infty = \cap_{p\geq 1, k\geq 1} \mathbb{D}^{k,p}$.

For $F \in (\mathbb{D}^\infty)^m$ the notation γ_F stands for the Malliavin covariance matrix of F. This a $(m \times m)$ matrix defined by

$$\forall (i,j) \in [\![1, m]\!]^2, \ \gamma_F^{i,j} = \langle D^i F, D^j F \rangle_H.$$

The inverse of γ_F, when it exists, is denoted by Γ_F.

Integration by parts and associated controls

Let $G \in \mathbb{D}^\infty$, and $F \in (\mathbb{D}^\infty)^d$ satisfy the non degeneracy condition $\Gamma_F \in \bigcap_{p\geq 1} \mathrm{L}_p$. For any smooth function φ with polynomial growth and any multi-index α there exists a family of random variables H_α s.t.

$$\mathbb{E}[\partial^\alpha \varphi(F) G] = \mathbb{E}[\varphi(F) H_\alpha(F, G)]. \tag{1.4.13}$$

Furthermore for any $p > 1$ and any multi-index α, there exist a constant C and integers $(q_i)_{i\in[\![1,5]\!]}, (j_i)_{i\in[\![1,2]\!]}$ depending on p, α s.t. for any measurable set $A \subset \Omega$ and F, G as above, one has

$$\mathbb{E}[|H_\alpha(F,G)|^p \mathbb{I}_A]^{1/p} \leq C \|\Gamma_F \mathbb{I}_A\|_{\mathrm{L}_{q_3}}^{q_4} \|F\|_{\mathbb{D}^{j_1, q_1}}^{q_5} \|G\|_{\mathbb{D}^{j_2, q_2}}. \tag{1.4.14}$$

This will help to establish a "local" version of (1.4.13) as stated in Proposition 4.4 in [BT96b]. Assume $G \in \mathbb{D}^\infty$, $F \in (\mathbb{D}^\infty)^d$ and satisfy the partial non-degeneracy condition $\mathbb{I}_{G\neq 0 \bigcup_{|\beta|\leq|\alpha|} D^\beta G \neq 0} \Gamma_F \in \bigcap_{p\geq 1} \mathrm{L}_p$. From Proposition 4.4 in [BT96b], (1.4.14) and the Cauchy-Schwarz inequality we get

$$|\mathbb{E}[\partial^\alpha \varphi(F) G]| \leq C \|\varphi(F)\|_{\mathrm{L}_2} \|\Gamma_F \mathbb{I}_{G\neq 0 \bigcup_{|\beta|\leq|\alpha|} D^\beta G \neq 0}\|_{\mathrm{L}_{q_3}}^{q_4} \|F\|_{\mathbb{D}^{j_1, q_1}}^{q_5} \|G\|_{\mathbb{D}^{j_2, q_2}} \tag{1.4.15}$$

for some universal constants (depending on α) and for any smooth function φ with polynomial growth.

Inequality (1.4.15) is a crucial tool in Sections 4.3 and 4.4.

A criterion for (local) absolute continuity of one dimensional random variable

The following theorem puts together Theorem 2.1.3 and Corollary 2.1.1 from Nualart, [Nua95], under stronger assumptions. It concerns one dimensional random variables.

Practically, we will apply this result to the minimum of the distance of the process to the boundary in order to show the convergence of $\mathrm{Err}(T,h,f,x)$ under weak assumptions, cf. Proposition 1.5.1 in Section 5.

Theorem 1.4.8 *Let F be a one dimensional random variable in $\mathbb{D}^{1,2}$. Suppose $\|DF\|_H > 0$ a.s. Then the law of F is absolutely continuous w.r.t. the Lebesgue measure of $\mathbb{R}$.*

From a local viewpoint, the measure $(\|DF\|_H.\mathbb{P})\circ F^{-1}$ is absolutely continuous w.r.t. the Lebesgue measure of $\mathbb{R}$.

4.3 Interior estimates

In this section we give estimates concerning the term $E_1^I(h)$ in (1.3.5).

Hypoelliptic case

In the whole paragraph we work under the assumptions **(C)**, **(D-∞)**, **(H')**, **(F1)**, **(S-∞)**.

Equation (1.2.3) gives a point-wise estimate on the derivatives of v close to the boundary. Lemma 1.4.9 below gives controls of the expectations of these derivatives far from the boundary.

From now on we denote by ψ a cut-off function near ∂D such that:

$$\psi \in C_b^\infty(\mathbb{R}^d,\mathbb{R}),\ \mathbb{I}_{V_{\partial D}(\varepsilon/2)} \le 1-\psi \le \mathbb{I}_{V_{\partial D}(\varepsilon)} \text{ and } \|\partial^\alpha \psi\|_\infty \le \frac{\mathrm{C}_{|\alpha|}}{1\wedge\varepsilon^\alpha}$$

for all multi-index α.

Lemma 1.4.9 (Expectation of the derivatives "far" from the boundary). *Under* **(C)**, **(D-∞)**, **(F1)**, **(H')**, **(S-∞)**, *for all multi-indices α, α', all function $g \in C_b^\infty(\mathbb{R}^d,\mathbb{R})$, there exist constants ξ and C such that $\forall i \in [\![0,N-1]\!], t\in[t_i,T-h]$:*

$$\left|\mathbb{E}_x\left[\mathbb{I}_{t<\tau^N\wedge\tau_R^N} g(X_t^N)\partial_x^\alpha(v\partial^{\alpha'}\psi)(t,X_t^N)\right]\right| \le C\frac{\|f\|_\infty}{1\wedge\varepsilon^\xi},$$
$$\left|\mathbb{E}_x\left[\mathbb{I}_{t<\tau^N\wedge\tau_R^N} g(X_{t_i}^N)\partial_x^\alpha(v\partial^{\alpha'}\psi)(t,X_t^N)\right]\right| \le C\frac{\|f\|_\infty}{1\wedge\varepsilon^\xi}.$$

We postpone the proof of this Lemma to the end of the section. We now state the main result of this paragraph.

Lemma 1.4.10 *Under* **(C)**, **(D-∞)**, **(F1)**, **(H')**, **(S-∞)**, *one has*

$$E_1^I(h) = \mathbb{E}_x\left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} \mathbb{I}_{X_t^N\in D}\left((\mathcal{L}_{X^N_{\phi(t)}} - L)v(t,X_t^N)\right)dt\right] = O(h).$$

Proof of Lemma 1.4.10.

Using ψ we isolate the values of X_t^N, $t\in[0,T-h]$ that are "far" from the boundary and those that are close to it. It comes

$$\begin{aligned}
E_1^I(h) &= \mathbb{E}_x\left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} dt\mathbb{I}_{X_t^N\in D}\left(-Lv+\mathcal{L}_z v\right)|_{z=X^N_{\phi(t)}}(t,X_t^N)\right] = \\
&\quad \mathbb{E}_x\left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} dt\mathbb{I}_{X_t^N\in D}\left(-L+\mathcal{L}_z\right)|_{z=X^N_{\phi(t)}}((1-\psi)v)(t,X_t^N)\right] \\
&\quad +\mathbb{E}_x\left[\int_0^{(T-h)\wedge\tau^N\wedge\tau_R^N} dt\mathbb{I}_{X_t^N\in D}\left(-L+\mathcal{L}_z\right)|_{z=X^N_{\phi(t)}}(\psi v)(t,X_t^N)\right] := A_C + A_F.
\end{aligned}$$

Control of A_C:
One can extend the solution v of problem (1.2.1) into a function $\bar v$ defined on the whole space so that the derivatives of $\bar v$ on D^C have the same properties as those of v on $V_{\partial D}(\varepsilon)$. For a justification of this statement see e.g. Theorem 4.1 Chapter 4 from [LSU68]. Hence, $\forall (t,x)\in[0,T-h]\times D^C\cup V_{\partial D}(\varepsilon)$ the controls of equation (1.2.3) hold true for $\bar v$.

It comes

$$\begin{aligned}
A_C &= \int_0^{T-h} dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X^N_{\phi(t)}}[\mathbb{I}_{\tau^N_R>t}\mathbb{I}_{X^N_t\in D}(-L+\mathcal{L}_z)\mid_{z=X^N_{\phi(t)}}((1-\psi)\bar v)(t,X^N_t)]]\\
&= \int_0^{T-h} dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X^N_{\phi(t)}}[(-L+\mathcal{L}_z)\mid_{z=X^N_{\phi(t)}}((1-\psi)\bar v)(t,X^N_t)]]+R_C=:A^1_C+R_C
\end{aligned}$$

where

$$\begin{aligned}
|R_C| \le &\left|\int_0^{T-h}\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X^N_{\phi(t)}}[\mathbb{I}_{\tau^N_R\le t}(L-\mathcal{L}_z)\mid_{z=X^N_{\phi(t)}}((1-\psi)\bar v)(t,X^N_t)]]\right|\\
&+\left|\int_0^{T-h}dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X^N_{\phi(t)}}[\mathbb{I}_{X^N_t\notin D}(L-\mathcal{L}_z)\mid_{z=X^N_{\phi(t)}}((1-\psi)\bar v)(t,X^N_t)]]\right|\\
&+\left|\int_0^{T-h}dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X^N_{\phi(t)}}[\mathbb{I}_{\tau^N_R\le t}\mathbb{I}_{X^N_t\notin D}(L-\mathcal{L}_z)\mid_{z=X^N_{\phi(t)}}((1-\psi)\bar v)(t,X^N_t)]]\right| := R^1_C+R^2_C+R^3_C.
\end{aligned}$$

Recalling $\mathrm{supp}(1-\psi)\subset V_{\partial D}(\varepsilon)$ we get

$$\exists C>0,\ \beta>0,\ \forall\alpha,\ |\alpha|\le 4,\ \forall(t,x)\in[0,T-h]\times\mathbb{R}^d,\ |\partial^\alpha((1-\psi(x))v(t,x))|\le\frac{C}{1\wedge\varepsilon^\beta}. \tag{1.4.16}$$

For R^1_C, equation (1.4.16) and Lemma 1.1.2 give $R^1_C=O_{pol}(h)$. From (1.4.16) and Lemma 1.4.5 we also derive $R^2_C\le\frac{Ch}{1\wedge\varepsilon^\beta}$. The term R^3_C can be handled just as R^1_C. Hence, $R_C=O(h)$.

Regarding A^1_C, let us write

$$A^1_C=\int_0^{T-h}dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\int_{\phi(t)}^t ds[\partial_s+\mathcal{L}_{X^N_{\phi(t)}}]\left(-L+\mathcal{L}_{X^N_{\phi(t)}}\right)[(1-\psi)\bar v](s,X^N_s)].$$

We note that $(\partial_t+\mathcal{L}^N_{X_{\phi(t)}})(-L+\mathcal{L}_{X^N_{\phi(t)}})[(1-\psi)\bar v](s,X^N_s)$ is a finite sum of terms that write

$$\partial^\alpha_x[(1-\psi)\bar v(s,X^N_s)]g(X^N_u),\ u\in\{s,\phi(s)\},\ |\alpha|\le 4$$

where g is bounded and only depends on the coefficients b,σ of the diffusion. Equation (1.4.16) readily gives $|A^1_C|\le\frac{Ch}{1\wedge\varepsilon^\beta}$. We finally obtain

$$|A_C|\le\frac{Ch}{1\wedge\varepsilon^\zeta}+O_{pol}(h). \tag{1.4.17}$$

Control of A_F:

$$\begin{aligned}
A_F &= \int_0^{T-h}dt\mathbb{E}_x[\mathbb{I}_{t<\tau^N\wedge\tau^N_R}\mathbb{I}_{X^N_t\in D}\left(-L+\mathcal{L}_{X^N_{\phi(t)}}\right)[\psi v](t\wedge\tau^N\wedge\tau^N_R,X^N_{t\wedge\tau^N\wedge\tau^N_R})]\\
&= \int_0^{T-h}dt\mathbb{E}_x[\mathbb{I}_{t<\tau^N\wedge\tau^N_R}\left(-L+\mathcal{L}_{X^N_{\phi(t)}}\right)[\psi v](t\wedge\tau^N\wedge\tau^N_R,X^N_{t\wedge\tau^N\wedge\tau^N_R})]\\
&= \int_0^{T-h}dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N\wedge\tau^N_R}\int_{\phi(t)}^t ds\mathbb{I}_{s<\tau^N\wedge\tau^N_R}[\partial_s+\mathcal{L}_{X^N_{\phi(t)}}]\left(-L+\mathcal{L}_{X^N_{\phi(t)}}\right)[\psi v](s,X^N_s)].
\end{aligned}$$

As for A_c we write:

$$(\partial_s+\mathcal{L}_{X^N_{\phi(s)}})(-L+\mathcal{L}_{X^N_{\phi(s)}})(\psi v)(s,X^N_s)=\sum_{|\alpha|\le4,|\alpha|+|\alpha'|\le4,u\in\{s,\phi(s)\}}g_{\alpha,\alpha'}(X^N_u)\partial^\alpha_x[v\partial^{\alpha'}\psi](s,X^N_s)$$

where the functions $g_{\alpha,\alpha'}$ only depend on b, σ and their derivatives. Lemma 1.4.9 yields:

$$|\mathbb{E}_x[\mathbb{I}_{s<\tau^N\wedge\tau_R^N}(\partial_s+\mathcal{L}_z)(-L+\mathcal{L}_z)(\psi v)(s,X_s^N)]| \quad \leq \quad \frac{\|f\|_\infty}{1\wedge\varepsilon^\zeta}\frac{C}{T^\zeta}.$$

Thus,

$$|A_F| \leq \frac{C\|f\|_\infty}{1\wedge\varepsilon^\zeta}h. \tag{1.4.18}$$

Equations (1.4.18), (1.4.17) give the result.

□

We conclude this paragraph giving the proof of Lemma 1.4.9.

Proof of Lemma 1.4.9.

We prove the first identity stated in the Lemma. One could derive the second one in a similar way. From (1.2.2) and the definition of ψ we get that for $t < T/2$:

$$\left|\mathbb{E}_x\left[\mathbb{I}_{t<\tau^N\wedge\tau_R^N}g(X_t^N)\partial_x^\alpha(v\partial^{\alpha'}\psi)(t,X_t^N)\right]\right| \leq \frac{\|f\|_\infty C}{1\wedge\varepsilon^{|\alpha'|+|\alpha|}}\frac{1}{T^\zeta}.$$

For $t \geq T/2$, following an idea from Cattiaux, cf. [Cat91], we write:

$$\begin{aligned}
&\mathbb{E}_x\left[\mathbb{I}_{t<\tau^N\wedge\tau_R^N}g(X_t^N)\partial_x^\alpha(v\partial^{\alpha'}\psi)(t,X_t^N)\right] = \\
&\mathbb{E}_x\left[g(X_t^N)\partial_x^\alpha(v\partial^{\alpha'}\psi)(t,X_t^N)\right] - \mathbb{E}_x\left[\mathbb{I}_{t\geq\tau^N\wedge\tau_R^N}g(X_t^N)\partial_x^\alpha(v\partial^{\alpha'}\psi)(t,X_t^N)\right] \\
&:= A_{US} + A_C.
\end{aligned}$$

The term A_{US} is "usual". Lemma 4.1 in [BT96a] gives the expected result. For A_C we compute a "conditional" Malliavin calculus after the exit time. As a consequence of assumption **(F1)** we also use Bernstein like arguments, cf. Lemma 1.1.2. We point out that the support condition is here crucial.

Put: $\bar\tau := \tau^N \wedge \tau_R^N, \tilde t_0 = 0, \tilde t_1 = \phi(\bar\tau) + h - \bar\tau,\ \forall i > 1, \tilde t_i = \tilde t_{i-1} + h$. We define the process $(Y_s^N)_{s\geq 0}$:

$$\begin{aligned}
Y_0^N &= X_{\bar\tau}^N, \\
\forall s \in [0,\tilde t_1],\ Y_s^N &= Y_0^N + b(X_{\phi(\bar\tau)}^N)s + \sigma(X_{\phi(\bar\tau)}^N)(W_{\bar\tau+s} - W_{\bar\tau})
\end{aligned} \tag{1.4.19}$$

and for $s \in [\tilde t_k, \tilde t_{k+1}],\ k \geq 1$, Y^N is defined as the usual Euler scheme. Put $(v\partial^{\alpha'}\psi)(.,.) := U(.,.)$. Conditioning in A_C w.r.t. $\mathcal{F}_{\bar\tau}$, for $u := t - \bar\tau$ it comes

$$A_C \quad = \quad \mathbb{E}_x[\mathbb{I}_{\bar\tau\leq t}\mathbb{E}_{X_{\bar\tau}^N, X_{\phi(\bar\tau)}^N}[\partial^\alpha U(t,Y_u^N)g(Y_u^N)]].$$

The two indices in $\mathbb{E}_{X_{\bar\tau}^N, X_{\phi(\bar\tau)}^N}$ respectively refer to the initial value of the process Y and to the value taken into consideration in the piecewise constant approximation of the drift and diffusion terms for the first time step, see (1.4.19).

In the following we discuss the case $u > 0$ since if $u = 0$ the definition of ψ yields $A_C = 0$. Let us denote by γ_u the Malliavin matrix of the diffusion $(Y_v)_{v\geq 0}$ starting from $X_{\bar\tau}^N$ at time 0. We write $\Gamma_u := \gamma_u^{-1}$, and $\hat\gamma_u := \det(\gamma_u)$. The same notations indexed by N stand for Y^N.

Following the localization techniques from Bally and Talay, [BT96a], we partition $\{\bar\tau < t\}$ into $\Omega^D(u) \bigcup \Omega^S(u)$ where $\Omega^D(u) := \{\omega \in \Omega : \bar\tau < t, |\hat\gamma_u^N - \hat\gamma_u| > \hat\gamma_u/2\}$, and $\Omega^S(u) = \{\omega \in \Omega : \bar\tau < t, |\hat\gamma_u^N - \hat\gamma_u| \leq \hat\gamma_u/2\}$.

We show $\mathbb{P}[\Omega^D(u)]$ is small enough to compensate the explosive terms in the derivatives of the Green kernel. For this we use two facts. First, the Euler scheme is a good approximation of the diffusion. Second, under our assumptions the norms $\|\hat\gamma_u^{-1}\|_{\mathrm{L}_p}$ are finite. On $\Omega^S(u)$ the Malliavin covariance matrix of the Euler scheme is somehow "close" to the one of the diffusion. In this case we use integrations by parts.

Let us introduce an even function $\eta \in C_b^\infty(\mathbb{R})$ s.t. for $x \geq 0$, $\mathbb{I}_{[0,1/4]} \leq \eta \leq \mathbb{I}_{[0,1/2]}$. Put $r_u^N := \dfrac{(\hat\gamma_u^N - \hat\gamma_u)}{\hat\gamma_u}$. We have

$$\begin{aligned}
A_C \quad &= \quad \mathbb{E}_x[\mathbb{I}_{\bar\tau<t}\mathbb{E}_{X_{\bar\tau}^N, X_{\phi(\bar\tau)}^N}[\partial^\alpha U(t,Y_u^N)g(Y_u^N)\eta(r_u^N)]] + \\
&\qquad \mathbb{E}_x[\mathbb{I}_{\bar\tau<t}\mathbb{E}_{X_{\bar\tau}^N, X_{\phi(\bar\tau)}^N}[\partial^\alpha U(t,Y_u^N)g(Y_u^N)(1-\eta(r_u^N))]] \\
&:= \quad A_C^{IBP} + A_C^D.
\end{aligned}$$

Under the assumptions of the Lemma we recall, cf. [BT96a], that

$$\forall p > 1, \forall j \geq 1, \exists Q \in \mathbb{N},\ \sup_N \|Y_u^{N,x}\|_{\mathbb{D}^{j,p}} < C(1+\|x\|^Q),\ \text{and}\ \sup_N \|Y_u^x - Y_u^{x,N}\|_{\mathbb{D}^{j,p}} \leq C\sqrt{h}.$$

$$\text{Furthermore, } \exists L \in \mathbb{N},\ \forall u > 0,\ \|\hat{\gamma}_u^{-1}\|_{\mathrm{L}_p} < \frac{C}{u^{dL}}. \tag{1.4.20}$$

Control of A_C^D

$$A_C^D = \mathbb{E}_x\left[\mathbb{I}_{\bar{\tau}<t}\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}\left[g(Y_u^N)\partial_x^\alpha U(t, Y_u^N)(1-\eta(r_u^N))(\mathbb{I}_{Y_u^N \in V_{\partial D}(\varepsilon/2)} + \mathbb{I}_{Y_u^N \in D\backslash V_{\partial D}(\varepsilon/2)})\right]\right] := A_C^{D,1} + A_C^{D,2}.$$

Since $\mathrm{supp}(\psi) \subset D\backslash V_{\partial D}(\varepsilon/2)$, we get $A_C^{D,1} = 0$. From Proposition 1.2.1 and Lemma 1.1.2 we obtain

$$\begin{aligned}
|A_C^D| &\leq \mathbb{E}_x\left[\mathbb{I}_{\bar{\tau}<t}\mathbb{E}^N_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}}\left[|g(Y_u^N)\partial_x^\alpha U(t, Y_u^N)|(1-\eta(r_u^N))\mathbb{I}_{Y_u^N \in D\backslash V_{\partial D}(\varepsilon/2)}\right]\right] \\
&\leq \mathbb{E}_x[\mathbb{I}_{\bar{\tau}<t}\exp\left(-c\frac{\varepsilon^2}{u}\right)\frac{C\|f\|_\infty}{(1\wedge\varepsilon^{|\alpha'|+|\alpha|})(T-t)^q}\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[|1-\eta(r_u^N)|^2]^{1/2}].
\end{aligned} \tag{1.4.21}$$

Recalling $|1-\eta(r_u^n)| \leq 1$ we write:

$$\begin{aligned}
\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[|1-\eta(r_u^N)|] &\leq \mathbb{P}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[|r_u^N| \geq 1/4] \\
&\leq \mathbb{P}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[\hat{\gamma}_u \leq N^{-1/4}] + \mathbb{P}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[|\hat{\gamma}_u^N - \hat{\gamma}_u| \geq 1/4N^{-1/4}] \\
&\leq N^{-p/4}\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[|\hat{\gamma}_u|^{-p}] + (4N^{1/4})^p\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[|\hat{\gamma}_u - \hat{\gamma}_u^N|^p] \\
&\leq N^{-p/4}\left(\frac{C}{u^{dLp}} + C\right)
\end{aligned}$$

where the last inequality is a consequence of (1.4.20). Plugging this control into (1.4.21) it comes:

$$|A_C^D| \leq \frac{C\|f\|_\infty}{(1\wedge\varepsilon^{|\alpha'|+|\alpha|})(T-t)^q}N^{-p/8}\mathbb{E}_x[\mathbb{I}_{\bar{\tau}<t}\exp\left(-C\frac{\varepsilon^2}{u}\right)u^{-dLp/2}]. \tag{1.4.22}$$

Note that for $t \leq T-h, (T-t)^{-q} \leq CN^q$. Hence, taking $p = 8q$ we deduce from (1.4.22)

$$|A_C^D| \leq \frac{\|f\|_\infty C}{1\wedge\varepsilon^{|\alpha'|+|\alpha|+\zeta(p)}}(N^{q-p/8}) \leq \frac{\|f\|_\infty C}{1\wedge\varepsilon^\zeta}. \tag{1.4.23}$$

Control of A_C^{IBP}.

To control this term we introduce a perturbed process $(Y^{N,\delta})_{s\geq 0} := (Y^N + \delta\tilde{W})_{s\geq 0}$, where $\tilde{W}$ is a d-dimensional BM independent from W and $\delta \in [0,1]$. As before $\gamma_s^{N,\delta}, \hat{\gamma}_s^{N,\delta}, \Gamma_s^{N,\delta}$ respectively denote the Malliavin covariance matrix, the determinant, the inverse of the Malliavin covariance matrix of $Y^{N,\delta}$ at time s. The process $Y^{N,\delta}$ satisfies the non-degeneracy condition $\Gamma_s^{N,\delta} \in \cap_{p\geq 1}\mathrm{L}_p$ for all $\delta > 0,\ s > 0$. Actually, one has $\|\Gamma_s^{N,\delta}\|_p \leq C(s^{1/2}\delta)^{-2d}$. This perturbation allows us to integrate by parts to get rid of the exploding derivatives of v. We show that all our estimates are uniform in δ and then take the limit. Let us point out that with the above perturbed process we do not need like in [BT96a] to make a time extension of the domain. The same perturbation technique is also used in the next paragraph to analyze the contribution of the last time step in the error.

Put

$$A_C^{IBP,\delta} := \mathbb{E}_x[\mathbb{I}_{t>\bar{\tau}}\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[\partial^\alpha U(t, Y_u^{N,\delta})g(Y_u^{N,\delta})\eta(r_u^{N,\delta})]].$$

The integration by parts formula (1.4.13) gives

$$\begin{aligned}
A_C^{IBP,\delta} &= \mathbb{E}_x[\mathbb{I}_{t>\bar{\tau}}\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[U(t, Y_u^{N,\delta})H_\alpha(Y_u^{N,\delta}, g(Y_u^{N,\delta})\eta(r_u^{N,\delta}))],\\
|A_C^{IBP,\delta}| &\leq \mathbb{E}_x[\mathbb{I}_{t>\bar{\tau}}\left(\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[U(t, Y_u^{N,\delta})^2]\right)^{1/2}\left(\mathbb{E}_{X_{\bar{\tau}}^N, X_{\phi(\bar{\tau})}^N}[|H_\alpha(Y_u^{N,\delta}, g(Y_u^{N,\delta})\eta(r_u^{N,\delta}))|^2]\right)^{1/2}].
\end{aligned} \tag{1.4.24}$$

Following the proof of the first inequality of Lemma 1.1.2 for the perturbed process we obtain

$$\forall \nu \geq 0,\ \mathbb{P}[\sup_{s\in[0,u]} \|Y_s^{N,\delta} - Y_0^N\| \geq \nu] \leq C \exp\left(-c\frac{\nu^2}{(\|\sigma\sigma^*\|_\infty + d\delta^2)u}\right)$$

which is uniform w.r.t. $\delta \in [0,1]$. Since $U(t, Y_u^{N,\delta}) = 0$ for $\|Y_u^{N,\delta} - Y_0^{N,\delta}\| \leq \varepsilon/2$, we have

$$\left(\mathbb{E}_{X_{\bar\tau}^N, X_{\phi(\bar\tau)}^N}[U(t, Y_u^{N,\delta})^2]\right)^{1/2} \leq \frac{\|f\|_\infty}{1 \wedge \varepsilon^{|\alpha'|}} C \exp\left(-c\frac{\varepsilon^2}{4u}\right). \tag{1.4.25}$$

By definition of η, if $|r_u^{N,\delta}| \notin [0,1/2]$ then $g(Y_u^{N,\delta})\eta(r_u^{N,\delta}) = 0$. The process H can be expressed as a sum of terms, each one being a product that includes a partial derivative of η taken at point $r_u^{N,\delta}$. From the definition of H_α, see equation (2.1) in [BT96a], and the local property of the Skorohod integral, cf. Proposition 1.3.6 from [Nua95], we have

$$H_\alpha(Y_u^{N,\delta}, g(Y_u^{N,\delta})\eta(r_u^{N,\delta})) = H_\alpha(Y_u^{N,\delta}, g(Y_u^{N,\delta})\eta(r_u^{N,\delta}))\mathbb{I}_{[0,1/2]}|r_u^{N,\delta}|.$$

Using now (1.4.14) and (1.4.20) it comes

$$\begin{aligned}\left(\mathbb{E}_{X_{\bar\tau}^N, X_{\phi(\bar\tau)}^N}[|H(Y_u^{N,\delta}, g(Y_u^{N,\delta})\eta(r_u^{N,\delta}))|^2]\right)^{1/2} &\leq \|\Gamma_u^{N,\delta}\mathbb{I}_{\hat\gamma_u^{N,\delta}\geq 1/2\hat\gamma_u}\|_{\mathrm{L}_{q_3}}^{q_4} \|Y_u^{N,\delta}\|_{\mathbb{D}^{p_1,q_1}}^{q_5} \|g(Y_u^{N,\delta})\eta(r_u^{N,\delta})\|_{\mathbb{D}^{p_2,q_2}} \\ &\leq C\|\Gamma_u\|_{\mathrm{L}_{q_3}}^{q_4} \frac{1+\|X_{\bar\tau}^N\|^\mu}{u^{dL}} \end{aligned} \tag{1.4.26}$$

where the upper-bound in (1.4.26) is uniform w.r.t. δ. Let us now plug (1.4.25), (1.4.26) into (1.4.24). Since v is continuous on $[0, T-h) \times \bar D$, the dominated convergence Theorem yields $\lim_{\delta\to 0} A_C^{IBP,\delta} = A_C^{IBP}$. Hence,

$$\begin{aligned}|A_C^{IBP}| &\leq \frac{\|f\|_\infty}{1 \wedge \varepsilon^{|\alpha'|}} C \exp\left(-c\frac{\varepsilon^2}{4u}\right) \mathbb{E}_x[\frac{1+\|X_{\bar\tau}^N\|^q}{u^{\zeta_1}}] \\ &\leq \frac{\|f\|_\infty}{1 \wedge \varepsilon^{|\alpha'|+2\zeta_1}} C. \end{aligned} \tag{1.4.27}$$

Relations (1.4.23) and (1.4.27) give the result.

□

Uniformly elliptic case

Lemma 1.4.11 *Assume* **(UE)**, **(D-3)**, **(S-$l+1$)**, **(F2-$l+2$)**, $l \in (0,1)$ *with f satisfying the compatibility conditions $f|_{\partial D} = Lf|_{\partial D} = 0$. One has*

$$E_1^I(h) = O(h^{1/2+l/2}).$$

Proof. From Proposition 1.2.2 and Theorem 4.1 Chapter 4 from [LSU68], one can extend v into a $H^{l/2+1,l+2}([0,T]\times\mathbb{R}^d) \subset C^{1,2}([0,T]\times\mathbb{R}^d)$ function $\bar v$ s.t. $\bar v|_{[0,T]\times\bar D} = v$. It comes

$$\begin{aligned}\mathbb{E}_x\left[\int_0^T dt \mathbb{I}_{\tau_R^N\wedge\tau^N > t}\mathbb{I}_{X_t^N\in D}\left(-Lv + \mathcal{L}_z v\right)|_{z=X_{\phi(t)}^N}(t, X_t^N)\right] &= \\ \int_0^T dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X_{\phi(t)}^N}[\mathbb{I}_{\tau_R^N>t}\mathbb{I}_{X_t^N\in D}\left(-L\bar v + \mathcal{L}_z\bar v\right)|_{z=X_{\phi(t)}^N}(t, X_t^N)]] &= \\ \int_0^T dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X_{\phi(t)}^N}[\mathbb{I}_{\tau_R^N>t}\left(-L\bar v + \mathcal{L}_z\bar v\right)|_{z=X_{\phi(t)}^N}(t, X_t^N)]] - \\ \int_0^T dt\mathbb{E}_x[\mathbb{I}_{\phi(t)<\tau^N}\mathbb{E}_{X_{\phi(t)}^N}[\mathbb{I}_{\tau_R^N>t}\mathbb{I}_{X_t^N\notin D}\left(-L\bar v + \mathcal{L}_z\bar v\right)|_{z=X_{\phi(t)}^N}(t, X_t^N)]] := A - B.\end{aligned}$$

Control of A: from Lemma 1.1.2 one has

$$A = \int_0^T dt\mathbb{E}_x[\mathbb{I}_{\tau^N>\phi(t)}\mathbb{E}_{X_{\phi(t)}^N}[(-L\bar v + \mathcal{L}_z\bar v)|_{z=X_{\phi(t)}^N}(t, X_t^N)]] + O_{pol}(h).$$

Write

$$\begin{aligned}
A &= \int_0^T dt\mathbb{E}_x\Bigg[\mathbb{I}_{\tau^N>\phi(t)}\Bigg\{\mathbb{E}_{X^N_{\phi(t)}}\big[(b(X^N_{\phi(t)})-b(X^N_t))\cdot\nabla\bar{v}(\phi(t),X^N_{\phi(t)})+\\
&\quad \frac{1}{2}\mathrm{tr}((\sigma\sigma^*(X^N_{\phi(t)})-\sigma\sigma^*(X^N_t))H_{\bar{v}}(\phi(t),X^N_{\phi(t)}))\big]+\\
&\quad \mathbb{E}_{X^N_{\phi(t)}}\big[(b(X^N_{\phi(t)})-b(X^N_t))\cdot(\nabla\bar{v}(t,X^N_t)-\nabla\bar{v}(\phi(t),X^N_{\phi(t)}))+\\
&\quad \frac{1}{2}\mathrm{tr}((\sigma\sigma^*(X^N_{\phi(t)})-\sigma\sigma^*(X^N_t))(H_{\bar{v}}(t,X^N_t)-H_{\bar{v}}(\phi(t),X^N_{\phi(t)})))\big]\Bigg\}\Bigg]+O_{pol}(h)\\
&:= \int_0^T dt\mathbb{E}_x[\mathbb{I}_{\tau^N>\phi(t)}(A_1(t,X^N_{\phi(t)})+A_2(t,X^N_{\phi(t)}))]+O_{pol}(h). \qquad (1.4.28)
\end{aligned}$$

For $A_1(t,X^N_{\phi(t)}),\ t\in[0,T]$ we do a Taylor expansion of the coefficients around $X^N_{\phi(t)}:=z$. We get

$$\begin{aligned}
A_1(t,z) &= \nabla\bar{v}(\phi(t),z)\cdot\Big(\mathbb{E}_z[b(X^N_t)-b(z)-Db(z)(X^N_t-z)]+Db(z)\mathbb{E}_z[X^N_t-z]\Big)+\\
&\quad \frac{1}{2}\mathrm{tr}\Big(H_{\bar{v}}(\phi(t),z)\left(\mathbb{E}_z[\sigma\sigma^*(X^N_t)-\sigma\sigma^*(z)-D\sigma\sigma^*(z)\otimes(X^N_t-z)]+\mathbb{E}_z[D\sigma\sigma^*(z)\otimes(X^N_t-z)]\right)\Big).
\end{aligned}$$

From assumption **(S-$l+1$)** the functions $Db, D\sigma\sigma^*$ are l-Hölder continuous. It comes

$$|A_1(t,z)| \le C(\mathbb{E}_z[\|X^N_t-z\|^{1+l}]+h)\le Ch^{1/2+l/2}$$

where we used Lemma 1.1.2 for the last inequality.

For $A_2(t,z),\ t\in[0,T]$ we use the l-Hölder continuity in space (resp. $l/2$-Hölder continuity in time) of $H_{\bar{v}}$ and the Lipschitz property of the coefficients. We obtain

$$|A_2(t,z)|\le C(\mathbb{E}_z[\|X^N_t-z\|^{1+l}]+h^{1/2+l/2})\le Ch^{1/2+l/2}.$$

Plugging the controls on $A_1(t,z),A_2(t,z)$ into (1.4.28) we finally derive

$$|A|\le Ch^{1/2+l/2}.$$

Control of B: Proposition 1.2.2 readily gives $|B|\le C\int_0^T dt\mathbb{P}_x[\phi(t)<\tau^N, X_t\notin D]$. From Lemma 1.4.5 we then derive $|B|\le Ch$.

The controls on A and B give the result.

□

4.4 Last step

Hypoelliptic case

Lemma 1.4.12 (Control of the last time step). *Under* **(C)**, **(D-∞)**, **(F1)**, **(H')**, **(S-∞)**, *there exist C,β such that:*

$$|\mathbb{E}_x[v(T\wedge\tau^N\wedge\tau^N_R,X^N_{T\wedge\tau^N\wedge\tau^N_R})-v((T-h)\wedge\tau^N\wedge\tau^N_R,X^N_{(T-h)\wedge\tau^N\wedge\tau^N_R})]|\le\frac{C}{1\wedge\varepsilon^\beta}h.$$

As in the previous section in that framework, the main ideas involved in the proof come from [Cat91], for the conditional Malliavin calculus, and [BT96a] for the localization techniques that allow the integrations by parts in order to get rid of the derivatives of v: nevertheless, the proof of Lemma 4.3 in [BT96a] seems to be incomplete[2]. We provide extra arguments that justify the result.

[2]The authors seem to assert that the non degeneracy of the Malliavin covariance matrix of the Euler scheme at time $T/2$ is enough to guarantee this property at time $s>T/2$. This property holds true for the diffusion but can fail for the Euler scheme. Actually, Lemma 4.3 in [BT96a] can be proved using the same procedure as for A_1 and A_2, with a perturbed process $X^{N,\lambda}$ which perturbation amplitude equals h^2 instead of h.

Proof. We denote $\psi v(.,x) := \psi(x)v(.,x)$ and recall $1-\psi(x) \neq 0 \Longrightarrow x \in V_{\partial D}(\varepsilon)$. As a consequence of **(F1)**, (1.2.3) and Lemma 1.1.2 the expectation in Lemma 1.4.12 writes:

$$\begin{aligned}
&\mathbb{E}_x[v(T\wedge\tau^N\wedge\tau_R^N, X^N_{T\wedge\tau^N\wedge\tau_R^N}) - v((T-h)\wedge\tau^N\wedge\tau_R^N, X^N_{(T-h)\wedge\tau^N\wedge\tau_R^N})]\\
=&\mathbb{E}_x[\mathbb{I}_{\tau^N>T-h}(\psi v(T,X_T^N) - \psi v(T-h, X^N_{T-h}))] + O_{pol}(h)\\
=&-\mathbb{E}_x[\mathbb{I}_{\tau^N\le T-2h}\mathbb{E}_{\tau^N,X^N_{\tau^N}}(\psi v(T,X_T^N) - \psi v(T-h, X^N_{T-h}))]\\
&+\mathbb{E}_x[\psi v(T,X_T^N) - \psi v(T-h, X^N_{T-h})] + O_{pol}(h) := -A_1 + A_2 + O_{pol}(h).
\end{aligned}$$

The choice of $T-2h$ in the last equation will be justified later on. We detail the control of A_1 that is the less usual term, we would treat A_2 in the same way. For sake of simplicity, denote $\mathbb{E}_. = \mathbb{E}_{\tau^N,X^N_{\tau^N}}$. In order to use classical expansion techniques for smooth functions, we write $\mathbb{E}_.[\psi v(T,X_T^N) - \psi v(T-h,X^N_{T-h})] = A_3(m) + R_m$ with

$$A_3(m) = \mathbb{E}_.[\psi v_m(T,X_T^N) - \psi v_m(T-h,X^N_{T-h})] \tag{1.4.29}$$

where we put

$$v_m(t,x) := \mathbb{E}_x[f_m(X_{T-t})\mathbb{I}_{\tau>T-t}] \text{ for } f_m \in C_0^\infty(\mathbb{R}^d)$$

and

$$R_m := \mathbb{E}_.[(\psi f - \psi f_m)(X_T^N)] + \mathbb{E}_.[(\psi v_m - \psi v)(T-h, X^N_{T-h})].$$

By a density argument, we can choose $(f_m)_{m\ge0}$ s.t. for all $m\ge 0$, $\|f_m\|_\infty \le 2\|f\|_\infty$, $d(\mathrm{supp}(f_m),\partial D)\ge 3/2\varepsilon$ and $f_m \underset{m\longrightarrow\infty}{\overset{L^1(\mu^N)}{\longrightarrow}} f$, where $\mu^N(dy) := \mathbb{E}_x[q_h(X^N_{T-h},y)]dy + \mathbb{P}\circ(X_T^{x,N})^{-1}(dy)$. Hence, for m large enough $\mathbb{E}_x[|R_m|] \le Ch$. It is enough to prove $|A_3(m)|\le Ch$ uniformly in m, τ^N and $X^N_{\tau^N}\in V_{\partial D}(R)$.
Since ψv_m is smooth, one would like to develop $A_3(m)$ with Itô's formula and then use standard Malliavin integrations by parts; this last step can not be so direct because the variables of interest may be degenerate in the Malliavin sense. To circumvent this difficulty, we introduce a family of perturbed processes $(X_s^{N,\lambda})_{s\in[\tau^N,T]} := (X_s^N + \lambda h(\tilde W_s - \tilde W_{\tau^N}))_{s\in[\tau^N,T]}$ ($\lambda\in[0,1]$) starting from $X^N_{\tau^N}$ at time τ^N, where $\tilde W$ is a standard d-dimensional BM independent from W. We also consider the diffusion $(X_s)_{s\ge\tau^N}$ starting at τ^N from $X^N_{\tau^N}$: in the following, estimates will be uniform in $\tau^N\le T-2h$ and $X^N_{\tau^N}\in V_{\partial D}(R)$ and we omit from now on to indicate this dependence.
The next Malliavin calculus computations will be performed w.r.t. the $(d'+d)-$dimensional BM $(W,\tilde W)$ after time τ^N: $\|Z\|_{\mathrm{L}_p,.}$ and $\|Z\|_{\mathbb{D}^{j,p},.}$ stand for the associated L_p and Sobolev norms of Z. We denote γ_s the Malliavin covariance matrix of X_s and $\hat\gamma_s := \det(\gamma_s)$ its determinant. The same notations indexed by N(resp. N,λ) stand for X_s^N(resp. $X_s^{N,\lambda}$). We now give an extension of the previous controls (1.4.20). Under the above assumptions, one has for any $p>1$ and $j\ge1$

$$\|X_s^{N,\lambda}\|_{\mathbb{D}^{j,p},.}\le C,\quad \|\hat\gamma_s^{-1}\|_{\mathrm{L}_p,.}\le\frac{C}{(s-\tau^N)^\zeta},\quad \|X_s - X_s^{N,\lambda}\|_{\mathbb{D}^{j,p},.}\le C\sqrt{h}, \tag{1.4.30}$$

for some constants, uniform in $\lambda\in[0,1]$, $\tau^N\le s\le T$ and $X^N_{\tau^N}\in V_{\partial D}(R)$.

From (1.4.29), $A_3(m)$ is equal to

$$\begin{aligned}
&\mathbb{E}_.[\psi v_m(T,X_T^N) - \psi v_m(T,X_T^{N,1})] + \mathbb{E}_.[\psi v_m(T,X_T^{N,1}) - \psi v_m(T-h,X^{N,1}_{T-h})]\\
&+\mathbb{E}_.[\psi v_m(T-h,X^{N,1}_{T-h}) - \psi v_m(T-h,X^N_{T-h})] := (A_4+A_5+A_6)(m).
\end{aligned}$$

For $A_4(m), A_6(m)$ we have to check that the difference between the Euler scheme and the perturbed process is negligible. For $A_5(m)$, since $X^{N,1}$ satisfies the non degeneracy condition we can use Itô's formula associated with integrations by parts techniques.

Control of $A_4(m), A_6(m)$. We only detail $A_4(m)$, the other term can be handled in the same way thanks to the restriction to $\tau^N\le T-2h$. Let η_T be a $\mathbb{D}^\infty$ $[0,1]$-valued random variable, satisfying **(C1)**: $\mathbb{P}_.[\eta_T\neq$

$1] \le C\frac{h^2}{(T-\tau^N)^\zeta}$ and **(C2)**: $\eta_T \neq 0 \Rightarrow \forall \lambda \in [0,1],\ \hat\gamma_T^{N,\lambda} \ge \hat\gamma_T/4$. It follows from **(C2)** and (1.4.30) that $\|(\hat\gamma_T^{N,\lambda})^{-1}\mathbb{I}_{\eta_T\neq 0}\|_{\mathrm{L}_{p,\cdot}} \le \frac{C}{(T-\tau^N)^\zeta}$ for $\lambda \in [0,1]$. A Taylor expansion yields:

$$\begin{aligned}A_4(m) :=& \mathbb{E}.[(\psi v_m(T,X_T^N) - \psi v_m(T,X_T^{N,1}))(1-\eta_T)]\\ &- h\int_0^1 \mathbb{E}.[\nabla \psi v_m(T,X_T^{N,\lambda}).(\tilde W_T - \tilde W_{\tau^N})\eta_T]d\lambda := (A_{41}+A_{42})(m).\end{aligned}$$

From the support property of f_m, Lemma 1.1.2 and **(C1)** we easily deduce $|A_{41}(m)| \le C\exp(-c\frac{\varepsilon^2}{T-\tau^N})\mathbb{E}.[1-\eta_T]^{1/2} \le C\frac{h}{1\wedge\varepsilon^\zeta}$. Taking additionally into account (1.4.15) and **(C2)** yields $|A_{42}(m)| \le C\frac{h}{1\wedge\varepsilon^\zeta}$.
We now turn to the construction of η_T. To satisfy **(C2)** we will choose η_T as a mollified indicator function of the sets where $\gamma_T^{N,\lambda}$ is close enough to γ_T uniformly in $\lambda\in[0,1]$. Remark that $A(\lambda) := \gamma_T^{N,\lambda} = \gamma_T^N + \lambda^2 h^2(T-\tau^N)I_d$ is $a.s$ invertible for $\lambda > 0$. The function $a(\lambda) := \det(A(\lambda))$ is differentiable in λ and its derivative is given by (see Theorem A.98 from [RT99]) $a'(\lambda) = \mathrm{tr}(\mathrm{Cof}(A(\lambda))A'(\lambda)) = 2\lambda h^2(T-\tau^N)\mathrm{tr}(\mathrm{Cof}(\gamma_T^{N,\lambda}))$. Simple computations yield $|a'(\lambda)|^2 \le Ch^4\Big(\int_{\tau^N}^T \{\|D_t X_T^N\|^2 + h^2\}dt\Big)^2 := R_N$ so that $|a(\lambda)-a(0)|^2 = |\hat\gamma_T^{N,\lambda} - \hat\gamma_T^N|^2 \le R_N$ for $\lambda\in[0,1]$. Introduce now an even function $\eta \in C_b^\infty(\mathbb{R})$ s.t. $\mathbb{I}_{[0,1/4]}(x) \le \eta(x) \le \mathbb{I}_{[0,1/2]}(x)$ for $x \ge 0$, and put $\eta_T^1 := \eta((\hat\gamma_T - \hat\gamma_T^N)/\hat\gamma_T), \eta_T^2 := \eta(8R_N/[\hat\gamma_T]^2)$: we set $\eta_T = \eta_T^1\eta_T^2$. Indeed, **(C2)** is fulfilled: $\eta_T^1\eta_T^2 \neq 0 \Rightarrow \hat\gamma_T^N \ge \hat\gamma_T/2,\ R_N \le [\hat\gamma_T]^2/16$ and thus $\hat\gamma_T^{N,\lambda} \ge \hat\gamma_T/4$ for $\lambda\in[0,1]$. To check **(C1)**, write: $\mathbb{E}.[1-\eta_T^1\eta_T^2] \le \mathbb{P}.[\eta_T^1 \neq 1] + \mathbb{P}.[\eta_T^2\neq 1]$. Using Markov's inequality and (1.4.30), one readily gets

$$\mathbb{P}.[\eta_T^1 \neq 1] \le \sqrt{\mathbb{E}.[4^q|\hat\gamma_T^N - \hat\gamma_T|^q]}\sqrt{\mathbb{E}.[\hat\gamma_T^{-q}]} \le C_q\frac{h^{q/4}}{(T-\tau^N)^{\zeta_q}} \tag{1.4.31}$$

for any q. An analogous estimate is available for $\mathbb{P}.[\eta_T^2 \neq 1]$.

Control of $A_5(m)$. Using Itô's formula $A_5(m)$ writes as a finite sum of terms

$$\int_{T-h}^T \mathbb{E}.[\partial_x^\alpha[(\partial_x^{\alpha'}\psi)v_m](s,X_s^{N,1})g_{\alpha,\alpha'}(X_s^{N,1},X_{T-h}^{N,1})]ds$$

where $|\alpha| \le 2, |\alpha|+|\alpha'| \le 2$ and $g_{\alpha,\alpha'}$ is a bounded function that only depends on b,σ in (1.0.1). Combining (1.4.15) with estimates (1.2.3) (written for f_m) and Lemma 1.1.2 give $|A_5(m)| \le C\frac{\|f\|_\infty}{1\wedge\varepsilon^2}\int_{T-h}^T \exp\big(-c\frac{\varepsilon^2}{s-\tau^N}\big)\|(\hat\gamma_s^{N,1})^{-1}\|_{\mathrm{L}_{p,\cdot}}^q ds$. To complete the proof, we assert that $\|(\hat\gamma_s^{N,1})^{-1}\|_{\mathrm{L}_{p,\cdot}} \le \|(\hat\gamma_s^{N,1})^{-1}\mathbb{I}_{\hat\gamma_s^{N,1}\ge\hat\gamma_s/2}\|_{\mathrm{L}_{p,\cdot}} + \|(\hat\gamma_s^{N,1})^{-1}\mathbb{I}_{\hat\gamma_s^{N,1}<\hat\gamma_s/2}\|_{\mathrm{L}_{p,\cdot}} \le \frac{C}{(s-\tau^N)^\zeta}$. Indeed, the first term readily satisfies the required upper bound if we apply (1.4.30). For the second, note that since $\hat\gamma_s^{N,1} \ge ((s-\tau^N)h^2)^d$, it is enough to get that $\mathbb{P}.[\hat\gamma_s^{N,1} < \hat\gamma_s/2] \le C_p\frac{h^p}{(s-\tau^N)^{\zeta_p}}$ for p large enough. This last estimate can be proved as (1.4.31). Indeed, $\mathbb{P}.[\hat\gamma_s^{N,1} < \hat\gamma_s/2] \le C\mathbb{P}.\Big[\Big|\frac{\hat\gamma_s^{N,1}-\hat\gamma_s}{\hat\gamma_s}\Big| > 1/2\Big] \le C\sqrt{\mathbb{E}.[2^q|\hat\gamma_s^{N,1}-\hat\gamma_s|^q]}\sqrt{\mathbb{E}.[\hat\gamma_s^{-q}]}$. Using the definition of R_N and (1.4.30) we get the result.

□

Uniformly Elliptic case

Lemma 1.4.13 *Under* **(UE)**, **(D-3)**, **(S-$l+1$)**, **(F2-$l+2$)**, $l\in(0,1)$ *with f satisfying the compatibility conditions* $f|_{\partial D} = Lf|_{\partial D} = 0$*:*

$$E_2(h) = \mathbb{E}_x[v(T\wedge\tau^N\wedge\tau_R^N, X^N_{T\wedge\tau^N\wedge\tau_R^N}) - v((T-h)\wedge\tau^N\wedge\tau_R^N, X^N_{(T-h)\wedge\tau^N\wedge\tau_R^N})] = o(h^{1/2}).$$

Proof. From Proposition 1.2.2, $u \in H^{l/2+1,l+2}([0,T]\times\bar D)$. Using this smoothness property and the semi-martingale decomposition of the projection of the Euler scheme, cf. equation (1.3.2), we write,

$$\begin{aligned}E_2(h) = \mathbb{E}_x\Bigg[&\mathbb{I}_{\tau^N>T-h}\int_{T-h}^{T\wedge\tau_R^N}\frac{1}{2}\frac{\partial v}{\partial n}(t,X_t^N)dL_t^0(F(X^N)) + \mathbb{I}_{X_t^N\in D}\Big((\mathcal{L}_{X^N_{T-h}} - L)v(t,X_t^N)\Big)dt\\ &+ \mathbb{I}_{X_t^N\notin D}\Big(\nabla v(t,\pi_{\partial D}(X_t^N))\cdot dX_t^{N,\partial D} + \frac{1}{2}\mathrm{tr}(H_v(t,\pi_{\partial D}(X_t^N))d\langle X^{N,\partial D}\rangle_t)\Big)\Bigg] := E_2^1(h)+E_2^2(h)+E_2^3(h).\end{aligned}$$

Since by Proposition 1.2.2 the integrands in $E_2^2(h), E_2^3(h)$ are bounded, we readily get from Lemma 1.4.5 $E_2^2(h) + E_2^3(h) = O(h)$. For $E_2^1(h)$, the boundedness of $\frac{\partial v}{\partial n}$ and Lemma 1.4.5 yield

$$
\begin{aligned}
&|E_2^1(h)| \le \mathbb{E}_x[\mathbb{I}_{\tau^N > T-h}(\mathbb{I}_{X_{T-h}^N \in V_{\partial D}(h^{1/4})} + \mathbb{I}_{X_{T-h}^N \notin V_{\partial D}(h^{1/4})})\mathbb{E}[F^-(X_{T\wedge \tau_R^N})|\mathcal{F}_{T-h}]] + O(h)\\
&:= E_2^{11}(h) + E_2^{12}(h) + O(h).
\end{aligned}
$$

Lemma 1.1.2 gives $E_2^{12}(h) = O_{pol}(h)$. On the other hand, following the proof of Lemma 1.4.2 we get

$$
E_2^{11}(h) \le C\sqrt{h}\mathbb{P}_x[X_{T-h}^N \in V_{\partial D}(h^{1/4})] \le Ch^{3/4}
$$

where the last inequality is a consequence of **(UE)**, under which one has Gaussian upper bounds for the marginal density of the Euler scheme.

□

5 Main results

Before stating our main results, we mention that Assumption **(C)** is sufficient to guarantee the convergence to 0 of the error.

Proposition 1.5.1 ***(Weak error convergence).*** *Under* **(C)**, *if D is of class C^2 with a compact boundary and the coefficients in (1.0.1) are Lipschitz continuous, for every continuous function f with polynomial growth we have:*

$$
\lim_{h \longrightarrow 0} \mathrm{Err}(T, h, f, x) = 0.
$$

Proof.

Let us first consider f is bounded. According to Proposition 1.1 in [Gob00] it suffices to satisfy the condition $\mathbb{P}_x[\exists t \in [0,T] : X_t \notin D;\ \forall t \in [0,T],\ X_t \in \bar{D}] = 0$. This condition can also be written $\mathbb{P}_x[M = 0] = 0$, where $M := \inf_{s\in[0,T]} F(X_s)$. From Proposition 2.1.3 in [Nua95] we derive $M \in \mathbb{D}^{1,2}$ and, cf. proof of Proposition 2.1.4 from the same reference, $DM = DF(X_s)$ on the set of $s \in \mathcal{M}_T := \{u \in [0,T] : F(X_u) = M\}$. According to Theorem 1.4.8 it is enough to prove that $\|DF(X_t)\|_{L^2[0,T]} > 0$ *a.s* for $t \in \mathcal{M}_T \subset]0,T]$ on the event $|M| \le \frac{R \wedge F(x)}{2}$. But for $t \in \mathcal{M}_T$ and $|M| \le R/2,\ X_t \in V_{\partial D}(R)$ and thus $\|D_tF(X_t)\|^2 = \alpha(X_t) \ge a_0 > 0$: by continuity of $s \in [0,t] \mapsto D_sF(X_t)$, it easily follows $\|DF(X_t)\|_{L^2([0,T])} > 0$ *a.s.*

If f has polynomial growth, from Lemma 1.1.2 one derives that $\forall \varepsilon > 0,\ \exists C := C(\varepsilon), \mathbb{P}_x[\|X_T\| + \|X_T^N\| \ge C] \le \varepsilon/2$. Choose now a function $\phi \in C_b^\infty(\mathbb{R}^d),\ \mathbb{I}_{\|x\|\le C} \le \phi(x) \le \mathbb{I}_{\|x\|\le C+1}$. The previous arguments then give that for N large enough

$$
|\mathbb{E}_x[f(X_T^N)\phi(X_T^N)\mathbb{I}_{\tau^N > T}] - \mathbb{E}_x[f(X_T)\phi(X_T)\mathbb{I}_{\tau > T}]| \le \varepsilon/2.
$$

□

The following counter example illustrates that **(C)** is also somehow minimal to ensure the convergence of $\mathrm{Err}(T, h, f, x)$.

For the one dimensional deterministic process $(X_s)_{s\in[0,1]} = (\exp(s))_{s\in[0,1]}$, corresponding to the choice of $b(y) = y,\ \sigma(y) = 0,\ x = 1$ in (1.0.1), and $D = (-\infty, e)$ one has $\tau = 1$. Since $\sigma = 0$, Assumption **(C)** is not satisfied.

Put $T = 1$. The Euler scheme of the previous process writes $\forall s \in [0,T],\ X_s^N = X_{\phi(s)}^N(1 + s - \phi(s)) = (1 + N^{-1})^{\phi(s)/h}(1 + s - \phi(s))$. By construction X^N is increasing on $[0,1]$ and $X_1^N = (1 + N^{-1})^N$. Hence $\tau^N > 1$ and for $f = 1$, $\mathrm{Err}(T, h, f, x) = 1$.

5.1 Statement of the main results

We briefly recall our two main working assumptions, see also Section 1 and Appendix B.

- Hypoellipticity conditions **(H)** or **(H')** associated to a non characteristic boundary condition (**C**), and to the assumptions **(D-∞)** for the domain, **(S-∞)** for the coefficients (the domain and the coefficients of the diffusion process X are infinitely differentiable), **(F1)** for the function (f is a bounded Borel function satisfying a support condition w.r.t. D).

- Uniform ellipticity conditions **(UE)** associated to the assumptions **(D-3)** for the domain, **(S-$l+1$)** for the coefficients (the domain is of class C^3 and the coefficients of the diffusion process X belong to the Hölder space $H^{l+1}(\bar{D})$), **(F2-$l+2$)** for the function (the function f belongs to the Hölder space $H^{l+2}(\bar{D})$).

Theorem 1.5.2 *Under* **(C)**, **(D-∞)**, **(F1)**, **(H')**, **(S-∞)** *(resp.* **(UE)**, **(D-3)**, **(S-l+1)**, **(F2-l+2)**, $l \in (0,1)$ *with f satisfying $f|_{\partial D} = Lf|_{\partial D} = 0$) we have*

$$\mathrm{Err}(T,h,f,x) = \frac{1}{2}\mathbb{E}_x\left[\int_0^T \frac{\partial v}{\partial n}(s, X_s^N) dL^0_{s\wedge\tau^N}(F(X^N))\right] + o(h^{1/2}).$$

Theorem 1.5.3 *Under* **(C)**, **(D-∞)**, **(F1)**, **(H')**, **(S-∞)** *(resp.* **(UE)**, **(D-3)**, **(S-l+1)**, **(F2-l+2)**, $l \in (0,1)$ *with f satisfying $f|_{\partial D} = Lf|_{\partial D} = 0$), for h small enough (depending on $d(x,\partial D) > 0$), we have*

$$|\mathrm{Err}(T,h,f,x)| \leq C_2\sqrt{h} \tag{1.5.1}$$

with $C_2 > 0$.

Under **(C)**, **(D-∞)**, **(F1)**, **(H)**, **(S-∞)** *(resp.* **(UE)**, **(D-3)**, **(S-$l+1$)**, **(F2-$l+2$)**, $l \in (0,1)$ *with f satisfying $f|_{\partial D} = Lf|_{\partial D} = 0$), for h small enough (depending on $d(x,\partial D) > 0$), we have*

$$C_1(x)\sqrt{h} \leq \mathrm{Err}(T,h,f,x) \tag{1.5.2}$$

with $C_1(x) > 0$. If the domain is bounded, $C_1(x)$ can be taken uniform w.r.t. the starting point $x \in D$.

Proposition 1.5.4 *Under* **(C)**, **(D-2)**, **(S-1)**, *for some $c_0 > 0$ one has:*

$$\sup_{N, s\in[0,T]} \mathbb{E}_x[\exp(c_0[h^{-1/2}F^-(X^N_{s\wedge\tau^N})]^2)] < \infty.$$

Hence, the sequence of random variables $(h^{-1/2}F^-(X^N_{s\wedge\tau^N}))_{N\geq 1}$ is uniformly tight on $[0,T]$.

The first theorem exhibits the relevant term of the error, i.e. the one that has to be developed in order to give an expansion of the error. The second one states that the leading term is really of order $\frac{1}{2}$. Moreover, the main term can be interpreted in terms of Tanaka's formula as a suitable average of the overshoot. Indeed, we show, see Lemma 1.4.5, that $\frac{1}{2}\mathbb{E}_x[L^0_{s\wedge\tau^N}(F(X^N))] = \mathbb{E}_x[F^-(X^N_{s\wedge\tau^N})] + O(h)$. Therefore, Theorem 1.5.2 and Proposition 1.5.4 are somehow the first step for a future expansion of the error.

Remark 1.5.1 To illustrate the necessity of assumption **(H)** to get a lower bound, we provide the following example suggested by P.Cattiaux, [Cat03]. Consider a domain $D =]-\pi/2, 3\pi/2[$ and a diffusion process following the dynamics of equation (1.0.1) with $X_0 = \pi/2$, $b(x) = \cos(x)$, $\sigma(x) = \sin(x)$. Assumption **(C)** is readily satisfied. With the notations of Section 1, or equivalently writing the above dynamic in Stratonovitch form, one has

$$A_0(x) = (1 - \frac{1}{2}\sin(x))\cos(x), A_1(x) = \sin(x).$$

The diffusion coefficient A_1 degenerates along the set $\{k\pi,\ k \in \mathbb{Z}\}$. Thus **(H)** is not satisfied. Anyhow, since

$$\begin{aligned}
[A_0, A_1](x) &= \left((1 - \frac{1}{2}\sin(x))\cos(x)\right)\cos(x) - \left(-\frac{1}{2}(\cos^2(x) - \sin^2(x)) - \sin(x)\right)\sin(x) \\
&= 1 - \frac{1}{2}\sin^3(x), \\
([A_1]^2 + [A_0, A_1])(x) &= \sin^2(x) + 1 - \frac{1}{2}\sin^3 x \geq \frac{1}{2}.
\end{aligned}$$

Hence, assumption **(H')** holds true. Note also that $(X_t)_{t\geq 0}$ lives in $[0,\pi]$. Indeed Feller's test for explosions is easily verified, see e.g. Proposition 5.22 Chapter 5 from [KS91]. We therefore derive that the killing boundary has no effect in that case for which $\mathrm{Err}(T,h,f,x)=O(h)$ (see [BT96a]). We fail to obtain a lower bound with rate $\sqrt{h}$.

5.2 Proof of the main results

Proof of Theorem 1.5.3. As a consequence of Theorem 1.5.2, the lower bound follows from Lemmas 1.2.4 and 1.4.6. Indeed, from Lemma 1.2.4 and Proposition 1.2.1(resp. Proposition 1.2.2) we derive that the function $\frac{\partial v}{\partial n}$ is positive and continuous on the compact set $[0,T/2]\times\partial D$. It is therefore uniformly positive. Lemma 1.4.6 then gives a lower-bound for the expectation of the local time.

The upper bound is derived from equation (1.2.3) in the hypoelliptic case (resp. Proposition 1.2.2 in the uniformly elliptic case) and Lemma 1.4.2.

□

Proof of Theorem 1.5.2.

To show Theorem 1.5.2, we use the technical results of Section 4 and the smoothness properties of v given in Section 2. This allows to control the terms $E_1^I(h), E_1^O(h), E_2(h)$ of the fundamental error decomposition (1.3.5).

Lemma 1.4.12 (resp. 1.4.13) states that $E_2(h)=O(h)$ (resp. $E_2(h)=o(h^{1/2})$). Estimates (1.2.3) (resp. Proposition 1.2.2) and Lemma 1.4.5 directly yield that the contribution $E_1^O(h)=O(h)$.

We then derive from Lemma 1.4.10 (resp. 1.4.11) that $E_1^I(h)=o(\sqrt{h})$. Note that in the hypoelliptic case we actually have $E_1^I(h)=O(h)$.

At last, the difference between the integrals w.r.t. the local time stopped in $(T-h)\wedge\tau^N\wedge\tau_R^N$ appearing in $E_1^{LT}(h)$, cf. (1.3.4), and $T\wedge\tau^N$ is a $O_{pol}(h)$. This is easy to prove using (1.2.3) (resp. Proposition 1.2.2), Lemma 1.1.2 and the local L_p boundedness of the local time, we omit details.

□

Proof of Proposition 1.5.4. We only have to prove that there exist constants $c>0$ and C s.t: $\forall A\geq 0, \sup_N \mathbb{P}_x[F^-(X^N_{t\wedge\tau^N})\geq Ah^{1/2}]\leq C\exp(-cA^2)$ for $t\in[0,T]$, then any choice of $c_0<c$ is valid. We write:

$$\begin{aligned}\mathbb{P}_x[F^-(X^N_{t\wedge\tau^N})\geq Ah^{1/2}] = &\sum_{i=1}^{\phi(t)/h}\mathbb{E}_x[\mathbb{I}_{\tau^N>t_{i-1}}\mathbb{I}_{\tau^N_{t_{i-1}}<t_i}\mathbb{P}[F^-(X^N_{t_i})\geq Ah^{1/2}|\mathcal{F}_{\tau^N_{t_{i-1}}}]]\\ &+\mathbb{P}_x[F^-(X^N_t)\geq Ah^{1/2},\tau^N>t]:=A_t+B_t\end{aligned}$$

where we define $\tau^N_{t_{i-1}}:=\inf\{t\geq t_{i-1}: X_t^N\notin D\}$. B_t is directly estimated applying Lemma 1.1.2. This Lemma also enables to develop A_t as follows:

$$\begin{aligned}A_t\leq &\sum_{i=1}^{\phi(t)/h}\mathbb{E}_x[\mathbb{I}_{\tau^N>t_{i-1}}\mathbb{I}_{\tau^N_{t_{i-1}}<t_i}C\exp\left(-c\frac{A^2h}{t_i-\tau^N_{t_{i-1}}}\right)]\\ \leq &C\exp(-cA^2)h^{-1}\int_0^T dt\mathbb{P}_x[\tau^N>\phi(t),\phi(t)+h>\tau^N_{\phi(t)}]\\ \leq &C\exp(-cA^2)h^{-1}\int_0^T dt\mathbb{E}_x[\mathbb{I}_{\tau^N>\phi(t)}\exp\left(-c\frac{d^2(X^N_{\phi(t)},\partial D)}{h}\right)].\end{aligned}$$

In the proof of Lemma 1.4.5, we have shown that the last integral is bounded by Ch, which completes the proof.

□

6 Numerical results

In this section we first assume that one of the sets of assumptions of Theorem 1.5.3 is in force. We present some numerical experiments that illustrate the results of the previous section. We detail in Subsection 6.1 the simulation parameters we choose in order to analyze the discretization error. We then describe in Subsection 6.2 a simulation procedure which avoids some numerical "variance" effects. Two examples are given in Subsection 6.3. Finally, in Subsection 6.4 we deal with a non-zero function on the boundary.

6.1 Simulation parameters

In this subsection we specify our choice of parameters in order to isolate the discretization error. The theorems of Section 5 give some bounds concerning expectations. From a practical viewpoint the expectations involving the killed Euler scheme are approximated by a Monte-Carlo procedure. It is therefore crucial to control that the statistical error induced by the Monte-Carlo approximation is negligible w.r.t. to the discretization error we want to characterize.

For $N_{MC} \in \mathbb{N}$ and $N \in \mathbb{N}^*$ define:

$$E_{N_{MC}}(f,N) \quad := \quad \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} f(X_T^{N,i}) \mathbb{I}_{\tau^{N,i}>T}$$

where N still stands for the total number of time steps and N_{MC} denotes the number of paths for the Monte-Carlo method. By $X_T^{N,i}, i \in [\![1, N_{MC}]\!]$ we denote independent realizations of the Euler scheme taken at time T. The notation $\tau^{N,i}, i \in [\![1, N_{MC}]\!]$ stands for the discrete exit time associated to the i^{th} path of the Euler scheme.

In what follows, to approximate $\text{Err}(T,h,f,x)$, we compute numerically the quantity $E_{N_{MC}}(f,N) - E_{N_{MCR}}(f,N_R)$. The second term in this difference is a "reference" approximation value of the quantity $\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]$. Namely, we take N_R, N_{MCR} so that the difference $\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] - E_{N_{MCR}}(f,N_R)$ is negligible w.r.t. $h^{1/2}$. One has,

$$\begin{aligned}
&E_{N_{MC}}(f,N) - E_{N_{MCR}}(f,N_R) = \text{Err}(T,h,f,x) + (E_{N_{MC}}(f,N) - \mathbb{E}_x[f(X_T^N)\mathbb{I}_{\tau^N>T}]) \\
&+(\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] - \mathbb{E}_x[f(X_T^{N_R})\mathbb{I}_{\tau^{N_R}>T}]) + (\mathbb{E}_x[f(X_T^{N_R})\mathbb{I}_{\tau^{N_R}>T}] - E_{N_{MCR}}(f,N_R)) \\
&:= \text{Err}(T,h,f,x) + D_1 + D_2 + D_3.
\end{aligned}$$

For a given $N \in \mathbb{N}^*$ representing the maximum number of steps for which we want to study the discretization error, it remains to calibrate N_{MC}, N_R, N_{MCR} so that $D_1 + D_2 + D_3 = o(h^{1/2})$ in some sense.

Theorem 1.5.3 readily gives $D_2 \le C\sqrt{1/N_R}$. A problem is anyhow that we know little about the constant C appearing in the bounds of $\text{Err}(T,h,f,x)$.

The terms D_1, D_3 correspond to a statistical error that can be controlled by the central limit theorem, see Lapeyre and *al.*, [LPS98] for details. Namely, $\forall n \in \mathbb{N}^*$

$$\lim_{N_{MC}} \mathbb{P}\left[\sqrt{\frac{N_{MC}}{\mathbb{V}[f(X_T^n)\mathbb{I}_{\tau^n>T}]}}\left(E_{N_{MC}}(f,n) - \mathbb{E}[f(X_T^n)\mathbb{I}_{\tau^n>T}]\right) \in [a,b]\right] = \mathbb{P}[\mathcal{N}(0,1) \in [a,b]].$$

Hence, replacing $\mathbb{V}[f(X_T^n)\mathbb{I}_{\tau^n>T}]$ by the usual unbiased estimator of the variance

$$\hat{\sigma}^2_{N_{MC}}(f,n) := \frac{1}{N_{MC}-1} \sum_{i=1}^{N_{MC}-1} f(X_T^{n,i})^2 \mathbb{I}_{\tau^{n,i}>T} - \left(\frac{N_{MC}}{N_{MC}-1}\right) E^2_{N_{MC}}(f,n)$$

we heuristically derive that for N_{MC} large enough

$$\mathbb{E}[f(X_T^n)\mathbb{I}_{\tau^n>T}] \in \left[E_{N_{MC}}(f,n) - 1.96\frac{\hat{\sigma}_{N_{MC}}(f,n)}{\sqrt{N_{MC}}}, E_{N_{MC}}(f,n) + 1.96\frac{\hat{\sigma}_{N_{MC}}(f,n)}{\sqrt{N_{MC}}}\right]$$

with probability 95%.

Thus, for N large enough one gets from the above controls that the choice of $N_R = N_{MC} = N_{MCR} = N^2$ yields $D_1 + D_2 + D_3 = O(h)$ with a high probability.

6.2 Simulation procedure

Assume the process X follows the dynamics of equation (1.0.1). Let $N_I \in \mathbb{N}^*$ be the minimal number of time steps for which we want to simulate $\mathbb{E}_{N_{MC}}(f,N)$. For $S \in \mathbb{N}$ define $N_{S,I} := N_I \times 2^S$. Let now $N_{S_0,I}, S_0 \in \mathbb{N}^*$ be the maximum number of time steps for which we want to simulate $\mathbb{E}_{N_{MC}}(f,N)$.

The following Algorithms describe our simulation procedure.

Algorithm 6.1 (Path simulation of $(X^{N_{S,I}})_{S\in[\![0,S_0]\!]}$ using common Gaussian increments.)

- *Simulate a $\mathbb{R}^{N_{S_0,I}\otimes d'}$ valued random vector V whose components are i.i.d. centered Gaussian vectors with covariance matrix $(T/N_{S_0,I})\mathbf{I}_{d'}$.*
- *For every $S \in [\![0,S_0]\!]$, define $h_S = T/N_{S,I}$, $R_S = N_{S_0,I}/N_{S,I} = 2^{S_0-S}$, $\forall k \in \mathbb{N}$, $t_k^S := kh_S$.*

For all $k \in [\![0, N_{S,I}-1]\!]$ introduce $\Delta_{t_k^S} := \sum_{i=kR_S}^{(k+1)R_S-1} V_i$. By construction $\Delta_{t_k^S} \sim \mathcal{N}(0, h_S\mathbf{I}_{d'})$.

We now define

$$X_0^{N_S} = x,\ \forall k \in [\![1, N_{S,I}]\!],\ X_{t_k^S}^{N_{S,I}} = X_{t_{k-1}^S}^{N_{S,I}} + b(X_{t_{k-1}^S}^{N_{S,I}})h_S + \sigma(X_{t_{k-1}^S}^{N_{S,I}})\Delta_{t_{k-1}^S}.$$

Note that to obtain a realization of V we can also use N_I initial i.i.d. Gaussian increments with covariance matrices $(T/N_I)\mathbf{I}_{d'}$ and then simulate conditional laws to get the intermediate increments. When S_0 goes to infinity, this procedure leads to Lévy's construction of BM, see e.g. Chapter 2.3 in [KS91].

Algorithm 6.2 (Approximation procedure)

- *Let $(V_j)_{j\in[\![1,N_{MC}]\!]}$ be a sequence of i.i.d. random variables with law V introduced in the previous Algorithm. For every $j \in [\![1, N_{MC}]\!]$, use Algorithm 6.1 with V_j to compute independent paths (in terms of j) $(X_{t_k}^{N_{S,I},j})$ for $S \in [\![0,S_0]\!]$, $k \in [\![0,N_S]\!]$, $j \in [\![1,N_{MC}]\!]$ and derive the associated expressions for $(E_{N_{MC}}(f,N_{S,I}))_{S\in[\![0,S_0]\!]}$.*

The key idea in the above Algorithm consists in using the components of $(V^j)_{j\in[\![1,N_{MC}]}$ to simulate all the paths $(X_{t_k}^{N_{S,I},j})_{k\in[\![0,N_{S,I}]\!]}$ of the Euler scheme for $S \in [\![0,S_0]\!]$. In this way, we diminish the variance effects in the estimation of $(E_{N_{MC}}(f,N_{S,I}))_{S\in[\![0,S_0]\!]}$.

In the Brownian case we observe that $(E_{N_{MC}}(f,N_{S,I}))_{S\in[\![0,S_0]\!]}$ is by construction decreasing. As a consequence of Theorem 1.5.3 we had this property for $\mathbb{E}_x[f(X_T^N)\mathbb{I}_{\tau^N>T}]$ that also always overestimates $\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]$. This is the "overshoot" effect associated to the discrete killing. In a Brownian framework this decreasing property is also satisfied by the empirical mean.

6.3 Results

Brownian Motion living in the unit ball of $\mathbb{R}^2$

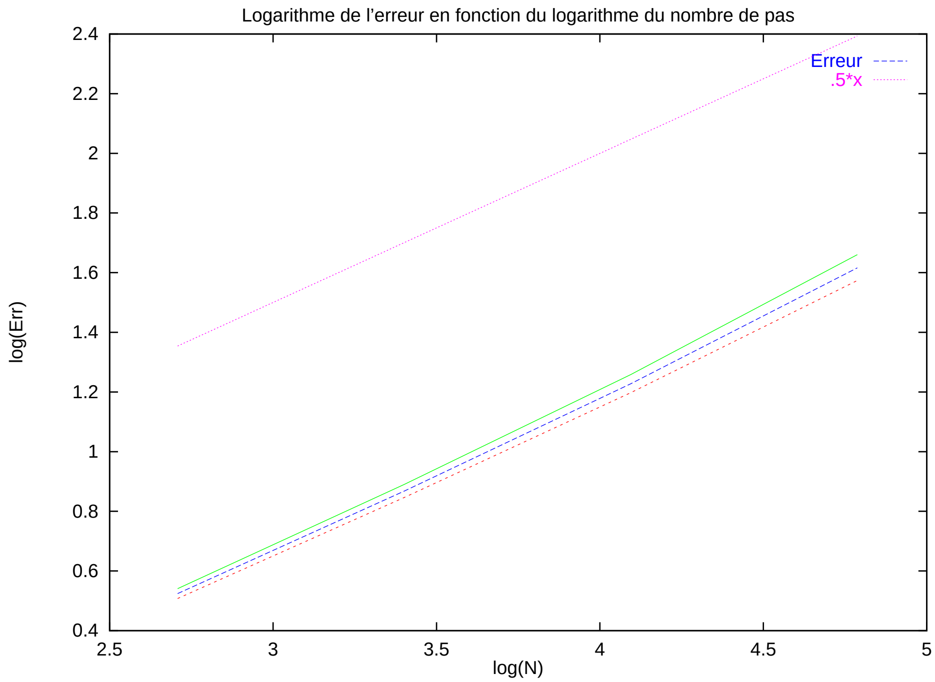
We consider for this example a standard BM in $\mathbb{R}^2$, killed when it leaves the unit ball. Put $T = 1, x = 0, D = \overset{\circ}{B}(0,1), f(y) = (\frac{1}{2} - \|y\|)^+$. The function f satisfies **(F1)**. We let the number of time steps N vary between 15 and 120 following the procedure of Subsection 6.2. We choose $N_{MC} = N_{MCR} = 10^6$, $N_R = 14400$ to compute $E_{N_{MCR}}(f,N^R)$ defined in Subsection 6.1.

We obtain $E_{N_{MCR}}(f,N^R) = .0204271$ and the associated 95% confidence interval writes

$$I_C(N_{MCR}) = [E_{N_{MCR}}(f,N^R) - 1.96\hat{\sigma}_{N_{MCR}}/\sqrt{N}, E_{N_{MCR}}(f,N^R) + 1.96\hat{\sigma}_{N_{MCR}}/\sqrt{N}] = [0.0203672, 0.0204871].$$

We have

	$N = 15$	$N = 30$	$N = 60$	$N = 120$	$N = 240$	$N = 480$
$E_{N_{MC}}(f,N)$	.032521	.0290039	.026411	.0244848	.0231462	.0222035
$\hat{\sigma}^2_{N_{MC}}(f,N)$	.0103078	.00933837	.00860405	.00804047	.00764966	.0073686
$E_{N_{MC}}(f,N) - 1.96\hat{\sigma}_{N_{MC}}(f,N)/\sqrt{N}$	.032322	.0288145	.0262292	.0243091	.0229748	.0220353
$E_{N_{MC}}(f,N) + 1.96\hat{\sigma}_{N_{MC}}(f,N)/\sqrt{N}$	.03272	.0291933	.0265928	.0246606	.0233177	.0223718



The curves bounding the logarithm of the relative error represent the image of the bounds of the confidence interval by the mapping $g(x) = -\log(|x - E_{N_{MCR}}(f, N^R)|/E_{N_{MCR}}(f, N^R))$.

Remark 1.6.1 *The striking point in the above graphic is that the slope obtained by simulation is very close to* $\frac{1}{2}$*. This is a first numerical evidence that emphasizes that an error expansion is likely under our working assumptions. Indeed, assume for a while the error writes* $\mathrm{Err}(T, h, f, x) = C\sqrt{h} + o(\sqrt{h})$ *for* h *small enough. In that case, recall from Theorem 1.5.3 the error is positive,* $\ln \mathrm{Err}(T, h, f, x) = \ln C + \frac{1}{2}\ln h + o(1)$*. That is exactly what we observe on the previous graphic. We also point out that since we represent the logarithm of the* ***relative*** *error in function of the logarithm of the time step the constant* C *can numerically be estimated by* $\exp(b)E_{N_{MCR}}(f, N^R)$ *where* b *is the ordinate to the origin associated to the right line in the previous figure.*

A special Hypoelliptic diffusion

Concerning our concluding remark in the previous paragraph, one could object that the Brownian case is very specific. This is the reason why we consider in this paragraph a general diffusion process whose diffusion matrix satisfies real hypoelliptic assumptions, i.e. the matrix itself can degenerate. The numerical results we obtain are qualitatively very similar to the previous ones.

We now present the discretization error associated to the $\mathbb{R}^2$-valued process

$$X_t = x + \int_0^t \sigma(X_s) dW_s$$

where

$$\forall x \in \mathbb{R}^2, \ \sigma(x) = \begin{pmatrix} 1 & 0 \\ 1 & \sin(x_2) \end{pmatrix}.$$

We set $T = 1, D := \{y \in \mathbb{R}^2 : \langle y, \mathbf{I} \rangle \geq 1\}, \ \mathbf{I} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ f(y) = (\langle y, \mathbf{I} \rangle - 1.2)^+$.

Let us point out that $\sigma\sigma^*(x) = \begin{pmatrix} 1 & 1 \\ 1 & 1 + \sin^2(x_2) \end{pmatrix}$ degenerates along $\{x \in \mathbb{R}^2 : \ x_2 = k\pi, \ k \in \mathbb{Z}\}$.

Anyhow $[\sigma_1,\sigma_2](x) = \begin{pmatrix} 0 \\ \cos(x_2) \end{pmatrix}$ so that, following the definitions of Section 1 we obtain

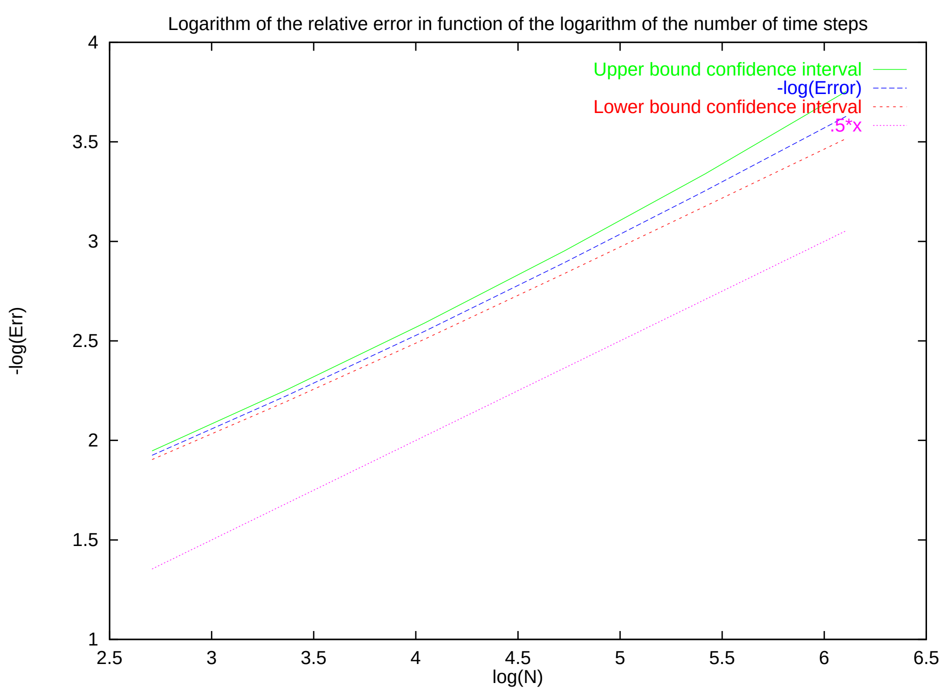
$$\begin{aligned} \forall x \in \mathbb{R}^2,\ \forall z \in \mathbb{R}^2,\ \sum_{y\in L_1(x)} \langle Y, z\rangle^2 &= \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, z \right\rangle^2 + \left\langle \begin{pmatrix} 0 \\ \sin(x_2) \end{pmatrix}, z \right\rangle^2 + \left\langle \begin{pmatrix} 0 \\ \cos(x_2) \end{pmatrix}, z \right\rangle^2 \\ &= \left\langle \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} z, z \right\rangle := \langle \tilde{A} z, z\rangle. \end{aligned}$$

Since $\tilde{A}$ is positive definite we have that **(H)** holds true with $M = 1$. The support condition of **(F1)** is satisfied. One could object that the boundary ∂D is unbounded. Anyhow, the boundedness condition for the boundary is essentially imposed to go from a local description of the boundary to a global one and is not needed for half-spaces. In particular the results of Section 5 remain true for half-spaces, see also Section 7.

We let the number of time steps N vary between 15 and 240 following the procedure of Subsection 6.2. We also chose $N_{MC} = N_{MCR} = 10^6,\ N_R = 14400$ to compute $E_{N_{MCR}}(f, N^R)$. For the reference value we obtain $E_{N_{MCR}}^{N_R} = .929443$ and the associated 95% confidence interval writes $[.925116, 93377]$.

It comes

	$N = 15$	$N = 30$	$N = 60$	$N = 120$	$N = 240$	$N = 480$
$E_{N_{MC}}(f,N)$	1.06498	1.02991	1.00212	0.980918	0.965342	0.954138
$\hat{\sigma}^2_{N_{MC}}(f,N)$	2.23472	2.25	2.23472	2.22405	2.21645	2.21190
$E_{N_{MC}}(f,N) - 1.96\hat{\sigma}_{N_{MC}}(f,N)/\sqrt{N}$	1.06205	1.0269	0.99919	0.977995	0.962424	0.951223
$E_{N_{MC}}(f,N) + 1.96\hat{\sigma}_{N_{MC}}(f,N)/\sqrt{N}$	1.06792	1.03284	1.00505	0.983841	0.96826	0.957052



6.4 Non-zero boundary conditions

Concerning the uniformly elliptic framework, put

$$\Theta(t,x) = \begin{cases} f(x),\ (t,x) \in \{T\} \times \bar{D} \\ g(t,x),\ (t,x) \in [0,T] \times \partial D \end{cases}$$

with $g \in H^{l/2+1,l+2}([0,T]\times\partial D),\ f \in H^{l+2}(\bar{D}),\ l \in (0,1),\ g \neq 0$.

Assume also that f, g satisfy the parabolic compatibility conditions $g(T,x) = f(x),\ \partial_t g(T,x) + Lf(x) = 0,\ x \in \partial D$. Under **(UE)**, **(D-3)**, **(S-$l+1$)**, we derive from Theorem 5.2 p. 320 in [LSU68] that the system

$$\begin{cases} (\partial_t + L)u(t,x) = 0,\ (t,x) \in [0,T]\times D, \\ u(t,.)_{\partial D} = \Theta(t,x), t \in [0,T] \end{cases}$$

has a unique solution $u \in H^{l/2+1,l+2}([0,T]\times\bar{D})$. As in Proposition 1.2.2 one gets $\forall(t,x) \in [0,T]\times\bar{D},\ u(t,x) = v(t,x) = \mathbb{E}_x[\Theta(T\wedge\tau^{t,x}, X^{t,x}_{T\wedge\tau^{t,x}})],\ \tau^{t,x} := \inf\{s \geq t : X^{t,x}_s \notin D\}$. Hence, we still have "good" smoothness properties on v. We now modify the definition of the error and consider

$$\mathrm{Err}_2(T,h,\Theta,x) = \mathbb{E}_x[\Theta(T\wedge\tau^N\wedge\tau^N_R, \pi_{\bar{D}}(X^N_{T\wedge\tau^N\wedge\tau^N_R}))] - \mathbb{E}_x[\Theta(T\wedge\tau, X_{T\wedge\tau})].$$

From a theoretical aspect, the upper bound stated in Theorem 1.5.3 remains valid for $|\mathrm{Err}_2(T,h,\Theta,x)|$. Concerning the lower bound, it is not that easy to derive in this case a positivity property for the normal derivative $\frac{\partial v}{\partial n}$. Indeed, Lemma 1.2.4 strongly relies on the positivity of v inside the domain and on the fact that it vanishes on the boundary.

From a numerical point of view we approximate $\mathbb{E}_x[\Theta(T\wedge\tau^N\wedge\tau^N_R, \pi_{\bar{D}}(X^N_{T\wedge\tau^N\wedge\tau^N_R}))]$ by the following Monte-Carlo procedure

$$E_{N_{MC}}(\Theta,N) := \frac{1}{N_{MC}}\sum_{i=1}^{N_{MC}} \Theta(T\wedge\tau^{N,i}, \pi_{\bar{D}}(X^{N,i}_{T\wedge\tau^{N,i}}))\mathbb{I}_{\tau^{N,i}_R > T}.$$

The last indicator function means that we do not take into account in our empirical mean the paths for which the Euler scheme has exited the region that allowed to uniquely define the projection on $\bar{D}$. Recall anyhow that the probability of such events is a $O_{pol}(h)$, cf. Lemma 1.1.2.

We also introduce

$$\hat{\sigma}^2_{N_{MC}}(\Theta,N) := \frac{1}{N_{MC}-1}\sum_{i=1}^{N_{MC}} \Theta(T\wedge\tau^{N,i}, \pi_{\bar{D}}(X^{N,i}_{T\wedge\tau^{N,i}}))^2\mathbb{I}_{\tau^{N,i}_R > T} - \left(\frac{N_{MC}}{N_{MC}-1}\right)E_{N_{MC}}(\Theta,N)^2$$

which is the usual unbiased estimator of the variance.

Let us now present some numerical results for

$$X_t = \sigma W_t,\ \sigma = \begin{pmatrix} .6 & 0 & 0 \\ .05 & .6 & 0 \\ .02 & .15 & .6 \end{pmatrix},\ W \text{ standard } 3-\text{dimensional BM},$$
$$D = B(0,1),\ \forall(t,x) \in [0,T]\times\bar{D},\ \Theta(t,x) = (T-t)^2 + x_1 + x_2 + x_3.$$

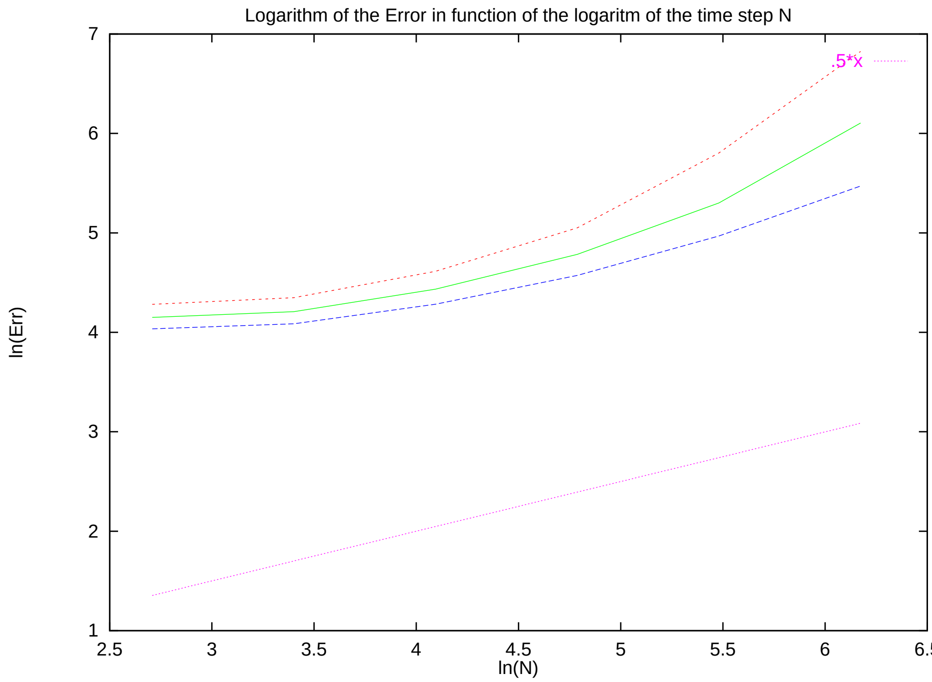
The parabolic compatibility condition $\partial_t\Theta(T,x) + \frac{1}{2}\mathrm{tr}(\sigma\sigma^* H_\Theta(T,x)) = 0$ is readily fulfilled since $\partial_t\Theta(T,x) = H_\Theta(T,x) = 0$.

We let the number of time steps N vary between 15 and 240 following the procedure of Subsection 6.2. We also choose $N_{MC} = N_{MCR} = 10^6,\ N_R = 14400$ to compute $E_{N_{MCR}}(\Theta, N^R)$.

The reference value is $E_{N_{MCR}}(\theta, N_R) = .1357$. The associated 95% confidence interval writes $[0.1341, 0.13766]$.

It comes

	$N=15$	$N=30$	$N=60$	$N=120$	$N=240$	$N=480$
$E_{N_{MC}}(\Theta,N)$	0.119944	0.120825	0.123838	0.127331	0.130715	0.133467
$\hat{\sigma}^2_{N_{MC}}(\Theta,N)$	0.965721	0.980102	0.991336	1.00032	1.00714	1.0123
$E_{N_{MC}}(\Theta,N) - 1.96\hat{\sigma}_{N_{MC}}(\Theta,N)/\sqrt{N}$	0.118017	0.118884	0.121887	0.125371	0.128748	0.131495
$E_{N_{MC}}(\Theta,N) + 1.96\hat{\sigma}_{N_{MC}}(\Theta,N)/\sqrt{N}$	0.12187	0.122765	0.125789	0.129292	0.132682	0.135439



Although we do not observe a "nice" right line as for the examples of Subsection 6.3, the above results do not give numerical evidence that an expansion is unlikely, even when the boundary data is non-zero.

To conclude the section we insist on the fact that numerically speaking, the results of Subsections 6.3 and 6.4 globally tend to show that an error expansion holds true (see also Remark 1.6.1).

All the computations presented in the section have been executed on a Personal Computer, with Pentium processor at **2GHz** and **1GO** of RAM memory. This last memory aspect is not very relevant for these examples since all we need is pure CPU force, and thus a powerful processor (or set of processors since Monte-Carlo approximations are well adapted to parallelism).

We used a Linux Operating System, C++ as programming language and g++ as C++ compiler. No external libraries have been used.

For every presented example, the simulation of the reference value $E_{N_{MCR}}(f, N_R), N_{MCR} = 10^6,\ N_R = 14400$ took about $1h30 - 2$ hours whereas the simulations of $(E_{N_{MC}}(f, N))_{N\le 480}, N_{MC} = 10^6$ took about $15 - 20$ minutes. Note the previous computing times are not proportional in terms of $N \times N_{MC}$. This is a consequence of the simulation procedure described in Subsection 6.2. Indeed, recall that since we keep the same Brownian increments to compute $E_{N_{MC}}(f, N)$ for various $N \le 480$ we have to do intermediary partial sums of these increments. The code could still be optimized but is not naive. The indicated simulation times are significant.

7 Extensions

We conclude giving some extensions of the previous results.

7.1 Half-space

As indicated in Subsection 6.3, the compactness assumption on ∂D may be removed in the half-space case. Indeed, the compactness is needed

- to go from a local description of the boundary to a global one, cf. Section 3.1.

- to give a lower-bound of the normal derivative $\frac{\partial v}{\partial n}(t,x)$ uniformly on $K_T \times \partial D$ where K_T is a given compact interval of $[0,T)$, cf. proof of Theorem 1.5.3.
- to build neigbourhoods that allow to apply the time-occupation formula, cf. proof of Lemma 1.4.1.

We now clarify each of these steps in the special case of the half-space. W.l.o.g., up to a change of coordinate we can consider $D = \mathbb{R}^d_+ := \{x \in \mathbb{R}^d : x_1 > 0\}$.

Concerning the first point, we do not have to straighten the boundary. We simply get $\forall x \in \mathbb{R}^d,\ F(x) = x_1, \pi_{\bar D}(x) = (x_1^+, x_2, ..., x_d)^*$.

For the lower bound we introduce $\tau_K^N := \inf\{t_i : X_{t_i}^N \notin B(x,K)\}$, where K will condition the lower bound for $\dfrac{\partial v}{\partial n}$ on $B(x,K) \cap \partial D$. We write

$$\begin{aligned}&\mathbb{E}_x[\int_0^{T\wedge\tau^N} \frac{\partial v(t,X_t^N)}{\partial n} dL_t^0(F(X^N))] \geq \mathbb{E}_x[\int_0^{T\wedge\tau^N\wedge\tau_K^N} \frac{\partial v(t,X_t^N)}{\partial n}(X_t^N) dL_t^0(F(X^N))]\\ &\geq C_K \mathbb{E}_x[L^0_{T/2\wedge\tau^N\wedge\tau_K^N}(F(X^N))].\end{aligned}$$

Following the proof of Lemma 1.4.6 we get

$$\mathbb{E}_x[L^0_{T/2\wedge\tau^N\wedge\tau_K^N}(F(X^N))] \quad\geq\quad C\int_{c_0/4}^{3c_0/4} dz g(z/\alpha_0')z\sqrt{h}\mathbb{P}_x[F(X^N_{T/4\wedge\tau^N\wedge\tau_K^N}) \leq 0] - Ch^{3/4}.$$

One has

$$\begin{aligned}\mathbb{P}_x[F(X^N_{T/4\wedge\tau^N\wedge\tau_K^N}) \leq 0] \quad&\geq\quad \mathbb{P}_x[\tau^N \leq T/4, \tau_K^N > T/4] \geq \mathbb{P}_x[\tau^N \leq T/4] - \mathbb{P}_x[\tau_K^N \leq T/4]\\ &\geq\quad \mathbb{P}_x[X^N_{T/4} \notin D] - C\exp(-cK^2/T)\end{aligned}$$

where the last inequality is a consequence of Lemma 1.1.2. From Theorem 3.1 in [BT96a] we have the pointwise convergence $\mathbb{P}_x[X^N_{T/4} \notin D] \underset{N}{\longrightarrow} \mathbb{P}[X_{T/4} \notin D]$. Choose K s.t. $C\exp(-cK^2/T) \leq 1/2\mathbb{P}[X_{T/4} \notin D]$. We conclude as in Lemma 1.4.6.

Concerning the last point, in the half-space case assumption **(C)** writes $\forall x \in \mathbb{R}^d,\ |x_1| \leq R,\ \langle \sigma\sigma^*(x)e_1, .e_1\rangle \geq a_0 > 0,\ e_1 = (1,0,...,0)^*$. Hence, we still locally have $dt \leq \dfrac{d\langle F(X^N)\rangle_t}{a_0}$ and can therefore apply the occupation times formula.

Remark that the previous arguments could be applied to a domain delimited by two hyperplanes.

7.2 Function f

The boundedness assumption on f in **(F1)** can also be relaxed into $f(x) \leq C\exp(c|x|)$. This is a direct consequence of the standard control $\mathbb{E}_x[\sup_{t\in[0,T]}\exp(r\|X_t^N\|)] \leq C\exp(r\|x\|)$ for the Euler scheme under **(S-1)**. The above control follows from Lemma 1.1.2.

Unfortunately, it seems difficult to get rid of the support assumption in **(F1)** in the general case, because we are not able to deal with exploding derivatives of v near ∂D.

Chapter 2

Discrete sampling of functionals of Itô processes

Introduction

In Chapter 1 we focused on the analysis of the weak error $\mathrm{Err}(T,h,f,x) = \mathbb{E}_x[f(X_T^N)\mathbb{I}_{\tau^N>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]$ where X was a diffusion process, X^N its Euler scheme , τ was the exit time of X from a given smooth domain D and τ^N the exit time of the Euler scheme along a regular time mesh with step h. In Theorem 1.5.3 we showed that $\mathrm{Err}(T,h,f,x)$ was of order $1/2$ w.r.t. h under suitable assumptions.

One can wonder whether the order $1/2$ stated in Theorem 1.5.3 comes from the discretization of the exit time, from the Euler discretization of the process or if it is related to the Markov property of diffusion processes. In [Gob00], Gobet proved in a uniformly elliptic Markovian setting and for smooth domains that the order $1/2$ was intrinsic to the discrete time killing.

In this chapter we extend this result in a possibly degenerate non-Markovian framework. Since we do not have in that context an underlying PDE we introduce some martingale techniques to analyze the error. This allows to take into account a wider class of functions than in the previous chapter and to handle a certain class of non-smooth domains.

We then turn to the control of the weak error associated to some special discretized functionals (integral, maximum of a process with Markovian diffusion coefficient).

Main results

Let $(X_t)_{0\le t\le T}$ be a d-dimensional Itô process, whose dynamics is given by

$$X_t = x + \int_0^t b_s ds + \int_0^t \sigma_s dW_s \tag{2.0.1}$$

with a fixed initial data x and a fixed terminal time T. Here, W is a d'-dimensional standard Brownian motion (BM in short) defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$, with the usual assumptions on the filtration $(\mathcal{F}_t)_{0\le t\le T}$. The progressively measurable coefficients $(b_s)_{s\ge 0}$ and $(\sigma_s)_{s\ge 0}$ are bounded.

We are more specifically interested in the law of this Itô process, stopped when it exits from some fixed domain D. Namely, for a measurable function g, we consider the quantity

$$\mathbb{E}[g(T\wedge\tau, X_{T\wedge\tau})] \tag{2.0.2}$$

with $\tau := \inf\{t \ge 0 : X_t \notin D\}$, and we focus on the impact of a discretization of the exit time in the above expectation.

Consider a regular mesh of the interval $[0,T]$ with N time steps $(t_i = ih)_{0\le i\le N}$ ($h = T/N$ being the step size), we put $\tau^N := \inf\{t_i \ge 0 : X_{t_i} \notin D\}$. The associated weak error is then defined by

$$\mathrm{Err}(T,h,g,x) := \mathbb{E}[g(T\wedge\tau^N\wedge\tau_R, \pi_{\bar D}(X_{T\wedge\tau^N\wedge\tau_R}))] - \mathbb{E}[g(T\wedge\tau, X_{T\wedge\tau})].$$

The definition of τ_R depends on the domain and is given later. For instance, with convex domains we may take $\tau_R = +\infty$. We should always keep in mind that $\tau^N < \tau_R$ with a very high probability.

For a smooth domain D with compact boundary, if $(\sigma_s)_{s\ge 0}$ satisfies $a.s$ a non-characteristic boundary condition **(C)**, some continuity in probability **(S)** and the function g is supposed to be either bounded measurable with a support condition, cf. **(G1)**, or to be smooth in both variables, cf. **(G2)**, we show that

$$\boxed{\mathrm{Err}(T,h,g,x) = O(\sqrt{h}).}$$

Then, for an intersection of smooth domains with compact boundaries we prove the previous control remains valid under **(G1)** if we additionally assume the process $(\sigma_s\sigma_s^*)_{s\ge 0}$ is $a.s$ uniformly elliptic in a neighborhood of the corners points resulting from the intersection. This is an interesting result even in the Markovian setting of Brownian Motion. Indeed, for non smooth domains it is a hard task to use the traditional error analysis techniques that require the smoothness of the derivatives of the solution of the underlying PDE up to the boundary, see also Chapter 3. We thus provide an alternative technique that points out that the main difficulty to upper-bound the weak error in the Brownian context does not lie in the lack of regularity of the domain.

Finally, we turn to the analysis of some other discretely sampled path-dependent functionals of the process X. Let ψ be a Lipschitz continuous function. Define for all $s \in [0,T]$, $\phi(s) := \inf\{s_i : s_i \le s < s_{i+1}\}$. Simple Fubini like arguments give

$$\boxed{\psi\left(\int_0^T X_s ds\right) - \psi\left(\int_0^T X_{\phi(s)} ds\right) \underset{\mathrm{L}_p(\mathbb{P})}{=} O(h),\ p \ge 1.}$$

For $d = d' = 1$, introduce $M_T := \sup_{s\in[0,T]} X_s,\ M_T^N := \sup_{s\in[0,T]} X_{\phi(s)}$. In the special case $\sigma_s = \sigma(X_s)$ in equation (2.0.1), with σ smooth and uniformly elliptic, we exploit some results from Asmussen and *al.*, cf. [AGP95], to show

$$\boxed{\exists C > 0,\ \mathbb{E}[M_T - M_T^N] \le C\sqrt{h}.}$$

Plan of the chapter

In section 1 we state our working assumptions and also recall some basic properties concerning Itô processes and smooth domains.

In section 2 we give the decomposition of the error $\mathrm{Err}(T,h,g,x)$ and state our first results for smooth domains. A crucial tool for the error analysis is the auxiliary process $V_t := \mathbb{E}[g(T\wedge\tau_t\wedge\tau_R, \pi_{\bar D}(X_{T\wedge\tau_t\wedge\tau_R}))|\mathcal{F}_t]$, where $\tau_t := \inf\{s \ge t : X_s \notin D\}$. This process corresponds to the generalization in a non-Markovian setting of the Feynman-Kac representation of the solution of the underlying PDE for a diffusion process. Namely, if X is a diffusion process with infinitesimal generator L, one has $\forall X_t \in \bar D,\ V_t := v(t, X_t)$ where v is the solution of $(\partial_t + L)v(t,x) = 0,\ \forall (t,x) \in [0,T)\times D,\ v(t,x) = g(t,x),\ \forall (t,x) \in [0,T)\times\partial D \cup \{T\}\times\bar D$. As in the Markovian case, we use V to write $\mathrm{Err}(T,h,g,x)$ as a sum of increments of V. Our strategy then consists in introducing some martingale techniques to control those increments that we can not anymore develop with Itô's formula on v.

To derive the main results of section 2 we need two auxiliary lemmas stated at the end of that section. Their proofs are postponed to section 3.

Section 4 is dedicated to the extension of the previous control of order $1/2$ for $\mathrm{Err}(T, h, g, x)$ in the special case of an intersection of smooth domains when g satisfies a support condition. The key idea is, for each domain of the intersection, to construct an auxiliary process satisfying the non-degeneracy conditions of section 2 in order to apply the results stated there.

We conclude in section 5 with the special cases of the weak errors associated respectively to a discretely sampled integral and to a discretely sampled maximum of a process with Markovian diffusion coefficient.

1 Notations and assumptions

1.1 About the process

We assume the coefficients $(b_s)_{s\in[0,T]}, (\sigma_s)_{s\in[0,T]}$ of (2.0.1) are bounded. Some mild smoothness property on σ (some continuity in probability) will be also needed: the condition stated below is not restrictive at all and is fulfilled for instance as soon as $(\sigma_s)_{0\le s\le T}$ satisfies a Hölder property in L_p-norm.

(S) For any $\delta > 0$, there is some function η_δ with $\lim_{h\to 0^+}\eta_\delta(h) = 0$ such that $a.s$, for $s \in]t_i, t_{i+1}[$ with $X_s \in \partial D$, one has $\mathbb{P}(|\int_s^{t_{i+1}}(\sigma_u - \sigma_s)dW_u| \ge \delta\sqrt{t_{i+1}-s}\,|\,\mathcal{F}_s) \le \eta_\delta(h)$.

1.2 About the domain

We introduce in this section the assumptions concerning "smooth" domains. Our working assumptions for an intersection of smooth domains are specified in section 4.

We define the hypothesis

(D-k) The domain D is of class C^k with bounded boundary ∂D.

For $x \in \partial D$, denote by $n(x)$ the unit inward normal vector at x. For $r \ge 0$, set $V_{\partial D}(r) := \{z \in \mathbb{R}^d : d(z, \partial D) \le r\}$ and $D(r) := \{z \in \mathbb{R}^d : d(z, D) \le r\}$. $B(z, r)$ stands for the closed ball with center z and radius r. The cone of origin $x \in \mathbb{R}^d$, direction $n \in \mathbb{R}^d$ $(n \neq 0)$ and angle $\theta \in (0, \pi)$ is denoted by $\mathcal{K}(x, n, \theta)$.

Assume D satisfies **(D-2)**. Following the notations of Proposition 1.1.1, we now introduce the non characteristic boundary condition

(C) $\exists a_0 > 0, \forall X_s \in V_{\partial D}(R), \forall s \in [0,T],\ \alpha_s := \nabla F(X_s).\sigma_s\sigma_s^*\nabla F(X_s) \ge a_0,\ a.s.$

Assumption **(C)** is somehow a minimal condition to ensure a convergent approximation. Indeed, it easy to imagine a deterministic path which hits ∂D only at time $\tau = \chi T$ where χ is an irrational number in $]0, 1[$: for this, $\tau^N > T$ for any $N \ge 1$ and $\mathrm{Err}_1(T, h, f, x) = f(X_T)$ is constant.

1.3 About the function

In the following, the function g in (2.0.2) satisfies one of the two assumptions below

(G1) g is a bounded Borel function and $\exists\varepsilon > 0,\ \forall t \in [0,T],\ d(\mathrm{supp}(g(t,.)), \partial D) \ge 2\varepsilon$.

(G2) $g \in C_b^{1,2}([0,T]\times\mathbb{R}^d)$ where $C_b^{1,2}([0,T]\times\mathbb{R}^d)$ is the space of real-valued bounded functions on $[0,T]\times\mathbb{R}^d$ with continuous and bounded derivatives up to order 1 in time, up to order 2 in space.

We denote by $\partial_x^\alpha g(t,x)$ the derivative of g w.r.t. x according to the multi-index α, whereas time derivatives of g are denoted by $\partial_t g(t,x), \partial_t^2 g(t,x), \ldots$. The notation $\nabla g(t,x)$ (resp. $H_g(t,x)$) stands for the gradient (resp. the Hessian matrix) of g w.r.t. x and $\frac{\partial g}{\partial n}(t,x) = \nabla g(t,x).n(x)$ is the normal derivative on the boundary.

1.4 Miscellaneous

The distribution function of the standard normal law is denoted by Φ.

We will keep the same notation C (or C') for all finite, non-negative constants which will appear in our computations: they may depend on D, T, b, σ, ε or g, but they will not depend on the number of time steps N and the initial value x. We reserve the notation c and c' for constants also independent of x, T and g.

In the following $O_{pol}(h)$ (resp. $O(h)$) stands for every quantity $R(h)$ such that $\forall n \in \mathbb{N}$, for some $C > 0$, one has $|R(h)| \le Ch^n$ (resp. $|R(h)| \le Ch$) (uniformly in x).

1.5 Usual controls

We recall some basic estimates that only exploit the boundedness of the coefficients.

Lemma 2.1.1 (Bernstein's type inequality) *Consider two stopping times S, S' upper bounded by T with $0 \le S' - S \le \Delta \le T$. Then for any $p \ge 1$ and $c' > 0$, there are some constants $c > 0$ and C, such that for any $\eta \ge 0$, one has a.s:*

$$\begin{aligned}
\mathbb{P}[\sup_{t\in[S,S']} \|X_t - X_S\| \ge \eta \mid \mathcal{F}_S] \le & C \exp\left(-c\frac{\eta^2}{\Delta}\right), \\
\mathbb{E}[\sup_{t\in[S,S']} \|X_t - X_S\|^p \mid \mathcal{F}_S] \le & C\Delta^{p/2}, \\
\mathbb{E}[\exp\left(-c'\frac{d^2(X_{S'}, \partial D)}{\Delta}\right) \mid \mathcal{F}_S] \le & C \exp\left(-c\frac{d^2(X_S, \partial D)}{\Delta}\right).
\end{aligned}$$

Proof. It is similar to the one of Lemma 1.1.2.

2 Main decomposition of the error and first results

In this section we assume **(D-2)** is in force. The constant R is the one of Proposition 1.1.1. In particular, on $D(R)$ the projection on $\bar{D}$ is uniquely defined and we recall that if the domain D is convex one can set $R = +\infty$. Put $\tau_R := \inf\{s \ge 0 : X_s \notin D(R)\}$, $\tau^{N,R} := \tau_R \wedge \tau^N$. We recall that from Lemma 2.1.1 $\mathbb{P}[\tau_R < \tau^N] \le C \exp\left(-c\frac{R^2}{h}\right)$. Thus, taking $\tau^{N,R}$ instead of τ^N has not a significant impact. This is only a technical restriction needed to keep the projection on $\bar{D}$ well defined.

Put $\forall (t,z) \in [0,T] \times D(R)$, $\tilde{g}(t,z) := g(t, \pi_{\bar{D}}(z))$. The main result of the section is the following

Theorem 2.2.1 Upper bound in the smooth case
Assume **(C)**, **(D-2)**, **(S)** *and either* **(G1)** *or* **(G2)**. *For some constant C, one has*

$$|\mathrm{Err}(T,h,g,x)| = |\mathbb{E}[\tilde{g}(T\wedge\tau^{N,R}, X_{T\wedge\tau^{N,R}}) - \tilde{g}(T\wedge\tau, X_{T\wedge\tau})]| \le C\sqrt{h}.$$

Remark 2.2.1 *Note that under* **(G1)**, *if g is non-negative one also has $\mathrm{Err}(T,h,g,x) \ge 0$. This readily derives from the inequality $\tau^{N,R} \ge \tau$ a.s.*

Proof of Theorem 2.2.1. The error writes

$$\begin{aligned}
\mathrm{Err}(T,h,g,x) &= \mathbb{E}[\tilde{g}(T\wedge\tau^{N,R}, X_{T\wedge\tau^{N,R}}) - \tilde{g}(T\wedge\tau, X_{T\wedge\tau})] \\
&= \mathbb{E}[\mathbb{I}_{\tau<T}\mathbb{E}[\tilde{g}(T\wedge\tau^{N,R}, X_{T\wedge\tau^{N,R}}) - \tilde{g}(\tau, X_\tau)|\mathcal{F}_\tau]].
\end{aligned}$$

Hence, to show Theorem 2.2.1 it is enough to derive

$$|\mathcal{E}| := |\mathbb{E}[\tilde{g}(T'\wedge\tau^{N,R'}, X_{T'\wedge\tau^{N,R'}}) - \tilde{g}(t,x)]| \le C\sqrt{h}, \tag{2.2.1}$$

for an initial point $x \in \partial D$, $t \in [0,T)$, for a shifted time mesh defined by $\{t_i : 0 \le i \le N'\}$ with $t_0 = 0$, $0 < t_1 \le h$, $t_{i+1} = t_i + h$ $(i \ge 1)$, for a new terminal time $T' = t_{N'}$ and a modified exit time $\tau^{N'} = \inf\{t_i \ge t_1 : X_{t_i} \notin D\}$. The constant C in (2.2.1) has to be uniform in T' in a compact set, in N', in x and in t. For notational convenience, we still write N for N', T for T' and take $t = 0$.

Introduce now for all $s \in [0,T]$, on the set $\{\tau^R \ge s\}$, $V_s := \mathbb{E}[\tilde{g}(T\wedge\tau_s, X_{T\wedge\tau_s})|\mathcal{F}_s]$ where $\tau_s := \inf\{u \ge s :$

$X_u \notin D\}$. For $x \in \partial D$, $\tau_0 = 0$ so $V_0 = g(0,x)$. On the other hand $V_{T\wedge\tau^{N,R}} = \tilde{g}(T\wedge\tau^{N,R}, X_{T\wedge\tau^{N,R}})$. Thus,

$$\begin{aligned}
\mathcal{E} &= \mathbb{E}[V_{T\wedge\tau^{N,R}}] - V_0 = \sum_{i=0}^{N-1}\mathbb{E}[V_{t_{i+1}\wedge\tau^{N,R}} - V_{t_i\wedge\tau^{N,R}}]\\
&= \sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^{N,R}>t_i}\left(V_{t_{i+1}\wedge\tau_R} - V_{t_i}\right)]\\
&= \sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^{N,R}>t_i}\left(V_{t_{i+1}\wedge\tau_R} - V_{t_{i+1}\wedge\tau_{t_i}}\right)] + \mathbb{E}[\mathbb{I}_{\tau^{N,R}>t_i}\left(V_{t_{i+1}\wedge\tau_{t_i}} - V_{t_i}\right)]\\
&:= \sum_{i=0}^{N-1}\mathcal{E}_1^i + \mathcal{E}_2^i := \mathcal{E}_1 + \mathcal{E}_2.
\end{aligned}$$

Term $\mathcal{E}_2$.

Let us prove that $\forall i \in [\![0, N-1]\!], \mathcal{E}_2^i = 0$. Write $\mathcal{E}_2^i = \mathbb{E}[\mathbb{I}_{\tau^{N,R}>t_i}\mathbb{E}[V_{t_{i+1}\wedge\tau_{t_i}} - V_{t_i}|\mathcal{F}_{t_i}]]$.

- On $\{t_{i+1} < \tau_{t_i}\}$, $V_{t_{i+1}\wedge\tau_{t_i}} = V_{t_{i+1}} = \mathbb{E}[\tilde{g}(T\wedge\tau_{t_{i+1}}, X_{T\wedge\tau_{t_{i+1}}})|\mathcal{F}_{t_{i+1}}]$.
- On $\{t_{i+1} \geq \tau_{t_i}\}$, $V_{t_{i+1}\wedge\tau_{t_i}} = V_{\tau_{t_i}} = \tilde{g}(\tau_{t_i}, X_{\tau_{t_i}})$.

Turning to the former definition of V it comes

$$\begin{aligned}
\mathbb{E}[V_{t_{i+1}\wedge\tau_{t_i}} - V_{t_i}|\mathcal{F}_{t_i}] &= \mathbb{E}[\tilde{g}(T\wedge\tau_{t_{i+1}\wedge\tau_{t_i}}, X_{T\wedge\tau_{t_{i+1}\wedge\tau_{t_i}}}) - \tilde{g}(T\wedge\tau_{t_i}, X_{T\wedge\tau_{t_i}})|\mathcal{F}_{t_i}]\\
&= \mathbb{E}[\mathbb{I}_{t_{i+1}<\tau_{t_i}}(\tilde{g}(T\wedge\tau_{t_{i+1}}, X_{T\wedge\tau_{t_{i+1}}}) - \tilde{g}(T\wedge\tau_{t_i}, X_{T\wedge\tau_{t_i}}))|\mathcal{F}_{t_i}]\\
&+ \mathbb{E}[\mathbb{I}_{t_{i+1}\geq\tau_{t_i}}(\tilde{g}(\tau_{t_i}, X_{\tau_{t_i}}) - \tilde{g}(\tau_{t_i}, X_{\tau_{t_i}}))|\mathcal{F}_{t_i}]\\
&= 0
\end{aligned}$$

since on the event $\{t_{i+1} < \tau_{t_i}\}$ one has $\tau_{t_i} = \tau_{t_{i+1}}$. Therefore we have

$$\mathcal{E} = \sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^{N,R}>t_i}\mathbb{I}_{\tau_{t_i}<t_{i+1}}(V_{t_{i+1}\wedge\tau_R} - V_{\tau_{t_i}})] := \mathcal{E}_1. \tag{2.2.2}$$

Term $\mathcal{E}_1$.

To control this term we state two auxiliary Lemmas whose proofs are postponed to section 3.

Lemma 2.2.2 *Assume* **(C)**, **(D-**2**)**, **(S)** *and either* **(G1)** *or* **(G2)**. *For all* $i \in [\![0, N-1]\!]$, *on the set* $\{\tau^{N,R} > t_i, \tau_{t_i} < t_{i+1}\}$ *one has*

$$|\mathbb{E}[V_{t_{i+1}\wedge\tau_R} - V_{\tau_{t_i}}|\mathcal{F}_{\tau_{t_i}}]| \leq C\sqrt{h}.$$

Lemma 2.2.3 *Assume* **(C)**, **(D-2)** *and* **(S)**. *There are some positive constants* C *and* N_0 *such that for* $N \geq N_0$, *for any* $i \in [\![0, N-1]\!]$, *one has for* $X_{t_i} \in D$

$$\mathbb{P}[\exists t \in [t_i, t_{i+1}] : X_t \notin D \mid \mathcal{F}_{t_i}] \leq C\ \mathbb{P}[X_{t_{i+1}} \notin D \mid \mathcal{F}_{t_i}].$$

Plugging the control of Lemma 2.2.2 into (2.2.2) we obtain

$$|\mathcal{E}| \leq C\sqrt{h}\sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^{N,R}>t_i}\mathbb{I}_{\tau_{t_i}<t_{i+1}}].$$

Using now Lemma 2.2.3 it comes

$$|\mathcal{E}| \leq C\sqrt{h}\sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^N>t_i}\mathbb{I}_{X_{t_{i+1}}\notin D}] = C\sqrt{h}\sum_{i=0}^{N-1}\mathbb{P}[\tau^N = t_{i+1}] \leq C\sqrt{h}$$

which completes the proof of Theorem 2.2.1.

$\square$

Note that under **(G1)**, because of the support condition one can define $\forall t \in [0,T]$, $\tilde V_t := \mathbb{E}[f(X_T)\mathbb{I}_{\tau_t>T}|\mathcal{F}_t]$ where $f(.) = g(T,.)$. On $\{t < \tau^R\}$ one has $\tilde V_t = V_t$. Introduce now

$$\begin{aligned}\mathrm{Err}(T,h,f,x) &:= \mathbb{E}[f(X_T)\mathbb{I}_{\tau^N>T}] - \mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}]\\ &= \mathrm{Err}(T,h,g,x) + \mathbb{E}[f(X_T)(\mathbb{I}_{\tau^N>T} - \mathbb{I}_{\tau^{N,R}>T})] := \mathrm{Err}(T,h,g,x) + R_N.\end{aligned}$$

Lemma 2.1.1 and **(G1)** yield $|R_N| \le \|f\|_\infty \mathbb{P}[\tau^N > T, \tau_R < T] = O_{pol}(h)$.

Hence, as a consequence of Theorem 2.2.1 we have the following

Corollary 2.2.4 *Assume* **(C)**, **(D-2)**, **(S)** *and* **(G1)**. *For some constant C, one has*

$$|\mathrm{Err}(T,h,f,x)| = |\mathbb{E}[f(X_T)\mathbb{I}_{\tau^N>T}] - \mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}]| \le C\sqrt{h}.$$

Remark 2.2.2 *Let us first recall that the results of Theorem 2.2.1 and Corollary 2.2.4 concern the impact of a discretization time in the quantity (2.0.2). They can therefore not be directly compared to the results of Theorem 1.5.3 of the previous chapter except in the special case of Brownian motion. Note anyhow that in that case we obtain the upper bound of the weak error with a much simpler proof.*

The next natural question, under **(G1)** *and when $f \ge 0$, concerns a possible lower bound of the same order for* $\mathrm{Err}(T,h,f,x)$ *as stated in Theorem 1.5.3 in a Markovian framework. We give a counter example that illustrates this property can fail under the sole assumption* **(C)**.

Define for all $t \ge 0$, the one dimensional diffusion process $X_t = \pi/2 + \int_0^t \cos(X_s)ds + \int_0^t \sin(X_s)dW_s$ and put $D :=]-\pi/2, 3\pi/2[$. **(C)** *is readily satisfied and by construction one has $X_s \in [0,\pi]$ a.s. Hence, $\mathbb{I}_{\tau^N>T} = \mathbb{I}_{\tau>T} = 1$ and $\mathrm{Err}(T,h,f,x) = 0$. A minimal necessary condition to have a lower bound of order $1/2$ w.r.t h is to reach the boundary on the interval $[0,T]$ with positive probability.*

3 Proof of the technical Lemmas

This section is devoted to the proof of Lemmas 2.2.2 and 2.2.3.

3.1 Proof of Lemma 2.2.2

For this proof we distinguish the cases associated to the assumptions **(G1)** and **(G2)**.

Proof under (G1)

In that case Lemma 2.2.2 is a direct consequence of the following

Lemma 2.3.1 *Assume* **(C)**, **(D-2)**, **(S)** *and* **(G1)**. *There is some constant C such that for any $t \in [0,T]$, on the set $\{\tau_R \ge t\}$ one has a.s*

$$|V_t| \le C\frac{\|f\|_\infty}{1\wedge\varepsilon}[F(X_t)]_+. \tag{2.3.1}$$

Indeed, $\forall i \in [\![0, N-1]\!]$, on $\{\tau^{N,R} > t_i, \tau_{t_i} \le t_{i+1}\}$ one has

$$\mathbb{E}[V_{t_{i+1}\wedge\tau_R} - V_{\tau_{t_i}}|\mathcal{F}_{\tau_{t_i}}] = \mathbb{E}[V_{t_{i+1}\wedge\tau_R}|\mathcal{F}_{\tau_{t_i}}]. \tag{2.3.2}$$

Under **(D-2)** we derive from Proposition 1.1.1 that $F \in C_b^2(\mathbb{R}^d)$. Hence, plugging the control (2.3.1) of Lemma 2.3.1 into (2.3.2) we get

$$\begin{aligned}|\mathbb{E}[V_{t_{i+1}\wedge\tau_R} - V_{\tau_{t_i}}|\mathcal{F}_{\tau_{t_i}}]| &\le C\frac{\|f\|_\infty}{1\wedge\varepsilon}\mathbb{E}[[F(X_{t_{i+1}\wedge\tau_R})]_+ - [F(X_{\tau_{t_i}})]_+|\mathcal{F}_{\tau_{t_i}}] \le C\frac{\|f\|_\infty}{1\wedge\varepsilon}\mathbb{E}[\|X_{t_{i+1}\wedge\tau_R} - X_{\tau_{t_i}}\||\mathcal{F}_{\tau_{t_i}}]\\ &\le C\sqrt{h}\frac{\|f\|_\infty}{1\wedge\varepsilon}\end{aligned}$$

where we used the Lipschitz property of F_+ for the second inequality, Lemma 2.1.1 for the last one.

It therefore remains to prove Lemma 2.3.1.

Proof of Lemma 2.3.1

W.l.o.g. we assume $f \geq 0$. Since $V_t = 0$ for $X_t \notin D$ on the set $\{\tau_R \geq t\}$, it is enough to prove the estimate for $X_t \in D \cap V_{\partial D}(R \wedge \varepsilon/2)$ for which $0 < F(X_t) \leq R \wedge \varepsilon/2$. Denote $\tau_t^R = \inf\{s \geq t : F(X_s) \geq R\}$ and split V into two parts $V_t = V_t^1 + V_t^2$ with $V_t^1 = \mathbb{E}\big[\mathbb{I}_{T<\tau_t}\mathbb{I}_{T<\tau_t^R} f(X_T) \mid \mathcal{F}_t\big]$ and $V_t^2 = \mathbb{E}\big[\mathbb{I}_{T<\tau_t}\mathbb{I}_{T\geq\tau_t^R} f(X_T) \mid \mathcal{F}_t\big]$ where $f(.) = g(T,.)$.

Before estimating separately each contribution, we set some standard notations related to time-changed Brownian martingales.

Define the increasing continuous process $\mathcal{A}_s = \int_t^s \alpha_u du$ (from $[t, +\infty[$ into $\mathbb{R}^+$) and its increasing right-continuous inverse $\mathcal{C}_s = \inf\{u \geq t : \mathcal{A}_u > s\}$ (from $\mathbb{R}^+$ into $[t, +\infty[$) (see section V.1 in Revuz-Yor [RY99]) and put $M_s = \int_t^{\mathcal{C}_s} \nabla F(X_u).\sigma_u dW_u$, $Z_s = F(X_{\mathcal{C}_s})$. From the Dambis-Dubins-Schwarz theorem, M coincides with a standard BM β (defined on a possibly enlarged probability space) for $s < \int_t^\infty \alpha_u du$ and it is easy to check that β is independent of $\mathcal{F}_t$ (see the arguments in the proof of Theorem V.1.7 in [RY99]).
Owing to the assumption **(C)**, $\mathcal{A}$ and $\mathcal{C}$ are strictly increasing on $[t, \tau_t^R]$ and $[0, \int_t^{\tau_t^R} \alpha_u du]$. Thus, for $s \in [0, \int_t^{\tau_t^R} \alpha_u du]$, one easily obtains

$$Z_s = F(X_t) + \beta_s + \int_0^s \lambda_v dv$$

where $\lambda_v = \{[\nabla F(X_u).b_u + \frac{1}{2}\mathrm{tr}(H_F(X_u)\sigma_u\sigma_u^*)]|_{u=\mathcal{C}_v}\}\frac{1}{\alpha_{\mathcal{C}_v}}$ is bounded by a constant C_λ. Define

$$Z'_s = F(X_t) + \beta_s + C_\lambda s \geq Z_s. \tag{2.3.3}$$

Finally, put $\tau_0^Z = \inf\{s \geq 0 : Z_s \leq 0\}$, $\tau_R^Z = \inf\{s \geq 0 : Z_s \geq R\}$ and analogously $\tau_0^{Z'}$, $\tau_R^{Z'}$ for Z'.

<u>*Estimation of* V^1.</u> Let us first prove that for any stopping time $S \in [t, T]$, one has

$$\begin{aligned}\mathbb{E}\big[f(X_T) \mid \mathcal{F}_S\big] \leq& \|f\|_\infty \mathbb{P}[F(X_T) \geq 2\varepsilon \mid \mathcal{F}_S] \\ \leq& C\|f\|_\infty \exp\big(-c\frac{(2\varepsilon - F(X_S))_+^2}{T-S}\big) \ a.s.\end{aligned} \tag{2.3.4}$$

The first inequality simply results from the support of f included in $D\backslash V_{\partial D}(2\varepsilon)$ (assumption **(G1)**). To justify the second one, note that $\{F(X_T) \geq 2\varepsilon\} \subset \{|F(X_T) - F(X_S)| \geq 2\varepsilon - F(X_S)\} \subset \{|F(X_T) - F(X_S)| \geq (2\varepsilon - F(X_S))_+\}$ and the proof of (2.3.4) is complete using Lemma 1.1.2 applied to the Itô process $(F(X_s))_{s\geq 0}$ with bounded coefficients.
We now turn to the evaluation of V_t^1. On $\{T < \tau_t^R\}$, using the notation with the time change above, one has $T = \mathcal{C}_{\mathcal{A}_T} \geq \mathcal{C}_{a_0(T-t)}$ and $a_0(T - \mathcal{C}_{a_0(T-t)}) \leq \displaystyle\int_{\mathcal{C}_{a_0(T-t)}}^T \alpha_u du = \mathcal{A}_T - \mathcal{A}_{\mathcal{C}_{a_0(T-t)}}$. Hence, $T - \mathcal{C}_{a_0(T-t)} \leq \frac{1}{a_0}(\mathcal{A}_T - a_0(T-t)) \leq \frac{\|\alpha\|_\infty}{a_0}(T-t)$. Thus, one obtains

$$\begin{aligned}V_t^1 &\leq \mathbb{E}\big[\mathbb{I}_{\mathcal{C}_{a_0(T-t)}<\tau_t}\mathbb{I}_{\mathcal{C}_{a_0(T-t)}<\tau_t^R}\mathbb{I}_{T-\mathcal{C}_{a_0(T-t)}\leq\frac{\|\alpha\|_\infty}{a_0}(T-t)}\mathbb{E}\big[f(X_T) \mid \mathcal{F}_{\mathcal{C}_{a_0(T-t)}}\big] \mid \mathcal{F}_t\big] \\ &\leq C\|f\|_\infty \mathbb{E}\big[\mathbb{I}_{\mathcal{C}_{a_0(T-t)}<\tau_t}\mathbb{I}_{\mathcal{C}_{a_0(T-t)}<\tau_t^R} \exp\big(-c'\frac{(2\varepsilon - F(X_{\mathcal{C}_{a_0(T-t)}}))_+^2}{T-t}\big) \mid \mathcal{F}_t\big] \\ &\leq C\|f\|_\infty \mathbb{E}\big[\mathbb{I}_{a_0(T-t)<\tau_0^{Z'}}\mathbb{I}_{\mathcal{C}_{a_0(T-t)}<\tau_t^R} \exp\big(-c'\frac{(2\varepsilon - Z'_{a_0(T-t)})_+^2}{T-t}\big) \mid \mathcal{F}_t\big]\end{aligned}$$

where one has applied at the second line the estimate (2.3.4) with $S = \mathcal{C}_{a_0(T-t)}$ (here $c' = c\frac{a_0}{\|\alpha\|_\infty}$), at the third one $\{\mathcal{C}_{a_0(T-t)} < \tau_t\} = \{\forall s \in [t, \mathcal{C}_{a_0(T-t)}] : F(X_s) > 0\} = \{\forall u \in [0, a_0(T-t)] : Z_u > 0\} = \{a_0(T-t) < \tau_0^Z\} \subset \{a_0(T-t) < \tau_0^{Z'}\}$ and $(2\varepsilon - F(X_{\mathcal{C}_{a_0(T-t)}}))_+ = (2\varepsilon - Z_{a_0(T-t)})_+ \geq (2\varepsilon - Z'_{a_0(T-t)})_+$. Reminding the law of β, one finally gets that $V_t^1 \le C\|f\|_\infty \Phi_1(a_0(T-t), F(X_t))$ with $\Phi_1(r,z) = \mathbb{E}\big(\mathbb{I}_{\forall u\in[0,r]:z+\beta_u+C_\lambda u>0} \exp\big(-a_0 c' \frac{(2\varepsilon - z-\beta_r - C_\lambda r)_+^2}{r}\big)\big)$. With clear notations involving the smooth transition density of the killed drifted BM (see Chapters 1 and 3) and Gaussian type estimates of its gradient (see [LSU68] Theorem 16.3), one has $\Phi_1(r,z) = \int_0^\infty q_r(z,y)\exp\big(-a_0c'\frac{(2\varepsilon-y)_+^2}{r}\big)dy$ and

$$|\partial_z \Phi_1(r,z)| \le C\int_0^\infty \frac{1}{r}\exp(-c\frac{(z-y)^2}{r})\exp\big(-a_0c'\frac{(2\varepsilon-y)_+^2}{r}\big)dy.$$

We now justify that $|\partial_z \Phi_1(r,z)| \le \frac{C}{1\wedge\varepsilon}$ for $0\le z\le \epsilon/2$ and for this, we may split the domain of integration into two parts. For $y<\varepsilon$, $(2\varepsilon-y)_+^2 \ge \varepsilon^2$ and the corresponding contribution for the integral is bounded by $\int_0^\infty \frac{1}{\sqrt{r}}\exp(-c\frac{(z-y)^2}{r})\big[\frac{1}{\sqrt{r}}\exp\big(-a_0c'\frac{\varepsilon^2}{r}\big)\big]dy \le \frac{C}{1\wedge\varepsilon}$. For $y\ge\varepsilon$ and $0\le z\le\varepsilon/2$, $(z-y)^2\ge\varepsilon^2/4$ and the integral is bounded by $\int_0^\infty \frac{1}{\sqrt{r}}\exp(-\frac{c}{2}\frac{(z-y)^2}{r})\frac{1}{\sqrt{r}}\exp(-\frac{c}{2}\frac{\varepsilon^2}{4r})dy \le \frac{C}{1\wedge\varepsilon}$.
Since $\Phi_1(r,0)=0$, one gets $\Phi_1(r,z)\le\frac{C}{1\wedge\varepsilon}z$ for $z\in[0,\varepsilon/2]$ and this proves that $V_t^1 \le C\frac{\|f\|_\infty}{1\wedge\varepsilon}F(X_t)$.

Estimation of V^2. Clearly, one has $V_t^2 \le \|f\|_\infty \mathbb{P}\big[\tau_t^R < \tau_t \mid \mathcal{F}_t\big]$. Note that $\{\tau_t^R<\tau_t\} = \{\tau_R^Z < \tau_0^Z\} \subset \{\tau_R^{Z'} < \tau_0^{Z'}\}$ because of (2.3.3). Hence, one has $V_t^2 \le \|f\|_\infty \Phi_2(F(X_t))$ where $\Phi_2(z) = \mathbb{P}[(z+\beta_u + C_\lambda u)_{u\ge0}$ hits R before $0]$. It is well-known that $\Phi_2(z) = \frac{1-\exp(-2C_\lambda z)}{1-\exp(-2C_\lambda R)} \le Cz$ (see Section 5.5 in [KS91] e.g.) and this proves that $V_t^2 \le C\|f\|_\infty F(X_t)$.

Combining estimates for V^1 and V^2 gives the result of Lemma 2.3.1.

□

Proof under (G2)

In this case, we use the smoothness of g. Since we also stop X_t in τ_R the semi-martingale decomposition stated in Proposition 1.3.2 remains valid for $(\pi_{\bar{D}}(X_{t\wedge\tau_R}))_{t\ge0}$. Hence, $\forall i\in [\![0,N-1]\!]$, on the set $\{\tau_{t_i}\le t_{i+1}\}$ we write

$$\begin{aligned}
&\tilde{g}(T\wedge\tau_{t_{i+1}\wedge\tau_R}, X_{T\wedge\tau_{t_{i+1}\wedge\tau_R}}) - \tilde{g}(\tau_{t_i}, X_{\tau_{t_i}})\\
&= \int_{\tau_{t_i}}^{T\wedge\tau_{t_{i+1}\wedge\tau_R}} \partial_u g(u,\pi_{\bar{D}}(X_u))du + \nabla g(u,\pi_{\bar{D}}(X_u))\cdot d(\pi_{\bar{D}}(X_u)) + \frac{1}{2}\mathrm{tr}(H_g(u,\pi_{\bar{D}}(X_u))d\langle\pi_{\bar{D}}(X)\rangle_u)\\
&:= (M_{T\wedge\tau_{t_{i+1}\wedge\tau_R}} - M_{\tau_{t_i}}) + (V_{T\wedge\tau_{t_{i+1}\wedge\tau_R}} - V_{\tau_{t_i}}) + \int_{\tau_{t_i}}^{T\wedge\tau_{t_{i+1}\wedge\tau_R}} \frac{\partial g}{\partial n}(u,X_u)dL_u^0(F(X))
\end{aligned}$$

where M is a local martingale and V a finite variation process. From the boundedness of the derivatives of g and of the coefficients b_s, σ_s, we derive that M is a true martingale and that $a.s$ $|V_{T\wedge\tau_{t_{i+1}\wedge\tau_R}} - V_{\tau_{t_i}}| \le C(T\wedge\tau_{t_{i+1}\wedge\tau_R} - \tau_{t_i})$.

It comes

$$\begin{aligned}
&|\mathbb{E}[\tilde{g}(T\wedge\tau_{t_{i+1}\wedge\tau_R}, X_{T\wedge\tau_{t_{i+1}\wedge\tau_R}}) - \tilde{g}(\tau_{t_i}, X_{\tau_{t_i}})|\mathcal{F}_{\tau_{t_i}}]| \le C\Big\{\mathbb{E}[L^0_{T\wedge\tau_{t_{i+1}\wedge\tau_R}}(F(X)) - L^0_{\tau_{t_i}}(F(X))|\mathcal{F}_{\tau_{t_i}}]\\
&+\mathbb{E}[(T\wedge\tau_{t_{i+1}\wedge\tau_R} - \tau_{t_i})|\mathcal{F}_{\tau_{t_i}}]\Big\} := C\left(A^1_{\tau_{t_i}} + A^2_{\tau_{t_i}}\right).
\end{aligned}$$

Term $A^1_{\tau_{t_i}}$: control of the local time.

Since the measure $dL_t^0(F(X))$ is $a.s$ carried by the set $\{t : F(X_t)=0\}$ we write

$$\begin{aligned}
A^1_{\tau_{t_i}} &= \mathbb{E}[L^0_{t_{i+1}\wedge\tau_R}(F(X)) - L^0_{\tau_{t_i}}(F(X))|\mathcal{F}_{\tau_{t_i}}]\\
&\le 2\left(\mathbb{E}[F^-(X_{t_{i+1}\wedge\tau_R}) - F^-(X_{\tau_{t_i}}) + \int_{\tau_{t_i}}^{t_{i+1}\wedge\tau_R}\mathbb{I}_{F(X_s)<0}dF(X_s)|\mathcal{F}_{\tau_{t_i}}]\right)\\
&\le C(\mathbb{E}[\|X_{t_{i+1}\wedge\tau_R} - X_{\tau_{t_i}}\||\mathcal{F}_{\tau_{t_i}}] + h) \le C\sqrt{h}.
\end{aligned} \tag{2.3.5}$$

The last two inequalities are a consequence of the boundedness of F and its derivatives, the boundedness of the coefficients of X and Lemma 2.1.1.

Term $A^2_{\tau_{t_i}}$: time-change techniques

Write

$$A^2_{\tau_{t_i}} = (T-\tau_{t_i})\mathbb{P}[\tau_{t_{i+1}\wedge\tau_R} > T|\mathcal{F}_{\tau_{t_i}}] + \mathbb{E}[(\tau_{t_{i+1}\wedge\tau_R} - \tau_{t_i})\mathbb{I}_{\tau_{t_{i+1}\wedge\tau_R}\le T}|\mathcal{F}_{\tau_{t_i}}] := A^{21}_{\tau_{t_i}} + A^{22}_{\tau_{t_i}}.$$

The key idea is now, as in the proof of Lemma 2.3.1, to use time-changes in order to apply well known results for hitting times in a Brownian framework.

We rewrite

$$A^{21}_{\tau_{t_i}} = (T-\tau_{t_i})\mathbb{E}[\mathbb{I}_{\tau_R>t_{i+1}}\mathbb{I}_{X_{t_{i+1}}\in D}\mathbb{E}[\mathbb{I}_{\tau_{t_{i+1}}>T}|\mathcal{F}_{t_{i+1}}]|\mathcal{F}_{\tau_{t_i}}].$$

Put $C_{t_{i+1}} := \mathbb{P}[\tau_{t_{i+1}} > T|\mathcal{F}_{t_{i+1}}]$ and define $\tau^R_t := \inf\{s\ge t : F(X_s)\ge R\}$. We decompose $C_{t_{i+1}} = \mathbb{P}[\tau_{t_{i+1}} > T, \tau^R_{t_{i+1}} \le T|\mathcal{F}_{t_{i+1}}] + \mathbb{P}[\tau_{t_{i+1}} > T, \tau^R_{t_{i+1}} > T|\mathcal{F}_{t_{i+1}}] := C^1_{t_{i+1}} + C^2_{t_{i+1}}$. Since $C^1_{t_{i+1}} \le \mathbb{P}[\tau_{t_{i+1}} > \tau^R_{t_{i+1}}|\mathcal{F}_{t_{i+1}}]$, we can control this term in the same way we did for V^2 in the proof of Lemma 2.3.1. Namely, we get

$$\mathbb{E}[\mathbb{I}_{\tau_R>t_{i+1}}\mathbb{I}_{X_{t_{i+1}}\in D}C^1_{t_{i+1}}|\mathcal{F}_{\tau_{t_i}}] \le C\mathbb{E}[F(X_{t_{i+1}})^+|\mathcal{F}_{\tau_{t_i}}]. \tag{2.3.6}$$

In the following we use the notation introduced in the proof of Lemma 2.3.1 for time-changed martingales with $t = t_{i+1}$. For all $i\in[\![0,N-2]\!]$, on the set $\{X_{t_{i+1}}\in D\}$ we write

$$\begin{aligned}
C^2_{t_{i+1}} &= \mathbb{P}[\inf_{s\in[t_{i+1},T]} F(X_{t_{i+1}}) + \beta_{\mathcal{A}_s} + \lambda_{\mathcal{A}_s} > 0, \tau^R_{t_{i+1}} > T|\mathcal{F}_{t_{i+1}}]\\
&\le \mathbb{P}[\inf_{s\in[0,\mathcal{A}_T]} F(X_{t_{i+1}}) + \beta_s + C_\lambda s > 0, \tau^R_{t_{i+1}} > T|\mathcal{F}_{t_{i+1}}]\\
&\le \mathbb{P}[\inf_{s\in[0,a_0(T-t_{i+1})]} F(X_{t_{i+1}}) + \beta_s + C_\lambda s > 0, \tau^R_{t_{i+1}} > T|\mathcal{F}_{t_{i+1}}]\\
&\le \int_{a_0(T-t_{i+1})}^{\infty} dt\frac{F(X_{t_{i+1}})}{(2\pi t^3)^{1/2}}\exp(-\frac{(F(X_{t_{i+1}})+C_\lambda t)^2}{2t}) \le \frac{CF(X_{t_{i+1}})}{(T-t_{i+1})^{1/2}}
\end{aligned} \tag{2.3.7}$$

exploiting the explicit density for the hitting times of the drifted BM, see e.g. [KS91] section 3.5.C, for the last but one inequality.

From (2.3.6) and (2.3.7) we derive that $\forall i\in[\![0,N-2]\!]$

$$A^{21}_{\tau_{t_i}} \le C(T-\tau_{t_i})\mathbb{E}[F(X_{t_{i+1}})^+(1+\frac{1}{(T-t_{i+1})^{1/2}})|\mathcal{F}_{\tau_{t_i}}].$$

Observing that $\forall i\in[\![0,N-2]\!], T-t_{i+1}\ge \dfrac{T-t_i}{2}$ we derive similarly to (2.3.5)

$$A^{21}_{\tau_{t_i}} \le C\mathbb{E}[F(X_{t_{i+1}})^+|\mathcal{F}_{\tau_{t_i}}] \le C\sqrt{h}. \tag{2.3.8}$$

Since for $i = N-1$ we also have $A^{21}_{\tau_{t_i}} \le (T-\tau_{t_i}) \le h$ we finally obtain that equation (2.3.8) is valid for all $i\in[\![0,N-1]\!]$.

We now turn to the control of $A^{22}_{\tau_{t_i}}$.

$$A^{22}_{\tau_{t_i}} = \mathbb{E}[(\tau_R-\tau_{t_i})\mathbb{I}_{\tau_R<t_{i+1}}|\mathcal{F}_{\tau_{t_i}}] + \mathbb{E}[(\tau_{t_{i+1}}-\tau_{t_i})\mathbb{I}_{\tau_R\ge t_{i+1},\tau_{t_{i+1}}\le T}|\mathcal{F}_{\tau_{t_i}}] := O(h) + A^{221}_{\tau_{t_i}}.$$

We reintroduce the events $\{\tau^R_{t_{i+1}} > \tau_{t_{i+1}}\}, \{\tau^R_{t_{i+1}} < \tau_{t_{i+1}}\}$ in $A^{221}_{\tau_{t_i}}$. It comes

$$\begin{aligned}
A^{221}_{\tau_{t_i}} &= \mathbb{E}[(\tau_{t_{i+1}}-\tau_{t_i})\mathbb{I}_{\tau_R\ge t_{i+1},\tau_{t_{i+1}}\le T}\mathbb{I}_{\tau_{t_{i+1}}>t_{i+1}}(\mathbb{I}_{\tau^R_{t_{i+1}}>\tau_{t_{i+1}}} + \mathbb{I}_{\tau^R_{t_{i+1}}<\tau_{t_{i+1}}})|\mathcal{F}_{\tau_{t_i}}] + O(h)\\
&:= A^{2211}_{\tau_{t_i}} + A^{2212}_{\tau_{t_i}} + O(h).
\end{aligned}$$

Conditioning w.r.t. $\mathcal{F}_{t_{i+1}}$ and using the same arguments as for $C^1_{t_{i+1}}$ we readily get $A^{2212}_{\tau_{t_i}} \le C\mathbb{E}[F(X_{t_{i+1}})^+|\mathcal{F}_{\tau_{t_i}}] \le C\sqrt{h}$. For $A^{2211}_{\tau_{t_i}}$ write

$$\begin{aligned}
A^{2211}_{\tau_{t_i}} &\le h + \mathbb{E}[\mathbb{I}_{X_{t_{i+1}}\in D}\mathbb{E}[(\tau_{t_{i+1}} - t_{i+1})\mathbb{I}_{\tau_{t_{i+1}}\le T}\mathbb{I}_{\tau^R_{t_{i+1}}>\tau_{t_{i+1}}}|\mathcal{F}_{t_{i+1}}]|\mathcal{F}_{\tau_{t_i}}] \\
&:= h + \mathbb{E}[\mathbb{I}_{X_{t_{i+1}}\in D}Q_{t_{i+1}}|\mathcal{F}_{\tau_{t_i}}].
\end{aligned}$$

Regarding $Q_{t_{i+1}}$, one has

$$\begin{aligned}
Q_{t_{i+1}} &\le \int_0^{T-t_{i+1}} ds\mathbb{P}[\tau_{t_{i+1}} - t_{i+1} \ge s, \tau^R_{t_{i+1}} > \tau_{t_{i+1}}|\mathcal{F}_{t_{i+1}}] \\
&\le \int_0^{T-t_{i+1}} ds\mathbb{P}[\inf_{u\in[0,\mathcal{A}_{s+t_{i+1}}]} F(X_{t_{i+1}}) + \beta_u + C_\lambda u > 0, \tau^R_{t_{i+1}} > \tau_{t_{i+1}}|\mathcal{F}_{t_{i+1}}] \\
&\le \int_0^{T-t_{i+1}} ds\mathbb{P}_y[\tau_0^{\tilde\beta} \ge a_0 s]
\end{aligned}$$

where we denote $y = F(X_{t_{i+1}})$, $\tilde\beta_u = y + \beta_u + C_\lambda u$, $\tau_0^{\tilde\beta} := \inf\{s \ge 0 : \tilde\beta_s = 0\}$. Thus, recalling that $y > 0$ on the set $\{X_{t_{i+1}} \in D\}$, it comes

$$\begin{aligned}
Q_{t_{i+1}} &\le a_0^{-1}\int_0^{(T-t_{i+1})a_0} ds\mathbb{P}_y[\tau_0^{\tilde\beta} \ge s] = a_0^{-1}\mathbb{E}_y[\tau_0^{\tilde\beta}\mathbb{I}_{\tau_0^{\tilde\beta}\le a_0(T-t_{i+1})}] \\
&\le a_0^{-1}\int_0^{(T-t_{i+1})a_0} dt\frac{ty}{(2\pi t^3)^{1/2}}\exp(-\frac{(y+C_\lambda t)^2}{2t}) \le Cy(T-t_{i+1})^{1/2}.
\end{aligned}$$

From this last estimate and the previous controls we derive

$$A^{2211}_{\tau_{t_i}} \le h + C\mathbb{E}[\mathbb{I}_{X_{t_{i+1}}\in D}F(X_{t_{i+1}})|\mathcal{F}_{\tau_{t_i}}] \le C\sqrt{h},\ A^{221}_{\tau_{t_i}} \le C\sqrt{h}.$$

Hence, for all $i \in [\![0, N-1]\!]$,

$$A^{22}_{\tau_{t_i}} \le C\sqrt{h}. \tag{2.3.9}$$

We conclude the proof of Lemma 2.2.2 under **(G2)** putting together the controls (2.3.5), (2.3.8), (2.3.9) .

□

3.2 Proof of Lemma 2.2.3

We adapt ideas from [Gob00]: in the quoted paper, a uniform ellipticity condition was assumed, and this enabled to use a Gaussian type lower bound for the conditional density of $X_{t_{i+1}}$ w.r.t. the Lebesgue measure, together with some computations related to a cone exterior to D. Here, under **(C)**, the conditional law of $X_{t_{i+1}}$ may degenerate and our proof relies on the scaling invariance of the cone and of the Brownian increments.
It is enough to prove that $a.s$ on $\{t_i < \tau_{t_i} < t_{i+1}\}$, one has

$$\mathbb{P}[X_{t_{i+1}} \notin D \mid \mathcal{F}_{\tau_{t_i}}] \ge \frac{1}{C}. \tag{2.3.10}$$

Indeed, it follows that $\mathbb{P}[X_{t_{i+1}} \notin D \mid \mathcal{F}_{t_i}] = \mathbb{E}[\mathbb{I}_{\tau_{t_i}\le t_{i+1}}\mathbb{P}[X_{t_{i+1}} \notin D \mid \mathcal{F}_{\tau_{t_i}}] \mid \mathcal{F}_{t_i}] \ge \frac{\mathbb{P}[\tau_{t_i}\le t_{i+1} \mid \mathcal{F}_{t_i}]}{C}$ and Lemma 2.2.3 is proved.
To get (2.3.10), write $X_{t_{i+1}} = X_{\tau_{t_i}} + \sigma_{\tau_{t_i}}(W_{t_{i+1}} - W_{\tau_{t_i}}) + R_i$ where $R_i = \int_{\tau_{t_i}}^{t_{i+1}} b_u du + \int_{\tau_{t_i}}^{t_{i+1}}(\sigma_u - \sigma_{\tau_{t_i}})dW_u$. The domain D is of class C^2, and thus satisfies a uniform exterior sphere condition with radius $R/2$ (R defined in Proposition 1.1.1): for any $z \in \partial D$, $B(z - \frac{R}{2}n(z), \frac{R}{2}) \subset D^c$. In particular, if we define for $\theta \in]0, \pi/2[$ the cone $\mathcal{K}(\theta, z) = \{y \in \mathbb{R}^d : (y-z).[-n(z)] \ge \|y-z\| \cos(\theta)\}$, then one has $\mathcal{K}(\theta, z) \cap B(z, R(\theta)) \subset B(z - \frac{R}{2}n(z), \frac{R}{2}) \subset D^c$

for some appropriate choice of the *positive* function $R(.)$. Then, it follows that

$$\begin{aligned}
&\mathbb{P}[X_{t_{i+1}} \notin D \mid \mathcal{F}_{\tau_{t_i}}] \geq \mathbb{P}[X_{t_{i+1}} \in \mathcal{K}(\theta, X_{\tau_{t_i}}) \cap B(X_{\tau_{t_i}}, R(\theta)) \mid \mathcal{F}_{\tau_{t_i}}] \\
&\geq \mathbb{P}[X_{t_{i+1}} \in \mathcal{K}(\theta, X_{\tau_{t_i}}) \mid \mathcal{F}_{\tau_{t_i}}] - \mathbb{P}[X_{t_{i+1}} \notin B(X_{\tau_{t_i}}, R(\theta)) \mid \mathcal{F}_{\tau_{t_i}}] \\
&\geq \mathbb{P}[(X_{t_{i+1}} - X_{\tau_{t_i}}).(-n(X_{\tau_{t_i}})) \geq \sqrt{\alpha_{\tau_{t_i}}(t_{i+1} - \tau_{t_i})} \geq \|X_{t_{i+1}} - X_{\tau_{t_i}}\| \cos(\theta) \mid \mathcal{F}_{\tau_{t_i}}] \\
&\quad - \mathbb{P}[X_{t_{i+1}} \notin B(X_{\tau_{t_i}}, R(\theta)) \mid \mathcal{F}_{\tau_{t_i}}] \geq A_1 - A_2(\theta) - A_3(\theta),
\end{aligned} \tag{2.3.11}$$

$$\begin{aligned}
\text{where } \; A_1 &= \mathbb{P}[(X_{t_{i+1}} - X_{\tau_{t_i}}).(-n(X_{\tau_{t_i}})) \geq \sqrt{\alpha_{\tau_{t_i}}(t_{i+1} - \tau_{t_i})} \mid \mathcal{F}_{\tau_{t_i}}], \\
A_2(\theta) &= \mathbb{P}[\sqrt{\alpha_{\tau_{t_i}}(t_{i+1} - \tau_{t_i})} < \|X_{t_{i+1}} - X_{\tau_{t_i}}\| \cos(\theta) \mid \mathcal{F}_{\tau_{t_i}}], \\
A_3(\theta) &= \mathbb{P}[X_{t_{i+1}} \notin B(X_{\tau_{t_i}}, R(\theta)) \mid \mathcal{F}_{\tau_{t_i}}].
\end{aligned}$$

Term A_1. Clearly, one has $A_1 \geq \mathbb{P}[(-n(X_{\tau_{t_i}})).\sigma_{\tau_{t_i}}(W_{t_{i+1}} - W_{\tau_{t_i}}) \geq 2\sqrt{\alpha_{\tau_{t_i}}(t_{i+1} - \tau_{t_i})} \mid \mathcal{F}_{\tau_{t_i}}] - \mathbb{P}[|n(X_{\tau_{t_i}}).R_i| \geq \sqrt{\alpha_{\tau_{t_i}}(t_{i+1} - \tau_{t_i})} \mid \mathcal{F}_{\tau_{t_i}}] := A_{11} - A_{12}$. The random variable $(-n(X_{\tau_{t_i}})).\sigma_{\tau_{t_i}}(W_{t_{i+1}} - W_{\tau_{t_i}})$ is conditionally to $\mathcal{F}_{\tau_{t_i}}$ a centered Gaussian variable with variance $\alpha_{\tau_{t_i}}(t_{i+1} - \tau_{t_i})$, and thus $A_{11} = \Phi(-2) > 0$. Owing to the condition **(S)** and since $\alpha_{\tau_{t_i}} \geq a_0$ *a.s*, it is easy to see that the contribution A_{12} converges uniformly to 0 when h goes to 0, and thus for $h = T/N$ enough small, one has $A_1 \geq \frac{A_{11}}{2} > 0$.

Term $A_2(\theta)$. From Markov's inequality, $A_2(\theta) \leq \frac{\mathbb{E}[\|X_{t_{i+1}} - X_{\tau_{t_i}}\|^2 \cos^2(\theta) \mid \mathcal{F}_{\tau_{t_i}}]}{\alpha_{\tau_{t_i}}(t_{i+1} - \tau_{t_i})} \leq C\cos^2(\theta)$ using **(C)** and estimates of Lemma 1.1.2. In particular, taking θ close to $\pi/2$ ensures that $A_2(\theta) \leq \frac{A_{11}}{6}$.

Term $A_3(\theta)$. Using Lemma 1.1.2, one readily gets $A_3(\theta) \leq C\exp\big(-c\frac{R^2(\theta)}{h}\big) \leq \frac{A_{11}}{6}$ for h small enough $(R(\theta) > 0)$.

Putting together estimates for $A_1, A_2(\theta)$ and $A_3(\theta)$ into (2.3.11) give $\mathbb{P}[X_{t_{i+1}} \notin D \mid \mathcal{F}_{\tau_{t_i}}] \geq \frac{A_{11}}{6}$. This proves (2.3.10).

3.3 A simple extension

From the previous controls we easily derive the following

Theorem 2.3.2 *Assume* **(C)**, **(D-**2**)**, **(S)** *and that g is bounded, uniformly Hölder continuous with index $\alpha \in (0, 1/2]$ in time and Hölder continuous with index 2α in space. For some constant C, one has*

$$|\mathrm{Err}(T, h, g, x)| \leq Ch^{\alpha/2}.$$

Proof. Starting from (2.2.2) we write

$$\begin{aligned}
|\mathcal{E}| &\leq C\sum_{i=0}^{N-1} \mathbb{E}[\mathbb{I}_{\tau^{N,R} > t_i} \mathbb{I}_{\tau_{t_i} \leq t_{i+1}} \mathbb{E}[(T \wedge \tau_{t_{i+1} \wedge \tau_R} - \tau_{t_i})^\alpha + \|X_{T \wedge \tau_{t_{i+1} \wedge \tau_R}} - X_{\tau_{t_i}}\|^{2\alpha} | \mathcal{F}_{\tau_{t_i}}]] \\
&\leq C\sum_{i=0}^{N-1} \mathbb{E}[\mathbb{I}_{\tau^{N,R} > t_i} \mathbb{I}_{\tau_{t_i} \leq t_{i+1}} \mathbb{E}[(T \wedge \tau_{t_{i+1} \wedge \tau_R} - \tau_{t_i})^\alpha | \mathcal{F}_{\tau_{t_i}}]]
\end{aligned}$$

using Lemma 2.1.1 for the last inequality. We controlled the term $\mathbb{E}[(T \wedge \tau_{t_{i+1} \wedge \tau_R} - \tau_{t_i}) | \mathcal{F}_{\tau_{t_i}}] \leq C\sqrt{h}$ in the proof of Lemma 2.2.2 under **(G2)**. Hence, the result is then a consequence of Hölder's inequality and Lemma 2.2.3.

□

4 Extension to an intersection of smooth domains

4.1 Additional notations and assumptions

In this section we allow the domain to be singular in the sense of the following Assumption

(D') The domain $D = \bigcap_{j=1}^m D_j,\ m \geq 2$. For all $j \in [\![1,m]\!]$, D_j satisfies **(D-2)**. We denote its boundary by $\Gamma_j := \partial D_j$.

For $r \geq 0$, we set $\forall j \in [\![1,m]\!],\ V_{\Gamma_j}(r) := \{z \in \mathbb{R}^d : d(z,\Gamma_j) \leq r\},\ V_{\partial D}(r) := \{z \in \mathbb{R}^d : d(z,\partial D) \leq r\},\ D(r) := D \cup V_{\partial D}(r)$. Since the Γ_j are C^2, we recall from Proposition 1.1.1 that $\exists R_j > 0$ s.t. on $V_{\Gamma_j}(R_j)$ the projection on Γ_j is uniquely defined. For all $x \in \Gamma_j$, the notation $n_j(x)$ stands for the inner normal unit of D_j. In the following, F_j denotes the signed distance to Γ_j which is C^2 on $V_{\Gamma_j}(R_j)$ and can be extended into a C^2 function on $\mathbb{R}^d$ with bounded derivatives (see once again Proposition 1.1.1 for details). Set $R := \wedge_{j=1}^m R_j$. Our non degeneracy assumption on the domain D is stated as follows:

(C') $\exists a_0 > 0$ such that $a.s.$ $\big(X_s \in V_{\Gamma_j}(R) \cap V_{\partial D}(R), s \in [0,T], j \in [\![1,m]\!] \Longrightarrow \nabla F_j(X_s).\sigma_s\sigma_s^*\nabla F_j(X_s) \geq a_0\big)$.

This corresponds to a non characteristic boundary condition w.r.t. every hypersurface in a neighbourhood of the domain D.

4.2 Main result

We are now in a position to state the main result of the section.

Theorem 2.4.1 (Upper Bound for an intersection of smooth domains in the killed case) *Assume* **(C')**, **(D')**, **(S)** *and let f be as in Theorem 2.2.1. For some constant $C := C(m)$, one has*

$$|\mathrm{Err}(T,h,f,x)| = |\mathbb{E}[f(X_T)\mathbb{I}_{\tau^N>T}] - \mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}]| \leq C\frac{\|f\|_\infty}{1\wedge\varepsilon}\sqrt{h}.$$

We restrict ourselves to the killed case for simplicity because we do not need to project X_{τ^N} on the boundary to define our approximation.

Remark 2.4.1 *The result of Theorem 2.4.1 is very interesting even in the Markovian setting of Brownian Motion. Indeed, for non smooth domains it is a hard task to use the traditional error analysis techniques that require the smoothness of the derivatives of the solution of the underlying PDE (1.2.1) up to the boundary, see also Chapter 3. We thus provide an alternative technique that points out that the main difficulty to upper-bound the weak error in the Brownian context does not lie in the lack of regularity of the domain.*

4.3 Proof of Theorem 2.4.1

Without modifying the rate of convergence, we can assume $X_t \in D(R)$ $a.s.$ Indeed, from Lemma 2.1.1, $\mathbb{P}[\tau_R \leq \tau^N] = O_{pol}(h)$.

Using the above definition of $(\tilde{V}_t)_{t\in[0,T]}$, i.e. $\forall t \in [0,T],\ \tilde{V}_t = \mathbb{E}[f(X_T)\mathbb{I}_{\tau_t>T}|\mathcal{F}_t]$, and for an initial point $x \in \partial D$, we derive in a similar way than for the proof of Theorem 2.2.1

$$\mathcal{E} := \mathbb{E}[f(X_T)\mathbb{I}_{\tau^N>T}] = \sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^N>t_i}\mathbb{I}_{\tau_{t_i}<t_{i+1}}\tilde{V}_{t_{i+1}}].$$

Recall that, to prove Theorem 2.4.1, it is enough to show $|\mathcal{E}| \leq C\sqrt{h}$ controlling that C is uniform w.r.t. $x \in \partial D$.

Put $\tau_t^j := \inf\{s > t : X_s \notin D_j\}$ and note that $\tau_t = \wedge_{j=1}^m \tau_t^j$. From **(C')**, we then derive that X satisfies our previous assumption **(C)** w.r.t. $D_j, \forall j \in [\![1,m]\!]$. Hence, as a consequence of Lemma 2.3.1 it comes

$$\begin{aligned} |\tilde{V}_{t_{i+1}}| &= |\mathbb{E}[f(X_T)\mathbb{I}_{\tau_{t_{i+1}}>T}|\mathcal{F}_{t_{i+1}}]| \leq \mathbb{E}[|f(X_T)|\mathbb{I}_{\tau^j_{t_{i+1}}>T}|\mathcal{F}_{t_{i+1}}] \\ &\leq \frac{C\|f\|_\infty}{1\wedge\varepsilon}[F_j(X_{t_{i+1}})]_+,\ \forall j \in [\![1,m]\!]. \end{aligned}$$

Thus,

$$\begin{aligned}|\mathcal{E}| &\le \sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^N>t_i,\tau_{t_i}<t_{i+1}}|\tilde{V}_{t_{i+1}}|] = \sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^N>t_i,\cup_{j=1}^m\{\tau_{t_i}^j<t_{i+1}\}}|\tilde{V}_{t_{i+1}}|]\\ &\le \frac{C\|f\|_\infty}{1\wedge\varepsilon}\sum_{j=1}^m\sum_{i=0}^{N-1}\mathbb{E}[\mathbb{I}_{\tau^{N,j}>t_i,\tau_{t_i}^j<t_{i+1}}[F_j(X_{t_{i+1}})]_+]\end{aligned}$$

where $\tau^{N,j} := \inf\{s_i \ge 0 : X_{s_i} \notin D_j\}$. Conditioning w.r.t. $\tau_{t_i}^j$ and applying Lemma 2.1.1 we derive that

$$|\mathcal{E}| \le \sqrt{h}\frac{C\|f\|_\infty}{1\wedge\varepsilon}\sum_{j=1}^m\sum_{i=0}^{N-1}\mathbb{P}[\tau^{N,j} > t_i, \tau_{t_i}^j < t_{i+1}].$$

We conclude the proof using Lemma 2.2.3 for all $j \in [\![1, m]\!]$.

□

5 Discrete sampling of some other path dependent functionals

For a regular mesh of $[0,T]$ with N time steps $(t_i = ih)_{i\in[\![0,N]\!]}$, h being the step size, we define $\forall s \in [0,T]$, $\phi(s) := \sup\{(t_i)_{i\in[\![0,N]\!]} : t_i \le s < t_{i+1}\}$.

5.1 Discretely sampled integral

We state the following

Theorem 2.5.1 *Let X be an Itô process following the dynamics of equation (2.0.1). Assume the coefficients b and σ are bounded and that ψ is a Lipschitz continuous function from $\mathbb{R}^d$ into $\mathbb{R}$. For $p \ge 1$ one has*

$$\psi(\int_0^T X_s ds) - \psi(\int_0^T X_{\phi(s)}ds) \underset{\mathrm{L}_p(\mathbb{P})}{=} O(h).$$

Note that in the above theorem we do not impose any continuity in probability or non-degeneracy conditions on σ.

Proof of Theorem 2.5.1. The idea of the proof is quite simple and relies on Fubini like arguments. For all $p \ge 1$ write

$$Q^p := \mathbb{E}[|\psi(\int_0^T X_s ds) - \psi(\int_0^T X_{\phi(s)}ds)|^p] \le C\mathbb{E}[|\int_0^T\left(\int_0^T \mathbb{I}_{t\in[\phi(s),s]}dX_t\right)ds|^p].$$

The assumptions of the theorem allow to apply Fubini's theorem for stochastic integrals, see [RY99] Chapter IV.5. We get

$$\begin{aligned}Q^p &\le C\mathbb{E}[|\int_0^T\left(\int_0^T\mathbb{I}_{t\in[\phi(s),s]}ds\right)dX_t|^p] = C\mathbb{E}[|\int_0^T(\phi(t)+h-t)dX_t|^p]\\ &\le C\left(\mathbb{E}[|\int_0^T b_t(\phi(t)+h-t)dt|^p] + \mathbb{E}[|\int_0^T(\phi(t)+h-t)\sigma_t dW_t|^p]\right).\end{aligned}$$

Owing to the BDG inequality it comes

$$Q^p \le C(h^p + \mathbb{E}[(\int_0^T\|\sigma_t\sigma_t^*\|^2(\phi(t)+h-t)^2dt)^{p/2}] \le Ch^p.$$

The proof is complete.

□

Remark 2.5.1 *An application of this result in a Mathematical finance framework concerns Asian options. Consider the case of a Black-Scholes model for $d = d' = 1$. The price dynamics of the underlying asset, $(S_t)_{t\in[0,T]}$, follows a Geometric Brownian Motion: $S_t = S_0 \exp(\sigma W_t + (\mu - \sigma^2/2)t)$, $\sigma > 0$.*

For a given strike level $K > 0$, the price of the Asian put option is then essentially given by $\mathbb{E}[\psi(\int_0^T S_s ds)]$ where $\psi(x) = (K - x)^+$, we do not detail the change of probability measure and the discounting factor.

Even though $(S_t)_{t\in[0,T]}$ does not satisfy the assumptions of Theorem 2.5.1, since $\forall p < \infty$, $\mathbb{E}[S_t^p] < \infty$ the result of that theorem holds true.

Hence, we obtain that a time discretization of step h in the above integral yields an error of order 1 w.r.t. h.

5.2 Discrete sampling for the maximum of some special processes

In this section we assume the dimension $d = d' = 1$. For a process X following the dynamics of equation (2.0.1), we define $M_T := \max_{s\in[0,T]} X_s$, $M_T^N := \max_{s\in[0,T]} X_{\phi(s)}$.

In the following we say we are in a Brownian setting if

$$X_s = x + \mu s + \sigma W_s,\ (\mu,\sigma) \in \mathbb{R} \times \mathbb{R}^{+*} \tag{2.5.1}$$

where W is a standard one-dimensional BM.

In that context, Asmussen, Glynn and Pitman, cf. [AGP95], used some path-decomposition techniques to obtain the limit distribution as well as the uniform integrability of the non-negative random variable $\sqrt{N}\varepsilon_N(T) := \sqrt{N}(M_T - M_T^N)$. Precisely, they showed the following

Theorem 2.5.2 *Assume X has the dynamics of equation (2.5.1). Then,*

$$(\sqrt{h})^{-1}\varepsilon_N(T) \underset{\mathcal{L},N}{\longrightarrow} \sigma\sqrt{T}Z,$$

where $Z = \min_{n\in\mathbb{Z}} \tilde{R}(U + n)$, U is uniform on $(0,1)$ and $\tilde{R}(t) = R_1(t)\mathbb{I}_{t\geq 0} + R_2(-t)\mathbb{I}_{t\leq 0}$, R_1, R_2 being two independent copies of a three dimensional Bessel process $(R_t)_{t\geq 0} = (\|\bar{W}_t\|)_{t\geq 0}$. The process $\bar{W}$ is the standard BM of $\mathbb{R}^3$.

They also proved

Proposition 2.5.3 *Assume X has the dynamics of equation (2.5.1). For any $\beta < \infty$, the sequence of random variables $(\exp(\beta\sqrt{N}\varepsilon_N))_{N\in\mathbb{N}^*}$ is uniformly integrable. In particular, for $p < \infty$, the sequence $((\sqrt{N}\varepsilon_N)^p)_{N\in\mathbb{N}^*}$ is uniformly integrable.*

The proof of the above results strongly relies on the decomposition in terms of 3 dimensional Bessel Bridges of the Brownian path around its maximum over a finite time interval.

Namely, in a Brownian setting for X with $x = 0, \sigma = 1$, and a given μ, let τ be the a.s. unique time in $[0,T]$ at which X attains its maximum M_T. Conditionally on $\tau = s, M_T = m, X_T = m - y$, the process $(m - X_{s-u})_{u\in[0,s]}$ is a three dimensional Bessel Bridge $\mathbf{BB}(3, s, m)$ independent of $(m - X_{s+u})_{u\in[0,T-s]}$ which is a $\mathbf{BB}(3, T - s, y)$.

The above identities are consequences of William's path decomposition, cf. Theorem VII.4.9 in Revuz and Yor [RY99]. For an intuitive description in terms of transition density see also Chapter XI.3 of the same reference.

Proposition 2.5.3 readily gives that there exists a constant $C > 0$ s.t.

$$\mathrm{Err}_{\max}(h) := \mathbb{E}[M_T - M_T^N] \leq C\sqrt{h}. \tag{2.5.2}$$

We provide another proof of this result using the techniques of Section 2 in Appendix 1.

We now show how the results of Asmussen and *al.* can be exploited to show (2.5.2) holds true under the assumption

(MID)

$$X_s = x + \int_0^s b_u du + \int_0^s \sigma(X_s) dW_s,$$

where $(b_u)_{u\geq 0}$ is a bounded progressively measurable coefficient and $\sigma \in C_b^1(\mathbb{R})$, i.e. σ is once continuously differentiable, bounded with bounded first derivative, and s.t. $\exists \sigma_0 > 0,\ \forall x \in \mathbb{R},\ \sigma(x) \geq \sigma_0$.

The key idea is to use a Lamperti transform to go from the dynamics of Assumption **(MID)** back to the case of a drifted Brownian motion. Since the drift we obtain is not constant we then use a Girsanov transformation to annihilate it and control the expectation of $\sqrt{N}\varepsilon_N$ relying on the previous uniform integrability results for the driftless BM.

The limiting factor in our approach is the use of Lamperti's transformation that imposes to have a Markovian diffusion term. Unfortunately, our main attempts to derive (2.5.2) for general non-degenerated Itô processes failed. We mainly tried to use Dambis-Dubins-Schwarz time changes techniques. The main difficulty is that little is known on the dependence between the time-changed Brownian Motion and the time-change itself.

Let us state the following

Theorem 2.5.4 *Assume* **(MID)**. *There exists a constant $C > 0$ s.t.*

$$\mathrm{Err}_{\max}(h) = \mathbb{E}[M_T - M_T^N] \leq C\sqrt{h}.$$

Proof. Define $\forall x \in \mathbb{R},\ \varphi(x) = \displaystyle\int_0^x \frac{dy}{\sigma(y)},\ \forall s \geq 0,\ Y_s := \varphi(X_s)$. Under **(MID)**, the function φ is $C^2(\mathbb{R})$. Hence, for all $t \geq 0$ Itô's formula yields

$$\begin{aligned} Y_t &= \varphi(x) + \int_0^t \varphi'(X_s)dX_s + \frac{1}{2}\int_0^t \varphi''(X_s)d\langle X\rangle_s \\ &= \varphi(x) + W_t + \int_0^t \left(\frac{b_s}{\sigma(X_s)} - \frac{1}{2}\sigma'(X_s)\right) ds := \varphi(x) + W_t + \int_0^t \tilde{b}_s ds. \end{aligned}$$

This can be found in Chapter VIII Section 6 of Gikhman and Skorokhod [GS69] or in Rogers [Rog85]. Note that under **(MID)**, the function φ is strictly increasing on $\mathbb{R}^+$, let us denote φ^{-1} its inverse on this set. On $\mathbb{R}^+$ we also have $\frac{x}{|\sigma|_\infty} \leq \varphi(x) \leq \frac{x}{\sigma_0}$. From this last property we derive that $\forall x > y > 0,\ \varphi^{-1}(x) - \varphi^{-1}(y) \leq |\sigma|_\infty (x-y)$.

We can suppose w.l.o.g. that $x = 0$ in **(MID)**. Since $M_T \geq M_T^N \geq 0$, it comes

$$\mathrm{Err}_{\max}(h) = \mathbb{E}[\varphi^{-1}(\varphi(M_T)) - \varphi^{-1}(\varphi(M_T^N))] \leq |\sigma|_\infty \mathbb{E}[\varphi(M_T) - \varphi(M_T^N)] = |\sigma|_\infty \mathbb{E}[M_T^Y - M_T^{Y,N}]$$

with $M_T^Y := \max_{s\in[0,T]} Y_s,\ M_T^{Y,N} := \max_{s\in[0,T]} Y_{\phi(s)}$.

Put $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = \exp(-\int_0^t \tilde{b}_s dW_s - \frac{1}{2}\int_0^t \tilde{b}_s^2 ds) := \mathcal{E}(-\int_0^t \tilde{b}_s dW_s)$. Under $\mathbb{Q}$, we derive from the Girsanov Theorem that Y is a standard BM. We obtain

$$\mathrm{Err}_{\max}(h) = \sqrt{h}\mathbb{E}^{\mathbb{Q}}\left[\mathcal{E}(\int_0^T \tilde{b}_s dY_s) N^{1/2}\varepsilon_N\right] \leq \sqrt{h}\mathbb{E}^{\mathbb{Q}}[N\varepsilon_N^2]^{1/2}\mathbb{E}^{\mathbb{Q}}[\exp(2\int_0^T \tilde{b}_s dY_s - \int_0^T \tilde{b}_s^2 ds)]^{1/2}.$$

Proposition 2.5.3 readily gives $\mathbb{E}^{\mathbb{Q}}[N\varepsilon_N^2] \leq C$. On the other hand, recall that under **(MID)**, $\tilde{b}$ is bounded. Thus,

$$\mathbb{E}^{\mathbb{Q}}[\exp(2\int_0^T \tilde{b}_s dY_s - \int_0^T \tilde{b}_s^2 ds)] \leq \exp(T|\tilde{b}|_\infty^2)\underbrace{\mathbb{E}^{\mathbb{Q}}[\mathcal{E}(2\int_0^T \tilde{b}_s dY_s)]}_{=1,\ \text{Novikov}}.$$

This gives the result.

$\square$

Remark 2.5.2 *In Mathematical Finance, the result of Theorem 2.5.4 has applications for the evaluation of look-back options. If the price dynamics of the underlying asset, $(S_t)_{t\in[0,T]}$, is given by $S_t = S_0\exp(X_t)$, where X satisfies* **(MID)**, *omitting the discounting factor the price of such options writes $\mathbb{E}[\psi(\exp(X_T),\exp(M_T))]$ for a given pay-off function ψ.*

For Lipschitz continuous functions ψ we have

$$\begin{aligned}\mathrm{Err}(T,h,\psi,x) &:= \mathbb{E}[\psi(\exp(X_T),\exp(M_T))] - \mathbb{E}[\psi(\exp(X_T),\exp(M_T^N))],\\ |\mathrm{Err}(T,h,\psi,x)| &\le C\mathbb{E}[\exp(M_T)(M_T - M_T^N)].\end{aligned}$$

Thus, following the proof of Theorem 2.5.4 one obtains $\mathrm{Err}(T,h,\psi,x) = O(\sqrt{h})$. For instance, the call options on the maximum with strike level $K > 0$ or w.r.t. to the terminal value S_T enter in that framework with respectively $\psi(x,y) = (y-k)^+$ and $\psi(x,y) = y - x$.

6 Conclusion

In this chapter we first emphasized that for killed/stopped Itô processes, under suitable assumptions, the order $1/2$ for $\mathrm{Err}(T,h,f,x)$ is really intrinsic to the discrete time killing.

To obtain our main result, i.e. Theorem 2.2.1, we introduced some martingale techniques that allow to go beyond the Markovian framework and also to control $\mathrm{Err}(T,h,f,x)$ at the expected rate for a certain class of non-smooth domains.

As a matter of fact, few technical tools are needed for the error analysis we present. This is promising since even in a Brownian setting, for non-smooth domains the PDE approach for the error analysis is rather technical or fails.

In terms of financial applications, the results of Sections 2 and 4 provide an upper bound for the error associated to a discrete time observation for barrier options. The results of Section 5.2 can be seen as preliminary controls to deal with the impact of a time discretization for Asian or look-back options.

Chapter 3

Error expansion and correction: The Brownian case

Introduction

Let $(X_t)_{t\in[0,T]}$ be a d-dimensional Brownian Motion (BM in short) with dynamics

$$X_t = x + \mu t + \sigma W_t \tag{3.0.1}$$

where W is a standard d-dimensional BM defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ with the usual assumptions on $(\mathcal{F}_t)_{t\in[0,T]}$. We assume $\sigma\sigma^*$ to be positive definite.

Let D be a domain of $\mathbb{R}^d$. Define $\tau := \inf\{t \geq 0 : X_t \notin D\}$. For a regular time mesh of the interval $[0,T]$ with N time steps, $(t_i = ih)_{i\in[\![0,N]\!]}$, $h = T/N$ being the step size, introduce now $\tau^N := \inf\{t_i \geq 0 : X_{t_i} \notin D\}$. For a measurable function f, and an initial point $x \in D$ we refer to the quantity

$$\mathrm{Err}(T,h,f,x) = \mathbb{E}_x\left[f(X_T)\mathbb{I}_{\tau^N>T}\right] - \mathbb{E}_x\left[f(X_T)\mathbb{I}_{\tau>T}\right]$$

as the weak error associated to the discrete time killing of X w.r.t. the domain D. This is the object studied in this chapter.

Note that from the two previous chapters we already have some controls on the above quantity under suitable assumptions on the functions and domains. We proved in Chapter 1 that for smooth domains and functions f satisfying either some support condition w.r.t. D or some smoothness properties and compatibility conditions one had that $\mathrm{Err}(T,h,f,x)$ was upper and lower bounded at order $1/2$ w.r.t. h. We also showed in Chapter 2 that the upper bound holds true for an intersection of smooth domains.

Main Results

Let $D \subset \mathbb{R}^d$ be a domain of the form $D = \cap_{i=1}^m D^i,\ m \in [\![1,d]\!]$ where the $(D^i)_{i\in[\![1,m]\!]}$ are d-dimensional half spaces with non-empty intersection. If we have "good" smoothness properties up to the boundary for the function $v(t,x) := \mathbb{E}_x\left[f(X_{T-t})\mathbb{I}_{\tau>T-t}\right]$, we obtain that for h small enough

$$\boxed{\mathrm{Err}(T,h,f,x) = C\sqrt{h} + o(\sqrt{h})}$$

for some constant C depending on D, f, x.

As emphasized in Theorem 1.5.2, the leading term in the error is still the one associated to the local time of the process on the boundary. According to Tanaka's formula and to Lemma 1.4.5 we also derived that the overshoot of the killed process above the boundary was the dominant part in the local time (the overshoot being defined as the distance to the boundary of the process when it exits the domain).

In the special case of Brownian Motion, for half spaces or intersections of half spaces forming a cone, we are able to obtain the asymptotic distribution of the overshoot, and therefore an equivalent for the local time, uniformly in time. To derive the error expansion we then use usual techniques based on Itô's formula. The smoothness of v is needed for this last step. From a theoretical point of view, the main difficulty is analytical and consists in having good smoothness properties for the Green kernel of the heat operator in non-smooth domains.

In practice we have been able to show the required smoothness conditions for v in two special cases.

- In the half space case, i.e. when $m = 1$, imposing either that f is smooth and satisfies some compatibility conditions, or that f is borelian and $d(\mathrm{supp}(f), \partial D) \geq 2\varepsilon > 0$.
- In the bidimensional cone, imposing f is smooth, satisfies some compatibility conditions and has a support at a strictly positive distance from the origin of the cone.

In all cases we assume that f is bounded and vanishes on the boundary.

A half space is a smooth domain for which we have well known results concerning the derivatives of v.

For an intersection of half spaces forming a cone, the control of those derivatives is not easy even for $d = 2$. For some special cases, i.e. when for $\sigma = \mathrm{I}_2$ and the angle of the cone writes $\alpha = \pi/m_0,\ m_0 \in \mathbb{N}^*$, the transition density of the killed process (Green kernel of the underlying PDE) has the same smoothness properties as in the case of smooth domains. For general angles, the transition density writes as an infinite sum of modified Bessel functions and it is a hard task to obtain tractable controls on its derivatives up to the boundary. We get rid of the problem imposing that the support of f is at a strictly positive distance from the origin of the cone.

From a numerical point of view, the error expansion is the preliminary step for a procedure that aims to improve the convergence rate. A standard one in this framework is the Romberg extrapolation, see Talay and Tubaro, [TT90], and Section 5 for details.

We propose an alternative correction method based on the recent work of Costantini, El Karoui and Gobet, see [CKG03], concerning the sensitivity of the Dirichlet problem w.r.t. the domain.

Unlike in the Romberg extrapolation we do not need to refine the time step and thus the procedure is computationally cheaper. We simply proceed to the simulation w.r.t. a more constrained domain. Namely, instead of killing the process when it exits from D at one of the discretization times, we kill it when it leaves $D_h := \cap_{i=1}^m D_h^i,\ D_h^i := \{x \in \mathbb{R}^d : x - C_i\sqrt{h}n_i \in D^i\}$ where n_i denotes the inner normal associated to the half-space D^i. The positive constant C_i depends on D_i, on the coefficient σ in (3.0.1) and on a universal constant that comes from the renewal techniques employed to obtain the equivalent of the local time. Define $\tau_{D^h}^N := \inf\{s_i \geq 0 : X_{s_i} \notin D_h\}$. We get

$$\boxed{\mathrm{Err}'(T,h,f,x) = \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D_h}^N > T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau > T}] = o(\sqrt{h}),}$$

removing the main error term.

We mention that in a one dimensional setting, both the expansion result and the correction procedure could be derived by direct computations from the work of Broadie, Glasserman and Kou, [BGK99], who established $\mathbb{E}[f(X_T)\mathbb{I}_{\tau^N > T}] = \mathbb{E}[f(X_T)\mathbb{I}_{\tau_{\tilde{D}_h} > T}] + o(\sqrt{h}),\ \tau_{\tilde{D}_h} := \inf\{s \geq 0 : X_s \notin \tilde{D}_h\}, D \subset \tilde{D}_h := \{y \in \mathbb{R} : y - \tilde{C}\sqrt{h} \in D\}$ for $f(x) = e^{-rT}(\exp(x) - K)^+,\ K \in \mathbb{R}^+$. The domain $\tilde{D}_h$ is slightly extended in order to compensate the overestimation due to the discrete exit time when $f \geq 0$.

Plan of the Chapter

We specify our current notations and assumptions in Section 1.

In Section 2 we detail the various steps that lead to the error expansion assuming that v is smooth enough.

Section 3 is dedicated to the correction procedure.

In Section 4 we introduce sufficient conditions that guarantee the required smoothness assumptions on v when D is a half space or a bidimensional cone.

Numerical results are presented in Section 5. We first deal with the half-space and the bidimensional cone for which the results confirm the correction procedure is rather accurate. Then, we describe how our shifting boundary condition can be adapted in a diffusion framework and give an example that illustrates this procedure.

We finally conclude in Section 6 evoking possible extensions and open problems.

1 Notations and assumptions

1.1 Reduction to a centered correlated Brownian motion in an orthant

Let $(X_s)_{s\geq 0}$ be a d-dimensional process with dynamics $X_s := x + \mu s + \sigma W_s$, where W is a standard d-dimensional BM and $\sigma\sigma^*$ is assumed to be positive definite. Let $D = \cap_{j=1}^m D^j$, with $m \in [\![1, d]\!]$, and for all $j \in [\![1, m]\!]$, $D^j := \{y \in \mathbb{R}^d : a_j \cdot y > b_j\}$, $a_j \in \mathbb{R}^d$, $a_j \neq 0$, $b_j \in \mathbb{R}$. The matrix $A := \begin{pmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_m^* \end{pmatrix} \in \mathbb{R}^{m\otimes d}$ is assumed to be of rank m. For notational convenience, we write $D := \{x \in \mathbb{R}^d : Ax > b\}$.

The next proposition illustrates how, for a given function f, one can rewrite $\mathrm{Err}(T, h, f, x)$ as the weak error associated to a centered, correlated BM killed when it leaves an orthant of the form $\{y \in \mathbb{R}^d : y_i > b_i,\ i \in [\![1, m]\!]\}$. This reformulation of the problem then turns out to be more tractable for the error analysis.

Namely, we have

Proposition 3.1.1 *Under the above assumptions on X, D and for a borelian bounded function f*

$$\mathrm{Err}(T, h, f, x) = \mathbb{E}_0[f_0(\check{W}_T)(\mathbb{I}_{\tau_{D_0^N} > T} - \mathbb{I}_{\tau_{D_0} > T})]$$

where $\check{W}$ is a centered d-dimensional Brownian motion with covariance matrix $\sigma_0\sigma_0^* = \begin{pmatrix} \Sigma & 0 \\ 0 & \mathbf{I}_{d-m} \end{pmatrix}$, $\Sigma \in \mathbb{R}^{m\otimes m}$ and $\forall (i,j) \in [\![1, m]\!]$, $\Sigma_{ij} = \langle \sigma^* a_i, \sigma^* a_j\rangle/(\|\sigma^* a_i\|\|\sigma^* a_j\|)$.

The domain $D_0 := \cap_{j=1}^m D_0^j$, where for all $j \in [\![1, m]\!]$, $D_0^j := \{y \in \mathbb{R}^d :\ y_j > b_0^j\}$, $b_0^j = \frac{b_j - a_j\cdot x}{\|\sigma^* a_j\|}$. Denoting $\tau_{D_0^j} := \inf\{s \geq 0 : \check{W}_s \notin D_0^j\}$, $\tau_{D_0^j}^N := \inf\{s_i \geq 0 : \check{W}_{s_i} \notin D_0^j\}$, we have $\tau_{D_0} := \wedge_{j=1}^m \tau_{D_0^j}, \tau_{D_0}^N := \wedge_{j=1}^m \tau_{D_0^j}^N$. The function f_0 writes $f_0(y) = \exp(\sigma^{-1}\mu\cdot\Lambda^{-1}y - \frac{\|\sigma^{-1}\mu\|^2}{2}T)f(x + \sigma\Lambda^{-1}y)$ with $\Lambda = \left(\frac{\sigma^* a_1}{\|\sigma^* a_1\|} \cdots \frac{\sigma^* a_m}{\|\sigma^* a_m\|}\ c_1 \ \ldots\ c_{d-m}\right)^*$. The $(c_i)_{i\in[\![1,d-m]\!]}$ form an orthonormal basis of $\{\mathrm{Span}((\sigma^* a_j)_{j\in[\![1,m]\!]})\}^\perp$.

Note that in the half space case, i.e. for $m = 1$, this transformation illustrates that the problem is essentially one-dimensional.

Proof. The change of probability measure $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} = \exp(-\sigma^{-1}\mu\cdot W_t - \frac{\|\sigma^{-1}\mu\|^2}{2}t) := \mathcal{E}(-\sigma^{-1}\mu\cdot W_t)$ yields

$$\mathrm{Err}(T, h, f, x) \quad = \quad \mathbb{E}^{\mathbb{Q}}[\mathcal{E}(\sigma^{-1}\mu\cdot\tilde{W}_T)f(x + \sigma\tilde{W}_T)(\mathbb{I}_{\forall s\in[0,T]:A(x+\sigma\tilde{W}_{\phi(s)})>b} - \mathbb{I}_{\forall s\in[0,T]:A(x+\sigma\tilde{W}_s)>b})]$$

where $\tilde{W}$ is a standard $\mathbb{Q}$-BM and $\phi(s) := \inf\{s_i \geq 0 : s_i \leq s < s_{i+1}\}$. Define now $\check{W}$ by setting

$$\check{W} = \left(\frac{\sigma^* a_1}{\|\sigma^* a_1\|} \cdots \frac{\sigma^* a_m}{\|\sigma^* a_m\|}\ c_1 \ \ldots\ c_{d-m}\right)^* \tilde{W} := \Lambda\tilde{W}$$

where the $(c_i)_{i\in[\![1,d-m]\!]}$ form an orthonormal basis of $\{\mathrm{Span}((\sigma^* a_j)_{j\in[\![1,m]\!]})\}^\perp$. By construction, $\check{W}$ is a correlated $\mathbb{Q}$-BM with covariance matrix

$$\sigma_0\sigma_0^* = \begin{pmatrix} \Sigma & 0 \\ 0 & \mathbf{I}_{d-m} \end{pmatrix},\ \forall (i,j) \in [\![1, m]\!]^2,\ \Sigma_{ij} = \frac{\langle \sigma^* a_i, \sigma^* a_j\rangle}{\|\sigma^* a_i\|\|\sigma^* a_j\|} := \rho_{ij}.$$

With the above definition it comes

$$\begin{aligned}\{\forall s \in [0, T] : A(x + \sigma\tilde{W}_s) > b\} \quad &= \quad \left\{\forall (s, j) \in [0, T] \times [\![1, m]\!] :\ \frac{(A\sigma\tilde{W}_s)_j}{\|\sigma^* a_j\|} > \frac{(b_j - a_j\cdot x)}{\|\sigma^* a_j\|}\right\} \\ &:= \quad \left\{\forall (s, j) \in [0, T] \times [\![1, m]\!] :\ \check{W}_s^j > \frac{(b_j - a_j\cdot x)}{\|\sigma^* a_j\|}\right\}.\end{aligned}$$

Since we assumed A is of rank m, Λ is invertible. Hence,

$$\mathrm{Err}(T,h,f,x) = \mathbb{E}\big[\exp(\sigma^{-1}\mu\cdot\Lambda^{-1}\check{W}_T - \frac{\|\sigma^{-1}\mu\|^2T}{2})f(x+\sigma\Lambda^{-1}\check{W}_T)\times$$
$$(\mathbb{I}_{\forall(s,j)\in[0,T]\times[\![1,m]\!],\ \check{W}^j_{\phi(s)}>\frac{(b_j-a_j\cdot x)}{\|\sigma^* a_j\|}} - \mathbb{I}_{\forall(s,j)\in[0,T]\times[\![1,m]\!],\ \check{W}^j_s>\frac{(b_j-a_j\cdot x)}{\|\sigma^* a_j\|}})\big].$$

With the notations of the proposition we finally write

$$\begin{aligned}\mathrm{Err}(T,h,f,x) &= \mathbb{E}_0[\exp(\sigma^{-1}\mu\cdot\Lambda^{-1}\check{W}_T - \frac{\|\sigma^{-1}\mu\|^2T}{2})f(x+\sigma\Lambda^{-1}\check{W}_T)(\mathbb{I}_{\tau_{D_0^N}>T} - \mathbb{I}_{\tau_{D_0}>T})]\\ &= \mathbb{E}_0[f_0(\check{W}_T)(\mathbb{I}_{\tau_{D_0^N}>T} - \mathbb{I}_{\tau_{D_0}>T})].\end{aligned}$$

The proof is complete.

□

According to the previous proposition, in the following of the chapter we restrict ourselves to the case of a correlated BM for the process X and to an orthant for the domain D. We introduce the following assumptions.

1.2 About the domain

We suppose our domain satisfies Assumption

(D) $D=\cap_{j=1}^m D^j,\ \forall j\in[\![1,m]\!],\ D^j:=\{y\in\mathbb{R}^d: y_j>b_0^j\}$, where $m\in[\![1,d]\!]$.

1.3 About the process

We introduce

(BM) The d-dimensional process $(X_s)_{s\geq 0}$ has the form $X_s:=x+\sigma_0 W_s$, where W is a standard d-dimensional BM and $\sigma_0\sigma_0^* = \begin{pmatrix}\Sigma & 0\\ 0 & \mathbf{I}_{d-m}\end{pmatrix}$ is assumed to be positive definite and Σ is a correlation matrix with coefficients $(\rho_{ij})_{(i,j)\in[\![1,m]\!]^2}$. The integer $m\in[\![1,d]\!]$ is the same as in assumption **(D)**.

1.4 About the function

Assume **(BM)**, **(D)**. For a function $f\in C_b^0(\bar{D})$ vanishing on the boundary, define now $\forall(t,y)\in[0,T]\times\mathbb{R}^d$, $v(t,y):=\mathbb{E}_y[f(X_{T-t})\mathbb{I}_{\tau>T-t}]$, $\tau:=\inf\{s\geq 0: X_s\notin D\}$. The error writes

$$\mathrm{Err}(T,h,f,x) = \mathbb{E}_x[v(T\wedge\tau^N, X_{T\wedge\tau^N})] - v(0,x) \tag{3.1.1}$$

where we recall $\tau^N:=\inf\{t_i\geq 0: X_{t_i}\notin D\}$.

From Theorem 16.1 Chapter 4 in [LSU68] we deduce that $v\in C^{1,2}([0,T)\times D)\cap C_b^0([0,T]\times\bar{D})$ and satisfies

$$\begin{cases}(\partial_t v+\frac{1}{2}\mathrm{tr}(H_v\sigma_0\sigma_0^*))(t,y)=0,\ (t,y)\in[0,T)\times D,\\ v(t,.)|_{\partial D}=0, t\in[0,T],\ v(T,y)=f(y),\ y\in\bar{D}\end{cases} \tag{3.1.2}$$

where $\partial_t v$ (resp. H_v) denotes the time derivative (resp. the Hessian matrix) of v.

In the following, under **(BM)**, **(D)**, we assume

(S) The function f vanishes on the boundary. The associated function v belongs to $C_b^{2,4}([0,T]\times\bar{D})$, i.e. there exists a constant $C>0$, s.t. for all multi-indices $\gamma,\ \zeta,\ |\gamma|\leq 2, |\zeta|\leq 4$

$$\forall(t,y)\in[0,T]\times\bar{D},\ |\partial_t^\gamma v(t,y)|+|\partial_y^\zeta v(t,y)|\leq C$$

and the partial derivatives of order 2 in time (resp. order 4 in space) of v are continuous on $[0,T]\times\bar{D}$.

In particular, this means that the function f is at least $C_b^4(\bar{D})$.

We specify in Section 4 sufficient conditions on f to obtain **(S)** in some special cases.

1.5 Miscellaneous

For smooth functions $g(t,x)$, the notation $\nabla g(t,x)$ stands for the gradient of g w.r.t. x.

The density of the standard normal law $\mathcal{N}(0,s\mathbf{I}_d)$ is denoted by $g_s(x) := \exp\left(-\frac{\|x\|^2}{2s}\right)(2\pi s)^{-d/2}$.

We will keep the same notation C (or C') for all finite, non-negative constants which will appear in our computations: they may depend on D, T, σ_0, or f, but they will not depend on the number of time steps N and the initial value x. We reserve the notation c and c' for constants also independent of x, T and f. Other possible dependences for the constants are explicitly indicated.

2 Error Expansion

The aim of this section is to present an expansion result of the error $\mathrm{Err}(T,h,f,x)$ under **(BM)**, **(D)** and **(S)**.

In Subsection 2.1, we decompose the error in terms of the expectation of an integral w.r.t. the increments of the local time on the boundary before the discrete exit time. Although the domain is non-smooth, the expression we obtain is very similar to the one stated in Theorem 1.5.2. We point out that in the Brownian case, since there is no discretization error on the process, we do not have any residual terms as in the cited theorem.

Then, in Subsection 2.2, we specify some asymptotic behaviours for the expectation of the local time of the discretely killed process on the boundary. Recall that under **(BM)**, we get from Tanaka's formula that the expectation of the local time is twice the one of the overshoot (the overshoot being defined as the distance to the boundary of the process when it exits the domain). Up to a rescaling, we can establish a connection between this overshoot and the ladder height of a Random walk with Gaussian increments. This provides an equivalent of the expectation of the local time.

Eventually, Subsection 2.3 is dedicated to the error expansion.

2.1 A first error decomposition

From (3.1.1) and the fact that f vanishes on ∂D we derive

$$\mathrm{Err}(T,h,f,x) = \mathbb{E}_x\left[v(T\wedge\tau^N, \Pi_{\bar D}(X_{T\wedge\tau^N}))\right] - v(0,x).$$

Under **(D)** we have a simple expression for $\Pi_{\bar D}$. Indeed, for $D := \{y\in\mathbb{R}^d : y_j > b_0^j,\ j\in[\![1,m]\!]\}$,

$$\forall y\in\mathbb{R}^d,\ \Pi_{\bar D}(y) = ((y_1-b_0^1)^+ + b_0^1, ..., (y_m-b_0^m)^+ + b_0^m, y_{m+1}, ..., y_d).$$

Since we have assumed **(S)**, Itô-Tanaka's formula gives

$$\begin{aligned}
\mathrm{Err}(T,h,f,x) &= \mathbb{E}_x\Bigg[\int_0^{T\wedge\tau^N}\partial_s v(s,\Pi_{\bar D}(X_s))ds + \int_0^{T\wedge\tau^N}\nabla v(s,\Pi_{\bar D}(X_s))d\Pi_{\bar D}(X_s)\\
&\quad +\frac12\int_0^{T\wedge\tau^N}\mathrm{tr}(H_v(s,\Pi_{\bar D}(X_s))d\langle\Pi_{\bar D}(X)\rangle_s)\Bigg]\\
&= \mathbb{E}_x\Bigg[\int_0^{T\wedge\tau^N}\mathbb{I}_{X_s\in D}\partial_s v(s,X_s)ds + \frac12\sum_{i=1}^m\int_0^{T\wedge\tau^N}\partial_{x_i}v(s,\Pi_{\bar D}(X_s))dL_s^{b_0^i}(X^i)\\
&\quad + \frac12\int_0^{T\wedge\tau^N}\mathrm{tr}(H_v(s,\Pi_{\bar D}(X_s))\Sigma_0(X_s))ds\Bigg]
\end{aligned}$$

where, recalling $\sigma_0\sigma_0^* = \begin{pmatrix}\Sigma & 0\\ 0 & \mathbf{I}_{d-m}\end{pmatrix}$, we have $\forall y\in\mathbb{R}^d,\ \forall (i,j)\in[\![1,m]\!]^2,\ (\Sigma_0(y))_{ij} := (\sigma_0\sigma_0^*)_{ij}\mathbb{I}_{y_i>b_0^i, y_j>b_0^j}$, and $\forall (i,j)\in[\![1,d]\!]^2\backslash[\![1,m]\!]^2,\ (\Sigma_0(y))_{ij} = \delta_{ij}$, δ being the Kronecker symbol.

Remark 3.2.1 *Note that $\forall y\in\mathbb{R}^d$ s.t. $\exists k\in[\![1,m]\!],\ y_k\le b_0^k$ one has $\forall s\in[0,T],\ v(s,\Pi_{\bar D}(y)) = 0$. Thus, $\forall (i,j)\in[\![1,m]\!]^2, i\neq k, j\neq k, \partial^2_{x_ix_j}v(s,\Pi_{\bar D}(y)) = 0$.*

From the above remark we derive

$$\begin{aligned}\mathrm{Err}(T,h,f,x) &= \mathbb{E}_x\Bigg[\int_0^{T\wedge\tau^N}\mathbb{I}_{X_s\in D}\Big(\partial_s v(s,X_s)+\frac{1}{2}\mathrm{tr}(H_v(s,X_s)\sigma_0\sigma_0^*)\Big)ds\\ &\quad+\frac{1}{2}\sum_{i=1}^m\int_0^{T\wedge\tau^N}\partial_{x_i}v(s,\Pi_{\bar D}(X_s))dL_s^{b_0^i}(X^i)\Bigg].\end{aligned}$$

Recalling that v satisfies (3.1.2) we finally write

$$\mathrm{Err}(T,h,f,x) = \frac{1}{2}\sum_{i=1}^m\mathbb{E}_x\left[\int_0^{T\wedge\tau^N}\partial_{x_i}v(s,\Pi_{\bar D}(X_s))dL_s^{b_0^i}(X^i)\right]. \tag{3.2.1}$$

2.2 Equivalent of the local time on the boundary

Preliminary one-dimensional results

We now introduce some notations related to one dimensional random walk techniques (see Siegmund [Sie79]), which will be used in the sequel. Let us define $s_0 := 0, \forall n\geq 1, s_n := \sum_{i=1}^n G^i$, where the G^i are i.i.d. standard centered normal variables. We introduce the stopping times $\bar\tau_a := \inf\{n>0 : s_n > a\}$ for $a\geq 0$, $\tau^+ := \bar\tau_0$ and define $H(x) := (\mathbb{E}_0[s_{\tau^+}])^{-1}\int_0^x dy\mathbb{P}_0[s_{\tau^+} > y]$.

Lemma 3.2.1 (Asymptotic independence of the overshoot and the discrete exit time - Equivalence of the expectation of the local time.) *Let W be a standard linear BM. Put $x>0$ and consider the domain $D :=]-\infty,x[$. We have for any $y\geq 0$*

$$\lim_{h\longrightarrow 0}\mathbb{P}_0[\tau^N\leq t,(W_{\tau^N}-x)\leq y\sqrt{h}] = \mathbb{P}_0[\tau\leq t]H(y) \tag{3.2.2}$$

$$\frac{1}{2}\mathbb{E}_0[L^x_{t\wedge\tau^N}(W)] = \sqrt{h}\frac{\mathbb{E}_0[s^2_{\tau^+}]}{2\mathbb{E}_0[s_{\tau^+}]}\mathbb{P}_0[\tau\leq t]+o(\sqrt{h}). \tag{3.2.3}$$

Both limits are uniform in $t\in[0,T]$.

One knows from [Sie79] and [AGP95] that $\dfrac{\mathbb{E}_0[s^2_{\tau^+}]}{2\mathbb{E}_0[s_{\tau^+}]} = -\dfrac{\zeta(1/2)}{\sqrt{2\pi}} = 0.5823\ldots$, where ζ denotes Riemann's Zeta function.

Proof. Equality (3.2.2) is a direct consequence of Lemma 3 in [Sie79] for a fixed t. We derive the uniformity on $[0,T]$ using Dini-like arguments noting that the l.h.s. of (3.2.2) defines a sequence of (discontinuous) increasing functions and that the simple limit is continuous (see e.g. problem 7.2.3 in [Die71]).
To prove (3.2.3), use Tanaka's formula to get

$$\frac{1}{2}\mathbb{E}_0[L^x_{t\wedge\tau^N}(W)] = \sqrt{h}\mathbb{E}_0[h^{-1/2}(W_{\tau^N}-x)\mathbb{I}_{\tau^N\leq t}]+\mathbb{E}_0[(W_t-x)^+\mathbb{I}_{\tau^N>t}] := A_1(t)+A_2(t).$$

For $A_2(t)$, introducing $\forall t\in[0,T],\ \tau_t := \inf\{s\geq t : W_s = x\}$ we get

$$A_2(t) \leq \mathbb{E}_0[\mathbb{I}_{\tau^N>t}\mathbb{I}_{\tau_{\phi(t)}\leq t}\mathbb{E}[|W_t-x||\mathcal{F}_{\tau_{\phi(t)}}]]\leq C\sqrt{h}\mathbb{E}_0[\exp(-c\frac{(W_{\phi(t)}-x)^2}{h})]$$

where the last inequality is a consequence of Lemma 1.1.2. Explicit computations give

$$A_2(t)\leq\frac{Ch}{(\phi(t)+h)^{1/2}}\exp(-c\frac{x^2}{\phi(t)+h})\leq\frac{C}{x}h.$$

To deal with $A_1(t)$, put $\Psi_N(y,t) = \mathbb{P}_0[h^{-1/2}(W_{\tau^N}-x)\geq y,\tau^N\leq t]$: it converges owing to (3.2.2) to $\Psi(y,t) := \mathbb{P}_0[\tau\leq t](1-H(y))$ uniformly on $[0,T]$. Proposition 1.5.4 guarantees that the sequence $(\Psi_N(.,t))_N$ is uniformly integrable, uniformly in $t\in[0,T]$. Thus, the dominated convergence theorem gives $\displaystyle\int_{\mathbb{R}^+}\Psi_N(y,t)dy\underset{N}{\longrightarrow}$ $\displaystyle\int_{\mathbb{R}^+}\Psi(y,t)dy = \frac{\mathbb{E}_0[s^2_{\tau^+}]}{2\mathbb{E}_0[s_{\tau^+}]}\mathbb{P}_0[\tau\leq t]$ for each t, and using again Dini-like arguments, uniformly on $[0,T]$.

□

Correlated BM in dimension $d > 1$

We now extend the result of the previous paragraph to the special case of our working assumptions **(BM)**, **(D)** when $m \geq 2$.

Lemma 3.2.2 *Assume* **(BM)**, **(D)**. *Put* $\forall i \in [\![1,m]\!],\ \tau^i := \inf\{t \geq 0 : X_t^i = b_0^i\},\ \tau^{N,i} := \inf\{t_j := jh \geq 0 : X_{t_j}^i \leq b_0^i\}$. *Sticking to the notations of the previous paragraph one has*

$$\forall y \in \mathbb{R}^{+,*},\ \mathbb{P}_x[\sqrt{h}^{-1}(X^1_{\tau^N} - b_0^1)^- \geq y, \tau^N \leq t, \tau^{N,1} \leq \wedge_{i=2}^m \tau^{N,i}] \underset{N}{\longrightarrow} (1 - H(y))\mathbb{P}_x[\tau^1 \leq t, \tau^1 < \wedge_{i=2}^m \tau^i]$$

and the limit is uniform on $[0, T]$.

Proof. Let us first show that $\forall (t,y) \in [0,T] \times \mathbb{R}^{+,*},\ \zeta_N(t) := \mathbb{P}_x[\sqrt{h}^{-1}(X^1_{\tau^{N,1}} - b_0^1)^- \geq y, \tau^N \leq t, \tau^{N,1} \leq \wedge_{i=2}^m \tau^{N,i}] \underset{N}{\longrightarrow} (1 - H(y))\mathbb{P}_x[\tau^1 \leq t, \wedge_{i=2}^m \tau^i > \tau^1] := \zeta(t)$. We write

$$\begin{aligned}
\zeta_N(t) &= \mathbb{P}_x[\sqrt{h}^{-1}(X^1_{\tau^{N,1}} - b_0^1)^- \geq y, \tau^{N,1} \leq t] - \mathbb{P}_x[\sqrt{h}^{-1}(X^1_{\tau^{N,1}} - b_0^1)^- \geq y, \tau^{N,1} \leq t, \tau^{N,1} > \wedge_{i=2}^m \tau^{N,i}] \\
&:= (\zeta_N^1 - \zeta_N^2)(t). \qquad (3.2.4)
\end{aligned}$$

From Lemma 3.2.1 one gets

$$\zeta_N^1(t) \underset{N}{\longrightarrow} \zeta^1(t) := (1 - H(y))\mathbb{P}_x[\tau^1 \leq t] \qquad (3.2.5)$$

uniformly on $[0,T]$. Let us turn to the control of ζ_N^2. As a consequence of the strong Markov property of X it comes

$$\begin{aligned}
\zeta_N^2(t) &= \mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} \leq t} \mathbb{I}_{\tau^{N,1} > \wedge_{i=2}^m \tau^{N,i}} \mathbb{P}[\sqrt{h}^{-1}(X^1_{\tau^{N,1}} - b_0^1)^- \geq y, \tau^{N,1} \leq t | \mathcal{F}_{\wedge_{i=2}^m \tau^{N,i}}]] \\
&= \mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} \leq t} \mathbb{I}_{\tau^{N,1} > \wedge_{i=2}^m \tau^{N,i}} \mathbb{P}_{X^1_{\wedge_{i=2}^m \tau^{N,i}}}[\sqrt{h}^{-1}(\tilde{X}^1_{\tilde{\tau}^{N,1}} - b_0^1)^- \geq y, \tilde{\tau}^{N,1} \leq t - \wedge_{i=2}^m \tau^{N,i}]] \\
&:= \mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} \leq t} \mathbb{I}_{\tau^{N,1} > \wedge_{i=2}^m \tau^{N,i}} \xi_N(X^1_{\wedge_{i=2}^m \tau^{N,i}}, \wedge_{i=2}^m \tau^{N,i}, t)]
\end{aligned}$$

where $(\tilde{X}_t^1)_{t \geq 0}$ is a standard BM with starting point $X^1_{\wedge_{i=2}^m \tau^{N,i}}$ and $\tilde{\tau}^{N,1} := \inf\{t_i \geq 0 : \tilde{X}^1_{t_i} \leq b_0^1\}$.

For a given arbitrary compact interval $\mathcal{K} := [\underline{\mathcal{K}}, \overline{\mathcal{K}}] \subset (b_0^1, +\infty)$ we split $\zeta_N^2(t)$ into two parts.

$$\begin{aligned}
\zeta_N^2(t) &= \mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} \leq t} \mathbb{I}_{\tau^{N,1} > \wedge_{i=2}^m \tau^{N,i}} \mathbb{I}_{X^1_{\wedge_{i=2}^m \tau^{N,i}} \in \mathcal{K}} \xi_N(X^1_{\wedge_{i=2}^m \tau^{N,i}}, \wedge_{i=2}^m \tau^{N,i}, t)] \\
&+ \mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} \leq t} \mathbb{I}_{\tau^{N,1} > \wedge_{i=2}^m \tau^{N,i}} \mathbb{I}_{X^1_{\wedge_{i=2}^m \tau^{N,i}} \notin \mathcal{K}} \xi_N(X^1_{\wedge_{i=2}^m \tau^{N,i}}, \wedge_{i=2}^m \tau^{N,i}, t)] := \zeta_N^{21}(t) + \zeta_N^{22}(t).
\end{aligned}$$

Fix $\varepsilon > 0$. We now show that one can choose $\mathcal{K}(\varepsilon),\ N_0 := N_0(\varepsilon, \mathcal{K}(\varepsilon))$ s.t. for $N \geq N_0$,

$$\zeta_N^2(t) = (1 - H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \leq t] + O(\varepsilon). \qquad (3.2.6)$$

Control of $\zeta_N^{21}(t)$.

Write first

$$\begin{aligned}
\zeta_N^{21}(t) &= \Big(\zeta_N^{21}(t) - (1 - H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \leq t, X^1_{\wedge_{i=2}^m \tau^i} \in \mathcal{K}]\Big) \\
&\quad + (1 - H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \leq t] - R(t, \mathcal{K})
\end{aligned}$$

where $R(t,\mathcal{K}) = (1 - H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \leq t, X^1_{\wedge_{i=2}^m \tau^i} \notin \mathcal{K}]$.

Note that

$$0 \leq R(t, \mathcal{K}) \leq \mathbb{P}_x[\wedge_{i=2}^m \tau^i \leq T, X^1_{\wedge_{i=2}^m \tau^i} \geq \overline{\mathcal{K}}] + \mathbb{P}_x[\wedge_{i=2}^m \tau^i \leq T, X^1_{\wedge_{i=2}^m \tau^i} \in (b_0^1, \underline{\mathcal{K}}]] := R_1(\overline{\mathcal{K}}) + R_2(\underline{\mathcal{K}}).$$

Lemma 1.1.2 readily gives $R_1(\overline{\mathcal{K}}) \leq C \exp(-c\frac{(\overline{\mathcal{K}} - x_1)^2}{T})$. On the other hand, $R_2(\underline{\mathcal{K}}) \underset{\underline{\mathcal{K}} \longrightarrow b_0^1}{\longrightarrow} 0$.

Hence, for $\varepsilon > 0$ we can choose $\mathcal{K} = \mathcal{K}(\varepsilon)$ s.t.

$$\begin{aligned}\zeta_N^{21}(t) &= \Big(\zeta_N^{21}(t) - (1-H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \le t, X^1_{\wedge_{i=2}^m \tau^i} \in \mathcal{K}]\Big) + (1-H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \le t] \\ &\quad + O(\varepsilon) := \delta_N(t) + (1-H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \le t] + O(\varepsilon).\end{aligned}$$

For the term $\delta_N(t)$ we introduce the following Lemma whose proof is postponed to the end of the section.

Lemma 3.2.3 *Let $\tilde X^1$ be a standard BM with starting point $\tilde x$ in a given compact interval $\mathcal{K} = [\underline{\mathcal{K}}, \overline{\mathcal{K}}] \subset (b_0^1, +\infty)$. Then*

$$\mathbb{P}_{\tilde x}[\sqrt{h}^{-1}(\tilde X^1_{\tilde\tau^{N,1}} - b_0^1)^- \ge y, \tilde\tau^{N,1} \le u] \underset{N}{\longrightarrow} (1-H(y))\mathbb{P}_{\tilde x}[\tilde\tau^1 \le u]$$

uniformly on $(\tilde x, u) \in \mathcal{K} \times [0,T]$.

From Lemma 3.2.3, $\forall \varepsilon > 0$, $\exists N_0 := N_0(\mathcal{K}(\varepsilon), \varepsilon)$, s.t. $N \ge N_0$

$$\begin{aligned}\delta_N(t) &= (1-H(y))\Big\{\mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} \le t}\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} < \tau^{N,1}}\mathbb{I}_{X^1_{\wedge_{i=2}^m \tau^{N,i}} \in \mathcal{K}}\mathbb{P}_{X^1_{\wedge_{i=2}^m \tau^{N,i}}}[\tilde\tau^1 \le t - \wedge_{i=2}^m \tau^{N,i}]] \\ &\quad - \mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \le t, X^1_{\wedge_{i=2}^m \tau^i} \in \mathcal{K}]\Big\} + O(\varepsilon) \\ &:= (1-H(y))\Big(\mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} \le t}\mathbb{I}_{\wedge_{i=2}^m \tau^{N,i} < \tau^{N,1}}\mathbb{I}_{X^1_{\wedge_{i=2}^m \tau^{N,i}} \in \mathcal{K}}\xi_t(X^1_{\wedge_{i=2}^m \tau^{N,i}}, \wedge_{i=2}^m \tau^{N,i})] \\ &\quad - \mathbb{E}_x[\mathbb{I}_{\wedge_{i=2}^m \tau^i \le t}\mathbb{I}_{\wedge_{i=2}^m \tau^i < \tau^1}\mathbb{I}_{X^1_{\wedge_{i=2}^m \tau^i} \in \mathcal{K}}\xi_t(X^1_{\wedge_{i=2}^m \tau^i}, \wedge_{i=2}^m \tau^i)]\Big) + O(\varepsilon).\end{aligned}$$

Note that $\xi_t(x,u) = \mathbb{P}_x[\tilde\tau \le t-u]$ is continuous in $(x,u) \in (b_0^1, +\infty) \times [0,t]$. Recall that $\tau^{N,i} \underset{N,\text{a.s}}{\longrightarrow} \tau^i$, $i \in [\![1,m]\!]$, and by continuity $X^1_{\wedge_{i=2}^m \tau^{N,i}} \underset{N,\text{a.s}}{\longrightarrow} X^1_{\wedge_{i=2}^m \tau^i}$. One can check that the law of $(\tau^1, \wedge_{i=2}^m \tau^i, X^1_{\wedge_{i=2}^m \tau^i})$ is absolutely continuous w.r.t. the Lebesgue measure. We thus derive by convergence in law that for N large enough

$$\begin{aligned}\delta_N(t) &= O(\varepsilon), \\ \zeta_N^{21}(t) &= (1-H(y))\mathbb{P}_x[\wedge_{i=2}^m \tau^i < \tau^1, \tau^1 \le t] + O(\varepsilon).\end{aligned} \tag{3.2.7}$$

Control of $\zeta_N^{22}(t)$.

The arguments we use to control this term are quite similar to those introduced to treat the terms $R_1(\overline{\mathcal{K}}), R_2(\underline{\mathcal{K}})$ above.

Indeed, since $\xi_N \in [0,1]$ one gets

$$\zeta_N^{22}(t) \le \mathbb{P}_x[\wedge_{i=2}^m \tau^{N,i} \le T, X^1_{\wedge_{i=2}^m \tau^{N,i}} \ge \overline{\mathcal{K}}] + \mathbb{P}_x[\wedge_{i=2}^m \tau^{N,i} \le T, X^1_{\wedge_{i=2}^m \tau^{N,i}} \in (b_0^1, \underline{\mathcal{K}}]] := R_1^N(\overline{\mathcal{K}}) + R_2^N(\underline{\mathcal{K}}).$$

From Lemma 1.1.2 we get $R_1^N(\overline{\mathcal{K}}) \le \mathbb{P}_{x_1}[\sup_{s\in[0,T]} X_s^1 \ge \overline{\mathcal{K}}] \le C\exp(-c\frac{(\overline{\mathcal{K}}-x_1)^2}{T})$. The previous choice of $\overline{\mathcal{K}}$ gives $R_1^N(\overline{\mathcal{K}}) = O(\varepsilon)$.

With the previous notations we write,

$$R_2^N(\underline{\mathcal{K}}) := (R_2^N(\underline{\mathcal{K}}) - R_2(\underline{\mathcal{K}})) + R_2(\underline{\mathcal{K}}).$$

On the one hand, the former choice of $\underline{\mathcal{K}}$ yields $R_2(\underline{\mathcal{K}}) = O(\varepsilon)$. On the other hand, for the difference $(R_2^N - R_2)(\underline{\mathcal{K}})$, since $\tau^{N,i} \underset{N,\text{ a.s}}{\longrightarrow} \tau^i$, $i \in [\![2,m]\!]$, $X^1_{\wedge_{i=2}^m \tau^{N,i}} \underset{N,\text{ a.s}}{\longrightarrow} X^1_{\wedge_{i=2}^m \tau^i}$, with the same arguments we employed to control $\delta_N(t)$, we derive by convergence in law

$$\exists N_0 := N_0(\underline{\mathcal{K}}, \varepsilon),\ N \ge N_0,\ |(R_2^N - R_2)(\underline{\mathcal{K}})| \le \varepsilon.$$

Hence, for $N := N(\mathcal{K}, \varepsilon)$ large enough, we write

$$\zeta_N^{22}(t) = O(\varepsilon). \tag{3.2.8}$$

Equations (3.2.7) and (3.2.8) give (3.2.6). From (3.2.6), (3.2.5) and (3.2.4) we derive the simple convergence of ζ_N to ζ for a fixed $t \in [0,T]$.

The uniformity in $t \in [0,T]$ derives from the fact that $\zeta_N(t)$ is a cumulative distribution function with continuous limit, see also the arguments at the beginning of the proof of Lemma 3.2.1.

$\square$

Proof of Lemma 3.2.3.

Let us define $\forall (x,u) \in \mathcal{K} = [\underline{\mathcal{K}}, \overline{\mathcal{K}}] \times [0,T]$, $\underline{\mathcal{K}} > b_0^1$, $\Psi_N(x,u) = \mathbb{P}_x[\sqrt{h}^{-1}(\tilde{X}^1_{\tilde{\tau}^{N,1}} - b_0^1)^- \geq y, \tilde{\tau}^{N,1} \leq u]$. For a fixed $x \in \mathcal{K}$, Lemma 3.2.1 yields that $\Psi_N(x,u) \underset{N}{\longrightarrow} (1-H(y))\mathbb{P}_x[\tau^1 \leq u] := \Psi(x,u)$ uniformly on $u \in [0,T]$. Let us now show that for a fixed $u \in [0,T]$ we have the uniform convergence w.r.t. $x \in \mathcal{K}$. Write

$$\Psi_N(x,u) \quad = \quad \mathbb{P}_x[\sqrt{h}^{-1}(\tilde{X}^1_{\tilde{\tau}^{N,1}} - b_0^1)^- \geq y] - \mathbb{P}_x[\sqrt{h}^{-1}(\tilde{X}^1_{\tilde{\tau}^{N,1}} - b_0^1)^- \geq y, \tilde{\tau}^{N,1} > u] := \Psi_N^1(x) - \Psi_N^2(x,u).$$

With the notations of the previous paragraph, $\Psi_N^1(x) = \mathbb{P}_x[\sqrt{h}^{-1}(\tilde{X}^1_{\tilde{\tau}^{N,1}} - b_0^1)^- \geq y] = \mathbb{P}_0[(s_{\bar{\tau}_{(x-b_0^1)/\sqrt{h}}} - (x - b_0^1)/\sqrt{h}) \geq y]$. Equation (19) from [Sie79] gives $\lim_{b\longrightarrow\infty} \mathbb{P}_0[s_{\bar{\tau}_b} - b \geq y] = 1 - H(y)$. Hence, $\Psi_N^1(x) \underset{N}{\longrightarrow} (1 - H(y))$ uniformly on $x \in \mathcal{K}$. We develop Ψ_N^2 like in the proof of Lemma 3 from the same reference, controlling that we can isolate uniform rests. We get

$$\begin{aligned}
\Psi_N^2(x,u) &= \mathbb{P}[(s_{\bar{\tau}_{(x-b_0^1)/\sqrt{h}}} - (x-b_0^1)/\sqrt{h}) \geq y, \bar{\tau}_{(x-b_0^1)/\sqrt{h}} > \phi(u)/h] \\
&= \int_0^\infty \mathbb{P}[\bar{\tau}_{(x-b_0^1)/\sqrt{h}} > \phi(u)/h, (x-b_0^1)/\sqrt{h} - s_{\phi(u)/h} \in [z, z+dz)]\mathbb{P}[s_{\bar{\tau}_z} - z \geq y].
\end{aligned}$$

We split the above integral into three terms $\Psi_N^{21}, \Psi_N^{22}, \Psi_N^{23}$ respectively associated to the intervals $(0, \varepsilon(x - b_0^1)/\sqrt{h})$, $(\varepsilon(x-b_0^1)/\sqrt{h}, (x-b_0^1)/(\varepsilon\sqrt{h}))$, $((x-b_0^1)/(\varepsilon\sqrt{h}), \infty)$. One has

$$\begin{aligned}
\Psi_N^{21}(x,u) &\leq \mathbb{P}[\frac{(1-\varepsilon)(x-b_0^1)}{\sqrt{h}\sqrt{\phi(u)/h}} \leq \mathcal{N}(0,1) \leq \frac{x-b_0^1}{\sqrt{h}\sqrt{\phi(u)/h}}] \\
&\leq \mathbb{P}[\frac{(1-\varepsilon)(x-b_0^1)}{T^{1/2}} \leq \mathcal{N}(0,1) \leq \frac{x-b_0^1}{T^{1/2}}] \\
&\leq \frac{C\varepsilon(\overline{\mathcal{K}} - b_0^1)}{T^{1/2}}
\end{aligned}$$

uniformly for $x \in \mathcal{K}$. We also have

$$\begin{aligned}
\Psi_N^{23}(x,u) &\leq \mathbb{P}[\mathcal{N}(0,1) \leq (1-\varepsilon^{-1})\frac{x-b_0^1}{\phi(u)^{1/2}}] \leq \mathbb{P}[\mathcal{N}(0,1) \leq (1-\varepsilon^{-1})\frac{\underline{\mathcal{K}}-b_0^1}{T^{1/2}}] \\
&\leq \frac{CT^{1/2}}{\underline{\mathcal{K}} - b_0^1}\frac{\varepsilon}{1-\varepsilon}
\end{aligned}$$

which is still uniform w.r.t. $x \in \mathcal{K}$. From these computations we derive that for N large enough, $\Psi_N^{22}(x,u) = (1-H(y))\mathbb{P}_x[\tilde{\tau}^{N,1} > u] + O(\varepsilon)$, where the rest is uniform w.r.t. $\mathcal{K}$. It therefore remains to show $\mathbb{P}_x[\tilde{\tau}^{N,1} > u] := \gamma_N(u,x) \underset{N}{\longrightarrow} \gamma(u,x) := \mathbb{P}_x[\tilde{\tau}^1 > u]$ uniformly on $\mathcal{K}$. We note that $1 - \gamma_N(u,x) = \mathbb{P}_0[\sup_{i \in [\![0,\phi(u)/h]\!]} \tilde{X}^1_{t_i} \geq (x-b_0^1)]$ is decreasing in x, so that $\gamma_N(u,.)$ is increasing. Since the simple limit is continuous, we derive the uniformity using the same arguments as in the proof of Lemma 3.2.1.

Now, we have shown that for a fixed parameter $x \in \mathcal{K}$, $u \in [0,T]$, we have the uniform convergence with respect to the other. Let us now show the joint uniform convergence. The limit Ψ is uniformly continuous on $\mathcal{K} \times [0,T]$. This reads

$$\forall \varepsilon > 0, \exists \eta := \eta(\varepsilon), \forall (x,x') \times (t,t') \in \mathcal{K}^2 \times [0,T]^2,\ |t-t'| + |x-x'| \leq \eta,\ |\Psi(x,t) - \Psi(x',t')| \leq \varepsilon. \quad (3.2.9)$$

In particular, $|t-t'| \leq \eta \Rightarrow \sup_{x\in\mathcal{K}} |\Psi(x,t) - \Psi(x,t')| \leq \varepsilon$. Let us now consider a regular grid $\Lambda := \{s_i\}_{i\in[\![1,a]\!]}$ of $[0,T]$ with step $s = s_{i+1} - s_i \leq \eta$. Since for a fixed $t \in [0,T]$ we have uniform convergence in space it comes

$$\forall \varepsilon > 0,\ \exists N_0 = \max_{i\in[\![1,a]\!]} N_0(s_i),\ N \geq N_0,\ \sup_{i\in[\![1,a]\!]} \sup_{x\in\mathcal{K}} |\Psi_N(x,s_i) - \Psi(x,s_i)| \leq \varepsilon. \quad (3.2.10)$$

Noting that both $\Psi_N(x,.), \Psi(x,.)$ are increasing functions we derive from (3.2.9), (3.2.10)

$$\begin{aligned}
&\forall t \in [s_i, s_{i+1}],\ \Psi(x,s_i) - \Psi(x,s_{i+1}) + \Psi(x,s_{i+1}) - \Psi_N(x,s_{i+1}) \\
&\le \Psi(x,t) - \Psi_N(x,t) \le \Psi(x,s_{i+1}) - \Psi(x,s_i) + \Psi(x,s_i) - \Psi_N(x,s_i), \\
&\forall \varepsilon > 0, \exists N_0,\ N \ge N_0,\ \sup_{t\in[0,T]} \sup_{x\in\mathcal{K}} |\Psi(x,t) - \Psi_N(x,t)| \le \varepsilon
\end{aligned}$$

which shows the joint uniformity and completes the proof.

□

2.3 Expansion result

Our main expansion result is given by the following

Theorem 3.2.4 (Error expansion for the correlated Brownian motion in an orthant). *Assume* **(BM)**, **(D)** *and* **(S)**. *For h small enough the error writes*

$$\mathrm{Err}(T,h,f,x) = C_1\sqrt{h} + o(\sqrt{h})$$

with $C_1 = C_0 \sum_{i=1}^m \mathbb{E}_x\left[\mathbb{I}_{\tau^i \le T, \wedge_{j\in[\![1,m]\!]\setminus\{i\}}\tau^j > \tau^i}(\partial_{y_i} v(\tau^i, X_{\tau^i}))\right],\ C_0 = \dfrac{\mathbb{E}_0[s^2_{\tau^+}]}{2\mathbb{E}_0[s_{\tau^+}]}$.

Proof.

We separate the case $m=1$ corresponding to the half-space, which is much easier, from the case $m>1$.

Case $m=1$.

From (3.2.1) the error writes $\mathrm{Err}(T,h,f,x) = \frac{1}{2}\mathbb{E}_x\left[\int_0^T Y_s dL^{b_0}_{s\wedge\tau^N}(X^1_.)\right]$, with

$$Y_s = \partial_{x_1} v(s,(b_0, X^2_s, \cdots, X^d_s)).$$

Note that $dY_s = y_s ds + dM_s$, where M is a martingale, and that Y (and $(y_s)_{s\in[0,T]}$) is independent of X^1 (and hence of $\tau = \tau^1$ and $\tau^N = \tau^{N,1}$). Exploiting these independence properties and using twice the integration by parts combined with equality (3.2.3), one obtains

$$\begin{aligned}
\mathbb{E}_x\left[\int_0^T Y_s dL^{b_0}_{s\wedge\tau^N}(X^1)\right] &= \mathbb{E}_x\left[Y_T L^{b_0}_{T\wedge\tau^N}(X^1)\right] - \mathbb{E}_x\left[\int_0^T L^{b_0}_{s\wedge\tau^N}(X^1) y_s ds\right] \\
&= \sqrt{h}\frac{\mathbb{E}_0[s^2_{\tau^+}]}{\mathbb{E}_0[s_{\tau^+}]}\left(\mathbb{P}_x[\tau \le T]\mathbb{E}_x[Y_T] - \int_0^T \mathbb{P}_x[\tau \le s]\mathbb{E}_x[y_s] ds\right) + o(\sqrt{h}) \\
&= \sqrt{h}\frac{\mathbb{E}_0[s^2_{\tau^+}]}{\mathbb{E}_0[s_{\tau^+}]}\mathbb{E}_x\left[Y_\tau \mathbb{I}_{\tau\le T}\right] + o(\sqrt{h})
\end{aligned}$$

and the result follows.

Case $m>1$.

In this case we lose the above nice independence property that made the problem essentially one-dimensional for the half-space case. From (3.2.1) write

$$\mathrm{Err}(T,h,f,x) = \frac{1}{2}\sum_{i=1}^m \mathbb{E}_x\left[\int_0^T \partial_{x_i} v(s, \Pi_{\bar{D}}(X_s)) dL^{b^i_0}_{s\wedge\tau^N}(X^i)\right] = \sum_{i=1}^m E_i.$$

Limit behaviour of E_i

We detail the term E_1, the other ones can be handled exactly in the same way. Since the measure $dL_s^{b_0^1}(X^1)$ is a.s. carried by the set $\{s \geq 0 : X_s^1 = b_0^1\}$ we write

$$E_1 = \frac{1}{2}\int_0^T \mathbb{E}_x[\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))dL_{s\wedge\tau^N}^{b_0^1}(X^1)]$$

where for simplicity we denote $\forall y \in \mathbb{R}^{d-1}$, $\Pi_{\bar{D}^{2,d}}(y) := ((y_1 - b_0^2)^+ + b_0^2, ..., (y_{m-1} - b_0^m)^+ + b_0^m, y_m, ..., y_{d-1})$.

Recall from Remark 3.2.1 that since v vanishes on ∂D, $\forall (s, y) \in [0, T] \times \mathbb{R}^{d-1}$, $\exists k \in [\![2, m]\!]$, $y_{k-1} \leq b_0^k$ we have

$$\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}_0^{2,d}}(y)) = 0.$$

In other words, there are no contributions associated to the corner points in E_1.

Under **(S)**, we can integrate by part in E_1. We obtain

$$\begin{aligned} E_1 &= \frac{1}{2}\Bigg\{\mathbb{E}_x[L_{T\wedge\tau^N}^{b_0^1}(X^1)\partial_{x_1} v(T, b_0^1, \Pi_{\bar{D}^{2,d}}(X_T^2, ..., X_T^d))] \\ &\quad -\mathbb{E}_x[\int_0^T L_{s\wedge\tau^N}^{b_0^1}(X^1)d\big(\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))\big)]\Bigg\}. \end{aligned}$$

Developing the last Itô differential with Itô-Tanaka's formula we get

$$\begin{aligned} &d\big(\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))\big) = dM_s + dy_s, \\ &dM_s = \sum_{i=2}^d \partial^2_{x_i,x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))(\mathbb{I}_{X_s^i > b_0^i, i\in[\![2,m]\!]} + \mathbb{I}_{i\in[\![m+1,d]\!]})dX_s^i, \\ &dy_s = \sum_{i=2}^m \partial^2_{x_i,x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))\frac{1}{2}dL_s^{b_0^i}(X^i) + \Bigg(\partial_s\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d)) \\ &+\frac{1}{2}\sum_{i,j=2}^d \partial^3_{x_1,x_i,x_j} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))(\rho_{ij}\mathbb{I}_{X_s^i > b_0^i, X_s^j > b_0^j}\mathbb{I}_{(i,j)\in[\![2,m]\!]^2} + \mathbb{I}_{i=j}\mathbb{I}_{i\in[\![m+1,d]\!]})\Bigg) ds. \end{aligned} \tag{3.2.11}$$

From Remark 3.2.1 we derive that for $(s, y) \in [0, T] \times \mathbb{R}^{d-1}$, $\exists j \in [\![2, m]\!]$, $y_{j-1} \leq b_0^j$, the crossed derivatives $\partial_{x_1,x_j} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(y)) = \partial^3_{x_1,x_i,x_j} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(y)) = 0$. Hence, in dy_s there is no contribution of the local time of the process on the corner regions and

$$\begin{aligned} dy_s &= \Bigg(\partial_s\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d)) + \frac{1}{2}\sum_{i,j=2}^d \partial^3_{x_1,x_i,x_j} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))(\rho_{ij}\mathbb{I}_{(i,j)\in[\![2,m]\!]^2} \\ &\quad +\mathbb{I}_{i=j}\mathbb{I}_{i\in[\![m+1,d]\!]})\Bigg)ds := \Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))ds. \end{aligned}$$

We point out that under **(S)**, $\Theta \in C_b^{1/2,1}([0, T] \times \bar{D}), \Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}(.))|_{\partial D^{2,d}} = 0$ and $(M_s)_{s\in[0,T]}$ is a true martingale.

Tanaka's formula then gives

$$\begin{aligned} E_1 &= \mathbb{E}_x[\Bigg((X_{T\wedge\tau^N}^1 - b_0^1)^- + \int_0^{T\wedge\tau^N} \mathbb{I}_{X_s^1 < b_0^1} dX_s^1\Bigg)\partial_{x_1} v(T, b_0^1, \Pi_{\bar{D}^{2,d}}(X_T^2, ..., X_T^d))] \\ &\quad -\mathbb{E}_x[\int_0^T ds\Bigg((X_{s\wedge\tau^N}^1 - b_0^1)^- + \int_0^{s\wedge\tau^N} \mathbb{I}_{X_u^1 < b_0^1} dX_u^1\Bigg)\Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))] \\ &= \mathbb{E}_x[(X_{\tau^N}^1 - b_0^1)^-\mathbb{I}_{\tau^N\leq T}\partial_{x_1} v(T, b_0^1, \Pi_{\bar{D}^{2,d}}(X_T^2, ..., X_T^d))] \\ &\quad -\mathbb{E}_x[\int_0^T ds(X_{\tau^N}^1 - b_0^1)^-\mathbb{I}_{\tau^N\leq s}\Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))] + R_1 + R_2 \end{aligned} \tag{3.2.12}$$

where

$$\begin{aligned}
R_1 &= \mathbb{E}_x[\tilde{M}_T \partial_{x_1} v(T, b_0^1, \Pi_{\bar{D}^{2,d}}(X_T^2, ..., X_T^d))] - \mathbb{E}_x[\int_0^T \tilde{M}_s d\big(\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))\big)],\\
\tilde{M}_s &:= \int_0^s \mathbb{I}_{X_u^1 < b_0^1, u < \tau^N} dX_u^1,\\
R_2 &= -\mathbb{E}_x[\int_0^T ds (X_s^1 - b_0^1)^- \mathbb{I}_{\tau^N > s} \Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2, ..., X_s^d))].
\end{aligned}$$

Note in (3.2.12) that $(X_{\tau^N}^1 - b_0^1)^- \neq 0 \iff \tau^N = \tau^{N,1}$. Now, our strategy consists in conditioning w.r.t. $\tau^{N,1}$ in (3.2.12). In order to isolate the overshoot $Z_N := \sqrt{h}^{-1}(X_{\tau^N}^1 - b_0^1)^-$ we rewrite the components $X^2, ..., X^m$, of the correlated part of X in terms of X^1 and an additional correlated $(m-1)$-dimensional BM $\tilde{X}$ independent of X^1 and $(X^i)_{i \in [\![m+1, d]\!]}$. Namely,

$$\forall i \in [\![2, m]\!],\ X_s^i = \rho_{1i} X_s^1 + (1 - \rho_{1i}^2)^{1/2} \tilde{X}_s^{i-1},\ \tilde{X}_0^{i-1} = \frac{x_0^i - \rho_{1i} x_0^1}{(1 - \rho_{1i}^2)^{1/2}}.$$

Introduce also for notational convenience

$$\forall (s,t) \in [0,T]^2,\ s \geq t,\ X_{t \to s} := X_s - X_t,$$

as well as, for all $s \in [0,T]$,

$$\begin{aligned}
(\rho_{1.} X_s^1 + (1 - \rho_{1.}^2)^{1/2} \tilde{X}_s^{\cdot -1})^{2,m} &:= (\rho_{12} X_s^1 + (1 - \rho_{12}^2)^{1/2} \tilde{X}_s^1, ..., \rho_{1m} X_s^1 + (1 - \rho_{1m}^2)^{1/2} \tilde{X}_s^{m-1})\\
&= (X_s^2, ..., X_s^m) := X_s^{2,m},\\
X_s^{m+1,d} &:= (X_s^{m+1}, ..., X_s^d).
\end{aligned}$$

With the above notations it comes

$$\begin{aligned}
E_1 = \sqrt{h}\Big\{ &\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1} \leq T, \wedge_{j=2}^m \tau^{N,j} \geq \tau^{N,1}} \mathbb{E}[\partial_{x_1} v(T, b_0^1, \Pi_{\bar{D}^{2,d}}((\rho_{1.} X_{\tau^{N,1}}^1 + (1 - \rho_{1.}^2)^{1/2} \tilde{X}_{\tau^{N,1}}^{\cdot -1})^{2,m} + X_{\tau^{N,1} \to T}^{2,m},\\
&X_{\tau^{N,1}}^{m+1,d} + X_{\tau^{N,1} \to T}^{m+1,d}))|\mathcal{F}_{\tau^{N,1}}]] - \int_0^T ds \mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1} \leq s, \wedge_{j=2}^m \tau^{N,j} \geq \tau^{N,1}} \times\\
&\mathbb{E}[\Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}((\rho_{1.} X_{\tau^{N,1}}^1 + (1 - \rho_{1.}^2)^{1/2} \tilde{X}_{\tau^{N,1}}^{\cdot -1})^{2,m} + X_{\tau^{N,1} \to s}^{2,m}, X_{\tau^{N,1}}^{m+1,d} + X_{\tau^{N,1} \to s}^{m+1,d}))|\mathcal{F}_{\tau^{N,1}}]]\Big\} + R_1 + R_2.
\end{aligned}$$

Introducing explicitly the overshoot Z_N in the conditional expectations of the former development we get

$$\begin{aligned}
&E_1 = \sqrt{h}\{\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1} \leq T, \wedge_{j=2}^m \tau^{N,j} \geq \tau^{N,1}} \times\\
&\mathbb{E}[\partial_{x_1} v(T, b_0^1, \Pi_{\bar{D}^{2,d}}((\rho_{1.}(-\sqrt{h} Z_N + b_0^1) + (1 - \rho_{1.}^2)^{1/2} \tilde{X}_{\tau^{N,1}}^{\cdot -1})^{2,m} + X_{\tau^{N,1} \to T}^{2,m}, X_{\tau^{N,1}}^{m+1,d} + X_{\tau^{N,1} \to T}^{m+1,d}))|\mathcal{F}_{\tau^{N,1}}]]\\
&- \int_0^T ds \mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1} \leq s, \wedge_{j=2}^m \tau^{N,j} \geq \tau^{N,1}} \mathbb{E}[\Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}((\rho_{1.}(-\sqrt{h} Z_N + b_0^1) + (1 - \rho_{1.}^2)^{1/2} \tilde{X}_{\tau^{N,1}}^{\cdot -1})^{2,m} + X_{\tau^{N,1} \to s}^{2,m},\\
&X_{\tau^{N,1}}^{m+1,d} + X_{\tau^{N,1} \to s}^{m+1,d}))|\mathcal{F}_{\tau^{N,1}}]]\} + R_1 + R_2. \qquad (3.2.13)
\end{aligned}$$

Thanks to the strong Markov property the above conditional expectations are functions of $Z_N, \tau^{N,1}, \tilde{X}_{\tau^{N,1}}, X_{\tau^{N,1}}^{m+1,d}$. Introduce now, $\forall s \in [0,T], \forall (z,u,w) \in \mathbb{R}^- \times [0,s] \times \mathbb{R}^{d-1}$,

$$\begin{aligned}
&\Phi_s^h(z,u,w) :=\\
&\mathbb{E}[\Theta(s, b_0^1, \Pi_{\bar{D}^{2,d}}((\rho_{1.}(\sqrt{h} z + b_0^1) + (1 - \rho_{1.}^2)^{1/2} w^{\cdot -1})^{2,m} + X_{u \to s}^{2,m}, w^{m,d-1} + X_{u \to s}^{m+1,d}))].
\end{aligned}$$

Similarly, we define $\forall (z,u,w) \in \mathbb{R}^- \times [0,T] \times \mathbb{R}^{d-1}$,

$$\Psi^h(z,u,w) := \mathbb{E}[\partial_{x_1} v\big(T, b_0^1, \Pi_{\bar{D}^{2,d}}((\rho_{1.}(\sqrt{h} z + b_0^1) + (1 - \rho_{1.}^2)^{1/2} w^{\cdot -1})^{2,m} + X_{u \to T}^{2,m}, w^{m,d-1} + X_{u \to T}^{m+1,d}))].$$

Using these functions in (3.2.13) we get

$$\begin{aligned} E_1 &= \sqrt{h}\Big\{\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1}\le T, \wedge_{j=2}^m \tau^{N,j}\ge\tau^{N,1}}\Psi^h(-Z_N,\tau^{N,1},(\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}}))] \\ &\quad -\int_0^T ds\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1}\le s, \wedge_{j=2}^m \tau^{N,j}\ge\tau^{N,1}}\Phi_s^h(-Z_N,\tau^{N,1},(\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}}))]\Big\} + R_1 + R_2. \end{aligned}$$

Let us set $\Phi_s(u,w)=\Phi_s^h(0,u,w)$, $\Psi(u,w)=\Psi^h(0,u,w)$, it comes

$$\begin{aligned} E_1 &= \sqrt{h}\Big\{\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1}\le T, \wedge_{j=2}^m \tau^{N,j}\ge\tau^{N,1}}\Psi(\tau^{N,1},\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}})] \\ &\quad -\int_0^T ds\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1}\le s, \wedge_{j=2}^m \tau^{N,j}\ge\tau^{N,1}}\Phi_s(\tau^{N,1},\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}})]\Big\} + R_1 + R_2 + R_3 \end{aligned} \tag{3.2.14}$$

where

$$\begin{aligned} R_3 &:= \sqrt{h}\Big\{\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1}\le T, \wedge_{j=2}^m \tau^{N,j}\ge\tau^{N,1}}(\Psi^h(-Z_N,\tau^{N,1},\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}}) - \Psi^h(0,\tau^{N,1},\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}}))] \\ &\quad -\int_0^T ds\mathbb{E}_x[Z_N \mathbb{I}_{\tau^{N,1}\le s, \wedge_{j=2}^m \tau^{N,j}\ge\tau^{N,1}}(\Phi_s^h(-Z_N,\tau^{N,1},\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}}) - \Phi_s^h(0,\tau^{N,1},\tilde{X}_{\tau^{N,1}},X^{m+1,d}_{\tau^{N,1}}))]\Big\}. \end{aligned}$$

Put $R := R_1 + R_2 + R_3$. Recall that the law of $(\tau^1, \wedge_{i=2}^m \tau^i)$ is absolutely continuous w.r.t. the Lebesgue measure and that under **(S)**, Φ_s, Ψ are continuous and bounded. Thus, from (3.2.14), Lemma 3.2.2 and uniform integrability arguments similar to those of the proof of Lemma 3.2.1, we derive by convergence in law that for h small enough

$$\begin{aligned} E_1 &= \sqrt{h}\Big\{\mathbb{E}_x[Z \mathbb{I}_{T>\tau^1, \wedge_{j=2}^m \tau^{j}>\tau^{1}}\Psi(\tau^1,\tilde{X}_{\tau^1},X^{m+1,d}_{\tau^1})] \\ &\quad -\int_0^T ds\mathbb{E}_x[Z \mathbb{I}_{s>\tau^1, \wedge_{j=2}^m \tau^{j}>\tau^{1}}\Phi_s(\tau^1,\tilde{X}_{\tau^1},X^{m+1,d}_{\tau^1})]\Big\} + R + o(\sqrt{h}) \end{aligned} \tag{3.2.15}$$

where Z is independent of τ^1 and with distribution function given by H defined in Section 2.2. From the limit laws, the asymptotic independence of the overshoot and the former semi-martingale expansion of $(\partial_{x_1} v(s, b_0^1, \Pi_{\bar{D}^{2,d}}(X_s^2,...,X_s^d)))_{s\ge 0}$ given in (3.2.11), one obtains

$$\begin{aligned} E_1 &= \sqrt{h}\mathbb{E}[Z]\Big\{\mathbb{E}_x[\mathbb{I}_{\tau^1\le T, \wedge_{j=2}^m \tau^{j}>\tau^{1}}\big(\partial_{x_1} v(0,b_0^1,x^2,...,x^d) + M_T + \int_0^T \Theta(s,b_0^1,\Pi_{\bar{D}^{2,d}}(X_s^2,...,X_s^d))ds\big)] \\ &\quad -\mathbb{E}_x[\mathbb{I}_{\tau^1\le T, \wedge_{j=2}^m \tau^{j}>\tau^{1}}\mathbb{E}[\int_{\tau^1}^T \Theta(s,b_0^1,\Pi_{\bar{D}^{2,d}}(X_s^2,...,X_s^d))ds|\mathcal{F}_{\tau^1}]]\Big\}. \end{aligned}$$

Thus,

$$\begin{aligned} E_1 &= \sqrt{h}\mathbb{E}[Z]\Big\{\mathbb{E}_x[\mathbb{I}_{\tau^1\le T, \wedge_{j=2}^m \tau^{j}>\tau^{1}}\big(\partial_{x_1} v(0,b_0^1,x^2,...,x^d) + M_{\tau^1} + (M_T - M_{\tau^1}) \\ &\quad + \int_0^{\tau^1} \Theta(s,b_0^1,\Pi_{\bar{D}^{2,d}}(X_s^2,...,X_s^d))ds\big)]\Big\}. \end{aligned}$$

Hence, since $(M_s)_{s\in[0,T]}$ is a martingale and recalling $\mathbb{E}[Z] = C_0 = \dfrac{\mathbb{E}[s^2_{\tau^+}]}{2\mathbb{E}[s_{\tau^+}]}$, we get

$$E_1 = \sqrt{h}C_0\mathbb{E}_x[\mathbb{I}_{\tau^1\le T, \wedge_{j=2}^m \tau^{j}>\tau^{1}}\partial_{x_1} v(\tau^1, b_0^1, X^2_{\tau^1},...,X^d_{\tau^1}))] + R + o(\sqrt{h}). \tag{3.2.16}$$

Control of the rests

For R_1 one gets

$$\begin{aligned} R_1 &= \mathbb{E}_x[\tilde{M}_T\partial_{x_1}v(T,b_0^1,\Pi_{\bar{D}^{2,d}}(X_T^2,...,X_T^d))] - \mathbb{E}_x[\int_0^T \tilde{M}_s d(\partial_{x_1}v(s,b_0^1,\Pi_{\bar{D}^{2,d}}(X_s^2,...,X_s^d)))] \\ &= \mathbb{E}_x[\int_0^T \partial_{x_1}v(s,b_0^1,\Pi_{\bar{D}^{2,d}}(X_s^2,...,X_s^d))d\tilde{M}_s] + \mathbb{E}_x[\langle\partial_{x_1}v(.,b_0^1,\Pi_{\bar{D}^{2,d}}(X_.^2,...,X_.^d)),\tilde{M}\rangle_T] \\ &= \mathbb{E}_x[\langle\partial_{x_1}v(.,b_0^1,\Pi_{\bar{D}^{2,d}}(X_.^2,...,X_.^d)),\tilde{M}\rangle_T] \\ &= \sum_{i=2}^m \rho_{1i}\mathbb{E}_x[\int_0^T ds\partial_{x_i}\partial_{x_1}v(s,b_0^1,\Pi_{\bar{D}^{2,d}}(X_s^2,...,X_s^d))\mathbb{I}_{s<\tau^N,X_s^1<b_0^1}]. \end{aligned}$$

We recall that under Assumption **(S)**,

$$\exists C>0,\ \forall\zeta,|\zeta|\le 4,\ \forall(s,x)\in[0,T]\times\bar{D},\ |\partial_x^\alpha v(s,x)|\le C \tag{3.2.17}$$

Hence, denoting $\tau_t^1 := \inf\{s\ge t : X_s^1 = b_0^1\}$ we get

$$\begin{aligned} |R_1|+|R_2| &\le C\int_0^T ds\mathbb{P}_x[\tau^N>s,X_s^1<b_0^1] + C\int_0^T ds\mathbb{E}_x[\mathbb{I}_{\tau^N>\phi(s)}\mathbb{I}_{\tau^1_{\phi(s)}\le s}\mathbb{E}[|X_s^1-b_0^1||\mathcal{F}_{\tau^1_{\phi(s)}}]] \\ &\le C\int_0^T ds\mathbb{P}_{x_1}[\tau^{N,1}>s,X_s^1<b_0^1] + Ch^{1/2}\int_0^T ds\mathbb{P}_{x_1}[\tau^{N,1}>\phi(s),\tau^1_{\phi(s)}\le s] \end{aligned}$$

where the last inequality is a consequence of Lemma 1.1.2. From Lemma 1.4.5 and its proof we derive that $R_1+R_2=O(h)$.

Owing to (3.2.17), we also have that $\Theta\circ\Pi_{\bar{D}}$ is uniformly Lipschitz continuous in space. We get

$$|R_3| \le Ch\left(\mathbb{E}_x[Z_N^2\mathbb{I}_{\tau^{N,1}\le T,\wedge_{j=2}^m\tau^{N,j}\ge\tau^{N,1}}] + \int_0^T ds\mathbb{E}_x[Z_N^2\mathbb{I}_{\tau^{N,1}\le s,\wedge_{j=1}^m\tau^{N,j}\ge\tau^{N,1}}]\right).$$

In the proof of Proposition 1.5.4, we showed $\exists c>0, C>0$, s.t. $\forall s\in[0,T]$, $\sup_N\mathbb{P}_x[Z_N\ge y,\ \tau^{N,1}\le s]\le C\exp(-cy^2)$. Hence, $|R_3|\le Ch$.

From the above controls on R_1,R_2,R_3 we get $R=O(h)$ which together with (3.2.16) gives

$$E_1 = \sqrt{h}C_0\mathbb{E}_x[\mathbb{I}_{\tau^1\le T,\wedge_{j=2}^m\tau^j>\tau^1}\partial_{x_1}v(\tau^1,X_{\tau^1})] + o(\sqrt{h})$$

for h small enough, which concludes the proof.

□

Remark 3.2.2 *The controls in the previous proof as well as the residual terms appearing in the computations are locally uniform w.r.t. the domain D.*

Remark 3.2.3 *In the above proof, for $m>1$, we strongly used that the derivatives up to order 4 of the function v were well controlled, the upper bounds being in that case given by Assumption* **(S)**. *One can wonder why we imposed such a regularity in* **(S)** *since the only quantity that appears eventually is the derivative of order one. Note that the correlation between the processes $X^1,X^2,...,X^m$ does not allow to separate the overshoot and the derivatives of v like in the proof with $m=1$. For the half space, this last fact is remarkable since it permits to get rid of the potentially degenerate derivatives of third order thanks to the second integration by part (see above proof). Thus, for the half space case all the previous computations hold true if we replace* **(S)** *by*

(S') *The function $f|_{\partial D}=0$.*

The function $v\in C^{2,4}([0,T)\times\bar{D})$ and there exist $C>0,c>0$, $\forall(t,y)\in[0,T]\times\partial D$, $\forall\xi$, $|\xi|\le 2$, $|\partial_y^\xi v(t,y)|\le C\exp(c|y|)$.

For a real intersection case, if v does not satisfy **(S)**, *we should introduce a mollified version of $\partial_{x_1} v$. Unfortunately, one can lose the uniformity of the residual terms w.r.t. the mollifying parameter from equation (3.2.14) to (3.2.15).*

Remark 3.2.4 *To conclude this section, we would like to emphasize that the main difficulty in order to apply the previous Theorem consists in finding conditions on f that guarantee* **(S)** *is fulfilled. We provide some sufficient conditions in Section 4 but in all generality this is far from being easy. Anyhow, this difficulty is essentially of analytical nature.*

3 Error Correction

3.1 Main result

Under our current assumptions **(BM)**, **(D)**, **(S)**, the next Theorem improves the accuracy of the numerical procedure by removing the term of order $\frac{1}{2}$ in the error.

For this, the simulation of $(X_{t_i})_{0\le i\le N}$ is performed in a modified domain, namely $D^h := \{y \in \mathbb{R}^d : \forall i \in [\![1,m]\!],\ y_i > b_0^i + C_0\sqrt{h}\}$ instead of $D := \{y \in \mathbb{R}^d : \forall i \in [\![1,m]\!],\ y_i > b_0^i\}$.

We denote $\tau^N_{D^h}$ (resp. τ_{D^h}) the discrete (resp. continuous) exit time from this domain D^h.

Theorem 3.3.1 *Assume* **(BM)**, **(D)**, **(S)**. *For h small enough we have*

$$\mathrm{Err}'(T,h,f,x) := \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau^N_{D^h}>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] = o(\sqrt{h}).$$

Remark 3.3.1 *Consider now the more general case $X_s = x + \mu s + \sigma W_s, D := \{x \in \mathbb{R}^d : A.x > b\}$ where $A = \begin{pmatrix} a_1^* \\ \vdots \\ a_m^* \end{pmatrix}$ is of rank m. If we have good smoothness properties for the function v_0 introduced in Proposition 3.1.1, then the former error expansion remains valid. We can as in the previous theorem remove the leading term of the error by simulating w.r.t. $D^h := \{x \in \mathbb{R}^d : A.x > b + C_0 e\sqrt{h}\},\ e = (\|\sigma^* a_1\|, ..., \|\sigma^* a_m\|)^*$.*

3.2 Proof of the Main Result

In this subsection we detail how the arguments from Costantini, El Karoui and Gobet, see [CKG03], can be employed to prove our correction result. We write

$$\begin{aligned}\mathrm{Err}'(T,h,f,x) &= \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau^N_{D^h}>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D^h}>T}] + \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D^h}>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]\\ &:= E_1 + E_2.\end{aligned}$$

From Remark 3.2.2 we derive that one could show just like in Theorem 3.2.4 that even though the domain depends on h we have $E_1 = C_1\sqrt{h} + o(\sqrt{h})$, where C_1 denotes the constant introduced in the cited theorem. For E_2 we adapt some ideas from [CKG03] concerning the sensitivity of the Dirichlet problem w.r.t. the domain.

For a given $c \in \mathbb{R}^d$, let us denote $\forall \eta > 0,\ D_\eta := \{y \in \mathbb{R}^d : y - \eta c \in D\}$. We define $\tau_{D_\eta} := \inf\{s > 0 : X_s \notin D_\eta\}$ and we introduce for all $x \in D$ the mapping $\mathcal{J}_c^x : \eta \longrightarrow \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D_\eta}>T}]$. We show below that under the assumptions of Theorem 3.3.1, the mapping $\mathcal{J}_c^x$ is differentiable in $\eta = 0$ and for $c = (\underbrace{1,...,1}_{m},\underbrace{0,...,0}_{d-m})$, one has

$$\partial_\eta \mathcal{J}_c^x(\eta)|_{\eta=0} = -\mathbb{E}_x[\nabla v(\tau, X_\tau).c\mathbb{I}_{\tau<T}] = -\sum_{i=1}^m \mathbb{E}_x[\partial_{x_i} v(\tau^i, X_{\tau^i})\mathbb{I}_{\tau^i\le T, \wedge_{j\in[\![1,m]\!]\setminus\{i\}}\tau^j>\tau^i}]. \tag{3.3.1}$$

From (3.3.1) we then derive that $E_2 = \mathcal{J}_c^x(C_0\sqrt{h}) - \mathcal{J}_c^x(0) = \partial_\eta \mathcal{J}_c^x(0)C_0\sqrt{h} + o(\sqrt{h}) = -C_1\sqrt{h} + o(\sqrt{h})$ which proves the Theorem.

□

Proof of (3.3.1)

Let us define $X_s^\eta := X_s - \eta c,\ \tau^{D,\eta} = \inf\{s > 0 : X_s^\eta \notin D\}$. Note that one has $\tau_{D_\eta} = \tau^{D,\eta}$. Denoting $\Delta_\eta := \mathbb{E}_x[f(X_T^\eta + \eta c)\mathbb{I}_{\tau^{D,\eta}>T}] - v(0,x)$, we have to identify the limit of Δ_η/η as $\eta \longrightarrow 0$. It comes

$$\begin{aligned}\Delta_\eta &= \mathbb{E}_x[f(X_T^\eta + \eta c)\mathbb{I}_{\tau^{D,\eta}>T}] - \mathbb{E}_x[f(X_T^\eta)\mathbb{I}_{\tau^{D,\eta}>T}] + \mathbb{E}_x[f(X_T^\eta)\mathbb{I}_{\tau^{D,\eta}>T} - v(T\wedge\tau^{D,\eta}, X_{T\wedge\tau^{D,\eta}})] \\ &\quad + \mathbb{E}_x[v(T\wedge\tau^{D,\eta}, X_{T\wedge\tau^{D,\eta}})] - v(0,x) := \Delta_{\eta,1} + \Delta_{\eta,2} + \Delta_{\eta,3}.\end{aligned}$$

Since $(M_t)_{t\in[0,T]} := (v(t\wedge\tau, X_{t\wedge\tau}^{0,x}))_{t\in[0,T]}$ is a martingale and $\tau^{D,\eta} < \tau$, we readily get $\Delta_{\eta,3} = 0$.

Note that $f(X_T^\eta)\mathbb{I}_{\tau^{D,\eta}>T} = v(T\wedge\tau^{D,\eta}, X_{T\wedge\tau^{D,\eta}}^\eta)$. One also has $\tau_{D_\eta} \underset{\eta\longrightarrow 0,\ \text{a.s}}{\longrightarrow} \tau$. From Assumption **(S)**, v is continuously differentiable. Thus, one gets $\lim_{\eta\longrightarrow 0}\Delta_{\eta,2}/\eta = -\mathbb{E}_x[\nabla v(T\wedge\tau, X_{T\wedge\tau}).c]$.

On the other hand, since we assumed f to be continuously differentiable we obtain $\lim_{\eta\longrightarrow 0}\Delta_{\eta,1}/\eta = \mathbb{E}_x[\nabla f(X_\tau).c\ \mathbb{I}_{\tau>T}]$. Recalling $\forall x\in\bar{D},\ v(T,x) = f(x)$ we write

$$\partial_\eta \mathcal{J}_c^x(\eta)|_{\eta=0} = -\mathbb{E}_x\left[\left(\nabla v(T\wedge\tau, X_{T\wedge\tau}) - \nabla f(X_T)\mathbb{I}_{\tau>T}\right).c\right] = -\mathbb{E}_x[\mathbb{I}_{\tau\le T}\nabla v(\tau, X_\tau).c]$$

which for $c = (\underbrace{1,...,1}_{m},\underbrace{0,...,0}_{d-m})$ proves (3.3.1).

□

3.3 Alternative proof in the half-space case

We give a more explicitly probabilistic proof of Theorem 3.3.1 when D is a half space, i.e. $m = 1$. Write as before

$$\mathrm{Err}'(T,h,f,x) = \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D^h}^N>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] = E_1 + E_2$$

where $D^h = \{y\in\mathbb{R}^d : y_1 > b_0 + \dfrac{\mathbb{E}_0[s_{\tau^+}^2]}{2\mathbb{E}_0[s_{\tau^+}]}\sqrt{h}\}$, $E_1 = \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D^h}^N>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D^h}>T}]$ and $E_2 = \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau_{D^h}>T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}]$.
As indicated in the previous proof, for h small enough

$$E_1 = \frac{\mathbb{E}_0[s_{\tau^+}^2]}{2\mathbb{E}_0[s_{\tau^+}]}\sqrt{h}\mathbb{E}_x[\mathbb{I}_{\tau\le T}(\partial_{y_1}v(\tau,X_\tau))] + o(\sqrt{h}).$$

For the estimation of E_2, note that it is enough to get

$$\partial_{b_0}v(0,x) = \partial_{b_0}\mathbb{E}_x[f(X_T)\mathbb{I}_{\tau>T}] = \mathbb{E}_x[\mathbb{I}_{\tau\le T}(-\partial_{y_1}v(\tau,X_\tau))].$$

To justify this equality, we exploit the explicit form of the killed transition densities for the linear BM (see [KS91] p.97-98). To simplify, put $X_t' = (X_t^2,\cdots,X_t^d)$, recall $g_t(z) = \frac{e^{-z^2/(2t)}}{\sqrt{2\pi t}}$ and define $a(x_1-b_0,t) = \frac{x_1-b_0}{t}g_t(x_1-b_0) = -g_t'(x_1-b_0)$ the density at time t of τ. Clearly, by independence of X^1 and X', one has $v(t,y) = \int_{b_0}^{+\infty}(g_{T-t}(z-y_1) - g_{T-t}(z+y_1-2b_0))\mathbb{E}[f(z,X_T')|X_t'^2 = y_2,\cdots,X_t'^d = y_d]dz$, from which it is easy to derive

$$\begin{aligned}-\partial_{y_1}v(t,(b_0,X_t')) &= -\int_{b_0}^{+\infty}2a(z-b_0,T-t)\mathbb{E}[f(z,X_T')|X_t']dz,\\ \partial_{b_0}v(0,x) &= -2\int_{b_0}^{+\infty}a(z+x_1-2b_0,T)\mathbb{E}_{x_2,\dots,x_d}[f(z,X_T')]dz,\\ \mathbb{E}_x[\mathbb{I}_{\tau\le T}(-\partial_{y_1}v(\tau,X_\tau))] &= \int_0^T a(x_1-b_0,t)\mathbb{E}_{x_2,\dots,x_d}[-\partial_{y_1}v(t,(b_0,X_t'))]dt,\\ &= -2\int_{b_0}^{+\infty}\mathbb{E}_{x_2,\dots,x_d}[f(z,X_T')]\Big(\int_0^T a(x_1-b_0,t)a(z-b_0,T-t)dt\Big)dz.\end{aligned}$$

Thanks to the Markov property on hitting times, see [KS91] p.197, the convolution integral w.r.t. t simply reduces to $a(z+x_1-2b_0,T)$. This proves our assertion.

□

4 Some sufficient conditions to fulfill (S)

In the whole section we assume **(BM)**, **(D)**.

4.1 Half-space

Main working assumption

We introduce Assumption

(F1) The function f belongs to $C_b^4(\bar{D})$, $f|_{\partial D} = 0$, $\mathrm{tr}(H_f\sigma_0\sigma_0^*)|_{\partial D} = 0$.

In the half space case $D := \{y \in \mathbb{R}^d : y_1 > b_0\}$. From the explicit law of the one dimensional BM killed at b_0, we derive $\forall (t,x) \in [0,T] \times D$

$$v(t,x) \quad := \quad \int_D f(y) q_{T-t}(x,y)dy = \int_{(b_0,+\infty)\times\mathbb{R}^{d-1}} f(y) q^1_{T-t}(x_1,y_1) \prod_{i=2}^d g_{T-t}(y_i - x_i) dy_1...dy_d \quad (3.4.1)$$

where $q_s^1(x_1,y_1) = g_s(y_1 - x_1) - g_s(y_1 + x_1 - 2b_0)$.

Under **(F1)**, we readily get **(S)** by simple differentiation after a clear change of variable in the above integral.

Alternative assumption in the half space case

In the special case of the half space, we now show that under **(BM)**, **(D)**, the following assumption on f is sufficient to satisfy **(S')** introduced in Remark 3.2.3.

(F1') The function f is borelian with support satisfying $d(\mathrm{supp}(f), \partial D) \geq 2\varepsilon > 0$ and has exponential growth, i.e. $\exists c > 0, C > 0,\ \forall x \in \mathbb{R}^d,\ |f(x)| \leq C\exp(c\|x\|)$.

From the explicit representation of the killed transition density $q_s(.,.)$ in (3.4.1) we derive by differentiation that $\exists c' > 0, C' > 0$ s.t. $\forall (s,x,y) \in (0,T] \times \bar{D} \times \bar{D},\ \forall \zeta,\ |\zeta| \leq 4$,

$$|\partial_x^\zeta q_s(x,y)| \leq \frac{C'}{s^{(d+|\zeta|)/2}} \exp\left(-c'\frac{\|x-y\|^2}{s}\right).$$

The function $v \in C^{2,4}([0,T) \times \bar{D})$. Owing to the support condition and to the exponential growth in Assumption **(F1')** we get from the above control that $\forall x \in \mathbb{R}^d, x_1 \in [b_0, b_0+\varepsilon), \forall t \in [0,T),\ \forall \zeta, |\zeta| \leq 4$,

$$|\partial_x^\zeta v(t,x)| \leq C' \int_D dy \exp\left(-c'\frac{\varepsilon^2}{2(T-t)}\right) \frac{\exp(c\|y\|)}{(T-t)^{(d+|\zeta|)/2}} \exp\left(-c'\frac{\|x-y\|^2}{2(T-t)}\right) \leq \frac{C\exp(c\|x\|)}{1 \wedge \varepsilon^{|\zeta|}}.$$

Thus, **(F1')** associated to **(BM)**, **(D)**, guarantees Assumption **(S')** introduced in Remark 3.2.3 is satisfied and the error expansion of Theorem 3.2.4 holds true. We could show as in Theorem 3.3.1 that the correction result remains valid under **(BM)**, **(D)**, **(S')**.

4.2 Bidimensional cone

In this subsection we provide sufficient conditions on f to fulfill **(S)** under **(BM)**, **(D)**, when the dimension $d = 2$.

We first give the explicit expression for the heat kernel in a bidimensional cone (up to a simple transformation this gives the expression of the density for the killed correlated BM in an orthant). In our two-dimensional context, this had been obtained by Carslaw and Jaeger extending the well known method of images, see [CJ59] and details below. For the sake of completeness, we present the slightly different proof of Iyengar, cf. [Iye85], justifying every step and correcting minor mistakes. During the proof we also state useful controls for the derivatives of the heat kernel.

Then, using some usual PDE results and the previous controls on the derivatives of the heat kernel we derive suitable assumptions on f to obtain **(S)**.

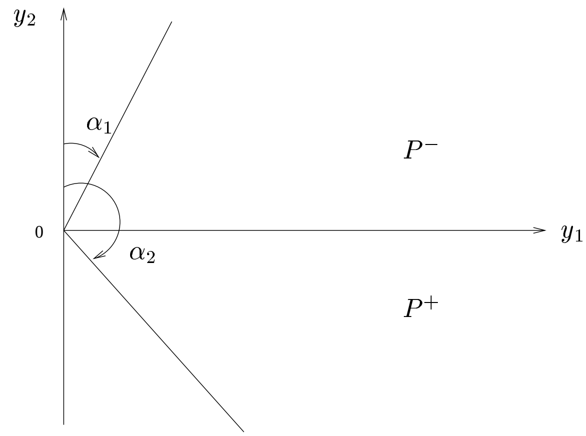


Figure 3.1: Connection between correlated BM in the positive quartant and standard BM in a wedge.

Explicit expression of the heat kernel in a bidimensional cone

In our current framework, assumption **(BM)** writes $X_s = x + \sigma_0 W_s$, where W is a standard BM, $\sigma_0\sigma_0^* = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \rho \in (-1,1)$. For the sake of simplicity we also assume in **(D)** w.l.o.g. that the origin of the cone $b_0 = 0$. We write $D := D_0 = \{y \in \mathbb{R}^2 : y_1 > 0, y_2 > 0\}$.

By simple transformations we derive that this density is linked to the one of the non-correlated Brownian Motion W killed when it leaves a wedge. Indeed, put $\forall i \in \{1,2\}$, $\tau^i_{D_0} := \inf\{s \geq 0 : X^i_s = 0\}$, one has $\tau_{D_0} = \tau^1_{D_0} \wedge \tau^2_{D_0}$. Using the Cholesky factorization of $\sigma_0\sigma_0^*$, i.e. $\sigma_0 = \begin{pmatrix} 1 & 0 \\ \rho & (1-\rho^2)^{1/2} \end{pmatrix}$, rewrite now $\tau^2_{D_0} := \inf\{s \geq 0 : x_2 + \sin(\beta)W^1_s + \cos(\beta)W^2_s = 0\}$ where $\rho = \sin(\beta), \beta \in (-\pi/2, \pi/2)$. One gets $\mathbb{P}_x[\tau_{D_0} > t] = \mathbb{P}_{\tilde{x}}[\tau_{\tilde{D}_0} > t]$ where $\tilde{x} = (x_2/\cos(\beta) - x_1\tan(\beta), x_1)$, $\tilde{D}_0 := \{y = (r\cos\theta, r\sin\theta) \in \mathbb{R}^2 : r > 0, 0 < \theta < \alpha\}$ with $\alpha = \frac{\pi}{2} + \arcsin(\rho) = \frac{\pi}{2} + \beta$, and $\tau_{\tilde{D}_0} := \inf\{s \geq: W_s \notin \tilde{D}_0\}$. Introducing the change of variable

$$z = Ay,\ A = \begin{pmatrix} -\tan(\beta) & \cos(\beta)^{-1} \\ 1 & 0 \end{pmatrix},\ \det(A^{-1}) = -\cos(\beta) \tag{3.4.2}$$

that maps D_0 onto $\tilde{D}_0$ we finally get

$$\mathbb{P}_x[\tau_{D_0} > t, X_t \in [y, y+dy)] = |\cos(\beta)^{-1}|\mathbb{P}_{\tilde{x}}[\tau_{\tilde{D}_0} > t, W_t \in [z, z+dz)]. \tag{3.4.3}$$

The previous transformation is illustrated in the Figure 3.1.

The wedges with respective angles α_1, α_2 in the above picture are associated to initial correlation coefficients $\rho_i = \sin(\alpha_i - \frac{\pi}{2}),\ i \in \{1,2\}$. If $\rho \in (-1,0)$, (resp. $\rho \in [0,1)$) $\{y \in \mathbb{R}^2 : y_1 > 0, y_2 = -\tan(\beta)y_1\} \subset P^+ := \{y \in \mathbb{R}^2 : y_1 > 0, y_2 \geq 0\}$ (resp. $P^- := \{y \in \mathbb{R}^2 : y_1 > 0, y_2 \leq 0\}$). This establishes a clear correspondence between the original correlation coefficient and the angle of the wedge. We inverted the coordinates in (3.4.3) in order to use a standard polar representation.

Method of Images

We present in this paragraph the so-called method of images that provides a simple expression for $\mathbb{P}_{\tilde{x}}[\tau_{\tilde{D}_0} > t, W_t \in dz]$ in terms of standard Gaussian kernels when the angle α from the above definition has the form $\alpha = \pi/m_0,\ m_0 \in \mathbb{N}^*$.

This method initially comes from physics, see Carslaw and Jaeger [CJ59], and consists in using successive reflections across the lines $x_2 = \tan(j\pi/m_0)x_1,\ j \in [\![0, 2m_0 - 1]\!]$ alternating heat sources and sinks along the orbit associated to a given starting point in the wedge.

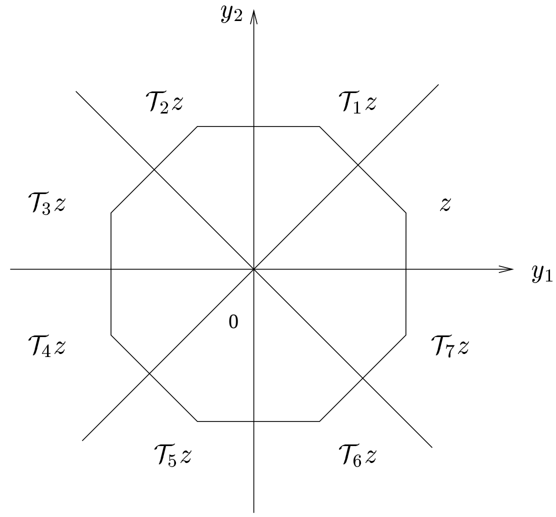


Figure 3.2: Orbit for $z \in \tilde{D}_0,\ \alpha = \frac{\pi}{4}$.

It provides a way to eliminate the difficulties associated to the stopping time $\tau_{\tilde{D}_0}$ in the determination of an explicit expression for the density.

Precisely, one has the following

Proposition 3.4.1 *Le W be a standard Brownian motion with starting point $x \in \tilde{D}_0 := \{y = (r\cos\theta, r\sin\theta) \in \mathbb{R}^2 : r > 0, 0 < \theta < \alpha\},\ \alpha = \pi/m_0, m_0 \in \mathbb{N}^*$. For all $t > 0$ one has*

$$\mathbb{P}_x[\tau_{\tilde{D}_0} > t, W_t \in dz] = \tilde{P}(t,x,z)dz = \sum_{j=0}^{2m_0-1} (-1)^j g_t(x - \mathcal{T}_j z)dz \tag{3.4.4}$$

with $\mathcal{T}_0 = I$ and $\forall j \in [\![1, 2m_0-1]\!],\ \mathcal{T}_j = R_j \mathcal{T}_{j-1}$ where the $(R_j)_{j\in[\![1,2m_0-1]\!]}$ are the matrices representing the reflection across the line $x_2 = \tan(j\pi/m_0)x_1$.

Remark 3.4.1 *From equation (3.4.4) we derive that in this special case we have the "usual" controls for the derivatives of $\tilde{P}$, that also corresponds to those of the Green kernel associated to problem (3.4.5) below. Namely, noting that $\forall (x,z) \in \tilde{D}_0 \times \tilde{D}_0,\ \forall j \in [\![0, 2m_0-1]\!],\ |x-z| \le |x - \mathcal{T}_j z|$, we get that for a given $T > 0$ and all multi-indices ζ, γ there exist constants $C := C(m_0, |\zeta|, |\gamma|, T) > 0,\ c := c(|\gamma|, |\zeta|) > 0$ s.t.*

$$\forall (t,x,z) \in (0,T] \times \tilde{D}_0 \times \tilde{D}_0,\ |\partial_t^\zeta \partial_x^\gamma \tilde{P}(t,x,z)| \le C \frac{\exp(-c\frac{|x-z|^2}{t})}{t^{1+|\zeta|+|\gamma|/2}}.$$

Proof.
Let h be a positive bounded continuous function defined on $\tilde{D}_0$ and vanishing on $\partial\tilde{D}_0$. Define for all $(t,x) \in [0,T] \times \bar{\tilde{D}}_0,\ w(t,x) = \int_{\tilde{D}_0} h(z)\mathbb{P}_x[\tau_{\tilde{D}_0} > T-t, W_{T-t} \in dz]$. The function w satisfies the mixed Cauchy-Dirichlet problem

$$\begin{cases} (\partial_t + \frac{1}{2}\Delta)w(t,x) = 0,\ (t,x) \in [0,T) \times \tilde{D}_0, \\ w(t,x)|_{\partial\tilde{D}_0} = 0,\ t \in [0,T], \\ w(T,x) = h(x),\ x \in \bar{\tilde{D}}_0. \end{cases} \tag{3.4.5}$$

With the above notations define $\forall y \in \mathbb{R}^2,\ \tilde{h}(y) = \sum_{j=0}^{2m_0-1} (-1)^j h(\mathcal{T}_j^{-1} y)\mathbb{I}_{y\in\mathcal{T}_j(\tilde{D}_0)}$ and introduce the Cauchy problem

$$\begin{cases} (\partial_t + \frac{1}{2}\Delta)u(t,x) = 0, \ (t,x) \in [0,T) \times \mathbb{R}^2, \\ u(T,x) = \tilde{h}(x), \ x \in \mathbb{R}^2. \end{cases} \tag{3.4.6}$$

From the standard Feynman-Kac representation we derive that the solution of (3.4.6) writes

$$u(t,x) = \mathbb{E}_x[\tilde{h}(W_{T-t})] = \int_{\mathbb{R}^2} dy \tilde{h}(y) g_{T-t}(y-x).$$

Computing for each $j \in [\![0, 2m_0 - 1]\!]$ the change of variable $z = \mathcal{T}_j^{-1} y$ for $y \in \mathcal{T}_j(\tilde{D}_0)$ it comes

$$u(t,x) = \int_{\tilde{D}_0} dz h(z) \sum_{j=0}^{2m_0-1} (-1)^j g_{T-t}(x - \mathcal{T}_j z).$$

Put $\forall t > 0, \ \tilde{P}(t,x,z) := \sum_{j=0}^{2m_0-1} (-1)^j g_t(x - \mathcal{T}_j z)$. By construction, since we have Gaussian kernels, $(\partial_t - \frac{1}{2}\Delta_x)\tilde{P}(t,x,z) = 0$. To identify $\tilde{P}(T-t,x,z)dz$ with $\mathbb{P}_x[\tau_{\tilde{D}_0} > T-t, W_{T-t} \in dz]$ for $x \in \tilde{D}_0$, it therefore remains to check that for all $t > 0$, $\tilde{P}(t,x,z)$ satisfies the boundary conditions w.r.t. $x \in \partial\tilde{D}_0$.

- For $x \in \{y \in \mathbb{R}^2 : y_2 = 0\}$, by construction $\forall z \in \tilde{D}_0, \ \forall j \in [\![0, 2m_0 - 1]\!], \ |x - \mathcal{T}_j z| = |x - \mathcal{T}_{2m_0-1-j} z|$. Thus, $\tilde{P}(t,x,z) = 0$ since $(-1)^j + (-1)^{2m_0-1-j} = 0$.
- For $x \in \{y \in \mathbb{R}^2 : y_2 = y_1 \tan\alpha\}$, we get $\forall z \in \tilde{D}_0, \ \forall j \in [\![0, 2m_0 - 1]\!], \ |x - \mathcal{T}_j z| = |x - \mathcal{T}_{2m_0+1-j} z|$. Thus, $\tilde{P}(t,x,z) = 0$ since $(-1)^j + (-1)^{2m_0+1-j} = 0$.

Observe $\partial\tilde{D}_0 = \{y \in \mathbb{R}^2 : y_2 = 0\} \cup \{y \in \mathbb{R}^2 : y_2 = y_1 \tan\alpha\}$, the proof is complete.

□

Remark 3.4.2 *From equation (3.4.4) in Proposition 3.4.1 we derive by simple differentiation that for angles $\alpha = \pi/m_0, \ m_0 \in \mathbb{N}^*$, if $h \in C_b^4(\bar{\tilde{D}}_0), \ h|_{\tilde{D}_0} = \Delta h|_{\tilde{D}_0} = 0$ then the solution w of problem (3.4.5) belongs to $C_b^{2,4}([0,T] \times \bar{\tilde{D}}_0)$.*

Thus, using the change of variable given in (3.4.2), we derive that under **(BM)**, **(D)**, *the solution v of problem (3.1.2), with $\sigma_0\sigma_0^* = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $\rho = \sin(\alpha - \frac{\pi}{2})$, $D_0 = \{x \in \mathbb{R}^2 : x_1 > b_0^1, x_2 > b_0^2\}$, $f(x) = h(A(x - b_0))$ satisfies Assumption* **(S)**.

Extension to the general case

In this paragraph we use Proposition 3.4.1 to derive an explicit expression for $\mathbb{P}_x[\tau_{\tilde{D}_0} > t, W_t \in dz]$ for general angles $\alpha \in (0, \pi)$.

Proposition 3.4.2 *Le W be a standard Brownian motion with starting point $x \in \tilde{D}_0 := \{y = (r\cos\theta, r\sin\theta) \in \mathbb{R}^2 : r > 0, 0 < \theta < \alpha\}$, $\alpha \in (0,\pi)$. For $t > 0$ one has*

$$\mathbb{P}_x[\tau_{\tilde{D}_0} > t, W_t \in dz] = \tilde{P}(t,x,z)dz = \frac{2}{t\alpha}\exp(-\frac{r^2 + r_0^2}{2t})\sum_{n=0}^{\infty} \sin(\nu_n\theta)\sin(\nu_n\theta_0) I_{\nu_n}\left(\frac{rr_0}{t}\right) r_0 dr_0 d\theta_0 \tag{3.4.7}$$

where $x = (r\cos\theta, r\sin\theta)$, $z = (r_0\cos\theta_0, r_0\sin\theta_0), \forall n \in \mathbb{N}$, $\nu_n = \frac{\pi n}{\alpha}$ and I_ν denotes the modified Bessel function of order ν.

Remark 3.4.3 *For some other explicit expressions associated to the killed/stopped bidimensional Brownian Motion in a wedge, such as* $\mathbb{P}_x[\tau_{\tilde{D}_0} \in dt, W_{\tau_{\tilde{D}_0}} \in da]$, $\mathbb{P}_x[\tau^1_{\tilde{D}_0} \in ds, \tau^2_{\tilde{D}_0} \in dt]$, *we refer to the article of Iyengar, [Iye85]. Concerning the expression of* $\mathbb{P}_x[\tau_{\tilde{D}_0} > t]$ *that one obtains by simple integration of (3.4.7), we correct the statement of equation (10) in Iyengar in Section 5.4.*

We also mention the paper of Bañuelos and Smits, [BS97], who extended the idea used for the proof of Proposition 3.4.2 to obtain an explicit expression for the heat kernel in generalized cones of $\mathbb{R}^d$. *In their context a general cone with origin 0 generated by a proper open connected subset* G *of* S^{n-1}, *where* S^{n-1} *is the unit sphere in* $\mathbb{R}^n$, *is the set of all rays emanating from 0 and passing through* G.

Proof. We adapt the proof proposed by Iyengar, [Iye85], specifying at some steps useful controls for the derivatives of the heat kernel in wedges with general angles.

Let us start from (3.4.4) in the special case $\alpha = \pi/m_0,\ m_0 \in \mathbb{N}^*$. Denote $(\theta_j)_{j\in[\![0,2m_0-1]\!]}$ the argument of $\mathcal{T}_j z$. Writing (3.4.4) in polar coordinates it comes

$$\tilde{P}(t,x,z) = \frac{1}{2\pi t}\exp(-\frac{r^2+r_0^2}{2t})\sum_{j=0}^{2m_0-1}(-1)^j \exp(\frac{rr_0\cos(\theta-\theta_j)}{t}).$$

From identity 9.6.34 in Abramowitz and Stegun [AS72] we also have $\forall z \in \mathbb{R}$

$$\exp(z\cos(\theta-\theta_j)) = I_0(z) + 2\sum_{n=1}^{\infty} I_n(z)\cos(n(\theta-\theta_j)) \tag{3.4.8}$$

where I_n is the modified Bessel function of order n. Note that equation (6) from [Iye85] concerning the expansion of the exponential in series of modified Bessel functions is incorrect, anyhow his following computations are not affected.

Observing that

$$\sum_{j=0}^{2m_0-1}(-1)^j\cos(n(\theta-\theta_j)) = 2m_0\sin(n\theta)\sin(n\theta_0)$$

if m_0 divides n and 0 otherwise, we finally rewrite

$$\tilde{P}(t,x,z) = \frac{2}{t\alpha}\exp(-\frac{r^2+r_0^2}{2t})\sum_{n=0}^{\infty}\sin\left(\frac{n\pi\theta}{\alpha}\right)\sin\left(\frac{n\pi\theta_0}{\alpha}\right)I_{n\pi/\alpha}\left(\frac{rr_0}{t}\right) \tag{3.4.9}$$

where $\alpha = \pi/m_0,\ m_0 \in \mathbb{N}^*$. Anyhow, identity (3.4.9) is meaningful for all $\alpha \in (0,\pi)$, and we claim that for a bounded continuous function h vanishing on $\partial\tilde{D}_0$, $w(t,x) = \int_{\tilde{D}_0} dzh(z)\tilde{P}(T-t,x,z),\ (t,x) \in [0,T]\times\bar{\tilde{D}}_0$, satisfies (3.4.5). The boundary conditions are readily fulfilled. It thus remains to show $(\partial_t + \frac{1}{2}\Delta)w(t,x) = 0$.

To this end we introduce a technical Lemma whose proof is postponed to the end of the section.

Lemma 3.4.3 *For all* $R>0, T>0$ *and all multi-indices* $\alpha,\zeta, |\alpha|\le 2,\ |\zeta|\le 5$ *there exist positive constants* $C := C(R,T,2,5), c := c(2,5), \xi := \xi(2,5)$ *s.t.* $\forall(t,x,z)\in(0,T]\times(\tilde{D}_0\cap B(0,R))\times\tilde{D}_0$, $x=(r\cos\theta, r\sin\theta), z = (r_0\cos\theta_0, r_0\sin\theta_0)$ *one has*

$$\exp\left(-\frac{r^2+r_0^2}{2t}\right)\sum_{n=0}^{\infty}\left|(\partial_t^\alpha+\partial_x^\zeta)\left\{\sin\left(\frac{n\pi\theta}{\alpha}\right)\sin\left(\frac{n\pi\theta_0}{\alpha}\right)I_{n\pi/\alpha}\left(\frac{rr_0}{t}\right)\right\}\right| \le \frac{C}{t^\xi}\exp\left(-c\frac{|r-r_0|^2}{t}\right).$$

Lemma 3.4.3 guarantees we can differentiate under the sum in (3.4.9). Using the recurrence relations $\frac{d}{dz}I_\nu(z) = (I_{\nu-1}+I_{\nu+1})(z)/2,\ \nu I_\nu(z) = \frac{z}{2}(I_{\nu-1}-I_{\nu+1})(z)$ for modified Bessel functions, see 9.6.26 in [AS72], we obtain

$$
\begin{aligned}
\partial_t \tilde{P}(t,x,z) &= \tilde{P}(t,x,z)\left(-\frac{1}{t}+\frac{r^2+r_0^2}{2t^2}\right)\\
&\quad -\frac{rr_0}{t^3\alpha}\exp\left(-\frac{r^2+r_0^2}{2t}\right)\sum_{n=0}^{\infty}\sin(\nu_n\theta)\sin(\nu_n\theta_0)\left(I_{\nu_n-1}+I_{\nu_n+1}\right)\left(\frac{rr_0}{t}\right),\\
\partial_{x_1}\tilde{P}(t,x,z) &= -\frac{x_1}{t}\tilde{P}(t,x,z)+\frac{r_0}{t^2\alpha}\exp\left(-\frac{r^2+r_0^2}{2t}\right)\sum_{n=0}^{\infty}\sin(\nu_n\theta_0)\Big\{\sin((\nu_n-1)\theta)I_{\nu_n-1}\left(\frac{rr_0}{t}\right)\\
&\quad +\sin((\nu_n+1)\theta)I_{\nu_n+1}\left(\frac{rr_0}{t}\right)\Big\} := -\frac{x_1}{t}\tilde{P}(t,x,z)+\tilde{F}_1(r,\theta,z),\\
\partial_{x_2}\tilde{P}(t,x,z) &= -\frac{x_2}{t}\tilde{P}(t,x,z)+\frac{r_0}{t^2\alpha}\exp\left(-\frac{r^2+r_0^2}{2t}\right)\sum_{n=0}^{\infty}\sin(\nu_n\theta_0)\Big\{\cos((\nu_n-1)\theta)I_{\nu_n-1}\left(\frac{rr_0}{t}\right)\\
&\quad -\cos((\nu_n+1)\theta)I_{\nu_n+1}\left(\frac{rr_0}{t}\right)\Big\} := -\frac{x_2}{t}\tilde{P}(t,x,z)+\tilde{F}_2(r,\theta,z),\\
\partial^2_{x_1}\tilde{P}(t,x,z) &= -\frac{\tilde{P}(t,x,z)}{t}-\frac{x_1(\partial_{x_1}\tilde{P}(t,x,z)+\tilde{F}_1(r,\theta,z))}{t}\\
&\quad +\frac{r_0^2}{2t^3\alpha}\exp\left(-\frac{r_0^2+r^2}{2t}\right)\sum_{n\geq 1}\sin(\nu_n\theta_0)\Big\{-\sin((\nu_n-2)\theta)I_{\nu_n-2}\left(\frac{rr_0}{t}\right)\\
&\quad +2\sin(\nu_n\theta)I_{\nu_n}\left(\frac{rr_0}{t}\right)+\sin((\nu_n+2)\theta)I_{\nu_n+2}\left(\frac{rr_0}{t}\right)\Big\},\\
\partial^2_{x_2}\tilde{P}(t,x,z) &= -\frac{\tilde{P}(t,x,z)}{t}-\frac{x_2(\partial_{x_2}\tilde{P}(t,x,z)+\tilde{F}_2(r,\theta,z))}{t}\\
&\quad +\frac{r_0^2}{2t^3\alpha}\exp\left(-\frac{r_0^2+r^2}{2t}\right)\sum_{n\geq 1}\sin(\nu_n\theta_0)\Big\{\sin((\nu_n-2)\theta)I_{\nu_n-2}\left(\frac{rr_0}{t}\right)\\
&\quad +2\sin(\nu_n\theta)I_{\nu_n}\left(\frac{rr_0}{t}\right)-\sin((\nu_n+2)\theta)I_{\nu_n+2}\left(\frac{rr_0}{t}\right)\Big\}
\end{aligned}
$$

From the above expressions, one checks $\forall (t,x,z)\in\mathbb{R}^{+,*}\times\tilde{D}_0^2,\ (\partial_t-\frac{1}{2}\Delta_x)\tilde{P}(t,x,z)=0$. This completes the proof.

□

Remark 3.4.4 *Note that Lemma 3.4.3 also gives a radial control on $\tilde{P}$ and its derivatives. In particular, for given $R>0,T>0$, there exist positive constants c,C s.t. for $x=(r\cos\theta,r\sin\theta)\in(\tilde{D}_0\cap B(0,R)), z=(r_0\cos\theta_0,r_0\sin\theta_0)\in\tilde{D}_0$, with $|r-r_0|\geq\varepsilon>0$ we have*

$$
\forall\zeta,|\zeta|\leq 5,\ \text{and}\ \forall s\in(0,T],\ |\partial_x^\zeta\tilde{P}(s,x,z)|\leq\frac{C}{1\wedge\varepsilon^\xi}\exp(-c\frac{\varepsilon^2}{s})\exp(-c\frac{|r-r_0|^2}{s}). \tag{3.4.10}
$$

This kind of estimates will be crucial in the following.

Proof of Lemma 3.4.3.

The key tool in the proof of the Lemma is the following identity

$$
\forall x>0,\ \forall\mu>\nu\geq 0,\ I_\mu(x)<I_\nu(x). \tag{3.4.11}
$$

Relation (3.4.11) was proved by Jones in [Jon68]. The arguments are rather simple and in order to be self contained we recall them at the end of our proof.

From (3.4.8) we get that $\exp(z)=I_0(z)+2\sum_{n=1}^{\infty}I_n(z)$. Hence, (3.4.11) yields

$$
\tilde{P}(t,x,z)\leq\frac{\exp(-\frac{r^2+r_0^2}{2t})}{t\alpha}\sum_{n=1}^{\infty}I_{\nu_n}\left(\frac{rr_0}{t}\right)\leq\frac{\exp(-\frac{r^2+r_0^2}{2t})}{t\alpha}\sum_{n=1}^{\infty}I_n\left(\frac{rr_0}{t}\right)\leq\frac{\exp(-\frac{|r-r_0|^2}{2t})}{2t\alpha}.
$$

Put $A_t := \sum_{n=1}^{\infty} \partial_t (\sin(\nu_n\theta)\sin(\nu_n\theta_0) I_{\nu_n}\left(\frac{rr_0}{t}\right))$. One gets

$$|A_t| \le \frac{rr_0}{2t^2} \sum_{n=1}^{\infty} (I_{\nu_n+1} + I_{\nu_n-1})\left(\frac{rr_0}{t}\right) \le C\frac{rr_0}{t^2}\exp\left(\frac{rr_0}{t}\right).$$

Thus,

$$\begin{aligned}\exp\left(-\frac{r^2+r_0^2}{2t}\right)|A_t| &\le C\frac{rr_0}{t^2}\exp\left(-\frac{|r-r_0|^2}{2t}\right) \le C(\frac{R^2}{t^2}+\frac{R|r-r_0|}{t^2})\exp\left(-c\frac{|r-r_0|^2}{2t}\right)\\ &\le (\frac{R^2}{t^2}+\frac{R}{t^{3/2}})\exp\left(-c\frac{|r-r_0|^2}{t}\right) \le \frac{C}{t^{\xi}}\exp\left(-c\frac{|r-r_0|^2}{t}\right)\end{aligned}$$

which gives the result for the time derivative. The upper bounds for the second time derivative and

$$\exp\left(-\frac{r^2+r_0^2}{2t}\right)\sum_{n=1}^{\infty}\left|\partial_x^{\alpha}(\sin(\nu_n\theta)\sin(\nu_n\theta_0)I_{\nu_n}\left(\frac{rr_0}{t}\right))\right|,\ |\alpha| \le 5$$

could be derived in a similar way.

$\square$

Proof of (3.4.11).

Turning to the definition, the modified Bessel function $I_\nu, \nu \ge 0$ is a solution, for all $x > 0$, of the following differential equation:

$$\frac{d}{dx}\left(xI_\nu'(x)\right) - \left(x + \frac{\nu^2}{x}\right)I_\nu(x) = 0. \tag{3.4.12}$$

For $\nu \notin \mathbb{N}$, the function $K_\nu(x) = \frac{\pi}{2}\frac{I_{-\nu}(x)-I_\nu(x)}{\sin(\nu\pi)}$ is also solution of (3.4.12).

From (3.4.12) one easily derives that $\forall(\mu,\nu) \in (\mathbb{R}^+)^2, \forall x > 0$

$$I_\nu(x)\frac{d}{dx}\left(xI_\mu'(x)\right) - I_\mu(x)\frac{d}{dx}\left(xI_\nu'(x)\right) = (\mu^2-\nu^2)\frac{I_\nu(x)I_\mu(x)}{x}.$$

Integrating the above equation we get

$$x[I_\nu(x)I_\mu'(x) - I_\nu'(x)I_\mu(x)] = (\mu^2-\nu^2)\int_0^x I_\mu(y)I_\nu(y)y^{-1}dy. \tag{3.4.13}$$

Define $\forall x > 0$, $G(x) = I_\mu(x)/I_\nu(x)$. For $\mu > \nu \ge 0$, equation (3.4.13) gives $G'(x) > 0$. Recalling the asymptotics for modified Bessel functions, i.e. $I_\nu(x) \underset{x\to 0}{\sim} (x/2)^\nu/\Gamma(1+\nu)$ and $I_\nu(x) \underset{x\to+\infty}{\sim} \exp(x)/\sqrt{2\pi x}$, see 9.6.7 and 9.7.1 in [AS72] for details, we also get

$$\forall\mu > \nu \ge 0,\ \lim_{x\to 0} G(x) = 0,\ \lim_{x\to+\infty} G(x) = 1.$$

Since in this case $G'(x) > 0$ for all $x > 0$ we have $G(x) < 1$ which completes the proof of (3.4.11).

$\square$

Some smoothness results for the underlying Dirichlet problem

We have already indicated, see Remark 3.4.2, that when $\alpha = \pi/m_0, m_0 \in \mathbb{N}^*$, under suitable assumptions on the final condition h one has the "usual" smoothness properties for the solution w of problem (3.4.5). We then derived that for some correlation coefficients, namely those of the form $\rho_{m_0} = \sin(\pi/m_0 - \pi/2)$, $m_0 \in \mathbb{N}^*$, and for a function $f \in C_b^{2,4}(\bar{D})$ satisfying some compatibility conditions, Assumption **(S)** was satisfied by the solution v of problem (3.1.2) that we now recall.

$$\begin{cases} (\partial_t v + \frac{1}{2}\mathrm{tr}(H_v \sigma_0\sigma_0^*))(t,x) = 0, \ [0,T) \times D, \\ v(t,.)|_{\partial D} = 0, \ t \in [0,T], \ v(T,x) = f(x), \ \forall x \in \bar{D}, \end{cases}$$

where $\sigma_0\sigma_0^* = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $\rho \in (-1,1)$, $D := \{x \in \mathbb{R}^2 : x_i > b_0^i\}$.

We now give a smoothness result for the solution v of the above system in dimension 2. We introduce Assumption

(F2) The function $f \in C_b^5(D)$, $f|_{\partial D} = \mathrm{tr}(H_f\sigma_0\sigma_0^*)|_{\partial D} = \mathrm{tr}(\sigma_0\sigma_0^* H_{\mathrm{tr}(H_f\sigma_0\sigma_0^*)})|_{\partial D} = 0$ and

$$d(\mathrm{supp}(f), b_0) \geq 2\varepsilon > 0.$$

Proposition 3.4.4 *Assume* **(F2)**. *For* $D := \{x \in \mathbb{R}^2 : x_1 > b_0^1, x_2 > b_0^2\}, \sigma_0\sigma_0^* = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $\rho \in (-1,1)$, *the unique solution* v *of (3.1.2) belongs to* $C_b^{2,4}([0,T] \times \bar{D})$. *In particular* $\exists C > 0, \ \forall (t,x) \in [0,T] \times \bar{D}, \ \forall \zeta, \ |\zeta| \leq 4, \ |\partial_x^\zeta v(t,x)| \leq C$.

Proof. From Proposition 3.4.2 we derive that problem (3.1.2) has a unique solution $v \in C^{1,2}([0,T) \times D) \cap C_b^0([0,T] \times \bar{D})$ and $\forall (t,x) \in [0,T] \times \bar{D}, \ v(t,x) = \mathbb{E}[f(X_{T\wedge\tau^{t,x}}^{t,x})], \ \forall s \geq t, \ X_s^{t,x} = x + \sigma_0(W_s - W_t), \ \tau^{t,x} := \inf\{s \geq t : X_s^{t,x} \notin D\}$.

Let us now note that as a consequence of the support condition in **(F2)** and the radial control of equation (3.4.10) there exist positive constants C, c, ξ s.t.

$$\forall (t,x) \in [0,T) \times (B(b_0,\varepsilon) \cap \bar{D}), \ \forall \zeta, |\zeta| \leq 5, \ |\partial_x^\zeta v(t,x)| \leq \frac{C}{1 \wedge \varepsilon^\xi} \exp\left(-c\frac{\varepsilon^2}{T-t}\right). \tag{3.4.14}$$

Choose now $\tilde{D}$ to be a C^5 domain[1] s.t. $d(\bar{\tilde{D}}, b_0) \geq \varepsilon/3 > 0$ and $\{x \in \mathbb{R}^2 : |x - b_0| \geq \varepsilon, x \in \partial D\} = \{x \in \mathbb{R}^2 : |x - b_0| \geq \varepsilon, x \in \partial\tilde{D}\}$. We point out that $\tilde{D}$ is an infinite domain (see Figure 3.3 below).

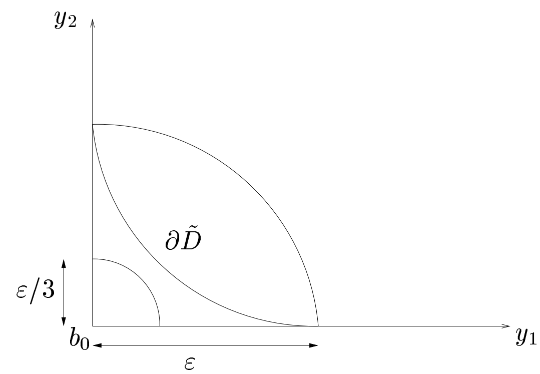


Figure 3.3: Covering strategy.

It therefore remains to show that under **(F2)**, $v \in C_b^{2,4}([0,T] \times \bar{\tilde{D}})$. Consider now the parabolic problem

$$\begin{cases} (\partial_t u + \frac{1}{2}\mathrm{tr}(H_u\sigma_0\sigma_0^*))(t,x) = 0, \ (t,x) \in [0,T) \times \tilde{D}, \\ u(t,x)|_{\partial\tilde{D}} = v(t,x), \ t \in [0,T), \ u(T,x) = f(x), \ x \in \bar{\tilde{D}}. \end{cases} \tag{3.4.15}$$

From (3.4.14) we derive that there exists $\tilde{v} \in C_b^{2+\alpha/2,4+\alpha}([0,T] \times \bar{\tilde{D}})$, $\alpha \in (0,1)$ s.t. $\tilde{v}|_{\partial\tilde{D}} = v$. Hence, Theorem 5.2 Chapter 4 from [LSU68] yields that there exists a unique solution u_1 to (3.4.15) and $u_1 \in$

[1] We mean by this that the diffeomorphism ψ from Definition 1.1.1 belongs to $C^5(\mathbb{R}^d)$.

$C_b^{2,4}([0,T] \times \tilde{\bar{D}})$. Denote $\tau_{\tilde{D}}^{t,x} := \inf\{s \geq t : X_s^{t,x} \notin \tilde{D}\}$. From the Feynman-Kac representation and the Markov property it then follows $u_1(t,x) = \mathbb{E}[\tilde{v}(T \wedge \tau_{\tilde{D}}^{t,x}, X_{T\wedge\tau_{\tilde{D}}^{t,x}}^{t,x})] = \mathbb{E}[f(X_{T\wedge\tau}^{t,x})] = v(t,x)$.

This completes the proof.

□

Hence, Assumption **(F2)** is sufficient to obtain **(S)** .

5 Numerical Results

In this section we provide some numerical tests and compare our correction method from Theorem 3.3.1 with the usual Romberg correction that we briefly recall.

From Theorem 3.3.1, we have that the discretization error writes $\mathrm{Err}(T,h,f,x) = C_1\sqrt{h} + o(\sqrt{h})$ for h small enough. Thus, $\mathrm{Err}(T,h/2,f,x) = C_1\frac{\sqrt{h}}{\sqrt{2}} + o(\sqrt{h})$. It comes

$$\mathbb{E}[f(X_T)\big((\sqrt{2}\mathbb{I}_{\tau^{2N}>T} - \mathbb{I}_{\tau^N>T}) - (\sqrt{2}-1)\mathbb{I}_{\tau>T}\big)] = o(\sqrt{h}),$$
$$\frac{1}{\sqrt{2}-1}\mathbb{E}[f(X_T)(\sqrt{2}\mathbb{I}_{\tau^{2N}>T} - \mathbb{I}_{\tau^N>T})] - \mathbb{E}[f(X_T)\mathbb{I}_{\tau>T}] = o(\sqrt{h}).$$

The above relation therefore suggests to modify the Monte-Carlo simulation in the following way. For a given $N \in \mathbb{N}^*, N_{MC} \in \mathbb{N}^*$ we define

$$\tilde{E}_{N_{MC}}(f,N) := \frac{1}{\sqrt{2}-1}\left(\sqrt{2}(\sum_{i=1}^{N_{MC}} f((X_T)^i)\mathbb{I}_{(\tau^{2N})^i>T}) - (\sum_{i=1}^{N_{MC}} f((X_T)^i)\mathbb{I}_{(\tau^N)^i>T})\right).$$

We point out that our correction is numerically less expensive than the Romberg procedure that requires to refine the time step.

5.1 Half-Space case

We take an example from financial applications. Consider a two-dimensional risky asset following the Black-Scholes-Merton dynamics, $S_t^1 = S_0^1\exp(\sigma_1 W_t^1 + (r - \frac{\sigma_1^2}{2})t), S_t^2 = S_0^2\exp(\sigma_2\rho W_t^1 + \sigma_2\sqrt{1-\rho^2}W_t^2 + (r-\frac{\sigma_2^2}{2})t)$, where $W = (W^1, W^2)$ is a standard two dimensional BM. For a fixed final time T, given level B and strike K, put $D := \{s \in \mathbb{R}^2 : s_1 > B\}$, we are interested in computing the quantity $\mathbb{E}[e^{-rT}\mathbb{I}_{\tau>T}\mathbb{I}_{(S_T^1\wedge S_T^2)\geq K}]$ associated to the price of a digital barrier option. Using the transformation of Proposition 3.1.1 we derive that assumption **(F1')** introduced in Section 4.1 is satisfied as soon as $K > B$. Since **(F1')** guarantees **(S')**, cf. Remark 3.2.3, the correction result of Theorem 3.3.1 holds true. For $r = .04, \sigma_1 = \sigma_2 = .3, \rho = .5, S_0^1 = S_0^2 = K = 100, B = 90, T = 1$ we compute the standard Monte-Carlo approximation, the Romberg extrapolation and the correction of Theorem 3.3.1 with 10^6 paths. The reference value has been computed with the usual Brownian bridge techniques for 10^8 paths. These techniques have been introduced by Lerche and Siegmund [LS89] in the case of BM in two-dimensional smooth domains, and later generalized in arbitrary dimensions by Baldi [Bal95]. Some numerical studies with financial applications have been developed by Andersen and Brotherton-Ratcliffe [ABR96], Beaglehole, Dybvig and Zhou [BDZ97], Baldi, Caramellino and Iovino [BCI99] among others.

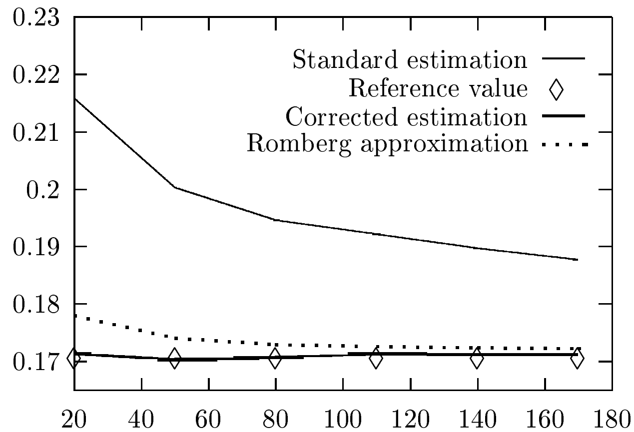
For the empirical mean we obtain

	MC	MC Shift	Romberg	Reference
$N = 20$	0.2159	0.1713	0.1779	0.1709
$N = 50$	0.2003	0.1702	0.1740	0.1709
$N = 80$	0.1946	0.1707	0.1729	0.1709
$N = 110$	0.1922	0.1713	0.1726	0.1709
$N = 140$	0.1897	0.1711	0.1723	0.1709
$N = 170$	0.1877	0.1712	0.1722	0.1709

For the usual unbiased estimator of the variance, we get

	MC	MC Shift	Romberg	Reference
$N = 20$	.174189	.146521	.285529	.07236
$N = 50$	.165018	.145789	.241693	.07236
$N = 80$	.161527	.146099	.224704	.07236
$N = 110$	.16006	.1465	.215754	.07236
$N = 140$	.158467	.146372	.208061	.07236
$N = 170$	.157189	.146422	.202833	.07236

The width of the 95%-confidence interval is essentially equal to 1.5 10^{-3}.



Convergence results w.r.t. the number of steps N.

Note first the positive bias for the standard procedure as proved in chapter 1. What also appears is that the shifting boundary correction is more accurate than the Romberg extrapolation: it is promising since the computational time is also lower. We finally point out that the empirical variance associated to our correction procedure is lower than the one of the Romberg extrapolation.

5.2 Cone case

We keep a two-dimensional risky asset following the Black-Scholes-Merton dynamics, $S_t^1 = S_0^1 \exp(\sigma_1 W_t^1 + (r - \frac{\sigma_1^2}{2})t)$, $S_t^2 = S_0^2 \exp(\sigma_2(\rho W_t^1 + (1-\rho^2)^{1/2} W_t^2) + (r - \frac{\sigma_2^2}{2})t)$, where W is a standard bidimensional BM. For a fixed final time T, a given strike K and threshold B, put $D := \{(s_1, s_2) \in \mathbb{R}^2 : s_1 > B, s_2 > B\}$, we are interested in computing $\mathbb{E}[e^{-rT}\mathbb{I}_{\tau>T} h(S_T)]$, where h is a smooth approximation of the indicator function that one expects in the case of a digital barrier option. We take $h(s) := \mathbb{I}^\star_{K,\varepsilon}(s_1)\mathbb{I}^\star_{K,\varepsilon}(s_2)$ with $\mathbb{I}^\star_{K,\varepsilon}(s_1) = 0$, if $s_1 \le K - \varepsilon$, $\mathbb{I}^\star_{K,\varepsilon}(s_1) = 1$ if $s_1 \ge K$, and in between we use the smooth interpolating function $\mathbb{I}^\star_{K,\varepsilon}(s_1) = 10\varepsilon^{-3}(s_1 - (K-\varepsilon))^3 - 15\varepsilon^{-4}(s_1 - (K-\varepsilon))^4 + 6\varepsilon^{-5}(s_1 - (K-\varepsilon))^5$. The previous function h satisfies the compatibility and support conditions from assumption **(F2)** as soon as $K > B + \varepsilon$. Let us mention that the associated function f_0 from Proposition 3.1.1 is not bounded, but has exponential growth. Thus, we are not exactly under our previous working assumptions. Anyhow, following the notations of Proposition 3.1.1 and using a Girsanov transform in the expression of v_0, we obtain $\forall (t,y) \in [0,T] \times \bar{D}_0$, $v_0(t,y) = \exp(\frac{\|\tilde{\mu}\|^2 T}{2} - \frac{\|\sigma^{-1}\mu\|^2 T}{2})\mathbb{E}_y[f(x + \sigma\Lambda^{-1}\sigma_0 X_{T-t})\mathbb{I}_{\tau_{D_0} > T-t}]$ where $X_t = y + \tilde{\mu}t + W_t$, $\tilde{\mu} = \sigma_0^*(\Lambda^{-1})^*\sigma^{-1}\mu$ and W is standard BM. Hence, we are reduced to adapt our previous arguments of Section 4.2 to the case of a drifted BM. In the sequel we freely use the notations of Section 4.2.

For the transition density, the drift affects the radial control of equation (3.4.10). Denoting by $\tilde{P}_b(t,x,z)dz = \exp(b \cdot (z-x) - \|b\|^2 t/2)\tilde{P}(t,x,z)dz$ the transition density of the killed BM with drift b in a wedge, we get that there exist positive constants c, C, c' s.t. $\forall x = (r\cos\theta, r\sin\theta) \in \tilde{D}_0 \cap B(0,R), z = (r_0\cos\theta_0, r_0\sin\theta_0)$ s.t. $|r - r_0| \geq \varepsilon$

$$\forall \zeta, |\zeta| \leq 5, \text{ and } \forall s \in (0,T],\ |\partial_x^\zeta \tilde{P}_b(s,x,z)| \leq \frac{C}{1 \wedge \varepsilon^\xi} \exp(c' r_0) \exp(-c\frac{\varepsilon^2}{s}) \exp(-c\frac{|r-r_0|^2}{s}).$$

On the other hand, when $K > B+\varepsilon$, since $d(f, \bar{D}_0) > 0$, the function f satisfies the compatibility conditions up to order two for the operator $\tilde{\mu}\cdot\nabla + \mathrm{tr}(\sigma_0\sigma_0^* H_.)$. Thus one could show as in Proposition 3.4.4 that we still have the required smoothness properties on the function v_0.

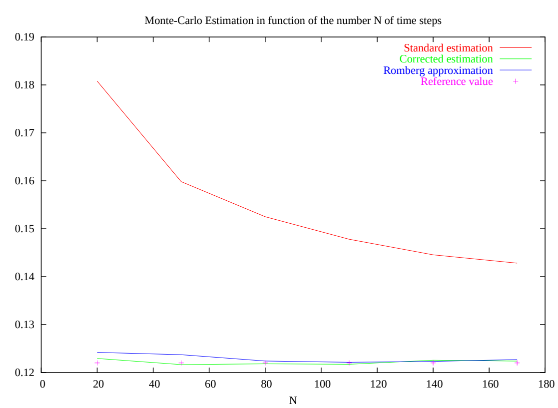
For $r = .04, \sigma_1 = \sigma_2 = .3, \rho = .5, S_0^1 = S_0^2 = 100 = K, B = 90, T = 1, \varepsilon = 5$ we compute the standard Monte-Carlo approximation, the Romberg approximation and the correction proposed in Theorem 3.3.1 for 10^6 paths. The reference value has been computed with 10^6 paths and 15000 times steps.

For the empirical mean we obtain

	MC	MC Shift	Romberg	Reference
$N = 20$	.180807	0.12295	0.124209	0.122017
$N = 50$	.159831	0.121661	0.123736	0.122017
$N = 80$	.15251	0.12184	0.122422	0.122017
$N = 110$	.147821	0.121735	0.122145	0.122017
$N = 140$	.144559	0.122581	0.122338	0.122017
$N = 170$	.142827	0.122393	0.122717	0.122017

For the usual unbiased estimator of the variance, we get

	MC	MC Shift	Romberg	Reference
$N = 20$	.146365	0.106928	.298985	.10413
$N = 50$	.132417	0.105984	.230244	.10413
$N = 80$	.127643	0.106117	.205762	.10413
$N = 110$	.12445	0.106048	.191531	.10413
$N = 140$	.122845	0.106683	.182411	.10413
$N = 170$	.121412	0.106549	.175806	.10413



The width of the 95%-confidence interval is essentially equal to 1.5 10^{-3}. We still observe that the correction procedure of Theorem 3.3.1 has a smaller empirical variance than the Romberg extrapolation.

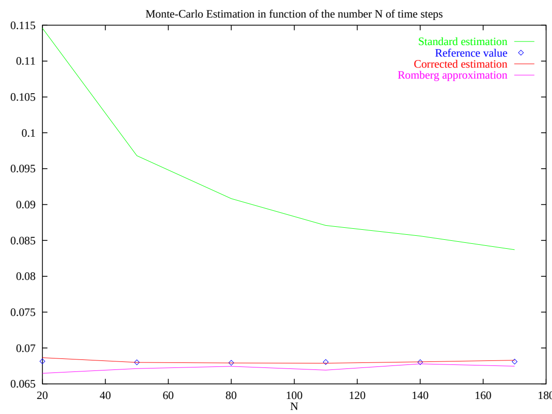
In the special case $\rho = 0$ we can use the Brownian bridges approximations to obtain the reference value for 10^8 paths.

For the empirical mean it comes

	MC	MC Shift	Romberg	Reference
$N = 20$	.114593	.0686465	.0664721	.0681463
$N = 50$	.0968083	.0679943	.0671348	.0680067
$N = 80$	.0908229	.06679159	.0674427	.0679375
$N = 110$	.0870839	.0678774	.0669212	.0680465
$N = 140$	.0856083	.0680759	.0677883	.0680259
$N = 170$	.0837043	.0682977	.0674671	.0681076

For the usual unbiased estimator of the variance, we get

	MC	MC Shift	Romberg	Reference
$N = 20$	.114593	.0632417	.22177	.062833
$N = 50$	.0968083	.062702	.161051	.062729
$N = 80$	.0908229	.0626483	.140756	.0626703
$N = 110$	.0870839	.0625886	.129253	.0627543
$N = 140$	.0856083	.0627679	.122527	.06275
$N = 170$	.0837043	.062994	.115846	.0628089



To conclude this section we mention that in the above examples, the convergence of our correction is always faster than the Romberg approximation. Furthermore we also observe that the associated empirical variance is lower than the one of the Romberg extrapolation.

From a numerical point of view a natural question, according to our previous results, concerns the behaviour of the $o(\sqrt{h})$ appearing in Theorem 3.3.1. We have not experimentally emphasized a constant exponent, anyhow it turns out that the numerical rest is smaller than $O(h^{3/2})$.

5.3 Intuitive extension of the correction in a non-Brownian setting

We introduce in this section an algorithm that aims to extend the correction method of Theorem 3.3.1 to the diffusion case.

Let $(X_s)_{s\geq 0}$ be a diffusion process with dynamics

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{3.5.1}$$

where b, σ are bounded and Lipschitz continuous.

We approximate it by its Euler scheme

$$X_t^N = x + \int_0^t b(X_{\phi(s)}^N)ds + \int_0^t \sigma(X_{\phi(s)}^N)dW_s. \tag{3.5.2}$$

For given domain D and function $f \geq 0$, we know from Theorem 1.5.3 that under suitable assumptions on b, σ, D, f there exists a constant $C > 0$ s.t.

$$0 \leq \mathrm{Err}(T, h, f, x) := \mathbb{E}_x[f(X_T^N)\mathbb{I}_{\tau^N > T}] - \mathbb{E}_x[f(X_T)\mathbb{I}_{\tau > T}] \leq C\sqrt{h}$$

where $\tau^N := \inf\{s_i \geq 0 : X_{s_i}^N \notin D\}$.

We would like as in Theorem 3.3.1 to have an alternative simulation procedure that increases the above speed of convergence.

In the following we freely use the notations of Proposition 1.1.1 and assume the domain D is at least C^3 (see Assumption **(D-k)** in Chapter 1 for a precise definition).

Note that the Euler scheme (3.5.2) is locally in time nothing else but a Brownian Motion with constant drift and diffusion coefficients. For a Brownian Motion in an intersection of half spaces, the correction of Theorem 3.3.1 consisted in constraining the domain in the directions $(n_i)_{i \in [\![1,m]\!]}$, n_i being the inner normal vector of the i^{th} half space, with intensities $C_0\sqrt{h}\|\sigma^* n_i\|$. Thus, on the set $\tau^N > t_i$, $i \in [\![0, N-1]\!]$, and for $X_{t_i}^N$ in a neighbourhood of the boundary s.t. $\Pi_{\partial D}$ is well defined, mimicking the previous procedure, we heuristically extend the correction to the Euler scheme of a diffusion and a more general smooth domain by killing the Euler scheme in t_{i+1} whenever it is outside $D_h(X_{t_i}^N) := \{y \in D : d(y, \partial D) \geq C_0\sqrt{h}\|\sigma^*(X_{t_i}^N)n(\Pi_{\partial D}(X_{t_i}^N))\|\}$.

From an algorithmic point of view, the computation of $n(\Pi_{\partial D}(X_{t_i}^N))$ can be very demanding. Lemma 1.1.2 provides a choice for the neighbourhood on which we compute the correction. Indeed, on $\{\tau^N > t_i\}$ for a given $\eta > 0$ s.t. $X_{t_i}^N \in D \backslash V_{\partial D}(h^{1/2-\eta})$ Lemma 1.1.2 gives $\mathbb{P}[\tau^N = t_{i+1}|\mathcal{F}_{t_i}] \leq C\exp(-ch^{-\eta}) = O_{pol}(h)$. It is therefore useless to refine the simulation procedure for those events.

We sum up this heuristic correction in the following

Algorithm 5.1 Empirical correction procedure in a diffusion framework

- Let the domain D be at least C^3 with bounded boundary.

- Assume X follows the dynamics of equation (3.5.1).

We provide a simulation procedure for the corresponding discretely killed Euler scheme by setting

i) $X_0^N = x$, *fix* $\eta > 0$.

ii) $\forall i \in [\![0, N-1]\!]$ *s.t.* $\tau^N > t_i$ *set* $X_{t_{i+1}}^N := X_{t_i}^N + b(X_{t_i}^N)h + \sigma(X_{t_i}^N)(W_{t_{i+1}} - W_{t_i})$.

- **If** $X_{t_i} \in V_{\partial D}(h^{1/2-\eta})$ *and* $X_{t_{i+1}} \notin D_h(X_{t_i}^N) = \{y \in D : d(y, \partial D) \geq C_0\|\sigma^*(X_{t_i}^N)n(\Pi_{\partial D}(X_{t_i}^N))\|\sqrt{h}\}$
  Kill the path.
  Else
  If $i + 1 \neq N$, *iterate step ii).*

- **If** $X_{t_i} \notin V_{\partial D}(h^{1/2-\eta})$ *and* $X_{t_{i+1}} \notin D$
  Kill the path (rare event).
  Else
  If $i + 1 \neq N$, *iterate step ii).*

We now present the results associated to $D = B(0,1) \subset \mathbb{R}^2$, when the process X is a standard BM and $f(y) = (\frac{1}{2} - \|y\|)^+$. We recall from Chapter 1 that the 95% confidence interval associated to the reference value computed for $N_{MCR} = 10^6, N_R = 14400$ is $I_C(N_{MCR}) = [.0203672, .0204871]$.

Denote by $E_{N_{MC}}(f, N)$ the empirical mean and by $\hat{\sigma}^2_{N_{MC}}(f, N)$ the usual unbiased estimator of the variance for N_{MC} paths and N discretization steps. Put $I^{\pm}_{N_{MC}}(f, N) = E_{N_{MC}}(f, N) \pm 1.96\hat{\sigma}_{N_{MC}}(f, N)/\sqrt{N_{MC}}$. For $N_{MC} = 10^6$, with the previous algorithm we get

	$N = 15$	$N = 30$	$N = 60$	$N = 120$	$N = 240$	$N = 480$
$E_{N_{MC}}(f, N)$	.0200716	.0201528	.0202128	.0202142	.0201588	.020114
$\hat{\sigma}^2_{N_{MC}}(f, N)$	.00671978	.0067683	.00678523	.00678666	.00677632	.00674031
$I^-_{N_{MC}}(f, N)$	.019911	.01999915	.0200514	.0200527	.0199975	.0199531
$I^+_{N_{MC}}(f, N)$	.0202323	.020314	.0203743	.0203756	.0203202	.0202749

Even though most of the intervals $I^{\pm}_{N_{MC}}(f,N)$ do not intersect $I_C(N_{MCR})$, they are quite close to it. This is promising because the computational time employed to get the above estimates is significantly reduced w.r.t. to the one needed to obtain $I_C(N_{MCR})$. Furthermore, the fact that the quantity to estimate is small brings additional numerical difficulty.

5.4 Numerical estimation of the law of the exit time

To conclude this numerical part, we present some results associated to the estimation of the distribution function of the exit time in the case of a bidimensional cone.

In the sequel we use freely the notations of Proposition 3.4.2. From the identities $\int_0^\infty e^{-\beta t^2} I_\nu(\alpha t)dt = \frac{1}{2}\sqrt{\frac{\pi}{\beta}}\exp(\alpha^2/(8\beta))I_{\nu/2}\left(\frac{\alpha^2}{8\beta}\right)$, $2I'_\nu(x) = I_{\nu-1}(x) + I_{\nu+1}(x)$, cf. 11.4.31 and 9.6.26 in [AS72], we obtain by direct integration of expression (3.4.7) that for an initial point $x = (r\cos\theta, r\sin\theta) \in \tilde{D}_0$

$$\mathbb{P}_x[\tau_{\tilde{D}_0} > t] = \frac{2r}{\sqrt{2\pi t}}\, e^{-\frac{r^2}{4t}} \sum_{n\geq 1,\ n \text{ odd}} \frac{1}{n}\sin(\nu_n\theta)\left\{ I_{(\nu_n-1)/2}\left(\frac{r^2}{4t}\right) + I_{(\nu_n+1)/2}\left(\frac{r^2}{4t}\right)\right\}. \tag{3.5.3}$$

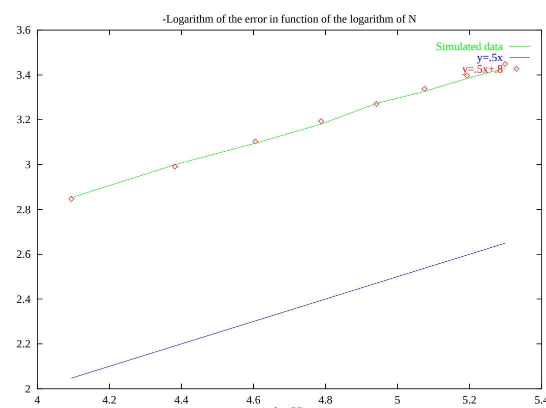
Equation (3.5.3) corrects equation (10) from Iyengar, [Iye85].

We estimate the above quantity with a standard Monte Carlo procedure using the correspondence of Section 4.2 between the standard BM in a wedge and the correlated BM in an orthant. We point out that concerning theoretical estimates for the speed of convergence, we are not under the assumptions of Theorem 3.2.4. Indeed, **(S)** is not satisfied since we take $f = 1$. Both $f|_{\partial D} = 0$ and the support condition of Assumption **(F2)** are violated.

Take $\alpha = \pi/4, t = 1, x = (2\ 1)^*$. Denoting $E_{N_{MC}}(N)$ the empirical mean and $\hat{\sigma}^2_{N_{MC}}(N)$ the usual unbiased estimator of the variance we obtain for $N_{MC} = 10^6$

N	$E_{N_{MC}}(N)$	$\hat{\sigma}^2_{N_{MC}}(N)$	True Value
60	0.321123	0.218003	0.263410762703
80	0.313229	0.215117	0.263410762703
100	0.308674	0.213395	0.263410762703
120	0.304988	0.211971	0.263410762703
140	0.301315	0.210524	0.263410762703
160	0.299338	0.209735	0.263410762703
180	0.297315	0.208919	0.263410762703
200	0.295794	0.2083	0.263410762703

Even though the assumptions needed for our previous expansion result fail, if we plot the logarithm of the error in function of the logarithm of the number of time steps, we still observe a nice right line with slope 1/2.



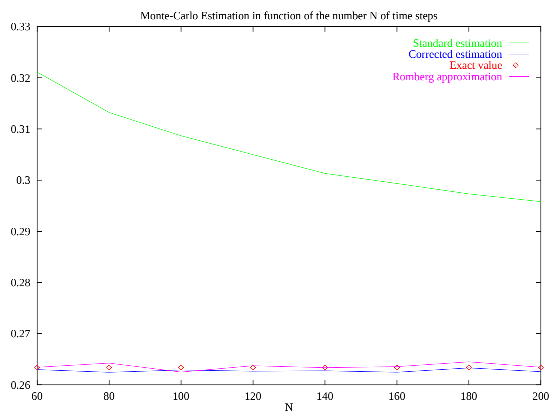
We now give the results obtained with the shifting boundary condition of Theorem 3.3.1 and with the Romberg extrapolation.

For the empirical mean we obtain

N	MC	MC Shift	Romberg	True Value
60	0.321123	0.262989	0.26344	0.263410762703
80	0.313229	0.262455	0.264257	0.263410762703
100	0.308674	0.262913	0.262495	0.263410762703
120	0.304988	0.262667	0.263737	0.263410762703
140	0.301315	0.262754	0.263364	0.263410762703
160	0.299338	0.262463	0.263568	0.263410762703
180	0.297315	0.263317	0.264505	0.263410762703
200	0.295794	0.262586	0.263436	0.263410762703

And for the associated empirical variances

N	MC	MC Shift	Romberg	True Value
60	0.218003	0.193826	0.333513	0
80	0.215117	0.193573	0.316301	0
100	0.213395	0.19379	0.302716	0
120	0.211971	0.193673	0.29338	0
140	0.210524	0.193715	0.287467	0
160	0.209735	0.193576	0.280433	0
180	0.208919	0.193981	0.273886	0
200	0.2083	0.193635	0.270653	0



The previous results are promising. Even though we can not say anymore that the shifting boundary correction is the closest approximation of the exact value, the computational cost of the Romberg extrapolation is a slight advantage for our method. Note also that we still observe that the estimator associated to the shifting boundary correction has the smallest empirical variance. Eventually, the above results emphasize that numerically speaking the shifting boundary correction turns out to be relevant in a wider context than the one of Theorem 3.3.1.

6 Conclusion

In this chapter we obtained expansion results for the weak error in the special case of a discretely killed Brownian Motion in either a half-space or a bidimensional cone. We exploited the explicit asymptotic distribution of the overshoot above the boundary that had previously been characterized as the leading term of the weak error, see Theorem 1.5.2. The correction method proposed to improve the convergence rate also gave promising results.

The main motivation that led us to deal with conical cases comes from Mathematical finance. Indeed, with multi assets, one often defines the domain of a barrier option as a product domain. For the moment we are only able to handle the case of bidimensional domains in a Black-Scholes framework.

Concerning further extensions in bigger dimensions, let us point out that the remaining efforts to be done concern the smoothness properties of the underlying function $v(t,x) = \mathbb{E}_x[f(X_{T-t})\mathbb{I}_{\tau>T-t}]$.

Anyhow, for some special angles, or equivalently for special correlation coefficients, we can extend the method of images presented in Section 4.2 and thus express the transition density as a sum of standard Gaussian kernels. In that case, assuming **(F2)** for f, we have the usual smoothness properties on v, Assumption **(S)** is fulfilled and both the expansion and correction results hold true.

For general angles, or correlations, Bañuelos and Smits, see [BS97], gave the explicit expression of the Green kernel for a generalized cone in arbitrary dimensions. It therefore remains to find assumptions on the function f that guarantee **(S)**. This will concern further investigations.

We also mention that our smoothness controls are not optimal. A possible framework consists in taking the function v in $C_b^{1+\alpha,3+\alpha}([0,T]\times\bar{D}),\ \alpha\in(0,1)$, instead of $C_b^{2,4}([0,T]\times\bar{D})$ in Assumption **(S)**. In this way we would get rid of the parabolic compatibility condition of order 2 in **(F2)**.

The boundedness assumption on f gave the boundedness of the derivatives $\partial_x^\alpha v, |\alpha|\le 4$. However, all we need for the proof of our main results is a control of the type

$$\exists c>0, C>0,\ \alpha,\ [\alpha|\le 4,\ |\partial_x^\alpha v(s,x)|\le C\exp(c\|x\|),\ \text{for } (s,x)\in[0,T]\times\bar{D}.$$

Hence, we can weaken the assumptions on f while this last condition is fulfilled (and take for instance f bounded and vanishing in a neighbourhood of the boundary in the half space case).

Anyhow, up to now, we are forced for the error analysis to consider functions f vanishing on the boundary of the cone. We recall however, see Section 5.4 for details, that from a numerical point of view our results seem to hold without these assumptions.

Chapter 4

A Forward-Backward Stochastic Algorithm for Quasi-Linear Parabolic PDEs

1 Introduction

Introduced first by Antonelli [Ant93], and then by Ma, Protter and Yong [MPY94], Forward-Backward Stochastic Differential Equations (FBSDEs in short) provide an extension of the Feynman-Kac representation to a certain class of quasi-linear parabolic PDEs. These equations also appear in a large number of application fields such as the Hamiltonian formulation of control problems or the option hedging problem with large investors in financial mathematics (i.e. when the wealth or strategy of an agent has an impact on the volatility). We refer to the monograph of Ma and Yong, [MY99] for details and further applications.

1.1 FBSDE Theory and Discretization Algorithm

Connection between FBSDEs and Quasi-linear parabolic PDEs. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a d-dimensional Brownian motion $(B_t)_{t\in[0,T]}$, where T denotes an arbitrarily prescribed nonnegative real. For a given initial condition $x_0 \in \mathbb{R}^d$, a Forward-Backward SDE strongly couples a diffusion process U to the solution (V, W) of a Backward SDE (as defined in the earlier work of Pardoux and Peng [PP90]):

$$(E) \begin{cases} \forall t \in [0,T], \\ U_t = x_0 + \displaystyle\int_0^t b(U_s, V_s, W_s)ds + \int_0^t \sigma(U_s, V_s)dB_s, \\ V_t = H(U_T) + \displaystyle\int_t^T f(U_s, V_s, W_s)ds - \int_t^T W_s dB_s. \end{cases}$$

In the whole paper, the coefficients b, f, σ and H are deterministic (and for simplicity also time independent). In this case, Ma, Protter and Yong [MPY94], Pardoux and Tang [PT99] and Delarue [Del02] have investigated in detail the link with the following quasi-linear PDE on $[0,T[\times\mathbb{R}^d$:

$$(\mathcal{E}) \begin{cases} \partial_t u(t,x) + \langle b(x, u(t,x), \nabla_x u(t,x)\sigma(x, u(t,x))), \nabla_x u(t,x)\rangle + \dfrac{1}{2}\mathrm{tr}(a(x, u(t,x))\nabla^2_{x,x} u(t,x)) \\ +f(x, u(t,x), \nabla_x u(t,x)\sigma(x, u(t,x))) = 0, \\ u(T,x) = H(x), \end{cases}$$

with $a(x,y) = (\sigma\sigma^*)(x,y),\ (x,y) \in \mathbb{R}^d \times \mathbb{R}$.

A Probabilistic Numerical Method for FBSDEs and Quasi-Linear PDEs. This paper aims to derive from the probabilistic theory of FBSDEs a completely tractable algorithm to approximate the solution of the equation $(\mathcal{E})$. As a bypass product, the procedure also provides a discretization of the triple (U, V, W).

Most of the available numerical methods proposed so far are purely analytic and involve finite-difference or finite-element techniques to approximate the solution u of $(\mathcal{E})$. For example, the discretization procedure for FBSDEs of type (E) given in Douglas, Ma and Protter [DMP96] consists in discretizing first the PDE $(\mathcal{E})$ and then in deriving an approximation of the underlying FBSDE.

At the opposite, we propose in this paper to derive from the FBSDE representation a numerical scheme for quasi-linear equations of type $(\mathcal{E})$. This strategy finds its origin in the earlier work of Chevance [Che97] who introduced a time-space discretization scheme in the decoupled or so-called "pure backward" case. In this latter frame, the coefficients b and σ do not depend on V and W and the forward equation reduces to a classical SDE. The process U then appears as an "objective diffusion". Note in this particular case that the time-space discretization scheme and the specific form of the system (E) permit to use a standard "dynamic programming principle".

From a numerical point of view, two other kinds of approaches have been developed in the backward case. The first one is based on Monte-Carlo simulations and Malliavin integration by parts, see Bouchard and Touzi [BT04]. The other one relies on quantization techniques for a discretization scheme of the underlying forward equation. Quantization consists in approximating a random variable by a suitable discrete law. It provides a cheap and numerically efficient alternative to usual Monte-Carlo methods to estimate expectations. In the works of Bally and Pagès [BP03] or Bally, Pagès and Printems [BPP02] on American options, the key idea is to perform an optimal quantization procedure of a discretized version of the underlying diffusion process in order to compute **once for all** by a Monte-Carlo method the corresponding semi-group. Then, the second step consists in doing a dynamic programming descent. For other applications of quantization, we refer to the works of Pagès, Pham and Printems, [PPP04] or Pagès and Printems, [PP04].

Discretization Strategy. In the coupled case, or quasi-linear framework, the diffusion U is not "objective" anymore. Indeed, due to the strong non-linearity of the equation $(\mathcal{E})$, the coefficients of the underlying forward diffusion depend on the solution and on its gradient.

In particular, we can not quantify a discretization scheme of the diffusion process as explained above. This is well understood: without approximating u, we do not have any *a priori* knowledge of the optimal shape of the associated grid. Hence, we just focus on the quantization of the Brownian increments appearing in the forward SDE and then choose to define the approximate diffusion on a sequence of truncated d-dimensional Cartesian grids. Note that the discretization procedure of U is now coupled to the approximation procedure of $(u, \nabla_x u)$ (denoted in a generic way by $(\overline{u}, \overline{v})$) which is computed along the same sequence of grids. The time-space discretization scheme allows to define $(\overline{u}, \overline{v})$ and the approximations of the transitions of U in order to recover a kind of "dynamic programming principle". Consider indeed a given regular time mesh $(t_i = ih)_{i \in \{0,\dots,N\}}$ of $[0, T]$, h being the step size. To every discretization time t_i, associate a spatial Cartesian grid $\mathcal{C}_i \equiv \{(x_k^i)_{k \in \mathcal{I}_i}\}$, $\mathcal{I}_i \subset \mathbb{N}^*$, such that $\forall i \in \{0, ..., N-1\}$, $\mathcal{C}_i \subset \mathcal{C}_{i+1}$. Starting from $t_N = T$ for which the solution of $(\mathcal{E})$ and its gradient are known, the transition of U from t_i to t_{i+1}, $i \in \{0, ..., N-1\}$, is then updated iteratively through the Brownian quantized increments and through the values of $\overline{u}(t_{i+1}, .)$ and $\overline{v}(t_{i+1}, .)$ on the grid $\mathcal{C}_{i+1}$. This permits to express the approximation $\overline{u}(t_i, .)$ through a discretized version of the Feynman-Kac formula.

At this stage, it remains to precise the way we update the approximation of the gradient of the solution u. We mention actually that the strategy aims to approximate the product $\nabla_x u(t_k, .)\sigma(., u(t_k, .))$ instead of $\nabla_x u(t_k, .)$ itself. This explains the specific writing of the PDE $(\mathcal{E})$. We then proceed in two different steps. A first approximation is performed through a martingale increment procedure as done in the discretization scheme of BSDEs explained in Bouchard and Touzi [BT04], or as used in Bally et *al.* [BPP02]. A second step consists in quantizing the Gaussian increments appearing in the former representation. This is an alternative solution to the usual techniques based on Monte-Carlo simulations or on Malliavin integration by parts as employed in [BT04]. Of course, if the matrix $\sigma\sigma^*$ is non-degenerate, the strategy still applies, up to an inversion procedure, to coefficients of the form $(b,f)(x,u(t,x),\nabla_x u(t,x))$.

Extra References. Some of the preliminaries of our approach can be found in Milstein and Tretyakov [MT99] in the specific case where $(b,f)(x,u(t,x),\nabla_x u(t,x)\sigma(t,x,u(t,x)))$ reduces to $(b,f)(x,u(t,x))$. Note however that the proof of the convergence of the underlying numerical scheme proposed in this reference just holds for so-called "equations with small parameter" (i.e. with a small diffusion matrix). Generally speaking, the authors have then to control the regularity properties of the solution of the transport problem associated to the equation $(\mathcal{E})$ (i.e. the same equation as $(\mathcal{E})$, but without any second order terms). Without discussing in detail the basic assumptions made in our paper, note that no condition of this type appears in the sequel: in particular, the matrix a is assumed to be uniformly elliptic. Hence, we feel that the work of Milstein and Tretyakov [MT99] applies to a different framework than ours. For this reason, we avoid any further comparisons between both situations. Add finally for the sake of completeness that Makarov [Mak03] has successfully applied the strategy of Milstein and Tretyakov [MT99] to the case $(b,f) \equiv (b,f)(x,u(t,x),\nabla_x u(t,x)\sigma(t,x,u(t,x)))$ under suitable smoothness properties on the coefficients. Of course, the small parameter condition is then still necessary.

1.2 Novelties Brought by the Paper

A purely probabilistic point of view. The proof of the convergence of our algorithm is somehow the first to be essentially of probabilistic nature, since we are able to adapt the usual stability techniques of BSDE theory to the discretized framework. Note in particular that we follow the proof of uniqueness in the *four step scheme* given in Ma, Protter and Yong [MPY94] to handle the strong coupling between the forward and backward components.

In the discretized framework, the gradient terms appearing in b and f bring additional difficulties. Indeed, our gradient approximation does not appear as a representation process given by the martingale representation theorem as the process W in (E). In particular, the strategy introduced by Pardoux and Peng [PP90] to estimate the L^2 norm of W over $[0,T]$ fails in the discretized setting. We then propose a specific probabilistic strategy to overcome this deep trouble and thus to handle the nonlinearities of order one, see Subsections 3.3 and 9.3 for details

Convergence under weak assumptions. In Douglas, Ma and Protter [DMP96], the authors handle the gradient terms by working under smoothness assumptions that allow them to study the gradient of u as the solution of the differentiated PDE.

Our strategy permits to avoid to differentiate the PDE and thus to really weaken the assumptions required both on the coefficients of (E) and on the smoothness of the solution u of $(\mathcal{E})$ in the above reference. In the previous paper, the coefficients are assumed to be smoothly differentiable and bounded. We just suppose that they are Lipschitz continuous and bounded in x. In Douglas and *al.*, the solution u of $(\mathcal{E})$ is at least bounded in $C^{2+\alpha/2,4+\alpha}([0,T]\times\mathbb{R}^d), \alpha \in]0,1[$. In our paper, we only impose u to belong to $C^{1,2}([0,T]\times\mathbb{R}^d)$ with bounded derivatives of order one in t and one and two in x.

A Completely Tractable Algorithm. Furthermore in [DMP96], the authors always take into consideration the case of infinite spatial grids. This turns out to be simpler for the convergence analysis, anyhow it does not provide in all generality a fully implementable algorithm. We discuss the impact of the truncation of the grids and analyze its contribution in the error.

Finally, a linear interpolation procedure is also used in Douglas et *al.* to define the algorithm. This can be heavy in large dimension. The algorithm we propose allows to define the approximate solution only at the nodes of the spatial grid. In this way, we feel that our method is simpler to implement and numerically cheaper. Note moreover that we avoid the inversion of large linear systems associated to "usual" numerical analysis techniques.

1.3 Organization of the Paper

In Section 2, we detail general assumption and notation as well as several smoothness properties of the solution u of $(\mathcal{E})$. We also precise the connection between the FBSDE (E) and the quasi-linear PDE $(\mathcal{E})$. Section 3 explains the main algorithmic choices. We present in particular the various steps that led us to the current discretization scheme. The main results are stated and discussed in Section 4. In particular, we give an estimate of the speed of convergence of the algorithm. As a probabilistic counterpart, we estimate the difference between the approximating processes and the initial solution (U, V, W) of (E). Numerical examples are presented in Section 5.

The end of the paper is then mainly devoted to the proof of the convergence results. The proof is divided into three parts. Various *a priori* controls of the discrete objects are stated and proved in Section 6. In Section 7, we adapt the FBSDE machinery to our setting to prove a suitable stability property. Section 8 is then devoted to the last step of the proof and more precisely to a specific refinement of Gronwall's Lemma.

As a conclusion, we compare in Section 9 our strategy to other methods and explain some technical points that motivated the choice of our current algorithm. We also indicate further conceivable extensions.

2 Non-Linear Feynman-Kac Formula

In this section, we first give the assumptions on the coefficients of the FBSDE and then briefly recall the connection with quasi-linear PDEs. As detailed later, under these assumptions, the underlying PDE admits a unique strong solution, whose partial derivatives of order one in t and one and two in x are controlled on the whole domain by known parameters. For the sake of simplicity, we also assume that the coefficients do not depend on time.

2.1 Coefficients of the Equation

For a given $d \in \mathbb{N}^*$, we consider the following coefficients:

$$b : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d,\ f : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R},\ \sigma : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^{d\times d},\ H : \mathbb{R}^d \to \mathbb{R}.$$

In the following, we denote by $|.|$ the Euclidean norm of $\mathbb{R}^n$, $n \geq 1$.

Assumption (A) We say that the functions b, f, H and σ satisfy Assumption **(A)** if there exist four constants $\alpha > 0$, κ, Λ and $\lambda > 0$ such that:

(A.1) $\forall (x, y, z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$, $\big|(b, f, \sigma, H)(x, y, z)\big| \leq \Lambda\big(1 + |y| + |z|\big)$.

(A.2) $\forall (x, y) \in \mathbb{R}^d \times \mathbb{R}$, $\forall \zeta \in \mathbb{R}^d$, $\langle \zeta, a(x, y)\zeta\rangle \geq \lambda|\zeta|^2$, where $\forall (x, y) \in \mathbb{R}^d \times \mathbb{R}$, $a(x, y) = \sigma\sigma^*(x, y)$.

(A.3) $\forall (x, y, z),\ (x', y', z') \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$:

$$\big|(b, f, \sigma, H)(x, y, z) - (b, f, \sigma, H)(x', y', z')\big| \leq \kappa\big(|x - x'| + |y - y'| + |z - z'|\big).$$

(A.4) The function H belongs to $\mathcal{C}^{2+\alpha}(\mathbb{R}^d)$ and its $\mathcal{C}^{2+\alpha}$ norm is bounded by κ.

From now on, Assumption **(A)** *is in force.*

2.2 Forward-Backward SDE

Consider now a given $T > 0$ and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a Brownian motion $(B_t)_{0\leq t\leq T}$ whose natural filtration, augmented with $\mathbb{P}$ null sets, is denoted by $\{\mathcal{F}_t\}_{0\leq t\leq T}$.

Fix an initial condition $x_0 \in \mathbb{R}^d$ and recall (see Ma, Protter and Yong [MPY94] and Delarue [Del02]) that there exists a unique progressively measurable triple (U, V, W), with values in $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$, such that:

$$\mathbb{E} \sup_{t \in [0,T]} \left(|U_t|^2 + |V_t|^2 \right) < +\infty, \ \mathbb{E} \int_0^T |W_t|^2 dt < +\infty,$$

and which satisfies $\mathbb{P}$ almost surely the couple of equations (E).

2.3 Quasi-Linear PDE

According to Ladyzhenskaya et *al.* [LSU68], Chapter VI, Theorem 4.1, and to [MPY94] (up to a regularization procedure of the coefficients), we claim that $(\mathcal{E})$ admits a solution $u \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$ satisfying:

Theorem 4.2.1 *There exists a constant $C_{4.2.1}$, only depending on T and on known parameters appearing in* **(A)**, *such that* $\forall (t,x) \in [0,T] \times \mathbb{R}^d$,

$$\begin{aligned} &|u(t,x)| + |\nabla_x u(t,x)| + |\nabla^2_{x,x} u(t,x)| + |\partial_t u(t,x)| \\ &\quad + \sup_{t' \in [0,T], t \neq t'} \left[|t-t'|^{-1/2} |\nabla u(t,x) - \nabla u(t',x)| \right] \leq C_{4.2.1}. \end{aligned}$$

Moreover, u is unique in the class of functions $\widetilde{u} \in \mathcal{C}([0,T] \times \mathbb{R}^d, \mathbb{R}) \cap \mathcal{C}^{1,2}([0,T[\times \mathbb{R}^d, \mathbb{R})$ which satisfy:

$$\sup_{(t,x) \in [0,T[\times \mathbb{R}^d} \left(|\widetilde{u}(t,x)| + |\nabla_x \widetilde{u}(t,x)| \right) < +\infty.$$

From Ma, Protter and Yong [MPY94], Pardoux and Tang [PT99] and Delarue [Del02], the FBSDE (E) is connected with the PDE $(\mathcal{E})$.

Set $\forall (t,x) \in [0,T[\times \mathbb{R}^d, \ v(t,x) = \nabla_x u(t,x) \sigma(x, u(t,x))$. The relationship between (E) and $(\mathcal{E})$ can be summed up as follows:

$$\forall t \in [0,T], \ V_t = u(t, U_t), \ W_t = v(t, U_t), \ V_t = \mathbb{E}[V_T | \mathcal{F}_t] + \mathbb{E}\left[\int_t^T f(U_s, V_s, W_s) ds | \mathcal{F}_t \right]. \qquad (4.2.1)$$

3 Approximation Procedure

In this section, we detail the construction of the approximation algorithm of the solution u of $(\mathcal{E})$. We explain how the final form of the discretization procedure can be derived step by step from the forward-backward representation (E). We also present the quantization techniques used in order to compute expectations related to Brownian increments and we discuss the choice of the underlying spatial grids which appear in the approximating scheme.

3.1 Rough Algorithms

Localization Procedure. As already seen in the introduction, the forward-backward equation (E) appears as the starting point of our discretization procedure. Indeed, this couple of stochastic equations provides a probabilistic representation of the quasi-linear PDE $(\mathcal{E})$ and summarizes in an integral form the local evolution of the solution u. Define now, for a given integer $N \geq 1$, a regular mesh of $[0,T]$ with step $h \equiv T/N$:

$$t_0 \equiv 0, \ t_1 \equiv h = T/N, \ t_2 \equiv 2h, ..., t_N \equiv T.$$

Using (E), for $k \in \{0, ..., N\}$, the evolution of the solution u along the interval $[t_k, t_{k+1}]$ writes from a probabilistic point of view:

$$\begin{cases} U_{t_{k+1}} = U_{t_k} + \displaystyle\int_{t_k}^{t_{k+1}} b(U_s, V_s, W_s) ds + \int_{t_k}^{t_{k+1}} \sigma(U_s, V_s) dB_s, \\ V_{t_k} = V_{t_{k+1}} + \displaystyle\int_{t_k}^{t_{k+1}} f(U_s, V_s, W_s) ds - \int_{t_k}^{t_{k+1}} W_s dB_s, \end{cases} \qquad (4.3.1)$$

where the initial condition U_{t_k} is a square integrable and $\mathcal{F}_{t_k}$-measurable random vector with values in $\mathbb{R}^d$. In particular, conditioning by $U_{t_k} = x \in \mathbb{R}^d$, we deduce:

$$U_{t_{k+1}}^{t_k,x} = x + \int_{t_k}^{t_{k+1}} b\big(U_s^{t_k,x}, V_s^{t_k,x}, W_s^{t_k,x}\big)ds + \int_{t_k}^{t_{k+1}} \sigma\big(U_s^{t_k,x}, V_s^{t_k,x}\big)dB_s, \tag{4.3.2}$$

and

$$\begin{aligned} V_{t_k}^{t_k,x} &= \mathbb{E}\Big[V_{t_{k+1}}^{t_k,x} + \int_{t_k}^{t_{k+1}} f\big(U_s^{t_k,x}, V_s^{t_k,x}, W_s^{t_k,x}\big)ds\Big], \\ \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} W_s^{t_k,x}ds\Big] &= \mathbb{E}\big[V_{t_{k+1}}^{t_k,x}\big(B_{t_{k+1}} - B_{t_k}\big)\big] + O(h^{3/2}), \end{aligned}$$

where the superscript (t_k, x) denotes the starting point of the diffusion process U. The remaining term $O(h^{3/2})$ is a consequence of Assumption **(A.1)**, (4.2.1) (relationships between V, W and u) and Theorem 4.2.1 (boundedness of u and $\nabla_x u$). Relation (4.2.1) also yields

$$\begin{aligned} u(t_k, x) &= \mathbb{E}\Big[u\big(t_{k+1}, U_{t_{k+1}}^{t_k,x}\big) + \int_{t_k}^{t_{k+1}} f\big(U_s^{t_k,x}, V_s^{t_k,x}, W_s^{t_k,x}\big)ds\Big], \\ \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} W_s^{t_k,x}ds\Big] &= \mathbb{E}\big[u\big(t_{k+1}, U_{t_{k+1}}^{t_k,x}\big)\big(B_{t_{k+1}} - B_{t_k}\big)\big] + O(h^{3/2}). \end{aligned} \tag{4.3.3}$$

In the following, the Brownian increment $B_{t_{k+1}} - B_{t_k}$ is denoted by ΔB^k. In particular, we derive from the above relation that, neglecting the rest, the best constant approximation of $(W_s)_{s\in[t_k,t_{k+1}]}$ in the $L^2([t_k, t_{k+1}] \times \Omega, ds \otimes d\mathbb{P})$ sense is given by:

$$\hat{W}_{t_k}^{t_k,x} \equiv h^{-1}\mathbb{E}[u(t_{k+1}, U_{t_{k+1}}^{t_k,x})\Delta B^k]. \tag{4.3.4}$$

Relationships (4.3.2), (4.3.3) and (4.3.4) provide a rough background to discretize the local evolution given in (4.3.1). However, this first form is not satisfactory from an algorithmic point of view. Indeed, because of the strong coupling between the forward and the backward equations, the transition of the diffusion depends on the solution itself, both in the drift term and in the martingale part. At the opposite, in the so-called "pure backward" case, or correspondingly for semi-linear equations, the underlying operator does not depend on the solution. In such a case, the classical Euler machinery applies to discretize the decoupled diffusion U.

Induction Principle. Recall that similar difficulties occur to establish the unique solvability of the FBSDE (E). In Delarue [Del02], the first author overcomes the strong coupling between the forward and backward equations by solving by induction the sequence of local FBSDEs (4.3.1), k running downwards from N to 0. By analogy with this approach, the discretization procedure of the forward component on a step $[t_k, t_{k+1}[$, $0 \le k \le N-1$, must take into account the issues of the former local discretizations of the backward equation, and more specifically the approximations of $u(t_{k+1}, .)$ and $v(t_{k+1}, .)$.

Predictors. Assume to this end that, at time t_{k+1}, some approximations $\overline{u}(t_{k+1}, .), \overline{v}(t_{k+1}, .)$ of $u(t_{k+1}, .)$, $v(t_{k+1}, .)$ are available on the whole space. These approximations appear as the "natural" predictors of the true solution and of its gradient on $[t_k, t_{k+1}[$. Introducing the forward approximating transition

$$\mathcal{T}(t_k, x) \equiv b(x, \overline{u}(t_{k+1}, x), \overline{v}(t_{k+1}, x))h + \sigma(x, \overline{u}(t_{k+1}, x))\Delta B^k, \tag{4.3.5}$$

we derive an associated updating procedure by setting:

$$\begin{aligned} \overline{u}(t_k, x) &\equiv \mathbb{E}\big[\overline{u}\big(t_{k+1}, x + \mathcal{T}(t_k, x)\big)\big] + hf\big(x, \overline{u}(t_{k+1}, x), \overline{v}(t_{k+1}, x)\big), \\ \overline{v}(t_k, x) &\equiv h^{-1}\mathbb{E}\big[\overline{u}\big(t_{k+1}, x + \mathcal{T}(t_k, x)\big)\Delta B^k\big]. \end{aligned} \tag{4.3.6}$$

Once the predictors are updated, the procedure can be iterated. Of course, at time $T = t_N$ we set: $\overline{u}(t_N, .) \equiv H(.)$ and $\overline{v}(t_N, .) \equiv \nabla_x H(.)\sigma(., H(.))$. Note in particular that the expectations appearing in (4.3.6) are correctly defined. Indeed, a simple induction procedure shows from Assumptions **(A.1)** and **(A.4)** that $\overline{u}$ and $\overline{v}$ are bounded on $\{t_0, ..., t_N\} \times \mathbb{R}^d$ (but the bound depends on the discretization parameters).

Spatial Discretization. In practice, it is anyhow impossible to define and to update $\overline{u}, \overline{v}$ on the whole space as done above. The most natural strategy consists in defining the approximations $\overline{u}(t_k, .)$ and $\overline{v}(t_k, .)$ of the true solution and its gradient on a discrete subset of $\mathbb{R}^d$. Those approximations could then be extended to the whole space with a linear interpolation procedure. However, in high dimension, this last operation can be computationally demanding. We thus prefer for simplicity to restrict the approximations to a given spatial grid $\mathcal{C}_k \equiv \{(x_j^k)_{j\in\mathcal{I}_k}, \mathcal{I}_k \subset \mathbb{N}^*\} \subset \mathbb{R}^d$, for $k \in \{0, ..., N\}$. This choice imposes to modify (4.3.6). Indeed, the "terminal" value $x + \mathcal{T}(t_k, x)$ must belong to the former grid $\mathcal{C}_{k+1}$.

Hence, denoting by Π_{k+1} a projection mapping on the grid $\mathcal{C}_{k+1}$, we replace (4.3.6) by:

$$\begin{aligned} \forall x \in \mathcal{C}_k,\ \overline{u}(t_k, x) &\equiv \mathbb{E}\big[\overline{u}\big(t_{k+1}, \Pi_{k+1}(x + \mathcal{T}(t_k, x))\big)\big] + hf\big(x, \overline{u}(t_{k+1}, x), \overline{v}(t_{k+1}, x)\big), \\ \overline{v}(t_k, x) &\equiv h^{-1}\mathbb{E}\big[\overline{u}\big(t_{k+1}, \Pi_{k+1}(x + \mathcal{T}(t_k, x))\big)\Delta B^k\big]. \end{aligned} \tag{4.3.7}$$

In the following, we suppose that $\forall (i,j) \in \{0, ..., N\}^2,\ j < i \Rightarrow \mathcal{C}_j \subset \mathcal{C}_i$, so that $\overline{u}(t_{k+1}, x), \overline{v}(t_{k+1}, x)$ are well defined for $x \in \mathcal{C}_k$. Note that if the cardinal of $\mathcal{C}_k$ is finite for every k, the above scheme is already implementable up to the computations of the underlying expectations.

We then need to detail the way the Gaussian integrals appearing in (4.3.7) are computed and to precise the choice of the grids. This is done in Subsections 3.2 and 3.4.

Global Updating. Our updating method using the predictors $\overline{u}(t_{k+1}, .), \overline{v}(t_{k+1}, .)$ is an alternative to the standard fixed point procedure. This latter consists in giving first some global predictors $\overline{u}^0(t_k, .), \overline{v}^0(t_k, .),\ k \in \{0, ..., N\}$. These are used to compute the transitions of the approximating forward process. In this way, we obtain a decoupled forward-backward system, whose solution may be computed by a standard dynamic programming algorithm. A complete descent of this algorithm from $k = N$ to $k = 0$ produces $\overline{u}^1(t_k, .), \overline{v}^1(t_k, .),\ k \in \{0, ..., N\}$, from which we can iterate the previous procedure. In this frame, the underlying distance used to describe the convergence of the fixed point procedure involves all the discretization times and all the spatial points. This strategy appears as a "global updating" one.

From a numerical point of view this seems unrealistic. Indeed, one would need to solve a large number of linear problems. This would either require to use massive Monte-Carlo simulations at each step of the algorithm or to apply, again at each step of the algorithm, a quantization procedure of the approximate diffusion process associated to the current linear problem. Furthermore, it seems intuitively clear that a local updating is far more efficient than a global one.

3.2 Quantization

Expectations Approximation. Two methods are conceivable to compute expectations appearing in (4.3.7).

The first one consists in applying the classical Monte-Carlo procedure for every $k \in \{0, ..., N-1\}$ and for every $x \in \mathcal{C}_k$, and therefore to repeat this argument $\sum_{k=0}^{N-1} |\mathcal{I}_k|$ times. From the central limit Theorem, such a strategy would lead to perform $\sum_{k=0}^{N-1} |\mathcal{I}_k| \times \varepsilon_{\mathrm{MC}}^{-2}$ elementary operations to compute underlying expectations up to the error term $\varepsilon_{\mathrm{MC}}$. This approach seems rather hopeless.

A more efficient method consists in replacing the Gaussian variables appearing in (4.3.7) by discrete ones with known weights. This procedure is known as "quantization". Consider to this end a probability measure on $\mathbb{R}^d$ with finite support $(y_i)_{i\in\{1,\dots,M\}}$ and denote by $(p_i)_{i\in\{1,\dots,M\}}$ the associated weights. Replace then the Gaussian increments in (4.3.7) by this law. For a given $x \in \mathcal{C}_k$, $0 \le k \le N$, the expectations appearing in the induction scheme (4.3.7) then write:

$$\overline{u}(t_k,x) \equiv \sum_{i=1}^{M}\Big[p_i\overline{u}\big(t_{k+1},\Pi_{k+1}(x+b(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x))h+\sigma(x,\overline{u}(t_{k+1},x))y_i)\big)\Big]$$
$$+ hf(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x)),$$
$$\overline{v}(t,x) \equiv h^{-1}\sum_{i=1}^{M}\Big[p_i\overline{u}\big(t_{k+1},\Pi_{k+1}(x+b(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x))h+\sigma(x,\overline{u}(t_{k+1},x))y_i)\big)y_i\Big],$$

and are explicitly computable.

Quantization Principle. We briefly recall the basic principle of quantization and refer to the monograph of Graf and Lushgy, cf. [GL00], for details. Generally speaking, for a given random variable Δ, the quantization procedure consists in replacing Δ by its projection on a finite grid $\Lambda(M) \equiv \{(y_i)_{i\in\{1,\dots,M\}}\} \subset \mathbb{R}^d,\ M \in \mathbb{N}^*$. The projection mapping $G_{\Lambda(M)}$ simply writes:

$$G_{\Lambda(M)}(y) = \sum_{i=1}^{M} y_i \mathbf{1}_{V_i}(y),$$

where, for every $i \in \{1,\dots,M\}$, $V_i \equiv \{y \in \mathbb{R}^d,\ |y-y_i| = \min_{j\in\{1,\dots,M\}} |y-y_j|\}$. In quantization theory, V_i is known as the Voronoi tessel of y_i.

Define now the quantization of Δ with respect to the grid $\Lambda(M)$ by: $\hat{\Delta} \equiv G_{\Lambda(M)}(\Delta)$. The law of $\hat{\Delta}$ then writes: $p_i \equiv \mathbb{P}\{\Delta \in V_i\}$, $i \in \{1,\dots,M\}$. For a given grid, one can compute these values once for all, using e.g. a Monte-Carlo method, so that we may assume to have the "exact" values of these weights.

The crucial step in the quantization procedure therefore lies in the choice of the grid $\Lambda(M)$. To this end, we introduce the so-called "distortion" in order to measure the error associated to the grid $\Lambda(M)$:

$$D_{\Delta,p}\big(\Lambda(M)\big) \equiv \|\Delta - \hat{\Delta}\|_{L^p(\mathbb{P})},\ p \geq 1. \tag{4.3.8}$$

Optimal Grids. The Bucklew-Wise Theorem, see Theorem 6.2 Chapter II in [GL00] for details, then gives for $\Lambda^*(M)$ achieving the minimum in (4.3.8):

$$M^{p/d} D^p_{\Delta,p}\big(\Lambda^*(M)\big) \longrightarrow C(p,d) \text{ as } M \to +\infty, \tag{4.3.9}$$

where $C(p,d)$ is a constant depending on p,d and the variable at hand.

Various algorithms are available to compute an optimal grid $\Lambda^*(M)$. In dimension 1, one may use Lloyd's algorithm, which is deterministic. For $d>1$, one usually uses the Kohonen algorithm which is a stochastic one, see Bally and Pagès [BP03]. We also recall that, for $d>1$, the optimal grid is not unique.

Up to a rescaling, the basic object associated to Brownian increments is a d-dimensional standard normal random variable. Hence, we assume in the following that an optimal grid $\Lambda^*(M)$ for $\Delta \sim \mathcal{N}(0,\mathbf{I}_d)$ as well as the associated weights $(p_i)_{i\in\{1,\dots,M\}}$ are given and "perfectly" computed. Let us remark that the Feynman-Kac formula gives an analytical interpretation of p_i. Indeed, one has $\forall i \in \{1,\dots,M\},\ p_i = u_i(0,0)$, where u_i is the solution of the backward heat equation

$$\partial_t u_i(t,x) + \frac{1}{2}\Delta u_i(t,x) = 0, (t,x) \in [0,1[\times\mathbb{R}^d, u_i(1,.) = \mathbf{1}_{V_i}(.).$$

Quantized Algorithm. We are now in a position to introduce a more tractable induction principle. To this end, for every $k \in \{0,\dots,N-1\}$, set $g(\Delta B^k) \equiv h^{1/2} G_{\Lambda^*(M)}(h^{-1/2}\Delta B^k)$. In particular, we note from (4.3.9) that, for every $p \geq 1$, there exists a constant $C_{\text{Quantiz}}(p,d)$ such that:

$$\mathbb{E}\big[|g(\Delta B^k) - \Delta B^k|^p\big]^{1/p} \leq C_{\text{Quantiz}}(p,d) h^{1/2} M^{-1/d}. \tag{4.3.10}$$

Turn now (4.3.5) and (4.3.7) into:

$$\forall x \in \mathcal{C}_k,\ \mathcal{T}(t_k,x) \equiv b(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x))h + \sigma(x,\overline{u}(t_{k+1},x))g(\Delta B^k), \tag{4.3.11}$$

and,

$$\forall x \in \mathcal{C}_k,\ \overline{u}(t_k,x) \equiv \mathbb{E}\big[\overline{u}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}(t_k,x))\big)\big] + hf\big(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x)\big),$$
$$\overline{v}(t_k,x) \equiv h^{-1}\mathbb{E}\big[\overline{u}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}(t_k,x))\big)g(\Delta B^k)\big]. \tag{4.3.12}$$

To sum up our strategy, the use of predictors allows to recover a kind of standard dynamic programming principle. The quantization provides an easy, cheap and computable algorithm.

3.3 Algorithms

In the following algorithms, we define approximation schemes of the solution u of $(\mathcal{E})$, on spatial grids $(\mathcal{C}_k)_{0\le k\le N}$ s.t. $\mathcal{C}_j \subset \mathcal{C}_k,\ j \le k$. We denote by $\Pi_k(x)$ the projection of $x \in \mathbb{R}^d$ on the grid $\mathcal{C}_k$.

We "decouple" the system plugging into the forward transition the available estimates of the solution and its gradient at the previous time step. We use quantized Gaussian increments to compute the conditional expectations deriving from the discretization of the local evolution of the FBSDE.

The algorithm associated to (4.3.11) and to (4.3.12) takes the following form:

Algorithm 3.1 *For a given sequence of spatial grids $(\mathcal{C}_k)_{0\le k\le N}$, we set:*

$$\begin{aligned}
&\forall x \in \mathcal{C}_N,\ \overline{u}(T,x) \equiv H(x),\ \overline{v}(T,x) \equiv \nabla_x H(x)\sigma(x,H(x)),\\
&\forall k \in \{0,...,N-1\},\ \forall x \in \mathcal{C}_k,\\
&\quad \mathcal{T}(t_k,x) \equiv b(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x))h + \sigma(x,\overline{u}(t_{k+1},x))g(\Delta B^k),\\
&\quad \overline{u}(t_k,x) \equiv \mathbb{E}\big[\overline{u}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}(t_k,x))\big)\big] + hf(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x)),\\
&\quad \overline{v}(t_k,x) \equiv h^{-1}\mathbb{E}\big[\overline{u}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}(t_k,x))\big)g(\Delta B^k)\big].
\end{aligned}$$

Up to the choice of the underlying grids, this form follows heuristics given in Subsection 3.1: the coefficients of the local transition $\mathcal{T}(t_k,x)$ are expressed in function of the former approximations $\overline{u}(t_{k+1},.)$ and $\overline{v}(t_{k+1},.)$.

For technical reasons detailed in Section 9 we consider for the convergence analysis a slightly different version of the above algorithm. Namely, we need to change, at a given time t_k, the discretization of b and f and in particular to replace $\overline{v}(t_{k+1},.)$ by a new predictor. Concerning the driver of the BSDE, we replace $f(x,\overline{u}(t_{k+1},x),\overline{v}(t_{k+1},x))$ by $f(x,\overline{u}(t_{k+1},x),\overline{v}(t_k,x))$: the definition of $\overline{v}(t_k,x)$ does not involve $\overline{u}(t_k,x)$.

The story is rather different for b. Indeed, the definition of $\overline{v}(t_k,x)$ relies on the choice of the underlying transition. In particular, putting $\overline{v}(t_k,x)$ in b as done in f would lead to an implicit scheme.

Nevertheless, for a given intermediate predictor $\hat{v}(t_k,.)$ of $v(t_k,.)$, we can put:

$$\mathcal{T}(t_k,x) \equiv b\big(x,\overline{u}(t_{k+1},x),\hat{v}(t_k,x)\big)h + \sigma\big(x,\overline{u}(t_{k+1},x)\big)g(\Delta B^k).$$

The whole difficulty is then hidden in the choice of $\hat{v}(t_k,x)$. Our strategy consists in choosing $\hat{v}(t_k,x)$ as the expectation of $\overline{v}(t_{k+1},.)$ with respect to the transition $\mathcal{T}^0(t_k,x) \equiv \sigma(x,\overline{u}(t_{k+1},x))g(\Delta B^k)$. This transition differs from $\mathcal{T}(t_k,x)$ in the drift b and leads to an explicit scheme. Namely, we set:

$$\hat{v}(t_k,x) \equiv \mathbb{E}\big[\overline{v}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}^0(t_k,x))\big)\big]. \tag{4.3.13}$$

The predictor $\hat{v}(t_k,.)$ in (4.3.13) appears as a "regularized" version of $\overline{v}(t_{k+1},.)$. Thanks to a Gaussian change of variable, the laws of the underlying transitions $\mathcal{T}^0(t_k,x)$ and $\mathcal{T}(t_k,x)$ are compared in Subsection 7.3.

Final Algorithm.

Algorithm 3.2 *The final algorithm writes:*

$$\begin{aligned}
&\forall x \in \mathcal{C}_N,\ \overline{u}(T,x) \equiv H(x),\ \overline{v}(T,x) \equiv \nabla_x H(x)\sigma(x,H(x)),\\
&\forall k \in \{0,...,N-1\},\ \forall x \in \mathcal{C}_k,\\
&\quad \mathcal{T}^0(t_k,x) \equiv \sigma\big(x,\overline{u}(t_{k+1},x)\big)g(\Delta B^k)\\
&\quad \hat{v}(t_k,x) \equiv \mathbb{E}\big[\overline{v}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}^0(t_k,x))\big)\big],\\
&\quad \mathcal{T}(t_k,x) \equiv b\big(x,\overline{u}(t_{k+1},x),\hat{v}(t_k,x)\big)h + \sigma\big(x,\overline{u}(t_{k+1},x)\big)g(\Delta B^k),\\
&\quad \overline{v}(t_k,x) \equiv h^{-1}\mathbb{E}\big[\overline{u}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}(t_k,x))\big)g(\Delta B^k)\big],\\
&\quad \overline{u}(t_k,x) \equiv \mathbb{E}\big[\overline{u}\big(t_{k+1},\Pi_{k+1}(x+\mathcal{T}(t_k,x))\big)\big] + f\big(x,\overline{u}(t_{k+1},x),\overline{v}(t_k,x)\big)h.
\end{aligned}$$

A Discrete Probabilistic Representation. Note that Algorithm 3.2 provides a discretization procedure of the FBSDE (E) just like $(\mathcal{E})$ provides an analytical counterpart to the stochastic system of equations (E). Consider to this end an initial condition $x_0 \in \mathcal{C}_0$ and define a Markov process on the grids $(\mathcal{C}_k)_{0\le k\le N}$ according to the transitions $(\mathcal{T}(t_k,x))_{k\in\{0,...,N-1\},x\in\mathcal{C}_k}$:

$$X_0 \equiv x_0,\ \forall k \in \{0,...,N-1\},\ X_{t_{k+1}} \equiv \Pi_{k+1}\big(X_{t_k} + \mathcal{T}(t_k,X_{t_k})\big). \tag{4.3.14}$$

Referring to the connection between U and (V,W), see e.g. (4.2.1), put now:

$$\forall k \in \{0,...,N\},\ Y_{t_k} \equiv \overline{u}(t_k,X_{t_k}),\ Z_{t_k} \equiv \overline{v}(t_k,X_{t_k}). \tag{4.3.15}$$

Note that Y and Z are correctly defined since X_{t_k} belongs to the grid $\mathcal{C}_k$. The couple (Y,Z) appears as a discrete version of the couple (V,W) in (E). More precisely, one can prove the following discrete Feynman-Kac formula (see Proposition 4.6.1 for a precise statement):

$$\forall 0 \le k \le N-1,\ Y_{t_k} = \mathbb{E}\Big[H(X_{t_N}) + h\sum_{i=k+1}^{N} f\big(X_{t_{i-1}},\overline{u}(t_i,X_{t_{i-1}}),Z_{t_{i-1}}\big)|\mathcal{F}_{t_k}\Big]. \tag{4.3.16}$$

Note anyhow that the process Z does not appear as the martingale part of the process Y. However, thanks to the martingale representation theorem, there exists a progressively measurable process $\overline{Z}$, with finite moment of order two, such that:

$$Y_{t_N} + h\sum_{i=1}^{N} f\big(X_{t_{i-1}},\overline{u}(t_i,X_{t_{i-1}}),Z_{t_{i-1}}\big) = Y_0 + \int_0^{t_N} \overline{Z}_s dB_s. \tag{4.3.17}$$

Of course, the process $\overline{Z}$ does not match exactly the process Z. However, for a given $k \in \{0,...,N-1\}$, it is readily seen from the above expression that the best $\mathcal{F}_{t_k}$-measurable approximation of $(\overline{Z}_s)_{s\in[t_k,t_{k+1}]}$ in $L^2([t_k,t_{k+1}]\times\Omega, ds\otimes d\mathbb{P})$ is given by:

$$h^{-1}\mathbb{E}\big[Y_{t_{k+1}}\Delta B^k\big].$$

Up to the quantization procedure, this term coincides with $\overline{v}(t_k,X_{t_k})$. In other words, the processes Z and $\overline{Z}$ may be considered as close.

3.4 Choice of the Grids

As indicated at the end of Subsection 3.1, it remains to precise the choice of the grids. Because of the strong coupling, little is *a priori* known on the behaviour of the paths of the forward process. Hence, we can not compute a kind of optimal grid for X. The most natural choice turns out to be the one of Cartesian grids.

Unbounded Cartesian Grids. Two different choices of grids are conceivable. First, we can treat the case of infinite Cartesian grids:

$$\forall k \in \{0,...,N\},\ \mathcal{C}_k \equiv \mathcal{C}_\infty,\ \mathcal{C}_\infty \equiv \delta\mathbb{Z}^d, \tag{4.3.18}$$

where $\delta > 0$ denotes a spatial discretization parameter. In this case, the projection mapping takes the following form:

$$\forall x \in \mathbb{R}^d,\ \Pi_\infty(x) \equiv \sum_{y \in \mathcal{C}_\infty} \Big[y \prod_{j=1}^{d} \mathbf{1}_{[-\delta/2,\delta/2[}(x_j - y_j) \Big]. \tag{4.3.19}$$

In other words, for every $j \in \{1, ..., d\}$, the coordinate j of $\Pi_\infty(x)$ is given by $(\Pi_\infty(x))_j = \delta \lfloor \delta^{-1} x_j + 1/2 \rfloor$.

This choice actually simplifies the convergence analysis and allows a direct comparison with the results from the existing literature, see Douglas et *al.* [DMP96]. Note however that it does not provide a fully implementable scheme since the set $\mathcal{C}_\infty$ is infinite.

Truncated Grids. We now discuss the case of truncated grids. Actually, several truncation procedures may be considered, but all need to take into account the specific geometry of a non-degenerate diffusion, or more simply, of the Brownian motion. Set for example, for a given $R > 0$, $\mathcal{C}_0 \equiv \mathcal{C}_\infty \cap \Delta_0$, where:

$$\Delta_0 \equiv \big\{ x \in \mathbb{R}^d,\ \forall 1 \le j \le d,\ -\delta \lfloor R\delta^{-1} \rfloor - \delta/2 \le x_j < \delta \lfloor R\delta^{-1} \rfloor + \delta/2 \big\}. \tag{4.3.20}$$

The particular choice of the bounds in the definition of Δ_0 ensures that for all $x \in \mathbb{R}^d$, $\Pi_\infty(x) \in \mathcal{C}_0 \Leftrightarrow x \in \Delta_0$.

Due to the drift part and to the diffusive part of the forward process, it is clear that we need to enlarge the spatial grids as time increases. To this end, fix $\rho > 0$ and define, for every $i \in \{1, ..., N\}$, the truncated grid $\mathcal{C}_i \equiv \mathcal{C}_\infty \cap \Delta_i$, where:

$$\Delta_i \equiv \big\{ x \in \mathbb{R}^d,\ \forall 1 \le j \le d,\ -\delta \lfloor (R+\rho)\delta^{-1} \rfloor - \delta/2 \le x_j < \delta \lfloor (R+\rho)\delta^{-1} \rfloor + \delta/2 \big\}. \tag{4.3.21}$$

Note in this way that the size of the grid $\mathcal{C}_i$, i.e. of the grid at time t_i, does not depend on t_i itself. In other words, the Hölder regularity of the paths of the Brownian motion does not interact with the definition of the grids. To take into account these pathwise properties, the following type of grids could also be used:

$$\Delta_i \equiv \big\{ x \in \mathbb{R}^d,\ \forall 1 \le j \le d,\ -\delta \lfloor (R+\rho t_i^\alpha)\delta^{-1} \rfloor - \delta/2 \le x_j < \delta \lfloor (R+\rho t_i^\alpha)\delta^{-1} \rfloor + \delta/2 \big\},$$

for a given $0 < \alpha < 1/2$. With this definition, the number of points involved in the discretization procedure is smaller. However, since the proof of the convergence of the algorithm is far from being trivial, we prefer, for the sake of simplicity, to keep the first definition of the truncated grids. Hence, for every $i \in \{0, ..., N\}$, Π_i writes:

$$\begin{aligned} \forall 0 \le i \le N, \forall x \in \Delta_i,\ \Pi_i(x) &\equiv \mathcal{Q}\big(R+\rho, \Pi_\infty(x)\big) \equiv \Pi_\infty(x), \\ \forall 1 \le i \le N, \forall x \notin \Delta_i,\ \Pi_i(x) &\equiv \mathcal{Q}\big(R+\rho, \Pi_\infty(x)\big), \\ \forall x \notin \Delta_0,\ \Pi_0(x) &\equiv \mathcal{Q}\big(R, \Pi_\infty(x)\big), \end{aligned} \tag{4.3.22}$$

where for a given $(r, y) \in \mathbb{R}^{+*} \times \mathbb{R}^d$, $\mathcal{Q}(r, y)$ denotes the orthogonal projection of y on the hypercube $[-\delta \lfloor r\delta^{-1} \rfloor, \delta \lfloor r\delta^{-1} \rfloor]^d$:

$$\mathcal{Q}(r, y) \equiv \big((y_i \vee (-\delta \lfloor r\delta^{-1} \rfloor)) \wedge (\delta \lfloor r\delta^{-1} \rfloor) \big)_{1 \le i \le d}.$$

Note finally that R is fixed by the reader once for all in function of the set on which u has to be approximated at the initial time. At the opposite, ρ appears as a discretization parameter chosen by the reader in function of the required precision and of the affordable complexity for Algorithm 3.2.

4 Convergence Results

This section is devoted to the convergence analysis of $\overline{u}$ to u. As stated in the following theorem, which is the main result of the paper, five different types of errors can be distinguished:

Theorem 4.4.1 *Let $p \geq 2$. There exist two constants $c_{4.4.1}$ and $C_{4.4.1}$, only depending on p, T and on known parameters appearing in* **(A)***, such that for $h < c_{4.4.1}$, $\delta^2 < h$, $M^{-2/d} < h$ and $\rho \geq 1$:*

$$\sup_{x \in \mathcal{C}_0} |u(0,x) - \overline{u}(0,x)|^2 \leq C_{4.4.1}\big[\mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{trunc}) + \mathcal{E}^2(\text{quantiz}) + \mathcal{E}^2(\text{gradient}, p)\big],$$

with:

$$\mathcal{E}(\text{time}) \equiv h^{1/2},\ \mathcal{E}(\text{space}) \equiv h^{-1}\delta,\ \mathcal{E}(\text{trunc}) \equiv R/(R+\rho),\ \mathcal{E}(\text{quantiz}) \equiv h^{-1/2}M^{-1/d},$$
$$\mathcal{E}(\text{gradient}, p) \equiv h^{p/2+d/4-1/2}M^{-p/d}\delta^{-p-d/2}$$

Remark 4.4.1 *The FBSDE counterpart of Theorem 4.4.1 is given in Subsection 4.3: see Theorems 4.4.2 and 4.4.3. Since additional notations are needed to detail these results in a relevant form, we prefer, for the sake of clarity, to postpone their statement (and thus the definition of the associated probabilistic tools) to a different part.*

4.1 Classification of Errors

We now detail the meaning of the different errors appearing in Theorem 4.4.1:

Temporal Discretization Error. $\mathcal{E}(\text{time})$ denotes the temporal discretization error due to the discretization of $[0,T]$ along a regular mesh. In a nutshell, the $1/2$ exponent appearing in the definition of $\mathcal{E}(\text{time})$ corresponds to the Hölder regularity of u and $\nabla_x u$ in time and to the $L^2(\mathbb{P})$ $1/2$-Hölder property of the Brownian increments.

Spatial Discretization Error. $\mathcal{E}(\text{space})$ denotes the spatial discretization error. This quantity highly depends on the ratio between the spatial and the temporal steps. This connection between δ and h can be explained as follows: the drift part of the transitions $(\mathcal{T}(t_k,.))_{0\leq k\leq N}$ is of order h and the diffusive one is of order $h^{1/2}$. Thus, to take into account the influence of the drift at the local level, the spatial discretization parameter must be smaller than h. In other words, δh^{-1} must be small.

Quantization Error. $\mathcal{E}(\text{quantiz})$ denotes the error due to the quantization procedure of the Brownian increments. This error depends on the ratio between the distortion and the temporal step. The quantity $\mathcal{E}(\text{quantiz})$ represents the typical bound between $\overline{v}(t_k,x)$ and the best constant approximation of the process $(\overline{Z}_s)_{s\in[t_k,t_{k+1}]}$, i.e. between $\overline{v}(t_k,x)$ and:

$$h^{-1}\mathbb{E}\big[\overline{u}\big(t_{k+1}, \Pi_{k+1}(x + \mathcal{T}(t_k,x))\big)\Delta B^k\big].$$

Note indeed that the distance between ΔB^k and $g(\Delta B^k)$ is of order $h^{1/2}M^{-1/d}$, see (4.3.10). Since the underlying expectation is divided by h^{-1}, this leads to a term in $h^{-1/2}M^{-1/d}$.

Truncation Error. $\mathcal{E}(\text{trunc})$ denotes the error associated to the truncation procedure. As written in Theorem 4.4.1, it depends on R and ρ, where R denotes the radius of the initial grid $\mathcal{C}_0$ and $R+\rho$ the radius of the grids $(\mathcal{C}_k)_{1\leq k\leq N}$. If ρ tends to $+\infty$, i.e. if the grids are not truncated, this error term reduces to zero.

Generally speaking, $\mathcal{E}(\text{trunc})$ appears as the Bienaymé-Chebychev estimate of the probability that the approximating process X stays inside the grids $(\mathcal{C}_k)_{0\leq k\leq N}$. The lack of relevant estimates of the discretized version of the drift b (recall that the function b is not bounded), and more specially of the discretized gradient $\overline{v}$ explains the reason why the Bienaymé-Chebychev estimate applies in this framework and not better ones (as Bernstein inequality). We also recall that the unboundedness of the coefficients is the most common case in the applications, see e.g. Burgers equation in Section 5.

Gradient Error. $\mathcal{E}(\text{gradient}, p)$ denotes an extra error generated by the lack of estimates of the discretized gradient $\overline{v}$. This term follows from the specific choice of the predictor $\hat{v}$ made in Subsection 3.3 and appears in the second step of the proof of Theorem 4.4.1, see more precisely Subsections 7.1 and 7.3.

The convergence of $\mathcal{E}(\text{gradient}, p)$ towards 0 relies on the term $h^{p/2}M^{-p/d}\delta^{-p}$, M being chosen large enough and p as large as necessary. In short, this reduced form represents the probability that the distance between the Gaussian increment and its quantization exceeds the spatial step δ. Note indeed from (4.3.10) that for every $p \geq 2$:

$$\mathbb{P}\{|\Delta B^k - g(\Delta B^k)| > \delta\} \leq C_{\mathrm{Quantiz}}(p,d) h^{p/2} M^{-p/d} \delta^{-p}. \tag{4.4.1}$$

Thus, the error term $\mathcal{E}(\mathrm{gradient}, p)$ depends on the ratio between the spatial discretization step and the quantization distortion of the underlying Gaussian increments.

The probability (4.4.1) appears in the control of the distance between the predictor $\hat{v}$ and the true gradient v. In this frame, the strategy consists in writing the predictor $\hat{v}$ as an expectation with respect to the Gaussian kernel and not to its quantized version. Generally speaking, this strategy holds when the quantized transition $\mathcal{T}(t_k, x)$ and its Gaussian counterpart belong to the same cell of the spatial grid, i.e. when the distance between the Brownian increment and the quantized one is of the same order as the length of a given cell. Since the spatial grid step is given by δ, we then need to control the probability that the difference between the increments exceeds δ.

Of course, when b does not depend on z, there is no reason to define $\hat{v}$. In such a case, $\mathcal{E}(\mathrm{gradient}, p)$ reduces to 0.

4.2 Comments on the Rate of Convergence

Error in function of h***.*** To detail in a more explicit way the rate of convergence given by Theorem 4.4.1, we give an example in which ρ ($\rho < +\infty$), δ and M are expressed as powers of h. Assume indeed that ρ, δ and M are chosen in the following way:

$$\rho = Rh^{-1/2},\ \delta \equiv h^{1+\alpha},\ M^{-2/d} \equiv h^{1+\beta}, \alpha, \beta \geq 0.$$

In such a case:

$$\mathcal{E}(\mathrm{gradient}, p) = h^{p/2+d/4-1/2} M^{-p/d} \delta^{-p-d/2} = \exp\big[\ln(h)\big[p(\beta/2 - \alpha) - (d/2 + 1 + \alpha d)/2\big]\big].$$

To ensure the convergence of the algorithm we then need to choose:

$$p(\beta/2 - \alpha) - (d/2 + 1 + \alpha d)/2 > 0 \Longleftrightarrow \beta > 2\alpha + (1/p)(d/2 + 1 + \alpha d).$$

Put finally $\beta = 2\alpha + (d/2 + 1 + \alpha d)/p + \eta$, $\eta > 0$. The rate of convergence of the fully implementable algorithm is given by:

$$\sup_{x \in \mathcal{C}_0} |u(0,x) - \overline{u}(0,x)|^2 \leq C_{4.4.1}\big[h + h^{2\alpha} + h^{\beta} + h^{p\eta}\big].$$

Taking $\alpha = 1/2$ and $\eta = 1/p$ then yields:

$$\sup_{x \in \mathcal{C}_0} |u(0,x) - \overline{u}(0,x)|^2 \leq C_{4.4.1} h.$$

In particular, for p large enough, the exponent β is close to 1 and the number M of points needed to quantify the Brownian increments is close to h^{-d}. Here is the limit of the method: for a large d and a small h, we need a rather large number of points for the Gaussian quantization. Recall anyhow that the Gaussian grids are computed once for all. Thus, the numerical effort to get sharp quantization grids can be made apart from our Algorithm.

Estimates of $\nabla_x u$***.*** The reader might wonder about the estimate of the gradient of u. Note in this framework that two strategies are conceivable.

First, the probabilistic counterpart of Theorem 4.4.1 given in Subsection 4.3 provides an L^2 estimate of the distance between $\overline{v}$ and the gradient of the true solution. Note however that the underlying L^2 norm is taken with respect to the distribution of the discrete process X (cf. (4.3.14)).

To get a joint estimate of the solution and of its gradient with respect to the *supremum* norm, the reader can apply the following strategy: differentiate if possible the PDE ($\mathcal{E}$) and apply, once again if possible, Algorithm 3.2 to $(u, \nabla_x u)$ seen as the solution of a system of parabolic quasi-linear PDEs. Such a strategy is applied in Section 5 to the solution of the porous media equation and to its gradient. Note that this approach coincides with the one followed by Douglas et *al.* [DMP96].

4.3 Estimates of the Discrete Processes

We now translate Theorem 4.4.1 in a more probabilistic way. Recall indeed that, in several situations (e.g. in financial mathematics), the knowledge of the triple (U, V, W) is as crucial as the knowledge of the couple $(u, \nabla_x u)$.

Recall from (4.3.14), (4.3.15) and (4.3.16) that the discrete process (X, Y, Z) given by:

$$\begin{aligned}
&X_0 \equiv x_0,\ \forall k \in \{0, ..., N-1\},\ X_{t_{k+1}} \equiv \Pi_{k+1}\big(X_{t_k} + \mathcal{T}(t_k, X_{t_k})\big),\\
&\forall k \in \{0, ..., N\},\ Y_{t_k} \equiv \overline{u}(t_k, X_{t_k}),\ Z_{t_k} \equiv \overline{v}(t_k, X_{t_k}),
\end{aligned}$$

provides a discretization of (U, V, W).

We then prove that (X, Y, Z) and (U, V, W) get closer in a suitable sense as h, δ, M^{-1} and ρ^{-1} vanish. Note however that we are not able to prove that the distance between (X, Y, Z) and (U, V, W) over the whole interval $[0, T]$ tends to zero. Indeed, since the projections $(\Pi_i)_{0 \le i \le N}$ map every point outside the sets $(\Delta_i)_{0 \le i \le N}$ onto the boundaries of $(\mathcal{C}_i)_{0 \le i \le N}$ (see e.g. (4.3.22)), we do not control efficiently the transition of the process X after the first hitting time of the boundaries of the grids by X. It is then well understood that we have to stop the triple (X, Y, Z) at this first hitting time. Put to this end:

$$\tau_\infty \equiv \inf\big\{t_k,\ 1 \le k \le N,\ X_{t_{k-1}} + \mathcal{T}(t_{k-1}, X_{t_{k-1}}) \notin \Delta_k\big\},\ \inf(\emptyset) = +\infty. \tag{4.4.2}$$

First, as a bypass product of the proof of Theorem 4.4.1, the function $\overline{v}$ provides an approximation of v in the following L^2 sense:

Theorem 4.4.2 *Let $p \ge 2$. Then, there exist two constants $c_{4.4.2}$ and $C_{4.4.2}$, only depending on p and on known parameters appearing in* **(A)**, *such that for $h < c_{4.4.2}$, $\delta^2 < h$, $M^{-2/d} < h$ and $\rho \ge 1$:*

$$\begin{aligned}
&h \sum_{i=0}^{N-1} \mathbb{E}\big[|\overline{v}(t_i, X_{t_i})\mathbf{1}_{\{t_i < \tau_\infty\}} - v(t_i, X_{t_i})|^2\big]\\
&\quad \le C_{4.4.2}\big[\mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz}) + \mathcal{E}^2(\text{trunc}) + \mathcal{E}^2(\text{gradient}, p)\big].
\end{aligned}$$

Moreover, the triple (X, Y, Z) stopped at time τ_∞ satisfies:

Theorem 4.4.3 *Let $p \ge 2$. Then, there exist two constants $c_{4.4.3}$ and $C_{4.4.3}$, only depending on p and on known parameters appearing in* **(A)**, *such that for $h < c_{4.4.3}$, $\delta^2 < h$, $M^{-2/d} < h$ and $\rho \ge 1$:*

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{i \in \{0,\dots,N\}} |X_{t_i \wedge \tau_\infty} - U_{t_i}|^2\Big] + \mathbb{E}\Big[\sup_{i \in \{0,\dots,N\}} |Y_{t_i \wedge \tau_\infty} - V_{t_i}|^2\Big] + h \sum_{i=0}^{N-1} \mathbb{E}\big[|Z_{t_i}\mathbf{1}_{\{t_i < \tau_\infty\}} - W_{t_i}|^2\big]\\
&\quad \le C_{4.4.3}\big[\mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz}) + \mathcal{E}^2(\text{trunc}) + \mathcal{E}^2(\text{gradient}, p)\big].
\end{aligned}$$

5 Numerical Examples

In this section, we illustrate the behaviour of the algorithm with numerical examples. To this end, we choose equations that admit an explicit solution. This permits to compare the results obtained with our algorithm to a reference value. In this frame, we focus on three examples: the one-dimensional Burgers equation, the deterministic KPZ equation in dimension two and the one-dimensional porous media equation.

5.1 One Dimensional Burgers Equation

Consider first the backward Burgers equation:

$$\begin{aligned}
&\partial_t u(t,x) - (u\partial_x u)(t,x) + \frac{\varepsilon^2}{2}\partial^2_{x,x} u(t,x) = 0, (t,x) \in [0,T[\times\mathbb{R},\ \varepsilon > 0,\\
&u(T,x) = H(x),\ x \in \mathbb{R},\ H \in C_b^{2+\alpha}(\mathbb{R}),\ \alpha \in]0,1[.
\end{aligned} \tag{4.5.1}$$

Using a non-linear transformation, one can derive an explicit expression of the solution of (4.5.1). This is known as the Cole-Hopf factorization, see Whitham [Whi73], Chapter IV, or Woyczyński [Woy98], Chapter III, for details. The solution of (4.5.1) then writes:

$$\forall (t,x) \in [0,T] \times \mathbb{R}, \ u(t,x) = \frac{\mathbb{E}[H(x+\varepsilon B_{T-t})\phi(x+\varepsilon B_{T-t})]}{\mathbb{E}[\phi(x+\varepsilon B_{T-t})]}, \tag{4.5.2}$$

where B is a standard Brownian motion and $\forall y \in \mathbb{R}, \ \phi(y) \equiv \exp\left(-\varepsilon^{-2}\int_0^y H(u)du\right)$.

From the explicit representation (4.5.2), we can derive numerically, using e.g. a Riemann sum, a Monte-Carlo method, or a quantized version of the expectation (4.5.2), a reference solution to test the algorithm.

The reader may object that Burgers equation is actually semi-linear and not quasi-linear. Actually, it depends on whether we consider the non-linear term as a drift or as a second member. We describe below the algorithms associated to these two points of view, even if the coupled case is the only one to fulfill Assumption **(A)**.

Moreover, in the forward-backward representation of Burgers equation, the estimation procedure of the gradient is not necessary to compute the approximate solution $\overline{u}$. Numerically, this case turns out to be the most robust. Finally, in both cases, the intermediate predictor $\hat{v}$ is useless: in the coupled case, the drift of the diffusion U reduces to V (and thus does not depend on W), and in the decoupled one, the drift vanishes.

Explicit Expression of the Algorithms

For a given final condition $H \in C_b^{2+\alpha}(\mathbb{R}), \ \alpha \in]0,1[$, we write:

Algorithm 5.1 (Coupled case)

$\forall x \in \mathcal{C}_N, \ \overline{u}(T,x) \equiv H(x),$

$\forall k \in \{0,...,N-1\}, \ \forall x \in \mathcal{C}_k, \ \overline{u}(t_k,x) \equiv \mathbb{E}[\overline{u}(t_{k+1}, \Pi_{k+1}(x - \overline{u}(t_{k+1},x)h + \varepsilon g(\Delta B^k)))],$

$\forall k \in \{0,...,N-1\}, \ \forall x \in \mathcal{C}_k, \ \overline{v}(t_k,x) \equiv h^{-1}\mathbb{E}[\overline{u}(t_{k+1}, \Pi_{k+1}(x - \overline{u}(t_{k+1},x)h + \varepsilon g(\Delta B^k)))g(\Delta B^k)].$

Algorithm 5.2 (Pure Backward case)

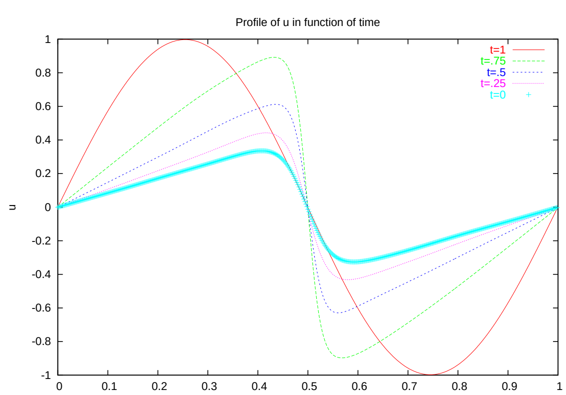
$\forall x \in \mathcal{C}_N, \ \overline{u}(T,x) \equiv H(x),$

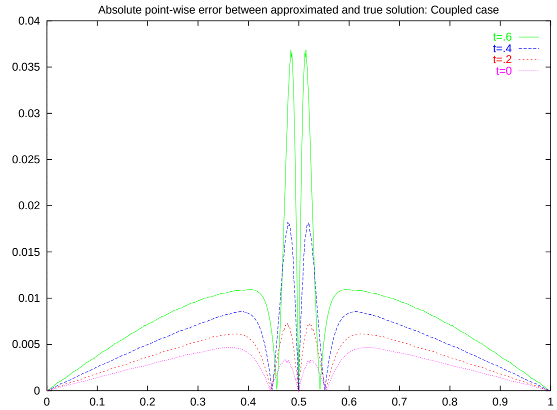
$\forall k \in \{0,...,N-1\}, \ \forall x \in \mathcal{C}_k, \ \overline{u}(t_k,x) \equiv \mathbb{E}[\overline{u}(t_{k+1}, \Pi_{k+1}(x + \varepsilon g(\Delta B^k)))] - h\varepsilon^{-1}\overline{u}(t_{k+1},x)\overline{v}(t_k,x),$

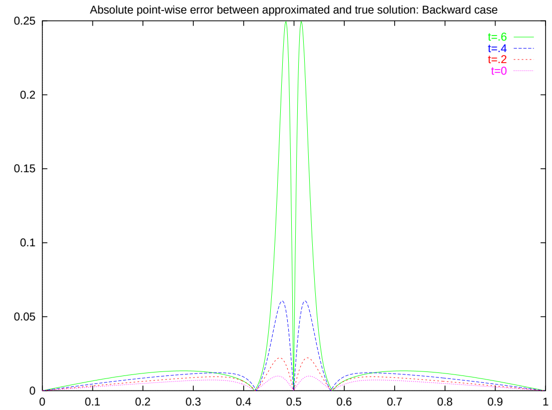
$\forall k \in \{0,...,N-1\}, \ \forall x \in \mathcal{C}_k, \ \overline{v}(t_k,x) \equiv h^{-1}\mathbb{E}[\overline{u}(t_{k+1}, \Pi_{k+1}(x + \varepsilon g(\Delta B^k)))g(\Delta B^k)].$

Numerical Results

In order to avoid first to truncate the grids, we choose a periodic initial solution. Put to this end $H(x) = \sin(2\pi x)$ and derive from (4.5.2) that u is 1-periodic. This allows to define $\overline{u}(t_k,.)$ on $\mathcal{C}_\infty$ by setting $\forall x \in \mathcal{C}_\infty, \ \overline{u}(t_k,x) \equiv \overline{u}(t_k, x - \lfloor x \rfloor)$. Hence, we can set $\mathcal{C}_k \equiv \mathcal{C}_\infty$ for $k \in \{0,...,N-1\}$. For $T = 1, \delta = 10^{-3}, h = .01, M = 160, \varepsilon = .15$, we present below the results of the previous algorithms. The explicit solution given by (4.5.2) is approximated by quantization techniques with a 500 points grid. We plot below some profiles of the reference value for various discretization times as well as the pointwise absolute error between this reference solution and the approximations obtained with our algorithms.







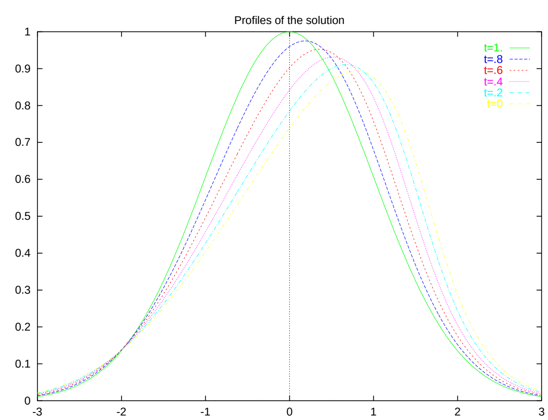
On the profiles of the explicit solution, the abscisses of the peaks of the initial sinusoidal wave are going closer to each other up to a given time t_0. This is a typical shocking wave behaviour. Because of the viscosity, i.e. ε is non zero, there is no shock and the amplitude of the wave decays when t goes to zero.

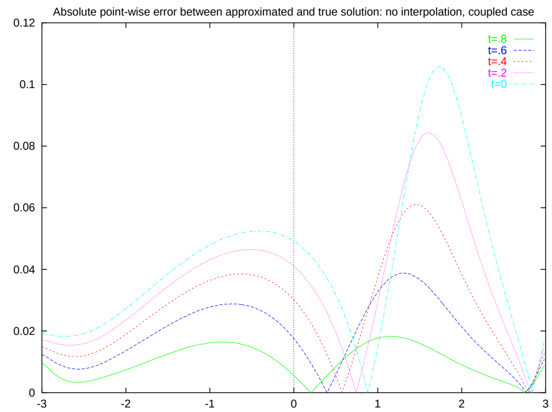
From a numerical point of view, the coupled case provides several advantages. First, the convergence of Algorithm 5.1 does not rely on the discretization procedure of the gradient. In short, there is no reason to update the gradient in order to obtain the approximate solution with the first algorithm. The computation of $\overline{v}$ just provides in this case an L^2 estimate of the gradient. At the opposite, this computation is necessary in Algorithm 5.2.

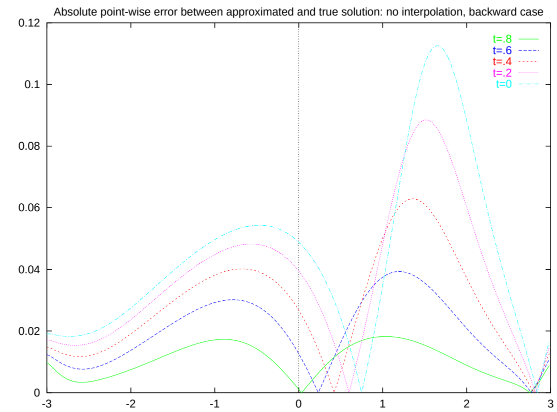
Moreover, since the coefficient $f(y,z) = \varepsilon^{-1}yz$ is not globally Lipschitz in the pure backward case, it is then another story to establish the convergence of Algorithm 5.2.

These theoritical remarks are confirmed by the above pictures. Even though Algorithm 5.2 does not behave too poorly, it is still less precise than Algorithm 5.1. The factor between the absolute pointwise errors of the two algorithms is approximately 5.

Truncation error. We now illustrate the effects of truncation and deal with a non periodic final data. Namely, we take $H(x) = \exp(-x^2/2)$, $T = 1$, $h = .02$, $\rho = 3$, $\delta = \rho/500$, $M = 250$. The reference value, see profiles below, is computed from the Cole-Hopf explicit solution by quantization techniques with a 500 points grid. We run Algorithm 3.2 with the previous parameters to obtain:



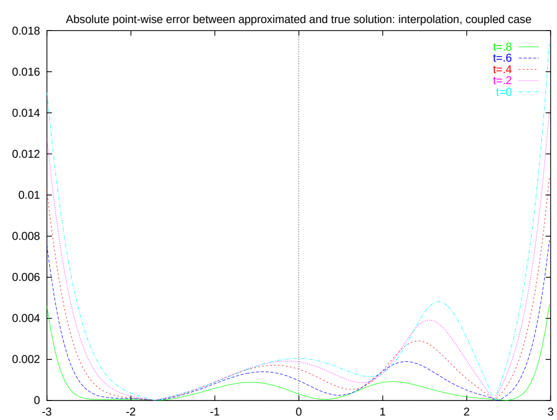


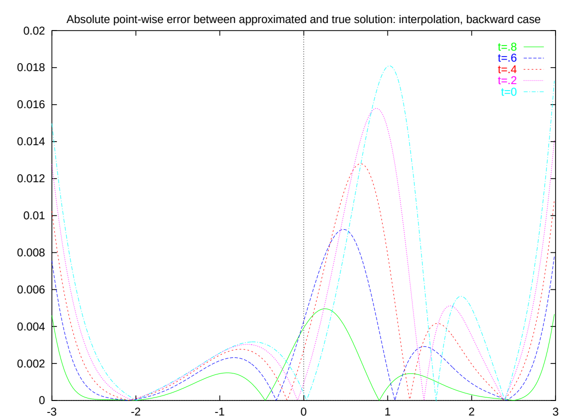


Choose now $R = 1$: the expected truncated error $\mathcal{E}(\text{trunc})$ is given by .25 whereas the absolute point-wise error between both solutions is bounded by .05 on $[-1, 1]$. This emphasizes the difficulty to control the truncation procedure in our algorithm. There are two possible arguments to explain this difference between .25 and .05. First, as explained in Subsection 9.1, our way to estimate $\mathcal{E}(\text{trunc})$ is suitable for unbounded drifts b, and more particulary for drifts depending on the gradient. In our case, the drift is bounded (since the solution is bounded by 1), and most relevant estimates could apply. Second, the fast decay of the final condition H may explain the low influence of distant points on the values of the solution on $[-1, 1]$.

Note also that the relative error is close to .1 on $[-1, 1]$. A possible strategy to decrease it would consist in refining the spatial mesh.

We also feel that the choice of the rough projection mappings $(\Pi_k)_{0 \le k \le N}$ deeply affects the global error. To investigate more precisely their influence, we replace them by standard linear interpolation procedures (which are defined in an obvious way since the underlying space is one dimensional). In short, this permits to extend continuously the approximated solution $\overline{u}$ to the whole space. With the same parameters as above, we then get:





Numerically, the interpolation can thus be really relevant to improve the convergence (see Subsection 9.2 for further details and explanations on this point). To obtain the same precision without interpolation we need to refine significantly the parameters (taking e.g. $\delta = 2 \times 10^{-4}$). Let us finally mention that the results obtained with the coupled representation and the linear interpolation are still more accurate than with the backward one.

5.2 Quadratic Backward Equation: Deterministic KPZ Equation

In this subsection, we focus on the so-called "deterministic KPZ" equation (see e.g. Kardar, Parisi and Zhang [KPZ86] and Woyczyński [Woy98], Chapter I, for a physical interpretation):

$$\begin{aligned} &\partial_t u(t,x) + \frac{1}{2}\mathrm{tr}(\sigma\sigma^* \nabla^2_{x,x} u(t,x)) + \frac{\nu}{2}\langle \sigma\sigma^* \nabla_x u(t,x), \nabla_x u(t,x)\rangle = 0,\ (t,x) \in [0,T[\times \mathbb{R}^d, \\ &u(T,x) = H(x),\ x \in \mathbb{R}^d, \end{aligned} \tag{4.5.3}$$

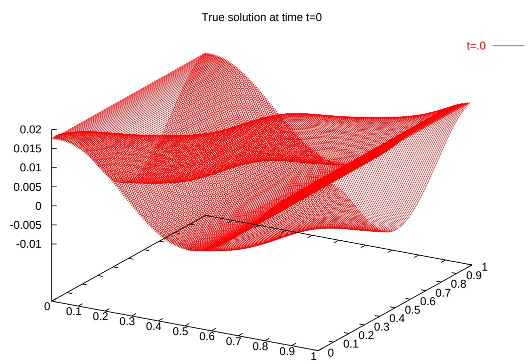
where $\nu \in \mathbb{R}^{+*}$ is a given parameter and σ a given constant matrix such that $\sigma\sigma^*$ is positive definite.

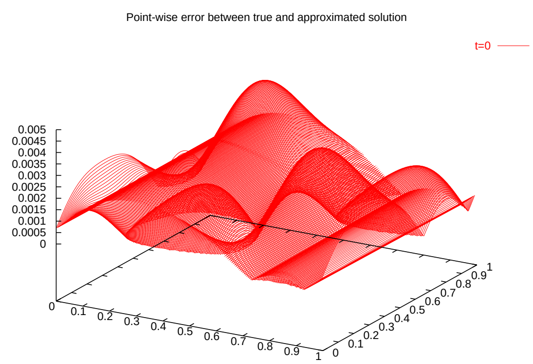
Such an equation admits too a "Cole-Hopf explicit solution", see again [KPZ86], that writes:

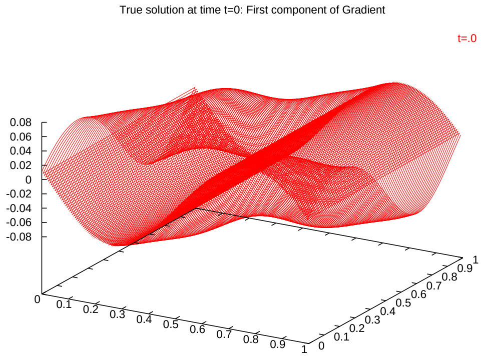
$$u(t,x) = \frac{\log(\mathbb{E}[\exp(\nu H(x + \sigma B_{T-t}))])}{\nu}. \tag{4.5.4}$$

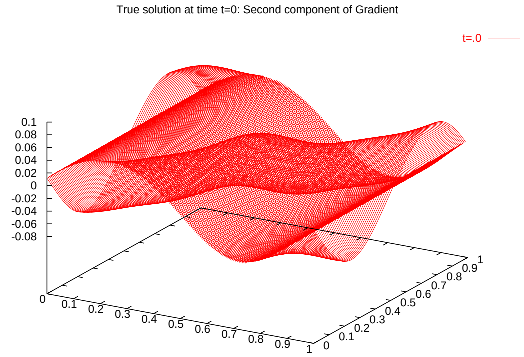
We then apply Algorithm 3.2 to Equation (4.5.3) seen as a true quasi-linear equation (so-called "coupled case" in the former subsection).

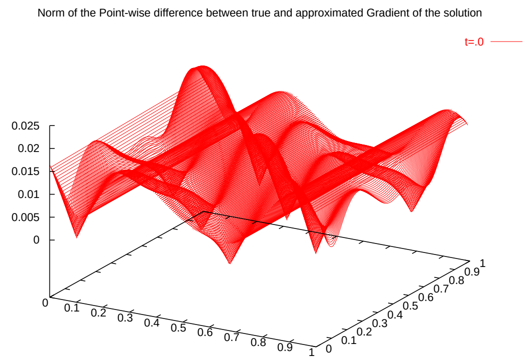
Concerning the initial condition, we choose $H(x) = \prod_{i=1}^d \sin(2\pi x_i)$. By construction, we have $\forall x \in \mathbb{R}^d,\ \forall k \in \mathbb{Z}^d,\ u(t,x+k) = u(t,x)$. Since the solution is periodic, $\overline{u}$ can be defined on the whole grid $\mathcal{C}_\infty$ (see also Paragraph 5.1). We now present the results for $d = 2,\ \nu = .3, T = .5,\ h = .02, \delta = 5 \times 10^{-4}, M = 160$ and $\sigma\sigma^* = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ with $\rho = .8$. The reference value and its gradient have been derived from (4.5.4) using quantization techniques with a 500 points grid. At $t = 0$, one has:











The relative error between the approximate and true solutions is at most .25. The explanation seems rather simple: the explicit solution quickly decays as time decreases. Anyway, we feel that our algorithm manages to catch this specific decreasing phenomenon.

Let us also mention that the last picture represents the pointwise difference of the true and approximated gradients, but the control given by Theorem 4.4.2 just holds in L^2.

5.3 Porous Media Equation

To conclude this section, we focus on the equation (this example is taken from Makarov [Mak03]):

$$\partial_t u(t,x) + (u\partial^2_{x,x}u)(t,x) + (\partial_x u)^2(t,x) + u^2(t,x) = 0,\ (t,x) \in]0,T] \times \mathbb{R},$$
$$u(T,x) = T^{-1}\frac{4}{3}\cos^2(\frac{\pi x}{L}),\ L = 2\sqrt{2}\pi, \tag{4.5.5}$$

which admits the L-periodic explicit solution $u(t,x) = t^{-1}\frac{4}{3}\cos^2(\frac{\pi x}{L})$.

Note that (4.5.5) does not fulfill Assumption **(A)**. In the sequel, we choose without any rigorous justifications to apply Algorithm 3.2 on $[T/2, T]$ (note however for a rough explanation that the quadratic growth of the coefficients ensures that Theorem 4.4.1 holds on a suitable interval $[t, T]$, for t close enough to T, and, in the same way, Theorem 4.2.1 applies away from 0).

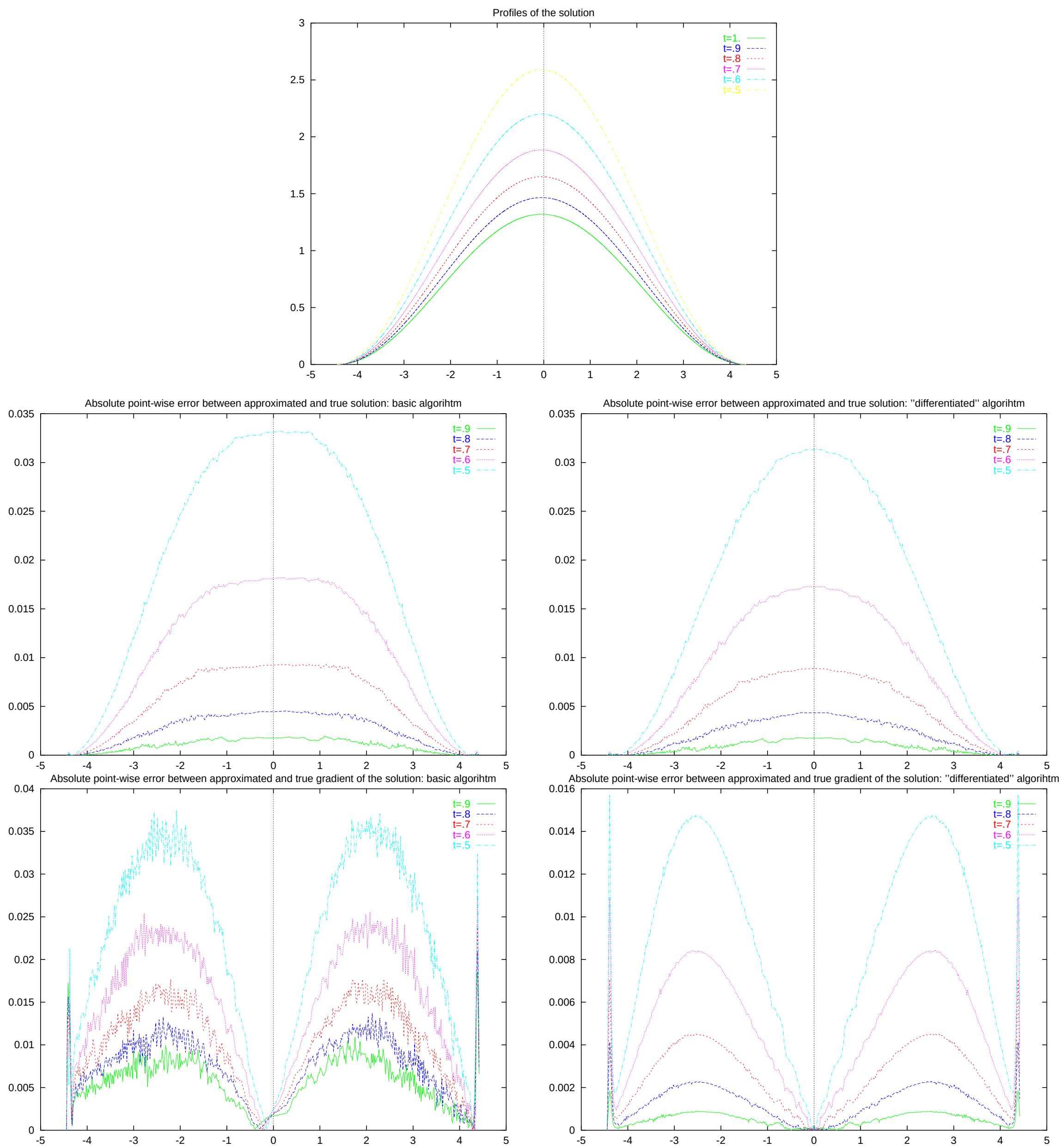
Nevertheless, as explained in Subsection 4.2, this procedure just provides an L^2-estimate of $\nabla_x u$. In this framework, we have decided to apply the so-called “differentiated" approach, described in Subsection 4.2, to obtain a pointwise estimate of $\nabla_x u$ (see Algorithm 5.3 below).

Note finally from the periodicity of u that $\overline{u}$ can be defined on the whole grid $\mathcal{C}_\infty$ as in the previous example (see also Paragraph 5.1).

Algorithm 5.3 (Differentiated Algorithm)

$$\forall x \in \mathcal{C}_N,\ \overline{u}(T,x) = T^{-1}\frac{4}{3}\cos^2(\frac{\pi x}{L}),\ \overline{w}(T,x) = T^{-1}\left(-\frac{8\pi}{3L}\cos(\frac{\pi x}{L})\sin(\frac{\pi x}{L})\right),$$
$\forall k \in \{0, ..., N-1\}, \forall x \in \mathcal{C}_k,$
$$\overline{u}(t_k,x) = \mathbb{E}\big[\overline{u}\big(t_{k+1}, \Pi_{k+1}(x + \overline{w}(t_{k+1},x)h + \sqrt{2\overline{u}(t_{k+1},x)}g(\Delta B^k))\big)\big] + h\overline{u}(t_{k+1},x)^2,$$
$$\overline{w}(t_k,x) = \mathbb{E}\big[\overline{w}\big(t_{k+1}, \Pi_{k+1}(x + 3\overline{w}(t_{k+1},x)h + \sqrt{2\overline{u}(t_{k+1},x)}g(\Delta B^k))\big)\big] + 2h\overline{u}(t_{k+1},x)\overline{w}(t_{k+1},x).$$

For $T = 1, h = .02, \delta = L/500, M = 160$, we present below the results obtained first with Algorithm 3.2 (the approximation of the gradient with this algorithm is undefined at $x = \pm L/2$ and we thus arbitrarily set it to zero) and then with Algorithm 5.3. On $[-L/2, L/2]$ it comes:



We first observe that the approximated solutions obtained with the two algorithms are not significantly different. The main advantage of the differentiated algorithm is, as expected, for the pointwise approximation of the gradient. Indeed, in that case there is a factor 4 between the absolute pointwise errors associated to the two methods. Let us also indicate that both methods present some "singularity" in the neigbourhood of $x = \pm L/2$ for the estimation of the gradient. This could be expected for Algorithm 3.2 since the estimation of the gradient is obtained by dividing $\overline{v}$ by a term that goes to 0 when $x \longrightarrow \pm L/2$. It is a bit more surprising for Algorithm 5.3.

6 Proof, First Step: *A Priori* Controls of the Discrete Objects

In this section, we give various *a priori* estimates of the couple (Y, Z) introduced in (4.3.15) and of the approximate diffusion X defined in (4.3.14). These controls are necessary to establish Theorems 4.4.1, 4.4.2 and 4.4.3.

About Constants. In the following, we keep the same notation C, C_α, c_α (or C', C'_α, c'_α) for all finite, non-negative constants which appear in our computations: they may depend on known parameters in **(A)**, on T and on p, but not on any of the discretization parameters. The index α in the previous notation refers to the numbering of the Proposition, Lemma, Theorem, ... where the constant appears.

Conditions on Parameters. Recall that p denotes in Theorems 4.4.1, 4.4.2 and 4.4.3 a real larger than 2. It is from now on fixed. Furthermore, we assume that the conditions of Theorem 4.4.1 on h, δ, M and ρ are fulfilled. Namely, the statements of the following Propositions and Lemmas hold for h, $h^{-1}\delta^2$, $h^{-1}M^{-2/d}$ and ρ^{-1} small enough even if these conditions are not explicitly written.

6.1 Discrete Backward Equation and Associated *a priori* Estimates

Discrete Feynman-Kac Formula.

Proposition 4.6.1 *With the notations of Algorithm 3.2 and Subsection 4.3, the sequence $(Y_{t_k})_{0\le k\le N}$ satisfies the discrete Feynman-Kac representation:*

$$\forall 0 \le k \le N-1,\ Y_{t_k} = \mathbb{E}\Big[H(X_{t_N}) + h\sum_{i=k+1}^{N} f\big(X_{t_{i-1}}, \overline{u}(t_i, X_{t_{i-1}}), Z_{t_{i-1}}\big)|\mathcal{F}_{t_k}\Big]. \tag{4.6.1}$$

Proof. Recall first that $\overline{u}$ and $\overline{v}$ are bounded (but estimates are not uniform with respect to the parameters): see Subsection 3.1, (4.3.6). In particular, the r.h.s. in the above expression is correctly defined. Note moreover, by definition of X, Y and Z, that for a given $k \in \{0, ..., N-1\}$:

$$Y_{t_k} = \mathbb{E}\Big[Y_{t_{k+1}} + hf\big(X_{t_k}, \overline{u}(t_{k+1}, X_{t_k}), Z_{t_k}\big)|\mathcal{F}_{t_k}\Big].$$

The proof of Proposition 4.6.1 follows by iteration of this last identity. $\square$

Recall now from the Martingale Representation Theorem, that there exists a progressively measurable process $\overline{Z}$, with finite moment of order two, such that:

$$Y_{t_N} + h\sum_{i=1}^{N} f\big(X_{t_{i-1}}, \overline{u}(t_i, X_{t_{i-1}}), Z_{t_{i-1}}\big) = Y_0 + \int_0^{t_N} \overline{Z}_s dB_s. \tag{4.6.2}$$

This representation permits to apply the BSDE machinery to our frame. However, as well-know in the literature devoted to SDEs (or equivalently to PDEs), several *a priori* estimates of the solution are necessary to apply this strategy.

Propositions 4.6.2 and 4.6.3 provide *a priori* estimates of the *supremum* norm of $\overline{u}$ and of the L^2 norms of Z and $\overline{Z}$, as well as a pointwise upper bound of the predictors $\overline{v}$ and $\hat{v}$ defined in Algorithm 3.2. Lemma 4.6.4 gives a crucial estimate of the difference between Z and $\overline{Z}$. The proofs are postponed to Subsection 6.3.

Proposition 4.6.2 *There exists a constant $C_{4.6.2}$ such that:*

$$\sup_{i=0\ldots N}\big[\sup_{x\in\mathcal{C}_i}|\overline{u}(t_i,x)|^2\big] \le C_{4.6.2}.$$

Proposition 4.6.3 *There exists a constant $C_{4.6.3}$ such that:*

$$\mathbb{E}\Big[\int_0^T |\overline{Z}_s|^2 ds\Big] + h\sum_{i=0}^{N-1}\mathbb{E}\big[|Z_{t_i}|^2\big] + h\sup_{i=0\ldots N}\big[\sup_{x\in\mathcal{C}_i}|\overline{v}(t_i,x)|^2\big] + h\sup_{i=0\ldots N-1}\big[\sup_{x\in\mathcal{C}_i}|\hat{v}(t_i,x)|^2\big] \le C_{4.6.3}.$$

The distance between Z and $\overline{Z}$ can be estimated as follows:

Lemma 4.6.4 *There exists a constant $C_{4.6.4}$ such that for $k \in \{1, ..., N\}$:*

$$\mathbb{E}\left| hZ_{t_{k-1}} - \mathbb{E}\left[\int_{t_{k-1}}^{t_k} \overline{Z}_s ds | \mathcal{F}_{t_{k-1}}\right]\right|^2 \leq C_{4.6.4} h^2 \mathcal{E}^2(\text{quantiz}).$$

6.2 Approximate Diffusion

Extension of the "Discrete Diffusion". Recall from (4.3.14) that, up to now, the approximate diffusion X is defined at the discretization times $(t_k)_{0 \leq k \leq N}$. For the proof, we need to extend the definition of X to the whole set $[0, T]$. Put for all $k \in \{0, ..., N-1\}$ and $t \in [t_k, t_{k+1}[$:

$$X_t \equiv X_{t_k} + b\big(X_{t_k}, \overline{u}(t_{k+1}, X_{t_k}), \hat{v}(t_k, X_{t_k})\big)(t - t_k) + \sigma\big(X_{t_k}, \overline{u}(t_{k+1}, X_{t_k})\big)\big[B_t - B_{t_k}\big]. \tag{4.6.3}$$

Hence, the extended process $(X_t)_{0 \leq t \leq T}$ is discontinuous at times $(t_k)_{1 \leq k \leq N-1}$. At a given time t_k, $1 \leq k \leq N$, the size of the jump performed by the process depends on the quantization error and on the spatial projection error. The first error is easily controlled by the distortion. Concerning the second one, the projection error is close to the spatial step δ when the grids are infinite. For truncated grids, the story is slightly different. In fact, as soon as the process stays inside $(\Delta_k)_{0 \leq k \leq N}$, the projection error is close to the step δ of the interior mesh of the grid $(\mathcal{C}_k)_{0 \leq k \leq N}$. At the opposite, outside $(\Delta_k)_{0 \leq k \leq N}$, the jump of the process may take large values.

Hitting Time of the Boundaries of the Grids. It is then well understood that we need to control the size of these jumps. In particular, we need to control the first hitting time of the boundaries of the sets $(\Delta_k)_{0 \leq k \leq N}$ by the discrete process $(X_{t_k})_{0 \leq k \leq N}$. This is the reason why the stopping time τ_∞, defined in (4.4.2), appears in the statement of Theorems 4.4.2 and 4.4.3.

Increments of the Forward Process.

Lemma 4.6.5 *There exists a constant $C_{4.6.5}$ such that for every $k \in \{0, ..., N-1\}$:*

$$\forall t \in [t_k, t_{k+1}[, \;\; \mathbb{E}\big[|X_t - X_{t_k}|^2 | \mathcal{F}_{t_k}\big] \leq C_{4.6.5} h.$$

Proof. The proof just follows from Assumption **(A.1)**, i.e. b and σ have linear growth, and from the boundedness of $\overline{u}$ and $h|\hat{v}|^2$ (see Propositions 4.6.2 and 4.6.3). □

Lemma 4.6.6 *For a given $k \in \{0, ..., N-1\}$, the norm of the increment $X_{t_{k+1}} - X_{t_k}$ is always bounded by $|\mathcal{T}(t_k, X_{t_k})| + \delta$. In particular, there exists a constant $C_{4.6.6}$ such that:*

$$\mathbb{E}\big[|X_{t_{k+1}} - X_{t_k}|^2 | \mathcal{F}_{t_k}\big] \leq C_{4.6.6}\big[h + \delta^2\big].$$

Proof. Since $X_{t_k} \in \mathcal{C}_\infty$, one has $\Pi_\infty(X_{t_k} + \mathcal{T}(t_k, X_{t_k})) = X_{t_k} + \Pi_\infty(\mathcal{T}(t_k, X_{t_k}))$ (invariance by translation of the grid $\mathcal{C}_\infty$). By definition of $X_{t_{k+1}}$, we get $X_{t_{k+1}} = \mathcal{Q}(R + \rho, \Pi_\infty(X_{t_k} + \mathcal{T}(t_k, X_{t_k}))) = \mathcal{Q}(R + \rho, X_{t_k} + \Pi_\infty(\mathcal{T}(t_k, X_{t_k})))$, where $\mathcal{Q}$ is defined in Section 3.4. Now, for every y in the image of the projection $\mathcal{Q}(R+\rho, .)$ and for every $z \in \mathbb{R}^d$, the distance $|\mathcal{Q}(R + \rho, y + z) - y|$ is bounded by $|z|$. Hence:

$$\begin{aligned} |X_{t_{k+1}} - X_{t_k}| &= \big|\mathcal{Q}\big(R + \rho, X_{t_k} + \Pi_\infty(\mathcal{T}(t_k, X_{t_k}))\big) - X_{t_k}\big| \\ &\leq |\Pi_\infty(\mathcal{T}(t_k, X_{t_k}))| \\ &\leq |\mathcal{T}(t_k, X_{t_k})| + \delta. \end{aligned} \tag{4.6.4}$$

Thanks to Propositions 4.6.2 and 4.6.3, we are able to bound the drift b appearing in the transition. Since $\mathbb{E}[|g(\Delta B^k)|^2] \leq Ch$, from Assumption **(A.1)** and Proposition 4.6.2, we also control the martingale part of the transition. This completes the proof. □

Auxiliary Controls of the Forward Process. The time continuous extension of X remains close to the discrete version of X up to time τ_∞:

Lemma 4.6.7 *There exists a constant $C_{4.6.7}$ such that:*

$$\sum_{i=0}^{N-1} \mathbb{E}\big[\mathbf{1}_{\{t_{i+1}<\tau_\infty\}}|X_{t_{i+1}} - X_{t_{i+1}-}|^2\big] \leq C_{4.6.7} h\big(\mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz})\big).$$

The proof is postponed to Subsection 6.3.

6.3 Proofs of the *A Priori* Controls

Discrete BSDE

This paragraph is devoted to the proof of Propositions 4.6.2, 4.6.3 and Lemma 4.6.4. We first give a control of the L^2 norm between $Z_{t_{k-1}}$ and the conditional expectation of $\int_{t_{k-1}}^{t_k} \overline{Z}_s ds$ appearing in Lemma 4.6.4. This preliminary estimate permits to prove Proposition 4.6.2. We then derive the complete proofs of Proposition 4.6.3 and Lemma 4.6.4.

Step One: Preliminary Control in Lemma 4.6.4.

The strategy just follows from the local BSDE writing explained in the proof of Proposition 4.6.1. Indeed, from (4.6.2), write for a given $k \in \{0, ..., N-1\}$:

$$Y_{t_{k+1}} + hf\big(X_{t_k}, \overline{u}(t_{k+1}, X_{t_k}), Z_{t_k}\big) = Y_{t_k} + \int_{t_k}^{t_{k+1}} \overline{Z}_s dB_s.$$

Multiplying this identity by ΔB^k and taking the conditional expectation w.r.t. $\mathcal{F}_{t_k}$ we obtain:

$$\mathbb{E}\big[Y_{t_{k+1}}\Delta B^k|\mathcal{F}_{t_k}\big] = \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} \overline{Z}_s ds|\mathcal{F}_{t_k}\bigg].$$

Recall now that:

$$\begin{aligned} hZ_{t_k} &= h\overline{v}(t_k, X_{t_k}) \\ &= \mathbb{E}\big[\overline{u}(t_{k+1}, X_{t_{k+1}})g(\Delta B^k)|\mathcal{F}_{t_k}\big] = \mathbb{E}\big[Y_{t_{k+1}}g(\Delta B^k)|\mathcal{F}_{t_k}\big]. \end{aligned}$$

Hence, we deduce that:

$$hZ_{t_k} - \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} \overline{Z}_s ds|\mathcal{F}_{t_k}\bigg] = \mathbb{E}\big[Y_{t_{k+1}}\big(g(\Delta B^k) - \Delta B^k\big)|\mathcal{F}_{t_k}\big]. \tag{4.6.5}$$

Recall now from (4.3.10) that there exists $C > 0$ s.t.:

$$\mathbb{E}[|g(\Delta B^k) - \Delta B^k|^2] \leq ChM^{-2/d}. \tag{4.6.6}$$

From (4.6.5) and (4.6.6) we derive:

$$\mathbb{E}\bigg[\bigg|hZ_{t_k} - \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} \overline{Z}_s ds|\mathcal{F}_{t_k}\bigg]\bigg|^2\bigg] \leq ChM^{-2/d}\mathbb{E}\big[Y_{t_{k+1}}^2\big]. \tag{4.6.7}$$

As already explained, this preliminary estimate (4.6.7) is necessary to prove Proposition 4.6.2 from which we will derive $\mathbb{E}[Y_{t_{k+1}}^2] \leq C$, and thus complete the proof of Lemma 4.6.4.

Step Two: Proof of Proposition 4.6.2 (Boundedness of the Approximate Solution)

To estimate the *supremum* norm of $\overline{u}$ over the grids $\mathcal{C}_0, ..., \mathcal{C}_N$, we follow the basic strategy of the BSDE theory and therefore apply a discrete version of Itô's formula to the discrete BSDE formula given in Proposition 4.6.1. Such a formula can be found in Shiryaev [Shi96], Chapter VII, Subsection 9:

Lemma 4.6.8 *Let $(A_k)_{0\le k\le n}$ be a sequence of vectors with values in $\mathbb{R}^q$, $q \ge 1$. Then :*

$$|A_k|^2 = |A_0|^2 + 2\sum_{i=1}^{k}\langle \Delta A_i, A_{i-1}\rangle + \sum_{i=1}^{k}|\Delta A_i|^2,$$

where $\forall 1 \le i \le n,\ \Delta A_i \equiv A_i - A_{i-1}$.

Apply now Lemma 4.6.8 to the sequence $(Y_{t_k})_{0\le k\le N}$. We obtain:

$$|Y_T|^2 = |Y_0|^2 + 2\sum_{i=1}^{N}\langle Y_{t_i} - Y_{t_{i-1}}, Y_{t_{i-1}}\rangle + \sum_{i=1}^{N}|Y_{t_i} - Y_{t_{i-1}}|^2.$$

Recall from Subsection 6.1 that:

$$Y_{t_i} - Y_{t_{i-1}} = -hf(X_{t_{i-1}}, \overline{u}(t_i, X_{t_{i-1}}), Z_{t_{i-1}}) + \int_{t_{i-1}}^{t_i}\overline{Z}_s dB_s.$$

We deduce that:

$$\begin{aligned}\mathbb{E}|Y_T|^2 = |Y_0|^2 &+ 2h\sum_{i=1}^{N}\mathbb{E}\langle -f(X_{t_{i-1}}, \overline{u}(t_i, X_{t_{i-1}}), Z_{t_{i-1}}), Y_{t_{i-1}}\rangle \\ &+ h^2\sum_{i=1}^{N}\mathbb{E}\big[f^2(X_{t_{i-1}}, \overline{u}(t_i, X_{t_{i-1}}), Z_{t_{i-1}})\big] + \mathbb{E}\int_0^T|\overline{Z}_s|^2 ds.\end{aligned}$$

Thanks to Assumption **(A.1)**, there exists a constant C such that:

$$|Y_0|^2 + \mathbb{E}\int_0^T|\overline{Z}_s|^2 ds \le \mathbb{E}|Y_T|^2 + Ch\sum_{i=1}^{N}\mathbb{E}\Big[|Y_{t_{i-1}}|\big(1 + |\overline{u}(t_i, X_{t_{i-1}})| + |Z_{t_{i-1}}|\big)\Big].$$

From Young's inequality, we derive for every $\eta > 0$:

$$\begin{aligned}|Y_0|^2 + \mathbb{E}\int_0^T|\overline{Z}_s|^2 ds \le \mathbb{E}|Y_T|^2 &+ C\eta^{-1}h\sum_{i=1}^{N}\mathbb{E}\big[|Y_{t_{i-1}}|^2\big] \\ &+ C\eta h\sum_{i=1}^{N}\mathbb{E}\big[1 + |\overline{u}(t_i, X_{t_{i-1}})|^2 + |Z_{t_{i-1}}|^2\big].\end{aligned} \tag{4.6.8}$$

From (4.6.7) we get that for every $i \in \{1, ..., N\}$:

$$h\mathbb{E}\big[|Z_{t_{i-1}}|^2\big] \le 2\Big[\mathbb{E}\int_{t_{i-1}}^{t_i}|\overline{Z}_s|^2 ds + CM^{-2/d}\mathbb{E}\big[Y_{t_i}^2\big]\Big].$$

Hence, from (4.6.8) and the above identity:

$$\begin{aligned}|Y_0|^2 + \frac{h}{2}\sum_{i=0}^{N-1}\mathbb{E}\big[|Z_{t_i}|^2\big] \le \mathbb{E}|Y_T|^2 &+ C(1+\eta^{-1})(h + M^{-2/d})\sum_{i=0}^{N}\mathbb{E}\big[|Y_{t_i}|^2\big] \\ &+ C\eta h\sum_{i=1}^{N}\mathbb{E}\big[1 + |\overline{u}(t_i, X_{t_{i-1}})|^2 + |Z_{t_{i-1}}|^2\big].\end{aligned} \tag{4.6.9}$$

Choose $\eta = (4C)^{-1}$ and deduce (with a new constant C) that:

$$|Y_0|^2 \le \mathbb{E}|Y_T|^2 + C(h + M^{-2/d})\sum_{i=0}^{N}\mathbb{E}\big[|Y_{t_i}|^2\big] + Ch\sum_{i=1}^{N}\mathbb{E}\big[1 + |\overline{u}(t_i, X_{t_{i-1}})|^2\big].$$

Recall that $|Y_T| \leq |H|_\infty$. Thus, we claim (recall that $M^{-2/d} < h$):

$$\sup_{x \in \mathcal{C}_0} |\overline{u}(0,x)|^2 \leq C + Ch \sum_{i=0}^{N} \sup_{x \in \mathcal{C}_i} |\overline{u}(t_i,x)|^2.$$

Therefore, there exists a constant $c > 0$ such that for $h < c$ (recall indeed that h is small):

$$\sup_{x \in \mathcal{C}_0} |\overline{u}(0,x)|^2 \leq C + Ch \sum_{i=1}^{N} \sup_{x \in \mathcal{C}_i} |\overline{u}(t_i,x)|^2.$$

As usual in BSDE theory, we could establish in a similar way that for every initial condition (t_k, x), $1 \leq k \leq N$:

$$\forall k \in \{0, ..., N-1\}, \ \sup_{x \in \mathcal{C}_k} |\overline{u}(t_k,x)|^2 \leq C + Ch \sum_{i=k+1}^{N} \sup_{x \in \mathcal{C}_i} |\overline{u}(t_i,x)|^2.$$

A discrete version of Gronwall's Lemma yields the result. □

Step Three: Proofs of Proposition 4.6.3 and Lemma 4.6.4.

Proof of Proposition 4.6.3. The L^2-estimate of Z follows from Proposition 4.6.2 and (4.6.9). Then, the L^2-estimate of $\overline{Z}$ follows from (4.6.8).

Finally, as a consequence of Proposition 4.6.2 and the definitions of $\overline{v}$ and $\hat{v}$, see Algorithm 3.2, we deduce the estimates of the *supremum* norms of $\overline{v}$ and $\hat{v}$. □

Proof of Lemma 4.6.4. Lemma 4.6.4 follows from (4.6.7) and Proposition 4.6.2. □

Approximate Forward Diffusion: Proof of Lemma 4.6.7.

Recall from (4.6.3) that the difference $X_{t_{i+1}} - X_{t_{i+1}-}$ writes:

$$\begin{aligned} X_{t_{i+1}} &- X_{t_{i+1}-} \\ &= \big[\Pi_{i+1}\big(X_{t_i} + \mathcal{T}(t_i, X_{t_i})\big) - \big(X_{t_i} + \mathcal{T}(t_i, X_{t_i})\big)\big] + \sigma\big(X_{t_i}, \overline{u}(t_{i+1}, X_{t_i})\big)\big[g(\Delta B^i) - \Delta B^i\big] \\ &\equiv E_1(i+1) + E_2(i+1). \end{aligned} \tag{4.6.10}$$

$E_1(i+1)$ appears as a projection error and $E_2(i+1)$ as a quantization one. There is no difficulty to estimate the second term: it is readily seen from (4.3.10) that $\mathbb{E}[|E_2(i+1)|^2 | \mathcal{F}_{t_i}] \leq ChM^{-2/d}$. To estimate the first term, note that $X_{t_i} + \mathcal{T}(t_i, X_{t_i})$ belongs to Δ_{i+1} on the event $\{t_{i+1} < \tau_\infty\}$. Thus, the distance between $X_{t_i} + \mathcal{T}(t_i, X_{t_i})$ and $\Pi_{i+1}(X_{t_i} + \mathcal{T}(t_i, X_{t_i}))$ is then bounded by the step δ. Deduce that the term $E_1(i+1)$ is bounded by δ on $\{t_{i+1} < \tau_\infty\}$. Finally,

$$\sum_{i=1}^{N} \mathbb{E}\big[\mathbf{1}_{\{t_{i+1} < \tau_\infty\}} |X_{t_{i+1}} - X_{t_{i+1}-}|^2\big] \leq Ch^{-1}\big[\delta^2 + hM^{-2/d}\big]. \tag{4.6.11}$$

This completes the proof. □

7 Proof, Second Step: Stability Properties

This section focuses on the second step of the proof of Theorems 4.4.1, 4.4.2 and 4.4.3, and aims to establish more specifically a suitable intermediate inequality, close to usual stability properties of FBSDEs.

Strategy. Recall first that two main strategies are conceivable in the theoretical framework to establish classical stability theorems for FBSDEs.

Denote to this end by (U', V', W') a solution of another FBSDE of type (E) with different coefficients. The associated PDE solution is just denoted by u'. In order to compare u' with u, recall that the following approaches have been employed in the literature devoted to FBSDE:

1. First, the recent induction principle given in Delarue [Del02] can be applied. In short, u and u' are compared on a neighborhood of the boundary T with classical arguments of stochastic analysis and the estimate of the difference between these solutions is then extended by induction from the final bound T to the initial bound 0. The local estimates consist in studying the distance between U and U' and between (V, W) and (V', W'). This strategy has been successfully applied to establish the existence and uniqueness of solutions to FBSDEs under a non-degeneracy assumption, see again [Del02], or to establish convergence properties arising in homogenization of quasilinear PDEs [Del04].

2. A second approach follows the earlier *Four Step Scheme* of Ma, Protter and Yong [MPY94]. In a nutshell, instead of studying the difference between U and U' and between (V, W) and (V', W'), the process $(u(t, U'_t))_{0 \le t \le T}$ is written with Itô's formula as the solution of a BSDE. This BSDE is then compared with the one satisfied by (V', W'). In particular, these BSDEs are both written with respect to the same diffusion U'. Generally speaking, this strategy holds when u is smooth enough (e.g. if u satisfies Theorem 4.2.1). It is then more direct than the previous one.

Thanks to Theorem 4.2.1, we will apply the second strategy: we will compare the process Y with the process $(u(t, X_t))_{0 \le t \le T \wedge \tau_\infty}$ (see (4.6.3) for the definition of the extension of X).

7.1 Statements of the Stability Results

First Stability Property. Applying the usual FBSDE machinery, we are able to establish in Subsection 7.2 the following first inequality:

Proposition 4.7.1 *There exists a constant $C_{4.7.1}$ such that for η small enough:*

$$\begin{aligned}
&|(\overline{u} - u)(0, x_0)|^2 + C_{4.7.1}^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{v} - v)(t_{j-1}, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big] \\
&\quad \le C_{4.7.1} \Bigg[\mathbb{P}\{\tau_\infty < +\infty\} + \mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz}) \\
&\qquad + \eta^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{u} - u)(t_j, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big] \\
&\qquad + \eta^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{u} - u)(t_{j-1}, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big] \\
&\qquad + (\eta + h) h \sum_{j=1}^{N} \mathbb{E}\big[|(\hat{v} - v)(t_{j-1}, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big] \Bigg].
\end{aligned} \tag{4.7.1}$$

When the drift b does not depend on z, the last term of the r.h.s. does not appear.

Estimates of the Gradient Increment. Assume for the moment that Proposition 4.7.1 holds. Note that the main problem then remains to estimate the last term in the r.h.s. of (4.7.1). Thanks to the specific choice of $\hat{v}$ in Subsection 3.3, we are able to establish in Section 7.3 the following control:

Proposition 4.7.2 *There exists a constant $C_{4.7.2}$ such that, for $k \in \{0, ..., N-1\}$, on $\{t_k < \tau_\infty\}$:*

$$|(\hat{v} - v)(t_k, X_{t_k})| \le C_{4.7.2} \Big[\mathcal{E}(\text{gradient}, p) + \mathcal{E}(\text{time}) + h\mathcal{E}(\text{space}) + \mathbb{E}\big[|(\overline{v} - v)(t_{k+1}, X_{t_{k+1}})|^2 | \mathcal{F}_{t_k}\big]^{1/2} \Big].$$

Main Stability Theorem. From Propositions 4.7.1 and 4.7.2, we claim:

Theorem 4.7.3 *There exists a constant $C_{4.7.3}$ such that for η small enough:*

$$\begin{aligned}
&|(\overline{u}-u)(0,x_0)|^2 + C_{4.7.3}^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{v}-v)(t_{j-1},X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big] \\
&\quad \le C_{4.7.3}\Bigg[\mathbb{P}\{\tau_\infty < +\infty\} + \mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz}) + \mathcal{E}^2(\text{gradient}, p) \\
&\qquad + \eta^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{u}-u)(t_j,X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big] \\
&\qquad + \eta^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{u}-u)(t_{j-1},X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big] \\
&\qquad + (\eta+h) h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{v}-v)(t_j,X_{t_j})|^2 \mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\Bigg].
\end{aligned} \tag{4.7.2}$$

Application of Theorem 4.7.3 to the proof of Theorems 4.4.1, 4.4.2 and 4.4.3 is given in Section 8.

7.2 Proof of Proposition 4.7.1

Starting Point: Time Continuous Backward Processes.

Recall that we aim to apply the second strategy exposed in the beginning of this section and thus to apply Itô's formula to $(u(t,X_t))_{0\le t\le T}$. Referring to the structure of the PDE $(\mathcal{E})$, set to this end for notational convenience:

$$\forall t \in [0,T],\ \overline{V}_t \equiv u(t,X_t),\ \overline{W}_t \equiv \nabla_x u(t,X_t)\sigma(X_t,\overline{V}_t). \tag{4.7.3}$$

Note moreover that the martingale part of $(\overline{V}_t)_{0\le t\le T}$ is driven by:

$$\forall t \in [0,T[,\ \hat{W}_t \equiv \nabla_x u(t,X_t)\sigma\big(X_{\phi(t)}, \overline{u}(\phi(t)+h, X_{\phi(t)})\big). \tag{4.7.4}$$

where $\phi(t) = t_k$ for $t_k \le t < t_{k+1},\ k \in \{0,...,N-1\}$.

From Theorem 4.2.1, we derive the following *a priori estimates* of $\overline{V},\overline{W}$:

$$\forall k \in \{0,...,N-1\},\ \forall s \in [t_k,t_{k+1}[,\ \mathbb{E}\big[|\overline{V}_s - \overline{V}_{t_k}| + |\overline{W}_s - \overline{W}_{t_k}|\ |\mathcal{F}_{t_k}\big] \le C\mathbb{E}\big[(s-t_k)^{1/2} + |X_s - X_{t_k}|\ |\mathcal{F}_{t_k}\big].$$

Hence, from Lemma 4.6.5, we get for $s \in [t_k,t_{k+1}[$ (recall that h is small) :

$$\mathbb{E}\big[|\overline{V}_s - \overline{V}_{t_k}| + |\overline{W}_s - \overline{W}_{t_k}|\ |\mathcal{F}_{t_k}\big] \le Ch^{1/2}. \tag{4.7.5}$$

Step One: Itô's formula for $\bar{V}$.

Using Itô's formula and the equation satisfied by u, we obtain for $t \in [t_i,t_{i+1}[,\ i \in \{0,...,N-1\}$:

$$\begin{aligned}
\overline{V}_t &= \overline{V}_{t_i} + \int_{t_i}^{t} \langle \nabla_x u(s,X_s), b\big(X_{t_i}, \overline{u}(t_{i+1},X_{t_i}), \hat{v}(t_i,X_{t_i})\big) - b(X_s,\overline{V}_s,\overline{W}_s)\rangle ds \\
&+ \frac{1}{2}\int_{t_i}^{t} \operatorname{tr}\Big(\big[a(X_{t_i},\overline{u}(t_{i+1},X_{t_i})) - a(X_s,\overline{V}_s)\big]\nabla^2_{x,x}u(s,X_s)\Big) ds \\
&- \int_{t_i}^{t} f\big(X_s,\overline{V}_s,\overline{W}_s\big) ds + \int_{t_i}^{t} \nabla_x u(s,X_s)\sigma\big(X_{t_i},\overline{u}(t_{i+1},X_{t_i})\big) dB_s.
\end{aligned}$$

Let t tend to t_{i+1} and deduce with obvious notation that:

$$\begin{aligned}
&\overline{V}_{t_{i+1}} - \overline{V}_{t_i} \\
&= \overline{V}_{t_{i+1}} - \overline{V}_{t_{i+1}-} + \int_{t_i}^{t_{i+1}} \big[F\big(s,X_s,X_{t_i},\overline{u}(t_{i+1},X_{t_i}),\hat{v}(t_i,X_{t_i})\big) - F\big(s,X_s,X_s,\overline{V}_s,\overline{W}_s\big)\big] ds \\
&- \int_{t_i}^{t_{i+1}} f\big(X_s,\overline{V}_s,\overline{W}_s\big) ds + \int_{t_i}^{t_{i+1}} \hat{W}_s dB_s.
\end{aligned}$$

Step Two: Difference of the Processes.

The strategy is well-known: we aim to make the difference between $\overline{V}$ and Y and then to apply the usual BSDE machinery to estimate the distance between these processes. Hence, we claim from (4.6.2):

$$\begin{aligned}
\overline{V}_{t_{i+1}} - Y_{t_{i+1}} - \big[\overline{V}_{t_i} - Y_{t_i}\big] &= \overline{V}_{t_{i+1}} - \overline{V}_{t_{i+1}-} \\
&+ \int_{t_i}^{t_{i+1}} \big[F\big(s, X_s, X_{t_i}, \overline{u}(t_{i+1}, X_{t_i}), \hat{v}(t_i, X_{t_i})\big) - F\big(s, X_s, X_s, \overline{V}_s, \overline{W}_s\big)\big] ds \\
&- \int_{t_i}^{t_{i+1}} \big[f\big(X_s, \overline{V}_s, \overline{W}_s\big) - f(X_{t_i}, \overline{u}(t_{i+1}, X_{t_i}), Z_{t_i})\big] ds \\
&+ \int_{t_i}^{t_{i+1}} \big[\hat{W}_s - \overline{Z}_s\big] dB_s \\
&\equiv \Delta E_{i+1}(1) + \Delta E_{i+1}(2) + \Delta E_{i+1}(3) + \Delta E_{i+1}(4).
\end{aligned}$$

As explained above, we now aim to estimate the distance between the processes $\overline{V}$ and Y up to time $\tau_\infty \wedge T$. Lemma 4.6.8 yields:

$$\begin{aligned}
\mathbb{E}|\overline{V}_{T\wedge\tau_\infty} - Y_{T\wedge\tau_\infty}|^2 &= |\overline{V}_0 - Y_0|^2 \\
&+ 2\mathbb{E}\sum_{j=1}^N \Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[\overline{V}_{t_{j-1}} - Y_{t_{j-1}}\big]\big[\Delta E_j(1) + \Delta E_j(2) + \Delta E_j(3) + \Delta E_j(4)\big]\Big] \\
&+ \mathbb{E}\sum_{j=1}^N \Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[\Delta E_j(1) + \Delta E_j(2) + \Delta E_j(3) + \Delta E_j(4)\big]^2\Big].
\end{aligned} \tag{4.7.6}$$

Denote $\forall j \in \{1, ..., N\},\ E_j \equiv \Delta E_j(1) + \Delta E_j(2) + \Delta E_j(3)$. From the above expression, we get:

$$\begin{aligned}
|\overline{V}_0 - Y_0|^2 &+ \sum_{j=1}^N \mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\Delta E_j(4)^2] = \mathbb{E}|\overline{V}_{T\wedge\tau_\infty} - Y_{T\wedge\tau_\infty}|^2 \\
&- 2\mathbb{E}\sum_{j=1}^N \Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[\overline{V}_{t_{j-1}} - Y_{t_{j-1}}\big]E_j\Big] - \sum_{j=1}^N \mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}(E_j^2 + 2E_j\Delta E_j(4))].
\end{aligned}$$

From Young's inequality we derive:

$$\begin{aligned}
|\overline{V}_0 - Y_0|^2 &+ \frac{1}{2}\sum_{j=1}^N \mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\Delta E_j(4)^2] \le \mathbb{E}|\overline{V}_{T\wedge\tau_\infty} - Y_{T\wedge\tau_\infty}|^2 \\
&- 2\mathbb{E}\sum_{j=1}^N \Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[\overline{V}_{t_{j-1}} - Y_{t_{j-1}}\big]E_j\Big] + \sum_{j=1}^N \mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}E_j^2].
\end{aligned} \tag{4.7.7}$$

Put:

$$\begin{aligned}
D(1) &\equiv -2\mathbb{E}\sum_{j=1}^N \Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[\overline{V}_{t_{j-1}} - Y_{t_{j-1}}\big]E_j\Big],\ D(2) \equiv \sum_{j=1}^N \mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}E_j^2], \\
D(3) &\equiv \sum_{j=1}^N \mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\Delta E_j(4)^2].
\end{aligned} \tag{4.7.8}$$

Step Three: Standard BSDE Techniques.

Following the BSDE techniques, we have to upper bound $D(1), D(2)$ (resp. lower bound $D(3)$) by terms appearing in the r.h.s. (resp. l.h.s.) of (4.7.1). The following lemmas whose proofs are postponed to the end of the subsection give the needed controls.

Lemma 4.7.4 *There exists a constant $C_{4.7.4}$ such that for $\eta \in]0,1]$:*

$$\begin{aligned}
|D(1)| + D(2) \le C\big[&\mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz}) + \mathbb{P}\{\tau_\infty < +\infty\}\big] \\
&+ Ch\sum_{j=1}^{N}\Big[\eta^{-1}\mathbb{E}\big[|(\overline{u}-u)(t_{j-1},X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big] + \mathbb{E}\big[|(\overline{u}-u)(t_j,X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\Big] \\
&+ h(\eta+h)\sum_{j=1}^{N}\Big[\mathbb{E}\big[|(\hat{v}-v)(t_{j-1},X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big] + \mathbb{E}\big[|(\overline{v}-v)(t_{j-1},X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\Big].
\end{aligned}$$

Lemma 4.7.5 *There exists a constant $C_{4.7.5} > 0$ such that:*

$$\begin{aligned}
D(3) \ge C_{4.7.5}^{-1} h\sum_{j=1}^{N}\mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|(\overline{v}-v)(t_{j-1},X_{t_{j-1}})|^2] - C_{4.7.5}\mathcal{E}^2(\text{quantiz}) - C_{4.7.5}\mathcal{E}^2(\text{time}) \\
- C_{4.7.5} h\sum_{j=1}^{N}\mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|(\overline{u}-u)(t_j,X_{t_{j-1}})|^2].
\end{aligned}$$

Note to conclude the proof of Proposition 4.7.1 that $Y_T = \overline{V}_T$. Hence, from Theorem 4.2.1 and Proposition 4.6.2 (boundedness of u and $\overline{u}$), $\mathbb{E}|\overline{V}_{T\wedge\tau_\infty} - Y_{T\wedge\tau_\infty}|^2 \le C\mathbb{P}\{\tau_\infty < T\} \le C\mathbb{P}\{\tau_\infty < +\infty\}$. Choose finally η small enough to obtain inequality (4.7.1) from (4.7.7), (4.7.8), and Lemmas 4.7.4 and 4.7.5. This completes, up to the proofs of Lemmas 4.7.4 and 4.7.5, the proof of Proposition 4.7.1. □

Proof of Lemma 4.7.4.

Recall that:

$$D(1) = -2\mathbb{E}\sum_{j=1}^{N}\Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[\overline{V}_{t_{j-1}} - Y_{t_{j-1}}\big]\big[\Delta E_j(1) + \Delta E_j(2) + \Delta E_j(3)\big]\Big]. \tag{4.7.9}$$

Note from Theorem 4.2.1 that $\Delta E(2)$ and $\Delta E(3)$ may be seen as "Lipschitz" differences since the partial derivatives of u of order one and two in x are bounded. Hence, it is readily seen that there exists a constant C, such that:

$$\begin{aligned}
|D(1)| \le C\mathbb{E}\sum_{j=1}^{N}\Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|\overline{V}_{t_{j-1}} - Y_{t_{j-1}}| \times \Big(&|\overline{V}_{t_j} - \overline{V}_{t_j-}| \\
&+ \int_{t_{j-1}}^{t_j}\big[|X_s - X_{t_{j-1}}| + |\overline{V}_s - \overline{u}(t_j,X_{t_{j-1}})| + |\overline{W}_s - \hat{v}(t_{j-1},X_{t_{j-1}})|\big]ds \\
&+ \int_{t_{j-1}}^{t_j}\big[|X_s - X_{t_{j-1}}| + |\overline{V}_s - \overline{u}(t_j,X_{t_{j-1}})| + |\overline{W}_s - Z_{t_{j-1}}|\big]ds\Big)\Big].
\end{aligned}$$

Recall that $\overline{V}_s = u(s,X_s)$. From Theorem 4.2.1 (Hölder regularity of u in t), (4.7.5) (regularity of $\overline{V}$ and $\overline{W}$) and Lemma 4.6.5 (control of the increments of X), we then deduce:

$$\begin{aligned}
|D(1)| \le C\mathbb{E}\sum_{j=1}^{N}\Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|\overline{V}_{t_{j-1}} - Y_{t_{j-1}}| \times \Big(&|\overline{V}_{t_j} - \overline{V}_{t_j-}| \\
&+ \int_{t_{j-1}}^{t_j}\big[h^{1/2} + |(\overline{u}-u)(t_j,X_{t_{j-1}})|\big]ds \\
&+ \int_{t_{j-1}}^{t_j}\big[|\overline{W}_{t_{j-1}} - \hat{v}(t_{j-1},X_{t_{j-1}})| + |\overline{W}_{t_{j-1}} - Z_{t_{j-1}}|\big]ds\Big)\Big].
\end{aligned} \tag{4.7.10}$$

Recall now that $\overline{W}_{t_{j-1}} = v(t_{j-1}, X_{t_{j-1}})$ and $Z_{t_{j-1}} = \overline{v}(t_{j-1}, X_{t_{j-1}})$. Thus, applying Young's inequality to (4.7.10), it comes for every $\eta \in]0,1]$:

$$\begin{aligned}
|D(1)| \leq C\mathcal{E}^2(\text{time}) &+ C\mathbb{E}\sum_{j=1}^{N}\Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[|\overline{V}_{t_{j-1}} - Y_{t_{j-1}}||\overline{V}_{t_j} - \overline{V}_{t_j-}|\big]\Big] \\
&+ Ch\sum_{j=1}^{N}\Big[\eta^{-1}\mathbb{E}\big[|(\overline{u}-u)(t_{j-1}, X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big] + \mathbb{E}\big[|(\overline{u}-u)(t_j, X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\Big] \quad (4.7.11) \\
&+ \eta h\sum_{j=1}^{N}\Big[\mathbb{E}\big[|(\hat{v}-v)(t_{j-1}, X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big] + \mathbb{E}\big[|(\overline{v}-v)(t_{j-1}, X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\Big].
\end{aligned}$$

It now remains to estimate the second term in the r.h.s. of (4.7.11). Note first that for all $j \in \{1, ..., N\}$, $\{t_{j-1} < \tau_\infty\} = \{t_j < \tau_\infty\} \cup \{t_j = \tau_\infty\}$. Hence, thanks to the boundedness of u and $\overline{u}$ (see Theorem 4.2.1 and Proposition 4.6.2):

$$\begin{aligned}
&\mathbb{E}\sum_{j=1}^{N}\Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[|\overline{V}_{t_{j-1}} - Y_{t_{j-1}}||\overline{V}_{t_j} - \overline{V}_{t_j-}|\big]\Big] \\
&\quad\leq \Big[\mathbb{E}\sum_{j=1}^{N}\big[\mathbf{1}_{\{t_j<\tau_\infty\}}|\overline{V}_{t_{j-1}} - Y_{t_{j-1}}|^2\big]\Big]^{1/2}\Big[\mathbb{E}\sum_{j=1}^{N}\big[\mathbf{1}_{\{t_j<\tau_\infty\}}|\overline{V}_{t_j} - \overline{V}_{t_j-}|^2\big]\Big]^{1/2} \qquad (4.7.12) \\
&\qquad + C\mathbb{P}\{\tau_\infty < +\infty\}.
\end{aligned}$$

Deduce from Lemma 4.6.7 (jumps of the process X) and the global Lipschitz property of u (see Theorem 4.2.1) that:

$$\begin{aligned}
&\mathbb{E}\sum_{j=1}^{N}\Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[|\overline{V}_{t_{j-1}} - Y_{t_{j-1}}||\overline{V}_{t_j} - \overline{V}_{t_j-}|\big]\Big] \\
&\quad\leq C\Big[\mathbb{E}\sum_{j=1}^{N}\big[\mathbf{1}_{\{t_j<\tau_\infty\}}|(\overline{u}-u)(t_{j-1}, X_{t_{j-1}})|^2\big]\Big]^{1/2} h^{1/2}\big(\mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz})\big)^{1/2} \\
&\qquad + C\mathbb{P}\{\tau_\infty < +\infty\} \qquad (4.7.13) \\
&\quad\leq C\big(\mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz})\big) + Ch\mathbb{E}\sum_{j=1}^{N}\big[\mathbf{1}_{\{t_j<\tau_\infty\}}|(\overline{u}-u)(t_{j-1}, X_{t_{j-1}})|^2\big] \\
&\qquad + C\mathbb{P}\{\tau_\infty < +\infty\}.
\end{aligned}$$

Plug (4.7.13) in (4.7.11) to derive the required control for $D(1)$.

Let us now turn to the estimation of $D(2)$. Recall that $D(2)$ is given by:

$$D(2) = \mathbb{E}\sum_{j=1}^{N}\Big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big[\Delta E_j(1) + \Delta E_j(2) + \Delta E_j(3)\big]^2\Big].$$

We first give a control of $\Delta E_j(1)$, for a given $j \in \{1, ..., N\}$. To this end, note again from Lemma 4.6.7 (jumps of the process X) that:

$$\begin{aligned}
\sum_{j=1}^{N}\mathbb{E}\big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}(\Delta E_j(1))^2\big] &= \sum_{j=1}^{N}\mathbb{E}\big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|u(t_j, X_{t_j}) - u(t_j, X_{t_j-})|^2\big] \\
&\leq C\sum_{j=1}^{N}\mathbb{E}\big[\mathbf{1}_{\{t_j<\tau_\infty\}}|X_{t_j} - X_{t_j-}|^2\big] + C\mathbb{P}\{\tau_\infty < +\infty\}. \qquad (4.7.14) \\
&\leq Ch\big(\mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz})\big) + C\mathbb{P}\{\tau_\infty < +\infty\}.
\end{aligned}$$

We recall that $\Delta E_j(2)$ and $\Delta E_j(3)$ are "Lipschitz" differences. Hence, as done in (4.7.10) to estimate $D(1)$, we can derive that:

$$
\begin{aligned}
&\sum_{j=1}^{N} \mathbb{E}\big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big(\Delta E_j(2)+\Delta E_j(3)\big)^2\big] \\
&\quad\le Ch^2\left(1+\sum_{j=1}^{N}\mathbb{E}\big[|(\overline{u}-u)(t_j,X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\right. \qquad (4.7.15)\\
&\quad\left.+\sum_{j=1}^{N}\Big[\mathbb{E}\big[|(\hat{v}-v)(t_{j-1},X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]+\mathbb{E}\big[|(\overline{v}-v)(t_{j-1},X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\Big]\right).
\end{aligned}
$$

Equations (4.7.14) and (4.7.15) give the required control for $D(2)$. This completes the proof of Lemma 4.7.4. □

Proof of Lemma 4.7.5.

Write first:

$$
\begin{aligned}
&h\sum_{j=1}^{N}\mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|\overline{v}(t_{j-1},X_{t_{j-1}})-v(t_{j-1},X_{t_{j-1}})|^2] \\
&\quad\le Ch\sum_{j=1}^{N}\left\{\mathbb{E}\left[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\left|\overline{v}(t_{j-1},X_{t_{j-1}})-\frac{1}{h}\mathbb{E}\Big[\int_{t_{j-1}}^{t_j}\overline{Z}_s ds\big|\mathcal{F}_{t_{j-1}}\Big]\right|^2\right]\right.\\
&\quad+\mathbb{E}\left[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\left|\frac{1}{h}\mathbb{E}\Big[\int_{t_{j-1}}^{t_j}[\overline{Z}_s-\hat{W}_s]ds\big|\mathcal{F}_{t_{j-1}}\Big]\right|^2\right]\\
&\quad\left.+\mathbb{E}\left[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\left|\frac{1}{h}\mathbb{E}\Big[\int_{t_{j-1}}^{t_j}[\hat{W}_s-v(t_{j-1},X_{t_{j-1}})]ds\big|\mathcal{F}_{t_{j-1}}\Big]\right|^2\right]\right\}\\
&\quad\equiv A(1)+A(2)+A(3).
\end{aligned}
$$

From Lemma 4.6.4 (distance between Z and $\overline{Z}$) we then derive $A(1)\le C\mathcal{E}^2(\text{quantiz})$. For the term $A(2)$, the Cauchy-Schwarz inequality yields:

$$
A(2)\le C\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\int_{t_{j-1}}^{t_j}|\overline{Z}_s-\hat{W}_s|^2ds\right]=C\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big(\Delta E_j(4)\big)^2\right]=CD(3).
$$

Concerning $A(3)$ it comes:

$$
\begin{aligned}
A(3)&\le C\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\int_{t_{j-1}}^{t_j}|\hat{W}_s-v(t_{j-1},X_{t_{j-1}})|^2ds\right]\\
&=C\sum_{j=1}^{N}\mathbb{E}\left[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\int_{t_{j-1}}^{t_j}|\nabla_x u(s,X_s)\sigma\big(X_{t_{j-1}},\overline{u}(t_j,X_{t_{j-1}})\big)\right.\\
&\qquad\left.-\nabla_x u(t_{j-1},X_{t_{j-1}})\sigma\big(X_{t_{j-1}},u(t_{j-1},X_{t_{j-1}})\big)|^2ds\right].
\end{aligned}
$$

Following the techniques employed in the previous proof, relying on the smoothness of the true solution, cf. Theorem 4.2.1, on the boundedness of the approximate solution, cf. Proposition 4.6.2, and on intermediate controls of the process X, cf. Lemma 4.6.5, we get:

$$
A(3)\le Ch\left[1+\sum_{j=1}^{N}\mathbb{E}\big[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|\overline{u}(t_j,X_{t_{j-1}})-u(t_j,X_{t_{j-1}})|^2\big]\right].
$$

The above estimates of $A(1), A(2), A(3)$ give:

$$D(3) \geq C^{-1}h\sum_{j=1}^{N}\mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|\overline{v}(t_{j-1}, X_{t_{j-1}}) - v(t_{j-1}, X_{t_{j-1}})|^2]$$
$$- Ch\sum_{j=1}^{N}\mathbb{E}[\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}|\overline{u}(t_j, X_{t_{j-1}}) - u(t_j, X_{t_{j-1}})|^2] - C\mathcal{E}^2(\text{quantiz}) - C\mathcal{E}^2(\text{time}).$$

This completes the proof. □

7.3 Proof of Proposition 4.7.2 (Difference of the Gradients)

Strategy.

In Proposition 4.7.2, we aim to control the quantity $|(\hat{v} - v)(t_k, X_{t_k})|$ for $t_k < \tau_\infty$. Recall to this end (see Algorithm 3.2):

$$\hat{v}(t_k, X_{t_k}) = \mathbb{E}\big[\overline{v}\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k}))\big)|\mathcal{F}_{t_k}\big].$$

We first write $v(t_k, X_{t_k})$ in a similar way to study the difference $(\hat{v}-v)(t_k, X_{t_k})$. From Theorem 4.2.1 (regularity of u) and from the proof of Lemma 4.6.6 (regularity of X), we claim:

$$\begin{aligned}
&\big|\mathbb{E}\big[v(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k})))|\mathcal{F}_{t_k}\big] - v(t_k, X_{t_k})\big| \\
&\quad\leq C\Big[h^{1/2} + \big|\mathbb{E}\big[v(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k})))|\mathcal{F}_{t_k}\big] - v(t_{k+1}, X_{t_k})\big|\Big] \\
&\quad\leq C\Big[h^{1/2} + \mathbb{E}\big[|\Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k})) - X_{t_k}|\ |\mathcal{F}_{t_k}\big]\Big] \\
&\quad\leq C\big[h^{1/2} + \delta\big].
\end{aligned}$$

Hence:

$$|(\hat{v} - v)(t_k, X_{t_k})| \leq C\mathbb{E}\big[\big|(\overline{v} - v)\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k}))\big)\big|\ |\mathcal{F}_{t_k}\big] + C(h^{1/2} + \delta). \tag{4.7.16}$$

Proposition 4.7.2 directly follows from (4.7.16) and from the next theorem:

Theorem 4.7.6 *There exists a constant $C_{4.7.6}$ such that on $\{t_k < \tau_\infty\}$:*

$$\begin{aligned}
&\mathbb{E}\big[\big|(\overline{v} - v)\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k}))\big)\big|\ |\mathcal{F}_{t_k}\big] \\
&\quad\leq C_{4.7.6}\mathcal{E}(\text{gradient}, p) + C_{4.7.6}\mathbb{E}\big[|(\overline{v} - v)(t_{k+1}, X_{t_{k+1}})|^2|\mathcal{F}_{t_k}\big]^{1/2}.
\end{aligned}$$

The main difficulty to prove Theorem 4.7.6 lies in the lack of regularity of $\overline{v}$. To overcome this point, note first that

$$\mathbb{E}\big[\big|(\overline{v} - v)\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k}))\big)\big|\ |\mathcal{F}_{t_k}\big] \tag{4.7.17}$$

and

$$\mathbb{E}\big[\big|(\overline{v} - v)\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}(t_k, X_{t_k}))\big)\big|^2|\mathcal{F}_{t_k}\big]^{1/2} \tag{4.7.18}$$

write as expectations of a given function with respect to two different kernels. We then aim to compare these underlying kernels. Recall that for a given $x \in \mathcal{C}_k$, both $\mathcal{T}^0(t_k, x)$ and $\mathcal{T}(t_k, x)$ are, up to a quantization procedure, Gaussian random variables with same covariance matrices but different means. The strategy then consists in applying a Gaussian change of variable to pass from the first kernel to the second one.

Step One: Proof of Theorem 4.7.6, Exhibition of Underlying Kernels.

We first write (4.7.17) with respect to the underlying kernel $\mathcal{T}^0$. Note in this frame, with the notations of

Subsection 3.4, that for every $x \in \mathbb{R}^d$, $\Pi_{k+1}(x) = \Pi_{k+1} \circ \Pi_\infty(x)$ since $\Pi_\infty(x) \in \Delta_k \Leftrightarrow x \in \Delta_k$. Thus, using the invariance by translation of $\mathcal{C}_\infty$ to pass from the second to the third line, (4.7.17) writes:

$$
\begin{aligned}
&\mathbb{E}\big[\big|(\overline{v}-v)\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k}))\big)\big| \,|\mathcal{F}_{t_k}\big] \\
&\quad = \mathbb{E}\big[\big|(\overline{v}-v)\big(t_{k+1}, \Pi_{k+1}(\Pi_\infty(X_{t_k} + \mathcal{T}^0(t_k, X_{t_k})))\big)\big| \,|\mathcal{F}_{t_k}\big] \\
&\quad = \mathbb{E}\big[\big|(\overline{v}-v)\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \Pi_\infty(\mathcal{T}^0(t_k, X_{t_k})))\big)\big| \,|\mathcal{F}_{t_k}\big] \\
&\quad = \sum_{y \in \mathcal{C}_\infty} \Big[\big|(\overline{v}-v)(t_{k+1}, \Pi_{k+1}(X_{t_k} + y))\big| \mathbb{P}\big\{\Pi_\infty(\mathcal{T}^0(t_k, X_{t_k})) = y | \mathcal{F}_{t_k}\big\}\Big].
\end{aligned}
\tag{4.7.19}
$$

In the same way, the square of (4.7.18) writes:

$$
\begin{aligned}
&\mathbb{E}\big[\big|(\overline{v}-v)\big(t_{k+1}, \Pi_{k+1}(X_{t_k} + \mathcal{T}(t_k, X_{t_k}))\big)\big|^2 \,|\mathcal{F}_{t_k}\big] \\
&\quad = \sum_{y \in \mathcal{C}_\infty} \Big[\big|(\overline{v}-v)(t_{k+1}, \Pi_{k+1}(X_{t_k} + y))\big|^2 \mathbb{P}\big\{\Pi_\infty(\mathcal{T}(t_k, X_{t_k})) = y | \mathcal{F}_{t_k}\big\}\Big].
\end{aligned}
\tag{4.7.20}
$$

Equations (4.7.19) and (4.7.20) provide relevant writings to estimate (4.7.17) and (4.7.18). Indeed, it is sufficient to bound for a given $x \in \mathcal{C}_k$ and a given $y \in \mathcal{C}_\infty$ the probability $\mathbb{P}\{\Pi_\infty(\mathcal{T}^0(t_k, x)) = y\}$ by (up to a multiplicative constant) the probability $\mathbb{P}\{\Pi_\infty(\mathcal{T}(t_k, x)) = y\}$. Recall to this end that:

$$
\begin{aligned}
&\mathcal{T}^0(t_k, x) = \sigma(x, \overline{u}(t_{k+1}, x)) g(\Delta B^k), \\
&\mathcal{T}(t_k, x) = b(x, \overline{u}(t_{k+1}, x), \hat{v}(t_k, x)) h + \sigma(x, \overline{u}(t_{k+1}, x)) g(\Delta B^k) \\
&\hat{v}(t_k, x) = \mathbb{E}[\overline{v}(t_{k+1}, \Pi_{k+1}(x + \mathcal{T}^0(t_k, x)))].
\end{aligned}
$$

For notational convenience, we set:

$$
\begin{aligned}
&\Sigma(t_{k+1}, x) = \sigma(x, \overline{u}(t_{k+1}, x)), \\
&\mu(t_{k+1}, x) = b(x, \overline{u}(t_{k+1}, x), \hat{v}(t_k, x)).
\end{aligned}
$$

We also introduce $\|\Sigma\|_\infty = \sup_{k \in \{0,\dots,N\}} \big[\sup_{x \in \mathcal{C}_k} |\Sigma(t_k, x)|\big]$ and $\|\mu\|_\infty = \sup_{k \in \{0,\dots,N\}} \big[\sup_{x \in \mathcal{C}_k} |\mu(t_k, x)|\big]$. From Assumption **(A.1)** and Propositions 4.6.2 and 4.6.3 (boundedness of $\overline{u}$ and $h^{1/2}\hat{v}$), $\|\Sigma\|_\infty + h^{1/2}\|\mu\|_\infty \leq C$.

Step Two: Proof of Theorem 4.7.6, Comparison of Kernels.

The following proposition, whose proof is postponed to the end of the section, establishes a connection between the previous probabilities $\mathbb{P}\big\{\Pi_\infty(\mathcal{T}^0(t_k, x)) = y\big\}$ and $\mathbb{P}\big\{\Pi_\infty(\mathcal{T}(t_k, x)) = y\big\}$.

Proposition 4.7.7 *There exists a constant $C_{4.7.7} > 0$ such that for every $y \in \mathcal{C}_\infty$:*

$$
\mathbb{P}\big\{\Pi_\infty(\mathcal{T}^0(t_k, x)) = y\big\} \leq \alpha_k(y) + \beta(y)(\mathbb{P}^{1/2}\big\{\Pi_\infty(\mathcal{T}(t_k, x)) = y\big\} + \eta_k)
$$

where

$$
\begin{aligned}
\alpha_k(y) &\equiv \mathbb{P}\big\{|\Sigma(t_{k+1}, x) g(\Delta B^k) - y|_\infty \leq \delta/2,\ |g(\Delta B^k) - \Delta B^k|_\infty > \delta/(2\|\Sigma\|_\infty)\big\}, \\
\beta(y) &\equiv C_{4.7.7} \delta^{d/2} h^{-d/4} \exp\big[-C_{4.7.7}^{-1} h^{-1} |y|^2\big],\ \eta_k \equiv \mathbb{P}^{1/2}\big\{|g(\Delta B^k) - \Delta B^k|_\infty > \delta/(4\|\Sigma\|_\infty)\big\}.
\end{aligned}
$$

In the above expression, for all $z \in \mathbb{R}^d$, $|z|_\infty \equiv \max_{i \in \{1,\dots,d\}} |z_i|$.

From Proposition 4.6.3, $h^{1/2}\overline{v}$ is bounded by a known constant. Owing to Proposition 4.7.7 and (4.7.19), we then get:

$$
\begin{aligned}
&\sum_{y\in\mathcal{C}_\infty}\Big[\big[|(\overline{v}-v)(t_{k+1},\Pi_{k+1}(x+y))|\big]\mathbb{P}\big\{\Pi_\infty(\mathcal{T}^0(t_k,x))=y\big\}\Big]\\
&\quad\le Ch^{-1/2}\mathbb{P}\big\{|g(\Delta B^k)-\Delta B^k|_\infty>\delta/(2\|\Sigma\|_\infty)\big\}\\
&\qquad+C\delta^{d/2}h^{-d/4-1/2}\mathbb{P}^{1/2}\big\{|g(\Delta B^k)-\Delta B^k|_\infty>\delta/(4\|\Sigma\|_\infty)\big\}\Big[\sum_{y\in\mathcal{C}_\infty}\exp\big[-C^{-1}h^{-1}|y|^2\big]\Big]\\
&\qquad+C\Big[\sum_{y\in\mathcal{C}_\infty}\big[\delta^d h^{-d/2}\exp\big[-C^{-1}h^{-1}|y|^2\big]\big]\Big]^{1/2}\\
&\qquad\quad\times\Big[\sum_{y\in\mathcal{C}_\infty}|(\overline{v}-v)(t_{k+1},\Pi_{k+1}(x+y))|^2\mathbb{P}\big\{\Pi_\infty(\mathcal{T}(t_k,x))=y\big\}\Big]^{1/2}\\
&\quad\equiv T(1)+T(2)+T(3).
\end{aligned}
\tag{4.7.21}
$$

Thanks to (4.3.10) and to the Bienaymé-Chebychev inequality:

$$
T(1)\le Ch^{p/2-1/2}\delta^{-p}M^{-p/d}. \tag{4.7.22}
$$

Thanks again to (4.3.10) (applied to the exponent $2p$):

$$
T(2)\le Ch^{p/2-d/4-1/2}\delta^{-p+d/2}M^{-p/d}\Big[\sum_{y\in\mathcal{C}_\infty}\exp\big[-C^{-1}h^{-1}|y|^2\big]\Big]. \tag{4.7.23}
$$

Note now from (4.7.20) that:

$$
\begin{aligned}
T(3)=C&\Big[\sum_{y\in\mathcal{C}_\infty}\big[\delta^d h^{-d/2}\exp\big[-C^{-1}h^{-1}|y|^2\big]\big]\Big]^{1/2}\\
&\times\mathbb{E}\big[|(\overline{v}-v)(t_{k+1},\Pi_{k+1}(x+\mathcal{T}(t_k,x)))|^2\big]^{1/2}.
\end{aligned}
\tag{4.7.24}
$$

Since $h^{-1}\delta^2$ is small, note now that:

$$
\begin{aligned}
&\sum_{y\in\mathcal{C}_\infty}(\delta h^{-1/2})^d\exp(-C^{-1}h^{-1}|y|^2)\\
&\quad=\left(\sum_{j\in\mathbb{Z}}\delta h^{-1/2}\exp(-C^{-1}j^2\delta^2h^{-1})\right)^d\le C'\left(\int_{\mathbb{R}}\delta h^{-1/2}\exp(-C^{-1}z^2\delta^2h^{-1})dz\right)^d\le C''.
\end{aligned}
$$

Hence:

$$
\sum_{y\in\mathcal{C}_\infty}\exp\big[-C^{-1}h^{-1}|y|^2\big]\le C''\delta^{-d}h^{d/2}. \tag{4.7.25}
$$

From (4.7.21), (4.7.22), (4.7.23), (4.7.24) and (4.7.25), we complete the proof of Theorem 4.7.6 (recall again that $h^{-1}\delta^2$ is small to dominate $T(1)$ by $\mathcal{E}(\text{gradient},p)$). □

Proof of Proposition 4.7.7, Gaussian Change of Variable.

It now remains to prove Proposition 4.7.7. Note first that:

$$
\begin{aligned}
&\mathbb{P}\big\{\Pi_\infty(\mathcal{T}^0(t_k,x))=y\big\}\\
&\quad\le\mathbb{P}\big\{|\mathcal{T}^0(t_k,x)-y|_\infty\le\delta/2\big\}\\
&\quad=\mathbb{P}\big\{|\Sigma(t_{k+1},x)g(\Delta B^k)-y|_\infty\le\delta/2\big\}\\
&\quad=\mathbb{P}\big\{|\Sigma(t_{k+1},x)g(\Delta B^k)-y|_\infty\le\delta/2,\ |g(\Delta B^k)-\Delta B^k|_\infty>\delta/(2\|\Sigma\|_\infty)\big\}\\
&\qquad+\mathbb{P}\big\{|\Sigma(t_{k+1},x)g(\Delta B^k)-y|_\infty\le\delta/2,\ |g(\Delta B^k)-\Delta B^k|_\infty\le\delta/(2\|\Sigma\|_\infty)\big\}\\
&\quad\equiv P(1)+P(2).
\end{aligned}
\tag{4.7.26}
$$

Note that $P(1)$ matches $\alpha_k(y)$ in Proposition 4.7.7. Hence, we just have to estimate $P(2)$:

$$\begin{aligned} P(2) &= \mathbb{P}\big\{|\Sigma(t_{k+1},x)g(\Delta B^k) - y|_\infty \le \delta/2,\ |g(\Delta B^k) - \Delta B^k|_\infty \le \delta/(2\|\Sigma\|_\infty)\big\} \\ &\le \mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k - y|_\infty \le \delta\big\}. \end{aligned} \tag{4.7.27}$$

The strategy is now clear. A Gaussian change of variable permits to introduce artificially the drift appearing in the definition of $\mathcal{T}(t_k, x)$. It comes:

$$\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k - y|_\infty \le \delta\big\} = \mathbb{E}^{\mathbb{Q}_k}\left[\frac{d\mathbb{P}}{d\mathbb{Q}_k}\mathbf{1}_{\{|\Sigma(t_{k+1},x)\Delta B^k - y|_\infty \le \delta\}}\right]$$

where $\frac{d\mathbb{Q}_k}{d\mathbb{P}} = \exp\left(\langle(\Sigma^{-1}\mu)(t_{k+1},x), \Delta B^k\rangle - \frac{|(\Sigma^{-1}\mu)(t_{k+1},x)|^2 h}{2}\right)$. Under $\mathbb{Q}_k$, ΔB^k is a d-dimensional Gaussian random variable with mean $(\Sigma^{-1}\mu)(t_{k+1},x)h$ and covariance matrix $h\mathbf{I}_d$. Hence,

$$\begin{aligned} &\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k - y|_\infty \le \delta\big\} \\ &\quad = \mathbb{E}\left[\exp\left(-\langle(\Sigma^{-1}\mu)(t_{k+1},x), \Delta B^k\rangle - \frac{|(\Sigma^{-1}\mu)(t_{k+1},x)|^2 h}{2}\right)\mathbf{1}_{\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta\}}\right] \\ &\quad \le C\exp\big[C(1 + h^{-1/2}|y|)\big]\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta\big\}, \end{aligned}$$

where the last identity follows from Assumption **(A)** and Propositions 4.6.2 and 4.6.3.

We deduce from (4.7.27) and the above control that:

$$P(2) \le C\exp\big[C(1 + h^{-1/2}|y|)\big]\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta\big\}. \tag{4.7.28}$$

It now remains to quantify the Gaussian increment in the latter quantity with the converse procedure. The first step is to replace the upper bound δ in (4.7.28) by $\delta/4$. Note to this end that, for a given $y \in \mathcal{C}_\infty$, we can find a finite subset $(y_i)_{1\le i\le N(d)}$ included in $\mathbb{R}^d$, $N(d) = 4^d$, such that:

$$\begin{aligned} &\forall i \in \{1, ..., N(d)\},\ |y - y_i|_\infty \le \delta, \\ &\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta\big\} = \bigcup_{i=1}^{N(d)}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y_i|_\infty \le \delta/4\big\}. \end{aligned}$$

Hence:

$$\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta\big\} \le \sum_{i=1}^{N(d)}\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y_i|_\infty \le \delta/4\big\}. \tag{4.7.29}$$

In order to recover y instead of y_i for each term of the sum in the r.h.s. of the above expression, we use once again a Gaussian change of variable. Namely, $\forall i \in \{1, ..., N(d)\}$, put:

$$\frac{d\mathbb{Q}_k(i)}{d\mathbb{P}} = \exp\left(\langle\Sigma^{-1}(t_{k+1},x)(y_i - y), \Delta B^k\rangle - \frac{|\Sigma^{-1}(t_{k+1},x)(y_i - y)|^2 h}{2}\right).$$

Set $G_k(i) = \Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y_i$. Under $\mathbb{Q}_k(i)$, $G_k(i) \sim \mathcal{N}(\mu(t_{k+1},x)h - y, \Sigma\Sigma^*(t_{k+1},x)h)$. Hence, from (4.7.29):

$$\begin{aligned} &\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta\big\} \\ &\quad \le \sum_{i=1}^{N(d)}\mathbb{E}\Bigg[\exp\left(-\langle\Sigma^{-1}(t_{k+1},x)(y_i - y), \Delta B^k\rangle - \frac{|\Sigma^{-1}(t_{k+1},x)(y_i - y)|^2 h}{2}\right) \\ &\qquad\qquad \times \mathbf{1}_{\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta/4\}}\Bigg]. \end{aligned}$$

Arguments similar to the ones used at the preceding Gaussian change of variable combined with the inequality $|y - y_i|_\infty \le \delta$ then give:

$$
\begin{aligned}
&\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta\big\} \\
&\quad \le CN(d)\exp(C\delta|y|)\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta/4\big\}.
\end{aligned}
\tag{4.7.30}
$$

Write now:

$$
\begin{aligned}
&\mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty \le \delta/4\big\} \\
&\quad = \mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty < \delta/4\big\} \\
&\quad \le \mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty < \delta/4,\ |\Delta B^k - g(\Delta B^k)|_\infty \le \delta/(4\|\Sigma\|_\infty)\big\} \\
&\qquad\quad + \mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty < \delta/4,\ |\Delta B^k - g(\Delta B^k)|_\infty > \delta/(4\|\Sigma\|_\infty)\big\} \\
&\quad \le \mathbb{P}\big\{|\Sigma(t_{k+1},x)g(\Delta B^k) + \mu(t_{k+1},x)h - y|_\infty < \delta/2,\ |\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty < \delta/4\big\} \\
&\qquad\quad + \mathbb{P}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty < \delta/4,\ |\Delta B^k - g(\Delta B^k)|_\infty > \delta/(4\|\Sigma\|_\infty)\big\} \\
&\quad \equiv P(3) + P(4).
\end{aligned}
\tag{4.7.31}
$$

Note first from the Cauchy-Schwarz inequality, Assumption **(A)** and the explicit expression of the Gaussian kernel:

$$
\begin{aligned}
P(3) &\le \mathbb{P}^{1/2}\big\{|\Sigma(t_{k+1},x)g(\Delta B^k) + \mu(t_{k+1},x)h - y|_\infty < \delta/2\big\} \\
&\qquad \times \mathbb{P}^{1/2}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty < \delta/4\big\} \\
&\le C\mathbb{P}^{1/2}\big\{\Pi_\infty(\mathcal{T}(t_k,x)) = y\big\} \\
&\qquad \times \left(\int_{B_\infty(0,\delta/4)} \exp\left(-h^{-1}|\Sigma^{-1}(t_{k+1},x)(z-(\mu(t_{k+1},x)h-y))|^2\right)\frac{dz}{(2\pi h)^{d/2}}\right)^{1/2} \\
&\le C\big[\delta^d h^{-d/2}\exp\big[-C^{-1}h^{-1}|y|^2\big]\big]^{1/2}\mathbb{P}^{1/2}\big\{\Pi_\infty(\mathcal{T}(t_k,x)) = y\big\}.
\end{aligned}
\tag{4.7.32}
$$

From the same arguments, we also derive:

$$
\begin{aligned}
P(4) &\le \mathbb{P}^{1/2}\big\{|\Sigma(t_{k+1},x)\Delta B^k + \mu(t_{k+1},x)h - y|_\infty < \delta/4\big\}\mathbb{P}^{1/2}\big\{|\Delta B^k - g(\Delta B^k)|_\infty > \delta/(4\|\Sigma\|_\infty)\big\} \\
&\le C\big[\delta^d h^{-d/2}\exp\big[-C^{-1}h^{-1}|y|^2\big]\big]^{1/2}\mathbb{P}^{1/2}\big\{|\Delta B^k - g(\Delta B^k)|_\infty > \delta/(4\|\Sigma\|_\infty)\big\}.
\end{aligned}
\tag{4.7.33}
$$

Thus, from (4.7.32) and (4.7.33):

$$
P(3) + P(4) \le \beta(y)\big(\mathbb{P}^{1/2}\big\{\Pi_\infty(\mathcal{T}(t_k,x)) = y\big\} + \eta_k\big). \tag{4.7.34}
$$

Apply Young's inequality to the product $h^{-1/2}|y|$ and $\delta|y|$ in the exponential terms in (4.7.28) and (4.7.30) and deduce finally from (4.7.28), (4.7.30), (4.7.31) and (4.7.34) that:

$$
\begin{aligned}
P(2) \le C\delta^{d/2}h^{-d/4}\exp\big[-C^{-1}h^{-1}|y|^2\big]\big[&\mathbb{P}^{1/2}\big\{\Pi_\infty(\mathcal{T}(t_k,x)) = y\big\} \\
&+ \mathbb{P}^{1/2}\big\{|\Delta B^k - g(\Delta B^k)|_\infty > \delta/(4\|\Sigma\|_\infty)\big\}\big].
\end{aligned}
$$

Thanks to (4.7.26), this completes the proof of Proposition 4.7.7. □

8 Proof, Third Step: Gronwall's Lemma

This section is devoted to the final step of the proof of Theorems 4.4.1, 4.4.2 and 4.4.3. For notational convenience, we set:

$$
\mathcal{E}^2(\text{global}) \equiv \mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{trunc}) + \mathcal{E}^2(\text{quantiz}) + \mathcal{E}^2(\text{gradient}, p).
$$

8.1 Proof of Theorem 4.4.1, Infinite Grids

We first explain how to derive Theorem 4.4.1 from Theorem 4.7.3 when $\rho = +\infty$, i.e. when $\tau_\infty = +\infty$ a.s.. In this framework, the term $\mathcal{E}^2(\text{trunc})$ in $\mathcal{E}^2(\text{global})$ reduces to 0. The general case is detailed in the next subsection. For infinite grids, we obtain from Theorem 4.7.3:

$$
\begin{aligned}
&|(\overline{u} - u)(0, x_0)|^2 + C^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{v} - v)(t_{j-1}, X_{t_{j-1}})|^2\big] \\
&\quad \leq C\Bigg[\mathcal{E}^2(\text{global}) + \eta^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{u} - u)(t_j, X_{t_{j-1}})|^2\big] + \eta^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{u} - u)(t_{j-1}, X_{t_{j-1}})|^2\big] \\
&\qquad + (\eta + h) h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{v} - v)(t_j, X_{t_j})|^2\big]\Bigg].
\end{aligned}
$$

Note that, for every $x \in \mathcal{C}_N$, $\overline{v}(T, x) = v(T, x)$. Hence, choose η and h small enough to deduce that:

$$
|(\overline{u} - u)(0, x_0)|^2 + C^{-1} h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{v} - v)(t_{j-1}, X_{t_{j-1}})|^2\big] \leq C\Bigg[\mathcal{E}^2(\text{global}) + h \sum_{j=0}^{N} \sup_{x \in \mathcal{C}_j} |(\overline{u} - u)(t_j, x)|^2\Bigg]. \tag{4.8.1}
$$

As usual in BSDE theory, note that the estimate (4.8.1) holds actually for any starting point (t_k, x), $0 \leq k \leq N$, $x \in \mathcal{C}_k$. Hence:

$$
\sup_{x \in \mathcal{C}_k} |(\overline{u} - u)(t_k, x)|^2 \leq C\Bigg[\mathcal{E}^2(\text{global}) + h \sum_{j=k}^{N} \sup_{x \in \mathcal{C}_j} |(\overline{u} - u)(t_j, x)|^2\Bigg].
$$

Since h is small, for every $k \in \{0, ..., N-1\}$:

$$
\sup_{x \in \mathcal{C}_k} |(\overline{u} - u)(t_k, x)|^2 \leq C\Bigg[\mathcal{E}^2(\text{global}) + h \sum_{j=k+1}^{N} \sup_{x \in \mathcal{C}_j} |(\overline{u} - u)(t_j, x)|^2\Bigg].
$$

Apply now Gronwall's Lemma and deduce that:

$$
\sup_{x \in \mathcal{C}_0} |\overline{u}(0, x) - u(0, x)|^2 \leq C\mathcal{E}^2(\text{global}).
$$

This completes the proof of Theorem 4.4.1 when $\rho = +\infty$. $\square$

8.2 Proof of Theorem 4.4.1, General Case

We now turn to the case of truncated grids. Generally speaking, most of the approach given in the former subsection still applies in the general framework. It is however impossible to mimic word for word the arguments given above and we need to refine the previous Gronwall argument.

First Step. We first aim to get rid of the difference $\overline{v} - v$ appearing in the r.h.s. of (4.7.2). Due to the functions $(\mathbf{1}_{\{t_{j-1} < \tau_\infty\}})_{j=1,\dots,N}$, the machinery used in the previous subsection does not apply. To overcome this difficulty, we write $\{t_{j-1} < \tau_\infty\} = \{t_j < \tau_\infty\} \cup \{t_j = \tau_\infty\}$. Indeed, since $\overline{v}(T, x) = v(T, x)$ for $x \in \mathcal{C}_N$ and $h|\overline{v} - v|^2$ is bounded (see Theorem 4.2.1 and Proposition 4.6.3), note that:

$$
\begin{aligned}
&h \sum_{j=1}^{N} \mathbb{E}\big[|(\overline{v} - v)(t_j, X_{t_j})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big] \\
&\quad = h \sum_{j=1}^{N-1} \mathbb{E}\big[|(\overline{v} - v)(t_j, X_{t_j})|^2 \mathbf{1}_{\{t_j < \tau_\infty\}}\big] + h \sum_{j=1}^{N-1} \mathbb{E}\big[|(\overline{v} - v)(t_j, X_{t_j})|^2 \mathbf{1}_{\{t_j = \tau_\infty\}}\big] \\
&\quad \leq h \sum_{j=2}^{N} \mathbb{E}\big[|(\overline{v} - v)(t_{j-1}, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big] + C\mathbb{P}\{\tau_\infty < +\infty\}.
\end{aligned} \tag{4.8.2}
$$

Hence, (4.7.2) and (4.8.2) give for η and h small enough:

$$
\begin{aligned}
&|(\overline{u}-u)(0,x_0)|^2 + C^{-1}h\sum_{j=1}^{N}\mathbb{E}\big[|(\overline{v}-v)(t_{j-1},X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\\
&\quad\le C\Big[\mathbb{P}\{\tau_\infty<+\infty\}+\mathcal{E}^2(\text{global})+h\sum_{j=1}^{N}\mathbb{E}\big[|(\overline{u}-u)(t_j,X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\\
&\qquad+h\sum_{j=2}^{N}\mathbb{E}\big[|(\overline{u}-u)(t_{j-1},X_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty\}}\big]\Big].
\end{aligned}
\tag{4.8.3}
$$

The term $\mathcal{E}^2(\text{time})+\mathcal{E}^2(\text{space})+\mathcal{E}^2(\text{quantiz})+\mathcal{E}^2(\text{gradient},p)$ appearing in (4.7.2) has been replaced, for notational convenience, by $\mathcal{E}^2(\text{global})$ in the above reference. Nevertheless, mention carefully that the origin of the term $\mathcal{E}^2(\text{trunc})$ has not been explained yet. It is in the following lines.

Second Step. Note that (4.8.3) still holds if X starts at a given time t_i, $i\in\{0,...N\}$, from an $\mathcal{F}_{t_i}$-measurable and square integrable random variable ξ with values in $\mathcal{C}_i$. In such a case:

$$
\begin{aligned}
&\mathbb{E}\big[|(\overline{u}-u)(t_i,\xi)|^2\big]\\
&\quad\le C\Big[\mathbb{P}\{\tau_\infty^{t_i,\xi}<+\infty\}+\mathcal{E}^2(\text{global})+h\sum_{j=i+1}^{N}\mathbb{E}\big[|(\overline{u}-u)(t_j,X^{t_i,\xi}_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty^{t_i,\xi}\}}\big]\\
&\qquad+h\sum_{j=i+2}^{N}\mathbb{E}\big[|(\overline{u}-u)(t_{j-1},X^{t_i,\xi}_{t_{j-1}})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty^{t_i,\xi}\}}\big]\Big],
\end{aligned}
\tag{4.8.4}
$$

where the superscript (t_i,ξ) denotes the initial condition of the process X. Due to the shift between t_{j-1} and t_j in the r.h.s., there is no possible choice of ξ to recover the same form of terms in the left and right hand sides. In particular, Gronwall's Lemma does not apply at this stage of the proof. Note in fact that the same problem occured in Subsection 8.1: this was the reason why the supremum was taken in the r.h.s. of (4.8.1). To mimic this strategy, choose $\xi=x\in\mathcal{C}_i$ and take the supremum over these x:

$$
\sup_{x\in\mathcal{C}_i}|(\overline{u}-u)(t_i,x)|^2\le C\Big[\sup_{x\in\mathcal{C}_i}\mathbb{P}\{\tau_\infty^{t_i,x}<+\infty\}+\mathcal{E}^2(\text{global})+h\sum_{j=i+1}^{N}\sup_{x\in\mathcal{C}_j}|(\overline{u}-u)(t_j,x)|^2\Big].
\tag{4.8.5}
$$

It is then readily seen that $\mathbb{P}\{\tau_\infty^{t_i,x}<+\infty\}$ is far from being negligible when x tends to the boundary of the grid $\mathcal{C}_i$. In particular, there is no hope to prove Theorem 4.4.1 in the case $\rho<+\infty$ with the arguments used in Subsection 8.1.

Strategy. The reason why the latter method fails is rather clear: the way the supremum is taken in (4.8.5) is too rough to be really efficient. Our strategy then consists in applying (4.8.4) to a suitable choice of ξ. We then have to estimate the probability $\mathbb{P}\{\tau_\infty^{t_i,\xi}<+\infty\}$ for a random initial condition (t_i,ξ), $\xi\in L^2(\Omega,\mathcal{F}_{t_i},\mathbb{P})$ with values in $\mathcal{C}_i$. To this end, we need to control efficiently the tails of the variables $(X^{t_i,\xi}_{t_j\wedge\tau_\infty^{t_i,\xi}})_{i\le j\le N}$. Since the drift b is not bounded, the best we can do consists in estimating the L^2 norms of these variables.

An L^2 Control of the Process X.

The following lemma is proved at the end of the subsection:

Lemma 4.8.1 *For all $k\in\{0,...,N\}$, put $\tau_k=\tau_\infty\wedge t_k$. Then, there exists a constant $C_{4.8.1}$ such that for all $i\in\{0,...,N\}$ and $\xi\in L^2(\Omega,\mathcal{F}_{t_i},\mathbb{P})$ with values in $\mathcal{C}_i$:*

$$
\forall k\in\{i,...,N\},\ \mathbb{E}\big[|X^{t_i,\xi}_{\tau_k}|^2\big]\le C_{4.8.1}\big[\mathbb{E}|\xi|^2+1+\mathcal{E}^2(\text{space})+\mathcal{E}^2(\text{gradient},p)\big].
$$

Estimate of the Probability of Hitting the Boundary.

Thanks to the previous lemma, we are now able to estimate the probability $\mathbb{P}\{\tau_\infty^{t_i,\xi}<+\infty\}$, with (i,ξ) as in

Lemma 4.8.1. Indeed, $\{\tau_\infty^{t_i,\xi} < +\infty\} \subset \{|X_{\tau_N}^{t_i,\xi}|_\infty + \delta \geq R + \rho\}$. Thanks to the Bienaymé-Chebychev inequality and to Lemma 4.8.1 (with $k = N$), it comes:

$$\begin{aligned}\mathbb{P}\{\tau_\infty^{t_i,\xi} < +\infty\} &\leq C(R+\rho)^{-2}\big[\mathbb{E}\big[|X_{\tau_N}^{t_i,\xi}|^2\big] + \delta^2\big] \\ &\leq C\big[(R+\rho)^{-2}\mathbb{E}[|\xi|^2] + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{trunc}) + \mathcal{E}^2(\text{gradient}, p)\big].\end{aligned} \tag{4.8.6}$$

Plug now (4.8.6) into (4.8.4) to obtain:

$$\begin{aligned}&\mathbb{E}\big[|(\overline{u} - u)(t_i, \xi)|^2\big] \\ &\leq C\Bigg[(R+\rho)^{-2}\mathbb{E}[|\xi|^2] + \mathcal{E}^2(\text{global}) + h\sum_{j=i+1}^{N}\mathbb{E}\big[|(\overline{u} - u)(t_j, X_{t_{j-1}}^{t_i,\xi})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty^{t_i,\xi}\}}\big] \\ &\quad + h\sum_{j=i+2}^{N}\mathbb{E}\big[|(\overline{u} - u)(t_{j-1}, X_{t_{j-1}}^{t_i,\xi})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty^{t_i,\xi}\}}\big]\Bigg].\end{aligned} \tag{4.8.7}$$

A Refined Gronwall Argument.

The key idea is to find by induction a sequence of constants $c_i(1)$, $c_i(2)$, $i \in \{0, ..., N\}$, such that for any $\xi \in L^2(\Omega, \mathcal{F}_{t_i}, \mathbb{P})$ with values in $\mathcal{C}_i$:

$$\mathbb{E}\big[|(\overline{u} - u)(t_i, \xi)|^2\big] \leq c_i(1)\mathcal{E}^2(\text{global}) + c_i(2)(R+\rho)^{-2}\mathbb{E}[|\xi|^2]. \tag{4.8.8}$$

If we can exhibit two sequences $c_i(1)$ and $c_i(2)$, $i \in \{0, ..., N\}$, satisfying (4.8.8) and uniformly bounded by a constant C, then (4.8.8) yields:

$$\mathbb{E}\big[|(\overline{u} - u)(t_i, \xi)|^2\big] \leq C\big(\mathcal{E}^2(\text{global}) + (R+\rho)^{-2}\mathbb{E}[|\xi|^2]\big). \tag{4.8.9}$$

Choosing $i = 0$ and $\xi = x_0 \in \mathcal{C}_0$, we then complete the proof of Theorem 4.4.1.

Construction of the Sequences $c_i(1), c_i(2)$.

For $i = N$, we put: $c_N(1) = c_N(2) = 0$.

Consider now a given $i \in \{0, ..., N-1\}$ and assume that $c_j(1)$ and $c_j(2)$ are known for $j = i+1, ..., N$. Note first from (4.8.8), with $i = j$ and $\xi = X_{t_{j-1}}^{t_i,\xi}$, and from Lemma 4.8.1, with $k = j-1$, that:

$$\begin{aligned}\mathbb{E}\big[|(\overline{u} - u)(t_j, X_{t_{j-1}}^{t_i,\xi})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty^{t_i,\xi}\}}\big] &\leq c_j(1)\mathcal{E}^2(\text{global}) + c_j(2)(R+\rho)^{-2}\mathbb{E}\big[|X_{\tau_{j-1}}^{t_i,\xi}|^2\big] \\ &\leq c_j(1)\mathcal{E}^2(\text{global}) + c_j(2)C_{4.8.1}(R+\rho)^{-2}\left[\mathbb{E}[|\xi|^2] + 1 + \mathcal{E}^2(\text{global})\right] \\ &\leq \big(c_j(1) + 2c_j(2)C_{4.8.1}\big)\mathcal{E}^2(\text{global}) + c_j(2)C_{4.8.1}(R+\rho)^{-2}\mathbb{E}[|\xi|^2].\end{aligned} \tag{4.8.10}$$

Note that the same holds for $\mathbb{E}[|(\overline{u} - u)(t_{j-1}, X_{t_{j-1}}^{t_i,\xi})|^2\mathbf{1}_{\{t_{j-1}<\tau_\infty^{t_i,\xi}\}}]$ with respect to $c_{j-1}(1)$ and $c_{j-1}(2)$, $j \in \{i+2, ..., N\}$. Thus, from (4.8.7) and (4.8.10):

$$\begin{aligned}\mathbb{E}\big[|(\overline{u} - u)(t_i, \xi)|^2\big] &\leq C\Bigg[(R+\rho)^{-2}\mathbb{E}[|\xi|^2] + \mathcal{E}^2(\text{global}) \\ &\quad + 2h\sum_{j=i+1}^{N}\big[\big(c_j(1) + 2c_j(2)C_{4.8.1}\big)\mathcal{E}^2(\text{global}) + c_j(2)C_{4.8.1}(R+\rho)^{-2}\mathbb{E}[|\xi|^2]\big]\Bigg]\end{aligned} \tag{4.8.11}$$

Set:

$$c_i(1) = C + 2Ch\sum_{j=i+1}^{N}\big[c_j(1) + 2C_{4.8.1}c_j(2)\big],\ c_i(2) = C + 2CC_{4.8.1}h\sum_{j=i+1}^{N}\big[c_j(2)\big]. \tag{4.8.12}$$

Gronwall's Lemma directly proves that the second sequence is bounded. This allows to derive the boundedness of the first sequence. □

It now remains to prove Lemma 4.8.1.

Proof of Lemma 4.8.1. We remove the superscript (t_i, ξ) in the writing of X. Then:

$$
\begin{aligned}
X_{\tau_k} &= \xi + \sum_{j=i}^{k-1} \big[(X_{t_{j+1}} - X_{t_j})\mathbf{1}_{\{t_j < \tau_\infty\}}\big] \\
&= \xi + \sum_{j=i}^{k-1} \big[\mathcal{T}(t_j, X_{t_j})\mathbf{1}_{\{t_j < \tau_\infty\}}\big] \\
&\quad + \sum_{j=i}^{k-1} \big[\big(\Pi_{j+1}(X_{t_j} + \mathcal{T}(t_j, X_{t_j})) - X_{t_j} - \mathcal{T}(t_j, X_{t_j})\big)\mathbf{1}_{\{t_{j+1} < \tau_\infty\}}\big] \\
&\quad + \sum_{j=i}^{k-1} \big[\big(\Pi_{j+1}(X_{t_j} + \mathcal{T}(t_j, X_{t_j})) - X_{t_j} - \mathcal{T}(t_j, X_{t_j})\big)\mathbf{1}_{\{t_{j+1} = \tau_\infty\}}\big] \\
&\equiv \xi + S(1) + S(2) + S(3).
\end{aligned}
\tag{4.8.13}
$$

Deal first with $S(2)$ and $S(3)$. Note to this end that the distance between $X_{t_j} + \mathcal{T}(t_j, X_{t_j})$ and $\Pi_{j+1}(X_{t_j} + \mathcal{T}(t_j, X_{t_j}))$ is bounded by δ on $\{t_{j+1} < \tau_\infty\}$. Deduce in particular that $\mathbb{E}[|S(2)|^2] \le \delta^2(k-i)^2 \le C\mathcal{E}^2(\text{space})$. Note moreover that the distance between $X_{t_j} + \mathcal{T}(t_j, X_{t_j})$ and $\Pi_{j+1}(X_{t_j} + \mathcal{T}(t_j, X_{t_j}))$ is always bounded by $2|\mathcal{T}(t_j, X_{t_j})| + \delta$ (see Lemma 4.6.6). Thus, applying the Cauchy-Schwarz inequality, it is readily seen that:

$$
\begin{aligned}
\mathbb{E}\big[|S(3)|^2\big] &\le C\delta^2 + C\sum_{j=i}^{k-1} \mathbb{E}\big[|\mathcal{T}(t_j, X_{t_j})|^2 \mathbf{1}_{\{t_{j+1} = \tau_\infty\}}\big] \\
&\le Ch^2\mathcal{E}^2(\text{space}) + C\mathbb{P}\{\tau_\infty < +\infty\} + C\sum_{j=i}^{k-1} \mathbb{E}\big[|\mathcal{T}(t_j, X_{t_j})|^4\big].
\end{aligned}
\tag{4.8.14}
$$

From Propositions 4.6.2 and 4.6.3, we can prove that $\mathbb{E}[|\mathcal{T}(t_j, X_{t_j})|^4] \le Ch^2$. We finally deduce ($h$ being small):

$$
\mathbb{E}\big[|S(2)|^2\big] + \mathbb{E}\big[|S(3)|^2\big] \le C\big[\mathbb{P}\{\tau_\infty < +\infty\} + \mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space})\big]. \tag{4.8.15}
$$

A simpler strategy in (4.8.14) consists in bounding $\mathbb{E}[|\mathcal{T}(t_j, X_{t_j})|^2]$ by Ch and thus $\mathbb{E}[|S(3)|^2]$ by $C(1 + \mathcal{E}^2(\text{space}))$. However, the control obtained in (4.8.14) turns out to be useful in the next subsection.

It now remains to deal with $S(1)$. The strategy is slightly different: thanks to Propositions 4.6.3 and 4.7.2, we are able to estimate the drift, and thanks to the independence of the Brownian increments, we easily bound the martingale part. From Assumption **(A)**, we deduce that there exists a constant C such that (recall that h is small):

$$
\mathbb{E}\big[|S(1)|^2\big] \le Ch(k-i)\left[1 + h\sum_{j=i}^{k-1} \mathbb{E}\big[|\hat{v}(t_j, X_{t_j})|^2 \mathbf{1}_{\{t_j < \tau_\infty\}}\big]\right].
$$

Apply now Proposition 4.7.2 concerning the difference $\hat{v} - v$. Since v is bounded, deduce that (recall that $\mathcal{E}^2(\text{time}) = h$ and $h^2\mathcal{E}^2(\text{space}) = \delta^2$ are small):

$$
\mathbb{E}\big[|S(1)|^2\big] \le Ch(k-i)\left[1 + \mathcal{E}^2(\text{gradient}, p) + h\sum_{j=i}^{k-1} \mathbb{E}\big[|\overline{v}(t_{j+1}, X_{t_{j+1}})|^2 \mathbf{1}_{\{t_j < \tau_\infty\}}\big]\right]. \tag{4.8.16}
$$

Recall now from Proposition 4.6.3 that $h\sum_{j=i}^{N-1} \mathbb{E}[|\overline{v}(t_{j+1}, X_{t_{j+1}})|^2]$ is bounded.

Thus, from (4.8.16), we claim:

$$
\mathbb{E}\big[|S(1)|^2\big] \le C\big[1 + \mathcal{E}^2(\text{gradient}, p)\big]. \tag{4.8.17}
$$

Thanks to (4.8.13), (4.8.15) and (4.8.17), we complete the proof of Lemma 4.8.1. □

8.3 Proofs of Theorems 4.4.2 and 4.4.3

We turn to the proof of Theorems 4.4.2 and 4.4.3. The initial condition of the process X is given by $X_0 = x_0$, $x_0 \in \mathcal{C}_0$, as in (4.3.14).

Proof of Theorem 4.4.2. From inequalities (4.8.3) (deriving from the stability theorem), (4.8.6) (probability of hitting the boundary of the grids) and (4.8.10) (estimate of $\overline{u} - u$, recall that $c_j(1)$, $c_j(2)$, $j \in \{0, ..., N\}$, are uniformly bounded), Theorem 4.4.2 holds with $v(t_i, X_{t_i})\mathbf{1}_{\{t_i < \tau_\infty\}}$ instead of $v(t_i, X_{t_i})$. Since v is bounded (see Theorem 4.2.1) and since the probability of hitting the boundaries of the grids is controlled (see again (4.8.6)), we easily complete the proof of Theorem 4.4.2. □

Proof of Theorem 4.4.3. It just remains to study the convergence of $(X_{t_k}, Y_{t_k}, Z_{t_k})_{0 \le t_k \le \tau_\infty \wedge T}$ towards the solution (U, V, W) of (E). Thanks to the Lipschitz properties of b and σ, we first deduce by standard computations the analogue of Proposition 4.7.1. Recall to this end that, for all $k \in \{0, ..., N\}$, $\tau_k = \tau_\infty \wedge t_k$.

Proposition 4.8.2 *There exists a constant $C_{4.8.2}$ such that for $k \in \{1, ..., N\}$:*

$$\begin{aligned}\mathbb{E}|X_{\tau_k} - U_{\tau_k}|^2 \le C_{4.8.2}\Bigg[&\mathbb{P}\{\tau_\infty < +\infty\} + \mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz}) \\ &+ h\sum_{j=0}^{k-1}\Big[\mathbb{E}\big[|X_{t_j} - U_{t_j}|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big] + \mathbb{E}\big[|(\overline{u} - u)(t_{j+1}, X_{t_j})|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big]\Big] \\ &+ h\sum_{j=0}^{k-1}\mathbb{E}\big[|(\hat{v} - v)(t_j, X_{t_j})|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big]\Bigg].\end{aligned} \tag{4.8.18}$$

Sketch of the Proof. Note that $X_{\tau_k} - U_{\tau_k}$ writes:

$$\begin{aligned}X_{\tau_k} - U_{\tau_k} &= \sum_{j=0}^{k-1}\big[[(X_{t_{j+1}} - U_{t_{j+1}}) - (X_{t_j} - U_{t_j})]\mathbf{1}_{\{t_j<\tau_\infty\}}\big] \\ &= \sum_{j=0}^{k-1}\big[[\mathcal{T}(t_j, X_{t_j}) - (U_{t_{j+1}} - U_{t_j})]\mathbf{1}_{\{t_j<\tau_\infty\}}\big] \\ &\quad + \sum_{j=0}^{k-1}\big[[\Pi_{j+1}(X_{t_j} + \mathcal{T}(t_j, X_{t_j})) - (X_{t_j} + \mathcal{T}(t_j, X_{t_j}))]\mathbf{1}_{\{t_{j+1}<\tau_\infty\}}\big] \\ &\quad + \sum_{j=0}^{k-1}\big[[\Pi_{j+1}(X_{t_j} + \mathcal{T}(t_j, X_{t_j})) - (X_{t_j} + \mathcal{T}(t_j, X_{t_j}))]\mathbf{1}_{\{t_{j+1}=\tau_\infty\}}\big] \\ &\equiv R(1) + R(2) + R(3).\end{aligned}$$

Note that $R(2)$ and $R(3)$ exactly match $S(2)$ and $S(3)$ in the proof of Lemma 4.8.1, with $i = 0$. In particular, (4.8.15) applies. Due to the Lipschitz properties of b and σ and to the smoothness of u (see Theorem 4.2.1), the term $R(1)$ is treated in a classical way. □

Recall now from Proposition 4.7.2 (estimate of $\hat{v} - v$), Theorem 4.4.2 (L^2 estimate of $\overline{v} - v$) and (4.8.6) (probability of hitting the boundary of the grids):

$$
\begin{aligned}
&h\sum_{j=0}^{k-1}\mathbb{E}\big[|(\hat{v}-v)(t_j,X_{t_j})|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big]\\
&\quad\le C\Big[\mathcal{E}^2(\text{time})+\mathcal{E}^2(\text{space})+\mathcal{E}^2(\text{gradient},p)+h\sum_{j=0}^{k-1}\mathbb{E}\big[|(\overline{v}-v)(t_{j+1},X_{t_{j+1}})|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big]\Big]\\
&\quad\le C\Big[\mathcal{E}^2(\text{time})+\mathcal{E}^2(\text{space})+\mathcal{E}^2(\text{gradient},p)+h\sum_{j=1}^{k}\mathbb{E}\big[|(\overline{v}-v)(t_j,X_{t_j})|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big]\\
&\qquad+h\sum_{j=1}^{k}\mathbb{E}\big[|(\overline{v}-v)(t_j,X_{t_j})|^2\mathbf{1}_{\{t_j=\tau_\infty\}}\big]\Big]\\
&\quad\le C\big[\mathcal{E}^2(\text{global})+\mathbb{P}\{\tau_\infty<+\infty\}\big]\le C\mathcal{E}^2(\text{global}).
\end{aligned}
\tag{4.8.19}
$$

Apply now inequality (4.8.10) (estimate of $\overline{u}-u$) and (4.8.19) to (4.8.18) and deduce from Gronwall's Lemma:

$$\sup_{k\in\{0,\dots,N\}}\mathbb{E}|X_{\tau_k}-U_{\tau_k}|^2\le C\mathcal{E}^2(\text{global}).\tag{4.8.20}$$

Finally, according to Theorem 4.2.1, to Theorem 4.4.2 (L^2 estimate of $\overline{v}-v$) and to (4.8.10) (estimate of $\overline{u}-u$), we deduce the following intermediate estimate:

$$\sup_{k\in\{0,\dots,N\}}\mathbb{E}|X_{\tau_k}-U_{\tau_k}|^2+\sup_{k\in\{0,\dots,N\}}\mathbb{E}|Y_{\tau_k}-V_{\tau_k}|^2+h\sum_{j=0}^{N-1}\mathbb{E}\big[|Z_{t_j}-W_{t_j}|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big]\le C\mathcal{E}^2(\text{global}).\tag{4.8.21}$$

Applying Doob's inequality, we easily obtain the upper bound for:

$$\mathbb{E}\big[\sup_{k\in\{0,\dots,N\}}|X_{\tau_k}-U_{\tau_k}|^2\big]+\mathbb{E}\big[\sup_{k\in\{0,\dots,N\}}|Y_{\tau_k}-V_{\tau_k}|^2\big]+h\sum_{j=0}^{N-1}\mathbb{E}\big[|Z_{t_j}-W_{t_j}|^2\mathbf{1}_{\{t_j<\tau_\infty\}}\big].\tag{4.8.22}$$

It finally remains to prove that (4.8.22) holds with $(U_{t_k},V_{t_k},W_{t_k})_{0\le k\le N}$ instead of $(U_{\tau_k},V_{\tau_k},W_{t_k}\mathbf{1}_{\{t_k<\tau_\infty\}})_{0\le k\le N}$. Since the same arguments apply for V and W, we just detail the case of U. Note indeed that for every $k\in\{0,...,N\}$:

$$\sup_{k\in\{0,\dots,N\}}|X_{\tau_k}-U_{t_k}|^2\le C\sup_{k\in\{0,\dots,N\}}|X_{\tau_k}-U_{\tau_k}|^2+C\sup_{k\in\{0,\dots,N\}}|U_{\tau_k}-U_{t_k}|^2.$$

Thanks to Burkholder, Davis and Gundy inequalities, it is readily seen that:

$$\mathbb{E}\big[\sup_{k\in\{0,\dots,N\}}|U_{\tau_k}-U_{t_k}|^2\big]\le C\mathbb{E}\big[t_N-\tau_\infty\big]\le CT\mathbb{P}\{\tau_\infty<+\infty\}.$$

Referring to (4.8.6), we easily complete the proof of Theorem 4.4.3. □

9 Conclusion

As a conclusion, we first give in Subsection 9.1 further comments on Theorem 4.4.1 and compare in particular the global error with the one obtained by Douglas et *al.* in [DMP96]. We then give some easy extensions in Subsection 9.2. Finally, we detail in Subsection 9.3 the technical difficulties associated to the natural Algorithm 3.1.

9.1 Comments and Comparisons with Other Methods

We discuss in this subsection the total complexity and the rate of convergence of Algorithm 3.2.

Complexity of the Algorithm. Note first that the order of the total complexity of the algorithm is $h^{-1} \times M \times (2\delta^{-1}(\rho + R))^d$.

Rate of Convergence. Recall also that the global error of the algorithm is given by Theorem 4.4.1. Comparing with the results in [DMP96], this global error is worse in our case. There are two reasons to explain this difference. The first one does not depend on the algorithm, but is a consequence of our working assumptions. Indeed, under suitable smoothness properties of the coefficients b, f, σ and of the solution u, standard Itô developments in (4.7.9) would lead to $\mathcal{E}^2(\text{time}) = h^2$ as in [DMP96].

At the opposite, the second reason for which the global error is worse in our case, depends on the specific structure of the algorithm. Indeed, our choice to avoid linear interpolation procedures induces a rather large projection error $\mathcal{E}^2(\text{space})$. To reach a term of order one with respect to h for $\mathcal{E}^2(\text{space})$, we then need to take $\delta \equiv h^{3/2}$. This choice is far from being satisfactory and highly increases the complexity when the dimension d increases. Intuitively, there is no specific reason for such a relationship between δ and h: as explained in Subsection 4.1, δ has just to be small in front of h to take into account the influence of the drift b at the local level. For this reason, we aim to study in further investigations the convergence analysis of the algorithm when using a suitable "smooth" interpolation operator instead of a rough projection mapping. This point is discussed in a detailed way in the next subsection.

Further Comments on Errors. To conclude this subsection, we investigate the three last error terms $\mathcal{E}(\text{trunc})$, $\mathcal{E}(\text{quantiz})$ and $\mathcal{E}(\text{gradient}, p)$.

The truncation error decays linearly when the grid size increases. This control may seem rather poor to the reader. Recall indeed that $\mathcal{E}(\text{trunc})$ appears, up to the discretization procedure, as the probability that a diffusion process leaves a given bounded set. In the case of elliptic diffusions with bounded coefficients, it is well known that this probability decays exponentially fast as the size of the underlying set increases. Recall in this frame from Theorem 4.2.1 that the coefficients of the elliptic diffusion U are bounded. Note however that this rough argument fails in the discretized setting since there is no *a priori* sharp estimate of the approximate gradient $\overline{v}$ and thus of the associated approximate drift. This explains why our strategy to estimate $\mathcal{E}(\text{trunc})$ lies on the Bienaymé-Chebychev inequality and thus provides the current form given by Theorem 4.4.1.

Note finally that the errors associated to the quantization procedure, $\mathcal{E}^2(\text{quantiz})$, and to the probabilistic approximation of the gradient, $\mathcal{E}^2(\text{gradient}, p)$, are explicitly controlled in terms of M, h, and δ. They emphasize the price to pay to weaken the assumptions: we have to assume that the quantization grid is rather small compared to the spatial discretization one. Obviously, this increases the number of elementary operations of the algorithm and thus its total complexity. However, this does not affect so much the discretization procedure of the Gaussian law itself since quantization grids can be computed once for all apart from the implementation procedure of the algorithm.

9.2 Extensions and Further Investigations

We now discuss some possible extensions of our work.

Interpolation Procedure. As announced in the latter subsection, we first investigate the assets and liabilities of a smooth interpolation procedure. One of the main advantages of the spatial discretization proposed in Section 3.4, and then used in Algorithm 3.2, lies in its simplicity of implementation. However, from a purely mathematical point of view, this procedure may be rather awkward, since it ignores more or less the deep smoothness of the true solution u.

Note in this framework that the function Π_∞ may be seen as an operator acting on functions from $\mathbb{R}^d$ into $\mathbb{R}$. For such a function, the operator provides a rough interpolation of order 0 depending on the values of the function on the spatial mesh $\mathcal{C}_\infty$. As mentioned above, this interpolation procedure does not preserve the smoothness properties of the underlying function: in any cases, except if the function is constant, the interpolation procedure induces jumps of size of order δ. As a consequence, the distance between the function and the interpolated one is also of order δ.

A relevant strategy would consist in replacing the projection Π_∞ by a smoother interpolation operator. In our framework, an interpolation operator is said to be smooth if the distance between a given function ℓ and the interpolated one decreases with the regularity order of ℓ. For example, in dimension 1, the linear interpolation operator:

$$\ell \mapsto \big(x \mapsto \delta^{-1}(\delta + \delta\lfloor\delta^{-1}x\rfloor - x)\ell\big(\delta\lfloor\delta^{-1}x\rfloor\big) + \delta^{-1}(x - \delta\lfloor\delta^{-1}x\rfloor)\ell\big(\delta\lfloor\delta^{-1}x\rfloor + \delta\big)\big),$$

maps a $\mathcal{C}^2(\mathbb{R},\mathbb{R})$ function into a piecewise smooth function and the distance between them is of order δ^2.

Algorithm 3.2 can be written with respect to this new choice, but we also believe that the proof would be more difficult to detail. Moreover, smooth interpolation procedures in higher dimension slow down the running of the underlying algorithm.

Weakening Assumption. Note to conclude this subsection that some assumptions could be weakened. First, Theorem 4.2.1 still holds if b and f are just Hölder in x: in such a case, usual estimates of the gradient of u hold and Schauder's theory still applies. In particular, the reader can verify that Theorems 4.4.1 and 4.4.2 are still valid in this case (but Theorem 4.4.3 given in Subsection 4.3 is not).

Moreover, Algorithm 3.2 still converges if b, f, and σ depend on t in a Hölder way.

Finally, the following extension is conceivable. For $H \in C^{1+\alpha}$, $\alpha \in]0,1[$, the partial derivatives of order two of u have an integrable singularity in the neighborhood of T. In this frame, it would be interesting to adapt the Gronwall arguments given in Section 8.

9.3 Justification of Algorithm 3.2

We finally explain why we are not able to show the convergence of Algorithm 3.1.

Convergence of Algorithm 3.1. Recall that the main difference between Algorithm 3.1 and Algorithm 3.2 lies in the definition of the forward transitions. Indeed, in Algorithm 3.1:

$$\begin{aligned}\mathcal{T}(t_k, x) &\equiv b(x, \overline{u}(t_{k+1}, x), \overline{v}(t_{k+1}, x))h + \sigma(x, \overline{u}(t_{k+1}, x))g(\Delta B^k),\\ X_0 &\equiv x_0,\ \forall k \in \{0, ..., N-1\},\ X_{t_{k+1}} = \Pi_{k+1}\big(X_{t_k} + \mathcal{T}(t_k, X_{t_k})\big).\end{aligned}$$

Unfortunately, in this case, the well known BSDE machinery fails under Assumption **(A)**. At first sight, this could seem rather amazing. Indeed, recall that very strong *a priori* estimates of the solution u and of its partial derivatives hold in our framework. In particular, we could expect the discretization procedure of u and of its gradient to converge under such smoothness properties.

The main difficulty encountered to establish the convergence of Algorithm 3.1 appears in Section 7. More precisely, the lack of *a priori* controls of the regularity of $\overline{u}$ and $\overline{v}$ makes the stability strategy fruitless. Note indeed that inequality (4.7.1) becomes in the frame of Algorithm 3.1:

$$\begin{aligned}&|(\overline{u} - u)(0, x_0)|^2 + C_{4.7.1}^{-1} h \sum_{j=1}^N \mathbb{E}\big[|(\overline{v} - v)(t_{j-1}, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big]\\ &\quad\le C_{4.7.1}\Bigg[\mathbb{P}\{\tau_\infty < +\infty\} + \mathcal{E}^2(\text{time}) + \mathcal{E}^2(\text{space}) + \mathcal{E}^2(\text{quantiz})\\ &\qquad + \eta^{-1} h \sum_{j=1}^N \mathbb{E}\big[|(\overline{u} - u)(t_j, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big]\\ &\qquad + \eta^{-1} h \sum_{j=1}^N \mathbb{E}\big[|(\overline{u} - u)(t_{j-1}, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big]\\ &\qquad + (\eta + h)h \sum_{j=1}^N \mathbb{E}\big[|(\overline{v} - v)(t_j, X_{t_{j-1}})|^2 \mathbf{1}_{\{t_{j-1} < \tau_\infty\}}\big]\Bigg].\end{aligned} \tag{4.9.1}$$

Inequalities (4.7.1) and (4.9.1) just differ in the last term: $\hat{v}(t_{j-1}, X_{t_{j-1}})$ becomes $\overline{v}(t_j, X_{t_{j-1}})$. Note to be complete that a similar shift occurs in v, but, due to Theorem 4.2.1, it can be removed without any difficulties.

To apply the strategy used in Section 7, and in particular to derive an equivalent of Theorem 4.7.3 from (4.9.1), we then need to investigate the regularity in space of $\overline{v}$. According to the definition of $\overline{v}$, a first step then consists in studying the regularity in space of $\overline{u}$.

Lipschitz Control of $\overline{u}$. Note that the natural strategy to control the oscillations of $\overline{u}$ would consist in applying the usual FBSDE machinery to the triples $(X^{t_k,x}, Y^{t_k,x}, Z^{t_k,x})$ and $(X^{t_k,y}, Y^{t_k,y}, Z^{t_k,y})$ for $k \in \{0, ..., N-1\}$ and $x, y \in \mathcal{C}_k$. Of course, superscripts (t_k, x) and (t_k, y) denote the initial conditions of the Markov process X.

Nevertheless, we are not able to apply the strategies used in [Del02] and [Del03] to derive from the forward-backward writing local and global estimates of the discrete gradient of $\overline{u}$. There are two reasons to explain this failure.

First, the rough projection mapping chosen in Algorithms 3.1 and 3.2 induces an irreducible error greater than δ when estimating the difference between $\overline{u}(t_k, x)$ and $\overline{u}(t_k, y)$ in function of the parameters appearing in **(A)**. The strategy to overcome this difficulty is well known: the projection mapping has to be replaced by a smoother interpolation operator.

Second, any probabilistic strategy to estimate the Lipschitz constant of $\overline{u}$ in x such as the one exposed in [Del02] leads one way or another to the same difficulty as the one encountered to apply the stability procedure to Algorithm 3.1. More precisely, studying the difference between the triples $(X^{t_k,x}, Y^{t_k,x}, Z^{t_k,x})$ and $(X^{t_k,y}, Y^{t_k,y}, Z^{t_k,y})$, for $k \in \{0, ..., N-1\}$ and $x, y \in \mathcal{C}_k$, leads to investigate the regularity of $\overline{v}$. In short, one needs to estimate first the regularity of $\overline{v}$ to derive the one of $\overline{u}$. Intuitively, it is well understood that this is hopeless.

Appendix A

Various auxiliary results

1 Discrete sampling for the maximum of a Brownian motion: an alternative proof

In this subsection, the process $(X_t)_{t\in[0,T]}$ is a one dimensional Brownian Motion with dynamics $X_t = x_0 + \mu t + \sigma W_t, \sigma > 0$. In the above expression W is a standard Brownian Motion. Define $M_T := \max_{s\in[0,T]} X_s$, $M_T^N := \max_{s\in[0,T]} X_{\phi_s}$. We are interested in controlling the quantity $\mathrm{Err}_{\max}(h) := \mathbb{E}[M_T - M_T^N]$ w.r.t. h. Note that by definition one has $\mathrm{Err}_{\max}(h) \geq 0$.

Recall from Section 5.2 that we have the following

Theorem A.1.1 *Let X be a one dimensional BM following the above dynamics. There exists a constant $C > 0$ s.t. for h small enough*

$$\mathrm{Err}_{\max}(h) = \mathbb{E}[M_T - M_T^N] \leq C\sqrt{h}.$$

We provide an alternative proof to the one of Asmussen and *al.*, see Section 5.2 and [AGP95], and show this result using some of the techniques developed in section 2. Furthermore this proof also emphasizes we can weaken some assumptions in Theorems 2.2.1 and 2.4.1 in the special case of Brownian Motion. In the following of the section we stick to the notations introduced in Chapter 2.

Note first that

$$\mathrm{Err}_{\max}(h) = \sigma\mathbb{E}[\max_{s\in[0,T]}(\mu\sigma^{-1}s + W_s) - \max_{s\in[0,T]}(\mu\sigma^{-1}\phi(s) + W_{\phi(s)})].$$

Hence, w.l.o.g. we prove the result for a standard drifted Brownian Motion, i.e. setting $x_0 = 0, \sigma = 1$ in the previous definition.

Error decomposition of $\mathrm{Err}_{\max}(h)$

Let $X_t = \mu t + W_t$. Note that $M_T \leq \mu T + \max_{s\in[0,T]} W_s$. From Levy's identity $\max_{s\in[0,T]} W_s \underset{\mathcal{L}}{=} |W_T|$ and $M_T \geq M_T^N$, *a.s.* we first derive that $\mathrm{Err}_{\max}(h)$ is well defined. Put $\forall a \in \mathbb{R}^{+*}, \tau^a := \inf\{s \geq 0 : X_s = a\}$, $\tau^{a,N} := \inf\{s_i \geq \phi(\tau^a) + h : X_{s_i} \geq a\}$. It comes

$$\begin{aligned}
\mathrm{Err}_{\max}(h) &= \int_0^\infty da\mathbb{P}[M_T \geq a] - \mathbb{P}[M_T^N \geq a] = \int_0^\infty da\mathbb{P}[\tau^{a,N} > T] - \mathbb{P}[\tau^a > T] \\
&= \int_0^\infty da\mathbb{P}[\tau^a \leq T, \tau^{a,N} > T].
\end{aligned}$$

To get rid of the drift term we use the usual Girsanov transformation. Put $\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(-\mu W_t - \frac{\mu^2}{2}t\right) := \mathcal{E}(-\mu W_t)$. Under $\mathbb{Q}$ the process X is a standard Brownian Motion. It comes

$$\mathrm{Err}_{\max}(h) = \int_0^\infty da\mathbb{E}^{\mathbb{Q}}[\mathcal{E}(\mu X_T)\mathbf{1}_{\tau^a\leq T, \tau^{a,N}>T}].$$

Define now $\forall(a,t)\in\mathbb{R}^{+*}\times[0,T]$, $\tau_t^a:=\inf\{s\geq t: X_s\geq a\}, V_t^a:=\mathbb{E}^{\mathbb{Q}}[\mathcal{E}(\mu X_T)\mathbf{1}_{\tau_t^a>T}|\mathcal{F}_t]$. Set $V^a_{\phi(\tau^a)}=0$. Following the strategy of section 2 we write

$$\begin{aligned}
\mathrm{Err}_{\max}(h) &= \int_0^\infty da\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^a\leq T}\sum_{i=\phi(\tau^a)/h}^{N-1}\mathbb{E}^{\mathbb{Q}}[V^a_{t_{i+1}\wedge\tau^{a,N}}-V^a_{t_i\wedge\tau^{a,N}}|\mathcal{F}_{\tau^a}]]\\
&= \int_0^\infty da\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^a\leq T}\sum_{i=\phi(\tau^a)/h}^{N-1}\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^{a,N}>t_i,\tau^a_{t_i}\leq t_{i+1}}\mathbb{E}^{\mathbb{Q}}[V^a_{t_{i+1}}|\mathcal{F}_{\tau^a_{t_i}}]|\mathcal{F}_{\tau^a}]].
\end{aligned}$$

Proof of Theorem A.1.1

We split the former decomposition of $\mathrm{Err}_{\max}(h)$ into two parts. Namely,

$$\begin{aligned}
\mathrm{Err}_{\max}(h) &= \int_0^\infty da\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^a<T-h}\sum_{i=\phi(\tau^a)/h}^{N-2}\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^{a,N}>t_i,\tau^a_{t_i}\leq t_{i+1}}\mathbb{E}^{\mathbb{Q}}[V^a_{t_{i+1}}|\mathcal{F}_{\tau^a_{t_i}}]|\mathcal{F}_{\tau^a}]]\\
&\quad+\int_0^\infty da\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^a<T}\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^{a,N}>T-h,\tau^a_{T-h}\leq T}\mathbb{E}^{\mathbb{Q}}[V^a_T|\mathcal{F}_{\tau^a_{T-h}}]|\mathcal{F}_{\tau^a}]]:=A_1+A_2.
\end{aligned}$$

For $i\in[\![0,N-2]\!]$ on the set $\{\tau^a_{t_i}\leq t_{i+1}\}$ we derive from the Markov property of Brownian Motion and the joint law of the Brownian Motion and its maximum

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}[V^a_{t_{i+1}}|\mathcal{F}_{\tau^a_{t_i}}]=\\
&\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{X_{t_{i+1}}<a}\int_0^{a-X_{t_{i+1}}}ds\int_{-\infty}^s du\frac{2(2s-u)}{(2\pi(T-t_{i+1})^3)^{1/2}}\exp(-\frac{(2s-u)^2}{2(T-t_{i+1})})\exp(\mu(X_{t_{i+1}}+u)-\frac{\mu^2}{2}T)|\mathcal{F}_{\tau^a_{t_i}}]\\
&=\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{X_{t_{i+1}}<a}\int_{-\infty}^{a-X_{t_{i+1}}}du\exp(\mu(X_{t_{i+1}}+u)-\frac{\mu^2}{2}T)\int_{0\vee u}^{a-X_{t_{i+1}}}ds\frac{2(2s-u)}{(2\pi(T-t_{i+1})^3)^{1/2}}\exp(-\frac{(2s-u)^2}{2(T-t_{i+1})})|\mathcal{F}_{\tau^a_{t_i}}]\\
&=\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{X_{t_{i+1}}<a}\int_{-(a-X_{t_{i+1}})}^{a-X_{t_{i+1}}}du\exp(\mu(X_{t_{i+1}}+u)-\frac{\mu^2}{2}T)\frac{1}{(2\pi(T-t_{i+1}))^{1/2}}\exp(-\frac{u^2}{2(T-t_{i+1})})|\mathcal{F}_{\tau^a_{t_i}}]\\
&\leq C\exp(\mu a-\frac{\mu^2}{2}T)\mathbb{E}^{\mathbb{Q}}[\frac{|X_{t_{i+1}}-a|}{(T-t_{i+1})^{1/2}}|\mathcal{F}_{\tau^a_{t_i}}]\leq C\sqrt{h}\frac{\exp(\mu a)}{(T-t_{i+1})^{1/2}}
\end{aligned}$$

where we used Lemma 2.1.1 for the last inequality. Plugging this control into A_1 and using Lemma 2.2.3 that readily holds true uniformly in $a\in\mathbb{R}^{+*}$ under our current assumptions we get

$$A_1\leq C\sqrt{h}\int_0^\infty da\exp(\mu a)\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^a<T-h}\sum_{i=\phi(\tau^a)/h}^{N-2}\mathbf{1}_{\tau^{N,a}=t_{i+1}}(T-t_{i+1})^{-1/2}].$$

On the other hand, on the set $\{\tau^a_{T-h}\leq T\}$, $\mathbb{E}^{\mathbb{Q}}[V^a_T|\mathcal{F}_{\tau^a_{T-h}}]=\mathbb{E}^{\mathbb{Q}}[\exp(\mu W_T-\frac{\mu^2}{2}T)\mathbf{1}_{W_T<a}|\mathcal{F}_{\tau^a_{T-h}}]\leq\exp(\mu a-\frac{\mu^2}{2}T)$. Thus,

$$A_2\leq C\int_0^\infty da\exp(\mu a)\sum_{i=0}^{N-1}\mathbb{Q}[\tau^a\in[t_i,t_{i+1}),\tau^{a,N}=T].$$

We now state a technical Lemma whose proof is postponed to the end of the section.

Lemma A.1.2 *For all $i\in[\![1,N]\!]$ on the set $\{\tau^a<t_i\}$ put $\tilde{i}:=(t_i-\phi(\tau^a))/h$. One has*

$$\mathbb{Q}[\tau^{a,N}=t_i|\mathcal{F}_{\tau_a}]\leq C\left(\frac{1}{2}\right)^2\left(\frac{\Gamma(\tilde{i}-1/2)}{\Gamma(\tilde{i}+1)\Gamma(1/2)}+\frac{\Gamma(\lfloor\tilde{i}/2\rfloor+\tilde{i}\%2-1/2)}{\Gamma(\lfloor\tilde{i}/2\rfloor+\tilde{i}\%2+1)\Gamma(1/2)}\right):=\tilde{\psi}(\tilde{i})$$

where % denotes the rest of the Euclidean division.

Term A_2

Put $\forall i \in [\![0, N-1]\!],\ \tilde{p}_i = \tilde{\psi}(N-i)$. From Lemma A.1.2 we get

$$\begin{aligned}
A_2 &\leq C\int_0^\infty da \exp(\mu a) \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} ds \frac{a}{(2\pi s^3)^{1/2}} \exp\left(-\frac{a^2}{2s}\right) \tilde{p}_i \\
&\leq C\int_0^\infty da \exp(\mu a) \sum_{i=0}^{\lfloor \frac{(N-1)}{2} \rfloor} \int_{t_i}^{t_{i+1}} ds \frac{a}{(2\pi s^3)^{1/2}} \exp\left(-\frac{a^2}{2s}\right) \tilde{p}_i \\
&+ C\int_0^\infty da \exp(\mu a) \sum_{i=\lfloor \frac{(N-1)}{2} \rfloor}^{N-1} \int_{t_i}^{t_{i+1}} ds \frac{a}{(2\pi s^3)^{1/2}} \exp\left(-\frac{a^2}{2s}\right) \tilde{p}_i \\
&:= A_{21} + A_{22}.
\end{aligned}$$

From Stirling's Formula, i.e. $\Gamma(x) \underset{x\to\infty}{\sim} \sqrt{2\pi} x^{x-1/2} e^{-x}$, we derive that for x large enough

$$\frac{1}{2}\sqrt{2\pi} x^{x-1/2} e^{-x} \leq \Gamma(x) \leq \frac{3}{2}\sqrt{2\pi} x^{x-1/2} e^{-x}.$$

Thus, for h small enough, $\forall i \in [\![0, \lfloor \frac{(N-1)}{2} \rfloor]\!]$,

$$\begin{aligned}
\tilde{p}_i &\leq C\left(\frac{(N-i-1/2)^{N-i-1} e^{-(N-i-1/2)}}{(N-i+1)^{N-i+1/2} e^{-(N-i+1)}} + \frac{((N-i)/2-1/2)^{(N-i)/2-1} e^{-((N-i)/2-1/2)}}{((N-i)/2+1)^{(N-i)/2+1/2} e^{-((N-i)/2+1)}}\right) \\
&\leq C(N-i)^{-3/2} \leq Ch^{3/2}. \qquad \text{(A.1.1)}
\end{aligned}$$

Hence,

$$A_{21} \leq Ch^{3/2} \int_0^\infty da \exp(\mu a) \mathbb{Q}[\tau^a \in [0, T/2]] \leq Ch^{3/2} \int_0^\infty da \mathbb{P}[M_{T/2} \geq a] \leq Ch^{3/2}. \qquad \text{(A.1.2)}$$

On the other hand for $i \in [\![\lfloor (N-1)/2 \rfloor, N]\!]$,

$$\int_{t_i}^{t_{i+1}} ds \frac{a}{(2\pi s^3)^{1/2}} \exp(-\frac{a^2}{2s}) \leq Ch \frac{a}{(2\pi (T/2)^3)^{1/2}} \exp(-\frac{a^2}{T}) \leq Ch \exp(-ca^2/T).$$

From (A.1.1) we derive that $\exists C > 0, \displaystyle\sum_{i=\lfloor (N-1)/2 \rfloor}^{N} \tilde{p}_i \leq C$. Thus,

$$A_{22} \leq Ch \int_0^\infty da \exp(\mu a) \exp(-ca^2/T) \leq Ch. \qquad \text{(A.1.3)}$$

From (A.1.2) and (A.1.3) we get that there exists a constant $C > 0$ s.t. for h small enough

$$A_2 \leq Ch. \qquad \text{(A.1.4)}$$

Term A_1

Put $A_1^I := \displaystyle\int_0^\infty da \exp(\mu a) \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau^a < T-h} \sum_{i=\phi(\tau^a)/h}^{N-2} \mathbf{1}_{\tau^{N,a} = t_{i+1}} (T - t_{i+1})^{-1/2}]$. We aim to show that this quantity is bounded. Using the Markov property and the stationarity of the increments of Brownian motion we write

$$\begin{aligned}
A_1^I &= \int_0^\infty da \exp(\mu a) \int_0^{T-h} ds \frac{a}{(2\pi s^3)^{1/2}} \exp(-\frac{a^2}{2s}) \sum_{j=1}^{(T-(\phi(s)+h))/h} (T - (\phi(s) + jh))^{-1/2} \mathbb{P}[\tau^{N,0,s} = jh] \\
&:= \int_0^\infty da \exp(\mu a) \int_0^{T-h} ds \frac{a}{(2\pi s^3)^{1/2}} \exp(-\frac{a^2}{2s}) \sum_{j=1}^{(T-(\phi(s)+h))/h} (T - (\phi(s) + t_j))^{-1/2} \psi_s(t_j)
\end{aligned}$$

putting $\tau^{N,0,s} := \inf\{s_i \geq \phi(s)+h-s : \tilde{W}_{s_i} > 0\}$, where $\tilde{W}$ is a standard BM s.t. $\tilde{W}_{\phi(s)+h-s} = 0$. From Lemma A.1.2 we get $\forall s \in [0, T-h),\ j \in [\![1, (T-(\phi(s)+h))/h]\!],\ \psi_s(t_j) \leq C\tilde{\psi}(t_j/h)$. Thus,

$$\begin{aligned}
A_1^I &\leq h^{-1}\int_0^\infty da \exp(\mu a)\int_0^{T-h} ds\frac{a}{(2\pi s^3)^{1/2}}\exp(-\frac{a^2}{2s})\int_0^{(T-(\phi(s)+h))} dt(T-(\phi(s)+\phi(t)))^{-1/2}\tilde{\psi}(\phi(t)/h)\\
&\leq Ch^{-1}\int_0^{T-h} dt\tilde{\psi}(\phi(t)/h)\int_0^{T-h-t} ds(T-\phi(s)-\phi(t))^{-1/2}s^{-1/2}\int_0^\infty da\exp(\mu a)\exp\left(-c\frac{a^2}{s}\right)\frac{1}{(2\pi s)^{1/2}}\\
&\leq Ch^{-1}\int_0^{T-h} dt\tilde{\psi}(\phi(t)/h)\int_0^{T-h-t} ds(T-\phi(s)-\phi(t))^{-1/2}s^{-1/2} \leq C\sum_{i=0}^{N-1}\tilde{\psi}(i).
\end{aligned}$$

From (A.1.1), we derive that this series converges. Hence, $\exists C > 0$ s.t.

$$A_1 \leq C\sqrt{h}. \tag{A.1.5}$$

From equations (A.1.4) and (A.1.5) the proof of Theorem A.1.1 is complete.

□

Proof of Lemma A.1.2. Let $(Y_i)_{i\in\mathbb{N}^*}$ be a sequence of i.i.d. standard Gaussian variables. Define, for all $a \in [0,1]$, a random walk $(S_n^a)_{n\in\mathbb{N}}$ starting from 0 and such that $S_1^a = aY_1, \forall j \geq 2,\ S_j^a = aY_1 + \sum_{i=2}^{j} Y_i$.

From the Markov property of Brownian Motion we get that for all $i \in [\![1, N]\!]$ on the set $\{\tau^a < t_i\}$, $\mathbb{Q}[\tau^{a,N} = t_i|\mathcal{F}_{\tau^a}]$ corresponds to the probability that the first ladder epoch of $S^a, a = \left(\frac{\phi(\tau^a)+h-\tau^a}{h}\right)^{1/2}$ equals $i - \phi(\tau^a)/h$.

Recall now that for centered random walks with i.i.d. increments the generating function of the ladder epochs can be derived from Theorem 1 Chapter 12.7 in Feller, [Fel66]. Using the above definition, S^1 enters in this framework.

Set by convention $\cap_{j=k+1}^{k} A_j = \Omega$ for an arbitrary sequence of events $(A_j)_{j\geq 1}$ and an arbitrary integer k. Put $\forall n \in \mathbb{N}^*,\ p_n := \mathbb{P}[\cap_{j=1}^{n-1} S_j^1 \leq 0, S_n^1 > 0],\ \forall s \in [0,1],\ G(s) := \sum_{n\geq 1} p_n s^n$. One has

$$G(s) = 1 - (1-s)^{1/2}.$$

From this last expression we also derive by differentiation that

$$p_1 = 1/2,\ \forall n \geq 2,\ p_n = \frac{1}{2}\frac{\Gamma(n-1/2)}{\Gamma(n+1)\Gamma(1/2)}. \tag{A.1.6}$$

In our case we have to manage the time inhomogeneity of the first increment. To this end we introduce $\forall n \in \mathbb{N}^*, \forall a \in [0,1],\ p_n(a) = \mathbb{P}[\cap_{j=1}^{n-1} S_j^a \leq 0, S_n^a > 0]$. It comes $\forall a \in (0,1],\ p_1(a) = \frac{1}{2},\ \forall n \geq 2,\ p_n(1) = p_n, p_n(0) = \frac{1}{2}p_{n-1}$. We now aim to upper-bound uniformly in $a \in [0,1]$, $p_n(a)$ in terms of $(p_j)_{j\geq 1}$. Let us denote $g_s(x) = (2\pi s)^{-1/2}\exp(-x^2/2s)$ the density of the centered Gaussian law with variance $s > 0$. In the following, we use the convention $\Pi_{j=k+1}^{k} e_j = 1$ for an arbitrary sequence $(e_j)_{j\geq 1}$ and an arbitrary integer k.

For $a \in (1/2, 1]$, we get for $n \geq 2$

$$\begin{aligned}
p_n(a) &= \int_{-\infty}^0 dy_1 g_a(y_1)\left(\prod_{j=2}^{n-1}\int_{-\infty}^{-\sum_{i=1}^{j-1} y_i} dy_j g_1(y_j)\right)\int_{-\sum_{i=1}^{n-1} y_i}^\infty dy_n g_1(y_n)\\
&\leq \sqrt{2}p_n.
\end{aligned} \tag{A.1.7}$$

For $a \in [0, 1/2], n \geq 2$,

$$\begin{aligned}
p_n(a) &= \int_{-\infty}^0 dy_1 g_1(y_1)\left(\mathbf{1}_{n\geq 3}\int_{-\infty}^0 dy_2 g_1(y_2 - a^{1/2}y_1)\left(\prod_{j=3}^{n-1}\int_{-\infty}^0 dy_j g_1(y_j - y_{j-1})\right)\int_0^\infty dy_n g_1(y_n - y_{n-1})\right.\\
&\quad \left. + \mathbf{1}_{n=2}\int_0^\infty dy_2 g_1(y_2 - a^{1/2}y_1)\right).
\end{aligned}$$

Since we are mainly concerned with asymptotic behaviours of p_n w.r.t. n we now focus on the case $n \geq 3$. For $n = 2$ we simply remark that $\forall a \in [0, 1/2]$, $p_n(a) \leq 1/2$. Using the simple inequality $|y_2 - a^{1/2} y_1|^2 \geq \frac{1}{2}|y_2|^2 - a|y_1|^2$ we obtain

$$\begin{aligned} p_n(a) &\leq \sqrt{2} \int_{-\infty}^{0} dy_1 g_1((1-a)y_1) \int_{-\infty}^{0} g_2(y_2) \left(\prod_{j=3}^{n-1} \int_{-\infty}^{0} dy_j g_1(y_j - y_{j-1}) \right) \int_0^{\infty} dy_n g_1(y_n - y_{n-1}) \\ &\leq C\mathbb{P}[\tilde{Y}_1 \leq 0, \cap_{j=3}^{n-1} \tilde{Y}_1 + \sum_{i=3}^{j} Y_i \leq 0, \tilde{Y}_1 + \sum_{i=3}^{n} Y_i > 0] \end{aligned}$$

where $\tilde{Y}_1 \underset{\mathcal{L}}{=} \mathcal{N}(0,2)$ and is independent of $(Y_i)_{i \geq 3}$. Define $\forall i \geq 2, \tilde{Y}_i = Y_{2i-1} + Y_{2i}$. By construction, the $(\tilde{Y}_i)_{i \geq 1}$ form a sequence of i.i.d. centered Random variables. Hence, the law of the ladder epochs of the associated random walk $\tilde{S}$ is also given by (A.1.6). It comes,

$$p_n(a) \leq C \left(\mathbb{P}[\cap_{j=1}^{\lfloor n/2 \rfloor - 1} \tilde{S}_j \leq 0, \tilde{S}_{\lfloor n/2 \rfloor} > 0] \mathbf{1}_{n\%2=0} + \mathbb{P}[\cap_{j=1}^{\lfloor n/2 \rfloor} \tilde{S}_j \leq 0, \tilde{S}_{\lfloor n/2 \rfloor} + Y_n > 0] \mathbf{1}_{n\%2=1} \right).$$

For odd n one has $\{\tilde{S}_{\lfloor n/2 \rfloor} \leq 0, \tilde{S}_{\lfloor n/2 \rfloor} + Y_n > 0\} \subset \{\tilde{S}_{\lfloor n/2 \rfloor} \leq 0, \tilde{S}_{\lfloor n/2 \rfloor} + \sqrt{2} Y_n > 0\}$. We finally derive

$$p_n(a) \leq C p_{\lfloor n/2 \rfloor + n\%2}. \tag{A.1.8}$$

The Lemma is now a simple consequence of (A.1.6), (A.1.7) and (A.1.8).

□

Remark A.1.1 *In Mathematical Finance, the result of Theorem A.1.1 has applications for the evaluation of look-back options. In a Black-Scholes framework, under which the price dynamics of the underlying asset,* $(S_t)_{t \in [0,T]}$*, follows a Geometric Brownian Motion:* $S_t = S_0 \exp(\sigma W_t + (\mu - \sigma^2/2)t)$*,* $\sigma > 0$*, the price of such options writes* $\mathbb{E}[\psi(\exp(X_T), \exp(M_T))]$ *for a given pay-off function* ψ*. For Lipschitzian functions* ψ*, the proof of Theorem A.1.1 provides an upper bound of order 1/2 w.r.t.* h *for the weak error associated to the discrete observation of the maximum of the price of the underlying asset. This had been observed in the Ph. Dissertation of Seumen-Tonou, [ST97], who employed rather heavy analytical techniques.*

Remark A.1.2 *The above proof of Theorem A.1.1 also emphasizes that, for the Brownian Motion, Theorems 2.2.1 and 2.4.1 remain valid without the support condition appearing in* **(G1)**.

Appendix B

Index of the assumptions

In this appendix, we sum up our various assumptions.

Chapter 1

Assumption	Meaning	Defined at page
(C)	Uniform non characteristic boundary condition	19
(D-k), $k \in \mathbb{N}$	The domain D is of class C^k and has compact boundary	19
(D-l), $l \notin \mathbb{N}$	The domain D is of class H^l and has compact boundary	20
(F1)	$f \geq 0$ is Borel and bounded, $d(\text{supp}(f), \partial D) \geq 2\varepsilon > 0$	20
(F2-l), $l \notin \mathbb{N}$	f is of class $H^l(\bar{D})$	20
(H')	"Usual" hypoellipticity condition on the coefficients b, σ	18
(H)	Strong hypoellipticity condition on σ	18
(S-k), $k \in \mathbb{N}$	The coefficients b, σ are bounded in $C^k(\bar{D})$	18
(S-l), $l \notin \mathbb{N}$	The coefficients b, σ are bounded in $H^l(\bar{D})$	20
(UE)	Uniform ellipticity on the diffusion coefficient σ	18

Chapter 2

Assumption	Meaning	Defined at page
(C)	Almost sure uniform non characteristic boundary condition	56
(C')	Almost sure uniform non characteristic boundary condition associated to **(D')**	65
(D-k), $k \in \mathbb{N}$	The domain D is of class C^k and has compact boundary	56
(D')	$D = \cap_{i=1}^m D_i, m \in \mathbb{N}^*$, where D_i satisfies **(D-2)**	65
(G1)	g is Borel and bounded, $\forall t \in [0,T],\ d(\text{supp}(g(t,.), \partial D) \geq 2\varepsilon > 0$	56
(G2)	g is bounded in $C^{1,2}([0,T] \times \mathbb{R}^d)$	56
(MID)	X is real valued and has a smooth u.e. Markovian diffusion term	68
(S)	"Continuity" condition on σ	56

Chapter 3

Assumption	Meaning	Defined at page
(D)	$D \subset \mathbb{R}^d$ is an orthant, i.e. $D = \cap_{i=1}^m \{y \in \mathbb{R}^d : y_i > b_0^i\}$	75
(BM)	X is a BM of the form $X = x + \sigma W$, W standard BM	75
(S)	$v \in C_b^{2,4}([0,T] \times \mathbb{R}^d)$	75
(S')	Alternative smoothness assumption for v,	85
(F1)	Smoothness and compatibility conditions on f	88
(F1')	Alternative smoothness assumption on f	88
(F2)	Alternative smoothness assumption on f	95

Chapter 4

Assumption	Meaning	Defined at page
(A.1)	Boundedness of (b, σ, f, H) in x, linear growth in y, z	108
(A.2)	Uniform ellipticity of $\sigma\sigma^*$	108
(A.3)	Lipschitz continuity in all variables of (b, σ, f, H)	108
(A.4)	Boundedness of H in $C^{2+\alpha}(\mathbb{R}^d)$	108

Bibliography

[ABR96] L. Andersen and R. Brotherton-Ratcliffe. Exact exotics. *Risk*, 9:85–89, 1996.

[AGP95] S. Asmussen, P. Glynn, and J. Pitman. Discretization error in simulation of one-dimensional reflecting Brownian motion. *Ann. Appl. Probab.*, 5(4):875–896, 1995.

[AK80] A. Azzam and E. Kreyszig. On parabolic equations in n space variables and their solutions in regions with edges. *Hokkaido Math. J.*, 9:140–154, 1980.

[AK81] A. Azzam and E. Kreyszig. Smoothness of solutions of parabolic equations in regions with edges. *Nagoya Math. J.*, 84:159–168, 1981.

[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.

[AS72] M. Abramowitz and I.A Stegun. *Handbook of mathematical functions.* Dover Publications, 1972.

[Azz85] A. Azzam. On mixed boundary value problems for parabolic equations in singular domains. *Osaka J. Math.*, 22:691–696, 1985.

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

[BCI99] P. Baldi, L. Caramellino, and M.G. Iovino. Pricing general barrier options: a numerical approach using sharp large deviations. *Math. Finance*, 9(4):293–322, 1999.

[BDZ97] D.R. Beaglehole, P.H. Dybvig, and G. Zhou. Going to extremes: Correcting simulation bias in exotic option valuation. *Fin. Anal. Jour.*, 53:62–68, 1997.

[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.

[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.

[BL94] P.P. Boyle and S.H. Lau. Bumping up against the barrier with the binomial method. *Jour. of Derivat.*, 1:6–14, 1994.

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

[BPP02] V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multi-dimensional American options. *Tech. report of the university Paris VI, Lab. Prob. et Mod. Al., To appear in Math. Fin.*, 753, 2002.

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

[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.

[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.

[Cat92] P. Cattiaux. Stochastic calculus and degenerate boundary value problems. *Ann. Inst. Fourier, Grenoble*, 42-3:541–624, 1992.

[Cat03] P. Cattiaux. Personal communication. 2003.

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

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

[CKG03] C. Costantini, N. El Karoui, and E. Gobet. Représentation de Feynman-Kac dans des domaines temps-espace et sensibilité par rapport au domaine. *C.R. Acad. Sci. Paris*, 337:337–342, 2003.

[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.

[Del03] F. Delarue. Estimates of the solutions of a system of quasilinear PDEs. A probabilistic scheme. *Séminaire de Probabilités*, XXXVII:290–332, 2003.

[Del04] F. Delarue. Auxiliary SDEs for homogenization of quasilinear PDEs with periodic coefficients. *Annals of Probability*, 32:2305–2361, 2004.

[Die71] J. Dieudonné. *Eléments d'analyse, vol. 1.* Gauthiers-Villars, 1971.

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

[Fel66] W. Feller. *An Introduction to Probability Theory and its Applications, vol. 2.* Wiley, 1966.

[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.

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

[GLW04] E. Gobet, J.P. Lemor, and X. Warin. A regression-based Monte-Carlo method to solve backward stochastic differential equations. *Tech. Report of the CMAPX*, 533, 2004.

[Gob98] E. Gobet. *Schémas d'Euler pour diffusion tuée. Application aux options barrière.* PhD Thesis, University Paris VII., 1998.

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

[GS69] I.I. Gikhman and A.V. Skorokhod. *Introduction to the theory of random processes.* Saunders Math. Books, 1969.

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

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

[Iye85] S. Iyengar. Hitting Lines with Two-Dimensional Brownian Motion. *SIAM J. Appl. Math.*, 45–6:983–989, 1985.

[Jon68] A.L. Jones. An Extension of an Inequality involving modified Bessel functions. *J. Math. Physics*, 48:220–221, 1968.

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

[KS91] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus.* Second Edition, Springer Verlag, 1991.

[KS98] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance.* Appl. of Math. 39, Springer Verlag, 1998.

[Lie85] G.M. Lieberman. Regularized distance and its applications. *Pacific J. Math.*, 117(2):329–352, 1985.

[LPS98] B. Lapeyre, E. Pardoux, and R. Sentis. *Méthodes de Monte-Carlo pour les équations de transport et diffusion*, volume 29. Mathématiques et Applications-Springer, 1998.

[LS89] H.R. Lerche and D. Siegmund. Approximate exit probabilities for a Brownian bridge on a short time interval, and applications. *Adv. in Appl. Probab.*, 21(1):1–19, 1989.

[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.

[Mak03] R. N. Makarov. Numerical solution of quasilinear parabolic equations and Backward Stochastic Differential Equations. *Russian J. Numer. Anal. Math. Modelling*, 18–5:397–412, 2003.

[MP90] D. Michel and E. Pardoux. An introuction to Malliavin calculus and some of its applications. *Recent Adv. in Stoch. Proc.*, pages 65–104, 1990.

[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.

[MT99] G. N. Milstein and M. V. Tretyakov. Numerical algorithms for semilinear parabolic equations with small parameter based on approximation of stochastic equations. *Math. Comp.*, 69–229:237–267, 1999.

[MY99] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their applications.* LNM-1702, Springer-Verlag, 1999.

[Nua95] D. Nualart. *Malliavin calculus and related topics.* Springer Verlag, 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.

[PP90] E. Pardoux and S.G. Peng. Adapted solution of a Backward Stochastic Differential Equation. *Systems Control Lett*, 14-1:55–61, 1990.

[PP04] G. Pagès and J. Printems. Functional quantization for numerics: pricing asian options. *Tech. report of the university Paris VI, Lab. Prob. et Mod. Al.*, 900, 2004.

[PPP04] G. Pagès, H. Pham, and J. Printems. Optimal quantization methods and applications to numerical problems in finance. To appear in Handbook on Numerical Methods in Finance (S. Rachev, ed.), Birkhauser, 2003. *Tech. report of the university Paris VI, Lab. Prob. et Mod. Al.*, 813, 2004.

[PT99] E. Pardoux and S. Tang. Forward-Backward Stochastic Differential Equations and quasilinear parabolic PDEs. *Probab. Theory Related Fields*, 114-2:123–150, 1999.

[Rog85] L.C.G. Rogers. Smooth transition densities for one-dimensional diffusions. *Bull. London Math. Soc.*, 17:157–161, 1985.

[RR91] M. Rubinstein and E. Reiner. Breaking down the barriers. *Risk*, 4-8:28–35, 1991.

[RT99] C.R Rao and H. Toutenburg. *Linear models.* Springer series in statistics. Springer-Verlag. New-York, second edition, 1999.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion. 3rd ed.* Grundlehren der Mathematischen Wissenschaften. 293. Berlin: Springer, 1999.

[Shi96] A.N. Shiryaev. *Probability, Second Edition.* Graduate Texts in Mathematics, 95. Springer-Verlag, New York., 1996.

[Sie79] D. Siegmund. Corrected diffusion approximations in certain random walk problems. *Adv. in Appl. Probab.*, 11(4):701–719, 1979.

[ST97] P. Seumen-Tonou. Méthodes Numériques Probabilistes pour la Résolution d'Equations du Transport et pour l'évaluation d'Options Exotiques. *Ph. Dissertation. Université de Provence.*, 1997.

[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.

[Whi73] N. Whitham. *Linear and Non linear waves.* Wiley, 1973.

[Woy98] W.A. Woyczyński. *Burgers-KPZ turbulence.* LNM-1700, Springer-Verlag, 1998.