Stratégies de couverture presque optimale : théorie et applications

Nicolas Landon

Ecole Polytechnique

ÉCOLE DOCTORALE : Mathématiques Appliquées

T H È S E

pour obtenir le titre de

Docteur de l'Ecole Polytechnique

Spécialité : MATHÉMATIQUES APPLIQUÉES

Présentée et soutenue par

Nicolas LANDON

Almost sure optimal stopping times : theory and applications.

Thèse préparée au CMAP et à GDF-SUEZ

soutenue le 04 Février 2013.

Jury :

Directeur:	Emmanuel GOBET	-	Ecole Polytechnique
Rapporteurs:	Damien LAMBERTON	-	Université Paris-Est Marne-La-Vallée
	Mathieu ROSENBAUM	-	Université Paris 6 - Pierre et Marie-Curie
Président de jury:	Nicole EL KAROUI	-	Université Paris 6 - Pierre et Marie-Curie
Examinateurs:	Nizar TOUZI	-	Ecole Polytechnique
	Peter TANKOV	-	Université Paris 7 - Denis-Diderot
	Oliver BARDOU	-	GDFSUEZ
Invité:	Céline JÉRUSALEM	-	GDFSUEZ

Remerciements

Mes premiers remerciements vont tout naturellement à mon directeur de thèse Emmanuel Gobet. Je le remercie pour son encadrement et pour m'avoir toujours laissé la liberté dans mes recherches. Son énergie et sa passion des mathématiques, véhiculés constamment à chacune de nos rencontres, est contagieuse. Je mesure aujourd'hui la chance que j'ai eue d'avoir été formé par une telle figure de la recherche.

Je remercie également tous les membres du jury pour le temps qu'ils m'ont consacré. Mes remerciements vont tout particulièrement à Damien Lamberton et à Mathieu Rosenbaum, qui ont accepté de rapporter ma thèse. Je leur suis reconnaissant pour la lecture très attentive du manuscrit et leurs nombreuses remarques constructives.

J'exprime toute ma gratitude à Nicole El Karoui, Nizar Touzi et Peter Tankov pour l'honneur qu'ils me font d'avoir acceptés de faire partie de mon jury. Je tiens aussi à remercier tous ceux qui font du Centre de Mathématiques Appliquées de l'Ecole Polytechnique un lieu exceptionnel, sans citer de nom, merci à tous.

Je remercie Olivier Bardou et Céline Jérusalem de faire partie du jury. Bien évidemment, je suis très reconnaissant à GDFSUEZ pour avoir accepté de financer ma thèse. Je salue aussi tous les membres de l'équipe de Céline Jérusalem avec qui j'ai partagé ses trois années.

Contents

Résumé : Cette thèse comporte 8 chapitres.

Le chapitre 1 est une introduction aux problématiques rencontrées sur les marchés énergétiques : fréquence d'intervention faible, coûts de transaction élevés, évaluation des options spread.

Le chapitre 2 étudie la convergence de l'erreur de couverture d'une option call dans le modèle de Bachelier, pour des coûts de transaction proportionnels (modèle de Leland-Lott) et lorsque la fréquence d'intervention devient infinie. Il est prouvé que cette erreur est bornée par une variable aléatoire proportionnelle au taux de transaction. Cependant, les démonstrations de convergence en probabilité demandent des régularités sur les sensibilités assez restrictives en pratique. Les chapitres suivants contournent ces obstacles en étudiant des convergences presque sûres.

Le chapitre 3 développe tout d'abord de nouveaux outils de convergence presque sûre. Ces résultats ont de nombreuses conséquences sur le contrôle presque sûr de martingales et de leur variation quadratique, ainsi que de leurs incréments entre deux temps d'arrêt généraux. Ces résultats de convergence trajectorielle sont connus pour être difficiles à obtenir sans information sur les lois. Par la suite, nous appliquons ces résultats à la minimisation presque sûre de la variation quadratique renormalisée de l'erreur de couverture d'une option de payoff général (cadre multidimensionnel, payoff asiatique, lookback) sur une large classe de temps d'intervention. Une borne inférieure à notre critère est trouvée et une suite minimisante de temps d'arrêt optimale est exhibée : il s'agit de temps d'atteinte d'ellipsoïde aléatoire, dépendant du gamma de l'option.

Le chapitre 4 étudie la convergence de l'erreur de couverture d'une option de payoff convexe (dimension 1) en prenant en compte des coûts de transaction à la Leland-Lott. Nous décomposons l'erreur de couverture en une partie martingale et une partie négligeable, puis nous minimisons la variation quadratique de cette martingale sur une classe de temps d'atteintes générales pour des Deltas vérifiant une certaine EDP non-linéaire sur les dérivées secondes. Nous exhibons aussi une suite de temps d'arrêt atteignant cette borne. Des tests numériques illustrent notre approche par rapport à une série de stratégies connues de la littérature.

Le chapitre 5 étend le chapitre 3 en considérant une fonctionnelle des variations discrètes d'ordre β_Y et de β_Z de deux processus d'Itô Y et Z à valeurs réelles, la minimisation étant sur une large classe de temps d'arrêt servant au calcul des variations discrètes. Borne inférieure et suite minimisant sont obtenues. Une étude numérique sur les coûts de transaction est faite.

Le chapitre 6 étudie la discrétisation d'Euler d'un processus multidimensionnel X dirigé par une semi-martingale d'Itô Y. Nous minimisons sur les temps de la grille de discrétisation un critère quadratique sur l'erreur du schéma. Nous trouvons une borne inférieure et une grille optimale, ne dépendant que des données observables.

Le chapitre 7 donne un théorème limite centrale pour des discrétisations d'intégrale stochastique sur des grilles de temps d'atteinte d'ellipsoïdes adaptées quelconque. La corrélation limite est conséquence d'asymptotiques fins sur les problèmes de Dirichlet.

Dans le chapitre 8, nous nous intéressons aux formules d'expansion pour les options sur spread, pour des modèles à volatilité locale. La clé de l'approche consiste à conserver la propriété de martingale de la moyenne arithmétique et à exploiter la structure du payoff call. Les tests numériques montrent la pertinence de l'approche.

Mots clés : Convergence presque sure, discrétisation d'intégrale, couverture d'option, coût de transaction, schéma d'Euler-Maruyama, convergence en loi, option sur spread, calcul d'expansion

Abstract: This thesis has 8 chapters.

The chapter 1 is an introduction to the issues encountered in the energy market : low frequency trading, high transaction costs, spread option pricing.

The chapter 2 studies the hedging error convergence of a call option in the Bachelier model, for proportional transaction costs (Leland-Lott's model) and when the intervention frequency becomes infinite. We prove that this error is bounded by a random variable proportional to the convergence rate. However, the proof of convergence in probability requires some restrictive regularities on the sensitivities. The following chapters avoid these difficulties by studying the almost sure convergence.

The chapter 3 develop new tools for the almost sure convergence. These results have many consequences on the control path by path of martingales and of their quadratic variations, as their increments between two general stopping times. These convergence results are well-known to be difficult to demonstrate without any information on the laws. Moreover, we apply these results to the almost sure minimization of the renormalized quadratic variation of the hedging error for a general payoff (multidimensional setting, Asian and Lookback option) for a broad class of trading dates. A lower bound for our criterion is found and an optimal sequence of stopping times is described, which is given by hitting times of random ellipsoids, depending only on the option gamma.

The chapter 4 studies the hedging error convergence of an option with convex payoff (dimension 1) taking into account Leland-Lott's transaction costs. We decompose the error into a martingale part and a negligible part, then we minimize the quadratic variation of this martingale on a class of hitting times for Deltas satisfying some non-linear EDP on the second derivative. Moreover, we find a minimizing sequence of hitting times. Numerical tests illustrate our approach w.r.t. a series of strategies from the literature.

The chapter 5 extends the chapter 3 by considering a discrete variation functional of order β_Y and β_Z for two Ito processes Y and Z; the minimization is on a broad class of stopping times. Lower bound and minimizing sequence are obtained. A numerical study on the transaction costs is done.

The chapter 6 studies the Euler discretization of a multidimensional process X, controlled by a semi-martingale Y. We lessen some quadratic criterion on the error scheme over the discretization time grid. We find a lower bound and an optimal grid, independent of the observable data.

The chapter 7 gives a Limit Central Theorem for the discretization of stochastic integrals on hitting times of any adapted ellipsoids. The asymptotic correlation is a consequence of sharp limits involving solutions to Dirichlet's problem.

In the chapter 8, we are interested to expansion formulas for spread options in local volatility model. The originality of our approach is to keep the martingale property for the approximation and to exploit the call payoff structure. Numerical tests show the relevance of our approach.

Keywords: Almost sure convergence, integral discretization, option hedging, transaction costs, Euler's scheme, convergence in law, spread option, expansion calculus.

Chapter 1

Introduction

Contents

1.1 Introduction aux problématiques des marchés des énergies.

La gestion des risques financiers pour GDF-SUEZ (et plus généralement pour tous les énergéticiens) est devenue une activité incontournable depuis l'ouverture à la concurrence des marchés énergétiques. Contrairement aux marchés financiers traditionnels, les marchés énergétiques se distinguent par plusieurs aspects mis en avant dans cette thèse.

1. La couverture d'un contrat financier sur l'énergie est plus complexe de par la nature des marchés des commodités : fréquence de réajustement hebdomadaire qui induit une erreur de couverture plus grande que sur les marchés monétaires, où la fréquence est beaucoup plus élevée. Ainsi, décider d'une date optimale d'intervention sur le marché revêt une importance toute particulière. La couverture se faisant en général à l'aide de futures avec différentes périodes de livraison et les modèles rendant compte de plusieurs facteurs de risque, le problème apparait nécessairement multidimensionnel.

2. Si l'on prend l'actif physique comme objet de couverture, alors des coûts de transport et d'injection/soutirage (cas d'un stockage gazier) apparaissent. De plus, la fourchette bid/ask parfois de l'ordre de 1% est "titanesque" comparé au 0.2% rencontré en moyenne sur les autres marchés. Toutes ces frictions rendent imparfaite la couverture et incitent à ne pas négliger leurs effets.

3. Les options dites "spread" sont très courantes sur les marchés des commodités, ce sont des options portant sur la partie positive d'une différence d'indices. Elles servent entre autres à modéliser le fonctionnement d'une centrale à gaz par exemple. Traditionnellement, les méthodes de Monte Carlo sont utilisées, mais mènent à de long temps de calcul. Des formules analytiques approchées sont donc d'un grand intérêt.

1.2 Étude de convergence presque sûre d'intégrales stochastiques.

1.2.1 Introduction informelle.

Depuis la construction du mouvement brownien de Wiener (1923) et la théorie de l'intégrale stochastique par Itô (1951), les processus à temps continu sont devenus indispensables dans beaucoup de disciplines, de la physique à la finance. Cependant, les processus continus ne sont souvent qu'une approximation d'un phénomène à temps discret. En finance, l'interprétation en terme d'erreur de couverture suite à la vente d'une option et à la couverture à temps discret offre un exemple flagrant. Sur les marchés, un investisseur ne peut pas se couvrir en continu, cela est dû d'une part aux cotations discrètes des prix et d'autre part à l'accumulation des coûts de transaction qui incite fortement le trader à laisser du risque dans sa position ! Évidemment, nous comprenons bien que les processus continus sont plus simples à utiliser et la théorie sous-jacente est beaucoup plus fournie.

Leland (1985) fût un des premiers auteurs à étudier le problème de convergence de l'erreur de couverture dans un marché avec coûts de transaction. Une suite abondante d'articles de [Hodges 1989], [Bensaid 1992], [Henrotte 1993], [Avellaneda 1994], [Grannan 1996], [Kabanov 1997], [Pergamenshchikov 2003], [Sekine 2008], [Denis 2010b] parmi d'autres étudient des stratégies en présence de coûts de transaction.

Ainsi, les chapitres 3, 4, 5, 6 et 7 de cette thèse s'intéressent à une partie délicate de la discrétisation de processus : étudier le comportement asymptotique d'intégrales stochastiques, discrétisées à des temps aléatoires (plus précisément des temps d'arrêt), mais contrairement à la littérature ambiante sur le sujet qui s'intéresse à la limite de l'erreur renormalisée en loi ou dans $L^2(\Omega)$, nous avons choisi la convergence *presque sûre*, exception faite au chapitre 7, qui est singulier par rapport aux autres chapitres, où l'on s'intéresse à la convergence en loi. La convergence presque sûre a tout son sens pour la couverture des risques, où l'on ne peut pas rajouter les scenarii de marché. Pour accomplir cette tâche, nous avons construit de nouveaux outils, simples et efficaces, qui nous permettent de contrôler aussi bien des martingales locales continues que leurs incréments entre deux temps aléatoires. Cette partie étend la littérature sur l'approximation d'intégrale stochastique presque sûrement par [Bichteler 1981] et [Karandikar 1981], [Karandikar 1989], [Karandikar 1991], [Karandikar 1995], [Karandikar 2006].

1.2.2 Résultat de convergence presque sûre d'intégrales stochastiques.

Dans les livres classiques sur le calcul stochastique (par exemple, [Karatzas 1988], [Protter 2004]), l'un des premiers résultats que l'on nous apprend est que l'on ne peut pas définir l'intégrale stochastique de manière trajectorielle. Si l'on considère le cas simple de l'intégrale stochastique par rapport au mouvement brownien, l'approche d'Itô est justement de regarder des convergences en moyenne quadratique d'intégrales stochastiques discrètes pour des intégrandes progressivement mesurables (qui est une classe assez large). L'idée naturelle pour définir l'intégrale stochastique trajectoire par trajectoire est de restreindre cette classe d'intégrande à celle des processus continus à droite avec une limite à gauche (càdlàg). A ma connaissance, Bichteler est le premier à énoncer des résultats de convergence presque sûre d'intégrales stochastiques discrétisées, le théorème ci-dessous permet de prouver la convergence de l'intégrale stochastique discrète vers l'intégrale stochastique trajectoriellement :

Theorem 1.2.1 ([Karandikar 1995]). *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé supportant un mouvement brownien W standard et $(\mathcal{F}_t)$ une filtration satisfaisant les conditions habituelles. Soit f un processus càdlàg adapté et pour $n \geq 1$, soit $\{\tau_i^n : i \geq 0\}$ définie par $\tau_0^n = 0$ et pour $i \geq 0$*

$$\tau_{i+1}^n := \inf\{t \geq \tau_i^n : |f_t - f_{\tau_i^n}| \geq 2^{-n}\}.$$

Soit (Y_t^n) défini par : pour tout $\tau_k^n \leq t < \tau_{k+1}^n$, $k \geq 0$,

$$Y_t^n = \sum_{i=0}^{k-1} f_{\tau_i^n}(W_{\tau_{i+1}^n} - W_{\tau_i^n}) + f_{\tau_k^n}(W_t - W_{\tau_k^n}).$$

Alors, pour tout $T < \infty$,

$$\sup_{0\le t\le T}\left|Y_t^n - \int_0^t f_s \mathrm{d}W_s\right| \to 0 \quad p.s.$$

Cet élégant résultat énonce que si l'on a un contrôle assez fort sur la différence entre l'intégrande et son approximation uniformément en temps, alors on peut avoir une convergence uniforme trajectoire par trajectoire sur les intégrales stochastiques. Ce résultat est dû initialement à Bichteler, mais l'approche a été pleinement simplifiée par Karandikar par la suite. Un théorème qui permet de mettre en perspective cette idée de contrôle est le suivant :

Theorem 1.2.2 ([Bichteler 1981])**.** *Soit X une semi-martingale et f^n, f deux processus prévisibles localement bornés tels que*

$$\sup_{0\le t\le T}|f_t^n - f_t| \le \varepsilon_n,$$

où ε_n est une suite de carré sommable (c.à.d. $\sum_{n\ge 0}\varepsilon_n^2 < +\infty$). Soit $Y_t^n = \int_0^t f_s^n \mathrm{d}X_s$ et $Y_t = \int_0^t f_s \mathrm{d}X_s$. Alors,

$$\sup_{0\le t\le T}|Y_t^n - Y_t| \to 0 \quad p.s. \quad \textit{pour tout } T < \infty.$$

Dans le chapitre 3, nous développons un lemme ayant des similitudes, qui permet d'avoir des résultats de convergence presque sûre dans un cadre très général. Nous ne contrôlons pas nécessairement les intégrandes, mais une semi-martingale d'Ito continue, qui induit un contrôle sur beaucoup de semi-martingales d'Ito continues. Plus précisément, considérant$S$ une semi-martingale d'Ito continue vérifiant une hypothèse d'ellipticité trajectorielle, nous étudions des suites de temps d'arrêt $\mathcal{T}^n := \{\tau_0^n = 0 < \tau_1^n < ... < \tau_i^n < ... \le \tau_{N_T^n}^n\}$, pour $n = 0, 1, \dots$, où nous définissons un cadre asymptotique, lorsque le paramètre n tend vers l'infini de la manière suivante : on se donne une suite $(\varepsilon_n)_{n\ge 0}$ de carré sommable, i.e. $\sum_{n\ge 0}\varepsilon_n^2 < +\infty$,

1. le paramètre $\varepsilon_n^{-2\rho_N}$ (pour un certain $\rho_N \in [1, \rho_N^{\max}[$, où $\rho_N^{\max}$ est une certaine constante) borne supérieurement le nombre de temps d'arrêt

$$\sup_{n\ge 0}\left(\varepsilon_n^{2\rho_N} N_T^n\right) < +\infty, \quad p.s.,$$

2. le paramètre ε_n contrôle la taille de la variation de S entre deux temps d'arrêt dans $\mathcal{T}^n$

$$\sup_{n\ge 0}\left(\varepsilon_n^{-2}\sup_{1\le i\le N_T^n}\sup_{t\in(\tau_{i-1}^n,\tau_i^n]}|S_t - S_{\tau_{i-1}^n}|^2\right) < +\infty, \quad p.s..$$

Les temps d'arrêt vérifiant ces deux conditions constituent la classe de temps d'arrêt de notre étude. On remarque que cette classe est assez large car elle inclue :

1. les temps d'arrêt déterministes,
2. les temps d'atteinte d'ensembles fermés aléatoires d'intérieur non vide,
3. les schémas de discrétisation restreint de [Jacod 2012][Chapitre 14].

En dépit de la généralité de notre approche, nous obtenons des estimations fines et élégantes de l'incrément entre deux temps d'arrêt :

$$\sup_{n\geq 0}\big(\varepsilon_n^{\rho-2}(\tau_i^n-\tau_{i-1}^n)\big)<+\infty,\quad \forall\rho>0,\quad p.s..$$

Ainsi, cette estimation est valable en particulier pour les trois types de temps d'arrêt mentionnés au-dessus.
De plus, si l'on se donne une martingale locale continue M, alors à partir d'un contrôle sur S, nous en déduisons

$$\sup_{n\geq 0}\left(\varepsilon_n^{\rho-1}\sup_{1\leq i\leq N_T^n}\sup_{\tau_{i-1}^n\leq t\leq\tau_i^n}|M_t-M_{\tau_{i-1}^n}|\right)<+\infty,\quad \forall\rho>0,\quad p.s..$$

Ce qui signifie que l'ordre de grandeur de l'incrément de M est presque identique à celui de S. Cette estimation est très fine car l'intuition nous dit que l'ordre de l'incrément de M ne peut pas dépasser 1.
Enfin, nous obtenons des estimations pour des semi-martingales d'Ito continues solutions d'équations différentielles stochastiques linéaires. Si Z_t^n satisfait

$$Z_t^n=\int_0^t Z_s^n(\tilde b_s^n\mathrm{d}s+\sum_{j=1}^d(\tilde\sigma_s^n)^j\mathrm{d}B_s^j)+H_t^n,$$

où $\tilde b^n$, $(\tilde\sigma^n)^j$ sont deux processus adaptés dans $(\mathbb{R}^q)^{\otimes 2}$, bornés uniformément par une constante M et $\sup_{n\geq 0}(\varepsilon_n^{-\theta_H}\sup_{0\leq t\leq T}|H_t^n|)<+\infty$. Alors

$$\sup_{n\geq 0}\left(\varepsilon_n^{\rho-\theta_H}\sup_{0\leq t\leq T}|Z_t^n|\right)<+\infty,\quad \forall\rho>0,\quad p.s..$$

Les conséquences de ces estimations sont multiples. Il fournit que lorsque l'on contrôle les incréments d'une certaine martingale locale continue, alors on a un contrôle p.s. sur beaucoup de martingales locales continues et sur les incréments entre deux temps d'arrêt consécutifs. Ce qui a priori n'est pas évident en soi et à prouver sans nouveaux outils, c'est le but du lemme fondamental ci-dessous qui permet de prouver des résultats de convergence presque sûre sous des conditions très faibles (pas besoin d'intégrabilité, ni de connaissance des lois). Par exemple, nous pouvons contrôler des martingales locales qui ne sont pas intégrables, alors que tous les résultats jusqu'à présent pour démontrer ce type de convergence s'appuient en général sur une hypothèse d'intégrabilité (inégalité de Markov, inégalité de Lenglart, argument de type Borel-Cantelli).

Lemma 1.2.1. *Soit $\mathcal{M}_0^+$ l'ensemble des processus mesurables positifs, s'annulant en $t = 0$. Soit $(U^n)_{n\geq 0}$ et $(V^n)_{n\geq 0}$ deux suites de processus dans $\mathcal{M}_0^+$. Suppose que*

i) La série $\sum_{n\geq 0} V_t^n$ converge pour tout $t \in [0,T]$, presque sûrement;

ii) La limite de la somme est bornée supérieurement par un processus $\bar{V} \in \mathcal{M}_0^+$ et que $\bar{V}$ est continu p.s.;

iii) Il existe une constante $c \geq 0$ telle que pour tout $n \in \mathbb{N}$, $k \in \mathbb{N}$ et $t \in [0,T]$, nous avons

$$\mathbb{E}[U^n_{t\wedge\theta_k}] \leq c\mathbb{E}[V^n_{t\wedge\theta_k}]$$

avec le temps aléatoire $\theta_k := \inf\{s \in [0,T] : \bar{V}_s \geq k\}$.

Alors, pour tout $t \in [0,T]$, la série $\sum_{n\geq 0} U_t^n$ converge presque sûrement et, donc, $U_t^n \xrightarrow{a.s.} 0$.

Une conséquence immédiate de ce lemme est l'équivalence entre contrôle de variation quadratique presque sûr et contrôle uniforme de martingales locales continues presque sûr.

Theorem 1.2.3. *Soit $p > 0$ et soit $\{(M_t^n)_{0\leq t\leq T} : n \geq 0\}$ une suite de martingales locales continues, s'annulant en zéro. Alors,*

$$\sum_{n\geq 0} \langle M^n\rangle_T^{p/2} \text{ converge p.s.} \iff \sum_{n\geq 0} \sup_{0\leq t\leq T} |M_t^n|^p \text{ converge p.s..}$$

1.3 Couverture à pas discret en marché complet.

Soit S le prix des actifs donné par

$$S_t = S_0 + \int_0^t b(s, S_s)\mathrm{d}s + \int_0^t \sigma(s, S_s)\mathrm{d}B_s,$$

où B est un d-mouvement brownien.
Notons

$$Z_s^n = \int_0^s D_x u(t, S_t) \cdot \mathrm{d}S_t - \sum_{\tau_{i-1}^n \leq s} D_x u(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n}),$$

qui s'interprète comme l'erreur de couverture d'une stratégie Delta-neutre pour une option européenne de sous-jacent S, de maturité T, de fonction de prix u et de payoff $g(S_T)$. Les instants de réajustement sont donnés par une suite des temps déterministes ou aléatoires $(\tau_i^n)_{1\leq i\leq N_T^n}$ et le nombre de dates est noté N_T^n, qui peut être aléatoire. Dans un marché complet, la thèse de [Zhang 1999] offre un début de réponse à l'étude de la convergence de l'erreur de couverture. Un autre résultat de [Bertsimas 2001] généralisé par [Hayashi 2005] donne une convergence en loi de l'erreur de couverture. Cependant, ces résultats dépendent beaucoup de la régularité du payoff de l'option à répliquer. En effet, [Gobet 2001] prouve que dans le cas de la couverture d'une option digitale à des temps uniformément répartis, le taux de convergence n'est plus en $n^{1/2}$ mais en $n^{1/4}$, si n est le nombre de dates. Ce phénomène a été longuement étudié par Geiss et ses co-auteurs [Geiss 2004]. Un résultat intéressant est que pour toute suite de temps d'arrêt de longueur fixe n, la vitesse de décroissance vers 0 en norme $L^2(\Omega)$ de l'erreur d'approximation ne peut pas être supérieure à $1/\sqrt{n}$, en dehors des cas triviaux. Il se trouve (voir [Fukasawa 2011b]) que cette vitesse reste maximale même pour une classe (particulière) de temps d'arrêt avec un nombre de pas de temps stochastique. Une revue complète de la littérature sur la régularité fractionnaire apparait dans l'article [Geiss 2011]. Un des résultats importants est donné dans l'article de [Geiss 2009], dont l'interprétation financière est que lorsque le payoff est irrégulier, plus l'on se rapproche de la maturité T du contrat, plus notre stratégie sera sensible au changement du sous-jacent et donc on s'attend à vouloir réajuster notre position plus souvent près de la maturité du contrat. Bien que cette solution fait sens en pratique, une question importante est de trouver des dates de réajustement non pas déterministe, mais stochastique (en fait des temps d'arrêt), qui donnerait un taux de convergence optimal comme les temps déterministes du résultat précédent, mais qui en plus minimiserait un certain critère sur l'erreur de couverture. Les premiers auteurs, à ma connaissance, à poser le problème furent Martini et Patry dans l'article [Martini 1999]. Ils résolvent le problème de minimisation de la variance de l'erreur de couverture sur le coût initial et les temps de réajustement, pour un nombre de réajustement fixe. Cependant, la résolution du problème est très compliquée (il faut résoudre une suite de problème d'arrêt optimal imbriquée) et nécessite le recours à des méthodes numériques délicates. Suite à cet article, Fukasawa introduit une approche asymptotique en le nombre de dates, dans [Fukasawa 2011a] et [Fukasawa 2011b], une approche plus simple et qui donne une stratégie sous forme explicite, mais sous des hypothèses contraignantes.

Theorem 1.3.1 ([Fukasawa 2011b]). *Pour toute suite de grille de temps d'arrêt* *satisfaisant certaines conditions, nous avons*

$$\liminf_{n\to+\infty} \mathbb{E}[N_T^n]\mathbb{E}[\langle Z^n\rangle_T] \geq \frac{1}{6}\mathbb{E}[\int_0^T D_{xx}u(t,S_t)\mathrm{d}\langle S\rangle_t]^2.$$

Le résultat est intéressant car il donne une borne inférieure indépendante de la classe de temps d'arrêt considérée, mais sous l'hypothèse que le payoff est convexe en dimension 1, afin de ne pas avoir de problème d'intégrabilité. Le but du prochain théorème est justement de trouver une stratégie de temps d'arrêt qui atteint cette borne. A priori, il n'est pas évident que la borne ne soit pas stricte.

Theorem 1.3.2 ([Fukasawa 2011b])**.** *Soit h_n une suite qui tend vers 0. Soit la suite de temps d'arrêt τ^n définie par $\tau_0^n = 0$ et*

$$\tau_{i+1}^n := \inf\{t \geq \tau_i^n : |S_t - S_{\tau_i^n}| > h_n/\sqrt{D_{xx}u(\tau_i^n, S_{\tau_i^n})}\}.$$

Alors,

$$\lim_{n\to+\infty} \mathbb{E}[N_T^n]\mathbb{E}[\langle Z^n\rangle_T] = \frac{1}{6}\mathbb{E}[\int_0^T D_{xx}u(t,S_t)\mathrm{d}\langle S\rangle_t]^2.$$

Nous comprenons l'avantage par rapport à la précédente approche : nous avons une formule explicite pour les temps d'atteintes. Cependant, nous remarquons que la barrière des temps d'atteintes optimaux ne permet pas l'annulation du gamma de l'option et donc l'hypothèse sur la convexité de g est cruciale.
Dans le chapitre 3, nous étendons cette approche dans plusieurs directions :

1. Nos résultats sont valables pour une large classe de temps d'arrêt regroupant les temps déterministes, les schémas de discrétisation restreint de [Jacod 2012][Chapitre 14] et les temps d'atteinte d'ensembles fermés aléatoires.

2. Nous considérons un cadre multidimensionnel.

3. Nous optimisons sur un critère trajectorielle, ce qui rend le résultat plus fort.

4. La classe des payoff européens admissibles est très générale, en fait nous n'avons pas réussi à exhiber un contre-exemple de payoff européen qui ne vérifie pas notre condition. Entre autres, nos sensibilités peuvent s'annuler et nous avons levé cette contrainte inhérente à l'approche de [Fukasawa 2011b] par du calcul trajectorielle.

5. Les options asiatiques et lookback peuvent être considérées.

Tout d'abord, en mettant en oeuvre l'ensemble des résultats de convergence p.s. mentionnés avant, nous avons obtenu une borne inférieure :

Theorem 1.3.3. *Soit X un processus matriciel symétrique et positif vérifiant une certaine équation matricielle, dépendant du modèle et du payoff. Alors,*

$$\liminf_{n\to+\infty} N_T^n \langle Z^n\rangle_T \geq \Big(\int_0^T \mathrm{Tr}\,(X_t)\,\mathrm{d}t\Big)^2, \quad p.s..$$

Remark 1.3.1. *On en déduit immédiatement une borne inférieure pour le critère $L_p(\Omega)$: en effet, par une application du lemme de Fatou et de l'inégalité de Cauchy-Schwarz, nous avons (pour tout $p>0$)*

$$\begin{aligned}\Big[\mathbb{E}(\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t)^p\Big]^2 &\leq \Big[\mathbb{E}\Big(\liminf_{n\to+\infty}\big(N_T^n\langle Z^n\rangle_T\big)^{p/2}\Big)\Big]^2 \leq \liminf_{n\to+\infty}\Big[\mathbb{E}\big(N_T^n\langle Z^n\rangle_T\big)^{p/2}\Big]^2\\ &\leq \liminf_{n\to+\infty}\mathbb{E}((N_T^n)^p)\mathbb{E}(\langle Z^n\rangle_T^p).\end{aligned}$$

Pour $p=1$, nous retrouvons l'approche de Fukasawa [Fukasawa 2011a].

Ensuite, nous avons exhibé les temps d'intervention optimaux, qui sont des temps d'atteintes d'ellipsoïdes aléatoires.

Theorem 1.3.4. *Pour un processus Λ à valeurs dans l'espace des matrices symétriques définies positives données de manière explicite en fonction de X, nous définissons la stratégie $\mathcal{T}^n$ par (pour une certaine suite $(\varepsilon_n)_{n\geq 0}$)*

$$\begin{cases}\tau_0^n := 0,\\ \tau_i^n = \inf\Big\{t\geq \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^*\Lambda_{\tau_{i-1}^n}(S_t - S_{\tau_{i-1}^n}) > \varepsilon_n^2\Big\}\wedge T,\ i\geq 1.\end{cases}$$

Alors, la suite de stratégie est dans notre classe d'admissibilité de grille aléatoire et est "presque" asymptotiquement optimale dans ce sens :

$$\limsup_{n\to+\infty}\Big|N_T^n\langle Z^n\rangle_T - (\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t)^2\Big|$$

est aussi petite que l'on souhaite, où la borne d'erreur est donnée explicitement.

Techniquement, les processus matriciels X et Λ sont approximativement des transformations croissantes des valeurs propres du gamma de l'option, c'est-à-dire que dans la décomposition spectrale de la matrice symétrique de $\Gamma_t = P_t D_t P_t^*$, X_t ou Λ_t sont égaux à $P_t f(D_t) P_t^*$, où f est une certaine fonction croissante (en les valeurs propres) connue. Ainsi, nos temps d'atteintes optimaux exhibent une règle en accord avec la pratique des traders : plus le gamma est élevé, plus la fréquence de réajustement des positions de couverture augmente. Ici, la règle est explicite et non triviale. Des tests numériques montrent la réduction importante de la variance de l'erreur de couverture en utilisant nos temps optimaux par rapport à des temps déterministes.

Dans le chapitre 5, nous avons généralisé le critère de coût (c'est-à-dire le terme N_T^n dans le produit $N_T^n \langle Z^n \rangle_T$) et l'erreur (c'est-à-dire $\langle Z^n \rangle_T$) par des sommes plus générales, impliquant deux semi-martingales d'Ito; cependant, le résultat est valable en dimension 1.

1.4 Couverture à temps discret et avec coût de transaction.

Notons

$$Z_T^n = g(S_T) - \Bigg(v(0, S_0) + \sum_{\tau_{i-1}^n < T} D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n})(S_{\tau_i^n} - S_{\tau_{i-1}^n}) \\ - k_n \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left| D_x v(\tau_i^n, S_{\tau_i^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \right| \Bigg).$$

Ici, k_n représente un coût de transaction proportionnel aux montants. La littérature académique sur les erreurs de couverture avec coûts de transaction est immense et nous allons nous restreindre au cas de l'étude de la convergence de l'erreur de couverture en probabilité, en loi et dans $L^2(\Omega)$. Un premier résultat fondateur est dû à Leland dans le cas d'un Call :

Theorem 1.4.1 ([Leland 1985]). *Soit $\mathcal{T}^n$ la suite de temps uniforme ($\tau_i^n = \frac{iT}{n}$) et $k_n = k_0 n^{-\alpha}$, où $k_0 > 0$ et $\alpha \in]0, \frac{1}{2}]$. Alors,*

$$\lim_{n \to +\infty} Z_T^n = 0, \quad \textit{en probabilité},$$

en prenant une volatilité modifiée dans les formules de Black-Scholes pour calculer v.

On peut voir que si $\alpha \neq 1/2$, les Deltas de la stratégie dépendent de n, ce qui n'est pas réaliste en pratique. Le cas $\alpha = 0$ est important car le coût de transaction proportionnel aux montants est constant ou en tout cas ne décroit pas vers 0 lorsque l'investisseur effectue une stratégie avec de plus en plus de réajustement. Ce cas a été résolu par [Kabanov 1997] dans le cas d'un Call et un biais apparait. Une question intéressante concerne le taux de convergence de l'erreur de couverture biaisé vers 0. La réponse a été donnée par [Pergamenshchikov 2003]. Dans [Denis 2010a], l'auteur exhibe une stratégie, en modifiant le Delta de couverture, telle que l'erreur de couverture converge vers 0 en probabilité pour le cas $\alpha = 0$. Cependant, les articles de [Levental 1997] et de [Soner 1995] mettent en lumière dans un cadre très général l'intuition suivante : quand le paramètre k_n est constant, le coût minimal d'une stratégie de couverture est le prix de sur-réplication et ainsi la stratégie optimale est la stratégie Buy and Hold. L'article de [Denis 2010a] exhibe simplement une stratégie qui tend plus rapidement vers la stratégie Buy and Hold, ce qui augmente grandement la variance de la limite de $\sqrt{n}Z_T^n$ et rend la stratégie pire (voir figure 4.1) que la stratégie initiale de Leland en pratique (c'est-à-dire pour n fixe). Un problème général est soulevé sur le choix de "l'asymptotique" dans les problèmes de couverture avec coûts de transaction. En pratique, on raisonne toujours pour un nombre de réajustement fixe. Donc, une stratégie asymptotiquement pertinente doit être robuste à des tests non-asymptotiques. Quelques simulations numériques faites au Chapitre 4 vont dans ce sens et incitent à approcher le problème par la méthode de Leland-Lott. Bien qu'en pratique le coefficient k_n ne tend pas vers 0, le cas $k_n = k_0/\sqrt{n}$ est intéressant car donne des Deltas qui ne dépendent pas de n. Les premiers résultats, à ma connaissance, à avoir répondu à cette question sont dûs à Lott. Le théorème ci-dessous donne le taux de convergence dans $L^2(\Omega)$ de la stratégie de Leland-Lott.

Theorem 1.4.2 ([Denis 2010b]). *Dans le Call de maturité* 1 *dans le modèle de Black-Scholes, on a*

$$\mathbb{E}(Z_1^n) = A_1 n^{-1} + o(n^{-1}), \quad n \to +\infty,$$

où le coefficient

$$A_1 := \int_0^1 \left[\frac{\sigma^4}{4} + \sigma^3 k_0 \sqrt{\frac{2}{2\pi}} + k_0^2 \sigma^2 (1 - \frac{2}{\pi}) \right] \Lambda_t \mathrm{d}t,$$

avec $\Lambda_t = \mathbb{E}[S_t^4 D_{xx}^2 u(t, S_t)^2]$.

Le chapitre 2 étudie la convergence (lorsque la fréquence d'intervention devient infinie) de l'erreur de couverture d'un Call pour une stratégie de Leland dans le modèle de Bachelier, lorsque le coût de transaction proportionnel est fixe ($k_n = k_0$). Il est prouvé que cette erreur est bornée par une variable aléatoire proportionnelle au taux de transaction.

Il existe aussi une large littérature sur l'approche par indifférence de prix par exemple, [Davis 1993]. En général, les auteurs cherchent les frontières d'exécution optimales à l'achat et à la vente , i.e. quand le Delta a suffisamment bougé. L'inconvénient majeur de cette approche est le recours systématique à des méthodes numériques pour calculer ces frontières d'exercice.

Dans le chapitre 4, nous nous plaçons en dimension 1 et considérons le modèle de Leland-Lott, c'est-à-dire quand le coût de transaction décroit vers 0 à un certain taux ε_n qui rend les sensibilités indépendantes de n. Nous avons trouvé dans ce cadre asymptotique des régions d'exercice optimales sous une forme explicite, ce sont des temps d'atteinte du prix d'une barrière aléatoire (dépendent du gamma de l'option). Tout d'abord, nous avons montré une décomposition qui permet de se ramener à l'étude d'une partie martingale locale de l'erreur de couverture, afin de comprendre ce que l'on minimise.

Proposition 1.4.1. *L'erreur de couverture Z_T^n peut être décomposée sous la forme*

$$Z_T^n = M_T^n + R_T^n,$$

avec $(N_T^n)^{1/2} R_T^n \xrightarrow{a.s.} 0$ *et* M^n *est une martingale locale spécifique.*

Le résultat nous dit simplement que dans la semi-martingale Z^n, le terme dominant de l'erreur est un terme martingale. Grossièrement, Z_T^n est du même ordre de grandeur que $1/(N_T^n)^{1/2}$. Ainsi, si l'on veut minimiser asymptotiquement la variance de Z_T^n, if suffit de minimiser asymptotiquement la variance de M^n et c'est l'idée du théorème suivant.

Theorem 1.4.3. *Pour des stratégies admissibles, nous avons*

$$\liminf_{n\to+\infty} N_T^n \langle M^n \rangle_T \geq \left(\int_0^T \frac{1}{\sqrt{6}} \left(1 + \frac{2kS_t}{\lambda(t,S_t)} \right) |D_{xx}^2 v(t,S_t)| \mathrm{d}\langle S \rangle_t \right)^2. \qquad (\star)$$

Soit $\lambda(t,x) := -kx + \sqrt{(kx)^2 + \frac{\sqrt{6}}{|D_{xx}v(t,x)|}}$, *avec* v *supposant vérifier l'équation*

$$D_t v(t,x) + \frac{\sigma(t,x)^2 x^2}{2} \left(1 + \frac{2kx\sqrt{|D_{xx}v(t,x)|}}{-kx\sqrt{|D_{xx}v(t,x)|} + \sqrt{(kx)^2 |D_{xx}v(t,x)| + \sqrt{6}}} \right) D_{xx}^2 v(t,x) = 0,$$

et

$$v(T,x) = g(x),$$

avec g *convexe. Alors, la suite de stratégie est optimale dans le sens :*

$$Z_T^n \xrightarrow{a.s.} 0$$

et atteint la borne inférieure $(\star)$ *sur tous les* λ.

L'avantage de notre approche asymptotique est d'avoir de manière explicite les bornes inférieures et supérieures des régions d'exercice. La restriction majeure est la dimension 1 du problème. Les résultats numériques non-asymptotiques sont excellents en terme de balance entre la moyenne et la variance ou la value-at-risk de l'erreur de couverture (voir Section 4.4 du chapitre 4).

1.5 Schéma d'Euler.

Depuis le travail pionnier de [Maruyama 1955] et, jusqu'à maintenant, l'étude de l'approximation d'équation différentielle stochastique a été un champs de recherche très actif à l'intersection entre l'analyse numérique et les processus stochastiques avec de nombreuses applications dans différents domaines (voir par exemple, [Kloeden 2010], [Milstein 1994], [Platen 1999] and [Talay 1995]). De plus, un large ensemble d'article de Müller-Gronbach et al. [Hofmann 2000], [Muller-Gronbach 2002], [Muller-Gronbach 2004], [Muller-Gronbach 2008] considère le problème dans une théorie générale.

Cependant, peu d'études ont traité le problème pour des processus multidimensionnels avec des grilles de discrétisations aléatoires. Dans le chapitre 6, nous permettons une vaste famille de grilles de discrétisation en spécifiant directement un contrôle uniforme sur un processus connu (par exemple, le mouvement brownien) et sur le nombre de temps d'arrêt de notre grille. Le but est de trouver une suite de temps d'arrêt minimisant presque sûrement un certain critère quadratique de l'erreur du schéma d'Euler. L'optimalité provient d'un choix judicieux de la norme de minimisation

$$0 < \liminf_{n\to+\infty} (N_T^n)^{1/2} \cdot ||X_T - X_T^n|| < +\infty, \quad p.s.,$$

où X^n est le schéma d'Euler sur une grille π. Le choix de la norme $||.||$ joue un rôle crucial et est guidé par sa simplicité qu'il induit dans les calculs et par sa pertinence par rapport aux problèmes spécifiques considérés. Par ailleurs, nous avons obtenu des estimations fines presque sûre d'approximation de processus, qui sont nouveaux et qui ne sont pas aisées à obtenir sans les outils développés au chapitre 3. Nous avons obtenu une borne inférieure pour la variation quadratique renormalisée de l'erreur du schéma d'Euler et nous avons exhibé une suite minimisante.

Theorem 1.5.1. *Sous certaines hypothèses, en notant $e^X_{n,t} = X^n_t - X_t$ et ∇X^{-1} l'inverse du processus tangent, nous avons*

$$\liminf_{n\to+\infty} \left\{ N_T^n \mathrm{Tr}\langle \nabla X^{-1} e^X_{n,.}\rangle_T \right\} \geq L_T,$$

où L_T est une borne inférieure indépendante de la grille de discrétisation.

Notons $\xi_{x_0,\Sigma} := \{x \in \mathbb{R}^d : (x - x_0)^*\Sigma(x - x_0) < 1\}$, pour l'ellipsoïde centré en x_0 et dont les axes sont décrits par Σ.

Theorem 1.5.2. *Sous certaines hypothèses, on peut définir un processus $(\Gamma^{n,\mu})$ basé* *sur le schéma d'Euler (et donc explicite et calculable), de telle sorte à pouvoir poser*

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n = \inf\left\{ t \geq \tau_{i-1}^n : Y_t \notin \xi_{Y_{\tau_{i-1}^n}, \varepsilon_n^{-2}\Gamma^{n,\mu}_{\tau_{i-1}^n}} \right\} \wedge T, \quad i \geq 1. \end{cases} \tag{1.5.1}$$

Alors, nous avons :

$$\limsup_{n \to +\infty} \left| N_T^n \mathrm{Tr}\langle \nabla X^{-1} e_{n,.}^X \rangle_T - L_T \right| \leq C_\mu,$$

où C_μ *est connu et tend vers* 0 *quand le paramètre de régularisation* μ *tend vers* 0.

1.6 Convergence en loi de processus discrétisés.

La littérature sur la convergence en loi de processus discrétisés est très abondante. La convergence en loi des schémas d'Euler apparait dans [Kurtz 1991], [Jacod 1998], où la discrétisation est une grille déterministe. La convergence de l'erreur de couverture est dans [Hayashi 2005], [Gobet 2001], [Geiss 2009] pour des temps d'intervention déterministe. Pour des temps d'arrêt satisfaisant certaines conditions de symétrie, [Fukasawa 2011b] prouve aussi des convergence en loi en dimension 1. En résumé, soit les auteurs prouvent des résultats de convergence dans un cadre multidimensionnel, mais avec des grilles déterministes, soit des convergences en loi avec des grilles de discrétisations aléatoires symétriques, mais en dimension 1. Techniquement, ce qui fait bien fonctionner leurs preuves est que l'on a dans tous ces cas $\mathbb{E}_{\tau_{i-1}^n}[(B_{\tau_i^n} - B_{\tau_{i-1}^n})^{2p+1}] = 0$ (où $p \in \mathbb{N}$ et B est le mouvement brownien "ambiant"), soit en utilisant la symétrie de la distribution gaussienne dans le cas des temps déterministes, soit par symétrie des barrières dans la définition des temps d'arrêt en dimension 1. Ainsi, le mouvement brownien limite est décorrélé du mouvement brownien ambiant. Nous traitons dans le chapitre 7 le cas général de convergence en loi de processus discrétisés dans un cadre multidimensionnel avec des temps d'atteinte d'ellipse, ce choix étant motivé par l'optimalité que cette famille satisfait en général dans les chapitres précédents. Dans notre cas, la variable limite est corrélée au mouvement brownien ambiant (principalement parce que $\mathbb{E}_{\tau_{i-1}^n}[(B_{\tau_i^n} - B_{\tau_{i-1}^n})^{2p+1}] \neq 0$) et la corrélation limite se calcule à l'aide d'une famille de problèmes de Dirichlet changée d'échelle. A ma connaissance, c'est la première fois que l'on est capable d'identifier ce type de corrélation.

1.7 Formule d'approximation pour les options sur Spread.

Les options sur spread sont communément traitées sur les marchés des énergies, car les spreads modélisent des actifs physiques et les options permettent de couvrir les risques inhérents à la gestion de ces actifs physiques (centrales à gaz, centrales nucléaires,...). Le recours à des méthodes de Monte Carlo pour évaluer ses contrats est assez systématique. L'intérêt pour des formules analytiques est bien réel.
La littérature sur les approximations des prix d'une option dans un modèle général étant abondante, nous allons nous restreindre à l'étude des approximations par paramétrisation (en ϵ) du processus initial (pour une revue voir [Bompis 2012]). A ma connaissance, trois approches dominent dans la littérature. Tout d'abord, [Yoshida 1992], [Uchida 2004], [Kunitomo 2001] supposent un cadre petit bruit, c'est-à-dire la volatilité est proportionnelle à ϵ qui tend vers 0. En utilisant la théorie principalement développée par [Watanabe 1987] basée sur le calcul de Malliavin, ils démontrent que les prix d'options sont développables en puissance de ϵ et rendent explicites les coefficients du développement. Souvent, à l'ordre 0, le modèle implicitement utilisé est gaussien (modèle de Bachelier). Cette approche asymptotique ne permet pas de voir le rôle des autres paramètres du modèle, ce qui peut conduire à des conclusions incorrectes sur la précision pratique des formules (voir [Bompis 2012]).
[Fouque 2003] et ses co-auteurs remarquent que la volatilité d'un actif a de plus faible variation que les variations de l'actif lui-même, ce qui mène à des modélisations multi-échelles. Par des techniques d'homogénéisation basées surtout sur les EDPs, ils parvient à faire des développements du prix de la forme $prix = prix_{BS} + Grecques_{BS}$, avec des formules simples pour des volatilités stochastiques revenant rapidement à la moyenne (comportement ergodique).
Une autre approche due à [Benhamou 2009], [Benhamou 2010a], [Benhamou 2010b], [Benhamou 2012] et [Gobet 2011b] considère des approximations non-asymptotiques autour d'un modèle "proxy" où l'on sait faire les calculs de manière explicite (normal, log-normal). Là aussi, le calcul de Malliavin est utilisé pour estimer l'erreur commise pour des payoff non-réguliers et pour calculer les termes d'erreurs correctifs; les estimations d'erreur non-asymptotique permettent de bien comprendre le rôle du temps, de la volatilité locale, de la fluctuation de la volatilité dans la précision des formules.
Maintenant, nous allons discuter plus particulièrement des approximations d'options sur spread. Une formule explicite existe pour une option d'échange donnée par la formule de Margrabe. Outre ce cas simple, une des formules d'approximation les

plus célèbres fût élaborée par [Kirk 1995]. D'autres formules plus précises sont données dans [Carmona 2003] et [Bjerksund 2006], cette dernière étant d'après des tests numériques la formule la plus précise. Ces formules approchent une option spread à deux actifs et un strike $K > 0$. Récemment, [Alos 2011] généralise la formule de [Kirk 1995] à trois actifs. Les inconvénients majeurs de ces formules sont l'absence d'estimation d'erreur, le modèle log-normal utilisé pour modéliser les actifs et la restriction sur le nombre d'actifs.

Le chapitre 8 développe des formules approchées non-asymptotiques du prix d'une option sur spread dans un modèle à volatilité locale multidimensionnel dans l'esprit de [Gobet 2011b], pour les options sur moyenne. Mais contrairement à l'article [Gobet 2011b], nous utilisons la structure linéaire du payoff d'une option spread pour effectuer des changements de numéraire et faire un développement non-asymptotique sous différentes probabilité de la région d'exercice; bien approcher la probabilité de la région d'exercice est au coeur de l'esprit de la formule de [Bjerksund 2006]. Nous approchons une somme convexe de martingales exponentielles par une martingale exponentielle, en préservant ainsi le moment d'ordre 1. Naturellement, notre approximation à l'ordre 0 est une généralisation de la formule de Bjerksund et Stensland en dimension quelconque, que l'on sait très précise pour deux actifs. Même dans le modèle log-normal, l'ordre 1 de notre approximation n'est pas nul et mène à une correction qui améliore la formule de Bjerksund et Stensland. Des tests numériques montrent que notre formule est bien plus précise que les formules de Kirk et de Bjerksund et Stensland.

Chapter 2

A toy model : transaction costs proportional to volume

Contents

2.1 Introduction

One of the most breathtaking notion in financial mathematics has been the concept of replication. In a complete market, the price is unique and the Delta hedging strategy is well-known to lead to the target price, when the number of rebalancing dates becomes infinite. However, if we relax the assumption made on the transaction costs, the theory is no longer simple : the price is not at all unique and depends on the risk criterion involved. In their pioneering paper of 1973, Black and Scholes noticed a discrepancy between the observed option price and the "theoretical" one given by the formula. Among other suggestions, they indicated that this may be due to the transaction costs. Indeed, though the percentage of the transaction volume paid as the brokerage fee individually can be considered as negligible, the total sum after hundreds and thousands portfolio revisions is far from being such: in the continuous trading, the Black-Scholes prescription leads to the explosion of the accumulated transaction cost payments. The situation is even more dramatic in the case of energy or agricultural market, where costs are proportional to the volume transported from one place to another and are no longer negligible.

The development of a mathematical framework to deal with transaction costs has proceeded along several different lines. The results in this chapter are related to the pioneering work of Leland (1985), in which a replicating portfolio is rebalanced at equal time intervals. By increasing the frequency of rebalancing while letting transaction costs vanish at an appropriate rate, a cost of replication is obtained as a solution to a Black-Scholes partial differential equation with an enhanced volatility. The input volatility is a multiple of the "true" volatility by a certain magnifying factor depending on the transaction constant coefficient, the volatility itself and the number n of the revision intervals. Leland's conclusion was very important for practitioners because it provides a reference point for pricing contingent claims in real-world markets. Its great advantage is an easy implementation.

Unfortunately, Leland could not provide a mathematically correct confirmation for his prescription. In his basic setting, he considered proportional transaction cost coefficient constant in n. His main theorem claimed that the terminal values of portfolios converge to the pay-off. This assertion is false: the convergence holds but not to the terminal pay-off indicated in the contract. There is a non-trivial discrepancy, which is proved by Kabanov and Safarian [Kabanov 1997], and then by [Pergamenshchikov 2003], ... Leland also made a remark, without providing arguments, that the convergence holds also in the model where the transaction cost coefficient is a function of the number of revisions decreasing as $n^{\frac{1}{2}}$. This conjecture is correct. It was proven in the thesis of Klaus Lott. This is the first rigorous explanation for why the Leland strategy does work in practical situations of small transaction costs and not very high frequencies of portfolio revisions. Several extensions have been following :

- The trading dates are no longer uniform but deterministic and design to get faster convergence result (see [Denis 2010b]).
- The trading dates become stochastic (see [Ahn 1998]).

In this chapter, we extend Leland-Lott strategy to transaction costs proportional to volume. That is we study the asymptotic behaviour (in probability) of the wealth process V_T^n rebalanced at discrete times $(t_i := \frac{i}{n})_{0 \le i \le n}$, when the transaction cost parameter κ is kept fixed. Actually, we show that the difference between the portfolio value and the desired target is bounded $a.s.$ by $\alpha\kappa$, where α is a finite and positive random variable, when the number of trading dates increases to infinity. Here, we consider the Bachelier model and small transaction costs proportional to the volume as we could encounter in commodity market. For instance, in the case of gas market, the underlying is stored in a storage and proportional transaction costs to volume injected or withdrawn appear. Under this toy setting, an unpleasant phenomenon unveils : the problems of explosion of moments of some Greeks when we approach the maturity of the contract. One way to tackle this issue is to consider another type of convergence, different from strong convergence in $L^2(\Omega)$ [Denis 2010b] or convergence in law [Pergamenshchikov 2003], as commonly assumed in the literature, for instance, the almost sure convergence and this is the topic of the remainder of this part (see Chapter 3); indeed, the hedging problem can be seen more naturally as a pathwise problem.

2.2 The model

Let $T > 0$ be the maturity of a contract. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a probability space. The price process $S = \{S_t, 0 \leq t \leq T\}$ is an $\mathbb{R}$-valued stochastic process defined by (Bachelier's model) :

$$dS_t = rS_t\mathrm{d}t + \sigma \mathrm{d}W_t, \quad 0 \leq t \leq T, \tag{2.2.1}$$

where W is a Brownian motion under the risk-neutral measure $\mathbb{P}$.

Remark 2.2.1. *Assume $\mathbb{P} \sim \mathbb{Q}$ on $\mathcal{F}_T$ and that ξ_N is $\mathcal{F}_T$-measurable.*

- *If $\xi_N \to_{\mathbb{Q}} 0$, then $\xi_N \to_{\mathbb{P}} 0$.*
- *If $\xi_N \to_{L^1(\mathbb{Q})} 0$ and $\frac{d\mathbb{P}}{d\mathbb{Q}}|_{\mathcal{F}_T} \in L^q(\mathbb{Q})$ for some $q > 1$, then $\xi_N \to_{L^{1/p}(\mathbb{P})} 0$ where $\frac{1}{p} + \frac{1}{q} = 1$.*

2.3 Self-financing strategy

We assume an agent sells a derivative security with pay-off $g(S_T)$, g being an $\mathbb{R}$-valued function, and immediately takes a position of δ_t shares of the asset S_t and δ_t^0 of bonds $B_t := e^{rt}$ (with constant risk-free interest rate $r \geq 0$). The value at time t of the self financing portfolio is given by

$$V_t = \delta_t^0 B_t + \delta_t S_t,$$

with the objective of having $V_T = g(S_T)$.
In order to well understand the self-financing condition, we write the portfolio change without and with transaction costs :

- The change in the portfolio from t_i to t_{i+1} without transaction costs is

$$V_{t_{i+1}} - V_{t_i} = \delta_{t_i}^0 (B_{t_{i+1}} - B_{t_i}) + \delta_{t_i}(S_{t_{i+1}} - S_{t_i}). \tag{2.3.1}$$

- The change in the portfolio from t_i to t_{i+1} with transaction costs is

$$V_{t_{i+1}} - V_{t_i} = \delta^0_{t_i}(B_{t_{i+1}} - B_{t_i}) + \delta_{t_i}(S_{t_{i+1}} - S_{t_i}) - \kappa|\delta_{t_{i+1}} - \delta_{t_i}|, \qquad (2.3.2)$$

where κ is the transaction cost (in $\$/m^3$, for example, if S is the spot gas).

The strategy is made up of a position in the asset and in the bond. Only, the variations of the price of asset and of the bond contribute to the variations of the portfolio, this is the first and the second terms on the right-hand side of the equation (2.3.2). But, the asset must be stored (oil, gas, agricultural products) or transported, the costs become proportional to the volume stored or transported (third term).
The following lemma deals with the case of discounted portfolio $V_t = V(t, S_t)$:

Lemma 2.3.1. *Let V_0 be the initial portfolio value, $\delta_{\varphi(t)} = \sum_{i=1}^n \delta_{t_{i-1}} 1_{]t_{i-1},t_i]}(t)$ be the piecewise constant strategy ($\varphi(t) := \sup\{t_i : t_i < t\}$).*
Then, the discounted portfolio value is

$$\tilde{V}_T = V_0 + \int_0^T \delta_{\varphi(t)} d\tilde{S}_t - k\sum_{i=1}^n e^{-rt_i}|\delta_{t_i} - \delta_{t_{i-1}}|.$$

Proof. By the equation (2.3.2)

$$\begin{aligned} V_{t_{i+1}} &= V_{t_i} + \frac{B_{t_{i+1}} - B_{t_i}}{B_{t_i}}(V_{t_i} - \delta_{t_i}S_{t_i}) + \delta_{t_i}(S_{t_{i+1}} - S_{t_i}) - k|\delta_{t_{i+1}} - \delta_{t_i}| \\ &= \frac{B_{t_{i+1}}}{B_{t_i}}V_{t_i} + \delta_{t_i}(S_{t_{i+1}} - \frac{B_{t_{i+1}}}{B_{t_i}}S_{t_i}) - k|\delta_{t_{i+1}} - \delta_{t_i}|. \end{aligned}$$

Then,

$$\tilde{V}_{t_{i+1}} = \tilde{V}_{t_i} + \delta_{t_i}(\tilde{S}_{t_{i+1}} - \tilde{S}_{t_i}) - \frac{k}{B_{t_{i+1}}}|\delta_{t_{i+1}} - \delta_{t_i}|.$$

□

Remark 2.3.1.

1. *The trading cost only occurs at time t_i when we change the weight of the portfolio.*

2. *Taking the expectation in both sides of the equation, we have*

$$\mathbb{E}[\tilde{V}^n_{t_j}] = V_0 - k\sum_{i=1}^j e^{-rt_i}\mathbb{E}|\delta_{t_i} - \delta_{t_{i-1}}|.$$

Thus, the discounted portfolio value decreases with t_j in expectation. An interesting model is given by Lott in his thesis where k is related to the number of rebalancing n by the relation $k = \frac{1}{\sqrt{n}}$. In this case, the premium does not tend to the super-replication price. Even if the hypothesis on k is dubious, that gives nevertheless a robust framework to deal with small transaction costs.

3. *In the previous lemma, we do not consider the costs at the beginning and at the end of the strategy. The cost at the inception of our strategy is* $-\kappa|\delta_0|$*. The portfolio value becomes*

$$\tilde{V}_T = V_0 - \kappa|\delta_0| + \int_0^T \delta_{\varphi(t)} d\tilde{S}_t - k\sum_{i=1}^n e^{-rt_i}|\delta_{t_i} - \delta_{t_{i-1}}|.$$

In the case of the settlement of the underlying, we must sell off our final position (for example, by emptying a storage), thus subtract $\kappa e^{-rT}|\delta_T|$ *:*

$$\tilde{V}_T = V_0 + \int_0^T \delta_{\varphi(t)} d\tilde{S}_t - k\sum_{i=1}^n e^{-rt_i}|\delta_{t_i} - \delta_{t_{i-1}}| - \kappa e^{-rT}|\delta_T|.$$

In the following, we do not write these constraints to simplify the presentation. To generalize the following results, we would just have to add the previous terms to the portfolio value V*.*

2.4 The Bachelier model

As we want to keep the setting as simple as possible, we consider the Bachelier model and a Call option in order to get closed formulas for the sensitivities of the price function. Here, the aim is not to be general in our model or in the pay-off function g, but to shed light on some issues inherently linked to the calculus and to put the approach of Chapter 3 into perspective.

2.4.1 Model and formulas

Consider a Call option with underlying S (gas, oil, coal,...), strike K and maturity T. We know that in the Bachelier model (see [Musiela 2000, p.122])

$$\begin{aligned}\mathcal{C}(t, S_t; T, K, \sigma) := C(t, S_t) &= \mathbb{E}_t\left[e^{-r(T-t)}(S_T - K)_+\right]\\ &= e^{rt}\mathbb{E}_t\left[(\tilde{S}_T - \tilde{K})_+\right],\end{aligned}$$

where $\tilde{S}_T = \tilde{S}_t + \sigma\int_t^T e^{-ru}\mathrm{d}W_u$ and $\tilde{K} = e^{-rT}K$. As usual, we note $\mathbb{E}_t[.] := \mathbb{E}_{\mathbb{P}}[.|\mathcal{F}_t]$. Let

$$\Sigma(T-t) = \sigma\sqrt{\frac{1 - e^{-2r(T-t)}}{2r}}$$

and

$$d(t, S) = \frac{S - Ke^{-r(T-t)}}{\Sigma(T-t)}.$$

On $\mathcal{F}_t$, $\tilde{S}_T \sim \mathcal{N}(\tilde{S}_t, (e^{-rt}\Sigma(T-t))^2) = \mathcal{N}\left(\tilde{S}_t, \frac{\sigma^2}{2r}(e^{-2rt} - e^{-2rT})\right)$, where $N(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$ and $N'(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ are respectively the cumulative distribution

function and the density of a standard Gaussian variable.
Then,

$$
\begin{aligned}
C(t,S_t) &= e^{rt}\mathbb{E}_t[(\tilde{S}_t + e^{-rt}\Sigma(T-t)W_1 - \tilde{K})_+] \\
&= e^{rt}\int_{\frac{Ke^{-r(T-t)}-S_t}{\Sigma(T-t)}}^{+\infty} (\tilde{S}_t + e^{-rt}\Sigma(T-t)x - \tilde{K})\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx \\
&= e^{rt}\left[(\tilde{S}_t - \tilde{K})N\left(\frac{S_t - Ke^{-r(T-t)}}{\Sigma(T-t)}\right) + e^{-rt}\Sigma(T-t)N'\left(\frac{S_t - Ke^{-r(T-t)}}{\Sigma(T-t)}\right)\right] \\
&= \Sigma(T-t)N'\left(d(t,S_t)\right) + (S_t - Ke^{-r(T-t)})N\left(d(t,S_t)\right).
\end{aligned}
$$

The following lemma gives useful inequalities on the sensitivities of C, denoting by $C_S(t,S) = \partial_S C(t,S)$, $C_{SS}(t,S) = \partial^2_{SS}C(t,S)$, $C_{tS}(t,S) = \partial^2_{tS}C(t,S)$, ...

Lemma 2.4.1. *The Greeks are*

$$
\begin{aligned}
C_S(t,S) &= N(d(t,S)), \\
C_{SS}(t,S) &= \frac{N'(d(t,S))}{\Sigma(T-t)}, \\
C_{SSS}(t,S) &= -\frac{S - Ke^{-r(T-t)}}{\Sigma(T-t)^3}N'(d(t,S)), \\
C_{St}(t,S) &= e^{-2r(T-t)}\sigma^2\frac{S - Ke^{r(T-t)}}{2\Sigma(T-t)^3}N'(d(t,S)), \\
C_{SSt}(t,S) &= \frac{\sigma^2}{2\Sigma(T-t)^3}e^{-2r(T-t)}\left(1 - \frac{(S - Ke^{r(T-t)})(S - Ke^{-r(T-t)})}{\Sigma(T-t)^2}\right)N'(d(t,S)), \\
C_{SSSS}(t,S) &= \frac{1}{\Sigma(T-t)^3}\left(-1 + \frac{(S - Ke^{-r(T-t)})^2}{\Sigma(T-t)^2}\right)N'(d(t,S)).
\end{aligned}
$$

Proof. See Appendix. □

Lemma 2.4.2. *The function $(t,S) \mapsto C(t,S)$ is solution to the equations*

$$
C_{tS} + rSC_{SS} + \frac{\sigma^2}{2}C_{SSS} = 0,
$$

$$
C_{tSS} + rC_{SS} + rSC_{SSS} + \frac{\sigma^2}{2}C_{SSSS} = 0.
$$

Proof. We know that

$$
C_t + rSC_S + \frac{\sigma^2}{2}C_{SS} - rC = 0.
$$

We just have to differentiate once and twice w.r.t. S. □

Modified price. In the following, we consider the "modified" Bachelier price C^n, where we replace the parameter $\Sigma(T-t)$ by the modified one

$$\Sigma_n(T-t) := \Sigma(T-t)\sqrt{1+\alpha_n},\ d^n(t,S) = \frac{d(t,S)}{\sqrt{1+\alpha_n}},\ \sigma_n = \sigma\sqrt{1+\alpha_n},$$

with $\alpha_n = \frac{\kappa}{\sigma}\sqrt{\frac{8}{\pi h_n}}$, We shall use the notation "$\leq_c$" for "$\leq$" up to a multiplicative constant, which does not depend on h_n and remains bounded when κ is bounded. We give some handy estimates on the moments of the Greeks

Lemma 2.4.3. *For all $\gamma > 0$, for any $0 \leq s < t \leq T$, we have*

$$\mathbb{E}_s|C^n_{SS}(t,S_t)|^\gamma \leq_c \frac{\sigma_n^{1-\gamma}(T-t)^{\frac{1-\gamma}{2}}}{\sqrt{\sigma_n^2(T-t)+\sigma^2(t-s)}},$$

$$\mathbb{E}_s|C^n_{SSS}(t,S_t)|^\gamma \leq_c \frac{\sigma_n^{1-2\gamma}(T-t)^{\frac{1-2\gamma}{2}}}{\sqrt{\sigma_n^2(T-t)+\sigma^2(t-s)}},$$

$$\mathbb{E}_s|C^n_{SSSS}(t,S_t)|^\gamma \leq_c \frac{\sigma_n^{1-3\gamma}(T-t)^{\frac{1-3\gamma}{2}}}{\sqrt{\sigma_n^2(T-t)+\sigma^2(t-s)}}.$$

In particular, for $s = 0$,

$$\mathbb{E}|C^n_{SS}(t,S_t)|^\gamma \leq_c \frac{\sigma_n^{1-\gamma}(T-t)^{\frac{1-\gamma}{2}}}{\sqrt{\sigma_n^2(T-t)+\gamma\sigma^2 t}},$$

$$\mathbb{E}|C^n_{SSS}(t,S_t)|^\gamma \leq_c \frac{\sigma_n^{1-2\gamma}(T-t)^{\frac{1-2\gamma}{2}}}{\sqrt{\sigma_n^2(T-t)+\sigma^2 t}},$$

$$\mathbb{E}|C^n_{SSSS}(t,S_t)|^\gamma \leq_c \frac{\sigma_n^{1-3\gamma}(T-t)^{\frac{1-3\gamma}{2}}}{\sqrt{\sigma_n^2(T-t)+\sigma^2 t}}.$$

Proof. See Appendix. □

Remark 2.4.1. *As we shall see in Chapter 3, the condition on the non-explosion of the integrated moments necessitates strong conditions on the final pay-off, even in the simple case of a Call option; indeed, we remark that the second-order moment of the speed C_{SSS} of a Call option is not integrable in time (see the proof in the Appendix), so we need to be careful in computations when we deal with such quantities. Chapter 3 will overcome these well-known problems.*

2.4.2 The main convergence result under the assumption "fixed small transaction costs"

Theorem 2.4.1. *The hedging strategy $(\delta^0,\delta) = (C^n - C^n_S S, C^n_S)$ yields a wealth process V^n_T satisfying : there exists a finite random variable U_0 such that*

$$\mathbb{P}\left(|V^n_T - (S_T-K)_+| \leq U_0.\kappa\right) \to_{n\to+\infty} 1.$$

Remark 2.4.2.
- *$\sigma_n \to +\infty \implies C^n(t,S_t) \to S_t$ and $C^n_S(t,S_t) \to 1$ (buy and hold strategy giving super-replication strategy).*
- *The analysis would be similar if $\kappa = \kappa_n \to 0$ with $\frac{\kappa_n}{\sqrt{h_n}} \to +\infty$.*
- *However, the case κ fixed is more meaningful for the applications of Gdf Suez.*

Proof. We adapt the proof of [Kabanov 1997].
Let $H = (S_T - K)_+$, $t_i = \frac{iT}{n}$ $(i \in \{0,...,n\})$, $\varphi(t) := \sup\{t_i : t_i \leq t\})$ and $C^n_S(\varphi(t), S_{\varphi(t)}) := \sum_{i=1}^n C^n_S(t_{i-1}, S_{t_{i-1}}) 1_{[t_{i-1},t_i)}(t)$.
Taking $\delta_{\varphi(t)} = C^n_S(\varphi(t), S_{\varphi(t)})$ in Lemma 2.3.1, we deduce that

$$\tilde{V}_T = C^n(0,S_0) + \int_0^T C^n_S(\varphi(t), S_{\varphi(t)}) \mathrm{d}\tilde{S}_t - k \sum_{i=1}^n e^{-rt_i} |C^n_S(t_i, S_{t_i}) - C^n_S(t_{i-1}, S_{t_{i-1}})|.$$

Lemma 2.4.4. *We have $\tilde{V}_T - \tilde{H} = F^n_1 + F^n_2$, where*

$$F^n_1 := \int_0^T (C^n_S(\varphi(t), S_{\varphi(t)}) - C^n_S(t,S_t)) \mathrm{d}\tilde{S}_t,$$

$$F^n_2 := \int_0^T e^{-rt} \frac{\sigma^2 \alpha_n}{2} C^n_{SS}(t,S_t) \mathrm{d}t - k \sum_{i=1}^n e^{-rt_i} |C^n_S(t_i, S_{t_i}) - C^n_S(t_{i-1}, S_{t_{i-1}})|.$$

Proof. Owing to Ito's lemma applied to $e^{-rt}C(t,S_t)$:

$$\tilde{H} := e^{-rT} H = e^{-rT} C(T,S_T) = C(0,S_0) + \int_0^T C_S(t,S_t) d\tilde{S}_t.$$

Then,

$$\begin{aligned}\tilde{V}_T - \tilde{H} = &\int_0^T (C^n_S(\varphi(t), S_{\varphi(t)}) - C^n_S(t,S_t)) d\tilde{S}_t + \int_0^T (C^n_S(t,S_t) - C_S(t,S_t)) d\tilde{S}_t \\ &+ (C^n(0,S_0) - C(0,S_0)) - k \sum_{i=1}^n e^{-rt_i} |C^n_S(t_i, S_{t_i}) - C^n_S(t_{i-1}, S_{t_{i-1}})|.\end{aligned} \tag{2.4.1}$$

Moreover, let $f(t,S_t) = C^n(t,S_t) - C(t,S_t)$, then

$$\begin{aligned}\mathrm{d}(e^{-rt} f(t,S_t)) &= e^{-rt}(f_t(t,S_t)\mathrm{d}t + f_S(t,S_t)\mathrm{d}S_t + \frac{\sigma^2}{2} f_{SS}(t,S_t)\mathrm{d}t) - re^{-rt} f(t,S_t)\mathrm{d}t \\ &= e^{-rt}((f_t(t,S_t) + rS_t f_S(t,S_t) + \frac{\sigma^2}{2} f_{SS}(t,S_t) - rf(t,S_t))\mathrm{d}t + \sigma f_S(t,S_t)\mathrm{d}W_t) \\ &= e^{-rt}\left(-\frac{\sigma^2 \alpha_n}{2} C^n_{SS}(t,S_t)\mathrm{d}t + \sigma f_S(t,S_t)\mathrm{d}W_t\right),\end{aligned}$$

using the equations of Lemma 2.4.2 to show $r(SC_S^n - C^n) + C_t^n + \frac{\sigma^2(1+\alpha_n)}{2} C_{SS}^n = 0$ and $r(SC_S - C) + C_t + \frac{\sigma^2}{2} C_{SS} = 0$. Moreover, $f(T,.) \equiv 0$ and then, we deduce that

$$f(0, S_0) = \int_0^T e^{-rt} \frac{\sigma^2 \alpha_n}{2} C_{SS}^n(t, S_t) \mathrm{d}t - \int_0^T \sigma e^{-rt} f_S(t, S_t) \mathrm{d}W_t$$

i.e.,

$$C^n(0, S_0) - C(0, S_0) = \int_0^T e^{-rt} \frac{\sigma^2 \alpha_n}{2} C_{SS}^n(t, S_t) \mathrm{d}t - \int_0^T (C_S^n(t, S_t) - C_S(t, S_t)) \mathrm{d}\tilde{S}_t. \tag{2.4.2}$$

Substituting equation (2.4.2) for equation (2.4.1), we conclude the proof. □

Lemma 2.4.5.

$$\lim_{n \to \infty} F_1^n = 0, \quad in\ L^2(\mathbb{P}).$$

Proof. The second moment of F_1^n writes

$$\mathbb{E}[(F_1^n)^2] = \mathbb{E}\left[\int_0^T \sigma^2 e^{-2rt} (C_S^n(\varphi(t), S_{\varphi(t)}) - C_S^n(t, S_t))^2 \mathrm{d}t\right]. \tag{2.4.3}$$

By definition (cf Lemma 2.4.2), $\mathrm{d}\mathbb{P}$ *a.s.*

$$\begin{aligned} C_S^n(t, S_t) - C_S^n(\varphi(t), S_{\varphi(t)}) &= \int_{\varphi(t)}^t \left(r S_u C_{SS}^n(u, S_u) + C_{tS}^n(u, S_u) + \frac{\sigma^2}{2} C_{SSS}^n(u, S_u) \right) \mathrm{d}u \\ &\quad + \int_{\varphi(t)}^t \sigma C_{SS}^n(u, S_u) \mathrm{d}W_u \\ &= - \int_{\varphi(t)}^t \frac{\sigma^2 \alpha_n}{2} C_{SSS}^n(u, S_u) \mathrm{d}u + \int_{\varphi(t)}^t \sigma C_{SS}^n(u, S_u) \mathrm{d}W_u. \end{aligned}$$

Then,

$$\mathbb{E}\left(C_S^n(t, S_t) - C_S^n(\varphi(t), S_{\varphi(t)})\right)^2 \leq 2h_n \int_{\varphi(t)}^t \frac{\sigma^4 \alpha_n^2}{4} \mathbb{E}\left(C_{SSS}^n(u, S_u)^2\right) \mathrm{d}u + 2 \int_{\varphi(t)}^t \sigma^2 \mathbb{E}\left(C_{SS}^n(u, S_u)^2\right) \mathrm{d}u.$$

To deal with integrals of this type, we state a handy result, which stems from a straightforward application of Ito's Lemma to $s \mapsto (Z_s - Z_{t_i})(t_{i+1} - s)$ between t_i and t_{i+1}; after summing from $i = 0$ to $n - 1$,

$$\int_0^T (Z_s - Z_{\varphi(s)}) ds = \int_0^T (\varphi(s) + h_n - s) dZ_s, \tag{2.4.4}$$

for all continuous semi-martingales Z.
Therefore,

$$\mathbb{E}\left(F_1^n\right)^2 \leq_c \int_0^T (\varphi(u) + h_n - u) \left[h_n \sigma^4 \alpha_n^2 \mathbb{E}\left(C_{SSS}^n(u, S_u)^2\right) + \sigma^2 \mathbb{E}\left(C_{SS}^n(u, S_u)^2\right)\right] \mathrm{d}u.$$

Owing to Lemma 2.4.3, we get

$$
\begin{aligned}
&\mathbb{E}\left(F_1^n\right)^2 \\
&\leq_c \int_0^{T-h_n} h_n\left[\frac{h_n\sigma^4\alpha_n^2}{\sigma_n^3(T-t)^{3/2}\sqrt{\sigma_n^2(T-t)+\sigma^2 t}}+\frac{\sigma^2}{\sigma_n(T-t)^{1/2}\sqrt{\sigma_n^2(T-t)+\sigma^2 t}}\right]\mathrm{d}t \\
&\quad+\int_{T-h_n}^{T}(T-t)\left[\frac{h_n\sigma^4\alpha_n^2}{\sigma_n^3(T-t)^{3/2}\sqrt{\sigma_n^2(T-t)+\sigma^2 t}}+\frac{\sigma^2}{\sigma_n(T-t)^{1/2}\sqrt{\sigma_n^2(T-t)+\sigma^2 t}}\right]\mathrm{d}t.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
\mathbb{E}\left(F_1^n\right)^2 &\leq_c \int_0^{T-h_n} h_n\left[\frac{h_n\sigma^4\alpha_n^2}{\sigma_n^3(T-t)^{3/2}\sigma\sqrt{T}}+\frac{\sigma^2}{\sigma_n\sqrt{T-t}\sigma\sqrt{T}}\right]\mathrm{d}t \\
&\quad+\int_{T-h_n}^{T}\left[\frac{h_n\sigma^4\alpha_n^2}{\sigma_n^3(T-t)^{1/2}\sigma\sqrt{T}}+\frac{\sigma^2(T-t)^{1/2}}{\sigma_n\sigma\sqrt{T}}\right]\mathrm{d}t \\
&\leq_c h_n\left[\frac{\sigma^3\alpha_n^2 h_n^{1/2}}{\sigma_n^3\sqrt{T}}+\frac{\sigma h_n}{\sigma_n}\right]+\left[h_n\sigma^4\alpha_n^2\frac{h_n^{1/2}}{\sigma_n^3\sigma\sqrt{T}}+\frac{\sigma^2 h_n^{3/2}}{\sigma_n\sigma\sqrt{T}}\right] \\
&\leq_c 2h_n^{5/4}\sqrt{\frac{\kappa}{T\sigma}}+h_n+\frac{h_n^{3/2}}{\sqrt{T}}\to 0,
\end{aligned}
$$

because κ is bounded. □

Lemma 2.4.6. *There exists a finite random variable U_0 such that*

$$\mathbb{P}\left(|F_2^n|\leq \kappa U_0\right)\to_{n\to+\infty} 1.$$

Proof. Let

$$
\begin{cases}
L_1^n = \frac{\sigma^2\alpha_n}{2}\int_0^T e^{-ru}C_{SS}^n(u,S_u)\mathrm{d}u-\frac{\sigma^2\alpha_n}{2}\int_0^T\sum_{i=1}^n e^{-rt_{i-1}}C_{SS}^n(t_{i-1},S_{t_{i-1}})1_{]t_{i-1},t_i]}(u)\mathrm{d}u, \\
L_2^n = \frac{\sigma^2\alpha_n}{2}\sum_{i=1}^n e^{-rt_{i-1}}C_{SS}^n(t_{i-1},S_{t_{i-1}})h_n-k\sigma\sum_{i=1}^n e^{-rt_i}C_{SS}^n(t_{i-1},S_{t_{i-1}})|W_{t_i}-W_{t_{i-1}}|, \\
L_3^n = k\sigma\sum_{i=1}^n e^{-rt_i}C_{SS}^n(t_{i-1},S_{t_{i-1}})|W_{t_i}-W_{t_{i-1}}|-k\sum_{i=1}^n e^{-rt_i}|M_{t_i}^n-M_{t_{i-1}}^n|, \\
L_4^n = k\sum_{i=1}^n e^{-rt_i}|M_{t_i}^n-M_{t_{i-1}}^n|-k\sum_{i=1}^n e^{-rt_i}|C_S^n(t_i,S_{t_i})-C_S^n(t_{i-1},S_{t_{i-1}})|.
\end{cases}
$$

where $M_t^n=\sigma\int_0^t C_{SS}^n(u,S_u)\mathrm{d}W_u$.
Then, F_2^n consists of $F_2^n=L_1^n+L_2^n+L_3^n+L_4^n$.

Lemma 2.4.7.

$$L_1^n\to_{n\to\infty} 0 \quad \textit{in } L^1(\mathbb{P}).$$

Proof. The norm of L_1^n is

$$\begin{aligned}
\mathbb{E}|L_1^n| &= \mathbb{E}\left|\frac{\sigma^2\alpha_n}{2}\int_0^T e^{-ru}C_{SS}^n(u,S_u)\mathrm{d}u - \frac{\sigma^2\alpha_n}{2}\int_0^T\sum_{i=1}^n e^{-rt_{i-1}}C_{SS}^n(t_{i-1},S_{t_{i-1}})1_{]t_{i-1},t_i]}(u)\mathrm{d}u\right| \\
&= \frac{\sigma^2\alpha_n}{2}\mathbb{E}\left|\int_0^T e^{-ru}C_{SS}^n(u,S_u) - e^{-r\varphi(u)}C_{SS}^n(\varphi(u),S_{\varphi(u)})\mathrm{d}u\right| \\
&= \frac{\sigma^2\alpha_n}{2}\mathbb{E}\left|\int_0^T(\varphi(t)+h_n-t)d(e^{-rt}C_{SS}^n(t,S_t))\right|.
\end{aligned}$$

But, the Ito formula gives

$$\begin{aligned}
d(e^{-rt}C_{SS}^n(t,S_t)) = e^{-rt}&\left(-rC_{SS}^n(t,S_t)+C_{SSt}^n(t,S_t)+rS_tC_{SSS}^n(t,S_t)+\frac{\sigma^2}{2}C_{SSSS}^n(t,S_t)\right)\mathrm{d}t \\
&+e^{-rt}\sigma C_{SSS}^n(t,S_t)\mathrm{d}W_t.
\end{aligned}$$

and Lemma 2.4.2 implies

$$-rC_{SS}^n(t,S_t)+C_{SSt}^n(t,S_t)+rS_tC_{SSS}^n(t,S_t)+\frac{\sigma^2}{2}C_{SSSS}^n(t,S_t) = -2rC_{SS}^n(t,S_t)-\frac{\sigma^2\alpha_n}{2}C_{SSSS}^n(t,S_t).$$

To sum up, one has

$$d(e^{-rt}C_{SS}^n(t,S_t)) = e^{-rt}\left(-2rC_{SS}^n(t,S_t)-\frac{\sigma^2\alpha_n}{2}C_{SSSS}^n(t,S_t)\right)\mathrm{d}t+e^{-rt}\sigma C_{SSS}^n(t,S_t)\mathrm{d}W_t.$$

So, owing to the Burkholder-Davis-Gundy inequality,

$$\mathbb{E}|L_1^n| \le_c A_1^n + A_2^n + A_3^n,$$

where

$$\begin{aligned}
A_1^n &= r\sigma^2\alpha_nh_n\mathbb{E}\left[\int_0^T|C_{SS}^n(t,S_t)|\,\mathrm{d}t\right], \\
A_2^n &= \frac{\sigma^2\alpha_nh_n}{2}\mathbb{E}\left[\int_0^{T-h_n}\left|\frac{\sigma^2\alpha_n}{2}C_{SSSS}^n(t,S_t)\right|\mathrm{d}t+\left(\int_0^{T-h_n}\sigma^2C_{SSS}^n(t,S_t)^2\mathrm{d}t\right)^{1/2}\right], \\
A_3^n &= \frac{\sigma^2\alpha_n}{2}\mathbb{E}\left[\int_{T-h_n}^T(T-t)\left|\frac{\sigma^2\alpha_n}{2}C_{SSSS}^n(t,S_t)\right|\mathrm{d}t+\left(\int_{T-h_n}^T\sigma^2(T-t)^2C_{SSS}^n(t,S_t)^2\mathrm{d}t\right)^{1/2}\right].
\end{aligned}$$

Using profusely Lemma 2.4.3, we have

$$\begin{aligned}
r\sigma^2\alpha_nh_n\mathbb{E}\left[\int_0^T|C_{SS}^n(t,S_t)|\mathrm{d}t\right] &\le_c r\sigma^2\alpha_nh_n\int_0^T\frac{\mathrm{d}t}{\sigma_n\sqrt{T-t}} \\
&\le_c r\sqrt{\sigma\kappa T}h_n^{3/4},
\end{aligned}$$

$$\begin{aligned}
\sigma^2\alpha_nh_n\mathbb{E}\left[\int_0^{T-h_n}\sigma^2\alpha_n|C_{SSSS}^n(t,S_t)|\mathrm{d}t\right] &\le_c \sigma^4\alpha_n^2h_n\int_0^{T-h_n}\frac{\mathrm{d}t}{\sigma_n^3(T-t)^{3/2}} \\
&\le_c \sqrt{\sigma\kappa}h_n^{1/4},
\end{aligned}$$

$$\sigma^2\alpha_n\mathbb{E}\left[\int_{T-h_n}^{T}(T-t)\sigma^2\alpha_n|C^n_{SSSS}(t,S_t)|\mathrm{d}t\right] \leq_c \sigma^4\alpha_n^2\int_{T-h_n}^{T}\frac{T-t}{\sigma_n^2(T-t)\sigma\sqrt{T}}\mathrm{d}t$$
$$\leq_c \frac{\kappa}{\sqrt{T}}\sqrt{h_n},$$

$$\sigma^3\alpha_n h_n\mathbb{E}\left[\int_{0}^{T-h_n}C^n_{SSS}(t,S_t)^2\mathrm{d}t\right]^{1/2} \leq_c \sigma^3\alpha_n h_n\left(\int_0^{T-h_n}\frac{\mathrm{d}t}{\sigma_n^4(T-t)^2}\right)^{1/2}$$
$$\leq_c \sigma\sqrt{h_n},$$

$$\sigma^3\alpha_n\mathbb{E}\left[\int_{T-h_n}^{T}(T-t)^2C^n_{SSS}(t,S_t)^2\mathrm{d}t\right]^{1/2} \leq_c \sigma^3\alpha_n\left(\int_{T-h_n}^{T}\frac{(T-t)^2}{\sigma_n^4(T-t)^2}\mathrm{d}t\right)^{1/2}$$
$$\leq_c \sigma\sqrt{h_n}.$$

As a consequence, all the previous moments converge to 0, when n tends to $+\infty$. □

Lemma 2.4.8.

$$L_2^n \to_{n\to\infty} 0 \quad in\ L^2(\mathbb{P}).$$

Proof. Using Lemma 2.4.1, we have

$$\left|\frac{\sigma^2\alpha_n}{2}\sum_{i=1}^{n}(e^{-rt_{i-1}}-e^{-rt_i})C^n_{SS}(t_{i-1},S_{t_{i-1}})h_n\right| \leq \sigma^2\alpha_n r h_n^2\sum_{i=1}^{n}\frac{1}{\sigma\sqrt{T-t_{i-1}}} \leq_c r\kappa h_n^{1/2}\sqrt{T}.$$

So, r plays no crucial role in the convergence of L_2^n, we take it equal to 0 to make computations much simpler.

The increments of a Brownian motion are orthogonal, so

$$\mathbb{E}[(L_2^n)^2] = \mathbb{E}\left[\frac{\sigma^2\alpha_n}{2}\sum_{i=1}^{n}C^n_{SS}(t_{i-1},S_{t_{i-1}})h_n - k\sigma\sum_{i=1}^{n}C^n_{SS}(t_{i-1},S_{t_{i-1}})|W_{t_i}-W_{t_{i-1}}|\right]^2$$
$$= (k\sigma)^2\sum_{i=1}^{n}\mathbb{E}\left[C^n_{SS}(t_{i-1},S_{t_{i-1}})^2\left(\sqrt{\frac{2h_n}{\pi}}-|W_{t_i}-W_{t_{i-1}}|\right)^2\right].$$

But,

$$\mathbb{E}_{t_{i-1}}\left[\left(\sqrt{\frac{2h_n}{\pi}}-|W_{t_i}-W_{t_{i-1}}|\right)^2\right] = \left(1-\frac{2}{\pi}\right)h_n.$$

From Lemma 2.4.3, we have

$$\mathbb{E}[(L_2^n)^2] \leq_c (k\sigma)^2 h_n\sum_{i=0}^{n-1}\frac{1}{\sigma_n\sqrt{T-t_i}\sigma\sqrt{T}}$$
$$\leq_c \frac{(k\sigma)^2}{\sigma_n\sigma}\frac{1}{\sqrt{T}}\int_0^T\frac{\mathrm{d}t}{\sqrt{T-t}} \leq_c \kappa^{3/2}\sigma^{1/2}h_n^{1/4}.$$

The result is proved. □

Lemma 2.4.9.

$$L_3^n \to_{n\to\infty} 0,$$

in probability.

Proof.

$$\begin{aligned}|L_3^n| &= \left| k\sigma \sum_{i=1}^n e^{-rt_i} C_{SS}^n(t_{i-1}, S_{t_{i-1}})|W_{t_i} - W_{t_{i-1}}| - k\sum_{i=1}^n e^{-rt_i}|M_{t_i}^n - M_{t_{i-1}}^n| \right| \\ &\le k\sigma \sum_{i=1}^n \left| \int_{t_{i-1}}^{t_i} (C_{SS}^n(t_{i-1}, S_{t_{i-1}}) - C_{SS}^n(t, S_t))\mathrm{d}W_t \right|, \qquad (C_{SS}^n \ge 0,\ r \ge 0).\end{aligned}$$

Set $\epsilon_i = \kappa \left| \int_{t_{i-1}}^{t_i} (C_{SS}^n(t_{i-1}, S_{t_{i-1}}) - C_{SS}^n(t, S_t))\mathrm{d}W_t \right| \in \mathcal{F}_{t_i}$, we aim at proving $\mathbb{E}\left(\sum_{i=1}^n \epsilon_i\right) \to_{n\to+\infty} 0$. First, we have the crude estimation

$$\begin{aligned}\mathbb{E}[\epsilon_n] &\le \sqrt{\mathbb{E}[\epsilon_n^2]} \\ &= \kappa \sqrt{\mathbb{E}\left[\int_{t_{n-1}}^T (C_{SS}^n(t_{n-1}, S_{t_{n-1}}) - C_{SS}^n(t, S_t))^2 \mathrm{d}t\right]} \\ &\le_c \kappa \left(\int_{t_{n-1}}^T \frac{4\mathrm{d}t}{\sigma_n\sqrt{T-t}\sigma\sqrt{T}}\right)^{1/2} \\ &\le_c \kappa \left(\frac{\sqrt{h_n}}{\sigma_n\sigma\sqrt{T}}\right)^{1/2} \to_{n\to+\infty} 0.\end{aligned}$$

Second, we handle $\sum_{i=1}^{n-1} \mathbb{E}[\epsilon_i]$. We have

$$d(C_{SS}^n(t, S_t)) = \left(-rC_{SS}^n(t, S_t) - \frac{\sigma^2\alpha_n}{2} C_{SSSS}^n(t, S_t)\right)\mathrm{d}t + \sigma C_{SSS}^n(t, S_t)\mathrm{d}W_t.$$

We remark that

$$\begin{aligned}\mathbb{E}[\epsilon_i] &= \kappa\mathbb{E}\left| \int_{t_{i-1}}^{t_i} (C_{SS}^n(t_{i-1}, S_{t_{i-1}}) - C_{SS}^n(t, S_t))\mathrm{d}W_t \right| \\ &= \kappa\mathbb{E}\left| W_{t_i}\int_{t_{i-1}}^{t_i} \mathrm{d}C_{SS}^n(t, S_t) - \int_{t_{i-1}}^{t_i} W_t \mathrm{d}C_{SS}^n(t, S_t) - \int_{t_{i-1}}^{t_i} \mathrm{d}\left\langle W, C_{SS}^n(., S_.)\right\rangle_t \right|.\end{aligned}$$

Then, we have

$$\begin{aligned}\mathbb{E}[\epsilon_i] = \kappa\mathbb{E}\Bigg| &\int_{t_{i-1}}^{t_i} (W_{t_i} - W_t)\left(-rC_{SS}^n(t, S_t)) - \frac{\sigma^2\alpha_n}{2} C_{SSSS}^n(t, S_t)\right)\mathrm{d}t \\ &+(W_{t_i} - W_{t_{i-1}})\int_{t_{i-1}}^{t_i} \sigma C_{SSS}^n(t, S_t)\mathrm{d}W_t - \int_{t_{i-1}}^{t_i} (W_t - W_{t_{i-1}})\sigma C_{SSS}^n(t, S_t)\mathrm{d}W_t \\ &- \int_{t_{i-1}}^{t_i} \sigma C_{SSS}^n(t, S_t)\mathrm{d}t \Bigg|. \end{aligned} \tag{2.4.5}$$

The independence of the increments of a Brownian motion and $\mathbb{E}_{t_{i-1}}|W_{t_i} - W_t| = \sqrt{\frac{2(t_i - t)}{\pi}}$ entail that

$$\begin{aligned}
\mathbb{E}[\epsilon_i] \leq_c \kappa \int_{t_{i-1}}^{t_i} & \left(\sqrt{t_i - t}\mathbb{E}|C^n_{SS}(t,S_t)| + \sigma^2\alpha_n\sqrt{t_i - t}\mathbb{E}|C^n_{SSSS}(t,S_t)|\right) \mathrm{d}t \\
& + 2\kappa h_n^{1/2}\left(\int_{t_{i-1}}^{t_i} \sigma^2\mathbb{E}\left(C^n_{SSS}(t,S_t)^2\right)\mathrm{d}t\right)^{1/2} \\
& + \kappa\sigma\mathbb{E}\left(\int_{t_{i-1}}^{t_i} (W_t - W_{t_{i-1}})^2 C^n_{SSS}(t,S_t)^2 \mathrm{d}t\right)^{1/2},
\end{aligned}$$

where the upper bound for the last term in 2.4.5 is similar to that of the third one. From Lemma 2.4.3, we have

$$\begin{aligned}
\kappa\sqrt{h_n}\sum_{i=1}^{n-1}\mathbb{E}\left[\int_{t_{i-1}}^{t_i} |C^n_{SS}(t,S_t)|\mathrm{d}t\right] &\leq_c \kappa\sqrt{h_n}\int_0^T \frac{\mathrm{d}t}{\sigma_n\sqrt{T-t}} \\
&\leq_c \frac{\kappa^{1/2}T^{1/2}}{\sigma^{1/2}} h_n^{3/4}.
\end{aligned}$$

We get the following estimates

$$\begin{aligned}
\kappa\alpha_n\sigma^2\sqrt{h_n}\sum_{i=1}^{n-1}\mathbb{E}\left[\int_{t_{i-1}}^{t_i} |C^n_{SSSS}(t,S_t)|\mathrm{d}t\right] &\leq_c \kappa\alpha_n\sigma^2\sqrt{h_n}\int_0^{T-h_n} \frac{\mathrm{d}t}{\sigma_n^3(T-t)^{3/2}} \\
&\leq_c \frac{\kappa^{1/2}}{\sigma^{1/2}} h_n^{1/4},
\end{aligned}$$

$$\begin{aligned}
\kappa\sum_{i=1}^{n-1}\sqrt{h_n\sigma^2\mathbb{E}\left[\int_{t_{i-1}}^{t_i} C^n_{SSS}(t,S_t)^2\mathrm{d}t\right]} &\leq \kappa\sigma\sqrt{h_n}\sum_{i=1}^{n-1}\sqrt{\int_{t_{i-1}}^{t_i}\frac{\mathrm{d}t}{\sigma_n^3(T-t)^{3/2}\sigma\sqrt{T}}} \\
&\leq_c \kappa\sigma h_n\sum_{i=1}^{n-1}\frac{1}{\sigma_n^{3/2}\sigma^{1/2}(T-t_i)^{3/4}T^{1/4}} \\
&\leq_c \frac{\kappa}{\sigma_n^{3/2}\sigma^{1/2}} \to_{n\to+\infty} 0,
\end{aligned}$$

$$\begin{aligned}
\kappa\sigma & \sum_{i=1}^{n-1}\mathbb{E}\left(\int_{t_{i-1}}^{t_i}(W_t - W_{t_{i-1}})^2 C^n_{SSS}(t,S_t)^2\mathrm{d}t\right)^{1/2} \\
&\leq_c \kappa\sigma\sum_{i=1}^{n-1}\left(\int_{t_{i-1}}^{t_i}\left[\mathbb{E}(W_t - W_{t_{i-1}})^6\right]^{1/3}\left[\mathbb{E}(C^n_{SSS}(t,S_t)^3)\right]^{2/3}\mathrm{d}t\right)^{1/2} \\
&\leq_c \kappa\sigma h_n\sum_{i=1}^{n-1}\left[\frac{1}{\sigma_n^5(T-t_i)^{5/2}\sigma\sqrt{T}}\right]^{2/6} \\
&\leq_c \frac{\kappa\sigma^{5/6}}{\sigma_n^{5/6}}.
\end{aligned}$$

□

Lemma 2.4.10. *There exists a positive and finite random variable U_0 such that*

$$\limsup_{n\to+\infty} |L_4^n| \le U_0 \kappa \quad \text{a.s..}$$

Proof. Ito's lemma gives

$$C_S^n(t, S_t) = C_S^n(0, S_0) + M_t^n + A_t^n,$$

where

$$M_t^n := \int_0^t \sigma C_{SS}^n(u, S_u) \mathrm{d}W_u$$

and

$$\begin{aligned} A_t^n &:= \int_0^t \left[C_{St}^n(u, S_u) + rS_u C_{SS}^n(u, S_u) + \frac{\sigma^2}{2} C_{SSS}^n(u, S_u) \right] \mathrm{d}u \\ &= \int_0^t \frac{-\sigma^2 \alpha_n}{2} C_{SSS}^n(u, S_u) \mathrm{d}u, \end{aligned}$$

by applying Lemma 2.4.2. Then,

$$\begin{aligned} |L_4^n| &= \left| k \sum_{i=1}^n e^{-rt_i} |M_{t_i}^n - M_{t_{i-1}}^n| - k \sum_{i=1}^n e^{-rt_i} |C_S^n(t_i, S_{t_i}) - C_S^n(t_{i-1}, S_{t_{i-1}})| \right| \\ &\le k \sum_{i=1}^n e^{-rt_i} |A_{t_i}^n - A_{t_{i-1}}^n| \\ &\le \kappa \sigma^2 \sum_{i=1}^n \left| \int_{t_{i-1}}^{t_i} \frac{\alpha_n}{2} C_{SSS}^n(u, S_u) \mathrm{d}u \right|. \end{aligned}$$

On $\mathcal{T} = \{S_T - K \neq 0\}$ (which has probability $\mathbb{P}(\mathcal{T}) = 1$), $\int_0^T \alpha_n |C_{SSS}^n(u, S_u)| \mathrm{d}u$ converges; indeed, let $\omega \in \mathcal{T}$,

$$\alpha_n |C_{SSS}^n(t, S_t(\omega))| = \alpha_n \frac{|S_t(\omega) - Ke^{-r(T-t)}|}{\Sigma_n (T-t)^3} N'(d^n(t, S_t(\omega))).$$

Now, let $x_t(\omega) = |S_t(\omega) - Ke^{-r(T-t)}|$ ($\omega \in \mathcal{T}$), in particular, *a.s.* $x_T > 0$ and the process x is continuous on $[0, T]$. Hence, for all $\omega \in \mathcal{T}$, there exists $0 < \epsilon(\omega) < T$ (independent of n) such that, $x_{|[T-\epsilon(\omega), T]}(\omega) > 0$ and (continuity on compact set) $\underline{x}(\omega) = \min_{T-\epsilon(\omega) \le t \le T} x_t(\omega) > 0$. We note $\overline{x}(\omega) = \max_{T-\epsilon(\omega) \le t \le T} x_t(\omega)$.
Then, on $[0, T - \epsilon(\omega)]$, we have

$$\alpha_n |C_{SSS}^n(t, S_t(\omega))| \to_{n\to+\infty} 0$$

and, using Lemma 2.5.2, we get

$$\alpha_n |C_{SSS}^n(t, S_t(\omega))| = \alpha_n \frac{|d^n(t, S_t)| N'(d^n(t, S_t))}{\Sigma_n (T-t)^2} \le_c \frac{1}{\sigma^2 (T-t)}.$$

Then, the dominated convergence theorem implies

$$\lim_{n\to+\infty}\int_0^{T-\epsilon(\omega)}\frac{\alpha_n}{2}|C^n_{SSS}(u,S_u(\omega))|\mathrm{d}u = 0 \quad a.s..$$

Moreover, on $[T-\epsilon(\omega),T]$, there exists $\mu>0$ such that

$$\alpha_n|C^n_{SSS}(t,S_t(\omega))| \le_c \overline{x}(\omega)\frac{e^{-\frac{\mu\underline{x}(\omega)^2}{2\sigma_n^2(T-t)}}}{\sqrt{\alpha_n}(T-t)^{3/2}}.$$

So, taking $y=\frac{\underline{x}(\omega)^2}{2\sigma_n^2(T-t)}$,

$$\alpha_n\int_{T-\epsilon(\omega)}^T|C^n_{SSS}(t,S_t(\omega))|\mathrm{d}t \le_c \frac{\overline{x}(\omega)}{\underline{x}(\omega)}\int_{\frac{\mu\underline{x}(\omega)^2}{2\sigma_n^2\epsilon(\omega)}}^{+\infty}\frac{e^{-y}}{\sqrt{y}}dy \le_c \frac{\overline{x}(\omega)}{\underline{x}(\omega)}.$$

Then,

$$\limsup_{n\to+\infty}\int_0^T\frac{\alpha_n}{2}|C^n_{SSS}(u,S_u(\omega))|\mathrm{d}u \le_c \frac{\overline{x}(\omega)}{\underline{x}(\omega)} \quad a.s.$$

So, we deduce the following result :

$$\limsup_{n\to+\infty}|L^n_4| \le U_0\kappa \qquad a.s.$$

where $U_0 :=_c \frac{\overline{x}(\omega)}{\underline{x}(\omega)} > 0$ $a.s.$. □

To sum up, by gathering Lemmas 2.4.7, 2.4.8, 2.4.9 and 2.4.10, we have showed the convergence stated in Lemma 2.4.6. □

Taking into account Lemma 2.4.5, we conclude the proof of Theorem 2.4.1. □

2.5 Appendix

Proof of Lemma 2.4.1. We have

$$
\begin{aligned}
\partial_{S_t} S_T &= e^{r(T-t)}, \\
\partial_S d(t,S) &= \frac{1}{\Sigma(T-t)}, \\
\partial_t \Sigma(T-t) &= -\frac{\sigma r}{\sqrt{2r}} \frac{e^{-2r(T-t)}}{\sqrt{1-e^{-2r(T-t)}}} \\
&= -\frac{\sigma^2}{2\Sigma(T-t)} e^{-2r(T-t)}, \\
\partial_t d(t,S) &= \frac{-rKe^{-r(T-t)}\Sigma(T-t) - (S-Ke^{-r(T-t)})\partial_t\Sigma(T-t)}{\Sigma(T-t)^2} \\
&= \frac{-rKe^{-r(T-t)}\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}} + (S-Ke^{-r(T-t)})\frac{\sigma}{2\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}e^{-2r(T-t)}}{\sigma^2\frac{1-e^{-2r(T-t)}}{2r}} \\
&= \frac{-Ke^{-r(T-t)}(1-e^{-2r(T-t)}) + (S-Ke^{-r(T-t)})e^{-2r(T-t)}}{2\sigma\left(\frac{1-e^{-2r(T-t)}}{2r}\right)^{3/2}} \\
&= e^{-2r(T-t)}\sigma^2\frac{S-Ke^{r(T-t)}}{2\Sigma(T-t)^3}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
C_S(t,S) &= \mathbb{E}_{t,S}\left[e^{-r(T-t)}1_{S_T\geq K}\partial_S S_T\right] \\
&= N(d(t,S)),
\end{aligned}
$$

$$
\begin{aligned}
C_{SS}(t,S) &= \partial_S d(t,S) N'(d(t,S)) \\
&= \frac{N'(d(t,S))}{\Sigma(T-t)},
\end{aligned}
$$

$$
\begin{aligned}
C_{SSS}(t,S) &= \partial_S \frac{N'(d(t,S))}{\Sigma(T-t)} \\
&= \frac{-d(t,S)N'(d(t,S))}{\Sigma(T-t)^2} \\
&= -\frac{S-Ke^{-r(T-t)}}{\Sigma(T-t)^3}N'(d(t,S)),
\end{aligned}
$$

$$
\begin{aligned}
C_{St}(t,S) &= \partial_t d(t,S) N'(d(t,S)) \\
&= e^{-2r(T-t)}\sigma^2\frac{S-Ke^{r(T-t)}}{2\Sigma(T-t)^3}N'(d(t,S)),
\end{aligned}
$$

$$
\begin{aligned}
C_{SSt}(t,S) &= \frac{\partial_t d(t,S)(-d(t,S))N'(d(t,S))\Sigma(T-t) - N'(d(t,S))\partial_t\Sigma(T-t)}{\Sigma(T-t)^2} \\
&= \frac{-e^{-2r(T-t)}\sigma^2\frac{(S-Ke^{r(T-t)})(S-Ke^{-r(T-t)})}{2\Sigma(T-t)^3} + \frac{\sigma^2}{2}\frac{e^{-2r(T-t)}}{\Sigma(T-t)}}{\Sigma(T-t)^2}N'(d(t,S)) \\
&= \frac{\sigma^2}{2\Sigma(T-t)^3}e^{-2r(T-t)}\left(1 - \frac{(S-Ke^{r(T-t)})(S-Ke^{-r(T-t)})}{\Sigma(T-t)^2}\right)N'(d(t,S)),
\end{aligned}
$$

$$
\begin{aligned}
C_{SSSS}(t,S) &= \partial_S\left(-\frac{S-Ke^{-r(T-t)}}{\Sigma(T-t)^3}N'(d(t,S))\right) \\
&= -\frac{1}{\Sigma(T-t)^3}N'(d(t,S)) - \frac{S-Ke^{-r(T-t)}}{\Sigma(T-t)^3}\frac{-d(t,S)N'(d(t,S))}{\Sigma(T-t)} \\
&= -\frac{1}{\Sigma(T-t)^3}N'(d(t,S)) - \frac{S-Ke^{-r(T-t)}}{\Sigma(T-t)^3}\left(-\frac{S-Ke^{-r(T-t)}}{\Sigma(T-t)^2}N'(d(t,S))\right) \\
&= \frac{1}{\Sigma(T-t)^3}\left(-1 + \frac{(S-Ke^{-r(T-t)})^2}{\Sigma(T-t)^2}\right)N'(d(t,S)).
\end{aligned}
$$

□

Let $N_{m,\sigma^2}(x) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-m)^2}{2\sigma^2}}dy$ and $N'_{m,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}}$ be respectively the cumulative distribution function and the density of a Gaussian variable with mean m and variance σ^2.

Lemma 2.5.1. *$\forall\gamma > 0$, $\forall 0 \leq s < t \leq T$, we have*

$$
\begin{aligned}
\mathbb{E}_s\left[N'(\gamma d^n(t,S_t))\right] &= \frac{\Sigma_n(T-t)}{\gamma}N'_{0,\frac{\Sigma_n(T-t)^2}{\gamma^2}+e^{2r(t-s)}\Sigma(t-s)^2}\left(S_s e^{r(t-s)} - Ke^{-r(T-t)}\right) \\
&\leq \frac{\Sigma_n(T-t)}{\sqrt{2\pi}\sqrt{\gamma^2\Sigma(t-s)^2 + \Sigma_n(T-t)^2}}.
\end{aligned}
$$

Proof. We observe that on $\mathcal{F}_s$, $S_t \sim \mathcal{N}(e^{r(t-s)}S_s, e^{2r(t-s)}\Sigma(t-s)^2)$.

$$
\begin{aligned}
\mathbb{E}_s\left[N'(\gamma d^n(t,S_t))\right] &= \frac{\Sigma_n(T-t)}{\gamma}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\frac{\Sigma_n(T-t)^2}{\gamma^2}}}e^{-\gamma^2\frac{\left(x-Ke^{-r(T-t)}\right)^2}{2\Sigma_n(T-t)^2}} \\
&\quad\times\frac{1}{\sqrt{2\pi e^{2r(t-s)}\Sigma(t-s)^2}}e^{-\frac{\left(x-e^{r(t-s)}S_s\right)^2}{2e^{2r(t-s)}\Sigma(t-s)^2}}dx \\
&= \frac{\Sigma_n(T-t)}{\gamma}\left(N'_{0,\frac{\Sigma_n(T-t)^2}{\gamma^2}} * N'_{0,e^{2r(t-s)}\Sigma(t-s)^2}\right)\left(S_s e^{r(t-s)} - Ke^{-r(T-t)}\right) \\
&= \frac{\Sigma_n(T-t)}{\gamma}N'_{0,\frac{\Sigma_n(T-t)^2}{\gamma^2}+e^{2r(t-s)}\Sigma(t-s)^2}\left(S_s e^{r(t-s)} - Ke^{-r(T-t)}\right)
\end{aligned}
$$

where "$*$" stands for the convolution product. □

Lemma 2.5.2. *For all $\alpha \geq 0$, there exists $\nu > 1$ such that*

$$u^\alpha N'(u) \leq N'(\nu u).$$

Proof. Let $\mu > 1$ and f be defined by

$$f(u) = u^\alpha N'((\mu-1)u) = u^\alpha \frac{e^{-(\mu-1)^2 \frac{u^2}{2}}}{\sqrt{2\pi}}.$$

Then, f is $\mathcal{C}^\infty$ and

$$f'(u) = (\alpha - (\mu-1)^2 u^2) u^{\alpha-1} \frac{e^{-(\mu-1)^2 \frac{u^2}{2}}}{\sqrt{2\pi}}.$$

So, the maximum of f is attained in $u_+ = \frac{\sqrt{\alpha}}{\mu-1}$ and

$$f(u_+) = \frac{\alpha^{\alpha/2}}{(\mu-1)^\alpha} \frac{e^{-\frac{\alpha}{2}}}{\sqrt{2\pi}}.$$

Taking μ such that $f(u_+) \leq 1$, we get the desired result. $\square$

Proof of Lemma 2.4.3. Let us use Lemma 2.5.1 and the inequality

$$\frac{1}{C} \leq \frac{\Sigma_n(t-s)}{\sigma_n \sqrt{t-s}} \leq C, \quad \forall 0 \leq s < t \leq T.$$

We successively obtain (here, the constant ν can change from line to line)

$$\begin{aligned}
\mathbb{E}_s |C^n_{SS}(t, S_t)|^{\gamma^2} &= \frac{\mathbb{E}_s[N'(\gamma d^n(t, S_t))]}{\Sigma_n(T-t)^{\gamma^2}} \\
&\leq_c \frac{\Sigma_n(T-t)^{1-\gamma^2}}{\sqrt{\Sigma_n(T-t)^2 + \Sigma(t-s)^2}} \\
&\leq_c \frac{\sigma_n^{1-\gamma^2} (T-t)^{\frac{1-\gamma^2}{2}}}{\sqrt{\sigma_n^2(T-t) + \sigma^2(t-s)}},
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_s |C^n_{SSS}(t, S)|^{\gamma^2} &\leq_c \frac{\mathbb{E}_s\left[|d^n(t, S_t)|^{\gamma^2} N'(\gamma d^n(t, S_t))\right]}{\Sigma_n(T-t)^{2\gamma^2}} \\
&\leq_c \frac{\mathbb{E}_s\left[N'(\nu d^n(t, S_t))\right]}{\Sigma_n(T-t)^{2\gamma^2}} \\
&\leq_c \frac{\sigma_n^{1-2\gamma^2} (T-t)^{\frac{1-2\gamma^2}{2}}}{\sqrt{\sigma_n^2(T-t) + \sigma^2(t-s)}},
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}_s|C^n_{SSSS}(t,S)|^{\gamma^2} &= \frac{1}{\Sigma_n(T-t)^{3\gamma^2}}\mathbb{E}_s\left(\left|-1+d^n(t,S_t)^2\right|^{\gamma^2} N'(\gamma d^n(t,S_t))\right) \\
&\leq_c \frac{1}{\Sigma_n(T-t)^{3\gamma^2}}\mathbb{E}_s\left(1+d^n(t,S_t)^{2\gamma^2}\right) N'(\gamma d^n(t,S_t)) \\
&\leq_c \frac{\mathbb{E}_s N'(\nu d^n(t,S_t))}{\Sigma_n(T-t)^{3\gamma^2}} \\
&\leq_c \frac{\sigma_n^{1-3\gamma^2}(T-t)^{\frac{1-3\gamma^2}{2}}}{\sqrt{\sigma_n^2(T-t)+\sigma^2(t-s)}}.
\end{aligned}
$$

□

Proof of Remark 2.4.1. Assume for the sake of simplicity that $r=0$.

$$
\begin{aligned}
\mathbb{E}\,|C_{SSS}(t,S_t)|^2 &= \mathbb{E}\left(\frac{(S_t-K)^2}{\sigma^6(T-t)^3}N'(2d(t,S_t))\right) \\
&= \int_{\mathbb{R}} \frac{x^2}{\sigma^6(T-t)^3}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{\sigma^2(T-t)}}\frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{(x-S_0+K)^2}{2\sigma^2 t}}\mathrm{d}x \\
&= \frac{\sqrt{\alpha_t}}{\sqrt{2\pi t}\sigma^7(T-t)^3}\int_{\mathbb{R}}\frac{x^2}{\sqrt{2\pi\alpha_t}}e^{-\frac{\left(x-\frac{S_0-K}{\sigma^2 t}\alpha_t\right)^2}{2\alpha_t}}\mathrm{d}x \\
&= \frac{\sqrt{\alpha_t}}{\sqrt{2\pi t}\sigma^7(T-t)^3}\left(\alpha_t+\left(\frac{S_0-K}{\sigma^2 t}\alpha_t\right)^2\right) \\
&\sim_{t\to T} \frac{\sqrt{\sigma^2(T-t)}}{2\sqrt{2\pi T}\sigma^7(T-t)^3}\left(\frac{\sigma^2(T-t)}{2}+\left(\frac{S_0-K}{\sigma^2 T}\frac{\sigma^2(T-t)}{2}\right)^2\right) \\
&\sim_{t\to T} \frac{1}{4\sqrt{2\pi T}\sigma^4(T-t)^{3/2}},
\end{aligned}
$$

where $\alpha_t := \frac{\sigma^4(T-t)t}{\sigma^2(T-t)+2\sigma^2 t}$. The last term is not integrable on $[0,T[$. □

Chapter 3

Almost sure optimal hedging strategy

Contents

In this chapter, we study the optimal discretization error of stochastic integrals, in the context of the hedging error in a multidimensional Itô model when the discrete rebalancing dates are stopping times. We investigate the convergence, in an *almost sure* sense, of the renormalized quadratic variation of the hedging error, for which we exhibit an asymptotic lower bound for a large class of stopping time strategies. Moreover, we make explicit a strategy which asymptotically attains this lower bound *a.s.* . Remarkably, the results hold under great generality on the payoff and the model. Our analysis relies on new results enabling to control *a.s.* processes, stochastic integrals and related increments. These results are further developed and applied in the following chapters.

3.1 Introduction

The problem We aim at finding a finite sequence of optimal stopping times $\mathcal{T}^n = \{\tau_0^n = 0 < \tau_1^n < \dots < \tau_i^n < \dots \leq \tau_{N_T^n}^n = T\}$ which minimizes the quadratic variation of the discretization error of the stochastic integral

$$Z_s^n = \int_0^s D_x u(t, S_t) \cdot \mathrm{d}S_t - \sum_{\tau_{i-1}^n \leq s} D_x u(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n}),$$

which interpretation is the hedging error [Bertsimas 2000] of the discrete Delta-hedging strategy of an European option with underlying asset S (multidimensional Itô process), maturity $T > 0$, price function u (for the ease of presentation, here u only depends on S) and payoff $g(S_T)$. The times $(\tau_i^n)_{1 \leq i \leq N_T^n}$ read as rebalancing dates (or trading dates) and their number N_T^n is a random variable which is finite *a.s.* . The exponent n refers to a control parameter introduced later on (see Section 3.2). The *a.s.* minimization of Z_T^n is hopeless since after a suitable renormalisation, it is known that it weakly converges to a mixture of Gaussian random variables (see [Bertsimas 2000, Gobet 2001, Hayashi 2005, Geiss 2009] when trading dates are deterministic and under some mild assumptions on the model and payoff; see [Fukasawa 2011b] for stopping times under stronger assumptions). Hence, it is more appropriate to investigate the *a.s.* minimization of the quadratic variation $\langle Z^n \rangle_T$ which, owing to the Lenglart inequality (resp. the Burkholder-Davis-Gundy inequality), allows the control of the distribution (resp. the L_p-moments, $p > 0$) of $\sup_{t \leq T} |Z_t^n|$ under martingale measure. To avoid trivial lower bounds by letting $N_T^n \to +\infty$, we reformulate our problem into the *a.s.* minimization of the product

$$N_T^n \langle Z^n \rangle_T. \tag{3.1.1}$$

Our Theorem 3.3.1 states that the above renormalized error has *a.s.* an asymptotic lower bound over the class of admissible strategies which consist (roughly speaking[1]) of deterministic times and of hitting times of random ellipsoids of the form

$$\tau_0^n := 0, \quad \tau_i^n := \inf\{t \geq \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n}) \cdot H_{\tau_{i-1}^n}^n (S_t - S_{\tau_{i-1}^n}) = 1\} \wedge T \tag{3.1.2}$$

[1] a precise definition is given in Section 3.2

where $(H_t^n)_{0\leq t\leq T}$ is a measurable adapted positive-definite symmetric matrix process. It includes the Karandikar scheme [Karandikar 1995] for discretization of stochastic integrals. In addition, in Theorems 3.3.2 and 3.3.3 we show the existence of a strategy of the hitting time form attaining the *a.s.* lower bound. The derivation of a Central Limit-type Theorem for Z^n is left to further research, in particular because the verification of the criteria in [Fukasawa 2011b] is difficult to handle in our general setting.

Literature background Our work extends the existing literature on discretization errors for stochastic integrals with deterministic time mesh, mainly considered with financial applications. Many works deal with hedging rebalancing at regular intervals of length $\Delta t_i = T/n$. In [Zhang 1999] and [Bertsimas 2000], the authors show that $\mathbb{E}[\langle Z^n\rangle_T]$ converges to 0 at rate n for payoffs smooth enough (this convergence rate originates to consider the product (3.1.1) as a minimization criterion). However, in [Gobet 2001] it is proved that the irregularity of the payoff may deteriorate the convergence rate: it becomes $n^{1/2}$ for digital call option. This phenomenon has been intensely analyzed by Geiss and his co-authors using the concept of fractional smoothness (see [Geiss 2004, Gobet 2010a, Geiss 2011, Geiss 2012] and references therein): by the choice of rebalancing dates suitably concentrated at maturity, we recover the rate n.

The first attempt to find optimal strategies with non-deterministic times goes back to [Martini 1999]: the authors allow a fixed number n of random rebalancing dates, which actually solve an optimal multiple-stopping problem. Numerical methods is required to compute the solution. In [Fukasawa 2011a], Fukasawa performs an asymptotic analysis for minimizing the product $\mathbb{E}(N_T^n)\mathbb{E}(\langle Z^n\rangle_T)$ (an extension to jump processes has been recently done in [Rosenbaum 2011]). Under regularity and integrability assumptions (and for a convex payoff on a single asset), Fukasawa derives an asymptotic lower bound and provides an optimal strategy. His contribution is the closest to our current work. But there are major differences.

1. We focus on *a.s.* results, which is probably more meaningful for hedging issues. We are not aware of similar works in this direction.

2. We allow a quite general model for the asset. It can be a multidimensional diffusion process (local volatility model), see the discussion in Subsection 3.5.4. As a comparison, in [Fukasawa 2011a] the analysis is carried out for a one-dimensional model (mainly Black-Scholes model).

3. We also allow a great generality on the payoff. In particular, the payoff can be discontinuous, the option can be exotic (Asian, lookback ...), see Subsection

3.5.4 for examples: for mathematical reasons, this is a major difference in comparison with [Fukasawa 2011a]. Indeed, in the latter reference, the payoff convexity is needed to ensure the positivity of the option Gamma (second derivative of price), which is a crucial property in the analysis. Also, for discontinuous payoff the L_p integrability of the sensitivities (Greeks) up to maturity may be not satisfied (see [Gobet 2011a]); thus, some quantities in the analysis (e.g. the integral of the second moment of the Gamma of digital call option) may become infinite. In our setting, we circumvent these issues by only requiring the sensitivities to be finite *a.s.* up to maturity: actually, this property is systematically satisfied by payoffs for which the discontinuity set has a zero-measure (see Subsection 3.5.4), which includes all the usual situations to our knowledge.

To achieve such a level of generality and an *a.s.* analysis, we design efficient tools to analyze the *a.s.* control and *a.s.* convergence of local martingales, of their increments and so forth. All these results represent another important theoretical contribution of this work. Other applications of these techniques are in the following chapters. At last, although the distribution of hitting time of random ellipsoid of the form (3.1.2) is not explicit, quite surprisingly we obtain tight estimates on the maximal increments of $\sup_{i \le N_T^n}(\tau_i^n - \tau_{i-1}^n)$, which may have applications in other areas (like stochastic simulation).

Outline of the chapter In the following, we present some notations and assumptions that will be used throughout the paper. Section 3.2 is aimed at defining our class of stopping time strategies and deriving some general theoretical properties in this class. For that, we establish new key results about *a.s.* convergence, which fit well our framework. All these results are not specifically related to financial applications. The main results about hedging error are stated and proved in Section 3.3. Numerical experiments are presented in Section 3.4, with a practical description of the algorithm to build the optimal sequence of stopping times (actually hitting times) and a numerical illustration regarding the exchange binary option (in dimension 2).

Notation used throughout the paper

- We denote by $x \cdot y$ the scalar product between two vectors x and y, and by $|x| = (x \cdot x)^{1/2}$ the Euclidean norm of x; the induced norm of a $m \times d$-matrix A is denoted by $|A| := \sup_{x \in \mathbb{R}^d : |x|=1} |Ax|$.

- A^* stands for the transposition of the matrix A; I_d stands for the identity matrix of size d; the trace of a square matrix A is denoted by $\mathrm{Tr}(A)$.

- $\mathcal{S}^d(\mathbb{R})$, $\mathcal{S}^d_+(\mathbb{R})$ and $\mathcal{S}^d_{++}(\mathbb{R})$ are respectively the set of symmetric, symmetric nonnegative-definite and symmetric positive-definite $d \times d$-matrices with coefficients in $\mathbb{R}$: $A \in \mathcal{S}^d_+(\mathbb{R})$ (resp. $\mathcal{S}^d_{++}(\mathbb{R})$) if and only if $x \cdot Ax \geq 0$ (resp. > 0) for any $x \in \mathbb{R}^d \backslash \{0\}$.

- For $A \in \mathcal{S}^d(\mathbb{R})$, $\Lambda(A) := (\lambda_1(A), ..., \lambda_d(A))$ stands for its spectrum (its $\mathbb{R}$-valued eigenvalues) and we set $\lambda_{\min}(A) := \min_{1 \leq i \leq d} \lambda_i(A)$.

- For the partial derivatives of a function $f : (t, x, y) \mapsto f(t, x, y)$, we write $D_t f(t, x, y) = \frac{\partial f}{\partial t}(t, x, y)$, $D_{x_i} f(t, x, y) = \frac{\partial f}{\partial x_i}(t, x, y)$, $D^2_{x_i x_j} f(t, x, y) = \frac{\partial^2 f}{\partial x_i \partial x_j}(t, x, y)$, $D^2_{x_i y_j} f(t, x, y) = \frac{\partial^2 f}{\partial x_i \partial y_j}(t, x, y)$ and so forth.

- When convenient, we adopt the short notation f_t in place of $f(t, S_t, Y_t)$ where f is a given function and $(S_t, Y_t)_{0 \leq t \leq T}$ is a continuous time process (introduced below).

- For a $\mathbb{R}^d$-valued continuous semimartingale M, $\langle M \rangle_t$ stands for the matrix of cross-variations $(\langle M^i, M^j \rangle_t)_{1 \leq i,j \leq d}$.

- The constants of the multidimensional version of the Burkholder-Davis-Gundy inequalities [Karatzas 1991, p. 166] are defined as follows: for any $p > 0$ there exists $c_p > 1$ such that for any vector $M = (M^1, \dots, M^d)$ of continuous local martingales with $M_0 = 0$ and any stopping time θ, we have
$$c_p^{-1} \mathbb{E} \Big| \sum_{j=1}^d \langle M^j \rangle_\theta \Big|^p \leq \mathbb{E} \Big(\sup_{t \leq \theta} |M_t|^{2p} \Big) \leq c_p \mathbb{E} \Big| \sum_{j=1}^d \langle M^j \rangle_\theta \Big|^p. \tag{3.1.3}$$

- For a given sequence of stopping times $\mathcal{T}^n$, the last time before $t \leq T$ is defined by $\varphi(t) = \max\{\tau^n_j ; \tau^n_j \leq t\}$: although dependent on n, we omit to indicate this dependency to alleviate notation. Furthermore, for a process $(f_t)_{0 \leq t \leq T}$, we write $\Delta f_t := f_t - f_{\varphi(t-)}$ (omitting again the index n for simplicity); in particular, we have $\Delta f_{\tau^n_i} = f_{\tau^n_i} - f_{\tau^n_{i-1}}$. Besides we set $\Delta_t = t - \varphi(t-)$ and $\Delta \tau^n_i := \tau^n_i - \tau^n_{i-1}$.

- We shortly write $X^n \xrightarrow{a.s.}$ if the random variables $(X^n)_{n \geq 0}$ converge almost surely as $n \to \infty$. We write $X^n \xrightarrow{a.s.} X^\infty$ to additionally indicate that the almost sure limit is equal to X^∞. We shall say that the sequence $(X^n)_{n \geq 0}$ is bounded if $\sup_{n \geq 0} |X^n| < +\infty$, *a.s.*.

- C_0 is a *a.s.* finite non-negative random variable, which may change from line to line.

Model Let $T > 0$ be a given terminal time (maturity) and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space, supporting a d-dimensional Brownian motion $B = (B^i)_{1\leq i\leq d}$ defined on $[0,T]$, where $(\mathcal{F}_t)_{0\leq t\leq T}$ is the $\mathbb{P}$-augmented natural filtration of B and $\mathcal{F} = \mathcal{F}_T$. This stochastic basis serves as a modeling of the evolution of d tradable risky assets without dividends, which price processes are denoted by $S = (S^i)_{1\leq i\leq d}$. Their dynamics are given by an Itô continuous semimartingale which solves

$$S_t = S_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s \tag{3.1.4}$$

with measurable and adapted coefficients b and σ. This is the usual framework of complete market, see [Musiela 2005]. Assumptions on σ are given below. Furthermore, for the sake of simplicity we directly assume that the return of the money market account $(r_t)_t$ is zero and that $b \equiv 0$. This simplification is not really a restriction (see [Musiela 2005] for details): indeed, first we can still re-express prices in the money market account numéraire; second, because we deal with *a.s.* results, we can consider dynamics under any equivalent probability measure, and we choose the martingale measure.

From now on, S is a *continuous local martingale* and σ satisfies the following assumption.

$(\mathbf{A}_\sigma)$ *a.s.* for any $t \in [0,T]$ σ_t is non zero; moreover σ satisfies the continuity condition : there exist a parameter $\theta_\sigma \in (0,1]$ and a non-negative *a.s.* finite random variable C_0 such that

$$|\sigma_t - \sigma_s| \leq C_0(|S_t - S_s|^{\theta_\sigma} + |t-s|^{\frac{\theta_\sigma}{2}}), \quad \forall 0 \leq s,t \leq T \quad a.s..$$

The above continuity condition is satisfied if $\sigma_t := \sigma(t, S_t)$ for a function $\sigma(.)$ which is θ_σ-Hölder continuous w.r.t. the parabolic distance. For some of our results, the above assumption is strengthened into

$(\mathbf{A}_\sigma^{\text{Ellip.}})$ Assume $(\mathbf{A}_\sigma)$ and that σ_t is elliptic in the sense

$$0 < \lambda_{\min}(\sigma_t \sigma_t^*), \quad \forall\, 0 \leq t \leq T \quad a.s..$$

The assumption $(\mathbf{A}_\sigma^{\text{Ellip.}})$ is undemanding, since we do not suppose any uniform (in ω) lower bound.

We consider an exotic option written on S with payoff $g(S_T, Y_T)$ where Y_T is a functional of $(S_t)_{0\leq t\leq T}$. In the subsequent asymptotic analysis, we assume that $Y = (Y^i)_{1\leq i\leq d'}$ is a vector of adapted continuous non-decreasing processes. Examples of such an option are given below: this illustrates that the current setting covers numerous relevant situations beyond the case of simple vanilla options (with payoff of form $g(S_T)$).

Example 3.1.1. *1. Asian options : $Y_t^j := \int_0^t S_s^j \mathrm{d}s$ and $g(x, y) := (\sum_{j=1}^q \pi_j y^j - K)_+$, for some weights π_j and a given $K \in \mathbb{R}$.*

2. Lookback options : $Y_t^j := \max_{0\le s\le t} S_s^j$ and $g(x, y) := \sum_{j=1}^q (\pi_j y^j - \pi_j' x^j)$.

Furthermore, we assume that the price at time t of such an option is given by $u(t, S_t, Y_t)$ where u is a $\mathcal{C}^{1,3,1}\big([0, T[\times\mathbb{R}^d \times \mathbb{R}^{d'}\big)$ function verifying

$$u(T, S_T, Y_T) = g(S_T, Y_T) \quad \text{and} \quad u(t, S_t, Y_t) = u(0, S_0, Y_0) + \int_0^t D_x u(s, S_s, Y_s) \cdot \mathrm{d}S_s \tag{3.1.5}$$

for any $t \in [0, T]$. The above set of conditions is related to probabilistic and analytical properties. First, although not strictly equivalent, it essentially means that the pair (S, Y) forms a Markov process and this originates why the randomness of the fair price $\mathbb{E}(g(S_T, Y_T)|\mathcal{F}_t)$ at time t only comes from (S_t, Y_t). Observe that this Markovian assumption about (S, Y) is satisfied in the above examples. Secondly, the regularity of the price function u is usually obtained by applying PDE results thanks to Feynman-Kac representations: it is known that the expected regularity can be achieved under different assumptions on the smoothness of the coefficients of S and Y, of the payoff g, combined with some appropriate non-degeneracy conditions on (S, Y). The pictures are multiple and it is not our current aim to list all the known related results; we refer to [Wilmott 1994] for various Feynman-Kac representations related to exotic options, and to [Pascucci 2011] for regularity results and references therein. See Subsection 3.5.4 for extra regularity results. Besides, we assume

$(\mathbf{A}_u)$ Let $\mathcal{A} \in \mathcal{D} := \Big\{D^2_{x_j x_k}, D^3_{x_j x_k x_l}, D^2_{t x_j}, D^2_{x_j y_m} : 1 \le j, k, l \le d, 1 \le m \le d'\Big\}$,

$$\mathbb{P}\left(\lim_{\delta\to 0}\sup_{0\le t<T}\sup_{|x-S_t|\le\delta, |y-Y_t|\le\delta}\big|\mathcal{A}u(t, x, y)\big| < +\infty\right) = 1.$$

Observe that the above assumption is really weak: this is a pathwise result and we do not require any L_p-integrability of the derivatives of u. In Subsection 3.5.4, we provide an extended list of payoffs (continuous or not) of options (vanilla, Asian, lookback) in log-normal or local volatility models, for which $(\mathbf{A}_u)$ holds. Even for the simple option payoff $g(S_T)$ in the simple log-normal model, we have not been able to exhibit a payoff function g for which $(\mathbf{A}_u)$ is not satisfied.

3.2 Class $\mathcal{T}^{\text{adm.}}$ of strategies and convergence results

In this section, we define the class of strategies under consideration, and establish some preliminary almost sure convergence results in connection with this class.

A strategy is a finite sequence of increasing stopping times $\{\tau_0 = 0 < \tau_1 < ... < \tau_i < ... \leq \tau_{N_T} = T\}$ (with $N_T < +\infty$ *a.s.*) which stand for the rebalancing dates. Furthermore, the number of risky assets held on each interval $[\tau_i, \tau_{i+1})$ follows the usual Delta-neutral rule $D_x u(\tau_i, S_{\tau_i}, Y_{\tau_i})$.

3.2.1 Assumptions

Now to derive *asymptotically* optimal results, we consider a sequence of strategies indexed by the integers $n = 0, 1, \dots$, i.e. writing

$$\mathcal{T}^n := \{\tau_0^n = 0 < \tau_1^n < ... < \tau_i^n < ... \leq \tau_{N_T^n}^n\}, \qquad \text{for } n = 0, 1, \dots,$$

and we define an appropriate *asymptotic framework*, as the convergence parameter n goes to infinity. Let $(\varepsilon_n)_{n\geq 0}$ be a sequence of positive deterministic real numbers converging to 0 as $n \to \infty$; assume that it is a square-summable sequence

$$\sum_{n\geq 0} \varepsilon_n^2 < +\infty. \tag{3.2.1}$$

On the one hand, the parameter $\varepsilon_n^{-2\rho_N}$ (for some $\rho_N \geq 1$) upper bounds (up to a constant) the number of rebalancing dates of the strategy $\mathcal{T}^n$, i.e.

($\mathbf{A}_N$) The following non-negative random variable is *a.s.* finite:

$$\sup_{n\geq 0} \left(\varepsilon_n^{2\rho_N} N_T^n\right) < +\infty$$

for a parameter ρ_N satisfying $1 \leq \rho_N < (1 + \frac{\theta_\sigma}{2}) \wedge \frac{4}{3}$.

On the other hand, the parameter ε_n controls the size of variations of S between two stopping times in $\mathcal{T}^n$.

($\mathbf{A}_S$) The following non-negative random variable is *a.s.* finite:

$$\sup_{n\geq 0} \left(\varepsilon_n^{-2} \sup_{1\leq i\leq N_T^n} \sup_{t\in(\tau_{i-1}^n, \tau_i^n]} |S_t - S_{\tau_{i-1}^n}|^2\right) < +\infty.$$

Observe that assumptions ($\mathbf{A}_N$) and ($\mathbf{A}_S$) play complementary (and not equivalent) roles. We are now ready to define the class of sequence of strategies in which we are seeking the optimal element.

Definition 3.2.1. *A sequence of strategies $\mathcal{T} := \{\mathcal{T}^n : n \geq 0\}$ is* **admissible** *if it fulfills the hypotheses* ($\mathbf{A}_N$) *and* ($\mathbf{A}_S$)*. The set of admissible sequences $\mathcal{T}$ is denoted by $\mathcal{T}^{\text{adm.}}$.*

The above definition depends on the sequence $(\varepsilon_n)_{n\geq 0}$, which is fixed from now on.

Remark 3.2.1.
- *The larger ρ_N, the wider the class of strategies under consideration. The choice $\rho_N = 1$ is allowed, but seemingly it rules out deterministic strategies; see the next remark.*
- *If $\rho_N > 1$, a strategy $\mathcal{T}^n$ consisting of $N_T^n = 1 + \lfloor \varepsilon_n^{-2\rho_N} \rfloor$ deterministic times with mesh size $\sup_{1\leq i\leq N_T^n} \Delta\tau_i^n \leq C\varepsilon_n^{2\rho_N}$ (this includes the cases of uniform and some non-uniform time grids) forms an admissible sequence of strategies, thanks to the $\frac{1}{2}^{-}$-Hölder property of the Dambis-Dubins-Schwarz Brownian motion of S^j $(1 \leq j \leq d)$ (under the additional assumption that σ is uniformly bounded to safely maintain the time-changes into a fixed compact interval).*
- *Our setting allows to consider stopping times satisfying the* strong predictability condition *(i.e. τ_i^n is $\mathcal{F}_{\tau_{i-1}^n}$-measurable), see [Jacod 2012, Chapter 14].*
- *We show in Proposition 3.2.4 that the strategy $\mathcal{T}^n$ of successive hitting times of ellipsoid of size ε_n forms a sequence in $\mathcal{T}^{\text{adm.}}$.*
- *In Subsections 3.2.3-3.2.4, we investigate properties of admissible sequences of strategies. Among others, we show that the mesh size of $\mathcal{T}^n$ shrinks* a.s. *to 0 and we establish tight* a.s. *upper bounds (see Corollary 3.2.2): namely for any $\rho \in (0, 2]$, there is a* a.s. *finite random variable C_ρ such that $\sup_{1\leq i\leq N_T^n} \Delta\tau_i^n \leq C_\rho \varepsilon_n^{2-\rho}$ for any $n \geq 0$.*

We require an extra technical condition on the non-decreasing process Y which is fulfilled in practical cases for an admissible sequence of strategies.

$(\mathbf{A}_Y)$ The following non-negative random variable is *a.s.* finite: for some $\rho_Y > 4(\rho_N - 1)$

$$\sup_{n\geq 0}\left(\varepsilon_n^{-\rho_Y} \sup_{1\leq i\leq N_T^n} |\Delta Y_{\tau_i^n}|\right) < +\infty.$$

Example 3.2.1. *Let $\mathcal{T} := \{\mathcal{T}^n : n \geq 0\}$ satisfy* $(\mathbf{A}_S)$*-*$(\mathbf{A}_N)$*.*

1. *Asian options : applying Corollary 3.2.2 (item* ii*) with $\rho = \frac{2}{3}$ and taking $\rho_Y = \frac{4}{3} > 4(\rho_N - 1)$ (since $\rho_N < \frac{4}{3}$) gives*

$$\sup_{n\geq 0}\left(\varepsilon_n^{-\rho_Y} \sup_{1\leq i\leq N_T^n} |\Delta Y_{\tau_i^n}|\right) \leq \sup_{0\leq t\leq T} |S_t| \sup_{n\geq 0}\left(\varepsilon_n^{\rho-2} \sup_{1\leq i\leq N_T^n} \Delta\tau_i^n\right) < +\infty \quad \text{a.s.}.$$

2. *Lookback options : clearly, we have*

$$\sup_{n\geq 0}\left(\varepsilon_n^{-1} \sup_{1\leq i\leq N_T^n} |\Delta Y_{\tau_i^n}|\right) \leq \sup_{n\geq 0}\left(\varepsilon_n^{-1} \sup_{0\leq t\leq T} |\Delta S_t|\right) < +\infty \quad \text{a.s.},$$

thus $(\mathbf{A}_Y)$ *is satisfied with $\rho_Y = 1$ provided that $\rho_N < 5/4$.*

3.2.2 Fundamental lemmas about almost sure convergence

This subsection is devoted to the main ingredient (Lemmas 3.2.1 and 3.2.2) about almost sure convergence, which is involved in the subsequent asymptotic analysis. We first recall some usual approaches to establish that a sequence $(U^n_T)_{n\geq 0}$ converges to 0 in probability or almost surely, as $n \to \infty$: it serves as a preparation for the comparative discussion we will have regarding our almost sure convergence results.

- *Convergence in probability.* It can be handled, for instance, by using the Markov inequality and showing that the L_p-moment (for some $p > 0$) of U^n_T converges to 0: for $p = 1$ and $\delta > 0$, it writes $\mathbb{P}(|U^n_T| \geq \delta) \leq \frac{\mathbb{E}|U^n_T|}{\delta} \to_{n\to\infty} 0$. Observe that this approach requires a bit of integrability of the random variable U^n_T.
  To achieve the uniform convergence in probability of $(U^n_t)_{0\leq t\leq T}$ to 0, Lenglart [Lenglart 1977] introduced an extra condition: the relation of domination. Namely, assume that $(U^n_t)_{0\leq t\leq T}$ is a non-negative continuous adapted process and that it is dominated by a non-decreasing continuous adapted process $(V^n_t)_{0\leq t\leq T}$ (with $V^n_0 = 0$) in the sense $\mathbb{E}(U^n_\theta) \leq \mathbb{E}(V^n_\theta)$ for any stopping time $\theta \in [0, T]$. Then, for any $c_1, c_2 > 0$ we have
  $$\mathbb{P}\left(\sup_{t\leq T} U^n_t \geq c_1\right) \leq \frac{1}{c_1}\mathbb{E}(V^n_T \wedge c_2) + \mathbb{P}(V^n_T \geq c_2).$$
  A standard application consists in taking U^n as the square of a continuous local martingales M^n; then, the convergence in probability of $\langle M^n, M^n\rangle_T$ to 0 implies the uniform convergence in probability of $(M^n_t)_{0\leq t\leq T}$ to 0. The converse is also true, the relation of domination deriving from BDG inequalities. This kind of result leads to useful tools for establishing the convergence in probability of triangular arrays of random variables: for instance, see [Genon-Catalot 1993, Lemma 9] in the context of parametric estimation of stochastic processes.

- *Almost sure convergence.* We may use a Borel-Cantelli type argument, assuming that $\sum_{n\geq 0}\mathbb{E}|U^n_T| < +\infty$. Fubini-Tonelli's theorem yields that the series $\sum_{n\geq 0}|U^n_T|$ converges *a.s.* , and in particular $U^n_T \xrightarrow{a.s.} 0$. Here again, the integrability of U^n_T is required.
  Bichteler and Karandikar leveraged this type of series argument to establish the *a.s.* convergence of stochastic integrals under various assumptions, with in view either approximation issues or pathwise stochastic integration; see [Bichteler 1981], [Karandikar 1989], [Karandikar 1995], [Karandikar 2006] and references therein.

Our result below (Lemma 3.2.1) is inspired by the above references, but its conditions of applicability are less stringent and it allows more flexibility in our framework. We assume a relation of domination, but:

1. not for all stopping times (as in Lenglart domination);
2. the processes $(U_t^n)_{0\le t\le T}$ are not assumed to be continuous (nor $(\sum_{n\ge 0} U_t^n)_{0\le t\le T}$);
3. the dominating process V^n is not assumed to be non-decreasing.

Thus, our assumptions are less demanding, but on the other way, we do not obtain any uniform convergence result. Moreover, we emphasize that we do not assume any integrability on U_T^n. This is crucial, because the typical applications of Lemma 3.2.1 are related to U_T^n defined as a (possibly stochastic) integral of the derivatives of u evaluated along the path $(S_t, Y_t)_{0\le t\le T}$: since usual payoff functions are irregular, it is known that the L_p-moments of related derivatives blow up as time goes to maturity, and it is hopeless to obtain the required integrability on U_T^n assuming only $(\mathbf{A}_u)$.

We are now ready for the statement of our *a.s.* convergence result.

Lemma 3.2.1. *Let $\mathcal{M}_0^+$ be the set of non-negative measurable processes vanishing at $t = 0$. Let $(U^n)_{n\ge 0}$ and $(V^n)_{n\ge 0}$ be two sequences of processes in $\mathcal{M}_0^+$. Assume that*

i) the series $\sum_{n\ge 0} V_t^n$ converges for all $t \in [0,T]$, almost surely;

ii) the above limit is upper bounded by a process $\bar V \in \mathcal{M}_0^+$ and that $\bar V$ is continuous a.s. *;*

iii) there is a constant $c \ge 0$ such that, for every $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $t \in [0,T]$, we have

$$\mathbb{E}[U^n_{t\wedge\theta_k}] \le c\,\mathbb{E}[V^n_{t\wedge\theta_k}]$$

with the random time $\theta_k := \inf\{s \in [0,T] : \bar V_s \ge k\}$[2].

Then for any $t \in [0,T]$, the series $\sum_{n\ge 0} U_t^n$ converges almost surely. As a consequence, $U_t^n \xrightarrow{a.s.} 0$.

Proof. First, observe that $(\theta_k)_{k\ge 0}$ well defines random times since $\bar V$ is continuous. Denote by $\mathcal{N}_V$ the $\mathbb{P}$-negligible set on which the series $(\sum_{n\ge 0} V_t^n)_{0\le t\le T}$ do not

[2] with the usual convention $\inf\emptyset = +\infty$.

converge and on which $\bar{V}$ and then $(\theta_k)_{k\geq 0}$ are not defined; observe that for $\omega \notin \mathcal{N}_V$, we have $\bar{V}_{t\wedge\theta_k}(\omega) \leq k$ for any $t \in [0,T]$ and $k \in \mathbb{N}$. Set $\bar{V}^p := \sum_{n=0}^{p} V^n$: we have $\bar{V}^p \leq \bar{V}$ on $\mathcal{N}_V^c$; thus, the localization of $\bar{V}$ entails that of $\bar{V}^p$ and we have $\bar{V}^p_{t\wedge\theta_k} \leq k$ for any k, p and t (on $\mathcal{N}_V^c$).

Moreover, for any n and k, the relation of domination writes

$$\mathbb{E}\left[\sum_{n=0}^{p} U^n_{t\wedge\theta_k}\right] \leq c\mathbb{E}\left[\sum_{n=0}^{p} V^n_{t\wedge\theta_k}\right] = c\mathbb{E}\left[\bar{V}^p_{t\wedge\theta_k}\right] \leq ck. \tag{3.2.2}$$

From Fatou's lemma, we get $\mathbb{E}\left[\sum_{n\geq 0} U^n_{t\wedge\theta_k}\right] < +\infty$: in particular, for any $k \in \mathbb{N}$, there is a $\mathbb{P}$-negligible set $\mathcal{N}_{k,t}$, such that $\sum_{n\geq 0} U^n_{t\wedge\theta_k}(\omega)$ converges for all $\omega \notin \mathcal{N}_{k,t}$. The set $\mathcal{N}_t = \bigcup_{k\in\mathbb{N}} \mathcal{N}_{k,t} \bigcup \mathcal{N}_V$ is $\mathbb{P}$-negligible and it follows that for $\omega \notin \mathcal{N}_t$, the series $\sum_{n\geq 0} U^n_{t\wedge\theta_k}(\omega)$ converges for all $k \in \mathbb{N}$. For $\omega \notin \mathcal{N}_t$, we have $\theta_k(\omega) = +\infty$ as soon as $k > \bar{V}_T(\omega)$; thus by taking such k, we complete the convergence of $\sum_{n\geq 0} U^n_t$ on $\mathcal{N}_t^c$. □

Observe that in our argumentation, we do not assume that the non-negative random variables U^n_t and V^n_t have a finite expectation (and in some examples, it is false, especially at $t = T$). However, note that in (3.2.2) we prove that $U^n_{t\wedge\theta_k}$ and $V^n_{t\wedge\theta_k}$ have a finite expectation: in other words, $(\theta_k)_{k\geq 0}$ serves as a common localization for U^n and V^n. In addition, Lemma 3.2.1 is general and thorough since we do not assume any adaptedness or regularity properties of the processes U^n and V^n. We provide a simpler version that can be customized for our further applications:

Lemma 3.2.2. *Let $\mathcal{C}_0^+$ be the set of non-negative continuous adapted processes, vanishing at $t = 0$. Let $(U^n)_{n\geq 0}$ and $(V^n)_{n\geq 0}$ be two sequences of processes in $\mathcal{C}_0^+$. Replace the two first items of Lemma 3.2.1 by*

i') $t \mapsto V^n_t$ is a non-decreasing function on $[0,T]$, almost surely;

ii') the series $\sum_{n\geq 0} V^n_T$ converges almost surely;

iii') there is a constant $c \geq 0$ such that, for every $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $t \in [0,T]$, we have

$$\mathbb{E}[U^n_{t\wedge\theta_k}] \leq c\mathbb{E}[V^n_{t\wedge\theta_k}] \tag{3.2.3}$$

with the stopping time $\theta_k := \inf\{s \in [0,T] : \bar{V}_s \geq k\}$ setting $\bar{V}_t = \sum_{n\geq 0} V^n_t$.

Then, the conclusion of Lemma 3.2.1 still holds.

Proof. We just have to prove that items *i') + ii')* entails items *i) + ii)* of Lemma 3.2.1 for U^n and V^n in $\mathcal{C}_0^+ \subset \mathcal{M}_0^+$. Since V^n is non-decreasing, the *a.s.* convergence of $\sum_{n\geq 0} V_T^n$ implies that of $\sum_{n\geq 0} V_t^n$. Moreover $\sum_{n\geq 0}\sup_{0\leq t\leq T} V_t^n = \sum_{n\geq 0} V_T^n < +\infty$ *a.s.*: therefore, *a.s.* the series associated with V^n is normally convergent on $[0,T]$ and $\bar{V} := \sum_{n\geq 0} V^n \in \mathcal{C}_0^+$: items *i) + ii)* are satisfied. Observe θ_k is a *stopping* time since $\bar{V}$ is continuous and adapted. $\square$

We apply Lemma 3.2.2 to derive a simple criterion for the convergence of continuous local martingales.

Corollary 3.2.1. *Let $p > 0$ and let $\{(M_t^n)_{0\leq t\leq T} : n \geq 0\}$ be a sequence of scalar continuous local martingales vanishing at zero. Then,*

$$\sum_{n\geq 0}\langle M^n\rangle_T^{p/2} \xrightarrow{a.s.} \quad \Longleftrightarrow \quad \sum_{n\geq 0}\sup_{0\leq t\leq T}|M_t^n|^p \xrightarrow{a.s.} .$$

Proof. We first prove the implication $\Rightarrow$. Set $U_t^n := \sup_{0\leq s\leq t}|M_s^n|^p$ and $V_t^n := \langle M^n\rangle_t^{p/2}$ and let us check the conditions of Lemma 3.2.2 : *i')* V^n is non-decreasing, *ii')* $\sum_{n\geq 0} V_T^n$ converges *a.s.* . The relation of domination (3.2.3) follows from the BDG inequalities (see the r.h.s. of (3.1.3)) and we are done. The implication $\Leftarrow$ is proved similarly, using the l.h.s. of (3.1.3) regarding the BDG inequalities. $\square$

3.2.3 Controls of $\Delta\tau^n$ and of the martingales increments

Being inspired by the scaling property of Brownian motion, we might intuitively guess that a sequence of strategy $(\mathcal{T}^n)_{n\geq 0}$ satisfying $(\mathbf{A}_S)$ yields stopping times increments of magnitude equal roughly to ε_n^2. Actually, thorough estimates are difficult to derive: for instance the exit times of balls by a Brownian motion define unbounded random variables.
To address these issues, we take advantage of Lemma 3.2.2 to establish estimates on the sequence $(\Delta\tau_i^n := \tau_i^n - \tau_{i-1}^n)_{1\leq i\leq N_T^n}$, which show that we almost recover the familiar scaling ε_n^2.

Proposition 3.2.1. *Assume $(\mathbf{A}_\sigma)$. Let $\mathcal{T}$ be a sequence of strategies satisfying $(\mathbf{A}_S)$ and let $p \geq 0$. Then*

i) the series $\sum_{n\geq 0}\varepsilon_n^{-(p-2)}\sup_{1\leq i\leq N_T^n}(\Delta\tau_i^n)^p \xrightarrow{a.s.}$.

ii) Assume moreover that $\mathcal{T} \in \mathcal{T}^{\text{adm.}}$: the series $\sum_{n\geq 0}\varepsilon_n^{-2(p-1)+2\rho_N}\sum_{\tau_{i-1}^n<T}(\Delta\tau_i^n)^p \xrightarrow{a.s.}$.

Proof. • Let us prove *i)*, assuming only $(\mathbf{A}_S)$. For $p=0$, this is trivial.
Now consider the case $p>0$. Since σ_t is non-zero for any t and continuous, $C_E := \inf_{t\in[0,T]}(\sum_{j=1}^d e_j.\sigma_t\sigma_t^* e_j) > 0$ *a.s.* , where e_j is the j-th element of the canonical basis in $\mathbb{R}^d$. Therefore, *a.s.* for any $0\le s\le t\le T$ we have

$$\begin{aligned} 0 \le t-s \le C_E^{-1}\int_s^t \Big(\sum_{j=1}^d e_j.\sigma_r\sigma_r^* e_j\Big)dr &= C_E^{-1}\sum_{j=1}^d [\langle S^j\rangle_t - \langle S^j\rangle_s] \qquad (3.2.4)\\ &= C_E^{-1}\sum_{j=1}^d \Big[(S_t^j - S_s^j)^2 - 2\int_s^t (S_r^j - S_s^j)\mathrm{d}S_r^j\Big], \end{aligned}$$

applying the Itô formula at the last equality. Take $s=\tau_{i-1}^n$, $t=\tau_i^n$ and use $(\mathbf{A}_S)$:

$$\begin{aligned} \Delta\tau_i^n &\le C_E^{-1}\Big(C_0\varepsilon_n^2 + 2\sum_{j=1}^d \Big|\int_{\tau_{i-1}^n}^{\tau_i^n} \Delta S_r^j \mathrm{d}S_r^j\Big|\Big)\\ &\le C_E^{-1}\Big(C_0\varepsilon_n^2 + 4\sum_{j=1}^d \sup_{0\le t\le T}\Big|\int_0^t \Delta S_r^j \mathrm{d}S_r^j\Big|\Big). \qquad (3.2.5) \end{aligned}$$

Now for $j=1,\dots d$, set $M_t^{j,n} := \varepsilon_n^{2/p-1}\int_0^t \Delta S_r^j \mathrm{d}S_r^j$ (recalling that $p>0$). Then

$$\sum_{n\ge0}\langle M^{j,n}\rangle_T^{p/2} = \sum_{n\ge0}\varepsilon_n^{2-p}\Big(\int_0^T |\Delta S_t^j|^2 \mathrm{d}\langle S^j\rangle_t\Big)^{p/2} \le C_0\sum_{n\ge0}\varepsilon_n^2 < +\infty \quad a.s..$$

Thus owing to Corollary 3.2.1 the terms $\big(\sup_{0\le t\le T}\big|M_t^{j,n}\big|^p\big)_{n\ge0}$ define an *a.s.* convergent series. Combining this with (3.2.5), we finally derive

$$\sum_{n\ge0}\Big[\varepsilon_n^{2/p-1}\sup_{1\le i\le N_T^n}|\Delta\tau_i^n|\Big]^p \le C_0\Big(\sum_{n\ge0}\big[\varepsilon_n^{2/p-1}\varepsilon_n^2\big]^p + \sum_{j=1}^d\sum_{n\ge0}\sup_{0\le t\le T}\big|M_t^{j,n}\big|^p\Big) < +\infty \ a.s..$$

• It remains to justify *ii)*. For $p=0$, the result directly follows from $(\mathbf{A}_N)$ and the inequality (3.2.1) . Now take $p>0$ and set

$$U_t^n := \varepsilon_n^{-2(p-1)+2\rho_N}\sum_{\tau_{i-1}^n<t}\Big|\sum_{j=1}^d \Delta\langle S^j\rangle_{\tau_i^n\wedge t}\Big|^p, \quad V_t^n := \varepsilon_n^{-2(p-1)+2\rho_N}\sum_{\tau_{i-1}^n<t}\sup_{s\in(\tau_{i-1}^n,\tau_i^n\wedge t]}|\Delta S_s|^{2p}.$$

If $\sum_{n\ge0}U_T^n \xrightarrow{a.s.}$, (3.2.4) immediately yields that $\sum_{n\ge0}\varepsilon_n^{-2(p-1)+2\rho_N}\sum_{\tau_{i-1}^n<T}(\Delta\tau_i^n)^p \xrightarrow{a.s.}$. Thus, it is sufficient to show $\sum_{n\ge0}U_t^n \xrightarrow{a.s.}$, for any $t\in[0,T]$, and this is achieved by an application of Lemma 3.2.2. The sequences of processes $(U^n)_{n\ge0}$ and $(V^n)_{n\ge0}$ are in $\mathcal{C}_0^+$. Then, V^n is non-decreasing and using $(\mathbf{A}_S)$-$(\mathbf{A}_N)$

$$\sum_{n\ge0}V_T^n \le C_0\sum_{n\ge0}\varepsilon_n^{-2(p-1)+2\rho_N}\varepsilon_n^{2p}N_T^n \le C_0\sum_{n\ge0}\varepsilon_n^2 < +\infty \quad a.s..$$

Then, we deduce that items *i')* and *ii')* of Lemma 3.2.2 are fulfilled. It remains to check the relation of domination (item *iii')*). Let $k \in \mathbb{N}$. On the set $\{\tau_{i-1}^n < t \wedge \theta_k\}$, from the multidimensional BDG inequality in a conditional version, we have

$$\mathbb{E}\Big(\Big|\sum_{j=1}^{d} \Delta\langle S^j\rangle_{\tau_i^n \wedge t \wedge \theta_k}\Big|^p |\mathcal{F}_{\tau_{i-1}^n}\Big) \leq c_p \mathbb{E}\Big(\sup_{\tau_{i-1}^n < s \leq \tau_i^n \wedge t \wedge \theta_k} |\Delta S_s|^{2p} \,|\mathcal{F}_{\tau_{i-1}^n}\Big). \tag{3.2.6}$$

Then, it follows

$$\mathbb{E}\left[U_{t\wedge\theta_k}^n\right] = \varepsilon_n^{-2(p-1)+2\rho_N} \sum_{i=1}^{+\infty} \mathbb{E}\Big(1_{\tau_{i-1}^n < t\wedge\theta_k} \mathbb{E}\Big[\Big|\sum_{j=1}^{d} \Delta\langle S\rangle_{\tau_i^n\wedge t\wedge\theta_k}\Big|^p |\mathcal{F}_{\tau_{i-1}^n}\Big]\Big) \leq c_p \mathbb{E}[V_{t\wedge\theta_k}^n].$$

The proof is complete. □

As a consequence of Proposition 3.2.1, the mesh size of $\mathcal{T}^n$, i.e. $\sup_{1\leq i\leq N_T^n} \Delta\tau_i^n$, converges *a.s.* to 0 as $n \to \infty$, with some explicit rates of convergence: this is the statement below.

Corollary 3.2.2. *With the same assumptions and notations as Proposition 3.2.1, we have the following estimates, for any $\rho > 0$:*

i) Under $(\mathbf{A}_S)$, $\sup_{n\geq 0} \Big(\varepsilon_n^{\rho-1} \sup_{1\leq i\leq N_T^n} \Delta\tau_i^n\Big) < +\infty$ a.s..

ii) Under $(\mathbf{A}_S)$-$(\mathbf{A}_N)$, $\sup_{n\geq 0} \Big(\varepsilon_n^{\rho-2} \sup_{1\leq i\leq N_T^n} \Delta\tau_i^n\Big) < +\infty$ a.s..

Proof. Item *i)*. Clearly, from Proposition 3.2.1 - *i)*, we obtain $\sup_{n\geq 0}\Big(\varepsilon_n^{-(p-2)} \sup_{1\leq i\leq N_T^n} (\Delta\tau_i^n)^p\Big) < +\infty$ *a.s.* for any $p \geq 0$ and the result follows by taking $p = 2/\rho$.
Item *ii)*. We proceed similarly by observing that Proposition 3.2.1 - *ii)* gives

$$\sup_{n\geq 0}\Big(\varepsilon_n^{-2(p-1-\rho_N)} \sup_{1\leq i\leq N_T^n} (\Delta\tau_i^n)^p\Big) \leq \sup_{n\geq 0}\Big(\varepsilon_n^{-2(p-1-\rho_N)} \sum_{\tau_{i-1}^n < T} (\Delta\tau_i^n)^p\Big) < +\infty \quad a.s..$$

□

We are now in a position to control the *a.s.* convergence of some stochastic integrals appearing in our further optimality analysis. The following proposition and corollary will play a crucial role in the estimations of the error terms appearing in the main theorems (see Section 3.3).

Proposition 3.2.2. *Assume* $(\mathbf{A}_\sigma)$. *Let $\mathcal{T} = (\mathcal{T}^n)_{n\geq 0}$ be a sequence of strategies, $((M_t^n)_{0\leq t\leq T})_{n\geq 0}$ be a sequence of $\mathbb{R}$-valued continuous local martingales such that*

$\langle M^n\rangle_t = \int_0^t \alpha_r^n \mathrm{d}r$ *for a non-negative measurable adapted* α^n *satisfying the following inequality: there exists a non-negative* a.s. *finite random variable* C_α *and a parameter* $\theta \geq 0$ *such that*

$$0 \leq \alpha_r^n \leq C_\alpha(|\Delta S_r|^{2\theta} + |\Delta r|^\theta), \quad \forall\, 0 \leq r < T, \forall n \geq 0, \quad \text{a.s..}$$

Then, the following convergences hold.

i) Assume $\mathcal{T}$ *satisfies* $(\mathbf{A}_S)$ *and let* $p \geq 2$*:*

$$\sum_{n\geq 0}\Big(\varepsilon_n^{3-\frac{1+\theta}{2}p} \sum_{\tau_{i-1}^n<T}\sup_{\tau_{i-1}^n\leq t\leq \tau_i^n} |\Delta M_t^n|^p\Big) < +\infty, \quad \text{a.s..}$$

ii) Assume furthermore that $\mathcal{T}$ *satisfies* $(\mathbf{A}_N)$ *(i.e.* $\mathcal{T}\in\mathcal{T}^{\text{adm.}}$*) and let* $p>0$*:*

$$\sum_{n\geq 0}\Big(\varepsilon_n^{2-(1+\theta)p+2\rho_N} \sum_{\tau_{i-1}^n<T}\sup_{\tau_{i-1}^n\leq t\leq \tau_i^n} |\Delta M_t^n|^p\Big) < +\infty, \quad \text{a.s..}$$

Proof. Let $p>0$. Let δ be the parameter standing for $\frac{1}{2}$ under $(\mathbf{A}_S)$ and 1 under $(\mathbf{A}_S)$-$(\mathbf{A}_N)$. Set

$$U_t^n := \varepsilon_n^{-2\delta\left(\frac{p(\theta+1)}{2}-2(1-\delta)\right)+2+2\rho_N(2\delta-1)} \sum_{\tau_{i-1}^n<t}\sup_{\tau_{i-1}^n\leq s\leq \tau_i^n\wedge t} |\Delta M_t^n|^p,$$
$$V_t^n := \varepsilon_n^{-2\delta\left(\frac{p(\theta+1)}{2}-2(1-\delta)\right)+2+2\rho_N(2\delta-1)} \sum_{\tau_{i-1}^n<t}\left|\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\alpha_r^n\mathrm{d}r\right|^{p/2}.$$

Observe that the announced result reads as $\sum_{n\geq 0} U_T^n \overset{a.s.}{\longrightarrow}$. To prove this convergence, it is enough to establish that $\sum_{n\geq 0} V_T^n \overset{a.s.}{\longrightarrow}$. Indeed, following the arguments of the proof of Proposition 3.2.1-*ii)*, we can apply Lemma 3.2.2 since $(U^n)_{n\geq 0}$ and $(V^n)_{n\geq 0}$ are two sequences of continuous adapted processes and

i') V^n is non-decreasing on $[0,T]$ *a.s.* ;

iii') the domination is satisfied thanks to the BDG inequalities, similarly to (3.2.6).

Now to prove $\sum_{n\geq 0} V_T^n \overset{a.s.}{\longrightarrow}$, write

$$\sum_{n\geq 0} V_T^n \leq \sum_{n\geq 0}\varepsilon_n^{-2\delta\left(\frac{p(\theta+1)}{2}-2(1-\delta)\right)+2+2\rho_N(2\delta-1)} \sum_{\tau_{i-1}^n<T}\left|C_0(\varepsilon_n^{2\theta}+(\Delta\tau_i^n)^\theta)\Delta\tau_i^n\right|^{p/2} \quad a.s..$$

First, consider the case $(\mathbf{A}_S)$ and set $D_n^{(q)} := \sup_{1\leq i\leq N_T^n}(\Delta\tau_i^n)^q$ for $q\geq 0$: Proposition 3.2.1-*i)* yields $\mathcal{D}^{(q)} := \sum_{n\geq 0}\varepsilon_n^{-(q-2)} D_n^{(q)} < +\infty$ *a.s.* . Using $p \geq 2$, it readily follows that

$$\sum_{n\geq 0} V_T^n \leq \sum_{n\geq 0} \varepsilon_n^{-(p(\theta+1)/2-3)} C_0^{p/2} \sum_{\tau_{i-1}^n<T} (\varepsilon_n^{2\theta} + (\Delta\tau_i^n)^\theta)^{p/2} (\Delta\tau_i^n)^{p/2-1}\Delta\tau_i^n$$
$$\leq \sum_{n\geq 0} \varepsilon_n^{-(p(\theta+1)/2-3)} C_0^{p/2} 2^{p/2-1} T\big(\varepsilon_n^{p\theta} D_n^{(p/2-1)} + D_n^{((\theta+1)p/2-1)}\big)$$
$$\leq C_0^{p/2} 2^{p/2-1} T\big((\sup_{n\geq 0}\varepsilon_n)^{p\theta/2}\mathcal{D}^{(p/2-1)} + \mathcal{D}^{((\theta+1)p/2-1)}\big) < +\infty \quad a.s..$$

Second for the case $(\mathbf{A}_S)$-$(\mathbf{A}_N)$, setting $D_n^{(q)} := \sum_{\tau_{i-1}^n<T} (\Delta\tau_i^n)^q$ for $q\geq 0$, we have $\mathcal{D}^{(q)} := \sum_{n\geq 0} \varepsilon_n^{-2(q-1)+2\rho_N} D_n^{(q)} < +\infty$ *a.s.* thanks to Proposition 3.2.1-*ii)*. Then we easily deduce (for any $p>0$)

$$\sum_{n\geq 0} V_T^n \leq C_0^{p/2} 2^{(p/2-1)_+} \sum_{n\geq 0} \varepsilon_n^{-2(p(\theta+1)/2-1)+2\rho_N} \sum_{\tau_{i-1}^n<T} \big(\varepsilon_n^{p\theta}(\Delta\tau_i^n)^{p/2} + (\Delta\tau_i^n)^{(\theta+1)p/2}\big)$$
$$= C_0^{p/2} 2^{(p/2-1)_+}\big(\mathcal{D}^{(p/2)} + \mathcal{D}^{((\theta+1)p/2)}\big) < +\infty \quad a.s..$$

□

A straightforward consequence of the aforementioned proposition is given by the following corollary, which proof is left to the reader.

Corollary 3.2.3. *Using the assumptions and notations of Proposition 3.2.2, we have the following estimates, for any $\rho>0$:*

i) Under $(\mathbf{A}_S)$, $\sup_{n\geq 0}\left(\varepsilon_n^{\rho-\frac{1+\theta}{2}} \sup_{1\leq i\leq N_T^n} \sup_{\tau_{i-1}^n\leq t\leq \tau_i^n} |\Delta M_t|\right) < +\infty$, a.s. .

ii) Under $(\mathbf{A}_S)$-$(\mathbf{A}_N)$, $\sup_{n\geq 0}\left(\varepsilon_n^{\rho-(1+\theta)} \sup_{1\leq i\leq N_T^n} \sup_{\tau_{i-1}^n\leq t\leq \tau_i^n} |\Delta M_t|\right) < +\infty$, a.s. .

Remark 3.2.2. *Observe that in the proofs of Subsection 3.2.3, we have not used the knowledge of the upper bound on ρ_N (stated in $(\mathbf{A}_N)$): it means that all the related results are true for any admissible sequence of strategies assuming only $\rho_N \geq 1$.*

3.2.4 Almost sure convergence of weighted discrete quadratic variation

Proposition 3.2.3. *Assume $(\mathbf{A}_\sigma)$ and let $\mathcal{T}$ be a sequence of strategies satisfying $(\mathbf{A}_S)$. Let $(H_t)_{0\leq t<T}$ be a continuous adapted $d\times d$-matrix process such that* $\sup_{t\in[0,T)}|H_t| < +\infty$ a.s. *and let $(M_t)_{0\leq t\leq T}$ be a $\mathbb{R}^d$-valued continuous local martingale such that $\langle M\rangle_t = \int_0^t \alpha_r \mathrm{d}r$ with* $\sup_{0\leq t\leq T}|\alpha_t| < +\infty$ a.s. . *Then*

$$\sum_{\tau_{i-1}^n<T} \Delta M_{\tau_i^n}^* H_{\tau_{i-1}^n} \Delta M_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}\left(H_t \mathrm{d}\langle M\rangle_t\right).$$

Proof. From Itô's lemma, $\sum_{\tau_{i-1}^n<T}\Delta M^*_{\tau_i^n} H_{\tau_{i-1}^n}\Delta M_{\tau_i^n}$ is equal to

$$\sum_{k,l=1}^{d}\sum_{\tau_{i-1}^n<T}\Delta M^k_{\tau_i^n} H^{k,l}_{\tau_{i-1}^n}\Delta M^l_{\tau_i^n} = \sum_{k,l=1}^{d}\int_0^T H^{k,l}_{\varphi(t)}(\Delta M^k_t \mathrm{d}M^l_t + \Delta M^l_t \mathrm{d}M^k_t + \mathrm{d}\langle M^k, M^l\rangle_t)$$
$$= \int_0^T \Delta M^*_t (H_{\varphi(t)} + H^*_{\varphi(t)})\mathrm{d}M_t + \int_0^T \mathrm{Tr}(H_{\varphi(t)}\mathrm{d}\langle M\rangle_t).$$

The second term in the above r.h.s. converges *a.s.* to $\int_0^T \mathrm{Tr}(H_t \mathrm{d}\langle M\rangle_t)$: indeed, the difference is bounded by $C_0\int_0^T |H_t - H_{\varphi(t)}|\mathrm{d}t$ and we conclude by an application of the dominated convergence theorem, invoking the continuity and boundedness of H and the convergence to 0 of the mesh size of $\mathcal{T}^n$ (see Corollary 3.2.2).
Thus, it remains to show that the stochastic integral w.r.t. $\mathrm{d}M_t$ converges *a.s.* to 0. Owing to Corollary 3.2.1, it is enough to study the series of quadratic variations, i.e. to show that $\sum_{n\geq 0}\left[\int_0^T \left(\Delta M^*_t (H_{\varphi(t)}+H^*_{\varphi(t)})\mathrm{d}\langle M\rangle_t (H_{\varphi(t)}+H^*_{\varphi(t)})\Delta M_t\right)\right]^3 \overset{a.s.}{\longrightarrow}$, and since α and H are *a.s.* bounded on $[0,T)$, it is sufficient to show

$$\sum_{n\geq 0}\left[\int_0^T |\Delta M_t|^2 \mathrm{d}t\right]^3 \overset{a.s.}{\longrightarrow} . \tag{3.2.7}$$

Clearly $\left[\int_0^T |\Delta M_t|^2\mathrm{d}t\right]^3$ is bounded by $d^3T^3 \sup_{1\leq j\leq d}\sup_{1\leq i\leq N_T^n}\sup_{\tau_{i-1}^n\leq t\leq \tau_i^n}|\Delta M^j_t|^6 \leq C_0\varepsilon_n^2$ owing to Corollary 3.2.3 (item *i*)) for $\theta = 0$ and $\rho = \frac{1}{6}$. The convergence (3.2.7) is proved and we are done. □

3.2.5 Verification of the hypothesis on a special family of hitting times

One of the more appealing result of the paper is that a very large family of hitting times fulfills the assumptions $(\mathbf{A}_N)$ and $(\mathbf{A}_S)$ with a threshold depending of ε_n.

Proposition 3.2.4. *Assume* $(\mathbf{A}_\sigma)$*. Let* $(H_t)_{0\leq t<T}$ *be a continuous adapted nonnegative-definite* $d\times d$*-matrix process, such that* a.s.

$$0 < \inf_{0\leq t<T}\lambda_{\min}(H_t) \leq \sup_{0\leq t<T}\lambda_{\max}(H_t) < +\infty.$$

The strategy $\mathcal{T}^n$ *given by*

$$\begin{cases}\tau_0^n := 0,\\ \tau_i^n := \inf\left\{t\geq \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^* H_{\tau_{i-1}^n}(S_t - S_{\tau_{i-1}^n}) > \varepsilon_n^2\right\}\wedge T,\end{cases}$$

defines a sequence of strategies satisfying assumptions $(\mathbf{A}_N)$ *(with* $\sup_{n\geq 0}(\varepsilon_n^2 N_T^n) < +\infty$ a.s. *) and* $(\mathbf{A}_S)$*, that is* $\{\mathcal{T}^n : n\geq 0\}\in\mathcal{T}^{\mathrm{adm.}}$.

The proof is postponed in Appendix 3.5.1. Observe that the above sequence of strategies is admissible even in the most constrained case $\rho_N = 1$. As we shall see later on, the optimal stopping times are given by the hitting times by the process S of an ellipsoid (corresponding to the case H symmetric).

3.3 Main results

3.3.1 Statements

We now go back to the hedging issue: at time $s \in [0, T]$, the fair value of the option is $u(s, S_s)$ and the hedging portfolio with discrete rebalancing dates $\mathcal{T}^n$ is $u(0, S_0) + \sum_{\tau_{i-1}^n \leq s} D_x u(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n})$, which yields an hedging error equal to

$$\begin{aligned} Z_s^n :&= u(s, S_s) - \Big(u(0, S_0) + \sum_{\tau_{i-1}^n \leq s} D_x u(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \cdot (S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n})\Big) \\ &= \int_0^s \big(D_x u_t - D_x u_{\varphi(t)}\big) \cdot \mathrm{d}S_t \qquad (3.3.1) \end{aligned}$$

using (3.1.5), where the integrand appears as the difference of Delta between τ_{i-1}^n and $t \in]\tau_{i-1}^n, \tau_i^n]$ for each $0 \leq i \leq N_T^n$.

One main result of the paper is a lower bound of the renormalized quadratic variation of the hedging error Z^n: it is partly derived from a *smart* representation of

$$\langle Z^n \rangle_T = \int_0^T \big(D_x u_t - D_x u_{\varphi(t)}\big)^* \mathrm{d}\langle S \rangle_t \big(D_x u_t - D_x u_{\varphi(t)}\big) \qquad (3.3.2)$$

as a sum of squared random variables and an application of the Cauchy-Schwarz inequality. To derive this suitable representation, we apply the Itô formula and identify the bounded variation term; it is straightforward in dimension one, much more intricate in a multidimensional setting, and this is equivalent to solve the following matrix equation.

Lemma 3.3.1. *Let $c \in \mathcal{S}^d(\mathbb{R})$. Then, the equation*

$$2\mathrm{Tr}(x)x + 4x^2 = c^2 \qquad (3.3.3)$$

admits exactly one solution $x(c) \in \mathcal{S}_+^d(\mathbb{R})$. In addition, $x(c)$ is positive-definite if and only if c^2 is positive-definite. Last, the mapping $c \mapsto x(c)$ is continuous.

The proof is given in Subsection 3.5.2. We are now in a position to give an explicit asymptotic lower bound for $N_T^n \langle Z^n \rangle_T$: this is the contents of the following theorem.

Theorem 3.3.1. *Assume the assumptions* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_u)$, $(\mathbf{A}_S)$, $(\mathbf{A}_N)$ *and* $(\mathbf{A}_Y)$ *are in force. Let* X *be the solution to* (3.3.3) *with* $c := \sigma^* D^2_{xx} u \sigma$. *Then,*

$$\liminf_{n\to+\infty} \ N_T^n \langle Z^n\rangle_T \geq \Big(\int_0^T \mathrm{Tr}\,(X_t)\,\mathrm{d}t\Big)^2, \quad \text{a.s..}$$

Let us comment a bit on the above lower bound.

- First, it is *a.s.* finite: indeed, $\sup_{t<T} |\sigma_t^* D^2_{xx} u_t \sigma_t| < +\infty$ *a.s.* and the continuity of $c \mapsto x(c)$ imply $\sup_{t<T} |X_t| < +\infty$ *a.s.* .

- Second, observe that *a.s.*

  $$\{\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t = 0\} = \{\forall t < T : \sigma_t^* D^2_{xx} u_t \sigma_t = 0\} \overset{\text{under } (\mathbf{A}_\sigma^{\text{Ellip.}})}{=} \{\forall t < T : D^2_{xx} u_t = 0\}$$

  using at the first equality that $\mathrm{Tr}(x(c)) > 0 \Leftrightarrow x(c) \neq 0 \Leftrightarrow c \neq 0$. Then we obtain that except in degenerate situations (where the Gamma matrix $D^2_{xx} u_t$ is zero at any time, assuming $(\mathbf{A}_\sigma^{\text{Ellip.}})$), the lower bound in Theorem 3.3.1 is non-zero.

- As a consequence , we immediately obtain a lower bound for the L_p-criterion: indeed, using the Fatou lemma and the Cauchy-Schwarz inequality, we derive (for any $p > 0$)

  $$\begin{aligned}\Big[\mathbb{E}(\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t)^p\Big]^2 &\leq \Big[\mathbb{E}\Big(\liminf_{n\to+\infty} \big(N_T^n\langle Z^n\rangle_T\big)^{p/2}\Big)\Big]^2 \leq \liminf_{n\to+\infty} \Big[\mathbb{E}\big(N_T^n\langle Z^n\rangle_T\big)^{p/2}\Big]^2 \\ &\leq \liminf_{n\to+\infty} \mathbb{E}((N_T^n)^p)\mathbb{E}(\langle Z^n\rangle_T^p).\end{aligned}$$

  For $p = 1$ we recover the Fukasawa approach [Fukasawa 2011a].

The next theorem tells us that along a suitable sequence $\mathcal{T}^n$ (the hitting times of some random ellipsoids) the lower bound of Theorem 3.3.1 is reached. Let $\chi(.)$ be a smooth function such that $\mathbf{1}_{]-\infty,1/2]} \leq \chi(.) \leq \mathbf{1}_{]-\infty,1]}$ and for $\mu > 0$, set $\chi_\mu(x) = \chi(x/\mu)$.

Theorem 3.3.2. *Assume the assumptions* $(\mathbf{A}_\sigma^{\text{Ellip.}})$, $(\mathbf{A}_u)$, $(\mathbf{A}_S)$, $(\mathbf{A}_N)$ *and* $(\mathbf{A}_Y)$ *are in force. Let* $\mu > 0$, *for* $t \geq 0$ *set* $\Lambda_t := (\sigma_t^{-1})^* X_t \sigma_t^{-1}$ *and* $\Lambda_t^\mu := \Lambda_t + \mu\chi_\mu(\lambda_{\min}(\Lambda_t))I_d$.
For a given $n \in \mathbb{N}$, *define the strategy* $\mathcal{T}_\mu^n$ *by*

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n = \inf\Big\{t \geq \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^* \Lambda^\mu_{\tau_{i-1}^n} (S_t - S_{\tau_{i-1}^n}) > \varepsilon_n^2\Big\} \wedge T. \end{cases} \tag{3.3.4}$$

Then, the sequence of strategies $\mathcal{T}_\mu = \{\mathcal{T}_\mu^n : n \geq 0\}$ *is admissible and it is* μ*-asymptotically optimal in the following sense:*

$$\limsup_{n\to+\infty} \Big| N_T^n \langle Z^n \rangle_T - (\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t)^2 \Big| \leq C_\mu \mu \int_0^T \chi_\mu(\lambda_{\min}(\Lambda_t))\mathrm{Tr}(\sigma_t\sigma_t^*)\mathrm{d}t$$

where the random variable $C_\mu := \int_0^T \big(4\mathrm{Tr}(X_t) + 3\mu\chi_\mu(\lambda_{\min}(\Lambda_t))\mathrm{Tr}(\sigma_t\sigma_t^*)\big)\mathrm{d}t$ *is* a.s. *finite (locally uniformly w.r.t.* $\mu \geq 0$*).*
In particular, on the event $\{\forall t \in [0,T] : \lambda_{\min}(\Lambda_t) \geq \mu\}$, $N_T^n\langle Z^n\rangle_T$ *converges* a.s. *to* $\big(\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t\big)^2$.

Observe that we require the ellipticity condition to hold. The proof is given in Subsection 3.3.3.
We can strengthen the above Theorem by allowing $\mu = 0$ under stronger assumptions.

Theorem 3.3.3. *Assume the assumptions of Theorem 3.3.2 and additionally that*

$$\mathbb{P}\big(\inf_{t\in[0,T[} \lambda_{\min}(D^2_{xx}u_t) > 0\big) = 1. \tag{3.3.5}$$

Then, the sequence of strategies $\mathcal{T}_0 = \{\mathcal{T}^n(0) : n \geq 0\}$ *defined in* (3.3.4) *with* $\mu = 0$ *is admissible and asymptotically optimal:*

$$\lim_{n\to+\infty} N_T^n\langle Z^n\rangle_T = (\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t)^2, \quad \text{a.s.}.$$

For the proof, see Subsection 3.3.4. The extra assumption (3.3.5) is satisfied in dimension one for call/put option in Black-Scholes model only if the hedging time horizon is strictly smaller than the option maturity. But it is not satisfied in digital call/put option. This discussion can be extented to higher multidimensional situations.

Remark 3.3.1. *In the one dimensional case, we have*

$$X_t = \frac{1}{\sqrt{6}}\sigma_t^2|D^2_{xx}u_t|, \quad \Lambda_t = \frac{1}{\sqrt{6}}|D^2_{xx}u_t|$$

and the μ*-optimal stopping times read*

$$\tau_i^n = \inf\left\{ t \geq \tau_{i-1}^n : |S_t - S_{\tau_{i-1}^n}| > \frac{\varepsilon_n}{\sqrt{|D^2_{xx}u_{\tau_{i-1}^n}|/\sqrt{6} + \mu\chi_\mu(|D^2_{xx}u_{\tau_{i-1}^n}|/\sqrt{6})}} \right\} \wedge T.$$

For $|D^2_{xx}u_t|$ *bounded from below, we can take* $\mu = 0$ *and the optimal strategy coincides with that of [Fukasawa 2011a, Theorem C].*
The threshold $\mu \neq 0$ *ensures that the hedging rebalancing occurs often enough, even if* $\Lambda_t \neq 0$ *for some time* t*: this interpretation is also valid in the multidimensional case.*

3.3.2 Proof of Theorem 3.3.1

It is split into several steps.

Step 1: Quadratic variation decomposition

We start from the hedging error (3.3.1). A natural idea consists in writing a Taylor expansion (regarding the S variable only) and showing that the residual terms converge to 0 fast enough as we could expect :

$$Z_s^n = \int_0^s (D_{xx}^2 u_{\varphi(t)} \Delta S_t) \cdot \mathrm{d}S_t + R_s^n, \tag{3.3.6}$$

where

$$R_s^n := \int_0^s \big(D_x u_t - D_x u_{\varphi(t)} - D_{xx}^2 u_{\varphi(t)} \Delta S_t\big) \cdot \mathrm{d}S_t, \quad s \le T. \tag{3.3.7}$$

Then passing to quadratic variation, we obtain

$$\langle Z^n \rangle_T = \int_0^T \Delta S_t^* D_{xx}^2 u_{\varphi(t)} \mathrm{d}\langle S \rangle_t D_{xx}^2 u_{\varphi(t)} \Delta S_t + e_{1,T}^n$$

where

$$e_{1,T}^n := \langle R^n \rangle_T + 2\Big\langle \int_0^\cdot (D_{xx}^2 u_{\varphi(t)} \Delta S_t) \cdot \mathrm{d}S_t, R_\cdot^n \Big\rangle_T. \tag{3.3.8}$$

Now, we wish an expression involving only the Brownian motion for ease of mathematical analysis: hence we replace ΔS_t by $\sigma_{\varphi(t)} \Delta B_t$ and $\mathrm{d}\langle S \rangle_t$ by $\sigma_{\varphi(t)} \sigma_{\varphi(t)}^* \mathrm{d}t$, leading to

$$\begin{aligned}
\langle Z^n \rangle_T &= \int_0^T \Delta B_t^* \big(\sigma_{\varphi(t)}^* D_{xx}^2 u_{\varphi(t)} \sigma_{\varphi(t)}\big)^2 \Delta B_t \mathrm{d}t + e_{1,T}^n + e_{2,T}^n, \\
e_{2,T}^n &:= \int_0^T \Delta S_t^* D_{xx}^2 u_{\varphi(t)} \Delta(\sigma_t \sigma_t^*) D_{xx}^2 u_{\varphi(t)} \Delta S_t \mathrm{d}t \\
&\quad + \int_0^T \big(\Delta S_t + \sigma_{\varphi(t)} \Delta B_t\big)^* D_{xx}^2 u_{\varphi(t)} \sigma_{\varphi(t)} \sigma_{\varphi(t)}^* D_{xx}^2 u_{\varphi(t)} \Big(\int_{\varphi(t)}^t \Delta \sigma_r \mathrm{d}B_r \Big) \mathrm{d}t.
\end{aligned} \tag{3.3.9}$$

As mentioned before, we seek a *smart* representation of the main term of $\langle Z^n \rangle_T$ in the form $\sum_{\tau_{i-1}^n < T} \big(\Delta B_{\tau_i^n}^* X_{\tau_{i-1}^n} \Delta B_{\tau_i^n}\big)^2$ plus a stochastic integral, where X is a measurable adapted $d \times d$-matrix process which has to be defined. Instead of directly giving the solution, let us discuss a bit on the expected properties of X. Applying Itô's formula on each interval $[\tau_{i-1}^n, \tau_i^n]$, we obtain

$$\begin{aligned}
\sum_{\tau_{i-1}^n < T} \big(\Delta B_{\tau_i^n}^* X_{\tau_{i-1}^n} \Delta B_{\tau_i^n}\big)^2 &= \int_0^T \Delta B_t^* \Big(2\mathrm{Tr}\big(X_{\varphi(t)}\big) X_{\varphi(t)} + \big(X_{\varphi(t)} + X_{\varphi(t)}^*\big)^2 \Big) \Delta B_t \mathrm{d}t \\
&\quad + 2 \int_0^T \Delta B_t^* X_{\varphi(t)} \Delta B_t \Delta B_t^* \big(X_{\varphi(t)} + X_{\varphi(t)}^*\big) \mathrm{d}B_t,
\end{aligned}$$

with the tentative identification

$$2\mathrm{Tr}\big(X_{\varphi(t)}\big)X_{\varphi(t)} + \big(X_{\varphi(t)} + X^*_{\varphi(t)}\big)^2 = \big(\sigma^*_{\varphi(t)} D^2_{xx} u_{\varphi(t)} \sigma_{\varphi(t)}\big)^2. \tag{3.3.10}$$

Mainly, two reasons prompt us to impose $X_{\varphi(t)} \in \mathcal{S}^d_+(\mathbb{R})$.

- Gathering the previous identities and anticipating a little bit on the following, the main contribution in $N^n_T \langle Z^n \rangle_T$ is

$$N^n_T \sum_{\tau^n_{i-1}<T} \big(\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big)^2 \geq \Bigg(\sum_{\tau^n_{i-1}<T} \big|\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big|\Bigg)^2$$

  using the Cauchy-Schwarz inequality. In general the limit of the above lower bound is not easy to handle because of the absolute values, but if the matrix $X_{\varphi(t)}$ is nonnegative-definite, we can remove them and conclude using a convergence result about discrete quadratic variations (Proposition 3.2.3).

- Once that we have restricted to nonnegative-definite matrices, let us prove that the solution to (3.3.10) (whenever it exists) is symmetric. If $\mathrm{Tr}(X_{\varphi(t)}) = 0$ then $X_{\varphi(t)} = 0$ (thus symmetric): indeed, $X_{\varphi(t)} + X^*_{\varphi(t)}$ is symmetric nonnegative-definite and has a null trace, thus it is the zero-matrix and consequently $X_{\varphi(t)} = -X^*_{\varphi(t)} = 0$ (since both $X_{\varphi(t)}$ and $X^*_{\varphi(t)}$ are nonnegative-definite). If $\mathrm{Tr}(X_{\varphi(t)}) > 0$ then taking the transposition of (3.3.10) readily gives $X_{\varphi(t)} = X^*_{\varphi(t)}$.

From Lemma 3.3.1, there exists exactly one adapted process X with values in $\mathcal{S}^d_+(\mathbb{R})$, solution to the equation $2\mathrm{Tr}(X)X + 4X^2 = (\sigma^* D^2_{xx} u \sigma)^2$. In addition, this solution is continuous *a.s.* because $C := \sigma^* D^2_{xx} u \sigma$ is continuous *a.s.* and the solution X is continuous as a function of C on $\mathcal{S}^d$. Gathering the previous identities, we have established a nice decomposition of the quadratic variation of the hedging error

$$\langle Z^n \rangle_T = \sum_{\tau^n_{i-1}<T} \Big(\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\Big)^2 + e^n_{1,T} + e^n_{2,T} + e^n_{3,T}, \tag{3.3.11}$$

$$e^n_{3,T} := -\,4 \int_0^T \Delta B^*_t X_{\varphi(t)} \Delta B_t \Delta B^*_t X_{\varphi(t)} \mathrm{d}B_t. \tag{3.3.12}$$

Step 2: lower bound for the renormalized quadratic variation

The Cauchy-Schwarz inequality yields that $N^n_T \sum_{\tau^n_{i-1}<T} \big(\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big)^2$ is bounded from below by

$$\Bigg(\sum_{\tau^n_{i-1}<T} \big|\Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\big|\Bigg)^2 = \Bigg(\sum_{\tau^n_{i-1}<T} \Delta B^*_{\tau^n_i} X_{\tau^n_{i-1}} \Delta B_{\tau^n_i}\Bigg)^2 \xrightarrow{a.s.} \Bigg(\int_0^T \mathrm{Tr}\,(X_t)\,\mathrm{d}t\Bigg)^2,$$

using that X is a nonnegative-definite matrix process and applying Proposition 3.2.3.

Step 3: the renormalized errors $\varepsilon_n^{-2\rho_N} e^n_{1,T}$, $\varepsilon_n^{-2\rho_N} e^n_{2,T}$ and $\varepsilon_n^{-2\rho_N} e^n_{3,T}$ converge to 0 *a.s.*

Observe that once these convergences are established, in view of (3.3.11) and $(\mathbf{A}_N)$ we easily complete the proof of Theorem 3.3.1.

• *Proof of* $\varepsilon_n^{-2\rho_N} e^n_{1,T} \xrightarrow{a.s.} 0$. We first state an intermediate result which is proved in Appendix (Subsection 3.5.3).

Lemma 3.3.2. *Assume the hypotheses* $(\mathbf{A}_\sigma)$, $(\mathbf{A}_u)$, $(\mathbf{A}_S)$, $(\mathbf{A}_N)$ *and* $(\mathbf{A}_Y)$ *are in force. Then* $\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T \xrightarrow{a.s.} 0$ *where* R^n *is defined in* (3.3.7).

Then, starting from (3.3.8), applying the Cauchy-Schwarz inequality to the cross-variation and using $(\mathbf{A}_\sigma)$-$(\mathbf{A}_u)$-$(\mathbf{A}_S)$, we derive

$$\begin{aligned}\varepsilon_n^{-2\rho_N}|e^n_{1,T}| &\leq \varepsilon_n^{-2\rho_N}\langle R^n\rangle_T + 2\big(\varepsilon_n^{-2}\int_0^T \Delta S_t^* D^2_{xx}u_{\varphi(t)}\mathrm{d}\langle S\rangle_t D^2_{xx}u_{\varphi(t)}\Delta S_t\big)^{1/2}\,\big(\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T\big)^{1/2}\\ &\leq \varepsilon_n^{2(\rho_N-1)}\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T + 2C_0\big(\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T\big)^{1/2} \xrightarrow{a.s.} 0.\end{aligned}$$

□

• *Proof of* $\varepsilon_n^{-2\rho_N} e^n_{2,T} \xrightarrow{a.s.} 0$. We analyze separately the two contributions in (3.3.9).

1. First, simple computations using $(\mathbf{A}_\sigma)$-$(\mathbf{A}_u)$-$(\mathbf{A}_S)$ and Corollary 3.2.2 directly give (for any given $\rho > 0$)

   $$\varepsilon_n^{-2\rho_N}\Big|\int_0^T \Delta S_t^* D^2_{xx}u_{\varphi(t)}\Delta(\sigma_t\sigma_t^*)D^2_{xx}u_{\varphi(t)}\Delta S_t\mathrm{d}t\Big| \leq C_0\varepsilon_n^{-2\rho_N+2}(\varepsilon_n^{\theta_\sigma} + \varepsilon_n^{\frac{\theta_\sigma}{2}(2-\rho)}).$$

   Since $\rho_N < 1+\theta_\sigma/2$ and ρ can be taken arbitrary small, we obtain that the above upper bound converges *a.s.* to 0.

2. Second, we apply twice Corollary 3.2.3-*ii)*, first taking $\theta = 0$ and second taking $\theta = \theta_\sigma$, so that we obtain,for any given $\rho > 0$, *a.s.* for any $n \geq 0$

   $$\sup_{1\leq i\leq N_T^n}\ \sup_{\tau_{i-1}^n\leq t\leq\tau_i^n}|\Delta S_t + \sigma_{\varphi(t)}\Delta B_t| \leq C_0\varepsilon_n^{1-\rho}, \tag{3.3.13}$$

   $$\sup_{1\leq i\leq N_T^n}\ \sup_{\tau_{i-1}^n\leq t\leq\tau_i^n}\Big|\int_{\varphi(t)}^t \Delta\sigma_r\mathrm{d}B_r\Big| \leq C_0\varepsilon_n^{1+\theta_\sigma-\rho}, \tag{3.3.14}$$

$$\varepsilon_n^{-2\rho_N}\Big|\int_0^T(\Delta S_t + \sigma_{\varphi(t)}\Delta B_t)^* D^2_{xx}u_{\varphi(t)}\sigma_{\varphi(t)}\sigma^*_{\varphi(t)}D^2_{xx}u_{\varphi(t)}\Big(\int_{\varphi(t)}^t \Delta\sigma_r\mathrm{d}B_r\Big)\mathrm{d}t\Big| \leq C_0\varepsilon_n^{2+\theta_\sigma-2\rho_N-2\rho}.$$

Owing to $\rho_N < 1+\theta_\sigma/2$, taking ρ small enough implies the *a.s.* convergence of the latter upper bound to 0. As a result, $\varepsilon_n^{-2\rho_N} e^n_{2,T} \xrightarrow{a.s.} 0$. □

• *Proof of* $\varepsilon_n^{-2\rho_N} e^n_{3,T} \xrightarrow{a.s.} 0$. It is a direct consequence of the following lemma. □

Lemma 3.3.3. *Assume* $(\mathbf{A}_\sigma)$. *Let* $\mathcal{T} = (\mathcal{T}^n)_{n\geq 0}$ *be an admissible sequence of strategies and let* $(H_t)_{0\leq t<T}$ *be a continuous adapted* $d \times d$*-matrix process such that* $\sup_{t\in[0,T)} |H_t| < +\infty$ a.s. . *Then for any* $p > \frac{2}{3-2\rho_N}$*, the series* $\sum_{n\geq 0} |\varepsilon_n^{-2\rho_N} \int_0^T \Delta B_t^* H_{\varphi(t)} \Delta B_t \Delta B_t^* H_{\varphi(t)} \mathrm{d}B_t|^p$ *converges almost surely.*

Proof. Set $\alpha_t^n := \Delta B_t^* H_{\varphi(t)} \Delta B_t \Delta B_t^* H_{\varphi(t)}$ and define the scalar continuous local martingale $M_t^n := \varepsilon_n^{-2\rho_N} \int_0^t \alpha_s^n \mathrm{d}B_s$. In view of Corollary 3.2.1, it is enough to check that $(\langle M^n\rangle_T^{p/2})_{n\geq 0}$ defines the terms of an *a.s.* convergent series. An application of Corollary 3.2.3-*ii)* with $\rho = \frac{(3-2\rho_N)p-2}{3p} > 0$ and $\theta = 0$ gives $\sup_{1\leq i\leq N_T^n} \sup_{\tau_{i-1}^n\leq t\leq \tau_i^n} |\Delta B_t| < C_0 \varepsilon_n^{1-\rho}$ and therefore

$$\langle M^n\rangle_T^{p/2} = \varepsilon_n^{-2p\rho_N} \big(\int_0^T |\alpha_t^n|^2 \mathrm{d}t\big)^{p/2} \leq C_0 \varepsilon_n^{-2p\rho_N} \sup_{1\leq i\leq N_T^n} \sup_{\tau_{i-1}^n\leq t\leq \tau_i^n} |\Delta B_t|^{3p} \leq C_0 \varepsilon_n^2 \quad a.s..$$

We are finished. □

3.3.3 Proof of Theorem 3.3.2

We first check the admissibility of $\mathcal{T}_\mu$, by applying Proposition 3.2.4. Indeed, owing to $(\mathbf{A}_u)$ and $(\mathbf{A}_\sigma^{\text{Ellip.}})$, $(\Lambda_t)_{0\leq t<T}$ is a continuous adapted nonnegative-definite $d \times d$-matrix process with $\sup_{0\leq t<T} |\Lambda_t| < +\infty$ *a.s.* . The same properties clearly hold for $(\Lambda_t^\mu)_{0\leq t<T}$. In addition, $\lambda_{\min}(\Lambda_t^\mu) \geq \mu/2 > 0$ and $\sup_{0\leq t<T} \lambda_{\max}(\Lambda_t^\mu) \leq \mu + \sup_{0\leq t<T} \lambda_{\max}(\Lambda_t) < +\infty$ *a.s.* . Therefore, $\mathcal{T}_\mu$ is admissible and in addition $\sup_{n\geq 0} \varepsilon_n^2 N_T^n < +\infty$ *a.s.* . Hence, it allows to re-use the computations of the proof of Theorem 3.3.1 in the case $\rho_N = 1$.

Now let us show the μ-optimality. Writing $N_T^n = 1 + \sum_{1\leq i\leq N_T^n-1} 1$, we point out

$$\begin{aligned} \varepsilon_n^2 N_T^n &= \varepsilon_n^2 + \sum_{1\leq i\leq N_T^n-1} \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n}^\mu \Delta S_{\tau_i^n} \\ &= \varepsilon_n^2 - \Delta S_T^* \Lambda_{\tau_{N_T^n-1}^n}^\mu \Delta S_T + \sum_{\tau_{i-1}^n<T} \Delta S_{\tau_i^n}^* \Lambda_{\tau_{i-1}^n}^\mu \Delta S_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}\big(\Lambda_t^\mu \sigma_t \sigma_t^*\big) \mathrm{d}t \end{aligned} \tag{3.3.15}$$

using the convergence of Proposition 3.2.3. On the other hand, starting from the decomposition (3.3.11) of the hedging error quadratic variation, we write

$$\langle Z^n\rangle_T = \sum_{1\le i\le N_T^n-1} \big(\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big)^2 + e^n_{1,T}+e^n_{2,T}+e^n_{3,T}+e^n_{4,T}+e^n_{5,T}+e^n_{6,T}, \tag{3.3.16}$$

$$\begin{aligned}
e^n_{4,T} &:= \sum_{\tau_{i-1}^n<T} \big(\Delta B^*_{\tau_i^n}X_{\tau_{i-1}^n}\Delta B_{\tau_i^n}\big)^2 - \big(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big)^2,\\
e^n_{5,T} &:= \sum_{\tau_{i-1}^n<T} \big(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big)^2 - \big(\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big)^2,\\
e^n_{6,T} &:= \big(\Delta S^*_T\Lambda^\mu_{\tau^n_{N_T^n-1}}\Delta S_T\big)^2.
\end{aligned}$$

In view of the definition of the strategy $\mathcal{T}^n_\mu$, (3.3.16) becomes

$$\varepsilon_n^{-2}\langle Z^n\rangle_T = \sum_{1\le i\le N_T^n-1} \Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n} + \varepsilon_n^{-2}\sum_{j=1}^{6} e^n_{j,T}. \tag{3.3.17}$$

Similarly to (3.3.15), we show that $\sum_{1\le i\le N_T^n-1}\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}\big(\Lambda^\mu_t\sigma_t\sigma_t^*\big)\mathrm{d}t$. Furthermore we have already established (see Step 3 of proof of Theorem 3.3.1) that $\varepsilon_n^{-2}e^n_{j,T}\xrightarrow{a.s.} 0$ for $j=1,2,3$ (remind that we can take $\rho_N=1$); the case $j=6$ is also fulfilled because $0\le e^n_{6,T}\le\varepsilon_n^4$.
To analyze $e^n_{4,T}$, set $D_{B,i} := \sigma_{\tau_{i-1}^n}\Delta B_{\tau_i^n}$ and $D_{S,i} := \Delta S_{\tau_i^n}$, write $X_{\tau_{i-1}^n} = \sigma^*_{\tau_{i-1}^n}\Lambda_{\tau_{i-1}^n}\sigma_{\tau_{i-1}^n}$ and

$$\begin{aligned}
\big(\Delta B^*_{\tau_i^n}X_{\tau_{i-1}^n}\Delta B_{\tau_i^n}\big)^2 - \big(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big)^2 &= \big(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i}\big)^2 - \big(D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\big)^2\\
&= \big(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i} - D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\big)\big(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i} + D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\big)\\
&= \big(D_{B,i}+D_{S,i}\big)^*\Lambda_{\tau_{i-1}^n}\big(D_{B,i}-D_{S,i}\big)\big(D^*_{B,i}\Lambda_{\tau_{i-1}^n}D_{B,i} + D^*_{S,i}\Lambda_{\tau_{i-1}^n}D_{S,i}\big).
\end{aligned}$$

Then, we deduce that $\varepsilon_n^{-2}|e^n_{4,T}|$ is bounded by

$$\begin{aligned}
&\varepsilon_n^{-2}N_T^n \sup_{1\le i\le N_T^n}\sup_{\tau_{i-1}^n\le t\le\tau_i^n]} |\Lambda_{\varphi(t)}|^2|\Delta S_t+\sigma_{\varphi(t)}\Delta B_t|\Big|\int_{\varphi(t)}^t \Delta\sigma_s dB_s\Big|\big(|\Delta S_t|^2+|\sigma_{\varphi(t)}\Delta B_t|^2\big)\\
&\le C_0\varepsilon_n^{-2}\varepsilon_n^{-2}\varepsilon_n^{1-\rho}\varepsilon_n^{(1+\theta_\sigma-\rho)}\varepsilon_n^{2(1-\rho)} = C_0\varepsilon_n^{\theta_\sigma/5}\xrightarrow{a.s.} 0
\end{aligned}$$

where we have used $(\mathbf{A}_N)$ (with $\rho_N=1$) and the estimates (3.3.13-3.3.14) with $\rho=\theta_\sigma/5$ (which are available for any sequence of admissible strategies). This proves $\varepsilon_n^{-2}e^n_{4,T}\xrightarrow{a.s.} 0$.
Finally regarding $e^n_{5,T}$, recalling that the matrix $\Lambda_{\tau_{i-1}^n}$ is nonnegative-definite, we

obtain that $|\varepsilon_n^{-2}e_{5,T}^n|$ is bounded by

$$\begin{aligned}
&\varepsilon_n^{-2}\sum_{\tau_{i-1}^n<T}\big|\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}-\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big|\big(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}+\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big),\\
&\le\sum_{\tau_{i-1}^n<T}\mu\chi_\mu(\lambda_{\min}(\Lambda_{\tau_{i-1}^n}))|\Delta S_{\tau_i^n}|^2\big[2\varepsilon_n^{-2}\Delta S^*_{\tau_i^n}\Lambda^\mu_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big]\\
&\le 2\mu\sum_{\tau_{i-1}^n<T}\chi_\mu(\lambda_{\min}(\Lambda_{\tau_{i-1}^n}))|\Delta S_{\tau_i^n}|^2
\end{aligned}$$

where we have used the definition of $\mathcal{T}_\mu$ at the last inequality. Thus Proposition 3.2.3 yields

$$\limsup_{n\to+\infty}|\varepsilon_n^{-2}e_{5,T}^n|\le 2\mu\int_0^T\chi_\mu(\lambda_{\min}(\Lambda_t))\mathrm{Tr}(\sigma_t\sigma_t^*)\mathrm{d}t,\quad a.s..$$

Let us summarize: setting $L_T:=\int_0^T\mathrm{Tr}\big(\Lambda_t\sigma_t\sigma_t^*\big)\mathrm{d}t=\int_0^T\mathrm{Tr}\big(X_t\big)\mathrm{d}t$ and $L_T^\mu:=\int_0^T\chi_\mu(\lambda_{\min}(\Lambda_t))\mathrm{Tr}(\sigma_t\sigma_t^*)\mathrm{d}t$ so that $\int_0^T\mathrm{Tr}\big(\Lambda_t^\mu\sigma_t\sigma_t^*\big)\mathrm{d}t=L_T+\mu L_T^\mu$, we have shown

$$\begin{aligned}
&\varepsilon_n^2N_T^n\xrightarrow{a.s.}L_T+\mu L_T^\mu,\qquad\limsup_{n\to+\infty}\big|\varepsilon_n^{-2}\langle Z^n\rangle_T-(L_T+\mu L_T^\mu)\big|\le 2\mu L_T^\mu,\qquad a.s.,\\
&\limsup_{n\to+\infty}|N_T^n\langle Z^n\rangle_T-(L_T)^2|\\
&\qquad\le\limsup_{n\to+\infty}|\varepsilon_n^{-2}\langle Z^n\rangle_T-L_T|\limsup_{n\to+\infty}\varepsilon_n^2N_T^n+L_T\limsup_{n\to+\infty}|\varepsilon_n^2N_T^n-L_T|\\
&\qquad\le 3\mu L_T^\mu(L_T+\mu L_T^\mu)+L_T\mu L_T^\mu=\mu L_T^\mu(4L_T+3\mu L_T^\mu),\quad a.s..
\end{aligned}$$

Theorem 3.3.2 is proved. □

3.3.4 Proof of Theorem 3.3.3

Here, arguments are simpler in all steps of the proof of Subsection 3.3.3, then we shall skip details; the admissibility of the strategy comes readily from the ad hoc assumption (3.3.5) and Proposition 3.2.4; the optimality follows as before from

$$\varepsilon_n^2N_T^n=\varepsilon_n^2+\sum_{1\le i\le N_T^n-1}\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\xrightarrow{a.s.}\int_0^T\mathrm{Tr}\big(X_t\big)\mathrm{d}t,$$

and from (setting $\bar e_{6,T}^n:=(\Delta S_T^*\Lambda_{\tau^n_{N_T^n-1}}\Delta S_T)^2$)

$$\varepsilon_n^{-2}\langle Z^n\rangle_T=\varepsilon_n^{-2}\sum_{1\le i\le N_T^n-1}\big(\Delta S^*_{\tau_i^n}\Lambda_{\tau_{i-1}^n}\Delta S_{\tau_i^n}\big)^2+\varepsilon_n^{-2}\sum_{j=1}^4e_{j,T}^n+\varepsilon_n^{-2}\bar e_{6,T}^n\xrightarrow{a.s.}\int_0^T\mathrm{Tr}\big(X_t\big)\mathrm{d}t$$

with the help of the convergence results already obtained. Theorem 3.3.3 is proved. □

3.4 Numerical experiments

3.4.1 Algorithm for the optimal stopping times

From the previous section (Theorem 3.3.2), the μ-optimal stopping times ($\mu > 0$) are iteratively given by $\tau_0^n := 0$ and

$$\tau_i^n := \inf\left\{t \geq \tau_{i-1}^n : (S_t - S_{\tau_{i-1}^n})^* \Lambda^\mu_{\tau_{i-1}^n}(S_t - S_{\tau_{i-1}^n}) \geq \varepsilon_n^2\right\} \wedge T$$

where for any t, $\Lambda_t^\mu := \Lambda_t + \mu\chi_\mu(\lambda_{\min}(\Lambda_t))I_d$, $\Lambda_t := (\sigma_t^{-1})^* X_t \sigma_t^{-1}$ and X_t solves (3.3.3) with $c_t = \sigma_t^* D^2_{xx} u_t \sigma_t$. Thus, τ_i^n is the first hitting time of an ellipsoid centered at $S_{\tau_{i-1}^n}$ with principal axes equal to the orthogonal eigenvectors of the symmetric positive-definite matrix $\Lambda^\mu_{\tau_{i-1}^n}$ (or equivalently those of $\Lambda_{\tau_{i-1}^n}$). We briefly recall (see Subsection 3.5.2) the main steps to compute the matrix $X_{\tau_{i-1}^n}$ ($i \geq 1$) from which we derive $\Lambda_{\tau_{i-1}^n}$ and $\Lambda^\mu_{\tau_{i-1}^n}$.

1. Diagonalize the symmetric matrix $c_{\tau_{i-1}^n} = \sigma^*_{\tau_{i-1}^n} D^2_{xx} u_{\tau_{i-1}^n} \sigma_{\tau_{i-1}^n} := P_{\tau_{i-1}^n} \mathrm{Diag}\big(\lambda_j(c_{\tau_{i-1}^n}) : 1 \leq j \leq d\big) P^*_{\tau_{i-1}^n}$, where $P_{\tau_{i-1}^n}$ is an orthogonal matrix.

2. Find the zero $y_{\tau_{i-1}^n} \in \mathbb{R}^+$ of the increasing function $y \mapsto (4+d)y - \sum_{j=1}^d \sqrt{y^2 + 4\lambda_j^2(c_{\tau_{i-1}^n})}$. This root lies in the interval $\big[0, d|\lambda(c_{\tau_{i-1}^n})|/\sqrt{4+2d}\big]$ (see the proof of Lemma 3.3.1).

3. From (3.5.4), we obtain

$$X_{\tau_{i-1}^n} = P_{\tau_{i-1}^n} \mathrm{Diag}\left(\frac{-y_{\tau_{i-1}^n} + \sqrt{y^2_{\tau_{i-1}^n} + 4\lambda_j^2(c_{\tau_{i-1}^n})}}{4} : 1 \leq j \leq d\right) P^*_{\tau_{i-1}^n}.$$

Last, we mention that even if $\Lambda^\mu_{\tau_{i-1}^n}$ is tractable, the exact simulation of τ_i^n is in general impossible and approximations are required (see [Gobet 2010b] and references therein).

3.4.2 Numerical tests

This section is dedicated to an application of Theorem 3.3.2 to the case of an exchange binary option $g(S_T) = \mathbf{1}_{S_T^1 \geq S_T^2}$. This example is relevant in our study (and improves the setting of [Fukasawa 2011a]) because this is a simple *bi-dimensional non-convex* function, for which the value function u and its sensitivities are available in the Black-Scholes model

$$\mathrm{d}\begin{pmatrix} S_t^1 \\ S_t^2 \end{pmatrix} = \begin{pmatrix} \sigma_1 S_t^1 & 0 \\ \rho\sigma_2 S_t^2 & \sqrt{1-\rho^2}\sigma_2 S_t^2 \end{pmatrix} \mathrm{d}\begin{pmatrix} B_t^1 \\ B_t^2 \end{pmatrix},$$

where (B^1, B^2) are two independent Brownian motions. The model parameters are set to $S_0^1 = 100$, $S_0^2 = 100$, $\sigma_1 = 0.3$, $\sigma_2 = 0.4$, $\rho = 0.5$ and $T = 1$.
We take $\varepsilon_n = 0.05$. In our different tests, we have not observed a significant difference by taking $\mu = 0$ or μ small; hence, we only report the values for $\mu = 0$. We generate 1000 experiments ω, independently. To compute the hitting times for each ω, we use a thin uniform time mesh $\pi_{\bar{n}} = (iT/\bar{n})_{0\le i\le \bar{n}}$ ($\bar{n} = 50000$ in our tests) : we draw $S^1(\omega)$ and $S^2(\omega)$ along $\pi_{\bar{n}}$ and compute (with the help of the previous algorithm) the hitting times $\tau_i^n(\omega) = \inf\big\{t \in \pi_{\bar{n}} \cap]\tau_{i-1}^n(\omega), T] : [(S_t - S_{\tau_{i-1}^n})^* \Lambda^{\mu}_{\tau_{i-1}^n}(S_t - S_{\tau_{i-1}^n})](\omega) \ge \varepsilon_n^2\big\} \wedge T$; at the end of the process, we get the number $N_T^n(\omega)$ of discrete times. The mesh $\pi_{\bar{n}}$ is also used to compute subsequent quadratic variations and time integrals.
We compare ω by ω the above strategy with that based on the *uniform* mesh $\pi_{N_T^n(\omega)}$ and with that based on the so-called *fractional* mesh[3] $\big(T\big[1 - (1 - i/N_T^n(\omega))^2\big]\big)_{1\le i\le N_T^n(\omega)}$: in that way, the comparison is done for the same number of times, which looks quite fair. We define $\beta_{\text{stochastic}}(\omega)$, $\beta_{\text{uniform}}(\omega)$, $\beta_{\text{fractional}}(\omega)$ where we compute $\beta_{\cdot}(\omega) := \frac{N_T^n \langle Z^n\rangle_T}{(\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t)^2}(\omega)$ according to each of these three strategies: in view of Theorem 3.3.2, this ratio is asymptotically greater than 1 and adimensional; moreover, the closer to 1 the ratio, the better the strategy.
Results. Figure 3.1 displays, for each ω, the couples

$$(\beta_{\text{stochastic}}(\omega), \beta_{\text{uniform}}(\omega)) \quad \text{and} \quad (\beta_{\text{stochastic}}(\omega), \beta_{\text{fractional}}(\omega)).$$

Most of the times, the points are above the diagonal, showing that the μ-optimal strategy lessens the quadratic variation ω-wise (remind that the strategies have got the same number of discrete times N_T^n), compared to the quadratic variation worked out over the deterministic time mesh. In addition, $\beta_{stochastic}$ is concentrated around 1, which means a convergence of $N_T^n\langle Z^n\rangle_T$ towards the lower bound $(\int_0^T \mathrm{Tr}(X_t)\mathrm{d}t)^2$.

Figure 3.2 displays $\langle Z^n\rangle_T$ as a function of N_T^n for the three strategies and for different ω: here again, we observe that the μ-optimal strategy outperforms deterministic strategies.

[3] According to [Geiss 2011], the fractional smoothness of $g(S_T)$ is $\frac{1}{2}$; thus, when $N_T^n(\omega)$ is deterministic, this choice of fractional mesh yields that $\mathbb{E}(\langle Z^n\rangle_T)$ is of order 1 w.r.t. the inverse of the number of times, instead of order $\frac{1}{2}$ with the uniform mesh.

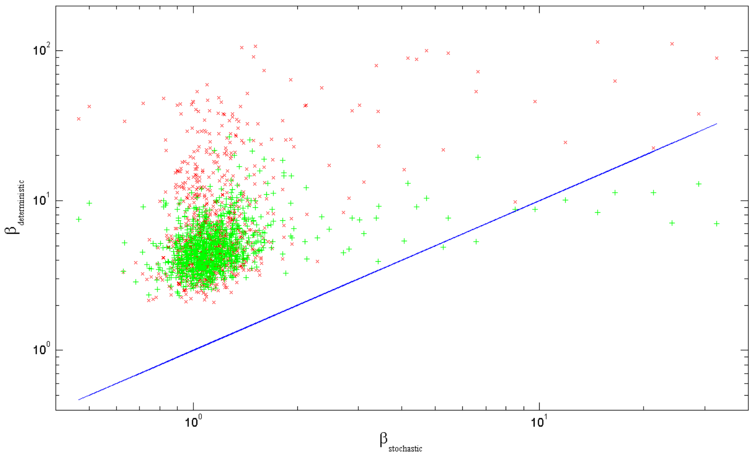

Figure 3.1: "$\times$", "+" and the blue line correspond respectively to "$(\beta_{\text{stochastic}}, \beta_{\text{uniform}})$", "$(\beta_{\text{stochastic}}, \beta_{\text{fractional}})$" and the identity function.

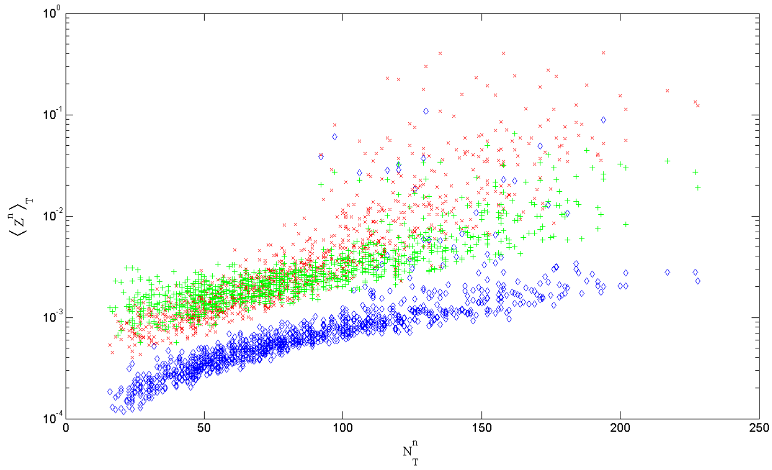

Figure 3.2: "$\times$", "+" and "$\diamond$" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$", "$\langle Z^n \rangle_{T,\text{fractional}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

3.5 Appendix

3.5.1 Proof of Proposition 3.2.4

It is standard to check that $\mathcal{T}^n$ is a sequence of increasing stopping times, we skip details. Let us justify that the size of $\mathcal{T}^n$ is *a.s.* finite, for any $n \geq 0$. For a given $n \geq 0$, define the event $\mathcal{N}^n := \{N^n_T = +\infty\}$. For $\omega \in \mathcal{N}^n$, the infinite sequence $(\tau^n_i(\omega))_{i \geq 0}$ converges, because increasing and bounded by T. Thus, on $\mathcal{N}^n \cap E_S$

with $E_S = \{(S_t)_{0\le t\le T}$ continuous and $\sup_{0\le t<T}\lambda_{\max}(H_t) < +\infty\}$, we have

$$0 < \varepsilon_n = (S_{\tau_i^n} - S_{\tau_{i-1}^n})^* H_{\tau_{i-1}^n}(S_{\tau_i^n} - S_{\tau_{i-1}^n}) \le \sup_{0\le t<T}\lambda_{\max}(H_t)|S_{\tau_i^n} - S_{\tau_{i-1}^n}|^2 \to_{i\to+\infty} 0,$$

which is impossible. Thus, $\mathcal{N}^n \subset E_S^c$ and $\mathbb{P}(\mathcal{N}^n) = 0$ since S is *a.s.* continuous and $\sup_{0\le t<T}\lambda_{\max}(H_t)$ is *a.s.* finite.
Besides, we have $C_H := \inf_{0\le t<T}\lambda_{\min}(H_t) > 0$ *a.s.* and we immediately get

$$\varepsilon_n^{-2}\sup_{1\le i\le N_T^n}\sup_{t\in(\tau_{i-1}^n,\tau_i^n]}|\Delta S_t|^2 \le C_H^{-1}\varepsilon_n^{-2}\sup_{1\le i\le N_T^n}\sup_{t\in(\tau_{i-1}^n,\tau_i^n]}(\Delta S_t^* H_{\tau_{i-1}^n}\Delta S_t) \le C_H^{-1}$$

which validates the assumption $(\mathbf{A}_S)$.

Then, writing $N_T^n = 1 + \sum_{1\le i\le N_T^n-1} 1$, we point out (for n large enough so that $\varepsilon_n \le 1$)

$$\varepsilon_n^{2\rho_N} N_T^n \le \varepsilon_n^2 N_T^n \le \varepsilon_n^2 + \sum_{1\le i\le N_T^n-1}\Delta S_{\tau_i^n}^* H_{\tau_{i-1}^n}\Delta S_{\tau_i^n} \le \varepsilon_n^2 + \sum_{\tau_{i-1}^n<T}\Delta S_{\tau_i^n}^* H_{\tau_{i-1}^n}\Delta S_{\tau_i^n},$$

using Moreover from Proposition 3.2.3, we know that under the assumption $(\mathbf{A}_S)$ only,

$$\sum_{\tau_{i-1}^n<T}\Delta S_{\tau_i^n}^* H_{\tau_{i-1}^n}\Delta S_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}(H_t \mathrm{d}\langle S\rangle_t) < +\infty.$$

This validates the assumption $(\mathbf{A}_N)$. □

Remark 3.5.1. *The structure of hitting times of ellipsoids with size ε_n has a specific feature compared to general admissible strategies: the assumption $(\mathbf{A}_S)$ entails the assumption $(\mathbf{A}_N)$.*

3.5.2 Proof of Lemma 3.3.1

We split the proof into several steps.
Let $h : \begin{cases} \mathbb{R}^d \times \mathbb{R}_+ & \to \mathbb{R} \\ (\lambda, y) & \mapsto (4+d)y - \sum_{i=1}^d \sqrt{y^2 + 4\lambda_i^2} \end{cases}$. Assume for a while that

($\star$) (a) for any $\lambda \in \mathbb{R}^d$, there exists a unique non-negative root y_λ satisfying $h(\lambda, y_\lambda) = 0$;

(b) $y_0 = 0$; $\lambda \ne 0 \Rightarrow y_\lambda > 0$;

(c) the mapping $\lambda \mapsto y_\lambda$ is continuous.

Necessary conditions on the spectrum of $x(c)$ Let $\mathcal{D}iag$ denote the set of $d \times d$ diagonal matrices. Take $c \in \mathcal{S}^d(\mathbb{R})$ and let $x(c) \in \mathcal{S}_+^d(\mathbb{R})$ be a solution (whenever it exists) to (3.3.3). Then by the spectral theorem, $x(c)$ is diagonalizable: there exists an orthogonal matrix $p_{x(c)}$ such that $p^*_{x(c)}x(c)p_{x(c)} \in \mathcal{D}iag$. Equation (3.3.3) is stable by unitary transformation:

$$2\mathrm{Tr}\Big(p^*_{x(c)}x(c)p_{x(c)}\Big)p^*_{x(c)}x(c)p_{x(c)}+4\Big(p^*_{x(c)}x(c)p_{x(c)}\Big)^2 = p^*_{x(c)}c^2p_{x(c)} \in \mathcal{D}iag. \quad (3.5.1)$$

The diagonal elements of $p^*_{x(c)}c^2p_{x(c)}$ must be the eigenvalues of c^2, that is the square of the eigenvalues of c (which is in $\mathcal{S}^d(\mathbb{R})$). Identifying the diagonal elements from (3.5.1) gives a relation between the spectra of c and $x(c)$:

$$2\mathrm{Tr}(x(c))\lambda_i(x(c)) + 4\lambda_i(x(c))^2 = \lambda_i(c)^2, \quad 1 \le i \le d.$$

Thus, the non-negative eigenvalues of $x(c)$ must satisfy $\lambda_i(x(c)) = (-\mathrm{Tr}(x(c)) + \sqrt{\mathrm{Tr}(x(c))^2 + 4\lambda_i(c)^2})/4$. By summing over $i = 1, \dots, d$, we obtain an implicit equation for $\mathrm{Tr}(x(c))$, which is $h(\lambda(c), \mathrm{Tr}(x(c))) = 0$. By $(\star)$, there is a unique solution and

$$\mathrm{Tr}(x(c)) = y_{\lambda(c)}. \quad (3.5.2)$$

Thus, we have proved that the eigenvalues of $x(c)$ must be

$$\lambda_i(x(c)) = \frac{-y_{\lambda(c)} + \sqrt{y^2_{\lambda(c)} + 4\lambda_i(c)^2}}{4}. \quad (3.5.3)$$

Existence/uniqueness of solution to (3.3.3) Take $c \in \mathcal{S}^d(\mathbb{R})$. Starting from (3.3.3), owing to (3.5.2) $x(c)$ must solve

$$\big(2x(c) + \frac{1}{2}y_{\lambda(c)}I_d\big)^2 = \frac{1}{4}y^2_{\lambda(c)}I_d + c^2.$$

The matrix $c^2 + \frac{1}{4}y^2_{\lambda(c)}I_d$ is symmetric nonnegative-definite, thus it has a unique square-root (symmetric nonnegative-definite matrix) [Horn 1990, Theorem 7.2.6 p.405] and we obtain

$$x(c) := -\frac{y_{\lambda(c)}}{4}I_d + \frac{1}{2}\Big(\frac{y^2_{\lambda(c)}}{4}I_d + c^2\Big)^{1/2}. \quad (3.5.4)$$

The uniqueness is proved. It is now easy to check that $x(c)$ given in (3.5.4) solves (3.3.3), using the implicit equation satisfied by $\mathrm{Tr}(x(c))$. Last, $\lambda_{\min}(c^2) > 0$ if and only if $\lambda_{\min}(x(c)) > 0$ (owing to (3.5.3)).

Continuity From Hoffman and Wielandt's theorem [Horn 1990, p.368], the function $c \mapsto \lambda(c)$ is continuous on $\mathcal{S}^d(\mathbb{R})$ into $\mathbb{R}^d$. Hence, combined with ($\star$ c), we obtain the continuity of $c \mapsto y_{\lambda(c)}$ on $\mathcal{S}^d(\mathbb{R})$ into $\mathbb{R}$.

Then, the continuity of $x(.)$ at $c_0 = 0$ easily follows since as $c \to 0$, $y_{\lambda(c)} \to y_0 = 0$ and $\lambda(x(c)) \to 0$ (using (3.5.3)): thus $x(c) \to 0 = x_0$. For $c_0 \neq 0$, we invoke the property that $c \mapsto c^{1/2}$ is locally lipschitz (and even analytic) on $\mathcal{S}^d_{++}(\mathbb{R})$ into $\mathcal{S}^d_{++}(\mathbb{R})$ [Stroock 2006, Lemma 5.2.1 p.131]: we use this with $\frac{y^2_{\lambda(c)}}{4} I_d + c^2 \in \mathcal{S}^d_{++}(\mathbb{R})$ for c close enough to c_0 (using $y_{\lambda(c)} > 0$ for $c \neq 0$). In view of (3.5.4), the continuity of $x(.)$ at $c_0 \neq 0$ follows.

Proof of ($\star$) h is continuous on $\mathbb{R}^d \times [0, \infty[$ into $\mathbb{R}$. Moreover,

- $h(\lambda, 0) = -2\sum_{i=1}^d |\lambda_i| \leq 0$ and $\lim_{y\to+\infty} h(\lambda, y) = +\infty$,
- h is continuously differentiable on $\mathbb{R}^d \times]0, \infty[$,
- $D_y h(\lambda, y) = 4 + d - \sum_{j=1}^q \frac{y}{\sqrt{y^2+4\lambda_i^2}} \geq 4$, implying that $y \mapsto h(\lambda, y)$ is (strictly) increasing.

Then, there is a unique $y_\lambda \in \mathbb{R}_+$ such that $h(\lambda, y_\lambda) = 0$. We point out at first glance, $\lambda \neq 0 \Leftrightarrow y_\lambda > 0$. The continuity of $y_.$ is proved on $\mathbb{R}^d_*$ on the one hand, and at 0 on the other hand.

- On $\mathbb{R}^d_* \times]0, +\infty[$: $D_y h(\lambda, y)$ exists and is non zero: then by the implicit function theorem, there exists an open set $U \subset \mathbb{R}^d_*$ containing λ and an open set $V \subset]0, +\infty[$ containing y_λ such that y is continuously differentiable from U to V. That proves the continuously differentiability of $y_.$ in $\mathbb{R}^d_*$.
- At $\lambda = 0$: $h((|\lambda|)_{1\leq i\leq d}, y) \leq h(\lambda, y)$ and $y \geq \frac{d|\lambda|}{\sqrt{4+2d}} \Leftrightarrow h((|\lambda|)_{1\leq i\leq d}, y) \geq 0$. It implies $0 \leq y_\lambda \leq \frac{d|\lambda|}{\sqrt{4+2d}}$ and $\lim_{|\lambda|\to 0} y_\lambda = 0$.

That concludes the continuity of $\lambda \mapsto y_\lambda$ on $\mathbb{R}^d$ and by the previous discussion, the proof of the lemma. □

3.5.3 Proof of Lemma 3.3.2

We have $\langle R^n \rangle_T = \int_0^T \left|\sigma_t^*(D_x u_t - D_x u_{\varphi(t)} - D^2_{xx} u_{\varphi(t)} \Delta S_t)\right|^2 \mathrm{d}t$: to prove the result, we aim at performing a Taylor expansion using $(\mathbf{A}_u)$, i.e. derivatives of u are *a.s.* finite in a small tube around $(t, S_t, Y_t)_{0\leq t\leq T}$. Because of this local assumption, a careful treatment is required, which we now detail. In view of $(\mathbf{A}_u)$, there exists $\Omega_{\mathcal{D}}$ such that $\mathbb{P}(\Omega_{\mathcal{D}}) = 1$ and for every $\omega \in \Omega_{\mathcal{D}}$ there is $\delta(\omega) > 0$ such that

$$|\mathcal{A}u|_{\delta}(\omega) := \sup_{0\leq t<T} \sup_{|x-S_t(\omega)|\leq\delta(\omega), |y-Y_t(\omega)|\leq\delta(\omega)} \left|\mathcal{A}u(t, x, y)\right| < +\infty$$

for any $\mathcal{A} \in \mathcal{D} := \left\{D^2_{x_jx_k}, D^3_{x_jx_kx_l}, D^2_{tx_j}, D^2_{x_jy_m} : 1 \le j,k,l \le d, 1 \le m \le d'\right\}$.
Since $\sup_{1\le i\le N^n_T} \Delta\tau^n_i \xrightarrow{a.s.} 0$ and $(S_t, Y_t)_{0\le t\le T}$ are *a.s.* continuous on the compact interval $[0,T]$, there exists $\Omega_{\mathcal{C}}$ with $\mathbb{P}(\Omega_{\mathcal{C}}) = 1$ such that for every $\omega \in \Omega_{\mathcal{C}}$, there is $p(\omega) \in \mathbb{N}$ such that $\forall n \ge p(\omega)$,

$$\Big(\sup_{0\le s,t\le T, |t-s|\le \sup_{1\le i\le N^n_T}\Delta\tau^n_i} |S_t - S_s| \vee |Y_t - Y_s|\Big)(\omega) \le \delta(\omega).$$

Hence for $\omega \in \Omega_{\mathcal{D}} \cap \Omega_{\mathcal{C}}$, let $n \ge p(\omega)$, $i \in \{1,\dots,N^n_T\}$ and $t \in [\tau^n_{i-1}, \tau^n_i]$, and write

$$\begin{aligned}
&D_x u(t, S_t, Y_t) - D_x u(\tau^n_{i-1}, S_{\tau^n_{i-1}}, Y_{\tau^n_{i-1}}) - D^2_{xx} u(\tau^n_{i-1}, S_{\tau^n_{i-1}}, Y_{\tau^n_{i-1}})\Delta S_t \\
&= [D_x u(t, S_t, Y_t) - D_x u(\tau^n_{i-1}, S_t, Y_t)] + [D_x u(\tau^n_{i-1}, S_t, Y_t) - D_x u(\tau^n_{i-1}, S_t, Y_{\tau^n_{i-1}})] \\
&\quad + \Big[D_x u(\tau^n_{i-1}, S_t, Y_{\tau^n_{i-1}}) - D_x u(\tau^n_{i-1}, S_{\tau^n_{i-1}}, Y_{\tau^n_{i-1}}) - D^2_{xx} u(\tau^n_{i-1}, S_{\tau^n_{i-1}}, Y_{\tau^n_{i-1}})\Delta S_t\Big].
\end{aligned}$$

Now apply Taylor theorem to the terms above, by observing that the involved derivatives of u are locally bounded by the (*a.s.* finite) random variable $C_u := \max_{\mathcal{A}\in\mathcal{D}} |\mathcal{A}u|_\delta$:

$$\begin{aligned}
&\big|D_x u(t, S_t, Y_t) - D_x u(\tau^n_{i-1}, S_{\tau^n_{i-1}}, Y_{\tau^n_{i-1}}) - D^2_{xx} u(\tau^n_{i-1}, S_{\tau^n_{i-1}}, Y_{\tau^n_{i-1}})\Delta S_t\big| \\
&\le \sqrt{d} C_u \Big((t - \tau^n_i) + \sqrt{d'}|Y_t - Y_{\tau^n_{i-1}}| + \frac{d}{2}|\Delta S_t|^2\Big).
\end{aligned}$$

Plugging this estimate in $\langle R^n\rangle_T$and using that Y is non-decreasing, we derive that *a.s.* , for n large enough,

$$\begin{aligned}
\varepsilon_n^{2-4\rho_N}\langle R^n\rangle_T \le 3dC_u^2 \sup_{0\le t\le T}|\sigma_t|^2 \varepsilon_n^{2-4\rho_N} \sum_{\tau^n_{i-1}<T} \Big((\Delta\tau^n_i)^3 + d'|\Delta Y_{\tau^n_i}|^2\Delta\tau^n_i \\
+ \frac{d^2}{4}\Delta\tau^n_i \sup_{\tau^n_{i-1}\le t\le\tau^n_i} |\Delta S_t|^4\Big).
\end{aligned}$$

To prove the *a.s.* convergence of the upper bound to 0, we separately analyze each of the three contributions.

- $\varepsilon_n^{2-4\rho_N} \sum_{\tau^n_{i-1}<T} (\Delta\tau^n_i)^3 \le \varepsilon_n^{2-4\rho_N} N^n_T \sup_{1\le i\le N^n_T} (\Delta\tau^n_i)^3 \le C_0 \varepsilon_n^{4-3\rho_N} \xrightarrow{a.s.} 0$ by Corollary 3.2.2-*ii)* with $\rho = \frac{4}{3} - \rho_N > 0$ (see $(\mathbf{A}_N)$).
- Combining $(\mathbf{A}_Y)$ and Corollary 3.2.2-*ii)* with $\rho = \frac{\rho_Y}{2} - 2(\rho_N - 1) > 0$, we easily obtain

$$\begin{aligned}
\varepsilon_n^{2-4\rho_N} \sum_{\tau^n_{i-1}<T} |\Delta Y_{\tau^n_i}|^2 \Delta\tau^n_i &\le \sum_{j=1}^{d'} (Y^j_T - Y^j_0)\varepsilon_n^{2-4\rho_N} \sup_{1\le i\le N^n_T} |\Delta Y^j_{\tau^n_i}| \sup_{1\le i\le N^n_T} \Delta\tau^n_i \\
&\le \sqrt{d'}|Y_T - Y_0| C_0 \varepsilon_n^{2-4\rho_N}\varepsilon_n^{\rho_Y}\varepsilon_n^{2-\rho} \le C_0 \varepsilon_n^{\rho_Y/2-2(\rho_N-1)} \xrightarrow{a.s.} 0.
\end{aligned}$$

- Using $(\mathbf{A}_S)$, $\varepsilon_n^{2-4\rho_N} \sum_{\tau^n_{i-1}<T} \Delta\tau^n_i \sup_{\tau^n_{i-1}\le t\le\tau^n_i} |\Delta S_t|^4 \le C_0 \varepsilon_n^{6-4\rho_N} T \xrightarrow{a.s.} 0$ since $\rho_N < \frac{3}{2}$.

All these convergences lead to the results. □

3.5.4 Assumption $(\mathbf{A}_u)$

We show that the assumption $(\mathbf{A}_u)$ is satisfied in most usual situations, even if the payoff g is not smooth. Actually, we have not been able to exhibit an example of g for which $(\mathbf{A}_u)$ does not hold. The following discussion should convince the reader that finding a counter-example is far from being straightforward, but we conjecture that it is possible.

Vanilla option in Black-Scholes model For *pedagogic reasons*, we start with the one-dimensional log-normal model $\mathrm{d}S_t = \sigma S_t \mathrm{d}B_t$ $(\sigma > 0)$. Consider first the Call option with strike $K > 0$: for $t < T$ we have $D_x u(t,x) = \mathcal{N}\big(\frac{\log(x/K)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}\big) \in [0,1]$ where $\mathcal{N}(.)$ is the cdf of the standard Gaussian law. The second derivative writes

$$D^2_{xx}u(t,x) = \frac{1}{\sigma x\sqrt{2\pi(T-t)}}\exp\Big(-\frac{1}{2}\Big[\frac{\log(x/K)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}\Big]^2\Big);$$

thus bounding the exponential term by 1, we have for any given $t_0 < T$ $\lim_{\delta\to 0}\sup_{0\le t\le t_0}\sup_{|x-S_t|\le\delta}|D^2_{xx}u(t,x)| \le \frac{1}{\sigma\inf_{0\le t\le T} S_t\sqrt{2\pi(T-t_0)}} < +\infty$. It shows that an *a.s.* finite bound on the second derivative is available provided that the time to maturity does not vanish. For the third derivative, this is similar: indeed using $\sup_{y\in\mathbb{R}} e^{y^2/4}|\partial_y(e^{-y^2/2})| = \sup_{y\in\mathbb{R}}|y|e^{-y^2/4} = \sqrt{2}e^{-1/2} \le 1$, we deduce

$$|D^3_{xxx}u(t,x)| \le \frac{1+\sigma\sqrt{T}}{x^2\sqrt{2\pi}\sigma^2(T-t)}\exp\Big(-\frac{1}{4}\Big[\frac{\log(x/K)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}\Big]^2\Big),$$

and as before $\lim_{\delta\to 0}\sup_{0\le t\le t_0}\sup_{|x-S_t|\le\delta}|D^3_{xxx}u(t,x)| < +\infty$ for any given $t_0 < T$. The next step consists in deriving *a.s.* upper bounds on derivatives for arbitrary small time to maturity. We take advantage of the property $\mathbb{P}(S_T \ne K) = 1$, which implies (by *a.s.* continuity of S) that for $\mathbb{P}$-a.e. ω there exists $t_0(\omega) \in [0,T[$ such that $\inf_{t_0(\omega)\le t\le T}|S_t(\omega) - K| \ge |S_T(\omega) - K|/2 := 2\delta_0(\omega) > 0$. Then, for $t \in [t_0(\omega), T]$ and $\delta \le \delta_0 \wedge [2^{-1}\inf_{0\le t\le T} S_t]$, we have $\inf_{|x-S_t|\le\delta}|\log(x/K)| \ge \inf_{u>0:|u-1|\ge\delta_0/K}|\log(u)| := c(\omega) > 0$ and $\inf_{|x-S_t|\le\delta} x \ge S_t/2$: therefore using the inequality $-(\alpha+\beta)^2 \le -\frac{\alpha^2}{2} + \beta^2$, we obtain, for $t \in [t_0(\omega), T[$

$$\sup_{|x-S_t|\le\delta}|D^2_{xx}u(t,x)| \le \frac{2}{\sigma S_t\sqrt{2\pi(T-t)}}\exp\Big(-\frac{c^2(\omega)}{4\sigma^2(T-t)} + \frac{1}{8}\sigma^2 T\Big).$$

Observe that $c(\omega) > 0$ implies that the above upper bound converges to 0 as $t \to T$: thus, we have completed the proof of $\lim_{\delta\to 0}\sup_{0\le t<T}\sup_{|x-S_t|\le\delta}\big|D^2_{xx}u(t,x)\big| < +\infty$ *a.s.* . For the third derivative, similarly we obtain for $t \in [t_0(\omega), T[$ and $\delta \le \delta_0(\omega) \wedge [2^{-1}\inf_{0\le t\le T} S_t(\omega)]$

$$\sup_{|x-S_t|\le\delta}|D^3_{xxx}u(t,x)| \le \frac{4(1+\sigma\sqrt{T})}{S_t^2\sqrt{2\pi}\sigma^2(T-t)}\exp\Big(-\frac{c^2(\omega)}{8\sigma^2(T-t)} + \frac{1}{16}\sigma^2 T\Big)$$

and we conclude as for the second derivative. To derive the property for $D^2_{tx}u$, we use the relation $D^2_{tx}u = -\frac{1}{2}\sigma^2 x^2 D^3_{xxx}u - \sigma^2 x D^2_{xx}u$. Finally, $(\mathbf{A}_u)$ is proved for the call option (and thus for the put option).
The same argumentation can be applied for the digital call option which payoff is of the form $g(x) = \mathbf{1}_{x\geq K}$: indeed, the derivatives of u blow up only at the discontinuity point K which has null probability for the law of S_T. $(\mathbf{A}_u)$ holds for digital options.

Vanilla option in general local volatility model The previous arguments are based on the explicit Black-Scholes formula for call and digital call options, but we can generalize them to more general models and payoffs and handle derivatives at any order. Denote by $X^j = \log(S^j)$ $(1 \leq j \leq d)$ the log-asset price in a diffusion model and assume that $\mathrm{d}X_t = b^X(t, X_t)\mathrm{d}t + \sigma^X(t, X_t)\mathrm{d}B_t$ for coefficients b^X and σ^X of class $\mathcal{C}^\infty_b\big([0,T] \times \mathbb{R}^d\big)$ (bounded with bounded derivatives). The price function in the log-variables is then $v(t,x) := u(t, \exp(x^1), \dots, \exp(x^d)) = \mathbb{E}(g(S_T)|S^j_t = \exp(x^j), 1 \leq j \leq d) := \mathbb{E}(G(X_T)|X_t = x)$. We first consider the simple case of $\mathcal{C}^\infty$-payoff G with exponentially bounded derivatives: for any $k \geq 0$, there is a constant $C^G_k \geq 0$ such that $|D^k_x G(x)| \leq C^G_k \exp(C^G_k|x|)$ for $x \in \mathbb{R}^d$. In that case, a direct differentiation of $\mathbb{E}(G(X_T)|X_t = x)$ using the smooth flow $x \mapsto X^{t,x}_T$ [Kunita 1984] shows the differentiability of v w.r.t. the space variable with derivatives bounded on compact subsets of $[0,T] \times \mathbb{R}^d$; in addition the time smoothness is obtained using Itô's formula; these arguments are standard and we skip details. $(\mathbf{A}_u)$ is proved for these smooth payoffs.
Now we tackle the case of discontinuous payoffs of the form $G(x) = \mathbf{1}_{x\in\mathcal{D}}\varphi(x)$ for a closed set $\mathcal{D} \subset \mathbb{R}^d$ and a $\mathcal{C}^\infty$-function φ with exponentially bounded derivatives : observe that by combining the analysis for smooth payoffs and that for discontinuous ones will allow to cover a quite large class of g satisfying $(\mathbf{A}_u)$ (such as call/put, digital call/put, exchange call, digital exchange call and so on). We assume that a uniform ellipticity assumption is satisfied: $\inf_{0\leq t\leq T, x\in\mathbb{R}^d} \inf_{|\xi|=1} \xi.[\sigma^X(\sigma^X)^*](t,x)\xi > 0$. In this setting, $v(t,x) = \int_{\mathbb{R}^d} \mathbf{1}_{z\in\mathcal{D}} p(t,x,T,z)\varphi(z)\mathrm{d}z$ where p is the transition density function of X, which is smooth and satisfies to Aronson-type estimates [Friedman 1964, Theorem 8 p. 263]: for any $i \geq 0$ and any differentiation index α, there exists a constant $C_{i,\alpha} = C_{i,\alpha}(T, b^X, \sigma^X) > 0$ such that

$$|D^{i,\alpha}_{tx} p(t,x,T,z)| \leq C_{i,\alpha}(T-t)^{-(d+2i+|\alpha|)/2} \exp(-|x-z|^2/[C_{i,\alpha}(T-t)])$$

for any $0 \leq t < T$, $x \in \mathbb{R}^d$, $z \in \mathbb{R}^d$. From the integral representation of v, it readily follows that

$$\begin{aligned}|D^{i,\alpha}_{tx} v(t,x)| &\leq C_{i,\alpha}(T-t)^{-(2i+|\alpha|)/2} \int_{\mathbb{R}^d} C_0^{\varphi} e^{C_0^{\varphi}|z|}(T-t)^{-d/2} e^{-|x-z|^2/[C_{i,\alpha}(T-t)]}\mathrm{d}z,\\ &\leq C_{i,\alpha}(T-t)^{-(2i+|\alpha|)/2} C_0^{\varphi} e^{C_0^{\varphi}|x|} \int_{\mathbb{R}^d} e^{C_0^{\varphi}\sqrt{T}|w|} e^{-|w|^2/C_{i,\alpha}}\mathrm{d}w,\end{aligned}$$

which proves locally uniform bounds on derivatives provided that the time to maturity remains bounded away from 0. To handle the case $t \to T$, we additionally assume that *the boundary $\partial\mathcal{D}$ of $\mathcal{D}$ is Lebesgue-negligible* (thus including usual situations but excluding Cantor like sets, see [DiBenedetto 2002, p. 114]): thus for $\mathbb{P}$-a.e. ω, the distance to the boundary (a closed set) is positive, i.e. $\delta_0(\omega) := \frac{1}{4} d(X_T(\omega), \partial\mathcal{D}) > 0$ and there exists $t_0(\omega) \in [0, T[$ such that $\inf_{t_0(\omega)\leq t\leq T} d(X_t(\omega), \partial\mathcal{D}) \geq 3\delta_0(\omega)$ (we recall that the distance function $x \mapsto d(x, \partial\mathcal{D})$ is Lipschitz continuous). Now, let ω be given as above; by the smooth version of the Urysohn lemma [Dieudonné 1990, p.90], there exists a smooth function ξ (depending on ω) such that $\mathbf{1}_{x\in\mathcal{D},\delta_0\leq d(x,\partial\mathcal{D})} \leq \xi(x) \leq \mathbf{1}_{x\in\mathcal{D}}$. Decompose the price function into two parts $v = v_1 + v_2$ with

$$v_1(t,x) := \int_{\mathbb{R}^d} \mathbf{1}_{z\in\mathcal{D}} p(t,x,T,z)\varphi(z)\xi(z)\mathrm{d}z,\ v_2(t,x) = \int_{\mathcal{D}} p(t,x,T,z)\varphi(z)(1-\xi(z))\mathrm{d}z.$$

We easily handle the derivatives of v_1 using the first case of smooth functions since $\mathbf{1}_{\mathcal{D}}\varphi\xi = \varphi\xi \in \mathcal{C}^\infty$ with exponentially bounded derivatives. Regarding v_2, observe that we integrate over the z such that $z \in \mathcal{D}$ and $d(z, \partial\mathcal{D}) < \delta_0$; for such z, for $t \in [t_0, T[$ and $|x - X_t| \leq \delta \leq \delta_0$, we have $|x-z| \geq d(X_t, \partial\mathcal{D}) - |x-X_t| - d(z, \partial\mathcal{D}) \geq \delta_0$ and thus

$$\begin{aligned}&\sup_{|x-X_t|\leq\delta} |D^{i,\alpha}_{tx} v_2(t,x)|\\ &\leq \sup_{|x-X_t|\leq\delta} \int_{\mathcal{D}} C_0^{\varphi} e^{C_0^{\varphi}|z|} C_{i,\alpha}(T-t)^{-(d+2i+|\alpha|)/2} e^{-|x-z|^2/[2C_{i,\alpha}(T-t)]} e^{-\delta_0^2/[2C_{i,\alpha}(T-t)]}\mathrm{d}z\\ &\leq C_{i,\alpha}(T-t)^{-(2i+|\alpha|)/2} e^{-\delta_0^2/[2C_{i,\alpha}(T-t)]} C_0^{\varphi} e^{C_0^{\varphi}(|X_t|+\delta_0)} \int_{\mathbb{R}^d} e^{C_0^{\varphi}\sqrt{T}|w|} e^{-|w|^2/[2C_{i,\alpha}]}\mathrm{d}w.\end{aligned}$$

The above upper bound converges to 0 as $t \to T$ and the proof of $(\mathbf{A}_u)$ is complete. Interestingly, we can weaken the ellipticity assumption into a hypoellipticity assumption: indeed, our analysis essentially relies on transition density estimates in small time and away from the diagonal. These estimates are available in the hypoelliptic homogeneous diffusion case [Kusuoka 1985, Corollary 3.25] and in the inhomogeneous case [Cattiaux 2002, Assumption (1.10)].

Asian option in general local volatility model The payoff is of the form $g(S_T, I_T)$ where $I_T = \int_0^T S_t \mathrm{d}t$ and S is a one-dimensional homogeneous diffusion $\mathrm{d}S_t = \sigma(S_t)\mathrm{d}B_t$. The analysis is reduced to the previous case of vanilla option by considering the 2-dimensional diffusion $(S_t, I_t)_{0\leq t\leq T}$: it is not elliptic but hypoelliptic [Kusuoka 1985] provided that σ is smooth and that $\sigma(x) > 0$ for $x \in I$ where $I \subset \mathbb{R}$ is given by $\mathbb{P}(\forall t \in [0,T] : X_t \in I) = 1$ (in usual cases, $I =]0, +\infty[$). It includes the Black-Scholes model and any model with local volatility bounded away from 0 and smooth. We skip details.

Lookback option in Black-Scholes model The payoff is of the form $S_T - m \wedge \min_{0\leq t\leq T} S_t$ or $M \vee \max_{0\leq t\leq T} S_t - S_T$ for lookback call or put, $(M \vee \max_{0\leq t\leq T} S_t - K)_+$ or $(K - m \wedge \min_{0\leq t\leq T} S_t)_+$ for call on maximum or on minimum, $(S_T - \lambda m \wedge \min_{0\leq t\leq T} S_t)_+$ (with $\lambda > 1$) or $(\lambda M \vee \max_{0\leq t\leq T} S_t - S_T)_+$ (with $\lambda < 1$) for partial lookback call or put. In all these cases, Black-Scholes type formulas are available in closed forms [Conze 1991]. Then it is straightforward to check that $(\mathbf{A}_u)$ is satisfied and this is essentially based on the property that under the assumption of non-zero volatility, the joint law $(S_T, \max_{0\leq t\leq T} S_t, \min_{0\leq t\leq T} S_t)$ has a density (derived from [Revuz 1999, Exercise 3.15]), implying that the events on which the derivatives may blow up (such as $\{S_T = \min_{0\leq t\leq T} S_t\} \dots$) have zero probability.

3.5.5 More numerical tests.

Here, we keep the same model and analysis as those in the numerical section, but we skip details. For an exchange option $g(S_T) = (S_T^1 - S_T^2)_+$, a call option $g(S_T) = (S_T^1 - K)_+$ and a binary option $g(S_T) = 1_{S_T^1 \geq K}$, we obtain respectively the graphs 3.3, (3.4,3.5,3.6) and (3.7,3.8,3.9) .

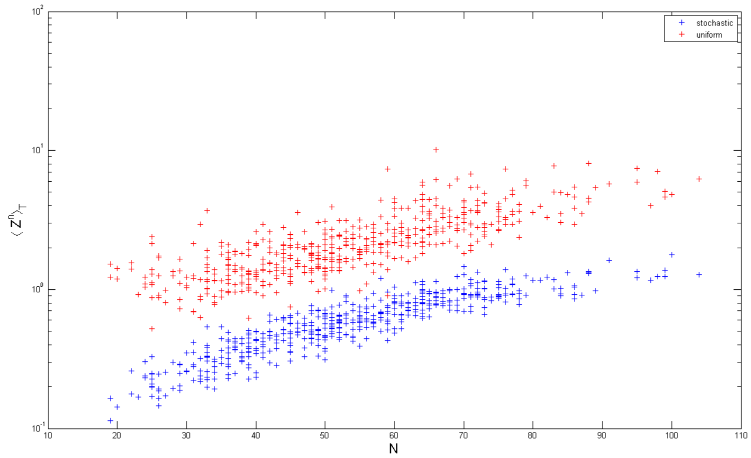

Figure 3.3: $g(S_T) = (S_T^1 - S_T^2)_+$, "+" and "+" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

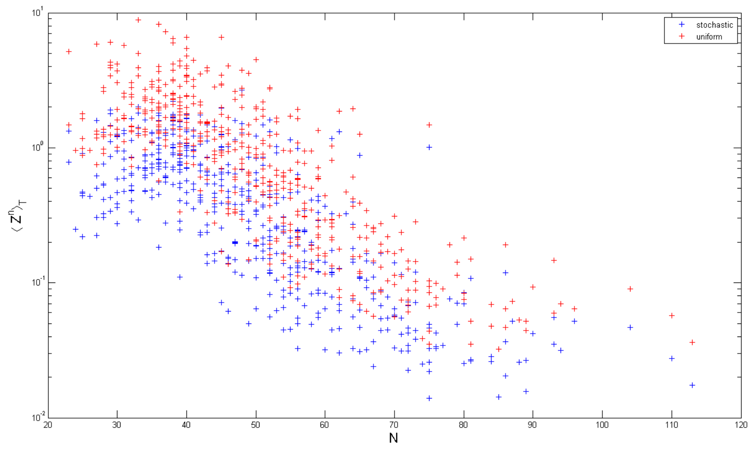

Figure 3.4: $g(S_T) = (S_T^1 - K)_+$, $K = 80$. "+" and "+" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

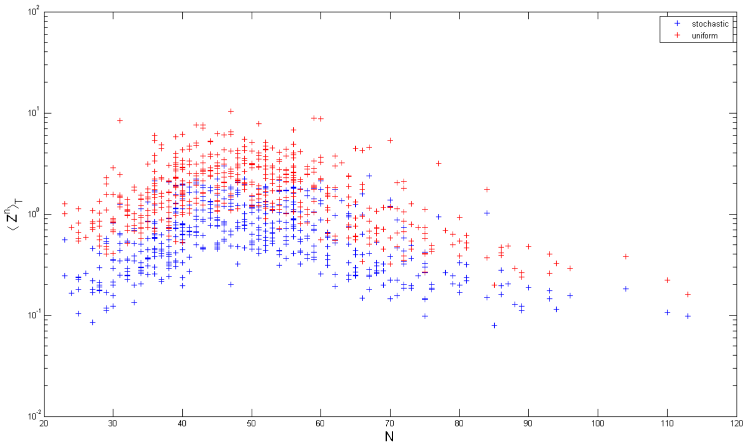

Figure 3.5: $g(S_T) = (S_T^1 - K)_+$, $K = 100$. "$+$" and "$+$" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

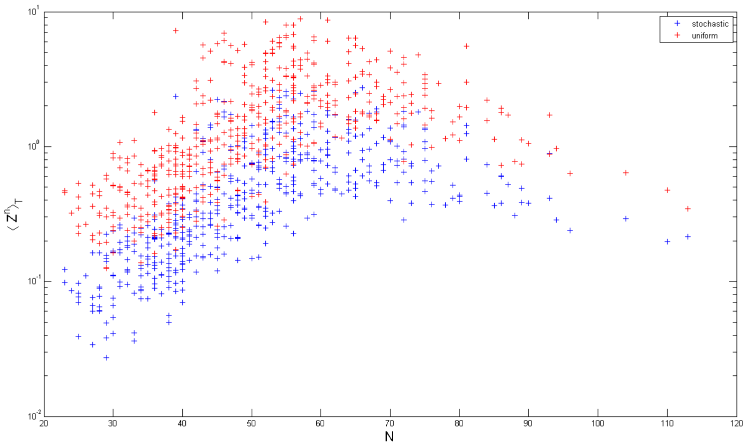

Figure 3.6: $g(S_T) = (S_T^1 - K)_+$, $K = 120$. "$+$" and "$+$" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

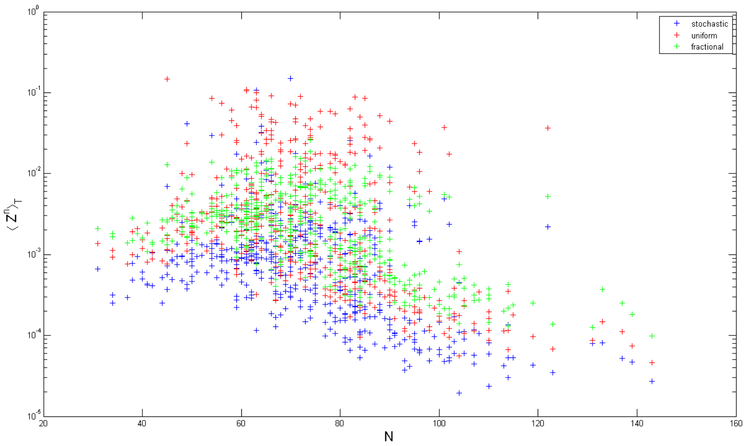

Figure 3.7: $g(S_T) = 1_{S_T^1 \geq K}$, $K = 80$. "+", "+" and "+" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$", "$\langle Z^n \rangle_{T,\text{fractional}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

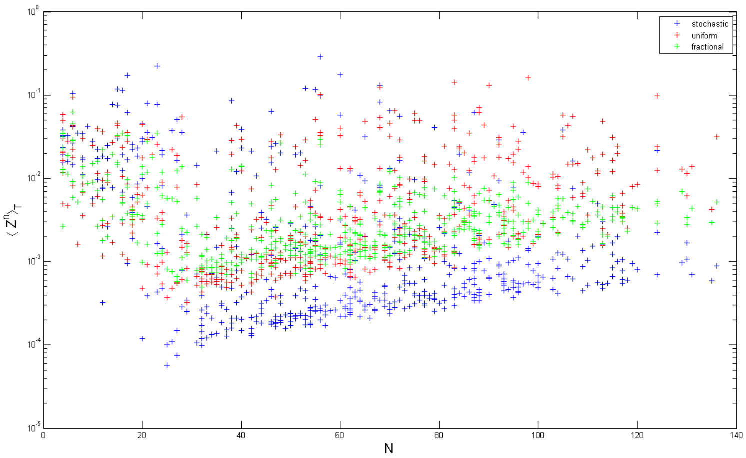

Figure 3.8: $g(S_T) = 1_{S_T^1 \geq K}$, $K = 100$. "+", "+" and "+" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$", "$\langle Z^n \rangle_{T,\text{fractional}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

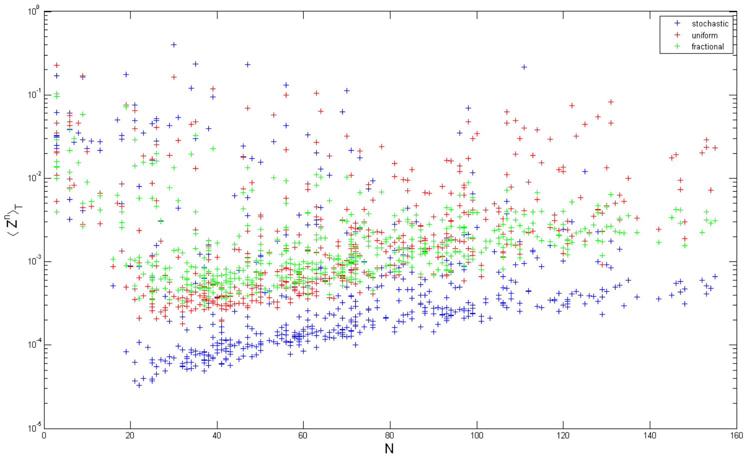

Figure 3.9: $g(S_T) = 1_{S_T^1 \geq K}$, $K = 120$. "+", "+" and "+" correspond respectively to "$\langle Z^n \rangle_{T,\text{uniform}}$", "$\langle Z^n \rangle_{T,\text{fractional}}$" and "$\langle Z^n \rangle_{T,\text{stochastic}}$".

Chapter 4

Almost sure optimal strategy and transaction costs

Contents

The work of this chapter is an application of the theoretical results proved in Chapter 3. We are going to exhibit almost sure hedging strategy when we take into account transaction costs.

4.1 Introduction

The problem The purpose is to extend the work already done to a market with transaction costs. Obviously, assuming only the hypotheses $(\mathbf{A}_N)$ and $(\mathbf{A}_S)$ of Chapter 3 is not enough and so it is virtually impossible to be as general as in a complete market about the class of stopping times involved, because we do not know *a priori* the number of asset units to hold, offsetting the transaction costs term. So, we *must* change the definition of admissible strategies and consider an optimal amount of underlying units which are held, noted $D_x v$, in addition to a finite sequence of optimal stopping times $\mathcal{T}^n = \{\tau_0^n = 0 < \tau_1^n < ... < \tau_i^n < ... \leq \tau_{N_T^n}^n = T\}$.

As in Chapter 3, the sequence of admissible strategies allows the *a.s.* convergence of the hedging error

$$Z_T^n = g(S_T) - \Bigg(v(0, S_0) + \sum_{\tau_{i-1}^n < T} D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n})(S_{\tau_i^n} - S_{\tau_{i-1}^n}) - k_n \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left| D_x v(\tau_i^n, S_{\tau_i^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \right| \Bigg),$$

to zero. Here, Z^n is understood as the hedging error [Denis 2010b] of a discrete Delta-hedging strategy of an European option with underlying asset S, maturity $T > 0$, price function v and pay-off $g(S_T)$, when transaction costs are small (in the following, k_n decays as ε_n). The times $(\tau_i^n)_{1 \le i \le N_T^n}$ read as rebalancing dates (or trading dates) and their number N_T^n is a random variable which is finite *a.s.* . The exponent n plays the same role as in Chapter 3. At first glance, it might seem by no way obvious that, here, there is a tradeoff between the hedging part and the transaction cost part of Z_T^n, which are of the same *order*, when k_n decreases to 0 as ε_n. So, we find a decomposition of Z_T^n in the form $M_T^n + R_T^n$, where M^n is a continuous local martingale and R^n is a continuous semi-martingale such that $\varepsilon_n^{-1} R_T^n$ tends to 0, when n increases to infinity. So, the sequence of optimal strategies minimizes the quadratic variation of M^n.

To our best knowledge, the problem of hedging under small transaction costs was first studied by Leland and more thoroughly by Klaus Lott in his thesis. In fact, when constant transaction costs are undergone, the strategy must tend (when the number of trading dates increases to infinity) to the Buy and Hold strategy, that is, we purchase the asset at the inception and sell it at the end. This is the cheapest super-replication strategy. To our point of view, the model suggested by Lott, in the absence of being realistic, gives nevertheless a robust setting : in the case of small transaction costs, this approach should give rather good outcomes. Furthermore, the independence of the option value w.r.t. the number of rebalancing dates seems to make sense in practice. So, the two directions ("k" kept constant or dwindling at some predefined rate) to cope with the problem of hedging under transaction costs seem to have each their favours and their drawbacks; Actually, the mathematical works in the literature always consider asymptotic results for tractability reasons; so when we perform a numerical non-asymptotic analysis of a strategy, the outcome may be mitigated. However, we obtain interesting outcomes − when the number of trading dates is fixed, the result with k decreasing to 0 seems to give more robust behaviour for the strategy involved than when we consider k constant. We shall study it in the numerical part of this chapter. Unlike the previous chapter, we cannot ensure the convergence of Z^n to 0, so it would be hopeless to get bound on any renormalisations. So, we must restrict even more the class of stopping times. However, looking at the results already obtained in chapter 2, the class of hitting times seems to be natural; indeed, under general assumptions on the stopping times, the optimal ones fall into the class of hitting times (cf Theorem 3.3.2). Then, the intuition is to take general hitting times of the process S and to optimize on the barrier with the help of the criterion

$$N_T^n \langle M^n \rangle_T, \tag{4.1.1}$$

designed in Chapter 3 (for the meaning of the criterion, we refer to the introduction of Chapter 3).

Outline of the chapter In the following, we fix the model and different assumptions. Section 4.2 is aimed at defining our new class of strategies under the presence of transaction costs. The convergence theorems are expressed and demonstrated in Section 4.3. Numerical experiments are given in Section 4.4.

Model Let $T > 0$ be a given terminal time (maturity) and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$ be a filtered probability space, supporting a 1-dimensional Brownian motion B defined on $[0, T]$, where $(\mathcal{F}_t)_{0\le t\le T}$ is the $\mathbb{P}$-augmented natural filtration of B and $\mathcal{F} = \mathcal{F}_T$. The price process S is given by an Itô continuous semi-martingale which solves

$$S_t = S_0 + \int_0^t b_s S_s \mathrm{d}s + \int_0^t \sigma(s, S_s) S_s \mathrm{d}B_s. \tag{4.1.2}$$

As in Chapter 3, we assume the return of the money market account $(r_t)_t$ to be zero, that $b \equiv 0$ and that σ fulfils the hypothesis $(\mathbf{A}_\sigma)$ or $(\mathbf{A}_\sigma^{\text{Ellip.}})$. From now on, S is a *continuous local martingale.*

Furthermore, the pay-off $g(S_T)$ satisfies the same kind of hypotheses as in Chapter 3. We do not mention the process Y for the sake of simplicity, but the computations would run in the same way.

4.2 A new class $\mathcal{T}^{\text{adm.}}$ of strategies

In this section, we define the class of strategies under consideration. A strategy is a finite sequence of increasing stopping times $\mathcal{T} = \{\tau_0 = 0 < \tau_1 < ... < \tau_i < ... \le \tau_{N_T} = T\}$ (with $N_T < +\infty$ *a.s.*) and a process δ which stand respectively for the rebalancing dates and the number of underlying units which are held in the replication portfolio (refers to Delta in the following).

Remark 4.2.1. *Unlike Chapter 3, the Deltas are not given by the usual Delta-neutral rule $D_x u(\tau_i, S_{\tau_i})$, but are part of the optimization scheme. This generalization comes from the fact that we wish optimal stopping times and convergence of the tracking error to* 0. *These two properties require two 'parameters' or quantities to adjust. Observe that we do not assume any link between the Deltas and the parameter n : this is Lott's approach.*

4.2.1 Assumptions

Let $(\varepsilon_n)_{n\geq 0}$ be a sequence of positive deterministic real numbers converging to 0 as $n \to \infty$; assume that it is a square-summable sequence

$$\sum_{n\geq 0} \varepsilon_n^2 < +\infty. \tag{4.2.1}$$

Now to derive *asymptotically* optimal results, we consider a family of hitting times $\mathcal{T}_n$ indexed by the integers $n = 0, 1, \dots$, i.e. writing

$$\tau_i^n := \inf\left\{t \geq \tau_{i-1}^n : \left|S_t - S_{\tau_{i-1}^n}\right| > \varepsilon_n.\lambda_{\tau_{i-1}^n}\right\} \wedge T, \qquad \text{for } i = 0, \dots, N_T^n, \tag{4.2.2}$$

where λ is a positive stochastic process, finite *a.s.* on $[0, T]$. We are now ready to define the class of sequence of strategies in which we are seeking the optimal element.

Definition 4.2.1. *A sequence of strategies $(\mathcal{T}_n, \delta)$ is* **admissible** *if there exist a smooth function λ and a sequence $(\varepsilon_n)_{n\geq 0}$ checking (4.2.1) such that for each n, $\mathcal{T}_n$ is in the form of (4.2.2) with $\lambda_t = \lambda(t, S_t) > 0$ (depending only on (t, S_t)) and there exists a function v such that $\delta = D_x v$ and v solves (k is a positive constant)*

$$D_t v(t,x) + \frac{\sigma(t,x)^2 x^2}{2}\left(1 + \frac{2kx}{\lambda(t,x)}\right) D_{xx}^2 v(t,x) = 0, \tag{4.2.3}$$

and

$$v(T, x) = g(x).$$

Besides, we assume g is convex and

$(\mathbf{A}_v)$ *Let $\mathcal{A} \in \mathcal{D} := \left\{D_{xx}^2, D_{xxx}^3, D_{tx}^2, D_{txx}^3, D_{xxxx}^4, D_{txxx}^4\right\}$,*

i) $\mathbb{P}\left(\lim_{\delta\to 0} \sup_{0\leq t<T} \sup_{|x-S_t|\leq\delta} \left|\mathcal{A}v(t,x)\right| < +\infty\right) = 1,$

ii) $\mathbb{P}\left(\inf_{t\in[0,T]} D_{xx}^2 v(t, S_t) > 0\right) = 1.$

The above definition depends on the sequence $(\varepsilon_n)_{n\geq 0}$, which is fixed from now on.

Remark 4.2.2.
- *The definition seems a bit tricky at first sight, but, as we shall see in the next section, these strategies are customized to deal with transaction costs. This insight is greatly inspired by [Ahn 1998].*

- *If we take $\lambda(t,x) = x$, the equation (4.2.3) becomes*

$$D_t v(t,x) + \frac{\sigma(t,x)^2 x^2}{2} (1 + 2k) D^2_{xx} v(t,x) = 0,$$

So, this is a constant change in the volatility similar to Leland's approach. We shall simulate this strategy and show that in practice, it works very well without any requisite numerical method.

- *Following Proposition 3.2.4, we have* $\sup_n \left(\varepsilon_n^2 N_T^n\right) < +\infty$ a.s. .

- ***Technical conditions (p.101-102 [El Karoui 1998])** : we assume that the function $\tilde{\sigma} : (t,x) \mapsto \sigma(t,x)\sqrt{1 + \frac{2kx}{\lambda(t,x)}}$ is continuous and bounded from above on $[0,T] \times (0,\infty)$, and that $\partial_x[x\tilde{\sigma}(t,x)]$ is continuous in (t,x) and Lipschitz continuous and bounded in x, uniformly in $t \in [0,T]$. Then, the convexity of g transfers to the solution v to the Cauchy problem (4.2.3) (whenever the solution exists). Nevertheless, the assumption of strict convexity in $(\mathbf{A}_v)$ is much more stringent : actually, it is a strong restriction on our model on g and S, which rules out the case of Call option for instance. The relaxation of this assumption is left to further research.*

4.3 Main results

4.3.1 Statements

We now go back to the hedging issue: the pay-off of the option is $g(S_T)$ and the admissible hedging portfolio with discrete rebalancing dates $\mathcal{T}^n$ and Delta $D_x v$ is $v(0,S_0) + \sum_{\tau^n_{i-1}<T} D_x v(\tau^n_{i-1}, S_{\tau^n_{i-1}})(S_{\tau^n_i} - S_{\tau^n_{i-1}}) - k_n \sum_{\tau^n_{i-1}<T} S_{\tau^n_i} \left| D_x v(\tau^n_i, S_{\tau^n_i}) - D_x v(\tau^n_{i-1}, S_{\tau^n_{i-1}})\right|$, which yields to an hedging error equal to

$$Z^n_T := g(S_T) - \left(v(0,S_0) + \sum_{\tau^n_{i-1}<T} D_x v(\tau^n_{i-1}, S_{\tau^n_{i-1}}) \Delta S_{\tau^n_i} - k_n \sum_{\tau^n_{i-1}<T} S_{\tau^n_i} \left|\Delta D_x v(\tau^n_i, S_{\tau^n_i})\right| \right).$$

As mentioned in the introduction, we choose

$$k_n = k\varepsilon_n,$$

for the transaction costs.
Owing to Ito's formula, we have

$$g(S_T) = v(0,S_0) + \int_0^T D_x v(t,S_t)\mathrm{d}S_t + \int_0^T \left(D_t v(t,S_t) + \frac{\sigma(t,S_t)^2 S_t^2}{2} D_{xx}^2 v(t,S_t) \right) \mathrm{d}t.$$

Then,

$$\begin{aligned} Z_T^n = & \int_0^T (D_x v(t,S_t) - D_x v(\varphi(t), S_{\varphi(t)}))\mathrm{d}S_t \\ & + k_n \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left| \Delta D_x v(\tau_i^n, S_{\tau_i^n}) \right| - \int_0^T \frac{k\sigma(t,S_t)^2 S_t^3 D_{xx}^2 v(t,S_t)}{\lambda(t,S_t)} \mathrm{d}t. \end{aligned} \tag{4.3.1}$$

4.3.2 Robust strategies under transaction costs.

Our first theorem consists in proving the convergence to 0 of Z_T^n :

Theorem 4.3.1. *Let $(\mathcal{T}, D_x v)$ be an admissible sequence of strategies. Then,*

$$Z_T^n \xrightarrow{a.s.} 0.$$

Proof. Owing to Corollary 3.2.1, for the first term of (4.3.1), it suffices to show that $\sum_{n=0}^{+\infty} \left(\int_0^T (D_x v(t,S_t) - D_x v(\varphi(t), S_{\varphi(t)}))^2 \mathrm{d}t \right)^2$ is finite. But, the Ito formula applied between $\varphi(t)$ and t gives

$$D_x v_t = D_x v_{\varphi(t)} + \int_{\varphi(t)}^t D_{xx}^2 v(s,S_s)\mathrm{d}S_s + \int_{\varphi(t)}^t \left(D_{tx}^2 v(s,S_s) + \frac{\sigma(s,S_s)^2 S_s^2}{2} D_{xxx}^2 v(s,S_s) \right) \mathrm{d}s.$$

We easily deduce the result from Corollary 3.2.2 and Corollary 3.2.3.
The second term of (4.3.1) needs more effort. Firstly,

$$k_n \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left(\left| \Delta D_x v(\tau_i^n, S_{\tau_i^n}) \right| - D_{xx}^2 v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) |\Delta S_{\tau_{i-1}^n}| \right) \xrightarrow{a.s.} 0.$$

Indeed,

$$\begin{aligned} & \left| \left| \Delta D_x v(\tau_i^n, S_{\tau_i^n}) \right| - D_{xx}^2 v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) |\Delta S_{\tau_{i-1}^n}| \right| \\ \leq & \left| \int_{\tau_{i-1}^n}^{\tau_i^n} \Delta D_{xx}^2 v(t,S_t)\mathrm{d}S_t \right| + \int_{\tau_{i-1}^n}^{\tau_i^n} \left| D_t v(t,S_t) + \frac{\sigma(t,S_t)^2 S_t^2}{2} D_{xxx}^2 v(t,S_t) \right| \mathrm{d}t. \end{aligned}$$

Applying Lemma 4.5.3 to the sequence $X = (S, D_{xx}^2 v(.,S), 1)$, the first term is bounded by $C_0 \varepsilon_n^{2-\rho}$ for any $\rho > 0$: so, taking $\rho = \frac{1}{2}$, we get the desired result. The second term above is easy to handle.

Last, using the form of the hitting times (4.2.2), we get,

$$
\begin{aligned}
k_n & \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} D_{xx}^2 v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) |\Delta S_{\tau_i^n}| \\
&= k \sum_{\tau_{i-1}^n < T} \frac{S_{\tau_i^n} D_{xx}^2 v(\tau_{i-1}^n, S_{\tau_{i-1}^n})}{\lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})} |\Delta S_{\tau_i^n}|^2 \\
&\quad - k \frac{S_T D_{xx}^2 v(\varphi(T), S_{\varphi(T)})}{\lambda(\varphi(T), S_{\varphi(T)})} |\Delta S_T|^2 + k_n S_T D_{xx}^2 v(\varphi(T), S_{\varphi(T)}) |\Delta S_T|.
\end{aligned}
$$

The two last terms readily converge to 0. The first term is analysed with Proposition 3.2.3, it gives

$$
k \sum_{\tau_{i-1}^n < T} \frac{S_{\tau_i^n} D_{xx}^2 v(\tau_{i-1}^n, S_{\tau_{i-1}^n})}{\lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})} |\Delta S_{\tau_i^n}|^2 \xrightarrow{a.s.} \int_0^T \frac{k\sigma(t, S_t)^2 S_t^3 D_{xx}^2 v(t, S_t)}{\lambda(t, S_t)} \mathrm{d}t.
$$

□

Remark 4.3.1. *The definition of the admissible sequence of strategies is clear from now on – v is the* good *function, which allows to compensate the transaction costs. The transaction costs term actually tends to that arising in the analysis of the robustness of Black-Scholes [El Karoui 1998]; indeed, with the notation $\tilde{\sigma}$ of Remark 4.2.2, the transaction cost term converges to $\frac{1}{2} \int_0^T (\tilde{\sigma}(t, S_t)^2 - \sigma(t, S_t)^2) S_t^2 D_{xx}^2 v(t, S_t) \mathrm{d}t$.*

4.3.3 Optimal strategies under transaction costs.

So far, we know that we handle *smart* strategies, which yield a tracking error as small as we want. Among them, we seek the optimal λ which lessens the renormalized tracking error.

Proposition 4.3.1. *The hedging error Z_T^n can be decomposed in the form*

$$Z_T^n = M_T^n + R_T^n,$$

with $\varepsilon_n^{-1} R_T^n \xrightarrow{a.s.} 0$ and M^n is a local martingale given in (4.3.12).

Theorem 4.3.2. *For any admissible strategies, we have*

$$
\liminf_{n \to +\infty} N_T^n \langle M^n \rangle_T \geq \left(\int_0^T \frac{1}{\sqrt{6}} \left(1 + \frac{2kS_t}{\lambda(t, S_t)} \right) |D_{xx}^2 v(t, S_t)| \mathrm{d}\langle S \rangle_t \right)^2. \tag{4.3.2}
$$

Set $\lambda(t, x) := -kx + \sqrt{(kx)^2 + \frac{\sqrt{6}}{|D_{xx} v(t,x)|}}$, with v is assumed to be a solution to the equation

$$
D_t v(t, x) + \frac{\sigma(t, x)^2 x^2}{2} \left(1 + \frac{2kx\sqrt{|D_{xx} v(t, x)|}}{-kx\sqrt{|D_{xx} v(t, x)|} + \sqrt{(kx)^2 |D_{xx} v(t, x)| + \sqrt{6}}} \right) D_{xx}^2 v(t, x) = 0,
$$

and

$$v(T,x)=g(x),$$

and we assume it defines a sequence of admissible strategies. Then, this sequence is optimal in the sense :

$$Z_T^n \xrightarrow{a.s.} 0$$

and

$$\begin{aligned} N_T^n\langle M^n\rangle_T &\xrightarrow{a.s.} \left(\int_0^T \frac{1}{\sqrt{6}}\left(1+\frac{2kS_t}{\lambda(t,S_t)}\right)|D_{xx}^2 v(t,S_t)|\mathrm{d}\langle S\rangle_t\right)^2 \\ &= \Bigg(\int_0^T \frac{|D_{xx}^2 v(t,S_t)|}{\sqrt{6}}\mathrm{d}\langle S\rangle_t \\ &\quad + \frac{k}{3}\int_0^T\left(kS_t^2D_{xx}^2v(t,S_t)^2+\sqrt{k^2S_t^4D_{xx}^2v(t,S_t)^4+\sqrt{6}|D_{xx}^2v(t,S_t)|^3S_t^2}\right)\mathrm{d}\langle S\rangle_t\Bigg)^2. \end{aligned}$$

Remark 4.3.2. *In the optimality result, observe that we do not prove that the threshold λ and the value function v exist and satisfy the admissibility assumptions. This is a non trivial issue that is left to further research.*

Remark 4.3.3. *For $k=0$, we acknowledge the optimal sequence of strategies of Theorem 3.3.2 (see Remark 3.3.1 for the one dimensional case).*

Proof of Proposition 4.3.1 and Theorem 4.3.2. The proof is really different from the remainder of the thesis. Here, we use heavily the dimension 1 of our problem; this demonstration cannot be extended to the multidimensional case.

Lower bound. We go back to the proof of Theorem 4.3.1, but making tighter estimates of the error terms. From (4.3.12), we show that

$$\begin{aligned} \langle M^n\rangle_T &= \int_0^T\left(1+\frac{2kS_{\varphi(t)}}{\lambda(\varphi(t),S_{\varphi(t)})}\right)^2 D_{xx}^2v(\varphi(t),S_{\varphi(t)})^2(\Delta S_t)^2\mathrm{d}\langle S\rangle_t \\ &= \frac{1}{6}\sum_{\tau_{i-1}^n<T}\left(1+\frac{2kS_{\tau_{i-1}^n}}{\lambda(\tau_{i-1}^n,S_{\tau_{i-1}^n})}\right)^2 D_{xx}^2v(\tau_{i-1}^n,S_{\tau_{i-1}^n})^2(\Delta S_{\tau_i^n})^4+R_n', \end{aligned}$$

where $R_n' := -\frac{2}{3}\int_0^T\left(1+\frac{2kS_{\varphi(t)}}{\lambda(\varphi(t),S_{\varphi(t)})}\right)^2 D_{xx}^2v(\varphi(t),S_{\varphi(t)})^2(\Delta S_t)^3\mathrm{d}S_t$. We prove easily that $N_T^nR_n' \xrightarrow{a.s.} 0$, owing to Corollary 3.2.1. Accepting it for a while, we have as in Chapter 3 from Cauchy-Schwarz's inequality,

$$N_T^n\langle M^n\rangle_T \geq \left(\sum_{\tau_{i-1}^n<T}\frac{1}{\sqrt{6}}\left(1+\frac{2kS_{\tau_{i-1}^n}}{\lambda(\tau_{i-1}^n,S_{\tau_{i-1}^n})}\right)|D_{xx}^2v(\tau_{i-1}^n,S_{\tau_{i-1}^n})|(\Delta S_{\tau_i^n})^2\right)^2+N_T^nR_n'.$$

Taking the infimum limit and we obtain the inequality (4.3.2) owing to Proposition 3.2.3.

Optimal strategy. Moreover, the thought remains the same, if one wants to have the equality at order 0, we need to force the equality, for each $i = 0, ..., N_T^n - 1$,

$$\frac{1}{6}\left(1 + \frac{2kS_{\tau_{i-1}^n}}{\lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})}\right)^2 D_{xx}^2 v(\tau_{i-1}^n, S_{\tau_{i-1}^n})^2 (\Delta S_{\tau_i^n})^4 = \varepsilon_n^4.$$

Remark 4.3.4. *We point out that the assumption* ii) *of* $(\mathbf{A}_v)$ *guarantees that* λ *remains bounded on* $[0, T]$ a.s. .

So, owing to the expression of the hitting times, we wish that $\lambda(t, S_t)$ solves the equation in y :

$$\left(y^2 + 2kS_t y\right)^2 D_{xx}^2 v(t, S_t)^2 = 6.$$

Among the roots, we seek the ones which exist independently of the parameters (e.g. when we take $k = 0$). That leads us to consider the equation

$$\left(y^2 + 2kS_t y\right) |D_{xx}^2 v(t, S_t)| - \sqrt{6} = 0.$$

A unique positive solution is given by

$$\lambda(t, S_t) = -kS_t + \sqrt{(kS_t)^2 + \frac{\sqrt{6}}{|D_{xx}^2 v(t, S_t)|}}. \tag{4.3.3}$$

Substituting λ in the equation (4.2.3), one gets

$$D_t v(t, x) + \frac{\sigma(t, x)^2 x^2}{2}\left(1 + \frac{2kx\sqrt{|D_{xx} v(t, x)|}}{-kx\sqrt{|D_{xx} v(t, x)|} + \sqrt{(kx)^2 |D_{xx} v(t, x)| + \sqrt{6}}}\right) D_{xx}^2 v(t, x) = 0,$$

and

$$v(T, x) = g(x).$$

Reciprocally, if we define λ by the equation (4.3.3) and v by a solution to the previous Cauchy problem, then we check that the associated admissible sequence of strategies is optimal. Indeed, for $i = 0, ..., N_T^n - 1$, the sequence of strategies is optimal by construction. The terms associated to the last date N_T^n need to be put apart as usual and carry on being negligible, we skip details, the proof is similar to that of Section 3.3.3 with $\mu = 0$.
So, the challenge of the proof is not to find the optimal sequence of admissible strategies, but to establish the *a.s.* convergence of $\varepsilon_n^{-1} R_T^n$ to 0. A similar statement has already been done in [Ahn 1998], but in the case of a geometric Brownian motion and $L^2(\Omega)$ convergence; estimates of the hitting times of Brownian motion were necessary to show their results.

Let us rewrite the expression (4.3.1)

$$\varepsilon_n^{-1} Z_T^n = \varepsilon_n^{-1} \int_0^T \left(1 + \frac{2kS_{\varphi(t)}}{\lambda(\varphi(t), S_{\varphi(t)})}\right) D^2_{xx} v(\varphi(t), S_{\varphi(t)}) \Delta S_t \mathrm{d}S_t \tag{4.3.4}$$

$$+ \varepsilon_n^{-1} \int_0^T \left(D_x v(t, S_t) - D_x v(\varphi(t), S_{\varphi(t)}) - D^2_{xx} v(\varphi(t), S_{\varphi(t)}) \Delta S_t\right) \mathrm{d}S_t \tag{4.3.5}$$

$$+ k \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left(\left| D_x v(\tau_i^n, S_{\tau_i^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \right| - \left| \alpha_{\tau_i^n} + \beta_{\tau_i^n} \right| \right) \tag{4.3.6}$$

$$+ k \sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n} \left(\left| \alpha_{\tau_i^n} + \beta_{\tau_i^n} \right| - \left| \alpha_{\tau_i^n} \right| \right) \tag{4.3.7}$$

$$+ k \sum_{\tau_{i-1}^n < T} \Delta S_{\tau_i^n} \left| \alpha_{\tau_i^n} \right| \tag{4.3.8}$$

$$+ k \sum_{\tau_{i-1}^n < T} S_{\tau_{i-1}^n} \left(\left| \alpha_{\tau_i^n} + \beta_{\tau_i^n} \right| - \left| \alpha_{\tau_i^n} \right| \right) \tag{4.3.9}$$

$$\begin{aligned} + \Bigg\{ & k \sum_{\tau_{i-1}^n < T} S_{\tau_{i-1}^n} D^2_{xx} v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \left| \Delta S_{\tau_i^n} \right| \\ & - 2k\varepsilon_n^{-1} \int_0^T \frac{S_{\varphi(t)}}{\lambda(\varphi(t), S_{\varphi(t)})} D^2_{xx} v(\varphi(t), S_{\varphi(t)}) \Delta S_t \mathrm{d}S_t \\ & - \varepsilon_n^{-1} \int_0^T \frac{k S_{\varphi(t)} D^2_{xx} v(\varphi(t), S_{\varphi(t)})}{\lambda(\varphi(t), S_{\varphi(t)})} \mathrm{d}\langle S \rangle_t \Bigg\} \end{aligned} \tag{4.3.10}$$

$$- k\varepsilon_n^{-1} \int_0^T \Delta \left(\frac{S_t D^2_{xx} v(t, S_t)}{\lambda(t, S_t)} \right) \mathrm{d}\langle S \rangle_t. \tag{4.3.11}$$

where $\begin{cases} \alpha_t := D^2_{xx} v(\varphi(t), S_{\varphi(t)}) \Delta S_t, \\ \beta_t := D_t v(\varphi(t), S_{\varphi(t)}) \Delta t + \frac{1}{2} D^3_{xxx} v(\varphi(t), S_{\varphi(t)}) (\Delta S_t)^2. \end{cases}$

We shall prove the convergence to 0 of (4.3.5), (4.3.6), (4.3.7), (4.3.8), (4.3.9), (4.3.10) and (4.3.11). So, the promised decomposition

$$Z_T^n = M_T^n + R_T^n$$

of the introduction holds with

$$M^n := \int_0^{\cdot} \left(1 + \frac{2kS_{\varphi(t)}}{\lambda(\varphi(t), S_{\varphi(t)})}\right) D^2_{xx} v(\varphi(t), S_{\varphi(t)}) \Delta S_t \mathrm{d}S_t \tag{4.3.12}$$

and R_T^n standing for the other terms times ε_n.

- The demonstration for the term (4.3.5) is already done in Lemma 3.3.2.

- For the term (4.3.6), we use the coarse triangle inequality :

$$
\begin{aligned}
&k\left|\sum_{\tau_{i-1}^n<T} S_{\tau_i^n}\left(\left|\Delta D_x v(\tau_i^n, S_{\tau_i^n})\right| - \left|\alpha_{\tau_i^n}+\beta_{\tau_i^n}\right|\right)\right| \\
&\leq k \sup_{0\leq t\leq T} S_t\left[\sum_{\tau_{i-1}^n<T}\left|\int_{\tau_{i-1}^n}^{\tau_i^n}\Delta D^2_{xx}v(t,S_t)\mathrm{d}S_t - D^3_{xxx}v(\tau_{i-1}^n, S_{\tau_{i-1}^n})\frac{(\Delta S_{\tau_i^n})^2-\Delta\langle S\rangle_{\tau_i^n}}{2}\right|\right.\\
&\quad\left.+\int_0^T\left|\Delta D_t v(t,S_t)\right|\mathrm{d}t + \int_0^T \frac{\left|\Delta D^3_{xxx}v(t,S_t)\right|}{2}\mathrm{d}\langle S\rangle_t\right].
\end{aligned}
$$

The latter terms converge to 0 owing to Lebesgue's theorem; indeed, for all $t\in[0,T]$, $\Delta D_t v(t,S_t)\xrightarrow{a.s.} 0$ and $\Delta D^3_{xxx}v(t,S_t)\xrightarrow{a.s.} 0$, the integrands are bounded respectively by $2\sup_{0\leq t\leq T}|D_t v(t,S_t)|$ and $2\sup_{0\leq t\leq T}|D^3_{xxx}v(t,S_t)|$, which are finite by assumption $(\mathbf{A}_v)$.
For the first term, we are going to use the Ito formula on $s\mapsto D^2_{xx}v(s,S_s)$ between $\varphi(t)$ and t :

$$
\begin{aligned}
&\left|\int_{\tau_{i-1}^n}^{\tau_i^n}\Delta D^2_{xx}v(t,S_t)\mathrm{d}S_t - D^3_{xxx}v(\tau_{i-1}^n, S_{\tau_{i-1}^n})\frac{(\Delta S_{\tau_i^n})^2-\Delta\langle S\rangle_{\tau_i^n}}{2}\right| \\
&=\left|\int_{\tau_{i-1}^n}^{\tau_i^n}\left(\int_{\tau_{i-1}^n}^t \Delta D^3_{xxx}v(s,S_s)\mathrm{d}S_s + \int_{\tau_{i-1}^n}^t\left(D^3_{xxt}v(s,S_s)+\frac{\sigma(s,S_s)^2S_s^2}{2}D^4_{xxxx}v(s,S_s)\right)\mathrm{d}s\right)\mathrm{d}S_t\right|.
\end{aligned}
$$

Applying Lemma 4.5.3 to the sequences $X=(S,S,D^3_{xxx}v(.,S),1)$ and $X=(S,\mathrm{Id},D^3_{xxt}v(.,S)+\frac{\sigma(.,S)^2S^2}{2}D^4_{xxxx}v(.,S))$, we conclude that the above term is bounded by $C\varepsilon_n^{3-\rho}$ (for any $\rho>0$) : thus the summation over i is bounded by $C\varepsilon_n^{1-\rho}$ and converges to 0 (take $\rho=1/2$).

- The convergence to 0 of the term (4.3.7) follows from :

$$
\left|\sum_{\tau_{i-1}^n<T}\Delta S_{\tau_i^n}\left(\left|\alpha_{\tau_i^n}+\beta_{\tau_i^n}\right|-\left|\alpha_{\tau_i^n}\right|\right)\right| \leq \sup_{0\leq t\leq T}|\Delta S_t|\sum_{\tau_{i-1}^n<T}\left|\beta_{\tau_i^n}\right| \xrightarrow{a.s.} 0.
$$

- The term (4.3.8) can be rewritten in the following form :

$$
\begin{aligned}
k\sum_{\tau_{i-1}^n<T}\Delta S_{\tau_i^n}\left|\alpha_{\tau_i^n}\right| &= k\varepsilon_n\int_0^T D^2_{xx}v(\varphi(t),S_{\varphi(t)})\lambda(\varphi(t),S_{\varphi(t)})\mathrm{d}S_t\\
&\quad - k\varepsilon_n D^2_{xx}v(\varphi(T),S_{\varphi(T)})\lambda(\varphi(T),S_{\varphi(T)})\Delta S_T + k\Delta S_T\left|\alpha_T\right|.
\end{aligned}
$$

The two last terms converge to 0. To conclude the convergence of the first term to 0, we apply Corollary 3.2.1 with $p=2$.

- We remark that so far we have not used in the proof of convergence of the residual terms to 0, the assumption of strict convexity of the value function v, but here it plays a crucial role to show the convergence of the term (4.3.9) to 0. Set $\xi_t := S_{\varphi(t)}\left(|\alpha_t+\beta_t| - |\alpha_t|\right)$. We remark that

$$\sup_{n\geq 0}\sum_{\tau_{i-1}^n<T}|\xi_{\tau_i^n}| \leq \sup_{n\geq 0}\sum_{\tau_{i-1}^n<T} S_{\tau_{i-1}^n}\left|\beta_{\tau_i^n}\right| < +\infty.$$

Then, a rough estimate does not lead to the convergence to 0 and we need to be more accurate. This is the purpose of the following identity : for any couple $(x,y)\in\mathbb{R}^2$ with $x\neq 0$,

$$|x+y|-|x| = sgn(x)y + (sgn(y)-sgn(x))(x+y)1_{|x|\leq|y|}.$$

Therefore, we have

$$\xi_{\tau_i^n} = sgn(\alpha_{\tau_i^n})\beta_{\tau_i^n} + (sgn(\beta_{\tau_i^n}) - sgn(\alpha_{\tau_i^n}))(\alpha_{\tau_i^n}+\beta_{\tau_i^n})1_{|\alpha_{\tau_i^n}|\leq|\beta_{\tau_i^n}|}.$$

Observe that owing to the assumption *ii)* of $(\mathbf{A}_v)$, we have

$$\begin{aligned}
&\sup_{1\leq i\leq N_T^n-1} 1_{|\alpha_{\tau_i^n}|\leq|\beta_{\tau_i^n}|}\\
&\leq 1_{0<\min_{t\in[0,T]} D^2_{xx}v(t,S_t)\lambda(t,S_t)\leq \max_{t\in[0,T]}|D_tv(t,S_t)|\varepsilon_n^{-1}\sup_{1\leq i\leq N_T^n}\Delta\tau_i^n+\frac{\varepsilon_n}{2}\max_{t\in[0,T]}|D^3_{xxx}v(t,S_t)|\lambda(t,S_t)^2}\\
&\xrightarrow{a.s.} 0,
\end{aligned}$$

because $\varepsilon_n^{-1}\sup_{1\leq i\leq N_T^n}\Delta\tau_i^n \xrightarrow{a.s.} 0$ by Corollary 3.2.2.
Then, we deduce the convergence to 0 of the second term because

$$\begin{aligned}
&\left|\sum_{\tau_{i-1}^n<T}[(sgn(\beta_{\tau_i^n})-sgn(\alpha_{\tau_i^n}))(\alpha_{\tau_i^n}+\beta_{\tau_i^n})]1_{|\alpha_{\tau_i^n}|\leq|\beta_{\tau_i^n}|}\right|\\
&\qquad\leq 4\sup_{1\leq i\leq N_T^n-1}1_{|\alpha_{\tau_i^n}|\leq|\beta_{\tau_i^n}|}\sum_{\tau_{i-1}^n<\varphi(T)}|\beta_{\tau_i^n}| + 4|\beta_T| \xrightarrow{a.s.} 0.
\end{aligned}$$

The first term is a little bit more delicate to handle : write $sgn(\alpha_{\tau_i^n}) = \frac{\alpha_{\tau_i^n}}{|\alpha_{\tau_i^n}|} = \frac{\Delta S_{\tau_i^n}}{\varepsilon_n\lambda(\tau_{i-1}^n,S_{\tau_{i-1}^n})}$ to obtain

$$\begin{aligned}
\sum_{\tau_{i-1}^n<T} sgn(\alpha_{\tau_i^n})\beta_{\tau_i^n} &= \varepsilon_n^{-1}\sum_{\tau_{i-1}^n<T}\frac{D_tv(\tau_{i-1}^n,S_{\tau_{i-1}^n})\Delta S_{\tau_i^n}\Delta\tau_i^n + \frac{1}{2}D^3_{xxx}v(\tau_{i-1}^n,S_{\tau_{i-1}^n})(\Delta S_{\tau_i^n})^3}{\lambda(\tau_{i-1}^n,S_{\tau_{i-1}^n})}\\
&\quad -\varepsilon_n^{-1}\frac{D_tv(\varphi(T),S_{\varphi(T)})\Delta S_T\Delta T + \frac{1}{2}D^3_{xxx}v(\varphi(T),S_{\varphi(T)})(\Delta S_T)^3}{\lambda(\varphi(T),S_{\varphi(T)})}\\
&\quad + sgn(\alpha_T)\beta_T.
\end{aligned}$$

The two last terms clearly converge to 0. We note that owing to Ito's formula, we have

$$\Delta S_{\tau_i^n}\Delta\tau_i^n = \frac{(\Delta S_{\tau_i^n})^3}{3\sigma^2_{\tau^n_{i-1}}} - \frac{1}{\sigma^2_{\tau^n_{i-1}}}\int_{\tau^n_{i-1}}^{\tau_i^n}(\Delta S_t)^2 \mathrm{d}S_t + \int_{\tau^n_{i-1}}^{\tau_i^n}\Delta S_t \Delta\left(\frac{1}{\sigma_t^2}\right)\mathrm{d}\langle S\rangle_t + \int_{\tau^n_{i-1}}^{\tau_i^n}(t-\tau^n_{i-1})\mathrm{d}S_t. \tag{4.3.13}$$

The convergence of the terms related to $(\Delta S_{\tau_i^n})^3$ is due to Lemma 4.5.2 with $p=3$. The other terms converge to 0, now classically, owing to Corollary 3.2.1.

- All the contributions (except $i=N_T^n$) in the summation (4.3.10) are equal to 0 by application of Ito's lemma to $t\mapsto(\Delta S_t)^2$ between τ^n_{i-1} and τ_i^n and by definition of the hitting times. The last term ($i=N_T^n$ and $\int_{\tau^n_{N_T^n-1}}^T\cdots$) converges to 0.

- For the term (4.3.11), we cannot apply directly Lemma 4.5.3 and need to be more careful on the estimation of the integral. So, we utilize Lemma 4.5.4 to the martingale part of $\frac{S_t D^2_{xx}v(t,S_t)}{\lambda(t,S_t)}$; for the finite variation part, the result is obtained by a rough estimation. The proof is done.

□

4.4 Numerical experiments

4.4.1 Algorithm

Theorem 4.3.2 states the optimal sequence of strategies − first of all, we need to solve the non-linear Cauchy problem

$$D_t v(t,x) + f(t,x,D^2_{xx}v(t,x)) = 0,$$

and

$$v(T,x) = g(x),$$

with

$$f(t,x,\Gamma) := \frac{\sigma(t,x)^2x^2}{2}\left(1+\frac{2kx\sqrt{\Gamma}}{-kx\sqrt{\Gamma}+\sqrt{(kx)^2\Gamma+\sqrt{6}}}\right)\Gamma.$$

Needless to say that an explicit solution is hopeless, so we choose to work out a sequence of equations using an explicit finite difference method : let $(\Delta t,\Delta x)\in\mathbb{R}^2_+$, let $(t_i = i\Delta t)_{1\le i\le n}$ be the uniform time mesh and let $(x_j = j\Delta x)_{1\le j\le m}$ be the uniform spacial mesh, we consider

$$\begin{cases} v(t_{i-1},x_j) = v(t_i,x_j) + \Delta t f\left(t_i, x_i, \frac{v(t_i,x_{j+1})+v(t_i,x_{j-1})-2v(t_i,x_j)}{\Delta x^2}\right), & 1 \le j \le m-1 \\ v(T,x_j) = g(x_j), \\ v(t_i,x_m) = g(x_m), \\ v(t_i,0) = g(0). \end{cases} \tag{4.4.1}$$

We need a CFL-type relation $\Delta t \le \Delta x^2$ for the convergence of the scheme.
Secondly, the approximated sequence of stopping times is given by

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n := \inf\left\{t_j > \tau_{i-1}^n : \left|S_{t_j} - S_{\tau_{i-1}^n}\right| > \varepsilon_n.\bar{\lambda}(\tau_{i-1}^n, S_{\tau_{i-1}^n})\right\} \wedge T, & \text{for } i = 1, \dots, N_T^n, \end{cases}$$

where

$$\bar{\lambda}(t_i, y) = -ky + \sqrt{(ky)^2 + \frac{\sqrt{6}\Delta x^2}{(v(t_i,\pi_x(y)+\Delta x) + v(t_i,\pi_x(y)-\Delta x) - 2v(t_i,\pi_x(y)))}},$$

and $\pi_x(y) := \arg\min(|x_j - y| : 1 \le j \le m)$ is the projection on the spacial mesh. In the simulation, we shall take $m = 400$, $x_{\min} = 0$, $x_{\max} = 200$, $\Delta x = 0.5$, $\Delta t = \Delta x^2$ and $n = 0.5/\Delta t$.

4.4.2 Example

For the application, we hedge the selling of an European option with strike $K = 100$, maturity $T = 0.5$. The model for the asset S is a geometric Brownian motion with initial price $S_0 = 100$, volatility $\sigma = 0.3$ and drift null. For the sake of simplicity, we take the rate $r = 0$. The transaction cost is fixed at $k_n = 100$ bps, which seems to be a realistic estimation in practice.
We note $C(t,x,\sigma)$ the price given by the Black-Scholes formula at (t,x) with the volatility σ.
We compare our strategy to six other strategies :

- Leland's strategy [Leland 1985] : plot an enhanced volatility $\sigma_n := \sigma\sqrt{1 + k\sqrt{\frac{8n}{\pi\Delta t}}}$ in Black-Scholes formula.
- Denis's strategy [Denis 2010a] : use Leland's modified volatility and Delta equal to $C_x(t_i, S_{t_i}, \sigma_n) - \sum_{1\le j\le i}\left[C_x(t_j, S_{t_{j-1}}, \sigma_n) - C_x(t_{j-1}, S_{t_{j-1}}, \sigma_n)\right]$.
- Whalley-Wilmott's strategy [Whalley 1997] : use the Black-Scholes Delta and rebalance the portfolio at stopping times given by

$$\tau_i^n := \inf\left\{t \ge \tau_{i-1}^n : |C_x(t, S_t, \sigma) - C_x(\tau_{i-1}^n, S_{\tau_{i-1}^n}, \sigma)| > \left(\frac{3kS_{\tau_{i-1}^n} C_{xx}(\tau_{i-1}^n, S_{\tau_{i-1}^n}, \sigma)}{2\gamma}\right)^{1/3}\right\}.$$

- Barles-Soner's strategy [Barles 1998] : let f be a unique solution of the non-linear equation
$$\frac{\mathrm{d}f(z)}{\mathrm{d}z} = \frac{f(z)+1}{2\sqrt{zf(z)}-z}, \quad z \neq 0, \quad f(0) = 0.$$
We shall sort it out by the function *ode45* of MATLAB software. The price function satisfies the following Cauchy problem
$$D_t w(t,x) + \frac{\sigma^2 x^2}{2} D^2_{xx} w(t,x) \left(1 + f(k^2\gamma^2 x^2 D^2_{xx} w(t,x))\right) = 0$$
and
$$v(T,x) = (x-K)_+.$$
We solve it by the same method as Section 4.4.1. At least, we define the volatility by
$$\sigma_{BaSo} := \sigma\sqrt{1 + f(k^2\gamma S^2_{\tau^n_{i-1}} D^2_{xx} w(\tau^n_{i-1}, S_{\tau^n_{i-1}}))}.$$
The strategy consists of Delta defined by Dw and rebalancing dates
$$\tau^n_i := \inf\left\{ t > \tau^n_{i-1} : \left| D_x w(t, S_t, \sigma_{BaSo}) - D_x w(\tau^n_{i-1}, S_{\tau^n_{i-1}}, \sigma_{BaSo}) \right| \right.$$
$$\left. > \frac{g(k^2\gamma S^2_{\tau^n_{i-1}} D^2_{xx} w(\tau^n_{i-1}, S_{\tau^n_{i-1}}, \sigma_{BaSo}))}{k\gamma S_{\tau^n_{i-1}}} \right\},$$
where $g(z) := \sqrt{zf(z)} - z$.

- The relative change strategy : this strategy is made up of two components – the first one seems to be well-known in the literature (see [Henrotte 1993]) and consists in rebalancing at times of the form
$$\tau^n_i := \inf\left\{ t > \tau^n_{i-1} : \frac{|S_t - S_{\tau^n_{i-1}}|}{S_{\tau^n_{i-1}}} > \varepsilon_n \right\};$$
the second one is new, to our knowledge : we consider Black-Scholes's deltas with the following change of volatility
$$\tilde{\sigma} := \sigma\sqrt{1+2k}.$$
To figure out why we choose such a volatility, we refer to the equation (4.2.3), where we take $\lambda(t,x) = x$. This choice enables us to get a similar equation for the price without transaction costs, but with a constant change in the volatility. The strategy is interesting because it is admissible (see Definition 4.2.1) and in the Black-Scholes model, we just have to substitute the volatility parameter by the new one as in Leland's strategy. However, the drawback is the non-optimality of the strategy, but as we shall see the strategy performs very well.

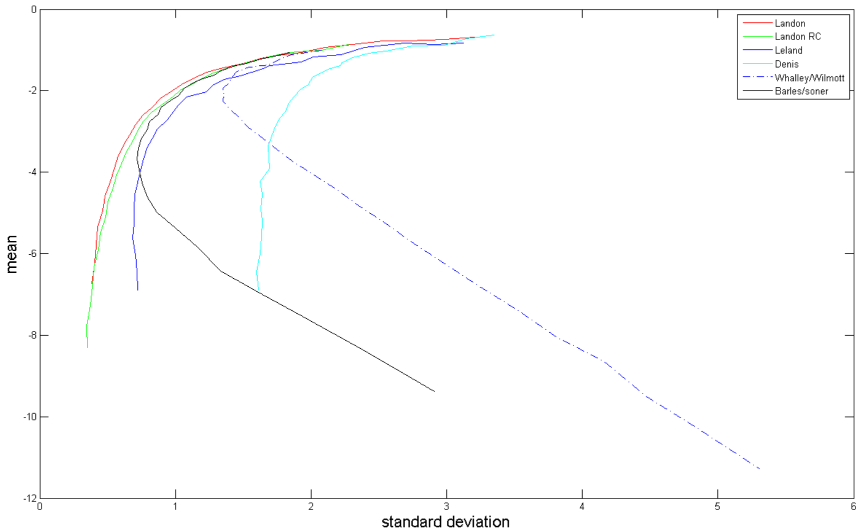

Figure 4.1: Mean-Variance of the hedging error for $k = 100$

In Figure 4.1 and Figure 4.2, we plot for each strategy, the couple mean-variance and mean-VaR (with 2000 draws) of the (new) hedging error :

$$
\begin{aligned}
& C(0, S_0) + \sum_{\tau_{i-1}^n < T} D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n})(S_{\tau_i^n \wedge s} - S_{\tau_{i-1}^n}) \\
& \quad - k \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left| D_x v(\tau_i^n, S_{\tau_i^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \right| - (S_T - K)_+ \\
& = \int_0^T (D_x v(\varphi(t), S_{\varphi(t)}) - C_x(t, S_t)) \mathrm{d}S_t - k_n \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left| D_x v(\tau_i^n, S_{\tau_i^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \right|.
\end{aligned}
$$

In particular, the mean of the hedging error is the mean of the cumulated transaction costs. In the simulation, we vary the step time Δt for Leland's strategy and Denis's strategy, the risk aversion γ for Whalley-Wilmott's strategy and Barles-Soner's strategy and the parameter ε_n for Landon's strategy and the relative change strategy (here, k and ε_n move simultaneously with the notations of Section 4.2, so that k_n is kept constant).
The following table gives the mean of minimum and maximum of the number of trading dates :

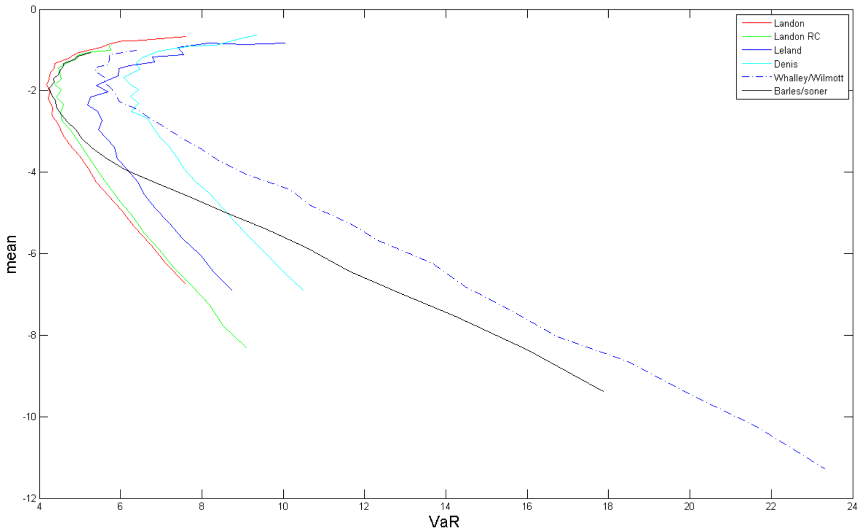

Figure 4.2: Mean-VaR of the hedging error for $k = 100$

Strategies	N_T^n	ε_n or γ
Landon	$[10, 652]$	$[0.12, 1]$
Relative Change	$[6, 843]$	$[0.007, 0.1]$
Leland	$[10, 700]$	
Denis	$[10, 700]$	
Whalley-Wilmott	$[3, 720]$	$[0.4, 150]$
Barles-Soner	$[8, 544]$	$[0.5, 50]$

We have chosen the parameters (ε_n or γ) such that the sets of values taken by N_T^n for each strategy have approximatively the same range. That enables us to compare fairly the different strategies.

4.5 Appendix

The dimension 1, as opposed to the multidimensional setting in Chapter 3, enables us more convergence results owing to the symmetry of the barrier of the hitting times as we shall describe in the following lemmas. We note

$$\tilde{\tau}_i^n := \inf\{t > \tau_{i-1}^n : |S_t - S_{\tau_{i-1}^n}| > \varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})\}.$$

Then, by definition, $\tau_i^n = \tilde{\tau}_i^n \wedge T$.

Lemma 4.5.1. *Let θ be a stopping time such that $\theta \le T$* a.s. . *For each $i = 1, ..., N_T^n$, for $p \in \mathbb{N}$,*

$$\mathbb{E}_{\tau_{i-1}^n}((\Delta S_{\tau_i^n \wedge \theta})^{2p+1})1_{\tau_{i-1}^n < \theta} = \mathbb{E}_{\tau_{i-1}^n}\Big([(\Delta S_\theta)^{2p+1} - (\Delta S_{\tilde{\tau}_{N_\theta^n}^n})^{2p+1}]1_{i=N_\theta^n}\Big).$$

Proof. For all $t \ge \tau_{i-1}^n$, $\mathbb{E}_{\tau_{i-1}^n}(\Delta S_{t\wedge\tilde{\tau}_i^n}) = 0$ and $|\Delta S_{t\wedge\tilde{\tau}_i^n})| \le \varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})$. Then, Lebesgue's theorem implies that

$$\mathbb{E}_{\tau_{i-1}^n}(\Delta S_{\tilde{\tau}_i^n}) = 0.$$

Then,

$$\varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})\mathbb{P}_{\tau_{i-1}^n}(\Delta S_{\tilde{\tau}_i^n} = \varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})) = \varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})\mathbb{P}_{\tau_{i-1}^n}(\Delta S_{\tilde{\tau}_i^n} = -\varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n}))$$

and then,

$$\mathbb{P}_{\tau_{i-1}^n}(\Delta S_{\tilde{\tau}_i^n} = \varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})) = \mathbb{P}_{\tau_{i-1}^n}(\Delta S_{\tilde{\tau}_i^n} = -\varepsilon_n \lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})) = \frac{1}{2}.$$

We deduce that

$$\mathbb{E}_{\tau_{i-1}^n}(\Delta S_{\tilde{\tau}_i^n})^{2p+1} = 0.$$

Moreover, by definition of $N_\theta^n := \inf\{i : \theta \le \tau_i^n\}$, we have, *a.s.*

$$\begin{aligned}(\Delta S_{\tau_i^n \wedge \theta})^{2p+1} 1_{\tau_{i-1}^n < \theta} &= (\Delta S_{\tilde{\tau}_i^n})^{2p+1} 1_{\tau_{i-1}^n < \theta} + [(\Delta S_\theta)^{2p+1} - (\Delta S_{\tilde{\tau}_i^n})^{2p+1}]1_{\tau_{i-1}^n < \theta, \tilde{\tau}_i^n \ge \theta} \\ &= (\Delta S_{\tilde{\tau}_i^n})^{2p+1} 1_{\tau_{i-1}^n < \theta} + [(\Delta S_\theta)^{2p+1} - (\Delta S_{\tilde{\tau}_{N_\theta^n}^n})^{2p+1}]1_{i=N_\theta^n}.\end{aligned}$$

Then, we have

$$\mathbb{E}_{\tau_{i-1}^n}((\Delta S_{\tau_i^n \wedge \theta})^{2p+1})1_{\tau_{i-1}^n < \theta} = \mathbb{E}_{\tau_{i-1}^n}\Big([(\Delta S_\theta)^{2p+1} - (\Delta S_{\tilde{\tau}_{N_\theta^n}^n})^{2p+1}]1_{i=N_\theta^n}\Big).$$

□

The next lemma utilizes those explicit expressions for the conditional moments :

Lemma 4.5.2. *Let $p \in \mathbb{N}$. Let H be an adapted continuous process, finite* a.s. *on $[0, T]$. Then,*

$$\varepsilon_n^{1-2p} \sum_{\tau_{i-1}^n < T} H_{\tau_{i-1}^n}(\Delta S_{\tau_i^n})^{2p+1} \xrightarrow{a.s.} 0.$$

Proof. For $p = 0$, it is a consequence of Corollary 3.2.1 with $p = 2$. Now, assume $p \ge 1$. Set

$$\begin{aligned} U_t^n &:= \varepsilon_n^{2-4p} \left| \sum_{\tau_{i-1}^n < t} H_{\tau_{i-1}^n}(\Delta S_{\tau_i^n \wedge t})^{2p+1} \right|^2, \\ V_t^n &:= 5\varepsilon_n^2 \sup_{0\le s\le t} |H_s|^2 \sup_{0\le s\le t} \lambda(s, S_s)^{4p} \sum_{\tau_{i-1}^n < t} (\Delta S_{\tau_i^n \wedge t})^2. \end{aligned}$$

We can apply Lemma 3.2.2 since $(U^n)_{n\ge0}$ and $(V^n)_{n\ge0}$ are two sequences of continuous adapted processes and i') V^n is non-decreasing on $[0,T]$ *a.s.* ;

ii') $\sum_{n\geq 0} V_T^n$ is finite *a.s.* , owing to Proposition 3.2.2 with $p=2$.

It remains to check the relation of domination (item *iii')*). Let $\theta_k := \inf\{s \in [0,T] : \bar{V}_s \geq k\}$ setting $\bar{V}_t = \displaystyle\sum_{n\geq 0} V_t^n$. We have

$$\mathbb{E}\left[U^n_{t\wedge\theta_k}\right] = \varepsilon_n^{2-4p}\mathbb{E}\left[\sum_{\tau^n_{i-1}<t\wedge\theta_k} \left|H_{\tau^n_{i-1}}(\Delta S_{\tau^n_i\wedge t\wedge\theta_k})^{2p+1}\right|^2\right]$$
$$+ 2\varepsilon_n^{2-4p}\sum_{i<j}\mathbb{E}\left[H_{\tau^n_{i-1}}(\Delta S_{\tau^n_i\wedge t\wedge\theta_k})^{2p+1}1_{\tau^n_{j-1}<t\wedge\theta_k}H_{\tau^n_{j-1}}\mathbb{E}_{\tau^n_{j-1}}\left((\Delta S_{\tau^n_j\wedge t\wedge\theta_k})^{2p+1}\right)\right].$$

But, we know from Lemma 4.5.2 that

$$\mathbb{E}_{\tau^n_{j-1}}\left((\Delta S_{\tau^n_j\wedge t\wedge\theta_k})^{2p+1}\right) = \mathbb{E}_{\tau^n_{j-1}}\left([(\Delta S_{t\wedge\theta_k})^{2p+1} - (\Delta S_{\tilde{\tau}^n_{N^n_{t\wedge\theta_k}}})^{2p+1}]1_{j=N^n_{t\wedge\theta_k}}\right).$$

Then, using $p\geq 1$,

$$\mathbb{E}\left[U^n_{t\wedge\theta_k}\right] = \varepsilon_n^{2-4p}\mathbb{E}\left[\sum_{\tau^n_{i-1}<t\wedge\theta_k} \left|H_{\tau^n_{i-1}}(\Delta S_{\tau^n_i\wedge t\wedge\theta_k})^{2p+1}\right|^2\right]$$
$$+ 2\varepsilon_n^{2-4p}\mathbb{E}\left[H_{\varphi(\tau^n_{N^n_{t\wedge\theta_k}})}\left[(\Delta S_{t\wedge\theta_k})^{2p+1} - (\Delta S_{\tilde{\tau}^n_{N^n_{t\wedge\theta_k}}})^{2p+1}\right]\sum_{\tau^n_{i-1}<t\wedge\theta_k} H_{\tau^n_{i-1}}(\Delta S_{\tau^n_i\wedge t\wedge\theta_k})^{2p+1}\right]$$
$$\leq 5\varepsilon_n^2\mathbb{E}\left[\sup_{0\leq s\leq t\wedge\theta_k}|H_s|^2\sup_{0\leq s\leq t\wedge\theta_k}\lambda(s,S_s)^{4p}\sum_{\tau^n_{i-1}<t\wedge\theta_k}(\Delta S_{\tau^n_i\wedge t\wedge\theta_k})^2\right]$$
$$= \mathbb{E}[V^n_{t\wedge\theta_k}].$$

Then, $\sum_{n=0}^{+\infty} U_T^n$ is *a.s.* finite and the proof is complete. □

For the sake of simplicity, we shall use the following notation for the iterated integrals : let X be a sequence of continuous Ito semi-martingales, we note

$$\begin{cases}\omega_t^1(X) := \int_{\varphi(t)}^t X_s^1 \mathrm{d}X_s^0,\\ \omega_t^j(X) := \int_{\varphi(t)}^t\int_{\varphi(t)}^{t_0}\cdots\int_{\varphi(t)}^{t_{j-2}} X^j_{t_{j-1}}\mathrm{d}X^{j-1}_{t_{j-1}}\cdots\mathrm{d}X^0_{t_0}.\end{cases}$$

Lemma 4.5.3. *Let X be a sequence of continuous Ito semi-martingales of the form $X_t^k = X_0^k + \int_0^t b_s^k\mathrm{d}s + \int_0^t a_s^k\mathrm{d}B_s$, with continuous adapted coefficients satisfying $\sup_{0\leq t\leq T}(|b_t^k| + |a_t^k|) < +\infty$. For any $\rho > 0$, we have*

$$\sup_n\left(\varepsilon_n^{\rho-k}\sup_{0\leq t\leq T}\omega_t^k(X)\right) < +\infty$$

Proof. Let $\rho > 0$, $\omega_t^1(X) := \int_{\varphi(t)}^t X_s^1 \mathrm{d}X_s^0 \leq C_0 \varepsilon_n^{1-\rho}$ by Corollaries 3.2.2 and 3.2.3, where $X_t^0 := X_0^0 + \int_0^t b_s^0 \mathrm{d}s + \int_0^t a_s^0 \mathrm{d}B_s$.
Let $H_t^n := \omega_t^k((X)_{n\geq 1})$. Assume that $\sup_{n\geq 0}\left(\varepsilon_n^{\rho-k}\sup_{0\leq t\leq T}|H_t^n|\right) < +\infty$ for some $k \geq 1$ and any $\rho > 0$. Then, $\omega_t^{k+1}(X) = \int_{\varphi(t)}^t H_s^n \mathrm{d}X_s^0$. We wish to show that $\sup_{n\geq 0}\left(\varepsilon_n^{\rho-(k+1)}\sup_{0\leq t\leq T}|\omega_t^{k+1}(X)|\right) < +\infty$ and the lemma will follow by induction. Owing to Corollary 3.2.2 (with $\rho = 1$), we have

$$\sup_{n\geq 0}\left(\varepsilon_n^{\rho-(k+1)}\sup_{0\leq t\leq T}\left|\int_{\varphi(t)}^t H_s^n b_s^0 \mathrm{d}s\right|\right) \leq \sup_{n\geq 0}\left(\varepsilon_n^{\rho-k}\sup_{0\leq t\leq T}|H_t^n|\sup_{0\leq t\leq T}|b_t^0|\varepsilon_n^{-1}\sup_{1\leq i\leq N_T^n}\Delta\tau_i^n\right) < +\infty.$$

Let $M_t^n := \int_0^t H_s^n a_s^0 \mathrm{d}B_s$. Let $p > 0$. Set

$$U_t^n := \varepsilon_n^{5-(k+1)p}\sum_{\tau_{i-1}^n < t}\sup_{\tau_{i-1}^n\leq s\leq \tau_i^n\wedge t}|\Delta M_s^n|^p,$$

$$V_t^n := \varepsilon_n^{5-(k+1)p}\sum_{\tau_{i-1}^n < t}\left|\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}(H_s^n a_s^0)^2 \mathrm{d}s\right|^{p/2}.$$

We can apply Lemma 3.2.2 and conclude that $\sum_{n=0}^{+\infty} U_T^n$ is finite *a.s.* since $(U^n)_{n\geq 0}$ and $(V^n)_{n\geq 0}$ are two sequences of continuous adapted processes and

i') V^n is non-decreasing on $[0, T]$ *a.s.* ;

ii')

$$\sum_{n=0}^{+\infty} V_T^n \leq \sup_{0\leq t\leq T}|a_t^0|^p\left(\sup_{n\geq 0}\varepsilon_n^{\rho-k}\sup_{0\leq t\leq T}|H_t^n|\right)^p$$
$$\times \sup_{n\geq 0}(\varepsilon_n^{\rho-2}\Delta\tau_i^n)^{p/2}\sup_{n\geq 0}(\varepsilon_n^{1-3\rho p/2})\sup_{n\geq 0}(\varepsilon_n^2 N_T^n)\sum_{n=0}^{+\infty}\varepsilon_n^2 < +\infty \qquad a.s.,$$

for $\rho = \frac{2}{3p}$;

iii') the domination is directly satisfied thanks to the BDG inequalities.

Clearly and now classically, we obtain $\sup_{n\geq 0}\left(\varepsilon_n^{5-(k+1)p}\sup_{0\leq s\leq T}|\Delta M_s^n|^p\right) < +\infty$ *a.s.* for any $p \geq 0$ and the result follows by taking $p = 5/\rho$. □

Lemma 4.5.4. *Let $\theta \in]0, 1]$ and let H be a θ-Holder function w.r.t. the parabolic distance (see definition in* $(\mathbf{A}_\sigma)$*). Then,* $\varepsilon_n^{-1}\int_0^T\int_{\varphi(t)}^t H(s, S_s)\mathrm{d}S_s \mathrm{d}\langle S\rangle_t \xrightarrow{a.s.} 0$.

Proof. Using the Ito formula, we have

$$\varepsilon_n^{-1}\int_0^T\int_{\varphi(t)}^t H(s, S_s)\mathrm{d}S_s\mathrm{d}\langle S\rangle_t$$
$$= \varepsilon_n^{-1}\sum_{\tau_{i-1}^n < T}\Delta\langle S\rangle_{\tau_i^n}\int_{\tau_{i-1}^n}^{\tau_i^n} H(s, S_s)\mathrm{d}S_s - \varepsilon_n^{-1}\int_0^T \Delta\langle S\rangle_t H(t, S_t)\mathrm{d}S_t.$$

The convergence to 0 of the second term is easy owing to Corollary 3.2.1 and Corollary 3.2.2, for instance. Regarding the first term, we have

$$\begin{aligned}
&\varepsilon_n^{-1} \sum_{\tau_{i-1}^n < T} \Delta\langle S\rangle_{\tau_i^n} \int_{\tau_{i-1}^n}^{\tau_i^n} H(s, S_s)\mathrm{d}S_s \\
&= \varepsilon_n^{-1} \sum_{\tau_{i-1}^n < T} \sigma(\tau_{i-1}^n, S_{\tau_{i-1}^n})^2 S_{\tau_{i-1}^n}^2 H(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \Delta\tau_i^n \Delta S_{\tau_i^n} \\
&\quad + \varepsilon_n^{-1} \sum_{\tau_{i-1}^n < T} \sigma(\tau_{i-1}^n, S_{\tau_{i-1}^n})^2 S_{\tau_{i-1}^n}^2 \Delta\tau_i^n \int_{\tau_{i-1}^n}^{\tau_i^n} \Delta H(s, S_s)\mathrm{d}S_s \\
&\quad + \varepsilon_n^{-1} \sum_{\tau_{i-1}^n < T} \int_{\tau_{i-1}^n}^{\tau_i^n} \Delta((\sigma(t, S_t)S_t)^2)\mathrm{d}t \int_{\tau_{i-1}^n}^{\tau_i^n} H(s, S_s)\mathrm{d}S_s.
\end{aligned}$$

Consequently, the convergence of the first term to 0 is proved with the help of the expression (4.3.13). Each of the two other terms tends to 0 owing to Corollary 3.2.2 and Corollary 3.2.3. □

CHAPTER 5

General almost sure control of semi-martingales

Contents

This chapter is a theoretical extension of Chapter 3 regarding the criterion with a focus on the one dimensional case.

5.1 Introduction

The problem Instead of minimizing $N_T^n\langle Z^n\rangle_T$ over a finite sequence of optimal stopping times $\mathcal{T}^n = \{\tau_0^n = 0 < \tau_1^n < ... < \tau_i^n < ... \leq \tau_{N_T^n}^n = T\}$, we design general costs $\sum_{\tau_{i-1}^n<T} w^Y_{\tau_{i-1}^n}|\Delta Y_{\tau_i^n}|^{\beta^Y}$ and general controls $\sum_{\tau_{i-1}^n<T} w^Z_{\tau_{i-1}^n}|\Delta Z_{\tau_i^n}|^{\beta^Z}$, which are identified to N_T^n and $\langle Z^n\rangle_T$, where $\beta^Y \in (0,2)$, $\beta^Z > 2$ and w^Y, w^Z and Y, Z are respectively two finite processes and two continuous semi-martingales (as opposed to continuous Ito semi-martingales encountered in Chapter 3).
Our Theorem 5.3.1 states that the product

$$\left(\sum_{\tau_{i-1}^n<T} w^Y_{\tau_{i-1}^n}|\Delta Y_{\tau_i^n}|^{\beta^Y}\right)^{1/p^*}\left(\sum_{\tau_{i-1}^n<T} w^Z_{\tau_{i-1}^n}|\Delta Z_{\tau_i^n}|^{\beta^Z}\right)^{1/q^*}$$

has *a.s.* a non-degenerate lower bound over the class of admissible strategies for suitable p^* and q^* (here, the definition of the admissible strategies differs sparsely from Chapter 3, see Section 5.2.2). In addition, in Theorems 5.3.2, we show the existence of a strategy of the hitting time form attaining the *a.s.* lower bound.

Literature background To our best knowledge, the first author dealing with this kind of criterion was Masaaki Fukasawa in [Fukasawa 2011b]. Extension to jump processes has recently been done in [Rosenbaum 2011]). We refer to the introduction of Chapter 3 for the advantages of our approach.

Outline of the chapter In the following, we present some notations and assumptions that will be used throughout the chapter. Section 5.2 is aimed at defining our class of stopping time strategies and deriving some general theoretical properties in this class. The main theorems are stated and proved in Section 5.3. All these results are not specifically related to financial applications. Section 5.4 is a new approach of the hedging issue under transaction costs. Some numerical experiments are postponed to section 5.5.

5.2 Model

To keep this chapter short, we list new bespoke notations.

5.2.1 Notations

- Let $(\alpha_n)_{n\geq 0}, (\beta_n)_{n\geq 0}$ be two sequences of random variables. We write $\alpha_n = O(\beta_n)$ (resp. $o(\beta_n)$), if $\sup_{n\geq 0}(|\beta_n^{-1}\alpha_n|) < +\infty$ *a.s.* (resp. $|\beta_n^{-1}\alpha_n| \xrightarrow{a.s.} 0$).

- For any càdlag process X, we define $|X|_* := \sup_{0\leq t\leq T}|X_t|$ and $|\Delta X|_* := \sup_{1\leq i\leq N_T^n}\sup_{\tau_{i-1}^n\leq t\leq \tau_i^n}|\Delta X_t|$, where $\Delta X_t := X_t - X_{\varphi(t)}$ as in Chapter 3. So, we say that X is finite *a.s.* if, and only if, $|X|_* < +\infty$ *a.s.* .

- X is a scalar Itô process $dX_t = b_t^X \mathrm{d}t + \sigma_t^X \mathrm{d}B_t$, called the control process, with σ^X satisfying the condition $(\mathbf{A}_\sigma^{\text{Ellip.}})$ in Chapter 3 with $\theta_\sigma \equiv \theta_X$. Take $b^X \equiv 0$ to simplify, which holds under any equivalent measure and in particular that which makes $B_t + \int_0^t (\sigma_s^X)^{-1} b_s \mathrm{d}s$ a Brownian motion (whenever this measure is well defined).

- A^Y and A^Z are two continuous adapted processes with finite variation satisfying $|\Delta A^Y|_* + |\Delta A^Z|_* = O(\varepsilon_n)$.

- Y, Z are processes controlled by X : $\mathrm{d}Y_t = b_t^Y \mathrm{d}A_t^Y + \sigma_t^Y \mathrm{d}X_t$, $\mathrm{d}Z_t = b_t^Z \mathrm{d}A_t^Z + \sigma_t^Z \mathrm{d}X_t$ with $b^Y, b^Z, \sigma^Y, \sigma^Z$ are continuous adapted processes where $\sigma^Y > 0$, $\sigma^Z > 0$, $|\Delta\sigma^Y|_* = O(\varepsilon_n^{\theta_Y})$ and $|\Delta\sigma^Z|_* = O(\varepsilon_n^{\theta_Z})$, with $\theta_Y, \theta_Z > 0$.

5.2.2 Class of stopping times

Let ε_n be a non-negative sequence such that $\sum_{n\geq 0} \varepsilon_n^2 < +\infty$.
We consider $\mathcal{T}^n := \{\tau_0^n < \tau_1^n < ... < \tau_{N_T^n}^n = T\}$ as a sequence of length N_T^n of stopping times. Instead of imposing some controls over the processes Y or Z, we choose to control the process X (which generalizes the assumption $(\mathbf{A}_S)$ in Chapter 3).
$(\mathbf{A}_X)$:

$$|\Delta X|_* := \sup_{1\leq i\leq N_T^n} \sup_{\tau_{i-1}^n \leq t \leq \tau_i^n} |X_t - X_{\tau_{i-1}^n}| = O(\varepsilon_n).$$

As in Chapter 3, we decide to bound the number of stopping times; that enables us to give sharp estimations of several process increments as we shall see later on.
$(\mathbf{A}_N)$: For $1 \leq \rho_N < (1 + \frac{\theta_X}{2}) \wedge \frac{4}{3}$,

$$N_T^n = O(\varepsilon_n^{-2\rho_N}).$$

Definition 5.2.1. *We say that $\mathcal{T} := \{\mathcal{T}^n; n \geq 0\}$ is an admissible class of strategies if the sequence of stopping times fulfils the assumptions $(\mathbf{A}_X)$ and $(\mathbf{A}_N)$.*

5.2.3 Almost sure convergence of semi-martingales

In this section, we provide a crucial and tailor-made result on convergence for general semi-martingales, easy to demonstrate with the help of Chapter 3.

Proposition 5.2.1. *Let $\mathcal{T} = (\mathcal{T}^n)_{n\geq 0}$ be an admissible class, $((M_t^n)_{0\leq t\leq T})_{n\geq 0}$ be a sequence of $\mathbb{R}$-valued continuous local martingales such that $|\Delta\langle M^n\rangle|_* = O(\varepsilon_n^{2(1+\theta_M)})$ for some $\theta_M \geq 0$ and A^n be a sequence of finite variation processes satisfying $|\Delta A^n|_* = O(\varepsilon_n^{\theta_A})$. Then, noting S^n for the semi-martingale $M^n + A^n$,* a.s. *, for any $\rho > 0$,*

$$|\Delta S^n|_* = O(\varepsilon_n^{(1+\theta_M-\rho)\wedge\theta_A}).$$

In particular, $|\Delta Y|_* = O(\varepsilon_n^{1-\rho})$ and $|\Delta Z|_* = O(\varepsilon_n^{1-\rho})$ for any $\rho > 0$.

Proof. We know bythe proof of Lemma 4.5.3 in Chapter 4 that $|\Delta M^n|_* = O(\varepsilon_n^{1+\theta_M-\rho})$. Moreover, from the rough estimation $|\Delta S^n|_* \leq |\Delta M^n|_* + |\Delta A^n|_*$, we get $|\Delta S^n|_* = O(\varepsilon_n^{(1+\theta_M-\rho)\wedge\theta_A})$. □

5.3 Main results

One of the strengths of theorems of Chapter 3 is the multidimensional setting. Indeed, few results exist in this case; most of the time, authors tackle the problem in dimension one. Obviously, proofs are easier, because we do not mind about matrix calculus, but also we can conceive broad controls over our desired process. The first result gives a lower bound for a generic criterion in an almost sure sense; this is the content of the theorem below :

Theorem 5.3.1. *Let $\beta^Y \in (0,2)$ and $\beta^Z \in (2,+\infty)$. Set $p^* = \frac{\beta^Z-\beta^Y}{\beta^Z-2}$ and $q^* = \frac{\beta^Z-\beta^Y}{2-\beta^Y}$. Let $\mathcal{T}$ be an admissible class of sequences of stopping times with $\rho_N <$*
$$1+\frac{1}{2}\begin{cases} \frac{\beta^Y}{p^*-\beta^Y} \wedge \frac{\theta_Y\beta^Y}{p^*} \wedge \frac{1}{q^*-1} \wedge \frac{\theta_Z}{q^*} \text{ if } \beta^Y \in (0,1) \\ \frac{1}{q^*-1} \wedge \frac{1}{p^*-1} \wedge \frac{\theta_Y}{p^*} \wedge \frac{\theta_Z}{q^*} \text{ if } \beta^Y \in [1,2) \end{cases}.$$
We define $\mathcal{Y}^n_T := \sum_{\tau^n_{i-1}<T} w^Y_{\tau^n_{i-1}} |\Delta Y_{\tau^n_i}|^{\beta^Y}$ and $\mathcal{Z}^n_T := \sum_{\tau^n_{i-1}<T} w^Z_{\tau^n_{i-1}} |\Delta Z_{\tau^n_i}|^{\beta^Z}$ with w^Y and w^Z being two non-negative continuous adapted processes. Then,

$$\liminf_{n\to+\infty} (\mathcal{Y}^n_T)^{1/p^*}(\mathcal{Z}^n_T)^{1/q^*} \geq \int_0^T \mathcal{L}_t \mathrm{d}t, \quad \text{a.s..}$$

where $\mathcal{L}_t := (\sigma^X_t)^2 (w^Y_t(\sigma^Y_t)^{\beta^Y})^{1/p^*}(w^Z_t(\sigma^Z_t)^{\beta^Z})^{1/q^*}$. Moreover, for any other conjugate powers p and q, the limit inf is, in general, trivial (0 or $+\infty$), it is precisely the case for $X = Y = Z$.

Remark 5.3.1.
- *We do not deal with the case $\beta^Y = 0$ because we have already treated it in Chapter 3 and in our setting, we would need to put apart this case.*
- *The cases $\beta^Y = 2$ or $\beta^Z = 2$ are uninteresting regarding the optimization of $\mathcal{T}$. Indeed, considering the convergence of the quadratic variation from Proposition 3.2.3, we easily get (if $\beta^Y = 2$) $\mathcal{Y}^n_T \xrightarrow{a.s.} \int_0^T w^Y_s \mathrm{d}\langle Y\rangle_s$, regardless the sequence $(\mathcal{T}^n)_{n\geq 0}$ and similarly for $\mathcal{Z}^n_T$ if $\beta^Z = 2$. Hence, in that case, there is no expected trade-off between $\mathcal{Y}^n_T$ and $\mathcal{Z}^n_T$ at the limit and we shall discard these situations.*

We now provide an optimal admissible class, that is the sequence of stopping times such that our criterion converges *a.s.* to the above lower bound.

Theorem 5.3.2. *Assume the assumptions of Theorem 5.3.1 and let $\mu > 0$, for $t \geq 0$, $\mathcal{L}^\mu_t := \mathcal{L}_t + \mu\chi_\mu\Big(\mathcal{L}_t\Big)$, where χ_μ is defined in Theorem 3.3.2.*

For a given $n \in \mathbb{N}$, define the strategy $\mathcal{T}^n_\mu$ by

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n = \inf\left\{ t \geq \tau_{i-1}^n : |X_t - X_{\tau_{i-1}^n}| > \varepsilon_n \left(\frac{(\sigma^X_{\tau_{i-1}^n})^2 w^Y_{\tau_{i-1}^n} (\sigma^Y_{\tau_{i-1}^n})^{\beta^Y}}{\mathcal{L}^\mu_{\tau_{i-1}^n}} \right)^{\frac{1}{2-\beta^Y}} \right\} \wedge T. \end{cases} \tag{5.3.1}$$

Then, the class $\mathcal{T}_\mu = \{\mathcal{T}^n_\mu : n \geq 0\}$ is admissible and μ-asymptotically optimal in the following sense:

$$\limsup_{n\to+\infty} \left| (\mathcal{Y}_T^n)^{1/p^*} (\mathcal{Z}_T^n)^{1/q^*} - \int_0^T \mathcal{L}_t^\mu \mathrm{d}t \right| = O\left(\left(\mu \int_0^T \mathcal{L}_t^\mu \mathrm{d}t\right)^{1/p^*} + \left(\mu \int_0^T \mathcal{L}_t^\mu \mathrm{d}t\right)^{1/q^*} \right).$$

In particular, on the event $\{\forall t \in [0,T] : \mathcal{L}_t \geq \mu\}$, $(\mathcal{Y}_T^n)^{1/p^*}(\mathcal{Z}_T^n)^{1/q^*}$ converges a.s. *to $\int_0^T \mathcal{L}_t \mathrm{d}t$.*

We can switch between $\sum_{\tau_{i-1}^n < T} w^Z_{\tau_{i-1}^n} |\Delta Z_{\tau_i^n}|^{\beta^Z}$ and the integral form $\frac{\beta^Z(\beta^Z-1)}{2} \int_0^T w_s^Z |\Delta Z_s|^{\beta^Z - 2} \mathrm{d}\langle Z\rangle_s$, in the previous theorems without changing the conclusion. This follows from the next proposition.

Proposition 5.3.1. *Under the notations and assumptions of Theorem 5.3.1, let $\mathcal{T}$ be an admissible class of sequences of stopping times with $\rho_N < 1 + \frac{\theta_Z^w}{2(q^*-1)}$ for some $\theta_Z^w \in (0,1]$. We consider $\mathcal{Z}_T^n := \sum_{\tau_{i-1}^n < T} w^Z_{\tau_{i-1}^n} |\Delta Z_{\tau_i^n}|^{\beta^Z}$, where w^Z is a non-negative continuous adapted process satisfying $|\Delta w^Z|_* = O(\varepsilon_n^{\theta_Z^w})$. Then,*

$$\lim_{n\to+\infty} \varepsilon_n^{(\beta^Y - 2\rho_N)(q^*-1)} \left| \mathcal{Z}_T^n - \frac{\beta^Z(\beta^Z-1)}{2} \int_0^T w_s^Z |\Delta Z_s|^{\beta^Z-2} (\sigma_s^Z)^2 \mathrm{d}s \right| = 0 \quad \text{a.s..}$$

Proof. From Itô's lemma applied to the twice continuously differentiable function $x \mapsto |x|^{\beta^Z}$, we have

$$\begin{aligned} &\mathcal{Z}_T^n - \frac{\beta^Z(\beta^Z-1)}{2} \int_0^T w_s^Z |\Delta Z_s|^{\beta^Z-2} (\sigma_s^Z)^2 \mathrm{d}s \\ &= \beta^Z \int_0^T w^Z_{\varphi(s)} sgn(\Delta Z_s) |\Delta Z_s|^{\beta^Z-1} \mathrm{d}Z_s - \frac{\beta^Z(\beta^Z-1)}{2} \int_0^T \Delta w_s^Z |\Delta Z_s|^{\beta^Z-2} (\sigma_s^Z)^2 \mathrm{d}s. \end{aligned}$$

For the first term, in view of Corollary 3.2.1, we just have to show that $\sum_{n\geq 0} \varepsilon_n^{p(\beta^Y - 2\rho_N)(q^*-1)} \left\langle \beta^Z \int_0^\cdot w^Z_{\varphi(s)} sgn(\Delta Z_s)|\Delta Z_s|^{\beta^Z-1} \mathrm{d}Z_s \right\rangle_T^{p/2}$ is finite for some $p > 0$ *a.s.* . But, by Proposition 5.2.1, $|\Delta Z|_*^{\beta^Z-1} = O(\varepsilon_n^{\beta^Z-1-\rho})$ for any $\rho > 0$. Thus the above series converges *a.s.* , if $p[(\beta^Y - 2\rho_N)(q^*-1) + \beta^Z - 1 - \rho] \geq 2$: by taking p large enough and ρ small enough it is sufficient to check that $(\beta^Y - 2\rho_N)(q^*-1) + \beta^Z - 1 > 0$. The second term is $O(\varepsilon_n^{\theta_Z^w + \beta^Z - 2 - \rho})$ and thus, it gives a negligible contribution provided that $0 < (\beta^Y - 2\rho_N)(q^*-1) + \theta_Z^w + \beta^Z - 2 = \theta_Z^w + 2(1-\rho_N)(q^*-1)$, and the assertion is proved. □

5.3.1 Proof of Theorem 5.3.1

Step 1: Decomposition of $\mathcal{Y}_T^n$ and of $\mathcal{Z}_T^n$

A simple idea consists in approximating the increments of the semi-martingales Y and Z by the increments of their local martingale components and showing that the residual terms (i.e. the increments of their finite variation parts) tend to 0 fast enough :

$$\sum_{\tau_{i-1}^n<T} w^Y_{\tau_{i-1}^n}|\Delta Y_{\tau_i^n}|^{\beta^Y} = \sum_{\tau_{i-1}^n<T} w^Y_{\tau_{i-1}^n}|\sigma^Y_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Y} + \mathcal{E}^n_{Y,1,T}, \tag{5.3.2}$$

where

$$\mathcal{E}^n_{Y,1,T} := \sum_{\tau_{i-1}^n<T} w^Y_{\tau_{i-1}^n}\Big(|\Delta Y_{\tau_i^n}|^{\beta^Y} - |\sigma^Y_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Y}\Big), \tag{5.3.3}$$

and as well,

$$\sum_{\tau_{i-1}^n<T} w^Z_{\tau_{i-1}^n}|\Delta Z_{\tau_i^n}|^{\beta^Z} = \sum_{\tau_{i-1}^n<T} w^Z_{\tau_{i-1}^n}|\sigma^Z_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Z} + \mathcal{E}^n_{Z,1,T}, \tag{5.3.4}$$

where

$$\mathcal{E}^n_{Z,1,T} := \sum_{\tau_{i-1}^n<T} w^Z_{\tau_{i-1}^n}\Big(|\Delta Z_{\tau_i^n}|^{\beta^Z} - |\sigma^Z_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Z}\Big). \tag{5.3.5}$$

Step 2: lower bound's proof

The aim of this step is to provide a simple demonstration of the lower bound stated in Theorem 5.3.1. The Holder inequality yields straightforwardly to a lower bound for the product of the two main terms of the equations (5.3.2) and (5.3.4), that is

$$\begin{aligned}&\Bigg(\sum_{\tau_{i-1}^n<T} w^Y_{\tau_{i-1}^n}|\sigma^Y_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Y}\Bigg)^{1/p}\Bigg(\sum_{\tau_{i-1}^n<T} w^Z_{\tau_{i-1}^n}|\sigma^Z_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Z}\Bigg)^{1/q}\\ &\geq \sum_{\tau_{i-1}^n<T}(w^Y_{\tau_{i-1}^n}(\sigma^Y_{\tau_{i-1}^n})^{\beta^Y})^{1/p}(w^Z_{\tau_{i-1}^n}(\sigma^Z_{\tau_{i-1}^n})^{\beta^Z})^{1/q}|\Delta X_{\tau_i^n}|^{\beta^Y/p+\beta^Z/q}.\end{aligned} \tag{5.3.6}$$

Remark 5.3.2. *This inequality remains true for $\mathcal{Y}_T^n$ and $\mathcal{Z}_T^n$ (i.e. taking into account the negligible terms) when $X = Y = Z$, because in this case $\mathcal{E}^n_{Y,1,T} = \mathcal{E}^n_{Z,1,T} = 0$.*

Then, for $\beta^Y/p^* + \beta^Z/q^* = 2$, we obtain a non trivial lower bound. That corresponds to $p^* = \frac{\beta^Z - \beta^Y}{\beta^Z - 2}$ and $q^* = \frac{\beta^Z - \beta^Y}{2 - \beta^Y}$ and then,

$$\left(\sum_{\tau_{i-1}^n < T} w^Y_{\tau_{i-1}^n} |\sigma^Y_{\tau_{i-1}^n} \Delta X_{\tau_i^n}|^{\beta^Y} \right)^{1/p^*} \left(\sum_{\tau_{i-1}^n < T} w^Z_{\tau_{i-1}^n} |\sigma^Z_{\tau_{i-1}^n} \Delta X_{\tau_i^n}|^{\beta^Z} \right)^{1/q^*}$$
$$\geq \sum_{\tau_{i-1}^n < T} \mathcal{L}_{\tau_{i-1}^n} (\sigma^X_{\tau_{i-1}^n})^{-2} (\Delta X_{\tau_i^n})^2 \xrightarrow{a.s.} \int_0^T \mathcal{L}_t \mathrm{d}t,$$

where $\mathcal{L}_t := (\sigma_t^X)^2 (w_t^Y (\sigma_t^Y)^{\beta^Y})^{1/p^*} (w_t^Z (\sigma_t^Z)^{\beta^Z})^{1/q^*}$.
Now, for $\beta^Y/p + \beta^Z/q < 2$ which corresponds to the cases $p < p^*$ and $q > q^*$, from (5.3.6) and $|\Delta X|_* \xrightarrow{a.s.} 0$, we deduce

$$\left(\sum_{\tau_{i-1}^n < T} w^Y_{\tau_{i-1}^n} |\sigma^Y_{\tau_{i-1}^n} \Delta X_{\tau_i^n}|^{\beta^Y} \right)^{1/p} \left(\sum_{\tau_{i-1}^n < T} w^Z_{\tau_{i-1}^n} |\sigma^Z_{\tau_{i-1}^n} \Delta X_{\tau_i^n}|^{\beta^Z} \right)^{1/q} \xrightarrow{a.s.} +\infty.$$

Last, for $\beta^Y/p + \beta^Z/q > 2$ which corresponds to the cases $p > p^*$ and $q < q^*$, we extract a sequence of non-optimal stopping times in the form of hitting times $\hat{\tau}_i^n := \inf\{t \geq \hat{\tau}_{i-1}^n : |\Delta X_t| = \varepsilon_n\}$ (which are embedded into an admissible class of sequences of stopping times with $\rho_N = 1$)

$$\left(\sum_{\tau_{i-1}^n < T} w^Y_{\hat{\tau}_{i-1}^n} |\sigma^Y_{\hat{\tau}_{i-1}^n} \Delta X_{\hat{\tau}_i^n}|^{\beta^Y} \right)^{1/p} \left(\sum_{\tau_{i-1}^n < T} w^Z_{\hat{\tau}_{i-1}^n} |\sigma^Z_{\hat{\tau}_{i-1}^n} \Delta X_{\hat{\tau}_i^n}|^{\beta^Z} \right)^{1/q} = O(\varepsilon_n^{\beta^Y/p + \beta^Z/q - 2}).$$

Then, in particular,

$$\liminf_{n \to +\infty} \left(\sum_{\tau_{i-1}^n < T} w^Y_{\tau_{i-1}^n} |\sigma^Y_{\tau_{i-1}^n} \Delta X_{\tau_i^n}|^{\beta^Y} \right)^{1/p} \left(\sum_{\tau_{i-1}^n < T} w^Z_{\tau_{i-1}^n} |\sigma^Z_{\tau_{i-1}^n} \Delta X_{\tau_i^n}|^{\beta^Z} \right)^{1/q} = 0.$$

Therefore, we have proved that the right conjugate powers for a non trivial limit are exactly (p^*, q^*). □

Step 3: the renormalized errors $\varepsilon_n^{(\beta^Z - 2\rho_N)(p^*-1)} \mathcal{E}^n_{Y,1,T}$ and $\varepsilon_n^{(\beta^Y - 2\rho_N)(q^*-1)} \mathcal{E}^n_{Z,1,T}$ converge to 0 *a.s.*

Assume the above convergences, then in view of (5.3.2), (5.3.4) and Step 2, we easily complete the proof of Theorem 5.3.1.
• *Proof of* $\varepsilon_n^{(\beta^Z - 2\rho_N)(p^*-1)} \mathcal{E}^n_{Y,1,T} \xrightarrow{a.s.} 0$. We actually distinguish two cases :

- $\beta^Y \geq 1$: use Taylor's theorem applied to the function $x \mapsto x^{\beta^Y}$ to get

$$
\begin{aligned}
&\varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)}|\mathcal{E}^n_{Y,1,T}| \\
&\leq \varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)} \sum_{\tau^n_{i-1}<T} w^Y_{\tau^n_{i-1}} \beta^Y \left(|\Delta Y_{\tau^n_i}|^{\beta^Y-1} \vee |\sigma^Y_{\tau^n_i}\Delta X_{\tau^n_i}|^{\beta^Y-1}\right)\left|\Delta Y_{\tau^n_i} - \sigma^Y_{\tau^n_i}\Delta X_{\tau^n_i}\right| \\
&\leq \varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)} \beta^Y |w^Y|_* \left(|\Delta Y|_*^{\beta^Y-1} \vee |\sigma^Y_{\varphi(.)}\Delta X|_*^{\beta^Y-1}\right) \\
&\quad \times \left(N^n_T \Big|\int_{\varphi(.)}^{.} \Delta\sigma^Y_s \mathrm{d}X_s\Big|_* + |b^Y|_* \sum_{\tau^n_{i-1}<T} |\Delta A^Y_{\tau^n_i}|\right).
\end{aligned}
$$

Owing to Proposition 5.2.1, the first term is $O\left(\varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)+\beta^Y-1-\rho-2\rho_N+1+\theta_Y-\rho}\right)$, for any $\rho > 0$. It converges *a.s.* to 0 as soon as

$$
\begin{aligned}
0 &< (\beta^Z - 2\rho_N)(p^*-1) + \beta^Y - 2\rho_N + \theta_Y \\
&= (\beta^Z - 2)(p^*-1) + \beta^Y - 2 + 2(1-\rho_N)(p^*-1) + \theta_Y + 2(1-\rho_N) \\
&= 2(1-\rho_N)p^* + \theta_Y,
\end{aligned}
$$

which holds by taking $\rho_N < 1 + \frac{\theta_Y}{2p^*}$. Similarly, the second term converges *a.s.* to 0 provided that $0 < (\beta^Z - 2\rho_N)(p^*-1) + \beta^Y - 1 = 2(1-\rho_N)(p^*-1) + 1$, i.e. $\rho_N < 1 + \frac{1}{2(p^*-1)}$.

- $\beta^Y \in (0,1)$: using that for $(a,b) \in \mathbb{R}^2_+$ and $p \in [0,1]$, $|b^p - a^p| \leq |b-a|^p$, we have

$$
\begin{aligned}
&\varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)}|\mathcal{E}^n_{Y,1,T}| \\
&\leq \varepsilon_n^{(\beta^Y-2\rho_N)(q^*-1)} \sum_{\tau^n_{i-1}<T} w^Y_{\tau^n_{i-1}} \left|\Delta Y_{\tau^n_i} - \sigma^Y_{\tau^n_i}\Delta X_{\tau^n_i}\right|^{\beta^Y} \\
&= \varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)} \sum_{\tau^n_{i-1}<T} w^Y_{\tau^n_{i-1}} \left|\int_{\tau^n_{i-1}}^{\tau^n_i} \Delta\sigma^Y_t \mathrm{d}X_t + \int_{\tau^n_{i-1}}^{\tau^n_i} b^Y_t \mathrm{d}A^Y_t\right|^{\beta^Y} \\
&\leq \varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)} |w^Y|_* \left(N^n_T \Big|\int_{\varphi(.)}^{.} \Delta\sigma^Y_s \mathrm{d}X_s\Big|_*^{\beta^Y} + |b^Y|_* \sum_{\tau^n_{i-1}<T} |\Delta A^Y_{\tau^n_i}|^{\beta^Y}\right).
\end{aligned}
$$

The first term is $O(\varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)-2\rho_N+(1+\theta_Y)\beta^Y-\rho})$, for any $\rho > 0$, thus it converges to 0 provided $0 < (\beta^Z - 2\rho_N)(p^*-1) - 2\rho_N + (1+\theta_Y)\beta^Y = 2(1-\rho_N)(p^*-1) + 2(1-\rho_N) + \theta_Y\beta^Y = 2(1-\rho_N)p^* + \theta_Y\beta^Y$, which holds under our assumptions. For the second term, we use Holder's inequality

$$
\begin{aligned}
|w^Y|_*|b^Y|_*\varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)} & \sum_{\tau_{i-1}^n<T} |\Delta A^Y_{\tau_i^n}|^{\beta^Y} \\
& \leq |w^Y|_*|b^Y|_*\varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)} \left(\sum_{\tau_{i-1}^n<T} |\Delta A^Y_{\tau_i^n}|\right)^{\beta^Y} N_T^{n\,1-\beta^Y} \\
& = O(\varepsilon_n^{(\beta^Z-2\rho_N)(p^*-1)-2\rho_N(1-\beta^Y)}).
\end{aligned}
$$

It converges to 0 since the exponent of ε_n is equal to $2(1-\rho_N)(p^*-1)-2\rho_N(1-\beta^Y)+2-\beta^Y = 2(1-\rho_N)(p^*-\beta_Y)+\beta^Y>0$.

• *Proof of* $\varepsilon_n^{(\beta^Y-2\rho_N)(q^*-1)}\mathcal{E}^n_{Z,1,T} \xrightarrow{a.s.} 0$. The computations are identical to those for $\mathcal{E}^n_{Y,1,T}$, when $\beta^Y\geq 1$, we skip details. □

5.3.2 Proof of Theorem 5.3.2

Firstly, we check the admissibility of $\mathcal{T}_\mu$: the verification of the assumption $(\mathbf{A}_X)$ is easy. For the assumption $(\mathbf{A}_N)$, we point out that

$$
\varepsilon_n^{2\rho_N}N_T^n = \varepsilon_n^{2\rho_N}+\varepsilon_n^{2(\rho_N-1)}\sum_{1\leq i\leq N_T^n-1}\left(\frac{\mathcal{L}^\mu_{\tau_{i-1}^n}}{(\sigma^X_{\tau_{i-1}^n})^2 w^Y_{\tau_{i-1}^n}(\sigma^Y_{\tau_{i-1}^n})^{\beta^Y}}\right)^{\frac{2}{2-\beta^Y}}(\Delta X_{\tau_{i-1}^n})^2 = O(\varepsilon_n^{2(\rho_N-1)}).
$$

Now, let us show the μ-optimality. Starting from the decomposition (5.3.2) of $\mathcal{Y}_T^n$ and in view of the definition of the strategy $\mathcal{T}_\mu^n$,

$$
\begin{aligned}
\varepsilon_n^{2-\beta^Y}\mathcal{Y}_T^n &= \sum_{1\leq i\leq N_T^n-1}\mathcal{L}^\mu_{\tau_{i-1}^n}(\sigma^X_{\tau_{i-1}^n})^{-2}(\Delta X_{\tau_i^n})^2+\varepsilon_n^{2-\beta^Y}\left(\mathcal{E}^n_{Y,1,T}+\mathcal{E}^n_{Y,2,T}+\mathcal{E}^n_{Y,3,T}\right), \\
\mathcal{E}^n_{Y,2,T} &:= \sum_{1\leq i\leq N_T^n-1}\Big[w^Y_{\tau_{i-1}^n}|\sigma^Y_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Y}-\varepsilon_n^{\beta^Y-2}\mathcal{L}^\mu_{\tau_{i-1}^n}(\sigma^X_{\tau_{i-1}^n})^{-2}(\Delta X_{\tau_i^n})^2\Big], \\
\mathcal{E}^n_{Y,3,T} &:= w^Y_{N_T^n-1}|\sigma^Y_{N_T^n-1}\Delta X_T|^{\beta^Y}.
\end{aligned}
$$

As well, from the decomposition (5.3.4) of $\mathcal{Z}_T^n$,

$$
\begin{aligned}
\varepsilon_n^{2-\beta^Z}\mathcal{Z}_T^n &= \sum_{1\leq i\leq N_T^n-1}\mathcal{L}^\mu_{\tau_{i-1}^n}(\sigma^X_{\tau_{i-1}^n})^{-2}(\Delta X_{\tau_i^n})^2+\varepsilon_n^{2-\beta^Z}\left(\mathcal{E}^n_{Z,1,T}+\mathcal{E}^n_{Z,2,T}+\mathcal{E}^n_{Z,3,T}\right), \\
\mathcal{E}^n_{Z,2,T} &:= \sum_{1\leq i\leq N_T^n-1}\Big[w^Z_{\tau_{i-1}^n}|\sigma^Z_{\tau_{i-1}^n}\Delta X_{\tau_i^n}|^{\beta^Z}-\varepsilon_n^{\beta^Z-2}\mathcal{L}^\mu_{\tau_{i-1}^n}(\sigma^X_{\tau_{i-1}^n})^{-2}(\Delta X_{\tau_i^n})^2\Big], \\
\mathcal{E}^n_{Z,3,T} &:= w^Z_{N_T^n-1}|\sigma^Z_{N_T^n-1}\Delta X_T|^{\beta^Z}.
\end{aligned}
$$

So far, we know that $\sum_{1\leq i\leq N_T^n-1}\mathcal{L}^\mu_{\tau_{i-1}^n}(\sigma^X_{\tau_{i-1}^n})^{-2}(\Delta X_{\tau_i^n})^2 \xrightarrow{a.s.} \int_0^T\mathcal{L}_t^\mu \mathrm{d}t$ (see Proposition 3.2.3). Furthermore, we have already established (see Step 3 of proof of

Theorem 5.3.1) that $\mathcal{E}^n_{Y,1,T} = o(\varepsilon_n^{-(\beta^Z-2\rho_N)(p^*-1)}) = o(\varepsilon_n^{\beta^Y-2})$ (using $\rho_N \geq 1$) and $\mathcal{E}^n_{Z,1,T} = o(\varepsilon_n^{-(\beta^Y-2\rho_N)(q^*-1)}) = o(\varepsilon_n^{\beta^Z-2})$ for any admissible sequence of strategies. Moreover, $\mathcal{E}^n_{Y,3,T} = O(\varepsilon_n^{\beta^Y}) = o(\varepsilon_n^{\beta^Y-2})$ and $\mathcal{E}^n_{Z,3,T} = O(\varepsilon_n^{\beta^Z}) = o(\varepsilon_n^{\beta^Z-2})$.
Finally, regarding $\mathcal{E}^n_{Z,2,T}$, we obtain that $|\varepsilon_n^{2-\beta^Z}\mathcal{E}^n_{Z,2,T}|$ is bounded by

$$\begin{aligned}
&\varepsilon_n^{2-\beta^Z} \sum_{1\leq i\leq N^n_T-1} \left| w^Z_{\tau^n_{i-1}} |\sigma^Z_{\tau^n_{i-1}} \Delta X_{\tau^n_i}|^{\beta^Z} - \varepsilon_n^{\beta^Z-2}\mathcal{L}^\mu_{\tau^n_{i-1}}(\sigma^X_{\tau^n_{i-1}})^{-2}(\Delta X_{\tau^n_i})^2 \right| \\
&= \sum_{1\leq i\leq N^n_T-1} \frac{|\mathcal{L}^{q^*}_{\tau^n_{i-1}} - (\mathcal{L}^\mu_{\tau^n_{i-1}})^{q^*}|}{(\mathcal{L}^\mu_{\tau^n_{i-1}})^{q^*-1}} (\sigma^X_{\tau^n_{i-1}})^{-2}(\Delta X_{\tau^n_i})^2 \\
&\leq \sum_{1\leq i\leq N^n_T-1} q^*\mu\chi_\mu(\mathcal{L}_{\tau^n_{i-1}})(\sigma^X_{\tau^n_{i-1}})^{-2}(\Delta X_{\tau^n_i})^2,
\end{aligned}$$

using Taylor's theorem applied to the function $x \mapsto x^{q^*}$. Thus,

$$\limsup_{n\to+\infty} |\varepsilon_n^{2-\beta^Z}\mathcal{E}^n_{Z,2,T}| \leq q^*\mu \int_0^T \chi_\mu(\mathcal{L}_t)\mathrm{d}t, \quad a.s..$$

Moreover, we have

$$\mathcal{E}^n_{Y,2,T} = 0.$$

Let us summarize: setting $L_T := \int_0^T \mathcal{L}_t \mathrm{d}t$ and $L^\mu_T := \int_0^T \chi_\mu(\mathcal{L}_t)\mathrm{d}t$ so that $\int_0^T \mathcal{L}^\mu_t \mathrm{d}t = L_T + \mu L^\mu_T$, we have shown

$$\begin{aligned}
\limsup_{n\to+\infty} \left|\varepsilon_n^{2-\beta^Z}\mathcal{Z}^n_T - (L_T + \mu L^\mu_T)\right| &\leq q^*\mu L^\mu_T, \\
\varepsilon_n^{2-\beta^Y}\mathcal{Y}^n_T &\xrightarrow{a.s.} L_T + \mu L^\mu_T.
\end{aligned}$$

Then, using that for $(a,b) \in \mathbb{R}^2_+$ and $p \in [0,1]$, $|b^p - a^p| \leq |b-a|^p$, we conclude that

$$\begin{aligned}
&\limsup_{n\to+\infty} \left|(\mathcal{Y}^n_T)^{1/p^*}(\mathcal{Z}^n_T)^{1/q^*} - L_T\right| \\
&\leq \limsup_{n\to+\infty} |\varepsilon_n^{2-\beta^Z}\mathcal{Z}^n_T - L_T|^{1/q^*} \lim_{n\to+\infty}\left(\varepsilon_n^{2-\beta^Y}\mathcal{Y}^n_T\right)^{1/p^*} + L_T^{1/q^*} \lim_{n\to+\infty}\left|\varepsilon_n^{2-\beta^Y}\mathcal{Y}^n_T - L_T\right|^{1/p^*} \\
&\leq \left[(1+q^*)\mu L^\mu_T\right]^{1/q^*}(L_T + \mu L^\mu_T)^{1/p^*} + L_T^{1/q^*}(\mu L^\mu_T)^{1/p^*} = O\left((\mu L^\mu_T)^{1/p^*} + (\mu L^\mu_T)^{1/q^*}\right).
\end{aligned}$$

Theorem 5.3.2 is proved. □

5.4 Back to the hedging problem under transaction costs

In this section, we choose another criterion, different from Chapter 4 to tackle the hedging problem under transaction costs. The main object of study is the tracking error :

$$Z_T^n = g(S_T) - \Bigg(u(0,S_0) + \sum_{\tau_{i-1}^n < T} D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n})(S_{\tau_i^n} - S_{\tau_{i-1}^n}) - k\varepsilon_n \sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left| D_x v(\tau_i^n, S_{\tau_i^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \right| \Bigg).$$

Here, Z^n is understood as the hedging error of a discrete Delta-hedging strategy of an European option with underlying asset S, maturity $T > 0$, initial endowment the Black-Scholes price $u(0,S_0)$, price function v and pay-off $g(S_T)$, when transaction costs are small.

Most of the time, researchers consider strategies whose initial endowments tend to the super-replication price. Indeed, we know by [Levental 1997] or [Soner 1995] that the cheapest super-replication price is given by the Buy-and-Hold strategy and we cannot expect to do better, when transaction costs are taken into account. For us, this is not a satisfactory answer to the hedging issue, that is why we shall work on the more realistic case, where investors want to hedge their risks against the sale of options at "realistic" prices (we mean not too high for obvious competitive reasons; indeed, on the market, nobody would pay for an option at the super-replication price).

Assuming that $(\mathcal{T}^n, D_x v)$ is an admissible sequence of strategies in the sense of Definition 4.2.1, we can be decomposed Z_T^n in $M_T^n + R_T^n$, where M^n is a local martingale and $\varepsilon_n^{-1} R_T^n \xrightarrow{a.s.} 0$ by Proposition 4.3.1. The quadratic variation of M^n is *approximatively* given by (see the proof of Proposition 4.3.1 and Theorem 4.3.1 in the **Lower bound** paragraph)

$$\langle M^n \rangle_T \approx \frac{1}{6} \sum_{\tau_{i-1}^n < T} \left(1 + \frac{2kS_{\tau_{i-1}^n}}{\lambda(\tau_{i-1}^n, S_{\tau_{i-1}^n})} \right)^2 D_{xx}^2 v(\tau_{i-1}^n, S_{\tau_{i-1}^n})^2 (\Delta S_{\tau_i^n})^4.$$

However, note that the definition of Z^n is a little bit different from that of Chapter 4 because, here, we choose the initial endowment as the Black-Scholes price. This is a constant, which does not play any role in the computation of the quadratic variation $\langle M^n \rangle_T$.

We want to build a "Markowitz Efficient Set"-like type curve, that is the optimal trade-off between the mean and the risk (here, standing for the variance or the VaR) of the tracking error, when we begin with the Black-Scholes price. Simple computations lead to

$$\mathbb{E}[Z_T^n] = -k\varepsilon_n \mathbb{E}\left[\sum_{\tau_{i-1}^n < T} S_{\tau_i^n} \left| D_x v(\tau_i^n, S_{\tau_i^n}) - D_x v(\tau_{i-1}^n, S_{\tau_{i-1}^n}) \right| \Big) \right].$$

Two remarks deserve to be considered :

1. $\mathbb{E}[Z_T^n]$ is non-positive.

2. $$k\sum_{\tau_{i-1}^n<T} S_{\tau_i^n}\left|D_xv(\tau_i^n,S_{\tau_i^n})-D_xv(\tau_{i-1}^n,S_{\tau_{i-1}^n})\right|$$

   appears as a natural cost process. We can verify easily the converge to $+\infty$ when n tends to $+\infty$.

Let $X_t = S_t$, $Y_t = D_xv(t,S_t)$, $Z_t = S_t$, $w_t^Y = kS_t$, $w_t^Z = \frac{1}{6}\left(1+\frac{2kS_t}{\lambda(t,S_t)}\right)^2(D_{xx}^2v(t,S_t))^2$, $\beta^Y = 1$, $\beta^Z = 4$. So, $\sigma_t^X = \sigma(t,S_t)S_t$, $\sigma_t^Y = D_{xx}^2v(t,S_t)$[1], $\sigma_t^Z = 1$, $p^* = 3/2$ and $q^* = 3$. Therefore, owing to Theorem 5.3.2, the optimal sequence of strategies which minimizes the criterion

$$\left(k\sum_{\tau_{i-1}^n<T} S_{\tau_i^n}\left|D_xv(\tau_i^n,S_{\tau_i^n})-D_xv(\tau_{i-1}^n,S_{\tau_{i-1}^n})\right|\right)^{2/3}$$
$$\times\left(\frac{1}{6}\sum_{\tau_{i-1}^n<T}\left(1+\frac{2kS_{\tau_{i-1}^n}}{\lambda(\tau_{i-1}^n,S_{\tau_{i-1}^n})}\right)^2 D_{xx}^2v(\tau_{i-1}^n,S_{\tau_{i-1}^n})^2(\Delta S_{\tau_i^n})^4\right)^{1/3}$$

is given by (taking $\mu=0$ to simplify the exposure)

$$\begin{cases}\tau_0^n := 0,\\ \tau_i^n = \inf\left\{t\geq\tau_{i-1}^n : |S_t-S_{\tau_{i-1}^n}|>\varepsilon_n\left(\frac{kS_{\tau_{i-1}^n}}{\frac{1}{6}\left(1+\frac{2kS_{\tau_{i-1}^n}}{\lambda(\tau_{i-1}^n,S_{\tau_{i-1}^n})}\right)^2|D_{xx}^2v(\tau_{i-1}^n,S_{\tau_{i-1}^n})|}\right)^{1/3}\right\}\wedge T.\end{cases}$$

Using the fact that, for each $i=1,...,N_T^n-1$, $|\Delta S_{\tau_i^n}| = \varepsilon_n\lambda(\tau_{i-1}^n,S_{\tau_{i-1}^n})$ by definition of an admissible sequence of strategies, we deduce that $\lambda(t,x)$ solves the equation in y :

$$y^3+4kxy^2+4k^2x^2y-\frac{6kx}{|D_{xx}^2v(t,x)|}=0.$$

Let $p = -\frac{4}{3}k^2x^2$, $q = -\frac{16k^3x^3}{27}-\frac{6kx}{|D_{xx}^2v(t,x)|}$ and $\Delta(t,x) = \frac{q^2}{4}+\frac{p^3}{27} = \frac{9k^2x^2}{D_{xx}^2v(t,x)^2}\left(1+\frac{16k^2x^2|D_{xx}^2v(t,x)|}{81}\right)>0$. Then, owing to Cardan's formula, there exists a unique solution given by

$$\lambda(t,x) = \left(-\frac{q}{2}+\sqrt{\Delta(t,x)}\right)^{1/3}+\left(-\frac{q}{2}-\sqrt{\Delta(t,x)}\right)^{1/3}-\frac{4}{3}kx. \tag{5.4.1}$$

[1] We are aware that σ^Y must be positive, actually this is not fulfilled at least we assume v convex else we have to change all the proof of this chapter.

Therefore, the optimal sequence of strategies is the hitting times given above with the λ just before and the Deltas given by $D_x v$ where v solves the non-linear Cauchy problem (we substitute λ in the equation (4.2.3))

$$D_t v(t,x) + \frac{\sigma(t,x)^2 x^2}{2}\left(1 + \frac{2kx}{\left(-\frac{q}{2}+\sqrt{\Delta(t,x)}\right)^{1/3} + \left(-\frac{q}{2}-\sqrt{\Delta(t,x)}\right)^{1/3} - \frac{4}{3}kx}\right) D^2_{xx} v(t,x) = 0, \tag{5.4.2}$$

and

$$v(T,x) = g(x).$$

5.5 Numerical experiments

Even if it seems more natural to take the sum of the transaction costs as cost function, the simulations give poor outcomes in figures 5.1 and 5.2, where the curves "*landon*" and "*landonTC*" stand respectively for the criterion with N_T^n and with the sum of transaction costs as cost function. We use the same method as in Section 4.4.1 to compute an approximation of a solution of the equation (5.4.2) (whenever the solution exists). We suggest as interpretation the lack of accuracy in the resolution of the equation (4.2.3) with λ given by (5.4.1), maybe an improvement in the method could turns out to be suitable.

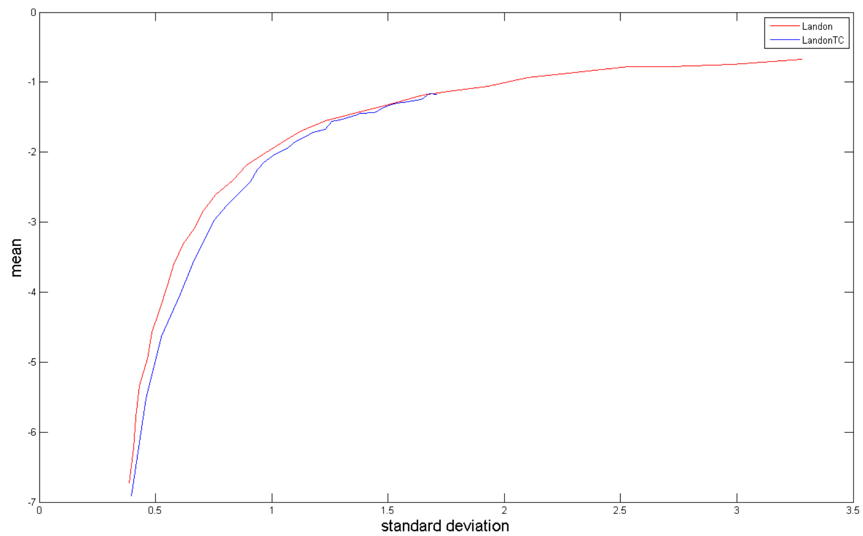

Figure 5.1: Mean-Variance of the hedging error for $k = 100$

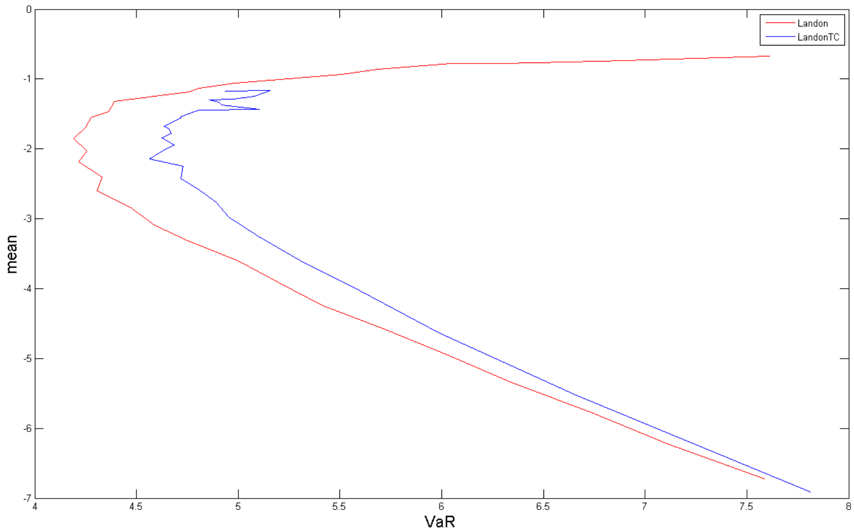

Figure 5.2: Mean-VaR of the hedging error for $k = 100$

Chapter 6

Strong optimality in Euler's method

Contents

Since the pioneering work of [Maruyama 1955] and by now, the study of approximation of stochastic differential equation has been a very active field of research at the intersection of numerical analysis and stochastic processes with numerous applications in different areas (see for instance, [Kloeden 2010], [Milstein 1994], [Platen 1999] and [Talay 1995]). Moreover, an impressive set of articles of Müller-Gronbach and al. [Hofmann 2000], [Muller-Gronbach 2002], [Muller-Gronbach 2004], [Muller-Gronbach 2008] investigates the problem in a general theory.

However, very few studies have dealt with the problem for general multidimensional processes or have got results when the discretization schemes are no longer deterministic. Here, we allow a rather large set of random algorithms by specifying directly an uniform control on a "benchmark" process and on the random number of stopping times involved in the discretization algorithm. More precisely, we consider a *continuous* stochastic differential equation

$$\mathrm{d}X_t = f(X_t)\mathrm{d}Y_t, \quad X_0 = x_0, \tag{6.0.1}$$

driven by the *continuous* semi-martingale Y. We aim at seeking an optimal approximation of the stochastic differential equation (6.0.1) by Euler's method based on the increment of Y, using adequate time mesh in a strong sense. We actually consider general (possibly stochastic) time grids of the form $\pi_n = \{\tau_0^n < \tau_1^n < ... < \tau_{N_T^n}^n\}$, where we restrict the times τ_i^n ($i = 0, ..., N_T^n$) to be stopping times (which is a mild assumption having in mind possible applications). The length of the stochastic grid N_T^n is also random and $\mathcal{F}_T$-adapted. The superscript "n" stands for a deterministic parameter which enables us to embed an asymptotic analysis into our setting; in other words, for almost all events ω, we let the length $N_T^n(\omega)$ become large and the step size $\sup_{1\leq i\leq N_T^n}(\tau_i^n - \tau_{i-1}^n)$ shrink to 0, when we take n large. This is similar to the point of view of Chapter 3, 4 and 5.

The optimality stems from a good choice of the mesh minimizing asymptotically a norm $||.||$ over all admissible grids :

$$0 < \liminf_{n\to+\infty}(N_T^n)^{1/2} \cdot ||X_T - X_T^n|| < +\infty, \quad a.s.$$

where X^n is the Euler approximation over the grid π. The choice of the norm $||.||$ plays a crucial role and is guided by simplicity and relevancy, it is discussed and defined in Section 6.4.

6.1 Notations

- For any $x \in \mathbb{R}^d$ (resp. $m \in (\mathbb{R}^d)^{\otimes 2}$ (the space of square matrices of rank d with I_d its identity element)), $|x|$ (resp. $|m|$) denotes the Euclidean norm of x (resp. the Frobenius norm $\sqrt{\mathrm{Tr}(mm^*)}$). If f is bounded, we note $|f|_\infty := \sup_x |f(x)|$. By convention, all the vectors in $\mathbb{R}^d$ are written in columns.

- $\mathcal{C}^m(U, V)$ is the set of functions m-times continuously differentiable from U to V. For any function $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^q)$ (resp. $\in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^q)$, the Jacobian $f'(x)$ (resp. the Hessian $f''(x)$) will stand as usual for the matrix $(\partial_{x_j} f(x)^i)_{1\leq i\leq q, 1\leq j\leq d} \in \mathbb{R}^q \otimes \mathbb{R}^d$ (resp. for the tensor $(\partial^2_{x_j,x_k} f(x)^i)_{1\leq i\leq q, 1\leq j,k\leq d} \in \mathbb{R}^q \otimes (\mathbb{R}^d)^{\otimes 2}$).

- For any càdlag process X, $|X|_*$ stands for the supremum $\sup_{0\leq t\leq T} |X_t|$.

- For a finite grid $\pi = \{\tau_0 = 0 < ... < \tau_i < ... < \tau_N = T\}$, $|\Delta\tau|$ stands for the largest increment of two consecutive stopping times in π and N_T for the cardinality of π (except the first time).

- Let $(\alpha_n)_{n\geq 0}$ and $(\beta_n)_{n\geq 0}$ be sequences of random variables. We write $\alpha_n = O(\beta_n)$ (resp. $o(\beta_n)$) if $\sup_{n\geq 0}(|\beta_n^{-1}\alpha_n|) < +\infty$ $a.s.$ (resp. $|\beta_n^{-1}\alpha_n| \xrightarrow{a.s.} 0$).

- The last time before $t \leq T$ is defined by $\varphi_n(t) := \max\{\tau_j^n : \tau_j^n \leq t\}$. Let a be a function defined on $[0,T]$, we note $\Delta a_s := a_s - a_{\varphi_n(s)}$.

6.2 Model

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a probability space, supporting a d-dimensional Brownian motion $(B^i)_{1\leq i\leq d}$ defined on $[0,T]$, where $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies to the "usual" assumptions. The classical Itô-type stochastic differential equation is of the form

$$\mathrm{d}X_t = \sum_{j=1}^{d} f_j(X_t)\mathrm{d}Y_t^j, \quad X_0 = x_0,$$

where $x_0 \in \mathbb{R}^q$ and $f_j \in \mathcal{C}^3(\mathbb{R}^q, \mathbb{R}^q)$ with $|f_j'|_\infty < +\infty$. The driving process Y is a continuous Itô semi-martingale in $\mathbb{R}^d$ such that

$$dY_t = b_t\mathrm{d}t + \sigma_t\mathrm{d}B_t, \quad Y_0 = y_0,$$

with b, σ adapted continuous processes in $\mathbb{R}^d$ and in $(\mathbb{R}^d)^{\otimes 2}$ such that $\sup_{0\leq t\leq T}|b_t| + \sup_{0\leq t\leq T}|\sigma_t| \leq M_{b,\sigma}$ *a.s.* for some constant $M_{b,\sigma} > 0$.
The above assumptions allow to well define a unique solution $(X_t)_{0\leq t\leq T}$ (see [Protter 2004][Theorem 6, Chapter V, p.249].)
From now on, we assume that σ satisfy the following ellipticity condition :

$$\lambda_{\min}(\sigma\sigma^*) > 0, \, a.s.$$

and the process σ is θ_σ-Holder ($\theta_\sigma \in (0,1]$), i.e.

$$|\sigma_t - \sigma_s| \leq C_0(|Y_t - Y_s|^{\theta_\sigma} + |t-s|^{\theta_\sigma/2}),$$

where C_0 is a non-negative finite random variable.
We denote by X^n the continuous Euler approximation of the equation (6.0.1) :

$$\mathrm{d}X_t^n = \sum_{j=1}^{d} f_j(X_{\varphi_n(t)}^n)\mathrm{d}Y_t^j, \quad X_0^n = x_0. \tag{6.2.1}$$

Then, the strong error of discretization by the Euler method is defined by

$$e_{n,T}^X := X_T - X_T^n = \sum_{j=1}^{d} \int_0^T \left(f_j(X_s) - f_j(X_{\varphi_n(s)}^n)\right) \mathrm{d}Y_s^j. \tag{6.2.2}$$

The further analysis relies on the representation of $(e_{n,t}^X)$ as solution of an affine equation (see [Kurtz 1991]) and to achieve this representation, we define the solutions

$\nabla X, \nabla X^n, \nabla X^{-1}, (\nabla X^{-1})^n$ to the stochastic differential equations onto $(\mathbb{R}^q)^{\otimes 2}$

$$\nabla X_t = I_q + \sum_{j=1}^{d} \int_0^t f_j'(X_s)\nabla X_s \mathrm{d}Y_s^j, \tag{6.2.3}$$

$$\nabla X_t^n = I_q + \sum_{j=1}^{d} \int_0^t f_j'(X_{\varphi_n(s)}^n)\nabla X_{\varphi_n(s)}^n \mathrm{d}Y_s^j, \tag{6.2.4}$$

$$\begin{aligned}\nabla X_t^{-1} = I_q &- \sum_{j=1}^{d} \int_0^t \nabla X_s^{-1} f_j'(X_s)\mathrm{d}Y_s^j \\ &+ \sum_{j=1}^{d}\sum_{k=1}^{d} \int_0^t \nabla X_s^{-1} f_j'(X_s) f_k'(X_s)\mathrm{d}\langle Y^j, Y^k\rangle_s,\end{aligned} \tag{6.2.5}$$

$$\begin{aligned}(\nabla X_t^{-1})^n = I_q &- \sum_{j=1}^{d} \int_0^t (\nabla X_{\varphi_n(s)}^{-1})^n f_j'(X_{\varphi_n(s)}^n)\mathrm{d}Y_s^j \\ &+ \sum_{j=1}^{d}\sum_{k=1}^{d} \int_0^t (\nabla X_{\varphi_n(s)}^{-1})^n f_j'(X_{\varphi_n(s)}^n) f_k'(X_{\varphi_n(s)}^n)(\sigma_{\varphi_n(s)}\sigma_{\varphi_n(s)}^*)^{jk}\mathrm{d}s.\end{aligned} \tag{6.2.6}$$

Remark 6.2.1.
- *Owing to [Protter 2004][Theorem 48, Chapter V, p.320], ∇X_t is invertible, with inverse ∇X_t^{-1}. However, $(\nabla X^{-1})^n$ is not the inverse of ∇X^n (which is in general not invertible), we have to be careful on the place of the exponents.*
- *Here, the processes ∇X^n and $(\nabla X^{-1})^n$ are directly computable knowing the increments of Y. Moreover, $|\nabla X - \nabla X^n|_* \xrightarrow{a.s.} 0$ and $|\nabla X^{-1} - (\nabla X^{-1})^n|_* \xrightarrow{a.s.} 0$ thanks to the estimates (6.3.2) and (6.3.3).*
- *In dimension one ($q = 1$), the process ∇X can be thought as a Doléans-Dade exponential.*

Let P^n be the solution to the linear stochastic differential equation :

$$P_t^n = I_q + \sum_{j=1}^{d} \int_0^t f_j'(X_{\varphi_n(s)}^n) P_s^n \mathrm{d}Y_s^j.$$

From [Protter 2004][Theorem 48, Chapter V, p.320], P^n has an inverse, which fulfils the SDE

$$\begin{aligned}(P_t^n)^{-1} = I_q &- \sum_{j=1}^{d} \int_0^t (P_s^n)^{-1} f_j'(X_{\varphi_n(s)}^n)\mathrm{d}Y_s^j \\ &+ \sum_{j=1}^{d}\sum_{k=1}^{d} \int_0^t (P_s^n)^{-1} f_j'(X_{\varphi_n(s)}^n) f_k'(X_{\varphi_n(s)}^n)\mathrm{d}\langle Y^j, Y^k\rangle_s.\end{aligned}$$

6.2.1 Class of stopping times

Let ε_n be a square summable series ($\sum_{n\geq 0}\varepsilon_n^2 < +\infty$). This parameter will be used for letting asymptotic calculus. More precisely, for a given stochastic grid π_n, we link the convergence rate of $|\Delta\tau^n|$ to 0 and of N_T^n to $+\infty$ to the convergence rate of ε_n to 0. So, if we wish to control the error $e_{n,T}^X$, we cannot use any information about the unknown process X, but only about the driving process Y (which is easier to simulate for instance) : in the standard setting we take $Y = B$. That is why we introduce a domination on the known semi-martingale Y. The purpose is to establish some necessary structure on the grid π_n without restraining too much the class of stopping times :
$(\mathbf{A}_Y)$:

$$\sup_{1\leq i\leq N_T^n}\ \sup_{\tau_{i-1}^n\leq t\leq\tau_i^n}|Y_t - Y_{\tau_{i-1}^n}| = O(\varepsilon_n).$$

As usual, when we deal with this kind of problem, a cost function will need to be defined as, for instance, in [Muller-Gronbach 2008]. The choice of an "acceptable" cost function is guided by realistic considerations, but also, for the sake of simplicity, to enable a tractable mathematical analysis. We select the number of stopping times as in Chapter 3 and in most of articles in the literature :
$(\mathbf{A}_N)$: For $1 \leq \rho_N < \frac{5}{4}\wedge(1+\frac{\theta_\sigma}{2})$,

$$N_T^n = O(\varepsilon_n^{-2\rho_N}).$$

Definition 6.2.1. *We say that $\mathcal{T} := \{\mathcal{T}^n; n \geq 0\}$ is an admissible class if the sequence of stopping times fulfils the assumptions $(\mathbf{A}_Y)$ and $(\mathbf{A}_N)$.*

6.3 Order of the error of discretization by Euler's method

In this section, we estimate the order of ΔX^n and of $e_{n,.}^X$. The first one is easy to assess because $\Delta X_t^n = f(X_{\varphi_n(t)}^n)\Delta Y_t$ and hence, $|\Delta X^n|_* = O(\varepsilon_n)$ owing to the assumption $(\mathbf{A}_Y)$ and to $\sup_{n\geq 0}|X^n|_* < \infty$ *a.s.* (see Remark 6.3.1). The second one is more delicate, because $e_{n,.}^X$ is not at all explicit, but is stated through a stochastic differential equation. To overcome this issue, we seek an accurate estimate of $e_{n,s}^X$ and this is done in the next proposition.

Proposition 6.3.1. *Under the hypothesis* $(\mathbf{A}_Y)$ *and* $(\mathbf{A}_N)$*, for any* $\rho > 0$*,*

$$|e^X_{n,.}|_* = |X - X^n|_* = O(\varepsilon_n^{1-\rho}), \tag{6.3.1}$$

$$|\nabla X - \nabla X^n|_* = O(\varepsilon_n^{1-\rho}), \tag{6.3.2}$$

$$|\nabla X^{-1} - (\nabla X^{-1})^n|_* = O(\varepsilon_n^{\theta_\sigma-\rho}), \tag{6.3.3}$$

$$|(\nabla X^{-1})^n - (P^n)^{-1}|_* = O(\varepsilon_n^{\theta_\sigma-\rho}), \tag{6.3.4}$$

$$|\nabla X^{-1} - (P^n)^{-1}|_* = O(\varepsilon_n^{1-\rho}). \tag{6.3.5}$$

Proof. Assume for a while the rough estimate

$$|X^n|_* = O(\varepsilon_n^{-\rho}), \tag{$\star$}$$

for any $\rho > 0$. Then, $|\Delta X^n|_* \le (|f(0)| + |f'|_\infty |X^n_{\varphi_n(.)}|_*)|\Delta Y|_* = O(\varepsilon_n^{1-\rho})$, $\forall \rho > 0$. We employ the notations of Lemma 3.2.2. The items *i)*, *ii)* and *iii)* will advert to this lemma throughout the demonstration.
Let $p > 2$. Set

$$U^n_t := \varepsilon_n^{3-p} \sup_{0\le s\le t} |e^X_{n,s}|^p,$$

$$V^n_t := \varepsilon_n^{3-p} \int_0^t |\Delta f(X^n_s)|^p \mathrm{d}s.$$

The first item *i)* is clearly fulfilled. For the second item *ii)*, we use the regularity property of f which leads to

$$V^n_T \le \varepsilon_n^{3-p} T |f'|^p_\infty |\Delta X^n|^p_* = O(\varepsilon_n^2),$$

using the estimate $(\star)$ with $\rho = \frac{1}{p}$. The convergence of the series immediately follows.
It remains to check the item *iii)* which is the toughest requirement. Let $t \in [0,T]$. Set $\theta_k := \inf\left\{ s \le T : \sum_{n\ge 0} V^n_s > k \right\}$. From the equation (6.2.2), we have

$$\begin{aligned}
&\mathbb{E}[\sup_{0\le s\le t\wedge\theta_k} |e^X_{n,s}|^p] \\
&\le 2^{p-1}\mathbb{E}\left[\sup_{0\le s\le t\wedge\theta_k} \left| \int_0^s (f(X_r) - f(X^n_{\varphi_n(r)}))b_r \mathrm{d}r \right|^p + \sup_{0\le s\le t\wedge\theta_k} \left| \int_0^s (f(X_r) - f(X^n_{\varphi_n(r)}))\sigma_r \mathrm{d}B_r \right|^p \right] \\
&\le 2^{p-1}\mathbb{E}\left[(t\wedge\theta_k)^{p-1} \int_0^{t\wedge\theta_k} |(f(X_s) - f(X^n_{\varphi_n(s)}))b_s|^p \mathrm{d}s + c_p \left(\int_0^{t\wedge\theta_k} |(f(X_s) - f(X^n_{\varphi_n(s)}))\sigma_s|^2 \mathrm{d}s \right)^{p/2} \right] \\
&\le 4^{p-1}\mathbb{E}\left[\left((t\wedge\theta_k)^{p-1} + (t\wedge\theta_k)^{p/2-1} c_p \right) M^p_{b,\sigma} \int_0^{t\wedge\theta_k} \left(|f(X_s) - f(X^n_s)|^p + |\Delta f(X^n_s)|^p \right) \mathrm{d}s \right],
\end{aligned}$$

where the second inequality comes from the Burkholder-Davis-Gundy inequality. Then, using again the regularity property of f, we unveil an integral-type inequality

$$\mathbb{E}[\sup_{0\le s\le t\wedge\theta_k} |e^X_{n,s}|^p] \le K_{p,T}\left\{\mathbb{E}\left[\int_0^{t\wedge\theta_k} |\Delta f(X^n_s)|^p \mathrm{d}s\right] + |f'|^p_\infty \int_0^t \mathbb{E}[\sup_{0\le r\le s\wedge\theta_k} |e^X_{n,r}|^p]\mathrm{d}s\right\},$$

where $K_{p,T} := 4^{p-1}M^p_{b,\sigma}(T^{p-1} + T^{p/2-1}c_p)$.
Moreover, it is a standard exercise to show that $|X^n|_* \in \mathcal{L}_p$ and thus $\mathbb{E}[\sup_{0\le s\le T} |e^X_{n,s}|^p] < +\infty$.
Thus, by Gronwall's inequality, we conclude

$$\mathbb{E}[\sup_{0\le s\le t\wedge\theta_k} |e^X_{n,s}|^p] \le K_{p,T}e^{K_{p,T}T|f'|_\infty}\mathbb{E}\left[\int_0^{t\wedge\theta_k} |\Delta f(X^n_s)|^p \mathrm{d}s\right],$$

that is,

$$\mathbb{E}U^n_{t\wedge\theta_k} \le K_{p,T}e^{K_{p,T}T|f'|_\infty}\mathbb{E}V^n_{t\wedge\theta_k}.$$

We conclude with the help of Lemma 3.2.2 that $\sum_{n\ge0} \varepsilon_n^{3-p}|e^X_{n,.}|^p_* < +\infty$ *a.s.* . So, $\sup_{n\ge0}\left(\varepsilon_n^{3-p}|e^X_{n,.}|^p_*\right) < +\infty$, in other words $|e^X_{n,.}|_* = O(\varepsilon_n^{1-\rho})$ with $\rho = 3/p$.
Now, it remains to show the estimate ($\star$). Actually, this is a simpler version of the proof we have done, but with $U^n_t := \varepsilon_n^2 \sup_{0\le s\le t} |X^n_s - X_0|^p$ and $V^n_t := \varepsilon_n^2(t^p + t^{p/2})$. The main step is the domination assumption, we have

$$\begin{aligned}
&\mathbb{E}[\sup_{0\le s\le t} |X^n_s - X_0|^p] \\
&\le 2^{p-1}\mathbb{E}\left[\sup_{0\le s\le t\wedge\theta_k}\left|\int_0^s f(X^n_{\varphi_n(r)})b_r\mathrm{d}r\right|^p + c_p\left(\int_0^{t\wedge\theta_k} |f(X^n_{\varphi_n(s)})\sigma_s|^2\mathrm{d}s\right)^{p/2}\right] \\
&\le 4^{p-1}M^p_{b,\sigma}\mathbb{E}\Bigg[t^{p-1}\int_0^t \sup_{0\le r\le s} |f(X^n_{\varphi_n(r)}) - f(X_0)|^p\mathrm{d}r + |f(X_0)|^p t^p \\
&\quad + c_p t^{p/2-1}\int_0^{t\wedge\theta_k} \sup_{0\le r\le s} |f(X^n_{\varphi_n(r)}) - f(X_0)|^p\mathrm{d}s + |f(X_0)|^p t^{p/2}\Bigg] \\
&\le 4^{p-1}M^p_{b,\sigma}\left\{(T^{p-1} + c_pT^{p/2-1})|f'|^p_\infty\int_0^t \sup_{0\le r\le s} |X^n_r - X_0|\mathrm{d}s + |f(X_0)|^p(t^p + c_pt^{p/2})\right\}.
\end{aligned}$$

Then, applying Gronwall's inequality, we get

$$\mathbb{E}[\sup_{0\le s\le t} |X^n_s - X_0|^p] \le |f(X_0)|^p(t^p + c_pt^{p/2})e^{Kt},$$

where $K := 4^{p-1}M^p_{b,\sigma}|f'|^p_\infty\{T^{p-1} + c_pT^{p/2-1}\}$. We skip details.

Remark 6.3.1. *We conclude the boundedness of X^n because* $\sup_{n\ge0}(|X^n|_*) \le \sup_{n\ge0}(|X - X^n|_*) + |X|_* < +\infty$.

Proof of (6.3.2), (6.3.3), (6.3.4), (6.3.5). First of all, we prove rough estimates on $|\nabla X^n|_*$ and $|(\nabla X^{-1})^n|_*$. Using Lemma 6.3.1 with $\tilde{b}^n_s = \sum_{j=1}^d f'_j(X^n_{\varphi_n(s)})b^j_s$, $(\tilde{\sigma}^n_s)^j = \sum_{k=1}^d f'_k(X^n_{\varphi_n(s)})\sigma^{kj}_s$ and $H = I_q$, so $\theta_H = 0$, we conclude that $|\nabla X^n|_* = O(\varepsilon_n^{-\rho})$, for any $\rho > 0$. The same argument holds for $|(\nabla X^{-1})^n|_* = O(\varepsilon_n^{-\rho})$, for any $\rho > 0$. The two differences $Z^{(1)}_t := (\nabla X^{-1}_t)^n - (P^n_t)^{-1}$ and $Z^{(2)}_t := \nabla X^{-1}_t - (P^n_t)^{-1}$ can be decomposed in the following form :

$$Z^{(1)}_t = \sum_{j=1}^d \int_0^t Z^{(1)}_s \left(-f'_j(X^n_{\varphi_n(s)})\mathrm{d}Y^j_s + \sum_{k=1}^d f'_j(X^n_{\varphi_n(s)})f'_k(X^n_{\varphi_n(s)})\mathrm{d}\langle Y^j, Y^k\rangle_s \right) + H^{(1)}_t,$$

$$H^{(1)}_t = \sum_{j=1}^d \int_0^t \Delta(\nabla X^{-1}_s)^n \left(f'_j(X^n_{\varphi_n(s)})\mathrm{d}Y^j_s - \sum_{k=1}^d f'_j(X^n_{\varphi_n(s)})f'_k(X^n_{\varphi_n(s)})\mathrm{d}\langle Y^j, Y^k\rangle_s \right)$$
$$- \sum_{j=1}^d \sum_{k=1}^d \int_0^t (\nabla X^{-1}_s)^n f'_j(X^n_{\varphi_n(s)})f'_k(X^n_{\varphi_n(s)})\Delta(\sigma_s\sigma^*_s)^{jk}\mathrm{d}s,$$

and

$$Z^{(2)}_t = \sum_{j=1}^d \int_0^t Z^{(2)}_s \left(-f'_j(X^n_{\varphi_n(s)})\mathrm{d}Y^j_s + \sum_{k=1}^d f'_j(X^n_{\varphi_n(s)})f'_k(X^n_{\varphi_n(s)})\mathrm{d}\langle Y^j, Y^k\rangle_s \right) + H^{(2)}_t,$$

$$H^{(2)}_t = - \sum_{j=1}^d \int_0^t \nabla X^{-1}_s (f'_j(X_s) - f'_j(X^n_{\varphi_n(s)}))\mathrm{d}Y^j_s$$
$$+ \sum_{j=1}^d \sum_{k=1}^d \int_0^t \nabla X^{-1}_s (f'_j(X_s)f'_k(X_s) - f'_j(X^n_{\varphi_n(s)})f'_k(X^n_{\varphi_n(s)}))\mathrm{d}\langle Y^j, Y^k\rangle_s.$$

Set $\tilde{b}^n_s := -\sum_{j=1}^d f'_j(X^n_{\varphi_n(s)})b^j_s + \sum_{j=1}^d \sum_{k=1}^d f'_j(X^n_{\varphi_n(s)})f'_k(X^n_{\varphi_n(s)})(\sigma_s\sigma^*_s)^{jk}$, $(\tilde{\sigma}^n_s)^j := -\sum_{k=1}^d f'_k(X^n_{\varphi_n(s)})\sigma^{kj}_s$ and $H^n = H^{(p)}$ $(p \in \{1,2\})$. Then, $\tilde{b}^n, \tilde{\sigma}^n$ are adapted processes in $(\mathbb{R}^q)^{\otimes 2}$, uniformly bounded. $H^{(1)}$ and $H^{(2)}$ are continuous semi-martingales. Moreover, we know that $|\Delta(\nabla X^{-1})^n|_* = O(\varepsilon_n^{1-\rho})$, owing to Lemma 6.3.2 and $|(\Delta(\sigma\sigma^*))^{jk}|_* \leq |((\Delta\sigma)\sigma^*)^{jk}|_* + |(\sigma_{\varphi_n(.)}\Delta\sigma^*)^{jk}|_* = O(\varepsilon_n^{\theta_\sigma})$, by assumption $(\mathbf{A}_\sigma)$. Then, using Lemma 6.3.2, we get

$$|H^{(1)}|_* = O(\varepsilon_n^{\theta_\sigma - \rho}).$$

Moreover, with probability 1, there exists a random variable $n_0 \in \mathbb{N}$ such that for any $n \geq n_0$, $|X - X^n_{\varphi_n(.)}|_* < 1$. The regularity property of f leads to

$$|f'_j(X) - f'_j(X^n_{\varphi_n(.)})|_* \leq \sup_{0\leq t\leq T} \sup_{|x - X_t|<1} |f''(x)|(|X - X_{\varphi_n(.)}|_* + |X_{\varphi_n(.)} - X^n_{\varphi_n(.)}|_*) = O(\varepsilon_n^{1-\rho}),$$

for $n \geq n_0$ *a.s.* , owing to Corollary 3.2.3 and Proposition 6.3.1. Then, using again Lemma 6.3.2, we have

$$|H^{(2)}|_* = O(\varepsilon_n^{1-\rho}).$$

Then, applying Lemma 6.3.1, we conclude that

$$|Z^{(1)}|_* = O(\varepsilon_n^{\theta_\sigma - \rho})$$

and

$$|Z^{(2)}|_* = O(\varepsilon_n^{1-\rho}).$$

We have proved (6.3.4) and (6.3.5). The estimate (6.3.3) readily follows from the two previous ones. To obtain (6.3.2), perform the same analysis for the simpler term $Z^{(3)} := \nabla X - \nabla X^n$ satisfying the equation :

$$\begin{aligned} Z_t^{(3)} &= \sum_{j=1}^{d} \int_0^t f_j'(X^n_{\varphi_n(s)}) Z_s^{(3)} \mathrm{d}Y_s^j + H_t^{(3)}, \\ H_t^{(3)} &= \sum_{j=1}^{d} \int_0^t f_j'(X^n_{\varphi_n(s)}) \Delta(\nabla X_s^n) \mathrm{d}Y_s^j + \sum_{j=1}^{d} \int_0^t (f_j'(X_s) - f_j'(X^n_{\varphi_n(s)})) \nabla X_s \mathrm{d}Y_s^j. \end{aligned}$$

□

In the above argumentation, we have made a frequent use of the above lemmas. We now justify them.

Lemma 6.3.1. *Let H^n be a sequence of continuous semi-martingales in $(\mathbb{R}^q)^{\otimes 2}$ such that $|H^n|_* = O(\varepsilon_n^{\theta_H})$, for some $\theta_H \in \mathbb{R}$. Let Z^n be the sequence of solutions to the linear systems of stochastic differential equation*

$$Z_t^n = \int_0^t Z_s^n (\tilde{b}_s^n \mathrm{d}s + \sum_{j=1}^{d} (\tilde{\sigma}_s^n)^j \mathrm{d}B_s^j) + H_t^n,$$

where $\tilde{b}^n, (\tilde{\sigma}^n)^j$ are two adapted processes in $(\mathbb{R}^q)^{\otimes 2}$, uniformly bounded by a positive constant M. Then, for any $\rho > 0$, the estimate

$$|Z^n|_* = O(\varepsilon_n^{\theta_H - \rho})$$

holds. The same result holds if Z^n is defined by $Z_t^n = \int_0^t (b_s^n \mathrm{d}s + \sum_{j=1}^{d} (\sigma_s^n)^j \mathrm{d}B_s^j) Z_s^n + H_t^n$.

Proof. We only prove the result for the first definition of Z^n. As the proof before, we use Lemma 3.2.2. Let $p > 2$. Set

$$\begin{aligned} U_t^n &:= \varepsilon_n^{2-p\theta_H} \sup_{0 \le s \le t} |Z_s^n|^p, \\ V_t^n &:= \varepsilon_n^{2-p\theta_H} \sup_{0 \le s \le t} |H_s^n|^p. \end{aligned}$$

The two first items *i)* and *ii)* are readily checked.
Let $t \in [0, T]$. Set $\theta_k := \inf \left\{ s \le T : \sum_{n \ge 0} V_s^n > k \right\}$. We aim at applying Gronwall's inequality to $\mathbb{E}[\sup_{0 \le s \le t \wedge \theta_k} |Z_s^n|^p]$.

1. $\mathbb{E}[\sup_{0\leq s\leq t\wedge\theta_k}|Z_s^n|^p] < +\infty\ \forall n,k$, because $\mathbb{E}[\sup_{0\leq s\leq t\wedge\theta_k}|H_s^n|^p] \leq k$ and $\tilde{b}^n, \tilde{\sigma}^n$ are bounded.

2. From the equation satisfied by Z^n, we easily have (as in the proof of Proposition 6.3.1)
$$\mathbb{E}[\sup_{0\leq s\leq t\wedge\theta_k}|Z_s^n|^p] \leq K\mathbb{E}\left[\int_0^{t\wedge\theta_k}|Z_s^n|^p\mathrm{d}s + \sup_{0\leq s\leq t\wedge\theta_k}|H_s^n|^p\right],$$
where the constant K depends on T, p, q, d and M (but not on k and n).

Gronwall's inequality yields
$$\mathbb{E}[\sup_{0\leq s\leq t\wedge\theta_k}|Z_s^n|^p] \leq Ke^{KT}\mathbb{E}[\sup_{0\leq s\leq t\wedge\theta_k}|H_s^n|^p],$$
i.e.
$$\mathbb{E}U^n_{t\wedge\theta_k} \leq Ke^{KT}\mathbb{E}V^n_{t\wedge\theta_k}.$$
We conclude (in the same way as in the proof of Proposition 6.3.1) that, for any $\rho > 0$, $|Z|_* = O(\varepsilon_n^{\theta_H-\rho})$. □

Lemma 6.3.2. *Assume* $(\mathbf{A}_N)$ *and* $(\mathbf{A}_Y)$. *Let* α^n, β^n *be two sequences of càdlàg adapted processes such that* $\sup_{n\geq 0}\varepsilon_n^{-\theta}(|\alpha^n|_* \vee |\beta^n|_*) < +\infty$ a.s. *for some* $\theta \in \mathbb{R}$. *Then, for any* $\rho > 0$,
$$\left|\int_0^\cdot \alpha_s^n\mathrm{d}B_s + \int_0^\cdot \beta_s^n\mathrm{d}s\right|_* = O(\varepsilon_n^{\theta-\rho})$$
and
$$\left|\int_{\varphi_n(.)}^\cdot \alpha_s^n\mathrm{d}B_s + \int_{\varphi_n(.)}^\cdot \beta_s^n\mathrm{d}s\right|_* = O(\varepsilon_n^{1+\theta-\rho}).$$

Proof. **First estimate** : We use Corollary 3.2.1 for the martingale part : let $p > 0$, we have
$$\sum_{n=0}^{+\infty}\left(\int_0^T|\varepsilon_n^{\rho-\theta}\alpha_s^n|^2\mathrm{d}s\right)^{p/2} < +\infty$$
for $\rho = 2/p$, which implies that
$$\sup_{n\geq 0}\left(\varepsilon_n^{\rho-\theta}\left|\int_0^\cdot \alpha_s^n\mathrm{d}B_s\right|_*^p\right) \leq \sum_{n=0}^{+\infty}\left|\int_0^\cdot \varepsilon_n^{\rho-\theta}\alpha_s^n\mathrm{d}B_s\right|_*^p < +\infty.$$
The term with β^n is obvious to deal with.

Second estimate : We use an *adaptation* of Lemma 4.5.3 component by component, for $k = 1$ to $X_t^0 = B_t^j$ (resp. $X_t^0 = t$) and $X_t^1 = \varepsilon_n^{-\theta}(\alpha_t^n)^j$ (resp. $X_t^1 = \varepsilon_n^{-\theta}\beta_t^n$) for any $j = 1, \cdots, d$. We emphasize the word *adaptation* because in the quoted lemma the process X^1 does not depend on n, but we can circumvent this problem using the *a.s.* boundedness of the sequence $\varepsilon_n^{-\theta}(|\alpha^n|_* \vee |\beta^n|_*)$, and then that of X^1, the rest of the proof of Lemma 4.5.3 being unchanged. □

6.4 Optimal grid for the Euler scheme method

6.4.1 Main results

It is time to have a discussion on the criterion we want to minimize and the sense of the minimization (in law or in $L^2(\Omega)$ or *a.s.*). A good criterion should be simple, for practical reasons, and relevant. Here, our choice is *by no way obvious*, but it is simple enough to have computations done explicitly. So, for reason given later on in Remark 6.4.2, we choose as *ad hoc* criterion the *a.s.* minimization of

$$N_T^n \mathrm{Tr}\langle \nabla X^{-1} e^X_{n,.}\rangle_T, \tag{6.4.1}$$

where ∇X is the tangent process defined in (6.2.3). Additionally, $\nabla X^{-1} e^X_{n,.}$ is the "main" stochastic integral term when we linearise the Euler scheme error (as in [Kurtz 1991]) and seek a Central Limit Theorem for the renormalized error. Beyond the tractability, this gives ground reasons to consider this criterion for the optimization. We give some convenient notations used in the presentation of the forthcoming results.

Notations

- Γ is a symmetric non-negative adapted matrix process satisfying :

$$2\mathrm{Tr}(\Gamma\sigma\sigma^*)\Gamma + 4\Gamma\sigma\sigma^*\Gamma = \Lambda, \tag{6.4.2}$$

  where Λ is given in (6.4.15).

- $(\Gamma^n)_{n\geq 0}$ is a sequence of symmetric non-negative adapted matrix processes satisfying :

$$2\mathrm{Tr}(\Gamma^n\sigma\sigma^*)\Gamma^n + 4\Gamma^n\sigma\sigma^*\Gamma^n = \Lambda^n, \tag{6.4.3}$$

  where Λ^n is given in (6.4.14).

The existence of Γ and $(\Gamma^n)_{n\in\mathbb{N}}$ follow from Lemma 3.3.1, as explained later. The content of the following theorem is to reveal a lower bound for the infimum limit of the criterion, independently of any admissible grid.

Theorem 6.4.1. *Under the assumptions* $(\mathbf{A}_Y)$ *and* $(\mathbf{A}_N)$ *with* $1 \leq \rho_N < \frac{5}{4} \wedge (1 + \frac{\theta_\sigma}{2})$*,*

$$\liminf_{n\to+\infty} \left\{N_T^n \mathrm{Tr}\langle \nabla X^{-1} e^X_{n,.}\rangle_T\right\} \geq \left\{\int_0^T \mathrm{Tr}(\Gamma_t\sigma_t\sigma_t^*)\mathrm{d}t\right\}^2.$$

Now, we explicitly build an optimal sequence of strategies. We note

$$\xi_{x_0,\Sigma} := \{x \in \mathbb{R}^d : (x - x_0)^*\Sigma(x - x_0) < 1\}.$$

Theorem 6.4.2. *Suppose that the assumptions of Theorem 6.4.1 are fulfilled. Set* $\Gamma^{n,\mu}$ *be the solution to the equation* (6.4.3)*, substituting* Λ^n *by* $\Lambda^n + \mu\chi_\mu(\lambda_{\min}(\Lambda^n))$ *with* $\mu > 0$*, where* χ_μ *is defined in Theorem 3.3.2. For a given* $n \in \mathbb{N}$*, define the strategy* $\mathcal{T}^n_\mu$ *by*

$$\begin{cases} \tau^n_0 := 0, \\ \tau^n_i = \inf\left\{ t \geq \tau^n_{i-1} : Y_t \notin \xi_{Y_{\tau^n_{i-1}}, \varepsilon_n^{-2}\Gamma^{n,\mu}_{\tau^n_{i-1}}} \right\} \wedge T, \quad i \geq 1. \end{cases} \tag{6.4.4}$$

Then, the sequence of strategies $\mathcal{T}_\mu = \{\mathcal{T}^n_\mu : n \geq 0\}$ *is admissible and* μ*-asymptotically optimal in the following sense:*

$$\limsup_{n\to+\infty} \left| N^n_T \mathrm{Tr}\langle \nabla X^{-1} e^X_{n,.}\rangle_T - (\int_0^T \mathrm{Tr}(\Gamma_t\sigma_t\sigma^*_t)\mathrm{d}t)^2 \right| \leq C_\mu,$$

where $C_\mu = \int_0^T \mathrm{Tr}\big(\Gamma^\mu_t\sigma_t\sigma^*_t\big)\mathrm{d}t|\Gamma^\mu + \Gamma|_* \int_0^T \frac{|\Gamma_t - \Gamma^\mu_t|}{\lambda_{\min}(\Gamma^\mu_t)}\mathrm{Tr}(\sigma_t\sigma^*_t)\mathrm{d}t + 2\int_0^T \mathrm{Tr}\big(\Gamma_t\sigma_t\sigma^*_t\big)\mathrm{d}t \left|\int_0^T \mathrm{Tr}\big((\Gamma^\mu_t - \Gamma_t)\sigma_t\sigma^*_t\big)\mathrm{d}t\right| + (\int_0^T \mathrm{Tr}\big((\Gamma^\mu_t - \Gamma_t)\sigma_t\sigma^*_t\big)\mathrm{d}t)^2$*. In particular, on the event* $\{\forall t \in [0,T], \lambda_{\min}(\Lambda_t) \geq \mu\}$*,* $N^n_T\mathrm{Tr}\langle\nabla X^{-1}e^X_{n,.}\rangle_T$ *converges* a.s. *to* $\big(\int_0^T \mathrm{Tr}(\Gamma_t\sigma_t\sigma^*_t)\mathrm{d}t\big)^2$.

Remark 6.4.1. *It is important to observe that the ellipsoid* $\xi_{Y_.,\varepsilon_n^{-2}\Gamma^{n,\mu}_.}$ *is computable by construction, that is it only relies on the observable processes* Y*,* X^n*,* $(\nabla X^{-1})^n$ *and the time mesh* $\mathcal{T}^n$*.*

6.4.2 Proof of Theorem 6.4.1

6.4.2.1 A direct demonstration

Starting from (6.2.2), a natural idea consists in approximating $f_j(X_s) - f_j(X^n_{\varphi_n(s)})$ using Taylor's theorem

$$\begin{aligned} f_j(X_s) - f_j(X^n_s) := f'_j(X^n_{\varphi_n(s)})e^X_{n,s} + \Delta f'_j(X^n_s)e^X_{n,s} \\ + \left[\int_0^1 \left(f'_j\left(X^n_s + \lambda e^X_{n,s}\right) - f'_j(X^n_s)\right)\mathrm{d}\lambda\right] e^X_{n,s}, \end{aligned} \tag{6.4.5}$$

and the variation $\Delta f_j(X^n_s)$ using Ito's formula

$$\begin{aligned} \Delta f_j(X^n_s) = f'_j(X^n_{\varphi_n(s)})\Delta X^n_s + \int_{\varphi_n(s)}^s \Delta f'_j(X^n_t)\mathrm{d}X^n_t \\ + \frac{1}{2}\int_{\varphi_n(s)}^s \mathrm{Tr}\left(f''_j(X^n_t)\mathrm{d}\langle X^n\rangle_t\right), \end{aligned} \tag{6.4.6}$$

where the formula is understood component by component.
Then, the equation (6.2.2) becomes

$$e^X_{n,T} = \sum_{j=1}^{d} \int_0^T \left(f'_j(X^n_{\varphi_n(s)}) e^X_{n,s} + f'_j(X^n_{\varphi_n(s)}) \Delta X^n_s \right) \mathrm{d}Y^j_s + e^n_{1,T}, \tag{6.4.7}$$

$$\begin{aligned} e^n_{1,t} := \sum_{j=1}^{d} \int_0^t \Bigg(& \left[\int_0^1 \left(f'_j \left(X^n_s + \lambda e^X_{n,s} \right) - f'_j(X^n_s) \right) \mathrm{d}\lambda \right] e^X_{n,s} \\ & + \Delta f'_j(X^n_s) e^X_{n,s} + \int_{\varphi_n(s)}^{s} \Delta f'_j(X^n_r) \mathrm{d}X^n_r \\ & + \frac{1}{2} \int_{\varphi_n(s)}^{s} \mathrm{Tr} \left(f''_j(X^n_r) \mathrm{d}\langle X^n \rangle_r \right) \Bigg) \mathrm{d}Y^j_s. \end{aligned} \tag{6.4.8}$$

The following result is a handy theorem [Protter 2004, Theorem 56 p.334] which gives an explicit alternative form for the solution to systems of linear equations

Theorem 6.4.3. *Let H be a column vector of q continuous semi-martingales and let A^j be an $q \times q$ matrices of adapted càdlàg processes. Let U be the solution to : $U_t = I_q + \sum_{j=1}^d \int_0^t A^j_s U_s \mathrm{d}Y^j_s$. Then, the solution X^H of the equation*

$$X_t = H_t + \sum_{j=1}^{d} \int_0^t A^j_s X_s \mathrm{d}Y^j_s$$

admits the following representation

$$X^H_t = U_t H_0 + U_t \int_0^t U_s^{-1} \left(\mathrm{d}H_s - \sum_{j=1}^{d} A^j_s \mathrm{d}\langle H, Y^j \rangle_s \right).$$

Then, taking

$$\begin{cases} H_t := \sum_{j=1}^d \int_0^t f'_j(X^n_{\varphi_n(s)}) \Delta X^n_s \mathrm{d}Y^j_s + e^n_{1,t}, \\ A^j_s := f'_j(X^n_{\varphi_n(s)}), \end{cases}$$

in Theorem 6.4.3 with U standing for (we emphasize the dependence on n)

$$P^n_t = I_q + \sum_{j=1}^{d} \int_0^t f'_j(X^n_{\varphi_n(s)}) P^n_s \mathrm{d}Y^j_s,$$

the equation (6.4.7) can be rewritten in

$$
\begin{aligned}
\nabla X_T^{-1} e_{n,T}^X = \int_0^T (\nabla X_s^{-1})^n \Bigg\{ & \sum_{j=1}^d f_j'(X_{\varphi_n(s)}^n) \Delta X_s^n \mathrm{d}Y_s^j \\
& - \sum_{j=1}^d \sum_{k=1}^d f_j'(X_{\varphi_n(s)}^n) f_k'(X_{\varphi_n(s)}^n) \Delta X_s^n \mathrm{d}\langle Y^j, Y^k \rangle_s \Bigg\} + e_{2,T}^n, \qquad (6.4.9)
\end{aligned}
$$

$$
\begin{aligned}
e_{2,t}^n := & \int_0^t (P_s^n)^{-1} \Bigg\{ \mathrm{d}e_{1,s}^n - \sum_{j=1}^d f_j'(X_{\varphi_n(s)}^n) \mathrm{d}\langle e_{1,.}^n, Y^j \rangle_s \Bigg\} \\
& + \int_0^T \left[(P_s^n)^{-1} - (\nabla X_s^{-1})^n \right] \Bigg\{ \sum_{j=1}^d f_j'(X_{\varphi_n(s)}^n) \Delta X_s^n \mathrm{d}Y_s^j \\
& - \sum_{j=1}^d \sum_{k=1}^d f_j'(X_{\varphi_n(s)}^n) f_k'(X_{\varphi_n(s)}^n) \Delta X_s^n \mathrm{d}\langle Y^j, Y^k \rangle_s \Bigg\} \\
& + (\nabla X_T^{-1} - (P_T^n)^{-1}) e_{n,T}^X. \qquad (6.4.10)
\end{aligned}
$$

Remark 6.4.2. *We mention that adding the intermediate term $(\nabla X_s^{-1})^n$ instead of $(P_s^n)^{-1}$ is useful to derive fully explicit optimal strategy. We note that the quantity $\nabla X_T^{-1} e_{n,T}^X$ is easier to handle than $e_{n,T}^X$, that is why we choose the criterion $N_T^n \mathrm{Tr} \langle \nabla X^{-1} e_{n,.}^X \rangle_T$ even if the interpretation of this quantity is less immediate. Actually, we have not succeeded in dealing with the simple term $e_{n,T}^X$, except in dimension one where ∇X_T^{-1} can be partially interpreted as a change of probability, tackling the problem with $e_{n,T}^X$ because the quadratic variation is invariant by drift transformation. Improvements are left to further research.*

As mentioned before, the error representation (6.4.9) gives the intuition that a good criterion for choosing the Euler scheme time grid consists in "minimizing" the stochastic integral contribution. Namely, with the help of equation (6.4.9), we can express the quadratic variation matrix of $\nabla X^{-1} e_{n,.}^X$:

$$
\begin{aligned}
\langle \nabla X^{-1} e_{n,.}^X \rangle_T = \sum_{j=1}^d \sum_{k=1}^d \int_0^T & (\nabla X_s^{-1})^n f_j'(X_{\varphi_n(s)}^n) \Delta X_s^n \\
& \times \left[(\nabla X_s^{-1})^n f_k'(X_{\varphi_n(s)}^n) \Delta X_s^n \right]^* \mathrm{d}\langle Y^j, Y^k \rangle_s + e_{3,T}^n, \qquad (6.4.11)
\end{aligned}
$$

$$
e_{3,t}^n := \int_0^t \sum_{j=1}^d (\nabla X_s^{-1})^n f_j'(X_{\varphi_n(s)}^n) \Delta X_s^n \mathrm{d}\langle Y^j, e_{2,.}^n \rangle_s + \langle e_{2,.}^n \rangle_t. \qquad (6.4.12)
$$

That is,

$$\mathrm{Tr}\langle \nabla X^{-1} e^X_{n,.}\rangle_T = \int_0^T \Delta Y_s^* \Lambda^n_{\varphi_n(s)} \Delta Y_s \mathrm{d}s + \int_0^T \Delta Y_s^* (\bar{\Lambda}^n_{\varphi_n(s),s} - \Lambda^n_{\varphi_n(s)}) \Delta Y_s \mathrm{d}s + \mathrm{Tr}(e^n_{3,T}), \tag{6.4.13}$$

$$\bar{\Lambda}^n_{r,s} := \left(\sum_{j=1}^d \sum_{k=1}^d \left[(\nabla X_s^{-1})^n f'_k(X^n_r) f_m(X^n_r) \right]^* (\nabla X_s^{-1})^n f'_j(X^n_r) f_l(X^n_r) (\sigma_s \sigma_s^*)^{jk} \right)_{1 \le l,m \le d},$$

$$\Lambda^n_s := \bar{\Lambda}^n_{s,s}. \tag{6.4.14}$$

The matrix Λ^n_s has three interesting properties :

1. $\Lambda^n_s \in \mathcal{S}^d(\mathbb{R})$.

2. $\forall x \in \mathbb{R}^d$, $x^* \Lambda^n_s x \ge 0$; indeed, set $z = \left(\left[(\nabla X_s^{-1})^n f'_j(X^n_s) \sum_{m=1}^d f_m(X^n_s) x^m \right]^k \right)_{1 \le j,k \le d}$, then $x^* \Lambda^n_s x = \mathrm{Tr}((\sigma^* z)^* \sigma^* z)) \ge 0$.

3. Owing to Lemma 6.3.2 and the estimate (6.3.3), we easily prove that, for any $\rho > 0$,

$$|\bar{\Lambda}^n_{\varphi_n(.),.} - \Lambda^n_{\varphi_n(.)}|_* = O(\varepsilon_n^{\theta_\sigma - \rho})$$

   and that the sequence Λ^n converges uniformly *a.s.* to

$$\Lambda := \left(\sum_{j=1}^d \sum_{k=1}^d \left[\nabla X^{-1} f'_k(X) f_m(X) \right]^* \nabla X^{-1} f'_j(X) f_l(X) (\sigma\sigma^*)^{jk} \right)_{1 \le l,m \le d}, \tag{6.4.15}$$

   by working component by component.

Now, as in Chapter 3, we seek a representation of the term $\int_0^T \Delta Y_s^* \Lambda_{\varphi_n(s)} \Delta Y_s \mathrm{d}s$ in the form $\sum_{\tau^n_{i-1} < T} \left(\Delta Y^*_{\tau^n_i} \Gamma_{\tau^n_{i-1}} \Delta Y_{\tau^n_i} \right)^2$ plus a stochastic integral, where Γ is a sequence of adapted $d \times d$-matrix process which has to be defined. Applying Itô's formula on each interval $[\tau_{i-1}^n, \tau_i^n]$, we obtain

$$\begin{aligned}
&\sum_{\tau_{i-1}^n<T} \big(\Delta Y_{\tau_i^n}^* \Gamma_{\tau_{i-1}^n} \Delta Y_{\tau_i^n}\big)^2 \\
&= \int_0^T \Delta Y_t^* \Big(2\mathrm{Tr}\big(\Gamma_{\varphi_n(t)}\sigma_{\varphi_n(t)}\sigma_{\varphi_n(t)}^*\big)\Gamma_{\varphi_n(t)} \\
&\quad + \big(\Gamma_{\varphi_n(t)} + (\Gamma_{\varphi_n(t)})^*\big)\sigma_{\varphi_n(t)}\sigma_{\varphi_n(t)}^*\big(\Gamma_{\varphi_n(t)} + (\Gamma_{\varphi_n(t)})^*\big)\Big)\Delta Y_t \mathrm{d}t \\
&\quad + \int_0^T \Delta Y_t^* \Big(2\mathrm{Tr}\big(\Gamma_{\varphi_n(t)}\Delta(\sigma_t\sigma_t^*)\big)\Gamma_{\varphi_n(t)} + \big(\Gamma_{\varphi_n(t)} \\
&\quad + (\Gamma_{\varphi_n(t)})^*\big)\Delta(\sigma_t\sigma_t^*)\big(\Gamma_{\varphi_n(t)} + (\Gamma_{\varphi_n(t)})^*\big)\Big)\Delta Y_t \mathrm{d}t \\
&\quad + 2\int_0^T \Delta Y_t^* \Gamma_{\varphi_n(t)}\Delta Y_t \Delta Y_t^*\big(\Gamma_{\varphi_n(t)} + (\Gamma_{\varphi_n(t)})^*\big)\mathrm{d}Y_t.
\end{aligned}$$

Then, noting $x^\Lambda = \sigma^*\Gamma\sigma$, we seek a solution to the non-linear system :

$$2\mathrm{Tr}\big(x_{\varphi_n(t)}^\Lambda\big)x_{\varphi_n(t)}^\Lambda + \big(x_{\varphi_n(t)}^\Lambda + (x_{\varphi_n(t)}^\Lambda)^*\big)^2 = \sigma_{\varphi_n(t)}^*\Lambda_{\varphi_n(t)}\sigma_{\varphi_n(t)}. \tag{6.4.16}$$

For each $t \in [0,T]$, $\sigma_t^*\Lambda_t\sigma_t$ is a symmetric non-negative-definite matrix, then it has a unique square-root (symmetric nonnegative-definite matrix) [Horn 1990, Theorem 7.2.6 p.405]. Then, owing to Lemma 3.3.1, there exists exactly one adapted process x^Λ with values in $\mathcal{S}_+^d(\mathbb{R})$, solution to the equation $2\mathrm{Tr}(x^\Lambda)x^\Lambda + 4(x^\Lambda)^2 = C^2$, where $C := (\sigma^*\Lambda\sigma)^{1/2}$. In addition, this solution is continuous as a function of Λ on $\mathcal{S}_+^d(\mathbb{R})$ and of t on $[0,T]$ *a.s.* because $\sigma^*\Lambda\sigma$ is continuous *a.s.* , $c \mapsto c^{1/2}$ is continuous on $\mathcal{S}_+^d(\mathbb{R})$ [Chen 1997] and the solution x^Λ is continuous as a function of C on $\mathcal{S}^d(\mathbb{R})$. Thanks to the ellipticity condition of σ and its continuity in time, $\Gamma = (\sigma^*)^{-1}x^\Lambda(\sigma^{-1})$ is also continuous as a function of Λ and time.To sum up, we have

$$\mathrm{Tr}\langle\nabla X^{-1}e_{n,.}^X\rangle_T = \sum_{\tau_{i-1}^n<T} (\Delta Y_{\tau_i^n}^*\Gamma_{\tau_{i-1}^n}\Delta Y_{\tau_i^n})^2 + e_{4,T}^n, \tag{6.4.17}$$

$$\begin{aligned}
e_{4,s}^n :=& \int_0^s \Delta Y_t^*\Big(2\mathrm{Tr}\big(\Gamma_{\varphi_n(t)}\Delta(\sigma_t\sigma_t^*)\big)\Gamma_{\varphi_n(t)} + 4\Gamma_{\varphi_n(t)}\Delta(\sigma_t\sigma_t^*)\Gamma_{\varphi_n(t)}\Big)\Delta Y_t\mathrm{d}t \\
&+ 4\int_0^s \Delta Y_t^*\Gamma_{\varphi_n(t)}\Delta Y_t\Delta Y_t^*\Gamma_{\varphi_n(t)}\mathrm{d}Y_t + \int_0^s \Delta Y_t^*(\Lambda_{\varphi_n(t)}^n - \Lambda_{\varphi_n(t)})\Delta Y_t\mathrm{d}t \\
&+ \int_0^s \Delta Y_t^*(\bar\Lambda_{\varphi_n(t),t}^n - \Lambda_{\varphi_n(t)}^n)\Delta Y_t\mathrm{d}t + \mathrm{Tr}(e_{3,s}^n),
\end{aligned} \tag{6.4.18}$$

and the Cauchy-Schwarz inequality implies the desired inequality :

$$N_T^n\mathrm{Tr}\langle\nabla X^{-1}e_{n,.}^X\rangle_T \geq \left\{\sum_{\tau_{i-1}^n<T}\Delta Y_{\tau_i^n}^*\Gamma_{\tau_{i-1}^n}\Delta Y_{\tau_i^n}\right\}^2 + N_T^n e_{4,T}^n.$$

Remark 6.4.3. *Here, we have used the non-negativity property of Γ, that induces an explicit asymptotic lower bound.*

Moreover, Proposition 3.2.3 gives us the convergence of the second variation of a stochastic integral to its quadratic variation for a bounded sequence of continuous adapted integrands (Γ satisfies these requirements) :

$$\sum_{\tau_{i-1}^n<T} \Delta Y_{\tau_i^n}^* \Gamma_{\tau_{i-1}^n} \Delta Y_{\tau_i^n} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}(\Gamma_s \sigma_s \sigma_s^*)\mathrm{d}s.$$

Thus, we deduce (admitting for a while the fast convergence to 0 of the error terms) that

$$\liminf_{n\to+\infty} \left\{ N_T^n \mathrm{Tr}\langle \nabla X^{-1} e_{n,.}^X \rangle_T \right\} \geq \left\{ \int_0^T \mathrm{Tr}(\Gamma_s \sigma_s \sigma_s^*)\mathrm{d}s \right\}^2.$$

6.4.2.2 Convergence of the residual terms

- We aim at proving $|e_{1,.}^n|_* = O(\varepsilon_n^{2-\rho})$: a Taylor expansion applied to the regular function f and combined to the estimates of Section 6.3 gives the following inequalities

$$\sup_{0\leq t\leq T} |f_j'\left(X_t^n + \lambda e_{n,t}^X\right) - f_j'(X_t)| \leq \sup_{0\leq t\leq T} \max_{0\leq\lambda\leq 1} |f_j''(X_t^n + \lambda e_{n,t}^X)||e_{n,t}^X| \leq C_0 \varepsilon_n^{1-\rho},$$

$$\sup_{0\leq t\leq T} |\Delta f_j'(X_t^n)| \leq \sup_{0\leq t\leq T} \sup_{0\leq\lambda\leq 1} |f_j''(X_t^n + \lambda \Delta X_t^n)||\Delta X_t^n| \leq C_0 \varepsilon_n^{1-\rho}$$

where ρ is any positive real number. So, we deduce that, for any $\rho > 0$,

$$\left.\begin{array}{c} \left|\int_0^1 \left[\left(f_j'\left(X^n + \lambda e_{n,.}^X\right) - f_j'(X^n)\right)\mathrm{d}\lambda\right] e_{n,.}^X\right|_* \leq C_0 \varepsilon_n^{2-\rho}, \\ \left|\Delta f_j'(X^n) e_{n,.}^X\right|_* \leq C_0 \varepsilon_n^{2-\rho}, \\ \left|\int_{\varphi_n(.)}^{.} \Delta f_j'(X_s^n)\mathrm{d}X_s^n\right|_* \leq C_0 \varepsilon_n^{2-\rho}, \\ \left|\frac{1}{2}\int_{\varphi_n(.)}^{.} \mathrm{Tr}\left(f_j''(X_s^n)\mathrm{d}\langle X^n\rangle_s\right)\right|_* \leq C_0 \varepsilon_n^{2-\rho}, \end{array}\right\} \qquad (\star)$$

applying Lemma 6.3.2 to the third term and Corollary 3.2.2 to the last term. We conclude that

$$|e_{1,.}^n|_* = O(\varepsilon_n^{2-\rho}), \tag{6.4.19}$$

with the help of Lemma 6.3.2. With the above arguments, we have also proved that

$$\langle e_{1,.}^n \rangle_T = O(\varepsilon_n^{4-\rho}),$$

because we know by Corollary 3.2.1 that a uniform control on a local martingale (here, the local martingale part of $e_{1,.}^n$) is equivalent to a control over its quadratic variation.

- We intend to show $|e^n_{2,.}|_* = O(\varepsilon_n^{1+\theta_\sigma-\rho})$: the Cauchy-Schwarz inequality for the norm induced by the quadratic variation implies
$$|\langle e^n_{1,.}, Y^j\rangle|_* \leq \langle e^n_{1,.}\rangle_T^{1/2}\langle Y^j\rangle_T^{1/2} = O(\varepsilon_n^{2-\rho}), \quad \forall \rho > 0. \tag{6.4.20}$$
Moreover, $|\int_0^\cdot (P^n_s)^{-1}\mathrm{d}e^n_{1,s}| = O(\varepsilon_n^{2-\rho})$ owing to the boundedness of $|(P^n)^{-1}|_*$ (coming from the estimates (6.3.5)), the estimates ($\star$) and Lemma 6.3.2. Putting together the estimates (6.4.19), (6.4.20), (6.3.1), (6.3.4), (6.3.5) and using Lemma 6.3.2, we conclude that, for any $\rho > 0$,
$$|e^n_{2,.}|_* = O(\varepsilon_n^{1+\theta_\sigma-\rho}).$$
- Let us prove $|e^n_{3,.}|_* = O(\varepsilon_n^{2+\theta_\sigma-\rho})$: using the same argument as for the control of $\langle e^n_{1,.}\rangle_T$ before, we immediately deduce that, for any $\rho > 0$,
$$\langle e^n_{2,.}\rangle_T = O(\varepsilon_n^{2(1+\theta_\sigma)-\rho}).$$
Then, the Cauchy-Schwarz inequality and the assumption $(\mathbf{A}_Y)$ imply that, for any $\rho > 0$,
$$|e^n_{3,.}|_* = O(\varepsilon_n^{2+\theta_\sigma-\rho}).$$
- We are going to establish $|e^n_{4,.}|_* = O(\varepsilon_n^{2+\theta_\sigma-\rho})$: as already mentioned, $|\bar\Lambda^n_{\varphi_n(.),.} - \Lambda^n_{\varphi_n(.)}|_* = O(\varepsilon_n^{\theta_\sigma-\rho})$, for any $\rho > 0$, then
$$\left|\int_0^\cdot \Delta Y_s^*(\bar\Lambda^n_{\varphi_n(s),s} - \Lambda^n_{\varphi_n(s)})\Delta Y_s \mathrm{d}s\right|_* = O(\varepsilon_n^{2+\theta_\sigma-\rho}).$$
Moreover, owing to the estimates (6.3.1), (6.3.3) and Lemma 6.3.2, for any $\rho > 0$, we get
$$\left|\int_0^\cdot \Delta Y_s^*(\Lambda^n_{\varphi_n(s)} - \Lambda_{\varphi_n(s)})\Delta Y_s \mathrm{d}s\right|_* = O(\varepsilon_n^{2+\theta_\sigma-\rho}).$$
As far as that goes, σ is θ_σ-Holder, then
$$\left|\int_0^\cdot \Delta Y_t^*\Big(2\mathrm{Tr}\big(\Gamma_{\varphi_n(t)}\Delta(\sigma_t\sigma_t^*)\big)\Gamma_{\varphi_n(t)} + 4\Gamma_{\varphi_n(t)}\Delta(\sigma_t\sigma_t^*)\Gamma_{\varphi_n(t)}\Big)\Delta Y_t \mathrm{d}t\right|_* = O(\varepsilon_n^{2+\theta_\sigma}).$$
It remains us to estimate $\left|\int_0^\cdot \Delta Y_t^*\Gamma_{\varphi_n(t)}\Delta Y_t\Delta Y_t^*\Gamma_{\varphi_n(t)}\mathrm{d}Y_t\right|_*$. But, this directly follows from Lemma 6.3.2 and we have for any $\rho > 0$,
$$\left|\int_0^\cdot \Delta Y_t^*\Gamma_{\varphi_n(t)}\Delta Y_t\Delta Y_t^*\Gamma_{\varphi_n(t)}\mathrm{d}Y_t\right|_* = O(\varepsilon_n^{3-\rho}).$$
We deduce that, with the help of the estimate of $|e^n_{3,.}|_*$,
$$|e^n_{4,.}|_* = O(\varepsilon_n^{2+\theta_\sigma-\rho})$$
holds, for any $\rho > 0$.

Then, we conclude that $N^n_T e^n_{4,T} \xrightarrow{a.s.} 0$ as soon as $\rho_N < 1 + \theta_\sigma/2$.

6.4.3 Proof of Theorem 6.4.2

We note Γ^μ the unique solution to the equation (6.4.2), substituting Λ for $\Lambda + \mu\chi_\mu(\lambda_{\min}(\Lambda))I_d$. For any given non-negative definite symmetric matrix Λ_t, the matrix $\Gamma_t \in S^d_+(\mathbb{R})$ solving $2\mathrm{Tr}(\Gamma_t\sigma_t\sigma_t^*) + 4\Gamma_t\sigma_t\sigma_t^*\Gamma_t = \Lambda_t$ is given by $\Gamma_t = (\sigma_t^*)^{-1}x^{\Lambda_t}\sigma_t^{-1}$, where x^{Λ_t} is the solution to

$$2\mathrm{Tr}(x)x + 4x^2 = [(\sigma_t^*\Lambda_t\sigma_t)^{1/2}]^2,$$

i.e. within the notation of Proposition 6.5.1, $x^{\Lambda_t} = x([\sigma_t^*\Lambda_t\sigma_t]^{1/2})$. Hence,

$$\Gamma_t = (\sigma_t^*)^{-1}x([\sigma_t^*\Lambda_t\sigma_t]^{1/2})\sigma_t^{-1}.$$

Thus, applying Proposition 6.5.1, we have

$$\begin{aligned}
|\Gamma_t^\mu - \Gamma_t^{n,\mu}| &\le |(\sigma_t^*)^{-1}||\sigma_t^{-1}||x([\sigma_t^*\Lambda_t^\mu\sigma_t]^{1/2}) - x([\sigma_t^*\Lambda_t^{n,\mu}\sigma_t]^{1/2})| \\
&\le K|(\sigma_t^*)^{-1}||\sigma_t^{-1}|\left(1 + \frac{\sqrt{\mathrm{Tr}(\sigma_t^*\Lambda_t^\mu\sigma_t + \sigma_t^*\Lambda_t^{n,\mu}\sigma_t)}}{\sqrt{\mathrm{Tr}(\sigma_t^*\Lambda_t^\mu\sigma_t) \wedge \mathrm{Tr}(\sigma_t^*\Lambda_t^{n,\mu}\sigma_t)}}\right)|\sigma_t^*\Lambda_t^\mu\sigma_t - \sigma_t^*\Lambda_t^{n,\mu}\sigma_t|^{1/2} \\
&\le K'|(\sigma_t^*)^{-1}||\sigma_t^{-1}|\left(1 + \frac{\sqrt{|\sigma^*\Lambda^\mu\sigma|_* + |\sigma^*\Lambda^{n,\mu}\sigma|_*}}{\sqrt{\mu\inf_{t\in[0,T]}\lambda_{\min}(\sigma_t\sigma_t^*)}}\right)|\sigma|_* \\
&\quad \left(|\Lambda - \Lambda^n|_* + \mu|\chi_\mu(\lambda_{\min}(\Lambda)) - \chi_\mu(\lambda_{\min}(\Lambda^n))|_*|I_d|\right)^{1/2},
\end{aligned}$$

where K, K' are independent of n. Finally, we have proved

$$|\Gamma^\mu - \Gamma^{n,\mu}|_* \le K^n|\Lambda - \Lambda^n|_*^{1/2} = O(\varepsilon_n^{\theta_\sigma/2-\rho}), \quad \forall\rho > 0, \quad a.s. \tag{6.4.21}$$

where K^n is a bounded (in n) random variable. The admissibility of $\mathcal{T}_\mu$ is guaranteed by the same argument as in Proposition 3.2.4. Now, let us show the μ-optimality. Writing $N_T^n = 1 + \sum_{1\le i\le N_T^n-1} 1$, we point out

$$\begin{aligned}
\varepsilon_n^2 N_T^n &= \varepsilon_n^2 + \sum_{1\le i\le N_T^n-1} \Delta Y_{\tau_i^n}^*\Gamma_{\tau_{i-1}^n}^{n,\mu}\Delta Y_{\tau_i^n} \\
&= \varepsilon_n^2 - \Delta Y_T^*\Gamma_{\tau_{N_T^n-1}^n}^{n,\mu}\Delta Y_T + \sum_{\tau_{i-1}^n<T} \Delta Y_{\tau_i^n}^*\Gamma_{\tau_{i-1}^n}^{n,\mu}\Delta Y_{\tau_i^n} \\
&\xrightarrow{a.s.} \int_0^T \mathrm{Tr}\big(\Gamma_t^\mu\sigma_t\sigma_t^*\big)\mathrm{d}t,
\end{aligned} \tag{6.4.22}$$

using the estimate (6.4.21).

Owing to (6.4.17), we write

$$\begin{aligned}
\mathrm{Tr}\langle\nabla X^{-1}e_{n,\cdot}^X\rangle_T = &\sum_{1\le i\le N_T^n-1}\big(\Delta Y_{\tau_i^n}^*\Gamma_{\tau_{i-1}^n}^{n,\mu}\Delta Y_{\tau_i^n}\big)^2 + \big(\Delta Y_T^*\Gamma_{\tau_{N_T^n-1}^n}^{n,\mu}\Delta Y_T\big)^2 \\
&+ \sum_{\tau_{i-1}^n<T}\Big\{\big(\Delta Y_{\tau_i^n}^*\Gamma_{\tau_{i-1}^n}\Delta Y_{\tau_i^n}\big)^2 - \big(\Delta Y_{\tau_i^n}^*\Gamma_{\tau_{i-1}^n}^{n,\mu}\Delta Y_{\tau_i^n}\big)^2\Big\} + e_{4,T}^n.
\end{aligned} \tag{6.4.23}$$

First of all, we have

$$\big(\Delta Y_T^* \Gamma^{n,\mu}_{\tau^n_{N^n_T-1}} \Delta Y_T\big)^2 = O(\varepsilon_n^4) = o(\varepsilon_n^2).$$

The third quantity in equation (6.4.23) can be written in the form

$$\sum_{\tau^n_{i-1}<T} \Big(\big(\Delta Y^*_{\tau^n_i} \Gamma_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2 - \big(\Delta Y^*_{\tau^n_i} \Gamma^{\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2\Big)$$
$$+ \sum_{\tau^n_{i-1}<T} \Big(\big(\Delta Y^*_{\tau^n_i} \Gamma^{\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2 - \big(\Delta Y^*_{\tau^n_i} \Gamma^{n,\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2\Big).$$

Owing to (6.4.21) and to Proposition 3.2.3, we write the last sum

$$\begin{aligned}
&\varepsilon_n^{-2} \left| \sum_{\tau^n_{i-1}<T} \Big(\big(\Delta Y^*_{\tau^n_i} \Gamma^{\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2 - \big(\Delta Y^*_{\tau^n_i} \Gamma^{n,\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2\Big) \right| \\
&\leq \varepsilon_n^{-2} \sum_{\tau^n_{i-1}<T} \big|\Delta Y^*_{\tau^n_i}(\Gamma^{\mu}_{\tau^n_{i-1}} - \Gamma^{n,\mu}_{\tau^n_{i-1}})\Delta Y_{\tau^n_i}\big| \big|\Delta Y^*_{\tau^n_i}(\Gamma^{\mu}_{\tau^n_{i-1}} + \Gamma^{n,\mu}_{\tau^n_{i-1}})\Delta Y_{\tau^n_i}\big|, \\
&\leq |\Gamma^{n,\mu} + \Gamma^{\mu}|_* |\Gamma^{n,\mu} - \Gamma^{\mu}|_* \sum_{\tau^n_{i-1}<T} \left(\frac{|\Delta Y_{\tau^n_i}|^2}{\lambda_{\min}(\Gamma^{\mu}_{\tau^n_{i-1}})} + \frac{\lambda_{\min}(\Gamma^{\mu}_{\tau^n_{i-1}}) - \lambda_{\min}(\Gamma^{n,\mu}_{\tau^n_{i-1}})}{\lambda_{\min}(\Gamma^{n,\mu}_{\tau^n_{i-1}})\lambda_{\min}(\Gamma^{\mu}_{\tau^n_{i-1}})} |\Delta Y_{\tau^n_i}|^2 \right) \\
&= O(\varepsilon_n^{\theta_\sigma/2-\rho}),
\end{aligned}$$

for any $\rho > 0$, where we have used the definition of $\mathcal{T}_\mu$ at the last inequality. Moreover, $\varepsilon_n^{-2} \left|\sum_{\tau^n_{i-1}<T} \Big(\big(\Delta Y^*_{\tau^n_i} \Gamma_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2 - \big(\Delta Y^*_{\tau^n_i} \Gamma^{\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2\Big)\right|$ is bounded by

$$\begin{aligned}
&\varepsilon_n^{-2} \sum_{\tau^n_{i-1}<T} \big|\Delta Y^*_{\tau^n_i}(\Gamma_{\tau^n_{i-1}} - \Gamma^{\mu}_{\tau^n_{i-1}})\Delta Y_{\tau^n_i}\big| \big|\Delta Y^*_{\tau^n_i}(\Gamma_{\tau^n_{i-1}} + \Gamma^{\mu}_{\tau^n_{i-1}})\Delta Y_{\tau^n_i}\big|, \\
&\leq |\Gamma + \Gamma^{\mu}|_* \sum_{\tau^n_{i-1}<T} |\Gamma_{\tau^n_{i-1}} - \Gamma^{\mu}_{\tau^n_{i-1}}| \left(\frac{|\Delta Y_{\tau^n_i}|^2}{\lambda_{\min}(\Gamma^{\mu}_{\tau^n_{i-1}})} + \frac{\lambda_{\min}(\Gamma^{\mu}_{\tau^n_{i-1}}) - \lambda_{\min}(\Gamma^{n,\mu}_{\tau^n_{i-1}})}{\lambda_{\min}(\Gamma^{n,\mu}_{\tau^n_{i-1}})\lambda_{\min}(\Gamma^{\mu}_{\tau^n_{i-1}})} |\Delta Y_{\tau^n_i}|^2 \right) \\
&\xrightarrow{a.s.} |\Gamma + \Gamma^{\mu}|_* \int_0^T \frac{|\Gamma_t - \Gamma^{\mu}_t|}{\lambda_{\min}(\Gamma^{\mu}_t)} \mathrm{Tr}(\sigma_t \sigma_t^*) \mathrm{d}t,
\end{aligned}$$

where the last limit comes from an application of Proposition 3.2.3. Then, we have

$$\begin{aligned}
\limsup_{n\to+\infty} &\left| \varepsilon_n^{-2} \left\{ \sum_{\tau^n_{i-1}<T} \big(\Delta Y^*_{\tau^n_i} \Gamma_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2 - \big(\Delta Y^*_{\tau^n_i} \Gamma^{n,\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i}\big)^2 \right\} \right| \\
&\leq |\Gamma^{\mu} + \Gamma|_* \int_0^T \frac{|\Gamma_t - \Gamma^{\mu}_t|}{\lambda_{\min}(\Gamma^{\mu}_t)} \mathrm{Tr}(\sigma_t \sigma_t^*) \mathrm{d}t, \quad a.s..
\end{aligned}$$

Similarly to (6.4.22), we show that

$$\sum_{1\leq i\leq N^n_T-1} \Delta Y^*_{\tau^n_i} \Gamma^{n,\mu}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i} \xrightarrow{a.s.} \int_0^T \mathrm{Tr}\big(\Gamma^{\mu}_t \sigma_t \sigma_t^*\big) \mathrm{d}t.$$

Let us summarize: setting $L_T := \int_0^T \mathrm{Tr}\big(\Gamma_t\sigma_t\sigma_t^*\big)\mathrm{d}t$ and $L_T^\mu := \int_0^T \mathrm{Tr}\big((\Gamma_t^\mu - \Gamma_t)\sigma_t\sigma_t^*\big)\mathrm{d}t$, we have shown

$$\varepsilon_n^2 N_T^n \xrightarrow{a.s.} L_T + L_T^\mu,$$

$$\limsup_{n\to+\infty} \big|\varepsilon_n^{-2}\mathrm{Tr}\langle\nabla X^{-1}e^X_{n,.}\rangle_T - L_T\big| \le |\Gamma^\mu+\Gamma|_* \int_0^T \frac{|\Gamma_t-\Gamma_t^\mu|}{\lambda_{\min}(\Gamma_t^\mu)}\mathrm{Tr}(\sigma_t\sigma_t^*)\mathrm{d}t + |L_T^\mu|, \qquad a.s..$$

Then, we get

$$\begin{aligned}
&\limsup_{n\to+\infty}\Big|N_T^n\mathrm{Tr}\langle\nabla X^{-1}e^X_{n,.}\rangle_T - L_T^2\Big| \\
&\le \limsup_{n\to+\infty}|\varepsilon_n^{-2}\mathrm{Tr}\langle\nabla X^{-1}e^X_{n,.}\rangle - L_T|\limsup_{n\to+\infty}\varepsilon_n^2N_T^n + L_T\limsup_{n\to+\infty}|\varepsilon_n^2N_T^n - L_T| \\
&\le (L_T+L_T^\mu)|\Gamma^\mu+\Gamma|_*\int_0^T\frac{|\Gamma_t-\Gamma_t^\mu|}{\lambda_{\min}(\Gamma_t^\mu)}\mathrm{Tr}(\sigma_t\sigma_t^*)\mathrm{d}t + 2L_T|L_T^\mu| + (L_T^\mu)^2, \quad a.s..
\end{aligned}$$

Theorem 6.4.2 is proved. □

6.5 Appendix

Proposition 6.5.1. *Let $c\in\mathcal{S}^d(\mathbb{R})$. Let $x(c)$ be the unique solution in $\mathcal{S}^d_+(\mathbb{R})$ of the equation*

$$2\mathrm{Tr}(x)x + 4x^2 = c^2.$$

Then,

$$\frac{2}{4+d}\mathrm{Tr}([c^2]^{1/2}) \le \mathrm{Tr}(x(c)) \le \frac{1}{2}\mathrm{Tr}([c^2]^{1/2}).$$

In addition, there exists $K>0$ such that for any $(c_1,c_2)\in\mathcal{S}^d(\mathbb{R})/\{0\}$,

$$|x(c_1)-x(c_2)| \le K\left(1+\frac{\sqrt{\mathrm{Tr}(c_1^2+c_2^2)}}{\sqrt{\mathrm{Tr}(c_1^2)\wedge\mathrm{Tr}(c_2^2)}}\right)|c_1^2-c_2^2|^{1/2}.$$

Proof. First of all , the existence and uniqueness of the equation are guaranteed by Lemma 3.3.1. Throughout the proof, we shall borrow the notations to Section 3.5.2. We know that $x(c)$ has the form

$$x(c) := -\frac{y_{\lambda(c)}}{4}I_d + \frac{1}{2}\Big(\frac{y^2_{\lambda(c)}}{4}I_d + c^2\Big)^{1/2},$$

where $y_{\lambda(c)} = \mathrm{Tr}(x(c))$ solves $h(\lambda(c), y_{\lambda(c)}) = 0$, where

$$h : \begin{cases} \mathbb{R}^d\times\mathbb{R}_+ & \to \mathbb{R} \\ (\lambda, y) & \mapsto (4+d)y - \sum_{i=1}^d\sqrt{y^2+4\lambda_i^2}. \end{cases}$$

Let us prove the upper and lower bounds on $\mathrm{Tr}(x(c))$. We have $D_y h(\lambda, y) = 4 + d - \sum_{j=1}^{q} \frac{y}{\sqrt{y^2+4\lambda_i^2}} \geq 4$, thus

$$4 \leq D_y h(\lambda, y) \leq 4 + d, \tag{$\star$}$$

for any $\lambda \in \mathbb{R}^d$ and $y \geq 0$. Thus, writing

$$h(\lambda(c), y_{\lambda(c)}) - h(\lambda(c), 0) = 2\sum_{j=1}^{d} |\lambda(c)_j| = 2\sum_{j=1}^{d} \sqrt{\lambda(c^2)_j} = 2\mathrm{Tr}([c^2]^{1/2}).$$

We conclude to the upper and lower bounds for $y_{\lambda(c)} = \mathrm{Tr}(x(c))$ using ($\star$).

Let $(c_1, c_2) \in \mathcal{S}^d(\mathbb{R})/\{0\}$. Then, $y_{\lambda(c_1)} > 0$ and $y_{\lambda(c_2)} > 0$, in view of bounds on $\mathrm{Tr}(x(c))$. Owing to [Stroock 2006][Theorem 5.2.2, p.131], we have

$$\begin{aligned}
|x(c_1) - x(c_2)| &\leq K\left(|y_{\lambda(c_1)} - y_{\lambda(c_2)}| + \left|\left(\frac{y^2_{\lambda(c_1)}}{4} I_d + c_1^2\right)^{1/2} - \left(\frac{y^2_{\lambda(c_2)}}{4} I_d + c_2^2\right)^{1/2}\right|\right) \\
&\leq K\left(|y_{\lambda(c_1)} - y_{\lambda(c_2)}| + \frac{K'}{y_{\lambda(c_1)} \wedge y_{\lambda(c_2)}}\left|\frac{y^2_{\lambda(c_1)} - y^2_{\lambda(c_2)}}{4} I_d + c_1^2 - c_2^2\right|\right) \\
&\leq K\left(\left(1 + \frac{K'(y_{\lambda(c_1)} + y_{\lambda(c_2)})|I_d|}{4(y_{\lambda(c_1)} \wedge y_{\lambda(c_2)})}\right)|y_{\lambda(c_1)} - y_{\lambda(c_1)}| + \frac{K'}{y_{\lambda(c_1)} \wedge y_{\lambda(c_2)}}|c_1^2 - c_2^2|\right),
\end{aligned}$$

where K, K' are two constants (independent of c_1, c_2). Using the bounds on $y_{\lambda(c)}$ and the obvious inequality $d\mathrm{Tr}(c^2) \geq \left(\mathrm{Tr}([c^2]^{1/2})\right)^2 \geq \mathrm{Tr}(c^2)$, we deduce, for a new constant K'' (independent of c_1 and c_2), that

$$|x(c_1) - x(c_2)| \leq K''\left\{\left(1 + \frac{\sqrt{\mathrm{Tr}(c_1^2 + c_2^2)}}{\sqrt{\mathrm{Tr}(c_1^2) \wedge \mathrm{Tr}(c_2^2)}}\right)|y_{\lambda(c_2)} - y_{\lambda(c_1)}| + \frac{|c_1^2 - c_2^2|}{\sqrt{\mathrm{Tr}(c_1^2) \wedge \mathrm{Tr}(c_2^2)}}\right\}.$$

It remains to get an estimation of $|y_{\lambda(c_1)} - y_{\lambda(c_2)}|$.
Writing the identity

$$h(\lambda(c_1), y_{\lambda(c_1)}) - h(\lambda(c_1), y_{\lambda(c_2)}) = h(\lambda(c_2), y_{\lambda(c_2)}) - h(\lambda(c_1), y_{\lambda(c_2)}),$$

and applying Taylor's theorem with ($\star$), we have

$$|h(\lambda(c_1), y_{\lambda(c_1)}) - h(\lambda(c_1), y_{\lambda(c_2)})| \geq 4|y_{\lambda(c_1)} - y_{\lambda(c_2)}|$$

and

$$\begin{aligned}
|h(\lambda(c_2), y_{\lambda(c_2)}) - h(\lambda(c_1), y_{\lambda(c_2)})| &= \left|\sum_{i=1}^{d} \sqrt{y^2_{\lambda(c_2)} + 4\lambda(c_2)_i^2} - \sum_{i=1}^{d} \sqrt{y^2_{\lambda(c_2)} + 4\lambda(c_1)^2_{\eta(i)}}\right| \\
&\leq 2d^{3/4}\left(\sum_{i=1}^{d} \left|\lambda(c_2^2)_i - \lambda(c_1^2)_{\eta(i)}\right|^2\right)^{1/4},
\end{aligned}$$

where η is any permutation of $\{1, ..., d\}$. From Hoffman and Wielandt's theorem [Horn 1990, p.368], there exists a permutation $\tilde{\eta}$ such that

$$\left(\sum_{i=1}^{d} \left| \lambda(c_2^2)_i - \lambda(c_1^2)_{\tilde{\eta}(i)} \right|^2 \right)^{1/2} \leq |c_2^2 - c_1^2|_2 \leq K'''|c_2^2 - c_1^2|,$$

where K''' is a constant coming from the equivalent norm.
Then, we have

$$|y_{\lambda(c_1)} - y_{\lambda(c_2)}| \leq \frac{d^{3/4}\sqrt{K'''}}{2}|c_2^2 - c_1^2|^{1/2},$$

which concludes the proof. □

CHAPTER 7

Central Limit Theorem for ellipsoid based time grid

Contents

7.1 Introduction.

In this Chapter, we study the convergence in law of the renormalized error (arising from discretization of stochastic processes or approximation of stochastic integrals), in the general form like those of previous chapters for instance, when the time grid is made of a sequence of hitting times of random ellipsoids. Indeed, these ellipsoid based time grids play a central role in our problems since they are optimal regarding the minimization of quadratic variation criterion. To our knowledge, this is the first attempt to study the convergence of discretization error in a multidimensional setting and when the time grid is given by a sequence of hitting times of random ellipsoids. To show this kind of result, we classically apply a mere version of [Jacod 2003, Theorem IX.7.3]. See also [Jacod 2012] for a recent account on the subject.

Theorem 7.1.1. *Let M^n be a continuous $\mathcal{F}$-local martingale on $\mathbb{R}^d$ starting from $M_0^n := 0$, defined on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the $\mathbb{P}$-augmentation of the filtration generated by a Brownian motion $(B_t)_{0\leq t\leq T}$. Suppose that* *a)* $\langle M^n \rangle_t \to_{n\to+\infty} \int_0^t k_t dt$ *in probability for* $t \in [0,T]$ *and* k *is an* $\mathbb{R}^{d\times d}$*-valued predictable process;*

b) $\langle M^n, B\rangle_t \to_{n\to+\infty} 0$ *in probability for* $t \in [0,T]$.

Then, M^n *converges* $\mathcal{F}$*-stably to* M*, where* M *is a continuous local martingale, with the representation*

$$M = \int_0^{\cdot} z_s \mathrm{d}W_s$$

where z *is an* $\mathbb{R}^{d\times d}$*-valued square-root matrix process (i.e.* $zz^* = k$*) and* W *is an* d*-dimensional Brownian motion independent of* B *defined on an extended probability space* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$.

In [Kurtz 1991], [Kurtz 1996], [Jacod 1998] regarding Euler scheme error or in [Rootzen 1980], [Gobet 2001], [Hayashi 2005] or [Geiss 2009] regarding discretization of stochastic integrals or hedging error in finance, the Central Limit Theorem holds with a limit having zero correlation with the initial Brownian motion, as soon as the time grid is deterministic. This choice of grid leads to the centering property $(\star)$ $\mathbb{E}[(B^j_{\tau^n_{i+1}} - B^j_{\tau^n_i})^{2p+1}|\mathcal{F}_{\tau^n_i}] = 0$ for the odd moments of the increments of the j-th component of B. When $(\tau^n_i)_i$ are hitting times of symmetric boundaries in dimension 1 (thus the property $(\star)$ still holds), [Fukasawa 2011b] shows also a Central Limit Theorem with an independent Brownian motion at the limit.

Our study is devoted to the multidimensional setting where $(\tau^n_i)_i$ are hitting times of random ellipsoids. In that case, there is no reason to have the centering property $(\star)$ and indeed, it is necessary to identify the correlation at the limit. This is the original contribution of our work, and we show that the limit correlation of the renormalized error depends on a family of elliptic Dirichlet problem.

7.2 Model.

For a vector $v \in \mathbb{R}^d$, v^k stands for k-th coefficient and v^* for the transposition of v. For a matrix $M \in (\mathbb{R}^d)^{\otimes 2}$, we note $M^{jk}, M^{j\cdot}, M^{\cdot k}$ respectively for the coefficient (j,k), the jth-row and the kth-column of M and M^* for the transposition of M. We assume that the process Y solves the SDE

$$dY_t = b(t, Y_t)\mathrm{d}t + \sigma(t, Y_t)\mathrm{d}B_t, \quad Y_0 = y_0,$$

with b, σ are continuous bounded functions. Moreover, we assume that σ satisfies the following uniform ellipticity condition :

$$\lambda_{\min}(\sigma\sigma^*)(t,x) \geq \sigma^2_{\min} > 0, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^d, \qquad (H1)$$

and b, σ is $\theta_{b,\sigma}$-Holder ($\theta_{b,\sigma} \in (0,1]$)) with respect to the parabolic distance, i.e.

$$|b(t,x) - b(t',x')| + |\sigma(t,x) - \sigma(t',x')| \leq C_{b,\sigma}(|x-x'|^{\theta_{b,\sigma}} + |t-t'|^{\theta_{b,\sigma}/2}), \quad \text{(H2)}$$

where $C_{b,\sigma}$ is a positive constant.
Let $(\Sigma_t)_{0 \leq t \leq T}$ be an adapted continuous process taking values in the set of positive definite matrices, such that there exists a constant $C_\Sigma > 0$ such that

$$\frac{1}{C_\Sigma} \leq \lambda_{\min}(\Sigma_t) \leq \lambda_{\max}(\Sigma_t) \leq C_\Sigma, \qquad \forall t \in [0,T] \quad a.s.. \qquad (\star\Sigma)$$

This process serves to build random ellipsoids, from which we construct hitting times for the time grid. Namely, for a given $n \in \mathbb{N}$, define the sequence $\mathcal{T}^n = (\tau_0^n = 0 < \tau_1^n < ... < \tau_{N_T^n}^n = T)$ by

$$\begin{cases} \tau_0^n := 0, \\ \tau_i^n = \inf\left\{t \geq \tau_{i-1}^n : (Y_t - Y_{\tau_{i-1}^n})^* \Sigma_{\tau_{i-1}^n} (Y_t - Y_{\tau_{i-1}^n}) > \varepsilon_n^2\right\} \wedge T, \ i \geq 1. \end{cases} \tag{7.2.1}$$

As before, $(\varepsilon_n)_{n \geq 0}$ is deterministic and $\sum_{n \geq 0} \varepsilon_n^2 < +\infty$.

From Chapter 3 and 6, we have seen that the error of discretization for the hedging error (3.3.6) or for the Euler-Maruyama scheme (6.4.9) can be essentially written in the form

$$M_T^n + A_T^n + o(\varepsilon_n^2),$$

where $M^n := \int_0^\cdot \beta_{\varphi_n(s)} \Delta Y_s \gamma_{\varphi_n(s)}^* \mathrm{d}B_s$ and $A^n := \int_0^\cdot \alpha_{\varphi_n(s)} \Delta Y_s \mathrm{d}s$, where α, β, γ are three continuous adapted processes with $\alpha_t, \beta_t \in (\mathbb{R}^d)^{\otimes 2}$ and $\gamma_t \in \mathbb{R}^d$. In fact, in our applications, α, β, γ and Σ may be additionally approximated, but this extra approximation is presumably negligible for the next limit in distribution.

7.3 Main result

From Chapter 3, we have

$$N_T^n = O(\varepsilon_n^{-2}), \quad |\Delta Y|_* = O(\varepsilon_n^2), \qquad (H_{Y,N})$$

(i.e. $(\mathbf{A}_Y)$ and $(\mathbf{A}_N)$ with $\rho_N = 1$) and more precisely

$$\varepsilon_n^2 N_T^n \xrightarrow{a.s.} \int_0^T \mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s, Y_s))\mathrm{d}s.$$

The main result below is a Central Limit-type theorem.

Theorem 7.3.1. *Let $\mathcal{E}_T^n := M_T^n + A_T^n$, where $M^n := \int_0^\cdot \beta_{\varphi_n(s)} \Delta Y_s \gamma_{\varphi_n(s)}^* \mathrm{d}B_s$ and $A^n := \int_0^\cdot \alpha_{\varphi_n(s)} \Delta Y_s \mathrm{d}s$, with α, β, γ are three continuous adapted processes with*

$\alpha_t, \beta_t \in (\mathbb{R}^d)^{\otimes 2}$ and $\gamma_t \in \mathbb{R}^d$. Then, there exists an d-dimensional Brownian motion W defined on an extended probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and independent of B such that

$$\sqrt{N_T^n}\mathcal{E}_T^n \xrightarrow{\mathcal{L}} \sqrt{\int_0^T \mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s, Y_s))\mathrm{d}s}\left(A_T + \int_0^T Q_s \mathrm{d}B_s + \int_0^T z_s \mathrm{d}W_s\right).$$

where $zz^* := k$, Q and A are given in (7.4.4), (7.4.1) and (7.4.5) and depend on the values $(u_{kkk}^{\sigma\sigma^*(t,Y_t),\Sigma_t}(0))_{1\leq k\leq d}$ and $(u_{jklm}^{\sigma\sigma^*(t,Y_t),\Sigma_t}(0))_{1\leq j,k,l,m\leq d}$ at $y = 0$ of solutions to elliptic Dirichlet problems :

$$\begin{cases} \frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^*)(t, Y_t)^{ij}\partial_{i,j}^2 u_{kkk}^{\sigma\sigma^*(t,Y_t),\Sigma_t}(y) = 0, & y \in \xi_{0,\Sigma_t}, \\ u_{kkk}^{\sigma\sigma^*(t,Y_t),\Sigma_t}(y) = y_k^3, & y \in \partial\xi_{0,\Sigma_t}, \end{cases}$$

$$\begin{cases} \frac{1}{2}\sum_{i,i'=1}^n (\sigma\sigma^*)(t, Y_t)^{ii'}\partial_{i,i'}^2 u_{jklm}^{\sigma\sigma^*(t,Y_t),\Sigma_t}(y) = 0, & y \in \xi_{0,\Sigma_t}, \\ u_{jklm}^{\sigma\sigma^*(t,Y_t),\Sigma_t}(y) = y^j y^k y^l y^m, & y \in \partial\xi_{0,\Sigma_t}. \end{cases}$$

7.4 Proof.

7.4.1 Proof of the convergence in law.

We apply Theorem 7.1.1 to the local martingale $M^n - \int_0^\cdot Q_s \mathrm{d}B_s$, where

$$Q_s := \sum_{k=1}^d \beta_s^{\cdot k}[(\sigma\sigma^*)(s, Y_s)^{kk}]^{-1} u_{kkk}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\gamma_s^* \mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s, Y_s)). \tag{7.4.1}$$

• ***Item b)*** : $\langle M^n - \int_0^\cdot Q_s \mathrm{d}B_s, B\rangle_t \xrightarrow{a.s.} 0$ because

$$\left\langle \int_0^\cdot \varepsilon_n^{-1}\beta_{\varphi_n(s)}\Delta Y_s \gamma_{\varphi_n(s)}^* \mathrm{d}B_s, B\right\rangle_t = \sum_{k=1}^d \int_0^t \varepsilon_n^{-1}\beta_{\varphi_n(s)}^{\cdot k}\Delta Y_s^k \gamma_{\varphi_n(s)}^* \mathrm{d}s \xrightarrow{a.s.} \int_0^t Q_s \mathrm{d}s,$$

using Lemma 7.5.2.

• ***Item a)*** : we decompose the quadratic variation of $M^n - \int_0^\cdot Q_s \mathrm{d}B_s$ at time t into three terms, as follows,

$$\begin{aligned} \langle M^n - \int_0^\cdot Q_s \mathrm{d}B_s\rangle_t &= Q_{1,t}^n + Q_{2,t}^n + Q_{3,t}, \\ Q_{1,t}^n &= \varepsilon_n^{-2}\int_0^t (\gamma_{\varphi_n(s)}^*\gamma_{\varphi_n(s)})\beta_{\varphi_n(s)}\Delta Y_s \Delta Y_s^* \beta_{\varphi_n(s)}^* \mathrm{d}s, \\ Q_{2,t}^n &= -2\sum_{k=1}^d \int_0^t \varepsilon_n^{-1}\beta_{\varphi_n(s)}^{\cdot k}\gamma_{\varphi_n(s)}^* Q_s^* \Delta Y_s^k \mathrm{d}s, \\ Q_{3,t} &= \int_0^t Q_s Q_s^* \mathrm{d}s. \end{aligned}$$

Owing to Lemma 7.5.3, for any $(p,q) \in \{1,...,q\}^2$, we have

$$\begin{aligned}(Q^n_{1,t})^{pq} \xrightarrow{a.s.} \frac{1}{4+d} \sum_{j,k,l,m=1}^{d} \int_0^t (\gamma_s^*\gamma_s)\Bigg(& (\sigma\sigma^*)^{-1}(s,Y_s)^{lm}\beta_s^{pj}\beta_s^{kq} \\ & - \frac{1}{4+2d}\beta_s^{p\cdot}(\sigma\sigma^*)(s,Y_s)\beta_s^{\cdot q}(\sigma\sigma^*)^{-1}(s,Y_s)^{jk}(\sigma\sigma^*)^{-1}(s,Y_s)^{lm}\Bigg) \\ & \times u_{jklm}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s,Y_s))\mathrm{d}s \\ := \int_0^t Q_{1,s}^{pq}\mathrm{d}s. \end{aligned} \tag{7.4.2}$$

Owing to Lemma 7.5.2, we have

$$\begin{aligned} Q^n_{2,t} \xrightarrow{a.s.} & -2\sum_{k=1}^d \int_0^t \beta_s^{\cdot k}\gamma_s^* Q_s^*[(\sigma\sigma^*)(s,Y_s)^{kk}]^{-1}u_{kkk}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s,Y_s))\mathrm{d}s \\ := & -2\int_0^t Q_sQ_s^*\mathrm{d}s. \end{aligned} \tag{7.4.3}$$

Putting all the convergences together the convergences (7.4.2) and (7.4.3), we obtain

$$k_s := Q_{1,s} - Q_sQ_s^*. \tag{7.4.4}$$

Then, we get the convergence involved in *item a)*

$$\langle M^n - \int_0^\cdot Q_s\mathrm{d}B_s\rangle_t \xrightarrow{a.s.} \int_0^t k_s\mathrm{d}s.$$

• Last, it remains to compute the limit of the finite variation part A^n_T. Owing to Lemma 7.5.2, we have

$$\begin{aligned} A^n_T &:= \varepsilon_n^{-1}\sum_{k=1}^d \int_0^T \alpha^k_{\varphi_n(s)}\Delta Y_s^k\mathrm{d}s \\ &\xrightarrow{a.s.} \sum_{k=1}^d \int_0^T \alpha_s^k[(\sigma\sigma^*)(s,Y_s)^{kk}]^{-1}u_{kkk}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s,Y_s))\mathrm{d}s := A_T. \end{aligned} \tag{7.4.5}$$

• **Convergence** : moreover, $\varepsilon_n^2 N_T^n \xrightarrow{a.s.} \int_0^T \mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s,Y_s))\mathrm{d}s$ using the structure of the stopping times (see ($H_{Y,N}$)). Thus, using the stable convergence, we get

$$\sqrt{N_T^n}\mathcal{E}_T^n \xrightarrow{\mathcal{L}} \sqrt{\int_0^T \mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s,Y_s))\mathrm{d}s}\left(A_T + \int_0^T Q_s\mathrm{d}B_s + \int_0^T z_s\mathrm{d}W_s\right).$$

The proof is complete.

7.4.2 Technical results.

The following lemma gives the asymptotic behaviour of renormalized conditional expectations for the process Y at hitting times of random ellipsoids, when the ellipsoid radius converges to 0.

Lemma 7.4.1. *Let $(t_0, x_0) \in [0, T) \times \mathbb{R}^d$. Set $u(t, x) := \mathbb{E}\left(f(Y_{\tau\wedge T} - x_0)|Y_t = x\right)$ for $(t, x) \in [t_0, T] \times \mathbb{R}^d$, where f is a three times continuously differentiable, α-homogeneous function (i.e. $f \in \mathcal{C}^3(\mathbb{R}^d)$, $f(\lambda y) = \lambda^\alpha f(y)$, for any $\lambda > 0$ and $y \in \mathbb{R}^d$), with $\alpha > 0$, and $\tau := \inf\{s \geq t : Y_s \notin \xi_{x_0,\varepsilon_n^{-2}\Sigma_{t_0}}\}$ deterministic matrix for a given Σ_{t_0} satisfying $(\star\Sigma)$. Then,*

$$|\varepsilon_n^{-\alpha} u(t_0, x_0) - u_f^{\sigma\sigma^*(t_0,x_0),\Sigma_{t_0}}(0)| \leq c(\varepsilon_n^{\theta_{b,\sigma}} + e^{-c\varepsilon_n^{-2}(T-t_0)}),$$

where the constant c depends only on $C_\Sigma, d, |b|_\infty, \sigma^2_{\min}, \theta_{b,\sigma}, |\sigma|_\infty, C_{b,\sigma}, f$ and where $u_f^{\sigma\sigma^*(t_0,x_0),\Sigma_{t_0}}(.)$ is the solution to the elliptic Dirichlet problem (with constant coefficients):

$$\begin{cases} \frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^*)(t_0, x_0)^{ij} \partial^2_{i,j} u_f^{\sigma\sigma^*(t_0,x_0),\Sigma_{t_0}}(y) = 0, & y \in \xi_{0,\Sigma_{t_0}}, \\ u_f^{\sigma\sigma^*(t_0,x_0),\Sigma_{t_0}}(y) = f(y), & y \in \partial\xi_{0,\Sigma_{t_0}}. \end{cases}$$

Remark 7.4.1. *It is known that $u_f^{\sigma\sigma^*(t_0,x_0),\Sigma_{t_0}}(0)$ is continuous in $\sigma\sigma^*(t_0, x_0)$ (see [Gilbarg 1977][Chapter 6]) and in Σ_{t_0} (see [Simon 1980]).*

Proof. We know that u solves the parabolic Dirichlet problem (see [Costantini 2006][Proposition 2.1])

$$\begin{cases} \partial_t u(t, x) + \sum_{i=1}^d b(t, x)^i \partial_i u(t, x) + \frac{1}{2}\sum_{i,j=1}^d (\sigma\sigma^*)(t, x)^{ij} \partial^2_{i,j} u(t, x) = 0, \ \forall (t, x) \in (t_0, T) \times \xi_{x_0,\varepsilon_n^{-2}\Sigma_{t_0}}, \\ u(T, z) = f(x - x_0), \quad x \in \xi_{x_0,\varepsilon_n^{-2}\Sigma_{t_0}}, \\ u(t, z) = f(x - x_0), \quad (t, x) \in [t_0, T] \times \partial\xi_{x_0,\varepsilon_n^{-2}\Sigma_{t_0}}. \end{cases}$$

Set $v(t, y) = \varepsilon_n^{-\alpha} u(t_0 + \varepsilon_n^2(t - t_0), x_0 + \varepsilon_n y)$ for $t \in [t_0, T_n] \times \xi_{0,\Sigma_{t_0}}$ with $T_n = t_0 + \varepsilon_n^{-2}(T - t_0)$. Then, we have

$$\begin{aligned} \partial_t v(t, y) &= \varepsilon_n^{2-\alpha} \partial_t u(t_0 + \varepsilon_n^2(t - t_0), x_0 + \varepsilon_n y), \\ \partial_i v(t, y) &= \varepsilon_n^{1-\alpha} \partial_i u(t_0 + \varepsilon_n^2(t - t_0), x_0 + \varepsilon_n y), \\ \partial^2_{i,j} v(t, y) &= \varepsilon_n^{2-\alpha} \partial^2_{i,j} u(t_0 + \varepsilon_n^2(t - t_0), x_0 + \varepsilon_n y), \end{aligned}$$

and v solves the PDE : $\forall (t,y) \in [t_0, T_n) \times \xi_{0,\Sigma_{t_0}}$,

$$\begin{aligned}
&\partial_t v(t,y) + \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^*)(t_0 + \varepsilon_n^2(t-t_0), x_0 + \varepsilon_n y)^{ij} \partial^2_{i,j} v(t,y) \\
&= \varepsilon_n^{2-\alpha} \left[\partial_t u + \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^*)^{ij} \partial^2_{i,j} u \right] (t_0 + \varepsilon_n^2(t-t_0), x_0 + \varepsilon_n y) \\
&= -\varepsilon_n \sum_{i=1}^{d} b(t_0 + \varepsilon_n^2(t-t_0), x_0 + \varepsilon_n y)^i \partial_i v(t,y),
\end{aligned}$$

with the boundary condition :

$$v(T_n, y) = \varepsilon_n^{-\alpha} u(T, x_0 + \varepsilon_n y) = \varepsilon_n^{-\alpha} f(\varepsilon_n y) = f(y),$$

and for $(t,y) \in [t_0, T_n] \times \partial\xi_{0,\Sigma_{t_0}}$,

$$v(t,y) = \varepsilon_n^{-\alpha} u(t_0 + \varepsilon_n^2(t-t_0), x_0 + \varepsilon_n y) = f(y).$$

Additionally, we have the Feynman-Kac representation

$$v(t,y) = \mathbb{E}(f(Y^n_{\tau^{Y^n} \wedge T_n}) | Y^n_t = y),$$

where

$$\mathrm{d}Y^n_t = \varepsilon_n b(t_0 + \varepsilon_n^2(t-t_0), x_0 + \varepsilon_n Y^n_t) + \sigma(t_0 + \varepsilon_n^2(t-t_0), x_0 + \varepsilon_n Y^n_t)\mathrm{d}B_t,$$

and $\tau^{Y^n} := \inf\{s \geq t_0 : Y^n_s \notin \xi_{0,\Sigma_{t_0}}\}$.
Let z be the solution to the elliptic Dirichlet problem :

$$\begin{cases} \frac{1}{2}\sum_{i,j=1}^{n} (\sigma\sigma^*)(t_0, x_0)^{ij} \partial^2_{i,j} z(y) = 0, \quad y \in \xi_{0,\Sigma_{t_0}}, \\ z(y) = f(y), \quad y \in \partial\xi_{0,\Sigma_{t_0}}. \end{cases}$$

With these notations at hand, the lemma is proved if we show $|v(t_0, 0) - z(0)| \leq c(\varepsilon_n^{\theta_{b,\sigma}} + e^{-c\varepsilon_n^{-2}(T-t_0)})$. Therefore, consider the solution Y^n starting from $Y^n_{t_0} = 0$ and write

$$\begin{aligned}
v(t_0, 0) &= \mathbb{E}\left[v(T_n \wedge \tau^{Y^n}, Y^n_{T_n \wedge \tau^{Y^n}}) \right] \\
&= \mathbb{E}\left[z(Y^n_{T_n \wedge \tau^{Y^n}}) 1_{\tau^{Y^n} < T_n} + v(T_n, Y^n_{T_n}) 1_{\tau^{Y^n} \geq T_n} \right] \\
&= \mathbb{E}\left[z(Y^n_{T_n \wedge \tau^{Y^n}}) + (v(T_n, Y^n_{T_n}) - z(Y^n_{T_n \wedge \tau^{Y^n}})) 1_{\tau^{Y^n} \geq T_n} \right].
\end{aligned}$$

First of all, we have

$$\left| \mathbb{E}\left[(v(T_n, Y^n_{T_n}) - z(Y^n_{T_n \wedge \tau^{Y^n}})) 1_{\tau^{Y^n} \geq T_n} \right] \right| \leq 2 \sup_{x \in \xi_{0,\Sigma_{t_0}}} |f(x)| \mathbb{P}(\tau^{Y^n} \geq T_n).$$

Secondly, we have

$$\begin{aligned}
&\mathbb{E}\left[z(Y^n_{T_n\wedge\tau^{Y^n}})\right]\\
&= z(0)+\mathbb{E}\left[\int_{t_0}^{T_n\wedge\tau^{Y^n}}\left(\frac{1}{2}\sum_{i,j=1}^d(\sigma\sigma^*)(t_0+\varepsilon_n^2(s-t_0),x_0+\varepsilon_nY^n_s)^{ij}\partial^2_{i,j}z(Y^n_s)\right.\right.\\
&\quad\left.\left.+\varepsilon_n\sum_{i=1}^d b(t_0+\varepsilon_n^2(s-t_0),x_0+\varepsilon_nY^n_s)^i\partial_iz(Y^n_s)\right)\mathrm{d}s\right]\\
&= z(0)+\mathbb{E}\left[\int_{t_0}^{T_n\wedge\tau^{Y^n}}\left(\frac{1}{2}\sum_{i,j=1}^d\left\{(\sigma\sigma^*)(t_0+\varepsilon_n^2(s-t_0),x_0+\varepsilon_nY^n_s)^{ij}\right.\right.\right.\\
&\quad\left.\left.\left.-(\sigma\sigma^*)(t_0,x_0)^{ij}\right\}\partial^2_{i,j}z(Y^n_s)+\varepsilon_n\sum_{i=1}^d b(t_0+\varepsilon_n^2(s-t_0),x_0+\varepsilon_nY^n_s)^i\partial_iz(Y^n_s)\right)\mathrm{d}s\right].
\end{aligned}$$

Then, we get the estimate

$$\begin{aligned}
\left|\mathbb{E}\left[z(Y^n_{T_n\wedge\tau^{Y^n}})-z(0)\right]\right| &\leq \mathbb{E}\left[\int_{t_0}^{T_n\wedge\tau^{Y^n}}\left(\frac{1}{2}\sum_{i,j=1}^d\sup_{0\leq t\leq T,x\in\xi_{0,\Sigma_{t_0}}}|\sigma(t,x_0+\varepsilon_nx)^{ij}|C_d\right.\right.\\
&\quad\varepsilon_n^{\theta_{b,\sigma}}\left\{(s-t_0)^{\theta_{b,\sigma}/2}+|Y^n_s|^{\theta_{b,\sigma}}\right\}|\partial^2_{i,j}z|_\infty\\
&\quad\left.\left.+\varepsilon_n\sum_{i=1}^d\sup_{0\leq t\leq T,x\in\xi_{0,\Sigma_{t_0}}}|b(t,x_0+\varepsilon_nx)^i||\partial_iz|_\infty\right)\mathrm{d}s\right]\\
&\leq C\varepsilon_n^{\theta_{b,\sigma}}\mathbb{E}\left((T^n\wedge\tau^{Y^n}-t_0)^{1+\theta_{b,\sigma}/2}+1\right),
\end{aligned}$$

where the derivatives of z are bounded owing to [Gilbarg 1977][Theorem 6.6] and $|Y^n_s|^{\theta_{b,\sigma}}\leq\left(\frac{(Y^n_s)^*\Sigma_{t_0}Y^n_s}{\lambda_{\min}(\Sigma_{t_0})}\right)^{\theta_{b,\sigma}/2}\leq C_\Sigma^{\theta_{b,\sigma}/2}$.
Then, we deduce that

$$|v(t_0,0)-z(0)|\leq 2\sup_{x\in\xi_{0,\Sigma_{t_0}}}|f(x)|\mathbb{P}(\tau^{Y^n}\geq T_n)+C\varepsilon_n^{\theta_{b,\sigma}}\mathbb{E}\left((T^n\wedge\tau^{Y^n}-t_0)^{1+\theta_{b,\sigma}/2}+1\right).$$

Moreover, by [Freidlin 1985][Chapter 3.3], we have

$$\mathbb{P}(\tau^{Y^n}>t_0+\lambda)\leq ce^{-c\lambda},$$

where c is uniform in n and depends only on $C_\Sigma,d,|b|_\infty,\sigma^2_{\min}$.
Then, we have

$$\begin{aligned}
\sup_{n\geq0}\mathbb{E}[T_n\wedge\tau^{Y^n}-t_0]^p &= p\sup_{n\geq0}\int_0^{+\infty}\lambda^{p-1}\mathbb{P}(\tau^{Y^n}-t_0>\lambda)\mathrm{d}\lambda\\
&\leq cp\int_0^{+\infty}\lambda^{p-1}e^{-c\lambda}\mathrm{d}\lambda:=K(p)<+\infty.
\end{aligned}$$

We conclude that

$$\begin{aligned}|\varepsilon_n^{-\alpha}u(t_0,x_0)-u_f^{\sigma\sigma^*(t_0,x_0),\Sigma_{t_0}}(0)| \leq & 2\sup_{x\in\xi_{0,\Sigma_{t_0}}}|f(x)|ce^{-c\varepsilon_n^{-2}(T-t_0)}\\ &+C\varepsilon_n^{\theta_{b,\sigma}}(K(1+\theta_{b,\sigma}/2)+1).\end{aligned}$$

□

The following martingale type result is useful in our analysis to replace a function of the Y-increments by its conditional expectations.

Lemma 7.4.2. *Let f be a continuous α-homogeneous function (i.e. $\forall y\in\mathbb{R}^d,\forall\lambda>0, f(\lambda y)=\lambda^\alpha f(y)$) with $\alpha>0$. Let H be a continuous adapted scalar process. Then,*

$$\varepsilon_n^{2-\alpha}\sum_{\tau_{i-1}^n<t}H_{\tau_{i-1}^n}\left(f(\Delta Y_{\tau_i^n\wedge t})-\mathbb{E}_{\tau_{i-1}^n}(f(\Delta Y_{\tau_i^n\wedge t}))\right)=o(1),$$

a.s. , *for any $t\in[0,T]$.*

Proof. First assume that $|H|_*\in L_p$ for any $p>0$, so that $\sum_{\tau_{i-1}^n<t}\sup_{\tau_{i-1}^n\leq s\leq\tau_i^n\wedge t}H_{\varphi_n(s)}|f(\Delta Y_s)-\mathbb{E}_{\varphi_n(s)}(f(\Delta Y_s))|\in L_p$ for any $p\geq 1$ (because $|H|_*\in L_p$, $N_T^n\in L_p$ and (7.4.6)) and all Fubini manipulations are justified. We show the convergence to 0 leveraging the fundamental lemma 3.2.2; set

$$\begin{aligned}U_t^n &:= \varepsilon_n^{4-2\alpha}\Big|\sum_{\tau_{i-1}^n<t}H_{\tau_{i-1}^n}\left\{f(\Delta Y_{\tau_i^n\wedge t})-\mathbb{E}_{\tau_{i-1}^n}f(\Delta Y_{\tau_i^n\wedge t})\right\}\Big|^2,\\ V_t^n &:= \varepsilon_n^{4-2\alpha}\sum_{\tau_{i-1}^n<t}H_{\tau_{i-1}^n}^2\sup_{\tau_{i-1}^n\leq s\leq\tau_i^n\wedge t}\left|f(\Delta Y_s)-\mathbb{E}_{\varphi_n(s)}f(\Delta Y_s)\right|^2.\end{aligned}$$

The sequences of processes $(U^n)_{n\geq0}$ and $(V^n)_{n\geq0}$ are in $\mathcal{C}_0^+$. Then, *i)* V^n is non-decreasing and *ii)* using ($H_{Y,N}$), we have *a.s.* , for any $s\in[0,T]$,

$$\begin{aligned}&|f(\Delta Y_s)-\mathbb{E}_{\varphi_n(s)}f(\Delta Y_s)|\\ &=\left||\Delta Y_s|^\alpha f\left(\frac{\Delta Y_s}{|\Delta Y_s|}\right)-\mathbb{E}_{\varphi_n(s)}\left(|\Delta Y_s|^\alpha f\left(\frac{\Delta Y_s}{|\Delta Y_s|}\right)\right)\right|\\ &\leq C_f\left|\left(\frac{(Y_s-Y_{\varphi_n(s)})^*\Sigma_{\varphi_n(s)}(Y_s-Y_{\varphi_n(s)})}{\lambda_{\min}(\Sigma_{\varphi_n(s)})}\right)^{\alpha/2}+\mathbb{E}_{\varphi_n(s)}\left(\frac{(Y_s-Y_{\varphi_n(s)})^*\Sigma_{\varphi_n(s)}(Y_s-Y_{\varphi_n(s)})}{\lambda_{\min}(\Sigma_{\varphi_n(s)})}\right)^{\alpha/2}\right|\\ &\leq 2C_fC_\Sigma^{\alpha/2}\varepsilon_n^\alpha,\end{aligned}\tag{7.4.6}$$

where $C_f:=\sup_{|x|=1}|f(x)|$. We easily get $\sum_{n\geq0}V_T^n<+\infty$ *a.s.*. The first two items are satisfied. Now, we have to check the relation of domination (item *iii)*). Let $k\in\mathbb{N}$. On the set $\{\tau_{i-1}^n<t\wedge\theta_k\}$, we are going to develop the square in the

expression of U^n and we use that θ_k is a stopping time to remove the cross product terms :

$$\begin{aligned}
\mathbb{E}\left[U^n_{t\wedge\theta_k}\right] &= \varepsilon_n^{4-2\alpha}\mathbb{E}\left[\sum_{\tau^n_{i-1}<t\wedge\theta_k} H^2_{\tau^n_{i-1}}\left\{f(\Delta Y_{\tau^n_i\wedge t\wedge\theta_k}) - \mathbb{E}_{\tau^n_{i-1}} f(\Delta Y_{\tau^n_i\wedge t\wedge\theta_k})\right\}^2\right] \\
&+ 2\varepsilon_n^{4-2\alpha}\sum_{1\le i<i'<+\infty}\mathbb{E}\Bigg(1_{\tau^n_{i'-1}<t\wedge\theta_k} H_{\tau^n_{i-1}}\left\{f(\Delta Y_{\tau^n_i\wedge t\wedge\theta_k}) - \mathbb{E}_{\tau^n_{i-1}} f(\Delta Y_{\tau^n_i\wedge t\wedge\theta_k})\right\}\times \\
&\qquad H_{\tau^n_{i'-1}}\mathbb{E}_{\tau^n_{i'-1}}\left[f(\Delta Y_{\tau^n_{i'}\wedge t\wedge\theta_k}) - \mathbb{E}_{\tau^n_{i'-1}} f(\Delta Y_{\tau^n_{i'}\wedge t\wedge\theta_k})\right]\Bigg) \\
&\le \mathbb{E}[V^n_{t\wedge\theta_k}].
\end{aligned}$$

The result follows and we are done.

Now, we remove the assumption $|H|_* \in L_p$, by considering $H^M_t = -M \vee H_t \wedge M$ for $M > 0$ for which $|H^M|_*$ is obviously in L_p. If we write $L^{M,n}_t = \varepsilon_n^{2-\alpha}\sum_{\tau^n_{i-1}<t} H^M_{\tau^n_{i-1}}\Big(f(\Delta Y_{\tau^n_i\wedge t}) - \mathbb{E}_{\tau^n_{i-1}}(f(\Delta Y_{\tau^n_i\wedge t}))\Big)$, we have proved that for any $M\in\mathbb{N}$, $L^{M,n}_t \xrightarrow{a.s.} 0$. Thus, *a.s.* for any $M\in\mathbb{N}$, $L^{M,n}_t \to 0$. But, the sequence $(L^{M,n}_t)_{M\in\mathbb{N}}$ is stationary for M large enough, uniformly in n and t : we are finished. $\square$

7.5 Appendix

Lemma 7.5.1. *Let $(t_0, x_0) \in (0,T)\times\mathbb{R}^d$ and $t \in]t_0, T]$. Set $u(r,x) := \mathbb{E}\left(f(Y_{\tau\wedge t} - x_0)|Y_r = x\right)$, where f is a three times continuously differentiable, α-homogeneous function ($\alpha > 0$) and $\tau := \inf\{s \ge r : Y_s \notin \xi_{x_0,\varepsilon_n^{-2}\Sigma_{t_0}}\}$. Let H be a scalar continuous adapted process. Then,*

$$\sum_{\tau^n_{i-1}<t} H_{\tau^n_{i-1}}\varepsilon_n^{2-\alpha}u(\tau^n_{i-1}, Y_{\tau^n_{i-1}}) \xrightarrow{a.s.} \int_0^t H_s u_f^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\mathrm{Tr}(\Sigma_s \mathrm{d}\langle Y\rangle_s),$$

where $u_f^{\sigma\sigma^*(s,Y_s),\Sigma_s}(.)$ is the solution to the elliptic Dirichlet problem :

$$\begin{cases} \frac{1}{2}\sum_{i,j=1}^n(\sigma\sigma^*)(s,Y_s)^{ij}\partial^2_{i,j}u_f^{\sigma\sigma^*(s,Y_s),\Sigma_s}(y) = 0, \quad y\in\xi_{0,\Sigma_s}, \\ u_f^{\sigma\sigma^*(s,Y_s),\Sigma_s}(y) = f(y), \quad y\in\partial\xi_{0,\Sigma_s}. \end{cases}$$

Proof. Using the form of the hitting times, we have

$$\begin{aligned}
&\sum_{\tau^n_{i-1}<t} H_{\tau^n_{i-1}}\varepsilon_n^{2-\alpha}u(\tau^n_{i-1}, Y_{\tau^n_{i-1}}) \\
&= \sum_{\tau^n_{i-1}<t} H_{\tau^n_{i-1}}\varepsilon_n^{-\alpha}u(\tau^n_{i-1}, Y_{\tau^n_{i-1}})\Delta Y^*_{\tau^n_i\wedge t}\Sigma_{\tau^n_{i-1}}\Delta Y_{\tau^n_i\wedge t} \\
&\quad - H_{\varphi_n(t)}\varepsilon_n^{-\alpha}u(\varphi_n(t), Y_{\varphi_n(t)})\Delta Y^*_t\Sigma_{\varphi_n(t)}\Delta Y_t + H_{\varphi_n(t)}\varepsilon_n^{2-\alpha}u(\varphi_n(t), Y_{\varphi_n(t)}).
\end{aligned}$$

The two last terms converge to 0. We write the first term as

$$
\begin{aligned}
&\left|\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\varepsilon_n^{-\alpha}u(\tau_{i-1}^n, Y_{\tau_{i-1}^n})\Delta Y^*_{\tau_i^n\wedge t}\Sigma_{\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t}\right| \\
&= \int_0^t H_{\varphi_n(s)} u_f^{\sigma\sigma^*(\varphi_n(s),Y_{\varphi_n(s)}),\Sigma_{\varphi_n(s)}}(0)\mathrm{Tr}(\Sigma_{\varphi_n(s)}\mathrm{d}\langle Y\rangle_s) \\
&\quad + \int_0^t H_{\varphi_n(s)}\left(\varepsilon_n^{-\alpha}u(\varphi_n(s), Y_{\varphi_n(s)}) - u_f^{\sigma\sigma^*(\varphi_n(s),Y_{\varphi_n(s)}),\Sigma_{\varphi_n(s)}}(0)\right)\mathrm{Tr}(\Sigma_{\varphi_n(s)}\mathrm{d}\langle Y\rangle_s) \\
&\quad + \sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}\varepsilon_n^{-\alpha}u(\tau_{i-1}^n, Y_{\tau_{i-1}^n})(\Delta Y^*_{\tau_i^n\wedge t}\Sigma_{\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t} - \mathrm{Tr}(\Sigma_{\tau_{i-1}^n}\Delta\langle Y\rangle_{\tau_i^n\wedge t})).
\end{aligned}
$$

The first term converges to $\int_0^t H_s u_f^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\mathrm{Tr}(\Sigma_s\mathrm{d}\langle Y\rangle_s)$ by Lebesgue's domination theorem, because of the continuity and boundedness of $(r,x) \mapsto u_f^{\sigma\sigma^*(r,x),\Sigma_r}(0)$, H and the convergence to 0 of the mesh size $\mathcal{T}^n$. Using Ito's lemma, the third term is equal to

$$
2\int_0^t H_{\varphi_n(s)}\varepsilon_n^{-\alpha}u(\varphi_n(s), Y_{\varphi_n(s)})\Sigma_{\varphi_n(s)}\Delta Y_s\mathrm{d}Y_s.
$$

We conclude the convergence to 0 owing to Lemma 6.3.2. For the second term, we apply the estimate derived in Lemma 7.4.1 :

$$
\begin{aligned}
&\left|\int_0^t H_{\varphi_n(s)}\left(\varepsilon_n^{-\alpha}u(\varphi_n(s), Y_{\varphi_n(s)}) - u_f^{\sigma\sigma^*(\varphi_n(s),Y_{\varphi_n(s)}),\Sigma_{\varphi_n(s)}}(0)\right)\mathrm{Tr}(\Sigma_{\varphi_n(s)}\mathrm{d}\langle Y\rangle_s)\right| \\
&\le c\int_0^t |H_{\varphi_n(s)}|(e^{-c\varepsilon_n^{-2}(t-\varphi_n(s))} + \varepsilon_n^{\theta_{b,\sigma}})\mathrm{Tr}(\Sigma_{\varphi_n(s)}\mathrm{d}\langle Y\rangle_s) \\
&\le c\int_0^t |H_{\varphi_n(s)}|(e^{-c\rho(t-\varphi_n(s))} + \varepsilon_n^{\theta_{b,\sigma}})\mathrm{Tr}(\Sigma_{\varphi_n(s)}\mathrm{d}\langle Y\rangle_s),
\end{aligned}
$$

for any $\rho > 0$, provided that $\rho \le \varepsilon_n^{-2}$ (n large enough). The latter upper bound converges *a.s.* to $\int_0^t |H_s|e^{-c\rho(t-s)}\mathrm{Tr}(\Sigma_s(\sigma\sigma^*)(s,Y_s))\mathrm{d}s \le c'/\rho$. Since ρ is arbitrary large, we conclude to the convergence to 0.

□

Lemma 7.5.2. *Let H be a scalar continuous adapted process. Then, $\forall t \in [0,T]$, $\forall k = 1, ..., d$,*

$$
\varepsilon_n^{-1}\int_0^t H_{\varphi_n(s)}\Delta Y_s^k\mathrm{d}s \overset{a.s.}{\longrightarrow} \int_0^t H_s[(\sigma\sigma^*)(s,Y_s)^{kk}]^{-1}u_{kkk}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\mathrm{Tr}(\Sigma_s\mathrm{d}\langle Y\rangle_s),
$$

where $u_{kkk}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(.)$ is the solution to the stationary Dirichlet problem :

$$
\begin{cases}
\frac{1}{2}\sum_{i,j=1}^n(\sigma\sigma^*)(s,Y_s)^{ij}\partial^2_{i,j}u_{kkk}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(y) = 0, \quad y \in \xi_{0,\Sigma_s}, \\
u_{kkk}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(y) = y_k^3, \quad y \in \partial\xi_{0,\Sigma_s}.
\end{cases}
$$

Proof. To simplify the presentation, we note $\sigma_s := \sigma(s, Y_s)$. Firstly, we use Ito's formula and get

$$\varepsilon_n^{-1}\int_0^t H_{\varphi_n(s)}\Delta Y_s^k \mathrm{d}s = \varepsilon_n^{-1}\sum_{\tau_{i-1}^n<t} H_{\tau_{i-1}^n}[(\sigma\sigma^*)_{\tau_{i-1}^n}^{kk}]^{-1}\Bigg[\frac{1}{3}\mathbb{E}_{\tau_{i-1}^n}(\Delta Y^k_{\tau_i^n\wedge t})^3 - \int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}(\Delta Y_s^k)^2\mathrm{d}Y_s^k$$
$$+\frac{1}{3}\left[(\Delta Y^k_{\tau_i^n\wedge t})^3 - \mathbb{E}_{\tau_{i-1}^n}(\Delta Y^k_{\tau_i^n\wedge t})^3\right] - \int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta Y_s^k\Delta((\sigma\sigma^*)_s^{kk})\mathrm{d}s\Bigg].$$

The second term converges to 0 thanks to Lemma 6.3.2. The last term clearly gives a global contribution like $O(\varepsilon_n^{\theta_{b,\sigma}})$ by the assumption on σ. For the third term, we use Lemma 7.4.2 with $f(x) = x^3$ and get directly the convergence to 0. The first term converges and gives the announced limit, by Lemma 7.5.1 applied to $f(z) := z_k^3$, which concludes the proof. □

Lemma 7.5.3. *Let $H_{1,.}, H_{2,.}$ be two continuous adapted processes with values in $\mathbb{R}^d$. Then, $\forall t\in[0,T]$,*

$$\varepsilon_n^{-2}\int_0^t H^*_{1,\varphi_n(s)}\Delta Y_s\Delta Y_s^* H_{2,\varphi_n(s)}\mathrm{d}s$$
$$\xrightarrow{a.s.}\sum_{j,k,l,m=1}^d\int_0^t\Bigg((\sigma\sigma^*)^{-1}(s,Y_s)^{lm}H^j_{1,s}H^k_{2,s} - \frac{1}{4+2d}H^*_{1,s}(\sigma\sigma^*)(s,Y_s)H_{2,s}$$
$$(\sigma\sigma^*)^{-1}(s,Y_s)^{jk}(\sigma\sigma^*)^{-1}(s,Y_s)^{lm}\Bigg)u_{jklm}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(0)\mathrm{Tr}(\Sigma_s\mathrm{d}\langle Y\rangle_s),$$

where $u_{jklm}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(.)$ is the solution to the elliptic Dirichlet problem :

$$\begin{cases}\frac{1}{2}\sum_{i,i'=1}^n(\sigma\sigma^*)(s,Y_s)^{ii'}\partial^2_{i,i'}u_{jklm}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(y) = 0, & y\in\xi_{0,\Sigma_s},\\ u_{jklm}^{\sigma\sigma^*(s,Y_s),\Sigma_s}(y) = y^jy^ky^ly^m, & y\in\partial\xi_{0,\Sigma_s}.\end{cases}$$

Proof. For the sake of convenience, we note $\sigma_s := \sigma(s, Y_s)$. Owing to Ito's formula, we have

$$\sum_{j,k=1}^d\sum_{\tau_{i-1}^n<t}H^j_{1,\tau_{i-1}^n}H^k_{2,\tau_{i-1}^n}\Delta Y^*_{\tau_i^n\wedge t}(\sigma\sigma^*)^{-1}_{\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t}\Delta Y^j_{\tau_i^n\wedge t}\Delta Y^k_{\tau_i^n\wedge t}$$
$$=\sum_{j,k=1}^d\int_0^t H^j_{1,\varphi_n(s)}H^k_{2,\varphi_n(s)}\Delta Y_s^j\Delta Y_s^k\left\{2\Delta Y_s^*(\sigma\sigma^*)^{-1}_{\varphi_n(s)}\mathrm{d}Y_s + \mathrm{Tr}((\sigma\sigma)^{-1}_{\varphi_n(s)}(\sigma\sigma^*)_s)\mathrm{d}s\right\}$$
$$+\sum_{j,k=1}^d\int_0^t H^j_{1,\varphi_n(s)}H^k_{2,\varphi_n(s)}\Delta Y_s^*(\sigma\sigma^*)^{-1}_{\varphi_n(s)}\Delta Y_s\left\{\Delta Y_s^j\mathrm{d}Y_s^k + \Delta Y_s^k\mathrm{d}Y_s^j + (\sigma\sigma^*)_s^{jk}\mathrm{d}s\right\}$$
$$+2\sum_{j,k,l,m=1}^d\int_0^t H^j_{1,\varphi_n(s)}H^k_{2,\varphi_n(s)}((\sigma\sigma^*)^{-1}_{\varphi_n(s)})^{lm}\Delta Y_s^l\left\{(\sigma\sigma^*)_s^{km}\Delta Y_s^j + (\sigma\sigma^*)_s^{jm}\Delta Y_s^k\right\}\mathrm{d}s.$$

All the stochastic integrals involving $\mathrm{d}Y_s$ give a global contribution like $O(\varepsilon_n^{3-\rho})$ for any $\rho > 0$ owing to Lemma 6.3.2. In all the Lebesgue-Stieltjes integrals, we can replace $(\sigma\sigma^*)_s$ by $(\sigma\sigma)_{\varphi_n(s)}$, because the error induced has a contribution like $O(\varepsilon_n^{2+\theta_\sigma})$. Then, all these terms contribute like $o(\varepsilon_n^2)$ and using that $\sum_{m=1}^d((\sigma\sigma^*)^{-1}_{\varphi_n(s)})^{lm}(\sigma\sigma^*)^{jm}_{\varphi_n(s)} = 1_{j=m}$ by definition, we have

$$\begin{aligned}
&\sum_{\tau_{i-1}^n<t} \Delta Y^*_{\tau_i^n\wedge t}(\sigma\sigma^*)^{-1}_{\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t}H^*_{1,\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t}\Delta Y^*_{\tau_i^n\wedge t}H_{2,\tau_{i-1}^n}\\
&= (4+d)\int_0^t H^*_{1,\varphi_n(s)}\Delta Y_s\Delta Y_s^* H_{2,\varphi_n(s)}\mathrm{d}s\\
&\quad+\int_0^t \Delta Y_s^*(\sigma\sigma^*)^{-1}_{\varphi_n(s)}\Delta Y_s H^*_{1,\varphi_n(s)}(\sigma\sigma^*)_{\varphi_n(s)}H_{2,\varphi_n(s)}\mathrm{d}s + o(\varepsilon_n^2).
\end{aligned}$$

Classically now, as in the previous chapters 3 and 6, we seek a representation of the term $\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta Y_s^*(\sigma\sigma^*)^{-1}_{\varphi_n(s)}\Delta Y_s\mathrm{d}s$ in the form $\big(\Delta Y^*_{\tau_i^n\wedge t}\Gamma_{\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t}\big)^2$ plus a stochastic integral, where Γ is a sequence of adapted $d\times d$-matrix process. Applying Itô's formula on each interval $[\tau_{i-1}^n, \tau_i^n\wedge t]$, $i=1,...,N_T^n$, we obtain

$$\begin{aligned}
\big(\Delta Y^*_{\tau_i^n\wedge t}\Gamma_{\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t}\big)^2 = &\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta Y_s^*\Big(2\mathrm{Tr}\big(\Gamma_{\varphi_n(s)}\sigma_{\varphi_n(s)}\sigma^*_{\varphi_n(s)}\big)\Gamma_{\varphi_n(s)}\\
&+\big(\Gamma_{\varphi_n(s)}+(\Gamma_{\varphi_n(s)})^*\big)\sigma_{\varphi_n(s)}\sigma^*_{\varphi_n(s)}\big(\Gamma_{\varphi_n(s)}+\Gamma^*_{\varphi_n(s)}\big)\Big)\Delta Y_s\mathrm{d}s\\
&+\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta Y_s^*\Big(2\mathrm{Tr}\big(\Gamma_{\varphi_n(s)}\Delta(\sigma_s\sigma_s^*)\big)\Gamma_{\varphi_n(s)}\\
&+\big(\Gamma_{\varphi_n(s)}+\Gamma^*_{\varphi_n(s)}\big)\Delta(\sigma_s\sigma_s^*)\big(\Gamma_{\varphi_n(s)}+\Gamma^*_{\varphi_n(s)}\big)\Big)\Delta Y_s\mathrm{d}s\\
&+2\int_{\tau_{i-1}^n}^{\tau_i^n\wedge t}\Delta Y_s^*\Gamma_{\varphi_n(s)}\Delta Y_s\Delta Y_s^*\big(\Gamma_{\varphi_n(s)}+\Gamma_{\varphi_n(s)}\big)\mathrm{d}Y_s.
\end{aligned}$$

Then, noting $x^{I_d} = \sigma^*\Gamma\sigma$, we seek a solution to the non-linear system :

$$2\mathrm{Tr}\big(x^{I_d}_{\varphi_n(t)}\big)x^{I_d}_{\varphi_n(t)} + \big(x^{I_d}_{\varphi_n(t)}+(x^{I_d}_{\varphi_n(t)})^*\big)^2 = I_d. \tag{7.5.1}$$

We easily check that $x^{I_d} := \frac{1}{\sqrt{4+2d}}I_d$ is solution to (7.5.1) (and this is the only one in $S_+^d(\mathbb{R})$ by Lemma 3.3.1). Then, we can take

$$\Gamma = \frac{1}{\sqrt{4+2d}}(\sigma\sigma^*)^{-1}.$$

Moreover, following the error analysis of Chapter 6, the estimate

$$\begin{aligned}
&\frac{1}{4+2d}\sum_{\tau_{i-1}^n<t} H^*_{1,\tau_{i-1}^n}(\sigma\sigma^*)_{\tau_{i-1}^n}H_{2,\tau_{i-1}^n}\big(\Delta Y^*_{\tau_i^n\wedge t}(\sigma\sigma^*)^{-1}_{\tau_{i-1}^n}\Delta Y_{\tau_i^n\wedge t}\big)^2\\
&=\int_0^t H^*_{1,\varphi_n(s)}(\sigma\sigma^*)_{\varphi_n(s)}H_{2,\varphi_n(s)}\Delta Y_s^*(\sigma\sigma^*)^{-1}_{\varphi_n(s)}\Delta Y_s\mathrm{d}s + o(\varepsilon_n^2)
\end{aligned}$$

holds.
Therefore, we have

$$
\begin{aligned}
&\int_0^t H^*_{1,\varphi_n(s)} \Delta Y_s \Delta Y^*_s H_{2,\varphi_n(s)} \mathrm{d}s \\
&= \frac{1}{4+d} \sum_{\tau^n_{i-1}<t} \Bigg(\Delta Y^*_{\tau^n_i \wedge t} (\sigma\sigma^*)^{-1}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i \wedge t} H^*_{1,\tau^n_{i-1}} \Delta Y_{\tau^n_i \wedge t} \Delta Y^*_{\tau^n_i \wedge t} H_{2,\tau^n_{i-1}} \\
&\qquad - \frac{1}{4+2d} H^*_{1,\tau^n_{i-1}} (\sigma\sigma^*)_{\tau^n_{i-1}} H_{2,\tau^n_{i-1}} \big(\Delta Y^*_{\tau^n_i \wedge t} (\sigma\sigma^*)^{-1}_{\tau^n_{i-1}} \Delta Y_{\tau^n_i \wedge t} \big)^2 \Bigg) + o(\varepsilon_n^2).
\end{aligned}
$$

Actually, we can take conditional expectation inside the sum of the main term, because $\varepsilon_n^{-2} \sum_{\tau^n_{i-1}<t} ((...) - \mathbb{E}_{\tau^n_{i-1}}(...)) \xrightarrow{a.s.} 0$ owing to Lemma 7.4.2 with $\alpha = 4$, where (...) stands for the term inside the sum. Now, applying Lemma 7.5.1 to $f(z) = z_j z_k z_l z_m$, we conclude that

$$
\begin{aligned}
&\sum_{j,k,l,m=1}^{d} \sum_{\tau^n_{i-1}<t} \Bigg(((\sigma\sigma^*)^{-1}_{\tau^n_{i-1}})^{lm} H^j_{1,\tau^n_{i-1}} H^k_{2,\tau^n_{i-1}} \\
&\qquad - \frac{1}{4+2d} H^*_{1,\tau^n_{i-1}} (\sigma\sigma^*)_{\tau^n_{i-1}} H_{2,\tau^n_{i-1}} ((\sigma\sigma^*)^{-1}_{\tau^n_{i-1}})^{jk} ((\sigma\sigma^*)^{-1}_{\tau^n_{i-1}})^{lm} \Bigg) \\
&\qquad \times \varepsilon_n^{-2} \mathbb{E}_{\tau^n_{i-1}} [\Delta Y^j_{\tau^n_i \wedge t} \Delta Y^k_{\tau^n_i \wedge t} \Delta Y^l_{\tau^n_i \wedge t} \Delta Y^m_{\tau^n_i \wedge t}] \\
&\xrightarrow{a.s.} \sum_{j,k,l,m=1}^{d} \int_0^t \Bigg(((\sigma\sigma^*)^{-1}_s)^{lm} H^j_{1,s} H^k_{2,s} - \frac{1}{4+2d} H^*_{1,s} (\sigma\sigma^*)_s H_{2,s} ((\sigma\sigma^*)^{-1}_s)^{jk} ((\sigma\sigma^*)^{-1}_s)^{lm} \Bigg) \\
&\qquad u^{\sigma\sigma^*(s,Y_s),\Sigma_s}_{jklm}(0) \mathrm{Tr}(\Sigma_s \mathrm{d}\langle Y \rangle_s),
\end{aligned}
$$

where $u^{\sigma\sigma^*(s,Y_s),\Sigma_s}_{jklm}(.)$ is the solution to the elliptic Dirichlet problem :

$$
\begin{cases}
\frac{1}{2} \sum_{i,i'=1}^n (\sigma\sigma^*)(s,Y_s)^{ii'} \partial^2_{i,i'} u^{\sigma\sigma^*(s,Y_s),\Sigma_s}_{jklm}(y) = 0, \quad y \in \xi_{0,\Sigma_s}, \\
u^{\sigma\sigma^*(s,Y_s),\Sigma_s}_{jklm}(y) = y^j y^k y^l y^m, \quad y \in \partial\xi_{0,\Sigma_s}.
\end{cases}
$$

□

Chapter 8

Price expansion for spread option

Contents

8.1 Introduction.

Spread options are becoming increasingly important in energy and commodity markets. We present a fast approximative (but accurate), analytic pricing formula as in [Gobet 2011b]. The formula is based on a new stochastic calculus technique introduced in [Gobet 2011b] and is a direct generalisation to several assets and local volatility models of the famous Bjerksund and Stensland's formula and Kirk's formula. [Alos 2011] generalizes the formula of Kirk for three assets, under log-normal assumptions. So, we go one step further, regarding the number of assets and the models.
Taking advantage of the approach given in [Gobet 2011b] and of the linearity of the payoff, we derive an expansion of option price involving a difference of averages of stochastic processes. The implicit strategy of the Kirk formula is to approximate each average by a log-normal random variable and then use Margrabe formula for

exchange options. Actually, we expand around a generalization of the Bjerksund and Stensland formula. The accurate improvement comes from the expansion of the probability of the exercise region under different probabilities arising in the changes of numeraire. Then, the exercise region is approximated by "a log-normal variable exceeds another one", in the spirit of Kirk's formula. This approach enables to approximate a convex sum of exponential martingales by an exponential martingale, which is a nice property, preserving first moment. Our contribution is to extend the above analysis to many assets and to local volatility models. Numerical investigations indicate that our formula is extremely accurate. The accuracy is much higher than the Kirk formula and the Bjerksund and Stensland formula. Actually, under log-normal assumptions, the Bjerksund and Stensland formula is the zero order term of our approximation and we incorporate extra correction terms. In the general local volatility case, volatility corrections are also computed.

8.2 Generalized spread option.

8.2.1 Model.

For $x \in \mathbb{R}^d$, we write down $|x| := \sqrt{x \cdot x}$ for the Euclidean norm of a vector $x \in \mathbb{R}^d$. By convention, all vectors are written in column and 1 stands for the vector with all the coefficients equal to one. $\mathcal{C}^m(U,V)$ is the set of functions m-times continuously differentiable from U to V. For any function $f \in \mathcal{C}^1(\mathbb{R}^d,\mathbb{R}^q)$ (resp. $\in \mathcal{C}^2(\mathbb{R}^d,\mathbb{R}^q)$, the Jacobian $f'(x)$ (resp. the Hessian $f''(x)$) will stand as usual for the matrix $(\partial_{x_j} f(x)^i)_{1\le i\le q, 1\le j\le d} \in \mathbb{R}^q \otimes \mathbb{R}^d$ (resp. for the tensor $(\partial^2_{x_j,x_k} f(x)^i)_{1\le i\le q,1\le j,k\le d} \in \mathbb{R}^q \otimes (\mathbb{R}^d)^{\otimes 2}$.
Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{0\le t\le T})$ be a filtered probability space supporting a q-Brownian motion $\bar{B}$; $\mathbb{P}$ is the risk-neutral measure, i.e. such that the discounted price process of a tradable asset $e^{-\int_0^t r_s \mathrm{d}s} S_t$ is a $\mathbb{P}$-martingale (here, r is the risk-free rate, that can be stochastic). $(\mathcal{F}_t)_{0\le t\le T}$ is the $\mathbb{P}$-completed Brownian filtration and $\mathcal{F} = \mathcal{F}_T$. Regarding the risky tradable assets, we consider the multidimensional Ito process $\bar{S} = (\bar{S}^1, ..., \bar{S}^d)$ given by :

$$\frac{\mathrm{d}\bar{S}^j_t}{\bar{S}^j_t} = r_t \mathrm{d}t + \bar{\sigma}_{j,t} \mathrm{d}\bar{B}_t$$

where $(\bar{\sigma}_{j,t})^*$ is a càdlàg adapted $\mathbb{R}^q$-valued process (standing for the volatility process of $\bar{S}^j$) and r is a càdlàg adapted stochastic process. Precise assumptions on $(\bar{\sigma}^j)_j$ and $(r_t)_t$ are given later. So far , we only assume that $(e^{-\int_0^t r_s \mathrm{d}s} \bar{S}^j_t)_t$ is a $\mathbb{P}$-martingale.

8.2.2 General spread option.

Let $\bar{\alpha}$ and $\bar{\beta}$ be two vectors of $\mathbb{R}_+^d$: our study focuses on the general spread

$$\sum_{k=1}^{d} \bar{\alpha}^k \bar{S}_T^k - \sum_{k=1}^{d} \bar{\beta}^k \bar{S}_T^k := (\bar{\alpha} - \bar{\beta}) \cdot \bar{S}_T.$$

With the above representation, it may happen that we take into account a given asset twice, first with $\bar{\alpha}^k$ and second with $\bar{\beta}^k$. A natural simplification would be to impose that either $\bar{\alpha}^k$ or $\bar{\beta}^k$ is zero. We do not make this assumption and choose to keep a general form for the spread : more generally, replacing $(\bar{\alpha}^k, \bar{\beta}^k)$ by $(\bar{\alpha}^k + \delta^k, \bar{\beta}^k + \delta^k)$ (for $\delta^k \in \mathbb{R}$) does not modify the spread but modifies the further approximative formula. It gives several degrees of freedom to optimize the formulas, which have not been exploited so far.
We assume the spread price is given by the risk-neutral rule, that is

$$\mathbb{E}\left((\bar{\alpha} - \bar{\beta}) \cdot e^{-\int_0^T r_s \mathrm{d}s} \bar{S}_T\right)_+ .$$

Finally, observe that a Call spread can be easily embedded in the above framework by adding the zero-coupon bond $(B(t,T))_{0 \le t \le T}$ with maturity T as an asset : indeed, write

$$\bar{\alpha} \cdot \bar{S}_T - \bar{\beta} \cdot \bar{S}_T - K = [\bar{\alpha}, 0] \cdot [\bar{S}_T, B(T,T)] - [\bar{\beta}, 0] \cdot [\bar{S}_T, B(T,T)].$$

8.3 Reduction of the problem.

We are inspired by the Bjerksund and Stensland approach. We linearise the payoff and use changes of probability measure (given by the numéraires $\bar{S}^k$).

$$\begin{aligned}
\mathbb{E}\left[\left((\bar{\alpha} - \bar{\beta}) \cdot e^{-\int_0^t r_s \mathrm{d}s} \bar{S}_T\right)_+\right] &= \sum_{k=1}^{d} (\bar{\alpha}^k - \bar{\beta}^k) \bar{S}_0^k \mathbb{E}\left[\frac{\bar{S}_T^k}{\bar{S}_0^k e^{\int_0^T r_s ds}} 1_{\bar{\alpha} \cdot \bar{S}_T \ge \bar{\beta} \cdot \bar{S}_T}\right] \\
&= \sum_{k=1}^{d} (\bar{\alpha}^k - \bar{\beta}^k) \bar{S}_0^k \mathbb{P}_k\left(\bar{\alpha} \cdot \bar{S}_T / \bar{S}_T^k \ge \bar{\beta} \cdot \bar{S}_T / \bar{S}_T^k\right),
\end{aligned}$$

where $\mathbb{P}_k$ is defined by the Radon-Nikodym derivative $\frac{\mathrm{d}\mathbb{P}_k}{\mathrm{d}\mathbb{P}}{}_{|\mathcal{F}_T} = \frac{\bar{S}_T^k}{\bar{S}_0^k e^{\int_0^T r_s \mathrm{d}s}}$. Then, the fundamental quantities to approximate are the probabilities of the exercise regions $\{\bar{\alpha} \cdot \bar{S}_T / \bar{S}_T^k \ge \bar{\beta} \cdot \bar{S}_T / \bar{S}_T^k\}$, where we observe that $\bar{S}/\bar{S}^k$ is a martingale under the probability $\mathbb{P}_k$. To sum up, we have reduced the problem to compute quantities of the form

$$\mathbb{Q}(\alpha \cdot S_T \ge K \beta \cdot S_T),$$

where $K > 0$, $\alpha \in \mathbb{R}^d_+$, $\beta \in \mathbb{R}^d_+$, $\alpha \cdot 1 = \beta \cdot 1 = 1$, S is a $\mathbb{R}^d$-valued $\mathbb{Q}$-martingale with dynamics

$$\frac{\mathrm{d}S^j_t}{S^j_t} = \sigma_{j,t}\mathrm{d}B_t, \quad S^j_0 = 1,$$

for a certain $\mathbb{Q}$-Brownian motion $(B_t)_{0\leq t\leq T}$. Indeed, regarding on previous notations and with the previous change of numéraire and related calculus rules, we have (for any fixed k)

$$\begin{cases} \mathbb{Q} = \mathbb{P}_k, \\ S^j_t = \bar{S}^j_t/\bar{S}^k_t, \\ \sigma_{j,t} = \bar{\sigma}_{j,t} - \bar{\sigma}_{k,t}, \\ \alpha^j = \frac{\bar{\alpha}^j \bar{S}^j_0/\bar{S}^k_0}{\sum_{l=1}^d \bar{\alpha}^l \bar{S}^l_0/\bar{S}^k_0}, \\ \beta^j = \frac{\bar{\beta}^j \bar{S}^j_0/\bar{S}^k_0}{\sum_{l=1}^d \bar{\beta}^l \bar{S}^l_0/\bar{S}^k_0}, \\ K = \frac{\sum_{l=1}^d \bar{\beta}^l \bar{S}^l_0/\bar{S}^k_0}{\sum_{l=1}^d \bar{\alpha}^l \bar{S}^l_0/\bar{S}^k_0}. \end{cases}$$

In the next section, we tackle the general problem of approximating

$$\mathbb{Q}\left(\ln\left(\frac{\alpha \cdot e^{X_T}}{\beta \cdot e^{X_T}}\right) \geq \ln(K)\right), \tag{8.3.1}$$

assuming a diffusion model for $S_t := (e^{X^1_t}, ..., e^{X^d_t})^* := e^{X_t}$, with $X_0 = 0$, $K > 0$, $\alpha \cdot 1 = \beta \cdot 1 = 1$. Namely, we assume

$$X_t = \int_0^t b(s, X_s)\mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_j(s, X_s)\mathrm{d}B^j_s,$$

where σ_j, b is a function in $\mathcal{C}^\infty_b([0,T] \times \mathbb{R}^d, \mathbb{R}^d)$ and B is a Brownian motion under $\mathbb{Q}$. The function b is chosen in the form

$$b_i(s, X_s) = -\frac{1}{2}\sum_{j=1}^d |\sigma_{i,j}(s, X_s)|^2, \tag{8.3.2}$$

such that e^X is an exponential martingale under $\mathbb{Q}$ starting from 1. In the following, we shall write expansion specifying only the expansion on the volatility.

8.4 Expansion of the probability of the exercise region.

8.4.1 Convex parametrization.

We define the parametrization X^ϵ by :

$$X^\epsilon_t = \int_0^t b^\epsilon(s, X^\epsilon_s)\mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma^\epsilon_j(s, X^\epsilon_s)\mathrm{d}B^j_s,$$

with $\sigma_j^\epsilon(t,x) = \sigma_j(t, \epsilon x + (1-\epsilon)y_t)$, $b_i^\epsilon(s, X_s) = -\frac{1}{2}\sum_{j=1}^d |\sigma_{i,j}^\epsilon(s, X_s)|^2$ and y is some deterministic function. In particular, for $\epsilon = 0$ setting $b_s = b(s, y_s)$ and $\sigma_s = \sigma(s, y_s)$, we acknowledge the Gaussian process

$$\bar{X}_t = \int_0^t b_s \mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_{j,s} \mathrm{d}B_s^j.$$

This parametrization interpolates between the initial model X ($\epsilon = 1$) and the proxy $\bar{X}$ ($\epsilon = 0$). In the literature [Gobet 2011b], we usually take $y_t = X_0 = 0$. However, we guess that taking $y_t \neq X_0$ may be interesting and this is left to further investigation. In the following, we take

$$y_t = 0.$$

8.4.2 Benchmark process.

For $\gamma \in \mathbb{R}_+^d$ such that $\gamma \cdot 1 = 1$, we approximate $\gamma \cdot e^X$ by the proxy $e^{\bar{X}^\gamma}$:

$$\bar{X}_t^\gamma = \int_0^t b_s^\gamma \mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_{j,s}^\gamma \mathrm{d}B_s^j,$$

where $b_s = b(s,0)$, $\sigma_{j,s} = \sigma_j(s,0)$, $\sigma_{j,t}^\gamma = \gamma \cdot \sigma_{j,t}$ and $b_t^\gamma = -\frac{1}{2}\sum_{j=1}^d (\gamma \cdot \sigma_{j,t})^2 = -\frac{1}{2}\sum_{j=1}^d (\sigma_{j,t}^\gamma)^2$. We know that $\gamma \cdot e^X$ is a martingale under $\mathbb{Q}$ and the first moment is 1. This property is satisfied for the approximation $e^{\bar{X}^\gamma}$. Furthermore, $e^{\bar{X}_T^\gamma}$ is a log-normal random variable.

8.4.2.1 Derivatives computation.

For the sake of simplicity, we write $\dot{X}_t = (\partial_\epsilon X_t^\epsilon)_{|\epsilon=0}$ and $\ddot{X}_t = (\partial_\epsilon^2 X_t^\epsilon)_{|\epsilon=0}$. Then, for $\epsilon = 0$, we get

$$\dot{X}_t = \int_0^t b_s^{(1)} \bar{X}_s \mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_{j,s}^{(1)} \bar{X}_s \mathrm{d}B_s^j,$$

where $b_s^{(1)} = b'(s,0)$ and $\sigma_{j,s}^{(1)} = \sigma_j'(s,0)$. Taylor's formula gives

$$X_t - \bar{X}_t = X_t^{\epsilon=1} - X_t^{\epsilon=0} = \dot{X}_t + \int_0^1 (1-\epsilon)\partial_\epsilon^2 X_t^\epsilon \mathrm{d}\epsilon. \tag{8.4.1}$$

8.4.2.2 Probability expansion.

Theorem 8.4.1. *Let $\varphi \in \mathcal{C}_b^\infty(\mathbb{R}^d, \mathbb{R})$, then*

$$\mathbb{E}_\mathbb{Q}\left[\varphi\left(\ln\left(\frac{\alpha \cdot e^{X_T}}{\beta \cdot e^{X_T}}\right)\right)\right] = \mathbb{E}_\mathbb{Q}\left[\varphi\left(\bar{X}_T^\alpha - \bar{X}_T^\beta\right)\right] + \sum_{i=0}^2 \lambda_i \partial_\eta^i \mathbb{E}_\mathbb{Q}\left[\varphi(\bar{X}_T^\alpha - \bar{X}_T^\beta + \eta)\right]_{|\eta=0} + \mathcal{E}_T(\varphi), \tag{8.4.2}$$

where

$$\lambda_0 = \int_0^T ((\alpha-\beta)\cdot b_s - (b_s^\alpha - b_s^\beta))\mathrm{d}s + (\alpha-\beta)\cdot \int_0^T b_s^{(1)} \int_0^s b_r \mathrm{d}r\mathrm{d}s,$$

$$\lambda_1 = (\alpha-\beta)\cdot \int_0^T \int_0^s b_s^{(1)} \sum_{j=1}^d \sigma_{j,r}(\sigma_{j,r}^\alpha - \sigma_{j,r}^\beta)\mathrm{d}r\mathrm{d}s$$
$$+ (\alpha-\beta)\cdot \int_0^T \int_0^s \sum_{j=1}^d \sigma_{j,s}^{(1)}(\sigma_{j,r}^\alpha - \sigma_{j,r}^\beta) \int_0^s b_r \mathrm{d}r\mathrm{d}s,$$

$$\lambda_2 = (\alpha-\beta)\cdot \int_0^T \sum_{j=1}^d \sigma_{j,s}^{(1)}(\sigma_{j,s}^\alpha - \sigma_{j,s}^\beta) \int_0^s \sum_{k=1}^d \sigma_{k,r}(\sigma_{k,r}^\alpha - \sigma_{k,r}^\beta)\mathrm{d}r\mathrm{d}s.$$

and $\mathcal{E}_T(\varphi)$ is an error term.

Remark 8.4.1. *1. An estimation of the error $\mathcal{E}_T$ is already done in a similar context in [Gobet 2011b], even when φ is only locally Holder continuous.*

2. Our formula (8.4.2) is explicit because $\bar{X}_T^\alpha - \bar{X}_T^\beta$ is a Gaussian random variable, with mean $\int_0^T (b^\alpha - b^\beta)\mathrm{d}s$ and variance $\int_0^T |\sigma_s^\alpha - \sigma_s^\beta|^2 \mathrm{d}s$.

3. The term $\mathbb{E}_{\mathbb{Q}}\left[\varphi\left(\bar{X}_T^\alpha - \bar{X}_T^\beta\right)\right]$ is the Bjerksund and Stensland formula (matching when S follows a log-normal model and when the payoff is $(S_T^1 - S_T^2 - K)_+$, $K > 0$ (strike)), $\int_0^T ((\alpha-\beta)\cdot b_s - (b_s^\alpha - b_s^\beta))\mathrm{d}s$ is the correction of the Bjerksund and Stensland term and the other terms are corrections coming from the local volatility model.

Proof. Let $\gamma \in \mathbb{R}_+^d$ such that $\gamma \cdot 1 = 1$. Let

$$f^\gamma : \lambda \mapsto \ln\left(\sum_{j=1}^d \gamma^j e^{\bar{X}_T^\gamma + \lambda(X_T^j - \bar{X}_T^\gamma)}\right).$$

Then, the first and second derivatives of f^γ read as follow

$$\partial_\lambda f^\gamma(\lambda) = \frac{\sum_{j=1}^d \gamma^j (X_T^j - \bar{X}_T^\gamma) e^{\bar{X}_T^\gamma + \lambda(X_T^j - \bar{X}_T^\gamma)}}{\sum_{j=1}^d \gamma^j e^{\bar{X}_T^\gamma + \lambda(X_T^j - \bar{X}_T^\gamma)}}$$

and

$$\partial_\lambda^2 f^\gamma(\lambda) = \frac{\sum_{j=1}^d \gamma^j (X_T^j - \bar{X}_T^\gamma)^2 e^{\bar{X}_T^\gamma + \lambda(X_T^j - \bar{X}_T^\gamma)}}{\sum_{j=1}^d \gamma^j e^{\bar{X}_T^\gamma + \lambda(X_T^j - \bar{X}_T^\gamma)}} - \left(\frac{\sum_{j=1}^d \gamma^j (X_T^j - \bar{X}_T^\gamma) e^{\bar{X}_T^\gamma + \lambda(X_T^j - \bar{X}_T^\gamma)}}{\sum_{j=1}^d \gamma^j e^{\bar{X}_T^\gamma + \lambda(X_T^j - \bar{X}_T^\gamma)}}\right)^2.$$

In particular, for $\lambda = 0$, we have

$$f^\gamma(1) = \ln(\gamma \cdot e^{X_T}),$$
$$f^\gamma(0) = \bar{X}_T^\gamma,$$
$$\partial_\lambda f^\gamma(0) = \gamma \cdot X_T - \bar{X}_T^\gamma.$$

Owing to Taylor's theorem applied to $\lambda \mapsto \varphi(f^\alpha(\lambda) - f^\beta(\lambda))$ between 0 and 1, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[\varphi\left(\ln\left(\frac{\alpha \cdot e^{X_T}}{\beta \cdot e^{X_T}}\right)\right)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\varphi\left(\bar{X}_T^\alpha - \bar{X}_T^\beta\right)\right] + \mathbb{E}_{\mathbb{Q}}\left[\left((\alpha-\beta)\cdot X_T - (\bar{X}_T^\alpha - \bar{X}_T^\beta)\right)\varphi'(\bar{X}_T^\alpha - \bar{X}_T^\beta)\right] + \mathcal{E}_T^1(\varphi),
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{E}_T^1(\varphi) := \int_0^1 \mathbb{E}_{\mathbb{Q}}\Big[&(\partial_\lambda^2 f^\alpha(\lambda) - \partial_\lambda^2 f^\beta(\lambda))\varphi'(f^\alpha(\lambda) - f^\beta(\lambda)) \\
&+ (\partial_\lambda f^\alpha(\lambda) - \partial_\lambda f^\beta(\lambda))^2 \varphi''(f^\alpha(\lambda) - f^\beta(\lambda))\Big](1-\lambda)\mathrm{d}\lambda.
\end{aligned}$$

However, using the parametrization X^ϵ and (8.4.1), we get

$$\begin{aligned}
\gamma \cdot X_T - \bar{X}_T^\gamma &= \gamma \cdot (\bar{X}_T + \dot{X}_T) - \bar{X}_T^\gamma + \mathcal{E}_T^2 \\
&= \int_0^T (\gamma \cdot b_s - b_s^\gamma)\mathrm{d}s + \gamma \cdot \int_0^T b_s^{(1)} \bar{X}_s \mathrm{d}s + \gamma \cdot \sum_{j=1}^d \int_0^T \sigma_{j,s}^{(1)} \bar{X}_s \mathrm{d}B_s^j + \mathcal{E}_T^2,
\end{aligned}$$

where $\mathcal{E}_T^2 := \int_0^1 (1-\epsilon)(\gamma \cdot \partial_\epsilon^2 X_T^\epsilon)\mathrm{d}\epsilon$.
Then, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[\left((\alpha-\beta)\cdot X_T - (\bar{X}_T^\alpha - \bar{X}_T^\beta)\right)\varphi'(\bar{X}_T^\alpha - \bar{X}_T^\beta)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\Bigg[\Bigg(\int_0^T ((\alpha-\beta)\cdot b_s - (b_s^\alpha - b_s^\beta))\mathrm{d}s + (\alpha-\beta)\cdot \int_0^T b_s^{(1)} \bar{X}_s \mathrm{d}s \\
&\quad + (\alpha-\beta)\cdot \sum_{j=1}^d \int_0^T \sigma_{j,s}^{(1)} \bar{X}_s \mathrm{d}B_s^j\Bigg)\varphi'(\bar{X}_T^\alpha - \bar{X}_T^\beta)\Bigg].
\end{aligned}$$

Owing to Lemma 8.6.1 after some cumbersome computations, we get

$$\mathbb{E}_{\mathbb{Q}}\left[\left((\alpha-\beta)\cdot X_T - (\bar{X}_T^\alpha - \bar{X}_T^\beta)\right)\varphi'(\bar{X}_T^\alpha - \bar{X}_T^\beta)\right] = \sum_{i=0}^2 \lambda_i \mathbb{E}_{\mathbb{Q}}\left[\varphi^{(i+1)}(\bar{X}_T^\alpha - \bar{X}_T^\beta)\right],$$

To sum up, we have

$$\mathbb{E}_{\mathbb{Q}}\left[\varphi\left(\ln\left(\frac{\alpha \cdot e^{X_T}}{\beta \cdot e^{X_T}}\right)\right)\right] = \mathbb{E}_{\mathbb{Q}}\left[\varphi\left(\bar{X}_T^\alpha - \bar{X}_T^\beta\right)\right] + \sum_{i=0}^2 \lambda_i \partial_\eta^i \mathbb{E}_{\mathbb{Q}}\left[\varphi(\bar{X}_T^\alpha - \bar{X}_T^\beta + \eta)\right]_{|\eta=0} + \mathcal{E}_T(\varphi),$$

where

$$\mathcal{E}_T(\varphi) := \mathbb{E}\left[\mathcal{E}_T^2 \varphi'(\bar{X}_T^\alpha - \bar{X}_T^\beta)\right] + \mathcal{E}_T^1(\varphi).$$

□

We now indicate how to handle the case $\varphi(x) = 1_{x\geq \ln(K)}$. We follow the approach [Gobet 2011b] by smoothing the payoff and then let the smoothness go to 0. Let φ_δ be a $\mathcal{C}_b^\infty(\mathbb{R})$-regularization function of the function $1_{x\geq\ln(K)}$, so that we have $\varphi_\delta\left(\ln\left(\frac{\alpha\cdot e^{X_T}}{\beta\cdot e^{X_T}}\right)\right)$ instead of (8.4.2).

Theorem 8.4.2. *The probability of the exercise region can be expanded as follows*

$$\mathbb{Q}\left(\ln\left(\frac{\alpha\cdot e^{X_T}}{\beta\cdot e^{X_T}}\right)\geq \ln(K)\right)$$

$$=\mathcal{N}(d_T)+\left(\frac{\lambda_0}{\sqrt{\int_0^T|\sigma_s^\alpha-\sigma_s^\beta|^2\mathrm{d}s}}-\frac{d_T\lambda_1}{\int_0^T|\sigma_s^\alpha-\sigma_s^\beta|^2\mathrm{d}s}+\frac{(d_T^2-1)\lambda_2}{\left(\int_0^T|\sigma_s^\alpha-\sigma_s^\beta|^2\mathrm{d}s\right)^{3/2}}\right)\mathcal{N}'(d_T)+\mathcal{E}_T(\varphi_\delta),$$

where $d_T=\frac{\int_0^T(b_s^\alpha-b_s^\beta)\mathrm{d}s-\ln(K)}{\sqrt{\int_0^T|\sigma_s^\alpha-\sigma_s^\beta|^2\mathrm{d}s}}$, $\mathcal{N}(x)=\int_{-\infty}^x\frac{e^{-t^2/2}}{\sqrt{2\pi}}\mathrm{d}t$, $\mathcal{N}'(x)=\frac{e^{-x^2/2}}{\sqrt{2\pi}}$ *and* $\mathcal{E}_T(\varphi_\delta)$ *is an error term.*

Sketch of the proof in progress. Owing to Theorem 8.4.1, it remains to estimate the new error term $\mathcal{E}_T(\varphi_\delta)$, which can be split into

1. The contribution of the random variable X_T is $\mathbb{E}\left[1_{\ln\left(\frac{\alpha\cdot e^{X_T}}{\beta\cdot e^{X_T}}\right)\geq\ln(K)}-\varphi_\delta\left(\ln\left(\frac{\alpha\cdot e^{X_T}}{\beta\cdot e^{X_T}}\right)\right)\right]$: the idea is to use Malliavin calculus to perform an integration by part to get a sharp estimation of this error. We wish that when δ tends to 0, the error term converges to 0 as well. The difficulty comes from the possibly Malliavin degeneracy of the random variable $\ln\left(\frac{\alpha\cdot e^{X_T}}{\beta\cdot e^{X_T}}\right)$ and the irregularity of the indicator function. However, had we found a good assumption under which the law of the random variable $\ln\left(\frac{\alpha\cdot e^{X_T}}{\beta\cdot e^{X_T}}\right)$ has a bounded density, then we could use the estimation in [Avikainen 2009][Theorem 2.4 (i)]; but the sufficient condition is non-trivial and by no way obvious to explicit.

2. The contribution of the proxies $\bar{X}_T^\alpha$ and $\bar{X}_T^\beta$ is $\mathbb{E}\left[1_{\bar{X}_T^\alpha-\bar{X}_T^\beta\geq\ln(K)}-\varphi_\delta\left(\bar{X}_T^\alpha-\bar{X}_T^\beta\right)\right]$: here, the idea remains the same, but the study is greatly simplified by the Malliavin non-degeneracy of the Gaussian random variable $\bar{X}_T^\alpha-\bar{X}_T^\beta$. The proof of [Gobet 2011b][Lemma 4.1] can be adapted.

3. The contribution of the sensibilities is $\partial_\eta^i\mathbb{E}_{\mathbb{Q}}\left[1_{\bar{X}_T^\alpha-\bar{X}_T^\beta+\eta\geq\ln(K)}\right]_{|\eta=0}-$ $\partial_\eta^i\mathbb{E}_{\mathbb{Q}}\left[\varphi(\bar{X}_T^\alpha-\bar{X}_T^\beta+\eta)\right]_{|\eta=0}$: the proof of [Gobet 2011b][Lemma 4.2] can be adapted.

4. The remainder of the error term depends on φ_δ and its study is already done for regular functions regular functions, but the error depends on the bounds on derivatives of φ_δ. For the non smooth function like indicator function, we shall adapt the Malliavin calculus arguments in [Gobet 2011b][Section 5] to rid of the derivatives of φ_δ.

To sum up, the originality of the proof compared to [Gobet 2011b] is to prove the items 1 and 4. This is an open problem we are investigating. $\square$

8.5 Applications.

8.5.1 Examples.

Many examples of traded contracts in the energy market are spread options.

- Clean spark spread $= E_T - G_T/\rho_G - Ngg \times Pcc$, where E, G, Pcc are respectively the price of electricity, of gas and of a carbon credit, ρ_G is the efficiency of the gas power plant and Ngq is the number of carbon credits necessary to cover gas operations.

- Clean dark spread $= E_T - C_T/\rho_C - Nc \times Pcc$, where E, C, Pcc are respectively the price of electricity, of coal and of a carbon credit, ρ_C is the efficiency of the coal power plant and Nc is the number of carbon credits necessary to cover coal operations.

- Crack spread $= (\sum_{j=1}^d \alpha^j S^j)_+$, where $\alpha^j \in \mathbb{R}$ and S^j are the prices of many products from crude oil, including gasoline, kerosene, diesel, heating oil, aviation fuel, asphalt and others, coming from refinery productions.

8.5.2 Spread option with two assets.

The most famous spread option remains the spread option with two assets, where many analytical formulas are well-known in the literature [Kirk 1995] and [Bjerksund 2006], under a log-normal model. This numerical section compares the Bjerksund and Stensland formula to our formula. We choose the model and its parameters as follows:

$$\begin{aligned}
&dS_t^1 = \sigma_1 S_t^1 dB_t, S_0^1 = 100, \sigma_1 = 0.15,\\
&dS_t^2 = \sigma_2 S_t^2(\rho dB_t + \sqrt{1-\rho^2} dW_t), S_0^2 = 100, \sigma_2 = 0.20,\\
&\text{payoff} = (S_T^1 - S_T^2 - K)_+ \qquad T = 1.
\end{aligned}$$

The first row of the array below is the price given by Monte Carlo's method for 50000 drawings. The second and the third row are respectively the difference of the Bjerksund and Stensland price and our formula with Monte Carlo's price.

ρ/K	-20	-10	-5	0	5	10	20
-1	26.3121	19.5636	16.5871	13.8931	11.4876	9.3707	5.9727
	0.079	0.0135	$3.48e-005$	-0.0011	0.0087	0.026	0.072
	-0.0019	-0.0017	-0.0014	-0.0011	-0.00059	$-6.49e-005$	0.00013
-0.75	25.5908	18.7349	15.7271	13.0204	10.6223	8.53265	5.24055
	0.0832	0.0179	0.0036	0.0014	0.0097	0.0255	0.0664
	0.0022	0.0017	0.0016	0.0014	0.0012	0.0011	0.0010
-0.5	24.8374	17.8516	14.8015	12.0822	9.69246	7.63569	4.47307
	0.086	0.022	0.011	0.0046	0.0119	0.02639	0.0615
	0.0053	0.0050	0.0084	0.0046	0.0043	0.0040	0.0029
-0.25	24.0589	16.9121	13.8161	11.0719	8.69215	6.67651	3.67374
	0.0778	0.0154	0.0010	-0.0016	0.0055	0.0193	0.0506
	-0.0029	-0.0024	-0.0020	-0.0016	-0.0013	-0.0011	-0.0009
0	23.2311	15.8797	12.7195	9.9481	7.58276	5.62309	2.83136
	0.0789	0.0180	0.0030	-0.0005	0.0057	0.0177	0.0429
	-0.0012	-0.0007	-0.0006	-0.0005	-0.0005	-0.0007	-0.0010
0.25	22.3657	14.7392	11.4883	8.67756	6.33241	4.45233	1.95282
	0.0776	0.0195	0.0040	$7.7e-006$	0.0057	0.0166	0.0347
	-0.0017	-0.0004	-0.0001	$7.7e-006$	$6.9e-005$	$3.2e-005$	-0.0009
0.5	21.4693	13.4477	10.0562	7.18258	4.87081	3.11962	1.06241
	0.0765	0.0217	0.0048	-0.0003	0.0047	0.0141	0.0240
	-0.0014	-0.0001	-0.0001	-0.0003	-0.0005	-0.0008	-0.0018
0.75	20.5944	11.9339	8.27519	5.27319	3.03126	1.54493	0.267503
	0.0712	0.0269	0.0073	0.0005	0.0056	0.0133	0.0113
	-0.0017	0.0003	0.0005	0.0005	$7.5e-005$	-0.0006	-0.0017
1	20.0128	10.1778	5.6262	1.99434	0.20768	$1.84192e-005$	0
	0.0271	0.0425	0.0254	0.0002	0.0224	0.0013	$6.1e-029$
	-0.0071	-0.0006	0.0003	0.0002	$-1.0e-005$	-0.0015	$-2.6e-026$

We remark that : more the option is out or in the money, better is our approximation for any correlation ρ. The formulas degenerate to the Margrabe formula for $K = 0$.

8.6 Appendix

We remind us a handy lemma from [Gobet 2011b].

Lemma 8.6.1. *For $f, a, h \in \mathbb{R} \otimes \mathbb{R}^d$ and $e, g \in \mathbb{R}$, we have*

$$\mathbb{E}\left[l\left(\int_0^t a_s \mathrm{d}B_s\right)\int_0^t \left\{\int_0^s g_r \mathrm{d}r + \int_0^s h_r \mathrm{d}B_r\right\}(e_s \mathrm{d}s + f_s \mathrm{d}B_s)\right] = \sum_{i=0}^{2} \lambda_i \partial_\epsilon^i \mathbb{E}\left[l\left(\int_0^t a_s \mathrm{d}B_s + \epsilon\right)\right]_{|\epsilon=0},$$

where

$$\lambda_0 = \int_0^t \left(\int_0^s g_r \mathrm{d}r \right) e_s \mathrm{d}s,$$
$$\lambda_1 = \int_0^t \int_0^s (g_r a_s \cdot f_s + e_s a_r \cdot h_r) \mathrm{d}r \mathrm{d}s,$$
$$\lambda_2 = \int_0^t \left(\int_0^s a_r \cdot h_r \mathrm{d}r \right) a_s \cdot f_s \mathrm{d}s.$$

Bibliography

[Ahn 1998] Hyungsok Ahn, Mohit Dayal, Eric Grannan and Glen Swindle. *Option replication with transaction costs: General diffusion limits.* Ann. Appl. Probab., vol. 8, no. 3, pages 676–707, 1998. (Cited on pages 18, 83 and 87.)

[Alos 2011] E. Alos, A. Eydeland and P. Laurence. *A Kirk's and a Bachelier's formula for three-asset spread options.* Cutting Edge, risk.net/energy-risk, 2011. (Cited on pages 16 and 153.)

[Avellaneda 1994] M. Avellaneda and A. Paras. *Optimal hedging portfolios for derivative securities in the presence of large transaction costs.* Appl. Math. Finance, vol. 2, pages 165–193, 1994. (Cited on page 2.)

[Avikainen 2009] Rainer Avikainen. *On irregular functionals of SDEs and the Euler scheme.* Finance Stoch., vol. 13, no. 3, pages 381–401, 2009. (Cited on page 160.)

[Barles 1998] Guy Barles and Halil Mete Soner. *Option pricing with transaction costs and a nonlinear Black-Scholes equation.* Finance Stoch., vol. 2, no. 4, pages 369–397, 1998. (Cited on page 93.)

[Benhamou 2009] Eric Benhamou, Emmanuel Gobet and Mohammed Miri. *Smart expansion and fast calibration for jump diffusions.* Finance Stoch., vol. 13, no. 4, pages 563–589, 2009. (Cited on page 15.)

[Benhamou 2010a] Eric Benhamou, Emmanuel Gobet and Mohammed Miri. *Expansion formulas for European options in a local volatility model.* Int. J. Theor. Appl. Finance, vol. 13, no. 4, pages 603–634, 2010. (Cited on page 15.)

[Benhamou 2010b] Eric Benhamou, Emmanuel Gobet and Mohammed Miri. *Time dependent Heston model.* SIAM J. Financ. Math., vol. 1, pages 289–325, 2010. (Cited on page 15.)

[Benhamou 2012] Eric Benhamou, Emmanuel Gobet and Mohammed Miri. *Analytical formulas for a local volatility model with stochastic rates.* Quant. Finance, vol. 12, no. 2, pages 185–198, 2012. (Cited on page 15.)

[Bensaid 1992] Bernard Bensaid, Jean-Philippe Lesne, Henri Pagès and José Scheinkman. *Derivative asset pricing with transaction costs.* Math. Finance, vol. 2, no. 2, pages 63–86, 1992. (Cited on page 2.)

[Bertsimas 2000] Dimitris Bertsimas, Leonid Kogan and Andrew W. Lo. *When is time continuous?* Journal of Financial Economics, vol. 55, pages 173–204, 2000. (Cited on pages 38 and 39.)

[Bertsimas 2001] Dimitris Bertsimas, Leonid Kogan and Andrew W. Lo. *When is time continuous?* Avellaneda, Marco (ed.), Quantitative analysis in financial markets. Collected papers of the New York University mathematical finance seminar. Vol. II. Singapore: World Scientific. 71-102., 2001. (Cited on page 7.)

[Bichteler 1981] Klaus Bichteler. *Stochastic integration and L_p−theory of semimartingales.* Ann. Probab., vol. 9, pages 49–89, 1981. (Cited on pages 3, 4 and 46.)

[Bjerksund 2006] P. Bjerksund and G. Stensland. *Closed form spread option valuation.* preprint, `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1145206`, 2006. (Cited on pages 16 and 161.)

[Bompis 2012] R. Bompis and E. Gobet. *Asymptotic and non asymptotic approximations for option valuation.* Computational finance, Thomas Gerstner and Peter Kloeden (Ed.) (2012) 80, 2012. (Cited on page 15.)

[Carmona 2003] René Carmona and Valdo Durrleman. *Pricing and hedging spread options.* SIAM Rev., vol. 45, no. 4, pages 627–685, 2003. (Cited on page 16.)

[Cattiaux 2002] P. Cattiaux and L. Mesnager. *Hypoelliptic non-homogeneous diffusions.* Probab. Theory Related Fields, vol. 123, no. 4, pages 453–483, 2002. (Cited on page 73.)

[Chen 1997] Zizhen Chen and Zhongdan Huan. *On the continuity of the mth root of a continuous nonnegative definite matrix-valued function.* J. Math. Anal. Appl., vol. 209, no. 1, pages 60–66, 1997. (Cited on page 130.)

[Conze 1991] A. Conze and Viswanathan. *Path dependent options: the case of lookback options.* Journal of Finance, vol. 46, no. 5, pages 1893–1907, 1991. (Cited on page 74.)

[Costantini 2006] C. Costantini, E. Gobet and N. El Karoui. *Boundary sensitivities for diffusion processes in time dependent domains.* Applied Mathematics and Optimization, vol. 54, no. 2, pages 159–187, 2006. (Cited on page 144.)

[Davis 1993] Mark H.A. Davis, Vassilios G. Panas and Thaleia Zariphopoulou. *European option pricing with transaction costs.* SIAM J. Control Optimization, vol. 31, no. 2, pages 470–493, 1993. (Cited on page 11.)

[Denis 2010a] E. Denis. *Modified Leland's strategy for constant transaction costs rate.* preprint, http://hal.archives-ouvertes.fr/docs/00/60/31/72/PDF/Emmanuel.DENIS-Publication-MF31_03_01.pdf, 2010. (Cited on pages 10, 11 and 92.)

[Denis 2010b] Emmanuel Denis and Yuri Kabanov. *Mean square error for the Leland-Lott hedging strategy: convex pay-offs.* Finance Stoch., vol. 14, no. 4, pages 625–667, 2010. (Cited on pages 2, 11, 18, 19 and 80.)

[DiBenedetto 2002] Emmanuele DiBenedetto. Real analysis. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Boston Inc., Boston, MA, 2002. (Cited on page 73.)

[Dieudonné 1990] J. Dieudonné. Eléments d'analyse. Jacques Gabay, Paris, 1990. Tome 1, (chapitres I à XI), Fondements de l'Analyse moderne. (Cited on page 73.)

[El Karoui 1998] Nicole El Karoui, Monique Jeanblanc-Picqué and Steven E. Shreve. *Robustness of the Black and Scholes formula.* Math. Finance, vol. 8, no. 2, pages 93–126, 1998. (Cited on pages 83 and 85.)

[Fouque 2003] J.-P. Fouque, G. Papanicolaou, R. Sircar and K. Solna. *Singular perturbations in option pricing.* SIAM J. Appl. Math., vol. 63, no. 5, pages 1648–1665, 2003. (Cited on page 15.)

[Freidlin 1985] M. Freidlin. Functional integration and partial differential equations. Annals of Mathematics Studies - Princeton University Press, 1985. (Cited on page 146.)

[Friedman 1964] A. Friedman. Partial differential equations of parabolic type. Prentice-Hall, 1964. (Cited on page 72.)

[Fukasawa 2011a] M. Fukasawa. *Asymptotically efficient discrete hedging.* Stochastic analysis with Financial Applications, Progress in Probability, vol. 65, pages 331–346, 2011. (Cited on pages 7, 9, 39, 40, 56, 57 and 64.)

[Fukasawa 2011b] M. Fukasawa. *Discretization error of stochastic integrals.* Annals of Applied Probability, vol. 21, no. 4, pages 1436–1465, 2011. (Cited on pages 7, 8, 14, 38, 39, 102 and 140.)

[Geiss 2004] C. Geiss and S. Geiss. *On approximation of a class of stochastic integrals and interpolation.* Stoch. Stoch. Rep., vol. 76, no. 4, pages 339–362, 2004. (Cited on pages 7 and 39.)

[Geiss 2009] Stefan Geiss and Anni Toivola. *Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces.* Bernoulli, vol. 15, no. 4, pages 925–954, 2009. (Cited on pages 7, 14, 38 and 140.)

[Geiss 2011] Stefan Geiss and Emmanuel Gobet. *Fractional smoothness and applications in finance.* Di Nunno, Giulia (ed.) et al., Advanced mathematical methods for finance. Berlin: Springer. 313-331, 2011. (Cited on pages 7, 39 and 65.)

[Geiss 2012] Christel Geiss, Stefan Geiss and Emmanuel Gobet. *Generalized fractional smoothness and L_p-variation of BSDEs with non-Lipschitz terminal condition.* Stochastic Processes Appl., vol. 122, no. 5, pages 2078–2116, 2012. (Cited on page 39.)

[Genon-Catalot 1993] V. Genon-Catalot and J. Jacod. *On the estimation of the diffusion coefficient for multidimensional diffusion processes.* Ann. Inst. H. Poincaré (Probab. Statist.), vol. 29, pages 119–151, 1993. (Cited on page 46.)

[Gilbarg 1977] D. Gilbarg and N.S. Trudinger. Elliptic partial differential equations of second order. Springer Verlag, first edition, 1977. (Cited on pages 144 and 146.)

[Gobet 2001] Emmanuel Gobet and Emmanuel Teman. *Discrete time hedging errors for options with irregular payoffs.* Finance Stoch., vol. 5, no. 3, pages 357–367, 2001. (Cited on pages 7, 14, 38, 39 and 140.)

[Gobet 2010a] E. Gobet and A. Makhlouf. *L_2-time regularity of BSDEs with irregular terminal functions.* Stochastic Processes and their Applications, vol. 120, pages 1105–1132, 2010. (Cited on page 39.)

[Gobet 2010b] E. Gobet and S. Menozzi. *Stopped diffusion processes: boundary corrections and overshoot.* Stochastic Processes and Their Applications, vol. 120, pages 130–162, 2010. (Cited on page 64.)

[Gobet 2011a] E. Gobet and A. Makhlouf. *The tracking error rate of the Delta-Gamma hedging strategy.* Forthcoming in Mathematical Finance, 2011. (Cited on page 40.)

[Gobet 2011b] E. Gobet and M. Miri. *Weak approximation of averaged diffusion processes.* In revison for Stochastic Process and their Applications., 2011. (Cited on pages 15, 16, 153, 157, 158, 160, 161 and 162.)

[Grannan 1996] E.R. Grannan and G.H. Swindle. *Minimizing transaction costs of option hedging strategies.* Math. Finance, vol. 6, no. 4, pages 341–364, 1996. (Cited on page 2.)

[Hayashi 2005] Takaki Hayashi and Per A. Mykland. *Evaluating hedging errors: an asymptotic approach.* Math. Finance, vol. 15, no. 2, pages 309–343, 2005. (Cited on pages 7, 14, 38 and 140.)

[Henrotte 1993] P. Henrotte. *Transaction costs and duplication strategies.* Working paper, Stanford University and HEC, 1993. (Cited on pages 2 and 93.)

[Hodges 1989] S. Hodges and A. Neuberger. *Optimal replication of contingent claims under transaction costs.* Rev. Futures Mark., no. 8, pages 222–239, 1989. (Cited on page 2.)

[Hofmann 2000] Norbert Hofmann, Thomas Muller-Gronbach and Klaus Ritter. *Optimal approximation of stochastic differential equations by adaptive step-size control.* Math. Comput., vol. 69, no. 231, pages 1017–1034, 2000. (Cited on pages 13 and 115.)

[Horn 1990] Roger A. Horn and Charles R. Johnson. Matrix analysis. Reprinted with corrections. Cambridge etc.: Cambridge University Press. xiii, 561 p., 1990. (Cited on pages 68, 69, 130 and 137.)

[Jacod 1998] Jean Jacod and Philip Protter. *Asymptotic error distributions for the Euler method for stochastic differential equations.* Ann. Probab., vol. 26, no. 1, pages 267–307, 1998. (Cited on pages 14 and 140.)

[Jacod 2003] Jean Jacod and Albert N. Shiryaev. Limit theorems for stochastic processes. 2nd ed. Grundlehren der Mathematischen Wissenschaften. 288. Berlin: Springer. xx, 661 p., 2003. (Cited on page 139.)

[Jacod 2012] J. Jacod and P. Protter. *Discretization of Processes.* Stochastic Modelling and Applied Probability, 67. Springer-Verlag, 2012. (Cited on pages 5, 8, 45 and 139.)

[Kabanov 1997] Yuri M. Kabanov and Mher M. Safarian. *On Leland's strategy of option pricing with transactions costs.* Finance Stoch., vol. 1, no. 3, pages 239–250, 1997. (Cited on pages 2, 10, 18 and 24.)

[Karandikar 1981] Rajeeva L. Karandikar. *Pathwise solutions of stochastic differential equations.* Sankhya : The Indian Journal of Statistic, vol. 43, pages 121–132, 1981. (Cited on page 3.)

[Karandikar 1989] Rajeeva L. Karandikar. *On Métivier-Pellaumail inequality, Emery topology and pathwise formulae in stochastic calculus.* Sankhyā: The Indian Journal of Statistics, Series A., vol. 51, no. 2, pages 121–143, 1989. (Cited on pages 3 and 46.)

[Karandikar 1991] Rajeeva L. Karandikar. *On almost sure convergence of modified Euler-Peano approximation of solution to an S.D.E. driven by a semimartingale.* Séminaire de probabilités, Lect. Notes Math. 1485, 113-120 (1991)., 1991. (Cited on page 3.)

[Karandikar 1995] Rajeeva L. Karandikar. *On pathwise stochastic integration.* Stochastic Processes Appl., vol. 57, no. 1, pages 11–18, 1995. (Cited on pages 3, 39 and 46.)

[Karandikar 2006] Rajeeva L. Karandikar. *On almost sure convergence results in stochastic calculus.* Émery, Michel (ed.) et al., In memoriam Paul-André Meyer. Séminaire de probabilités XXXIX. Berlin: Springer. Lecture Notes in Mathematics 1874, 137-156 (2006)., 2006. (Cited on pages 3 and 46.)

[Karatzas 1988] I. Karatzas and S.E. Shreve. Brownian Motion and Stochastic Calculus. Springer-Verlag, 1988. (Cited on page 3.)

[Karatzas 1991] Ioannis Karatzas and Steven E. Shreve. Brownian motion and stochastic calculus. Graduate Texts in Mathematics, 113. New York etc.: Springer-Verlag. xxiii, 470 p., second édition, 1991. (Cited on page 41.)

[Kirk 1995] E. Kirk. *Correlation in the energy markets.* Managing Energy Price Risk (First Edition), London: Risk Publications and Enron, pages 71–78, 1995. (Cited on pages 16 and 161.)

[Kloeden 2010] Peter E. Kloeden and Eckhard Platen. Numerical solution of stochastic differential equations. 4th corrected printing. Applications of Mathematics 23. Berlin: Springer. xxxvi, 636 p. , 2010. (Cited on pages 13 and 115.)

[Kunita 1984] H. Kunita. *Stochastic differential equations and stochastic flows of diffeomorphisms.* Ecole d'Eté de Probabilités de St-Flour XII, 1982 - Lecture Notes in Math. 1097 - Springer Verlag, pages 144–305, 1984. (Cited on page 72.)

[Kunitomo 2001] Naoto Kunitomo and Akihiko Takahashi. *The asymptotic expansion approach to the valuation of interest rate contingent claims.* Math. Finance, vol. 11, no. 1, pages 117–151, 2001. (Cited on page 15.)

[Kurtz 1991] Thomas G. Kurtz and Philip Protter. *Wong-Zakai corrections, random evolutions, and simulation schemes for SDE's.* Stochastic analysis, Proc. Conf. Honor Moshe Zakai 65th Birthday, Haifa/Isr. 1991, 331-346 (1991)., 1991. (Cited on pages 14, 117, 125 and 140.)

[Kurtz 1996] Thomas G. Kurtz and Philip E. Protter. *Weak convergence of stochastic integrals and differential equations.* Graham, C. (ed.) et al., Probabilistic models for nonlinear partial differential equations. Lectures given at the 1st session of the Centro Internazionale Matematico Estivo, Montecatini Terme, Italy, May 22-30, 1995. Berlin: Springer. Lect. Notes Math. 1627, 1-41 (1996)., 1996. (Cited on page 140.)

[Kusuoka 1985] S. Kusuoka and D. Stroock. *Applications of the Malliavin calculus II.* J. Fac. Sci. Univ. Tokyo, vol. 32, pages 1–76, 1985. (Cited on pages 73 and 74.)

[Leland 1985] HE. Leland. *Option Pricing and Replication With Transaction Cost.* Journal of Finance, vol. 40, pages 1283–1301, 1985. (Cited on pages 10 and 92.)

[Lenglart 1977] E. Lenglart. *Relation de domination entre deux processus.* Ann. Inst. Henri Poincaré, Sect. B, vol. 13, no. 2, pages 171–179, 1977. (Cited on page 46.)

[Levental 1997] Shlomo Levental and Anatolii V. Skorohod. *On the possibility of hedging options in the presence of transaction costs.* Ann. Appl. Probab., vol. 7, no. 2, pages 410–443, 1997. (Cited on pages 10 and 111.)

[Martini 1999] C Martini and C Patry. *Variance optimal hedging in the Black-Scholes model for a given number of transactions.* INRIA Rapport de recherche, no. 3767, `http://hal.inria.fr/docs/00/07/28/95/PDF/RR-3767.pdf`, 1999. (Cited on pages 7 and 39.)

[Maruyama 1955] Gisirō Maruyama. *Continuous Markov processes and stochastic equations.* Rend. Circ. Mat. Palermo, II. Ser., vol. 4, pages 48–90, 1955. (Cited on pages 13 and 115.)

[Milstein 1994] G.N. Milstein. Numerical integration of stochastic differential equations. Transl. from the Russian. Mathematics and its Applications (Dordrecht). 313. Dordrecht: Kluwer Academic Publishers. vii, 169 p. , 1994. (Cited on pages 13 and 115.)

[Muller-Gronbach 2002] Thomas Muller-Gronbach. *The optimal uniform approximation of systems of stochastic differential equations.* Ann. Appl. Probab., vol. 12, no. 2, pages 664–690, 2002. (Cited on pages 13 and 115.)

[Muller-Gronbach 2004] Thomas Muller-Gronbach. *Optimal pointwise approximation of SDEs based on Brownian motion at discrete points.* Ann. Appl. Probab., vol. 14, no. 4, pages 1605–1642, 2004. (Cited on pages 13 and 115.)

[Muller-Gronbach 2008] Thomas Muller-Gronbach and Klaus Ritter. *Minimal errors for strong and weak approximation of stochastic differential equations.* Keller, Alexander (ed.) et al., Monte Carlo and quasi-Monte Carlo methods 2006. Selected papers based on the presentations at the 7th international conference 'Monte Carlo and quasi-Monte Carlo methods in scientific computing', Ulm, Germany, August 14–18, 2006. Berlin: Springer. 53-82 (2008)., 2008. (Cited on pages 13, 115 and 119.)

[Musiela 2000] M. Musiela and M. Rutkowski. Martingale Methods in Financial Modelling. Springer, 2000. (Cited on page 21.)

[Musiela 2005] Marek Musiela and Marek Rutkowski. Martingale methods in financial modelling. Stochastic Modelling and Applied Probability 36. Berlin: Springer. xvi , second édition, 2005. (Cited on page 42.)

[Pascucci 2011] A. Pascucci. PDE and martingale methods in option pricing. Bocconi & Springer Series, 2011. (Cited on page 43.)

[Pergamenshchikov 2003] S. Pergamenshchikov. *Limit theorem for Leland's strategy.* Ann. Appl. Probab., vol. 13, no. 3, pages 1099–1118, 2003. (Cited on pages 2, 10, 18 and 19.)

[Platen 1999] Eckhard Platen. *An introduction to numerical methods for stochastic differential equations.* Acta Numerica 8, 197-246., 1999. (Cited on pages 13 and 115.)

[Protter 2004] Philip E. Protter. Stochastic integration and differential equations. 2nd ed. Applications of Mathematics 21. Berlin: Springer. xiii, 415 p., 2004. (Cited on pages 3, 117, 118 and 127.)

[Revuz 1999] D. Revuz and M. Yor. Continuous Martingales and Brownian Motion. Springer, 1999. (Cited on page 74.)

[Rootzen 1980] H. Rootzen. *Limit distributions for the error in approximations of stochastic integrals.* The Annals of Probability, 1980. (Cited on page 140.)

[Rosenbaum 2011] M. Rosenbaum and P. Tankov. *Asymptotically optimal discretization of hedging strategies with jumps.* Submitted, 2011. (Cited on pages 39 and 102.)

[Sekine 2008] J. Sekine and J. Yano. *Hedging errors of Leland's strategies with time-inhomogeneous rebalancing.* Preprint, 2008. (Cited on page 2.)

[Simon 1980] J. Simon. *Differentiation with respect to the domain in boundary value problems.* Numer. Funct. Anal. Optim., vol. 2, no. 7-8, pages 649–687, 1980. (Cited on page 144.)

[Soner 1995] H.M. Soner, S.E. Shreve and J. Cvitanić. *There is no nontrivial hedging portfolio for option pricing with transaction costs.* Ann. Appl. Probab., vol. 5, no. 2, pages 327–355, 1995. (Cited on pages 10 and 111.)

[Stroock 2006] Daniel W. Stroock and S. R. Srinivasa Varadhan. Multidimensional diffusion processes. Reprint of the 2nd correted printing (1997). Classics in Mathematics. Berlin: Springer. xii, 338 p. , 2006. (Cited on pages 69 and 136.)

[Talay 1995] Denis Talay. *Simulation of stochastic differential systems.* Krée, Paul (ed.) et al., Probabilistic methods in applied physics. Berlin: Springer-Verlag. Lect. Notes Phys. 451, 54-96 (1995)., 1995. (Cited on pages 13 and 115.)

[Uchida 2004] Masayuki Uchida and Nakahiro Yoshida. *Asymptotic expansion for small diffusions applied to option pricing.* Stat. Inference Stoch. Process., vol. 7, no. 3, pages 189–223, 2004. (Cited on page 15.)

[Watanabe 1987] Shinzo Watanabe. *Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels.* Ann. Probab., vol. 15, pages 1–39, 1987. (Cited on page 15.)

[Whalley 1997] A.E. Whalley and P. Wilmott. *An asymptotic analysis of an optimal hedging model for option pricing with transaction costs.* Math. Finance, vol. 7, no. 3, pages 307–324, 1997. (Cited on page 92.)

[Wilmott 1994] P. Wilmott, J. Dewynne and S. Howison. Option pricing: mathematical models and computation. Oxford Financial Press, 1994. (Cited on page 43.)

[Yoshida 1992] Nakahiro Yoshida. *Asymptotic expansion for statistics related to small diffusions.* J. Jap. Stat. Soc., vol. 22, no. 2, pages 139–159, 1992. (Cited on page 15.)

[Zhang 1999] R. Zhang. *Couverture approchée des options européennes.* PhD thesis, Ecole Nationale des Ponts et Chaussées, http://cermics.enpc.fr/theses/99/zhangruotao.ps.gz, 1999. (Cited on pages 7 and 39.)