



HAL
open science

Problèmes numériques en mathématiques financières et en stratégies de trading

Julien Baptiste

► **To cite this version:**

Julien Baptiste. Problèmes numériques en mathématiques financières et en stratégies de trading. Mathématiques générales [math.GM]. Université Paris sciences et lettres, 2018. Français. NNT : 2018PSLED009 . tel-04477872

HAL Id: tel-04477872

<https://theses.hal.science/tel-04477872v1>

Submitted on 26 Feb 2024

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

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

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Université Paris-Dauphine

Résolution de problèmes numériques en finance :
évaluation d'options et méthodes d'apprentissage
automatique appliquées au trading algorithmique

École Doctorale de Dauphine — ED 543

Spécialité **Sciences**

**Soutenu le 21.06.2018
par Julien BAPTISTE**

Dirigée par **Emmanuel LÉPINETTE**

COMPOSITION DU JURY :

Bruno BOUCHARD
Université Paris Dauphine
Président du jury

Laurence CARASSUS
Université Reims Champagne Ardenne
Membre du jury

Denis GERBER
Das Bot
Membre du jury

Idris KHARROUBI
Sorbonne Université
Rapporteur

Emmanuel LÉPINETTE
Université Paris Dauphine
Directeur de thèse

Serge PERGAMENCHTCHIKOV
Université de Rouen
Rapporteur

Xiaolu TAN
Université Paris-Dauphine
Membre du jury

Table des matières

1	Introduction	1
I	Évaluation du prix d'options	4
2	Pricing without martingale measure	5
2.1	Introduction	6
2.2	The one-period framework	8
2.3	The multi-period framework	14
2.3.1	Multi-period super-hedging prices	14
2.3.2	Multi-period (AIP)	15
2.3.3	Explicit pricing of a convex payoff under (AIP)	17
2.4	Comparison between the (AIP) condition and classical no-arbitrage conditions	19
2.5	Numerical experiments	22
2.5.1	Calibration	22
2.5.2	Super-hedging	24
2.6	Appendix	26
2.6.1	Conditional support of a vector-valued random variable	26
2.6.2	Conditional essential supremum	27
2.6.3	Link between two notions	29
2.7	Conclusion	31
3	Diffusion equations : convergence of the functional scheme derived from	

the Binomial tree with local volatility for non smooth payoff functions	33
3.1 Introduction	34
3.2 Functional scheme for diffusion equations	36
3.3 Numerical examples	38
3.3.1 The case where the diffusion coefficient only depends on time.	38
3.3.2 The case where the diffusion coefficient only depends on space.	41
3.3.3 The case where the diffusion coefficient depends on time and space.	44
3.4 Comparison to the Finite Element method.	46
3.4.1 The case of a bounded payoff function.	46
3.4.2 The case of an unbounded payoff function.	47
3.5 Proof	50
3.6 Appendix	56
3.7 Conclusion	57
4 Approximation of non-Lipschitz SDEs by Picard iterations	59
4.1 Introduction	59
4.2 Model	61
4.2.1 Notations	61
4.2.2 The stochastic differential equation	61
4.3 Approximation of the stochastic process	64
4.3.1 Volatility truncation	64
4.3.2 Picard iterations	66
4.3.3 Volatility discretization based on Picard iterations.	67
4.3.4 Main result	68
4.4 Numerical example : the CEV model	68
4.4.1 The CEV model with elasticity $\beta > 1$	69
4.4.2 The CEV model with elasticity $\beta \in (0, 1)$	74
4.5 The Aït-Sahalia model	76
4.6 Example of mean reverting SDEs	77
4.7 Proofs	78

4.8	Conclusion	91
II Méthodes d'apprentissage automatique appliquées au trading algorithmique		93
5	Application de méthodes d'apprentissage automatique à une stratégie intraday d'Open Range Breakout sur Future DAX	94
5.1	Introduction	94
5.2	Open Range Breakout sur Future DAX	96
5.2.1	L'Open Range Breakout	96
5.2.2	Le contrat future FDAX	97
5.2.3	Première application de la stratégie d'Open Range Breakout	99
5.2.4	Significativité des signaux de trade	100
5.2.5	Optimisation des paramètres	101
5.2.6	Impact du slippage sur la stratégie finale	106
5.3	Open Range Breakout et apprentissage automatique	108
5.3.1	Indicateurs et variable à expliquer	108
5.3.2	Méthodes d'apprentissage et méthodologie	109
5.3.3	Premiers tests	110
5.3.4	Optimisation des hyper-paramètres	112
5.4	Combinaison des méthodes d'apprentissage et stratégie finale	115
5.4.1	Combinaison sur l'ensemble des classifieurs non calibrés	115
5.4.2	Combinaison sur l'ensemble des classifieurs calibrés	116
5.4.3	Impact des classifieurs sur la performance de la stratégie à risque équivalent et choix final	117
5.5	Stratégie finale, limites et perspectives	118
5.5.1	Une gamme de stratégies 2C	119
5.5.2	Performances depuis Janvier 2016	120
5.5.3	Limites et perspectives d'études	121
5.6	Conclusion	122

TABLE DES MATIÈRES

6 Conclusion	125
Bibliographie	128

Remerciements

Je tiens tout d'abord à remercier Denis Gerber de m'avoir proposé ce projet chez Das Bot et sans qui cette thèse n'aurait jamais vu le jour. Je le remercie également de m'avoir partagé son expérience et sa vision des marchés financiers et de me faire confiance pour poursuivre ensemble notre collaboration après la thèse.

Je remercie tout particulièrement Emmanuel Lépinette, mon directeur de thèse, pour sa générosité, son soutien, sa patience et son enthousiasme permanents. Je tiens à souligner la qualité de l'encadrement qu'il m'a fourni. Toujours disponible, je n'aurais pas pu imaginer meilleur accompagnement tout au long de ces trois années et demie passionnantes. Cela a été un réel plaisir de travailler ensemble et d'aboutir à plusieurs travaux en cours de soumission et à des résultats concluants sur les stratégies de trading.

Je suis reconnaissant au personnel de l'université Paris-Dauphine et du CEREMADE pour leur aide dans toutes les démarches administratives.

Je remercie Stéphane Elard pour sa gentillesse et son aide précieuse pour la gestion et l'utilisation des données de marché. Je remercie Laurence Carassus pour sa collaboration fructueuse. Je remercie également Victor Lora et Léo Valentin de l'espace de co-working GII où j'ai travaillé un peu plus d'un an. L'accueil et l'ambiance y ont été formidables.

Je remercie Fabrice Moullé-Berteaux, Arnauld Chamelot, Etienne Touraille, Guillaume Nicoulaud et Jean Raoul-Duval pour leur accueil et la bonne humeur régnant au sein de la société Sagara Financière qui héberge les locaux de Das Bot depuis Mai 2017.

Je termine en remerciant l'ensemble de mes proches pour leurs encouragements. Je remercie mes chats Hiro et Wendy pour leur soutien indéfectible lors des journées passées à travailler à la maison, bien que leur contribution se résume à de longues siestes sur mes genoux. Enfin je remercie ma compagne Sömayya, à qui je dédie cette thèse, qui a toujours cru en moi et qui m'a soutenu tout au long de cette aventure.

Résumé

Cette thèse CIFRE est composée d'une partie théorique abordant des problèmes d'évaluation de produits dérivés et d'une partie pratique portant sur la construction d'une stratégie de trading algorithmique basée sur des méthodes d'apprentissage automatique.

Nous nous demandons s'il est possible d'évaluer le prix d'une option sans avoir recours à une probabilité de risque neutre. Nous proposons une nouvelle approche nous menant à une condition faible d'absence d'opportunité d'arbitrage qui stipule que le prix d'un call de payoff nul devrait être nul. Dans un second travail, nous étudions la convergence du schéma fonctionnel dérivé du modèle Binomial et plus généralement du modèle introduit précédemment avec volatilité locale lorsque la fonction de payoff n'est pas régulière. Nous illustrons les résultats par des exemples et comparons les performances avec d'autres méthodes de résolution. La partie théorique s'achève avec l'étude d'une méthode d'approximation de solutions d'équations aux dérivées stochastiques par itérations de Picard déduite de la formule de Doléans-Dade. La méthode est appliquée au pricing d'options asiatiques dont le sous-jacent a pour dynamique un processus CEV.

Dans la partie pratique, nous revisitons la stratégie Open Range Breakout, qui repose sur des éléments techniques visant à tirer parti d'une tendance des prix d'un actif à évoluer dans une certaine direction plutôt qu'à évaluer et prédire le prix d'un actif. Nous appliquons des méthodes d'apprentissage supervisé et montrons qu'il est possible de créer une stratégie viable et rentable prenant en compte les coûts d'exécution. Nous terminons par la présentation d'une gamme de stratégies répondant à plusieurs profils d'investisseurs.

Abstract

This thesis consists of a theoretical section addressing derivatives pricing problems followed by an empirical section focused on designing a machine learning oriented algorithmic trading strategy.

We question whether it is possible to price a contingent claim option without using a risk-neutral probability measure. We establish a new approach leading to a weak condition of absence of arbitrage opportunity which asserts that the price of the zero claim should be zero. We then study the convergence of the functional scheme derived from the Binomial tree financial model and more generally from the previously introduced model with local volatility when the payoff function is not smooth. We illustrate our results with numerical examples and compare performances with other resolution methods. We conclude this theoretical part by studying an approximation method based on Picard iterations deduced from the Doleans-Dade exponential formula. We illustrate this by using this method to price an Asian-style contingent claims in the CEV model.

In this empirical section, we revisit the Open Range Breakout strategy whose trade signal relies on a set of technical rules aimed at profiting from a directional trend in the prices of an asset (as opposed to establishing a valuation of that assets price). We apply supervised learning methods to show that it is possible to create a profitable and realistic strategy taking execution costs into account. We conclude this section by presenting a family of strategies that can fit various investors profiles.

Chapitre 1

Introduction

L'essor des systèmes d'information et des volumes de données depuis le début des années 2000 a permis à de nouveaux acteurs d'accéder aux marchés financiers. Désormais, il n'est plus obligatoire d'être une grande banque ou un fond d'investissement de premier plan pour recevoir et analyser les cours en temps réel, développer des systèmes automatisés de trading et investir sur des périodes très courtes allant de quelques secondes à quelques heures.

C'est le chemin qu'a pris Das Bot, jeune entreprise innovante fondée en 2012 par Denis Gerber. Das Bot est un laboratoire de recherche aux frontières de la finance quantitative, la finance comportementale, l'analyse temps réel de gros volumes de données et le trading algorithmique. Ses travaux portent sur le développement de stratégies de trading reposant sur la détection d'anomalies de marchés permettant d'anticiper les variations à court terme du cours d'un actif. C'est dans le cadre de ses activités de recherche que Denis m'a proposé une collaboration CIFRE.

Après une première tentative de thèse avortée début 2013, nous avons repris notre projet à l'automne 2014. Entre temps, je suis parti pour un VIE d'un an en Belgique où j'ai été analyste quantitatif junior à la validation de modèles chez BNP Paribas. Au sein d'une équipe indépendante située entre la recherche quantitative et le front office, j'ai été en charge de passer en revue les méthodologies de calcul des valeurs de risque (VaR, Expected Effective Positive Exposure, Credit Value Adjustment ...) pour les dérivés de crédit. Cette expérience sera fondamentale car j'y découvrirai la recherche en entreprise et une méthodologie de travail que j'applique toujours aujourd'hui.

Lorsque nous rencontrons Emmanuel Lépinette, nous formons rapidement un projet de thèse proposant à la fois un travail académique sur l'évaluation de produits dérivés et une partie pratique sur le développement de stratégies de trading algorithmique repo-

sant sur des indicateurs techniques. L'objectif est d'explorer différentes facettes de la finance quantitative d'aujourd'hui et de tisser un lien entre la recherche universitaire et l'industrie financière.

Le contenu de la thèse présente donc deux parties bien distinctes. Une première partie correspond à trois articles en cours de soumission. Il s'agit de traiter différents aspects numériques du problème d'évaluation d'options. Dans une seconde partie, nous revisitons une stratégie nommée Open Range Breakout à laquelle nous appliquons des méthodes d'apprentissage supervisé.

Le chapitre 2 se pose la question de l'évaluation du prix d'une option sans avoir recours à une probabilité de risque neutre. Nous proposons une nouvelle approche basée sur la dualité convexe plutôt que la dualité de mesures de martingales. Les prix sont alors exprimés en utilisant la conjuguée et la bi-conjuguée de Fenchel. Cela nous mène à une condition faible d'absence d'opportunité d'arbitrage stipulant que le prix minimum d'un call de payoff nul devrait être nul. Nous montrons dans un modèle à deux dates que, sous cette nouvelle condition, le prix minimal de sur-réplication est déterminé par l'enveloppe concave de la fonction de payoff et que la condition est équivalente à la non-négativité des prix de sur-réplication pour les call. Avec plus de dates, nous proposons un schéma récursif pour le calcul du prix de sur-réplication d'une option à fonction de payoff convexe. Nous terminons par des exemples numériques. Ces résultats ont été obtenus en collaboration avec Laurence Carassus pour un papier en cours de soumission.

Le chapitre 3 prolonge le travail du chapitre 2. Nous reprenons le schéma récursif présenté précédemment, dont le modèle binomial est un cas particulier, et nous généralisons les résultats du papier [55] au cas d'une condition terminale non régulière ce qui se révèle important dans certains cas comme par exemple la fonction de payoff d'un call $g(x) = (x - K)^+$ qui n'est pas continûment dérivable. La fonction solution du schéma fonctionnel dérivé du modèle binomial avec volatilité locale converge vers la solution d'une équation de diffusion du type $h_t(t, x) + \frac{x^2 \sigma^2(t, x)}{2} h_{xx}(t, x) = 0$ lorsque le nombre de dates $n \rightarrow \infty$. Nous établissons la convergence uniforme en temps du schéma - dont la résolution est uniquement basée sur une discrétisation en temps du problème - et donnons le taux de convergence lorsque la fonction de payoff n'est pas aussi régulière qu'en finance. Nous illustrons ces résultats par des exemples et comparons les performances du schéma fonctionnel avec celles des méthodes aux différences finies et aux éléments finis.

Le chapitre 4 reprend un projet inachevé, proposé par Emmanuel Lépinette à Julien Grépat dans le cadre de sa thèse. Ce travail fait suite au papier [54] proposant une méthode basée sur l'utilisation de la forme Doléans-Dade du prix S_t d'un sous-jacent pour une option européenne de payoff $(S_T - K)^+$ et dont la dynamique est définie par

$\frac{dS_t}{dt} = \sigma(S_t, t)dW_t + \mu(S_t, t)dt$ ou encore $S_t = S_0 \exp\left(\int_0^t \sigma(S_u, u)dW_u + \int_0^t \mu(S_u, u)du\right)$. Ainsi, en pensant au théorème sur point fixe, nous sommes amenés à construire le schéma $S_t^{n+1} = S_0 \exp\left(\int_0^t \sigma(S_u^n, u)dW_u + \int_0^t \mu(S_u^n, u)du\right)$. Ma contribution a été d'obtenir une convergence du schéma proposé sous des conditions plus faibles sur les coefficients (volatilité locale) afin d'inclure le plus de modèles possible puis de tester la convergence sur plusieurs modèles par simulations en C++. Nous proposons une méthode d'approximation de solutions d'EDS par itérations de Picard déduite de la formule de l'exponentielle de Doléans-Dade. La méthode permet d'approximer les trajectoires d'une large classe de processus de Markov, i.e. solutions d'EDS à coefficients non-Lipschitz. Nous terminons par une application au pricing d'options asiatiques dont le sous-jacent a pour dynamique un processus CEV. Les résultats et les temps de calculs sont comparés à d'autres méthodes de la littérature.

Le chapitre 5, présente la stratégie d'Open Range Breakout détectant l'initiation de tendances sur la base d'indicateurs techniques. Nous l'appliquons sur le contrat future DAX et montrons qu'il est possible d'obtenir une stratégie profitable basée uniquement sur l'analyse de la donnée passée tout en prenant en compte les coûts d'exécution (frais de transaction et slippage). Nous appliquons dans un second temps des méthodes d'apprentissage supervisé et montrons que l'impact sur les performances de la stratégie sont statistiquement significatifs. Nous terminons par la présentation d'une gamme de stratégies réalistes adaptées à différents profils d'investisseurs.

En conclusion, nous dressons un résumé des principaux résultats obtenus. Nous en profitons également pour dresser un bilan des contributions effectués chez Das Bot lors de cette thèse et qui ne sont pas forcément visibles dans les travaux présentés précédemment.

Première partie

Évaluation du prix d'options

Chapitre 2

Pricing without martingale measure

Durant des décennies, les conditions d'absence d'opportunité d'arbitrage - ou no-arbitrage (NA) condition - liées aux mesures de martingale ont joué un rôle majeur dans la théorie d'évaluation des produits dérivés. Seulement il est difficile en pratique de vérifier l'existence de la probabilité de martingale à partir de l'observation des prix. La question que nous nous sommes posée peut se formuler ainsi : peut-on évaluer le prix d'une option sans avoir recours à une probabilité de risque neutre ? Ou de manière équivalente sous une condition d'absence d'opportunité d'arbitrage plus faible ?

Nous proposons dans cette partie une nouvelle approche basée sur la dualité convexe plutôt que la dualité de mesures de martingales : nos prix seront exprimés en utilisant la conjuguée et la bi-conjuguée de Fenchel. Cela nous mène naturellement à une condition faible d'absence d'opportunité d'arbitrage, appelée Absence de Profit Immédiat - Absence of Immediate Profit (AIP) - qui stipule que le prix minimum d'un call de payoff nul devrait être nul. Nous étudions le lien entre (AIP), (NA) et la condition No Free Lunch. Nous montrons dans un modèle à deux dates que, sous (AIP), le prix minimal de sur-réplication est déterminé par l'enveloppe concave de la fonction de payoff et que la condition (AIP) est équivalente à la non-négativité des prix de sur-réplication pour les call. Avec plus de dates, nous proposons un schéma récursif pour le calcul du prix de sur-réplication d'une option à fonction de payoff convexe. Nous terminons par des exemples numériques.

2.1 Introduction

The problem of giving a price to a financial asset G is central in the economic and financial theory. A selling price should be an amount which is enough to initiate a hedging strategy for G , i.e. a strategy whose value at maturity is always above G . It seems also natural to ask for the infimum of such amount. This is the so called super-replication price and it has been introduced in the binomial setup for transaction costs by [4]. Characterising and computing the super-replication price has become one of the central issue in mathematical finance theory. Until now it was intimately related to the No-Arbitrage (NA) condition. This condition asserts that starting from a zero wealth it is not possible to reach a positive one (non negative almost surely and strictly positive with strictly positive probability measure). Characterizing the (NA) condition or, more generally, the No Free Lunch condition leads to the Fundamental Theorem of Asset Pricing (FTAP in short). This theorem proves the equivalence between those absence of arbitrage conditions and the existence of risk-neutral probability measures (also called martingale measures or pricing measures) which are equivalent probability measures under which the (discounted) asset price process is a martingale. This was initially formalised in [33], [34] and [47] while in [16] the FTAP is formulated in a general discrete-time setting under the (NA) condition. The literature on the subject is huge and we refer to [20] and [43] for a general overview. Under the (NA) condition, the super-replication price of G is equal to the supremum of the (discounted) expectation of G computed under the risk-neutral probability measures. This is the so called dual formulation of the super-replication price or Superhedging Theorem. We refer to [29] and the references therein.

In this paper a super-hedging or super-replicating price is the initial value of some super-hedging strategy. We propose an innovating approach : we analyse from scratch the set of super-replicating prices and its infimum value, which will be called the infimum super-replication cost. Note that this cost is is not automatically a super-replicating price. Under mild assumptions, we show that the one-step set of super-replication prices can be expressed using Fenchel-Legendre conjugate and the infimum super-replication cost is obtained by the Fenchel-Legendre biconjugate. So we use here the convex duality instead of the usual financial duality based on martingale measures under the (NA) condition. We then introduce the condition of Absence of Immediate Profit (AIP). An Immediate Profit is the possibility of super-hedging 0 at a negative cost. We prove that (AIP) is equivalent to the fact that the stock value at the beginning of the period belongs to the convex envelop of the conditional (with respect to the information of the beginning

of the period) support of the stock value at the end of the period. Using the notion of conditional essential supremum, it is equivalent to say that the initial stock price is between the conditional essential infimum and supremum of the stock value at the end of the period. Under (AIP) condition we show that the one-step infimum super-replication cost is the concave envelop of the payoff relatively to the convex envelop of the conditional support. We also show that (AIP) is equivalent to the non-negativity of the super-hedging prices of any fixed call option. We then study the multiple-period framework. We show that the global (AIP) condition and the local ones are equivalent. We then focus on a particular, but still general setup, where we propose a recursive scheme for the computation of the super-hedging prices of a convex option. We obtain the same computative scheme as in [10] and [11] but here it is obtained by only assuming (AIP) instead of the stronger (NA) condition. We also give some numerical illustrations ; we calibrate historical data of the french index CAC 40 to our model and implement the super-hedging strategy for a call option.

Finally, we study the link between (AIP), (NA) and the weak no-free lunch (WNFL) conditions. We show that the (AIP) condition is the weakest-one and we also provide conditions for the equivalence between the (AIP) and the (WNFL) conditions.

The paper is organized as follows. In Section 2.2, we study the one-period framework while in Section 2.3 we study the multi-period one. Section 2.4 is devoted to the comparison between (AIP), (NA) and (WNFL) conditions. Section 2.5 contains the numerical experiments. Finally, Section 2.6 collects the results on conditional support and conditional essential supremum.

In the remaining of this introduction we introduce our framework and recall some results that will be used without further references in the sequel. Let $(\Omega, (\mathcal{F}_t)_{t \in \{0, \dots, T\}}, \mathcal{F}_T, P)$ be a filtered probability space where T is the time horizon. We consider a $(\mathcal{F}_t)_{t \in \{0, \dots, T\}}$ -adapted, real-valued, non-negative process $S := \{S_t, t \in \{0, \dots, T\}\}$, where for $t \in \{0, \dots, T\}$, S_t represents the price of some risky asset in the financial market in consideration. Trading strategies are given by $(\mathcal{F}_t)_{t \in \{0, \dots, T\}}$ -adapted processes $\theta := \{\theta_t, t \in \{0, \dots, T-1\}\}$ where for all $t \in \{0, \dots, T-1\}$, θ_t represents the investor's holding in the risky asset between time t and time $t+1$.

We assume that trading is self-financing and that the riskless asset's price is constant equal to 1. The value at time t of a portfolio θ starting from initial capital $x \in \mathbb{R}$ is given by

$$V_t^{x, \theta} = x + \sum_{u=1}^t \theta_{u-1} \Delta S_u.$$

For any σ -algebra \mathcal{H} and any $k \geq 1$, we denote by $L^0(\mathbf{R}^k, \mathcal{H})$ the set of \mathcal{H} -measurable

and \mathbf{R}^k -valued random variables. Let $h : \Omega \times \mathbf{R}^k \rightarrow \mathbf{R}$. The effective domain of $h(\phi, \cdot)$ is $\text{dom } h(\phi, \cdot) = \{x \in \mathbf{R}^k, h(\phi, x) < \infty\}$ and $h(\phi, \cdot)$ is proper if $\text{dom } h(\phi, \cdot) \neq \emptyset$ and $h(\phi, x) > -\infty$ for all $x \in \mathbf{R}^k$. Then h is \mathcal{H} -normal integrand (see Definition 14.27 in [62]) if and only if h is $\mathcal{H} \otimes \mathcal{B}(\mathbf{R}^k)$ -measurable and is lower semi-continuous (l.s.c. in the sequel) in x , see [62, Corollary 14.34]. Let $Z \in L^0(\mathbf{R}^k, \mathcal{H})$, we will use the notation $h(Z) : \phi \rightarrow h(Z(\phi)) = h(\omega, Z(\omega))$ and if h is $\mathcal{H} \otimes \mathcal{B}(\mathbf{R}^k)$ -measurable, $h(Z) \in L^0(\mathbf{R}^k, \mathcal{H})$. Let \mathcal{K} be a \mathcal{H} -measurable (see Definition 14.1 of [62]) and closed-valued random set of \mathbf{R}^k then \mathcal{K} admits a Castaing representation $(\eta_n)_{n \in \mathbf{N}}$ (see Theorem 14.5 in [62]) : $\mathcal{K}(\phi) = \text{cl}\{\eta_n(\phi), n \in \mathbf{N}\}$ for all $\phi \in \text{dom } \mathcal{K} = \{\phi \in \Omega, \mathcal{K}(\phi) \cap \mathbf{R}^k \neq \emptyset\}$, where the closure is taken in \mathbf{R}^k .

2.2 The one-period framework

For ease of notation, we consider two complete sub- σ -algebras of \mathcal{F}_T : $\mathcal{H} \subseteq \mathcal{F}$ and two random variables $y \in L^0(\mathbf{R}_+, \mathcal{H})$ and $Y \in L^0(\mathbf{R}_+, \mathcal{F})$. The setting will be applied in Section 2.3 with the choices $\mathcal{H} = \mathcal{F}_t$, $\mathcal{F} = \mathcal{F}_{t+1}$, $Y = S_{t+1}$, $y = S_t$.

Section's objective is to obtain a characterisation of the one-step set of super-hedging or super-replicating prices of $g(Y)$ under suitable assumptions on $g : \Omega \times \mathbf{R} \rightarrow \mathbf{R}$.

In the following, the notion of conditionnal support ($\text{supp}_{\mathcal{H}} Y$), conditional essential infimum ($\text{ess inf}_{\mathcal{H}}$) or supremum ($\text{ess sup}_{\mathcal{H}}$) will be in force, see Section 2.6.

Definition 2.2.1. *The set $\mathcal{P}(g)$ of super-hedging prices of the contingent claim $g(Y)$ consists in the initial values of super-hedging strategies θ :*

$$\mathcal{P}(g) = \{x \in L^0(\mathbf{R}, \mathcal{H}), \exists \theta \in L^0(\mathbf{R}, \mathcal{H}), x + \theta(Y - y) \geq g(Y) \text{ a.s.}\}.$$

The infimum super-hedging cost of $g(Y)$ is defined as $p(g) := \text{ess inf}_{\mathcal{H}} \mathcal{P}(g)$.

Notice that the infimum super-hedging cost is not a priori a price, i.e. an element of $\mathcal{P}(g)$, as the later may be an open interval.

Remark 2.2.2. As $P(Y \in \text{supp}_{\mathcal{H}} Y) = 1$ (see [1, definition of support on page 441, Theorems 12.7 and 12.14]), we have that $\text{supp}_{\mathcal{H}} Y$ is a.s. non-empty. Moreover since $0 \leq Y < \infty$, $\text{Dom } \text{supp}_{\mathcal{H}} Y = \Omega$. We could easily include the case $P(0 \leq Y < \infty) = 1$ by replacing Y by 0 on the complementary of $\{0 \leq Y < \infty\}$.

Lemma 2.2.3.

$$\mathcal{P}(g) = \{\text{ess sup}_{\mathcal{H}} (g(Y) - \theta Y) + \theta y, \theta \in L^0(\mathbf{R}, \mathcal{H})\} + L^0(\mathbf{R}_+, \mathcal{H}). \quad (2.2.1)$$

Suppose that g is a \mathcal{H} -normal integrand. Then

$$\operatorname{ess\,sup}_{\mathcal{H}}(g(Y) - \theta Y) = \sup_{z \in \operatorname{supp}_{\mathcal{H}} Y} (g(z) - \theta z) = f^*(-\theta), \quad (2.2.2)$$

where f^* is the Fenchel-Legendre conjugate of f i.e.

$$\begin{aligned} f^*(\phi, x) &= \sup_{z \in \mathbf{R}} (xz - f(\phi, z)), \\ f(\phi, z) &= -g(\phi, z) + \delta_{\operatorname{supp}_{\mathcal{H}} Y}(\phi, z), \end{aligned} \quad (2.2.3)$$

where $\delta_C(\phi, z) = 0$ if $z \in C(\phi)$ and $+\infty$ else. Both $f^*(\phi, \cdot)$ and $x \rightarrow f^*(\phi, -x)$ are a.s. proper, convex, l.s.c., f^* is $\mathcal{H} \otimes \mathcal{B}(\mathbf{R})$ -measurable and f^* is a \mathcal{H} -normal integrand. Moreover, we have that

$$p(g) = -f^{**}(y),$$

where f^{**} is the Fenchel-Legendre biconjugate of f i.e.

$$f^{**}(\phi, x) = \sup_{z \in \mathbf{R}} (xz - f^*(\phi, z)).$$

Proof. As $x \in \mathcal{P}(g)$ if and only if there exists $\theta \in L^0(\mathbf{R}, \mathcal{H})$ such that $x - \theta y \geq g(Y) - \theta Y$ a.s., we get by definition of the conditional essential supremum (see Definition 2.6.3) that (2.2.1) holds true. Then (2.2.2) follows from Lemma 2.6.8 (see Remark 2.2.2). Since the graph of $\operatorname{supp}_{\mathcal{H}} Y$ belongs to $\mathcal{H} \otimes \mathcal{B}(\mathbf{R})$ (see Lemma 2.6.2), we easily deduce that $\delta_{\operatorname{supp}_{\mathcal{H}} Y}$ is $\mathcal{H} \otimes \mathcal{B}(\mathbf{R})$ -measurable and it is clear that it is also l.s.c. As $\operatorname{dom} f = \operatorname{supp}_{\mathcal{H}} Y$ is a.s. non-empty (see again Remark 2.2.2) $f^*(\phi, \cdot)$ is convex and l.s.c. as the supremum of affine functions. Hence $x \rightarrow f^*(\phi, -x)$ is also a.s. l.s.c. and convex. Moreover, using Lemma 2.6.6 (and Remark 2.2.2), f^* is $\mathcal{H} \otimes \mathcal{B}(\mathbf{R})$ -measurable.

$$\begin{aligned} p(g) &= \operatorname{ess\,inf}_{\mathcal{H}} \{f^*(-\theta) + \theta y, \theta \in L^0(\mathbf{R}, \mathcal{H})\} \\ &= -\operatorname{ess\,sup}_{\mathcal{H}} \{\theta y - f^*(\theta), \theta \in L^0(\mathbf{R}, \mathcal{H})\} \\ &= -\sup_{z \in \mathbf{R}} (zy - f^*(z)) = -f^{**}(y). \end{aligned}$$

The first equality is a direct consequence of (2.2.1), the second one is trivial. In order to obtain the third one, we want to apply Lemma 2.6.9. First remark that $\operatorname{ess\,sup}_{\mathcal{H}} \{\theta y - f^*(\theta), \theta \in L^0(\mathbf{R}, \mathcal{H})\} = \operatorname{ess\,sup}_{\mathcal{H}} \{\theta y - f^*(\theta), \theta \in L^0(\mathbf{R}, \mathcal{H}) \cap \operatorname{Dom} f^*\}$. Now since f^* is $\mathcal{H} \otimes \mathcal{B}(\mathbf{R})$ -measurable, we deduce that $\operatorname{graph\,dom} f^* = \{(\omega, x) \in \Omega \times \mathbf{R}, f^*(\omega, x) < \infty\}$ is a $\mathcal{H} \otimes \mathcal{B}(\mathbf{R})$ -measurable set and $\operatorname{dom} f^*$ is also \mathcal{H} -measurable (see [62, Theorem 14.8]). Since $(\phi, z) \rightarrow zy(\phi) - f^*(\phi, z)$ is a $\mathcal{H} \otimes \mathcal{B}(\mathbf{R})$ -measurable function and $f^*(\phi, \cdot)$ is convex and thus u.s.c. on $\operatorname{dom} f^*(\phi)$, we can apply Lemma 2.6.9 and we obtain that

$$\begin{aligned} \text{ess sup}_{\mathcal{H}}\{\theta y - f^*(\theta), \theta \in L^0(\mathbf{R}, \mathcal{H}) \cap \text{Dom } f^*\} &= \sup_{z \in \text{Dom}(f^*)} (zy - f^*(z)) \\ &= \sup_{z \in \mathbf{R}} (zy - f^*(z)). \end{aligned}$$

□

Let $\text{conv } h$ be the convex envelop of h i.e. the greatest convex function dominated by h

$$\text{conv } h(x) = \sup\{u(x), u \text{ convex and } u \leq h\}.$$

The concave envelop is defined symmetrically and denoted by $\text{conc } h$. We also define the (lower) closure \underline{h} of h as the greatest l.s.c. function which is dominated by h i.e. $\underline{h} = \liminf_{y \rightarrow x} h(y)$. The upper closure is defined symmetrically : $\bar{h} = \limsup_{y \rightarrow x} h(y)$. It is easy to see that

$$\underline{\text{conv}} f(y) = \sup\{\alpha y + \beta, \alpha, \beta \in \mathbf{R}, f(x) \geq \alpha x + \beta, \forall x \in \mathbf{R}\}.$$

It is well-known (see for example [62, Theorem 11.1]) that

$$f^* = (\text{conv } f)^* = (\underline{f})^* = (\underline{\text{conv}} f)^*. \quad (2.2.4)$$

Moreover, if $\text{conv } f$ is proper, f^{**} is also proper, convex and l.s.c. and

$$f^{**} = \underline{\text{conv}} f. \quad (2.2.5)$$

So in order to compute $p(g)$, we need to compute $\text{conv } f$ and $\underline{\text{conv}} f$. To do so, we introduce the notion of relative concave envelop of g with respect to $\text{supp}_{\mathcal{H}} Y$:

$$\text{conc}(g, \text{supp}_{\mathcal{H}} Y)(x) = \inf\{v(x), v \text{ is concave and } v(z) \geq g(z), \forall z \in \text{supp}_{\mathcal{H}} Y\}.$$

In the following, we use the convention $0 \times (\pm\infty) = 0$ and $(+\infty) \times 0 = 0$.

Lemma 2.2.4. *Suppose that g is a \mathcal{H} -normal integrand. Then, we have :*

$$\begin{aligned} \text{conv } f(x) &= +\infty \mathbf{1}_{\mathbf{R} \setminus \text{conv supp}_{\mathcal{H}} Y}(x) - \mathbf{1}_{\text{conv supp}_{\mathcal{H}} Y}(x) \times \text{conc}(g, \text{supp}_{\mathcal{H}} Y)(x) \\ \underline{\text{conv}} f(x) &= +\infty \mathbf{1}_{\mathbf{R} \setminus \text{conv supp}_{\mathcal{H}} Y}(x) - \mathbf{1}_{\text{conv supp}_{\mathcal{H}} Y}(x) \times \overline{\text{conc}}(g, \text{supp}_{\mathcal{H}} Y)(x) \\ &= +\infty \mathbf{1}_{\mathbf{R} \setminus \text{conv supp}_{\mathcal{H}} Y}(x) - \mathbf{1}_{\text{conv supp}_{\mathcal{H}} Y}(x) \times \\ &\quad \inf\{\alpha x + \beta, \alpha, \beta \in \mathbf{R}, \alpha z + \beta \geq g(z), \forall z \in \text{supp}_{\mathcal{H}} Y\}, \end{aligned}$$

where $\text{conv supp}_{\mathcal{H}} Y$ is the convex envelop of $\text{supp}_{\mathcal{H}} Y$, i.e. the smallest convex set that contains $\text{supp}_{\mathcal{H}} Y$.

Remark 2.2.5. Note that $\text{conv } f$ is proper if and only if $\text{conc}(g, \text{supp}_{\mathcal{H}}Y)(x) < +\infty$ for all $x \in \text{convsupp}_{\mathcal{H}}Y$, since $\text{convsupp}_{\mathcal{H}}Y$ is non-empty (see Remark 2.2.2). So $\text{conv } f$ is proper if there exists some concave function φ such that $g \leq \varphi$ on $\text{supp}_{\mathcal{H}}Y$ and $\varphi < \infty$ on $\text{convsupp}_{\mathcal{H}}Y$ (by definition, $\text{conc}(g, \text{supp}_{\mathcal{H}}Y) \leq \varphi$).¹ As for all $x \in \text{convsupp}_{\mathcal{H}}Y$, $\text{conc}(g, \text{supp}_{\mathcal{H}}Y)(x) \geq g(x) > -\infty$, we get that $\text{conc}(g, \text{supp}_{\mathcal{H}}Y)(x) \in \mathbf{R}$ and also $\overline{\text{conc}}(g, \text{supp}_{\mathcal{H}}Y)(x) \in \mathbf{R}$, one may write that

$$\begin{aligned} \text{conv } f &= -\text{conc}(g, \text{supp}_{\mathcal{H}}Y) + \delta_{\text{convsupp}_{\mathcal{H}}Y}. \\ \underline{\text{conv}} f &= -\overline{\text{conc}}(g, \text{supp}_{\mathcal{H}}Y) + \delta_{\text{convsupp}_{\mathcal{H}}Y}. \end{aligned}$$

Proof. One can rewrite the convex envelop of f as follows (see [62, Proposition 2.31]) :

$$\text{conv } f(x) = \inf \left\{ \sum_{i=1}^n l_i f(x_i), n \geq 1, (l_i)_{i \in \{1, \dots, n\}} \in \mathbf{R}_+^n, (x_i)_{i \in \{1, \dots, n\}} \in \mathbf{R}^n, \right. \\ \left. x = \sum_{i=1}^n l_i x_i, \sum_{i=1}^n l_i = 1 \right\}.$$

The convex envelop of $\text{supp}_{\mathcal{H}}Y$ is given by (see [62, Proposition 2.27, Theorem 2.29]).

$$\text{convsupp}_{\mathcal{H}}Y = \left\{ \sum_{i=1}^n l_i x_i, n \geq 1, (l_i)_{i \in \{1, \dots, n\}} \in \mathbf{R}_+^n, \sum_{i=1}^n l_i = 1, x_i \in \text{supp}_{\mathcal{H}}Y \right\}.$$

Assume that $x \notin \text{convsupp}_{\mathcal{H}}Y$. Then if $x = \sum_{i=1}^n l_i x_i$ for some $n \geq 1, (l_i)_{i \in \{1, \dots, n\}} \in \mathbf{R}_+^n, (x_i)_{i \in \{1, \dots, n\}} \in \mathbf{R}^n$ such that $\sum_{i=1}^n l_i = 1$, at least one $x_i \notin \text{supp}_{\mathcal{H}}Y$ and $f(x_i) = \infty$ and also $\text{conv } f(x) = \infty = +\infty \mathbf{1}_{\mathbf{R} \setminus \text{convsupp}_{\mathcal{H}}Y}(x) - \mathbf{1}_{\text{convsupp}_{\mathcal{H}}Y}(x) \times \text{conc}(g, \text{supp}_{\mathcal{H}}Y)(x)$. If $x \in \text{convsupp}_{\mathcal{H}}Y$, $\text{conv } f(x) = -\text{conc}(g, \text{supp}_{\mathcal{H}}Y)(x)$. One can also remark that if $x \in \text{convsupp}_{\mathcal{H}}Y$,

$$\underline{\text{conv}} f(x) = -\inf \{ \alpha x + \beta, \alpha, \beta \in \mathbf{R}, g(z) \leq \alpha z + \beta, \forall z \in \text{supp}_{\mathcal{H}}Y \}.$$

□

So we have the following representation of the infimum super-hedging cost :

Proposition 2.2.6. *Suppose that g is a \mathcal{H} -normal integrand and that there exists some concave function φ such that $g \leq \varphi$ on $\text{supp}_{\mathcal{H}}Y$ and $\varphi < \infty$ on $\text{convsupp}_{\mathcal{H}}Y$. Then,*

$$p(g) = -\underline{\text{conv}}(f)(y) = \overline{\text{conc}}(g, \text{supp}_{\mathcal{H}}Y)(y) - \delta_{\text{convsupp}_{\mathcal{H}}Y}(y).$$

We see that the fact that y belongs to $\text{convsupp}_{\mathcal{H}}Y$ or not is important. In particular, in some cases, the infimum price of a European claim may be $-\infty$. This is related to the notion of absence of immediate profit that we present now. We say that there is

1. This is equivalent to assume that there exists $\alpha, \beta \in \mathbf{R}$, such that $g(x) \leq \alpha x + \beta$ for all $x \in \text{supp}_{\mathcal{H}}Y$.

an immediate profit when it is possible to super-replicate the contingent claim 0 at a negative price p . This implies that we may immediately make the positive profit $-p$ and then start a portfolio process ending up with a non negative wealth. On the contrary case, we say that the Absence of Immediate Profit (AIP) condition holds. We will see that (AIP) is strictly weaker than (NA).

Definition 2.2.7. *There is an immediate profit (IP) if there exists a non null element of $\mathcal{P}(0) \cap L^0(\mathbf{R}_-, \mathcal{H})$ or equivalently if $p(0) \leq 0$ with $P(p(0) < 0) > 0$.*

Notice that the (AIP) condition may be seen as a particular case of the utility based No Good Deal condition introduced by Cherny, see [12, Definition 3]. In the definition above, let us explain why $p(0) \leq 0$ with $P(p(0) < 0) > 0$ implies the existence of an immediate profit (IP). To see it, recall that $\mathcal{P}(0)$ is directed downward so that $p(0) = \lim_n \downarrow p_n$ where $p_n \in \mathcal{P}(0)$. Since $P(p(0) < 0) > 0$, we deduce that there exists n such that $P(p_n < 0) > 0$. Let us define $\tilde{p} = p_n 1_{p_n < 0}$. Then, $\tilde{p} \in \mathcal{P}(0) \cap L^0(\mathbf{R}_-, \mathcal{H})$ and $\tilde{p} \neq 0$, i.e. \tilde{p} generates an immediate profit.

Proposition 2.2.8. *(AIP) holds if and only if $y \in \text{convsupp}_{\mathcal{H}} Y$ a.s.*

Notice that, from Lemma 2.6.10, we get that

$$\text{convsupp}_{\mathcal{H}} Y = [\text{ess inf}_{\mathcal{H}} Y, \text{ess sup}_{\mathcal{H}} Y] \cap \mathbf{R}.$$

Proof. The assumptions of Proposition 2.2.6 are satisfied for $g = 0$ and we get that $p(0) = -\delta_{\text{convsupp}_{\mathcal{H}} Y}(y)$. Hence, there is no immediate profit if and only if $y \in \text{convsupp}_{\mathcal{H}} Y$ a.s. \square

Corollary 2.2.9. *The (AIP) condition holds true if and only if $p(g) \geq 0$ a.s. for some non-negative \mathcal{H} -normal integrand g such that there exists some concave function φ verifying that $g \leq \varphi < \infty$.*

So in particular the (AIP) condition holds true if and only the infimum super-hedging cost of a european call option is non-negative.

Proof. Assume that (AIP) condition holds true. Then from Definition 2.2.7, we get that $p(0) = 0$ a.s. As $g \geq 0$, it is clear that $p(g) \geq p(0) = 0$ a.s. Conversely, assume that there exists some (IP). From Proposition 2.2.6, we get that

$$p(g) = \overline{\text{conc}}(g, \text{supp}_{\mathcal{H}} Y)(y) - \delta_{\text{convsupp}_{\mathcal{H}} Y}(y).$$

From Proposition 2.2.8, we get that $P(y \in \text{convsupp}_{\mathcal{H}} Y) < 1$ and as $\overline{\text{conc}}(g, \text{supp}_{\mathcal{H}} Y)(y) \leq \varphi < \infty$, $P(p(g) = -\infty) > 0$ and the converse is proved. \square

Remark 2.2.10. Assume that the \mathcal{H} -measurable set $\Gamma = \{\text{ess sup}_{\mathcal{H}} Y < y\}$ has a non null probability. Then, on this set, from the zero initial capital, taking the physical position $(y, -1)$ while keeping the zero position otherwise, one get at time 1 the terminal wealth $y - Y \geq y - \text{ess sup}_{\mathcal{H}} Y > 0$ on Γ and zero otherwise, i.e. an arbitrage opportunity. Thus if $y \notin \text{convsupp}_{\mathcal{H}} Y$ a.s., one gets an Arbitrage Opportunity and (AIP) is weaker than (NA).

We provide some examples where (AIP) holds true and is strictly weaker than (NA). This is the case if there exists $Q_1, Q_2 \ll P$ such that S is a Q_1 -super martingale (resp. Q_2 -sub martingale), see Remark 2.6.4. This is of course true if $\text{ess inf}_{\mathcal{H}} Y = 0$ and $\text{ess sup}_{\mathcal{H}} Y = \infty$. Finally, this is also the case for a model of the form $Y = yZ$ where $Z > 0$ is such that $\text{supp}_{\mathcal{H}} Z = [0, 1]$ a.s. or $\text{supp}_{\mathcal{H}} Z = [1, \infty)$ a.s. and $y > 0$. Indeed (recall Lemma 2.6.10), if $\text{supp}_{\mathcal{H}} Z = [0, 1]$, $\text{ess inf}_{\mathcal{H}} Y = y \text{ess inf}_{\mathcal{H}} Z = 0 \leq y$ and $\text{ess sup}_{\mathcal{H}} Y = y \text{ess sup}_{\mathcal{H}} Z = y \geq y$. The same holds if $\text{supp}_{\mathcal{H}} Z = [1, \infty)$ a.s. Nevertheless, this kind of model does not admit a risk-neutral probability measure. Indeed, in the contrary case, there exists a density process i.e. a positive martingale $(\rho_t)_{t=0,1}$ with $\rho_0 = 1$ such that ρS is a P-martingale : $\mathbb{E}_P(\rho_1 Y | \mathcal{H}) = \rho_0 y$. We get that $\mathbb{E}_P(\rho_1 Z | \mathcal{H}) = \rho_0$. Since we also have $\rho_0 = \mathbb{E}_P(\rho_1 | \mathcal{H})$, we deduce that $\mathbb{E}_P(\rho_1(1 - Z) | \mathcal{H}) = 0$. Since $Z \leq 1$ a.s. or $Z \geq 1$ a.s., this implies that $\rho_1(1 - Z) = 0$ hence $Z = 1$ which yields a contradiction.

Corollary 2.2.11. *Suppose that (AIP) holds true. Let g be a \mathcal{H} -normal integrand, such that there exists some concave function φ verifying that $g \leq \varphi$ on $\text{supp}_{\mathcal{H}} Y$ and $\varphi < \infty$ on $\text{convsupp}_{\mathcal{H}} Y$. Then,*

$$\begin{aligned} p(g) &= \overline{\text{conc}}(g, \text{supp}_{\mathcal{H}} Y)(y) \\ &= \inf \{ \alpha y + \beta, \alpha, \beta \in \mathbf{R}, \alpha x + \beta \geq g(x), \forall x \in \text{supp}_{\mathcal{H}} Y \}. \end{aligned} \quad (2.2.6)$$

So in the case where g is concave and u.s.c., we get under (AIP) that $p(g) = g(y)$. If g is convex and $\lim_{x \rightarrow \infty} x^{-1}g(x) = M \in \mathbf{R}$, the relative concave envelop of g with respect to $\text{supp}_{\mathcal{H}} Y$ is the affine function that coincides with g on the extreme points of the interval $\text{convsupp}_{\mathcal{H}} Y$ i.e.

$$p(g) = \theta^* y + \beta^* = g(\text{ess inf}_{\mathcal{H}} Y) + \theta^* (y - \text{ess inf}_{\mathcal{H}} Y), \quad (2.2.7)$$

$$\theta^* = \frac{g(\text{ess sup}_{\mathcal{H}} Y) - g(\text{ess inf}_{\mathcal{H}} Y)}{\text{ess sup}_{\mathcal{H}} Y - \text{ess inf}_{\mathcal{H}} Y}, \quad (2.2.8)$$

where we use the conventions $\theta^* = \frac{0}{0} = 0$ in the case $\text{ess sup}_{\mathcal{H}} Y = \text{ess inf}_{\mathcal{H}} Y$ and $\theta^* = \frac{g(\infty)}{\infty} = M$ if $\text{ess inf}_{\mathcal{H}} Y < \text{ess sup}_{\mathcal{H}} Y = +\infty$. Moreover, using (2.2.6), we get that

$\theta^*Y + \beta^* \geq g(Y)$ a.s. (recall that $Y \in \text{supp}_{\mathcal{H}}Y$ a.s., see Remark 2.2.2) and this implies using (2.2.7) that

$$p(g) + \theta^*(Y - y) \geq g \text{ a.s.} \quad (2.2.9)$$

and $p(g) \in \mathcal{P}(g)$.

2.3 The multi-period framework

2.3.1 Multi-period super-hedging prices

For every $t \in \{0, \dots, T\}$ the set \mathcal{R}_t^T of all claims that can be super-replicated from the zero initial endowment at time t is defined by

$$\mathcal{R}_t^T := \left\{ \sum_{u=t+1}^T \theta_{u-1} \Delta S_u - \epsilon_T^+, \theta_{u-1} \in L^0(\mathbf{R}, \mathcal{F}_{u-1}), \epsilon_T^+ \in L^0(\mathbf{R}_+, \mathcal{F}_T) \right\}. \quad (2.3.10)$$

The set of (multi-period) super-hedging prices and the (multi-period) infimum super-hedging cost of some contingent claim $g \in L^0(\mathbf{R}, \mathcal{F}_T)$ at t are given by

$$\begin{aligned} \Pi_{T,T}(g) &:= \{g\} \\ \pi_{T,T}(g) &:= g \\ \Pi_{t,T}(g) &:= \{x_t \in L^0(\mathbf{R}, \mathcal{F}_t), \exists R \in \mathcal{R}_t^T, x_t + R = g \text{ a.s.}\}, t \in \{0, \dots, T-1\} \\ \pi_{t,T}(g) &:= \text{ess inf}_{\mathcal{F}_t} \Pi_{t,T}(g). \end{aligned} \quad (2.3.11)$$

As in the one-period case, it is clear that the infimum super-hedging cost is not necessarily a price in the sense that $\pi_{t,T}(g) \notin \Pi_{t,T}(g)$ when $\Pi_{t,T}(g)$ is not closed. Alternatively, we may define sequentially

$$\begin{aligned} \mathcal{P}_{T,T}(g) &= \{g\} \\ \mathcal{P}_{t,T}(g) &= \{x_t \in L^0(\mathbf{R}, \mathcal{F}_t), \exists \theta_t \in L^0(\mathbf{R}, \mathcal{F}_t), \exists p_{t+1} \in \mathcal{P}_{t+1,T}(g), x_t + \theta_t \Delta S_{t+1} \geq p_{t+1} \text{ a.s.}\}. \end{aligned}$$

The set $\mathcal{P}_{t,T}(g)$ contains all prices at time t super-replicating some price $p_{t+1} \in \mathcal{P}_{t+1,T}(g)$ at time $t+1$. First we show that for all $t \in \{0, \dots, T\}$

$$\Pi_{t,T}(g) = \mathcal{P}_{t,T}(g). \quad (2.3.12)$$

It is clear at time T . Let $x_t \in \Pi_{t,T}$. Then there exists for all $u \in \{t, \dots, T-1\}$, $\theta_u \in L^0(\mathbf{R}, \mathcal{F}_u)$ such that $x_t + \sum_{u=t+1}^{T-1} \theta_{u-1} \Delta S_u + \theta_{T-1} \Delta S_T \geq g$ a.s. As $g \in \mathcal{P}_{T,T}(g)$, $x_t + \sum_{u=t+1}^{T-1} \theta_{u-1} \Delta S_u \in \mathcal{P}_{T-1,T}(g)$ a.s. As $x_t + \sum_{u=t+1}^{T-2} \theta_{u-1} \Delta S_u + \theta_{T-2} \Delta S_{T-1} =$

$x_t + \sum_{u=t+1}^{T-1} \theta_{u-1} \Delta S_u$, it follows that $x_t + \sum_{u=t+1}^{T-2} \theta_{u-1} \Delta S_u \in \mathcal{P}_{T-2,T}(g)$ and recursively $x_t \in \mathcal{P}_{t,T}$. Conversely, let $x_t \in \mathcal{P}_{t,T}$, then there exists $\theta_t \in L^0(\mathbf{R}, \mathcal{F}_t)$ and $p_{t+1} \in \mathcal{P}_{t+1,T}(g)$, such that $x_t + \theta_t \Delta S_{t+1} \geq p_{t+1}$ a.s. Then as $p_{t+1} \in \mathcal{P}_{t+1,T}(g)$, there exists $\theta_{t+1} \in L^0(\mathbf{R}, \mathcal{F}_{t+1})$ and $p_{t+2} \in \mathcal{P}_{t+2,T}(g)$, such that $p_{t+1} + \theta_{t+1} \Delta S_{t+2} \geq p_{t+2}$ a.s. and going forward until T since $\mathcal{P}_{T,T}(g) = \{g\}$, $p_{T-1} + \theta_{T-1} \Delta S_T \geq g$ a.s., we get that $x_t + \sum_{u=t+1}^T \theta_{u-1} \Delta S_u \geq g$ a.s. and $x_t \in \Pi_{t,T}$ follows.

We now define a local version of super-hedging prices. Let $g_{t+1} \in L^0(\mathbf{R}, \mathcal{F}_{t+1})$, then the set of one-step super-hedging prices of g_{t+1} and its associated infimum super-hedging cost are given by

$$\begin{aligned}
 \mathcal{P}_{t,t+1}(g_{t+1}) &= \{x_t \in L^0(\mathbf{R}, \mathcal{F}_t), \exists \theta_t \in L^0(\mathbf{R}, \mathcal{F}_t), x_t + \theta_t \Delta S_{t+1} \geq g_{t+1} \text{ a.s.}\} \\
 \pi_{t,t+1}(g_{t+1}) &= \text{ess inf}_{\mathcal{F}_t} \mathcal{P}_{t,t+1}(g_{t+1}).
 \end{aligned}$$

The following lemma makes the link between local and global super-hedging under the assumption that the infimum (global) super-replication cost is a price. It provides a dynamic programming principle.

Lemma 2.3.1. *Let $g \in L^0(\mathbf{R}, \mathcal{F}_T)$ and $t \in \{0, \dots, T-1\}$. Then $\mathcal{P}_{t,T}(g) \subset \mathcal{P}_{t,t+1}(\pi_{t+1,T}(g))$ and $\pi_{t,T}(g) \geq \pi_{t,t+1}(\pi_{t+1,T}(g))$. Moreover if $\pi_{t+1,T}(g) \in \Pi_{t+1,T}(g)$, then $\mathcal{P}_{t,T}(g) = \mathcal{P}_{t,t+1}(\pi_{t+1,T}(g))$ and $\pi_{t,T}(g) = \pi_{t,t+1}(\pi_{t+1,T}(g))$.*

Remark 2.3.2. So under (AIP), if at each step, $\pi_{t+1,T}(g) \in \Pi_{t+1,T}(g)$ and if $\pi_{t+1,T}(g) = g_{t+1}(S_{t+1})$ for some \mathcal{F}_t -normal integrand g_{t+1} , we will get from Corollary 2.2.6 that $\pi_{t,T}(g) = \overline{\text{conc}}(g_{t+1}, \text{supp}_{\mathcal{F}_t} S_{t+1})(S_t)$. We will propose in Section 2.3.3 a quite general setting where this holds true.

Proof. Let $x_t \in \mathcal{P}_{t,T}(g)$, then there exists $\theta_t \in L^0(\mathbf{R}, \mathcal{F}_t)$ and $p_{t+1} \in \mathcal{P}_{t+1,T}(g)$ such that (recall (2.3.12))

$$x_t + \theta_t \Delta S_{t+1} \geq p_{t+1} \geq \text{ess inf}_{\mathcal{F}_t} \Pi_{t+1,T}(g) = \pi_{t+1,T}(g) \text{ a.s.}$$

and the first statement follows. The second one follows from $\pi_{t+1,T}(g) \in \mathcal{P}_{t+1,T}(g)$. \square

2.3.2 Multi-period (AIP)

We now define the notion of global and local immediate profit at time t . The first one says that it is possible to super-replicate at a negative cost from time t the claim 0 paid at time T and the local one the claim 0 paid at time $t+1$. We will see that they are equivalent.

Definition 2.3.3. Fix some $t \in \{0, \dots, T\}$. A global immediate profit (IP) at time t is a non null element of $\mathcal{P}_{t,T}(0) \cap L^0(\mathbf{R}_-, \mathcal{F}_t)$.

A local immediate profit at time t is a non null element of $\mathcal{P}_{t,t+1}(0) \cap L^0(\mathbf{R}_-, \mathcal{F}_t)$.

We say that the (AIP) condition holds if there is no global IP at any instant t , i.e. if $\mathcal{P}_{t,T}(0) \cap L^0(\mathbf{R}_-, \mathcal{F}_t) = \{0\}$ for all $t \in \{0, \dots, T\}$.

Using Proposition 2.2.8, we get the equivalence between the absence of local IP at time t and the fact that $S_t \in \text{convsupp}_{\mathcal{F}_t} S_{t+1}$ a.s. So Theorem 2.3.4 below will show that there is an equivalence between the absence of global IP and the absence of local one.

Theorem 2.3.4. (AIP) holds if and only if one of the the following assertions holds :

- 1) $S_t \in \text{convsupp}_{\mathcal{F}_t} S_{t+1}$ a.s., for all $t \in \{0, \dots, T-1\}$.
- 2) $\text{ess inf}_{\mathcal{F}_t} S_{t+1} \leq S_t \leq \text{ess sup}_{\mathcal{F}_t} S_{t+1}$ a.s., for all $t \in \{0, \dots, T-1\}$.
- 3) $\text{ess inf}_{\mathcal{F}_t} S_u \leq S_t \leq \text{ess sup}_{\mathcal{F}_t} S_u$ a.s. for all $u \in \{t, \dots, T\}$.
- 4) $\pi_{t,T}(0) = 0$ a.s. for all $t \leq T-1$.

Proof. Let $A_T = \Omega$ and for all $t \in \{0, \dots, T-1\}$

$$A_t := \{\text{ess sup}_{\mathcal{F}_t} \Delta S_{t+1} \geq 0\} \cap \{\text{ess inf}_{\mathcal{F}_t} \Delta S_{t+1} \leq 0\}.$$

We show by induction that $0 \in \mathcal{P}_{t,T}(0)$ and that under (AIP) at time $t+1$

$$\pi_{t,T}(0) = 0 \text{ a.s.} \Leftrightarrow P(A_t) = 1 \Leftrightarrow (\text{AIP}) \text{ holds at time } t.$$

The third assertion follows from Lemma 2.6.5.

We proceed by backward recursion. At time T , $\mathcal{P}_{T,T}(0) = \{0\}$, thus (AIP) holds at T and $\pi_{T,T}(0) = 0$. Fix some $t \in \{0, \dots, T-1\}$, assume that the induction hypothesis holds true at $t+1$ and that (AIP) holds at time $t+1$. As $\pi_{t+1,T}(0) = 0 \in \mathcal{P}_{t+1,T}(0)$, we can apply Lemma 2.3.1 and $\mathcal{P}_{t,T}(0) = \mathcal{P}_{t,t+1}(0)$. So we can apply Lemma 2.2.3 and

$$\begin{aligned} \mathcal{P}_{t,T}(0) &= \mathcal{P}_{t,t+1}(0) = \left\{ \sup_{z \in \text{supp}_{\mathcal{F}_t} S_{t+1}} (-\theta z) + \theta S_t, \theta \in L^0(\mathbf{R}, \mathcal{F}_t) \right\} + L^0(\mathbf{R}_+, \mathcal{F}_t) \\ &= \left\{ -\theta (\text{ess sup}_{\mathcal{F}_t} \Delta S_{t+1} 1_{\theta < 0} + \text{ess inf}_{\mathcal{F}_t} \Delta S_{t+1} 1_{\theta \geq 0}), \theta \in L^0(\mathbf{R}, \mathcal{F}_t) \right\} + L^0(\mathbf{R}_+, \mathcal{F}_t). \end{aligned}$$

Note that $0 \in \mathcal{P}_{t,T}(0)$. Moreover, (AIP) holds at time t if and only if $P(A_t) = 1$ (this also a direct consequence of Proposition 2.2.8). We also obtain that $\pi_{t,T}(0) = \text{ess inf}_{\mathcal{H}} \mathcal{P}_{t,T}(0) = (0)1_{A_t} + (-\infty)1_{\Omega \setminus A_t}$ and equivalently (AIP) holds at time t if and only if $\pi_{t,T}(0) = 0$ a.s. In particular, under (AIP) at time t , the infimum super-hedging cost at time t is a price for 0 : $\pi_{t,T}(0) = 0 \in \mathcal{P}_{t,T}(0)$. \square

Remark 2.3.5. Fix some $t \leq T-1$. If $\text{ess sup}_{\mathcal{F}_{t-1}} \Delta S_t < 0$ on a non null measure set, then as in Remark 2.2.10 there is an arbitrage opportunity at time t .

2.3.3 Explicit pricing of a convex payoff under (AIP)

The aim of this section is to obtain some results in a particular model where $\text{ess inf}_{\mathcal{F}_{t-1}} S_t = k_{t-1,t}^d S_{t-1}$ and $\text{ess sup}_{\mathcal{F}_{t-1}} S_t = k_{t-1,t}^u S_{t-1}$ for every $t \in \{1, \dots, T\}$ where $(k_{t-1,t}^d)_{t \in \{1, \dots, T\}}$, $(k_{t-1,t}^u)_{t \in \{1, \dots, T\}}$ and S_0 are deterministic non-negative numbers. We obtain the same computational scheme (see (2.3.13)) as in [11] but it is obtained here assuming only (AIP) and not (NA).

Theorem 2.3.6. *Suppose that the model is defined by $\text{ess inf}_{\mathcal{F}_{t-1}} S_t = k_{t-1,t}^d S_{t-1}$ and $\text{ess sup}_{\mathcal{F}_{t-1}} S_t = k_{t-1,t}^u S_{t-1}$ where $(k_{t-1,t}^d)_{t \in \{1, \dots, T\}}$, $(k_{t-1,t}^u)_{t \in \{1, \dots, T\}}$ and S_0 are deterministic non-negative numbers.*

- *The (AIP) condition holds at every instant t if and only if the super-hedging prices of some European call option are non-negative or equivalently if $k_{t-1,t}^d \in [0, 1]$ and $k_{t-1,t}^u \in [1, +\infty]$ for all $t \in \{1, \dots, T\}$.*
- *Suppose that the (AIP) condition holds. If $h : \mathbf{R} \rightarrow \mathbf{R}$ is a convex function with $\text{Dom } h = \mathbf{R}$, $h(z) \geq 0$ for all $z \geq 0$ and $\lim_{z \rightarrow +\infty} \frac{h(z)}{z} \in [0, \infty)$, the infimum super-hedging cost of the European contingent claim $h(S_T)$ is a price given by $\pi_{t,T}(h) = h(t, S_t) \in \mathcal{P}_{t,T}(h)$ where*

$$\begin{aligned} h(T, x) &= h(x) \\ h(t-1, x) &= \lambda_{t-1,t} h(t, k_{t-1,t}^d x) + (1 - \lambda_{t-1,t}) h(t, k_{t-1,t}^u x), \end{aligned} \tag{2.3.13}$$

where $\lambda_{t-1,t} = \frac{k_{t-1,t}^u - 1}{k_{t-1,t}^u - k_{t-1,t}^d} \in [0, 1]$ and $1 - \lambda_{t-1,t} = \frac{1 - k_{t-1,t}^d}{k_{t-1,t}^u - k_{t-1,t}^d} \in [0, 1]$, with the following conventions. When $k_{t-1,t}^d = k_{t-1,t}^u = 1$ or $S_{t-1} = 0$, $\lambda_{t-1,t} = \frac{0}{0} = 0$ and $1 - \lambda_{t-1,t} = 1$ and when $k_{t-1,t}^d < k_{t-1,t}^u = \infty$,

$$\begin{aligned} \lambda_{t-1,t} &= \frac{\infty}{\infty} = 1 \\ (1 - \lambda_{t-1,t}) h(t, (+\infty)x) &= (1 - k_{t-1,t}^d) x \frac{h(t, (+\infty)x)}{(+\infty x)} \\ &= (1 - k_{t-1,t}^d) x \lim_{z \rightarrow +\infty} \frac{h(z)}{z}. \end{aligned} \tag{2.3.14}$$

Moreover, for every t , $\lim_{z \rightarrow +\infty} \frac{h(z)}{z} = \lim_{z \rightarrow +\infty} \frac{h(t,z)}{z}$ and $h(\cdot, x)$ is non-increasing for all $x \geq 0$.

In the proof, the strategy associated to the minimal price is given and, in section 2.5, this result is illustrated through a numerical experiment.

Proof. The conditions $k_{t-1,t}^d \in [0, 1]$ and $k_{t-1,t}^u \in [1, +\infty]$ are equivalent to the (AIP) conditions by Theorem 2.3.4. We denote $M = \frac{h(\infty)}{\infty}$ and $M_t = \lim_{z \rightarrow +\infty} \frac{h(t,z)}{z}$. We prove the second statement. Assume that (AIP) holds true. We establish the recursive

formulation $\pi_{t,T}(h) = h(t, S_t)$ given by (2.3.13), that $h(t, \cdot) \geq h(t+1, \cdot)$ and that $M_t = M_{t+1}$. The case $t = T$ is immediate. As $h : \mathbf{R} \rightarrow \mathbf{R}$ is a convex function with $\text{Dom } h = \mathbf{R}$, h is clearly a \mathcal{F}_{T-1} -normal integrand, we can apply Proposition 2.2.6 and its consequence for convex functions (see (2.2.7) and (2.2.8)) and we get that

$$\begin{aligned} \pi_{T-1,T}(h) &= h(k_{T-1,T}^d S_{T-1}) + \theta_{T-1}^* \left(S_{T-1} - k_{T-1,T}^d S_{T-1} \right), \\ \theta_{T-1}^* &= \frac{h(k_{T-1,T}^u S_{T-1}) - h(k_{T-1,T}^d S_{T-1})}{k_{T-1,T}^u S_{T-1} - k_{T-1,T}^d S_{T-1}}, \end{aligned} \quad (2.3.15)$$

where we use the conventions $\theta_{T-1}^* = \frac{0}{0} = 0$ if either $S_{T-1} = 0$ or $k_{T-1,T}^u = k_{T-1,T}^d = 1$ and $\theta_{T-1}^* = \frac{h(\infty)}{\infty} = M$ if $k_{T-1,T}^d < k_{T-1,T}^u = +\infty$. Moreover, using (2.2.9), we obtain that $\pi_{T-1,T}(h) + \theta_{T-1}^* \Delta S_T \geq h$ a.s. i.e. $\pi_{T-1,T}(h) \in \mathcal{P}(h)$. So, using Lemma 2.3.1, we get that $\mathcal{P}_{T-2,T}(h) = \mathcal{P}_{T-2,T-1}(\pi_{T-1,T}(h))$ and $\pi_{T-2,T}(h) = \pi_{T-2,T-1}(\pi_{T-1,T}(h))$ and we may continue the recursion as soon as $\pi_{T-1,T}(h) = h(T-1, S_{T-1})$ where $h(T-1, \cdot)$ satisfies (2.3.13), is convex with domain equal to \mathbf{R} , is such that $h(T-1, z) \geq 0$ for all $z \geq 0$ and $M_{T-1} = M \in [0, \infty)$. To see that we distinguish three cases. If either $S_{T-1} = 0$ or $k_{T-1,T}^u = k_{T-1,T}^d = 1$, $\pi_{T-1,T}(h) = h(S_{T-1})$ and $h(T-1, z) = h(z) = h(T, z)$ satisfies all the required conditions. If $k_{T-1,T}^d < k_{T-1,T}^u = +\infty$, $\pi_{T-1,T}(h) = h(k_{T-1,T}^d S_{T-1}) + M \left(S_{T-1} - k_{T-1,T}^d S_{T-1} \right) = h(T-1, S_{T-1})$ with

$$\begin{aligned} h(T-1, z) &= h(k_{T-1,T}^d z) + Mz \left(1 - k_{T-1,T}^d \right) \\ &= \lim_{k^u \rightarrow +\infty} \left(\frac{k^u - 1}{k^u - k_{T-1,T}^d} h(k_{T-1,T}^d z) + \frac{1 - k_{T-1,T}^d}{k^u - k_{T-1,T}^d} h(k^u z) \right), \end{aligned}$$

using (2.3.14). The term in the r.h.s. above is larger than $h(z) = h(T, z)$ by convexity since $\frac{k^u - 1}{k^u - k_{T-1,T}^d} k_{T-1,T}^d z + \frac{1 - k_{T-1,T}^d}{k^u - k_{T-1,T}^d} k^u z = z$. As $k_{T-1,T}^d \in [0, 1]$ and $M \in [0, \infty)$, $h(T-1, z) \geq 0$ for all $z \geq 0$, we get that $h(T-1, \cdot)$ is convex function with domain equal to \mathbf{R} since h is so. The function $h(T-1, \cdot)$ also satisfies (2.3.13) (see (2.3.14)). Finally $M_{T-1} = \lim_{z \rightarrow +\infty} k_{T-1,T}^d \frac{h(k_{T-1,T}^d z)}{k_{T-1,T}^d z} + M \left(1 - k_{T-1,T}^d \right) = M$.

The last case is when $S_{T-1} \neq 0$ and $k_{T-1,T}^u \neq k_{T-1,T}^d$ and $k_{T-1,T}^u < +\infty$. It is clear that (2.3.15) implies (2.3.13). Moreover as $k_{T-1,T}^d \in [0, 1]$ and $k_{T-1,T}^u \in [1, +\infty)$, $\lambda_{T-1,T} = \frac{k_{T-1,T}^u - 1}{k_{T-1,T}^u - k_{T-1,T}^d} \in [0, 1]$ and $1 - \lambda_{T-1,T} = \frac{1 - k_{T-1,T}^d}{k_{T-1,T}^u - k_{T-1,T}^d} \in [0, 1]$ and (2.3.13) implies that $h(T-1, z) \geq 0$ for all $z \geq 0$, $h(T-1, \cdot)$ is convex with domain equal to \mathbf{R} since h is so. Moreover,

$$M_{T-1} = \lambda_{T-1,T} k_{T-1,T}^d \lim_{z \rightarrow +\infty} \frac{h(k_{T-1,T}^d z)}{k_{T-1,T}^d z} + (1 - \lambda_{T-1,T}) k_{T-1,T}^u \lim_{z \rightarrow +\infty} \frac{h(k_{T-1,T}^u z)}{k_{T-1,T}^u z} = M,$$

since

$$\lambda_{T-1,T} k_{T-1,T}^d + (1 - \lambda_{T-1,T}) k_{T-1,T}^u = 1.$$

If $h(x) = (x - K)^+$, for some $K \in \mathbf{R}$, $h(S_T)$ is a European contingent claim and $h : \mathbf{R} \rightarrow \mathbf{R}$ a convex function with $\text{Dom } h = \mathbf{R}$, $h \geq 0$ and $\lim_{z \rightarrow +\infty} \frac{h(z)}{z} = 1 \in [0, \infty)$. We have just seen that under (AIP) the infimum super-hedging cost of $h(S_T)$ is a price $\pi_{t,T}(h) \geq 0$. Reversely if (AIP) does not hold true, Proposition 2.2.6 implies

$$\begin{aligned} \pi_{T-1,T}(h) &= \inf \{ \alpha S_{T-1} + \beta, \alpha, \beta \in \mathbf{R}, (z - K)^+ \leq \alpha z + \beta, \forall z \in \text{supp}_{\mathcal{F}_{T-1,T}} S_T \} \\ &\quad - \delta_{[k_{T-1,T}^d S_{T-1}, k_{T-1,T}^u S_{T-1}] \cap \mathbf{R}}(S_{T-1}). \end{aligned}$$

As (AIP) does not hold true, either $k_{T-1,T}^d > 1$ or $k_{T-1,T}^u < 1$ and in both cases, $S_{T-1} \notin [k_{T-1,T}^d S_{T-1}, k_{T-1,T}^u S_{T-1}] \cap \mathbf{R}$ and $\pi_{T-1,T}(h) = -\infty$ since

$$\inf \{ \alpha S_{T-1} + \beta, \alpha, \beta \in \mathbf{R}, (z - K)^+ \leq \alpha z + \beta, \forall z \in \text{supp}_{\mathcal{F}_{T-1,T}} S_T \} \leq S_{T-1}$$

Thus the convex subset $\mathcal{P}_{T-1,T}(h)$ is equal to $L^0(\mathbf{R}, \mathcal{F}_{T-1})$. Similarly $\pi_{t,T}(h) = -\infty$ for all $t \in \{0, \dots, T-3\}$. This allows to conclude about the first statement. \square

Remark 2.3.7. *The infimum price of the European contingent claim $h(S_T)$ in our model is a price, precisely the same than the price we get in a binomial model $S_t \in \{k_{t-1,t}^d S_{t-1}, k_{t-1,t}^u S_{t-1}\}$ a.s., $t = 1, \dots, T$.*

2.4 Comparison between the (AIP) condition and classical no-arbitrage conditions

Examples have already show that (AIP) condition can be weaker than the classical absence of arbitrage opportunity (NA) characterized by the fundamental theorem of asset pricing (FTAP), see the Dalang-Morton-Willinger theorem in [16]. The goal of this section is compare the (AIP) condition with a weaker form of the classical No Free Lunch condition.

Recall that the set of all prices for the zero claim at time t is given by $\mathcal{P}_{t,T}(0) = (-\mathcal{R}_t^T) \cap L^0(\mathbf{R}, \mathcal{F}_t)$ (see (2.3.10), (2.3.11) and (2.3.12)). It follows that (AIP) reads as $\mathcal{R}_t^T \cap L^0(\mathbf{R}_+, \mathcal{F}_t) = \{0\}$. Recall that the (NA) condition is $\mathcal{R}_t^T \cap L^0(\mathbf{R}_+, \mathcal{F}_T) = \{0\}$. We also study a stronger condition than (AIP), i.e. $\overline{\mathcal{R}_t^T} \cap L^0(\mathbf{R}_+, \mathcal{F}_t) = \{0\}$ for all $t \in \{0, \dots, T\}$, where the closure of \mathcal{R}_t^T is taken with respect to the convergence in probability. Note that this condition is a weak form of the classical No Free Lunch condition $\overline{\mathcal{R}_t^T} \cap L^0(\mathbf{R}_+, \mathcal{F}_T) = \{0\}$ for all $t \in \{0, \dots, T\}$; we call it (WNFL) for Weak No Free Lunch. The following result implies that (WNFL) may be equivalent to (AIP) condition under an extra closedness condition. It also provides a characterization through (absolutely continuous) martingale measures.

Theorem 2.4.1. *The following statements are equivalent :*

- (WNFL) holds.
- For every $t \in \{0, \dots, T\}$, there exists $Q \ll P$ with $\mathbb{E}(dQ/dP|\mathcal{F}_t) = 1$ such that $(S_u)_{u \in \{t, \dots, T\}}$ is a Q -martingale.
- (AIP) holds and $\overline{\mathcal{R}_t^T} \cap L^0(\mathbf{R}, \mathcal{F}_t) = \mathcal{R}_t^T \cap L^0(\mathbf{R}, \mathcal{F}_t)$ for every $t \in \{0, \dots, T\}$.

Proof. Suppose that (WNFL) holds and fix some $t \in \{0, \dots, T\}$. We may suppose without loss of generality that the process S is integrable under P . Under (WNFL), we then have $\overline{\mathcal{R}_t^T} \cap L^1(\mathbf{R}_+, \mathcal{F}_t) = \{0\}$ where the closure is taken in L^1 . Therefore, for every nonzero $x \in L^1(\mathbf{R}_+, \mathcal{F}_t)$, there exists by the Hahn-Banach theorem a non-zero $Z_x \in L^\infty(\mathbf{R}_+, \mathcal{F}_T)$ such that (recall that \mathcal{R}_t^T is a cone) $\mathbb{E}Z_x x > 0$ and $\mathbb{E}Z_x \xi \leq 0$ for every $\xi \in \mathcal{R}_t^T$. Since $-L^1(\mathbf{R}_+, \mathcal{F}_T) \subseteq \mathcal{R}_t^T$, we deduce that $Z_x \geq 0$ and we may renormalise Z_x so that $\|Z_x\|_\infty = 1$. Let us consider the family $\mathcal{G} = \{\mathbb{E}(Z_x|\mathcal{F}_t) > 0\}$, $x \in L^1(\mathbf{R}_+, \mathcal{F}_t) \setminus \{0\}$. Consider any non null set $\Gamma \in \mathcal{F}_t$. Taking $x = 1_\Gamma \in L^1(\mathbf{R}_+, \mathcal{F}_t) \setminus \{0\}$, since $\mathbb{E}(Z_x 1_\Gamma) > 0$, we deduce that Γ has a non null intersection with $\{\mathbb{E}(Z_x|\mathcal{F}_t) > 0\}$. By [43, Lemma 2.1.3], we deduce an at most countable subfamily $(x_i)_{i \geq 1}$ such that the union $\bigcup_i \{\mathbb{E}(Z_{x_i}|\mathcal{F}_t) > 0\}$ is of full measure. Therefore, $Z = \sum_{i=1}^\infty 2^{-i} Z_{x_i} \geq 0$ is such that $\mathbb{E}(Z|\mathcal{F}_t) > 0$ and we define $Q \ll P$ such that $dQ = (Z/\mathbb{E}(Z|\mathcal{F}_t))dP$. As the subset $\{\sum_{u=t+1}^T \theta_{u-1} \Delta S_u, \theta_{u-1} \in L(\mathbf{R}, \mathcal{F}_{u-1})\}$ is a linear vector space contained in \mathcal{R}_t^T , we deduce that $(S_u)_{u \in \{t, \dots, T\}}$ is a Q -martingale.

Suppose that for every $t \in \{0, \dots, T\}$, there exists $Q \ll P$ such that $(S_u)_{u \in \{t, \dots, T\}}$ is a Q -martingale with $\mathbb{E}(dQ/dP|\mathcal{F}_t) = 1$. Let us define for $u \in \{t, \dots, T\}$, $\rho_u = \mathbb{E}_P(dQ/dP|\mathcal{F}_u)$ then $\rho_u \geq 0$ and $\rho_t = 1$. Consider $\gamma_t \in \mathcal{R}_t^T \cap L^0(\mathbf{R}_+, \mathcal{F}_t)$, i.e. γ_t is \mathcal{F}_t -measurable and is of the form $\gamma_t = \sum_{u=t}^{T-1} \theta_u \Delta S_{u+1} - \epsilon_T^+$. Since θ_u is \mathcal{F}_u -measurable, $\theta_u \Delta S_{u+1}$ admits a generalized conditional expectation under Q knowing \mathcal{F}_u and, by assumption, we have $\mathbb{E}_Q(\theta_u \Delta S_{u+1}|\mathcal{F}_u) = 0$. We deduce by the tower law that

$$\gamma_t = \mathbb{E}_Q(\gamma_t|\mathcal{F}_t) = \sum_{u=t}^{T-1} \mathbb{E}_Q(\mathbb{E}_Q(\theta_u \Delta S_{u+1}|\mathcal{F}_u)|\mathcal{F}_t) - \mathbb{E}_Q(\epsilon_T^+|\mathcal{F}_t) = -\mathbb{E}_Q(\epsilon_T^+|\mathcal{F}_t).$$

Hence $\gamma_t = 0$, i.e. (AIP) holds. It remains to show that $\overline{\mathcal{R}_t^T} \cap L^0(\mathbf{R}, \mathcal{F}_t) \subseteq \mathcal{R}_t^T \cap L^0(\mathbf{R}, \mathcal{F}_t)$.

Consider first a one step model, where $(S_u)_{u \in \{T-1, T\}}$ is a Q -martingale with $\rho_T \geq 0$ and $\rho_{T-1} = 1$. Suppose that $\gamma^n = \theta_{T-1}^n \Delta S_T - \epsilon_T^{n+} \in L^0(\mathbf{R}, \mathcal{F}_{T-1})$ converges in probability to $\gamma^\infty \in L^0(\mathbf{R}, \mathcal{F}_{T-1})$. We need to show that $\gamma^\infty \in \mathcal{R}_{T-1}^T$. On the \mathcal{F}_{T-1} -measurable set $\Lambda_{T-1} := \{\liminf_n |\theta_{T-1}^n| < \infty\}$, by [43, Lemma 2.1.2], we may assume w.l.o.g. that θ_{T-1}^n is convergent to some θ_{T-1}^∞ hence ϵ_T^{n+} is also convergent and we can conclude. Otherwise, on $\Omega \setminus \Lambda_{T-1}$, we use the normalized sequences $\tilde{\theta}_{T-1}^n := \theta_{T-1}^n / (|\theta_{T-1}^n| + 1)$, $\tilde{\epsilon}_T^{n+} := \epsilon_T^{n+} / (|\theta_{T-1}^n| + 1)$. By [43, Lemma 2.1.2], we may assume that $\tilde{\theta}_{T-1}^n \rightarrow \tilde{\theta}_{T-1}^\infty$, $\tilde{\epsilon}_T^{n+} \rightarrow \tilde{\epsilon}_T^{\infty+}$ and $\tilde{\theta}_{T-1}^\infty \Delta S_T - \tilde{\epsilon}_T^{\infty+} = 0$. As $|\tilde{\theta}_{T-1}^\infty| = 1$ a.s., first consider the subset $\Lambda_{T-1}^2 :=$

$\{\tilde{\theta}_{T-1}^\infty = 1\} \in \mathcal{F}_{T-1}$. We then have $\Delta S_T \geq 0$ on Λ_{T-1}^2 . Since $\mathbb{E}_Q(\Delta S_T 1_{\Lambda_{T-1}^2} | \mathcal{F}_{T-1}) = 0$, we get that $\rho_T \Delta S_T 1_{\Lambda_{T-1}^2} = 0$ a.s. Hence $\rho_T \gamma^n 1_{\Lambda_{T-1}^2} = -\rho_T \epsilon_T^{n+} 1_{\Lambda_{T-1}^2} \leq 0$. Taking the limit, we get that $\rho_T \gamma^\infty 1_{\Lambda_{T-1}^2} \leq 0$ and, since $\gamma^\infty \in L^0(\mathbf{R}, \mathcal{F}_{T-1})$, we deduce that $\rho_{T-1} \gamma^\infty 1_{\Lambda_{T-1}^2} \leq 0$. Recall that $\rho_{T-1} = 1$ hence $\gamma^\infty 1_{\Lambda_{T-1}^2} \leq 0$ and $\gamma^\infty 1_{\Lambda_{T-1}^2} \in \mathcal{R}_{T-1}^T$. On the subset $\{\tilde{\theta}_{T-1}^\infty = -1\}$ we may argue similarly and the conclusion follows in the one step model.

Fix some $s \in \{t, \dots, T-1\}$. We show that $\overline{\mathcal{R}}_{s+1}^T \cap L^0(\mathbf{R}, \mathcal{F}_{s+1}) \subseteq \mathcal{R}_{s+1}^T \cap L^0(\mathbf{R}, \mathcal{F}_{s+1})$ implies the same property for s instead of $s+1$. By assumption $(S_u)_{u \in \{s, \dots, T\}}$ is a Q -martingale with $\rho_u \geq 0$ for $u \in \{s, \dots, T\}$ and $\rho_s = 1$. Suppose that $\gamma^n = \sum_{u=s}^{T-1} \theta_u^n \Delta S_{u+1} - \epsilon_T^{n+} \in L^0(\mathbf{R}, \mathcal{F}_s)$ converges to $\gamma^\infty \in L^0(\mathbf{R}, \mathcal{F}_s)$. If $\gamma^\infty = 0$ there is nothing to prove. On the \mathcal{F}_s -measurable set $\Lambda_s := \{\liminf_n |\theta_s^n| < \infty\}$, by [43, Lemma 2.1.2], we may assume w.l.o.g. that θ_s^n converges to θ_s^∞ . Therefore, by the induction hypothesis, $\sum_{u=s+1}^{T-1} \theta_u^n \Delta S_{u+1} - \epsilon_T^{n+}$ is also convergent to an element of $\mathcal{R}_{s+1}^T \cap L^0(\mathbf{R}, \mathcal{F}_{s+1})$ and we conclude that $\gamma^\infty \in \mathcal{R}_s^T$. On $\Omega \setminus \Lambda_{s-1}$, we use the normalisation procedure, and deduce the equality $\sum_{u=s}^{T-1} \tilde{\theta}_u^\infty \Delta S_{u+1} - \tilde{\epsilon}_T^{\infty+} = 0$ for some $\tilde{\theta}_u^\infty \in L^0(\mathbf{R}, \mathcal{F}_u)$, $u \in \{s, \dots, T-1\}$ and $\tilde{\epsilon}_T^{\infty+} \geq 0$ such that $|\tilde{\theta}_s^\infty| = 1$ a.s. We then argue as in the one step model on $\Lambda_s^2 := \{\tilde{\theta}_s^\infty = 1\} \in \mathcal{F}_s$ and $\Lambda_s^3 := \{\tilde{\theta}_s^\infty = -1\} \in \mathcal{F}_s$ respectively. When $\tilde{\theta}_s^\infty = 1$, we deduce that $\Delta S_{s+1} + \sum_{u=s+1}^{T-1} \tilde{\theta}_u^\infty \Delta S_{u+1} - \tilde{\epsilon}_T^{\infty+} = 0$, i.e. $\Delta S_{s+1} \in \mathcal{P}_{s+1, T}(0)$ hence $\Delta S_{s+1} \geq 0$ under (AIP), see Theorem 2.3.4. Since $\mathbb{E}_Q(\Delta S_{s+1} 1_{\Lambda_s^2} | \mathcal{F}_s) = 0$, $\rho_{s+1} \Delta S_{s+1} 1_{\Lambda_s^2} = 0$ a.s. So, $\rho_{s+1} \gamma^n 1_{\Lambda_s^2} \in \mathcal{R}_{s+1}^T \cap L^0(\mathbf{R}, \mathcal{F}_{s+1})$ hence $\rho_{s+1} \gamma^\infty 1_{\Lambda_s^2} \in \mathcal{R}_{s+1}^T \cap L^0(\mathbf{R}, \mathcal{F}_{s+1})$ by induction. As $\rho_{s+1} \gamma^\infty 1_{\Lambda_s^2}$ admits a generalized conditional expectation knowing \mathcal{F}_s , we deduce from (AIP) that $\mathbb{E}_Q(\rho_{s+1} \gamma^\infty 1_{\Lambda_s^2} | \mathcal{F}_s) \leq 0$ hence $\rho_s \gamma^\infty 1_{\Lambda_s^2} \leq 0$. Recall that $\rho_s = 1$ hence $\gamma^\infty 1_{\Lambda_s^2} \leq 0$ so that $\gamma^\infty 1_{\Lambda_s^2} \in \mathcal{R}_s^T \cap L^0(\mathbf{R}, \mathcal{F}_s)$.

Finally, notice that the (AIP) condition implies (WNFL) as soon as the equality $\overline{\mathcal{R}}_t^T \cap L^0(\mathbf{R}_+, \mathcal{F}_t) = \mathcal{R}_t^T \cap L^0(\mathbf{R}_+, \mathcal{F}_t)$ holds for every $t \in \{0, \dots, T-1\}$. \square

Proposition 2.4.2. *Suppose that $P(\text{ess inf}_{\mathcal{F}_t} S_{t+1} = S_t) = P(\text{ess sup}_{\mathcal{F}_t} S_{t+1} = S_t) = 0$ for all $t \in \{0, \dots, T-1\}$. Then, (WNFL) is equivalent to (AIP) and, under these equivalent conditions, \mathcal{R}_t^T is closed in probability for every $t \in \{0, \dots, T-1\}$.*

Proof. It suffices to show that \mathcal{R}_t^T is closed in probability for every $t \in \{0, \dots, T-1\}$ under (AIP). Consider first the one step model, i.e. suppose that $\gamma^n = \theta_{T-1}^n \Delta S_T - \epsilon_T^{n+} \in \mathcal{R}_{T-1}^T$ is a convergent sequence to $\gamma^\infty \in L^0(\mathbf{R}, \mathcal{F}_T)$. It is then sufficient to show that the \mathcal{F}_{T-1} -measurable set $\Lambda_{T-1} := \{\liminf_n |\theta_{T-1}^n| < \infty\}$ satisfies $P(\Lambda_{T-1}) = 1$. Following the normalization procedure of proof of Theorem 2.4.1 on $\Omega \setminus \Lambda_{T-1}$, we get that $\tilde{\theta}_{T-1}^\infty \Delta S_T$ where $|\tilde{\theta}_{T-1}^\infty| = 1$ a.s. First consider the subset $\Lambda_{T-1}^2 := \{\tilde{\theta}_{T-1}^\infty = 1\} \in \mathcal{F}_{T-1}$. We have $\Delta S_T \geq 0$ and hence $\text{ess inf}_{\mathcal{F}_{T-1}} S_T \geq S_{T-1}$ on Λ_{T-1}^2 . By (AIP) (see Theorem 2.3.4), we

deduce that $\text{ess inf}_{\mathcal{F}_{T-1}} S_T = S_{T-1}$ on Λ_{T-1}^2 . The assumption implies that $P(\Lambda_{T-1}^2) = 0$. On the remaining subset $\Lambda_{T-1}^3 := \{\tilde{\theta}_{T-1}^\infty = -1\} \in \mathcal{F}_{T-1}$, we obtain similarly that $\text{ess sup}_{\mathcal{F}_{T-1}} S_T = S_{T-1}$ and thus that $P(\Lambda_{T-1}^3) = 0$.

By induction, assume that \mathcal{R}_{t+1}^T is closed in probability and let us show that \mathcal{R}_t^T is also closed in probability. To do so, suppose that $\gamma^n = \sum_{u=t+1}^T \theta_{u-1}^n \Delta S_u - \epsilon_T^{n+} \in \mathcal{R}_t^T$ converges to $\gamma^\infty \in L^0(\mathbf{R}, \mathcal{F}_T)$. On the \mathcal{F}_t -measurable set $\Lambda_t := \{\liminf_n |\theta_t^n| < \infty\}$, by [43, Lemma 2.1.2], we may assume w.l.o.g. that θ_t^n is convergent to θ_t^∞ . Therefore, by the induction hypothesis,

$\sum_{u=t+2}^T \theta_{u-1}^n \Delta S_u - \epsilon_T^{n+}$ is also convergent to an element of \mathcal{R}_{t+1}^T and we conclude that $\gamma^\infty \in \mathcal{R}_t^T$. On $\Omega \setminus \Lambda_{t-1}$, we use the normalization procedure, and deduce an equality $\sum_{u=t+1}^T \tilde{\theta}_{u-1}^\infty \Delta S_u - \tilde{\epsilon}_T^{\infty+} = 0$ where $\tilde{\theta}_{u-1}^\infty \in L(\mathbf{R}, \mathcal{F}_{u-1})$, $u \in \{t, \dots, T-1\}$ and $\tilde{\epsilon}_T^{\infty+} \geq 0$ such that $|\tilde{\theta}_t^\infty| = 1$ a.s. We then argue on $\Lambda_t^2 := \{\tilde{\theta}_t^\infty = 1\} \in \mathcal{F}_t$ and $\Lambda_t^3 := \{\tilde{\theta}_t^\infty = -1\} \in \mathcal{F}_t$ respectively. On Λ_t^2 , we obtain that $\Delta S_{t+1} \in \mathcal{P}_{t+1, T}(0)$ hence under (AIP), with Theorem 2.3.4, we obtain that $\Delta S_{t+1} \geq 0$ and $\text{ess inf}_{\mathcal{F}_t} S_{t+1} = S_t$ on Λ_t^2 . This implies that $P(\Lambda_t^2) = 0$ and similarly $P(\Lambda_t^3) = 0$. The conclusion follows. \square

Remark 2.4.3. *Under the assumption of Proposition 2.4.2, the infimum super-hedging cost is a price.*

Lemma 2.4.4. *The (AIP) condition is not necessarily equivalent to (WNFL).*

Proof. Let us consider a positive process $(\tilde{S}_t)_{t \in \{0, \dots, T\}}$ which is a P -martingale. We suppose that $\text{ess inf}_{\mathcal{F}_0} \tilde{S}_1 < \tilde{S}_1$ a.s., which holds in particular if \tilde{S} a geometric Brownian motion as $\text{ess inf}_{\mathcal{F}_0} \tilde{S}_1 = 0$ a.s. Let us define $S_t := \tilde{S}_t$ for $t \in \{1, \dots, T\}$ and $S_0 := \text{ess inf}_{\mathcal{F}_0} S_1$. We have $\text{ess inf}_{\mathcal{F}_0} S_1 \leq S_0$ and $\text{ess sup}_{\mathcal{F}_0} S_1 \geq S_1 \geq \text{ess inf}_{\mathcal{F}_0} S_1 = S_0$ hence (AIP) holds at time 0 (see Theorem 2.3.4). Moreover, by the martingale property, (AIP) also holds at any time $t \in \{1, \dots, T\}$ (see Remark 2.6.4). Let us suppose that (WNFL) holds. Then, there exists $\rho_T \geq 0$ with $\mathbb{E}(\rho_T) = 1$ such that S is a Q -martingale where $dQ = \rho_T dP$. Therefore, $\mathbb{E}(\rho_T \Delta S_1) = 0$. Since $\Delta S_1 > 0$ by assumption, we deduce that $\rho_T = 0$ hence a contradiction. \square

2.5 Numerical experiments

2.5.1 Calibration

In this section, we suppose that the discrete dates are given by $t_i^n = \frac{iT}{n}$, $i \in \{0, \dots, n\}$ where $n \geq 1$. We assume that $k_{t_{i-1}^n}^u = 1 + \sigma_{t_{i-1}^n} \sqrt{\Delta t_i^n}$ and $k_{t_{i-1}^n}^d = 1 - \sigma_{t_{i-1}^n} \sqrt{\Delta t_i^n} \geq 0$ where $t \mapsto \sigma_t$ is a positive Lipschitz-continuous function on $[0, T]$. This model implies that

$\text{ess inf}_{\mathcal{F}_{t_{j-1}^n}} S_{t_i^n} = k_{t_{j-1}^n, t_i^n}^d S_{t_{j-1}^n}$ and $\text{ess sup}_{\mathcal{F}_{t_{j-1}^n}} S_{t_i^n} = k_{t_{j-1}^n, t_i^n}^u S_{t_{j-1}^n}$, where for all $j \leq i$,

$$k_{t_{j-1}^n, t_i^n}^u = \prod_{r=j}^i k_{t_{r-1}^n, t_r^n}^u, \quad k_{t_{j-1}^n, t_i^n}^d = \prod_{r=j}^i k_{t_{r-1}^n, t_r^n}^d.$$

By Theorem 2.3.6, we deduce that the (minimal) price of the European Call option $(S_T - K)^+$ is given by $h^n(t, S_t)$ defined by (2.3.13) with terminal condition $h^n(T, x) = h(x) := (x - K)^+$. We extend the function h^n on $[0, T]$ in such a way that h^n is constant on each interval $[t_i^n, t_{i+1}^n[$, $i \in \{0, \dots, n\}$. Such a scheme is proposed by Milstein [55] where a convergence theorem is proven when the terminal condition, i.e. the payoff function, is smooth. Precisely, the sequence of functions h^n converges uniformly to $h(t, x)$, solution to the diffusion equation :

$$\partial_t h(t, x) + \sigma_t^2 \frac{x^2}{2} \partial_{xx} h(t, x) = 0, \quad h(T, x) = h(x).$$

In [55], it is supposed that the successive derivatives of the P.D.E.'s solution h are uniformly bounded. This is not the case for the Call payoff function $g(x) = (x - K)^+$. On the contrary the successive derivatives of the P.D.E.'s solution explode at the horizon date, see [49]. In [2], it is proven that the uniform convergence still holds when the payoff function is not smooth provided that the successive derivatives of the P.D.E. solution do not explode *too much*.

Supposing that Δt_i^n is closed to 0, we identify the observed prices of the call option with the limit theoretical prices $h(t, S_t)$ at any instant t to deduce an evaluation of the deterministic function $t \mapsto \sigma_t$. Note that the assumptions on the multipliers $k_{t_{i-1}^n}^u$ and $k_{t_{i-1}^n}^d$ mean that

$$\left| \frac{S_{t_{i+1}^n}}{S_{t_i^n}} - 1 \right| \leq \sigma_{t_i} \sqrt{\Delta t_i^n}, \text{ a.s.} \quad (2.5.16)$$

We propose to verify (2.5.16) on real data. The data set is composed of historical values of the french index CAC 40 and European call option prices of maturity 3 months from the 23rd of October 2017 to the 19th of January 2018. For several strikes, matching the observed prices to the theoretical ones derived from the Black an Scholes formula with time-dependent volatility, we deduce the associated *implied* volatility $t \mapsto \sigma_t$ and we compute the proportion of observations satisfying (2.5.16) :

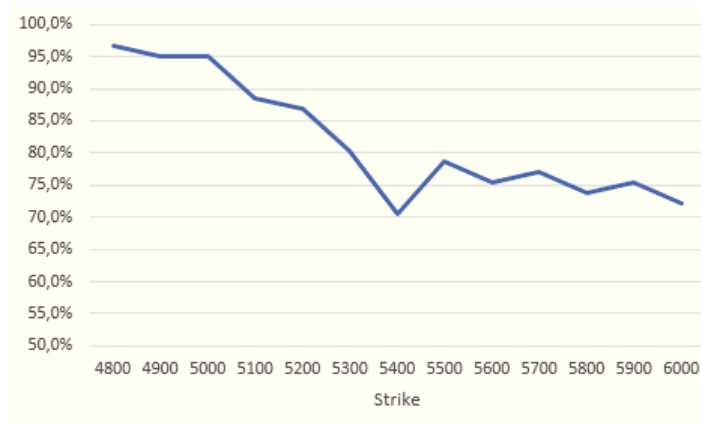


FIGURE 2.1 – Ratio of observations satisfying (2.5.16) as a function of the strike.

Strike	4800	4900	5000	5100	5200	5300	5400	5500	5600	5700	5800	5900	6000
Ratio	96,7%	95,1%	95,1%	88,5%	86,9%	80,3%	70,5%	78,7%	75,4%	77,0%	73,8%	75,4%	72,1%

2.5.2 Super-hedging

We test the infimum super-hedging cost deduced for Theorem 2.3.6 on some data set composed of historical daily closing values of the french index CAC 40 from the 5th of January 2015 to the 12th of March 2018. The interval $[0, T]$ we choose corresponds to one week composed of 5 days so that the number of discrete dates is $n = 5$. We first evaluate $\sigma_{t_i}^2$, $i = 0, \dots, 3$, as

$$\sigma_{t_i} = \overline{\max} \left(\left| \frac{S_{t_{i+1}}}{S_{t_i}} - 1 \right| / \sqrt{\Delta t_{i+1}^n} \right) \quad i = 0, \dots, 3,$$

where $\overline{\max}$ is the empirical maximum taken over a one year sliding sample window of 52 weeks. We then implement the super-hedging strategy on each of the 112 weeks following the sliding samples, i.e. every week from the 11th of January 2016 to the 5th of March 2018. We observe the empirical average $\mathbb{E}(S_{t_0}) = 4044$. The payoff function is $h(x) = (x - K)^+$.

Case where $K = 4700$.

We implement the strategy associated to the super-hedging cost given by Theorem 2.3.6. We deduce the distribution of the super-hedging error $\varepsilon_T := V_T - (S_T - K)^+$, see Figure 2.4 :

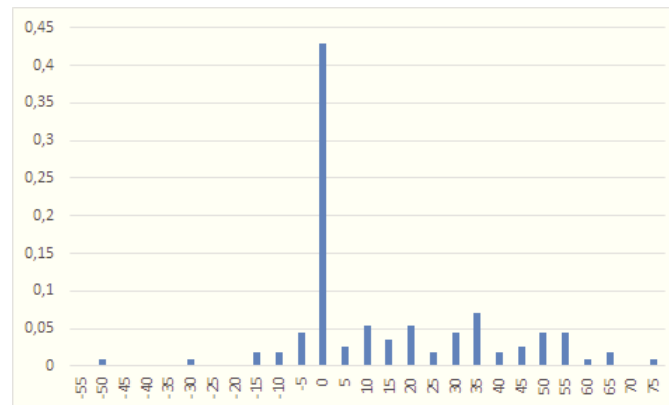


FIGURE 2.2 – Distribution of the super-hedging error $\varepsilon_T = V_T - (S_T - K)^+$.

The empirical average of the error ε_T is 12.63 and its standard deviation is 21.65. This result is rather satisfactory in comparison to the large value $\mathbb{E}(S_{t_0}) = 4044$. This empirically confirms the efficiency of our suggested method.

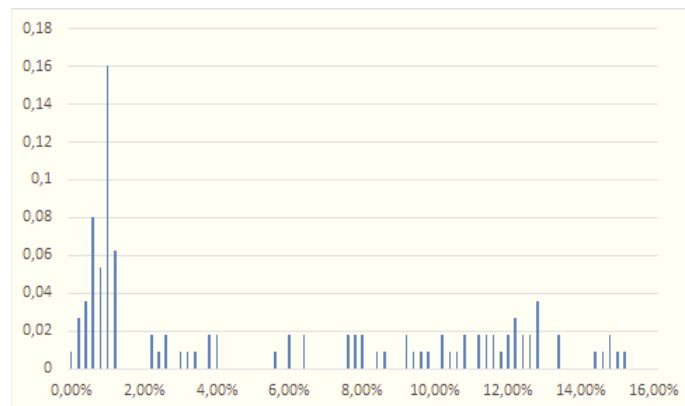


FIGURE 2.3 – Distribution of the ratio V_{t_0}/S_{t_0} .

The empirical average of V_{t_0}/S_{t_0} is 5.63% and its standard deviation is 5.14%. Notice that, in the discrete case with $k^d = 0$ and $k^u = \infty$, in particular when the dynamics of S is modeled by a (discrete) geometric Brownian motion, then the theoretical minimal initial price is $V_{t_0} = S_{t_0}$.

Case where $K = S_{t_0}$.

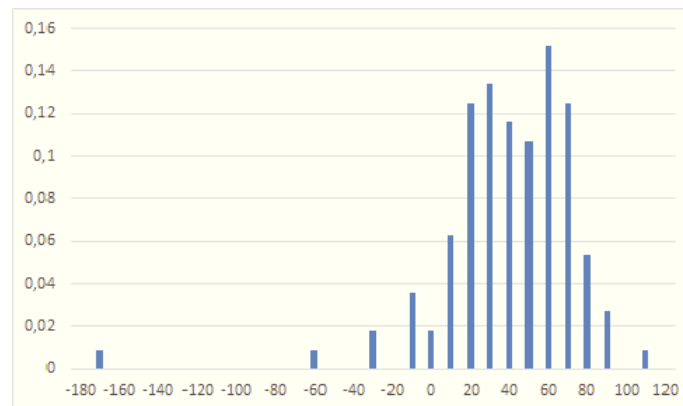


FIGURE 2.4 – Distribution of the super-hedging error ε_T .

The empirical average of $\varepsilon_T = V_T - (S_T - K)^+$ is 8.1 and its standard deviation is 30.78. Once again, this is rather satisfactory despite the possible loss of 170 which represents 4.2% of $\mathbb{E}(S_{t_0}) = 4044$.

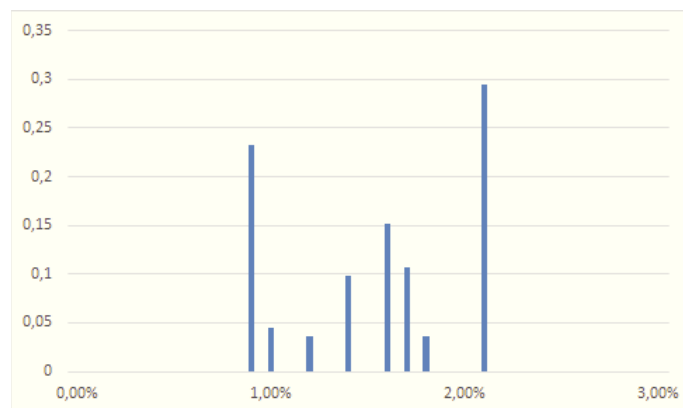


FIGURE 2.5 – Distribution of the ratio V_{t_0}/S_{t_0} .

The empirical average of V_0/S_0 is 2.51% and its standard deviation is 0.53%.

2.6 Appendix

2.6.1 Conditional support of a vector-valued random variable

We consider a random variable X defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in \mathbf{R}^d , $d \geq 1$, endowed with the Borel σ -algebra. The goal of this section is to define the conditional support of X with respect to a sub σ -algebra $\mathcal{H} \subseteq \mathcal{F}$. This notion

is very well known in the case where \mathcal{H} is the trivial sigma-algebra. Precisely, this is the usual support of X , i.e. the intersection of all closed deterministic subsets F of \mathbf{R}^d such that $P(X \in F) = 1$.

Definition 2.6.1. *Let (Ω, \mathcal{F}, P) be a probability space and \mathcal{H} be a sub- σ -algebra of \mathcal{F} . Let μ be a \mathcal{H} -stochastic kernel (i.e. for all $\omega \in \Omega$, $\mu(\cdot, \omega)$ is a probability on $\mathcal{B}(\mathbf{R}^d)$ and $\mu(A, \cdot)$ is \mathcal{H} -measurable, for all $A \in \mathcal{B}(\mathbf{R}^d)$). We define the random set $D_\mu : \Omega \rightarrow \mathbf{R}^d$:*

$$D_\mu(\omega) := \bigcap \left\{ A \subset \mathbf{R}^d, \text{ closed, } \mu(A, \omega) = 1 \right\}. \quad (2.6.17)$$

For $\omega \in \Omega$, $D_\mu(\omega) \subset \mathbf{R}^d$ is called the support of $\mu(\cdot, \omega)$. We will also call $\text{supp}_{\mathcal{H}} X$ the set defined in (2.6.17) when $\mu(A, \omega) = P(X \in A | \mathcal{H})(\omega)$ is the regular version of the conditional law of X knowing \mathcal{H} and call it the conditional support of X with respect to \mathcal{H} .

Using Theorems 12.7 and 12.14 of [1], we have that $\mu(\cdot, \omega)$ admits a unique support $D_\mu(\omega) \subset \mathbf{R}^d$ and that $\mu(D_\mu(\omega), \omega) = 1$ (see also the definition of support in [1] on page 441).

Lemma 2.6.2. *D_μ is non-empty, closed-valued, \mathcal{H} -measurable and graph-measurable random set (i.e. $\text{Graph}(D_\mu) \in \mathcal{H} \otimes \mathcal{B}(\mathbf{R}^d)$).*

Proof. It is clear from the definition (2.6.17) that for all $\omega \in \Omega$, $D_\mu(\omega)$ is a non-empty and closed subset of \mathbf{R}^d . We now show that D_μ is \mathcal{H} -measurable. Let O be a fixed open set in \mathbf{R}^d and $\mu_O : \omega \in \Omega \rightarrow \mu_O(\omega) := \mu(O, \omega)$. As μ is a stochastic kernel, μ_O is \mathcal{H} -measurable. By definition of $D_\mu(\omega)$ we get that $\{\omega \in \Omega, D_\mu(\omega) \cap O \neq \emptyset\} = \{\omega \in \Omega, \mu_O(\omega) > 0\} \in \mathcal{H}$, and D_μ is \mathcal{H} -measurable. Now using Theorem 14.8 of [62], $\text{Graph}(D_\mu) \in \mathcal{H} \otimes \mathcal{B}(\mathbf{R}^d)$ (recall that D_μ is closed-valued) and D_μ is \mathcal{H} -graph-measurable. \square

2.6.2 Conditional essential supremum

A very general concept of conditional essential supremum of a family of vector-valued random variables is defined in [42] with respect to a random partial order. In the real case, a generalization of the definition of essential supremum (see [43, Section 5.3.1] for the definition and the proof of existence of the classical essential supremum and Definition 3.1 and Lemma 3.9 in [42] for its conditional generalization as well as the existence, see also [3] where the conditional supremum is defined in the case where I is a singleton) is given by the following result :

Proposition 2.6.3. *Let $\mathcal{H} \subseteq \mathcal{F}$ be two σ -algebras on a probability space. Let $\Gamma = (\gamma_i)_{i \in I}$ be a family of real-valued \mathcal{F} -measurable random variables. There exists a unique \mathcal{H} -measurable random variable $\gamma_{\mathcal{H}} \in L^0(\mathbf{R} \cup \{\infty\}, \mathcal{H})$ denoted $\text{ess sup}_{\mathcal{H}} \Gamma$ which satisfies the following properties :*

- (i) *For every $i \in I$, $\gamma_{\mathcal{H}} \geq \gamma_i$ a.s.*
- (ii) *If $\zeta \in L^0(\mathbf{R} \cup \{\infty\}, \mathcal{H})$ satisfies $\zeta \geq \gamma_i$ a.s. $\forall i \in I$, then $\zeta \geq \gamma_{\mathcal{H}}$ a.s.*

Proof. The proof is given for sake of completeness and pedagogical purpose. The authors thanks T. Jeulin who suggested this (elegant) proof. Considering the homeomorphism \arctan we can restrict ourself to γ_i taking values in $[0, 1]$. We denote by $P_{\gamma_i|\mathcal{H}}$ a regular version of the conditional law of γ_i knowing \mathcal{H} . Let $\zeta \in L^0(\mathbf{R} \cup \{\infty\}, \mathcal{H})$ such that $\zeta \geq \gamma_i$ a.s. $\forall i \in I$. We have that

$$\begin{aligned} \zeta \geq \gamma_i \text{ a.s.} &\Leftrightarrow E(P(\zeta < \gamma_i|\mathcal{H})) = 0 \Leftrightarrow P(\zeta < \gamma_i|\mathcal{H}) = 0 \text{ a.s.} \\ &\Leftrightarrow P(\zeta \geq \gamma_i|\mathcal{H}) = P_{\gamma_i|\mathcal{H}}([-\infty, x])|_{x=\zeta} = 1 \text{ a.s.} \end{aligned}$$

From Definition 2.6.1, $\text{supp}_{\mathcal{H}} \gamma_i \subset]-\infty, \zeta]$ a.s. Let $\Lambda_{\gamma_i|\mathcal{H}} = \sup\{x \in [0, 1], x \in \text{supp}_{\mathcal{H}} \gamma_i\}$ then $\Lambda_{\gamma_i|\mathcal{H}} \leq \zeta$ a.s. For any $c \in \mathbf{R}$, $\{\Lambda_{\gamma_i|\mathcal{H}} \leq c\} = \{P_{\gamma_i|\mathcal{H}}([-\infty, c]) = 1\} \in \mathcal{H}$ since we have chosen for $P_{\gamma_i|\mathcal{H}}$ a regular version of the conditional law of γ_i knowing \mathcal{H} . It follows that $\Lambda_{\gamma_i|\mathcal{H}}$ is \mathcal{H} -measurable. So taking the classical essential supremum, we get that $\text{ess sup}_i \Lambda_{\gamma_i|\mathcal{H}} \leq \zeta$ a.s. and that $\text{ess sup}_i \Lambda_{\gamma_i|\mathcal{H}}$ is \mathcal{H} -measurable. We conclude that $\gamma_{\mathcal{H}} = \text{ess sup}_i \Lambda_{\gamma_i|\mathcal{H}}$ a.s. since for every $i \in I$, $P(\gamma_i \in \text{supp}_{\mathcal{H}} \gamma_i|\mathcal{H}) = 1$ and thus $\text{ess sup}_i \Lambda_{\gamma_i|\mathcal{H}} \geq \gamma_i$ a.s. \square

Remark 2.6.4. Let Q be an absolutely continuous probability measure with respect to P . Let $Z = dQ/dP$ and \mathbb{E}_Q be the expectation under Q . As for every $i \in I$, $\text{ess sup}_{\mathcal{H}} \Gamma \geq \gamma_i$ a.s. and $\text{ess sup}_{\mathcal{H}} \Gamma$ is \mathcal{H} -measurable,

$$\text{ess sup}_{\mathcal{H}} \Gamma \geq \frac{E(Z\gamma_i|\mathcal{H})}{E(Z|\mathcal{H})} = \mathbb{E}_Q(\gamma_i|\mathcal{H}).$$

Inspired by Theorem 2.8 in [3], we may easily show the following tower property :

Lemma 2.6.5. *Let $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \mathcal{F}$ be σ -algebras on a probability space and let $\Gamma = (\gamma_i)_{i \in I}$ be a family of real-valued \mathcal{F} -measurable random variables. Then,*

$$\text{ess sup}_{\mathcal{H}_1} (\text{ess sup}_{\mathcal{H}_2} \Gamma) = \text{ess sup}_{\mathcal{H}_1} \Gamma.$$

2.6.3 Link between two notions

Our goal is to extend the the fact that (see the proof of Proposition 2.6.3)

$$\text{ess sup}_{\mathcal{H}} X = \sup_{x \in \text{supp}_{\mathcal{H}} X} x \text{ a.s.}$$

First we show two useful lemmata on the measurability of the supremum and infimum.

Lemma 2.6.6. *Let $\mathcal{K} : \Omega \rightarrow \mathbb{R}^d$ be a \mathcal{H} -measurable and closed random set such that $\text{dom } \mathcal{K} = \{\emptyset \in \Omega, \mathcal{K}(\emptyset) \cap \mathbb{R}^d \neq \emptyset\} = \emptyset$ and let $h : \Omega \times \mathbb{R}^k \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable function, such that $h(\emptyset, x, \cdot)$ is either l.s.c. or u.s.c., for all $(\emptyset, x) \in \Omega \times \mathbb{R}^k$. Let for all $(\emptyset, x) \in \Omega \times \mathbb{R}^k$*

$$s(\emptyset, x) = \sup_{z \in \mathcal{K}(\emptyset)} h(\emptyset, x, z) \text{ and } i(\emptyset, x) = \inf_{z \in \mathcal{K}(\emptyset)} h(\emptyset, x, z).$$

Then i and s are $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}^k)$ -measurable.

Proof. Let $(\eta_n)_{n \in \mathbb{N}}$ be a Castaing representation of $\mathcal{K} : \mathcal{K}(\emptyset) = \text{cl}\{\eta_n(\emptyset), n \in \mathbb{N}\}$ where the closure is taken in \mathbb{R}^d and $\eta_n(\emptyset) \in \mathcal{K}(\emptyset)$ for all n . Note that η_n is defined in the whole space Ω since $\text{dom } \mathcal{K} = \Omega$. Fix some $c \in \mathbb{R}$. Then, we get that

$$\{(\omega, x) \in \Omega \times \mathbb{R}^d, s(\emptyset, x) \leq c\} = \bigcap_n \{(\omega, x) \in \Omega \times \mathbb{R}^d, h(\emptyset, x, \eta_n(\omega)) \leq c\}.$$

Indeed the first inclusion follows from the fact that $\eta_n(\omega) \in \mathcal{K}(\emptyset)$ for all n and all \emptyset . For the reverse inclusion, fix some $(\emptyset, x) \in \bigcap_n \{(\omega, x), h(\emptyset, x, \eta_n(\omega)) \leq c\}$. For any $z \in \mathcal{K}(\emptyset)$ one gets that $z = \lim_n \eta_n(\omega)$. Then from $h(\emptyset, x, \eta_n(\omega)) \leq c$ we get that $h(\emptyset, x, z) = \liminf h(\emptyset, x, \eta_n(\omega)) \leq c$ in the case where $h(\emptyset, x, \cdot)$ is l.s.c. and $h(\emptyset, x, z) = \limsup h(\emptyset, x, \eta_n(\omega)) \leq c$ in the case where $h(\emptyset, x, \cdot)$ is u.s.c. Now recalling that h is $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable and that η_n is \mathcal{H} -measurable, $(\omega, x) \rightarrow h(\emptyset, x, \eta_n(\omega))$ is $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}^k)$ -measurable,

$\{(\omega, x), h(\emptyset, x, \eta_n(\omega)) \leq c\} \in \mathcal{H} \otimes \mathcal{B}(\mathbb{R}^k)$ and we deduce that s is $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}^k)$ -measurable.

Then we apply the same arguments for i replacing $\leq c$ by $\geq c$. \square

Lemma 2.6.7. *Let $\mathcal{K} : \Omega \rightarrow \mathbb{R}^d$ be a \mathcal{H} -measurable and closed random set such that $\text{dom } \mathcal{K} = \emptyset$ and $h : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}$ be is l.s.c. in x . Then,*

$$\sup_{x \in \mathcal{K}} h(x) = \sup_n h(\eta_n), \tag{2.6.18}$$

where $(\eta_n)_n$ be a Castaing representation of \mathcal{K} .

Proof. As $(\eta_n)_n \subset \mathcal{K}$, $h(\eta_n) \leq \sup_{x \in \mathcal{K}} h(x)$ and thus $\sup_n h(\eta_n) \leq \sup_{x \in \mathcal{K}} h(x)$. Let $x \in \mathcal{K}$ and $\eta_n \rightarrow x$. By lower semicontinuity of h , $h(x) = \liminf_n h(\eta_n) \leq \sup_n h(\eta_n)$ and $\sup_{x \in \mathcal{K}} h(x) \leq \sup_n h(\eta_n)$ and (2.6.18) is proved. \square

Lemma 2.6.8. *Let $X \in L^0(\mathbf{R}^d, \mathcal{F})$ such that $\text{dom supp}_{\mathcal{H}} X = \Omega$ and $h : \Omega \times \mathbf{R}^d \rightarrow \mathbf{R}$ be a \mathcal{H} -normal integrand. Then,*

$$\text{ess sup}_{\mathcal{H}} h(X) = \sup_{x \in \text{supp}_{\mathcal{H}} X} h(x) = \sup_n h(\gamma_n) \text{ a.s.}, \quad (2.6.19)$$

where $(\gamma_n)_{n \in \mathbb{N}}$ is a Castaing representation of $\text{supp}_{\mathcal{H}} X$.

Proof. As $P(X \in \text{supp}_{\mathcal{H}} X | \mathcal{H}) = 1$ we have that $\sup_{x \in \text{supp}_{\mathcal{H}} X} h(x) \geq h(X)$ a.s. and by definition of $\text{ess sup}_{\mathcal{H}} h(X)$, we get that $\sup_{x \in \text{supp}_{\mathcal{H}} X} h(x) \geq \text{ess sup}_{\mathcal{H}} h(X)$ since $\sup_{x \in \text{supp}_{\mathcal{H}} X} h(x)$ is \mathcal{H} -measurable by Lemma 2.6.6 (recall that $\text{supp}_{\mathcal{H}} X$ is \mathcal{H} -measurable and closed, see lemma 2.6.2).

By definition of the essential supremum we also get that $\text{ess sup}_{\mathcal{H}} h(X) \geq h(X)$ a.s. Let $(\gamma_n)_n$ the Castaing representation of $\text{supp}_{\mathcal{H}} X(\emptyset)$, Lemma 2.6.7 implies that $\sup_{x \in \text{supp}_{\mathcal{H}} X} h(x) = \sup_n h(\gamma_n)$ a.s.. Fix some $\varepsilon > 0$ and set $Z_\varepsilon = 1_{B(\gamma_n, \varepsilon)}(X)$, where $B(\gamma_n, \varepsilon)$ is the closed ball of center γ_n and radius ε . Note that $E(Z_\varepsilon | \mathcal{H}) = P(X \in B(\gamma_n, \varepsilon) | \mathcal{H}) > 0$. Indeed if it does not hold true $P(X \in \mathbf{R}^d \setminus B(\gamma_n, \varepsilon) | \mathcal{H}) = 1$ on some $H \in \mathcal{H}$ such that $P(H) > 0$ and by definition 2.6.1, $\text{supp}_{\mathcal{H}} X \subset \mathbf{R}^d \setminus B(\gamma_n, \varepsilon)$ on H , which contradicts $\gamma_n \in \text{supp}_{\mathcal{H}} X$ a.s. As $\text{ess sup}_{\mathcal{H}} h(X)$ is \mathcal{H} -measurable we get that

$$\begin{aligned} \text{ess sup}_{\mathcal{H}} h(X) &\geq \frac{\mathbb{E}(Z_\varepsilon h(X) | \mathcal{H})}{E(Z_\varepsilon | \mathcal{H})} = \frac{\int 1_{B(\gamma_n, \varepsilon)}(x) h(x) P_{X | \mathcal{H}}(dx)}{E(Z_\varepsilon | \mathcal{H})} \\ &\geq \frac{\int (\inf_{y \in B(\gamma_n, \varepsilon)} h(y) 1_{B(\gamma_n, \varepsilon)}(x)) P_{X | \mathcal{H}}(dx)}{E(Z_\varepsilon | \mathcal{H})} \\ &\geq \inf_{y \in B(\gamma_n, \varepsilon)} h(y), \end{aligned}$$

since $\inf_{y \in B(\gamma_n, \varepsilon)} h(y)$ is \mathcal{H} -measurable (see Lemma 2.6.6). Since h is l.s.c., we have that $\lim_{\varepsilon \rightarrow 0} \inf_{x \in B(\gamma_n, \varepsilon)} h(x) = \liminf_{x \rightarrow \gamma_n} h(x) = h(\gamma_n)$ and it follows that $\text{ess sup}_{\mathcal{H}} h(X) \geq h(\gamma_n)$. Taking the supremum over all n , we get that $\text{ess sup}_{\mathcal{H}} h(X) \geq \sup_n h(\gamma_n) = \sup_{x \in \text{supp}_{\mathcal{H}} X} h(x)$.

□

We have the easy extension

Lemma 2.6.9. *Let $\mathcal{X} \subset L^0(\mathbf{R}^d, \mathcal{F})$ such that $\text{dom supp}_{\mathcal{H}} X = \Omega$ for all $X \in \mathcal{X}$ and $\cup_{X \in \mathcal{X}} \text{supp}_{\mathcal{H}} X$ is \mathcal{H} -measurable and closed valued. Let $h : \Omega \times \mathbf{R}^d \rightarrow \mathbf{R}$ be a \mathcal{H} -normal integrand. Then,*

$$\text{ess sup}_{\mathcal{H}} \{h(X), X \in \mathcal{X}\} = \sup_{x \in \cup_{X \in \mathcal{X}} \text{supp}_{\mathcal{H}} X} h(x), \text{ a.s.} \quad (2.6.20)$$

Note that if \mathcal{X} is countable, $\cup_{X \in \mathcal{X}} \text{supp}_{\mathcal{H}} X$ is clearly \mathcal{H} -measurable. If $\mathcal{X} = L^0(\mathbf{R}^d, \mathcal{F})$, then $\cup_{X \in \mathcal{X}} \text{supp}_{\mathcal{H}} X = \mathbf{R}^d$, which is clearly \mathcal{H} -measurable and closed valued.

Proof. For all $X \in \mathcal{X}$, $\text{ess sup}_{\mathcal{H}}\{h(X), X \in \mathcal{X}\} \geq h(X)$ and $\text{ess sup}_{\mathcal{H}}\{h(X), X \in \mathcal{X}\}$ is \mathcal{H} -measurable, so we get that by definition of $\text{ess sup}_{\mathcal{H}}h(X)$ that $\text{ess sup}_{\mathcal{H}}\{h(X), X \in \mathcal{X}\} \geq \text{ess sup}_{\mathcal{H}}h(X)$ and also $\text{ess sup}_{\mathcal{H}}\{h(X), X \in \mathcal{X}\} \geq \sup_{X \in \mathcal{X}} \text{ess sup}_{\mathcal{H}}h(X)$. Conversely, for all $X \in \mathcal{X}$, $\sup_{X \in \mathcal{X}} \text{ess sup}_{\mathcal{H}}h(X) \geq \text{ess sup}_{\mathcal{H}}h(X) \geq h(X)$ and if $\sup_{X \in \mathcal{X}} \text{ess sup}_{\mathcal{H}}h(X)$ is \mathcal{H} -measurable, we obtain by definition of $\text{ess sup}_{\mathcal{H}}\{h(X), X \in \mathcal{X}\}$ that $\sup_{X \in \mathcal{X}} \text{ess sup}_{\mathcal{H}}h(X) \geq \text{ess sup}_{\mathcal{H}}\{h(X), X \in \mathcal{X}\}$. Using Lemma 2.6.8, we get hat

$$\sup_{X \in \mathcal{X}} \text{ess sup}_{\mathcal{H}}h(X) = \sup_{X \in \mathcal{X}} \sup_{x \in \text{supp}_{\mathcal{H}}X} h(x) = \sup_{x \in \cup_{X \in \mathcal{X}} \text{supp}_{\mathcal{H}}X} h(x)$$

Since $\cup_{X \in \mathcal{X}} \text{supp}_{\mathcal{H}}X$ is \mathcal{H} -measurable and closed valued, Lemma 2.6.6 implies that $\sup_{x \in \cup_{X \in \mathcal{X}} \text{supp}_{\mathcal{H}}X} h(x)$ is \mathcal{H} -measurable and the proof is complete. \square

Lemma 2.6.10. *Consider $X \in L^0(\mathbf{R}_+, \mathcal{F})$. Then, we have a.s. that*

$$\begin{aligned} \text{ess inf}_{\mathcal{H}}X &= \inf \text{supp}_{\mathcal{H}}X, & \text{ess sup}_{\mathcal{H}}X &= \sup \text{supp}_{\mathcal{H}}X, \\ \text{ess inf}_{\mathcal{H}}X &\in \text{supp}_{\mathcal{H}}X, & \text{on the set } \{\text{ess inf}_{\mathcal{H}}X > -\infty\}, \\ \text{ess sup}_{\mathcal{H}}X &\in \text{supp}_{\mathcal{H}}X, & \text{on the set } \{\text{ess sup}_{\mathcal{H}}X < \infty\}. \end{aligned}$$

Proof. The two first statements are deduced from the construction of $\text{ess sup}_{\mathcal{H}}X$ in Proposition 2.6.3. Suppose that $\text{ess inf}_{\mathcal{H}}X \notin \text{supp}_{\mathcal{H}}X$ on some non null measure subset $\Lambda \in \mathcal{H}$ of $\{\text{ess inf}_{\mathcal{H}}X > -\infty\}$. By a measurable selection argument, we deduce the existence of $r \in L^0(\mathbf{R}_+, \mathcal{H})$ such that $r > 0$ and $[\text{ess inf}_{\mathcal{H}}X - r, \text{ess inf}_{\mathcal{H}}X + r] \subseteq \mathbf{R} \setminus \text{supp}_{\mathcal{H}}X$ on Λ . As $X \in \text{supp}_{\mathcal{H}}X$ a.s. and $X \geq \text{ess inf}_{\mathcal{H}}X$ a.s., we deduce that $X \geq \text{ess inf}_{\mathcal{H}}X + r$ on Λ , which contradicts the definition of $\text{ess inf}_{\mathcal{H}}X$. The last statement is similarly shown. \square

2.7 Conclusion

Les résultats de la stratégie de sur-réplication testée sur 112 semaines de données de l'indice CAC 40 sont convaincants. La distribution de l'erreur de sur-réplication est concentrée autour de zéro et la moyenne observée de l'erreur est très faible comparée au cours de l'indice dans les différentes configurations expérimentées. Ils confirment l'efficacité de la méthode présentée. De plus, l'indicateur défini par $\sup \left| \frac{S_{t_{i+1}}}{S_{t_i}} - 1 \right|$ semble pouvoir remplacer la volatilité lorsque celle-ci n'est pas disponible (absence de call sur le marché par exemple).

Une question que nous pouvons nous poser est : qu'est-ce qu'il se passe lorsque le pas de temps $\Delta t \rightarrow 0$? C'est ce à quoi nous répondons dans la partie suivante.

Chapitre 3

Diffusion equations : convergence of the functional scheme derived from the Binomial tree with local volatility for non smooth payoff functions

Ce travail fait suite à celui de la partie précédente. Nous nous intéressons ici au schéma [2.3.13](#) dont le modèle binomial est un cas particulier. Nous généralisons le travail du papier [\[55\]](#) au cas d'une condition terminale non régulière ce qui se révèle important dans certains cas comme par exemple la fonction de payoff d'un call $g(x) = (x - K)^+$ qui n'est pas continûment dérivable.

La fonction solution du schéma fonctionnel dérivé du modèle binomial avec volatilité locale converge vers la solution d'une équation de diffusion du type $h_t(t, x) + \frac{x^2 \sigma^2(t, x)}{2} h_{xx}(t, x) = 0$ lorsque le nombre de dates $n \rightarrow \infty$. Contrairement aux méthodes classiques, et notamment les différences finies, la résolution est uniquement basée sur une discrétisation en temps du problème. Nous établissons la convergence uniforme en temps du schéma et donnons le taux de convergence lorsque la fonction de payoff n'est pas aussi régulière qu'en finance. Nous illustrons ces résultats par des exemples et comparons les performances du schéma fonctionnel avec celles des méthodes aux différences finies et aux éléments finis.

3.1 Introduction

There are three major methods to simulate a diffusion equation of type

$$h_t(t, x) + \frac{x^2 \sigma^2(t, x)}{2} h_{xx}(t, x) = 0, \quad t \in [0, T], \quad (3.1.1)$$

with the boundary condition $h(T, x) = g(x)$. The first one is to use the Monte Carlo methods as the solution h admits a probabilistic representation, see [30]. The second one is to numerically compute the solution directly from the PDE. In particular, the very well known finite difference (FD) technique is based on approximations of the successive derivatives. It requires a discretization, both in time and in the space variable, of some compact subset $[0, T] \times [a, b]$, $a \leq b$. Moreover, as we need to fix extra boundary conditions, when $x = a$ or $x = b$, that are a priori approximations of h , a second type of approximation errors may appear. For instance, if b is large enough, we generally set the condition $h(t, b) = \lim_{x \rightarrow \infty} h(t, x)$ if the latter may be estimated while the condition $h(t, 0) = g(0)$ is chosen if $a = 0$ under some mild conditions on σ . There is a lot of articles in the literature focusing on this technique, see for example [8] and [60]. The finite element (FE) method is more sophisticated, see [13], but it is based on a discretization of the space variable as well as other techniques as finite volume [26] or spectral methods [59], [5].

The third one is based on binomial trees. Such a scheme is proposed by Milstein [55] where a convergence theorem is proven when the terminal condition, i.e. the payoff function in finance, is smooth. In particular, it is supposed that the successive derivatives of the solution h of the P.D.E. are uniformly bounded. In finance, e.g. for the Call payoff function $g(x) = (x - K)^+$, this is not the case. On the contrary the successive derivatives of the P.D.E. solution explode at the horizon date [49]. Motivated by pricing in finance, we consider non smooth payoff functions and we study the convergence of this very simple alternative scheme derived on the binomial tree with local volatility financial model, which is only based on a discretization in time of the interval $[0, T]$ and allows to simulate $t \mapsto h(t, x)$ for a fixed x . The advantage of this method is the ease of implementation of the functional scheme which is basically defined. The numerical scheme we study is defined by two functions depending on the number n of discretization dates. We focus here on a particular choice of such functions which leads to a uniform approximation in time of the diffusion equation under some mild conditions. A convergence theorem which generalizes [55] to non smooth terminal conditions is proved. We present some examples where the functional scheme appears to outperform the finite difference method and, moreover, it is well adapted to parallel computing. An open problem is to study the suggested functional scheme more generally.

Motivation and interpretation in finance. Let us consider a binomial model in finance with T steps defined as follows : we consider on a stochastic basis $(\Omega, (\mathcal{F}_t)_{t=0, \dots, T}, \mathbb{P})$ a price process $(S_t)_{t=0, \dots, T}$ where S_0 is given at time 0. Moreover, we suppose that $\mathbb{P}(S_{t+1} = S_t k_t^u | \mathcal{F}_t) = \mathbb{P}(S_{t+1} = S_t k_t^d | \mathcal{F}_t) = 1/2$ where $0 \leq k_t^d < 1 < k_t^u$ are two function-valued processes such that $k^d + k^u = 2$. By the assumptions, we easily deduce that $(S_t)_{t=0, \dots, T}$ is a \mathbb{P} -martingale.

Let us define a portfolio process V as an adapted process satisfying $\Delta V_t := V_t - V_{t-1} = \theta_{t-1} \Delta S_t$ for all $t = 1, \dots, T$, where θ_t is \mathcal{F}_t -measurable for all $t = 0, \dots, T-1$. Moreover, we say that it replicates the contingent claim $g(S_T)$, for some function g , if $V_T = g(S_T)$. If $g(S_T)$ is integrable, we easily deduce that V is a martingale and, by induction, we get that $V_t = C(t, S_t)$ for some measurable function C . Using the martingale property $V_{t-1} = \mathbb{E}(V_t | \mathcal{F}_t)$, we finally deduce that

$$C(t-1, S_{t-1}) = \frac{1}{2}C(t, S_{t-1}k_{t-1}^d) + \frac{1}{2}C(t, S_{t-1}k_{t-1}^u), \quad t = 1, \dots, T. \quad (3.1.2)$$

The natural idea is then to consider this scheme as the number $n+1$ of dates $(Ti/n)_{i=0, \dots, n}$ tends to ∞ when we discretize the continuous-time interval $[0, T]$. This leads to the scheme we propose to study in the next section when $k_t^d(x) = 1 - \sigma(t, x)\sqrt{T/n}$ and $k_t^u(x) = 1 + \sigma(t, x)\sqrt{T/n}$. Actually, the asymptotic behaviour of such binomial model is well known, at least for the price process. Indeed, let us consider the continuous-time price process S^n of the binomial model defined by $S_t^n = S_{T(i-1)/n}^n$ if $t \in [T(i-1)/n, Ti/n)$ where $S_{T(i-1)/n}^n$ is defined from k^d and k^u when starting from S_0 at $t = 0$ as explained above. Then, by [58, Proposition 3.2.1] under mild conditions on σ , the sequence S^n weakly converges to the diffusion process S satisfying the stochastic differential equation

$$dS_t = S_t \sigma(t, S_t) dW_t, \quad (3.1.3)$$

where W is a standard Brownian motion. We also deduce the convergence of the price functions $(C^n)_{n \geq 1}$ given by (3.1.2) towards the limit price C of the continuous-time model given by (3.1.3), i.e. the solution to the diffusion equation (3.1.1). This is indeed the case when σ is constant for a large class of payoff functions g as proven in [8, Chapter 4] where it is also proven that the best rate we can get in general is $\frac{1}{\sqrt{n}}$, see also [66] and the papers [14], [37], [23], and [65] among others and, finally, [55] for smooth terminal conditions and local volatility function. Actually, our main contribution is to show that the result is still valid for local diffusion coefficients even if the solution h of the diffusion equation admits unbounded derivatives, e.g. when the terminal condition is not smooth. Moreover, the numerical experiments we present confirm the accuracy of the functional scheme in comparison to the (FD) and (FE) methods.

3.2 Functional scheme for diffusion equations

Let us consider a bounded diffusion function $\sigma : [0, T] \times \mathbf{R}_+$ and the associated backward parabolic equation

$$h_t(t, x) + \frac{x^2 \sigma^2(t, x)}{2} h_{xx}(t, x) = 0, \quad t \in [0, T], \quad h(T, x) = g(x), \quad (3.2.4)$$

where the terminal condition is defined by a Lipschitz function g with Lipschitz constant L_g and h_t and h_{xx} are respectively the first and the second derivatives of h with respect to time t and space variable $x \in \mathbf{R}$. This equation is very well known in physics but also in mathematical finance for models without friction when the risky asset S is driven by a standard Brownian motion W so that it satisfies the stochastic differential equation $dS_t = \sigma(t, S_t) S_t dW_t$ under a risk neutral probability measure P^1 . In that case, $h(t, S_t)$ is the value at time t of the unique self-financing portfolio process V composed of a fraction of the risk-less bond $B = 1$ and the risky asset S such that it replicates the European option payoff $g(S_T)$, i.e. $V_T = g(S_T)$. The quantity $\sigma(t, S_t)$ is then interpreted as a local volatility coefficient.

In the following, we consider the uniform grid on $[0, T]$ given by $t_i^n = (T/n)i$, $i = 0, \dots, n$ where $n \geq 1$. We then define the following functions :

$$\begin{aligned} k_t^{n+}(x) &:= 1 + \sigma(t, x) \sqrt{\frac{T}{n}}, & k_t^{n-}(x) &:= 1 - \sigma(t, x) \sqrt{\frac{T}{n}}, & (3.2.5) \\ \lambda_t^n(x) &:= \frac{1 - k_t^{n-}(x)}{k_t^{n+}(x) - k_t^{n-}(x)} = \frac{1}{2}, & \mu_t^n(x) &:= 1 - \lambda_t^n(x) = \frac{1}{2}. \end{aligned}$$

Let h^n be the piecewise constant function defined on $[0, T]$ by $h^n(t, x) = h^n(t_{i-1}, x)$ if $t \in [t_{i-1}^n, t_i^n[$, $i \leq n$, where the sequence $(h^n(t_i, x))_{i \leq n}$ is defined recursively by the following functional scheme :

$$\begin{aligned} h^n(t_{i-1}, x) &= \lambda_{t_{i-1}}^n(x) h^n(t_i, k_t^{n-}(x)x) + \mu_{t_{i-1}}^n(x) h^n(t_i, k_t^{n+}(x)x), \\ h^n(T, x) &= g(x), \quad i \leq n. \end{aligned} \quad (3.2.6)$$

Our main result states that the sequence of functions $(h^n)_{n \in \mathbb{N}}$ uniformly converges to h as $n \rightarrow \infty$ under some mild conditions (called Condition D below) satisfied by the successive derivatives of h , solution to (3.2.4). Of course, the functional scheme (3.2.6) may also be considered for other choices of functions k^{n-} and k^{n+} . In this paper, we restrict ourselves to the functions defined by (3.2.5), i.e. when the coefficients defining

1. If we suppose that the risk-free interest rate is $r = 0$

the binomial model are symmetric, but it is an interesting problem to study such a scheme in its full generality, i.e. for more general coefficients k^{n-} and k^{n+} .

Recall that $h(t, x) = \mathbb{E}g(S_{x,t}(T))$ where, for $t \leq T$, $S_{x,t}$ is the unique solution to the stochastic differential equation (3.6.14) on the interval $[t, T]$ with initial condition $S_{x,t}(t) = x$. We deduce that $h(t, 0) = g(0)$. In the following, we suppose that σ is locally Lipschitz and bounded so that existence and uniqueness holds for (3.6.14), see for instance [30, Theorem 2.2, p104]. The conditions below hold in particular when the payoff function g is continuous with linear growth at infinity, e.g. $g(x) = (x - K)^+$, $K \in \mathbf{R}$, and the diffusion coefficient σ is smooth enough and positive as proven in [49, Theorem 4.1] for a more general case², see also [21, Section 5]. Observe that we do not suppose that the successive derivatives are bounded contrarily to [55].

Condition D :

There exists a constant $C > 0$ independent of n such that :

$$\begin{aligned} |h_x(t, x)| &\leq C, & |h_{xx}(t, x)| &\leq \frac{C}{|x|\sqrt{T-t}}, \\ |h_{xxx}(t, x)| &\leq \frac{C}{x^2(T-t)}, & |h_{xt}(t, x)| &\leq \frac{C}{T-t}, \\ |h_{xxt}(t, x)| &\leq \frac{C}{|x|(T-t)^{3/2}}, & |h_{tt}(t, x)| &\leq \frac{C|x|}{(T-t)^{\frac{3}{2}}}, \\ |h_{ttt}(t, x)| &\leq \frac{C|x|}{(T-t)^{\frac{5}{2}}}, & |h_{xtt}(t, x)| &\leq \frac{C}{(T-t)^2}, \end{aligned}$$

Theorem 3.2.1. *Suppose that σ is a bounded Lipschitz function and assume that the solution to Equation (3.2.4) satisfies Condition (D). Then, there exists a constant $C(x) > 0$ depending on x , g and σ such that*

$$\sup_{t \in [0, T]} |h^n(t, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{n}}, \quad n > 0.$$

The proof is deduced from the lemmas and the corollary of Section 3.5. Note that, by choosing specific discretization dates, see [18], we should improve the convergence rate of the functional scheme. This is left for future research.

2. Here, the coefficient γ_n of [49] appears to be a constant independent of n .

3.3 Numerical examples

In the following, we compare the performance of the functional scheme to the Crank-Nicolson finite difference method one on some examples.

3.3.1 The case where the diffusion coefficient only depends on time.

Suppose that $\sigma(t, x) = \sigma(t)$, $t \in [0, T]$, only depends on time. In that case, we may show (see [42]) that

$$h(t, x) = \int_{-\infty}^{\infty} g\left(xe^{\rho_t y - \frac{\rho_t^2}{2}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy,$$

where $\rho_t \geq 0$ is defined by $\rho_t^2 = \int_t^T \sigma_u^2 du$. Moreover, for some particular functions g , e.g. $g(x) = (x - K)^+$, it is trivial to derive an explicit formula of $h(t, x)$. As k^{n-} and k^{n+} do not depend on the space variable, we also deduce an explicit expression of h^n , solution to (3.2.6) :

$$h(t_{n-i}^n, x) = \frac{1}{2^i} \sum_{z \in \mathbb{E}_{n-i}} g(xz), \quad i = 0, \dots, n, \quad (3.3.7)$$

where the sets $(E_i)_{i=0, \dots, n}$ does not depend on x and are recursively defined as $\mathbb{E}_n = \{1\}$ and $\mathbb{E}_{n-i-1} = \left(k_{n-i-1}^{n-} \mathbb{E}_{n-i}\right) \cup \left(k_{n-i-1}^{n+} \mathbb{E}_{n-i}\right)$ for all $i = 0, \dots, n-1$.

Regarding the implementation, we may see each E_i as a vector with 2^{n-i} components. Once these vectors computed for all, we may simulate the scheme (3.2.6) for any x and any arbitrary terminal condition g . Indeed, just compute the vectors G_i whose components are $(g(xz))_{z \in \mathbb{E}_i}$ and deduce $h(t_{n-i}^n, x)$ as the scalar product of G_i with $\mathbf{1} = (1, \dots, 1) \in \mathbf{R}^{2^{n-i}}$.

Let us now consider the case $\sigma(t) = \sigma(2 + \cos(t))$. Then,

$$\rho_t^2 = 4\sigma^2 [(T + \sin T) - (t + \sin t)] + \frac{\sigma^2}{2} \left[(T + \frac{1}{2} \sin(2T)) - (t + \frac{1}{2} \sin 2t) \right]$$

and for $g(x) = (x - K)^+$, $h(t, x)$ is given explicitly by

$$\begin{aligned}
 h(t, x) &= \int_{-\infty}^{\infty} g\left(xe^{\rho_t y - \frac{\rho_t^2}{2}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
 &= \int_{-\infty}^{\infty} \left(xe^{\rho_t y - \frac{\rho_t^2}{2}} - K\right)^+ \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
 &= \int_{\rho_t^{-1}\left(\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}\right)}^{\infty} \left(xe^{\rho_t y - \frac{\rho_t^2}{2}} - K\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
 &= \int_{\rho_t^{-1}\left(\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}\right)}^{\infty} x \frac{e^{-\frac{(y-\rho_t)^2}{2}}}{\sqrt{2\pi}} dy - K \left(1 - \Phi\left(\frac{\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}}{\rho_t}\right)\right) \\
 &= x \left(1 - \Phi\left(\frac{\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}}{\rho_t} - \rho_t\right)\right) - K \Phi\left(\frac{\ln\left(\frac{x}{K}\right) - \frac{\rho_t^2}{2}}{\rho_t}\right) \\
 &= x \Phi\left(\frac{\ln\left(\frac{x}{K}\right) + \frac{\rho_t^2}{2}}{\rho_t}\right) - K \Phi\left(\frac{\ln\left(\frac{x}{K}\right) - \frac{\rho_t^2}{2}}{\rho_t}\right).
 \end{aligned}$$

Therefore, we are in position to numerically compare the convergence error between the explicit solution given above and the functional scheme as well as the finite difference approximations. We first consider the parameters $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$ and 30 discrete dates. As we may observe in Figure 3.1, with a space discretization of 400 points, the finite difference method does not provide a good approximation and, moreover, the monotonicity we expect to is not satisfied.

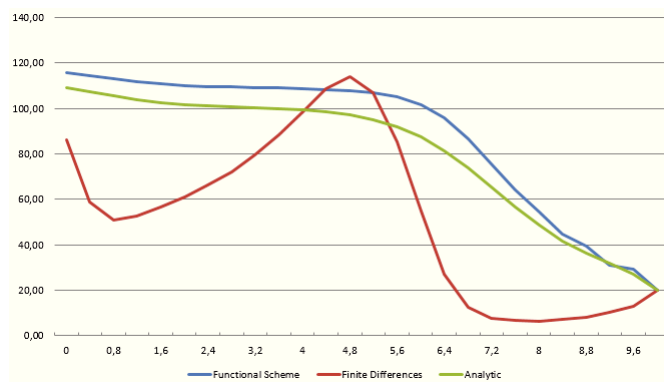


FIGURE 3.1 – Numerical example with $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

With the parameters $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$ and the same discretization in time and in space (see Figure 3.22), the finite difference method provides an approximating function which explodes as time tends to 0.

At last, with the parameters $\sigma = 0.1$, $T = 5$, $K = 115$, we consider 25 discrete dates and 400 points for the discretization in time. We observe that the approximating function given by the finite difference method is not very efficient as well.

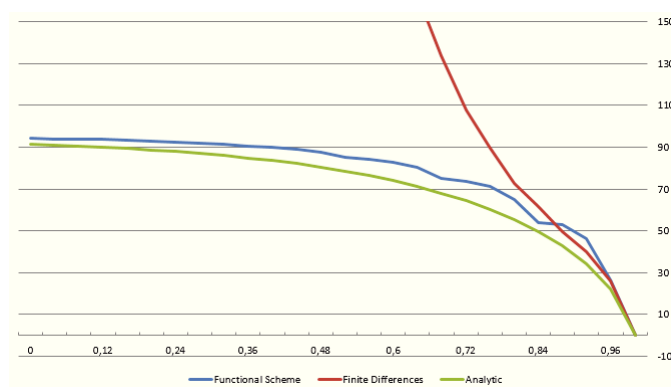


FIGURE 3.2 – Numerical example with $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

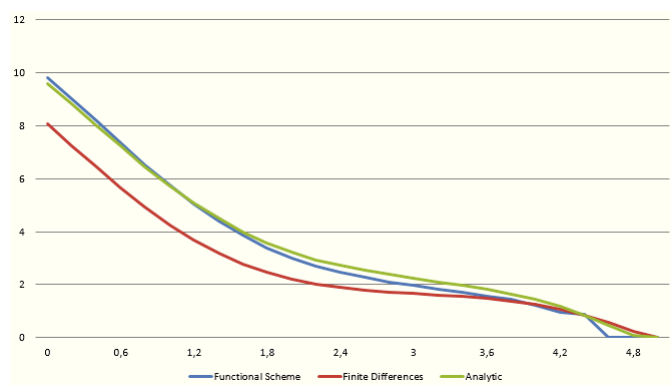


FIGURE 3.3 – Numerical example with $\sigma = 0.1$, $T = 5$, $K = 115$, $x = 115$.

In order to better compare the two methods, we have computed the needed time for several strikes to obtain a reasonable convergence error for the value of $h(0, x)$ with respect to the analytic expression of the later. Initially, the time discretization is done with 10 dates while the space is discretized in 400 points. If the error is larger than 5%, we add as many discrete dates as necessary (see Tables 3.4, 3.5, 3.6 and 3.6).

Strike	Result	Error	Time Discretization	Computation time (s)
50	117.74	4.78%	23	2.36
60	117.04	4.91%	24	4.90
70	116.18	4.83%	26	19.62
80	115.66	4.99%	27	76.45
90	115.00	4.99%	28	152.14
100	114.19	4.80%	30	323.16
110	113.76	4.93%	31	1273.74

FIGURE 3.4 – Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.5$, $T = 10$, $x = 120$.

Clearly, the results are bad for the finite difference method, see Table 3.5.

CHAPITRE 3. DIFFUSION EQUATIONS : CONVERGENCE OF THE FUNCTIONAL SCHEME DERIVED FROM THE BINOMIAL TREE WITH ...

Strike	Result	Error	Time Discretization	Computation time (s)
50	105.34	-6.26%	4000	> 1400
60	103.84	-6.93%	4000	> 1400
70	102.63	-7.40%	4000	> 1400
80	101.59	-7.78%	4000	> 1400
90	100.81	-7.97%	4000	> 1400
100	100.12	-8.11%	4000	> 1400
110	99.60	-8.12%	4000	> 1400

FIGURE 3.5 – Values at $t = 0$ by finite differences with parameters $\sigma = 0.5$, $T = 10$, $x = 120$.

With the parameters $\sigma = 1.5$, $T = 1$, $x = 95$, we get the following :

Strike	Result	Error	Time Discretization	Computation time (s)
50	95.84	3.34%	20	0.29
60	95.84	3.58%	20	0.29
70	95.83	3.80%	20	0.29
80	95.83	4.01%	20	0.29
90	95.82	4.20%	20	0.29
100	95.81	4.38%	20	0.29
110	95.81	4.54%	20	0.29

FIGURE 3.6 – Values at $t = 0$ by the functional scheme with parameters $\sigma = 1.5$, $T = 1$, $x = 95$.

Notice that the results given by the finite difference method are not acceptable, e.g. when K is too large :

Strike	Result	Error	Time Discretization	Computation time (s)
50	3876.89	> 4000%	100	> 20
60	3874.93	> 4000%	100	> 20
70	3873.15	> 4000%	100	> 20
80	3871.49	> 4000%	100	> 20
90	3869.98	> 4000%	100	> 20
100	3868.58	> 4000%	100	> 20
110	3867.27	> 4000%	100	> 20

FIGURE 3.7 – Values at $t = 0$ by the functional scheme with parameters $\sigma = 1.5$, $T = 1$, $x = 95$.

3.3.2 The case where the diffusion coefficient only depends on space.

Let us consider the diffusion coefficient $\sigma(t, x) = \sigma(x) = \sigma * (1 + e^{-x^2})$. As $h(t, x)$ does not admit any analytic expression, we first evaluate it by using its probabilistic representation through a Monte-Carlo discretization of the associated diffusion process (3.6.14). The discretization in time is composed of 10000 dates with 50000 samples for the Monte-Carlo method and 30 dates for the two other methods while we discretize the space variable in 400 points. As we may observe in Figure 3.8, the finite difference method

does not provide a non increasing function in time as expected and the approximation is not accurate.

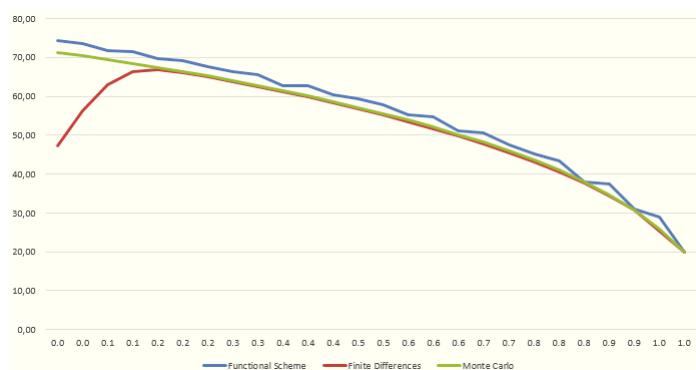


FIGURE 3.8 – Numerical example with $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	91.25	4.93%	12	< 0.01
60	87.17	4.81%	13	< 0.01
70	83.51	4.68%	12	< 0.01
80	80.47	4.91%	10	< 0.01
90	76.94	4.13%	19	0.13
100	73.96	3.73%	19	0.13
110	72.16	4.72%	19	0.13

FIGURE 3.9 – Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	60.40	-30.56%	30	12.46
60	57.21	-31.22%	30	12.10
70	54.39	-31.82%	30	12.34
80	51.80	-32.46%	30	12.27
90	49.57	-32.92%	30	12.19
100	47.46	-33.44%	30	11.99
110	45.62	-33.80%	30	12.24

FIGURE 3.10 – Values at $t = 0$ by the finite differences with parameters $\sigma = 0.5$, $T = 10$, $x = 120$.

In Table 3.11, we see that refining more the space variable grid does not significantly improve the results while the computation time increases.

CHAPITRE 3. DIFFUSION EQUATIONS : CONVERGENCE OF THE FUNCTIONAL SCHEME DERIVED FROM THE BINOMIAL TREE WITH ...

Number of points	Result	Error	Time discretization	Computation time (s)
400	93.45	-13.78%	30	11.58
600	94.57	-12.75%	20	112.23
800	94.96	-12.39%	30	305.76
1000	95.12	-12.24%	30	573.46
1200	95.23	-12.14%	30	1071.47
1400	95.23	-12.10%	30	1913.85
1600	95.32	-12.05%	30	2938.44

FIGURE 3.11 – Values at $t = 0$ by the finite differences with parameters $\sigma = 0.5$, $T = 10$, $x = 120$ and $K = 100$ when refining the space variable grid.

With the parameters $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$ and only 25 discrete dates, the function scheme provides a rather good approximation contrarily to the finite difference method :

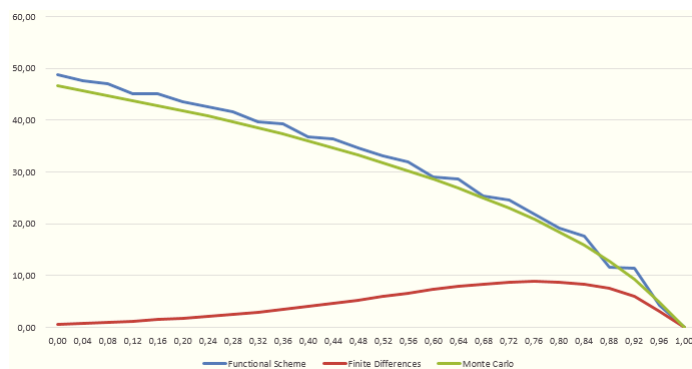


FIGURE 3.12 – Numerical example with $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	66.41	4.36%	11	< 0.01
60	62.89	4.52%	12	< 0.01
70	59.20	3.66%	12	< 0.01
80	56.66	4.24%	17	0.03
90	53.64	3.41%	17	0.03
100	51.98	4.77%	17	0.03
110	49.47	4.04%	20	0.26

FIGURE 3.13 – Values at $t = 0$ by the functional scheme with parameters $\sigma = 1.5$, $T = 1$, $x = 95$.

At last, we consider the parameters $\sigma = 2$, $T = 5$, $K = 10$, $x = 95$ with only 20 discrete dates for the functional scheme. We get satisfactory results in less than 0.26 seconds while with 30 discrete dates for the finite difference methods, the computation time is large. In particular, even if the finite difference method provides acceptable results with only 10 discrete dates with an error less than 5% it takes more than 3.60 secondes to computes them.

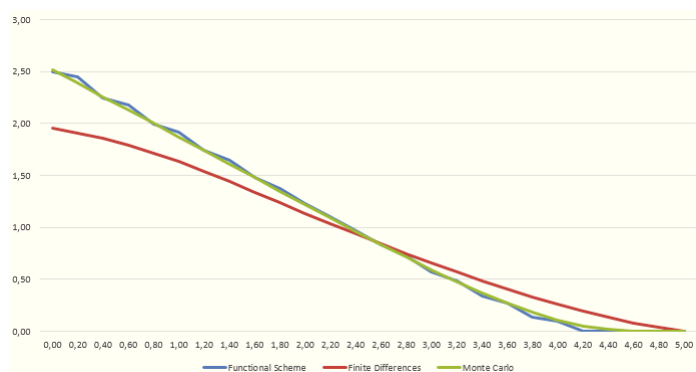


FIGURE 3.14 – Numerical example with $\sigma = 2$, $T = 5$, $K = 10$, $x = 95$

3.3.3 The case where the diffusion coefficient depends on time and space.

We consider the volatility function $\sigma(t, x) = \sigma(1 + \frac{0.01x^2}{1+0.01x^2} + t)$. The number of discrete dates is 25 for the functional scheme as well as the finite difference method and 400 points for the space variable. The Monte Carlo discretization is based on 10000 discrete dates and 50000 trajectories. When the parameters are $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$, the finite difference method provides bad results and it takes a long time to compute them as shown below.

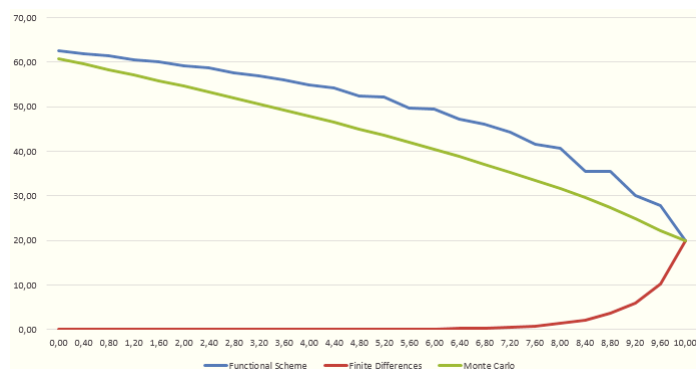


FIGURE 3.15 – Numerical example with $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

For the parameters $\sigma = 0.1$, $T = 5$, $K = 110$, $x = 95$, the functional scheme over evaluates the Monte Carlo method while monotonicity is not satisfied for the finite difference method.

CHAPITRE 3. DIFFUSION EQUATIONS : CONVERGENCE OF THE FUNCTIONAL SCHEME DERIVED FROM THE BINOMIAL TREE WITH ...

Strike	Result	Error	Time Discretization	Computation time (s)
50	83.21	4.36%	11	< 0.01
60	78.26	4.52%	11	< 0.01
70	73.64	3.66%	11	< 0.01
80	69.61	4.24%	11	< 0.01
90	66.18	3.41%	11	< 0.01
100	63.15	4.77%	11	< 0.01
110	60.32	4.04%	11	< 0.01

FIGURE 3.16 – Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.5$, $T = 1$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	0.02	-99.98%	30	11.42
60	0.01	-99.99%	30	13.00
70	0.00	-100.00%	30	11.82
80	0.00	-100.00%	30	12.23
90	0.00	-100.00%	30	11.77
100	0.00	-100.00%	30	12.42
110	0.00	-100.00%	30	11.77

FIGURE 3.17 – Values at $t = 0$ by finite differences with parameters $\sigma = 0.5$, $T = 1$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	56.13	4.36%	10	< 0.01
60	51.10	4.52%	10	< 0.01
70	46.69	3.66%	10	< 0.01
80	42.79	4.24%	10	< 0.01
90	39.33	3.41%	10	< 0.01
100	36.23	4.77%	10	< 0.01
110	33.46	4.04%	10	< 0.01

FIGURE 3.19 – Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.1$, $T = 5$, $x = 95$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	27.21	-50.66%	30	11.69
60	22.88	-54.42%	30	11.58
70	19.37	-57.79%	30	11.55
80	16.38	-61.11%	30	11.56
90	14.03	-63.83%	30	11.66
100	12.04	-66.40%	30	11.60
110	10.41	-68.65%	30	11.59

FIGURE 3.20 – Values at $t = 0$ for finite differences with parameters $\sigma = 0.1$, $T = 5$, $x = 95$.

Notice that it is possible to execute parallel computations by means of programming interfaces (e.g. OpenMP or MPI) i.e. we may use several calculus units simultaneously and reduce computation time. Using $n \geq 1$ processors, we should theoretically divide the needed time by n . In practice, that depends on how the executed parallel calculus are interconnected. In the finite difference method, when computing the value function at

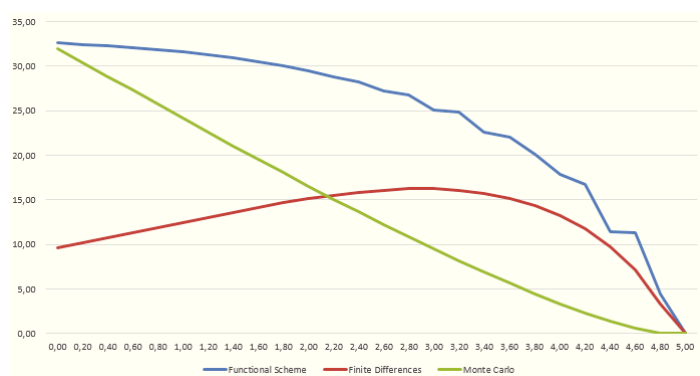


FIGURE 3.18 – Numerical example with $\sigma = 0.1$, $T = 5$, $K = 110$, $x = 95$.

some point of the grid, we need to estimate some expressions depending on the adjacent nodes that may be calculated and used by other processors. This should increase the computation time, as it is necessary to wait for some computations to be executed before starting new ones. On the contrary, for the functional scheme, we do not face this problem as the nodes of the binomial tree at the same level are independent. This is confirmed by the realized time savings we observe to be very closed to the theoretical ones (e.g. 50% with two cores and 75% with 4 cores), see Table 3.21 with the example of the current subsection and the parameters $\sigma = 0.5$, $T = 10$, $K = 100$, $S_0 = 120$. Therefore, even with parallel computing, the functional scheme should outperform the finite difference method in the examples we have presented.

Number of cores	Computation time (s)	Theoretical time saving	Realized time saving
1	388.203	0%	0%
2	194.492	50%	49.90%
4	102.717	75%	73.54%

FIGURE 3.21 – Time savings with the parallel functional scheme when $\sigma = 0.5$, $T = 10$, $K = 100$, $S_0 = 120$.

3.4 Comparison to the Finite Element method.

3.4.1 The case of a bounded payoff function.

Let us consider the payoff function $h(x) = (K - x)^+$. The case when $h(x) = (x - K)^+$ may be deduced by the Call-Put parity $(x - K)^+ - (K - x)^+ = x - K$.

When the volatility function only depends on time (see Section 3.3.1), we obtain very good approximations, see Figure 3.22 and Table 3.23, when the space variable interval is $[0, 10^4]$ discretised in 2000 points. This is confirmed by other sets of parameters.

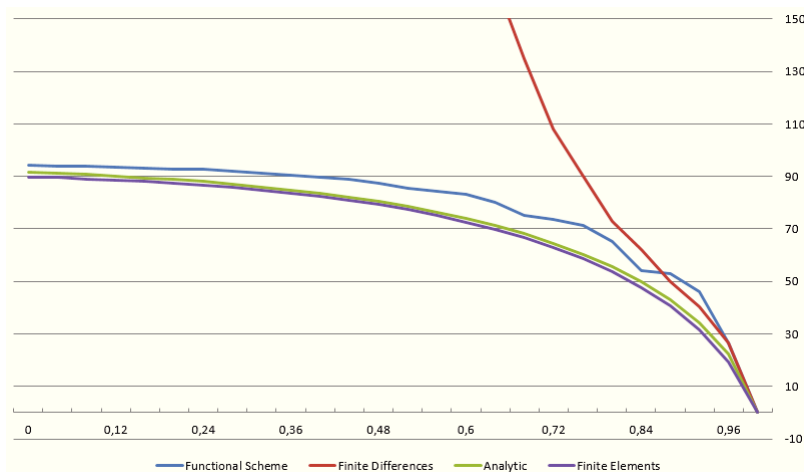


FIGURE 3.22 – Numerical example with $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	90.94	-1.95%	10	0.005
60	90.62	-2.06%	10	0.005
70	90.32	-2.17%	10	0.005
80	90.04	-2.27%	10	0.005
90	89.78	-2.37%	10	0.005
100	89.53	-2.47%	10	0.004
110	89.29	-2.57%	10	0.005

FIGURE 3.23 – Values at $t = 0$ by the FEM with parameters $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

3.4.2 The case of an unbounded payoff function.

Let us consider the payoff function $h(x) = (x - K)^8$ and let us fix the following parameters : $\sigma = 0.25$, $T = 1$ and $K = 0.5$. We then compute $t \mapsto h(t, x)$ for $x = 1$. We compare the obtained results to the approach based on the probabilistic representation of $h(t, x)$, i.e. by implementing the Monte Carlo (MC) technique with a time discretization of 50000 dates and a sample of 100000 trajectories.

First, we observe that the (FE) method is very performant when $\sigma(t, x) = \sigma$ is a constant function and the interval space $[0, 10]$ is discretised in 500 points. To see it, we estimate the relative error of the (FE) approximation with respect to the (MC) method, see Table 3.24.

In this example, the functional scheme seems to be more unstable even if it only takes 3 m.s. to get an error less than 5%. Nevertheless it takes too much time to obtain an error less than 1%, see Table 3.4.2.

number of dates	(FE) method	Time (s)	Error
5	1.1201	0.001	215.32%
10	0.5969	0.001	68.02%
20	0.4508	0.002	26.91%
50	0.3855	0.005	8.52%
100	0.3666	0.009	3.19%
200	0.3576	0.017	0.66%
300	0.3547	0.025	0.16%

FIGURE 3.24 – Approximation of $h(0, x)$ with the (FE) method.

Number of dates	Functional Scheme	Time (s)	Error
5	0.1831	0.000	93.96%
6	0.3331	0.000	6.66%
7	0.3344	0.000	6.23%
8	0.2280	0.000	55.81%
9	0.2382	0.000	49.15%
10	0.3378	0.001	5.16%
11	0.2543	0.001	39.70%
12	0.2608	0.001	36.23%
13	0.3402	0.003	4.41%
14	0.3409	0.006	4.21%
15	0.3415	0.013	4.03%
16	0.2799	0.012	26.91%
17	0.2835	0.024	25.29%
18	0.2868	0.049	23.87%
19	0.3433	0.197	3.48%
20	0.2925	0.196	21.47%
21	0.2949	0.400	20.44%
22	0.3443	1.571	3.17%
23	0.3446	3.135	3.08%
24	0.3449	6.285	3.00%
25	0.3031	6.237	17.20%
26	0.3454	24.917	2.86%
27	0.3456	51.146	2.80%
28	0.3458	107.718	2.74%
29	0.3460	213.635	2.68%

FIGURE 3.25 – Approximation of $h(0, x)$ with the (FS) method and constant volatility.

We now consider the volatility function

$$\sigma(t, x) = \sigma \left(1 + \frac{x^2}{100 \left(1 + \frac{x^2}{100} \right)} + t \right).$$

In that case, it is difficult to obtain the desired convergence as the result is very sensitive to the chosen space variable and its discretization as well. It appears that we reduce the error by increasing the discretization in time but it is not possible to lower the error as much as desired, see Table 3.26.

Surprisingly, if we reduce the number of discrete points of the space variable to 100 points, the error is smaller as observed in Table 3.27. Nevertheless, even by increasing the number of discrete dates to 10^7 and more, it is not possible to reduce as much as desired the convergence error. If we enlarge the space variable to $[0, 25]$ the error is 837% for 2000 discrete dates so that we need to increase more the number of discrete dates.

nbIterations	Finite Elements Method	Time (s)	Error
5	110,8036	0.001	644.37%
10	63.7537	0.001	328.29%
20	39.5068	0.003	165.40%
50	26.7455	0.006	79.67%
100	23.0440	0.011	54.81%
200	21.3239	0.022	43.25%
300	20.7713	0.034	39.54%

FIGURE 3.26 – Approximation of $h(0, x)$ with the (FE) method and non constant volatility

nbIterations	Finite Elements Method	Time (s)	Error
5	103,3949	0,000	594,60%
10	56,7661	0,000	281,35%
20	33,6140	0,001	125,81%
50	21,9113	0,001	47,20%
100	18,6310	0,002	25,16%
200	17,1327	0,005	15,10%
300	16,6556	0,008	11,89%
500	16,2822	0,012	9,38%
1000	16,0070	0,03	7,53%
2000	15,8710	0,049	6,62%
1000000	15,7363	25,581	5,71%
10000000	15,7360	254,609	5,71%

FIGURE 3.27 – Approximation of $h(0, x)$ with the (FE) method and non constant volatility for 100 discrete points.

On the contrary, the functional scheme appears to be efficient, see Table 3.28. In particular, it is possible to reduce as much as desired the convergence error. It takes about 145 seconds to obtain an error less than 5%. Moreover, a code parallelisation should significantly reduce the needed computation time.

nbIterations	Functional Scheme	Time (s)	Error
5	1,4608	0,000	919,00%
6	6,9680	0,000	113,63%
7	7,6439	0,000	94,74%
8	3,0234	0,000	392,35%
9	3,5389	0,001	320,63%
10	9,3410	0,001	59,36%
11	4,5235	0,000	229,07%
12	4,9881	0,001	198,43%
13	10,6514	0,01	39,75%
14	11,0229	0,008	35,04%
15	11,3674	0,017	30,95%
16	6,6509	0,017	123,81%
17	7,0200	0,032	112,04%
18	7,3720	0,066	101,92%
19	12,5264	0,261	18,83%
20	8,0281	0,261	85,42%
21	8,3339	0,527	78,61%
22	13,2189	2,098	12,61%
23	13,4236	4,167	10,89%
24	13,6171	8,331	9,32%
25	9,4280	8,436	57,89%
26	13,9739	34,286	6,52%
27	14,1387	71,794	5,28%
28	14,2953	144,877	4,13%
29	14,4445	291,819	3,05%

FIGURE 3.28 – Approximation of $h(0, x)$ with the (FS) method and non constant volatility.

3.5 Proof

From a line to the next one, we shall use the same notation C for some distinct constants independent of n . Also, notice that there is a constant $M > 0$ such that $|k_t^{n-}| + |k_t^{n+}| \leq M$ for all n and $t \in [0, T]$. We also use the notation $\Delta t_i^n = t_i^n - t_i^{n-1} = T/n$, $i \leq n$. At last, observe that the results below are trivial when $x = 0$ hence we only provide the proofs when $x \neq 0$.

Lemma 3.5.1. *Suppose that the conditions of Theorem (3.2.1) hold. Then, there exists a constant $C > 0$ which does not depend on x such that*

$$\sup_{t \in [t_{n-1}^n, T]} |h^n(t, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{n}}.$$

Proof. Let us define $\delta_t^n = |h(t, x) - h^n(t, x)|$. Recall that $h(t, x) = \mathbb{E}g(S_{x,t}(T))$ where $S_{x,t}$ is the solution to (3.6.14). We have :

$$\begin{aligned}
 \delta_t^n &= |h(t, x) - \frac{1}{2}g(k_{t_{n-1}}^{n-} x) - \frac{1}{2}g(k_{t_{n-1}}^{n+} x)| \\
 &\leq |h(t, x) - g(x)| + \frac{1}{2} |g(k_{t_{n-1}}^{n-} x) - g(x)| + \frac{1}{2} |g(k_{t_{n-1}}^{n+} x) - g(x)| \\
 &\leq \sup_{t_{n-1}^n \leq t \leq T} |h(t, x) - g(x)| + \frac{\mathbf{L}g}{2}(1 - k_{t_{n-1}}^d) |x| + \frac{\mathbf{L}g}{2}(k_{t_{n-1}}^u - 1) |x| \\
 &\leq \mathbf{L}g \sup_{t_{n-1}^n \leq t \leq T} \mathbb{E} |S_{x,t}(T) - x| + \frac{\mathbf{L}g}{2} \sigma(t, x) \sqrt{\frac{T}{n}} |x| + \frac{\mathbf{L}g}{2} \sigma(t, x) \sqrt{\frac{T}{n}} |x| \\
 &\leq C|x| \sqrt{\frac{T}{n}} + \frac{\mathbf{L}g}{2} \sigma^* \sqrt{\frac{T}{n}} |x|,
 \end{aligned}$$

where the last inequality is deduced from Lemma 3.6.1 and $\sigma^* = \sup_{t \in [0, T], x \in \mathbf{R}} |\sigma(t, x)|$. \square

Lemma 3.5.2. *Suppose that the conditions of Theorem (3.2.1) hold. Then, there exists a constant $C > 0$ such that*

$$|h^n(t_{n-2}^n, x) - h(t_{n-2}^n, x)| \leq \frac{C|x|}{\sqrt{n}}.$$

Proof. Let us introduce $\delta_{t-2}^n = |h^n(t_{n-2}^n, x) - h(t_{n-2}^n, x)|$. We have

$$\begin{aligned}
 \delta_{t-2}^n &= \left| \frac{1}{2}h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) + \frac{1}{2}h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, x) \right| \\
 &\leq \frac{1}{2} |h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-2}^n, x)| + \frac{1}{2} |h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, x)| \\
 &\leq \frac{1}{2} |h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x)| + \frac{1}{2} |h(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n-} x)| \\
 &\quad + \frac{1}{2} |h(t_{n-2}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-2}^n, x)| + \frac{1}{2} (|h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x)| \\
 &\quad + \frac{1}{2} |h(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n+} x)| + \frac{1}{2} |h(t_{n-2}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, x)|).
 \end{aligned}$$

Under Condition D, by the mean value theorem, we deduce that

$$\begin{aligned}
 |h(t_{n-2}^n, x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n-} x)| &\leq C |x - k_{t_{n-2}^n}^{n-} x| \leq C |x| \cdot |1 - k_{t_{n-2}^n}^{n-}| \\
 &\leq C |x| \sqrt{\frac{T}{n}}, \\
 |h(t_{n-2}^n, x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n+} x)| &\leq C |x - k_{t_{n-2}^n}^{n+} x| \leq C |x| \cdot |1 - k_{t_{n-2}^n}^{n+}| \\
 &\leq C |x| \sqrt{\frac{T}{n}}.
 \end{aligned}$$

Note that, under Condition D, we have

$$|h_t(t, x)| = \frac{x^2 \sigma^2(t, x)}{2} |h_{xx}(t, x)| \leq \frac{C |x|}{\sqrt{T-t}}. \quad (3.5.8)$$

Therefore, by the mean value theorem, since $k_{t_{n-2}}^{n-}, k_{t_{n-2}}^{n+} \leq M$,

$$\begin{aligned} |h(t_{n-1}^n, k_{t_{n-2}}^{n-} x) - h(t_{n-2}^n, k_{t_{n-2}}^{n-} x)| &\leq \frac{CM |x|}{\sqrt{T-t_{n-1}^n}} \Delta t_{n-1}^n \leq CM |x| \sqrt{\frac{T}{n}}, \\ |h(t_{n-1}^n, k_{t_{n-2}}^{n+} x) - h(t_{n-2}^n, k_{t_{n-2}}^{n+} x)| &\leq \frac{CM |x|}{\sqrt{T-t_{n-1}^n}} \Delta t_{n-1}^n \leq CM |x| \sqrt{\frac{T}{n}}. \end{aligned}$$

At last, by Lemma 3.5.1,

$$\begin{aligned} \gamma_-^n(t_{n-1}^n, x) &:= |h^n(t_{n-1}^n, k_{t_{n-2}}^{n-} x) - h(t_{n-1}^n, k_{t_{n-2}}^{n-} x)| \leq C k_{t_{n-2}}^{n-} |x| \sqrt{\frac{T}{n}}, \\ \gamma_+^n(t_{n-1}^n, x) &:= |h^n(t_{n-1}^n, k_{t_{n-2}}^{n+} x) - h(t_{n-1}^n, k_{t_{n-2}}^{n+} x)| \leq C k_{t_{n-2}}^{n+} |x| \sqrt{\frac{T}{n}}. \end{aligned}$$

As $k_{t_{n-2}}^{n-} + k_{t_{n-2}}^{n+} = 2$, we deduce that

$$\frac{1}{2} \gamma_-^n(t_{n-1}^n, x) + \frac{1}{2} \gamma_+^n(t_{n-1}^n, x) \leq C |x| \sqrt{\frac{T}{n}}.$$

The conclusion follows. \square

Lemma 3.5.3. *Suppose that the conditions of Theorem (3.2.1) hold. Then, there exists a constant $C > 0$ such that*

$$|h^n(t_i^n, x) - h(t_i^n, x)| \leq \frac{C|x|}{\sqrt{n}}, \quad \text{for all } i \leq n-2.$$

Proof. We first prove the result for $i = n-2$ and, then, we generalize the result by induction. Let us introduce the function

$$F_i(x) = \frac{1}{2} h(t_i^n, k_{t_{i-1}}^{n-} x) + \frac{1}{2} h(t_i^n, k_{t_{i-1}}^{n+} x).$$

By the Taylor formula, we get that

$$\begin{aligned} F_i(x) &= \frac{1}{2} [h(t_{i-1}^n, x) + R_i^{11}(x) + R_i^{12}(x) + R_i^{13}(x)] \\ &+ \frac{1}{2} [h(t_{i-1}^n, x) + R_i^{21}(x) + R_i^{22}(x) + R_i^{23}(x)]. \end{aligned}$$

where, for some constants α_i , $i = 1, \dots, 4$, and $\tilde{t}_{i-1}^n \in [t_{i-1}^n, t_i^n]$, $\tilde{x} = x + \tilde{\theta}x(k_{t_{i-1}^n}^n - 1)$, $\tilde{\theta} \in [0, 1]$, $\hat{t}_{i-1}^n \in [t_{i-1}^n, t_i^n]$, $\hat{x} = x + \hat{\theta}x(k_{t_{i-1}^n}^{n+} - 1)$, $\hat{\theta} \in [0, 1]$, we have

$$\begin{aligned}
 R_i^{11}(x) &= h_t(t_{i-1}^n, x) \frac{T}{n} - h_x(t_{i-1}^n, x) x \sigma(t_{i-1}^n, x) \sqrt{\frac{T}{n}}, \\
 R_i^{12}(x) &= -h_{tx}(t_{i-1}^n, x) x \sigma(t_{i-1}^n, x) \left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{1}{2} h_{xx}(t_{i-1}^n, x) x^2 \sigma^2(t_{i-1}^n, x) \frac{T}{n} \\
 &\quad + \frac{1}{2} h_{tt}(t_{i-1}^n, x) \left(\frac{T}{n}\right)^2, \\
 R_i^{13}(x) &= \alpha_1 h_{ttt}(\tilde{t}_{i-1}^n, \tilde{x}) \left(\frac{T}{n}\right)^3 - \alpha_2 h_{ttx}(\tilde{t}_{i-1}^n, \tilde{x}) \sigma(t_{i-1}^n, x) x \left(\frac{T}{n}\right)^{5/2} \\
 &\quad + \alpha_3 h_{txx}(\tilde{t}_{i-1}^n, \tilde{x}) \sigma^2(t_{i-1}^n, x) x^2 \left(\frac{T}{n}\right)^2 - \alpha_4 h_{xxx}(\tilde{t}_{i-1}^n, \tilde{x}) \sigma^3(t_{i-1}^n, x) x^3 \left(\frac{T}{n}\right)^{3/2}, \\
 R_i^{21}(x) &= h_t(t_{i-1}^n, x) \frac{T}{n} + h_x(t_{i-1}^n, x) x \sigma(t_{i-1}^n, x) \sqrt{\frac{T}{n}}, \\
 R_i^{22}(x) &= h_{tx}(t_{i-1}^n, x) x \sigma(t_{i-1}^n, x) \left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{1}{2} h_{xx}(t_{i-1}^n, x) x^2 \sigma^2(t_{i-1}^n, x) \frac{T}{n} \\
 &\quad + \frac{1}{2} h_{tt}(t_{i-1}^n, x) \left(\frac{T}{n}\right)^2, \\
 R_i^{13}(x) &= \alpha_1 h_{ttt}(\hat{t}_{i-1}^n, \hat{x}) \left(\frac{T}{n}\right)^3 + \alpha_2 h_{ttx}(\hat{t}_{i-1}^n, \hat{x}) \sigma(t_{i-1}^n, x) x \left(\frac{T}{n}\right)^{5/2} \\
 &\quad + \alpha_3 h_{txx}(\hat{t}_{i-1}^n, \hat{x}) \sigma^2(t_{i-1}^n, x) x^2 \left(\frac{T}{n}\right)^2 + \alpha_4 h_{xxx}(\hat{t}_{i-1}^n, \hat{x}) \sigma^3(t_{i-1}^n, x) x^3 \left(\frac{T}{n}\right)^{3/2}.
 \end{aligned}$$

We deduce that $F_i(x) = h(t_{i-1}^n, x) + R_i^1(x) + R_i^2(x) + R_i^3(x)$ where $R_i^1(x) = \frac{1}{2} (R_i^{11}(x) + R_i^{21}(x))$, $R_i^2(x) = \frac{1}{2} (R_i^{12}(x) + R_i^{22}(x))$ and $R_i^3(x) = \frac{1}{2} (R_i^{13}(x) + R_i^{23}(x))$. As h is the solution to (3.2.4), we may simplify $F_i(x)$ as

$$F_i(x) = h(t_{i-1}^n, x) + \frac{1}{2} h_{tt}(t_{i-1}^n, x) \left(\frac{T}{n}\right)^2 + R_i^3(x).$$

Under Condition D, there exists a constant C independant of i and n such that

$$\frac{1}{2} |h_{tt}(t_{i-1}^n, x)| \leq \frac{C |x|}{(T - t_{i-1}^n)^{\frac{3}{2}}}. \quad (3.5.9)$$

On the other hand, since $\tilde{t}_{i-1}^n, \hat{t}_{i-1}^n \in [t_{i-1}^n, t_i^n]$ and $|\hat{x}| + |\tilde{x}| \leq C|x|$ where $C > 0$ is independent of i and n , the residual error $R_i^3(x)$ is the sum of terms which may be dominated under Condition D using the following inequalities :

$$|\alpha_1 h_{ttt}(\tilde{t}_{i-1}^n, \tilde{x})| \leq \frac{C |x|}{(T - t_i^n)^{\frac{5}{2}}}. \quad (3.5.10)$$

$$|\alpha_2 h_{xtt}(\hat{t}_{i-1}^n, \hat{x}) \sigma(t_{i-1}^n, x) x| \leq \frac{C |x|}{(T - t_i^n)^2}. \quad (3.5.11)$$

$$|\alpha_3 h_{txx}(\hat{t}_{i-1}^n, \hat{x}) \sigma^2(t_{i-1}^n, x) x^2| \leq \frac{C |x|}{(T - t_i^n)^{\frac{3}{2}}}. \quad (3.5.12)$$

The inequalities above are also satisfied if we replace \hat{t}_{i-1}^n by \tilde{t}_{i-1}^n and \hat{x} by \tilde{x} . Recall that the constant C may denote distinct constants that change for a line to the next one but this constant does not depend on i and n . At last, by the Taylor formula, we rewrite the last term as

$$\begin{aligned} \epsilon_i &= \alpha_4 (h_{xxx}(\hat{t}_{i-1}^n, \hat{x}) \sigma^3(t_{i-1}^n, x) - h_{xxx}(\tilde{t}_{i-1}^n, \tilde{x}) \sigma^3(t_{i-1}^n, x)) x^3 \left(\frac{T}{n}\right)^{3/2}, \\ &= \alpha_4 h_{xxx}(\bar{t}_{i-1}^n, \bar{x}) x^3 (\hat{t}_{i-1}^n - \tilde{t}_{i-1}^n) \left(\frac{T}{n}\right)^{3/2} + \alpha_4 h_{xxxx}(\bar{t}_{i-1}^n, \bar{x}) x^3 (\hat{x} - \tilde{x}) \left(\frac{T}{n}\right)^{3/2}, \end{aligned}$$

where $\bar{t}_{i-1}^n \in [\tilde{t}_{i-1}^n, \hat{t}_{i-1}^n]$ and $\bar{x} = \alpha \tilde{x} + (1 - \alpha) \hat{x}$ for some $\alpha \in [0, 1]$. It is easily seen that $|\hat{t}_{i-1}^n - \tilde{t}_{i-1}^n| \leq CT/n$ for some constant C independent of i and n . Similarly, by definition of \tilde{x} and \hat{x} , there exists a constant C such that $|\hat{x} - \tilde{x}| \leq C|x|\sqrt{T/n}$. Therefore,

$$|\epsilon_i| \leq \frac{C |x|}{(T - t_i^n)^2} \left(\frac{T}{n}\right)^{5/2} + \frac{C |x|}{(T - t_i^n)^{3/2}} \left(\frac{T}{n}\right)^2. \quad (3.5.13)$$

Let us introduce $\delta_{t_{i-1}}^n(x) := |h^n(t_{i-1}^n, x) - h(t_{i-1}^n, x)|$, $i \leq n - 2$. From above, we get that

$$\begin{aligned} \delta_{t_{i-1}}^n(x) &= \left| \frac{1}{2} h^n(t_i^n, k_{t_{i-1}}^{n-} x) + \frac{1}{2} h^n(t_i^n, k_{t_{i-1}}^{n+} x) - h(t_{i-1}^n, x) \right| \\ &= \left| \frac{1}{2} h^n(t_i^n, k_{t_{i-1}}^{n-} x) + \frac{1}{2} h^n(t_i^n, k_{t_{i-1}}^{n+} x) - F_i(x) + \frac{1}{2} h_{tt}(t_{i-1}^n, x) \left(\frac{T}{n}\right)^2 + R_i^3(x) \right| \\ &\leq \frac{1}{2} \delta_{t_i}^n(k_{t_{i-1}}^{n-} x) + \frac{1}{2} \delta_{t_i}^n(k_{t_{i-1}}^{n+} x) + \frac{1}{2} \left(\frac{T}{n}\right)^2 |h_{tt}(t_{i-1}^n, x)| + |R_i^3(x)|. \end{aligned}$$

Using the inequalities (3.5.9), (3.5.10), ..., (3.5.13), we then deduce a constant $C > 0$ independent of i and n such that

$$\begin{aligned} \delta_{t_{i-1}}^n(x) &\leq \frac{1}{2} \delta_{t_i}^n(k_{t_{i-1}}^{n-} x) + \frac{1}{2} \delta_{t_i}^n(k_{t_{i-1}}^{n+} x) \\ &\quad + \frac{C|x|}{(T - t_i^n)^{\frac{1}{2}}} \left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{C|x|}{(T - t_i^n)^{\frac{3}{2}}} \left(\frac{T}{n}\right)^2 \\ &\quad + \frac{C|x|}{(T - t_i^n)^2} \left(\frac{T}{n}\right)^{\frac{5}{2}} + \frac{C|x|}{(T - t_i^n)^{\frac{5}{2}}} \left(\frac{T}{n}\right)^3. \end{aligned}$$

Recall that, by Lemma 3.5.2,

$$\frac{1}{2} \left(\delta_{t_{n-2}}^n(k_{t_{n-3}}^{n-} x) + \delta_{t_{n-2}}^n(k_{t_{n-3}}^{n+} x) \right) \leq \frac{1}{2} C|x| (k_{t_{n-3}}^{n-} + k_{t_{n-3}}^{n+}) \sqrt{\frac{T}{n}} = C|x| \sqrt{\frac{T}{n}}.$$

We deduce that

$$\begin{aligned} |\delta_{t_{n-3}}^n(x)| &\leq C|x| \sqrt{\frac{T}{n}} + \frac{C|x|}{(T - t_{n-2})^{\frac{1}{2}}} \left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{C|x|}{(T - t_{n-2})^{\frac{3}{2}}} \left(\frac{T}{n}\right)^2 \\ &\quad + \frac{C|x|}{(T - t_{n-2})^2} \left(\frac{T}{n}\right)^{\frac{5}{2}} + \frac{C|x|}{(T - t_{n-2})^{\frac{5}{2}}} \left(\frac{T}{n}\right)^3. \end{aligned}$$

Repeating the reasoning, given that $\frac{1}{2}(k_{t_{i-1}}^{n-} + k_{t_{i-1}}^{n+}) = 1$, we deduce by induction that, for every $i \leq n - 3$,

$$|\delta_{t_i}^n(x)| \leq C|x| \sqrt{\frac{T}{n}} + S_{i,n}^1 + S_{i,n}^2 + S_{i,n}^3 + S_{i,n}^4,$$

where, for $i \leq n - 3$,

$$\begin{aligned} S_{i,n}^1 &:= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_j^n)^{\frac{1}{2}}} \left(\frac{T}{n}\right)^{\frac{3}{2}}, & S_{i,n}^2 &= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_j^n)^{\frac{3}{2}}} \left(\frac{T}{n}\right)^2, \\ S_{i,n}^3 &= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_j^n)^2} \left(\frac{T}{n}\right)^{\frac{5}{2}}, & S_{i,n}^4 &= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_{n-2}^n)^{\frac{5}{2}}} \left(\frac{T}{n}\right)^3. \end{aligned}$$

Since $(T - u)^{-1} \geq (T - t_j^n)^{-1}$ if $u \in [t_j^n, t_{j+1}^n]$, we deduce that

$$\begin{aligned} S_{i,n}^1 &\leq C|x| \sqrt{\frac{T}{n}} \int_0^T \frac{1}{\sqrt{T-t}} dt \leq C|x| \sqrt{\frac{T}{n}}, \\ S_{i,n}^2 &\leq C|x| \frac{T}{n} \int_0^{T-T/n} \frac{1}{(T-t)^{3/2}} dt \leq C|x| \sqrt{\frac{T}{n}}, \\ S_{i,n}^3 &\leq C|x| \left(\frac{T}{n}\right)^{3/2} \int_0^{T-T/n} \frac{1}{(T-t)^2} dt \leq C|x| \sqrt{\frac{T}{n}}, \\ S_{i,n}^4 &\leq C|x| \left(\frac{T}{n}\right)^2 \int_0^{T-T/n} \frac{1}{(T-t)^{5/2}} dt \leq C|x| \sqrt{\frac{T}{n}}. \end{aligned}$$

Both with Lemma 3.5.1 and 3.5.2 and the inequalities above, we may conclude. \square

Corollary 3.5.4. *Suppose that the conditions of Theorem (3.2.1) hold. Then, there exists a constant $C > 0$ such that*

$$\sup_{t \in [0, T]} |h^n(t, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{n}}.$$

Proof. Let $t \in [0, T]$ be such that $t \in [t_{i-1}^n, t_i^n)$ for some $i \geq 1$. By Lemma 3.5.1, we may suppose that $i \leq n - 2$. Then,

$$\begin{aligned} |h^n(t, x) - h(t, x)| &= |h^n(t_{i-1}^n, x) - h(t, x)| \\ &\leq |h^n(t_{i-1}^n, x) - h(t_{i-1}^n, x)| + |h(t_{i-1}^n, x) - h(t, x)|. \end{aligned}$$

By the mean value theorem and Inequality (3.5.8), since $T - t_i^n \geq T/n$, we deduce that

$$|h(t_{i-1}^n, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{T - t_i^n}} \Delta t_i^n \leq \frac{C|x|}{\sqrt{n}}.$$

We then conclude by Lemma 3.5.3. \square

3.6 Appendix

Let us consider the unique solution $S_{x,t}$, $t \in [0, T]$, to the stochastic differential equation

$$dS_{t,x}(u) = S_{t,x}(u)\sigma(u, S_{t,x}(u))dW_u, \quad u \in [t, T], \quad S_{t,x}(t) = x \in \mathbf{R}, \quad (3.6.14)$$

where W is a standard Brownian motion.

Lemma 3.6.1. *Suppose that $t \in [0, T]$. Let $S_{x,t}$ be the solution to the stochastic differential equation (3.6.14). If σ is bounded by a constant $\sigma^* > 0$, there exists a constant C such that*

$$\mathbb{E} \sup_{t \leq u \leq T} S_{t,x}^2(u) \leq Cx^2, \quad \mathbb{E}(S_{t,x}(T) - x)^2 \leq Cx^2(T - t).$$

Proof. By the Doob's inequality, we obtain that for every $t \leq r \leq T$:

$$\phi(r) := \mathbb{E} \left| \sup_{t \leq u \leq r} S_{x,t}(u) \right|^2 \leq 4\mathbb{E} |S_{x,t}(r)|^2$$

As $S_{x,t}(r) = x + \int_t^r \sigma(u, S_{x,t}(u))S_{x,t}(u)dW_u$, using the inequality $(a + b)^2 \leq 2(a^2 + b^2)$ and the Ito isometry, we get that :

$$\begin{aligned} \phi(r) &\leq 8x^2 + 8\mathbb{E} \left(\int_t^r \sigma(u, S_{x,t}(u))S_{x,t}(u)dW_u \right)^2 \\ &\leq 8x^2 + 8\mathbb{E} \left(\int_t^r \sigma^2(u, S_{x,t}(u))S_{x,t}^2(u)du \right) \\ &\leq 8x^2 + 8(\sigma^*)^2 \int_t^r \mathbb{E}S_{x,t}^2(u)du \\ &\leq 8x^2 + 8(\sigma^*)^2 \int_t^r \phi(u)du. \end{aligned}$$

Applying the Gronwall lemma, we deduce that :

$$\mathbb{E} | S_{x,t}(r) |^2 \leq 8x^2 \exp(8\sigma^*(T-t)) \leq Cx^2,$$

where C does not depend on x . By the Ito isometry, we then deduce that

$$\mathbb{E}(S_{x,t}(T) - x)^2 = \mathbb{E} \int_t^T \sigma^2(u, S_{x,t}(u)) S_{x,t}^2(u) du$$

Using the inequality above, we may conclude. \square

3.7 Conclusion

Le schéma fonctionnel proposé dans cette partie se démarque par sa facilité d'implémentation comparé aux autres méthodes employées. Pour être le plus honnête possible sur les comparaisons des temps de calculs, l'implémentation de la méthode des différences finies a été largement inspirée de la librairie QuantLib et les objets utilisés sont optimisés pour le problème (conteneur spécial pour opérateur tridiagonal, méthode d'inversion de matrice adaptée, ...). Il en est de même pour la méthode des éléments finis qui n'est pas proposée dans la librairie mais qui est très proche dans son implémentation des différences finies.

Les résultats sont globalement bons pour le schéma fonctionnel qui converge toujours vers la solution bien que lentement parfois. La méthode des différences finies a souvent montré ses limites dans les exemples parcourus. Les éléments finis ont montré de très bonnes performances aussi bien en terme de précision que de temps de calculs sur de nombreux cas testés. Cependant, ils ont été moins performants que le schéma fonctionnel dans certains cas difficiles.

Enfin, le schéma fonctionnel est facilement parallélisable ce qui lui permet de diminuer les temps de calcul avec un ratio quasiment égal au nombre de coeurs utilisés. Avec une carte graphique appropriée contenant des milliers de coeurs, il est tout a fait possible de faire chuter les temps de calculs sans peine. A l'inverse, les deux autres méthodes vues sont plus difficiles à paralléliser du fait des dépendances en temps et en espace des schémas. L'impact sur la performance en serait donc plus modeste.

Chapitre 4

Approximation of non-Lipschitz SDEs by Picard iterations

Ce travail fait suite au papier [54] proposant une méthode basée sur l'utilisation de la forme Doléans-Dade du prix S_t d'un sous-jacent pour une option européenne de payoff $(S_T - K)^+$ et dont la dynamique est définie par $\frac{dS_t}{dt} = \sigma(S_t, t)dW_t + \mu(S_t, t)dt$ ou encore $S_t = S_0 \exp\left(\int_0^t \sigma(S_u, u)dW_u + \int_0^t \mu(S_u, u)du\right)$. Ainsi, en pensant au théorème sur point fixe, nous sommes amenés à construire le schéma $S_t^{n+1} = S_0 \exp\left(\int_0^t \sigma(S_u^n, u)dW_u + \int_0^t \mu(S_u^n, u)du\right)$.

Initialement, ce projet a été proposé par Emmanuel Lépinette à Julien Grépat dans le cadre de sa thèse. Projet inachevé, ma contribution a été d'obtenir une convergence du schéma proposé sous des conditions plus faibles sur les coefficients (volatilité locale) afin d'inclure le plus de modèles possible puis de tester la convergence sur plusieurs modèles par simulations en C++.

Dans cette partie, nous proposons une méthode d'approximation de solutions d'EDS par itérations de Picard déduite de la formule de l'exponentielle de Doléans-Dade. Notre méthode permet d'approximer les trajectoires d'une large classe de processus de Markov, i.e. solutions d'EDS à coefficients non-Lipschitz. La méthode est ensuite appliquée au pricing d'options asiatiques dont le sous-jacent a pour dynamique un processus CEV. Les résultats et les temps de calculs sont comparés à d'autres méthodes de la littérature.

4.1 Introduction

In frictionless models without arbitrage opportunities, it is possible to define the price of a contingent claim as the initial value of the unique portfolio process whose terminal value coincides with the payoff, see [17, Section 3.2.2]. The price of a replicable contingent

claim may be deduced as an expectation as soon as the price process is a local martingale and the replicating portfolio is a martingale under some equivalent probability measure. If the price dynamics is simple enough, e.g. in the Black and Scholes model, it is possible to obtain an explicit expression for the Call or Put option prices. When sophisticated models are considered for a sake of realism [9, 25, 38], one can not expect to compute easily option prices which do not admit pricing formulas in closed form. For this purpose, Monte-Carlo methods have been widely developed to simulate trajectories of diffusion processes and deduce approximations of European option prices [7, 46].

An alternative way is to study the density function of the terminal price value as in the Black and Scholes model [6]. This is also possible for the CEV model [50] where analytic expressions of European Call and Put option prices are deduced. Nevertheless, it is not possible to evaluate the price of an Asian option since a knowledge of the whole trajectory is required. Moreover, the Monte Carlo discretization method seems to be useless as the diffusion coefficient is not Lipschitz so that the Euler scheme should explode. In [31, 45], the asset price process is alternatively approximated through a Picard iteration scheme followed by approximations based on Wiener–Ito chaos expansions. Only the first three terms in the expansions are considered as significative and the associated density functions are approximated. The accuracy of the suggested numerical method illustrated by numerical implementations is highlighted even if it is not confirmed by mathematical arguments. A direct discretization scheme based on Doléans-Dade formula is also used (without mathematical proof of the convergence) in [54] to simulate trajectories in the CEV model. Inspired by these works, our goal is to study the method based on Picard iterations to approximate Asian option prices when the risky asset dynamics is a stochastic differential equation with non Lipschitz-coefficients.

The outline of our work is as follows. By a truncation, we first approximate the stochastic process of interest by a diffusion process with Lipschitz coefficients. Second, we introduce a Picard iteration technique that approximates the process as a Doléans-Dade exponential. Finally, a Monte Carlo discretization scheme is implemented. Comparatively to the classical Euler scheme, we obtain a convergence order n instead of \sqrt{n} where n is the number of discrete dates. The method may be applied to the pricing of a large class of European and Asian options. We illustrate it by computing prices of average Put options in the CEV model where the diffusion coefficient is not Lipschitz. Comparing with the exact formulas and other techniques, our method appears to provide accurate results. Moreover, as soon as we consider an average Put option with more than three dates, our method is less time-consuming than the method combining Monte Carlo simulations and analytic expressions. Nevertheless, in some cases, a simple truncated Euler scheme

appears to be rather efficient whereas the direct discretization scheme based on the Doléans-Dade formula without truncation does not converge. For some other cases, e.g. the CEV model with $\beta > 1$, we clearly observe that the simple Euler scheme explodes but, in that case the backward Euler–Maruyama scheme is the most efficient.

4.2 Model

4.2.1 Notations

We denote by \mathbf{R} the set of all real numbers.

The smallest σ -algebra containing a family of subsets \mathcal{A} is denoted by $\sigma(\mathcal{A})$.

The set of all bounded real-valued functions on $[0, T]$ is denoted by $\mathcal{B}([0, T])$.

Let (Ω, \mathcal{F}, P) be a stochastic basis with a complete filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in [0, T]}$ satisfying the usual hypothesis and $T > 0$ is fixed. We consider a standard Brownian motion W adapted to \mathcal{F} . For any predictable process H , we consider the stochastic integral of H with respect to W denoted by

$$H \cdot W_t := \int_0^t H_u dW_u, \quad t \in [0, T],$$

when existence holds. The Doleans-Dade exponential is denoted by

$$\mathcal{E}_t(H) = \exp \left(H \cdot W_t - \frac{1}{2} \int_0^t H_u^2 du \right), \quad t \in [0, T].$$

In the following, S_0 is the deterministic initial value at time 0 of a process S modelling an asset price in a financial market model. In the following, we consider uniform partitions $(\tau_n)_{n \geq 0}$ of the interval $[0, T]$. The grid τ_n is composed of $n + 1$ dates where n is called the order of the discretization scheme. It is defined by

$$\tau_n = \{t_0 = 0, t_1 = T/n, t_2 = 2T/n, \dots, t_n = T\}.$$

We use the abuse of notation $t_i^n = t_i = (T/n)i$, $i = 0, \dots, n$, for a sake of simplicity. If X is a stochastic process defined on $[0, T]$, we denote by $X^{(n)}$ the piecewise constant process defined by

$$X_t^{(n)} = X_{t_i}, \quad t \in [t_i, t_{i+1}), \quad X_T^{(n)} = X_T.$$

4.2.2 The stochastic differential equation

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}, P)$ be a continuous-time stochastic basis satisfying the usual assumptions and supporting a standard Brownian motion W , i.e. $\mathcal{F}_t := \sigma(W_s : s \leq t) \vee \mathcal{N}$

where \mathcal{N} designates the P-negligible sets. In the case of a financial market model, there is a risk-free bond given by the interest rate r^0 and a risky asset price stochastic process S .

By [30, Theorem 2.2 p 104], recall that we have the following.

Proposition 4.2.1. *Let $\sigma : [0, T] \times \mathbf{R} \mapsto \mathbf{R}$ and $r : [0, T] \times \mathbf{R} \mapsto \mathbf{R}$ be Lipschitz functions. The stochastic differential equation (s.d.e.)*

$$dS(t) = S(t)\sigma(t, S(t))dW_t + S(t)r(t, S(t))dt, \quad S(0) = S_0 > 0, \quad (4.2.1)$$

admits a unique strong solution S . Moreover, it is positive.

In the sequel, we shall consider locally Lipschitz coefficients r and σ and we assume that r is bounded. We suppose that $S = 0$ on $[\tau \wedge T, T]$ where $\tau := \inf\{t : S_t = 0\}$, meaning that 0 is an absorbing point. The s.d.e. (4.2.1) admits a unique non negative solution under the conditions of the proposition above but we only suppose one of the following Conditions 4.2.2 or 4.2.3.

Condition 4.2.2. *There exists constants $C_\kappa > 0$ for every $\kappa > 0$, such that*

$$\begin{aligned} |\sigma(t, x \wedge \kappa) - \sigma(u, x \wedge \kappa)| &\leq C_\kappa |t - u|, & (t, u, x) \in (\mathbf{R}_+)^2 \times (0, \infty), \\ |r(t, x \wedge \kappa) - r(u, x \wedge \kappa)| &\leq C_\kappa |t - u|, & (t, u, x) \in (\mathbf{R}_+)^2 \times (0, \infty), \\ |\sigma(t, x \wedge \kappa) - \sigma(t, y \wedge \kappa)| &\leq C_\kappa |x - y|, & (t, x, y) \in \mathbf{R}_+ \times (0, \infty)^2, \\ |r(t, x \wedge \kappa) - r(t, y \wedge \kappa)| &\leq C_\kappa |x - y|, & (t, x, y) \in \mathbf{R}_+ \times (0, \infty)^2, \\ |r(t, x \wedge \kappa)| + |\sigma(t, x \wedge \kappa)| &\leq C_\kappa, & (t, x, y) \in \mathbf{R}_+ \times (0, \infty)^2. \end{aligned}$$

Condition 4.2.3. *There exists constants $C_\kappa > 0$ for every $\kappa > 0$, such that*

$$\begin{aligned} |\sigma(t, x \vee \kappa) - \sigma(u, x \vee \kappa)| &\leq C_\kappa |t - u|, & (t, u, x) \in (\mathbf{R}_+)^2 \times (0, \infty), \\ |r(t, x \vee \kappa) - r(u, x \vee \kappa)| &\leq C_\kappa |t - u|, & (t, u, Y) \in (\mathbf{R}_+)^2 \times (0, \infty), \\ |\sigma(t, e^X \vee \kappa^{-1}) - \sigma(t, e^Y \vee \kappa^{-1})| &\leq C_\kappa |X - Y|, & (t, X, Y) \in \mathbf{R} \times \mathbf{R}^2, \\ |r(t, e^X \vee \kappa^{-1}) - r(t, e^Y \vee \kappa^{-1})| &\leq C_\kappa |X - Y|, & (t, X, Y) \in \mathbf{R} \times \mathbf{R}^2, \\ |r(t, x \vee \kappa)| + |\sigma(t, y \vee \kappa)| &\leq C_\kappa, & (t, x, y) \in \mathbf{R}_+ \times (0, \infty)^2. \end{aligned}$$

Moreover, when $P(\tau < T) > 0$, we suppose that the mappings $x \mapsto \sigma(t, x)$ and $x \mapsto r(t, x)$ are differentiable on $(0, \infty)$ for all $t \in [0, T]$ and $x\nabla_x \sigma(t, x)$, $x\nabla_x r(t, x)$ are bounded on $[\kappa^{-1}, \infty)$ by C_κ .

Remark 4.2.4. *Under Condition 4.2.2, S admits moments at any order, see the proof in Section 4.7. Notice that this assumption is satisfied by the CEV model $dS_t = \sigma S_t^\beta dW_t$*

and, more generally, when σ is of the form $\sigma(t, x) = \nu(t)x^{\beta(t)-1}$, $x > 0$, where ν and $\beta \geq 1$ are bounded functions and ν is Lipschitz.

Condition 4.2.3 is satisfied by the CEV model, e.g. when σ is of the form $\sigma(t, x) = \nu(t)x^{\beta(t)-1}$ where ν is Lipschitz and bounded and $1 - \alpha \leq \beta \leq 1$, $\alpha \in \mathbf{R}_+$. This also includes the CIR model and more generally mean-reverting SDEs as used in finance, see numerical simulations below.

In finance, under a risk-neutral probability measure for S , e.g. when the market is complete, we are interested in approximating the price of a contingent claim whose payoff $G(S)1_{\tau > T}$ depends on the whole trajectory $S = (S_t)_{t \in [0, T]}$. Note that the delivered payoff $G(S)1_{\tau > T}$ is 0 when $S_T = 0$, i.e. when the firm having issued the risky asset failed.

We suppose that the functional $G : \mathcal{B}([0, T]) \rightarrow \mathbf{R}_+$ satisfies the Lipschitz condition :

$$|G(\alpha) - G(\beta)| \leq C\|\alpha - \beta\|, \quad \alpha, \beta \in \mathcal{B}([0, T]), \quad (4.2.2)$$

where $C > 0$ and $\|\cdot\|$ is the norm of uniform convergence on $[0, T]$, i.e.

$$\|f\| := \sup_{t \in [0, T]} |f(t)|, \quad f \in \mathcal{B}([0, T]).$$

If we assume that \mathbf{P} is the unique risk-neutral probability measure in a complete market, the smallest price v of the contingent claim $G(S)1_{\tau > T}$ is evaluated under \mathbf{P} by $v = \mathbf{E}e^{-r^0 T}G(S)1_{\tau > T}$ where \mathbf{E} is the expectation under \mathbf{P} and r^0 is the constant risk-free interest rate.

In the following, a stochastic process $r : u \mapsto r_u$ models an interest rate of the market. The price of the associated bond with initial value 1 is

$$F_t(r) := \exp\left(\int_0^t r_u du\right), \quad t \in [0, T]. \quad (4.2.3)$$

We consider a class of payoff functions G adapted to the numerical scheme we propose. For a sake of simplicity, we make the abuse of notation $r_u = r(u, S_u)$ as well as $\sigma_u = \sigma(u, S_u)$ so that the notations $F(r)$ and $\mathcal{E}(\sigma)$ make sense.

Definition 4.2.5. We say that the payoff function G is adapted to the discretization scheme defined by the partition τ_n if there exists $\alpha(n) \geq 1$, $p, q \in \mathbf{N}$ such that, for all Lipschitz functions σ and r satisfying

$$\int_0^T \mathbf{E}F_u^{2p}(r)\mathcal{E}_u^{2p}(\sigma)du < \infty,$$

we have, for some function f ,

$$|\mathbf{E}G(S_{r, \sigma}) - \mathbf{E}G(S_{r, \sigma}^{(n)})| \leq \frac{f(G, \sigma, r, p, q)}{\alpha(n)} \left(\int_0^T [1 + \mathbf{E}S_{r, \sigma}^{2p}(u)du] \right)^{1/q},$$

where $S_{r,\sigma} = F(r)\mathcal{E}(\sigma)$ and f depends on σ, r, p through the quantities $\max_x |\sigma(t, x)x|^p / (1 + |x|^{2p}) < \infty$ and $\max_x |r(t, x)x|^p / (1 + |x|^{2p}) < \infty$.

Recall that $F(r)$ is given by (4.2.3) and, in the definition above, the process $S_{r,\sigma} = F(r)\mathcal{E}(\sigma)$ is interpreted as a price. Observe that the European call payoff function is adapted to the discretization whatever $\alpha(n)$ is. Indeed, the payoff function of the European call option with strike K is $G(S)$ where $G(\alpha) = (\alpha(T) - K)^+$, $\alpha \in \mathcal{B}([0, T])$ and, since G only depends on the terminal value of α , we have $G(S_{r,\sigma}) = G(S_{r,\sigma}^{(n)})$.

The payoff of the average strike Asian call option with strike K is $G(S)$ where

$$G(\alpha) = \left(\frac{1}{T} \int_0^T \alpha(t) dt - K \right)^+, \quad \alpha \text{ is } \mathcal{B}([0, T])\text{-measurable.}$$

Lemma 4.2.6. *The average strike Asian call option function G with strike K is adapted to the discretization with $\alpha(n) = n$.*

All the proofs are postponed to Section 4.7.

4.3 Approximation of the stochastic process

The approximation scheme we propose is obtained in three steps. By a truncation, we first approximate the price process by a diffusion process S_κ , $\kappa > 0$, with bounded and Lipschitz volatility. By Picard iterations, a recursive scheme is introduced to approximate S_κ by $S_{\kappa,m}$, $m \geq 1$, Doléans–Dade solutions to stochastic differential equations defined by iterated and uniformly bounded volatilities. At last, we discretize $S_{\kappa,m}$ in time.

4.3.1 Volatility truncation

Let $\kappa \in \mathbf{R}_+$ be such that $\kappa \geq \sup_{(t,x) \in \mathbf{R}^2} r(t, x)$ and $\kappa > S_0 \vee 1$ when Condition 4.2.2 holds and $\kappa^{-1} < \min(\max r, S_0 \vee 1)$ when Condition 4.2.3 is satisfied. Consider Y_κ the unique solution to the s.d.e.

$$\begin{aligned} Y_\kappa(0) &= 0, \\ dY_\kappa(t) &= \sigma_\kappa(t, S_0 e^{Y_\kappa(t)}) dW_t + \gamma_\kappa(t, S_0 e^{Y_\kappa(t)}) dt, \end{aligned} \tag{4.3.4}$$

where $\sigma_\kappa(u, x) := \sigma(u, x \wedge \kappa)$ and $r_\kappa(u, x) := r(u, x \wedge \kappa)$ when Condition 4.2.2 holds. If Condition 4.2.3 holds, we define $\sigma_\kappa(u, x) := \sigma(u, x \vee \kappa^{-1})$ and $r_\kappa(u, x) := r(u, x \vee \kappa^{-1})$. At last, we define $\gamma_\kappa(t, x) = r_\kappa(t, x) - \frac{1}{2}\sigma_\kappa^2(t, x)$. Observe that $S_\kappa := S_0 e^{Y_\kappa}$ satisfies, for

all $t \in [0, T]$,

$$\begin{aligned} S_\kappa(0) &= S_0, \\ S_\kappa(t) &:= S_0 \exp \left[\int_0^t \sigma_\kappa(u, S_\kappa(u)) dW_u + \int_0^t \gamma_\kappa(u, S_\kappa(u)) du \right]. \end{aligned}$$

Lemma 4.3.1. *Suppose that Conditions 4.2.2 and 4.2.3 hold respectively. Then, S_κ converges pointwise to S as $\kappa \rightarrow \infty$ on the interval $[0, T]$ (on the interval $[0, \tau)$ respectively).*

In the following, we use the following assumption :

Condition 4.3.2. *There exists $p > 0$ and a constant C_p such that*

$$\mathbf{E} \sup_{t \in [0, T]} S_t^{2p} + \sup_{\kappa} \mathbf{E} \sup_{t \in [0, T]} S_\kappa^{2p}(t) \leq C_p, \quad \forall \kappa > 0. \quad (4.3.5)$$

By the Lemmas 4.3.4 and 4.3.5 proven in appendix, we deduce that Condition 4.3.2 is satisfied for the large class of models satisfying the following Condition 4.3.3 above even if the coefficients of (4.2.1) are unbounded.

Condition 4.3.3. *There exists $L > 0$ and $\alpha \in (0, 1]$ such that the following conditions are satisfied under Conditions 4.2.2 and 4.2.3 respectively :*

$$(I) : 0 < \sigma^2(t, x) \leq L(1 + |\ln(\ln(x))|1_{x>1}), \quad (t, x) \in \mathbf{R} \times \mathbf{R},$$

$$|r(t, x)| \leq L, \quad (t, x) \in \mathbf{R} \times \mathbf{R},$$

$$(II) : |\sigma(t, x)| + |r(t, x)| \leq \frac{L}{|x|^\alpha} + L, \quad x > 0.$$

Lemma 4.3.4. *Assume that Condition 4.3.3 is satisfied. Then, for all $p \geq 0$, there exists a constant C_p independent of κ such that $\mathbf{E} \sup_{u \leq T} (S_\kappa(u))^p \leq C_p$.*

Lemma 4.3.5. *Assume that Condition 4.3.3 is satisfied. Then, for all $p \geq 0$, there exists a constant C such that $E \sup_{u \leq T} (S_u)^p \leq C_p$.*

Lemma 4.3.6. *Suppose that Condition 4.3.2 is satisfied for some $p > 0$. Moreover, assume that Condition 4.3.3 (II) holds if $\sigma_\kappa = \sigma(\cdot, \cdot \vee \kappa^{-1})$. Then, for all $l \geq 1$, $l \in \mathbf{N}$, there exists a constant $C_{l,p}$ such that*

$$\begin{aligned} \sigma_\kappa = \sigma(\cdot, \cdot \wedge \kappa) &\Rightarrow \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \leq \frac{C_{l,p}}{\kappa^l}, \\ \sigma_\kappa = \sigma(\cdot, \cdot \vee \kappa^{-1}) &\Rightarrow \mathbf{E} \sup_{t \in [0, \tau \wedge T]} |S_\kappa(t) - S_t|^p \leq C_p \alpha(\tau, \kappa), \end{aligned}$$

where $\alpha(\tau, \kappa) \rightarrow 0$, as $\kappa \rightarrow \infty$.

This result shows that S may be approximated by S_κ in the L^p sense, $p > 0$, as $\kappa \rightarrow \infty$.

4.3.2 Picard iterations

In this section, we construct iterated Doleans-Dade processes $S_{\kappa,m}$, $m \in \mathbf{N}$, such that $S_{\kappa,m}$ converge to S_κ when $m \rightarrow \infty$. To do so, we introduce the following scheme :

$$\begin{aligned} S_{\kappa,0} &:= S_0, \\ S_{\kappa,m+1}(t) &:= F_t(r_{\kappa,m+1})\mathcal{E}_t(\sigma_{\kappa,m+1}), \quad t \in [0, T], \end{aligned}$$

where

$$\begin{aligned} \sigma_{\kappa,m+1}(t) &:= \sigma_\kappa(t, S_{\kappa,m}(t)), \quad m \geq 0, \quad t \in [0, T], \\ r_{\kappa,m+1}(t) &:= r_\kappa(t, S_{\kappa,m}(t)), \quad m \geq 0, \quad t \in [0, T]. \end{aligned}$$

Notice that we have

$$S_{\kappa,m+1}(t) = S_0 \exp \left[\int_0^t \sigma_\kappa(u, S_{\kappa,m}(u)) dW_u + \int_0^t \gamma_\kappa(u, S_{\kappa,m}(u)) du \right].$$

Let us introduce

$$Y_{\kappa,m}(t) := \log S_{\kappa,m}(t) - \log S_0, \quad t \in [0, T].$$

This process satisfies the following s.d.e.

$$dY_{\kappa,m+1}(t) = \sigma_\kappa(t, S_{\kappa,m}(t)) dW_t + \gamma_\kappa(t, S_{\kappa,m}(t)) dt.$$

We shall prove the convergence of $S_{\kappa,m}$ to S_κ in L^2 as $m \rightarrow \infty$.

Lemma 4.3.7. *The sequence $Y_{\kappa,m}$ converges in L^4 to $Y_{\kappa,\infty}$ and*

$$\left\| \sup_{u \leq T} |Y_{\kappa,m+1}(u) - Y_{\kappa,m}(u)| \right\|_2 \leq C(1 + S_0)(C_\kappa)^{m+1} \frac{\sqrt{T}^{m+1}}{\sqrt{(m+1)!}}, \quad (4.3.6)$$

$$\left\| \sup_{u \leq T} |Y_{\kappa,m+1}(u) - Y_{\kappa,m}(u)| \right\|_4 \leq C(1 + S_0)(C_\kappa)^{m+1} \frac{(T^{1/4})^{m+1}}{(m+1)!^{1/4}}, \quad (4.3.7)$$

where C is a constant which does not depend on κ and C_κ is a constant which depends on κ .

Corollary 4.3.8. *There exists constants C, C_κ independent of m such that*

$$\begin{aligned} \mathbf{E} \sup_{u \leq t} |Y_{\kappa,\infty}(u) - Y_{\kappa,m}(u)|^2 &\leq C(1 + S_0^2) \sum_{j=m+1}^{\infty} \frac{(C_\kappa)^j}{j!}, \\ \mathbf{E} \sup_{u \leq t} |Y_{\kappa,\infty}(u) - Y_{\kappa,m}(u)|^4 &\leq C(1 + S_0^4) \sum_{j=m+1}^{\infty} \frac{(C_\kappa)^j}{j!}. \end{aligned}$$

Lemma 4.3.9. *For every κ , there exists constants C, C_κ such that*

$$\begin{aligned} \mathbf{E} \sup_{t \in [0, T]} (S_{\kappa, m+1}(t) - S_{\kappa, m}(t))^2 &\leq C_\kappa \sqrt{\mathbf{E} \sup_{t \in [0, T]} (Y_{\kappa, m}(t) - Y_{\kappa, m-1}(t))^4} \\ &\leq C_\kappa \frac{(C_\kappa)^m}{\sqrt{m!}}. \end{aligned}$$

By the lemma above, we deduce that $S_{\kappa, m}$ converges to $S_{\kappa, \infty}$ in L^2 with respect to the uniform convergence.

Lemma 4.3.10. *For every κ , there exists a constant C_κ such that*

$$\mathbf{E} \sup_{t \in [0, T]} (S_{\kappa, \infty}(t) - S_{\kappa, m}(t))^2 \leq C_\kappa \sum_{j=m}^{\infty} \frac{(C_\kappa)^j}{\sqrt{j!}}.$$

Note that the limit processes satisfy

$$dY_{\kappa, \infty}(t) = \sigma_\kappa(t, S_{\kappa, \infty}(t))dW_t + \gamma_\kappa(t, S_{\kappa, \infty}(t))dt, \quad (4.3.8)$$

$$dS_{\kappa, \infty}(t) = S_{\kappa, \infty}(t)\sigma_\kappa(t, S_{\kappa, \infty}(t))dW_t + S_{\kappa, \infty}(t)r_\kappa(t, S_{\kappa, \infty}(t))dt. \quad (4.3.9)$$

It follows that $S_{\kappa, \infty} = S_\kappa$ and $Y_{\kappa, \infty} = Y_\kappa$ since uniqueness holds.

4.3.3 Volatility discretization based on Picard iterations.

In the following, we suppose without loss of generality that $T/n \leq 1$ if n is large enough. We recursively define the processes $S_{\kappa, m, n}, \sigma_{\kappa, m+1, n}, m \geq 0$, as follows :

$$\begin{aligned} S_{\kappa, 0, n} &:= S_0, \\ \sigma_{\kappa, 1, n}(t) &:= \sigma_\kappa(t_i, S_0), \quad r_{\kappa, 1, n}(t) := r_\kappa(t_i, S_0), \quad t_i \leq t < t_{i+1}, \\ S_{\kappa, m, n}(t) &:= F_t(r_{\kappa, m, n})\mathcal{E}_t(\sigma_{\kappa, m, n}), \quad m \geq 1, \quad t \in [0, T], \\ \sigma_{\kappa, m+1, n}(t) &= \sigma_\kappa(t^{(n)}, S_{\kappa, m, n}^{(n)}(t)), \quad m \geq 1, \quad t \in [0, T], \\ r_{\kappa, m+1, n}(t) &= r_\kappa(t^{(n)}, S_{\kappa, m, n}^{(n)}(t)), \quad m \geq 1, \quad t \in [0, T]. \end{aligned}$$

where we use the notation $t^{(n)} = t_i$ if $t_i \leq t < t_{i+1}$ and we recall that

$$\begin{aligned} \mathcal{E}_t(\sigma_{\kappa, m, n}) &= \exp \left[\int_0^t \sigma_\kappa(u^{(n)}, S_{\kappa, m-1, n}^{(n)}(u))dW_u - \frac{1}{2} \int_0^t \sigma_\kappa^2(u^{(n)}, S_{\kappa, m-1, n}^{(n)}(u))du \right], \\ F_t(r_{\kappa, m, n}) &= S_0 \exp \left[\int_0^t r_\kappa(u^{(n)}, S_{\kappa, m-1, n}^{(n)}(u))du \right], \quad m \geq 1, \quad t \in [0, T]. \end{aligned}$$

We also define $Y_{\kappa, m, n} = \ln(S_{\kappa, m, n}/S_0)$.

Lemma 4.3.11. *For all $\kappa > 0$, there exists constants $\alpha_\kappa, C_\kappa > 0$ such that for all $m, n \geq 1$:*

$$\mathbf{E} \sup_{t \in [0, T]} |S_{\kappa, m}(t) - S_{\kappa, m, n}(t)|^2 \leq \frac{\alpha_\kappa C_\kappa^m}{n}.$$

4.3.4 Main result

In the following, we follow the convention that $\tilde{\tau}^{\kappa,m,n} = +\infty$ if $P(\tau \leq T) = 0$. Otherwise, $\tilde{\tau}^{\kappa,m,n}$ is defined as

$$\tilde{\tau}^{\kappa,m,n} := \inf\{t : S_{\kappa,m,n}^{(n)}(t) \leq \kappa^{-1}\}.$$

Theorem 4.3.12. *Assume that Conditions 4.2.2 or 4.2.3 hold and suppose that G is Lipschitz and consistent of order n with respect to discretization. Suppose that $\tau = +\infty$ a.s. under Condition 4.2.2. Then, for $\kappa > 0$ and $m, n \in \mathbb{N}$, there are some positive constants α_κ, C_κ and sequences $\varepsilon_\kappa > 0$, $\varepsilon_{\kappa,m} > 0$, and $\varepsilon_{\kappa,m,n}$ such that*

$$|\mathbf{E}G(S)1_{\tau>T} - \mathbf{E}G(S_{\kappa,m,n}^{(n)})1_{\tilde{\tau}^{\kappa,m,n}>T}| \leq \varepsilon_\kappa + \varepsilon_{\kappa,m} + \frac{\alpha_\kappa(C_\kappa)^m}{n} + \varepsilon_{\kappa,m,n},$$

where $\varepsilon_\kappa \rightarrow 0$ as $\kappa \rightarrow \infty$, $\varepsilon_{\kappa,m} \rightarrow 0$ as $m \rightarrow \infty$, $\varepsilon_{\kappa,m,n} \rightarrow 0$ as $n \rightarrow \infty$. Moreover, $\varepsilon_{\kappa,m,n} = 0$ if $P(\tau \leq T) = 0$.

Remark 4.3.13. *In practice, we first fix κ large enough and we fix m large enough so that $\varepsilon_{\kappa,m}$ is sufficiently small. At last, we choose n large enough. The algorithm's complexity allowing to compute this scheme is then mn . Notice that in the case where $G(S)$ only depends on S_T and $G(0) = 0$, then $\mathbf{E}G(S) = \mathbf{E}G(S)1_{\tau>T}$ which is approximated by $\mathbf{E}G(S_{\kappa,m,n}^{(n)})1_{\tilde{\tau}^{\kappa,m,n}>T}$.*

4.4 Numerical example : the CEV model

The Constant Elasticity of Variance (CEV) model is defined by the dynamics

$$dS_t = \sigma S_t^\beta dW_t, \quad t \in [0, T], \quad (4.4.10)$$

where the parameter β is called the elasticity. We suppose that $\sigma = \sigma_{LN} S_0^{1-\beta}$ where σ_{LN} is the effective lognormal volatility. This implies that the volatility depends on β as in [50], see also [19, 64] for more details on this model.

Consider the financial market model defined by a Bond $B \equiv 1$ and the risky asset price S . We recall that S is a martingale if $\beta \leq 1$. If $\beta > 1$, S is a strict local martingale, i.e. is not a martingale. In both cases, there is no arbitrage opportunity when using predictable strategies generating portfolio processes which are bounded from below, see [17, Section 3.2]. Indeed, any self-financing portfolio process which is bounded from below is a supermartingale.

Let h_T be a European or Asian contingent claim. Suppose that h_T is bounded, which is the case for a Put option. Then, by the predictable representation theorem, the martingale $V_t = \mathbf{E}(h_T | \mathcal{F}_t)$, $t \in [0, T]$, is a self-financing portfolio process such that $V_T = h_T$.

We deduce that the unique price for the payoff h_T is well defined as the expectation $V_0 = \mathbf{E}h_T$. In the following, we illustrate our method by estimating $V_0 = \mathbf{E}h_T$ when $\beta > 1$ and $\beta < 1$ respectively. The price for the Call option may be easily deduced from the Put-Call parity and the analytic expression of the expectations $\mathbf{E}S_t$, $t \in [0, T]$, see [50].

4.4.1 The CEV model with elasticity $\beta > 1$

Let us introduce $\delta = (1 - 2\beta)(1 - \beta)^{-1}$. We suppose that $\delta > 2$ or equivalently $\beta > 1$ so that S is a strictly positive process. The price of the Call option with strike $K > 0$ is then given in [50] by

$$p^{Call}(S_0, K, \sigma, \delta, T) = S_0 \left(\Gamma\left(\frac{\delta}{2} - 1, \frac{S_0}{2T}\right) - \chi'^2\left(\frac{S_0}{T}, \delta - 2, \frac{\tilde{K}}{T}\right) \right) - K \chi'^2\left(\frac{\tilde{K}}{T}, \delta, \frac{S_0}{T}\right),$$

where $\tilde{K} := \frac{K^2(1-\alpha)}{\sigma^2(1-\alpha)^2}$, Γ is the normalised incomplete gamma function

$$\Gamma(z, x) = \frac{1}{\Gamma(z)} \int_0^x t^{z-1} e^{-t} dt,$$

and $\chi'^2(x, k, \lambda)$ is the cumulative density function of the non-central chi-squared distribution with degrees of freedom k and non-centrality parameter λ . The price of the Put option with strike K is deduced as

$$p^{Put}(S_0, K, \sigma, \delta, T) = p^{Call}(S_0, K, \sigma, \delta, T) - S_0 \Gamma\left(1 - \frac{\delta}{2}, \frac{S_0}{2T}\right) + K.$$

In [50], the price of the Put option is computed analytically for several values of $\beta > 1$ when $T = 1$, $S_0 = 100$, $\sigma_{LN} = 20\%$ and $K = 90, 100, 110$. These analytical values are compared to the ones obtained by simulating the terminal value $X_T := \frac{S_T^{2(1-\beta)}}{\sigma^2(1-\beta)^2}$. In the following, we compare these results to the values we obtain when using our method. Actually, we can not expect to obtain the convergence of S_κ towards S in the L^p sense, $p \geq 1$, since Condition 4.3.2 does not a priori hold for $p \geq 1$, see also Condition 4.3.3 (I). Indeed, $\mathbf{E}S_\kappa(t) = S_0$ for all t since S_κ is a square integrable martingale. On the other hand, $\mathbf{E}S_t < S_0$ for all $t > 0$ as shown in [50] since S is a strict local martingale. Therefore, $\mathbf{E}S_\kappa(t)$ does not converge to $\mathbf{E}S_t$ as $\kappa \rightarrow \infty$. Nevertheless, when the payoff function $g \geq 0$ is bounded, the dominated convergence theorem yields that $g(S_\kappa) \rightarrow g(S)$ in the L^p sense as $\kappa \rightarrow \infty$, $p \geq 1$, which is enough for our purpose. Precisely, since $S = S_\kappa$ on $\tau^\kappa \geq T$, we get that

$$\mathbf{E}|g(S_\kappa) - g(S)|^p \leq (\max g)^p \mathbb{P} \left(\sup_{t \in [0, T]} S_t \geq \kappa \right) \rightarrow 0, \text{ as } \kappa \rightarrow \infty.$$

This implies that Theorem 4.3.12 still hold.

Let us compare the Picard method to the analytic one and to the backward Euler–Maruyama scheme (BEM), see [56]. In the following, we choose $m = 3$ and $\sigma_\kappa(u, x) := \sigma(u, x \wedge \kappa)$ where

$$\kappa = S_0 \left(1 + \frac{5}{1 + \beta^2} \right).$$

This choice has been empirically chosen to improve the performance of our results for the Put option prices and then implemented for the Asian options. Recall that the volatility of the model also depends on β . If $\beta = 1.5$, we observed that the approximation error is not smaller if we increase to much κ . The reason is very simple : choosing κ very large is equivalent to suppose that $\kappa = +\infty$ as the maximum of S over all simulated trajectories is finite since we fix a finite number of dates $n + 1$ and a finite number of simulations N . Theoretically, if κ is very large, we need to increase the number of dates n as the Lipschitz constant is larger, which is more time consuming.

We choose $N = 15000$ simulations and $n = 1000$ so that the discretization step is T/n . We consider the case $K = 100$. The presented results are averages we obtain by repeating more than 1000 times the same simulations so that the relative standard deviation of the results is smaller than 5% of the estimated averages.

β	Picard	Time in s.	BEM	Time in s.	Analytic	Err. (P)	Err. (BEM)
1.5	7.8602	6.38	7.8839	0.78	7.9696	1.4%	1.08%
2	7.8995	6.19	7.8930	0.76	7.9789	1.01%	1.08%
2.5	7.9948	6.09	7.9087	0.79	7.9960	0.02%	1.09%
3	8.1332	6.08	8.9329	0.76	8.0211	1.4%	1.10%
3.5	8.2779	6.06	7.9686	0.80	8.0560	2.76%	1.09%
4	8.4006	6.06	8.0200	0.81	8.1033	3.67%	1.03%
4.5	8.4886	6.13	8.0841	0.80	8.1639	3.98%	0.98%
5	8.5348	6.08	8.1583	0.78	8.2345	3.65%	0.93%
5.5	8.5344	5.936	8.2365	0.78	8.3084	2.72%	0.87%
6	8.5015	5.906	8.3086	0.80	8.3786	1.47%	0.84%
6.5	8.4499	6.096	8.3715	0.78	8.4396	0.13%	0.81%
7	8.3821	6.401	8.4228	0.82	8.4882	1.26%	0.77%

FIGURE 4.1 – Put option prices using Picard iterations and the analytic expression respectively for different values of $\beta > 1$.

Naturally, the computation time needed for the analytic method (< 0.01) is negligible contrary to the needed time for other methods which require to simulate all the trajectories.

In the following, we compare our method to the Monte Carlo one of [50] and the BEM method as well (see [56, Section 2, Section 3.4]), requiring that the radius of the 95% confidence interval is less than 0.025 when applying the law of large number. The simple Monte Carlo simulation is less time consuming. Nevertheless, we shall see that the two

CHAPITRE 4. APPROXIMATION OF NON-LIPSCHITZ SDES BY PICARD ITERATIONS

β	Picard	Comput. Time	BEM	Time	MC [32]	Time	Error
1.5	7.961482891	268.748	7.9572	31.73	7.958596698	20.756	0.03%
2	8.001945176	276.356	7.9692	28.28	7.968447144	12.571	0.42%
2.5	8.096442435	286.754	7.9841	26.05	7.981469166	13.68	1.44%
3	8.231139717	298.65	8.0099	23.60	8.002688168	13.098	2.86%
3.5	8.382673524	317.32	8.0483	22.09	8.050143346	10.78	4.14%
4	8.507583601	338.079	8.0948	20.57	8.091346416	7.908	5.15%
4.5	8.594539064	345.449	8.1543	19.06	8.164748506	5.855	5.17%
5	8.631951141	351.436	8.2231	17.13	8.234170007	3.698	4.84%

FIGURE 4.2 – Put option prices using Picard iterations, BEM and Monte Carlo methods respectively for different values of $\beta > 1$ so that the radius of the 95% confidence interval is less than 0.025.

other methods outperform the Monte Carlo simulations as soon as we consider an Asian Put option with several dates.

It is also natural to compare our method to a simple Doléans-Dade scheme deduced from the dynamics (4.4.10) where the volatility function $\sigma(x) = \sigma x^{\beta-1}$ is replaced by $\sigma_\kappa(x) = \sigma(x \wedge \kappa)^{\beta-1}$. Inspired by the Doléans-Dade formula, we consider the Doléans-Dade scheme

$$S_{t_{i+1}}^n = S_{t_i}^n \exp \left\{ \sigma_\kappa(S_{t_i}^n) \Delta W_{t_{i+1}} - \frac{\sigma_\kappa^2(S_{t_i}^n)}{2} \Delta t_{i+1} \right\}, S_0^n = S_0, \quad i = 0, \dots, n-1.$$

In the following table, we compare the approximate prices to the analytic ones for several values of β by implementing the scheme until the error is smaller than the one observed in Table 4.1 by increasing N with the bound $N \leq 5.10^6$. We choose $n = 1000$ so that the time step is T/n . The strike is $K = 100$ and $\kappa = 400$. The error is still computed as the ratio between the difference of the two prices and the smallest one.

β	Price (truncated Doléans-Dade scheme)	Comput. time in seconds	N	Error
1.5	7.1412	535	5.10^6	11.6%
2	7.1978	555	5.10^6	10.85%
2.5	7.2544	553	5.10^6	10.22%
3	7.3134	549	5.10^6	9.68%
3.5	7.3779	556	5.10^6	9.19%
4	7.4524	564	5.10^6	8.73%
4.5	7.5999	541	5.10^6	7.42%
5	8.0123	0.014	136	3.65%
5.5	8.4954	0.013	136	2.72%
6	10.7683	523	5.10^6	28.5%
6.5	11.934	518	5.10^6	41.4%
7	12.9042	518	5.10^6	52%

FIGURE 4.3 – Put option prices with truncated Doléans-Dade scheme for different values of β .

We observe that the results are similar for smaller $\kappa = 120$ as the errors are around 10% and computation times are about 520 seconds. If $\kappa = 800$, we get similar results. The

performance is still good for $\beta = 5$ but also $\beta = 4$ while it is bad for $\beta = 5.5$ contrary to the case $\kappa = 400$. The best results we can get seem to be when $\kappa = +\infty$, i.e. without any truncation, but remains bad. With a larger $n = 10000$ and $N = 15000$, errors go from 2.5% to 10.24% but computational times are larger than 14 seconds. With only $n = 1000$, computation times are smaller, precisely about 1.5 seconds, but the errors are similar i.e. from 2% to 20%. So, seemingly, it is more efficient to increase N rather than n if we aim to decrease the error. In any case, comparing with the results of Table 4.1, it is clear that our methods outperforms the simple truncated Euler scheme.

As to a direct discretization from (4.4.10), i.e. $S_{t_{i+1}}^n = S_{t_i}^n + \Sigma(S_{t_i}^n)\Delta W_{t_{i+1}}$, the Euler scheme with $n = 1000$ and $N = 15000$ explodes without truncation for $\beta \geq 3$ as we may expect. For $\beta \in \{1.5, 2, 2.5\}$, the scheme appears to be efficient as the error is small ($\leq 1.4\%$) and the computation time is also smaller ($\simeq 1.6$ seconds). With the truncation $\Sigma(x) = \sigma(x \wedge \kappa)^\beta$ where $\kappa = 300$, the errors are about 1.5% for the first values of $\beta > 1$ while the errors go to 4.6% for $\beta = 5$ to 769% for $\beta = 7$. Therefore, a truncation from the simple Euler scheme seems to be conceivable only for values of β smaller than 5.

At last, we shall see that our method outperforms the classical one based on Monte Carlo simulations as soon as we consider an Asian Put option of the form $(K - (S_{T_1} - \dots - S_{T_J})/J)^+$ where $J \geq 3$ and $T_1 < T_2 < \dots < T_J = T$.

Let us first consider the case $J = 2$. In that case, by conditioning, we rewrite $\mathbf{E}(2K - S_{T_1} - S_{T_2})^+ = \mathbf{E}P(S_{T_1})$ where, by the Markov property, $P(s)$ is the price of the Put option of strike $2K - s$ in the CEV model with initial value s and horizon date $T_2 - T_1$. As shown in [50], it is possible to directly simulate the distribution of S_{T_1} . We then analytically deduce the price $P(S_{T_1})$ for each simulated value of S_{T_1} . By the law of large numbers, we then deduce an approximation of $\mathbf{E}P(S_{T_1})$. With the same parameters as before, we obtain the following results when $T_1 = T/2$ and $T_2 = T = 1$.

CHAPITRE 4. APPROXIMATION OF NON-LIPSCHITZ SDES BY PICARD ITERATIONS

β	P	Time	BEM (B)	Time	MC	Time	Error P-C	Euler (E)	Error E-C	Error B-C
1.5	6.2284	6.02	6.3342	0.52	6.3147	0.328	1.39%	6.3983	1.32%	0.31%
2	6.2549	6.4	6.3379	0.52	6.3198	0.133	1.04%	6.4052	1.35%	0.29%
2.5	6.3142	6.178	6.3466	0.53	6.3297	0.121	0.25%	6.4165	1.37%	0.27%
3	6.4027	6.159	6.3603	0.52	6.3438	0.113	0.93%	6.4321	1.39%	0.26%
3.5	6.5011	6.398	6.3780	0.53	6.3625	0.118	2.18%	6.4520	1.41%	0.24%
4	6.5950	6.235	6.3998	0.53	6.3864	0.101	3.27%	6.4772	1.42%	0.21%
4.5	6.6737	6.25	6.4269	0.53	6.4166	0.098	4%	6.5162	1.55%	0.16%
5	6.7267	6.188	6.4623	0.52	6.4544	0.101	4.22%	6.7340	4.33%	0.12%
5.5	6.7465	5.967	6.5060	0.53	6.5003	0.102	3.79%	7.5778	16.58%	0.09%
6	6.7411	5.968	6.5571	0.53	6.5530	0.102	2.87%	12.7244	94.18%	0.06%
6.5	6.7192	5.997	6.6099	0.53	6.6102	0.100	1.65%	25.4794	285.46%	0.00%
7	6.6820	6.000	6.6638	0.53	6.6687	0.100	0.20%	51.2649	668.74	0.07%

FIGURE 4.4 – Asian Put option prices with two dates using Picard iterations (P), the Monte Carlo methods combined with the analytic expression (MC) and a truncated Euler scheme (E) respectively. The errors are computed between the Picard method (P) and the Euler scheme (E) respectively with respect to the combined method (C).

β	P	Time in s.	BEM (B)	Time	MC	Time	Error P-C	Error B-C
1.5	6.293809919	171.191	6.2947	20.49	6.307557882	10.375	0.22%	0.20%
2	6.314878756	173.881	6.3012	18.93	6.311235573	3.961	0.06%	0.16%
2.5	6.373160059	179.582	6.3126	17.68	6.320525289	3.429	0.84%	0.12%
3	6.460785842	186.615	6.3216	16.33	6.331600424	3.015	2.04%	0.16%
3.5	6.559108888	198.54	6.3398	15.48	6.347881289	2.524	3.33%	0.13%
4	6.655375191	205.6	6.3670	14.41	6.366022368	2.317	4.55%	-0.02%
4.5	6.734410841	216.076	6.3994	13.56	6.390785176	2.13	5.38%	-0.13%
5	6.781869529	219.877	6.4358	12.62	6.42642745	2.009	5.54%	-0.15%

FIGURE 4.5 – Asian Put option prices with two dates using Picard iterations and the Monte Carlo methods combined with the analytic expression respectively such that the 95% confidence interval radius is smaller than 0.025.

As we may observe, it is still advantageous to combine the Monte Carlo method with the analytic one in the case of an Asian option with only two dates. Moreover, we still observe that the simple truncated Euler scheme is efficient for $\beta \leq 5$ while it produces a large convergence error if $\beta > 5$ contrary to the Picard iterations.

Let us now suppose that $J = 3$. By the Markov property, we rewrite after conditioning $\mathbf{E}(3K - S_{T_1} - S_{T_2} - S_{T_3})^+ = \mathbf{E}P(S_{T_1})$ where $P(s)$ is the price of the Asian Put option of strike $3K - s$ in the CEV model with initial value s , intermediate date $T_2 - T_1$ and horizon date $T_3 - T_1$. By the Monte Carlo method, we simulate S_{T_1} so that we deduce $P(S_{T_1})$ using the pricer for two dates we have introduced before. In that case, the complexity explodes as we need to run two successive Monte Carlo schemes. With the same parameters but only with $\beta = 1.5$ or $\beta = 7$ and $T_1 = T/3$, $T_2 = 2T/3$ and $T_3 = T = 1$, we get the following results :

Conclusion : our method appears to outperform the technique based on the Monte Carlo methods combined with the analytic expression for Asian options with more than

CHAPITRE 4. APPROXIMATION OF NON-LIPSCHITZ SDES BY PICARD ITERATIONS

β	P	Time	BEM	Time	MC	Time	Error P-C	Euler (E)	Error E-C	Error B-C
1.5	5.3788	61.73	5.6768	7.87	5.2698	8333	2.07%	5.8172	10.39%	7.72%
7	5.6278	61.83	5.9535	8.36	5.5188	2306	1.98%	10.2290	85.35%	7.88 %

FIGURE 4.6 – Asian Put option price with three dates when $\beta = 1.5$ or $\beta = 7$ using Picard iterations, the Monte Carlo methods combined with the analytic expression and the truncated Euler scheme respectively. The errors are computed between the Picard method (P) and the Euler scheme (E) respectively with respect to the combined method (C).

β	P	Time	BEM	Time	CI	MC	Time	Error P-C	Error B-C
1.5	5.345	1322	5.7254	17.724	0.05	5.26	372089	1.62%	8.85%

FIGURE 4.7 – Asian Put option price with three dates using Picard iterations and the Monte Carlo methods combined with the analytic expression respectively such that the 95% confidence interval radius is smaller than 0.05.

three dates as well as the truncated Euler scheme. Notice that computational times may be shortened in the Picard scheme by replacing the time-consuming computation of the exponential function by the first terms of its Taylor expansion. Nevertheless, the BEM method appears to be the most efficient even if, with three dates, the error in Table 4.7 is smaller for the Picard method. In the case where $\beta = 2$, the simple truncated Euler scheme seems to be efficient provided that $\beta \leq 5$ as convergence errors and computation times are small. In the following section, we implement the CEV model for $\beta < 1$. In that case, the BEM method does not work well.

4.4.2 The CEV model with elasticity $\beta \in (0, 1)$

Recall that Condition 4.2.3 holds when $\beta \in (0, 1)$. Nevertheless, the required property that there exists an absorbing point only holds when $\beta \geq 0.5$ [50]. In this section, the parameters are $S_0 = 100$, $\sigma_{LN} = 0.5$, $T = 4$, $K = 100$, $N = 15000$, $n = 20$ and $m = 2$. The parameter κ is empirically fixed to $\kappa = S_0 (0.78 - 0.4 * \beta^2)$ for the Picard scheme and $\kappa = 300$ for the truncated Euler scheme.

When implementing the BEM method, the implicit scheme requires to solve an equation of the form $g(X) = 0$, e.g. using the Newton method. To do so, we fix an interval $[a, b]$ we expect the solution X^* to belong to. Unfortunately, the algorithm stops as soon as a is too large or b too small as $[a, b]$ does not contain X^* . We observe this phenomenon even with $a = 10^{-8}$ and $b = 10^8$ so that we can not obtain any result.

We consider the Put option.

CHAPITRE 4. APPROXIMATION OF NON-LIPSCHITZ SDES BY PICARD ITERATIONS

The methodology we follow consists in increasing the number of simulations N until the estimated price belongs to the interval of radius 1% and centred at the analytic expression of the price. Otherwise, we stop N to 2.10^5 . Repeating 1000 times the procedure, we deduce an average of the estimated price for each method and the needed time in milliseconds as well.

β	Truncated Euler scheme	Time (ms)	Analytic	Time (ms)	Error	# of simulations
0	37.5155	18.254	39.0452	< 0.01	3.92%	8618
0.1	38.8543	4.662	39.0089	< 0.01	0.40%	1891
0.2	38.9200	1.806	38.9307	< 0.01	0.03%	745
0.3	38.8460	1.106	38.8210	< 0.01	0.06%	460
0.4	38.7139	1.164	38.6962	< 0.01	0.05%	487
0.5	38.5853	1.099	38.5753	< 0.01	0.03%	456
0.6	38.4624	2.287	38.4724	< 0.01	0.03%	930
0.7	38.2547	4.537	38.3928	< 0.01	0.36%	1866
0.8	37.4803	12.511	38.3368	< 0.01	2.23%	5198
0.9	36.9813	17.533	38.3041	< 0.01	3.45%	7479

FIGURE 4.8 – Put option prices using truncated Euler Scheme and the analytic expression respectively for different values of $\beta < 1$.

Let us now present the approximated prices obtained by the Picard method with $n = 20$.

β	Picard	Comp. Time (ms)	Analytic	Error	# of simulations
0	39.0421	1.299	39.0451	0.01%	218
0.1	38.9987	1.608	39.0088	0.03%	269
0.2	38.9235	1.759	38.9306	0.02%	292
0.3	38.8090	1.331	38.8209	0.03%	219
0.4	38.7007	1.477	38.6961	0.01%	239
0.5	38.5871	2.22	38.5753	0.03%	367
0.6	38.5036	3.112	38.4724	0.08%	520
0.7	38.4055	2.136	38.3928	0.03%	357
0.8	38.3374	1.336	38.3368	0.00%	216
0.9	38.1972	6.028	38.3041	0.28%	991

FIGURE 4.9 – Put option prices using Picard iterations and the analytic expression respectively for different values of $\beta < 1$.

As we may observe, the Picard method outperforms the truncated euler method, i.e. provides smaller error with less computation time. For two dates Asian options we choose $n = 40$ in order to work with the same mesh.

CHAPITRE 4. APPROXIMATION OF NON-LIPSCHITZ SDES BY PICARD ITERATIONS

β	(P)	Time (ms)	MC (C)	Time (ms)	Error P-C	Euler (E)	Time (ms)	Error E-C
0,5	28,8773	266,07	32,5530	0,103	12,73%	28,0664	171,11	15,99%
0,6	31,3302	11,391	31,3744	0,124	0,14%	30,9665	18,045	1,32%
0,7	30,7832	5,097	30,7640	0,137	0,06%	30,7794	2,305	0,05%
0,8	30,6667	2,84	30,6624	0,149	0,01%	30,6813	3,402	0,06%
0,9	30,6594	4,208	30,6532	0,194	0,02%	30,6334	3,826	0,06%

FIGURE 4.10 – Asian Put option prices with two dates using Picard iterations (P), truncated Euler scheme (E) and the Monte Carlo methods combined with the analytic expression (C). The errors are computed between the Picard method (P) and the Euler scheme (E) respectively with respect to the combined method (C).

Clearly, the combined method derived from [50] remains the most performant for two dates. Notice that the Picard method outperforms the truncated one. At last, **with the combined method, it takes several hours** to estimate the price of an Asian option with three dates $Ti/3$, $i = 1, 2, 3$ and $T = 4$. It only takes 78 seconds for the Picard method and the estimated price is 27.8456.

Conclusion : in the case where $\beta \in [0, 1)$, the Picard method seems to be the best method comparatively to Euler truncated scheme, the BEM method or the combined method based on Monte Carlo simulations of the terminal price for more than three dates as the computation time is smaller and the errors are satisfactory. Otherwise, for one or two dates, the combined method derived from [50] appears to be efficient.

4.5 The Aït-Sahalia model

The Aït-Sahalia model (see [56, Section 3.6]) satisfies the s.d.e.

$$dS_t = \sigma S_t^\rho dW_t + (\alpha_{-1} S_t^{-1} - \alpha_0 + \alpha_1 S_t - \alpha_2 S_t^r) dt,$$

where $\rho, r > 1$, $\alpha_i \geq 0$, $i = -1, \dots, 2$ and $\sigma > 0$. The initial value is $S_0 = 0.2$ and the other parameters are $T = 1$, $K = 0$ for the Call option so that we shall estimate $E(S_T)$. We choose $\alpha_{-1} = 0.1$ and $\alpha_i = 0$, if $i = 0, \dots, 2$. We fix $\sigma = 0.2$ and $r = \rho = 2$. The BEM method provides $E(S_T) \simeq 0.4906674$ in 6.689 m.s. The simple Euler scheme compute the value $E(S_T) \simeq 0.49050227$ in 1.227 m.s. while the Picard method estimates $E(S_T) \simeq 0.4905902$ in 19.517 m.s. for $m = 10$. With $m = 5$, it takes 9.832 m.s. but we get $E(S_T) \simeq 0.5041992$.

If we choose the parameters $\alpha_{-1} = 0.1$, $\alpha_0 = 1$, $\alpha_1 = 5$ and $\alpha_2 = 2$, by the BEM method, we get $E(S_T) \simeq 1.59841981$ in 7.531 m.s. while the simple Euler scheme provides $E(S_T) \simeq 1.6022864$ in 1.245 m.s. At last, it takes 18.747 m.s. for the Picard method

to estimate $E(S_T) \simeq 1.6042199$ with $m = 10$. On these examples, we see that the simple Euler scheme provides a rather accurate estimation of ES_T in comparison to the evaluation given by the BEM method that requires more time. Clearly, the Picard method is not efficient on this model in the sense that the needed time to obtain a small error is significantly larger.

4.6 Example of mean reverting SDEs

Let us consider the dynamics $dS_t = (1-S_t)dt + \gamma S_t^\beta dW_t$, where W is a standard Brownian motion and $\beta \in [0.5, 1)$. Here we choose $\beta = 1$. Using the Lamperti transformation, we simulate the process $X_t = S_t^{-\beta+1}$ when implementing the BEM method. We noticed that the BEM scheme explodes because of the inversion problem that needs to be solved and the simple Euler scheme explodes in particular when β is closed to 0.5. Otherwise, an improved convergent Milstein scheme is proposed by [22]. As mentioned in [22], the discretization step needs to be smaller than γ^{-2} for the approximation process to be positive. This is indeed what we observed so that the needed time to compute appears to be significantly large for γ large enough. In that case, the Picard method appears to be more efficient. Here we compare the two methods with $\gamma = 1$ so that it is possible to get accurate estimations with the improved Milstein scheme in a reasonable computation time. The goal here is to present the accuracy of the Picard method. With the discrete dates $t_k^n = (T/n)k$, $k = 0, \dots, n$, the improved Milstein is given by :

$$\begin{aligned} S_{t_{k+1}}^n &= S_{t_k}^n + (1 - S_{t_{k+1}}^n)\Delta t_{k+1} + [S_{t_k}^n]^\beta \Delta W_{t_k}^n \\ &\quad + \frac{1}{2}(\beta - 1)[S_{t_k}^n]^{2\beta-1} ((\Delta W_{t_k})^2 - \Delta t_k). \end{aligned}$$

We propose to compare the results obtained by implementing the Picard method without truncature, i.e. $\kappa = +\infty$, to the Milstein scheme as a reference. To do so, we first implement the Milstein scheme. We then compute the averaged needed time over 1000 experiments for the Picard method to provide an approximation at a relative distance from the Milstein estimation less than 1%. As shown in the following table, the estimations are very accurate.

β	Milstein	Picard	Error	Time (s)
0.5	0.9968	0.9980	0.12%	1.103
0.55	0.9969	0.9974	0.05%	0.631
0.6	0.9969	0.9977	0.07%	0.633
0.65	0.9969	0.9976	0.07%	0.643
0.7	0.9968	0.9977	0.09%	0.656
0.75	0.9967	0.9979	0.12%	0.760
0.8	0.9965	0.9972	0.07%	0.835
0.85	0.9961	0.9974	0.13%	0.886
0.9	0.9955	0.9965	0.10%	0.979
0.95	0.9946	0.9960	0.14%	1.129
1	0.9932	0.9945	0.12%	0.853

FIGURE 4.11 – Estimation of ES_T with $T = 6$ and $S_0 = 1$ for several values of β .

4.7 Proofs

Throughout the paper, we shall use the following convention : from a line to the next one, constants K, κ or C may designate different constants which are independent of any other current variables except maybe the fixed parameters of the model as the maturity date T or Lipschitz constants. In the contrary case, we shall use notations like C_m to stress the dependence on some parameter m .

Proof of Lemma 4.2.6. Let us denote $S = S_{r,\sigma} = F(r)\mathcal{E}(\sigma)$. Using the integration by part formula, we obtain that

$$S_t t = \int_0^t S_u du + \int_0^t u dS_u, \quad t \in [0, T].$$

On the other hand, a direct computation shows that

$$\int_0^T S_u^{(n)} du = - \int_0^T u dS_u^{(n)} + TS_T^{(n)}.$$

Since $S_T^{(n)} = S_T$, we deduce that

$$\int_0^T S_u du - \int_0^T S_u^{(n)} du = \int_0^T u d(S_u^{(n)} - S_u).$$

Observe that

$$\begin{aligned} \int_0^T u dS_u^{(n)} &= \sum_{i=1}^n t_i^n \Delta S_{t_i^n}, \\ &= \sum_{i=1}^n t_i^n \int_{t_{i-1}^n}^{t_i^n} \sigma(u, S_u) S_u dW_u + \sum_{i=1}^n t_i^n \int_{t_{i-1}^n}^{t_i^n} r(u, S_u) S_u du, \\ &= \int_0^T I^{(n)}(u) \sigma(u, S_u) S_u dW_u + \int_0^T I^{(n)}(u) r(u, S_u) S_u du, \end{aligned}$$

where $I^{(n)}(u) = t_i$ if $u \in [t_{i-1}, t_i)$. It follows that

$$\int_0^T S_u du - \int_0^T S_u^{(n)} du = \int_0^T (I^{(n)}(u) - u) \sigma(u, S_u) S_u dW_u + \int_0^T (I^{(n)}(u) - u) r(u, S_u) S_u du$$

As $|u^{(n)}(u) - u| \leq T/n$ and r and σ are supposed to be Lipschitz, we deduce by the Cauchy-Schwarz inequality and the Ito isometry the following sequence of inequalities :

$$\begin{aligned}
\mathbf{E}|G(S) - G(S^{(n)})| &\leq C_T \mathbf{E} \left| \int_0^T (I^{(n)}(u) - u) \sigma(u, S_u) S_u dW_u \right| \\
&\quad + C_T \mathbf{E} \left| \int_0^T (I^{(n)}(u) - u) r(u, S_u) S_u du \right|, \\
\mathbf{E}|G(S) - G(S^{(n)})| &\leq C_T \sqrt{C(\sigma) \int_0^T (I^{(n)}(u) - u)^2 \mathbf{E}(1 + S_u^4) du} \\
&\quad + C_T(r) \int_0^T |I^{(n)}(u) - u| \mathbf{E}(1 + S_u^2) du, \\
&\leq \frac{C_G \sqrt{C(\sigma)}}{n} \sqrt{\int_0^T (1 + \mathbf{E}S_u^4) du} \\
&\quad + \frac{C_T(r)}{n} \int_0^T \mathbf{E}(1 + S_u^2) du, \\
&\leq \frac{C_G \sqrt{C(\sigma)}}{n} \sqrt{\int_0^T (1 + \mathbf{E}S_u^4) du} \\
&\quad + \frac{C_T(r)}{n} \sqrt{\int_0^T \mathbf{E}(1 + S_u^2)^2 du}.
\end{aligned}$$

The conclusion follows. \square

Lemma 4.7.1. *Suppose that Condition 4.2.3 holds. Let $\kappa > 0$. Then, there exists a constant C_κ depending on κ such that for all $x, y > 0$ and $t \in \mathbf{R}_+$,*

$$|\sigma_\kappa(t, x) - \sigma_\kappa(t, y)| \leq C_\kappa \sqrt{|x - y|}.$$

Proof. Under Conditions 4.2.3, we obtain the following chain of inequalities

$$\begin{aligned}
|\sigma_\kappa(t, x) - \sigma_\kappa(t, y)| &= |\sigma_\kappa(t, x \vee \kappa^{-1}) - \sigma_\kappa(t, y \vee \kappa^{-1})|, \\
&\leq C_\kappa |\ln(x \vee \kappa^{-1}) - \ln(y \vee \kappa^{-1})|, \\
&\leq C_\kappa |\ln(x) \vee \ln(\kappa^{-1}) - \ln(y) \vee \ln(\kappa^{-1})|, \\
&\leq 2C_\kappa |\ln(\sqrt{x}) \vee \ln(\kappa^{-1/2}) - \ln(\sqrt{y}) \vee \ln(\kappa^{-1/2})|, \\
&\leq C_\kappa |\sqrt{x} - \sqrt{y}| \leq C_\kappa \sqrt{|x - y|}.
\end{aligned}$$

The last inequalities are due to the fact that the mapping $z \mapsto \ln(z) \vee \ln(\kappa^{-1/2})$ is Lipschitz and the square root function is Holder continuous of order 1/2. \square

Lemma 4.7.2. *Suppose that Condition 4.2.2 holds. Let $\kappa > 0$. Then, there exists a constant C_κ depending on κ such that for all $X, Y \in \mathbf{R}^2$, $x, y > 0$ and $t \in \mathbf{R}_+$,*

$$\begin{aligned} |\sigma_\kappa(t, e^X) - \sigma_\kappa(t, e^Y)| &\leq C_\kappa |X - Y|, \\ |\sigma_\kappa(t, x) - \sigma_\kappa(t, y)| &\leq C_\kappa \sqrt{|x - y|}. \end{aligned}$$

Proof.

Using the assumption and the fact that the mapping $X \mapsto e^X \wedge \kappa$ is Lipschitz, we get that

$$\begin{aligned} |\sigma_\kappa(t, e^X) - \sigma_\kappa(t, e^Y)| &= |\sigma_\kappa(t, e^X \wedge \kappa) - \sigma_\kappa(t, e^Y \wedge \kappa)|, \\ &\leq C_\kappa |e^X \wedge \kappa - e^Y \wedge \kappa| \leq C_\kappa |X - Y|. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} |\sigma_\kappa(t, x) - \sigma_\kappa(t, y)| &= |\sigma_\kappa(t, x \wedge \kappa) - \sigma_\kappa(t, y \wedge \kappa)|, \\ &\leq C_\kappa |x \wedge \kappa - y \wedge \kappa| \leq C_\kappa \sqrt{|x \wedge \kappa - y \wedge \kappa|}, \\ &\leq C_\kappa \sqrt{|x - y|}. \end{aligned}$$

The last inequalities hold since $\sqrt{|x \wedge \kappa - y \wedge \kappa|}$ is bounded by $\sqrt{2\kappa}$ and the mapping $x \mapsto x \wedge \kappa$ is Lipschitz. \square

Proof of Lemma 4.3.1. Suppose that Condition 4.2.2 holds. Consider the sequence of stopping times

$$\tau^\kappa := \inf\{t : S_t \geq \kappa\}. \quad (4.7.11)$$

Since the stopped processes S^{τ^κ} and $(S_\kappa)^{\tau^\kappa}$ satisfy the same s.d.e., we deduce that $S_t = S_\kappa(t)$ on $t \in [0, \tau^\kappa]$. As $\tau^\kappa \rightarrow \infty$, we then conclude. Under Condition 4.2.3, we follow the same arguments with

$$\tau^\kappa := \inf\{t : S_t \leq \kappa^{-1}\}. \quad (4.7.12)$$

Note that $\tau^\kappa \rightarrow \tau$ as $\kappa \rightarrow \infty$. \square

Lemma 4.7.3. *Assume that Condition 4.3.3 (I) is satisfied. Then, there exists a constant C independent of κ such that $\sup_{u \leq T} \mathbf{E}S_\kappa(u) \leq C$.*

Proof. By definition, we have $S_\kappa(u)/F_u(r_\kappa(\cdot, S_\kappa(\cdot))) \leq M_u$ where $M_0 = 1$ and M is a local martingale which is solution to the s.d.e.

$$dM_u = M_u \sigma_\kappa(u, S_\kappa(u)) dW_u.$$

CHAPITRE 4. APPROXIMATION OF NON-LIPSCHITZ SDES BY PICARD ITERATIONS

As r is supposed to be bounded, we deduce that $S_\kappa(u) \leq CS_0M_u$. Let (τ^n) be a sequence of stopping times such that M^{τ^n} is a true martingale. Then, $\mathbf{E}S_\kappa(u \wedge \tau^n) \leq CS_0$ and finally $\mathbf{E}S_\kappa(u) \leq CS_0$ by virtue of by Fatou's lemma. \square

Lemma 4.7.4. *Assume that Assumption 4.3.3 (I) is satisfied. Then, for all p such that $p \leq 0$ or $p \geq 1/2$, there exists a constant C_p independent of κ such that $\sup_{u \leq T} \mathbf{E}(S_\kappa(u))^p \leq C_p$.*

Proof. Since r is bounded, the following inequality holds for all fixed $q \geq 0$:

$$(S_\kappa(u))^p \leq C_q S_0^p N_u e^{\frac{1}{2}(q-p) \int_0^u \sigma_\kappa^2(r, S_\kappa(r)) dr},$$

where

$$N_u := \exp\left\{p \int_0^u \sigma_\kappa(r, S_\kappa(r)) dW_r - \frac{1}{2}q \int_0^u \sigma_\kappa^2(r, S_\kappa(r)) dr\right\}.$$

Using the inequality $0 \leq ab \leq (a^2 + b^2)$, we get that

$$(S_\kappa(u))^p \leq CS_0^p \tilde{N}_u + CS_0^p e^{(q-p) \int_0^u \sigma^2(r, S_\kappa(r) \wedge \kappa) dr},$$

where $\tilde{N} = N^2$ is a local martingale if we choose $q = 2p^2$. Moreover, as the function $x \mapsto e^{(q-p)ux}$ is convex, the Jensen inequality and the hypothesis yields that

$$\begin{aligned} e^{(q-p) \int_0^u \sigma^2(r, S_\kappa(r) \wedge \kappa) dr} &\leq \frac{1}{u} \int_0^u e^{(q-p)r\sigma^2(r, S_\kappa(r) \wedge \kappa)} dr \\ &\leq C_p + \frac{C}{u} \int_0^u \left((\log(S_\kappa(r)))^{k(q-p)T} 1_{S_\kappa(r) \wedge \kappa \geq 1} + 1_{S_\kappa(r) \wedge \kappa \leq 1} \right) dr, \end{aligned}$$

where k is a constant. Using the property

$$(\log(x \vee 1))^{k(q-p)T} \leq C_p x, \quad \forall x \geq 0,$$

and Lemma 4.7.3, we deduce that $E(S_\kappa(u))^p \leq C_p$. \square

Lemma 4.7.5. *Assume that Condition 4.3.3 (II) is satisfied. Then, for all p such that $p \geq 0$, there exists a constant C_p independent of κ such that $\sup_{u \leq T} \mathbf{E}(S_\kappa(u))^p \leq C_p$.*

Proof. Let us consider $N = N(p) \in \mathbf{N}$ such that $0 \leq p \leq 2N\alpha$. Observe that, for $p \geq 0$, $0 \leq (S_\kappa(u))^p \leq 1 + (S_\kappa(u))^{2N\alpha} 1_{S_\kappa(u) \geq 1} \leq 1 + (S_\kappa(u))^{2N\alpha}$. Therefore, it is enough to show the lemma when $p = 2N\alpha$. We show by induction that the statement holds for every $p = 2j\alpha$, $0 \leq j \leq N$. This is trivial if $j = 0$. Applying the Ito formula when $q = 2(j+1)\alpha$, $j \geq 0$, we get that

$$\begin{aligned} (S_\kappa(u))^q &= (S_0)^q + \int_0^u q(S_\kappa(r))^q \sigma_\kappa(r, S_\kappa(r)) dW_r + \int_0^r q(S_\kappa(r))^q r_\kappa(r, S_\kappa(r)) dr, \\ &\quad + \frac{1}{2}q(q-1) \int_0^u (S_\kappa(r))^q \sigma_\kappa^2(r, S_\kappa(r))(r) dr. \end{aligned}$$

Recall that σ_κ is bounded. Using the Doleans-Dade representation of $S_\kappa(u)$, we deduce that $S_\kappa(u)$ admits finite moments at any order. It follows that the Ito integral in the inequality above is a square integrable martingale. Therefore,

$$\begin{aligned} \mathbf{E}(S_\kappa(u))^q &= (S_0)^q + \int_0^u q \mathbf{E}(S_\kappa(r))^q r_\kappa(r, S_\kappa(r)) dr, \\ &\quad + \frac{1}{2} q(q-1) \int_0^u \mathbf{E}(S_\kappa(r))^q \sigma_\kappa^2(r, S_\kappa(r)) dr. \end{aligned}$$

We then use the inequality $r \leq 1 + r^2$ and by Assumption 4.3.3 (II), we deduce that $g^{j+1}(u) := \mathbf{E}(S_\kappa(u))^{2(j+1)\alpha}$ satisfies

$$\begin{aligned} g^{j+1}(u) &\leq \max_{j \leq N} (S_0)^{2(j+1)\alpha} + C(N^2\alpha^2 + N\alpha) \int_0^u g^j(r) dr \\ &\quad + C(N\alpha + N^2\alpha^2) \int_0^u g^{j+1}(r) dr \end{aligned}$$

Suppose that $g^j \leq C^j$ where C^j does not depend on κ . Then, by Gronwall's lemma, we deduce that $g^{j+1} \leq C^{j+1}$ where C^{j+1} does not depend on κ . The conclusion follows with $j = N$ since the number of steps $N(p)$ is finite and only depends on p . \square

Corollary 4.7.6. *Assume that Condition 4.3.3 is satisfied. Then, for all $p \geq 0$, there exists a constant C_p independent of κ such that $\mathbf{E} \sup_{u \leq T} (S_\kappa(u))^p \leq C_p$.*

Proof. Since $\|\cdot\|_{L^{p_1}} \leq \|\cdot\|_{L^{p_2}}$ if $1 \leq p_1 \leq p_2$, and $(S_\kappa(u))^p \leq 1 + S_\kappa(u)$ when $p \in [0, 1]$, we may assume without loss of generality that $p \in 2\mathbb{N}$. Let us first suppose that Condition (I) holds. We recall that

$$S_\kappa(t) = S_0 + \int_0^t \sigma(u, S_\kappa(u) \wedge \kappa) S_\kappa(u) dW_u + \int_0^t r(u, S_\kappa(u) \wedge \kappa) S_\kappa(u) du.$$

Since $0 \leq \sigma^2(t, x) \leq C(1+x)$ and r is bounded, we easily deduce that $E \sup_{u \leq T} (S_\kappa(u))^p \leq C_p$ by the Burkholder-Davis-Gundy and Jensen inequalities as well as Lemma 4.7.4.

Let us now suppose that Condition (II) holds. In this case,

$$0 \leq S_\kappa^p(u) (\sigma^p(u, S_\kappa(u) \vee \kappa^{-1}) + r^p(u, S_\kappa(u) \vee \kappa^{-1})) \leq C(S_\kappa^p(u) + S_\kappa^{(1-\alpha)p}(u)).$$

Since $1 - \alpha \geq 0$, we use Lemma 4.7.5 and conclude as previously. \square

Since S_κ converges pointwise on $[0, \tau]$ to S , $S = S_\kappa$ on $[0, \tau^\kappa]$ where $\tau^\kappa \rightarrow \tau$, and $S_t = 0$ on $t \geq \tau$, we apply Fatou's Lemma and deduce the following.

Corollary 4.7.7. *Assume that Condition 4.3.3 is satisfied. Then, for all $p \geq 0$, there exists a constant C such that $E \sup_{u \leq T} (S_u)^p \leq C_p$.*

Remark 4.7.8. Recall that

$$dY_t = \sigma(t, S_t)dW_t + r(t, S_t)dt - \frac{1}{2}\sigma^2(t, S_t)dt.$$

Under Condition 4.3.3 (I), since $0 \leq \sigma^2(t, x) \leq C(1+x)$, we then deduce that $E \sup_{u \leq T} Y_u^2 < \infty$ as the coefficient r is bounded. In particular, the process $\int_0^\cdot \sigma(t, S_t)dW_t$ is a true martingale.

Proof of Lemma 4.3.6. If $\sigma_\kappa = \sigma(\cdot, \cdot \wedge \kappa)$, we consider the stopping time τ^κ defined by (4.7.11). We obtain that

$$\begin{aligned} \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p &= \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \mathbf{1}_{\tau^\kappa < t} \\ &\leq \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \mathbf{1}_{\tau^\kappa < T}. \end{aligned}$$

In the case where τ^κ is defined by (4.7.11), we deduce that

$$\begin{aligned} \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p &\leq \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \mathbf{1}_{\sup_{t \in [0, T]} S_t \geq \kappa} \\ &\leq \sqrt{\mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^{2p}} \sqrt{\frac{\mathbf{E} (\sup_{t \in [0, T]} S_t)^{2l}}{\kappa^{2l}}}. \end{aligned}$$

The conclusion follows. In the case where $\sigma_\kappa = \sigma(\cdot, \cdot \vee \kappa^{-1})$, let τ^κ be defined by (4.7.12). Then,

$$\begin{aligned} \mathbf{E} \sup_{t \in [0, \tau \wedge T]} |S_\kappa(t) - S_t|^p &= \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \mathbf{1}_{\tau^\kappa < t \leq \tau \wedge T} \\ &\leq \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \mathbf{1}_{\tau^\kappa < t \leq \tau} \mathbf{1}_{\tau \leq T} \\ &\quad + \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \mathbf{1}_{\inf_{t \in [0, T]} S_t \leq \kappa^{-1}} \mathbf{1}_{\tau > T}. \end{aligned}$$

Notice that the second term of the last inequality above is bounded as follows :

$$\mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p \mathbf{1}_{\inf_{t \in [0, T]} S_t \leq \kappa^{-1}} \mathbf{1}_{\tau > T} \leq \sqrt{\mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^{2p}} \gamma(\tau, \kappa),$$

where $\gamma(\tau, \kappa) := P(\inf_{t \in [0, T]} S_t \leq \kappa^{-1}; \tau > T) \rightarrow 0$ as $\kappa \rightarrow \infty$. If $t \in [\tau^\kappa, \tau]$, where $\tau < T$, then

$$\begin{aligned} S_\kappa(t) &= \kappa^{-1} + \int_{\tau^\kappa}^t S_\kappa(u) \sigma_\kappa(u, S_\kappa(u)) dW_u + \int_{\tau^\kappa}^t S_\kappa(u) r_\kappa(u, S_\kappa(u)) du, \\ S_t &= \kappa^{-1} + \int_{\tau^\kappa}^t S(u) \sigma(u, S(u)) dW_u + \int_{\tau^\kappa}^t S(u) r(u, S(u)) du. \end{aligned}$$

Under Condition (II) of Condition 4.3.3, we have

$$\begin{aligned} |S_\kappa(u)\sigma_\kappa(u, S_\kappa(u))| + |S_\kappa(u)r_\kappa(u, S_\kappa(u))| &\leq C(|S_\kappa(u)| + |S_\kappa(u)|^{1-\alpha}), \\ |S_\kappa(u)\sigma_\kappa(u, S_\kappa(u))| + |S_\kappa(u)r_\kappa(u, S_\kappa(u))| &\leq C(|S_\kappa(u)| + |S_\kappa(u)|^{1-\alpha}). \end{aligned}$$

Using the Burkholder-Davis-Gundy inequalities and the Cauchy-Schwarz inequality, we deduce that the term $\delta(\tau, \kappa) = \mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t) - S_t|^p 1_{\tau^\kappa < t < \tau} 1_{\tau \leq T}$ is bounded as follows :

$$\delta(\tau, \kappa) \leq C \left(\sup_\kappa \sqrt{\mathbf{E} \sup_{t \in [0, T]} |S_\kappa(t)|^{2p}} + \sqrt{\mathbf{E} \sup_{t \in [0, T]} |S(t)|^{2p}} \right) \sqrt{\mathbf{E} |\tau^\kappa \wedge T - \tau \wedge T|^p}.$$

By the Lebesgue theorem, since $\tau^\kappa \rightarrow \tau$, we deduce that $\mathbf{E} |\tau^\kappa \wedge T - \tau \wedge T|^p \rightarrow 0$ as $\kappa \rightarrow \infty$. The conclusion follows. \square

Lemma 4.7.9. *Let $(g^m)_m$ be a sequence of positive functions defined on an interval $[0, T]$, $T > 0$, such that for some $C > 0$, we have :*

$$g^{m+1}(t) \leq C \int_0^t g^m(u) du, \quad 0 \leq g^0 \leq C.$$

Then

$$\sup_{t \in [0, T]} g^m(t) \leq C^m \frac{t^m}{m!} \sup_{t \in [0, T]} g^0(t).$$

Proof. The proof relies on the following recursive reasoning. Set

$$C_g := \sup_{t \in [0, T]} g^0(t) \leq C.$$

Suppose that

$$\sup_{t \in [0, T]} g^{n-1}(t) \leq C^{n-1} \frac{t^{n-1}}{(n-1)!} C_g.$$

Then, we have

$$g^n(t) \leq C \int_0^t C^{n-1} \frac{u^{n-1}}{(n-1)!} C_g du.$$

The claim follows. \square

Proof of Lemma 4.3.7. If $\alpha \geq 2$, recall that for all $x, y \geq 0$:

$$|x^{\alpha/2} - y^{\alpha/2}| \leq \frac{\alpha}{2} (\max(|x|; |y|))^{\alpha/2-1} |x - y|.$$

We deduce by Conditions 4.2.2 and 4.2.3 that

$$\left(\int_0^t \sigma_{\kappa, m+1}^2(s) - \sigma_{\kappa, m}^2(s) ds \right)^2 \leq CC_\kappa \int_0^t (\sigma_{\kappa, m+1}(s) - \sigma_{\kappa, m}(s))^2 ds,$$

by virtue of Jensen's inequality. Together with the Burkholder–Davis–Gundy inequalities, we obtain that

$$\begin{aligned} \mathbf{E} \sup_{u \leq t} |Y_{\kappa, m+1}(u) - Y_{\kappa, m}(u)|^2 &\leq CC_\kappa \mathbf{E} \int_0^t \sup_{u \leq s} |\sigma_{\kappa, m+1}(u) - \sigma_{\kappa, m}(u)|^2 ds \\ &\quad + CC_\kappa \mathbf{E} \int_0^t \sup_{u \leq s} |r_{\kappa, m+1}(u) - r_{\kappa, m}(u)| ds. \end{aligned}$$

By Conditions 4.2.2 or 4.2.3, with Lemma 4.7.2,

$$|\sigma_\kappa(t, e^X) - \sigma(t, e^Y)| + |r_\kappa(t, e^X) - r(t, e^Y)| \leq C_\kappa |X - Y|,$$

hence

$$|\sigma_{\kappa, m+1}(u) - \sigma_{\kappa, m}(u)| + |r_{\kappa, m+1}(u) - r_{\kappa, m}(u)| \leq C_\kappa |Y_{\kappa, m}(u) - Y_{\kappa, m-1}(u)|.$$

We deduce that

$$\mathbf{E} \sup_{u \leq t} |Y_{\kappa, m+1}(u) - Y_{\kappa, m}(u)|^2 \leq C_\kappa \int_0^t \mathbf{E} \sup_{u \leq s} |Y_{\kappa, m}(u) - Y_{\kappa, m-1}(u)|^2 ds,$$

where C_κ is a constant depending on κ . Using Lemma 4.7.9, we get that

$$\begin{aligned} \mathbf{E} \sup_{u \leq t} |Y_{\kappa, m+1}(u) - Y_{\kappa, m}(u)|^2 &\leq (C_\kappa)^{m+1} \mathbf{E} \sup_{u \leq t} |Y_{\kappa, 1}(u) - Y_{\kappa, 0}(u)|^2 \frac{t^m}{m!}, \\ &\leq C(C_\kappa)^{m+1} \frac{t^{m+1}}{(m+1)!} (1 + S_0^2), \end{aligned} \quad (4.7.13)$$

where C does not depend on m and κ . We deduce Inequality (4.3.6). Similarly, we obtain Inequality (4.3.7) and the claim follows. \square

Lemma 4.7.10. *Let a_1, \dots, a_k be real numbers. we have the following inequality*

$$\left(\sum_{l=1}^k a_l \right)^2 \leq \sum_{l=1}^k 2^l a_l^2, \quad \left(\sum_{l=1}^k a_l \right)^4 \leq \sum_{l=1}^k 8^l a_l^4.$$

Proof. For any real numbers a, b , the inequality $2ab \leq a^2 + b^2$ implies that $(a+b)^2 \leq 2a^2 + 2b^2$. Therefore, $(a+b)^4 \leq (2a^2 + 2b^2)^2 \leq 8a^4 + 8b^4$. We conclude by induction. \square

Proof of Corollary 4.3.8. Using Lemma 4.7.10, we obtain the inequality

$$\mathbf{E} \sup_{u \leq t} |Y_{\kappa, \infty}(u) - Y_{\kappa, m}(u)|^2 \leq \sum_{j=m}^{\infty} 2^j \mathbf{E} \sup_{u \leq t} |Y_{\kappa, j+1}(u) - Y_{\kappa, j}(u)|^2.$$

We deduce the first inequality from (4.7.13) and, similarly, the second one. The conclusion follows. \square

Proof of Lemma 4.3.9. Applying the Burkholder–Davis–Gundy and Jensen inequalities, we get that

$$\begin{aligned} \mathbf{E} \sup_{u \in [0, t]} (S_{\kappa, m+1}(u) - S_{\kappa, m}(u))^2 &\leq C \mathbf{E} \int_0^t (S_{\kappa, m+1}(s) \sigma_{\kappa, m+1}(s) - S_{\kappa, m}(s) \sigma_{\kappa, m}(s))^2 ds, \\ &\quad + 2 \mathbf{E} \int_0^t (S_{\kappa, m+1}(s) r_{\kappa, m+1}(s) - S_{\kappa, m}(s) r_{\kappa, m}(s))^2 ds, \end{aligned}$$

where we recall that $\sigma_{\kappa, m}(s) := \sigma_{\kappa}(s, S_{\kappa, m-1}(s))$. The first term in the r.h.s. of the inequality is bounded from above as follows :

$$\begin{aligned} \mathbf{E} \int_0^t (S_{\kappa, m+1}(s) \sigma_{\kappa, m+1}(s) - S_{\kappa, m}(s) \sigma_{\kappa, m}(s))^2 ds &\leq \\ 2 \mathbf{E} \int_0^T S_{\kappa, m+1}^2(s) [\sigma_{\kappa, m+1}(s) - \sigma_{\kappa, m}(s)]^2 ds &+ 2 \mathbf{E} \int_0^T (\sigma_{\kappa, m}(s) (S_{\kappa, m+1}(s) - S_{\kappa, m}(s)))^2 ds. \end{aligned}$$

The second one is similarly bounded. Recall the inequalities $|\sigma_{\kappa}| + |r_{\kappa}| \leq C_{\kappa}$, $\kappa > 0$, and recall that, by assumption or by Lemma 4.7.2,

$$|\sigma_{\kappa}(t, e^X) - \sigma_{\kappa}(t, e^Y)| + |r_{\kappa}(t, e^X) - r_{\kappa}(t, e^Y)| \leq C_{\kappa} |X - Y|. \quad (4.7.14)$$

Moreover, using Gronwall's lemma, since σ_{κ} and r_{κ} are bounded by a constant depending on κ , we deduce the existence of $C_{p, \kappa} > 0$ such that for all p

$$\sup_m \mathbf{E} \sup_{t \in [0, T]} S_{\kappa, m}^{2p}(t) \leq C_{p, \kappa}.$$

Using the stochastic integral representation of $S_{\kappa, m+1}$ and $S_{\kappa, m}$ respectively, we obtain by means of Cauchy-Schwarz's inequality, Jensen's inequality and the Ito isometry that

$$\begin{aligned} \mathbf{E} \sup_{u \in [0, t]} (S_{\kappa, m+1}(u) - S_{\kappa, m}(u))^2 &\leq C_{\kappa} \sqrt{\mathbf{E} \sup_{u \leq T} |Y_{\kappa, m}(u) - Y_{\kappa, m-1}(u)|^4} \\ &\quad + C_{\kappa} \int_0^t \mathbf{E} \sup_{u \leq s} (S_{\kappa, m+1}(u) - S_{\kappa, m}(u))^2 ds. \end{aligned}$$

Using Gronwall's Lemma, we deduce that

$$\mathbf{E} \sup_{t \in [0, T]} (S_{\kappa, m+1}(t) - S_{\kappa, m}(t))^2 \leq C_{\kappa} \sqrt{\mathbf{E} \sup_{t \leq T} |Y_{\kappa, m}(t) - Y_{\kappa, m-1}(t)|^4}.$$

The conclusion follows. \square

Proof of Lemma 4.3.10. By virtue of Lemma 4.7.10, we get the inequality

$$\begin{aligned} \mathbf{E} \sup_{t \in [0, T]} (S_{\kappa, \infty}(t) - S_{\kappa, m}(t))^2 &\leq \sum_{j=m}^{\infty} 2^j \mathbf{E} \sup_{u \in [0, t]} (S_{\kappa, j+1}(u) - S_{\kappa, j}(u))^2 \\ &\leq C_{\kappa} \sum_{j=m}^{\infty} \frac{(C_{\kappa})^j}{\sqrt{j!}}. \end{aligned}$$

□

Proof of Lemma 4.3.11. Since σ_κ and r_κ are bounded by a constant depending on κ , we deduce by the Burkholder-Davis-Gundy inequalities and the Gronwall's lemma a constant $c_\kappa > 0$ such that for all m, n ,

$$\mathbf{E} \sup_{t \in [0, T]} |S_{\kappa, m}(t) + S_{\kappa, m, n}(t)|^4 \leq c_\kappa. \quad (4.7.15)$$

For a sake of simplicity, we suppose that $r = 0$ as the proof is similar for the general case. Therefore,

$$\begin{aligned} S_{\kappa, 1}(t) &= S_0 + \int_0^t S_{\kappa, 1}(u) \sigma_\kappa(u, S_0) dW_u, \\ S_{\kappa, 1, n}(t) &= S_0 + \int_0^t S_{\kappa, 1, n}(u) \sigma_\kappa(u^{(n)}, S_0) dW_u. \end{aligned}$$

Let us introduce $g_m(t) := \mathbf{E} \sup_{u \in [0, t]} |S_{\kappa, m}(u) - S_{\kappa, m, n}(u)|^2$, $m \geq 1$. By the Doob's inequality,

$$\begin{aligned} g_1(t) &\leq 4 \int_0^t \left(S_{\kappa, 1, n}(u) \sigma_\kappa(u^{(n)}, S_0) - S_{\kappa, 1}(u) \sigma_\kappa(u, S_0) \right)^2 du, \\ g_1(t) &\leq 8 \int_0^t \sigma_\kappa^2(u^{(n)}, S_0) (S_{\kappa, 1, n}(u) - S_{\kappa, 1}(u))^2 du \\ &\quad + 8 \int_0^t S_{\kappa, 1}^2(u) \left(\sigma_\kappa(u^{(n)}, S_0) - \sigma_\kappa(u, S_0) \right)^2 du. \end{aligned}$$

Since σ_κ is bounded by a constant C_κ and $|\sigma_\kappa(u^{(n)}, S_0) - \sigma_\kappa(u, S_0)| \leq T/n$, we deduce by (4.7.15) that

$$g_1(t) \leq \frac{C_\kappa}{n^2} + C_\kappa \int_0^t g(u) du, \quad t \in [0, T].$$

By the Gronwall's lemma, we deduce that $g_1(T) \leq C_\kappa/n^2$ for some constant C_κ which does not depend on n . More generally, if $m \geq 2$, we have

$$\begin{aligned} S_{\kappa, m}(t) &= S_0 + \int_0^t S_{\kappa, m}(u) \sigma_\kappa(u, S_{\kappa, m-1}(u)) dW_u, \\ S_{\kappa, m, n}(t) &= S_0 + \int_0^t S_{\kappa, m, n}(u) \sigma_\kappa(u^{(n)}, S_{\kappa, m-1, n}^{(n)}(u)) dW_u. \end{aligned}$$

Therefore, by the Doob's inequality,

$$\begin{aligned} g_m(t) &\leq 4 \int_0^t \left(S_{\kappa, m, n}(u) \sigma_\kappa(u^{(n)}, S_{\kappa, m-1, n}^{(n)}(u)) - S_{\kappa, m}(u) \sigma_\kappa(u, S_{\kappa, m-1}(u)) \right)^2 du, \\ g_m(t) &\leq 8 \int_0^t \sigma_\kappa^2(u^{(n)}, S_{\kappa, m-1, n}^{(n)}(u)) (S_{\kappa, m, n}(u) - S_{\kappa, m}(u))^2 du \\ &\quad + 8 \int_0^t S_{\kappa, m}^2(u) \left(\sigma_\kappa(u^{(n)}, S_{\kappa, m-1, n}^{(n)}(u)) - \sigma_\kappa(u, S_{\kappa, m-1}(u)) \right)^2 du \\ &\quad + 8 \int_0^t S_{\kappa, m}^2(u) \left(\sigma_\kappa(u, S_{\kappa, m-1, n}(u)) - \sigma_\kappa(u, S_{\kappa, m-1}(u)) \right)^2 du. \end{aligned}$$

CHAPITRE 4. APPROXIMATION OF NON-LIPSCHITZ SDES BY PICARD ITERATIONS

Recall that σ_κ is bounded by a constant C_κ and

$$|\sigma_\kappa(u^{(n)}, S_{\kappa, m-1, n}(u)) - \sigma_\kappa(u, S_{\kappa, m-1, n}(u))| \leq C_\kappa/n.$$

Therefore, we may obtain upper bounds for the two first terms of the r.h.s. in the inequality above as we do for $m = 1$. Precisely, we also need to estimate $\mathbf{E}(S_{\kappa, m-1, n}^{(n)}(u) - S_{\kappa, m-1, n}(u))^2$. To do so, it suffices to recall that $\mathbf{E} \sup_{t \in [0, T]} S_{\kappa; m}^4(t) \leq C_\kappa$ for some constant C_κ which does not depend on m hence, if $u \in [t_i^n, t_{i+1}^n)$, $i = 0, \dots, n$, we deduce by the Ito isometry and the Cauchy-Schwarz inequality that $\mathbf{E}(S_{\kappa, m-1, n}^{(n)}(u) - S_{\kappa, m-1, n}(u))^2 \leq C_\kappa n^{-1}$ where C_κ only depends on κ .

At last, an upper bound for the third term is obtained using the Cauchy-Schwarz inequality, (4.7.15) and the following inequality (see Lemmas 4.7.2 and 4.7.1) :

$$|\sigma_\kappa(u, S_{\kappa, m-1, n}(u)) - \sigma_\kappa(u, S_{\kappa, m-1}(u))| \leq C_\kappa \sqrt{|S_{\kappa, m-1, n}(u) - S_{\kappa, m-1}(u)|}.$$

Therefore,

$$g_m(t) \leq \frac{C_\kappa}{n} + C_\kappa g_{m-1}(t) + C_\kappa \int_0^T g_m(u) du.$$

By the Gronwall's lemma, we then deduce a constant C_κ which only depends on κ such that

$$g_m(T) \leq \frac{C_\kappa}{n} + C_\kappa g_{m-1}(T).$$

By induction, we deduce that, for all $m \geq 2$,

$$g_m(T) \leq n^{-1} \sum_{j=1}^m (C_\kappa)^j \leq \alpha_\kappa (C_\kappa)^m n^{-1},$$

for some constant α_κ . \square

Lemma 4.7.11. *Let S be a positive stochastic process such that $P(\inf_{t \in [0, T]} S_t = a) = 0$ for all $a \in (0, S_0)$. Let $(S^n)_{n \in \mathbb{N}}$ be a sequence of stochastic processes such that $\mathbf{E} \sup_{t \in [0, T]} \|S_t - S_t^n\|^2 \rightarrow 0$ as $n \rightarrow \infty$. Let us define $\tau^S := \inf\{t : S_t < a\}$ and $\tau^{S^n} := \inf\{t : S_t^n < a\}$. Then, $\mathbf{E}|1_{\tau^S > T} - 1_{\tau^{S^n} > T}| \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. Suppose that $\mathbf{E}|1_{\tau^S > T} - 1_{\tau^{S^n} > T}| \rightarrow l > 0$ for a subsequence. Since $|\inf_{t \in [0, T]} S_t - \inf_{t \in [0, T]} S_t^n| \leq \sup_{t \in [0, T]} \|S_t - S_t^n\|$, we may also suppose that $\inf_{t \in [0, T]} S_t^n$ converges to $\inf_{t \in [0, T]} S_t$ a.s. for a common subsequence. In the case where $\tau > T$, i.e. $\inf_{t \in [0, T]} S_t \geq a$, as $P(\inf_{t \in [0, T]} S_t = a) = 0$, we may assume that $\inf_{t \in [0, T]} S_t > a$ hence $\inf_{t \in [0, T]} S_t^n > a$ for n large enough and, finally $\tau^{S^n} > T$. If $\tau \leq T$, $\inf_{t \in [0, T]} S_t^n < a$ and then, similarly, we obtain that $\tau^{S^n} \leq T$ for n large enough. We deduce that $|1_{\tau^S > T} - 1_{\tau^{S^n} > T}| \rightarrow 0$ a.s. in contradiction with $l > 0$. \square

Lemma 4.7.12. *Let S be a positive stochastic process such that $P(\inf_{t \in [0, T]} S_t = a) = 0$ for all $a > 0$. Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of stochastic processes such that $\mathbf{E} \sup_{t \in [0, T]} |S_t - S_n(t)|^2 \rightarrow 0$ as $n \rightarrow \infty$. Let us denote by $S_n^{(n)}$ its discrete version given in Section 4.2.1. Let us define $\tau^S := \inf\{t : S_t < a\}$ and $\tau^{S_n^{(n)}} := \inf\{t : S_n^{(n)}(t) < a\}$. Then, $\mathbf{E}|1_{\tau^S > T} - 1_{\tau^{S_n^{(n)}} > T}| \rightarrow 0$.*

Proof. Recall that $\tau^S > T$ if and only if $\inf_{t \in [0, T]} S_t \geq a$. Let us first show that, for any random interval $I \subseteq [0, T]$, $\sup_{t \in I} S_n(t)$ converges in L^2 to $\sup_{t \in I} S_t$ as $n \rightarrow \infty$. To see it, note that for all $t \in [0, T]$, $S_t \leq S_n(t) + \sup_{t \in [0, T]} |S_t - S_n(t)|$ hence $\sup_{t \in I} S_t \leq \sup_{t \in I} S_n(t) + \sup_{t \in [0, T]} |S_t - S_n(t)|$. By symmetry, we get that

$$\left| \sup_{t \in I} S_n(t) - \sup_{t \in I} S_t \right| \leq \sup_{t \in [0, T]} |S_t - S_n(t)|$$

and the conclusion follows. Note that the same property holds if we replace the infimum by the supremum of S over I .

On the set $\{\inf_{t \in [0, T]} S_t < a\}$, as S is continuous, we may construct by a measurable selection argument a measurable interval $I = [t_0 - \varepsilon, t_0 + \varepsilon]$, $\varepsilon > 0$ a.s., such that $\sup_{t \in I} S_t < a$. Otherwise, we arbitrarily set $I = [0, T]$.

Suppose that for a subsequence $\mathbf{E}|1_{\tau^S > T} - 1_{\tau^{S_n^{(n)}} > T}| \rightarrow l > 0$. Consider another subsequence such that we also have $\sup_{t \in [0, T]} |S_t - S_n(t)| \rightarrow 0$ a.s. By the first step for the infimum, if $\inf_{t \in [0, T]} S_t > a$, then $\inf_{t \in [0, T]} S_n(t) > a$ if n is large enough. As $\inf_{t \in [0, T]} S_n^{(n)}(t) \geq \inf_{t \in [0, T]} S_n(t)$, we deduce that $\inf_{t \in [0, T]} S_n^{(n)}(t) > a$ and, finally, $|1_{\tau^S > T} - 1_{\tau^{S_n^{(n)}} > T}| = 0$ if n is large enough.

On the set $\inf_{t \in [0, T]} S_t < a$, i.e. when $\tau^S \leq T$, we have $\sup_{t \in I} S_t < a$. By the first step, $\sup_{t \in I} S_n(t) < a$ if n is large enough. Choose n sufficiently large so that $T/n < \varepsilon(\omega)$. Then, the random discrete date t_i^n such that $t_0(\omega) \in [t_i^n, t_{i+1}^n)$ satisfies $t_i^n \in I$ hence $S_n^{(n)}(t_0) = S_n(t_i^n) < a$. It follows that $\tau^{S_n^{(n)}} \leq T$. We finally deduce by the Lebesgue theorem that $\mathbf{E}|1_{\tau^S > T} - 1_{\tau^{S_n^{(n)}} > T}| \rightarrow 0$ in contradiction with $l > 0$. \square

Lemma 4.7.13. *Consider the case where $\sigma_\kappa = \sigma(\cdot, \cdot \vee \kappa^{-1})$ and $r_\kappa = r(\cdot, \cdot \vee \kappa^{-1})$ under Condition 4.2.3 with $P(\tau \leq T) > 0$. The random variables $Z_\kappa = \inf_{t \in [0, T]} Y_\kappa(t)$, $\kappa > 0$, and $Z_{\kappa, m} = \inf_{t \in [0, T]} Y_{\kappa, m}(t)$, $m \geq 1$, admit a density with respect to the Lebesgue measure.*

Proof. The proof is based on [57][Proposition 2.1.11]. To do so, we need verify the conditions of [57][Proposition 2.1.10]. As the mapping $h(x) = x \vee \kappa^{-1}$ is not differentiable at

point κ^{-1} , we first approximate h by

$$\begin{aligned} h^n(x) &= \kappa^{-1}1_{(-\infty, \kappa^{-1}-n^{-1}]}(x) + x1_{[\kappa^{-1}+n^{-1}, \infty)}(x) \\ &+ \left(\kappa^{-1} + \frac{n}{4}(x - \kappa^{-1} + n^{-1})^2 \right) 1_{[\kappa^{-1}-n^{-1}, \kappa^{-1}+n^{-1}]}(x). \end{aligned}$$

We have the following properties :

$$0 \leq h^n(x) - h(x) \leq \frac{1}{4n}, \quad |(h^n)'(x) - h'(x)| \leq 1_{[\kappa^{-1}-n^{-1}, \kappa^{-1}+n^{-1}]}(x).$$

We then consider the solution Y_κ^n of the s.d.e. (4.3.4) where we replace σ_κ by

$$\sigma_\kappa^n(t, x) := \sigma(t, h^n(x)).$$

We may show that $\sup_{t \in [0, T]} |Y_\kappa^n(t) - Y_\kappa(t)| \rightarrow 0$ in L^2 as $n \rightarrow \infty$. Therefore, since Y_κ^n admits a derivative in the Malliavin sense, we also deduce that $Y_\kappa \in D^{1,2}$. Moreover, for all $s \leq t$, making $n \rightarrow \infty$, we get that

$$\begin{aligned} D_s Y_\kappa(t) &= \sigma_\kappa(s, S_0 e^{Y_\kappa(s)}) + \int_s^t D_s Y_\kappa(u) h'(S_0 e^{Y_\kappa(u)}) S_0 e^{Y_\kappa(u)} \nabla_x \sigma_\kappa(t, S_0 e^{Y_\kappa(u)}) dW_u, \\ &+ \int_s^t D_s Y_\kappa(u) h'(S_0 e^{Y_\kappa(u)}) S_0 e^{Y_\kappa(u)} \nabla_x \gamma_\kappa(t, S_0 e^{Y_\kappa(u)}) du, \end{aligned}$$

where $\gamma_\kappa = r_\kappa - \sigma_\kappa^2/2$. By the assumptions, e.g. $X \nabla_X \sigma(t, X)$ is bounded if $X \geq \kappa^{-1}$ and $h'(X) = 0$ if $X \leq \kappa^{-1}$, and the Gronwall's lemma, we deduce that

$$\mathbf{E} \sup_{t \in [0, T]} |D_s Y_\kappa(t)|^p < \infty,$$

whatever $p \geq 2$. By the Ito formula, we obtain an expression of $[D_s Y_\kappa(t)]^2$ for $t \geq s$ as a stochastic integral. By the stochastic Fubini theorem, we then also deduce a representation of $t \mapsto \int_0^t [D_s Y_\kappa(t)]^2 ds$ as a stochastic integral. Using the Gronwall's lemma, we finally deduce that $\mathbf{E} \sup_{t \leq T} \int_0^t [D_s Y_\kappa(t)]^2 ds < \infty$, i.e. the conditions of [57][Proposition 2.1.10] hold. Moreover, as $(D_s Y_\kappa(t))_{t \geq s}$ is a Doléans-Dade exponential, it is strictly positive and we finally deduce that [57][Proposition 2.1.11] applies, i.e. $Z_\kappa = \inf_{t \in [0, T]} Y_\kappa(t)$ admits a density with respect to the Lebesgue measure. The second statement is shown similarly by induction for $m = 0, 1, \dots$. \square

Proof of Theorem 4.3.12. Notice that we have

$$|\mathbf{E}G(S)1_{\tau > T} - \mathbf{E}G(S_{\kappa, 1, n}^{(n)})1_{\tilde{\tau}^{\kappa, 1, n} > T}| \leq \mathbf{E}|G(S) - G(S_\kappa)|1_{\tau > T} \quad (4.7.16)$$

$$+ \mathbf{E}|G(S_\kappa) - G(S_{\kappa, m})| \quad (4.7.17)$$

$$+ \mathbf{E}|G(S_{\kappa, m}) - G(S_{\kappa, m, n})| \quad (4.7.18)$$

$$+ \mathbf{E}|G(S_{\kappa, m, n}) - G(S_{\kappa, m, n}^{(n)})| \quad (4.7.19)$$

$$+ \mathbf{E}|G(S)| |1_{\tau > T} - 1_{\tilde{\tau}^{\kappa, 1, n} > T}|, \quad (4.7.20)$$

Since G is supposed to be Lipschitz, the first term (4.7.16) is bounded from above by ε_κ as stated in Lemma 4.3.6. The second term (4.7.17) is bounded from above by the error $\varepsilon_{\kappa,m}$ given by Lemma 4.3.10. The term (4.7.18) is bounded from above according to Lemma 4.3.11. The consistency property with respect to discretization satisfied by G allows us to estimate the fourth term (4.7.19). As to the last term, notice that G is of linear growth since it is Lipschitz. By Corollary 4.7.7, we deduce from the Cauchy-Schwarz inequality a constant C such that

$$\mathbf{E}|G(S)||1_{\tau>T} - 1_{\tilde{\tau}^{\kappa,1,n}>T}| \leq C\sqrt{\mathbf{E}|1_{\tau>T} - 1_{\tilde{\tau}^{\kappa,1,n}>T}|}.$$

It remains to show that $\mathbf{E}|1_{\tau>T} - 1_{\tilde{\tau}^{\kappa,m,n}>T}| \rightarrow 0$ a.s. as $\kappa, m, n \rightarrow \infty$. Let us introduce $\tilde{\tau}^\kappa := \inf\{t : S_\kappa(t) < \kappa^{-1}\}$. Since S and S_κ coincides on $[0, \tau^\kappa]$ and $\tilde{\tau}^\kappa = \tau^\kappa > T$ if κ is large enough and $\tau > T$, we deduce that $|1_{\tau>T} - 1_{\tilde{\tau}^\kappa>T}| \rightarrow 0$ if $\tau > T$ as $\kappa \rightarrow \infty$. If $\tau \leq T$, we have $\tau^\kappa \leq \tau$ so that $|1_{\tau>T} - 1_{\tau^\kappa>T}| = 0$. Therefore, $\mathbf{E}|1_{\tau>T} - 1_{\tilde{\tau}^\kappa>T}| \rightarrow 0$ as $\kappa \rightarrow +\infty$. It remains to show that $\mathbf{E}|1_{\tilde{\tau}^\kappa>T} - 1_{\tilde{\tau}^{\kappa,m,n}>T}| \rightarrow 0$ a.s. as $\kappa, m, n \rightarrow \infty$. Let us introduce $\tilde{\tau}^{\kappa,m} := \inf\{t : S_{\kappa,m}(t) < \kappa^{-1}\}$. By Proposition 4.3.10, Lemma 4.7.13 and Lemma 4.7.11, for fixed κ , $\mathbf{E}|1_{\tilde{\tau}^\kappa>T} - 1_{\tilde{\tau}^{\kappa,m}>T}| \rightarrow 0$ as $m \rightarrow \infty$. Therefore, it remains to show that $\mathbf{E}|1_{\tilde{\tau}^{\kappa,m}>T} - 1_{\tilde{\tau}^{\kappa,m,n}>T}| \rightarrow 0$ a.s. as $n \rightarrow \infty$. By virtue of Lemma 4.3.11, Lemma 4.7.13 and Lemma 4.7.12, $\mathbf{E}|1_{\tilde{\tau}^{\kappa,m}>T} - 1_{\tilde{\tau}^{\kappa,m,n}>T}| \rightarrow 0$ a.s. as $n \rightarrow \infty$. \square

4.8 Conclusion

En terme de résultats, la méthode proposée possède deux points faibles qui sont une convergence plutôt faible et l'absence d'estimation du taux de convergence. Toutefois, elle s'avère performante dans certains cas. Il est difficile de déterminer une technique de simulation globalement meilleure que les autres. Comme nous l'avons vu dans les exemples rencontrés, une même méthode peut être très efficace ou au contraire inappropriée.

La méthode des itérations de Picard semble performante pour les modèles où le coefficient de diffusion est sous-linéaire - ce qui est intéressant pour le pricing d'options asiatiques - tandis que la méthode BEM est plus efficace pour les termes de diffusion à croissance sur-linéaire.

Bien que ce domaine ne soit pas en rapport avec les travaux de recherche de Das Bot, il a permis le développement et l'intégration de plusieurs modules de pricing d'options à la librairie de R&D de l'entreprise. Une large partie du travail a concerné l'implémentation en C++ de l'ensemble des dynamiques et des pricers étudiés. Le paradigme utilisé a été la programmation orientée objet ce qui a permis l'intégration d'outils de la librai-

rie QuantLib permettant de gérer des paramètres non utilisés dans les exemples mais indispensables dans la réalité (calendriers, term structure des taux d'intérêts et des dividendes, ...) rendant les pricers très flexibles à l'utilisation. Enfin, l'optimisation du code a été travaillée avec l'utilisation du multi-threading et une gestion optimale de la mémoire.

Deuxième partie

Méthodes d'apprentissage
automatique appliquées au
trading algorithmique

Chapitre 5

Application de méthodes d'apprentissage automatique à une stratégie intraday d'Open Range Breakout sur Future DAX

Nous nous intéressons à la construction d'une stratégie de trading basée sur des éléments d'analyse technique et améliorée par des méthodes d'apprentissage automatique. Sur la base d'une stratégie d'Open Range Breakout appliquée sur le marché des Futures DAX, nous montrons qu'il est possible de construire une stratégie rentable exploitant uniquement de la donnée passée. De plus, intégrer une combinaison de méthodes d'apprentissage supervisé permet de définir une famille de stratégies d'investissement viables et attractives pour des clients privés et institutionnels.

5.1 Introduction

L'analyse technique est une discipline dont les principes entrent de manière flagrante en contradiction avec l'hypothèse d'efficience faible des marchés comme énoncée par Eugène Fama dans [27] et [28]. Selon cette dernière, l'ensemble de l'information des prix passés d'un actif est contenu dans son prix actuel. Les travaux de Paul Samuelson dans [63] ont quant à eux conclu que les prix suivaient un marché aléatoire. Il est alors inutile de rechercher un avantage informationnel en manipulant de la donnée passée. Bien que cette hypothèse soit adoptée par une large partie des acteurs et des observateurs des marchés, beaucoup de professionnels ont recours à l'analyse technique et de nombreuses

études ont cherché à remettre en cause ses conclusions. Les travaux les plus connus sont ceux de Andrew W. Lo et A. Craig MacKinlay dans [51] et [52].

En plus de cette contradiction, l'analyse technique a souvent été décriée pour son caractère subjectif, et c'est un fait qui ne peut être nié. De nombreuses stratégies présentées sont souvent incomplètes - voire irréalistes - et ne permettent pas d'être exploitées. Certaines figures d'analyse technique peuvent être interprétées de différentes manières et aboutir à des conclusions différentes. Au final, la chance semble guider certains et mener à la perte la majorité.

Avec l'avènement du trading algorithmique, une opportunité s'est offerte à l'analyse technique : celle de définir de manière précise des indicateurs et des règles d'investissement puis de laisser opérer un robot qui prendra des décisions en toute objectivité. C'est le pari qu'a fait l'entreprise Das Bot. Aujourd'hui, la recherche est tournée vers la détection d'indicateurs techniques permettant d'exploiter des anomalies de marchés de manière systématique.

Mais au-delà de la question académique, il y a une réalité économique : celle de présenter des produits d'investissements compétitifs à des clients qui ne sont pas forcément connaisseurs tout en se démarquant de la concurrence. Cela fait naître une double problématique. Il y a d'abord plus de 600 acteurs sur le marché de la gestion privée en France. Comment proposer une approche des marchés différenciante au client alors que quasiment tous les champs de l'investissement sont déjà couverts ? En second lieu, comment décider un client à investir quand les mots "trading algorithmique" font peur ? Car c'est une réalité, le trading algorithmique, confondu avec les mathématiques financières, évoque pour de nombreux clients la crise de 2008 et ses dérivées. A cela s'ajoute la peur de voir leur argent géré par des robots.

De manière tout aussi irrationnelle, le public et les médias projettent sur l'intelligence artificielle les fantasmes d'une technologie qui remplacera à terme les humains dans les chaînes de production et les processus de décision. De manière plus raisonnable, ces technologies investissent le quotidien, ce qui fait dire à Cédric Villani : "L'intelligence artificielle sera partout, comme l'électricité".

C'est donc avec l'apprentissage automatique - un champ de l'intelligence artificielle - que nous avons l'ambition de nous démarquer. Combiner analyse technique, apprentissage automatique et trading algorithmique est très rare - voire inédit - ce qui nous

singularise auprès des concurrents. En second lieu, l'apport de l'intelligence artificielle en terme purement marketing est important. Enfin, cela ouvre la voie à un vaste domaine de recherche inexploré. Nous présentons donc ici une première stratégie basée sur des éléments d'analyse technique à laquelle nous avons appliqué des méthodes d'apprentissage automatique.

Dans une première partie nous présenterons le FDAX, qui est l'actif sur lequel nous travaillons, et la stratégie d'Open Range Breakout. Nous montrerons comment nous pouvons l'optimiser afin de proposer une première stratégie rentable prenant en compte tous les coûts d'exécution. Nous introduirons dans une seconde partie les méthodes d'apprentissage utilisées et étudierons l'impact de ces méthodes sur la performance de la stratégie. Dans une troisième partie nous construirons notre stratégie finale combinant une famille de méthodes d'apprentissage. Enfin nous terminerons par une discussion sur les résultats de la stratégie sur la période 2016-2018, sur ses limites et les pistes de recherche à poursuivre.

5.2 Open Range Breakout sur Future DAX

Nous commençons par présenter la stratégie d'Open Range Breakout puis le contrat Future DAX FDAX. Nous effectuons ensuite un premier test de la stratégie sur la période Janvier 2007 à Avril 2015 et nous montrons que les signaux d'entrée en position sont pertinents car l'espérance des rendements est significativement différente de zéro dans les minutes qui suivent un signal. Nous poursuivons par une optimisation des paramètres de la stratégie et montrons qu'en prenant en compte tous les coûts d'exécution, le système reste profitable.

La donnée de marché utilisée pour l'ensemble des tests est de la donnée tick by tick Level 1. Elle inclut le meilleur prix au bid, le meilleur prix à l'ask et le dernier prix échangé ainsi que leurs volumes correspondants. Nous travaillons donc avec l'historique complet des échanges effectués et les simulations sont faites avec les prix proposés dans le carnet d'ordre.

5.2.1 L'Open Range Breakout

La stratégie d'Open Range Breakout (ORB) est une stratégie de momentum (donc de suivi de tendance). C'est un système traditionnel en analyse technique d'abord utilisé en trading journalier ([15]). Avec le développement des systèmes d'informations, il a été

facile de la transposer au trading intra-day (positions ouvertes de quelques minutes à quelques heures) et sa faculté à générer du profit a été discutée dans plusieurs études ([24], [36] et [53] notamment) avec des résultats divers.

La stratégie est basée sur la détection de deux seuils déclencheurs de tendance (mouvement directionnel). La stratégie peut être décomposée en trois temps. Le premier temps est d'une durée fixe et débute dès l'ouverture des marchés actions. Il a pour objectif de définir l'Open Range soit l'intervalle des prix parcourus sur cette durée. A son terme, le prix négocié le plus haut S^{max} est considéré comme le seuil de déclenchement d'une tendance à la hausse et le prix le plus bas S^{min} , le seuil de déclenchement d'une tendance à la baisse.

La seconde phase de la stratégie concerne le signal d'ouverture de position. Dès lors qu'un seuil est franchi, une position est prise. Si le prix franchit S^{max} par le dessous, alors la position est LONG car nous estimons qu'une tendance haussière est amorcée. A l'inverse, si le prix franchit S^{min} par le dessus, une position SHORT est prise. Enfin le dernier temps concerne la fermeture de la position. Celle-ci est clôturée à la fermeture du marché.

Bien que simple et apparemment claire, la définition de cette stratégie souffre des lacunes trop souvent pointées du doigt par les détracteurs de l'analyse technique à savoir un trop grand flou. La durée de la fenêtre de détection de l'Open Range n'est pas donnée par exemple, ni les actifs à traiter pour servir de support d'investissement à la stratégie. Avant de préciser les paramètres de la stratégie, nous présentons l'actif sur lequel nous travaillerons car ses propriétés influent sur la construction de notre ORB.

5.2.2 Le contrat future FDAX

Nous travaillons sur l'actif FDAX, contrat future de l'indice boursier allemand DAX basé sur le cours des actions des 30 plus importantes entreprises de la Bourse de Francfort. Ce produit a été choisi pour plusieurs de ses caractéristiques.

Il y a d'abord sa liquidité. Le FDAX est le future indiciel le plus échangé en Europe offrant une contrepartie à toute heure de la journée de marché et un spread faible.

Par construction, le pas de cotation du FDAX est de 0.5 point, et un point de cota-

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

tion vaut 25 euros. A cela s'ajoute des frais de transaction très faibles. Le broker facture 1.5 euros pour un passage d'ordre. Le gain d'un seul tick permet donc de largement compenser les coûts d'entrée et sortie sur un trade.

Enfin, le niveau de cotation du contrat ainsi que l'amplitude de ses variations au cours d'une journée permet de générer des gains intéressants. Le FDAX est depuis plusieurs années au dessus des 10 000 points. Actuellement autour des 12 700, le prix d'un contrat avoisine les 320 000 euros. Une variation quotidienne, même faible, se compte donc en dizaines voire centaines de points. Or le prix d'un contrat FDAX est fixé comme le produit du nombre de points par la valeur d'un point (25 euros). Les frais d'exécutions sont là encore rendus minimales par les variations du contrat.

Par comparaison, le future sur l'indice CAC 40 a côté entre 4000 et 5000 ces dernières années. Le pas de cotation minimal est un point et il vaut 10 euros. Une variation d'un pourcent de ce produit à 4000 points revient à 400 euros tandis qu'une variation d'un pourcent du future DAX à 10 000 points revient à 2 500 euros ce qui est plus intéressant en terme de performance et d'absorption des coûts d'exécution.

Le contrat FDAX cote quotidiennement de 8h00 à 22h00 tandis que son sous-jacent cote de 9h00 à 17h30. De ce fait, l'activité est moindre sur les périodes où le DAX est fermé. Dernière précision, les contrats existent pour plusieurs maturités différentes, sont échangés sur le marché règlementé Eurex et expirent quatre fois dans l'année (troisième vendredi des mois de Mars, Juin, Septembre et Décembre). Le contrat dont la maturité est la plus proche est le plus traité et offre le plus de liquidité. Nous utilisons le contrat avec la maturité la plus courte. Dès lors qu'un contrat expire, nous travaillons avec le contrat encore en cours ayant la date d'expiration la plus proche.

En prenant en compte ces caractéristiques, nous complétons les paramètres de notre stratégie. La phase de détection de l'Open Range commence à 9h00 avec l'ouverture du DAX et dure 30 minutes (il s'agit de la durée la plus communément admise en pratique). A 9h30 les seuils sont fixés et des ordres limites sont posés sur ceux-ci. Dès lors qu'un seuil est franchi, une position est prise. Si le cours vient franchir le seuil de détection opposé, la première position est alors clôturée avant d'en ouvrir une seconde dans le sens opposé. Il n'y a en effet aucune logique à vouloir suivre une tendance si deux positions opposées sont ouvertes. Ce changement de direction de la tendance est appelé un reverse et un seul reverse peut être effectué par jour. Au-delà, nous considérons qu'aucune tendance ne se dégage, ou que celle-ci ne peut être exploitée de façon rentable par la

stratégie.

Après 17h30, plus aucune position ne peut être prise. En effet, les volumes et les variations du FDAX chutent après la fermeture du DAX. Il n'est donc plus judicieux de chercher une tendance après cette heure. Enfin, les positions sont clôturées à 21h30 et non à la clôture de 22h00 pour éviter tout problème de liquidité et d'exécution ce qui pourrait entrainer une position over-night et donc un risque de marché non prévu ainsi que des appels de marge.

5.2.3 Première application de la stratégie d'Open Range Breakout

Nous appliquons de manière stricte la stratégie présentée sur un portefeuille de capital un million d'euros. Lors de chaque trade, nous investissons toujours dans un même nombre de contrats future fixé à quatre. Ainsi, un point gagné ou perdu correspond à un différentiel de 100 euros. Il est tout a fait possible d'investir plus que la valeur du portefeuille étant donné que nous déposons une provision auprès du broker et non la totalité du montant des contrats. Un tel portefeuille est donc réaliste pour notre stratégie. Enfin, les coûts de transaction sont pris en compte et représentent 3 euros par trade (1.50 euros à l'entrée et à la sortie).

Le taux sans risque utilisé pour le calcul de la sur-performance est le taux OAT (Obligation Assimilable du Trésor) à 5 ans. Bien que le produit traité soit allemand, le taux sans risque utilisé est celui du pays de résidence de l'investisseur. Or une majorité des clients sont français d'où ce choix.

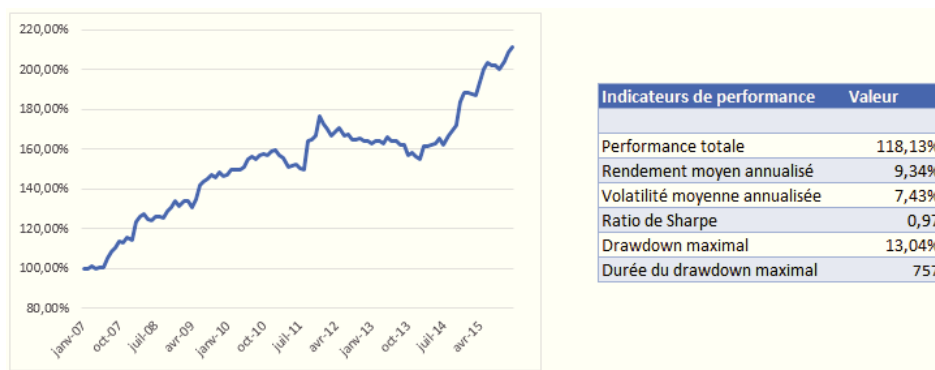


FIGURE 5.1 – Performance de la stratégie entre Janvier 2007 et Décembre 2015

Les premiers résultats sont globalement bons et le capital initial est plus que doublé en

moins de 10 ans (Fig. 5.1). Le rendement moyen annualisé de 9.34% est bon et la volatilité est acceptable avec une valeur de 7.43%. Le ratio de Sharpe, légèrement inférieur à 1, est moyen. Mais les principales réserves viennent du drawdown maximal supérieur à 13%. Le drawdown maximal est une mesure de risque qui correspond au pire rendement que l'actif subit sur la période observée. Cela signifierait qu'un client entrant le premier jour du drawdown maximal pourrait perdre jusqu'à 13% du capital qu'il a investi. De plus cette baisse durerait 757 jours soit plus de deux ans sans performance positive. Très rares sont les clients capables de soutenir une telle séquence et encore plus dans la clientèle privée.

C'est ici que réside la principale faiblesse de l'Open Range Breakout dans cette configuration. Si la stratégie dégage une performance correcte avec une volatilité importante mais raisonnable et un risque mesurée, elle est très difficilement concevable de la promouvoir et de la vendre à des clients dont l'objectif est de trouver un placement sûr à défaut d'être performant.

5.2.4 Significativité des signaux de trade

Avant de nous intéresser à l'optimisation de la stratégie, étudions la qualité du signal d'entrée en trade. Etant donné que nous ne travaillons que sur une seule trajectoire de prix, il est tout à fait possible que les gains constitués ne soient que le fruit du hasard, tout comme une série de lancés de dé pourraient augmenter la fortune d'un joueur. Pour cela, nous effectuons un test de Student (aussi appelé t-statistic) chaque minute après le signal de trade afin de déterminer si statistiquement, le décalage des prix est significatif.

Nous définissons l'instant $t = 0$ comme l'instant où un seuil de déclenchement de trade est franchi et notons μ_t la moyenne des rendements absolus à l'instant $t > 0$, l'hypothèse nulle testée ici est $H_0 : \mu_t = 0$. La statistique de test Z_t est définie par $Z_t = \frac{\bar{x}_t}{s_{n,t}^*} \sqrt{n_t}$ avec n_t la taille de l'échantillon, \bar{x}_t la moyenne empirique et $s_{n,t}^*$ l'estimateur sans biais de la variance à l'instant t . L'évolution de cette statistique de test est présentée pour les trades LONG (à l'achat) et SHORT (à la vente) sur le graphique de la Figure 5.2.

La significativité des résultats est immédiate. Dès la première demi-heure, la statistique de test dépasse le seuil de significativité de 99% aussi bien pour les trades LONG que pour les trades SHORT. L'hypothèse nulle est donc rejetée et la moyenne des rendements des prix 30 minutes après le signal de trade est différente de zéro au seuil de confiance de 99%.



FIGURE 5.2 – Evolution de la statistique de test sur les 600 minutes après le signal

De manière moins forte l'hypothèse nulle peut être rejetée sur la majorité de la première heure au seuil de 95%. Il y a donc bien un biais sur le comportement des prix lorsque les seuils de l'Open Range sont dépassés et ce biais perdure plusieurs dizaines de minutes.

A partir de la seconde heure, l'hypothèse nulle est conservée. Il n'y a alors plus de véritable signal. Cela ne changera pas pour les trades LONG. Quant aux trades SHORT, la statistique de test dépasse de nouveau le seuil de 95% près de 500 minutes après le signal et ce pendant une centaine de minute. Une première différence majeure apparait entre les tendances haussières et baissières. Si le dépassement d'un seuil de l'Open Range est significatif, la tendance initiée à la hausse sera très brève et les prix reprendront rapidement un comportement normal. A l'inverse, une tendance baissière aura un fort biais au départ, un comportement plus erratique par la suite, mais une tendance se construira à mesure que la journée avance pour être à terme significative.

Il y a donc au terme de cette étude un signal existant dès lors que les seuils de l'Open Range sont dépassés. Le comportement des prix est différent selon que la tendance est haussière ou baissière. Une anomalie est présente et nous allons dans la partie suivante proposer plusieurs améliorations afin de l'exploiter au mieux.

5.2.5 Optimisation des paramètres

L'optimisation des paramètres de la stratégie est une partie délicate car il y a un risque de sur-optimisation, c'est à dire de calibrer une stratégie afin qu'elle fonctionne très bien sur l'échantillon de test et réaliser par la suite qu'elle s'écroule dès lors qu'on l'utilise sur de la nouvelle donnée. Pour cette étape, nous nous limiterons à deux paramètres tout

en justifiant nos choix.

Amplitude de l'Open Range Le premier paramètre optimisé est la taille de l'Open Range. Intuitivement un Open Range trop petit ou trop grand parat problématique. En effet, s'il est trop petit, les variations du prix peuvent être de l'ordre de grandeur de l'Open Range et franchir les bornes sans pour autant qu'une tendance se dessine. Nous aurons alors l'apparition de faux signaux de trade. A l'inverse, si l'Open Range est trop important, la distance que le prix aura parcouru pour atteindre un seuil sera telle que la tendance sera terminée lors du départ en trade. Il faut donc encadrer l'amplitude.

Pour cela nous traçons le P&L de la stratégie en fonction de l'amplitude maximale de l'Open Range autorisée afin de déterminer sur quelles valeurs de l'amplitude sont concentrés les gains (Fig. 5.3). Les amplitudes inférieures à 30 points ne semblent pas pertinentes, notamment pour les trades LONG. La performance est nulle, voire négative et ces observations semblent en accord avec notre hypothèse à savoir que briser un Open Range de faible amplitude est plutôt dû aux variations du prix qu'à l'émergence d'une tendance.

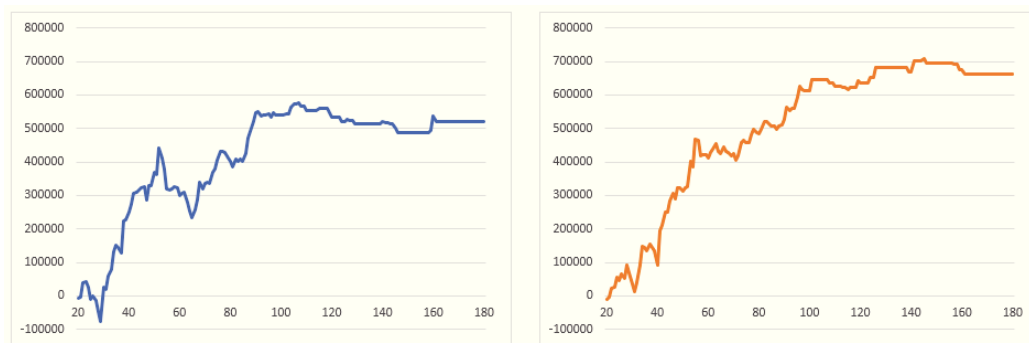


FIGURE 5.3 – Profits des trades LONG (à gauche) et SHORT (à droite) en fonction de l'amplitude maximale de l'Open Range autorisée

La performance apparait à partir de 30 points d'amplitude et se construit jusqu'au delà des 100 points. Nous notons cependant un fléchissement des gains pour les trades LONG aux alentours de 50 à 65 points. Il y a clairement ici un effet que nous ne retrouvons pas pour les trades SHORT. Cependant il est difficile de déterminer les causes de cette détérioration du profit et en l'absence d'explications, nous considérons que retirer cet intervalle d'amplitude mènerait à de la sur-optimisation.

La performance des trades LONG atteint un palier après les 90 points et diminue même

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

par la suite. Pour les trades SHORT, ce palier est atteint aux alentours des 150 points. Nous choisissons donc ici de déterminer un seuil maximale de l'amplitude différent selon le type de trade et ce choix est motivé par les résultats précédents. Nous avons observé une différence sur les tendances haussières et baissières. Les tendances à la baisse sont plus importantes et durent plus longtemps tandis que les tendances haussières ne sont plus significatives après la première heure. Les tendances baissières permettent donc d'aller chercher de la performance malgré une amplitude d'Open Range supérieure à 100 points ce que n'autorisent pas leur équivalent haussier.

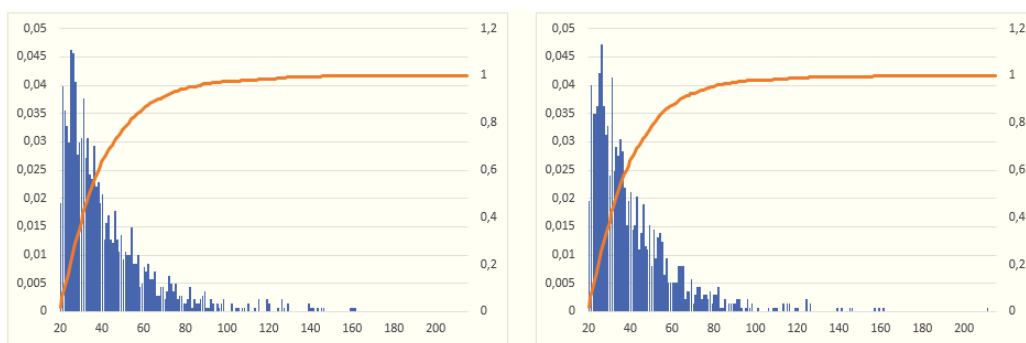


FIGURE 5.4 – Distribution et cumulatives des amplitudes de l'Open Range pour les trades LONG (à gauche) et SHORT (à droite)

Les graphiques de la Figure 5.4 présentent les distribution des amplitudes de l'Open Range. Il existe un nombre important de trades dont l'amplitude de l'Open Range est inférieur à 30 points. Ils représentent approximativement 38% du total. C'est à dire que plus du tiers des trades effectués constituent du bruit. Les supprimer n'enlève pas de performance et évite des frais inutiles.

Supprimer les trades dont l'amplitude d'Open Range est supérieure à 90 points pour les trades LONG et supérieure à 150 points pour les trades SHORT produit un impact beaucoup plus limité en nombre de trades supprimés. Cela permet donc de limiter les pertes potentielles tout en minimisant le risque d'opportunité.

Fenêtre horaire de trading Le deuxième paramètre que nous allons optimiser est l'étendue de la fenêtre de trading. Bien que les trades soient déjà stoppés à 17h30 avec la fermeture du DAX, nous pouvons nous demander s'il existe un effet heure sur les résultats de la stratégie.

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

Après avoir filtré les trades qui ne convenaient pas aux règles établies sur l'amplitude d'Open Range, les graphiques de la Figure 5.5 présentent le gain moyen et la distribution horaire des trades par type.

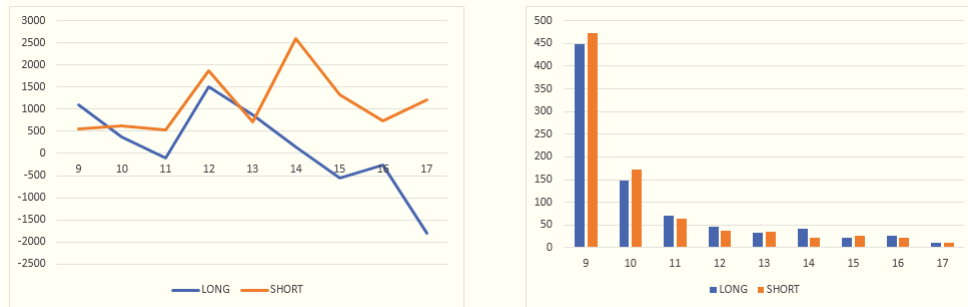


FIGURE 5.5 – Gain moyen horaire (à gauche) et distribution horaire (à droite) des trades LONG et SHORT

Les distributions horaires des trades sont similaires avec une forte activité à l'ouverture du DAX puis une décroissance continue jusqu'à la fermeture. Trois quart des trades ont lieu avant 11 heure et les trades débutant l'après-midi sont marginaux (de l'ordre de 10%).

Le gain moyen par trade pour les SHORT est nettement positif et augmente au cours de la journée avec de fortes variations. Le maximum atteint entre 14 heures et 15 heures concide avec l'heure d'ouverture des marchés américains. A l'inverse, les résultats des trades LONG sont plus irréguliers avec un premier fléchissement entre 11 heure et midi puis un gain moyen par trade négatif après 15 heures.

Afin de vérifier que les résultats négatifs des trades LONG ne sont pas l'effet d'un évènement ponctuel, nous nous intéressons à l'évolution des gains moyens sur les tranches horaires 15h-16h, 16h-17h et 17h-17h30 présentés Figure 5.6. Pour les trois tranches horaires, les valeurs sont dans l'ensemble négatives. Les résultats les plus mauvais concernent la tranche 17h-17h30 qui ne génère qu'un gain moyen par trade positif en 2008. La tranche horaire 15h-16h présentent deux années intéressantes en 2011 et 2014 mais les résultats sont fluctuants et non convaincants.

La tranche la plus intéressante est celle de 16h-17h avec deux belles années en 2011 et 2015. Les gains sont plombés par les années 2007, 2008 et 2011. Or il s'agit de trois

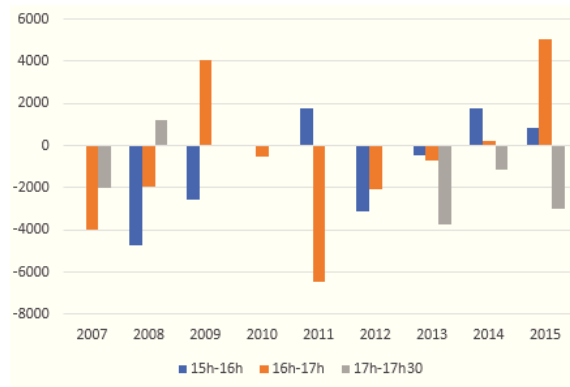


FIGURE 5.6 – Evolution des gains moyens par trades pour 3 tranches horaires entre 2007 et 2015

année qui ont connu des crises. Et si nous regardons les résultats plus en détails, les pertes se constituent en Juillet 2007, en Septembre 2008 et en Août-Septembre 2011 ce qui concident avec le blocage du marché interbancaire, la chute de Lehman Brothers et la crise de la dette européenne respectivement. Ces pertes sont donc le résultat de marchés très volatiles, marqués par l'ouverture des bourses américaines, où les seuils de l'Open Range ont été franchis sans véritable tendance. Cette tranche horaire est donc très sensible aux épisodes de crise.

En conclusion, la tranche horaire pour les trades SHORT reste inchangée et s'étend de 9h30 à 17h30. Pour les trades LONG, la fenêtre est raccourcie de 9h30 à 15h. Bien que les mauvais résultats de la tranche 16h-17h soient explicables, nous ne la retenons pas mais il n'est pas exclu de revenir sur cette décision dans le future.

Comparaison des résultats après optimisation La stratégie est nettement améliorée après optimisation. Le nombre de trade est passé de 2779 à 1654 soit une diminution de 40%. La nouvelle stratégie sous-performe la précédente sur la période 2007 à 2014 (Figure 5.7). Le départ est même très tardif car le P&L reste nul sur la première année. A terme, les deux stratégies se rejoignent et la version optimisée performe mieux sur l'année 2015.

La performance totale s'élève désormais à 126.84% Le rendement moyen annualisé a légèrement augmenté à 9.77% tandis que la volatilité moyenne annualisée est légèrement inférieure à 6.86% ce qui porte le ratio de Sharpe à 1.11. L'impact le plus spectaculaire concerne le drawdown qui chute à 6.94% contre 13.04% auparavant et la durée du drawdown maximal est désormais de 50 jours.

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX



FIGURE 5.7 – Performance de la stratégie optimisée entre Janvier 2007 et Décembre 2015

Si les optimisations apportées n'ont que peu touché à la performance de la stratégie, c'est la gestion du risque qui a été grandement améliorée. Désormais, elle varie entre des phases de gains et des plateaux où le portefeuille reste stable sans perte prononcée.

Avant d'appliquer des algorithmes d'apprentissage automatique à la stratégie, nous allons étudier l'impact du slippage. Il s'agit du dernier phénomène pouvant altérer la performance et sa prise en compte donnera des résultats reflétant avec exactitude la réalité du marché.

5.2.6 Impact du slippage sur la stratégie finale

Le slippage désigne la différence de cours d'exécution entre le cours souhaité au moment où l'ordre est envoyé au broker et le cours auquel l'ordre est réellement exécuté. Le slippage a des origines diverses. Il peut être dû à une latence (le prix proposé a déjà été exécuté mais l'information n'est pas encore arrivée ou le prix souhaité n'est plus disponible lorsque l'ordre arrive sur la place de marché) ou bien une position défavorable dans la file d'attente du carnet d'ordre.

Il est difficile d'anticiper le slippage mais nous pouvons le simuler en reprenant l'historique des trades effectués chez le broker. Pour cela nous nous appuyons sur 770 trades passés entre Mars 2015 et Novembre 2016. La distribution des valeurs du slippage observées est présentée Figure 5.8.

Seul un tiers des trades sont exécutés sans slippage. Celui-ci peut monter jusqu'à 18 points dans des cas extrêmes ce qui en fait un évènement non négligeable. Cependant,

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

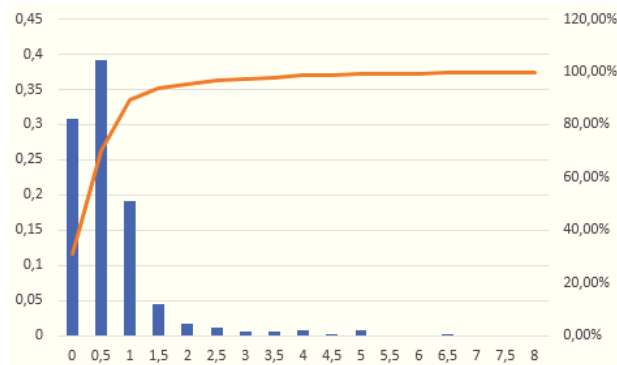


FIGURE 5.8 – Distribution des valeurs du slippage observées auprès du broker entre Mars 2015 et Novembre 2016

il est inférieur ou égal à 2 points dans 95% des cas et la moyenne observée est de 0.65 points. Afin d'être conservateur, nous utiliserons un slippage systématique d'un point à l'entrée pour tous les trades. Pour la sortie, nous compterons un point de slippage uniquement lors de l'utilisation d'un stop loss (il n'y a pas de slippage pour un ordre au marché utilisé pour clôturer les positions à 21h30).

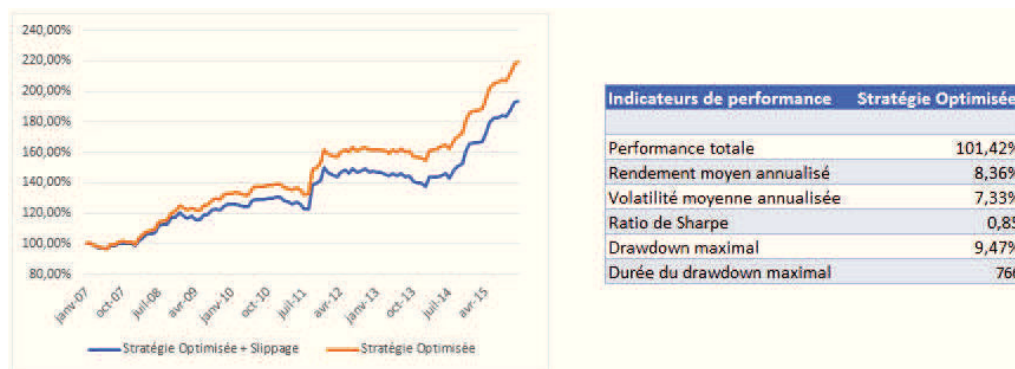


FIGURE 5.9 – Performance de la stratégie optimisée avec un slippage de 1 point entre Janvier 2007 et Décembre 2015

L'impact est très important sur toutes les caractéristiques de la stratégie. La performance totale chute à 101.42%. Le rendement moyen annualisé se replie tandis que la volatilité moyenne annualisée augmente. Le ratio de Sharpe est désormais inférieur à 1, le drawdown augmente à 9.47% et la durée du drawdown maximal explose à 766 jours.

Les résultats sont beaucoup moins bons. Seuls le drawdown maximal et la volatilité ont été améliorés par rapport à la stratégie initiale. Le slippage est véritablement l'élément qui peut détruire la performance d'une stratégie. Malgré tout, avec un slippage fortement

surestimé, la stratégie offre toujours des caractéristiques intéressantes.

Au terme de cette partie, nous avons vu qu'il était possible de construire une stratégie basée sur l'Open Range qui est un indicateur technique. Cette stratégie génère une performance convaincante sur 9 années de données et résiste aux coûts de transaction ainsi qu'au slippage. Cependant, elle est encore difficilement utilisable pour la gestion privée. Nous allons voir comment il est possible de l'améliorer par l'utilisation de méthodes d'apprentissage automatique.

5.3 Open Range Breakout et apprentissage automatique

L'apprentissage automatique sera utilisé comme sur-couche à la stratégie d'Open Range Breakout développée. Indépendamment du fonctionnement de l'ORB, au moment où un seuil est franchi et un ordre est sur le point d'être envoyé au broker, un algorithme analysera la configuration du marché (représentée par une collection d'indicateurs) et sur la base de configurations déjà observées, validera ou non l'envoi de l'ordre pour exécution.

Nous présentons dans un premier temps les indicateurs qui modéliseront la configuration de marché. Les algorithmes d'apprentissage seront ensuite abordés et nous déterminerons la méthodologie de tests. Nous terminerons par une première utilisation des méthodes d'apprentissage.

5.3.1 Indicateurs et variable à expliquer

L'apprentissage automatique traite principalement deux catégories de problèmes : la régression et la classification. Nous pourrions très bien utiliser les deux approches pour améliorer notre stratégie et il y aurait de nombreuses façons de le faire. Ici nous avons choisi d'adopter une approche de classification dans sa forme la plus basique : séparer les trades gagnants des trades perdants. La variable à expliquer est donc une variable binaire prenant la valeur 1 lorsque le gain est strictement positif et 0 sinon.

Choisir les indicateurs (ou variables explicatives) est la tâche la plus compliquée car il n'y a aucune règle prédéfinie pour cela. Le choix des variables sera fonction de l'actif suivi. De plus, tester trop d'indicateurs introduit un biais de data mining : à trop chercher on finit toujours pas trouver mais le résultat est certainement trop adapté à la donnée d'entraînement.

D'expérience, il s'est trouvé qu'utiliser trop de variables explicatives ne donnait pas de bons résultats. Notre choix s'est porté sur trois variables : l'amplitude de l'Open Range, la distance du seuil de déclenchement au prix d'ouverture de la journée et l'heure de départ en trade. L'utilisation de données sur les volumes et la volatilité n'ont pas permis d'obtenir des résultats convaincants.

Un élément important à noter également est que les valeurs des variables utilisées sont disponibles au moment où la décision doit être prise. Nous évitons ainsi le biais consistant à utiliser de la donnée future (Look-Ahead bias). Dès 9h30, deux des trois variables sont connues et l'heure de départ en trade correspond à l'heure ou l'action de classification est faite. Les variables choisies sont donc compatibles avec une utilisation en temps réel du système de trading.

5.3.2 Méthodes d'apprentissage et méthodologie

Nous nous limitons dans notre étude à des méthodes d'apprentissage supervisé pour la classification. Nous avons donc choisi cinq méthodes différentes pour nos tests.

Méthodes d'apprentissage Nous utilisons les méthodes les plus connues : la machine à vecteur de support (SVM), la méthode des k-plus proches voisins (KNN), la classification naïve bayésienne (GNB), les forêts d'arbres décisionnels (RF) et le perceptron multi-couches (réseau neuronal formel noté MLP). L'ensemble des tests ont été effectués avec les implémentations de la librairie Python scikit-learn (<http://www.scikit-learn.org>).

Le classifieur GNB est le seul ne demandant pas d'hyper-paramètres. Pour les autres, les valeurs utilisées lors des premiers tests sont pour la plupart celles données par défaut dans les constructeurs des objets de la librairie. Le classifieur KNN utilise cinq voisins et une fonction de pondération uniforme (par défaut). Le classifieur RF utilise 10 arbres dans la forêt (par défaut).

Le perceptron multi-couches est légèrement modifié par rapport aux valeurs par défaut de la librairie : il possède trois couches avec 100, 75 puis 50 neurones (de base, il n'y a qu'une couche de 100 neurones). Le SVM utilise le noyau gaussien (radial basis function kernel) par défaut. Seul le paramètre de pénalité est fixé à 3.5 au lieu de 1.0 dans le constructeur.

Méthodologie de tests Il existe de nombreuses méthodes pour valider la qualité d'un modèle de classification. Nous avons choisi ici une méthode de validation croisée consistant à séparer l'échantillon de données en deux afin de réserver la première partie pour l'apprentissage et la seconde partie pour le test. Ce choix est motivé par le fait que nous travaillons sur une série chronologique. Il est intuitivement plus intéressant d'apprendre sur le passé et de tester sur le futur. Ce point sera rediscuté par la suite.

Certains algorithmes demandent que la donnée soit centrée et réduite. Elle le sera donc pour tous les tests effectués. Là encore, la plus grande attention a été portée afin qu'aucune information future ne soit utilisée lorsqu'une décision est prise : les valeurs de moyenne et d'écart-type pour la normalisation de la donnée sont calculées uniquement sur l'échantillon d'apprentissage, puis utilisées sur l'échantillon de test.

Lors des étapes d'optimisation des classifieurs, deux paramètres sont utilisés. Le paramètre principal est la précision car nous voulons que le classifieur choisi soit celui qui se trompe le moins parmi les trades qu'il a considéré comme gagnants. On y associe également un rappel minimal. Il arrive en effet qu'un classifieur fasse un sans faute mais sur trop peu de trades, ce qui le rend inintéressant. La valeur du rappel minimal dépendra de la méthode d'apprentissage. En général le rappel minimal est fixé haut, puis nous faisons décroître sa valeur tant qu'aucun classifieur ne répond aux critères.

5.3.3 Premiers tests

Tout d'abord, nous détaillons quelques statistiques de la stratégie restreinte à l'échantillon de test (Figure 5.10). Le nombre de trades est de 827 sur cette période pour un pourcentage de gains de 41.96%. Si la stratégie est perdante près de 6 fois sur 10, elle se rattrape avec un profit moyen bien supérieur à la perte moyenne en valeur absolue. Au final l'espérance de gain est de 613.38 euros. Avec une t-statistic de 2.37, l'espérance de gain est significativement différente de zéro au seuil de 97.5%.

Statistiques générales	
Nombre de trades	827
Nombre de trades gagnants	347
Ratio de gains	41.96%
Profit moyen	7734.18
Perte moyenne	-4534.35
Espérance de gain	613.38
t-Statistic	2.37**

FIGURE 5.10 – Statistiques des trades sur l'échantillon de test

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

Nous appliquons maintenant les méthodes d'apprentissage non calibrées à l'échantillon de test après les avoir entrainées sur l'échantillon d'entrainement. Les résultats sont présentés dans le tableau de la Figure 5.11.

	SVM	KNN	GNB	MLP	RF
Probabilité de bien classés	59.85%	57.32%	59.98%	59.13%	58.28%
Précision	64.91%	49.12%	58.18%	51.85%	50.99%
Rappel	10.60%	31.81%	18.34%	44.13%	29.51%
Rendement moyen	4743.49	1581.74	2731.09	1753.06	1629.92
Statistique de comparaison des rendements moyens	2.656***	1.602	2.055*	2.106*	1.606
Trades conservés	6.89%	27.33%	13.30%	35.91%	24.43%
Performance conservée	53.30%	70.47%	59.22%	102.64%	64.91%
Diminution du drawdown maximal	80.68%	61.63%	68.34%	55.68%	48.01%

FIGURE 5.11 – Résultats de l'application des méthodes d'apprentissage non calibrées à la stratégie ORB sur l'échantillon de test

Pour les cinq classifieurs, la probabilité de bien classés avoisine les 58% ce qui est un résultat moyen, mais n'est pas pénalisant dans notre cas. La précision est supérieure à 49% et atteint même 65% pour la machine à vecteur de support en nette progression par rapport au pourcentage de gains de la stratégie initiale. La classification naïve bayésienne obtient également une précision intéressante avec plus de 58%. Les cinq classifieurs améliorent le pourcentage de gains de la stratégie dans des proportions non négligeables.

Le rappel est par contre pénalisant. Il s'agit du paramètre le plus difficile à améliorer et il est capital car une stratégie exécutant peu d'ordres s'expose à une dégradation importante de la performance si plusieurs pertes s'enchaînent. Il est donc important de le maintenir élevée. C'est le perceptron multi-couches qui offre les meilleurs résultats avec un rappel à 42%, ce qui est un bon score ici. La machine à vecteur de support présente un rappel très faible à 10.60%.

L'impact sur le rendement moyen de la stratégie est frappant. Il a au moins doublé pour les cinq classifieurs et il est presque huit fois supérieur pour le SVM. En appliquant un test de comparaison de moyennes, le rendements moyen a été significativement amélioré pour trois classifieurs : au seuil de 95% pour la classification naïve bayésienne et le perceptron multi-couches et au seuil de 99% pour le SVM. Sans calibration, et en évitant tout biais de data mining, ces deux méthodes d'apprentissage permettent d'améliorer le rendement de la stratégie en utilisant uniquement de la donnée passée, ce qui est un premier résultat fort.

Les classifieurs sont également impressionnants dans leur capacité à conserver la performance de la stratégie. Le SVM est celui qui en capte le moins. Cependant il réalise 53.30% de la performance totale avec moins de 7% des trades initiaux. Le perceptron à multi-couches conserve le plus de trades avec 35.91% pour une performance qui atteint 102% de la performance initiale.

Le drawdown maximal est très largement réduit. Tous les classifieurs le diminuent au moins du tiers, trois le diminuent pratiquement de moitié et le SVM le diminue par cinq. Le risque de la stratégie en est fortement diminué. Encore une fois, les résultats des classifieurs sont très bons.

Après ce premier test, les résultats sont sans appel. A la simple application des méthodes d'apprentissage, les performances de la stratégie ont été améliorées de manière prononcée. Statistiquement, trois méthodes ont eu un impact significatif et permettent d'affirmer qu'utiliser l'information du passé a amélioré le rendement de la stratégie sur l'intervalle de test.

Maintenant, un certain dilemme s'offre à nous : faut-il conserver des méthodes choisies "au hasard" bien que les résultats soient statistiquement significatifs ? Ou bien devons-nous optimiser les hyper-paramètres au risque de tomber dans le biais de data mining et donc de sur-optimiser les classifieurs ? Sans pour autant mettre de côté ces premiers résultats, nous allons optimiser de manière contrôlée les hyper-paramètres afin de déterminer si de meilleurs classifieurs existent.

5.3.4 Optimisation des hyper-paramètres

Seuls quatre classifieurs sont optimisables étant donné que la classification naïve bayésienne ne possède pas d'hyper-paramètres. Pour les autres méthodes d'apprentissage, nous générons des grilles de paramètres sur lesquelles nous testons tous les combinaisons (Figure 5.12).

Au final, nous testons 108 combinaisons pour la machine à vecteurs de support, 18 pour les k-plus proches voisins, 288 pour le perceptron à multi-couches et 336 pour la forêt d'arbres décisionnels. Les résultats sont présentés dans le tableau de la Figure 5.13.

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

```

# Support Vector Machine
param_grid_SVM = ParameterGrid({'C'      : [.1, .5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 7.0, 10.0],
                               'gamma'  : ['auto', 0.01, 0.1, 0.5, 1.0, 2.0, 2.5, 5.0, 10.0],
                               'kernel'  : ['rbf']})

# K-Nearest Neighbors
param_grid_KNN = ParameterGrid({'n_neighbors' : [5, 10, 15, 20, 25, 50, 75, 100, 150],
                               'weights'      : ['uniform', 'distance'],
                               'algorithm'    : ['auto']})

# Baussion Naïve Bayesian
param_grid_GNB = ParameterGrid({'priors' : [None]})

# Random Forest
param_grid_RF = ParameterGrid({'n_estimators' : [5, 10, 15, 20, 25, 50, 75, 100],
                               'criterion'    : ['gini', 'entropy'],
                               'max_features' : ['sqrt', 'log2', None],
                               'random_state' : [1],
                               'max_depth'    : [5, 10, 25, 50, 100, None]})

# Multi-Layers Perceptron
param_grid_MLP = ParameterGrid({'activation' : ['identity', 'logistic', 'tanh', 'relu'],
                               'solver'      : ['lbfgs', 'sgd', 'adam'],
                               'hidden_layer_sizes' : [(100,), (100, 100),
                                                       (100, 100, 100),
                                                       (100, 75, 75),
                                                       (100, 75, 50),
                                                       (100, 75, 100, 50),
                                                       (100, 75, 100, 75, 100, 50),
                                                       (100, 50, 100, 50, 100, 50, 100, 50)],
                               'random_state' : [1],
                               'alpha'       : [1e-5, 1e-4, 1e-3, 1e-2]
                               })

```

FIGURE 5.12 – Grilles de paramètres utilisées pour l'optimisation

	SVM	KNN	MLP	RF
Probabilité de bien classés	59.73%	57.32%	59.73%	60.10%
Précision	54.00%	49.12%	52.99%	54.77%
Rappel	30.95%	31.81%	40.69%	31.23%
Rendement moyen	2208.75	1581.74	2080.21	2376.40
Statistique de comparaison des rendements moyens	2.382**	1.602	2.504**	2.728***
Trades conservés	24.18%	27.33%	32.41%	24.06%
Performance conservée	87.08%	70.47%	109.90%	93.23%
Diminution du drawdown maximal	72.34%	61.63%	47.76%	69.96%

FIGURE 5.13 – Résultats de l'application des méthodes d'apprentissage optimisés à la stratégie ORB sur l'échantillon de test

La première information qui ressort est que le classifieur k-plus proches voisins renvoie les mêmes résultats après optimisation. Plus précisément, les mêmes résultats ont été obtenus pour des hyper-paramètres différents. Il s'agit du classifieur pour lequel nous avons testé le moins de combinaisons. Nous nous contenterons de ces résultats pour éviter toute sur-optimisation.

Les trois autres classifieurs voient leur probabilité de bien classés légèrement augmenter sans que cela soit flagrant. La précision est elle aussi légèrement meilleure hormis pour le SVM où elle passe à 54% ce qui reste correct. Cependant le rappel est triplé pour cette méthode ce qui est une très bonne amélioration. Le rappel des deux autres classifieurs reste proche des valeurs précédentes.

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

Le rendement moyen est en hausse pour la forêt d'arbres décisionnels et pour le perceptron à multi-couches. Désormais leur différence avec le rendement de la stratégie initiale est significative au seuil de 99% pour le premier et 97.5% pour le second. La significativité des résultats pour le SVM est légèrement moins bonne que précédemment avec un seuil de confiance passé à 97.5% ce qui reste toujours très correct.

Grce à un meilleur rappel, le SVM conserve désormais 24.18% des trades et 87.08% de la performance. Avec le même ratio de trades conservés, la méthode RF atteint 93.23% de performance conservée. Cela atteint 110% pour le perceptron avec environ 32% des trades, soit moins que précédemment.

Le drawdown maximal de la forêt d'arbres décisionnels est désormais réduit de 70% contre 48% précédemment. Avec une réduction de 72.34% également, le SVM fait moins bien qu'au premier test. Le perceptron voit son drawdown maximal plus important que lors du premier test.

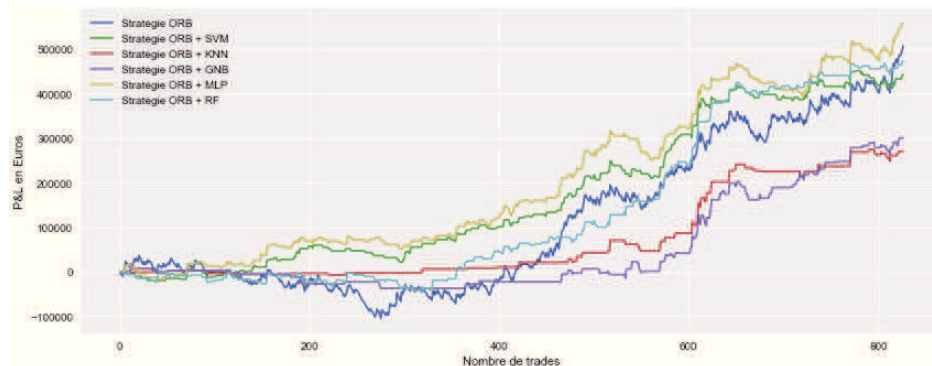


FIGURE 5.14 – Comparaison des gains cumulés des classifieurs sur l'échantillon de test

Ces tests sont concluants en terme de significativité. En testant un nombre mesuré de classifieurs, les résultats ont encore été améliorés. Il existe très certainement des hyperparamètres permettant de meilleurs résultats. Mais il ne serait pas judicieux d'aller plus loin.

Une autre piste d'amélioration est envisageable lorsque nous regardons les performances des divers classifieurs sur le graphique de la Figure 5.14. Deux groupes se détachent : trois classifieurs qui suivent de près la courbe de la stratégie initiale voire même la surpassent

et deux autres classifieurs qui restent en sommeil sur près des deux tiers de l'échantillon avant de performer.

L'idée serait de combiner les méthodes et deux voies sont possibles. Nous pourrions souhaiter qu'au moins une ou deux méthodes renvoient un résultat positif afin de valider un nombre plus important de trades en espérant augmenter le rappel au risque de peser sur la précision. Ou alors demander à ce qu'un maximum de classifieurs renvoient un résultat positif pour augmenter la précision au détriment du rappel.

5.4 Combinaison des méthodes d'apprentissage et stratégie finale

Nous allons donc ici tester la combinaison des résultats des classifieurs en utilisant les deux ensembles de cinq classifieurs utilisés précédemment. Pour chaque ensemble, nous allons calculer les performances de la stratégie en fonction du nombre de classifieurs requis pour classer un trade comme gagnant.

5.4.1 Combinaison sur l'ensemble des classifieurs non calibrés

Comme prévu, le tableau de la Figure 5.15 montre qu'augmenter le nombre de classifieurs pour prendre une décision augmente la précision mais diminue le rappel. Prendre une décision avec au moins un classifieur n'est pas suffisant. Malgré un bon rappel, la précision est trop moyenne et l'augmentation du rendement moyen n'est pas significative. Il en est de même avec au moins deux classifieurs malgré un drawdown maximal diminué de 63%, une performance conservée de 84.22% et 7 trades sur 10 éliminés.

	1	2	3	4	5
Probabilité de bien classés	54.05%	58.89%	62.52%	60.10%	59.01%
Précision	46.71%	51.78%	67.26%	75.68%	77.78%
Rappel	63.04%	37.54%	21.78%	8.02%	4.01%
Rendement moyen	971.95	1688.70	4124.43	7542.95	8222.00
Statistique de comparaison des rendements moyens	0.804	1.809	3.679***	3.708***	2.415**
Trades conservés	56.95%	30.59%	13.66%	4.47%	2.18%
Performance conservée	90.25%	84.22%	91.88%	55.02%	29.18%
Diminution du drawdown	36.40%	62.89%	80.96%	86.21%	88.30%

FIGURE 5.15 – Résultats des stratégies en fonction du nombre de classifieurs non calibrés requis pour classer un trade comme gagnant

Les résultats les plus intéressants sont obtenus avec au moins 3 classifieurs. Le rappel est

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

un peu bas à 21.78% mais la précision est de 67.26% ce qui est très bon. Le drawdown maximal est divisé par cinq et avec seulement 13.66% des trades initiaux, 91.88% de la performance est conservée. Le rendement moyen est désormais significativement différent du rendement moyen initial au seuil de 99.95%. Nous avons ici un très bon prétendant pour la stratégie finale.

Demander au moins quatre classifieurs ou plus n'est plus pertinent du fait du rappel trop bas. Néanmoins les résultats sont intéressants à regarder. La précision dépasse les 75%, le rendement moyen est désormais dix fois supérieur au rendement moyen initial de la stratégie et le drawdown maximal est divisé par huit. Un point remarquable parmi d'autres, avec seulement 2.18% des trades, nous conservons 29.18% de la performance lorsque nous demandons aux cinq classifieurs de classer le trade comme gagnant.

5.4.2 Combinaison sur l'ensemble des classifieurs calibrés

Nous recommençons le même travail avec les classifieurs calibrés. Les résultats sont présentés dans le tableau de la Figure 5.16.

	1	2	3	4	5
Probabilité de bien classés	55.38%	61.19%	60.34%	60.58%	59.37%
Précision	47.83%	54.70%	56.52%	67.69%	84.21%
Rappel	63.04%	46.70%	26.07%	12.61%	4.58%
Rendement moyen	1205.91	2120.83	2793.27	4645.46	10086.47
Statistique de comparaison des rendements moyens	1.299	2.698***	2.858***	3.139***	3.438***
Trades conservés	55.62%	36.03%	19.47%	7.86%	2.30%
Performance conservée	109.35%	124.59%	88.65%	59.53%	37.78%
Diminution du drawdown	47.46%	59.37%	66.21%	85.84%	91.95%

FIGURE 5.16 – Résultats des stratégies en fonction du nombre de classifieurs calibrés requis pour classer un trade comme gagnant

Seule la stratégie demandant au moins un classifieurs classant un trade comme gagnant n'a pas un rendement moyen significativement différent du rendement moyen initial. C'est la conséquence d'une précision trop faible car autrement le rappel atteint 63%, le nombre de trade est divisé par deux, le drawdown maximal est diminué de 47% et la performance finale dépasse celle de la stratégie initiale.

Les meilleurs résultats sont obtenus lorsqu'au moins deux classifieurs sont demandés. La précision est correcte avec 55% mais le rappel atteint 47%. Seulement 36% des trades initiaux sont conservés et la performance finale dépasse de 24.59% celle de la stratégie initiale pour un drawdown maximal réduit de 60%. Cette configuration a l'avantage de conserver un nombre plus important de trades ce qui la rend moins vulnérable à une

éventuelle série de pertes. Demander trois classifieurs fait chuter le rappel à 26%, ce qui est trop juste.

Les résultats sont impressionnants pour les derniers tests mais la contrepartie est un rappel trop faible. Lorsque les cinq classifieurs sont demandés, malgré le très faible nombre de trades, le rendement moyen est tel que la différence avec le rendement moyen initial est significative au seuil de 99.9%. La précision dépasse les 80%, le drawdown maximal est diminué de 92% et avec seulement 18 trades, 37.78% de la performance générée par les 827 trades initiaux est conservée. Il est ici indéniable que les classifieurs ont profité d'un surplus d'information contenu dans la base d'entraînement pour obtenir de tels résultats.

5.4.3 Impact des classifieurs sur la performance de la stratégie à risque équivalent et choix final

Nous retenons deux combinaisons de classifieurs des tests effectués précédemment. Une première combinaison demande au moins trois classifieurs classant un trade comme gagnant sur l'ensemble des cinq classifieurs non calibrés. Nous la noterons stratégie 3NC. La seconde demande au moins deux classifieurs classant un trade comme gagnant sur l'ensemble des cinq classifieurs calibrés. Elle sera notée stratégie 2C.

Nous avons jusqu'ici remarqué qu'une large partie de la performance initiale était conservée tandis que le drawdown maximal - que nous avons déterminé comme notre mesure de risque - était fortement diminué. Afin d'estimer au mieux l'impact de la classification sur les résultats de la stratégie, nous comparons les résultats à risque équivalent. C'est à dire que nous multiplions les résultats après classification par un coefficient de telle sorte que le drawdown maximal soit égal à ceux sans classification.

Le graphique de la Figure 5.17 compare les gains cumulés des stratégies à risque équivalent. Les résultats sont une fois de plus excellents. La performance finale est plus que doublée avec 2C, et même triplée avec 3NC. Les stratégies avec classification sur-performent catégoriquement la stratégie initiale dans des grandeurs impressionnantes. Même avec très peu d'optimisation, l'apprentissage automatique a permis d'extraire de la performance de la base d'apprentissage.

La t-statistic du P&L est de 4.490 pour la stratégie 3NC et de 4.283 pour la stratégie

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

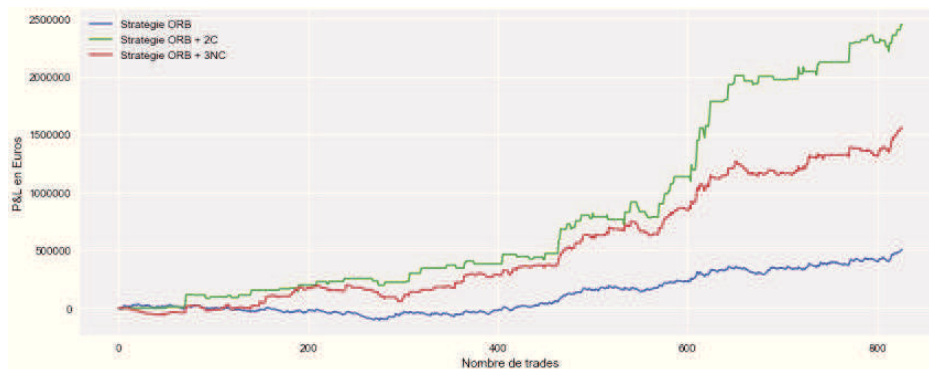


FIGURE 5.17 – Comparaison des gains cumulés des stratégies à risque équivalent

2C ce qui signifie que le rendement moyen des stratégies est significativement différent de zéro au seuil de 99.99%. Ce résultat est bien supérieur à celui de la stratégie initiale dont la significativité était obtenue au seuil de 97.5%.

Il s'agit maintenant d'effectuer un choix final entre les deux dernières stratégies. Ce choix est difficile car la stratégie 3NC possède l'avantage d'être non optimisée et donc moins sujette au biais de data mining. Mais son rappel est faible et peut l'exposer à des pertes consécutives difficile à gérer dans le temps. La stratégie 3NC possède un meilleur rappel, mais a été optimisée. Nous sommes au stade où aucune véritable règle n'existe et c'est un choix subjectif qui aura nécessairement lieu. Nous avons préféré conserver la stratégie 2C car il est préférable de conserver un nombre plus important de trades pour construire la performance. De plus, le graphique comparant les stratégies à risque équivalent montre que la sur-performance de 3NC se fait sur une courte période. Les deux stratégies sont confondues jusqu'au trade 600 et terminent sur la même dynamique.

5.5 Stratégie finale, limites et perspectives

Nous présentons succinctement une gamme de stratégies issue de la stratégie 2C et leur résultats sur la période allant d'Avril 2012 à Décembre 2015. Nous poursuivons avec des résultats sur la période allant de Janvier 2016 à Septembre 2017. Il s'agit là de données qui n'ont jamais été utilisées et qui permettront véritablement de valider notre travail. Enfin nous discuterons des limites de notre étude ainsi que des perspectives qui s'ouvrent pour des recherches futures.

5.5.1 Une gamme de stratégies 2C

La stratégie 2C développée précédemment est définie de telle sorte que les ordres émis portent sur quatre contrats FDAX. Il est facile de modifier le couple rentabilité/risque de la stratégie puisqu'il est proportionnel au nombre de contrats traités. Il nous est tout naturellement venu l'idée de proposer plusieurs stratégies dérivées offrant un rendement et un risque différents. Pour cela nous avons décidé de proposer les stratégies 2C-2, 2C-4 et 2C-8 traitant respectivement 2, 4 et 8 contrats.

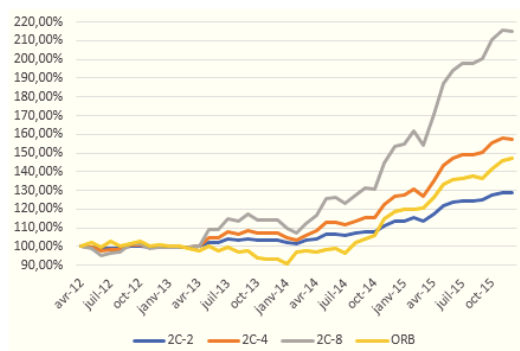


FIGURE 5.18 – Performance de la gamme de stratégies 2C et de la stratégie ORB optimisée entre Avril 2012 et Décembre 2015

Le graphique de la Figure 5.18 présente les courbes de rendement de la gamme de la stratégie 2C comparées à la stratégie ORB optimisée initiale tandis que le tableau de la Figure 5.19 en détaille les performances.

	2C-2	2C-4	2C-8	ORB Optimisée
Performance totale	32.47%	65.03%	130.15%	54.84%
Rendement moyen annualisé	7.88%	14.61%	25.93%	12.83%
Volatilité moyenne annualisée	4.30%	7.76%	13.38%	9.42%
Ratio de Sharpe	1.66	1.78	1.88	1.28
Drawdown maximal	2.75%	5.27%	9.77%	12.43%
Durée du drawdown maximal	156	156	156	605

FIGURE 5.19 – Performance de la gamme de stratégies 2C et de la stratégie ORB optimisée entre Avril 2012 et Décembre 2015

La stratégie 2C-2 est destinée à des clients souhaitant un placement peu risqué avec un rendement correct. Elle propose un rendement annualisé de 7.88% avec une volatilité annualisée bien maîtrisée à 4.30%. Le drawdown maximal est de 2.75% contre 12.43% initialement ce qui est excellent. La stratégie 2C-4 offre un bon rendement avec une volatilité annualisée qui avoisine les 8%. Il s'agit là d'un placement plus risqué mais

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

le drawdown maximal est encore bien contenu avec une valeur de 5.27%. La stratégie 2C-8 peut être considérée d'agressive avec un rendement moyenne annualisé de 25.93% pour une volatilité annualisée de 13.38%. Cependant le drawdown maximal est toujours inférieur à celui de la stratégie initiale.

Les trois nouvelle stratégies proposent un ratio de Sharpe supérieur à l'ORB optimisée. Celui-ci est compris entre 1.66 et 1.88 ce qui est très acceptable. Si la volatilité annualisée est prise comme valeur de risque, seule la stratégie 2C-8 est plus risquée que la stratégie initiale. Elles sont toutes moins risquées si le drawdown maximal est considéré comme valeur de risque.

5.5.2 Performances depuis Janvier 2016

Nous utilisons l'échantillon de données allant de Janvier 2016 à Décembre 2017 pour valider notre stratégie. Cet échantillon n'a jamais été utilisé et un unique essai a été effectué dessus. Nous sommes donc exactement dans les conditions d'une stratégie qui aurait été choisie fin 2015 pour tourner les deux années suivantes sur un portefeuille d'un million d'euros. Nous rappelons que les coûts de transaction sont pris en compte et le slippage est toujours estimé à un point pour chaque contrat. Les résultats sont résumés dans le tableau de la Figure 5.20.

	Stratégie 2C-2	Stratégie 2C-4	Stratégie 2C-8
Performance totale	6.12%	12.29%	24.64%
Rendement moyen annualisé	3.03%	6.05%	11.95%
Volatilité moyenne annualisée	2.04%	3.98%	7.61%
Ratio de Sharpe	1.56	1.56	1.59
Drawdown maximal	1.85%	3.59%	6.89%
Durée du drawdown maximal	42	42	42

FIGURE 5.20 – Performances de la gamme de stratégies 2C entre Janvier 2016 et Décembre 2017

La stratégie a continué de bien performer. Les rendements sont moins importants que précédemment, ce qui est certainement dû à une sur-optimisation inévitable du système sur l'échantillon 2007-2015. Néanmoins ils restent tout à fait correct. La bonne nouvelle vient également des mesures de risque. La volatilité et le drawdown ont diminué pour les trois stratégies. De plus, la durée du drawdown est nettement raccourcie. Ces résultats valident tout à fait la stratégie.

Le graphique de la Figure 5.21 présente les performances des stratégies au cours du temps. Les résultats sont très concluants. La performance se construit de manière linéaire. Le R^2 des trois courbes est de 96%. Le système a traversé de manière solide les deux

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

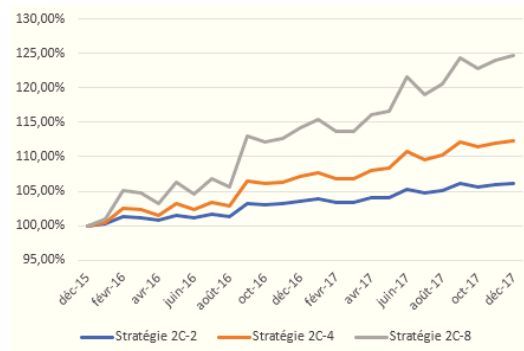


FIGURE 5.21 – Graphiques des performances de la gamme de stratégies 2C entre Janvier 2016 et Décembre 2017

années d'échantillon. Cette constance est un atout car il est appréciable pour un client de voir son capital peu fluctuer. Ces derniers résultats achèvent notre travail.

5.5.3 Limites et perspectives d'études

Tout au long de l'étude, nous avons fait de notre mieux pour éviter les biais cognitifs faussant la recherche de stratégies de trading. Nous avons limité le nombre de paramètres optimisés, limité le nombre de classifieurs testés, effectué une validation croisée pour valider les méthodes d'apprentissage et évité d'employer de la donnée future. Cependant, il y a encore des points sur lesquels nous pourrions améliorer les choses.

Choix des variables explicatives Nous nous sommes restreints à trois variables explicatives dont deux étaient déjà des variables préalablement optimisées. Cela n'a pas été un choix délibéré mais il est apparu que ces variables étaient les plus significatives pour les classifieurs alors qu'elles étaient déjà utilisées pour l'optimisation bien avant l'utilisation de l'apprentissage automatique. Un travail de recherche reste à faire dans la détection d'indicateurs pertinents pour la classification et dans la gestion du biais de data mining.

Méthode de validation croisée Notre démarche a consisté à apprendre sur 50% de l'échantillon et tester sur le reste. Il s'agit de la méthode la plus simple mais pas forcément la meilleure. Les méthodes de type "k-fold cross-validation" sont intéressantes car elles permettent de tester un classifieur sur toute la donnée de l'échantillon. Seulement, elles ne respectent pas le caractère chronologique de la donnée étudiée. Il existe une variante développée pour les séries temporelles où le classifieur est entraîné sur les k premières valeurs de l'échantillon puis est évalué sur sa capacité à bien classer la $(k + 1)$ -ème valeur.

Mais cette opération est très coûteuse en temps de calculs et fait se poser la question de la métrique de score.

Métriques de scores et optimisation Nous avons développé notre propre métrique de score pour l'optimisation de nos classifieurs, basée sur la précision et le rappel. Seulement ces métriques sont purement relatives au domaine de l'apprentissage automatique. Peut-être serait-il plus judicieux de comparer des classifieurs sur les performances ou le risque qu'ils génèrent. Comme nous l'avons vu, il n'est pas nécessaire d'avoir un excellent rappel ou la meilleure précision pour être la meilleure stratégie.

Régression ou classification Nos premières recherches se sont portées sur la capacité d'un classifieur à éviter les trades perdants. C'est pour cela que nous avons choisi la classification dans sa forme la plus simple. Il y aurait cependant d'autres manières de faire, en marquant les pertes anormalement grandes par exemple. Mais nous pourrions aussi nous porter sur des méthodes de régression et chercher à expliquer l'intensité des rendements d'une stratégie.

Apprentissage non-supervisé Nous avons fait le choix de l'apprentissage supervisé car il est plus accessible de prime abord et nous a permis de chercher dans une direction bien déterminée. Cependant, l'apprentissage non-supervisé a un vrai potentiel pour les recherches futures. En s'affranchissant de notre regard, les classifieurs pourraient trouver des relations que nous ne soupçonnons pas. Mais la tâche est plus compliquée. Nous avons beaucoup travaillé sur l'utilisation d'ACP sans succès.

Comprendre pourquoi ça marche Il s'agit là d'un point que nous n'avons pas abordé dans notre étude et qui pourtant est central. Pourquoi des classifieurs entraînés sur de la donnée entre Janvier 2007 et Avril 2012 arrive à améliorer significativement les résultats d'une stratégie encore aujourd'hui ? Il est déjà très difficile de comprendre les résultats d'un classifieur. Comprendre ceux d'une combinaison de méthodes est un travail nécessaire pour la suite.

5.6 Conclusion

Nous avons dans cette étude développé une gamme de trois stratégies performantes, viables et pertinentes pour de la gestion privée. Au-delà de ce développement, nous avons démontré deux choses.

CHAPITRE 5. APPLICATION DE MÉTHODES D'APPRENTISSAGE À UNE STRATÉGIE INTRADAY D'OPEN RANGE BREAKOUT SUR FUTURE DAX

Premièrement, il est possible de construire une stratégie de trading basée sur des éléments d'analyse technique. De plus cette stratégie résiste aux coûts d'exécution (transaction et slippage). L'Open Range Breakout est performante sur les 9 années d'étude malgré un risque important.

En second lieu, l'utilisation de méthodes d'apprentissage automatique améliore significativement de la stratégie ORB par un entrainement sur de la donnée passée. Il s'agit là d'une seconde évidence d'un apport informationnel de la part des prix passés.

Cette première expérience positive de l'utilisation de l'apprentissage automatique au trading algorithmique chez Das Bot ouvre les porte d'un vaste champ de recherche pour le développement de stratégies dans le futur.

Chapitre 6

Conclusion

Les travaux de cette thèse ont menés à plusieurs résultats. Du côté académique, nous avons proposé une nouvelle approche pour l'évaluation du prix d'une option sans avoir recours à une probabilité de risque neutre. Cela nous a amené à une condition faible d'absence d'opportunité d'arbitrage stipulant que le prix d'un call de payoff nul devrait être nul. Partant du modèle introduit dans ce travail, nous avons ensuite étudié la convergence du schéma fonctionnel avec volatilité locale qui en dérive lorsque la fonction de payoff n'est pas régulière. Enfin, nous avons proposé une méthode d'approximation de solutions d'équations aux dérivées stochastiques à coefficients non-Lipschitz par itérations de Picard déduite de la formule de Doléans-Dade.

La partie pratique a permis de développer une gamme de stratégies pour la gestion privée adaptée à différents profils d'investisseurs. Nous sommes partis d'une stratégie reposant sur des éléments techniques dont nous avons statistiquement démontré la pertinence. Puis à l'aide de méthodes d'apprentissage supervisé, nous avons construit trois stratégies viables et rentables aux couples rentabilité/risque différents.

Au-delà de ces résultats, je souhaite profiter de cette conclusion pour mettre en avant les contributions non visibles de cette thèse CIFRE au laboratoire de recherche de la société Das Bot. Lorsque nous avons débuté fin 2014, Das Bot était encore une jeune entreprise et avait concentré son travail sur le développement d'un système de trading automatisé entièrement propriétaire. Mon travail a été de mettre en place un ensemble d'outils permettant d'analyser le marché, de tester de nouvelles stratégies, de valider statistiquement les résultats et de gérer le risque.

Ces différents objectifs ont abouti à la création d'une librairie indépendante du robot de trading, écrite en C++ et interfacée avec Excel. La première étape a consisté à intégrer les outils indispensables du calcul numérique en C++ avec les bibliothèques Boost (et son

implémentation uBLAS pour l'algèbre linéaire entre autres), ALGLIB (algèbre linéaire et analyse de données) et CDFLIB (bibliothèque de fonctions de répartition notamment utilisée pour le travail sur les itérations de Picard). Nous avons également intégré l'excellente bibliothèque QuantLib pour les outils de finance quantitative. Dans un second temps nous avons implémenté un système de détection d'indicateurs et de patterns afin d'étudier le marché et de déceler les éventuels biais exploitables.

Un très gros effort a été fourni sur le développement d'un système de backtest de stratégies intra-day. Partant d'une structure générique, il est désormais possible de composer une stratégie basée sur la détection d'un ou plusieurs indicateurs. Entièrement orienté objet et intégrant des solutions de design pattern, le système est hautement réutilisable et limite au maximum les erreurs de programmation lorsqu'on manie le code.

Un second système plus classique pour la gestion de portefeuille a été implémenté lorsque Das Bot a développé des stratégies pour la société de gestion Sagara Financière. Nous avons notamment effectué un travail sur des stratégies de type Risk Parity ce qui a mené à la production d'un modèle inédit et propriétaire. Plus généralement, ces travaux ont abouti au lancement de deux fonds d'investissement depuis Septembre 2017 et un troisième est à venir d'ici cet été.

Enfin de nombreux outils ont été développés pour étudier le risque et calculer les valeurs de risque essentielles. Les outils sont intégrés directement dans Excel ce qui les rend facilement utilisables et intuitifs.

Au cours des recherches, nous nous sommes dirigés vers l'apprentissage automatique ce qui nous a amené à utiliser le langage Python et ses bibliothèques Pandas (analyse de données), NumPy (mathématiques) et scikit-lib (machine learning). La visualisation de la donnée a été faite avec l'application Glue et nous avons développé une série de Notebook avec le logiciel Jupyter permettant de manier les méthodes d'apprentissage même sans connaissances préalables en programmation.

En conclusion, je pense que les objectifs qui avaient été fixés par Das Bot pour cette thèse CIFRE ont tous été remplis. En près de quatre ans, l'entreprise a grandi et a su montrer sa capacité à développer, implémenter et exécuter des stratégies avec des outils et des modèles propriétaires. De plus, les résultats obtenus ouvrent de nouvelles portes et laissent entrevoir pour le futur de belles perspectives d'évolution.

Bibliographie

- [1] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis : A Hitchhiker's Guide, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 3rd edition, 2006.
- [2] Baptiste, J. and E. Lépinette (2018). Diffusion equations : convergence of the functional scheme derived from the Binomial tree with local volatility for non smooth payoff functions. <https://hal.archives-ouvertes.fr/hal-01507267>.
- [3] Barron, E. N., Cardaliaguet, P. and R. Jensen.(2003). Conditional Essential Suprema with Applications. *Appl Math Optim (48)* 229-253.
- [4] Bensaid, B., Lesne J.P., Pagès H. and J. Scheinkman.(1992). Derivative asset pricing with transaction costs . *Math.Fin.* **2**, 63-86.
- [5] Bernardi C., Maday Y. Spectral methods. In Handbook of numerical analysis, Vol. V, Handb. Numer. Anal., V, 209-485. North-Holland, Amsterdam (1997).
- [6] Black F., Scholes M. The pricing of options and corporate liabilities. The Journal of Political Economy, 81 , 637-659 (1973).
- [7] Bouleau N., Lépinette D. Numerical methods for stochastic processes. John Wiley and Sons Inc.(1994).
- [8] Brenner P., Thome V. and Wahlbin Lars B. Besov spaces and applications to difference methods for initial value problems. Lecture Notes in Mathematics 434, Springer-Verlag, Berlin-Heidelberg-New York 1975.
- [9] Brigo D., Mercurio F. A mixed-up smile. Risk, 13(9), 123-126 (2000).
- [10] Carassus, L., Gobet, E. and E. Temam. (2006). A class of financial products and models where super-replication prices are explicit “*International Symposium, on Stochastic Processes and Mathematical Finance*” at Ritsumeikan University, Kusatsu, Japan, March.

- [11] Carassus, L. and T. Vargiolu (2017). Super-replication price : it can be ok. Submitted.
- [12] Cherny A.S. Pricing with coherent risk. *Theory Probab. Appl.*, 52 (3), 389-415.
- [13] Ciarlet P.G. The finite element method for elliptic problems. North- Holland, Amsterdam, 1978.
- [14] Cox J., Ross S.A., Rubinstein M. Option pricing : a simplified approach, *Journal of Financial Economics*, 7 (1979), 229-263.
- [15] Crabel T. (1990). Day Trading With Short Term Price Patterns and Opening Range Breakout. Traders Press.
- [16] Dalang E.C., Morton A. and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics and Stochastic Reports*, 1990, 29, 185-201.
- [17] Dana R.A., Jeanblanc Monique. Financial markets in continuous time. Springer-Verlag Berlin Heidelberg (2003).
- [18] Darses S., Lépinette E. Limit theorem for a modified Leland hedging strategy under constant transaction costs rate. Inspired by Finance, The Musiela Festschrift, Eds. Yu. Kabanov, M. Rutkowski, T. Zariphopoulou. Springer, 159-199 (2014), Springer.
- [19] Delbaen F., Shirakawa H. A note on option pricing for constant elasticity of variance model. *Asia-Pacific Financial Markets*, 2, 2, 85-99(2002).
- [20] Delbaen F. and W. Schachermayer. *The Mathematics of Arbitrage*. Springer Finance, 2006.
- [21] Denis E. Marchés avec coûts de transaction : approximation de Leland et arbitrage. <http://indexation.univ-fcomte.fr/nuxeo/site/esupversions/8d5694b7-6409-413b-882a-37f2d7295a07>.
- [22] Desmond J. H., Xuerong M. Convergence, non-negativity and stability of a new Milstein scheme with applications to finance. *Discrete and continuous dynamical systems, series B*, 18, 8, 2083-2100 (2013).
- [23] Diener F., Diener M. Asymptotics of the price oscillations of a European call option in a tree model. *Mathematical Finance* 14(2004), 271-293.

- [24] Donninger C. (2014). An Investigation of Simple Intraday Trading Strategies.
- [25] Dupire B. Pricing and hedging with smiles. *Mathematics of Derivative Securities*. Cambridge Univ. Press, 15, 103-111 (1997).
- [26] Eymard R., Gallouet T., and Herbin R. Finite volume methods. In *Handbook of numerical analysis*, Vol. VII, *Handb. Numer. Anal.*, North-Holland, Amsterdam (2000), VII, 713-1020.
- [27] Fama E. F. (1965). The Behavior of Stock-Market Prices, *The Journal of Business* 38, 34-105.
- [28] Fama E. F. (1969). Efficient Capital Markets : A Review of Theory and Empirical Work, *The Journal of Finance* 25, 383-417.
- [29] Föllmer H. and D. Kramkov (1997). Optional Decompositions under Constraints, *Probability Theory and Related Fields* 109, 1-25.
- [30] Friedman A. *Stochastic Differential Equations And Applications*. Dover Books on Mathematics (1975).
- [31] Funahashi H. A chaos expansion approach under hybrid volatility models. *Quantitative Finance*, 14, 11, 1923-1936 (2014).
- [32] Giles Michael B. Improved multilevel Monte Carlo convergence using the Milstein scheme. *Monte Carlo and quasi-Monte Carlo methods*, 343-358, Springer, Berlin 2008 (2006).
- [33] Harrison J.M. and D.M. Kreps (1979). Martingale and Arbitrage in Multiperiods Securities Markets, *Journal of Economic Theory* 20, 381-408.
- [34] Harrison J.M. and S. Pliska (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading, *Stochastic Processes and their Applications* 11, 215-260.
- [35] C. Hess. Set-valued integration and set-valued probability theory : An overview. E. Pap, editor, *Handbook of Measure Theory*, Elsevier, 2002, 14, 617-673.
- [36] Holmberg U., Lönnbark C., Lundström C.(2012). Assessing the profitability of intraday opening range breakout strategies. *Finance Research Letters* 10.
- [37] Hsia C-C. On binomial option pricing. *Journal of Financial Research*, 6 (1983), 41-5.

- [38] Hull J., White A. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42, 281-300(1987).
- [39] Hutzenthaler M., Jentzen A. On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. arXiv :1401.0295, 1-41(2014).
- [40] Hutzenthaler M., Jentzen A. and Kloeden Peter E. Strong and weak divergence in finite time of Euler method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 467 :1563-1576 (2011).
- [41] Hutzenthaler M., Jentzen A., and Kloeden Peter E. Divergence of the multilevel Monte Carlo Euler method for nonlinear stochastic differential equations. *Ann. Appl. Probab.*, 23(5) :1913-1966 (2013).
- [42] Kabanov Y. and Denis E. Mean square error for the Leland-Lott hedging strategy : convex pay-off. *Finance and Stochastics*. 14 (2010),4, 626-667.
- [43] Kabanov Y. and M. Safarian. *Markets with transaction costs*. Mathematical Theory. Springer-Verlag, 2009.
- [44] Kabanov Y. and E. Lépinette. Essential supremum with respect to a random partial order. *Journal of Mathematical Economics*, 49 (2013), 6, 478-487.
- [45] Kijima M., Funahashi H. A chaos expansion approach for the pricing of contingent claims. *Journal of Computational Finance*, 18, 3, 27-58 (2015).
- [46] Kloeden P. E., Platen E. *Numerical solution of stochastic differential equations*. Springer-Verlag Berlin Heidelberg, *Stochastic Modelling and Applied Probability* (23, 1992).
- [47] Kreps D. (1981) Arbitrage and equilibrium in economies with infinitely many commodities, *Journal of Mathematical Economics* 8, 15-35.
- [48] Lépinette, E. and I. Molchanov. Conditional cores and conditional convex hulls of random sets. <https://arxiv.org/abs/1711.10303>
- [49] Lépinette, E. and T. Tran. Approximate hedging in a local volatility model with proportional transaction costs. *Applied Mathematical Finance*, 21 (2014), 4, 313-341.

- [50] Lindsay A.E., Brecher D.R. Simulation of the CEV process and the local martingale property. *Mathematics and Computers in Simulation*, 82, 5, 868878 (2012).
- [51] Lo A. W. and MacKinlay A. C.(1988). Stock market prices do not follow random walks : evidence from a simple specification test.
- [52] Lo A. W. and MacKinlay A. C.(2002). *A Non-Random Walk Down Wall Street*. Princeton University Press.
- [53] Lundström C.(2013). Day Trading Profitability across Volatility States : Evidence of Intraday Momentum and Mean Reversion.
- [54] Mehrdoust F., Babaei S. and Fallah S. Efficient Monte Carlo option pricing under CEV model. *Communications in Statistics - Simulation and Computation*, DOI 10.1080/03610918.2015.1040497 (2015).
- [55] Milstein G.N. The Probability Approach to Numerical Solution of Nonlinear Parabolic Equations. *Numerical methods for partial differential equations*, 18, 4 (2002), 490-522.
- [56] Neuenkirch A. and Szpruch L. First order strong approximations of scalar SDEs defined in a domain. *Numerische Mathematik*, 1-34 (2014).
- [57] Nualart D. *The Malliavin Calculus and Related Topics*. Probability, its applications (2006). Springer Berlin Heidelberg.
- [58] Prigent J.L. *Weak convergence of financial markets*. Springer-Verlag Berlin-Heidelberg (2003).
- [59] Quarteroni A. An introduction to spectral methods for partial differential equations. In *Advances in numerical analysis, I* (Lancaster, 1990), Oxford Sci. Publ., 96-146. Oxford Univ. Press, New York (1991).
- [60] Richtmyer R.D., Morton K.W. *Difference methods for initial value problems*. Robert E. Krieger Publishing Co. Inc., Malabar, FL, second edition (1994).
- [61] Rockafellar, R.T. (1972) : *Convex analysis*, Princeton University Press.
- [62] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998. ISBN 3-540-62772-3.
- [63] Samuelson P. A. (1965). Proof That Properly Anticipated Prices Fluctuate Randomly, *Industrial Management Review* 6, 41.

BIBLIOGRAPHIE

- [64] Schroder M. Computing the constant elasticity of variance option pricing formula. *The Journal of Finance*, 44, 1, 211-219 (1989).
- [65] Viitasaari L. Rate of convergence for discrete approximation of option prices. <https://arxiv.org/abs/1207.6756>.
- [66] Walsh J.B. The rate of convergence of the binomial tree scheme. *Finance and Stochastics* 7 (2003), 337-361.

Résumé

Cette thèse CIFRE est composée d'une partie théorique abordant des problèmes d'évaluation de produits dérivés et d'une partie pratique portant sur la construction d'une stratégie de trading algorithmique basée sur des méthodes d'apprentissage automatique.

Nous nous demandons s'il est possible d'évaluer le prix d'une option sans avoir recours à une probabilité de risque neutre. Nous proposons une nouvelle approche nous menant à une condition faible d'absence d'opportunité d'arbitrage qui stipule que le prix d'un call de payoff nul devrait être nul. Dans un second travail, nous étudions la convergence du schéma fonctionnel dérivé du modèle Binomial et plus généralement du modèle introduit précédemment avec volatilité locale lorsque la fonction de payoff n'est pas régulière. Nous illustrons les résultats par des exemples et comparons les performances avec d'autres méthodes de résolution. La partie théorique s'achève avec l'étude d'une méthode d'approximation de solutions d'équations aux dérivées stochastiques par itérations de Picard déduite de la formule de Doléans-Dade. La méthode est appliquée au pricing d'options asiatiques dont le sous-jacent a pour dynamique un processus CEV.

Dans la partie pratique, nous revisitons la stratégie Open Range Breakout qui repose sur des éléments techniques visant à tirer parti d'une tendance des prix d'un actif à évoluer dans une certaine direction plutôt qu'à évaluer et prédire le prix d'un actif. Nous appliquons des méthodes d'apprentissage supervisé et montrons qu'il est possible de créer une stratégie viable et rentable prenant en compte les coûts d'exécution. Nous terminons par la présentation d'une gamme de stratégies répondant à plusieurs profils d'investisseurs.

Mots Clés

Apprentissage automatique, Apprentissage supervisé, Conditions de non-arbitrage, Equations aux dérivées partielles, Modèles de marchés financiers, Modèle Binomial, Pricing d'options, Prix de sur-réplication, Trading algorithmique

Abstract

This thesis consists of a theoretical section addressing derivatives pricing problems followed by an empirical section focused on designing a machine learning oriented algorithmic trading strategy.

We question whether it is possible to price a contingent claim option without using a risk-neutral probability measure. We establish a new approach leading to a weak condition of absence of arbitrage opportunity which asserts that the price of the zero claim should be zero. We then study the convergence of the functional scheme derived from the Binomial tree financial model and more generally from the previously introduced model with local volatility when the payoff function is not smooth. We illustrate our results with numerical examples and compare performances with other resolution methods. We conclude this theoretical part by studying an approximation method based on Picard iterations deduced from the Doléans-Dade exponential formula. We illustrate this by using this method to the price an Asian-style contingent claims in the CEV model.

In this empirical section, we revisit the Open Range Breakout strategy whose trade signal relies on a set of technical rules aimed at profiting from a directional trend in the prices of an asset (as opposed to establishing a valuation of that asset's price). We apply supervised learning methods to show that it is possible to create a profitable and realistic strategy taking execution costs into account. We conclude this section by presenting a family of strategies that can fit various investors' profiles.

Keywords

Algorithmic Trading, Binomial tree model, Diffusion partial differential equations, Financial market models, Machine learning, No-arbitrage condition, Option pricing, Super-hedging prices, Supervised learning