Department of Applied Mathematics,
Computer Science and Statistics
Ghent University

School of Mathematics and Statistics
Central South University

Quantifying Model Uncertainty in Financial Markets

Xianming SUN

Promotors: Prof. dr. Michèle Vanmaele
Prof. dr. Siqing Gan

Thesis submitted in fulfillment of the requirements for the degree of
Doctor of Science: Mathematics
Academic year 2015-2016

Department of Applied Mathematics,
Computer Science and Statistics
Ghent University

School of Mathematics and Statistics
Central South University

Quantifying Model Uncertainty in Financial Markets

Xianming SUN

Promotors: Prof. dr. Michèle Vanmaele
Prof. dr. Siqing Gan

Examination Committee:

Prof. dr. Songhe Song (chair)
Prof. dr. Marnix Van Daele (chair)
Prof. dr. Haibo Chen
Prof. dr. Jan Dhaene
Prof. dr. Asma Khedher
Prof. dr. Qingsong Xu

Public defense on March 30, 2016

Acknowledgement

First and foremost, I am deeply indebted to my supervisors, Professor Michèle Vanmaele and Professor Siqing Gan, for their invaluable guidance, encouragement and support.

During my Ph.D. research, Professor Vanmaele always gave her feedback on my work in time. The notes and comments that she made on the working papers greatly promoted me to make progress. It is my greatest honour to have her as my first audience whenever I was going to give a presentation. She also did a lot of favours for me in daily life when I was in Ghent. I owe her so much.

When I was a master student, Professor Gan introduced me to the research area of numerical analysis, especially the numerical solution of stochastic differential equations. He also encouraged me to do interdisciplinary researches. His guidance and encouragement paved the way for me to do the Ph.D. research on some topics in computational finance. His insightful comments inspired me to come up with some ideas in this thesis. I greatly appreciate all what he did for me.

I appreciate the comments from all the members of my examination committee. I would like to thank my co-authors: Professor Asma Khedher, Thorsten Schulz and Dorien Haesen for the papers we completed together. I also would like to thank Professor Wim Schoutens and Professor Florence Guillaume for sharing data with us.

I would like to thank Catherine, a member of our financial maths group at the UGent. She helped me to find a comfortable living place in Ghent, to register at the city hall and so on. She also took care of the travelling details when we attended conferences in Paris, Brussels and Saint Petersburg. Many thanks also go to other members in our group: Professor David Vyncke, Hilmar and Mina.

I enjoy working with all the members in the research group of Professor Siqing Gan. Thanks to all of you for your company and interesting discussions, especially Lin Hu, Wenhao Li, Qiyong Li, Xiaojie Wang, Xu Yang, Jinran Yao and Zhengwei Yin. Many thanks are also addressed to the members in the stochastic analysis research group at the CSU, especially Professor Zaiming Liu and Professor Xiang Lin (ZJSU) who used to work at the CSU. I learned so much from your courses.

I am grateful to my colleagues in the UGent and the CSU, especially Ann, Herbert, Hilde, Wouter and Pei (CSU) for their kind help in the office.

My sincere gratitude also goes to all my friends. It would be a long list of their names, but I would like to list those who used to join me for lunch: Bashir, Beatriz, Binwei, Catherine, Diego, Guorong, Guoxiang, Hilmar, Holger, Jens, Jian, Johan, Kainan, Mushthofa, Serge, Zhenxing, Zhifei and Zhiyue.

I acknowledge the financial support by the China Scholarship Council under the CSC-grant (No. 201306370129) and by the Ghent University under the BOF-cofunding (No. 01SC0414) for Chinese Ph.D. students holding a CSC-grant.

I am thankful to my family. My parents have taught me that hardwork, perseverance and honesty pave the way for the future. I always keep these words in my mind. They gave me full support to study abroad. I am indebted to my sister and brother-in-law for their support and encouragement. They kept on saying that "do not worry, we are at home", and they did take care of the whole family. Mingming, my girl friend, always stands behind me. Thanks for her encouragement and support.

Xianming Sun
Changsha
March 29, 2016

Table of Contents

List of Figures

List of Tables

1 Introduction

This chapter presents my motivations for working on numerical methods involving model uncertainty in financial markets. These motivations are accompanied with a brief introduction on the state-of-the-art theory of financial risk management and derivative pricing under model uncertainty. A thorough literature review will be given at the beginning of each chapter. The topics of the thesis are introduced at the end of this chapter.

1.1 State of the art and motivations

Probabilistic models for financial markets play a crucial role in financial economics. These models can take the form of probability distributions or stochastic processes. The first continuous-time model for a stock price process can be traced back to the Brownian motion model proposed by Louis Bachelier in his PhD thesis in 1900. Arguing that prices should remain positive, Samuelson [150] proposed to model a stock price process with a geometric Brownian motion in 1965. Working with this model, Black and Scholes [24] and Merton [132] opened the door for derivative pricing with the no-arbitrage approach in 1973. Thereafter, the last more than four decades has witnessed tremendous advance in financial modelling, which both shaped and were shaped by

the revolutionary changes in the structure of world financial markets and institutions. Nowadays, there are various kinds of probabilistic models for risky asset prices, such as jump-diffusion models (see e.g. [133]), stochastic volatility models (see e.g. [103]) and so on. A *natural* question is how to select a probabilistic model to *quantitatively* predict the future states of financial markets or an asset price process. The *philosophical* story behind this *natural* question is whether we can base a financial decision on the quantitative analysis with a single probabilistic model.

Due to the immutability of nature's laws, pure mathematical logic can often yield useful insights about physical phenomena, and mathematical models can often precisely predict the evolvement of a physical system. On the contrary, mutable financial markets are mixed up with heterogeneous believes of investors and their decisions. Investors make financial markets so complex that any single probabilistic model can *hardly* be capable of quantitatively characterizing the *whole* picture of financial markets in the future. Otherwise, the single model characterizing the whole picture of financial markets allows the investor to fully control her risk position, and financial markets would have become much less complex than what they are. A prudential and experienced risk manager can easily acknowledge this limitation of using any single probabilistic model. Because of the conflict between the limited capability of any single model and the complexity of financial markets, model ambiguity comes on stage. Philosophically speaking, model ambiguity is ubiquitous whenever a probabilistic model is used in financial economics.

Except for the philosophical arguments for model ambiguity in financial markets, both the *empirical* facts and the *simulation* issues in applications of stochastic models do highlight model ambiguity in financial economics. For a given kind of stochastic model, its parameters can be calibrated by getting the model to match the market prices of liquidly traded derivatives (also called benchmark derivatives). This procedure is usually referred to as model calibration (see the Appendix of Chapter 6). There are two issues with model calibration: first, different kinds of models can be perfectly calibrated to the same market data [151]; second, different calibration methods may yield different estimations for each parameter of a specific model [91]. These empirical facts, together with a limited knowledge of the market dynamics, confront an agent with ambiguity about which model is the best one to value a target derivative, especially when the well-calibrated models may lead to quite different values for the target derivative. In addition, in the simulation of semimartingale models, different approaches can be used to approximate the stochastic systems driven by jumps with infinite mass in the tail [115].

In this setting, it is necessary to investigate whether the hedging strategies are robust with respect to the selection of approximation approaches.

Although philosophical arguments, empirical facts and simulation issues show that model ambiguity is ubiquitous in financial economics, it is not the impasse of probabilistic models for risk management and derivative pricing. One of the key ideas to select a probabilistic model is, not to expect a precise prediction, but to capture more uncertainties of interest, because any candidate model typically exhibits multiple levels of uncertainty. The premise of successfully implementing any model is to recognize its key characteristics and the boundaries of its validity. From this point of view, a set of plausible models, rather than a single model, can be simultaneously employed to make a robust investment decision. As for model ambiguity characterized by a set of plausible models, Knight [117] further categorized model ambiguity into two groups: model risk and model uncertainty. Model risk is associated with the setting in which the probabilities of the candidate models to be true are known, while model uncertainty arises from a lack of knowledge of the probabilistic information on these models. Parameter risk and parameter uncertainty can be defined in the similar way. In this thesis, we focus on model (parameter) uncertainty in the sense of Knight [117].

In academia, model uncertainty has been pioneered by several intellectual giants. Lars Peter Hansen, a Nobel laureate in 2013, and his coauthors proposed to use entropy to select a set of probabilities to capture the uncertainty of returns when investigating macroeconomics (see, e.g. [96–98]). Lyons [126], Avellaneda, Levy and Parás [11] proposed the famous volatility uncertainty model in the context of financial derivative design. The nonlinear expectation theory (g-expectation and G-expectation [143]) proposed by Shige Peng is a general framework to investigate volatility uncertainty as well as drift uncertainty. Especially, the G-Brownian motion defined on a sublinear expectation space can be used to characterize volatility uncertainty, including the volatility uncertainty model in [11, 126] as special cases. Denis and Martini [60] proposed the capacity-based framework for volatility uncertainty. We refer to [58] for the strong link between the capacity-based framework and the G-expectation framework. These frameworks all involve the functional $\tilde{\mathbb{E}}$ defined on some special space $\mathcal{L}$ of random variables,

$$\tilde{\mathbb{E}}[X] := \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[X], X \in \mathcal{L},$$

where $\mathbb{E}^Q$ is the classical mathematical expectation under the probability measure Q. The set $\mathcal{Q}$ of probability measures represent different models for the risk factor X. Compared with employing one model Q for the risk factor X, selecting a set $\mathcal{Q}$ of

plausible models for X captures its model uncertainty. Different approaches can be used to select the set of models to account for model uncertainty. Hansen and Sargent [97] employed relative entropy to select $\mathcal{Q}$, the elements of which are not far from a reference model $\mathbb{P}$. The probability measures in $\mathcal{Q}$ are absolutely continuous with respect to $\mathbb{P}$ in this setting, which is similar as the framework of g-expectation accounting for drift uncertainty of a risk factor [41]. However, in the framework of G-expectation, there is no reference probability measure $\mathbb{P}$ that dominates all probability measures in $\mathcal{Q}$. The elements in $\mathcal{Q}$ are singular with each other [72]. Due to the limited information about risk factors, volatility uncertainty is a realistic situation in financial economics [72, 173]. The singularity among probability measures in the setting of volatility uncertainty makes the related stochastic analysis more difficult than that in the setting of drift uncertainty. The theory of stochastic analysis in a nonlinear expectation space is attracting more and more attention in academia. In a more abstract setting than the G-expectation framework, module-based approach may provide another approach for stochastic analysis under model uncertainty (see e.g. [79, 93]).

Quantifying model uncertainty is not only an interesting topic in academia, but also is of great importance in financial industry. The financial crisis of 2007–2009 has aroused the financial industry and regulation sectors to take model uncertainty into account, see e.g. the regulations set by the Federal Reserve Board [78] and the Basel II market risk framework [23]. The interplay between theory and practice stimulates a fast growing literature on risk management and derivative design under model uncertainty (see e.g. [17,18,22,31,42,48,54,59,71–73,98,100,101,114,121,138,139,173]). In this literature, a financial analysis is usually carried out with relatively simple models accounting for model uncertainty, in which setting the corresponding financial problems admit analytic solutions. However, in some sophisticated models with parameter uncertainty (see e.g. [91, 151]), the financial problems may no longer admit analytical solutions. One issue is how to quantify model uncertainty in more realistic and practical models, in which it may be difficult to derive some analytical results.

Except for directly selecting a set of plausible models for the risk factors with the aforementioned entropy method [97], drift and volatility method [143] and calibration-based method [91, 151], distorted expectation provides an indirect way to account for model uncertainty. Distorted expectation is defined by

$$\mathbb{E}_\Psi[X] := \int_{-\infty}^{+\infty} x \, \mathrm{d}\Psi(F_X(x)) = \inf_{Q \in \mathcal{Q}_\Psi} \mathbb{E}^Q[X],$$

where $\mathcal{Q}_\Psi$ is a set of probability measures associated with a concave distortion function

Ψ. Note that $-\mathbb{E}_\Psi[-X] = \sup_{Q \in \mathcal{Q}_\Psi} \mathbb{E}^Q[X]$. Hence, an agent can work with $-\mathbb{E}_\Psi[-X]$, if she has to use the supreme of the expectation of X over a set of probability measures. $\mathcal{Q}_\Psi$ accounts for model uncertainty. In this setting, calculating distorted expectations is the key point to quantify model uncertainty. However, there is no analytical formula for the distorted expectation because the distorted probability density function does not admit an analytical formula in the general case. Hence, we have to rely on numerical methods to calculate distorted expectations. An efficient numerical method for distorted expectations can find extensive applications, because the distorted expectation is not only associated with model uncertainty but also risk measures [80], conic finance [68, 69, 127–131]), behavioural finance [99] and so on.

Up to now, we have introduced the background on model uncertainty in financial markets, and the key issues to quantify model uncertainty. Note that we presumed that the agent had a reference probabilistic model for the asset price process in the aforementioned settings. From the contrary point of view, if no parametric model is used, we will not come across model uncertainty at all. A few literature investigates the model-free pricing and hedging method for financial derivatives (see, e.g. [39, 40, 52, 84, 105–107]). These approaches assume that an agent has enough data of the market prices of derivatives written on the same or correlated underlying asset as the derivative of interest. The hedging strategies usually correspond to the solution of an optimization problem, which also provides value bounds for the target derivative. If the spread between the upper and lower value bounds is small enough, it is reasonable to use these value bounds to approximate the target derivative value. Otherwise, these value bounds have some limitation in applications. Hence, an interesting question is what we can get from the comparison between the model-based quantities and their model-free counterparts.

Motivated by the aforementioned computational issues related to model uncertainty in financial markets, we will focus on numerical methods to quantify uncertainty in financial markets. This theme relates the thesis to the following literature [5, 13, 49, 89, 94], as well as the related literature on the theories involving model uncertainty.
Arai and Suzuki [5] derived an integral representation, without implementable numerical method, of the local risk-minimizing (LRM) strategy for vanilla options, Asian options in a class of semimartingale models. In Chapter 3, under similar models, we formulate the LRM strategies with backward stochastic differential equations and propose an implementable numerical method to simulate these LRM strategies. In addition, we will investigate the robustness of the LRM strategies with respect to model uncertainty.

Cont [49] proposed a coherent uncertainty measure on model uncertainty embedded in a derivative price in terms of its value bounds. We are interested in how to calculate these values bounds in general cases. Chapter 4 presents an efficient method which can be used to calculate not only the value bounds of a derivative in general cases, but also the entropy of the derivative value. In [13, 94], the authors proposed the so-called robust calibration methods based on pricing error and Bayesian approach, respectively. These two calibration methods generate a large ensemble of model parameters, and the derivative price should be calculated with each realisation of model parameters. This procedure is usually time-consuming. However, it can be accelerated by our method proposed in Chapter 4.
Glasserman and Xu [89] proposed a Monte Carlo method to calculate risk measures under the worst-case probability measure within a constrained set of probability measures. In Chapter 5 we propose an analytical approximation method for distortion risk measures as long as the density function of the risk factor can be calculated under the probability measure of interest, such as the worst-case probability measure. We show that our method is more efficient than the Monte Carlo method when calculating distortion expectations.

More precisely, this thesis covers the following topics on numerical methods for quantifying model uncertainty, reducing model uncertainty and detecting model mis-specifications:

1. Quantifying model uncertainty when a parametric stochastic model is used.

   - Discretising locally risk-minimising (LRM) strategies for some derivatives in a class of semimartingale models, and investigating the robustness of the LRM strategies with respect to model uncertainty.
   - Quantifying model uncertainty embedded in financial derivatives in the general setting.
   - Calculating distorted expectations in an efficient way.

2. Reducing the impact of model uncertainty on the derivative value when the Monte Carlo method has to be used in derivative pricing.

3. Detecting model mis-specification by comparing model-free quantities and the model-based quantities of Asian options.

1.2 Outline of the thesis

The previous section presents the motivations for selecting model uncertainty as the theme of my thesis. A thorough literature review and motivations concerning the studied topics can be found at the beginning of each chapter. The topics of the following chapters can be summarized as follows.

Chapter 2 introduces some preliminaries on stochastic calculus, Malliavin calculus and derivative pricing.

Chapter 3 formulates, using backward stochastic differential equations, local risk-minimizing (LRM) strategies for Asian options, spread options and basket options in a class of semimartingale models. We propose a discretising method for the LRM strategies and investigate its convergence. We also investigate the robustness of the LRM strategies with respect to model uncertainty.

Chapter 4 proposes an efficient numerical method to quantify uncertainty embedded in financial derivatives. We focus on parameter uncertainty, which is characterized by intervals for the model parameters. Each parameter can take any value in its interval. Based on the Monte Carlo method and the Smolyak algorithm, the proposed method can be used to efficiently calculate an ensemble of the derivative prices associated with a given ensemble of the model parameters. The ensemble of the derivative prices can be used to calculate a coherent uncertainty measure [49] and the entropy measure. The entropy measure works as an auxiliary measure of uncertainty in this context.

Chapter 5 proposes an efficient analytical approximation method for distorted expectations, the so-called T-COS method. It is about 2500 times more efficient than the standard Monte Carlo method. Since distorted expectations are crucial tools in model uncertainty, risk management, insurance, behavioural finance and so on, the proposed T-COS method could find extensive applications in these areas.

Chapter 6 empirically investigates whether the weighted Monte Carlo method can reduce the impact of model uncertainty. Given different calibrated Heston models, we compare the weighted Monte Carlo method with the standard Monte Carlo method in terms of the implied volatilities of vanilla options and the prices of exotic options. The results confirm that the weighted Monte Carlo method can effectively reduce the impact of model uncertainty on the derivative prices.

Chapter 7 proposes an efficient numerical method to accelerate the calculation of the comonotonocity upper bounds for discrete arithmetic Asian options. Then, we empirically investigate how to detect model mis-specification by comparing the model-based (optimal) value bounds with their model-free counterparts of a discrete arithmetic

Asian option. We address some practical issues, such as the non-uniqueness of the marginal distribution for the underlying asset in the model-free setting.

Chapter 8 concludes the main contents of this thesis and looks into the future of some promising topics on uncertainty in financial markets.

Summaries in English and Dutch are presented at the end of this thesis.

2 Preliminaries

This chapter presents some preliminaries on stochastic calculus, Malliavin calculus and derivative pricing. The readers familiar with these theories may skip this chapter. We refer to [64,110,159] for the details of these theories. Most of the notations and theorems in this chapter are based on these three books, unless it is stated otherwise.

2.1 Stochastic calculus

We assume as given a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F} := (\mathcal{F}_t)_{0 \le t \le T \le +\infty}$, which consists of a family of σ-algebras satisfying $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \le t$.

Definition 2.1. *A filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is said to satisfy the **usual hypotheses** if*

(i) $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets of $\mathcal{F}$;

(ii) $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$, all t, $0 \le t < \infty$; that is, the filtration $\mathbb{F}$ is right continuous.

We always assume that the usual hypotheses hold in this thesis.

A stochastic process X on $(\Omega, \mathcal{F}, \mathbb{P})$ is a family $(X_t)_{0\leq t\leq\infty}$ of $\mathbb{R}$-valued or $\mathbb{R}^d$-valued random variables. The process is said to be adapted if X_t is $\mathcal{F}_t$-measurable for each t. One property of the process X is said to hold almost surely, abbreviated as $a.s.$ if it holds with probability one. A stochastic process X is said to be **càdlàg** if all its paths are right-continuous ($\lim_{s\to t, s>t} X_s = X_t$) and have left limits, a.s. ($\lim_{s\to t, s<t} X_s = X_{t-}$). For a càdlàg process X, we define the jump at time t as $\Delta X_t := X_t - X_{t-}$.

On the set $\Omega \times [0,T]$, we define two σ-fields $\mathcal{O}$ and $\mathcal{P}$ generated by the adapted and càdlàg processes and the adapted and continuous process, respectively. On the set $\tilde{\Omega} := \Omega \times [0,T] \times \mathbb{R}^d$, we consider the σ-field $\tilde{\mathcal{O}} = \mathcal{O} \otimes \mathcal{B}(\mathbb{R}^d)$ (resp. $\tilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$), where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-field on $\mathbb{R}^d$.

Definition 2.2. *A process or a random set is called an **optional** (resp. predictable) **process** (resp. random set) if the process (resp. random set) is $\mathcal{O}$ (resp. $\mathcal{P}$)-measurable.*

Definition 2.3. *A function W on $\tilde{\Omega}$ that is $\tilde{\mathcal{O}}$ (resp. $\tilde{\mathcal{P}}$)-measurable is called an **optional** (resp. predictable) **function**.*

Definition 2.4. *A random variable $\tau : \Omega \to [0,\infty]$ of the filtration $\mathbb{F}$ is a **stopping time** if the event $\{\tau \leq t\} \in \mathcal{F}_t$ for $t \in [0,\infty]$.*

2.1.1 (Semi)martingales

Denote by $L^p(\Omega, \mathcal{F}, \mathbb{P})$ the set of real-valued random variables X satisfying $\mathbb{E}[|X|^p] < \infty$, for $p \in [1,\infty)$.

Definition 2.5. *A real-valued, adapted process $X = (X_t)_{0\leq t<\infty}$ is called a **martingale** (resp. supermartingale, submartingale) with respect to the filtration $\mathbb{F}$ if*

(i) $X_t \in L^1(\Omega, \mathcal{F}, \mathbb{P})$;

(ii) if $s \leq t$, then $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$, a.s. ($\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$, $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$).

Definition 2.6. *A process M is a **local martingale** if there exists an increasing sequence (τ_n) of stopping times such that $\lim_n \tau_n = \infty$ a.s. and such that each stopped process $M^{\tau_n} := M_{t\wedge\tau_n}$ is a martingale.*

We denote by $\mathcal{L}$ all local martingales and by $\mathcal{V}$ the set of all real-valued processes A with $A_0 = 0$ that are càdlàg, adapted and for which each path $t \to A_t(\omega)$ has a finite variation over each finite interval $[0,t]$. The variation of A is given by $\int |dA_s|$.

Definition 2.7. *A process X is called to be a **semimartingale** if it has the following decomposition*

$$X = X_0 + M + A,$$

where X_0 is finite-valued and $\mathcal{F}_0$-measurable, $M \in \mathcal{L}$ and $A \in \mathcal{V}$. The decomposition is called the canonical decomposition.

Definition 2.8. *A **special semimartingale** X is a semimartingale which admits the unique decomposition $X = X_0+M+A$ such that $M \in \mathcal{L}$ and that $A \in \mathcal{V}$ is predictable.*

Due to Theorem I 4.18 in [110] on the decomposition of local martingales, a semimartingale X with decomposition $X = X_0 + M + A$ admits an unique decomposition of the following form

$$X = X_0 + M^c + M^d + A,$$

where M^c and M^d are the continuous part and pure jump part of the local martingale M, respectatively. In this setting, we denote by $X^c := M^c$ and $X^d := M^d$ the continuous local martingale part and the discontinuous local martingale part of a semimartingale X.

We introduce the following sets to be used in the following sections:

- $\mathcal{H}^2$: all square-integrable martingales M such that $\sup_{t\in\mathbb{R}^+} \mathbb{E}[M_t^2] < \infty$.
- $\mathcal{H}^2_{loc}$: all locally square-integrable martingales.
- $\mathcal{S}$: the set of all semimartingales.
- $\mathcal{S}^2$: the set of all square-integrable semimartingales.
- $\mathcal{V}^+$: the set of all real-valued adapted and increasing processes A with $A_0 = 0$.
- $\mathcal{A}^+$: the subset of all $A \in \mathcal{V}^+$ that are integrable: $\mathbb{E}[A_\infty] < \infty$.
- $\mathcal{A}$: the subset of all $A \in \mathcal{V}$ that have integrable variation: $\mathbb{E}[Var(A)_\infty] < \infty$.
- $\mathcal{A}_{loc}$ (resp. $\mathcal{A}^+_{loc}$): the localized class constructed from $\mathcal{A}$ (resp. $\mathcal{A}^+$).

Definition 2.9. *Let $A \in \mathcal{A}_{loc}$. The **compensator** of A under a probability measure $\mathbb{P}$, denoted by A^p, is the unique predictable process such that $A - A^p$ is a $\mathbb{P}$-local martingale.*

Definition 2.10. *A **random measure** on $\mathbb{R}^+ \times \mathbb{R}^d$ is a family $\mu = (\mu(\omega; dt, dx : \omega \in \Omega))$ of non-negative measures on the Blackwell space $(\mathbb{R}^+ \times \mathbb{R}^d, \mathcal{R}^+ \otimes \mathcal{E})$ satisfying $\mu(\omega; \{0\} \times \mathbb{R}^d) = 0$.*

The integral $W * \mu(\omega)$ of an optional function W with respect to a random measure μ can be defined as follows

$$W * \mu(\omega) = \int_0^t \int_{\mathbb{R}^d} W(\omega, s, x)\mu(\omega; \mathrm{d}s, \mathrm{d}x),$$

if the right hand is finite; otherwise $W * \mu(\omega) = +\infty$.

Theorem 2.11 ([110]). *Let μ be an optional $\tilde{\mathcal{P}}$-σ-finite random measure. There exists a random measure, called the **compensator** of μ and denoted by μ^p, which is unique up to a $\mathbb{P}$-null set, and which is characterized as being a predictable random measure satisfying either one of the two following equivalent properties:*

(i) $\mathbb{E}[W * \mu^p_\infty] = \mathbb{E}[W * \mu_\infty]$ for every nonnegative $\tilde{\mathcal{P}}$-measurable function W on $\tilde{\Omega}$.

(ii) For every $\mathcal{P}$-measurable function W on $\tilde{\Omega}$ such that $|W| * \mu \in \mathcal{A}^+_{loc}$, then $|W| * \mu^p$ belongs to $\mathcal{A}^+_{loc}$, and $W * \mu^p$ is the compensator of the process $W * \mu$.

Moreover, there exists a predictable $A \in \mathcal{A}^+$ and a kernel $K(\omega, t; \mathrm{d}x)$ from $(\Omega \times \mathbb{R}^+, \mathcal{P})$ into $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that

$$\mu^p(\omega; \mathrm{d}t, \mathrm{d}x) = \mathrm{d}A_t(\omega)K(\omega, t; \mathrm{d}x).$$

Now we are ready to define an integer-valued random measure μ associated with the jumps of X

$$\mu(\mathrm{d}t, \mathrm{d}x) = \sum_s \mathbf{1}_{\{\Delta X_s \neq 0\}} \delta_{(s, \Delta X_s)}(\mathrm{d}t, \mathrm{d}x), \tag{2.1}$$

where $\mathbf{1}$ is the indicator function and δ_a denotes the Dirac measure at point a.

An important random measure is the Poisson measure, which is used to count the jumps of specified size associated with a stochastic process.

Definition 2.12. *A **Poisson measure** on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^d_0)$ ($\mathbb{R}^d_0 := \mathbb{R}^d - \{0\}$) , relative to a filtration $\mathbb{F}$, is an integer-valued random measure N such that*

(i) the positive measure m on $\mathbb{R}^+ \times \mathcal{B}(\mathbb{R}^d_0)$ defined by $m(A) = \mathbb{E}[\mu(A)]$ is σ-finite;

(ii) for every $s \in \mathbb{R}^+$ and every $A \in \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{B}(\mathbb{R}_0^d)$ such that $A \subset (s, \infty) \times \mathbb{R}_0^d$ and $m(A) < \infty$, the variable $\mu(\cdot, A)$ is independent of the σ-field $\mathcal{F}_s$.

(iii) $m(\{t\} \times \mathbb{R}_0^d) = 0$ for each $t \in \mathbb{R}^+$.

*The measure m is called the **intensity measure** of N. If $m(\mathrm{d}t, \mathrm{d}x) = \ell(\mathrm{d}x)\mathrm{d}t$, where ℓ is a positive σ-finite measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}_0^d))$, then N is called a **homogeneous Poisson measure**.*

Proposition 2.13 ([110]). *Let N be a Poisson measure on $\mathbb{R}^+ \times \mathbb{R}_0^d$, relative to the filtration $\mathbb{F}$, with intensity measure m. Then its compensator $N^p(\omega, \cdot) = m(\cdot)$.*

2.1.2 Stochastic integral

We denote by $\mathfrak{S}$ the set of all processes of the form:

$$\begin{cases} \text{either} & H = Y\mathbf{1}_{[\![0]\!]}, & Y \text{ is bounded } \mathcal{F}_0\text{-measurable}, \\ \text{or} & H = Y\mathbf{1}_{]\!]r,s]\!]},\ r < s, & Y \text{ is bounded } \mathcal{F}_r\text{-measurable}, \end{cases}$$

where $]\!]r, s]\!] := \{t \in \mathbb{R}^+ : r < t \leq s\}$, $[\![0]\!] = \{0\}$ and $\mathbf{1}_B$ is an indicator function on a set B.

Given a process $H \in \mathfrak{S}$ and a semimartingale X, we can define the integral process $H \cdot X$ as

$$(H \cdot X)_t \triangleq \int_0^t H_s \, \mathrm{d}X_s := \begin{cases} 0 & \text{if } H = Y\mathbf{1}_{[\![0]\!]}, \\ Y(X_{s\wedge t} - X_{r\wedge t}) & \text{if } H = Y\mathbf{1}_{]\!]r,s]\!]}. \end{cases} \tag{2.2}$$

Theorem 2.14 ([110]). *Let X be a semimartingale. The map $H \rightsquigarrow H \cdot X$ defined on $\mathfrak{S}$ as (2.2) has an extension, still denoted by $H \rightsquigarrow H \cdot X$, to the space of all locally bounded predictable processes H. $H \cdot X$ is called the stochastic integral of H with respect to X.*

Definition 2.15. *Let X, Y be semimartingales. The **quadratic variation process** of X, denoted be $[X, X] = ([X, X]_t)_{t \geq 0}$, is defined by*

$$[X, X] = X^2 - 2\int X_- \, \mathrm{d}X = X^2 - 2X_- \cdot X,$$

*(recall that $X_{0-} = 0$). The **quadratic covariation** of X, Y, also called the bracket process of X, Y, is defined by*

$$[X, Y] = XY - \int X_- \, \mathrm{d}Y - \int Y_- \, \mathrm{d}X = XY - X_- \cdot Y - Y_- \cdot X.$$

Definition 2.16. *The **predictable quadratic covariation** of two semimartingales X, Y is the compensator of the quadratic covariation $[X, Y]$. It is denoted by $\langle X, Y\rangle$ and therefore also called the angle brackets of X and Y . The short hand notation $\langle X\rangle$ will be used for the angle bracket $\langle X, X\rangle$.*

For a given local $\mathbb{P}$-martingale X, we denote by $L^2(X)$ the space of all $\mathbb{R}^d$-valued predictable processes S such that

$$||S||_{L^2(X)} := \left(E\left[\int_0^T S_t^\star \, \mathrm{d}[X]_t S_t \right] \right)^{1/2} < \infty,$$

where $(\cdot)^\star$ is the transpose of a matrix.

Proposition 2.17 ([156]). *Let X is a local $\mathbb{P}$-martingale. For any $S \in L^2(X)$, the process $\int S \, \mathrm{d}X$ is well-defined and in the space $\mathcal{M}_0^2(\mathbb{P})$ of square-integrable $\mathbb{P}$-martingales null at 0.*

Definition 2.18. *Two $\mathbb{P}$-semimartingales X and Y are called **orthogonal** under a measure $\mathbb{P}$ if $[X, Y]$ is a local martingale under $\mathbb{P}$. Hence the angle bracket $\langle X, Y\rangle = 0$.*

Denote by $\mathfrak{H}_t^d$ the class of all functions $h : \mathbb{R}^d \to \mathbb{R}^d$ which are bounded and satisfy $h(x) = x$ in a neighbourhood of 0. For a given semimartingale X and $h \in \mathfrak{H}_t^d$, we can define a special semimartingale

$$\begin{cases} \tilde{X}(h)_t = \sum_{s \le t} [\Delta X_s - h(\Delta X_s)] \\ X(h) = X - \tilde{X}(h). \end{cases}$$

The canonical decomposition of $X(h)$ is

$$X(h) = X_0 + M(h) + B(h), \text{ where } M(h) \in \mathcal{L}^d, \quad B(h) \in \mathcal{V}^d. \tag{2.3}$$

Definition 2.19. *Let $h \in \mathcal{L}^d$ be fixed. the **characteristicsof a semimartingale** X is the triplet (B, C, ν) which are defined as follows:*

(i) $B = (B^i)_{i \le d}$ *is a predictable process defined in (2.3).*

(ii) $C = (C^{ij})_{i,j \le d}$ *is a continuous process in $\mathcal{V}^{d \times d}$, namely,*

$$C^{ij} = \langle X^{i,c}, X^{j,c}\rangle,$$

where X^c is the continuous martingale part of X.

(iii) ν *is the compensator of the random measure* μ (2.1) *associated with the jump of a semimartingale* X.

Theorem 2.20 (Itô's formula [110])**.** *Let* X *be a* d*-dimensional semimartingale, and* f *is a function of class* C^2 *on* $\mathbb{R}^d$*. Then* $f(X)$ *is a semimartingale and*

$$\begin{aligned} f(X_t) =& f(X_0) + \sum_{i\le d} f_i^{'}(X_-)\cdot X^i + \frac{1}{2}\sum_{i,j\le d} f_{ij}^{''}(X_-)\cdot\langle X^{i,c}, X^{j,c}\rangle \\ &+ \sum_{s\le t}\left[f(X_s) - f(X_{s-}) - \sum_{i\le d} f_i^{'}(X_{s-})\Delta X_s^i\right]. \end{aligned}$$

In terms of the characteristic triplet (B, C, ν) *of a semimartingale* X*, Itô's formula takes the following form*

$$\begin{aligned} f(X_t) =& f(X_0) + \sum_{i\le d} f_i^{'}(X_-)\cdot X^i + \frac{1}{2}\sum_{i,j\le d} f_{ij}^{''}(X_-)\cdot C \\ &+ \int_0^t \left[f(X_{s-}+x) - f(X_{s-}) - \sum_{i\le d} f_i^{'}(X_{s-})x\right]\mu_s(\mathrm{d}x). \end{aligned}$$

2.2 Malliavin calculus

The Malliavin derivative of a random variable $F = F(\omega)$, $\omega \in \Omega$, on a completed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ can be regarded as the derivative with respect to ω. It plays a critical role in investigating the regularity of solutions of stochastic differential equations. We introduce the preliminaries of Malliavin calculus in the Brownian motion case. In this case, there are several ways to define the Malliavin derivative. Here we follow the approach based on the chaos expansion. One may refer to [64, 137, 144] for more details.

2.2.1 The Wiener–Itô chaos expansion

Fix $T > 0$. Let $W(t) = W(t,\omega)$, $t \in [0,T]$, $\omega \in \Omega$ be an one-dimensional Wiener process (Brownian motion) on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $W(0,\omega) = 0$, $\mathbb{P}$-a.s..

For any $t \in [0,T]$, Let $\mathcal{F}_t$ be the σ-algebra generated by $W(s)$, $0 \le s \le t$, augmented by all the $\mathbb{P}$-zero measure events. The corresponding filtration is denoted

by $\mathbb{F}$,

$$\mathbb{F} = \{\mathcal{F}_t, t \in [0,T]\},$$

which satisfies the usual conditions.

Definition 2.21. *A real function $f : [0,T]^n \to \mathbb{R}$ is called **symmetric** if*

$$f(t_{\sigma_1}, t_{\sigma_2}, \ldots, t_{\sigma_n}) = f(t_1, t_2, \ldots, t_n)$$

for all permutations $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ of (1,2,...,n).

Denote by $L^2([0,T]^n)$ the standard space of square integrable Borel real functions on $[0,T]^n$ such that

$$||g||^2_{L^2([0,T]^n)} := \int\limits_{[0,T]^n} g^2(t_1, t_2, \ldots, t_n)\, \mathrm{d}t_1 \mathrm{d}t_2 \cdots \mathrm{d}t_n < \infty.$$

Let $\tilde{L}^2([0,T]^n) \subset L^2([0,T]^n)$ be the space of symmetric square integrable Borel real functions on $[0,T]^n$.

Let

$$S_n = \{(t_1, t_2, \ldots, t_n) \in [0,T]^n : 0 \le t_1 \le t_2 \le \cdots \le t_n \le T\}.$$

Note that the set S_n occupies the friction $\frac{1}{n!}$ of the whole n-dimensional box $[0,T]^n$. Therefore, if $g \in \tilde{L}^2([0,T]^n)$, we have $g|_{S_n} \in L^2(S_n)$ and

$$||g||^2_{L^2([0,T]^n)} = n! \int\limits_{S_n} g^2(t_1, t_2, \ldots, t_n)\, \mathrm{d}t_1 \mathrm{d}t_2 \cdots \mathrm{d}t_n = n! ||g||^2_{S_n},$$

where $||\cdot||_{L^2(S_n)}$ is the norm induced by $L^2([0,T]^n)$ on $L^2(S_n)$, the space of the square integrable functions on S_n.

Definition 2.22. *Let f be a deterministic function defined on S_n $(n \ge 1)$, such that*

$$||g||^2_{S_n} := \int\limits_{S_n} g^2(t_1, t_2, \ldots, t_n)\, \mathrm{d}t_1 \mathrm{d}t_2 \cdots \mathrm{d}t_n < \infty.$$

*Then we can define the n-folder **iterated Itô integral** as as*

$$J_n(f) := \int\limits_0^T \int\limits_0^{t_n} \cdots \int\limits_0^{t_3} \int\limits_0^{t_2} f(t_1, \ldots, t_n)\, \mathrm{d}W(t_1) \mathrm{d}W(t_2) \cdots \mathrm{d}W(t_{n-1}) \mathrm{d}W(t_n).$$

The n-folder iterated Itô integral is well-defined because at each Itô integration with respect to $dW(t_i)$ the integrand is $\mathcal{F}_t$-adapted and square integrable with respect to $d\mathbb{P} \times dt_i$, $1 \leq i \leq n$.

If $g \in \tilde{L}^2([0,T]^n)$, we can define

$$I_n(g) := \int_{[0,T]^n} g(t_1, \ldots, t_n)\, \mathrm{d}W(t_1)\mathrm{d}W(t_2)\cdots \mathrm{d}W(t_n) = n! J_n(g).$$

Theorem 2.23 (The Wiener–Itô Chaos Expansion [64])**.** *Let ξ be an $\mathcal{F}_T$ measurable random variable in $L^2(\mathbb{P})$. Then, there exists an unique sequence $\{f_n\}_{n=0}^{\infty}$ of functions $f_n \in \tilde{L}^2([0,T]^n)$ such that*

$$\xi = \sum_{n=0}^{\infty} I_n(f_n),$$

where the convergence is in $L^2(\mathbb{P})$. Moreover, we have the isometry

$$||\xi||^2_{L^2(\mathbb{P})} = \sum_{n=0}^{\infty} n! ||f_n||^2_{L^2([0,T]^n)}.$$

2.2.2 The Malliavin derivative

Although several ways can be used to define the Malliavin derivative (see e.g. [64]), we introduce the Malliavin derivative with the chaos expansion, because this approach can also be used to define the Malliavin derivative in the Lévy spaces [144].

Definition 2.24. *Let $F \in L^2(\mathbb{P})$ be $\mathcal{F}_T$-measurable with the chaos expansion*

$$F = \sum_{n=0}^{\infty} I_n(f_n),$$

where $f_n \in \tilde{L}^2([0,T]^n)$, $n = 1, 2, \ldots$.

1. *We say that $F \in \mathbb{D}_{1,2}$ if*

$$||F||^2_{\mathbb{D}_{1,2}} = \sum_{n=0}^{\infty} n n! ||f_n||^2_{L^2([0,T]^n)} < \infty.$$

2. *If $F \in \mathbb{D}_{1,2}$, we define the **Malliavin derivative** $D_t F$ of F at time t as the expansion*

$$D_t F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)), \quad t \in [0,T],$$

where $I_{n-1}(f_n(\cdot,t))$ is the $(n-1)$-fold iterated integral of $f_n(t_1,\ldots,t_{n-1},t)$ with respect to the first $n-1$ variables $t_1,\ldots,t_{n-1}$ and $t_n = t$ left as parameter.

Example 2.25 ([64]).

$$D_t \int_0^T f(s)\,\mathrm{d}W(s) = f(t), \quad f \in L^2([0,T]). \tag{2.4}$$

Theorem 2.26 (Chain rule [64]). *Let $F_1,\ldots,F_m \in \mathbb{D}_{1,2}$. Suppose that $\phi \in C^1(\mathbb{R}^m)$, $D_tF_i \in L^2(\mathbb{P})$, for all $t \in \mathbb{R}$, and $\frac{\partial\phi}{\partial x_i}(F)D_tF_i \in L^2(\mathbb{P}\times\lambda)$ for $i = 1,\ldots,m$, where $F = (F_1,\ldots,F_m)$ and λ is the Lebesgue measure on $[0,T]$. Then, $\phi(F) \in \mathbb{D}_{1,2}$ and*

$$D_t\phi(F) = \sum_{i=1}^m \frac{\partial\phi}{\partial x_i}(F)D_tF_i.$$

Let $0 = t_1 \le t_2 \le \cdots \le t_{m+1} = T$ and

$$F_i = \int_0^T f_i(t)\,\mathrm{d}W(t), \quad i = 1,\ldots,m,$$

where

$$f_i(t) = \begin{cases} 1 & t \in [t_i, t_{i+1}), \\ 0 & \text{others.} \end{cases}$$

That is,

$$F_i = W(t_{i+1}) - W(t_i) := \Delta W_i, \quad i = 1,\ldots,m.$$

Suppose that $\phi(F) \in \mathbb{D}_{1,2}$, by the chain rule and (2.4), we have

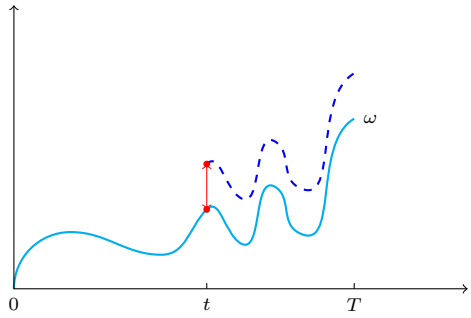
$$D_t\phi(F) = D_t\phi(\Delta W_1,\ldots,\Delta W_m) = \sum_{i=1}^m \frac{\partial\phi}{\partial x_i}(\Delta W_1,\ldots,\Delta W_m)f_i(t). \tag{2.5}$$

Remark 2.27.

1. *From (2.5), we can see that $D_t\phi(F)$ is a random variable, which depends on the whole path of the Wiener process W. That is, $(D_t\phi(F))_{0\le t\le T}$ is a stochastic process, but not an adapted one.*

2. *From an intuitive point of view, (2.5) shows that the Malliavin derivative $D_t\phi$ represents the derivative of ϕ with respect to the increments of W at time t, if we*

take the Wiener process as the canonical process, i.e., $W(t,\omega)=\omega_t$. Actually, the Malliavin derivative characterizes the change of a random variable with respect to the change of the elementary event $\omega\in\Omega$ at time t in the general case. Taking the limiting case $(m\to\infty)$ in (2.5), we can intuitively depict the change of the elementary event in the following figure:



2.3 Financial market models

The solution of a financial problem involving capital markets mostly relies on the prediction of the states of nature. The word "nature" is referred to as a generic and abstract description of the combination of all factors which may affect the evolvement of financial markets. The relations among these factors are so complex that we cannot exactly know what will happen in the future financial markets even in a short-time horizon. To account for this ambiguity about the future, one may use probability theory to describe the predictions of the nature states. In the simplified setting that the nature has finite states, the nature state ω and its probability p_ω constitute a pair (ω,p_ω) to characterize one prediction. So, the prediction of nature can be summarized as a finite probability space $(\Omega,\mathbb{P})$, where Ω is the space of finite states and $\mathbb{P}$ is a probability measure. In the setting of infinite states, we cannot define probabilities for each state, but define probabilities for events and construct a probability space $(\Omega,\mathcal{F},\mathbb{P})$.

Mathematically speaking, the factors in a financial problem are nothing but functions defined on $(\Omega,\mathcal{F},\mathbb{P})$, the so-called random variables. Hence, whenever a random variable is used in financial models, we implicitly select a probability space $(\Omega,\mathcal{F},\mathbb{P})$. One should keep it in mind that we assume that we have enough information to determine the probability of each event under $\mathbb{P}$. Otherwise, the story will be quite different[1]. In

[1]One may refer to Peng [143] for an introduction on stochastic calculus in a nonlinear expectation space

this section, we restrict ourself to the probability framework. Obvious, when a stochastic process is used to model a process in financial problems, we formulate the information as the filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$, where $\mathcal{F}_t$ represents the information available up to time t and $0 \le T \le +\infty$. In mathematical finance, stochastic models are usually defined on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Stochastic processes defined on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ can be used to model the price processes of stocks or underlying assets of derivatives. It is well-known that Louis Bachelier made the first attempt to describe the evolution of a stock price $S = (S_t)_{t\ge 0}$ by a Brownian motion in his thesis in 1900. Arguing that prices should remain positive, Samuelson [150] proposed to model a stock price process with a geometric Brownian motion in 1965. Working with this model, Black and Sholes [24] and Merton [132] opened the door to calculate a fair value for options with the arbitrage arguments in 1973. This pricing method is the so-called no-arbitrage approach.

2.3.1 Arbitrage-free market

Absence of arbitrage is a fundamental assumption for financial markets. Arbitrage means that an investor can make some strictly positive wealth out of nothing by a nonnegative total portfolio of investable securities. The nonnegativity of the investor's portfolio is legally enforced by his or her limited liability. The economic rationale behind this assumption is that in a liquid financial market, there is no such opportunity as a "free lunch". Hence, a mathematical model of a financial market should be designed in such a way that it does not permit arbitrage.

Denote by $S = (S^0, S^1, \dots, S^n)$ the tradable assets in the financial market consisting of one risk-free asset S^0 (bank account) and n risky assets $(S^1, \dots, S^n)$.

A model for S is assumed to be defined on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Given a fixed investment horizon $T \in \mathbb{R}^+$, we call a $\mathbb{R}^{n+1}$-valued predictable process $\delta = (\delta_t^0, \delta_t^1, \dots, \delta_t^n)_{0\le t\le T}$ a **strategy** if the Itô integral

$$\int_0^t \delta_s^\star \, \mathrm{d}S_s$$

is well defined for $t \in [0, T]$. δ_t^j, $j = 0, \dots, n$, is the number of units of the jth asset held at time t in the portfolio $S^\delta = (S_t^\delta)_{0\le t\le T}$. The value S^δ of the portfolio at time t

$(\Omega, \mathcal{F}, \hat{\mathbb{E}})$, which is tailored to financial applications under model uncertainty.

is given by

$$S_t^\delta = \sum_{i=0}^{n} \delta_t^i S_t^i.$$

A strategy δ and the corresponding portfolio S^δ are said to be **self-financing** if

$$\mathrm{d}S_t^\delta = \sum_{i=0}^{n} \delta_t^i \, \mathrm{d}S_t^i$$

for $t \in [0, T]$. This means that all changes in portfolio value are due to gains or losses from trading assets.

Definition 2.28. *A self-financing portfolio is an **arbitrage** if the value process S^δ satisfying*

1. $S_0^\delta = 0$;
2. $\mathbb{P}(S_T^\delta \geq 0) = 1, \quad \mathbb{P}(S_T^\delta > 0) > 0.$

One condition to guarantee that a financial market model does not permit an arbitrage is usually called the first fundamental asset pricing theorem, which involves the concept of risk-neutral probability measure.

Definition 2.29. *Let $(\hat{D}_t := 1/S_t^0)_{0 \leq t \leq T}$ be a discounting process. A probability measure $\hat{\mathbb{P}}$ is said to be **risk-neutral** if*

1. *$\hat{\mathbb{P}}$ and $\mathbb{P}$ are equivalent, i.e., for every $A \in \mathcal{F}$, $\mathbb{P}(A) = 0$ if and only if $\hat{\mathbb{P}}(A) = 0$.*
2. *under $\hat{\mathbb{P}}$, the discounted risky asset price process $(\hat{D}_t S_t^i)_{0 \leq t \leq T}$, $i = 1, \ldots, n$, is a martingale.*

For a given mathematical model for S, it is easy to see that under a risk-neutral probability measure $\hat{\mathbb{P}}$, the risky asset has the same average rate of return as that of the risk-free asset. Hence, under $\hat{\mathbb{P}}$, the average rate of return of the portfolio is also the same as that of the risk-free asset, and the discounted portfolio value is a martingale with respect to $\hat{\mathbb{P}}$. This fact plays a key role in the theory of risk-neutral pricing.

Theorem 2.30 (First fundamental asset pricing theorem)**.** *If a market model has a risk-neutral probability measure, then it does not admit arbitrage.*

We refer to [159] for the proof of this theorem, and [57] for other versions of this fundamental theorem in more general cases.

2.3.2 Market completeness

Let V_T be an $\mathcal{F}_T$-measurable random variable, which represents the payoff of a financial derivative. We wish to choose the amount of initial capital S_0^δ and a portfolio process δ in order to hedge a short position in the derivative, i.e.

$$S_T^\delta = V_T, \quad a.s. \tag{2.6}$$

The key objective of risk-neutral valuation is to calculate the initial capital S_0^δ and a portfolio process δ such that (2.6) holds. Once it has been done, the fact that $(\hat{D}_t S_t)_{0\le t\le T}$ is a martingale under a risk-neutral probability measure $\hat{\mathbb{P}}$ implies

$$\hat{D}_t S_t^\delta = \mathbb{E}^{\hat{\mathbb{P}}}[\hat{D}_T S_T^\delta] = \mathbb{E}^{\hat{\mathbb{P}}}[\hat{D}_T V_T].$$

The value S_t^δ of the hedging portfolio is the capital needed at time t to completely hedge the short position of V_T. Hence, we can define the value V_t of the derivative with payoff V_T at time t as

$$\hat{D}_t V_t = \mathbb{E}^{\hat{\mathbb{P}}}[\hat{D}_T V_T], \quad \text{for } 0 \le t \le T. \tag{2.7}$$

That is, $(\hat{D}_t V_t)_{0\le t\le T}$ is a martingale under $\hat{\mathbb{P}}$. It is this insight that provides the idea to prove the existence of the hedging portfolio and to calculate the hedging strategy (see [159] for the details).

The formula (2.7) is the so-called **risk-neutral pricing formula**. This formula requires a premise that the market model should admit a risk-neutral probability measure. When the asset prices are driven by Brownian motions, the existence of risk-neutral probability measure amounts to the existence of the market price of risk equations (see [159]).

The second fundamental theorem for asset pricing concerns the uniqueness of risk-neutral probability measures. It is related to the completeness of the market model.

Definition 2.31. *A market model is **complete** if every derivative security can be hedged.*

Theorem 2.32 (Second fundamental asset pricing theorem [159])**.** *Consider a market model that has a risk-neutral probability measure. The model is complete if and only if the risk-neutral probability measure is unique.*

If a financial market model permits multiple risk-neutral measures, a financial derivative has more than one value, and the position in this derivative cannot be fully hedged. One may select other hedging strategies for a financial derivative, such as the (locally) risk-minimising strategies investigated in the next chapter.

3

Model uncertainty and discretisation of locally risk-minimising strategies

In an incomplete market model, the locally risk-minimising (LRM) strategy is an important hedging strategy for financial derivatives. The goal of the LRM strategies is to minimise the variance of the cost process at any time t, while the value of the hedging portfolio at time T recovers the payoff of a contingent claim at time T.

Considering vanilla options, Asian options, spread options and basket options in a class of jump-diffusion models, we formulate the corresponding LRM strategies as solutions of backward differential equations (BSDEs). Using the discretisation scheme and the convergence results on BSDEs as studied in Khedher and Vanmaele [115], we can not only simulate the LRM strategies but also show that the LRM strategies are robust towards the choice of the model.

3.1 Introduction

In this chapter, we investigate the LRM strategies for vanilla options, Asian options, spread options and basket options in a class of jump-diffusion models. The LRM strategies are formulated as the solution of a forward-backward stochastic differential

equations of the following form

$$\begin{cases} S(t) = S(0) + \int_0^t S(s)a(s)\,\mathrm{d}s + \int_0^t S(s-)\,\mathrm{d}\zeta(s)\,, \\ V(t) = h(S(T)) + \int_t^T \varphi(s, S(s), V(s), \Upsilon(s))\,\mathrm{d}\langle\zeta\rangle_s \\ \qquad\quad - \int_t^T \Upsilon(s)\,\mathrm{d}\zeta(s) - L(T) + L(t)\,, \end{cases} \tag{3.1}$$

where $(\zeta(t))_{0\le t\le T}$ is a càdlàg martingale, $\langle\zeta\rangle_t$ is the predictable compensator of the quadratic variation of $(\zeta(t))_{0\le t\le T}$, and $(L(t))_{0\le t\le T}$ is a *martingale orthogonal* to $(\zeta(t))_{0\le t\le T}$. h, φ, and a have to fulfil certain conditions that we specify in the following sections. For the existence and uniqueness of the solution of such FBSDE, we refer to Carbone et al. [32].

In the sequel, the forward SDE in (3.1) models the asset price process, where we choose $(M(t))_{0\le t\le T}$ to be given by a sum of continuous and jump noises. The solution of the backward SDE which is given by a triplet $(V(t), \Upsilon(t), L(t))_{0\le t\le T}$ models respectively the value of the portfolio, the amount of wealth invested in the risky assets, and the remaining risk in a locally risk-minimising (LRM) strategy. We refer to [63, 111] for a rigorous study of the relation of FBSDEs to quadratic hedging strategies.

We note that quite few papers investigated the LRM strategies for Asian options or spread options in incomplete market models. Wang and Wang [176] investigated the LRM hedging strategy for discretely monitored geometric Asian call options under exponential Lévy models. Benth *et al.* [20] derived an integral representation of the variance-optimal hedging strategy for Asian-type options on energy with trading constricts in the BN-S model and an exponential additive model. Under a similar model as the one that we consider in this chapter, Arai and Suzuki [5] derived an integral representation of the LRM strategy for vanilla call options, Asian options, and lookback options. However, they did not provide an implementable method to calculate the integral representation of the LRM strategy derived therein.

Using the discretisation scheme and the convergence results on BSDEs as studied by Khedher and Vanmaele [115], we will simulate the LRM strategies and show that these LRM strategies are robust towards the choice of the model.

3.2 Preliminaries of locally risk-minimising Strategies

In an incomplete market model, a position of a derivative cannot be hedged with self-financing strategies (see e.g. [159] for a proof). The locally risk-minimising (LRM) strategy is a popular criterion for hedging and pricing in incomplete financial markets. We present the preliminaries of the LRM strategy for a contingent claim written on one risky asset and multiple assets, respectively. These results are mainly provided in a series of papers such as Föllmer and Sondermann [82], Schweizer [153, 157], Choulli et al. [46] and so on. The LRM strategies are characterized by the position in the underlying asset(s). If there is one underlying asset, the corresponding LRM strategy is called a one-dimensional LRM strategy in this thesis. Otherwise, it is called a multi-dimensional LRM strategy.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Fix $T > 0$. Denote by S the asset price which is adapted to a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ satisfying the *usual conditions*.

3.2.1 One-dimensional LRM strategies

Föllmer and Sondermann [82] proposed the so-called risk-minimising strategy in the sense that it minimises the remaining risk in terms of the cost process. Under the assumption that the discounted asset price is a square-integrable martingale under the original probability measure, Föllmer and Sondermann [82] proved the existence and uniqueness of the risk-minimisation strategy for a contingent claim $H \in L^2(\mathbb{P})$.

However, if the underlying asset is a semimartingale, the model may not admit a risk-minimising strategy, as illustrated by an example in [153]. Schweizer [153] extended the theory of risk minimisation to the setting where the underlying risky asset S is a semimartingale. The semimartingale S is assumed to have a decomposition

$$S = S(0) + M + A,$$

where M is a square-integrable martingale with $M(0) = 0$ and A is a predictable process of finite variation with $A(0) = 0$. In this setting, the resulting strategy is called the locally risk-minimising strategy in the sense that the risk is formulated in a local manner.

By the assumptions of M in the decomposition of S, we can define a measure $\mathbb{P}_M := \mathbb{P} \times \langle M \rangle$ on the product space $\Omega \times [0, T]$ with the product σ-algebra of predictable sets. An expectation $\mathbb{E}_M$ associated with $\mathbb{P}_M$ is defined by

$$\mathbb{E}_M[\alpha] = \frac{\mathbb{E}[(\alpha \cdot \langle M \rangle)_T]}{\mathbb{E}[\langle M \rangle_T]},$$

where α is an integrable process with respect to $\langle M \rangle$.

Definition 3.1. *Assume $S = S(0) + M + A$ is a semimartingale under the measure $\mathbb{P}$. A couple $\varphi = (\xi, \eta)$ is called a **trading strategy** if*

- *ξ is a predictable process,*
- *ξ satisfies*
$$\mathbb{E}\left[\int_0^T \xi_u^2 \mathrm{d}\langle M\rangle_u + \left(\int_0^T |\xi_u \mathrm{d}S_u|\right)^2\right] < \infty,$$
- *η is adapted,*
- *the process $V(\varphi) := \xi S + \eta$ has right-continuous paths and $\mathbb{E}[V_t^2] < \infty$ for every $t \in [0, T]$.*

*The process $V(\varphi)$ is called the **value process** of the trading strategy φ. A strategy is called **admissible** with respect to a contingent claim H or H-**admissible**, if its value process has terminal value H, $\mathbb{P}$-a.s..*

Definition 3.2. *Given a trading strategy φ and the value process $V(\varphi)$, we can define the **cost process** $C(\varphi)$ as the right-continuous square integrable process given by*

$$C_t(\varphi) := V_t(\varphi) - \int_0^t \xi_u \mathrm{d}S_u, \quad 0 \le t \le T. \tag{3.2}$$

*If the cost process is a square-integrable martingale, the strategy is called a **mean-self-financing** strategy. The **risk process** $R(\varphi)$ is the conditional mean square error process*

$$R_t(\varphi) := \mathbb{E}[(C_T(\varphi) - C_t(\varphi))^2 | \mathcal{F}_t].$$

To formulate what is locally risk-minimising strategy, we first explain the notion of small perturbation [152].

Definition 3.3. *A trading strategy $\Delta = (\delta, \varepsilon)$ is called a **small perturbation** if it satisfies the following conditions:*

- *δ is bounded,*
- *$\int_0^t |\delta_u \mathrm{d}A_u|$ is bounded,*

- $\delta_T = \varepsilon_T = 0$.

To investigate the local behaviour of a trading strategy, we consider partitions $\tau = (t_i)_{0 \le i \le N}$ of the interval $[0, T]$. These partitions satisfy

$$0 = t_0 < t_1 < \cdots < t_N = T,$$

and their mesh $|\tau| := \max_{t_i, t_{i+1} \in \tau} (t_{i+1} - t_i)$. A sequence $(\tau_n)_{n \in \mathbb{N}}$ is called increasing if $\tau_n \subset \tau_{n+1}$ for all n; it is called 0-convergent if $\lim_{n \to \infty} |\tau_n| = 0$. If $\Delta = (\delta, \varepsilon)$ is a small perturbation and $(s, T]$ a subinterval of $[0, T]$, we define the small perturbation

$$\Delta|_{(s,t]} := (\delta \mathbf{1}_{(s,t]}, \varepsilon \mathbf{1}_{[s,t)}).$$

Now we are ready to define what is a locally risk-minimising strategy.

Definition 3.4. *For a trading strategy φ, a small perturbation Δ and a partition τ of $[0, T]$, the **risk quotient** $r^\tau[\varphi, \Delta]$ is defined as*

$$r^\tau[\varphi, \Delta] := \sum_{t_i, t_{i+1} \in \tau} \frac{R_{t_i}(\varphi + \Delta|_{(t_i, t_{i+1}]}) - R_{t_i}(\varphi)}{\mathbb{E}[\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i} | \mathcal{F}_{t_i}]} 1_{(t_i, t_{i+1}]}(t).$$

*A trading strategy φ is called **locally risk-minimising** if*

$$\liminf_{n \to \infty} r^{\tau_n}(\varphi, \Delta) \ge 0, \quad \mathbb{P}_M\text{-}a.s.$$

for every small perturbation Δ and every increasing 0-convergent sequence (τ_n) of partitions of $[0, T]$.

The risk quotient is a measure for the total change of riskiness if the trading strategy φ is locally perturbed by the perturbation Δ along the partition τ. To calculate the locally risk-minimising strategy, we need additional assumptions on S.

Assumptions 3.5.

(A1) For $\mathbb{P}$-almost all ω, the measure on $[0, T]$ induced by $\langle M \rangle(\omega)$ has the whole interval $[0, T]$ as its support.

(A2) A is continuous.

(A3) A is absolutely continuous with respect to $\langle M \rangle$ with a density λ satisfying

$$\mathbb{E}_M[|\lambda| \log^+ |\lambda|] < \infty.$$

Under the Assumptions 3.5, the next proposition furnishes us with the key step to calculate the locally risk-minimising strategies.

Proposition 3.6 (see Schweizer [153]). *Let the special semimartingale S satisfies Assumptions 3.5. Let $H \in L^2(\mathbb{P})$ be a contingent claim and φ an H-admissible trading strategy. The following statements are equivalent:*

(1) φ is locally risk-minimising.

(2) φ is mean-self-financing and the martingale $C(\varphi)$, the cost process defined in (3.2), is orthogonal to the martingale part M of the special semimartingale S.

Proposition 3.6 allows us to calculate the locally risk-minimising strategy with the famous Föllmer-Schweizer decomposition.

Definition 3.7 (Föllmer-Schweizer decomposition [81]). *An $\mathcal{F}_T$-measurable and square-integrable random variable H admits a Föllmer-Schweizer decomposition (FS-decomposition hereafter) if there exist a constant H_0, an S-integrable process ξ^{FS} and a square-integrable martingale L^{FS}, such that $[L^{FS}, M]$ is a local martingale and*

$$H = H_0 + (\xi^{FS} \cdot S)_T + L_T^{FS}.$$

We refer to Choulli et al. [46] for the proof of the existence and uniqueness of the FS-decomposition in the general setting. On the basis of some conditions and arguments concerning the minimal martingale measure, the FS-decomposition directly provides the locally risk-minimising strategy φ in the following way:

$$\varphi_t = (\xi_t^{FS}, H_0 + \int_0^t \xi_u^{FS} \mathrm{d}S_u + L_t^{FS} - \xi_t^{FS} S_t).$$

We refer to Vandaele [170] and Choulli et al. [47] for a thorough analysis and calculation methods of the FS-decomposition. In some specific settings, the LRM strategy could be computed using the Fourier transform techniques as described in Cont and Tankov [51] and Hubalek et al. [108].

3.2.2 Multi-dimensional LRM strategies

Schweizer [157] generalized the theory of one-dimensional LRM strategies to the setting with multiple underlying assets. Let S be a d-dimensional square integrable

semimartingale under $\mathbb{P}$ adapted to a filtration $\mathbb{F}$. Then S can be decomposed as $S = S(0) + M + A$, where M is a d-dimensional locally square integrable $\mathbb{P}$-martingale starting in 0 and A is a d-dimensional predictable process of finite variation also starting in 0. We also assume that S satisfies the following structure conditions:

Assumptions 3.8.

(A) A is absolutely continuous with respect to $\langle M \rangle$ in the sense that

$$A(t) := \int_0^t \mathrm{d}\langle M \rangle_s \, \lambda(s) \,, \tag{3.3}$$

where λ is $\mathbb{R}^d$-valued predictable process.

(B) The mean variance tradeoff (MVT) process

$$K(t) := \int_0^t \lambda^\star(s) \, \mathrm{d}\langle M \rangle_s \, \lambda(s) \,, \tag{3.4}$$

is uniformly bounded $\mathbb{P}$-a.s. for all $t \in [0, T]$. $(\cdot)^\star$ denotes the transpose of a vector or matrix.

These structural conditions are related to an absence-of-arbitrage condition [154, 157]. We denote by $L(S)$, the S-integrable processes. That is the class of predictable processes for which we can determine the stochastic integral w.r.t. S. We define the space Θ by

$$\Theta := \left\{ \theta \in L(S) \;\middle|\; \mathbb{E}\left[\int_0^T \theta^\star(s) \, \mathrm{d}\langle M \rangle_s \, \theta(s) + \left(\int_0^T |\theta^\star(s) \, \mathrm{d}A(s)| \right)^2 \right] < \infty \right\} .$$

Let ξ be a square integrable contingent claim and denote by V the value process of the LRM strategy to hedge ξ.

Under Assumptions 3.8 we have the following unique decomposition for the value process written under the real world measure $\mathbb{P}$ (see Schweizer [157] for a proof)

$$V(t) = V(0) + \int_0^t \left(\chi^{FS}\right)^\star(s) \, \mathrm{d}S(s) + L^{FS}(t), \qquad 0 \le t \le T \,, \tag{3.5}$$

where L^{FS} is a square integrable $\mathbb{P}$-martingale such that $[L^{FS}, M]$ is a $\mathbb{P}$-martingale and $\chi^{FS} \in \Theta$. For $t = T$, the latter equation is the FS-decomposition of ξ. In fact, the components $\chi^{FS}(t)$ and $C(t) := L^{FS}(t) + V(0)$ represent respectively the number of risky assets and the cost process in an LRM strategy at time $t \in [0, T]$ (see Proposition 3.4 in Schweizer [156]).

To conclude the section, we would like to outline some notations to be used in the following sections.

NOTATIONS. Let $W(t) = (W^{(1)}, \dots, W^{(d)})(t)$, $t \in [0, T]$ be a d-dimensional standard Wiener process with $W^{(i)}$ and $W^{(j)}$ being independent for $i \neq j$ and $\tilde{N} = \tilde{N}(\mathrm{d}t, \mathrm{d}z)$, $(t, z) \in [0, T] \times \mathbb{R}$ be a d-dimensional vector consisting of independent one-dimensional centered Poisson random measures, i.e. $\tilde{N}(\mathrm{d}t, \mathrm{d}z) = N(\mathrm{d}t, \mathrm{d}z) - \ell(\mathrm{d}z)\mathrm{d}t$, where ℓ is the Lévy measure and $N(\mathrm{d}t, \mathrm{d}z)$ is the Poisson random measure such that $\mathbb{E}[N(\mathrm{d}t, \mathrm{d}z)] = \ell(\mathrm{d}z)\,\mathrm{d}t$. Define $\mathcal{B}(\mathbb{R})$ as the σ-algebra generated by the Borel sets $\bar{U} \subset \mathbb{R}$. We assume that the Lévy measure has finite activity, i.e. $\ell_j(|z| \leq 1) < \infty$, $j = 1, \dots, d$ and that the jump measure has a finite second moment, i.e. $\int_{\mathbb{R}} z^2 \ell_j(\mathrm{d}z) < \infty$. We specify the $\mathbb{P}$-augmented filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ by

$$\mathcal{F}_t = \sigma \left\{ W(s), \int_0^s \int_A \tilde{N}(\mathrm{d}u, \mathrm{d}z), \quad s \leq t, \quad A \in \mathcal{B}(\mathbb{R}) \right\} \vee \mathcal{N}\,, \tag{3.6}$$

where $\mathcal{N}$ represents the set of $\mathbb{P}$-null events in $\mathcal{F}$. An element $x \in \mathbb{R}^d$ will be identified with a column vector with ith component x_i and Euclidean norm $|x|$. We define the following spaces:

- L^2_T: the space of all $\mathcal{F}_T$-measurable random variables $X : \Omega \to \mathbb{R}$ such that

$$\|X\|^2 = \mathbb{E}[X^2] < \infty\,.$$

- $S^2_{[0,T]}$: the space of all $\mathbb{F}$-adapted, càdlàg processes $\gamma : \Omega \times [0, T] \to \mathbb{R}$ such that

$$\|\gamma\|^2_{S^2_{[0,T]}} = \mathbb{E}[\sup_{0 \leq t \leq T} |\gamma^2(t)|] < \infty\,.$$

- $H^2_{[0,T]}$: the space of all $\mathbb{F}$-predictable processes $\phi : \Omega \times [0, T] \to \mathbb{R}^d$, such that

$$\|\phi\|^2_{H^2_{[0,T]}} = \mathbb{E}\left[\int_0^T |\phi(t)|^2\,\mathrm{d}t\right] < \infty\,.$$

- $\hat{H}^2_{[0,T]}$: the space of all $\mathbb{F}$-predictable mappings $\theta : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}^d$, such that
$$\|\theta\|^2_{\hat{H}^2_{[0,T]}} = \mathbb{E}\left[\int_0^T \sum_{j=1}^d \int_{\mathbb{R}} |\theta_j(t,z)|^2 \ell_j(\mathrm{d}z)\,\mathrm{d}t\right] < \infty.$$
- $C^1_b(\mathbb{R}^d)$: the space of continuously differentiable functions with bounded derivatives in all the variables.

3.3 Discretisation of LRM strategies for vanilla and Asian options

3.3.1 Continuous-time model: one-dimensional jump-diffusion

In the sequel we specify a one-dimensional continuous-time model for the stock price process and we compute the LRM strategy in this setting. We consider the following continuous-time model

$$X(t) := X(0) + \int_0^t \delta(s)\,\mathrm{d}s + \int_0^t b(s)\,\mathrm{d}W(s) + \int_0^t \int_{\mathbb{R}} \gamma(s,z)\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,, \tag{3.7}$$

where $\delta(t), b(t), \gamma(t,z) \in \mathbb{R}$, for $t \geq 0$, $z \in \mathbb{R}$. X is a process with independent increments. Its parameters are assumed to satisfy the following assumptions.

Assumptions 3.9.

(A) $\gamma(t,z)$ *is assumed to be of the following form,*

$$\gamma(t,z) = g(z)\widetilde{\gamma}(t)\,,$$

for which we impose

$$\int_{\mathbb{R}} g(z)\,\ell(dz) < \infty,$$

and for some $\varepsilon > 0$*,*

$$G^2(\varepsilon) := \int_{|z|\leq\varepsilon} g^2(z)\,\ell(dz) < \infty\,. \tag{3.8}$$

(B) δ*,* b *and* $\tilde{\gamma}$ *are Lipschitz.*

We need the parameters of the process X to be Lipschitz on the one hand to guarantee the existence of the LRM strategy and on the other hand for our convergence study of the discretisation scheme. We define the process $S^{(1)}$ to be the stochastic exponential of X (see, e.g., Chapter II in Protter [146]). That is

$$S^{(1)}(t) := S^{(1)}(0)\mathcal{E}(X)_t = S^{(1)}(0) + \int_0^t S^{(1)}(s-)\,\mathrm{d}X(s)\,,$$

where $S^{(1)}(0)$ is a positive constant. For $S^{(1)}$ to be positive, we assume $\gamma(t,z) > -1$ for $(t\,,z) \in [0,T] \times \mathbb{R}$. Notice that the process $S^{(1)}$ is a square integrable semimartingale.

Let $S^{(0)}$ denote the riskless asset. It is given by

$$\mathrm{d}S^{(0)}(t) = S^{(0)}(t)r(t)\,\mathrm{d}t, \qquad S^{(0)}(0) = 1\,,$$

where the short rate r is Lipschitz and $r(0) = 0$. Let $S^{(1)}$ be the undiscounted price process of the risky asset. We assume that $S = S^{(1)}/S^{(0)}$ is the discounted stock price process. Its dynamics is thus given explicitly by

$$\mathrm{d}S(t) = S(t-)\left[a(t)\,\mathrm{d}t + b(t)\,\mathrm{d}W(t) + \int_{\mathbb{R}} \gamma(t,z)\,\tilde{N}(\mathrm{d}t,\mathrm{d}z)\right], \tag{3.9}$$

where $a(t) = \delta(t) - r(t)$, $t \in [0,T]$. Since δ and r are Lipschitz, a is also Lipschitz. Define

$$\kappa(t) := b^2(t) + \int_{\mathbb{R}} \gamma^2(t,z)\,\ell(\mathrm{d}z)\,, \qquad t \in [0,T]. \tag{3.10}$$

The assumption that b and $\tilde{\gamma}$ are Lipschitz, together with (3.8), implies that

$$C_1 \le \kappa(t) \le C_2\,, \quad t \in [0,T]\,. \tag{3.11}$$

where C_1 and C_2 are positive constants.

We can decompose S into a square integrable martingale M starting at zero in zero and a predictable finite variation process A, with $A(0) = 0$, where

$$M(t) = \int_0^t S(s-)b(s)\,\mathrm{d}W(s) + \int_0^t \int_{\mathbb{R}} S(s-)\gamma(s,z)\,\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,.$$

Thus the predictable process λ defined in (3.3) and associated with the process S is given by

$$\lambda(t) = \frac{a(t)}{S(t-)\kappa(t)}\,, \qquad 0 \le t \le T$$

and the MVT process defined in (3.4) and associated with the process S is given by

$$K(t) = \int_0^t \frac{a^2(s)}{\kappa(s)}\,\mathrm{d}s\,, \qquad 0 \le t \le T\,.$$

Notice that condition (3.11) and the boundedness of a imply that the MVT process is uniformly bounded in t by a constant C.

Let the contingent claim be given by ξ whose discounted value is $\tilde{\xi} = \xi/S^{(0)}(T)$, an $\mathcal{F}_T$-measurable square integrable random variable. Recall the FS-decomposition in (3.5). Substituting the dynamics (3.9) of S in (3.5) and considering the discounted value of the portfolio $\tilde{V} = V/S^{(0)}$, we get for all $t \in [0,T]$

$$\begin{cases} \mathrm{d}\tilde{V}(t) = \tilde{\Upsilon}(t)a(t)\,\mathrm{d}t + \tilde{\Upsilon}(t)b(t)\,\mathrm{d}W(t) + \displaystyle\int_{\mathbb{R}} \tilde{\Upsilon}(t)\gamma(t,z)\,\tilde{N}(\mathrm{d}t,\mathrm{d}z) \\ \qquad\qquad + \mathrm{d}L^{FS}(t), \\ \tilde{V}(T) = \tilde{\xi}\,, \end{cases} \tag{3.12}$$

where the predictable process $\tilde{\Upsilon}$ defined as $\tilde{\Upsilon}(t) = \chi^{FS}(t)S(t-)$ is the amount of wealth to invest in an LRM strategy. From the FS-decomposition in the LRM strategy we know that $\tilde{\Upsilon} \in \Theta$ and that L^{FS} is a square integrable $\mathbb{P}$-martingale such that $[L^{FS}, M]$ is a $\mathbb{P}$-martingale. The integral form of (3.12) is given by

$$\tilde{V}(t) = \tilde{\xi} - \int_t^T \frac{\tilde{\Upsilon}(s)a(s)}{\kappa(s)}\,\mathrm{d}\langle\zeta\rangle_s - \int_t^T \tilde{\Upsilon}(s)\,\mathrm{d}\zeta(s) - L^{FS}(T) + L^{FS}(t)\,, \tag{3.13}$$

where

$$\zeta(t) = \int_0^t b(s)\,\mathrm{d}W(s) + \int_0^t \int_{\mathbb{R}} \gamma(s,z)\,\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,. \tag{3.14}$$

Note that (3.13) is a BSDE of the type studied in Khedher and Vanmaele [115]. To study the time-discretisation of the LRM strategies, we use the approach in this latter paper. Before invoking the conclusions within that framework, we have to check whether all the conditions therein are satisfied for our model. We also have to adapt some proofs therein to our setting. Hereto we first use the representation theorem (see Kunita and Watanabe [119]) for the $\mathcal{F}_T$-measurable square integrable random variable $L^{FS}(T)$ and recall that L^{FS} is a $\mathbb{P}$-martingale being zero at zero to get the following representation

for the process L^{FS}

$$L^{FS}(t) = \int_0^t Y^{FS}(s)\,\mathrm{d}W(s) + \int_0^t \int_{\mathbb{R}} Z^{FS}(s,z)\,\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,, \tag{3.15}$$

where $Y^{FS} \in H^2_{[0,T]}$ and $Z^{FS} \in \hat{H}^2_{[0,T]}$ (see Section 3 in Di Nunno et al. [63] for more details). Next, we introduce the following classical BSDE

$$\begin{cases} -\mathrm{d}\tilde{V}(t) = f\left(t, Y(t), \Gamma(t)\right)\mathrm{d}t - Y(t)\,\mathrm{d}W(t) - \displaystyle\int_{\mathbb{R}} Z(t,z)\,\tilde{N}(\mathrm{d}t,\mathrm{d}z)\,, \\ \tilde{V}(T) = \tilde{\xi}\,, \end{cases} \tag{3.16}$$

where $\Gamma(t) = \int_{\mathbb{R}} Z(t,z)\gamma(t,z)\ell(\mathrm{d}z)$ and $\tilde{\xi} \in L^2_T$ is $\mathcal{F}_T$-measurable. Under the standard assumptions on $f : [0,T] \times \mathbb{R}^2 \to \mathbb{R}$, Tang and Li (1994) prove that the BSDE (3.9) admits an unique solution $(\tilde{V}, Y, Z) \in S^2_{[0,T]} \times H^2_{[0,T]} \times \hat{H}^2_{[0,T]}$. In the following proposition, we relate the BSDE (3.12) to a BSDE of the classical form. We do not present the proof since it follows similar lines as the proof of Lemma 4.1 in Di Nunno et al. [63].

Proposition 3.10. *Let $(\tilde{V}, \tilde{\Upsilon}, L^{FS})$ be the solution of the BSDE (3.12). Let the Assumptions 3.9 hold. Then, the BSDE (3.12) with L^{FS} represented by (3.15) can be rewritten as a classical BSDE (3.16), where*

$$\begin{aligned} Y(t) &= \tilde{\Upsilon}(t)b(t) + Y^{FS}(t)\,, \\ Z(t,z) &= \tilde{\Upsilon}(t)\gamma(t,z) + Z^{FS}(t,z)\,, \\ f(t,y,z) &= -\frac{a(t)}{\kappa(t)}\left(b(t)y + z\right)\,. \end{aligned} \tag{3.17}$$

We recall that Assumptions 3.9 imply the estimate (3.11). Further we notice that in the proof of the latter proposition, it is shown that f is Lipschitz to guarantee the existence and uniqueness of the solution to the BSDE.

We consider vanilla and Asian options in which cases the discounted payoff is formulated as follows

$$\tilde{\xi} = \begin{cases} \dfrac{h(S^{(1)}(T))}{S^{(0)}(T)}\,, & \text{vanilla options,} \\ \dfrac{h(S_A(T))}{S^{(0)}(T)}\,, \quad \text{with } S_A(T) = \dfrac{1}{T}\displaystyle\int_0^T S^{(1)}(s)\,\mathrm{d}s\,, & \text{Asian options,} \end{cases} \tag{3.18}$$

where h is Lipschitz on $\mathbb{R}$. In the sequel, without loss of generality, we assume that $h(0) = 0$. Otherwise, we can work with $\hat{h} = h - C$, for $h(0) = C$, where C is a constant.

Denote by π_n the regular time grid

$$\pi_n := \{0 = t_0, \dots, t_i, \dots, t_n = T\}, \qquad t_i = \frac{iT}{n}, \qquad n \in \mathbb{N}. \tag{3.19}$$

The statement of Proposition 3.10 will allow us to exploit estimates on solutions of classical BSDEs and to translate them to the solution of our BSDE (3.12). This is the purpose of the next theorem. Notice that these estimates are important for our robustness study later.

Theorem 3.11. *Let $(\tilde{V}, \tilde{\Upsilon}, L^{FS})$ be the solution of the BSDE (3.12) and $\tilde{\xi}$ be as in (3.18) with h being Lipschitz and $h(0) = 0$. Under Assumptions 3.9, we have*

$$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |\tilde{V}(t) - \tilde{V}(t_i)|^2\right] + \mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |\tilde{\Upsilon}(s) - \tilde{\Upsilon}(t_i)|^2 \,\mathrm{d}s\right] \leq \frac{C}{n},$$

where C is a positive constant.

Proof. The estimate for $\tilde{V}$ follows immediately from Theorem 2.1 in Bouchard and Elie [26] and Proposition 3.10 in this chapter. Here we focus on the estimate for $\tilde{\Upsilon}$.

We first show that f is in $C_b^1(\mathbb{R}^2)$. The derivative of f defined in (3.17) w.r.t. u is given by

$$\partial_u f(u, \Gamma(\cdot)) = -a(t)\frac{b(t)}{\kappa(t)}.$$

Since a, b, and κ are bounded above and below, it follows that $\partial_u f$ is bounded above by a positive constant. On the other hand, the partial derivative of f w.r.t. z is given by

$$\partial_z f(t, y, z) = -\frac{a(t)}{\kappa(t)}.$$

The boundedness below and above of a and κ imply that f_z is uniformly bounded above by a positive constant. That is, f has continuous and bounded partial derivatives in (y, z), and this property holds uniformly in t. To simplify notations, we abbreviate $f(t, y, z)$ to $f_t(y, z)$, and define $\nabla f_t(y, z) := (\partial_y f, \partial_z f)$.

Assuming that $h \in C_b^1(\mathbb{R})$, the statement of the theorem for *vanilla options* follows the statement of Theorem 2.3 in Khedher and Vanmaele [115]. To prove that the

statement holds for h being Lipschitz, we use an approximation argument that we present later.

Now we proceed to prove the statement in the setting of *Asian* options. We first derive estimates for Y and Z. Then, by using an approximation argument as in the proof of Theorem 2.1 in Bouchard and Elie [26], we can prove that these estimates hold when h is Lipschitz. Based on these estimates for Y and Z, the estimate for $\tilde{\Upsilon}$ will be derived in the end.

Integrating (3.16) and considering $\tilde{\xi} = h(S_A(T))/S^{(0)}(T)$, we have

$$\begin{aligned}\tilde{V}(t) = \frac{h(S_A(T))}{S^{(0)}(T)} + \int_t^T f_s(Y(s),\Gamma(s))\mathrm{d}s - \int_t^T Y(s)\mathrm{d}W(s) \\ - \int_t^T \int_{\mathbb{R}} Z(s,z)\tilde{N}(\mathrm{d}s,\mathrm{d}z),\end{aligned} \tag{3.20}$$

where h is assumed to be in $C_b^1(\mathbb{R})$ at the moment. The aim is to obtain first an estimate for the process Y. For this purpose, we consider a representation for the process Y using the Malliavin derivative as it is done in Bouchard and Elie [26] (see also Zhang [184] who first studied this type of representations for solutions of continuous BSDEs). Let D denote the Malliavin derivative operator w.r.t. the Brownian motion W where we consider the Malliavin operator introduced in Petrou [144] (see also Nualart [137] for an overview about the Malliavin derivative for continuous diffusions). Then taking the Malliavin derivative on both sides of (3.20), we get

$$\begin{aligned}D_s\tilde{V}(t) = \frac{h_x(S_A(T))}{S^{(0)}(T)} \cdot D_s S_A(T) + \int_t^T \nabla f_r(Y(r),\Gamma(r)) \cdot D_s(Y(r),\Gamma(r))\mathrm{d}r \\ - \int_t^T D_s Y(r)\mathrm{d}W(r) - \int_t^T \int_{\mathbb{R}} D_s Z(r,z)\tilde{N}(\mathrm{d}r,\mathrm{d}z).\end{aligned} \tag{3.21}$$

Moreover, considering the variation of V w.r.t. the initial condition, we have

$$\begin{aligned}\nabla\tilde{V}(t) = \frac{h_x(S_A(T))}{S^{(0)}(T)} \cdot \nabla S_A(T) + \int_t^T \nabla f_r(Y(r),\Gamma(r)) \cdot \nabla(Y(r),\Gamma(r))\mathrm{d}r \\ - \int_t^T \nabla Y(r)\mathrm{d}W(r) - \int_t^T \int_{\mathbb{R}} \nabla Z(r,z)\tilde{N}(\mathrm{d}r,\mathrm{d}z).\end{aligned} \tag{3.22}$$

Observe that the Malliavin derivative for the process S_A is given by (see Theorem 3 and Proposition 7 in Petrou [144] for more details)

$$D_s S_A(T) = \frac{1}{T}\int_0^T D_s S^{(1)}(r)\mathrm{d}r = \frac{1}{T}\int_0^T \frac{b(s)S^{(1)}(s)}{\nabla S^{(1)}(s)} \cdot \nabla S^{(1)}(r) I_{s\leq r}\mathrm{d}r$$
$$= \frac{b(s)S^{(1)}(s)}{\nabla S^{(1)}(s)} \cdot \frac{1}{T}\int_s^T \nabla S^{(1)}(r)\mathrm{d}r.$$

On the other hand,

$$\nabla S_A(T) = \frac{1}{T}\int_0^T \nabla S^{(1)}(r)\,\mathrm{d}r.$$

Hence, (3.22) is equivalent to

$$\nabla\tilde{V}(t) - \left(\frac{1}{T}\int_0^s \nabla S^{(1)}(r)\mathrm{d}r\right) \cdot \frac{h_x(S_A(T))}{S^{(0)}(T)}$$
$$= \frac{h_x(S_A(T))}{S^{(0)}(T)} \cdot \frac{1}{T}\int_s^T \nabla S^{(1)}(r)\mathrm{d}r + \int_t^T \nabla f_r(Y(r), Z(r,\cdot)) \cdot \nabla(Y(r), Z(r,\cdot))\mathrm{d}r$$
$$- \int_t^T \nabla Y(r)\mathrm{d}W(r) - \int_t^T\int_{\mathbb{R}} \nabla Z(r,z)\tilde{N}(\mathrm{d}r,\mathrm{d}z). \tag{3.23}$$

From (3.21) and (3.23) we have

$$D_s\tilde{V}(t) = \frac{b(s)S^{(1)}(s)}{\nabla S^{(1)}(s)}\left[\nabla\tilde{V}(t) - \left(\frac{1}{T}\int_0^s \nabla S^{(1)}(r)\mathrm{d}r\right) \cdot \frac{h_x(S_A(T))}{S^{(0)}(T)}\right]. \tag{3.24}$$

Noting that

$$\tilde{V}(t) = \tilde{V}(0) - \int_0^t f_s(Y(s),\Gamma(s))\mathrm{d}s + \int_0^t Y(s)\mathrm{d}W(s) + \int_0^t\int_{\mathbb{R}} Z(s,z)\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,,$$

we obtain

$$D_t\tilde{V}_t = Y(t).$$

Hence, (3.24) leads to

$$
\begin{aligned}
Y(t) &= \frac{b(t)S^{(1)}(t)}{\nabla S^{(1)}(t)} \left[\nabla \tilde{V}(t) - \left(\frac{1}{T} \int_0^t \nabla S^{(1)}(r) \mathrm{d}r \right) \cdot \frac{h_x(S_A(T))}{S^{(0)}(T)} \right] \\
&= \frac{b(t)S^{(1)}(t)}{\nabla S^{(1)}(t)} \left[\frac{h_x(S_A(T))}{S^{(0)}(T)} \left(\frac{1}{T} \int_t^T \nabla S^{(1)}(r) \mathrm{d}r \right) \right. \\
&\quad + \int_t^T \nabla f_r(Y(r), \Gamma(r)) \cdot \nabla (Y(r), \Gamma(r)) \mathrm{d}r \\
&\quad \left. - \int_t^T \nabla Y(r) \mathrm{d}W(r) - \int_t^T \int_{\mathbb{R}} \nabla Z(r,z) \tilde{N}(\mathrm{d}r, \mathrm{d}z) \right].
\end{aligned}
$$

Then following the same arguments as in the proof of Proposition 4.5 in Bouchard and Elie [26], we get

$$
\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |Y(s) - Y(t_i)|^2 \, \mathrm{d}s \right] \leq \frac{C}{n}. \tag{3.25}
$$

The estimate for the process Z follows as well the same arguments as in Bouchard and Elie [26] in which the authors used the relation of the BSDE of type (3.20) to partial integro-differential equations and showed

$$
\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} |Z(s,z) - Z(t_i,z)|^2 \, \ell(\mathrm{d}z) \, \mathrm{d}s \right] \leq \frac{C}{n}. \tag{3.26}
$$

If h is Lipschitz continuous with a Lipschitz constant K such that $h(0) < K$, we can use a so-called approximation argument to get the same estimations for Y and Z as in (3.25) and (3.26). We consider a $C_b^\infty(\mathbb{R})$ density q with compact support, and set

$$
h^k(x) = k \int_{\mathbb{R}} h(\bar{x}) q\left(k(x - \bar{x})\right) \mathrm{d}\bar{x} \,,
$$

where k is a positive real number. For large k, $h^k(0)$ is bounded by $2K$. Moreover, these functions h^k are K-Lipschitz in $C_b^1(\mathbb{R})$ and converge pointwise to h. We denote by (V^k, Y^k, Z^k) the solution of (3.16) where $\tilde{\xi}$ is given in (3.18) with h replaced by h^k. By Lemma A.2 in Bouchard and Elie [26], (V^k, Y^k, Z^k) converges to (V, Y, Z) in

$S^2_{[0,T]} \times H^2_{[0,T]} \times \hat{H}^2_{[0,T]}$. Since the estimation (3.25) for Y^k holds uniformly in k, then the estimate (3.25) holds for Y when h is Lipschitz continuous. Along the same lines of proof, we show that the estimate (3.26) for Z holds when h is Lipschitz continuous.

Finally, proceeding as in the proof of Theorem 2.3 in Khedher and Vanmaele [115], we prove the statement for the amount of wealth $\tilde{\Upsilon}$ in LRM strategies in the setting of Asian payoffs.

□

3.3.2 Discrete-time model

In this section, we consider a discrete-time version of the process S. Then we compute the LRM strategy related to the discrete-time model and write down the associated BSDE. Finally, we propose a backward iteration scheme to solve numerically the obtained BSDE.

We consider the FBSDE consisting of (3.9) and (3.12) but now in discrete time. Throughout this paper we shall use the notation

$$\Delta H(t_{i+1}) = H(t_{i+1}) - H(t_i)\,, \quad i = 0, \ldots, n-1\,,$$

for any process H.

Time-discretisation of the forward equation. The discrete-time version of the process S is denoted by $\hat{S}$ and defined as

$$\begin{aligned} \hat{S}(t_{i+1}) &:= \hat{S}(t_i) + \hat{S}(t_i)\left(a(t_i)\,\Delta t_i + \Delta\hat{\zeta}(t_i)\right)\,, \quad i = 0, \ldots, n-1\,, \\ \hat{S}(0) &= S(0)\,, \end{aligned} \tag{3.27}$$

where

$$\Delta\hat{\zeta}(t_{i+1}) := \int_{t_i}^{t_{i+1}} b(t_i)\,\mathrm{d}W(t) + \int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} g(z)\tilde{\gamma}(t_i)\,\tilde{N}(\mathrm{d}t, \mathrm{d}z)\,, \tag{3.28}$$

for a, b, and $\tilde{\gamma}$ as in (3.7) and to which we impose Assumptions 3.9. Recall κ in (3.10). The discrete-time specification of the predictable process λ defined in (3.3) is denoted by $\hat{\lambda}$ and is thus given by (see Schweizer [155])

$$\hat{\lambda}(t_{i+1}) = \frac{\Delta\hat{A}(t_{i+1})}{\Delta\langle\hat{M}\rangle_{t_{i+1}}} = \frac{a(t_i)}{\hat{S}(t_i)\kappa(t_i)}\,,$$

and the discrete-time specification of the MVT process K defined in (3.4) is given by

$$\hat{K}(t_{i+1}) = \sum_{0\le k\le i} \frac{a^2(t_k)\Delta t_{k+1}}{\kappa(t_k)}\,.$$

Note that (3.11) and the boundedness of a are sufficient conditions for the process $\hat{K}$ to be uniformly bounded in t_i by a constant.

Time-discretisation of the backward equation. Let $\hat{\xi}$ be the discrete approximation to the discounted payoff, i.e.,

$$\hat{\xi} = \begin{cases} h(\hat{S}^{(1)}(T))/\hat{S}^{(0)}(T)\,, & \text{vanilla options,} \\ h(\hat{S}_A(T))/\hat{S}^{(0)}(T)\,, \quad \text{with } \hat{S}_A(T) = \dfrac{1}{n}\displaystyle\sum_{i=0}^{n-1} \hat{S}^{(1)}(t_i)\,, & \text{Asian options,} \end{cases} \tag{3.29}$$

where h is Lipschitz on $\mathbb{R}$ and $h(0) = 0$. A discrete-time version of the process $\tilde{V}$ in (3.12) is given by, for all $i = 0, \dots, n-1$,

$$\begin{cases} \hat{V}(t_i) &= \hat{V}(t_{i+1}) - \hat{\Upsilon}(t_i)a(t_i)\,\Delta t_{i+1} - \hat{\Upsilon}(t_i)\Delta\hat{\zeta}(t_{i+1}) - \Delta\hat{L}^{FS}(t_{i+1})\,, \\ \hat{V}(T) &= \hat{\xi}\,, \end{cases} \tag{3.30}$$

where $\hat{\Upsilon}(t_i) = \hat{\chi}^{FS}(t_i)\hat{S}(t_i)$ is the amount of wealth to invest in an LRM strategy associated with the discrete-time model (3.27).

Note that $\hat{L}^{FS}$ in (3.30) is necessary for the existence of the solution since the predictable representation property does not hold in the discrete case (see e.g. Chapter 4 in Protter [146]). We cannot write the BSDE (3.30) as a time-discrete BSDE driven by a Brownian motion and jumps as we did in the continuous case in Proposition 3.10. Hence, we would stress that the BSDE (3.30) associated with a discrete-time model for the underlying asset price process is a discrete-time counterpart of (3.12), not just a numerical approximation for (3.12), although it is called the time-discretisation of the backward equation.

Backward iteration scheme. To obtain a backward iteration scheme for (3.30), we first take the expectation conditionally on $\mathcal{F}_{t_i}$ on both sides in (3.30) to arrive at the expression for $\hat{V}(t_i)$ (second equation in (3.31)). Then we multiply both sides in (3.30) by $\Delta\hat{\zeta}(t_{i+1})$, take conditional expectation w.r.t. $\mathcal{F}_{t_i}$, and solve for $\hat{\Upsilon}(t_i)$ using (3.30) and the fact that $[\hat{L}^{FS}, \hat{M}]$ is a $\mathbb{P}$-martingale, to find the system for $i = n-1, \dots, 0$

$$\begin{cases} \hat{\Upsilon}(t_i) = \dfrac{n}{T\kappa(t_i)}\mathbb{E}\left[\hat{V}(t_{i+1})\,\Delta\hat{\zeta}(t_{i+1})\Big|\,\mathcal{F}_{t_i}\right]\,, \\ \hat{V}(t_i) = \mathbb{E}[\hat{V}(t_{i+1})|\mathcal{F}_{t_i}] - \dfrac{Ta(t_i)}{n}\hat{\Upsilon}(t_i)\,. \end{cases} \tag{3.31}$$

As for $\hat{L}^{FS}$, being zero at zero, we have from (3.30)

$$\hat{L}^{FS}(T) = \hat{V}(T) - \sum_{i=0}^{n-1} \left[\hat{\Upsilon}(t_i) a(t_i)\, \Delta t_{i+1} + \hat{\Upsilon}(t_i) \Delta \hat{\zeta}(t_{i+1}) \right] - \hat{V}(0)\,. \tag{3.32}$$

Note that the process $\hat{L}^{FS}$ does not appear in the algorithm (3.31) and is computed afterwards. However the importance of $\hat{L}^{FS}$ is that it represents the remaining risk in the time-discretised version of the LRM hedging strategy which cannot be hedged away by trading in the underlying $\hat{S}$ and that it is needed to compute the cost process.

3.3.3 L^2-convergence of the discretisation scheme

In this section we study the convergence, in a space we specify, of the LRM strategy related to the time-discrete model to the LRM strategy related to the continuous-time model. Moreover, we compute rates for the convergence results. We will call this convergence study also a robustness study of the LRM strategy when using time discretisation since we are indeed comparing two models and two solutions of BSDEs.

In the following theorem we state the discretisation error of the approximation of (3.9) by (3.27). We refer to Platen [145] for a proof.

Theorem 3.12. *Recall the dynamics of S and $\hat{S}$ as in (3.9) and (3.27) respectively. Under Assumptions 3.9, we have*

$$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} \left| S(t) - \hat{S}(t_i) \right|^2 \right] \leq \frac{C}{n}\,,$$

for a positive constant C independent of the number of steps.

Using the Lipschitz property of the payoff function h we derive a convergence rate of the order of the time-step for the final value of the hedging portfolio.

Proposition 3.13. *Let $\tilde{V}$ and $\hat{V}$ be solutions of (3.12) and (3.30), and $\tilde{\xi}$ and $\hat{\xi}$ be as in (3.18) and (3.29) with h being Lipschitz and $h(0) = 0$. Under Assumptions 3.9, we have*

$$\mathbb{E}\left[|\tilde{V}(T) - \hat{V}(T)|^2 \right] = \mathbb{E}\left[|\tilde{\xi} - \hat{\xi}|^2 \right] \leq \frac{C}{n},$$

for a positive constant C independent of the number of steps.

Proof. For *vanilla options*, we get

$$\begin{aligned}\mathbb{E}[|\tilde{\xi}-\hat{\xi}|^2] &= \mathbb{E}\left[\left|\frac{h(S^{(1)}(T))}{S^{(0)}(T)}-\frac{h(\hat{S}^{(1)}(T))}{\hat{S}^{(0)}(T)}\right|^2\right]\\ &= \mathbb{E}\left[\left|\frac{h(S^{(1)}(T))\hat{S}^{(0)}(T)-h(\hat{S}^{(1)}(T))S^{(0)}(T)}{S^{(0)}(T)\hat{S}^{(0)}(T)}\right|^2\right].\end{aligned}$$

Since $0 < r(t) < C$, for all $t \in [0,T]$, where C is a constant, it holds that $S^{(0)}(T) > 1$ and $\hat{S}^{(0)}(T) > 1$. Hence, we find

$$\begin{aligned}\mathbb{E}[|\tilde{\xi}-\hat{\xi}|^2] &\leq \mathbb{E}\left[\left|h(S^{(1)}(T))\hat{S}^{(0)}(T)-h(\hat{S}^{(1)}(T))S^{(0)}(T)\right|^2\right]\\ &\leq C\left\{\mathbb{E}\left[\left|h(S^{(1)}(T))\right|^2\right]\mathbb{E}\left[\left|\hat{S}^{(0)}(T)-S^{(0)}(T)\right|^2\right]\right.\\ &\quad\left.+\,\mathbb{E}\left[\left|h(S^{(1)}(T))-h(\hat{S}^{(1)}(T))\right|^2\right]\mathbb{E}\left[\left|S^{(0)}(T)\right|^2\right]\right\}.\\ &\leq C\left\{\mathbb{E}\left[\left|S^{(1)}(T)\right|^2\right]\mathbb{E}\left[\left|\hat{S}^{(0)}(T)-S^{(0)}(T)\right|^2\right]\right.\\ &\quad\left.+\,\mathbb{E}\left[\left|S^{(1)}(T)-\hat{S}^{(1)}(T)\right|^2\right]\mathbb{E}\left[\left|S^{(0)}(T)\right|^2\right]\right\}.\end{aligned}$$

The last inequality results from the assumption that h is Lipschitz and $h(0) = 0$. The statement follows by using Lemma 3.2 in Benth et al. [21] and applying the analogue of Theorem 3.12 to $S^{(0)}$ and $S^{(1)}$.

In the setting of *Asian options*, the proof follows the same lines as for vanilla options. We obtain

$$\begin{aligned}\mathbb{E}[|\tilde{\xi}-\hat{\xi}|^2] \leq C\left\{\mathbb{E}\left[|S_A(T)|^2\right]\mathbb{E}\left[\left|\hat{S}^{(0)}(T)-S^{(0)}(T)\right|^2\right]\right.\\ \left.+\,\mathbb{E}\left[\left|S_A(T)-\hat{S}_A(T)\right|^2\right]\mathbb{E}\left[\left|S^{(0)}(T)\right|^2\right]\right\},\end{aligned}$$

thus we only have to prove that

$$\mathbb{E}[|S_A(T)|^2] = \mathbb{E}\left[\left(\frac{1}{T}\int_0^T S^{(1)}(t)\mathrm{d}t\right)^2\right] < \infty, \tag{3.33}$$

and

$$\mathbb{E}\left[\left|S_A(T) - \frac{1}{T}\sum_{i=0}^{n-1}\hat{S}^{(1)}(t_i)\,\Delta t_{i+1}\right|^2\right] \leq \frac{C}{n}\,. \tag{3.34}$$

On the one hand, since $S^{(1)}$ is a square integrable semimartingale and Fubini's theorem is applicable we find

$$\begin{aligned}\mathbb{E}\left[\left(\frac{1}{T}\int_0^T S^{(1)}(t)\mathrm{d}t\right)^2\right] \leq & C\mathbb{E}\left[\int_0^T \left(S^{(1)}(t)\right)^2\mathrm{d}t\right]\\ \leq & C\int_0^T \mathbb{E}\left[\left(S^{(1)}(t)\right)^2\right]\mathrm{d}t < \infty.\end{aligned}$$

Hence, (3.33) holds. On the other hand, by introducing the piecewise constant function

$$\overline{S}^{(1)}(s) = \sum_{i=0}^{n-1}\hat{S}^{(1)}(t_i)I_{[t_i,t_{i+1}]}(s),$$

and recalling (3.18), we find

$$\begin{aligned}&\mathbb{E}\left[\left|S_A(T) - \frac{1}{T}\sum_{i=0}^{n-1}\hat{S}^{(1)}(t_i)\,\Delta t_{i+1}\right|^2\right]\\ &= \frac{1}{T^2}\mathbb{E}\left[\left|\int_0^T S^{(1)}(s)\mathrm{d}s - \int_0^T \overline{S}^{(1)}(s)\mathrm{d}s\right|^2\right]\\ &\leq C\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\left|S^{(1)}(s) - \hat{S}^{(1)}(t_i)\right|^2\mathrm{d}s\right].\end{aligned}$$

Hence, the estimate (3.34) holds due to the analogue of Theorem 3.12 for $S^{(1)}$ and the statement follows for Asian options. □

To study the convergence of the time-discrete scheme, we consider a continuous-time version of the process $\hat{\zeta}$ (3.28) as follows

$$\zeta_1(t) = \zeta_1(t_i) + \int_{t_i}^t b(t_i)\mathrm{d}W(s) + \int_{t_i}^t\int_{\mathbb{R}}\gamma(t_i,z)\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,, \qquad \zeta_1(0) = 0\,.$$

Now consider the $\mathcal{F}_{t_{i+1}}$-measurable random variable

$$\xi(t_{i+1}) := \hat{V}(t_{i+1}) - \int_{t_i}^{t_{i+1}} \hat{\Upsilon}(t_i) a(t_i)\,\mathrm{d}s\,.$$

We know from the GKW decomposition (see, e.g., Ansel and Stricker [3]) that there exists a predictable process $\Upsilon_1 \in \tilde{H}^2_{[t_i,t_{i+1}]}$ such that

$$\xi(t_{i+1}) = \mathbb{E}\left[\xi(t_{i+1})\,|\mathcal{F}_{t_i}\right] + \int_{t_i}^{t_{i+1}} \Upsilon_1(s)\,\mathrm{d}\zeta_1(s) + \Delta L_1(t_{i+1})\,,$$

where L_1 is a square integrable $\mathbb{P}$-martingale such that $[\zeta_1, L_1]$ is a $\mathbb{P}$-martingale. From the latter equation and the second equality in (3.31), we deduce

$$\begin{aligned}\hat{V}(t_{i+1}) =& \hat{V}(t_i) + \int_{t_i}^{t_{i+1}} \hat{\Upsilon}(t_i) a(t_i)\,\mathrm{d}s + \int_{t_i}^{t_{i+1}} \Upsilon_1(s)\,\mathrm{d}\zeta_1(s) \\ &+ \Delta L_1(t_{i+1})\,. \end{aligned} \tag{3.35}$$

We define a continuous version of $\hat{V}$ as follows

$$\begin{aligned}\hat{V}(t) := \hat{V}(t_{i+1}) - &\int_{t}^{t_{i+1}} \hat{\Upsilon}(t_i) a(t_i)\,\mathrm{d}s - \int_{t}^{t_{i+1}} \Upsilon_1(s)\,\mathrm{d}\zeta_1(s) \\ &- L_1(t_{i+1}) + L_1(t)\,. \end{aligned} \tag{3.36}$$

The latter is an "intermediate" time-continuous BSDEJ which is needed for the convergence study later on. Since we are in a time-continuous setting, we can apply the classical martingale representation to L_1 to find

$$L_1(t_{i+1}) = L_1(t) + \int_{t}^{t_{i+1}} P_1(s)\,\mathrm{d}W(s) + \int_{t}^{t_{i+1}} \int_{\mathbb{R}} Q_1(s,z)\,\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,,$$

where $P_1 \in H^2_{[t_i,t_{i+1}]}$ and $Q_1 \in \hat{H}^2_{[t_i,t_{i+1}]}$. Substituting the latter in (3.36) for the continuous version of $\hat{V}$ leads to

$$\hat{V}(t) := \hat{V}(t_{i+1}) - \int_{t}^{t_{i+1}} \hat{\Upsilon}(t_i) a(t_i)\,\mathrm{d}s - \int_{t}^{t_{i+1}} Y_1(t_i,s)\,\mathrm{d}W(s)$$

$$- \int_t^{t_{i+1}} \int_{\mathbb{R}} Z_1(t_i, s, z)\, \tilde{N}(\mathrm{d}s, \mathrm{d}z)\,, \tag{3.37}$$

where

$$\begin{aligned} Y_1(t_i, s) = &\quad \Upsilon_1(s) b(t_i) + P_1(s)\,, \\ Z_1(t_i, s, z) = &\quad \Upsilon_1(s)\gamma(t_i, z) + Q_1(s, z)\,. \end{aligned} \tag{3.38}$$

Using the fact that $[\zeta_1, L_1]$ is a $\mathbb{P}$-martingale, we deduce, for $i = 0, ..., n-1$,

$$\Upsilon_1(s) = \frac{1}{\kappa(t_i)} \left(b(t_i) Y_1(t_i, s) + \int_{\mathbb{R}} Z_1(t_i, s, z)\gamma(t_i, z)\, \ell(\mathrm{d}z) \right)\,,\ t_i \le s \le t_{i+1}. \tag{3.39}$$

Multiplying by $\Delta\zeta_1(t_{i+1})$ in both sides in (3.35) and taking conditional expectation with respect to $\mathcal{F}_{t_i}$, we obtain

$$\mathbb{E}\left[\hat{V}(t_{i+1})\Delta\zeta_1(t_{i+1})\,|\mathcal{F}_{t_i}\right] = \kappa(t_i)\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Upsilon_1(s)\,\mathrm{d}s\,|\mathcal{F}_{t_i}\right]\,,$$

where κ is as in (3.10). Comparing the latter to the first equality in (3.31), we get

$$\hat{\Upsilon}(t_i) = \frac{n}{T}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Upsilon_1(s)\,\mathrm{d}s\,|\mathcal{F}_{t_i}\right]\,.$$

In the following two propositions, we compute estimates which we use later in the proofs of the convergence results.

Proposition 3.14 ([115]). *Let $\tilde{\Upsilon}$ and Υ_1 be respectively as in (3.12) and (3.35). Let Assumptions 3.9 hold and assume that a, b, and $\tilde{\gamma}$ are Lipschitz. Introduce the notation*

$$\delta Y(t) = Y(t) - Y_1(t_i, t)\,, \qquad \delta Z(t, z) = Z(t, z) - Z_1(t_i, t, z)\,, \tag{3.40}$$

where Y, Z are as in (3.16) and Y_1, Z_1 are as in (3.38). Then

$$\begin{aligned} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\tilde{\Upsilon}(s) - \Upsilon_1(s)|^2\right]\mathrm{d}s \le \frac{K}{n^2} &+ C\int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\delta Y(s)|^2\right]\mathrm{d}s \\ &+ C\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} \mathbb{E}\left[|\delta Z(s, z)|^2\right]\ell(\mathrm{d}z)\,\mathrm{d}s\,, \end{aligned}$$

where K and C are positive constants.

Proposition 3.15 ([115]). *Let $\tilde{\Upsilon}$ and $\hat{\Upsilon}$ be respectively as in (3.12) and (3.30). Let Assumptions 3.9 hold and assume that a, b, and $\tilde{\gamma}$ are Lipschitz. Then we have*

$$\begin{aligned}
&\int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\hat{\Upsilon}(t_i) - \tilde{\Upsilon}(s)|^2\right] \mathrm{d}s \\
&\leq \frac{K}{n^2} + C \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} \mathbb{E}\left[|\delta Z(s,z)|^2\right] \ell(\mathrm{d}z)\, \mathrm{d}s + C \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\delta Y(s)|^2\right] \mathrm{d}s \\
&\quad + C \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\tilde{\Upsilon}(s) - \tilde{\Upsilon}(t_i)|^2\right] \mathrm{d}s\,,
\end{aligned}$$

where K and C are positive constants and δY and δZ are as in (3.40).

Theorem 3.16. *Assume that the conditions of Theorem 3.11 hold. Let the triplets $(\tilde{V}, Y, Z)$ and $(\hat{V}, Y_1, Z_1)$ be respectively the solutions of (3.16) and (3.37). Then, it holds*

$$\begin{aligned}
&\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} \left|\tilde{V}(t) - \hat{V}(t)\right|^2\right] + \sum_{i=0}^{n-1} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |Y(s) - Y_1(t_i,s)|^2 \,\mathrm{d}s\right] \\
&\quad + \sum_{i=0}^{n-1} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} |Z(s,z) - Z_1(t_i,s,z)|^2 \,\ell(\mathrm{d}z)\,\mathrm{d}s\right] \leq \frac{C}{n}, \qquad (3.41)
\end{aligned}$$

where C is a positive constant.

Proof. Set $\delta\tilde{V}(t) = \tilde{V}(t) - \hat{V}(t)$, $\delta a(t) = a(t) - a(t_i)$, $\delta\tilde{\Upsilon}(t) = \tilde{\Upsilon}(t) - \hat{\Upsilon}(t_i)$. Recall the notations δY and δZ in (3.40). In the sequel, C denotes a positive constant independent of i and n and may take different values from line to line. Applying Itô's Lemma, we get

$$\begin{aligned}
A(t) &:= \mathbb{E}[|\delta\tilde{V}(t)|^2] - \mathbb{E}[|\delta\tilde{V}(t_{i+1})|^2] + \mathbb{E}\left[\int_t^{t_{i+1}} |\delta Y(s)|^2 \,\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_t^{t_{i+1}} \int_{\mathbb{R}} |\delta Z(s,z)|^2 \,\ell(\mathrm{d}z)\,\mathrm{d}s\right] \qquad (3.42)
\end{aligned}$$

$$= \mathbb{E}\left[\int_t^{t_{i+1}} 2\delta\tilde{V}(s)\left(-a(s)\tilde{\Upsilon}(s) + a(t_i)\hat{\Upsilon}(t_i)\right)\mathrm{d}s\right]$$
$$\leq \mathbb{E}\left[\int_t^{t_{i+1}} |2\delta\tilde{V}(s)\delta\tilde{\Upsilon}(s)a(t_i)|\,\mathrm{d}s\right] + \mathbb{E}\left[\int_t^{t_{i+1}} |2\delta\tilde{V}(s)\tilde{\Upsilon}(s)\delta a(s)|\,\mathrm{d}s\right].$$

Using $2ab \leq \alpha a^2 + b^2/\alpha$, for some $\alpha > 0$, and a, b, $\tilde{\gamma}$ are Lipschitz, we get

$$\begin{aligned}
A(t) &\leq \alpha\mathbb{E}\left[\int_t^{t_{i+1}} |\delta\tilde{V}(s)|^2\,\mathrm{d}s\right] + \frac{CT}{n}\mathbb{E}\left[\int_t^{t_{i+1}} 2|\delta\tilde{V}(s)\tilde{\Upsilon}(s)|\,\mathrm{d}s\right] \\
&\quad + \frac{C}{\alpha}\mathbb{E}\left[\int_t^{t_{i+1}} |\tilde{\Upsilon}(s) - \hat{\Upsilon}(t_i)|^2\mathrm{d}s\right] \\
&\leq 2\alpha\mathbb{E}\left[\int_t^{t_{i+1}} |\delta\tilde{V}(s)|^2\,\mathrm{d}s\right] + \frac{C}{\alpha n^2}\mathbb{E}\left[\int_t^{t_{i+1}} |\tilde{\Upsilon}(s)|^2\,\mathrm{d}s\right] \\
&\quad + \frac{C}{\alpha}\mathbb{E}\left[\int_t^{t_{i+1}} |\tilde{\Upsilon}(s) - \hat{\Upsilon}(t_i)|^2\mathrm{d}s\right] \\
&\leq 2\alpha\mathbb{E}\left[\int_t^{t_{i+1}} |\delta\tilde{V}(s)|^2\,\mathrm{d}s\right] + \frac{C}{\alpha n^2}\mathbb{E}\left[\int_0^{T} |\tilde{\Upsilon}(s)|^2\,\mathrm{d}s\right] \\
&\quad + \frac{C}{\alpha}\mathbb{E}\left[\int_t^{t_{i+1}} |\tilde{\Upsilon}(s) - \hat{\Upsilon}(t_i)|^2\mathrm{d}s\right] \\
&\leq 2\alpha\mathbb{E}\left[\int_t^{t_{i+1}} |\delta\tilde{V}(s)|^2\,\mathrm{d}s\right] + \frac{C}{n^2} + \frac{C}{\alpha}\mathbb{E}\left[\int_t^{t_{i+1}} |\tilde{\Upsilon}(s) - \hat{\Upsilon}(t_i)|^2\mathrm{d}s\right] \qquad (3.43)
\end{aligned}$$

Recall the expression of A in (3.42). We deduce from (3.43)

$$\begin{aligned}
\mathbb{E}[|\delta\tilde{V}(t)|^2] \leq \mathbb{E}[|\delta\tilde{V}(t)|^2] &+ \mathbb{E}\left[\int_t^{t_{i+1}} |\delta Y(s)|^2\,\mathrm{d}s\right] \\
&+ \mathbb{E}\left[\int_t^{t_{i+1}}\int_{\mathbb{R}} |\delta Z(s,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right] \qquad (3.44)
\end{aligned}$$

$$\leq C\alpha \int_t^{t_{i+1}} \mathbb{E}\left[|\delta\tilde{V}(s)|^2\right] \mathrm{d}s + B_i\,,$$

where

$$B_i = \frac{C}{n^2} + \mathbb{E}\left[|\delta\tilde{V}(t_{i+1})|^2\right] + \frac{C}{\alpha}\mathbb{E}\left[|\tilde{\Upsilon}(s) - \hat{\Upsilon}(t_i)|^2\right] \mathrm{d}s\,.$$

Thus applying Gronwall's lemma to (3.44), we get

$$\mathbb{E}\left[|\delta\tilde{V}(t)|^2\right] \leq B_i \exp\left\{\frac{C\alpha}{n}\right\}, \qquad t_i \leq t < t_{i+1}\,,\ i = 0, ..., n-1\,,$$

which plugged in (3.44), implies

$$\begin{aligned}
&\mathbb{E}[|\delta\tilde{V}(t)|^2] + \int_t^{t_{i+1}} \mathbb{E}\left[|\delta Y(s)|^2\right] \mathrm{d}s + \int_t^{t_{i+1}} \int_{\mathbb{R}} \mathbb{E}\left[|\delta Z(s,z)|^2\right]\, \ell(\mathrm{d}z)\, \mathrm{d}s \\
&\leq B_i \left(1 + \alpha\frac{C}{n}\right)\,. \qquad (3.45)
\end{aligned}$$

Taking $t = t_i$ and applying Proposition 3.15, we get

$$\begin{aligned}
&\mathbb{E}[|\delta\tilde{V}(t_i)|^2] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2 \,\mathrm{d}s\right] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} |\delta Z(s,z)|^2\, \ell(\mathrm{d}z)\, \mathrm{d}s\right] \\
&\leq \left(1 + \alpha\frac{C}{n}\right)\left(\frac{C}{n^2} + \mathbb{E}\left[|\delta\tilde{V}(t_{i+1})|^2\right] + \frac{C}{\alpha}\int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\tilde{\Upsilon}(s) - \tilde{\Upsilon}(t_i)|^2\right] \mathrm{d}s\right. \\
&\left. + \frac{C}{\alpha}\int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\delta Y(s)|^2\right] \mathrm{d}s + \frac{C}{\alpha}\int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} \mathbb{E}\left[|\delta Z(s,z)|^2\right]\, \ell(\mathrm{d}z)\, \mathrm{d}s\right)\,.
\end{aligned}$$

For α sufficiently larger than C, we deduce

$$\begin{aligned}
&\mathbb{E}[|\delta\tilde{V}(t_i)|^2] + \frac{1}{2}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2 \,\mathrm{d}s\right] + \frac{1}{2}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} |\delta Z(s,z)|^2\, \ell(\mathrm{d}z)\, \mathrm{d}s\right] \\
&\leq \left(1 + \frac{C}{n}\right)\left(\frac{C}{n^2} + \mathbb{E}\left[|\delta\tilde{V}(t_{i+1})|^2\right] + C\int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\tilde{\Upsilon}(s) - \tilde{\Upsilon}(t_i)|^2\right] \mathrm{d}s\right)\,.
\end{aligned}$$

Iterating the last inequality we get

$$\mathbb{E}[|\delta\tilde{V}(t_i)|^2] + \frac{1}{2}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2\,\mathrm{d}s\right] + \frac{1}{2}\,\mathbb{E}\left[\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} |\delta Z(s,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right]$$
$$\leq C\left(1+\frac{C}{n}\right)^n\left(\frac{1}{n} + \mathbb{E}\left[|\delta\tilde{V}(T)|^2\right] + \sum_{j=i}^{n-1}\int_{t_j}^{t_{j+1}} \mathbb{E}\left[|\Upsilon(s)-\Upsilon(t_i)|^2\right]\mathrm{d}s\right). \quad (3.46)$$

Using the estimates in Lemma 3.11 and Proposition 3.13, we have

$$\mathbb{E}\left[|\delta\tilde{V}(t_i)|^2\right] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2\,\mathrm{d}s\right] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} |\delta Z(s,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right] \leq \frac{C}{n}. \quad (3.47)$$

Taking $t = t_i$ in (3.45), summing up, and using Proposition 3.15, we arrive at

$$\sum_{i=0}^{n-1}\left(\mathbb{E}[|\delta\tilde{V}(t_i)|^2] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2\,\mathrm{d}s\right] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} |\delta Z(s,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right]\right)$$
$$\leq \sum_{i=0}^{n-1}\left[\left(1+\alpha\frac{C}{n}\right)\left(\frac{C}{n^2} + \mathbb{E}\left[|\delta\tilde{V}(t_{i+1})|^2\right] + \frac{C}{\alpha}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2\,\mathrm{d}s\right]\right.\right.$$
$$\left.\left. + \frac{C}{\alpha}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} |\delta Z(s,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right] + \frac{C}{\alpha}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\tilde{\Upsilon}(s)-\tilde{\Upsilon}(t_i)|^2\,\mathrm{d}s\right]\right)\right],$$

which implies

$$\left[1-\tfrac{C}{\alpha}\left(1+\tfrac{\alpha C}{n}\right)\right]\textstyle\sum_{i=0}^{n-1}\left(\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2\,\mathrm{d}s\right] + \mathbb{E}\left[\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} |\delta Z(s,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right]\right)$$
$$\leq \left(1+\alpha\frac{C}{n}\right)\frac{C}{n} + \left(1+\alpha\frac{C}{n}\right)\mathbb{E}\left[|\delta\tilde{V}(T)|^2\right] - \mathbb{E}\left[\delta\tilde{V}(t_0)|^2\right]$$
$$+ \alpha\frac{C}{n}\sum_{i=1}^{n-1}\mathbb{E}\left[\delta\tilde{V}(t_i)|^2\right] + \left(1+\alpha\frac{C}{n}\right)\frac{C}{\alpha}\sum_{i=0}^{n-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\tilde{\Upsilon}(s)-\tilde{\Upsilon}(t_i)|^2\,\mathrm{d}s\right].$$

Using Proposition 3.13, the latter implies that for α sufficiently larger than C, we obtain

$$\sum_{i=0}^{n-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}} |\delta Y(s)|^2\,\mathrm{d}s\right] + \sum_{i=0}^{n-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} |\delta Z(s,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right]$$

$$\leq C\left[\frac{1}{n}+\frac{1}{n}\sum_{i=1}^{n-1}\mathbb{E}\left[\delta\tilde{V}(t_i)|^2\right]+\sum_{i=0}^{n-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}|\tilde{\Upsilon}(s)-\tilde{\Upsilon}(t_i)|^2\,\mathrm{d}s\right]\right] \tag{3.48}$$

and the statement for the last two terms in (3.41) follows using (3.47) and Lemma 3.11. Finally, observe that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t_i\leq t\leq t_{i+1}}|\tilde{V}(t)-\hat{V}(t)|^2\right]\\
&\leq K\left(\mathbb{E}\left[|\tilde{V}(t_{i+1})-\hat{V}(t_{i+1})|^2\right]+\mathbb{E}\left[\int_{t_i}^{t_{i+1}}\left|\tilde{\Upsilon}(s)a(s)-\hat{\Upsilon}(t_i)a(t_i)\right|^2\mathrm{d}s\right]\right.\\
&\quad+\mathbb{E}\left[\sup_{t_i\leq t\leq t_{i+1}}\left|\int_t^{t_{i+1}}\delta Y(s)\,\mathrm{d}W(s)\right|^2\right]\\
&\quad\left.+\mathbb{E}\left[\sup_{t_i\leq t\leq t_{i+1}}\left|\int_t^{t_{i+1}}\int_{\mathbb{R}}|\delta Z(s,z)|^2\,\tilde{N}(\mathrm{d}s,\mathrm{d}z)\right|^2\right]\right).
\end{aligned}$$

Then using Burkholder's inequality, and iterating as we did to get (3.46), we deduce the result applying (3.48). □

Theorem 3.16 implies that Theorem 4.4 in Khedher and Vanmaele [115] holds. Hence, the robustness results for the LRM strategy in Khedher and Vanmaele [115] can be applied in the current setting. The following robustness results are direct applications of Theorem 4.4, Theorem 4.5, and Theorem 4.6 in Khedher and Vanmaele [115].

Theorem 3.17. *Let $\tilde{V}$ and $\hat{V}$ be respectively defined by (3.12) and (3.30). Under the Assumptions 3.9 on the parameters of the model, and the assumptions that h is Lipschitz and $h(0)=0$, it holds that*

$$\max_{i<n}\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]}|\tilde{V}(t)-\hat{V}(t_i)|^2\right]\leq\frac{C}{n},$$

where C is a positive constant.

The robustness of the amount of wealth in an LRM strategy follows in the next theorem.

Theorem 3.18. *Let $\tilde{\Upsilon}$ and $\hat{\Upsilon}$ be the amount of wealth defined as in (3.12) and (3.30). Under the conditions of Theorem 3.17 it holds that*

$$\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}|\tilde{\Upsilon}(s)-\hat{\Upsilon}(t_i)|^2\,\mathrm{d}s\right]\leq\frac{C}{n}\,,$$

where C is a positive constant.

We state the robustness of remaining risk L^{FS} in the following theorem.

Theorem 3.19. *Let the processes L^{FS} and $\hat{L}^{FS}$ be as defined in (3.12) and (3.30). Under the conditions of Theorem 3.17, we have for all $0\leq i\leq n-1$,*

$$\mathbb{E}\left[\left|L^{FS}(t_{i+1})-\hat{L}^{FS}(t_{i+1})\right|^2\right]\leq\frac{C}{n}\,,$$

where C is a positive constant.

Notice that this latter result implies the robustness of the cost process in an LRM strategy. Indeed the cost processes in an LRM strategy related to the price processes S and $\hat{S}$ are respectively given by $C(t)=L^{FS}(t)+V(0)$, $t\in[0,T]$, and $\hat{C}(t_i)=\hat{L}^{FS}(t_i)+\hat{V}(0)$, $i=0,\ldots,n$. Then applying Theorems 3.17 and 3.19, we get for all $i=0,\ldots,n$,

$$\mathbb{E}[|C(t_i)-\hat{C}(t_i)|^2]\leq C\mathbb{E}[|L^{FS}(t_i)-\hat{L}^{FS}(t_i)|^2]+\mathbb{E}[|V(0)-\hat{V}(0)|^2]\leq\frac{C}{n}\,.$$

3.3.4 Robustness study

In the previous subsections, we imposed that the jumps deriving the asset price process has finite activity, i.e. $\ell(|z|\leq 1)<\infty$, where ℓ is the Lévy measure such that $\mathbb{E}[N(\mathrm{d}t,\mathrm{d}z)]=\ell(\mathrm{d}z)\,\mathrm{d}t$. It is easy to simulate the jumps in this setting. However, if the Lévy measure has infinite activity, i.e. $\ell(|z|\leq 1)=\infty$, simulation of these jumps is a demanding task. Asmussen and Rosinsky [7] proposed to shift from a model with small jumps to another where those variations are represented by some appropriately scaled continuous component. The idea is to approximate the small jumps in the martingale ζ (3.14) by a Brownian motion B which is scaled with the stand deviation of the small jumps. That is, we will approximate ζ by ζ_ε,

$$\zeta_\varepsilon(t)=\int_0^t b(s)\,\mathrm{d}W(s)+\int_0^t G(\varepsilon)\tilde{\gamma}(s)\,\mathrm{d}B(s)+\int_0^t\int_{|z|>\varepsilon}\gamma(s,z)\,\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,. \tag{3.49}$$

Then, we enlarge the filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ to $\mathbb{G} = (\mathcal{G}_t)_{0\leq t\leq T}$ with the information of the Brownian motion B, i.e.

$$\mathcal{G}_t = \sigma\left\{W(s), B(s), \int_0^s \int_A \tilde{N}(\mathrm{d}u, \mathrm{d}z), \quad s \leq t, \quad A \in \mathcal{B}(\mathbb{R})\right\} \vee \mathcal{N}\,, \tag{3.50}$$

where $\mathcal{N}$ represents the set of $\mathbb{P}$-null events in $\mathcal{F}$. We define the process S_ε as follows

$$\mathrm{d}S_\varepsilon(t) = S_\varepsilon(t-)\left(a(t)\,\mathrm{d}t + \mathrm{d}\zeta_\varepsilon(t)\right), \tag{3.51}$$

where a is the same as that in (3.9). If S_ε is selected as a model for the underlying asset price process, the LRM strategy can also be formulated as the solution to a BSDE

$$\begin{cases} \mathrm{d}\tilde{V}_\varepsilon(t) = \tilde{\Upsilon}_\varepsilon(t)a(t)\,\mathrm{d}t + \tilde{\Upsilon}_\varepsilon(t)\,\mathrm{d}\zeta_\varepsilon(t) + \mathrm{d}L_\varepsilon^{FS}(t), \\ \tilde{V}_\varepsilon(T) = \tilde{\xi}_\varepsilon\,, \end{cases} \tag{3.52}$$

where $\tilde{\xi}_\varepsilon$ is the discounted payoff and the $\mathbb{G}$-predictable process $\tilde{\Upsilon}_\varepsilon$ defined as $\tilde{\Upsilon}_\varepsilon(t) = \chi_\varepsilon^{FS}(t)S_\varepsilon(t-)$ is the amount of wealth to invest in the LRM strategy.

Recall that the SDE (3.9) can be re-written as

$$\mathrm{d}S(t) = S(t-)\left(a(t)\,\mathrm{d}t + \mathrm{d}\zeta(t)\right).$$

Notice that by scaling the Brownian motion with the standard deviation of the small jumps, both processes S and S_ε have the same variance when ε goes to zero. It means that $(S_\varepsilon(t))_{0\leq t\leq T}$ converges to $(S(t))_{0\leq t\leq T}$ in an L^2-sense when ε goes to zero. However, S and S_ε are indeed two different models for the underlying assets.

From a modelling point of view, the difference of the two models could lead to model uncertainty, since it is difficult to confirm whether the jump process has finite activity or infinite activity. Hence, the issue is whether the induced LRM hedging strategies from S_ε converge to the LRM strategies from S. From a simulation point view, the issue is whether the simulated LRM strategy $\hat{\Upsilon}_\varepsilon$ with the scheme proposed in Section 3.3.2 converges to $\tilde{\Upsilon}$. In this setting, we are actually confronted with model uncertainty in the discrete-time models.

Theorem 5.7 in Khedher and Vanmaele [115] states that the numerical solution of the BSDE (3.52) with the proposed scheme in Section 3.3.2 converges to the original solution of the BSDE (3.12) in an L^2-sense. It means that in the present model setting, the LRM strategies for vanilla and Asian options are robust with respect to model uncertainty in both continuous-time and discrete-time senses.

Remark 3.20. *Except for the approximation method* (3.49)*, an alternative approximation method is to ignore the small jumps (see Di Nunno et al. [63]). The numerical solution of the induced BSDE with the proposed scheme in Section 3.3.2 also converges to the solution of the BSDE* (3.12) *in an* L^2*-sense [115].*

3.4 Discretisation of LRM strategies for spread and basket options

Taking the correlation between different risk factors into account, we investigate the LRM strategies for spread and/or basket options and we propose an Euler discretisation method for the simulation of such strategies. For this purpose we first extend the approach by Bouchard and Elie [26] to allow for options written on two correlated asset prices. Then we exploit the results by Khedher and Vanmaele [115] for the study of the convergence rates.

In the analysis below we consider for the ease of the exposition $d = 2$ but it can be extended to the case of a general $d(\geq 2)$ using matrix notations. For the application of spread options $d = 2$ is sufficient but for basket options one could think of baskets of dimensions $d > 2$.

3.4.1 Continuous-time model: two-dimensional jump-diffusion

Let the riskless asset $S^{(0)}$ be as in (3.3.1) and $S^{(\cdot)} := (S^{(1)}, S^{(2)})$ be a two-dimensional price process with corresponding discounted price process $S := (S_1, S_2)$ having dynamics

$$\begin{cases} \mathrm{d}S_1(t) = S_1(t-)a_1(t)\,\mathrm{d}t + S_1(t-)\mathrm{d}\zeta_1(t)\,, & S_1(0) = s_1 \geq 0\,, \\ \mathrm{d}S_2(t) = S_2(t-)a_2(t)\,\mathrm{d}t + S_2(t-)\mathrm{d}\zeta_2(t)\,, & S_2(0) = s_2 \geq 0\,, \end{cases} \tag{3.53}$$

where $a_i(t) = \delta_i(t) - r(t)$, $t \in [0,T]$, $i = 1,2$ and

$$\mathrm{d}\zeta_i(t) = b_i(t)\mathrm{d}W_i(t) + \int_{\mathbb{R}} \gamma_i(t,z)\tilde{N}_i(\mathrm{d}t,\mathrm{d}z)\,, \quad i = 1,2\,. \tag{3.54}$$

We impose Assumptions 3.9 on the parameters δ_i, r, b_i, γ_i, $i = 1,2$. Moreover, we assume that W_1 and W_2 are two correlated standard Brownian motions such that $\mathrm{d}W_1(t)\,\mathrm{d}W_2(t) = \rho\,\mathrm{d}t$, $t \in [0,T]$, $\rho \in (-1,1)$, and $\tilde{N}_1$ and $\tilde{N}_2$ are two independent centred Poisson random measures. Notice that there exist two independent Brownian

motions $W^{(1)}$ and $W^{(2)}$ such that $W_1 = W^{(1)}$ and $W_2 = \rho W^{(1)} + \sqrt{1-\rho^2}W^{(2)}$. Thus the σ-algebra generated by S_1 and by S_2 is the σ-algebra $(\mathcal{F}_t)_{0\leq t\leq T}$ defined in (3.50).

From (3.53), it is easy to verify that S can be decomposed into the two-dimensional square integrable martingale M starting in 0 and a predictable finite variation process A, with $A(0) = 0$ and such that $\langle M\rangle$ is a 2×2-dimensional process with components $\langle M_i, M_j\rangle$, $i,j \in \{1,2\}$, given by

$$\begin{aligned}
\mathrm{d}\langle M_i\rangle_t &= \kappa_i(t)\,(S_i(\,t))^2\mathrm{d}t\,, \quad i = 1,2\\
\mathrm{d}\langle M_i, M_j\rangle_t &= \kappa_{12}(t)S_1(t)S_2(t)\mathrm{d}t\,, \quad i\neq j,\ i,j = 1,2
\end{aligned}$$

where

$$\kappa_i(t) := b_i^2(t) + \int_{\mathbb{R}} \gamma_i^2(t,z)\ell_i(\mathrm{d}z) \quad \text{and} \quad \kappa_{12}(t) := b_1(t)b_2(t)\rho\,. \tag{3.55}$$

Assumptions 3.9 for b_i and γ_i imply

$$\overline{\kappa}(t) := \kappa_1(t)\kappa_2(t) - \kappa_{12}^2(t) \geq C\,, \tag{3.56}$$

for a strictly positive constant C. Then the predictable process λ defined in (3.3) and associated with the process S (3.53) is given by

$$(\lambda_1(t), \lambda_2(t)) = \left(\frac{\kappa_2(t)a_1(t) - a_2(t)\kappa_{12}(t)}{\overline{\kappa}(t)S_1(t-)}\,, \frac{\kappa_1(t)a_2(t) - a_1(t)\kappa_{12}(t)}{\overline{\kappa}(t)S_2(t-)}\right), \tag{3.57}$$

for $0 \leq t \leq T$. The MVT process defined in (3.4) and associated with the process S (3.53) is given by

$$K(t) = \int_0^t \frac{\kappa_1(s)a_2^2(s) + \kappa_2(s)a_1^2(s) - 2a_1(s)a_2(s)\kappa_{12}(s)}{\overline{\kappa}(s)}\,\mathrm{d}s\,. \tag{3.58}$$

Assumptions 3.9 for a_i, b_i and γ_i imply condition (3.56) and hence that K is uniformly bounded in t by a constant.

We consider payoffs of the form $h(S^{(\cdot)}(T))$, where h is assumed to satisfy $h(0,0) = 0$ in the sequel. Otherwise, we can work with $\hat{h} = h - C$, for $h(0,0) = C$, where C is a constant. In addition, h is assumed to be a Lipschitz function written on the two-dimensional process $S^{(\cdot)} = S^{(0)}S$ determined by (3.53). h can be for example a payoff of a spread or a basket option (see Carmone and Durrleman [33] and Xu and Zheng

[179] for more about pricing such options respectively in continuous and jump-diffusion setting). Let the contingent claim be given by $\xi = h(S^{(\cdot)}(T))$ and its discounted value by $\tilde{\xi} = \xi/S^{(0)}(T)$, which is assume to be square integrable. The FS-decomposition (3.5) written under the real world measure $\mathbb{P}$ for the two-dimensional price process S leads to the following BSDE for $\tilde{V}$

$$\begin{cases} \tilde{V}(t) = \tilde{V}(T) - \displaystyle\int_t^T \left\{ \tilde{\Upsilon}_1(s) a_1(s) + \tilde{\Upsilon}_2(s) a_2(s) \right\} \mathrm{d}s - \int_t^T \tilde{\Upsilon}_1(s)\, \mathrm{d}\zeta_1(s) \\ \qquad\qquad - \displaystyle\int_t^T \tilde{\Upsilon}_2(s)\, \mathrm{d}\zeta_2(s) - L^{FS}(T) + L^{FS}(t)\,, \\ \tilde{V}(T) = \tilde{\xi}\,, \end{cases} \tag{3.59}$$

where $\tilde{\Upsilon}_i(t) = \chi_i^{FS}(t) S_i(t-)$ is the amount of wealth invested in the risky asset S_i, for $i = 1, 2$. Moreover, from the FS-decomposition in the LRM strategy we know that $\tilde{\Upsilon} = (\tilde{\Upsilon}_1, \tilde{\Upsilon}_2) \in \Theta$ and L^{FS} is a square integrable $\mathbb{P}$-martingale such that $[L^{FS}, M_1]$ and $[L^{FS}, M_2]$ are $\mathbb{P}$-martingales. Since $L^{FS}(T)$ is an $\mathcal{F}_T$-measurable square integrable random variable, then applying the representation theorem for the process L^{FS}, we get

$$\begin{aligned} L^{FS}(t) = &\int_0^t P(s) \mathrm{d}W^{(1)}(s) + \int_0^t Q(s) \mathrm{d}W^{(2)}(s) + \int_0^t \int_{\mathbb{R}} R_1(s,z) \tilde{N}_1(\mathrm{d}s, \mathrm{d}z) \\ &+ \int_0^t \int_{\mathbb{R}} R_2(s,z) \tilde{N}_2(\mathrm{d}s, \mathrm{d}z)\,, \end{aligned} \tag{3.60}$$

where $P, Q \in H^2_{[0,T]}$ and $R_1, R_2 \in \hat{H}^2_{[0,T]}$. We introduce the following classical BSDE

$$\begin{cases} -\mathrm{d}\tilde{V}(t) = f(t, Y(t), Z(t), \Gamma_1(t), \Gamma_2(t)) \mathrm{d}t - Y(t) \mathrm{d}W^{(1)}(t) - Z(t) \mathrm{d}W^{(2)}(t) \\ \qquad\qquad - \displaystyle\int_{\mathbb{R}} U_1(t,z) \tilde{N}_1(\mathrm{d}t, \mathrm{d}z) - \int_{\mathbb{R}} U_2(t,z) \tilde{N}_2(\mathrm{d}t, \mathrm{d}z)\,, \\ \tilde{V}(T) = \tilde{\xi}\,, \end{cases} \tag{3.61}$$

where $\Gamma_i(t) = \int_{\mathbb{R}} U_i(t,z) \gamma_i(t,z) \ell_i(\mathrm{d}z)$ $(i = 1, 2)$ and $\tilde{\xi} \in L^2_T$ is $\mathcal{F}_T$-measurable. Under the standard assumptions on $f : [0,T] \times \mathbb{R}^4 \to \mathbb{R}$, Tang and Li (1994) prove that the

BSDE (3.61) admits a unique solution $(\tilde{V}, Y, Z, U_1, U_2) \in S^2_{[0,T]} \times H^2_{[0,T]} \times H^2_{[0,T]} \times \hat{H}^2_{[0,T]} \times \hat{H}^2_{[0,T]}$. As we did in Section 3.3, we investigate the relation of the BSDE (3.59) to the classical BSDE (3.61) in the following proposition.

Proposition 3.21. *Let $(\tilde{V}, \tilde{\Upsilon}_1, \tilde{\Upsilon}_2, L^{FS})$ be the solution of the BSDE (3.59). Under the Assumptions 3.9, the BSDE (3.59) with L^{FS} given by (3.60) can be rewritten as the classical BSDE (3.61) where*

$$Y(t) = \tilde{\Upsilon}_1(t)b_1(t) + \tilde{\Upsilon}_2(t)b_2(t)\rho + P(t)\,, \quad Z(t) = \tilde{\Upsilon}_2(t)b_2(t)\sqrt{1-\rho^2} + Q(t)\,,$$
$$U_1(t,z) = \tilde{\Upsilon}_1(t)\gamma_1(t,z) + R_1(t,z)\,, \quad U_2(t,z) = \tilde{\Upsilon}_2(t)\gamma_2(t,z) + R_2(t,z)$$

and

$$\begin{aligned} &f(t,y,z,u_1,u_2) \\ &= -\left\{ a_1(t)\frac{\kappa_2(t)b_1(t) - \kappa_{12}(t)b_2(t)\rho}{\overline{\kappa}(t)} + a_2(t)\frac{\kappa_1(t)b_2(t)\rho - \kappa_{12}(t)b_1(t)}{\overline{\kappa}(t)} \right\} y \\ &\quad + \left\{ a_1(t)\frac{\kappa_{12}(t)b_2(t)\sqrt{1-\rho^2}}{\overline{\kappa}(t)} - a_2(t)\frac{\kappa_1(t)b_2(t)\sqrt{1-\rho^2}}{\overline{\kappa}(t)} \right\} z \\ &\quad - \frac{a_1(t)\kappa_2(t) - a_2(t)\kappa_{12}(t)}{\overline{\kappa}(t)}u_1 - \frac{a_2(t)\kappa_1(t) - a_1(t)\kappa_{12}(t)}{\overline{\kappa}(t)}u_2\,. \end{aligned}$$

Proof. Recall κ_1, κ_2, and κ_{12} in (3.55). The fact that $[L^{FS}, M_1]$ and $[L^{FS}, M_2]$ are $\mathbb{P}$-martingales yields

$$\begin{cases} \kappa_1(t)\tilde{\Upsilon}_1(t) + \kappa_{12}(t)\tilde{\Upsilon}_2(t) = b_1(t)Y(t) + \displaystyle\int_{\mathbb{R}} \gamma_1(t,z)U_1(t,z)\,\ell_1(\mathrm{d}z)\,, \\ \kappa_{12}(t)\tilde{\Upsilon}_1(t) + \kappa_2(t)\tilde{\Upsilon}_2(t) = b_2(t)\rho Y(t) + b_2(t)\sqrt{1-\rho^2}Z(t) \\ \qquad\qquad + \displaystyle\int_{\mathbb{R}} \gamma_2(t,z)U_2(t,z)\,\ell_2(\mathrm{d}z)\,. \end{cases} \tag{3.62}$$

Recall $\overline{\kappa}(t)$ in (3.56). Solving (3.62) for $\tilde{\Upsilon}_1(t)$ and $\tilde{\Upsilon}_2(t)$, we get

$$\begin{aligned} \tilde{\Upsilon}_1(t) = &\frac{\kappa_2(t)b_1(t) - \kappa_{12}(t)b_2(t)\rho}{\overline{\kappa}(t)}Y(t) - \frac{\kappa_{12}(t)b_2(t)\sqrt{1-\rho^2}}{\overline{\kappa}(t)}Z(t) \\ &+ \frac{\kappa_2(t)}{\overline{\kappa}(t)}\int_{\mathbb{R}} \gamma_1(t,z)U_1(t,z)\,\ell_1(\mathrm{d}z) - \frac{\kappa_{12}(t)}{\overline{\kappa}(t)}\int_{\mathbb{R}} \gamma_2(t,z)U_2(t,z)\,\ell_2(\mathrm{d}z)\,, \end{aligned} \tag{3.63}$$

$$\tilde{\Upsilon}_2(t) = \frac{\kappa_1(t)b_2(t)\rho - \kappa_{12}(t)b_1(t)}{\overline{\kappa}(t)} Y(t) + \frac{\kappa_1(t)b_2(t)\sqrt{1-\rho^2}}{\overline{\kappa}(t)} Z(t)$$
$$- \frac{\kappa_{12}(t)}{\overline{\kappa}(t)} \int_{\mathbb{R}} \gamma_1(t,z) U_1(t,z)\, \ell_1(\mathrm{d}z) + \frac{\kappa_1(t)}{\overline{\kappa}(t)} \int_{\mathbb{R}} \gamma_2(t,z) U_2(t,z)\, \ell_2(\mathrm{d}z) \tag{3.64}$$

and noting that $f(t, Y(t), Z(t), \Gamma_1(t), \Gamma_2(t)) = -a_1(t)\tilde{\Upsilon}_1(t) - a_2(t)\tilde{\Upsilon}_2(t)$ the expression for f follows. The fact that this function f is Lipschitz is easily seen. In fact, by Assumptions 3.9, the functions a_i, b_i, and γ_i are uniformly bounded in $[0,T]$ and hence also the functions κ_i, κ_{12} , $i = 1, 2$. Therefore the lower bound (3.56) holds. □

The estimates for the value of the portfolio and the amounts of wealth invested in the risky assets are derived in the following theorem for options written in the two-dimensional price process S. These results are the analogues of those in Theorem 3.11.

Theorem 3.22. *Let $(\tilde{V}, \tilde{\Upsilon}_1, \tilde{\Upsilon}_2, L^{FS})$ be the solution of the BSDE (3.59). Under Assumptions 3.9 and the assumptions that h is Lipschitz and $h(0,0) = 0$, we have*

$$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} |\tilde{V}(t) - \tilde{V}(t_i)|^2\right] + \sum_{j=1}^{2} \mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |\tilde{\Upsilon}_j(s) - \tilde{\Upsilon}_j(t_i)|^2 \,\mathrm{d}s\right]$$
$$\leq \frac{C}{n},$$

where C is a positive constant.

Proof. The estimate for $\tilde{V}$ follows immediately from Theorem 2.1 in Bouchard and Elie [26] and Proposition 3.21 in this chaper. Here we focus on the estimate for $\tilde{\Upsilon}_1$ and $\tilde{\Upsilon}_2$. We proceed to prove the statement in the setting of options written in the two-dimensional price process S given in (3.53). To simplify the notation, we define $\Lambda(s) := (Y(s), Z(s), \Gamma_1(s), \Gamma_2(s))$, $s \in [0,T]$, $f_s(y,z,u_1,u_2) := f(s,y,z,u_1,u_2)$ and $\nabla f_t(y,z,u_1,u_2) := (\partial_y f, \partial_z f, \partial_{u_1} f, \partial_{u_2} f)$. Integrating (3.61) we get

$$\tilde{V}(t) = \frac{h(S^{(\cdot)}(T))}{S^{(0)}(T)} + \int_t^T f_s(\Lambda(s))\,\mathrm{d}s - \int_t^T Y(s)\,\mathrm{d}W^{(1)}(s) - \int_t^T Z(s)\,\mathrm{d}W^{(2)}(s)$$
$$- \int_t^T \int_{\mathbb{R}} U_1(s,z)\,\tilde{N}_1(\mathrm{d}s,\mathrm{d}z) - \int_t^T \int_{\mathbb{R}} U_2(s,z)\,\tilde{N}_2(\mathrm{d}s,\mathrm{d}z). \tag{3.65}$$

Following the same steps as in the proof of Theorem 3.11, we derive estimates for the processes Y and Z using the Malliavin derivative under an additional assumption that $h \in C_b^1(\mathbb{R}^2)$, and generalize these estimates to the setting that h is Lipschitz using an approximation argument. Let $D^{(1)}$ denote the Malliavin derivative operator w.r.t. the Brownian motion $W^{(1)}$. Then taking the Malliavin derivative $D^{(1)}$ on both sides of (3.65), we get

$$\begin{aligned} D_s^{(1)}\tilde{V}(t) = \frac{\nabla h(S^{(\cdot)}(T))}{S^{(0)}(T)} \cdot D_s^{(1)}(S_1(T), S_2(T)) + \int_t^T \nabla f_r(\Lambda(r)) \cdot D_s^{(1)}(\Lambda(r))\,\mathrm{d}r \\ - \int_t^T D_s^{(1)} Y(r)\,\mathrm{d}W^{(1)}(r) - \int_t^T D_s^{(1)} Z(r)\,\mathrm{d}W^{(2)}(r) \\ - \int_t^T \int_{\mathbb{R}} D_s^{(1)} U_1(r,z)\,\tilde{N}_1(\mathrm{d}r, \mathrm{d}z) - \int_t^T \int_{\mathbb{R}} D_s^{(1)} U_2(r,z)\,\tilde{N}_2(\mathrm{d}r, \mathrm{d}z), \end{aligned} \tag{3.66}$$

where h is assumed to be in $C_b^1(\mathbb{R}^2)$ at the moment. Note that $f_t \in C_b^1(\mathbb{R}^4)$ for $t \in [0,T]$ since the partial derivatives of f_t can be obtained in a similar way as in the proof of Theorem 3.11. Observe that the Malliavin derivative for the process S is given by (see Theorem 3 and Proposition 7 in Petrou [144] and Theorem 2.2.1 in Nualart [137] for more details)

$$\begin{cases} D_s^{(1)} S_1(T) = \dfrac{b_1(s) S_1(s)}{\partial_{s_1} S_1(s)} \cdot \partial_{s_1} S_1(T) I_{s \le T} = b_1(s) S_1(T) I_{s \le T}\,, \\ \\ D_s^{(1)} S_2(T) = \dfrac{b_2(s) \rho S_2(s)}{\partial_{s_2} S_2(s)} \cdot \partial_{s_2} S_2(T) I_{s \le T} = b_2(s) \rho S_2(T) I_{s \le T}\,. \end{cases}$$

Hence, (3.66) is equivalent to

$$\begin{aligned} & D_s^{(1)}\tilde{V}(t) - \frac{h_y(S^{(\cdot)}(T))}{S^{(0)}(T)} b_2(s) \rho S_2(T) I_{s \le T} \\ & = \frac{h_x(S^{(\cdot)}(T))}{S^{(0)}(T)} b_1(s) S_1(T) I_{s \le T} + \int_t^T \nabla f_r(\Lambda(r)) \cdot D_s^{(1)}(\Lambda(r))\,\mathrm{d}r \\ & \quad - \int_t^T D_s^{(1)} Y(r)\,\mathrm{d}W^{(1)}(r) - \int_t^T D_s^{(1)} Z(r)\,\mathrm{d}W^{(2)}(r) \end{aligned}$$

$$- \int_t^T \int_{\mathbb{R}} D_s^{(1)} U_1(r,z)\, \tilde{N}_1(\mathrm{d}r, \mathrm{d}z) - \int_t^T \int_{\mathbb{R}} D_s^{(1)} U_2(r,z)\, \tilde{N}_2(\mathrm{d}r, \mathrm{d}z)\,. \tag{3.67}$$

Moreover, considering the variation of $\tilde{V}$ w.r.t. s_1, we get

$$\begin{aligned}\partial_{s_1}\tilde{V}(t) = {} & \frac{h_x(S^{(\cdot)}(T))}{S^{(0)}(T)} \cdot \partial_{s_1} S_1(T) + \int_t^T \nabla f_r(\Lambda(r)) \cdot \partial_{s_1}(\Lambda(r))\,\mathrm{d}r \\ & - \int_t^T \partial_{s_1} Y(r)\,\mathrm{d}W^{(1)}(r) - \int_t^T \partial_{s_1} Z(r)\,\mathrm{d}W^{(2)}(r) \\ & - \int_t^T \int_{\mathbb{R}} \partial_{s_1} U_1(r,z)\, \tilde{N}_1(\mathrm{d}r, \mathrm{d}z) - \int_t^T \int_{\mathbb{R}} \partial_{s_1} U_2(r,z)\, \tilde{N}_2(\mathrm{d}r, \mathrm{d}z)\,. \end{aligned} \tag{3.68}$$

From (3.67) and (3.68), we deduce

$$D_s^{(1)}\tilde{V}(t) = \frac{h_y(S^{(\cdot)}(T))}{S^{(0)}(T)} b_2(s)\rho S_2(T) I_{s\le T} + s_1 b_1(s)\partial_{s_1}\tilde{V}(t) I_{s\le T}\,. \tag{3.69}$$

Noticing that $D_t^{(1)}\tilde{V}(t) = Y(t)$, (3.69) leads for $t \in [0,T]$ to

$$\begin{aligned} Y(t) = {} & \frac{h_y(S^{(\cdot)}(T))}{S^{(0)}(T)} b_2(t)\rho S_2(T) I_{t\le T} + s_1 b_1(t)\partial_{s_1}\tilde{V}(t) I_{t\le T} \\ = {} & \frac{h_y(S^{(\cdot)}(T))}{S^{(0)}(T)} b_2(t)\rho S_2(T) + s_1 b_1(t) \left[\frac{h_x(S^{(\cdot)}(T))}{S^{(0)}(T)} \cdot \partial_{s_1} S_1(T) \right. \\ & + \int_t^T \nabla f_r(\Lambda(r)) \cdot \partial_{s_1}(\Lambda(r))\,\mathrm{d}r - \int_t^T \partial_{s_1} Y(r)\,\mathrm{d}W^{(1)}(r) - \int_t^T \partial_{s_1} Z(r)\,\mathrm{d}W^{(2)}(r) \\ & \left. - \int_t^T \int_{\mathbb{R}} \partial_{s_1} U_1(r,z)\, \tilde{N}_1(\mathrm{d}r, \mathrm{d}z) - \int_t^T \int_{\mathbb{R}} \partial_{s_1} U_2(r,z)\, \tilde{N}_2(\mathrm{d}r, \mathrm{d}z) \right]. \end{aligned}$$

Then following the same arguments as in the proof of Proposition 4.5 in Bouchard and Elie [26], we get

$$\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |Y(s) - Y(t_i)|^2\,\mathrm{d}s \right] \le \frac{C}{n}. \tag{3.70}$$

Considering the Malliavin derivative w.r.t. the Brownian motion $W^{(2)}$, we get by the same type of argument the following representation for the process Z for $t \in [0,T]$

$$Z(t) = s_2 b_2(t)\sqrt{1-\rho^2}\left[\frac{h_y(S^{(\cdot)}(T))}{S^{(0)}(T)} \cdot \partial_{s_2} S_2(T) + \int_t^T \nabla f_r(\Lambda(r)) \cdot \partial_{s_2}(\Lambda(r))\,\mathrm{d}r \right.$$
$$- \int_t^T \partial_{s_2} Y(r)\,\mathrm{d}W^{(1)}(r) - \int_t^T \partial_{s_2} Z(r)\,\mathrm{d}W^{(2)}(r)$$
$$\left. - \int_t^T \int_{\mathbb{R}} \partial_{s_2} U_1(r,z)\,\tilde{N}_1(\mathrm{d}r,\mathrm{d}z) - \int_t^T \int_{\mathbb{R}} \partial_{s_2} U_2(r,z)\,\tilde{N}_2(\mathrm{d}r,\mathrm{d}z)\right].$$

Thus the estimate for the process Z follows the same arguments as in Bouchard and Elie [26] and we get

$$\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |Z(s) - Z(t_i)|^2\,\mathrm{d}s\right] \leq \frac{C}{n}. \tag{3.71}$$

By relating the BSDE (3.61) to partial integro-differential equations as in Bouchard and Elie [26] we show the following estimates for U_1 and U_2

$$\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}} |U_j(s,z) - U_j(t_i,z)|^2\,\ell(\mathrm{d}z)\,\mathrm{d}s\right] \leq \frac{C}{n}, \quad j = 1,2. \tag{3.72}$$

By the same approximation arguments used in the proof of Theorem 3.3, we get that the estimates (3.70), (3.71) and (6.5) for Y, Z and U_j $(j = 1,2)$ hold when h is Lipschitz.

Using the expressions for $\tilde{\Upsilon}_1$ and $\tilde{\Upsilon}_2$ in terms of Y, Z, U_1, and U_2 (see (3.63) and (3.64)), we obtain the statement following the same lines of proof as in Theorem 2.3 in Khedher and Vanmaele [115]. □

3.4.2 Discrete-time model

In the sequel, we consider a time-discretisation of the two-dimensional price process S and we compute the associated LRM strategy.

Time-discretisation of the forward equation. The discrete time version of the process S defined in (3.53) is given for $i = 0,\ldots,n-1$ by

$$\begin{cases} \hat{S}_j(t_{i+1}) := \hat{S}_j(t_i) + \hat{S}_j(t_i)\left(a_j(t_i)\,\Delta t_i + \Delta\hat{\zeta}_j(t_i)\right), \quad j = 1,2, \\ \hat{S}_j(0) = S_j(0), \quad j = 1,2, \end{cases} \tag{3.73}$$

where for $j = 1, 2$,

$$\Delta\hat{\zeta}_j(t_{i+1}) := \int_{t_i}^{t_{i+1}} b_j(t_i)\,\mathrm{d}W_j(t) + \int_{t_i}^{t_{i+1}}\int_{\mathbb{R}} g(z)\,\tilde{\gamma}_j(t_i)\,\tilde{N}_j(\mathrm{d}t, \mathrm{d}z)\,,$$

for a_j, b_j, and $\tilde{\gamma}_j$ as in (3.53)-(3.54) and to which we impose Assumptions 3.9. The discrete-time version of the predictable process λ in (3.57) is denoted by $\hat{\lambda}$ and is given by

$$(\hat{\lambda}_1(t_{i+1}), \hat{\lambda}_2(t_{i+1})) = \left(\tfrac{\kappa_2(t_i)a_1(t_i) - a_2(t_i)\kappa_{12}(t_i)}{\overline{\kappa}(t_i)\hat{S}_1(t_i)}\,,\, \tfrac{\kappa_1(t_i)a_2(t_i) - a_1(t_i)\kappa_{12}(t_i)}{\overline{\kappa}(t_i)\hat{S}_2(t_i)}\right).$$

The discrete-time specification of the MVT process K defined in (3.58) is given by

$$\hat{K}(t_{i+1}) = \sum_{0\le k\le i} \frac{\kappa_1(t_k)a_2^2(t_k) + \kappa_2(t_k)a_1^2(t_k) - 2a_1(t_k)a_2(t_k)\kappa_{12}(t_k)}{\overline{\kappa}(t_k)}\,\Delta t_{k+1},$$

and due to the assumptions on the model parameters is uniformly bounded in t_i by a constant.

Time-discretisation of the backward equation. Considering a discretisation of the BSDE (3.59), we get

$$\begin{cases} \hat{V}(t_i) & = \hat{V}(t_{i+1}) - \left(\hat{\Upsilon}_1(t_i)a_1(t_i) + \hat{\Upsilon}_2(t_i)a_2(t_i)\right)\Delta t_{i+1} \\ & \quad -\hat{\Upsilon}_1(t_i)\Delta\hat{\zeta}_1(t_{i+1}) - \hat{\Upsilon}_2(t_i)\Delta\hat{\zeta}_2(t_{i+1}) - \Delta\hat{L}^{FS}(t_{i+1})\,, \\ \hat{V}(T) & = \hat{\xi}\,, \end{cases} \tag{3.74}$$

where $\hat{\xi} = h(\hat{S}^{(\cdot)}(T))/\hat{S}^{(0)}(T)$, $\hat{\Upsilon}_j(t_i) = \hat{\chi}_j^{FS}(t_i)\hat{S}_j(t_i)$, $j = 1, 2$, is the amount of wealth invested in the asset $\hat{S}_j$, $j = 1, 2$, in an LRM strategy.

Backward iteration scheme. Denote

$$\varsigma_j(t_i) := \mathbb{E}[\hat{V}(t_{i+1})\Delta\hat{\zeta}_j(t_{i+1}) \mid \mathcal{F}_{t_i}]\,, \quad j = 1, 2\,.$$

To obtain a backward iteration scheme for (3.74), we first take the expectation conditionally on $\mathcal{F}_{t_i}$ on both sides in (3.74) to arrive at the expression for $\hat{V}(t_i)$. Then for $j = 1, 2$ we multiply both sides in (3.74) by $\Delta\hat{\zeta}_j(t_{i+1})$ and take conditional expectation w.r.t. $\mathcal{F}_{t_i}$ to find the system for $i = n-1, \ldots, 0$

$$\begin{cases} \varsigma_1(t_i) = \left(\hat{\Upsilon}_1(t_i)\kappa_1(t_i) + \hat{\Upsilon}_2(t_i)\kappa_{12}(t_i)\right)\Delta t_{i+1}\,, \\ \varsigma_2(t_i) = \left(\hat{\Upsilon}_2(t_i)\kappa_2(t_i) + \hat{\Upsilon}_1(t_i)\kappa_{12}(t_i)\right)\Delta t_{i+1}\,, \\ \hat{V}(t_i) = \mathbb{E}[\hat{V}(t_{i+1}) \mid \mathcal{F}_{t_i}] - \left(a_1(t_i)\hat{\Upsilon}_1(t_i) + a_2(t_i)\hat{\Upsilon}_2(t_i)\right)\Delta t_{i+1}\,. \end{cases}$$

Solving the latter for $\tilde{\Upsilon}_1$ and $\tilde{\Upsilon}_2$, we get for $i = n-1, \dots, 0$

$$\begin{cases} \hat{\Upsilon}_1(t_i) = \dfrac{n}{T\overline{\kappa}(t_i)} \left(\kappa_2(t_i)\varsigma_1(t_i) - \kappa_{12}(t_i)\varsigma_2(t_i)\right) , \\ \hat{\Upsilon}_2(t_i) = \dfrac{n}{T\overline{\kappa}(t_i)} \left(\kappa_1(t_i)\varsigma_2(t_i) - \kappa_{12}(t_i)\varsigma_1(t_i)\right) , \\ \hat{V}(t_i) = \mathbb{E}[\hat{V}(t_{i+1}) \mid \mathcal{F}_{t_i}] - \Big(a_1(t_i)\hat{\Upsilon}_1(t_i) + a_2(t_i)\hat{\Upsilon}_2(t_i)\Big)\, \Delta t_{i+1} . \end{cases} \tag{3.75}$$

As for $\hat{L}^{FS}$, being zero at zero, we have from (3.74)

$$\begin{aligned} \hat{L}^{FS}(T) = \hat{V}(T) - \sum_{i=0}^{n-1} \Big(\hat{\Upsilon}_1(t_i)a_1(t_i) + \hat{\Upsilon}_2(t_i)a_2(t_i)\Big)\, \Delta t_{i+1} \\ - \sum_{i=0}^{n-1} \hat{\Upsilon}_1(t_i)\Delta\hat{\zeta}_1(t_{i+1}) - \sum_{i=0}^{n-1} \hat{\Upsilon}_2(t_i)\Delta\hat{\zeta}_2(t_{i+1}) - \hat{V}(0) . \end{aligned}$$

3.4.3 L^2-convergence of the discretisation scheme

In the sequel we study the L^2-convergence of the LRM strategy related to the time-discrete model to the LRM strategy related to the continuous-time model where we consider the two-dimensional setting. Moreover, we compute rates for the convergence results.

The result of Theorem 3.12 is applied to the two-dimensional setting providing us with a convergence rate of the order of the time step for the approximation of S_j in (3.53) by $\hat{S}_j$ in (3.73) for $j = 1, 2$.

The analogue of Proposition 3.13 in the present setting is

Proposition 3.23. *Let $\tilde{V}$ and $\hat{V}$ be solutions of (3.59) and (3.74), with a payoff function h being Lipschitz and $h(0,0) = 0$. Under Assumptions 3.9, we have*

$$\mathbb{E}\left[|\tilde{V}(T) - \hat{V}(T)|^2\right] = \mathbb{E}\left[|\tilde{\xi} - \hat{\xi}|^2\right] \leq \frac{C}{n},$$

for a positive constant C independent of the number of steps.

Proof. Following the lines of the proof of Proposition 3.13 we easily arrive at

$$\mathbb{E}[|\tilde{\xi} - \hat{\xi}|^2] = \mathbb{E}\left[\left|\frac{h(S^{(\cdot)}(T))}{S^{(0)}(T)} - \frac{h(\hat{S}^{(\cdot)}(T))}{\hat{S}^{(0)}(T)}\right|^2\right]$$

$$\leq C \left\{ \sum_{j=1}^{2} \mathbb{E}\left[\left| S^{(j)}(T) \right|^2 \right] \mathbb{E}\left[\left| \hat{S}^{(0)}(T) - S^{(0)}(T) \right|^2 \right] + \sum_{j=1}^{2} \mathbb{E}\left[\left| S^{(j)}(T) - \hat{S}^{(j)}(T) \right|^2 \right] \mathbb{E}\left[\left| S^{(0)}(T) \right|^2 \right] \right\} .$$

The statement follows by using Lemma 3.2 in Benth et al. [21] and applying the analogue of Theorem 3.12 to $S^{(0)}$, $S^{(1)}$, and $S^{(2)}$. □

We adapt the results of Khedher and Vanmaele [115] to BSDEs driven by two correlated martingales. That is to a value process $\tilde{V}$ solving

$$\tilde{V}(t) = \tilde{V}(T) - \sum_{j=1}^{2} \int_t^T \tilde{\Upsilon}_j(s) \frac{a_j(s)}{\kappa_j(s)} \, \mathrm{d}\langle \zeta_j \rangle_s - \sum_{j=1}^{2} \int_t^T \tilde{\Upsilon}_j(s) \, \mathrm{d}\zeta_j(s) - L^{FS}(T) + L^{FS}(t) .$$

In order to extend the corresponding proofs in Khedher and Vanmaele [115] to our setting, we need to show that the coefficients in the expressions (3.63)-(3.64) of $\tilde{\Upsilon}_j$, $j = 1, 2$ are Lipschitz. Under Assumptions 3.9 the functions b_j, $j = 1, 2$, are Lipschitz and (3.56) holds, hence it is sufficient to show that κ_j, $j = 1, 2$ and κ_{12} are Lipschitz. The latter also immediately follows from the Assumptions 3.9 for b_j and γ_j, $j = 1, 2$.

Theorem 3.22, the analogue of Theorem 3.12, and Proposition 3.23 guarantee that the conditions of Theorem 4.4 in Khedher and Vanmaele [115] adapted to our setting hold. This statement can be confirmed by following the same lines as the proof of Theorem 3.16. Hence, the corresponding robustness results for the LRM strategy can be applied in the current setting. The following robustness results are applications of Theorem 4.4, Theorem 4.5, and Theorem 4.6 in Khedher and Vanmaele [115].

Theorem 3.24. *Let $\tilde{V}$ and $\hat{V}$ be respectively defined by (3.59) and (3.74). Under the Assumptions 3.9 on the parameters of the model and the assumptions that h is Lipschitz and $h(0, 0) = 0$, it holds that*

$$\max_{i<n} \mathbb{E}\left[\sup_{t \in [t_i, t_{i+1}]} |\tilde{V}(t) - \hat{V}(t_i)|^2 \right] \leq \frac{C}{n} ,$$

where C is a positive constant.

The robustness of the amount of wealth in an LRM strategy follows in the next theorem.

Theorem 3.25. *Let $\tilde{\Upsilon}$ and $\hat{\Upsilon}$ be the amount of wealth defined as in (3.59) and (3.74). Under the conditions of Theorem 3.24 it holds that*

$$\sum_{j=1}^{2} \mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |\tilde{\Upsilon}_j(s) - \hat{\Upsilon}_j(t_i)|^2 \, \mathrm{d}s\right] \leq \frac{C}{n},$$

where C is a positive constant.

We state the robustness of the remaining risk L^{FS} in the following theorem.

Theorem 3.26. *Let the processes L^{FS} and $\hat{L}^{FS}$ be as defined in (3.59) and (3.74). Under the conditions of Theorem 3.24, we have for all $0 \leq i \leq n-1$,*

$$\mathbb{E}\left[\left|L^{FS}(t_{i+1}) - \hat{L}^{FS}(t_{i+1})\right|^2\right] \leq \frac{C}{n},$$

where C is a positive constant.

Notice that this latter result implies the convergence of the cost process in an LRM strategy. Indeed the cost processes in an LRM strategy related to the price processes S and $\hat{S}$ are respectively given by $C(t) = L^{FS}(t) + V(0)$, $t \in [0,T]$, and $\hat{C}(t_i) = \hat{L}^{FS}(t_i) + \hat{V}(0)$, $i = 0, \ldots, n$. Then applying Theorems 3.24 and 3.26, we get for all $i = 0, \ldots, n$,

$$\mathbb{E}[|C(t_i) - \hat{C}(t_i)|^2] \leq C\mathbb{E}[|L^{FS}(t_i) - \hat{L}^{FS}(t_i)|^2] + \mathbb{E}[|V(0) - \hat{V}(0)|^2] \leq \frac{C}{n}.$$

3.5 Numerical examples

In this section, we consider a specification of the model class presented in Section 3.3. Algorithm 3.1 presented in the sequel gives an outline how the hedging strategy can be calculated in our model setup.

For the Euler-scheme, we consider the setup from Section 3.3.2. Recall the equidistant time grid π_n as in (3.19) and the discrete version of the stock price process as in (3.27). The simulation of the forward Euler-scheme is well-known and easy to implement. However, the calculation of the hedging strategy is not that straightforward because one has to deal with the backward scheme (3.31).

Algorithm for the LRM strategy. The crucial part in solving (3.31) is to calculate the conditional expectations appearing in the formula. This cannot be done analytically

and should therefore be estimated in a sensible manner. There are different methods in the literature dealing with the problem of numerical computation of conditional expectations. One approach is to use Monte Carlo as in Bally et al. [12] and Daveloose et al. [55] where a Malliavin approach and a conditional density method were considered to express the conditional expectations in terms of regular expectations without any conditioning. Another method is to use a regression-based ansatz as in Longstaff and Schwartz [123]. For a nice overview of various regression-based approaches to solve numerically conditional expectations, we refer to Kohler [118]. In this chapter we use this latter approach.

By applying such a regression-based method, one implicitly assumes a Markovian property for the conditional expectations. In our case it means, that the conditional expectations

$$\mathbb{E}\left[\hat{V}(t_{i+1})\left(b(t_i)\,\Delta W(t_{i+1}) + \int_{\mathbb{R}} \gamma(t_i,z)\tilde{N}\left((t_i,t_{i+1}],\mathrm{d}z\right)\right)\middle|\,\mathcal{F}_{t_i}\right]$$

and

$$\mathbb{E}\left[\hat{V}(t_{i+1})\,|\mathcal{F}_{t_i}\right]$$

are estimated by functions of $\hat{S}(t_i)$, namely

$$\alpha_i(\hat{S}(t_i)) := \mathbb{E}\left[\hat{V}(t_{i+1})\left(b(t_i)\,\Delta W(t_{i+1}) + \int_{\mathbb{R}} \gamma(t_i,z)\tilde{N}\left((t_i,t_{i+1}],\mathrm{d}z\right)\right)\middle|\,\hat{S}(t_i)\right]$$

and

$$\beta_i(\hat{S}(t_i)) := \mathbb{E}\left[\hat{V}(t_{i+1})\middle|\,\hat{S}(t_i)\right].$$

In Algorithm 3.1, we choose α_i and β_i to be piecewise linear by performing more than only one regression in each time step.

Algorithm 3.1 (Calculation of the LRM strategy)**.**
Let m denote the number of Monte-Carlo-paths and let $w \in \mathbb{N}$.

1. *Draw $m \times n$ normal distributed random variables representing the increments of the Brownian motion and $m \times n$ random variables from a distribution corresponding to the increments of the jump process.*

2. *Derive m paths of the asset price process $\hat{S}$ according to equation (3.27).*
3. *Calculate for each Monte-Carlo-path the expressions $\hat{V}(t_n) = \hat{\xi}$, where $\hat{\xi}$ represents the payoff at time T, as defined in Section 3.3.2.*
4. *For i in $\{n-1, n-2, \dots, 1\}$ do:*
   (a) *Compute for each path $\hat{V}'(t_{i+1}) := \hat{V}(t_{i+1})\Delta\hat{\zeta}(t_{i+1})$.*
   (b) *Sort the paths by the realisations of $\hat{S}(t_i)$ in an ascending order and obtain a sorted list of path $s_1, s_2, \dots, s_m$.*
   (c) *For j in $\{0, 1, \dots, w-1\}$ do:*
      i. *Select a subset of paths, namely $I := \left\{ s_{\frac{m}{w}j+1}, s_{\frac{m}{w}j+2}, \dots, s_{\frac{m}{w}(j+1)} \right\}$.*
      ii. *Perform a linear regression with explanatory data $\{s_i\}_{s\in I}$ and the $\hat{V}'(t_{i+1})$-values corresponding to these paths as endogenous data. This yields a linear function $\alpha_{ij}(x) = \alpha_{ij}^{(0)} + \alpha_{ij}^{(1)}x$.*
      iii. *Define for all $s \in I$ the value for $\hat{\Upsilon}(t_i)$ by $\frac{\alpha_{ij}(s(t_i))n}{T\kappa(t_i)}$.*
      iv. *Perform a linear regression with explanatory data $\{s(t_i)\}_{s\in I}$ and the $\hat{V}(t_{i+1})$-values corresponding to these paths as endogenous data. This yields a linear function $\beta_{ij}(x) = \beta_{ij}^{(0)} + \beta_{ij}^{(1)}x$.*
      v. *Define for all $s \in I$ the value for $\hat{V}(t_i)$ by $\beta_{ij}(s(t_i)) - \frac{T}{n}a(t_i)\hat{\Upsilon}(t_i)$.*
5. *Define $\hat{\Upsilon}(t_0)$ as the average over all paths of the following values:*

$$\frac{\hat{V}(t_1)}{\kappa(t_1)}\frac{n}{T}\Delta\hat{\zeta}(t_1)\,.$$

6. *Define $\hat{V}(t_0)$ as the average over all paths of the following values:*

$$\hat{V}(t_1) - \frac{T}{n}\hat{\Upsilon}(t_0)a(t_0).$$

7. *Compute for each Monte-Carlo-path the path of the hedging error $\hat{L}^{FS}$ following equation (3.32).*

Due to the division into subsets in step 4c of Algorithm 3.1, the approximating functions α and β are piecewise linear. Hence, the structure of α and β is flexible in the sense that it can approximate an arbitrary function quite good as long as w is large enough. An alternative approach can be created by widening the set of basis functions

for the regression, e.g. one could use polynomials of higher order than just one. Then, if one can guess how the functions α and β look like and the basis functions are chosen reasonably, one can omit the division of the paths. However, it is often difficult to choose a sensible set of basis functions, especially in case of the function α, which is the reason for considering a piecewise linear ansatz to achieve a general setup.

In the sequel, we investigate numerical examples to show the accuracy of Algorithm 3.1 and to consolidate the theoretical results in this chapter. In this context, we consider vanilla and spread options for our numerical examples. Notice that Algorithm 3.1 can be immediately applied to vanilla options. In the case of spread options, the conditional expectations appearing in Equation (3.75) are approximated by conditioning on $\hat{S}_1(t_i) - \hat{S}_2(t_i)$. Hence, to apply Algorithm 3.1 to spread options, one has to simulate two different asset price processes and the linear regressions in step 4c are performed with respect to the difference of these two prices. To assume the Markovianity in the case of Asian options, we need to condition on the current asset value $\hat{S}(t_i)$ and on the average value $\hat{S}_A(T)$ as defined in (3.29). Therefore, the regression based method is not suited for Asian options, since a regression with two-dimensional explanatory data becomes numerically involved. Thus, other numerical methods for the simulation of the conditional expectations are called for. One for example can use Monte Carlo methods as in Daveloose et al. [63], where a Malliavin derivative approach is used to express the conditional expectations in terms of regular expectations without any conditioning. We choose not to present this approach in this chapter and focus on the regression based method. Next we present the stock price model that we consider for our numerical examples.

Model setup. We assume that the dynamics of the discounted stock price process S is given by a two-dimensional Merton model. That is S is as in (3.53), where $\zeta = (\zeta_1, \zeta_2)$ is given by

$$d\zeta_i(t) = b_i(t)\,\mathrm{d}W_i(t) + \mathrm{d}\left(\sum_{j=1}^{N_i(t)} J_i(j) - \frac{c_i t}{\eta_i}\right), \quad i = 1, 2\,. \tag{3.76}$$

We impose Assumptions 3.9 on the parameters of the model. Moreover, we assume that the Brownian motions W_1 and W_2 are correlated with correlation coefficient ρ, N_i, $i = 1, 2$, is a Poisson process with intensity c_i, and $\{J_i(j)\}_{j\in\mathbb{N}}$, $i = 1, 2$, is a sequence of independent exponentially distributed random variables with parameter η_i. The Poisson processes N_1 and N_2 are independent.

3.5.1 Vanilla option

In this subsection we compute the LRM strategy for hedging vanilla options. For this purpose, we consider a European call option with strike $K = 1.15$ and maturity $T = 0.5$ year, i.e.,

$$\xi = (S(0.5) - 1.15)^+ ,$$

where S is the one-dimensional version of (3.53) with ζ being as in (3.76). For simplicity, we assume the interest rate $r = 0$ and we consider the following parameter setting

$$\begin{aligned} a(t) &= 0.04, & w &= 20, & c &= 5, \\ b(t) &= 0.3, & \eta &= 40, & S(0) &= 1. \end{aligned}$$

To test the accuracy of Algorithm 3.1 which results from the discretisation scheme (3.31), we compare it to the Fourier approach as suggested in Hubalek et al. [108]. Thus we first present the expressions for the amount of wealth and the value of the portfolio using the Fourier approach. Then we perform our comparison study.

3.5.1.1 Fourier approach

Let S be as described above. Then according to Theorem 5.1.6 in Applebaum [4], there exists a Lévy process L such that $S(t) = S(0)\exp(L(t))$ and where the cumulant function of $L(1)$ is given by

$$\begin{aligned} \varrho(u) &= \log \mathbb{E}(e^{uL_1}) \\ &= u(a - \frac{c}{\eta} - \frac{1}{2}b^2) + \frac{1}{2}b^2u^2 - c + c\eta \int_0^\infty (1+y)^u e^{-\eta y}\, \mathrm{d}y. \end{aligned}$$

By Proposition 3.1 in Hubalek et al. [108] we know that the number of assets and the value of the hedging portfolio can be expressed as follows

$$\begin{aligned} \chi^{FS}(t) &= \frac{S^{R-1}(t-)K^{1-R}}{2\pi} \int_{\mathbb{R}} e^{\bar{\eta}(R-iu)(T-t)} \left(\frac{K}{S(t-)}\right)^{iu} \frac{\bar{\mu}(R-iu)}{(R-iu)(R-iu-1)}\, \mathrm{d}u, \\ \tilde{V}(t) = V(t) &= \frac{K^{1-R}S^R(t-)}{2\pi} \int_{\mathbb{R}} \left(\frac{K}{S(t-)}\right)^{iu} \frac{e^{\bar{\eta}(R-iu)(T-t)}}{(R-iu)(R-iu-1)}\, \mathrm{d}u, \end{aligned} \tag{3.77}$$

where $R > 1$ is a damping factor, K is the strike of a vanilla call option, and

$$\bar{\mu}(R - iu) := \frac{\varrho(R - iu + 1) - \varrho(R - iu) - \varrho(1)}{\varrho(2) - 2\varrho(1)},$$
$$\bar{\eta}(R - iu) := \varrho(R - iu) - \mu(R - iu)\varrho(1).$$

As one can observe from the formulas (3.77), the idea in the Fourier approach is to exploit the cumulant function of the driving process L which is known for a big class of Lévy processes. This explains our choice for constant parameters a and b. This is not necessary for the applicability of the BSDE approach. Later on when we apply our discretisation scheme to spread options we choose linear functions in t for the parameters of the model.

3.5.1.2 Accuracy of the BSDE method

We compare $\chi^{FS}(t_i)S(t_i-)$, $V(t_i-)$ calculated by the Fourier transform techniques (3.77), with $\hat{\Upsilon}(t_i)$, $\hat{V}(t_i)$ calculated by the BSDE technique (3.31), respectively. The integrals in (3.77) are calculated with the T-COS method as introduced in Sun et al. [162], which is a FFT-based method combined with the approach of Fang and Oosterlee [76] based on the Fourier-cosine expansions.

The integral interval $\mathbb{R}$ is truncated into $[-40, 40]$ and the integrand is approximated by a truncation of its Fourier cosine expansion with 2^{11} terms. We refer to Sun et al. [162] for the details on the implementation. Fifty thousands $(m = 50\,000)$ paths are used in the Algorithm 3.1 and the stepsize is set to be 0.05. The FFT-based method totally costs about 448 s while the Algorithm 3.1 totally costs 36 s. The implementations are done in MATLAB2014b (processor: Intel i5 2.5GHz, RAM:4GB). Unless the parallelization computation technique is used in the FFT-based algorithm, it is more efficient to use the Algorithm 3.1 to calculate the sample average of the four errors. However, if one only considers a specific path of the value process, the wealth process or the hedging error, then it is obviously efficient to use the FFT-based scheme as it only needs a specific path of the asset price. In the sequence, whenever the sample average is involved, we use Algorithm 3.1.

The comparison over all the simulated paths is summarized in Table 3.1. Since the T-COS method is an accurate numerical method to calculate an integral with a smooth integrand, this comparison confirms that the numerical results with the BSDE approach are consistent with the numerical results using the Fourier transform approach.

terms	value
$\max_{i=0,\dots,n} \mathbb{E}\left[\left(\tilde{V}(t_i) - \hat{V}(t_i)\right)^2\right]$	2.8457e-06
$\max_{i=0,\dots,n} \mathbb{E}\left[\left(\tilde{\Upsilon}(t_i) - \hat{\Upsilon}(t_i)\right)^2\right]$	5.4924e-04
$\max_{i=0,\dots,n} \mathbb{E}\left[\left(L^{FS}(t_i) - \hat{L}^{FS}(t_i)\right)^2\right]$	1.6712e-05

Table 3.1: Overview of comparison between the Fourier transform approach and the BSDE approach. The stepsize is set to be 0.05.

3.5.1.3 Convergence study

To consolidate the convergence results in Section 3.3.3, we have to do another numerical test. We treat the numerical results for a large n as the solution of the continuous BSDE-problem[1]. Indeed, we choose a daily step size to be seen as the exact solution, which means $n = 128$. Then, we compare this solution to numerical results coming from bigger step sizes, i.e. $n \in \{1, 2, 4, 8, 16, 32, 64\}$. Table 3.2 gives an overview of the approximation errors for the stock price, the value process, the amount of wealth process and the hedging error. These results are depicted in Figure 3.1 (left), from which one can see that the four errors are linear functions of the step-size. The slope of these linear functions is approximately one. This confirms our theoretical statements in Section 3.3 that the L^2-convergence of the proposed numerical method is of the order of the time step.

3.5.2 Spread option

In this subsection, we aim at hedging a call spread option with strike $K = 0.15$ and maturity $T = 0.5$ year, i.e

$$\xi = (S_1(0.5) - S_2(0.5) - 0.15)^+ .$$

The parameter setting is given by

$$a_1(t) = 0.02 + 0.01\,t, \qquad b_1(t) = 0.04 - 0.01\,t, \quad c_1 = 3, \qquad \eta_1 = 30,$$

[1]Due to the aforementioned lower efficiency of the FFT-based scheme in the setting of calculating the sample average, we did not use this scheme here.

number of steps	1	2	4	8	16	32	64
$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]}\left(S(t)-\hat{S}(t_i)\right)^2\right]$	512	259	129	62.5	29.2	14.1	6.51
$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]}\left(\tilde{V}(t)-\hat{V}(t_i)\right)^2\right]$	113	73.9	41.6	21.7	10.6	4.70	2.11
$\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\left(\tilde{\Upsilon}(s)-\hat{\Upsilon}(t_i)\right)^2 \mathrm{d}s\right]$	608	326	173	89.2	45.0	21.5	9.81
$\mathbb{E}\left[\left(L^{FS}(T)-\hat{L}^{FS}(T)\right)^2\right]$	38.6	21.1	11.9	6.31	3.25	1.59	0.78

Table 3.2: Overview of the errors for a vanilla call depending on the number of steps. All errors are given in basis points, i.e. a percent of a percent.

$$a_2(t) = 0.04 + 0.005\,t, \qquad b_2(t) = 0.03 - 0.05\,t, \qquad c_2 = 5, \qquad \eta_2 = 40,$$
$$T = 0.5 \qquad S_1(0) = S_2(0) = 1 \qquad w = 20, \qquad m = 1\,000\,000.$$

We consider a similar robustness study as for the vanilla option. Thus we choose a daily step size to be seen as the exact solution and we compare this solution to numerical results coming from bigger step sizes. The results are summarized in Table 3.3 and depicted in Figure 3.1(right). From Figure 3.1(right), we can see that for small step sizes the discretisation error related to the LRM strategy for the call spread option is a linear function of the time stepsize. Again this confirms an L^2-convergence of the order of the time step as claimed in our theoretical statements in Section 3.3.

3.6 Conclusion

In this chapter we studied the L^2-convergence of the discrete-time LRM strategy to the continuous-time version using BSDEs driven by càdlàg martingales. We obtained a convergence rate of the order of the time step for the value of the portfolio, the amount of wealth invested in the risky asset, and the risk process in a LRM strategy for vanilla, Asian, and spread options.

The proposed discretisation scheme provides an implementable numerical method to simulate the LRM strategies for the mentioned options. We presented two numerical examples. In these examples we considered LRM strategies related to vanilla and spread

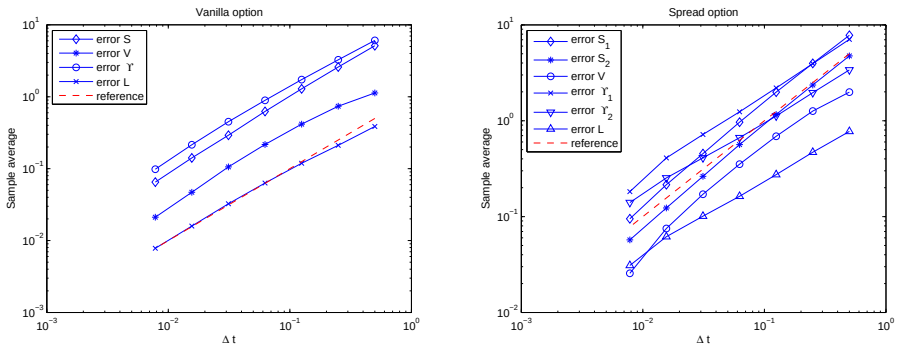


Figure 3.1: Convergence order for the asset price, the value process, the amount of wealth process and the hedging error for a vanilla call option (left) and a call spread option (right). The slop of the dashed red line is 1.

options written in a jump-diffusion model. Our theoretical results were consolidated with the numerical examples.

number of steps	1	2	4	8	16	32	64
$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} \left(S_1(t) - \hat{S}_1(t_i)\right)^2\right]$	778	397	199	96.6	45.4	21.5	9.49
$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} \left(S_2(t) - \hat{S}_2(t_i)\right)^2\right]$	474	236	116	56.4	26.3	12.3	5.71
$\max_{i<n} \mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]} \left(\tilde{V}(t_i) - \hat{V}(t_i)\right)^2\right]$	199	126	68.8	35.2	17.1	7.52	2.56
$\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \left(\tilde{\Upsilon}_1(s) - \hat{\Upsilon}_1(t_i)\right)^2 \mathrm{d}s\right]$	706	394	220	124	71.8	40.9	18.2
$\mathbb{E}\left[\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \left(\tilde{\Upsilon}_2(s) - \hat{\Upsilon}_2(t_i)\right)^2 \mathrm{d}s\right]$	339	196	112	66.5	41.0	25.3	14.0
$\mathbb{E}\left[\left(L^{FS}(T) - \hat{L}^{FS}(T)\right)^2\right]$	77.4	46.9	27.4	16.3	10.1	6.17	3.09

Table 3.3: Overview of the errors for a call spread depending on the number of steps. All errors are given in basis points, i.e. a percent of a percent.

4

Uncertainty quantification of derivative instruments

In this chapter, we focus on model uncertainty resulted from model calibration. Model uncertainty is characterized by a finite set of plausible models of different types, while parameter uncertainty is characterized by specifying each parameter in an interval. Each parameter can take any value from its value interval. Cont [49] proposed to quantify the uncertainty embedded in a derivative value in terms of its worst-case values. From this point of view, we only have to consider a special type of these models. Otherwise, we can calculate the worst-case values under the models of each type and find the bounds of these worst-case values. Hence, without loss of generality, we focus on parameter uncertainty in this chapter. In this setting, the key issue is how to efficiently calculate a large ensemble of the derivative value associated with a given large ensemble of the model parameters.

Combining the Monte Carlo method and the Smolyak interpolation algorithm, this chapter proposes an accurate and efficient numerical method to calculate a derivative value for any given realisation of the model parameters. The proposed method can be used to calculate the coherent uncertainty measure [49] and the entropy of the derivative value. It can also be used to accelerate the procedure of robust calibration (see e.g.

[13, 66, 94]).

4.1 Introduction

In the context of pricing and hedging of exotic, over-the-counter (OTC) derivatives, the price process of the underlying asset can be modeled by a set of parametric models of different types, such as local volatility models, stochastic volatility models, and Lévy models. The parameters of these models can be estimated by model calibration. However, there are two issues with model calibration: first, different kinds of models can be perfectly calibrated to the same market data [151]; second, different calibration methods may yield different estimations for each parameter of a specific model [91]. These empirical facts, together with a limited knowledge of the market dynamics, confront an agent with ambiguity about which model is the best one to value a target derivative, especially when the well-calibrated models may lead to quite different results. Model and parameter ambiguities are ubiquitous *whenever* a parametric model is employed. These ambiguities include two different categories according to Knight [117]: risk and uncertainty. Model (parameter) risk relates to the setting in which the probabilities of the candidate models (parameter realisations) are known, while model (parameter) uncertainty arises from a lack of knowledge of the probabilistic information on these models (parameter realisations). This chapter focuses on the model (parameter) uncertainty, and the aim is to numerically quantify its impact on a derivative value.

Uncertainty quantification is a growing concern from the point of view of regulation and risk management. According to the regulations set by the Federal Reserve Board [78] and the recent updates of the Basel II market risk framework [23], it is compulsory to assess and quantify the uncertainty embedded in a target derivative instrument. This procedure could provide financial institutions with the information to assess the need for valuation adjustments and mitigate the risk associated with their trading activities. Essentially, this information is very important for keeping agents informed when they bargain over the price of a derivative in the OTC market. Since the final price is based on bargaining (see e.g. [67]), we use the term 'value' rather than 'price' of a derivative in this chapter.

To account for model uncertainty, it is robust to consider a finite set of plausible models. At the parameter level, it is difficult to differentiate the well-calibrated models, under which the model prices of benchmarks lie between the corresponding bid-ask prices quoted in liquid markets. Hence, the confidence interval of each model parameter

can be employed for option pricing. To our knowledge, this idea was first proposed by [11] and [126]. In these settings, an agent may have no preference for the candidate models or the parameter realisations. As noted by the authors of [148], traders and institutions attack model uncertainty through a worst-case approach, for example, by stress testing of portfolios.

Following a worst-case approach, Cont [49] proposed an uncertainty measure, namely the spread between lower and upper bounds of the target derivative value when a set of plausible models are employed. This measure is not only consistent with the worst-case approach for option pricing, but also takes into account the hedgeable and unhedgeable risk. Although it satisfies a set of meaningful axioms, little attention has been paid on how to efficiently calculate it for exotic derivatives under complex parametric models. This chapter proposes an accurate and efficient numerical method to fill this gap.

The proposed method can be used to efficiently calculate a large ensemble of the derivative value with a given ensemble of the model parameters. The ensemble of the derivative value can be used to calculate the value bounds of the derivative value as well as the entropy, which is a complement to Cont's uncertainty measure.

4.2 Worst-case valuation and uncertainty quantification

This section is to introduce the worst-case value of a derivative instrument and the uncertainty measures, when a set of pricing models are employed to account for an agent's uncertainty about the pricing model.

4.2.1 Pricing models and uncertainty

Generally speaking, when the price process of an underlying asset is modeled by a stochastic system, this parametric model implicitly specifies a probability distribution on the future state of the underlying asset. Hence, a parametric model can be characterized by a probability measure $\mathbb{Q}_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta}$ represents the vector of model parameters. A pricing model corresponds to a risk-neutral measure $\mathbb{Q}_{\boldsymbol{\theta}}$, under which the discounted asset price $(S_t)_{t\in[0,T]}$ is a martingale with respect to its own history $(\mathcal{F}_t)_{t\in[0,T]}$.

The parameters $\boldsymbol{\theta}$ of a pricing model $\mathbb{Q}_{\boldsymbol{\theta}}$ can be estimated by calibrating the model to the prices of the benchmark instruments which are liquidly traded in markets. This procedure is the so-called model calibration. Denote the payoffs of benchmark instruments by $(f_i)_{i\in I}$ and their observed market prices by $(C_i)_{i\in I}$, where I denotes a

set of benchmarks. In most cases the market price of a benchmark instrument is quoted in terms of the bid-ask prices, that is, $(C_i^{\text{bid}}, C_i^{\text{ask}})$. When a parametric model $\mathbb{Q}_{\boldsymbol{\theta}}$ is given, the model price of a benchmark instrument is denoted by $P^{\mathbb{Q}_{\boldsymbol{\theta}}}(f_i)$. Model calibration is to estimate $\bar{\boldsymbol{\theta}}$ such that

$$C_i^{\text{bid}} \leq P^{\mathbb{Q}_{\bar{\boldsymbol{\theta}}}}(f_i) \leq C_i^{\text{ask}}, \quad \forall i \in I. \tag{4.1}$$

It is unrealistic to expect that there exists a unique solution to the calibration problem (4.1). In addition, different optimization methods may lead to different estimations for each parameter of a parametric model, even if an agent is to calibrate a parametric model to the middle prices of the bid-ask pairs $(C_i^{\text{bid}}, C_i^{\text{ask}})_{i\in I}$ [91]. To account for uncertainty about the model parameters, the agent could specify each model parameter in an interval, as long as the model prices of the benchmarks are consistent with $(C_i^{\text{bid}}, C_i^{\text{ask}})_{i\in I}$. These intervals could be the confidence intervals estimated with statistical methods from the market information (see e.g. [66, 74, 94]).

Let $\mathcal{Q} = \{\mathbb{Q}_{\boldsymbol{\theta}_1}, \ldots, \mathbb{Q}_{\boldsymbol{\theta}_J}\}$ be a finite set of plausible pricing models whose parameters are specified in terms of boxes, that is, $\boldsymbol{\theta}_j \in D_j$ for $j = 1, \ldots, J$. These boxes D_j account for parameter uncertainty while $\mathcal{Q}$ accounts for model uncertainty and parameter uncertainty. Since the size of $\mathcal{Q}$ is finite, we only have to consider one type of these models with parameters in a box. Otherwise, we can consider different types of these models one by one, and then integrate the results of uncertainty measures introduced in the next sections. Hence, we focus on parameter uncertainty in this chapter. Actually, if we interpret the models with different parameter values as different models, parameter uncertainty is also a kind of model uncertainty. Note that, in this setting, the elements in $\mathcal{Q}$ are not necessarily equivalent measures.

4.2.2 Worst-case valuation under parameter uncertainty

Intuitively, an uncertainty-aversion decision maker is concerned about what will happen in the worst and best cases when a set of candidate pricing models are available. If the difference of the results in two extreme cases is negligible, there is little uncertainty. Otherwise, the decision maker has to take uncertainty into account and make a robust decision. The idea of worst-case valuation is not only used in derivative pricing (see e.g. [11, 126]) but also in macroeconomics (see e.g. [96]) and other areas. We will introduce the worst-case approach and the numerical methods for the worst-case value in this subsection.

In fact, the worst case for the seller is the best case for the buyer, and vice versa. So, the worst-case approach consists in calculating the value bounds of a derivative. The upper and lower value bounds of a derivative instrument can be defined as

$$\overline{P}(f) = \sup_{\mathbb{Q}\in\mathcal{Q}} P^{\mathbb{Q}}(f) \quad \text{and} \quad \underline{P}(f) = \inf_{\mathbb{Q}\in\mathcal{Q}} P^{\mathbb{Q}}(f), \tag{4.2}$$

where f is the well-defined payoff function of the target derivative and $P^{\mathbb{Q}}$ is the value function under a pricing measure $\mathbb{Q}$. Note that $P^{\mathbb{Q}}(f)$ is a nonlinear function of the model parameters $\boldsymbol{\theta}$. We assume that $P^{\mathbb{Q}}(f)$ is a smooth function of $\boldsymbol{\theta}$ for each type of candidate models. In fact, this assumption should be a premise whenever a parametric model is employed to price exotic derivatives. Otherwise, there is an additional risk when the model parameters are not estimated with high accuracy.

If $\mathcal{Q}$ includes all the market-consistent martingale measures $\mathbb{Q}$, such that $P^{\mathbb{Q}}(f)$ is consist with the market price of each benchmark instrument with payoff f, the worst-case value calculation is the dual problem of searching for a semi-static super-hedging strategy when the payoff function satisfies some conditions (see e.g. [19,165]). Different from the prime-dual approach, our idea is to directly calculate the worst-case value of a derivative with a set of market-consistent martingale measures $\mathcal{Q}$.

The nonlinearity of the optimization problem (4.2) requires a global optimization method. In addition, an ensemble of derivative values is needed to quantify the uncertainty embedded in a target derivative. Hence, the Monte Carlo method can be used to solve the optimization problem (4.2). The basic idea of the direct Monte Carlo method is to search for the solution for the optimization problem by calculating the derivative value under each martingale measure $\mathbb{Q}$. It not only provides the global solution of the optimization problem, but also the ensemble of derivative values.

Since we are ambiguous about the model parameters and have no preferences over the realisations of the model parameters, a large ensemble of the model parameters can be generated according to the uniform distribution in each parameter direction. From another point of view, according to the principle of maximum entropy, if nothing is known about a distribution except that it belongs to a certain class, then the distribution with the largest entropy could be chosen as the default. The uniform distribution is the maximum entropy distribution among all continuous distributions which are supported in a finite interval [142]. Hence, we could choose the uniform distribution as an *auxiliary* distribution to generate a large ensemble of the model parameters on their support intervals. Thus, the upper and lower value bounds in (4.2) can be approximated by

$$\overline{P}(f) \approx \max_{\mathbb{Q}\in\bar{\mathcal{Q}}} P^{\mathbb{Q}}(f) \quad \text{and} \quad \underline{P}(f) \approx \min_{\mathbb{Q}\in\bar{\mathcal{Q}}} P^{\mathbb{Q}}(f), \tag{4.3}$$

where $\bar{\mathcal{Q}}$ is a set of the model parameter realisations. The larger the size of $\bar{\mathcal{Q}}$ is, the better (4.3) can approximate (4.2) because a small number of the realisations of model parameters may yield local optimal estimations for $\overline{P}(f)$ and $\underline{P}(f)$.

The basic idea of the Monte Carlo method for an optimization problem is straightforward, but it can be time-consuming if the original pricing problem has to be solved with a large ensemble of the model parameters, since the calculation of the exotic derivative value can be time-consuming under each realisation of the model parameters. If the value function is smooth enough with respect to the model parameters, it can be approximated by an interpolation formula. We can get the ensemble of the derivative values by substituting an ensemble of the model parameters into the interpolation formula. The interpolation-based method could greatly alleviate the computational burden. There is a rich body of literature on univariate interpolation. However, the choice of interpolation nodes in the multivariate case is more difficult. The sparse grid constructed in the Smolyak algorithm is one of the efficient ways to break the curse of dimensionality, a phrase that refers to the deterioration of the convergence rate and the explosion of computational effort as the dimension of the model parameters increases. This algorithm will be detailed in Section 4.3.

4.2.3 Uncertainty measures

Given the worst-case value of a target derivative, [49] proposes to measure the impact of parameter uncertainty on a derivative value by

$$\mu_{\mathcal{Q}}(f) = \overline{P}(f) - \underline{P}(f). \tag{4.4}$$

$\mu_{\mathcal{Q}}$ is a coherent uncertainty measure, verifying the following axioms [49]:

1. $\overline{P}$, $\underline{P}$ assign values to the benchmark derivatives compatible with their market bid-ask prices:

$$\forall i \in I, \quad C_i^{\text{bid}} \leq \underline{P}(f_i) \leq \overline{P}(f_i) \leq C_i^{\text{ask}}, \tag{4.5}$$

   where I is a set of benchmarks. Denote their payoffs as $(f_i)_{i\in I}$ and their observed market prices by $(C_i)_{i\in I}$. In most cases a unique price is not available; instead, we have a range of prices $C_i \in [C_i^{\text{bid}}, C_i^{\text{ask}}]$.

2. For liquid (benchmark) instruments, model uncertainty reduces to the uncertainty on market value:

$$\forall i \in I, \quad \mu_{\mathcal{Q}}(f_i) \leq \left| C_i^{\text{ask}} - C_i^{\text{bid}} \right|. \tag{4.6}$$

3. Effect of hedging with the underlying:

$$\forall \phi \in \mathcal{S}, \quad \mu_{\mathcal{Q}} \left(f + \int_0^T \phi_t \cdot dS_t \right) = \mu_{\mathcal{Q}}(f), \tag{4.7}$$

where $(S_t)_{0 \leq t \leq T}$ is the price process of the underlying and $\mathcal{S}$ is the set of well-defined admissible trading strategies. In particular, the value of a contingent claim which can be replicated in a model-free way by trading in the underlying has no model uncertainty:

$$\left[\exists f_0 \in \mathbb{R}, \exists \phi \in S, \forall \mathbb{Q} \in \mathcal{Q}, \quad \mathbb{Q} \left(f = f_0 + \int_0^T \phi_t \cdot dS_t \right) = 1 \right]$$
$$\Rightarrow \mu_{\mathcal{Q}}(f) = 0.$$

4. Convexity: model uncertainty can be decreased through diversification.

$$\mu_{\mathcal{Q}} \left(\lambda f_1 + (1 - \lambda) f_2 \right) \leq \lambda \mu_{\mathcal{Q}}(f_1) + (1 - \lambda) \mu_{\mathcal{Q}}(f_2), \tag{4.8}$$

for $\forall f_1, f_2 \in \mathcal{C}, \forall \lambda \in [0, 1]$, and where $\mathcal{C}$ is a set of well-defined contingent claims under all pricing models $\mathcal{Q}$.

5. Static hedging with traded options:

$$\forall f \in \mathcal{C}, \forall u \in \mathbb{R}^k, \mu_{\mathcal{Q}} \left(f + \sum_{i=1}^k u_i f_i \right) \leq \mu_{\mathcal{Q}}(f) + \sum_{i=1}^k \left| u_i \left(C_i^{\text{ask}} - C_i^{\text{bid}} \right) \right|. \tag{4.9}$$

In particular, for any payoff that can be statically replicated with traded options, model uncertainty reduces to the uncertainty on the cost of replication:

$$\left[\exists u \in \mathbb{R}^k, f = \sum_{i=1}^k u_i f_i \right] \Rightarrow \mu_Q \left(f \right) \leq \sum_{i=1}^k |u_i| \left| \left(C_i^{\text{ask}} - C_i^{\text{bid}} \right) \right|. \tag{4.10}$$

$\mu_{\mathcal{Q}}$ not only measures the impact of parameter uncertainty on derivative pricing, but also can be used to define the model risk ratio:

$$\text{MR}(f) = \frac{\mu_{\mathcal{Q}}(f)}{P_m(f)}, \tag{4.11}$$

where $P_m(f)$ could be the price of an option with the payoff function f or a position f in options. A high ratio $\text{MR}(f)$ indicates that the risk of model mis-specification is a

large component of the risk of a position in the option or the portfolio [49]. Hence, it's of great importance to explicitly calculate $\mu_{\mathcal{Q}}$ in the context of risk management.

The uncertainty measure $\mu_{\mathcal{Q}}$ defined by (4.4) only uses the maximum and minimum of an ensemble of the target derivative values. We propose to explore more information from the ensemble of the derivative values beyond the uncertainty measure $\mu_{\mathcal{Q}}$. In the field of information, it is broadly accepted that entropy represents the lack of information about the state of a system observed by an outsider [28]. The information is decreasing while entropy increases. Hence, to assess how much information we get from an ensemble of the option values, we will use entropy of the ensemble of the target derivative values to measure its information extent. This quantity is complementary to the coherent uncertainty measure $\mu_{\mathcal{Q}}$. If two derivatives have the same value of the uncertainty measure for a given ensemble of the model parameters , the one with smaller entropy embeds less uncertainty because its ensemble provides more information.

Definition 4.1 (Shannon Entropy [158])**.** *The Shannon entropy $H(X)$ of a discrete random variable X with distribution $p(x)$ is defined as:*

$$H(X) = \sum\nolimits_i p(x_i) \log \frac{1}{p(x_i)}. \tag{4.12}$$

In this chapter, we use the empirical distribution to calculate the entropy of a derivative value. We do not claim that the empirical distribution of the derivative value is the posterior distribution, because the uniform distribution is not the posterior distribution of the model parameters. This artificial distribution is only used to characterize the uncertainty of the derivative value when each combination of the model parameters is equally treated.

4.3 The Smolyak algorithm and Monte Carlo-based method

Based on the Smolyak algorithm, a Monte Carlo-based method is proposed to quantify the uncertainty embedded in the value of an exotic derivative.

4.3.1 The Smolyak algorithm

Let $D \subset \mathbb{R}^N$ be a N-dimensional bounded domain. Without loss of generality, hereafter we assume that D is a box

$$D = [-1, 1]^N, \quad N \geq 1. \tag{4.13}$$

Let $f : D \to \mathbb{R}$ be a smooth function to be interpolated by a set of orthogonal basis functions using a finite number of support interpolation nodes. In the setting of a one-dimensional interpolation ($N = 1$) or in each parameter direction $i \in \{1, \ldots, N\}$, the interpolation formula for the objective function f is given by

$$\mathcal{I}^i(f) = \sum_{m=1}^{m_i} f(y_m^i) a_m^i \tag{4.14}$$

based on nodal sets

$$\Theta^i = \{y_1^i, \ldots, y_{m_i}^i\} \subset [-1, 1], \tag{4.15}$$

where $\{a_m^i\}_{m=1,\ldots,m_i}$ are interpolation basis functions which satisfy the discrete orthogonality property $a_j^i(y_k^i) = \delta_{jk}$.

In the multivariate case ($N > 1$), a straightforward approach to approximate f on D is to construct a full tensor product interpolation by tensor products of the univariate ones. However, this method suffers from the the curse of dimensionality, because the number of interpolation nodes grows rapidly in high dimensions. An efficient approach is the Smolyak algorithm, which is given by

$$\begin{aligned} \mathcal{I}(f) &= \sum_{k+1 \le |\mathbf{i}| \le N+k} (-1)^{N+k-|\mathbf{i}|} \cdot \begin{pmatrix} N-1 \\ N+k-|\mathbf{i}| \end{pmatrix} \cdot \left(\mathcal{I}^{i_1} \otimes \cdots \otimes \mathcal{I}^{i_N}\right)(f) \\ &=: A(N+k, N)(f) \end{aligned} \tag{4.16}$$

where $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}^N$ is a multi-index with $|\mathbf{i}| = \sum_{l=1}^{N} i_l$, the integer k denotes the level of the construction and the tensor product rule is defined by

$$\left(\mathcal{I}^{i_1} \otimes \cdots \otimes \mathcal{I}^{i_N}\right)(f) = \sum_{k_1=1}^{m_{i_1}} \cdots \sum_{k_N=1}^{m_{i_N}} f(y_{k_1}^{i_1}, \ldots, y_{k_N}^{i_N}) \cdot (a_{k_1}^{i_1} \otimes \cdots \otimes a_{k_N}^{i_N}). \tag{4.17}$$

It is clear that the Smolyak algorithm is a linear combination of tensor product operators on the subsets (sparse grids) of the full grids. To set up $A(N+k, N)$, we only need to evaluate the objective function on the sparse grids

$$\Theta_N \equiv \mathcal{H}(N+k, N) = \bigcup_{k+1 \le |\mathbf{i}| \le N+k} (\Theta^{i_1} \times \cdots \times \Theta^{i_N}). \tag{4.18}$$

We refer to [177] for the detailed derivation of (4.16).

In the implementation of the Smolyak algorithm (4.16), different strategies can be used to select the interpolation basis functions and the interpolation nodes in each

direction. However, there is usually no explicit formula for the total number of interpolation nodes. But, when the nodal sets in each direction are nested, i.e., $\Theta^i \subset \Theta^j$, $\forall i < j$, the total number of interpolation nodes can reach a minimum. The Lagrange polynomials and the piecewise linear basis functions are two commonly used bases and the corresponding nodal sets are nested.

If the objective function to be interpolated is known to be very smooth, univariate Lagrange polynomial interpolation at the extrema of the Chebyshev polynomials is recommended to construct the Smolyak formula. The nodes are defined as

$$y_j^i = \begin{cases} 0 & m_i = 1,\ j = 1 \\ -\cos\dfrac{\pi(j-1)}{m_i - 1}, & m_i > 1,\ j = 1, \ldots, m_i. \end{cases} \tag{4.19}$$

where $m_1 = 1$ and $m_i = 2^{i-1} + 1$ for $i > 1$. The univariate Lagrange polynomial basis functions associated with Θ^i are

$$a_j^i(y) = \begin{cases} 1 & i = 1 \\ \displaystyle\prod_{\substack{k=1 \\ k \neq j}}^{m_i} \frac{y - y_k^i}{y_j^i - y_k^i} & i > 1,\ j = 1, \ldots, m_i. \end{cases} \tag{4.20}$$

Barthelemann et al. [16] derive the error bound of high-order polynomial interpolation with the Smolyak algorithm.

Theorem 4.2 ([16]). *In the space F_N^l defined by $\{f : [-1,1]^N \to \mathbb{R} \mid \partial^{|\mathbf{m}|} f$ is continuous, $m_i \leq l,\ \forall i\}$, where $\mathbf{m} \in \mathbb{N}_0^N$, $l \in \mathbb{N}_0$ and $\partial^{|\mathbf{m}|}$ is the usual N-variate partial derivative of order $|\mathbf{m}|$, the interpolation error is*

$$||f - A(N+k, N)(f)||_\infty \leq C_{N,l} \cdot M^{-l} \cdot (\log M)^{(l+2)(N-1)+1},$$

where $M = \dim(\mathcal{H}(N+k, N)) \approx \frac{2^k}{k!} N^k$ is the number of interpolation points.

Compared with the Lagrange polynomial which has a global support, the piecewise linear basis function has a local support. Hence, if one is not quite sure about the smoothness of the objective function, piecewise linear interpolation at equidistant support nodes is recommended. The equidistant support nodes are defined by

$$y_j^i = \begin{cases} 0, & m_i = 1,\ j = 1, \\ 2\dfrac{j-1}{m_i - 1} - 1, & m_i > 1,\ j = 1, \ldots, m_i, \end{cases} \tag{4.21}$$

where $m_1 = 1$ and $m_i = 2^{i-1} + 1$ for $i > 1$. Correspondingly, the piecewise basis functions are defined as $a_1^1(y) = 1$ for $i = 1$, and for $i > 1$

$$a_j^i(y) = \begin{cases} 1 - \frac{m_i}{2}|y - y_j^i|, & \text{if } |y - y_j^i| < \frac{2}{m_i} \\ 0, & \text{otherwise,} \end{cases} \tag{4.22}$$

where $j = 1, \ldots, m_i$.

The error bound of piecewise linear interpolation with the Smolyak algorithm was derived by Klimke and Wohlmuth [116].

Theorem 4.3 ([116]). *In the space F_N^2 defined by $\{f : [-1,1]^N \to \mathbb{R} \mid \partial^{|\mathbf{m}|} f$ is continuous, $m_i \leq 2, \forall i\}$, where $\mathbf{m} \in \mathbb{N}_0^N$ and $\partial^{|\mathbf{m}|}$ is the usual N-variate partial derivative of order $|\mathbf{m}|$, the piecewise linear interpolation error is*

$$||f - A(N+k, N)(f)||_\infty \leq C_N \cdot M^{-2} \cdot (\log_2 M)^{3\cdot(N-1)},$$

where $M = \dim(\mathcal{H}(N+k, N)) \approx \frac{2^k}{k!} N^k$ is the number of interpolation points.

Except for the Lagrange interpolation and the piecewise interpolation, one may refer to [134] and references therein for other strategies to choose the interpolation basis functions and the interpolation nodes.

4.3.2 Monte Calo-based method

Based on the Smolyak algorithm, the Monte Carlo(MC)-based method to quantify the uncertainty of a derivative value can be summarized as follows:

1. set up the interpolation formula (4.16) for the target value function,
2. get a huge ensemble of the target derivative value by substituting a huge ensemble of the model parameters into (4.16),
3. calculate the uncertainty measures (4.4) and (4.12).

In fact, once the interpolation formula (4.16) is set up, the computation will only involve interpolation basis functions (e.g. polynomials and piecewise linear functions). So, the computational cost is very small. Most of the attention goes to setting up the interpolation formula (4.16). The cost depends on the number of interpolation nodes and the computational cost of the value function on these nodes. Since the selection of the interpolation nodes is independent of the value function, we can select these nodes

beforehand and compute the value function on these nodes in a parallelized way. Note that either setting up the interpolation formula (4.16) or working with the classical Monte Carlo method requires calculations of the value function on a set of realisations of the model parameters. These calculations can be done in a parallelized way in each case. However, we calculate the derivative value on the fixed nodes in the former case while on the random nodes in the latter case. On the other hand, compared with the classical Monte Carlo method, the construction of the interpolation formula (4.16) requires far less nodes. Essentially, the MC-based method greatly reduces the computational cost since it avoids solving the original pricing problem for each combination of the model parameters, as is done in the standard Monte Carlo method for uncertainty quantification.

Since a large ensemble of the derivative values are calculated, the proposed Monte Carlo-based method is actually a global optimization method. It provides not only the maximum and minimum of the derivative value but also additional information to calculate the entropy measure. In addition, a trivial further application of the proposed method is to calculate the weighted average of the derivative value in the setting of parameter risk.

Remark 4.4. *The multivariate interpolation can be generalized to quantify the statistics of the solution of differential equations with random inputs, including random parameters, random boundary/initial conditions (see e.g. [178]).*

4.4 Numerical examples

In this section, we use the MC-based method to quantify the uncertainty embedded in Bermudan and Barrier options under the Heston model with parameter uncertainty. Parameter uncertainty is characterized by value intervals for each parameter.

4.4.1 Pricing models and target derivatives

In practice, the Heston model [103] is one of the most popular stochastic volatility models, especially in the equity market. Under this model, the asset price process (S) and the variance process (v) under a risk-neutral probability measure $\mathbb{Q}$ evolve according to the following system of stochastic differential equations,

$$\begin{cases} \mathrm{d}S_t = rS_t\,\mathrm{d}t + \sqrt{v_t}S_t\,\mathrm{d}W_t, & S_0 > 0, \\ \mathrm{d}v_t = \kappa(\eta - v_t)\,\mathrm{d}t + \sigma\sqrt{v_t}\,\mathrm{d}\widetilde{W}_t, & v_0 = \sigma_0^2 > 0, \end{cases} \tag{4.23}$$

where $(W_t)_{0\leq t\leq T}$ and $(\widetilde{W}_t)_{0\leq t\leq T}$ are correlated Brownian motions satisfying $\mathrm{d}W_t\mathrm{d}\widetilde{W}_t = \rho\mathrm{d}t$. The parameter η is the long-term average variance, while κ is the speed of the mean-reversion of the variance. The parameter σ is referred to as the volatility of variance since it scales the diffusion term of the variance process. r is the risk-free interest rate.

The price monotonicity of European vanilla options in the parameters of Heston's model has been investigated by Ould Aly [141]. However, it is difficult to know the monotonicity of the value function of an exotic derivative with respect to each parameter in the Heston model. For example, the monotonicity of an American or Bermudan put option value is one of the difficult cases. Assing et al. [8] proved the monotonicity of an American option value with respect to the initial values of the underlying price and the volatility process under two-dimensional models, including the Heston model. It is, however, unknown whether the value function of an American option is monotonic with respect to the other model parameters. Because of the nonlinearity of the value function with respect to the model parameters, it is difficult to check the corresponding monotonicity. In addition, to the best of our knowledge, the monotonicity of the value function of Barrier options under the Heston model is also unknown.

On the other hand, there is no analytical form for the value function of the American or Barrier options under the Heston model. The pricing problem has to be solved with numerical methods, such as PDE-based methods (see e.g. [45, 95, 109]), analytical approximation methods (see e.g. [77, 183]) and Monte Carlo methods (see e.g. [88]). It is too time-consuming to use these numerical methods to solve the pricing problem for thousands of times in the setting of uncertainty quantification.

In this section, we take Bermudan put options and the down-and-out put options as illustrating examples to show how to use the MC-based method to quantify the uncertainty of these options under the Heston model. When more candidate models are added, the worst-case value may become worse, i.e., lower bound becomes lower while upper bound becomes higher.

It is interesting and important to theoretically investigate the smoothness of Bermudan or Barrier option value functions under the Heston model. However, that is out of the scope of the thesis. Instead of checking the smoothness of the value function, we will check the accuracy of the interpolation formula for each option in the numerical tests. A PDE-based method will be employed to calculate the option value on the interpolation grids in the parameter space.

Let $P(s, v, t; \mathbf{Z})$ be the fair value of an option under the Heston model with a parameter set $\mathbf{Z}$ if at t time units before the given maturity time T the underlying asset

price equals $s \geq 0$ and the variance equals $v \geq 0$. The Heston spatial differential operator $\mathcal{A}$ is defined as

$$\mathcal{A} := \frac{1}{2}s^2 v\frac{\partial^2}{\partial s^2} + \rho\sigma sv\frac{\partial^2}{\partial s\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2} + rs\frac{\partial}{\partial s} + \kappa(\eta - v)\frac{\partial}{\partial v} - r. \tag{4.24}$$

$\mathcal{A}$ can be applied to P to define the valuation problem for American put options and down-and-out options.

1. American and Bermudan options

   The holder (buyer) of an American option has the right to exercise the option at any time before the expiration. The potential earlier exercise makes an American option more expensive than its counterpart of European style, which can only be exercised on a single date. Compared with an American option, a Bermudan option can be exercised only on predetermined dates.

   The valuation problem for American put options under the Heston model is reduced to solve the following so-called partial differential complementarity problem (PDCP):

$$\begin{cases} \dfrac{\partial P}{\partial t} \geq \mathcal{A}P, \quad P \geq \phi \\ (P - \phi)\left(\dfrac{\partial P}{\partial t} - \mathcal{A}P\right) = 0, \end{cases} \tag{4.25}$$

   for $s > 0$, $v > 0$, $0 < t \leq T$ and where $\phi = \max(K - s, 0)$, $\mathcal{A}$ is defined in (4.24) and $P(s, v, t)$, an abbreviation of $P(s, v, t; \mathbf{Z})$, is the option value. The initial condition and boundary conditions on the truncated domain $[0, S_{max}] \times [0, v_{max}]$ for the PDCP (4.25) are given by

$$P(s, v, 0) = \max(K - s, 0), \tag{4.26a}$$

$$P(0, v, t) = K, \tag{4.26b}$$

$$\frac{\partial P}{\partial s}(S_{\max}, v, t) = 0, \tag{4.26c}$$

$$\frac{\partial P}{\partial v}(s, v_{max}, t) = 0, \tag{4.26d}$$

   where $S_{\max}$ and $v_{\max}$ are set to be $14K$ and 5, respectively. The Heston PDCP problem (4.25) is implicitly assumed to be fulfilled at $v = 0$. Since this problem is solved on discrete temporal grid points, we actually calculate the Bermudan option value.

2. Down-and-out put options

   Given the lower bound B of a down-and-out put option, its value on the truncated domain $[B, S_{\max}] \times [0, v_{\max}]$ satisfies the following PDE

$$\frac{\partial P}{\partial t} = \mathcal{A}P \tag{4.27}$$

   with the initial and boundary conditions:

$$P(s, v, 0) = \max(K - s, 0), \tag{4.28a}$$
$$P(B, v, t) = 0, \tag{4.28b}$$
$$\frac{\partial P}{\partial s}(S_{\max}, v, t) = 0, \tag{4.28c}$$
$$\frac{\partial P}{\partial v}(s, v_{max}, t) = 0, \tag{4.28d}$$

   where $\mathcal{A}$ is defined in (4.24) and $P(s, v, t)$, an abbreviation of $P(s, v, t; \mathbf{Z})$, is the value of the down-and-out put option. It is assumed that (4.27) holds at $v = 0$, and $S_{\max}$ and $v_{\max}$ are set to be $14K$ and 5, respectively .

Assume r is a constant and let $\mathbf{Z} = (v_0, \lambda, \eta, \kappa, \rho)$ be the other model parameters in the Heston model. Without loss of generality, we assume that the other model parameters lie in the interval $[-1, 1]$, that is, $\mathbf{Z} \in D = [-1, 1]^5$. Otherwise, $[-1, 1]$ can be scaled into the confidence interval of each parameter. Then the worst-case values (4.2) are given by

$$\begin{aligned} \overline{P}(S_0, V_0, T; \mathbf{Z}, r) &= \sup_{\mathbf{Z} \in D} P(S_0, v_0, T; \mathbf{Z}, r), \\ \underline{P}(S_0, v_0, T; \mathbf{Z}, r) &= \inf_{\mathbf{Z} \in D} P(S_0, v_0, T; \mathbf{Z}, r), \end{aligned} \tag{4.29}$$

where P is the solution of (4.25) or (4.27) with corresponding conditions for American put options or down-and-out put options, respectively. In both cases, we presume that $P(S_0, V_0, T; \mathbf{Z}, r)$ is smooth in $\mathbf{Z}$. Thus, in (4.29) $\sup$ and $\inf$ are $\max$ and $\min$, respectively. Given the option parameters and r, $P(S_0, v_0, T; \mathbf{Z}, r)$ as a function of $\mathbf{Z}$ is approximated by the Smolyak interpolation formula, denoted by $\widetilde{P}(\mathbf{Z})$. To set up $\widetilde{P}(\mathbf{Z})$, we use the ADI scheme ([95, 109]) to solve the original pricing problem with the model parameters given by the interpolation nodes. Instead of the ADI scheme, one may use other numerical methods to calculate the fair values of the American or Barrier options. A comparison of these methods is out of the scope of this chapter.

The steps to quantify the uncertainty of a Bermudan put option (down-and-out put option) are as follows:

1. select a number (M) of interpolation nodes $\{z_j\}_{j=1,\ldots,M}$ in the parameter space,
2. calculate $P(S_0, v_0, T; z_j, r)$ for $j = 1, \ldots, M$ in a parallelized way with the ADI scheme,
3. code up $\widetilde{P}(\mathbf{Z})$ according to the approximation formula (4.16),
4. generate a huge ensemble with size M', $\{\hat{z}_i\}_{i=1,\ldots,M'}$ of the model parameters according to the uniform distribution[1] on the support interval in each parameter direction,
5. substitute $\{\hat{z}_i\}_{i=1,\ldots,M'}$ into $\widetilde{P}(\mathbf{Z})$ to generate the ensemble $\{\widetilde{P}(\hat{z}_i)\}_{i=1,\ldots,M'}$ of Bermudan put option (down-and-out put option) values,
6. find the maximum and minimum of $\{\widetilde{P}(\hat{z}_i)\}_{i=1,\ldots,M'}$,
7. measure the uncertainty of the value function with the coherent uncertainty measure (4.4) and the entropy measure (4.12).

4.4.2 Data set

The interval for each parameter represents the agent's uncertainty about this parameter. From this point of view, it is subjective to specify these intervals. Their specification relies on how much information the agent has. If the agent had perfect information to believe in a model and the estimated parameters, the spreads of these intervals would be zero. Otherwise, the parameter interval may be specified by the maximum and the minimum of the parameter estimations with different calibration methods, or by the Bayesian methods as in the references mentioned in Section 4.2.

According to the data of the calibrated Heston model to the market data in [91], κ and ρ vary over relatively big ranges compared with the ranges of the other parameters, if different calibration methods are used. So we assume the intervals corresponding to the model parameters' best estimations are given in the Table 4.1. We will use these model parameters to price Bermudan put options and down-and-out put options.

[1] One may refer to Section 4.2.2 for the explanations of our motivation of using the uniform distribution. M', the size of the ensemble, can be a very large positive integer. If the Bayesian method is used to calibrate a model to the market data, a large number of the realisations of the model parameters can be generated with the Markov Chain Monte Carlo method (see e.g. [94]). From this point of view, it is much clear that our method can be used in the setting of model risk.

The other parameters of Bermudan put options are given by

$$\begin{aligned} &r = 0.04, \qquad T = 0.25 \text{ year}, \quad S_0 = 100, \quad K = \{80, 100, 110\}, \\ &\sharp \text{ exercise times: } 60. \end{aligned} \tag{4.30}$$

which cover the in-the-money, out-of-the-money and at-the-money put options. The parameters of down-and-out put options are

$$r = 0.04, \quad T = 0.25 \text{ year}, \quad S_0 = 100, \quad K = \{80, 100, 110\}, \quad B = 70. \tag{4.31}$$

Z	min	max
v_0	0.015	0.025
κ	0.7	2
λ	0.15	0.25
η	0.01	0.02
ρ	-1	-0.5

Table 4.1: Heston's model parameters under uncertainty

4.4.3 Efficiency and accuracy

To set up the Smolyak interpolation formula, we could choose the Lagrange interpolation or the piecewise linear interpolation according to the smoothness of the objective function. In fact, they require the same number of interpolation nodes at the same level. However, the smoothness of a derivative value function usually is not available beforehand. Informally speaking, compared with the out-of-the-money options, the at-the-money and in-the-money option are more sensitive to volatility (see e.g. [38]). Thus, we select Lagrange polynomial interpolation to approximate the out-of-the-money options while the at-the-money and in-the-money options are approximated by piecewise linear interpolation. Theorem 4.2 and Theorem 4.3 show that it is better to choose a high level to set up the Smolyak interpolation formula if high accuracy is required. To select an accurate enough approximation formula, we set up the Smolyak interpolation formula at level $k = 2$ and $k = 3$ for the target options.

The number of interpolation nodes is outlined in Table 4.2. It means that we only have to solve the original pricing problem on the sparse grids in a parallelized way. So, it's quite practical and efficient. Of course, the total computational cost of setting up the interpolation formula relies on the computational cost of the pricing problem on each interpolation node.

Derivatives	Strike	Level2 (♯ grid points: 61)		Level3(♯ grid points: 241)	
		AbsError	RelError	AbsError	RelError
Bermudan put	80	$0.0928{\cdot}10^{-3}$	$0.5758{\cdot}10^{-2}$	$0.0047{\cdot}10^{-3}$	$0.3004{\cdot}10^{-3}$
	100	$2.7975{\cdot}10^{-3}$	$0.1194{\cdot}10^{-2}$	$0.6668{\cdot}10^{-3}$	$0.2839{\cdot}10^{-3}$
	110	$1.8079{\cdot}10^{-3}$	$0.1808{\cdot}10^{-3}$	$0.9100{\cdot}10^{-3}$	$0.0910{\cdot}10^{-3}$
Down-and-out put	80	$0.1322{\cdot}10^{-3}$	$0.2256{\cdot}10^{-2}$	$0.0044{\cdot}10^{-3}$	$0.3365{\cdot}10^{-3}$
	100	$3.3613{\cdot}10^{-3}$	$0.1470{\cdot}10^{-2}$	$0.7861{\cdot}10^{-3}$	$0.3418{\cdot}10^{-3}$
	110	$2.0383{\cdot}10^{-3}$	$0.2247{\cdot}10^{-3}$	$0.4477{\cdot}10^{-3}$	$0.0493{\cdot}10^{-3}$

Table 4.2: Number of grid points and the mean value of the error of the approximation formula for Bermudan and down-and-out put options

To check the accuracy of the interpolation formulas, we follow four steps:

1. Fifty realisations of the model parameters are randomly generated which are distributed according to independent uniform distributions on the candidate intervals of the model parameters. These model parameters will be used to calculate the value of all target options in the following two steps.

2. Each combination of the model parameters is substituted into the corresponding interpolation formula ($k = 2, 3$) for the target option value. That is, an ensemble of option values $\widetilde{P}_i$ ($i = 1, \ldots, 50$) can be calculated in an efficient way.

3. Correspondingly, a reference value $\hat{P}_i$ ($i = 1, \ldots, 50$) of a target option is calculated with the ADI scheme for each combination of the model parameters.

4. The accuracy can be measured by the the absolute error (AbsError, $|\widetilde{P}_i - \hat{P}_i|$) or the relative error (RelError, $|\widetilde{P}_i - \hat{P}_i|/\hat{P}_i$).

Table 4.2 summarizes the mean value of the absolute error and the relative error of the approximation formula for each target option. Compared with the interpolation formulas at level 2 ($k = 2$), the approximation formulas at level 3 ($k = 3$) are accurate enough for the target options, they will be selected to calculate the uncertainty measure in the next subsection.

4.4.4 Uncertainty measures

By the accurate interpolation formulas at level $k = 3$ for the Bermudan put options and the down-and-out put options, it is efficient to get a large ensemble of the target option values by substituting a huge ensemble of the model parameters into the approximation formula.

Different ensembles of the model parameters are firstly generated. The ensemble size of the model parameters starts from 2500, increases by 50 each time and ends with 7500. Since the computation of the option value with the approximation formula only involves the computation of simple basis functions, one may enlarge the ensemble size to get more accurate results. This is the advantage of the proposed MC-based method.

Each ensemble is substituted into the approximation formula for the target options, and the resulting ensemble of option values will be used to calculate the uncertainty measure. We first calculate the coherent uncertainty measure by searching for the maximum and minimum values of the target options in different ensembles.

In each ensemble, the interval between the minimum and maximum values is separated into 20 subintervals of the same size. The frequencies of the option value falling into these subintervals work as the probability distribution to calculate the entropy measure (4.12). For each option, we calculate the entropy and spread between the maximum and minimum option values of each ensemble (total 150).

1. Coherent uncertainty measure

   The left three sub-figures in Figure 4.1 show the impact of parameter uncertainty on the Bermudan put options. Qualitatively speaking, parameter uncertainty has much bigger impact on the at-the-money Bermudan put option than that on the Bermudan put options with $K = 80, 110$. This is also the case for down-and-out put options, as shown in the left three sub-figures in Figure 4.2.

   One may note that the value bounds of the options do not change too much when the size of ensemble increases. Informally speaking, if we only use two realisations of the model parameters, one of them may be accidently consistent with the global maximum and the other one may accidently correspond to the global minimum of the option value. Adding more realisations will not change the value bounds. However, we do not know beforehand what is the appropriate size of the ensemble to provide results of high accuracy. It is conservative to use a large ensemble, after all the computation mainly involves the simple basis functions after the interpolation formula (4.16) has been set up.

2. Entropy measure

   The sub-figure for entropy measure in Figure 4.1 shows that among the three Bermudan put options, the entropy value of the at-the-money option is the largest while that of the option with $K = 80$ is the smallest one.

In Figure 4.2, the sub-figure for entropy measure shows that at-the-money down-and-out put option has the largest value of the entropy measure while the entropy values of options with $K = 80, 110$ are similar.

Table 4.3 summarizes the average values of the coherent uncertainty measure (4.4) and the entropy measure (4.12) for the target options. These quantitative results are consistent with our qualitative observations in Figure 4.1 and Figure 4.2. For both, Bermudan put options and down-and-out put options, the values of the coherent uncertainty measure and of the entropy measure for the at-the-money option are larger than those for the other two options. It means that parameter uncertainty has more impact on the at-the-money Bermudan (down-and-out) put option. Based on the value of the coherent uncertainty measure (4.4), the model risk ratio (4.11) can be used to assess the risk of misusing a model when one takes a position in an option with a given price.

One may also compare the value of the uncertainty measure for different types of target options, such as Bermudan put option and down-and-out put option. According to the coherent uncertainty measure, parameter uncertainty under the Heston model has more impact on the Bermudan put options with strike $K = 80, 100$ than that on the down-and-out put options with the same strike as the former one. However, compared with the Bermudan put option with strike $K = 110$, the down-and-out put option with the same strike is much more impacted by the parameter uncertainty under the Heston model. On the other hand, we can use the entropy measure to characterize how much information we get from the ensemble of option values. It is a complementary quantity to measure the impact of parameter uncertainty if two options have the same value for the coherent uncertainty measure, In our numerical experiments, we do not find two options with the same value spread. Hence, in our illustrating experiments, we assess the impact of parameter uncertainty on options of different types according to the coherent uncertainty measure. Nevertheless, we can assess the extent of the information provided by the ensemble according to entropy. Although parameter uncertainty has more impact on the Bermudan put option with strike $K = 80$ than that on the down-and-out put option with the same strike according to the coherent uncertainty measure, the ensemble of the former provides more information than the information provided by the ensemble of the latter one according to the average value of entropy.

Overall, these numerical results provide an agent with the following information. First, if the Heston model is chosen as the candidate model to value Bermudan options and down-and-out put options, the bid-ask spread of the at-the-money option could be

	strike	spread	entropy
Bermudan put	80	0.0460	3.6935
	100	0.7323	4.1843
	110	0.0164	1.8967
Down-and-out put	80	0.0269	3.8468
	100	0.7226	4.1598
	110	0.4559	3.9207

Table 4.3: Mean value of uncertainty measures for Bermudan put options and down-and-out put options

higher than those of the options with the other moneyness. Second, compared with the Bermudan put option with strike $K = 110$, the bid-ask spread of the Bermudan put option with strike $K = 80$ could be higher, because more uncertainty is embedded in this option in terms of coherent uncertainty measure and the entropy measure. However, the bid-ask spread of the down-and-out put option with strike $K = 80$ could be smaller than that of the down-and-out put option with strike $K = 110$. Third, the bid-ask spread of Bermudan put options with strike $K = 80, 100$ could be bigger than their counterparts of down-and-out put options, whereas the value spread of the Bermudan put option with strike $K = 110$ could be smaller than that the down-and-out put option with the same strike. The final market price is left to a bargaining process in the OTC market.

4.5 Discussion

For a given stochastic model of a specific type, we proposed the interpolation approximation formula (4.16) for the calculation of a derivative value for any realisation of the model parameters.

The proposed method shares some similarity with the Monte Carlo method in the sense that both of them need to calculate the derivative value with some realisations of the model parameters. However, we should note their differences. When setting up the approximation formula (4.16), we only have to calculate the derivative value with a *small* number of pre-selected realisations of model parameters. Thereafter, a large ensemble of the derivative values can be calculated by substituting a large ensemble of parameters into the approximation formula (4.16). The Monte Carlo method provides a large ensemble of the derivative values by solving the *original* pricing problem with each realisation of the model parameters. Since solving the original pricing problem is

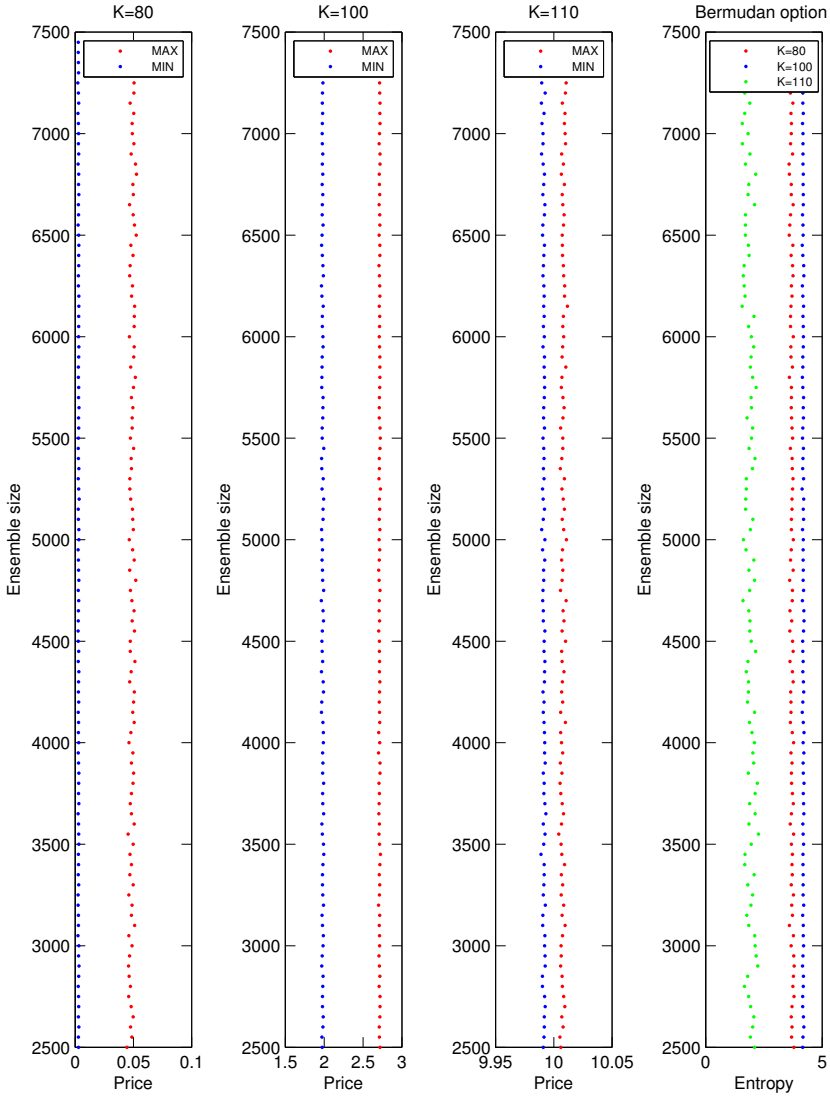


Figure 4.1: Coherent uncertainty measure and entropy measure of Bermudan put options.

usually time-consuming, our method is more efficient than the Monte Carlo method.

The proposed method is a reusable simulation infrastructure in the sense that

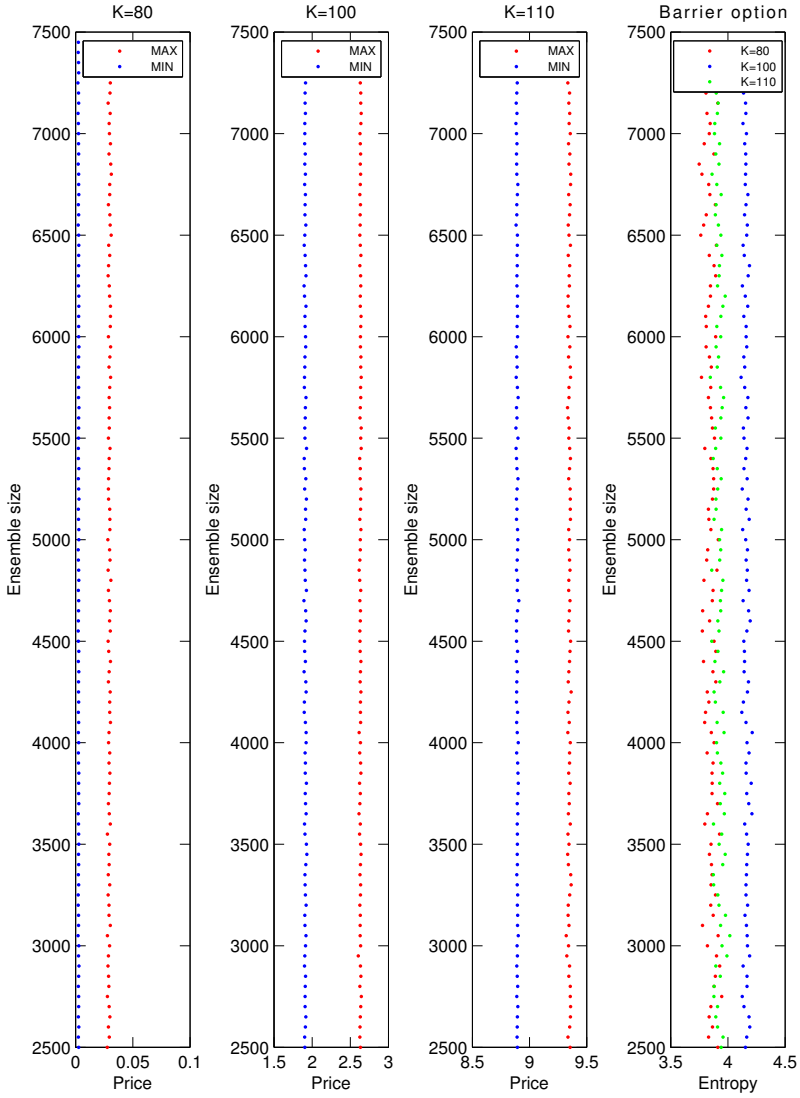


Figure 4.2: Coherent uncertainty measure and entropy measure of down-and-out put options

the interpolation formula (4.16) can be used in a reproducible manner if all model parameters are taken into consideration. Except for quantifying uncertainty embedded

in derivative value, it can also be used in a model validation procedure [102] and the robust calibration (see e.g. [13, 66, 74, 94]).

In 2015, Gaß et al. [87] proposed to interpolate the derivative value in the parameter space with the tensorized Chebyshev polynomials, while we set up the interpolation formula for a derivative value with the Smolyak algorithm. Independent of [87], we came up with this idea and presented our method at conferences and workshops[2]. Compared with the interpolation approximation with the tensorized Chebyshev polynomials, the Smolyak approximation is an interpolation approximation with a smaller number of interpolation nodes, especially in the high-dimensional setting (see e.g. [178]).

[2] (1)Stochastic calculus, Martingales, and Financial modeling, June 29 July 06, 2014, Saint Petersburg, Russia
(2)Summer School: Model Uncertainty in Economics and Finance - Advances in Stochastic Calculus July 7 July 18, 2014, Bielefeld University, Germany

5

Analytical approximation for distorted expectations

The distorted expectation of a random variable X is defined by

$$\mathbb{E}_{\Psi}[X] = \int_{-\infty}^{+\infty} x \, \mathrm{d}\Psi(F_X(x)),$$

where F_X is the cumulative distribution function of X, and Ψ is a distortion function. The distorted expectation is a non-linear expectation, which is extensively used in behavioural finance, conic finance, insurance, financial risk management (see e.g. [2, 15, 44, 62, 99, 125, 128, 168, 174, 175, 181], to name a few).

Generally speaking, a distorted expectation does not admit an analytical formula. One of the difficulties to derive such an formula lies in that there is no analytical formula for the distorted probability density function of X. Although the Monte Carlo method can be used to calculate distorted expectations, it can be time-consuming if sophisticated models have to be simulated. In this chapter, an efficient analytical approximation is proposed as an alternative to the Monte Carlo method. The proposed method can be about 2500 times more efficient than the Monte Carlo method. The main results of this chapter were published in Sun *et. al.* (2015).

5.1 Introduction

Distributions of risk factors are cornerstones in quantitative risk management, portfolio selection and pricing of financial derivatives. These distributions can be specified in the form of a probability distribution function or a parametric model. However, model uncertainty is ubiquitous whenever a parametric model (distribution) is used. Since decision making is the last step in risk management, portfolio selection and derivative pricing in an incomplete market model, model uncertainty in these fields can be interpreted from the point of view of decision theory under uncertainty.

Multiple priors, which is related to the Choquet expected utility, provide a powerful tool to model uncertainty in decision theory (see e.g. [75] for a review). This approach characterizes a random variable with a set of probability measures, which are candidate distributions of the random variable. There is no rule of thumb to select such a set of probability measures to account for model uncertainty. In Chapter 4, we have seen that these multiple priors can be selected through calibration. "Distance" measures such as relative entropy [97] can be used to specify multiple priors for a risk factor. In addition, distortion of a probability measure is another way to select a set of probability measures. It induces a distorted expectation, which is the worst-case expectation of a random variable under a set of probability measures, i.e.

$$\mathbb{E}_\Psi[X] = \int_{-\infty}^{+\infty} x \, \mathrm{d}\Psi(F_X(x)) = \inf_{Q \in \mathcal{Q}_\Psi} E^Q[X],$$

where $\mathcal{Q}_\Psi$ is a set of probability measures associated with a concave distortion function Ψ. These probability measures, termed as test measures in [34], represent the candidate distributions of X.

The distorted expectation is not only related to model uncertainty but also risk measures (see e.g. [62, 90, 174, 175, 181]). When the distortion function is concave, the distorted expectation $\mathbb{E}_\Psi[\cdot]$ can be used to define a coherent risk measure $\rho(\cdot) := -\mathbb{E}_\Psi[\cdot]$ (see e.g. [80] for the details). Employing the fact that a coherent risk measure is related to an acceptability set of risks, Cherny and Madan [44] proposed several acceptability indexes, which are further used to define the bid and ask prices in the framework of conic finance [68, 69, 127–131].

If the distortion function Ψ is of S-shape, the corresponding distortion function is related to the cumulative prospect theory (CPT) [169]. The CPT framework allows different weighting functions for gains and for losses. It corresponds to an agent who is more sensitive to losses than gains. Since this framework is consistent with experimental

evidence, it has been used in portfolio theory (see e.g. [99]), insurance (see e.g. [36]) and so on (see [14] for a literature review on applications of the CPT theory).

Although distorted expectations are extensively used in finance, insurance and many other areas, if any, quite few models for the risk factor X admit an analytical formula for the distorted expectation with respect to a distortion function in the general setting. The Monte Carlo (MC) method is a standard way to calculate the distorted expectations. If $\{x_1, \dots, x_N\}$ is a set of MC realisations of the risk factor X, or its sample from market data, the value of the distorted expectation $\mathbb{E}_\Psi[X]$ can be estimated by

$$\mathbb{E}_\Psi[X] \doteq \sum_{n=1}^{N} x_{(n)} \left(\Psi\left(\frac{n}{N}\right) - \Psi\left(\frac{n-1}{N}\right) \right), \tag{5.1}$$

where $x_{(1)}, \dots, x_{(N)}$ are the values of $x_1, \dots, x_N$ in the increasing order [44]. To achieve high accuracy, a large ordered sample of the risk factor is required. This procedure can be time-consuming under complex models for the risk factor, such as the Heston model [103]. On the other hand, some efficient numerical methods for the standard expectations, such as the COS method [76] and the Carr–Madan method [35], can hardly be used to calculate the distorted expectations, because there is no analytical formula for the *distorted* characteristic function in most of the realistic cases.

To overcome these difficulties, we propose an analytical approximation method, as an alternative to the MC method, for a class of distorted expectations. In applications, the parametric model for the risk factor X usually admits the density function f_X or the characteristic function ϕ_X in an analytical form. Then, the distribution function F_X can be numerically calculated or recovered from its characteristic function. In this chapter, we denote the distorted density function $\tilde{f}_X$ by $(\Psi' \circ F_X) \cdot f_X$ with $\tilde{f}_X(x) = \Psi'(F_X(x)) f_X(x)$, $x \in \mathbb{R}$. The fast Fourier transform (FFT) algorithm is used to set up an approximation for $\tilde{f}_X$ with a truncated sum of its Fourier-cosine series expansion on a finite interval. The resulting truncation approximation immediately provides an analytical approximation for the distorted expectations. The proposed method can be about 2500 times more efficient than the Monte Carlo method.

5.2 Preliminaries

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space. Denote by $L^2(\Omega, \mathcal{F}, P)$ the collection of square-integrable random variables representing the profits and losses of a financial position. The risk of a random variable $X \in L^2(\Omega, \mathcal{F}, P)$ can be quantified

with a risk measure ρ, a functional defined on $L^2(\Omega, \mathcal{F}, P)$. Probability distortion provides a large class of risk measures.

5.2.1 Distorted expectations

Definition 5.1. *Let $\Psi : [0,1] \to [0,1]$ be a continuous increasing function such that $\Psi(0) = 0$ and $\Psi(1) = 1$. The set function c_Ψ defined by*

$$c_\Psi(A) = \Psi(P(A)), \quad A \in \mathcal{F},$$

is called the distortion of the probability measure P with respect to the distortion function Ψ.

With a probability distortion function Ψ is associated another probability distortion function $\hat{\Psi}$ given by

$$\hat{\Psi}(x) = 1 - \Psi(1-x), \quad x \in [0,1].$$

Given a probability distortion Ψ, we can define the Choquet integral $\mathbb{E}_\Psi[X]$ of $X \in L^2(\Omega, \mathcal{F}, P)$ by

$$\mathbb{E}_\Psi[X] = \int_0^\infty (1 - c_\Psi(\{X \le x\}))\, \mathrm{d}x - \int_{-\infty}^0 c_\Psi(\{X \le x\})\, \mathrm{d}x. \tag{5.2}$$

We assume that $\int_0^1 \Psi(y) \frac{dy}{2y\sqrt{y}} < \infty$ and the probability space $(\Omega, \mathcal{F}, P)$ is atomless throughout this chapter. Then, according to [17], $\mathbb{E}_\Psi[X]$ is finite for any $X \in L^2(\Omega, \mathcal{F}, P)$ when Ψ is concave and continuous on $[0,1]$.

Let F_X be the distribution of X. We can rewrite $\mathbb{E}_\Psi[X]$ as

$$\mathbb{E}_\Psi[X] = \int_{-\infty}^{+\infty} x\, \mathrm{d}\Psi(F_X(x)). \tag{5.3}$$

$\mathbb{E}_\Psi$ refers to a distorted expectation associated with a distortion function Ψ. If $\Psi(x) = x$, $\mathbb{E}_\Psi$ is the standard expectation E.

For any $X \in L^2(\Omega, \mathcal{F}, P)$, the distorted expectation $\mathbb{E}_\Psi[X]$ associated with a concave and continuous distortion function Ψ admits a robust representation

$$\mathbb{E}_\Psi[X] = \inf_{Q \in \mathcal{Q}_\Psi} \mathbb{E}^Q[X], \tag{5.4}$$

where

$$\mathcal{Q}_\Psi = \left\{ Q \in \mathcal{M}_P : \hat{\Psi}(P(A)) \le Q(A) \le \Psi(P(A)) \quad \text{for all } A \in \mathcal{F} \right\},$$

with $\mathcal{M}_P$ being the collection of all probability measures absolutely continuous with respect to P [130].

5.2.2 Coherent risk measures

Definition 5.2. *A coherent risk measure on $L^2(\Omega, \mathcal{F}, P)$ is a map $\rho : L^2(\Omega, \mathcal{F}, P) \to \mathbb{R}$ satisfying the following properties:*

1. *(subadditivity) for $X, Y \in L^2(\Omega, \mathcal{F}, P)$, $\rho(X + Y) \leq \rho(X) + \rho(Y)$;*
2. *(monotonicity) If $X \leq Y$ $a.s.$, then $\rho(X) \leq \rho(Y)$;*
3. *(positive homogeneity) $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda \in \mathbb{R}_+$;*
4. *(translation invariance) $\rho(X + m) = \rho(X) - m$ for $m \in \mathbb{R}$;*

A coherent risk measure ρ defined on the space $L^\infty(\Omega, \mathcal{F}, P)$ of all bounded random variables $(a.s.)$, has a robust representation [6]

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[-X] = - \inf_{Q \in \mathcal{Q}} \mathbb{E}^Q[X], \quad \text{for } X \in L^\infty(\Omega, \mathcal{F}, P), \tag{5.5}$$

where $\mathcal{Q}$ is a subset of $\mathcal{M}_{1,f}(\Omega, \mathcal{F})$, the set of all finitely additive normalized set functions $Q : \mathcal{F} \to [0, 1]$. $\mathcal{Q}$ can be chosen as a convex set for which the supremum or infimum is attained (Proposition 4.15 of [80]). This robust representation can be generalized to the risk measures defined on the space $L^0(\Omega, \mathcal{F}, P)$ of all random variables [43]. Recalling the robust representation of the distorted expectation (5.4) defined on $L^2(\Omega, \mathcal{F}, P)$, we can identify the relation between risk measures and distorted expectations as follows:

$$\rho(X) := -\mathbb{E}_\Psi[X], \quad X \in L^2(\Omega, \mathcal{F}, P),$$

is a coherent risk measure if Ψ is concave and continuous on $[0, 1]$. The Average Value at Risk $(AV@R)$ is an example of the distortion risk measures [80].

Example 5.3 (Average Value at Risk)**.** *Let $\mathcal{Q}_\lambda$ be the class of all probability measures $Q \ll P$ whose density dQ/dP is bounded by $1/\lambda$ for some fixed parameter $\lambda \in (0, 1)$. The coherent risk measure*

$$AV@R_\lambda(X) := \sup_{Q \in \mathcal{Q}_\lambda} \mathbb{E}^Q[-X]$$

is so-called the Average Value at Risk. Actually,

$$AV@R_\lambda(X) = -\frac{1}{\lambda} \int_0^\lambda q_X(t)\, \mathrm{d}t,$$

where q_X is the quantile function of X. Induced by a distorted expectation, $AV@R_\lambda(X) = -\mathbb{E}_{\Psi^\lambda}[X]$ where

$$\Psi^\lambda(x) = \left(\frac{x}{\lambda}\right) \wedge 1, \quad x \in [0,1].$$

To complete this section, we outline some other risk measures induced by their corresponding distortion functions (see e.g. [44, 80] for details) .

1. For $\lambda > 0$, $MINV@R$ is induced by

$$\Psi^\lambda(x) = 1 - (1-x)^{\lambda+1}.$$

For any integer λ, we have $\hat{\mathbb{E}}[X] = E[Y]$, where

$$Y \overset{\text{law}}{=} \min\{X_1, \ldots, X_{\lambda+1}\}$$

and $\{X_1, \ldots, X_{\lambda+1}\}$ are independent draws of X.

2. For $\lambda > 0$, $MAXV@R$ is induced by

$$\Psi^\lambda(x) = x^{\frac{1}{1+\lambda}}.$$

For any integer λ, we have $\hat{\mathbb{E}}[X] = E[Y]$, where

$$\max\{Y_1, \ldots, Y_{\lambda+1}\} \overset{\text{law}}{=} X,$$

where $\{Y_1, \ldots, Y_{\lambda+1}\}$ are independent draws of Y.

3. For $\lambda > 0$, $MAXMINV@R$ [44] is induced by

$$\Psi^\lambda(x) = \left(1 - (1-x)^{\lambda+1}\right)^{\frac{1}{1+\lambda}}, \quad x \in [0,1]. \tag{5.6}$$

For any integer λ, we have $\mathbb{E}_\Psi[X] = E[Y]$, where Y is a random variable with the property

$$\max\{Y_1, \ldots, Y_{\lambda+1}\} \overset{\text{law}}{=} \min\{X_1, \ldots, X_{\lambda+1}\},$$

where $\{X_1, \ldots, X_{\lambda+1}\}$ being independent draws of X and $\{Y_1, \ldots, Y_{\lambda+1}\}$ being independent draws of Y.

4. For $\lambda > 0$, $MINMAXV@R$ [44] is induced by

$$\Psi^\lambda(x) = 1 - \left(1 - x^{\frac{1}{\lambda+1}}\right)^{\lambda+1}, \quad x \in [0,1]. \tag{5.7}$$

For any positive integer λ, we have $\mathbb{E}_\Psi[X] = E[Y]$, where Y is a random variable with the property

$$\begin{cases} Y \overset{\text{law}}{=} \min\{Z_1, \ldots, Z_{\lambda+1}\} \\ \max\{Z_1, \ldots, Z_{\lambda+1}\} \overset{\text{law}}{=} X \end{cases}$$

where $\{Z_1, \ldots, Z_{\lambda+1}\}$ are independent draws of Z.

5.2.3 Recovering the probability density function or the cumulative distribution function from the characteristic function

The probability density function (or density function, PDF), and the cumulative distribution function (or distribution function, CDF) of a random variable play a crucial role in the definition and the calculation of a distorted expectation (5.3). However, we may know the characteristic function of a random variable rather than its density function or distribution function in an analytical form. Hence, it is necessary to recover the density function or the distribution function from the corresponding characteristic function of a random variable.

Inspired by the idea of the COS method [76], we present an efficient approximation method for the distribution function of a random variable when its characteristic function is given in an analytical form. The derivation of this approximation method follows three steps.

1. Truncate the support of a density function into an finite interval.

   Let f be the density function of a random variable X whose characteristic function ϕ is defined by

   $$\phi(\omega) = \int_{\mathbb{R}} e^{i\omega x} f(x) dx. \tag{5.8}$$

   As a density function, f, decays to zero at $\pm\infty$, the integration range in (5.8) can be truncated in a large enough interval $[a, b] \subset \mathbb{R}$ such that

   $$\phi_1(\omega) = \int_a^b e^{i\omega x} f(x) dx \approx \phi(\omega). \tag{5.9}$$

2. Approximate the density function f with an approximation of an auxiliary function f_1, defined by $f_1(x) = f(x)\mathbf{1}_{[a,b]}(x)$, where $\mathbf{1}_{[a,b]}$ is an indicator function and $x \in \mathbb{R}$.

Since the function f_1 is supported on a finite interval $[a, b]$, its Fourier-cosine series expansion is

$$f_1(x) = {\sum}_{k=0}^{\prime\infty} A_k \cdot \cos\left(k\pi \frac{x-a}{b-a}\right), \tag{5.10}$$

where $\sum^{\prime}$ indicates that the first term in the summation is weighted by one-half, and the cosine series coefficients A_k are given by

$$\begin{aligned} A_k &= \frac{2}{b-a}\int_a^b f(x)\cos\left(k\pi\frac{x-a}{b-a}\right)dx \\ &\equiv \frac{2}{b-a}\mathrm{Re}\left\{\phi_1\left(\frac{k\pi}{b-a}\right)\cdot\exp\left(-i\frac{ka\pi}{b-a}\right)\right\} \end{aligned} \tag{5.11}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real part of the argument. Due to (5.9), A_k can be approximated by F_k,

$$F_k = \frac{2}{b-a}\mathrm{Re}\left\{\phi\left(\frac{k\pi}{b-a}\right)\cdot\exp\left(-i\frac{ka\pi}{b-a}\right)\right\}. \tag{5.12}$$

Hence, f_1 can be approximated by f_2,

$$f_2(x) = {\sum}_{k=0}^{\prime\infty} F_k \cdot \cos\left(k\pi \frac{x-a}{b-a}\right). \tag{5.13}$$

Then, f_2 can be approximated by a truncated summation f_3,

$$f_3(x) = {\sum}_{k=0}^{\prime n} F_k \cdot \cos\left(k\pi \frac{x-a}{b-a}\right). \tag{5.14}$$

The resulting error in $f_3(x)$ consists of two parts: a series truncation error from $f_2(x)$ to $f_3(x)$ and the error from approximating A_k by F_k. We will present the truncation error in the next section, while we refer to [76] for an error analysis related to the approximation error.

3. Approximate the distribution function

   Given the approximations in the previous steps, we approximate the distribution function F_X on the finite interval $[a, b]$ in the following way,

$$F_X(x) = \int_{-\infty}^{x} f(s)ds$$

$$
\begin{aligned}
&\approx \int_a^x f_3(s)ds \\
&= \int_a^x {\sum}'_{k=0}^{n} F_k \cdot \cos\left(k\pi \frac{s-a}{b-a}\right) ds \\
&= {\sum}'_{k=0}^{n} F_k \int_a^x \cos\left(k\pi \frac{s-a}{b-a}\right) ds \\
&= \frac{x-a}{b-a} + {\sum}_{k=1}^{n} F_k \cdot \frac{b-a}{k\pi} \cdot \sin(k\pi \frac{x-a}{b-a})
\end{aligned} \tag{5.15}
$$

where F_k is given in (5.12).

To conclude, if the characteristic function ϕ, rather than the density function f, of a random variable X is explicitly given in an analytical form, its distribution function F_X can be approximated by $\hat{F}_X$:

$$
\hat{F}_X(x) = \begin{cases} 1, & \text{if } x \geq b, \\ \dfrac{x-a}{b-a} + {\sum}_{k=1}^{n} F_k \cdot \dfrac{b-a}{k\pi} \cdot \sin(k\pi \dfrac{x-a}{b-a}), & \text{if } x \in (a,b), \\ 0, & \text{if } x \leq a, \end{cases} \tag{5.16}
$$

where a and b are chosen such that

$$
\int_a^b e^{i\omega x} f(x)dx \approx \int_{\mathbb{R}} e^{i\omega x} f(x)dx,
$$

and

$$
F_k = \frac{2}{b-a} \mathrm{Re}\left\{ \phi\left(\frac{k\pi}{b-a}\right) \cdot \exp\left(-i\frac{ka\pi}{b-a}\right) \right\}.
$$

In practice, a can be sufficiently small, while b is sufficiently large.

5.3 T-COS method for distorted expectations

Since the density function of the risk factor X can be estimated with statistical methods or recovered from its characteristic function, without loss of generality, we assume the density function f_X is given in its analytical form. Then, its distribution function can be calculated either by numerical integration methods or recovered from its characteristic function (see Section 5.2.3 or [140, 164]). Hence, we actually assume that both the density function and the distribution function are known when formulating the proposed approach.

Given the distribution function F_X and the distortion function Ψ, we can, theoretically, derive the distorted distribution function $\tilde{F}_X$ and the distorted density function $\tilde{f}_X$ as

$$\begin{cases} \tilde{F}_X(x) = \Psi\left(F_X(x)\right), & x \in \mathbb{R}, \\ \tilde{f}_X(x) = \Psi'\left(F_X(x)\right) f_X(x), & x \in \mathbb{R}. \end{cases} \tag{5.17}$$

However, due to the complexity of (5.17), it can hardly lead to an analytical formula for the distorted expectation (5.3). On the other hand, the COS method [76] cannot be used in this setting, because we do not know the analytical form of the distorted characteristic function of a random variable.

Truncating the support of a density function into a finite interval is one of the key ideas behind the COS method. Inspired by this idea, we first truncate the integration interval of the distorted expectation into a finite interval $[a, b]$ such that

$$\int_{(-\infty,a)\cup(b,+\infty)} x\tilde{f}_X(x)\,\mathrm{d}x < \varepsilon, \quad \text{for a given tolerance error } \varepsilon > 0.$$

Define $\hat{f}_X = \tilde{f}_X \mathbf{1}_{[a,b]}$, where $\mathbf{1}_{[a,b]}$ is an indicator function. For the convenience of formulating an algorithm in the following text, we present the Fourier-cosine series expansion of $\hat{f}_X$ as

$$\hat{f}_X(x) = \sum\nolimits_{k=0}^{\infty} \hat{F}_k \cdot \cos\left(k\pi \frac{x-a}{b-a}\right). \tag{5.18}$$

Note that this formulation is equivalent to (5.10). However, we cannot approximate the series coefficients in terms of the distorted characteristic function, as we do in the COS method (see (5.11) and (5.12)). Before giving the calculation method for these series coefficients in (5.18), we propose to approximate the $\hat{f}_X$ by the truncated sum $\bar{f}_X$ of a Fourier-cosine series expansion, i.e.,

$$\hat{f}_X(x) \approx \bar{f}_X(x) = \sum\nolimits_{k=0}^{K} \hat{F}_k \cdot \cos\left(k\pi \frac{x-a}{b-a}\right). \tag{5.19}$$

The following two propositions characterize the approximation error and convergence order of the truncation approximation (5.19). Their proofs can be found in Chapter 2 of [27].

Proposition 5.4. *The error in approximating $\hat{f}_X(x)$ (5.18) by the partial sum $\bar{f}_X(x)$ (5.19) is bounded by the sum of absolute values of all the neglected coefficients. That is,*

$$\left|\hat{f}_X(x) - \bar{f}_X(x)\right| \leq \sum\nolimits_{k=K+1}^{\infty} |\hat{F}_k|, \quad x \in [a, b].$$

Proposition 5.5. *If $\hat{f}_X$ is infinitely differentiable with nonzero derivatives on $[a, b]$, its Fourier-cosine series expansion has geometric convergence, i.e.,*

$$\hat{F}_k \sim O(k^{-n} \exp(-\gamma k)),$$

where γ is determined by the location in the complex plane of the singularities nearest to the expansion interval, and n is determined by the type and strength of the singularity. Otherwise, the convergence of Fourier cosine series is algebraic, i.e.,

$$\hat{F}_k \sim O(1/k^n),$$

where n is at least as large as the highest order of derivative that exists or is nonzero.

Proposition 5.5 implies that the right (left) truncation bound b (a) should not be too large (small). Otherwise, $\hat{f}_X$ may have zero derivatives, and the accuracy of the truncation approximation may be decreased. We suggest to check the tails of the original density function before selecting the truncation interval $[a, b]$.

The key procedure to set up the truncation approximation (5.19) is to calculate the coefficients $\hat{F}_k$, which can be calculated with the FFT algorithm as follows (see e.g. [86] for a theoretical derivation).

Algorithm 5.1 (The FFT algorithm for $\hat{F}_k$)**.**

1. *Calculate the value $\mathbf{Y} = (Y_0, \ldots, Y_K)$ of the distorted density function on the interpolation points $\mathbf{y} = (y_0, \ldots, y_K)$, i.e.,*

$$\mathbf{Y} = \Psi'\left(F_X(\mathbf{y})\right) f_X(\mathbf{y}), \quad \text{with } y_k = a + (b-a)k/K, \quad k \in \{0, 1, \ldots, K\}.$$

2. *Extend $\mathbf{Y}$ to $\tilde{\mathbf{Y}} = (Y_0, Y_1, \ldots, Y_{k-2}, Y_{K-1}, Y_K, Y_{K-1}, Y_{K-2}, \ldots, Y_2, Y_1)$.*

3. *Apply the FFT algorithm to $\tilde{\mathbf{Y}}$, and denote the Fourier transform of $\tilde{\mathbf{Y}}$ by $\bar{\mathbf{Y}}$, a $2K$-dimensional vector.*

4. *Then the coefficients are given by $\hat{F}_k = \bar{Y}_k$ if $k = 1, \ldots, K-1$; $\hat{F}_0 = 0.5\bar{Y}_0$; $\hat{F}_K = 0.5\bar{Y}_K$.*

After (5.19) is set up, the distorted expectation $\mathbb{E}_\Psi[X]$ can be approximated by

$$\mathbb{E}_\Psi[X] = \int_{-\infty}^{\infty} x \tilde{f}_X(x)\, \mathrm{d}x \approx \int_a^b x \bar{f}_X(x)\, \mathrm{d}x$$

$$= \int_a^b x \sum_{k=0}^{K} \hat{F}_k \cdot \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x = \sum_{k=0}^{K} \hat{F}_k V_k, \tag{5.20}$$

with

$$\begin{aligned} V_k &= \int_a^b x \cos\left(k\pi \frac{x-a}{b-a}\right) \mathrm{d}x = \frac{(b-a)^2}{(k\pi)^2}(\cos(k\pi) - 1), \quad 1 \le k \le K; \\ V_0 &= 0.5(b^2 - a^2). \end{aligned} \tag{5.21}$$

We call the truncated approximation (5.20) the T-COS approximation for the distorted expectation $\mathbb{E}_\Psi[X]$.

The truncation approximation (5.19) can be easily extended to calculate the distorted expectation of a combination of risk factors $(X_1, \ldots, X_n)$ with joint characteristic function $\phi(\boldsymbol{\omega})$,

$$\phi(\boldsymbol{\omega}) = E\left[e^{i(\omega_1 X_1 + \cdots + \omega_n X_n)}\right], \quad \text{where } \boldsymbol{\omega} = (\omega_1, \ldots, \omega_n).$$

In fact, the characteristic function of a portfolio $X = z_1 X_1 + \cdots + z_n X_n$ is

$$\phi_X(z) = E\left[e^{iz(z_1 X_1 + \cdots + z_n X_n)}\right] = \phi(\boldsymbol{\omega}), \quad \text{with } \boldsymbol{\omega} = (zz_1, \ldots, zz_n). \tag{5.22}$$

Hence, the density function and the distribution function can be recovered from the characteristic function ϕ_X with the COS method (see Section 5.2.3 or [140, 164]).

5.4 Numerical examples

The Heston model [103] is one of the most popular stochastic volatility models in the equity market, where the equity price $(S_t)_{0 \le t \le T}$ is modeled by

$$\begin{cases} \mathrm{d}S_t = rS_t\,\mathrm{d}t + \sqrt{v_t}S_t\left(\sqrt{1-\rho^2}\,\mathrm{d}W_t^{(1)} + \rho\,\mathrm{d}W_t^{(2)}\right), & S_0 > 0, \\ \mathrm{d}v_t = \kappa(\theta - v_t)\,\mathrm{d}t + \sigma\sqrt{v_t}\,\mathrm{d}W_t^{(2)}, & v_0 = \sigma_0^2 > 0, \end{cases} \tag{5.23}$$

where $W^{(1)}$ and $W^{(2)}$ are independent Brownian motions defined on $(\Omega, \mathcal{F}, \mathbb{F}, P)$[1]. Although the Heston model does not admit the analytical density function or distribution function, the log-asset price has the following characteristic function [103]

$$\phi(\omega, t) := E\left[\exp\left(i\omega \log(S_t)\right) \mid S_0, v_0\right] = \exp(A + B + C), \tag{5.24}$$

[1] Obviously, the Heston model formulated as (5.23) is equivalent to the form (4.23). To make it easier to read, we present this model again in this section. This model will be used in Section 6.3 and Section 7.4.1.2 where it will be presented again for the convenience in reading the text.

where

$$
\begin{aligned}
A &= i\omega\left(\log(S_0) + (r-q)t\right),\\
B &= \theta\kappa\sigma^{-2}\left((\kappa - \rho\sigma\omega i + d)\,t - 2\log\left(\left(1 - ge^{dt}\right)/(1-g)\right)\right),\\
C &= v_0\sigma^{-2}(\kappa - \rho\sigma\omega i + d)(1-e^{dt})/(1-ge^{dt}),\\
d &= \sqrt{(\rho\sigma\omega i - \kappa)^2 + \sigma^2(\omega i + \omega^2)},\\
g &= (\kappa - \rho\sigma\omega i + d)/(\kappa - \rho\sigma\omega i - d).
\end{aligned}
$$

Given the model parameters

$$
\begin{aligned}
&\kappa = 1.15, \quad \theta = 0.0348, \quad \sigma = 0.39, \quad r = 0.04,\\
&T = 1, \qquad v_0 = 0.0348, \quad S_0 = 10,
\end{aligned}
\tag{5.25}
$$

we first recover the density function $f_{\log S_1}$ and the distribution function $F_{\log S_1}$ from the characteristic function (7.17) with the COS method. The truncation interval is set to be $[-10, 5]$, and 2^{10} terms are used in the approximation formula for the density function. $f_{\log S_1}$ and $F_{\log S_1}$ work as the original density function and the distribution function, respectively. Sequentially, due to

$$
F_{S_1}(x) = F_{\log S_1}(\log x) \quad \text{and} \quad f_{S_1}(x) = \frac{1}{x} f_{\log S_1}(\log x), \tag{5.26}
$$

we can calculate the density function f_{S_1} and the distribution function F_{S_1} of S_1.

The truncation approximation (5.19) is set up for the distorted density function induced by the distortion functions (5.6) and (5.7) with $\lambda = 1$, respectively. The truncation interval for S_1 is set to be $[0.01, 40]$, and the number (K) of the terms in (5.19) varies among $\{2^5, 2^6, 2^8\}$. The computation time and the values of the distorted expectations are reported in Table 5.1. On the other hand, ten million paths of the Heston model are simulated with the balanced Milstein scheme [113] at the step-size 0.01. The simulation costs about 37 *seconds*. The ordered sample of the terminal value can be saved and reused to calculate distorted expectations under different distortion functions. However, it still costs about 0.1 *seconds* to calculate the distorted expectations with the saved paths, while the T-COS method costs less than 0.1 *milliseconds*. It means that the T-COS method is still far more efficient than the MC method with saved paths, when the distortion function or its parameters vary. When the number (K) of terms in the truncation approximation (5.19) increases, the increase in computation time is negligible[2]. Since the values of the distorted expectations corresponding to $K = 2^6$

[2]When the truncation interval for S_1 is enlarged to be $[0.01, 50]$, the changes in computation time and the distorted expectations are also negligible. We do not report them here.

and $K = 2^8$ are the same up to the current accuracy, we recommend to set $K = 2^6$ in this setting. In other applications of the T-COS method, one may do the same test to choose an appropriate value for K. Since the T-COS method and the MC method provide quite similar values for the distorted expectation, the T-COS method is recommended for its extraordinary efficiency. The implementations are done in MATLAB (2014b) (Processor: Intel Core(TM) i7-3770 CPU @ 3.4GHz, RAM: 8GB).

distortion	method		time	DE
MAXMIN (5.6)	MC		37 s	8.105682
	T-COS	$K = 2^5$	0.04 ms	8.104594
		$K = 2^6$	0.05 ms	8.105569
		$K = 2^8$	0.07 ms	8.105569
MINMAX (5.7)	MC		0.1 s †	7.588375
	T-COS	$K = 2^5$	0.04 ms	7.587122
		$K = 2^6$	0.04 ms	7.588247
		$K = 2^8$	0.07 ms	7.588247

†Calculated with the saved sample in the increasing order

Table 5.1: Distorted expectations $\mathbb{E}_{\Psi^1}[S_1]$ (DE)

To show the effect of the distortion function, we plot the distorted expectations under the distortion functions (5.6) and (5.7) with different values of λ. As shown in Figure 5.1, the distorted expectation decreases with the distortion parameter λ. The higher the value of λ is, the higher the stress level is. Note that different distortion functions lead to different approximations for the distorted expectation, especially at high stress levels.

5.5 Discussion

In this chapter, we proposed an analytical approximation, as an alternative to the Monte Carlo method, for the distorted expectations in the setting where the density function or the characteristic function of the risk factor is given in its analytical form.

The proposed method, the so-called T-COS method, is based on a truncated sum of the Fourier-cosine series expansion for the distorted density function on a finite interval. The coefficients of the terms in the truncation approximation can be calculated with Algorithm 5.1 based on the FFT algorithm. Then, the resulting approximation for the distorted density function provides an analytical approximation for the distorted expectation.

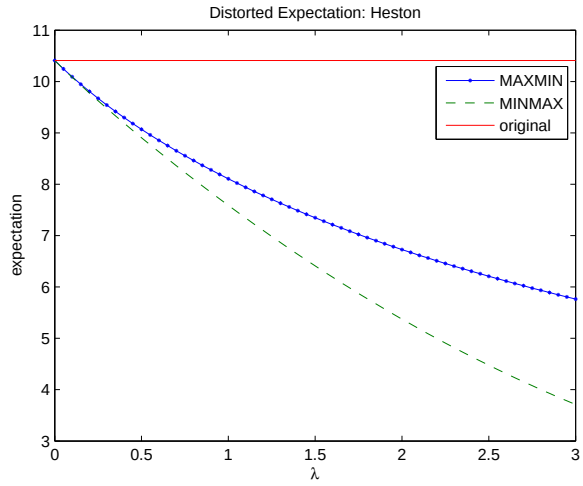


Figure 5.1: Distorted expectations $\mathbb{E}_{\Psi^\lambda}[S_1]$ associated with MINMAX and MAXMIN distortion functions at different stress levels λ.

Numerical examples show that the T-COS method is about 2500 times more efficient than the standard Monte Carlo method in the Heston model. This comparison is based on the saved and ordered pathes from the Monte Carlo simulation, and the assumption that the density function is known in an analytical form. From this point of view, the advantage of the T-COS method in efficiency is not restricted to the Heston model, but it holds in the general setting.

Distorted expectations are associated with distortion risk measures, which play an important role in risk management. The T-COS method can be used to quantify the risk of portfolios, as well as a single risk factor, with the risk measures beyond V@R and Expected Shortfall [140]. They also arise in portfolio optimization, conic finance, behavioural finance and so on. The problems involving distorted expectations in these areas can be solved with the T-COS method under realistic model assumptions.

6 Weighted Monte Carlo method and its application

As presented in the previous chapters, model uncertainty is ubiquitous whenever a parametric model is used in pricing and hedging of derivatives. In this chapter, we empirically confirm that the weighted Monte Carlo method can reduce the impact of model uncertainty on a derivative price.

6.1 Introduction

The current price of a risky asset should reflect the market belief (a sum of participants' subjective believes) about its future price performances. Economically, competitive market prices keep the market as a whole in the state of maximum uncertainty toward its future performances. Otherwise, it implies that some useful information could be utilized and the price will move up or down until the uncertainty is at its maximum. Since entropy could work as an uncertainty index of a random variable, there is an equivalence between maximizing entropy of the future price of a risky asset and maximizing market uncertainty towards its future price. Only if the market is efficient enough, calculating the maximum-entropy distribution is an alternative method for constructing risk-neutral

probabilities even in a incomplete market. This is the basic intuition of the entropy pricing theory which coincides with Efficient Market Hypothesis. For more details, we refer to [92] and references therein.

In the simulated market setting, searching for the maximum-entropy distribution reduces to assigning a probability to each simulated path. The weighted Monte Carlo (WMC) method is one of the ways to select such an entropy-based probability measure by minimizing relative entropy of the target probability measure with respect to a prior probability measure. This problem can be formulated as a convex optimization problem subject to some market constraints. When the prior probability measure is given by a uniform distribution, the optimization problem of minimizing relative entropy is equivalent to maximizing entropy. In other cases, the WMC method also has its economic grounds. There exists a one-to-one correspondence between the calibration of a model starting with a prior probability measure and using a penalization function on the space of probabilities and the calculation of state-prices via utility maximization [9]. Friedman et al. [83] proposed some new instances of the WMC method, employing relative $(U, \mathcal{O})$-*entropy*, a generalization of the relative entropy, as the measure of discrepancy between probability measures[1].

Under the assumption that the prior estimation remains fixed, its sensitivity to the benchmark prices has been investigated in [9]. In this chapter, we will fix the benchmark prices, and use different models to simulate the paths of the underlying asset. That is, we empirically investigate whether the weighted Monte Carlo can reduce the impact of model uncertainty in terms of the implied volatility of vanilla options and the prices of exotic options.

6.2 Entropy and the WMC method

Given the density function (g) of a random variable X, its uncertainty can be measured by the entropy $h(X)$ given by

$$h(X) = -\int_{-\infty}^{\infty} g(x) \log g(x) \, \mathrm{d}x. \tag{6.1}$$

Kullack–Leibler (KL) relative-entropy works as a measure of the similarity of two probability measures. The KL relative-entropy $H(\mathbb{Q} \mid \mathbb{P})$ of a probability measure $\mathbb{Q}$

[1] The generalized instances require to solve the corresponding connecting equation and are computationally more expensive. So we will use the instance based on relative entropy to investigate the objectives of this chapter.

with respect to $\mathbb{P}$ is defined by

$$H(\mathbb{Q} \mid \mathbb{P}) = \int_{-\infty}^{\infty} q(x) \log \frac{q(x)}{p(x)} \, \mathrm{d}x, \tag{6.2}$$

where p and q are the density functions of $\mathbb{P}$ and $\mathbb{Q}$, respectively.

From the information theory point of view, entropy can be used to select an Arrow–Debreu probability of a risky asset, either in a model-independent way or model-dependent way. There is a lot of literature on how to derive the maximum entropy distribution in a model-independent way (see, e.g. [25, 30, 135, 147]). Here we focus on the model-dependent way, especially in the setting of Monte Carlo simulation.

Usually the price of an underlying asset of a derivative is modeled by a stochastic differential system or a stochastic differential equation (SDE). Before the model can be used to price a target financial product, model parameters have to be estimated by statistical methods or the model has to be calibrated to the market data. In fact, once the model parameters are determined, a probability distribution of the underlying asset price is determined simultaneously. If one is not quite confident with the current estimations, adjustments could be made such that the estimations coincide with the market data and the new probability measure is 'close' to the old one. So, one may minimize KL entropy of a probability measure with respect to the prior probability measure to calibrate the model to the market data. Here the market data is used twice, once to estimate the prior probability and once to adjust it.

Assume that the price of an underlying asset follows the SDE

$$\frac{\mathrm{d}S_t}{S_t} = \sigma_t \, \mathrm{d}W_t + r \, \mathrm{d}t, \quad t \in [0, T], \tag{6.3}$$

where r is the risk-neutral drift and $(W_t)_{0 \le t \le T}$ is a Brownian motion defined on a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \le t \le T}, \mathbb{P})$. If we assume r is known and $\sigma_t \in [\sigma_{min}, \sigma_{max}]$, there exists volatility uncertainty in the model [11]. Any specification of σ determines a probability distribution, under which expectations are denoted by $\mathbb{E}^{\sigma}[\cdot]$. In this case, the market constraints can be written as

$$\mathbb{E}^{\sigma}[f_i(S_T)] = C_i, \quad i = 1, 2, \ldots, M,$$

where f_i and C_i $(i = 1, 2, \ldots, M)$ represent the discounted payoff functions of the market traded benchmark instruments and their prices, respectively. The objective of minimizing relative entropy is essentially equivalent to minimizing the functional

$$\mathbb{E}^{\sigma} \left[\int_0^T \eta(\sigma_t^2) \, \mathrm{d}t \right],$$

where η is a strictly convex function which vanishes at the volatility of the prior distribution [10, 149]. This is an example of the application of entropy in model calibration. In fact, entropy has been extensively used in model calibration (see, e.g. [50, 85]).

On the other hand, given a personal prior estimation on a model for $(S_t)_{0\le t\le T}$, for example, model (6.3), we can simulate its paths. Classical Monte Carlo methods assign the same weight to each simulated path while the WMC method assigns a special weight to each simulated path. That is, the sample of $(S_t)_{0\le t\le T}$ is uniformly distributed in the classical Monte Carlo setting while it follows a special distribution characterized by a probability distribution p in the WMC setting. Together with the market constraints, p can be selected by minimizing the discrete version of (6.2)

$$H(p \mid \hat{p}) = \sum_{i=1}^{N} p_i \log \frac{p_i}{\hat{p}_i}, \tag{6.4}$$

where N is the number of the simulated paths and $\hat{p}$ is a prior probability distribution of the sample. The resulting probability can be taken as an Arrow–Debreu probability. The Arrow–Debreu prices (state prices) coincide with the marginal utilities for consumption obtained by maximizing the expectation of the utility function $U(x) = -\exp(-\alpha x)$ $(\alpha > 0)$ by investing in a portfolio of the benchmark instruments [9].

Problem 1 (the WMC method [83])**.** *Assume there are a number* (N) *of simulated paths of a model for the underlying asset and a set* (M) *of benchmark European options with discounted payoff function* f_j, $j = 1, 2, \ldots, M$. *The bid and ask prices of these benchmark options are denoted by positive vectors* $b = (b_1, b_2, \ldots, b_M) \in \mathbb{R}_+^M$ *and* $a = (a_1, a_2, \ldots, a_M) \in \mathbb{R}_+^M$, *respectively. The WMC method can be formulated as follows*

$$(p^*, s^*) = \arg\min_{p,s} \left\{ H(p \mid \hat{p}) + \frac{d}{2} s's \right\}$$

subject to

$$\begin{aligned}
&b_j \le \mathbb{E}_p[f_j] \le a_j, && j = 1, 2, \ldots, M,\\
&\mathbb{E}_p[f_j] - \frac{b_j + a_j}{2} = s_j, && j = 1, 2, \ldots, M,\\
&\sum_{i=1}^{N} p_i = 1,\\
&p_i \ge 0, && i = 1, 2, \ldots, N,
\end{aligned}$$

where $\mathbb{E}_p[f_j] = \sum_{i=1}^{N} p_i f_j(S_T^i)$ *and* (b_j, a_j) *is the bid–ask price pair of the* j*-th benchmark option.* S_T^i *is the price of the underlying asset on the* i*-th path at* $t = T$*.* $\hat{p}$ *can be chosen as* $1/N$*, that's, the prior distribution of the sample is a uniform distribution. The objective function not only minimizes the relative entropy but also the difference* $s = (s_1, s_2, \ldots, s_M) \in \mathbb{R}^M$ *between the model prices and the market mid-prices of the benchmark options. The factor* d *works as a penalty factor.*

Due to bid–ask spreads in the market data, the conditions of the optimization problem only require the model prices of benchmarks to lie in the interval of the corresponding bid and ask prices. In addition, there exists a simulation error when we discretize a stochastic model. So, a weighted Monte Carlo calibration can lead to the maximum entropy distribution which can work as a risk-neutral probability, but it need not to insure that the re-weighted paths are completely consistent with a martingale measure. If the bid–ask spreads are very large, synthetic securities could be added to the benchmark data to enforce the martingale condition [70].

Even if there are a small number of benchmark options, it is not easy to directly solve Problem 1 when thousands of paths are simulated. It can be solved through its dual problem, which can be formulated as

Problem 2 (dual problem [83])**.** *The dual problem of Problem 1 is to find*

$$(\alpha^*, \beta^*, \rho^*) = \arg \max_{\alpha \geq 0, \beta \geq 0, \rho} \left\{ \beta' b - \alpha' a - \rho' m - \log(Z(\alpha, \beta, \rho)) - \frac{1}{2d} \rho' \rho \right\}$$

where

$$Z(\alpha, \beta, \rho) = \sum_{i=1}^{N} \hat{p}_i \exp((\beta - \alpha - \rho)' \bar{f}_i) \tag{6.5}$$

and $\bar{f}_i = \big(f_1(S_T^i), f_2(S_T^i), \ldots, f_M(S_T^i)\big) \in \mathbb{R}_+^M$*.* α*,* β*,* ρ *are* M*-dimensional column vectors. Once we solve this problem, the solution of Problem 1 is given by*

$$p_i^* = \frac{\hat{p}_i \exp((\beta^* - \alpha^* - \rho^*)' \bar{f}_i)}{Z(\alpha^*, \beta^*, \rho^*)}, \quad i = 1, 2, \ldots, N. \tag{6.6}$$

Problem 2 can be easily solved with the optimization toolbox in Matlab or other softwares. Now we are ready to price a target option with the simulated paths and the probability measure p^*.

6.3 Numerical Study

In this section, we work with the Heston model [103], in which the asset price process $(S_t)_{0\le t\le T}$ and the variance process $(v_t)_{0\le t\le T}$ follow the following risk-neutral processes respectively,

$$\begin{cases} \mathrm{d}S_t = S_t(r-q)\,\mathrm{d}t + \sqrt{v_t}S_t\,\mathrm{d}W_t, & S_0 > 0, \\ \mathrm{d}v_t = \kappa(\eta - v_t)\,\mathrm{d}t + \lambda\sqrt{v_t}\,\mathrm{d}\widetilde{W}_t, & v_0 = \sigma_0^2 > 0, \end{cases} \tag{6.7}$$

where $(W_t)_{0\le t\le T}$ and $(\widetilde{W}_t)_{0\le t\le T}$ are correlated Brownian motions defined on a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$. The correlation coefficient between W and $\widetilde{W}$ is ρ, i.e. $Cov(\mathrm{d}W_t, \mathrm{d}\widetilde{W}_t) = \rho\mathrm{d}t$. The parameter η is the long-term average variance, while κ is the speed of the mean-reversion of the variance. The parameter λ is referred to as the volatility of variance since it scales the diffusion term of the variance process. Many numerical methods have been proposed to simulate the variance process (see, e.g. [161]).

Calibration methods for the Heston model have been extensively investigated in [91]. Generally speaking, calibration is referred to minimizing the discrepancy $f(\{P_i\}, \{\hat{P}_i\}, A)$ between market prices P_i of benchmark instruments and the model prices $\hat{P}_i$ with model parameter set A. In practice, it is difficult to solve the optimization problem of calibrating a parametric model to the market data, because the discrepancy function may have some local minima. To overcome this problem, we may use some global optimization algorithms, such as genetic algorithms, simulated annealing and so on. Recently, Yang and Lee [180] proposed an efficient global calibration method with enhanced discrete local search strategy. Assuming an efficient global optimization method has been chosen, we are still confronted with model uncertainty if different discrepancy functions are used in the calibration procedure [91]. Guillaume and Schoutens [91] used three functions $(\mathrm{RMSE}, \mathrm{APE}, \mathrm{ARPE})$ to measure the discrepancy and calibrated the Heston model to the market data with a fully free parameter set $\{v_0, \kappa, \eta, \lambda, \rho\}$ and five $\left\{\mathrm{EWMA}, \mathrm{MW}(T^{\mathrm{VIX}} = 0.5, 3, 5), \mathrm{VIX}\right\}$ reduced parameter sets $\{\kappa, \lambda, \rho\}$. That is, eighteen methods were used in total. In the Appendix of this chapter, we will present two instances in which RMSE and ARPE work as the discrepancy functions accompanied with a fully free parameter set. We refer to [91] for more details of these eighteen calibration methods.

Firstly, we test the sensitivity of the WMC method with respect to the prior information of model parameters. The calibrated models in [91] can be taken as different sets of prior model information. Then we will check to what extent the WMC method

can reduce the calibration risk when pricing exotic options. The calibration risk can be measured by the global risk measure defined by

$$\frac{\max P_i - \min P_i}{\frac{1}{K}\sum_{i=1}^{K} P_i},$$

where P_i $(i = 1, \ldots, K)$ are the model implied target option prices with different calibration methods. Except for the quantitative standard, we can also get the first impression of the impact of model uncertainty from the figure of the model-based prices or their boxplot.

6.3.1 Sensitivity of the WMC method to the prior estimation

In general, all the eighteen methods in [91] can calibrate the Heston model well from the implied volatility point of view. But comparatively speaking, some of the calibrations are not good enough. We present the calibration results on two particular quoting days. The date $31/10/2006$ did not experience high market volatility while the date $11/12/2008$ was experiencing the subprime mortgage crisis.

We simulated 10 000 paths of the calibrated models (eighteen cases) in [91] on the chosen days using the Euler method for the price process and the balanced Milstein method for the variance process (see the Appendix at the end of this chapter for more information on simulation methods for the Heston model). The simulated period starts from the quoting date to the longest expiration date of the benchmark options and the stepsize is $1/365$. Then, we calculated the implied volatilities of the benchmark European calls with the simulated paths. They can be regarded as MC-implied volatilities. On the other hand, we re-calibrated the simulated paths with the WMC method and calculated the WMC-implied volatilities. In the procedure of re-calibrating, we didn't add synthetic securities to the market traded benchmarks, because the chosen benchmarks have quite small bid–ask spreads relative to their mid-prices. The results are shown in Figure 6.1 and Figure 6.2.

Figure 6.1(a) suggests that the combination of RMSE and EWMA leads to a relatively larger implied volatility error, although the parameter set is the optimal one calculated by this combination. The implied volatility errors from other combinations seem very small. But Figure 6.1(b) shows that the WMC method can reduce the implied volatility error. It can even correct the implied volatility from the parameter set of the combination of RMSE and EWMA to a sound level. It is also the case in other parameter

settings, as shown in Figure 6.1 and Figure 6.2. The most important observation is that the WMC-implied volatilities under different parameter settings closely gather at the market quoted implied volatility except the most deep out-of-money and in-the-money call options as shown in Figure 6.1 and Figure 6.2. These empirical results confirm that the WMC method is not sensitive to the prior estimation in terms of implied volatility error.

On the other hand, since the option price is a non-linear function of the implied volatility, a small deviation of the implied volatility from the market quoted implied volatility means that the model prices may be out of the corresponding market bid–ask spread. When 10 000 paths are used to price benchmarks with the classical MC method, the number of benchmark options whose model prices lie out of the bid–ask spread is more than half of the total number of benchmark options. Even when one million paths are used to price the benchmarks with the classical MC method, the number of benchmark options whose model prices lie out of the bid–ask spread is still more than its counterpart calculated by the WMC method with 10 000 paths. In fact, the formulation of the WMC method guarantees that the re-calibrated model is more market-consistent than the classical MC method. The numerical results also confirm that.

One may argue that the implied volatility error of the MC method may be reduced when more paths and kinds of control variate techniques are employed. Even if these strategies could reduce the implied volatility error, simulating millions of paths is much more time-consuming, compared with the time to solve the optimization problem for the WMC method. When we use 18 call options as benchmarks and simulate 10 000 paths for the calibrated models on $31/10/2006$, the optimization problems in the WMC method can be solved in about 1.5 seconds with the help of the optimization toolbox in Matlab(®$2012a$) (Processor: Intel Core(TM) i7-3770 CPU @ 3.4GHz). But it costs about 60 seconds to simulate and store one million paths. Here we want to point out that given the number of the benchmarks, the number of the decision variables in Problem 2 is then fixed. Compared with the time to find the direction of updating the variables in Problem 2, the time to calculate (6.5) and (6.6) is negligible. So, more simulated paths do not add a lot of computational burden. Even when 100 000 paths are used, the optimal weight can be found in comparable time as that used by the model with 10 000 paths. But additional paths add quite little extra value in reducing the implied volatility error. From this point of view, the WMC method is computationally much cheaper.

We don't make the comparison between the multi-level Monte Carlo method and the weighted Monte Carlo method in terms of the implied volatility error in this chapter.

The multi-level Monte Carlo method relies on the payoff function of the target payoff. It is useless to store the multi-level paths when pricing European options, because the simulated multi-level paths cannot be used to price exotic options, such as lookback options and Asian options. But the calibrated model by the WMC method can be directly used to price these exotic options.

6.3.2 Pricing exotic options with the WMC method

Since the WMC method can generate very similar implied volatilities under different prior information settings, we are wondering whether it can lead to similar prices of exotic options, i.e., whether it can reduce the impact of model uncertainty. To investigate this problem, we use the re-calibrated model to price Asian options, one-touch barrier options and lookback options in the Heston model.

The payoffs of continuous arithmetic average Asian calls and puts are given by

$$\hat{f}_1(S) = \left(\frac{1}{T} \int_0^T S(t)dt - K \right)^+ \quad \text{and} \quad \hat{f}_2(S) = \left(K - \frac{1}{T} \int_0^T S(t)dt \right)^+.$$

The discrete versions are given by

$$f_1(S) = \left(\frac{1}{N} \sum_{i=1}^{N} S\left(t_i\right) - K \right)^+ \quad \text{and} \quad f_2(S) = \left(K - \frac{1}{N} \sum_{i=1}^{N} S\left(t_i\right) \right)^+.$$

Under the risk-neutral measure $\mathbb{Q}$, the prices of the Asian call and put can be calculated by $\exp(-rT)\mathbb{E}^{\mathbb{Q}}[f_i]$ for $i = 1, 2$. We monitored at 10 time points between $31/10/2006$ and its expiration date ($T = 781/365$ year) and 14 time points from $11/12/2008$ to its expiration date ($T = 737/365$ year)).

The initial prices of the lookback call and put options are given by

$$\text{LC} = \exp(-rT)\mathbb{E}^{\mathbb{Q}}\left[(S_T - m_T^S)^+\right] \quad \text{and} \quad \text{LP} = \exp(-rT)\mathbb{E}^{\mathbb{Q}}\left[(M_T^S - S_T)^+\right],$$

respectively where m_t^S and M_t^S denote the minimum and maximum processes of the process $(S_t)_{0 \le t \le T}$, i.e.

$$m_t^S = \inf\left\{S_h, 0 \le h \le t\right\} \quad \text{and} \quad M_t^S = \sup\left\{S_h, 0 \le h \le t\right\}.$$

The payoff of a one-touch barrier option depends on whether the underlying stock price reaches the barrier H during the lifetime of the option. The prices of down-and-in put (DIBP) and up-and-in call (UIBC) are given by

$$\text{DIBP} = \exp(-rT)\mathbb{E}^{\mathbb{Q}}\left[(K - S_T)^+ \mathbf{1}(m_T^S \le H)\right]$$

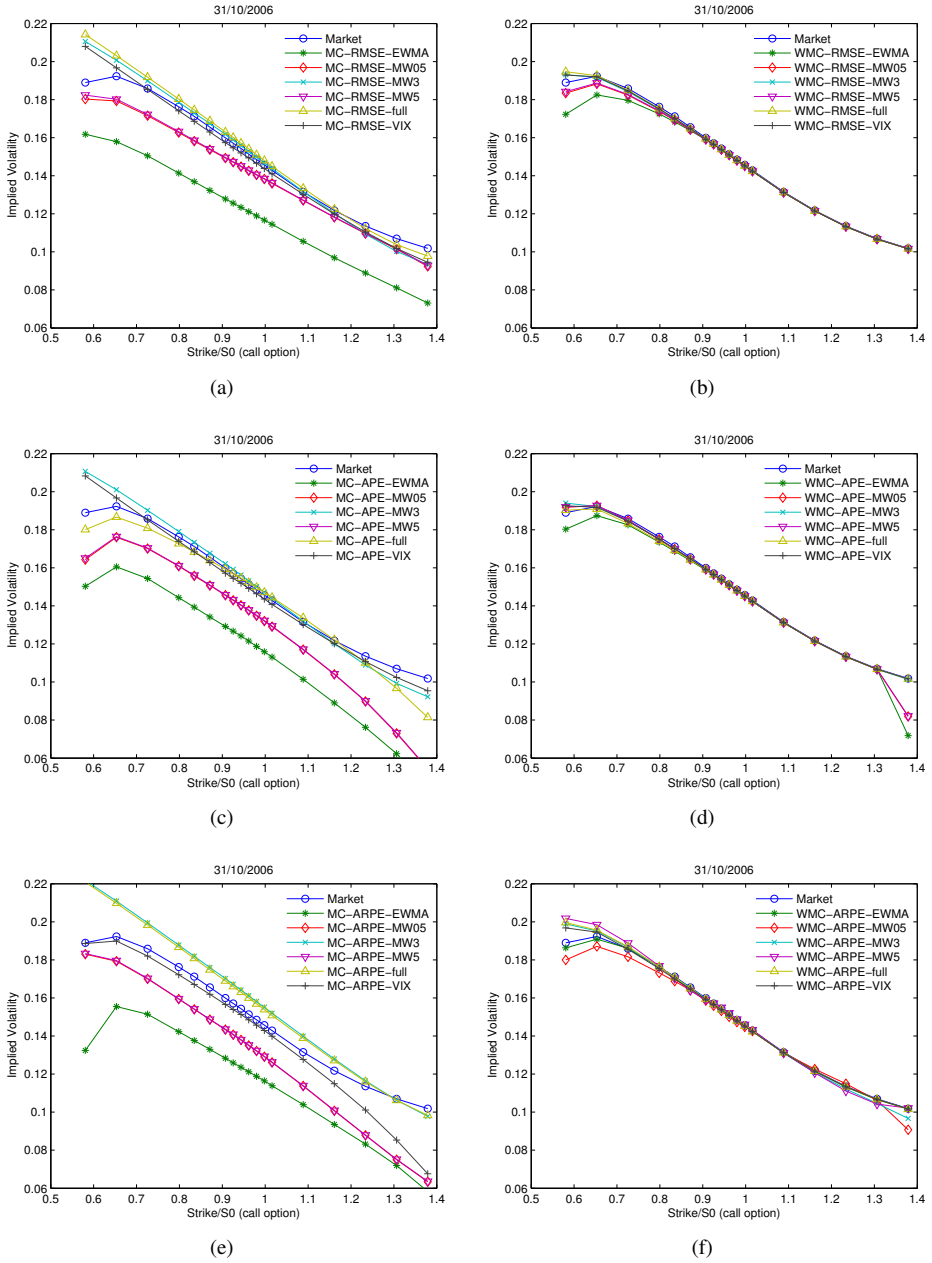


Figure 6.1: Model implied volatility and market quoted volatility of call option expired on 20/12/2008 *and quoted on* 31/10/2006

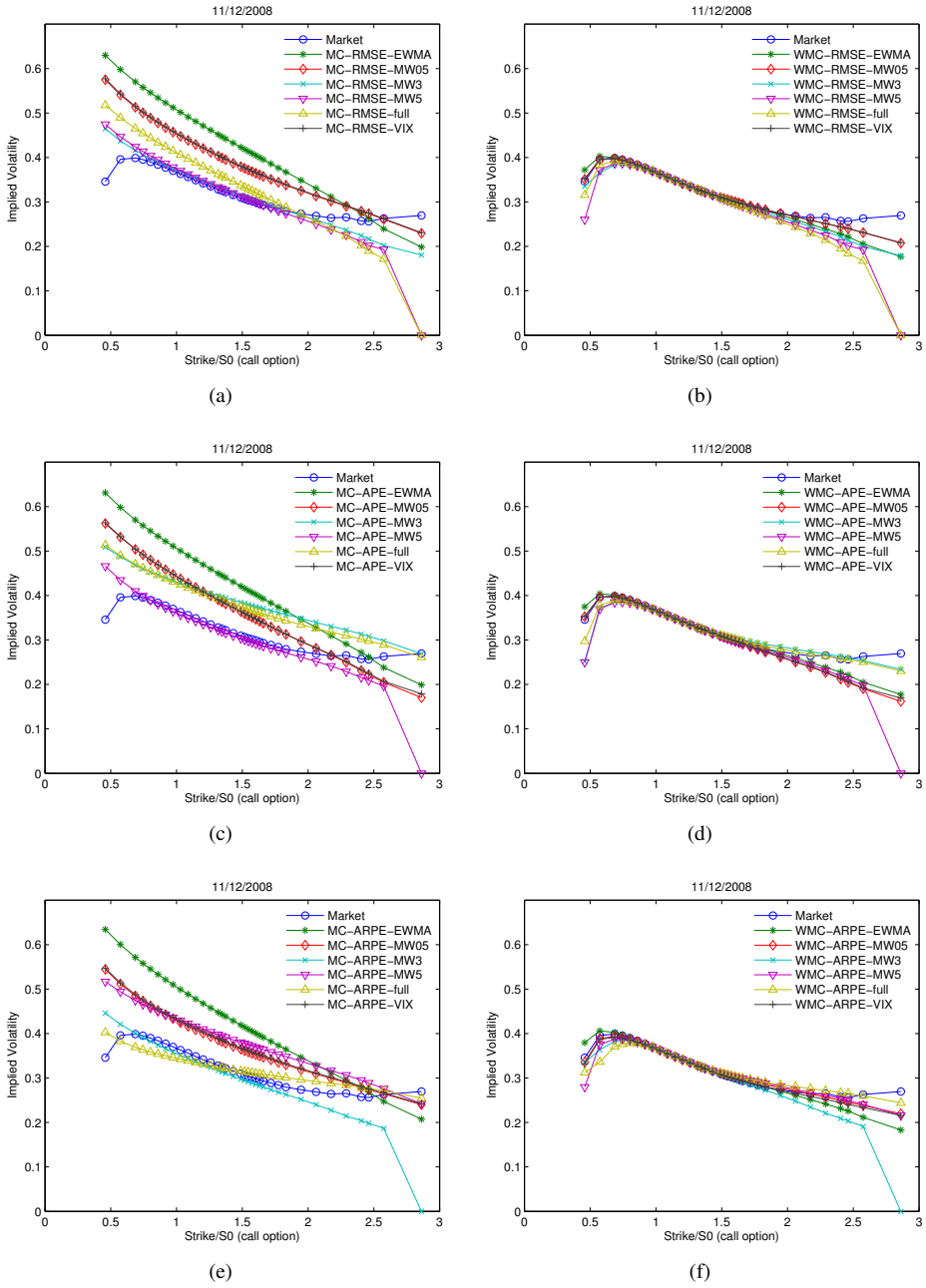


Figure 6.2: Model implied volatility and market quoted volatility of call option expired on 18/12/2010 *and quoted on* 11/12/2008

and

$$\text{UIBC} = \exp(-rT)\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+ \mathbf{1}(M_T^S \geq H)\right].$$

Figure 6.3 suggests that the WMC method can reduce the calibration risk when pricing Asian options on the two quoting dates $31/10/2006$ and $11/12/2008$. But compared with Figure 6.4, its advantage is not very big in this case.

Figure 6.4 shows that the WMC method performs definitely better than the classical MC method when pricing one-touch options, especially on $11/12/2008$. The boxplots in Figure 6.5 also show the capability of the WMC method in reducing the impact of model uncertainty on the value of lookback options.

To quantify to what extent the WMC method can reduce the impact of model uncertainty, we calculate the global risk for each option. The results can be found in Table 6.1, Table 6.2 and Table 6.3. Even when the global risks from the WMC method are relatively larger than those from the MC method in some cases of Asian call options on $31/10/2006$, they are all less than 1.1%. Except for these Asian call options, the Asian put with strike 1000 on $31/10/2006$ is the only case where the global risk from the WMC method is relatively bigger than that from the MC method. But both of them are larger than 100%. From Figure 6.3(b), we observe that the prices of the Asian put with strike 1000 are nearly zero, no matter whether the WMC method or the classical MC method is used. In fact, it holds in all of the cases in the present chapter that the impact of model uncertainty is increasing when the option price tends to zero. In practice, such options with price zero may not be traded at all.

If we ignore the cases where the classical MC calibration risk is less than 1%, the WMC method can reduce the calibration risk by 78.00% on average for one-touch barrier options, 38.53% for lookback options and 34.66% for Asian options quoted on $31/10/2006$. These quantities are 94.25%, 24.23% and 31.71% for one-touch barrier options, lookback options and Asian options quoted on $11/12/2008$.

All of these quantities confirm that the WMC method could effectively reduce the impact of model uncertainty on the derivative value, especially for the one-touch barrier options.

31/10/2006	K	1000	1050	1100	1150	1200	1250	1300	1350	1400	1450	1500
Call	WMC	0.005282	0.006047	0.007171	0.008585	0.010703	0.016731	0.029981	0.054505	0.103639	0.206809	0.412960
	MC	0.002054	0.002453	0.003021	0.004631	0.009403	0.019159	0.038691	0.078485	0.155716	0.323787	0.715368
Put	WMC	1.360817	0.954056	0.767202	0.626086	0.477155	0.344500	0.239087	0.153265	0.093550	0.052610	0.019805
	MC	1.241863	1.153246	0.987679	0.797831	0.682998	0.558222	0.436606	0.309671	0.190510	0.102810	0.041221
11/12/2008	K	700	750	800	850	900	950	1000	1050	1100	1150	1200
Call	WMC	0.052270	0.059546	0.076599	0.104413	0.14150	0.196692	0.295800	0.424496	0.591750	0.797755	1.057842
	MC	0.080920	0.118308	0.164853	0.219558	0.280625	0.347511	0.414693	0.524850	0.672888	0.858381	1.127192
Put	WMC	0.250088	0.208505	0.168724	0.143380	0.124273	0.106754	0.088456	0.070464	0.054477	0.040882	0.032284
	MC	0.69214	0.537267	0.409251	0.309706	0.228915	0.165284	0.114959	0.079403	0.057488	0.040412	0.027693

Table 6.1: Global risks of Asian options

UIBC	$H/S0$	0.5	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
31/10/2006	WMC	0.004533	0.011973	0.041121	0.084771	0.186476	0.224005	0.461389	1.363310	1.662961	2.284115
	MC	0.185644	0.204732	0.274577	0.433105	0.716268	1.160372	1.692950	2.102302	2.351611	3.013293
11/12/2008	WMC	0.002338	0.002817	0.004407	0.007989	0.013598	0.025527	0.039225	0.046356	0.063483	0.086354
	MC	0.353711	0.354427	0.356926	0.361537	0.369814	0.382748	0.396988	0.414866	0.438560	0.469529
DIBP	$H/S0$	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45	1.50
31/10/2006	WMC	0.367853	0.282557	0.199333	0.123883	0.079149	0.051471	0.041171	0.019281	0.011086	0.005970
	MC	1.459195	1.276770	1.099600	0.925454	0.762564	0.636381	0.533817	0.464873	0.421702	0.413146
11/12/2008	WMC	0.117173	0.072864	0.037760	0.018642	0.009318	0.003997	0.001950	0.001719	0.001780	0.001805
	MC	0.671605	0.589739	0.518573	0.470262	0.437084	0.417260	0.405307	0.399170	0.397389	0.396842

Table 6.2: Global risks of one-touch options

	LC		LP	
	WMC	MC	WMC	MC
31/10/2006	0.123910	0.197294	0.246191	0.409411
11/12/2008	0.261358	0.329305	0.253876	0.351767

Table 6.3: Global risks of looback call (LC) and put options (LP)

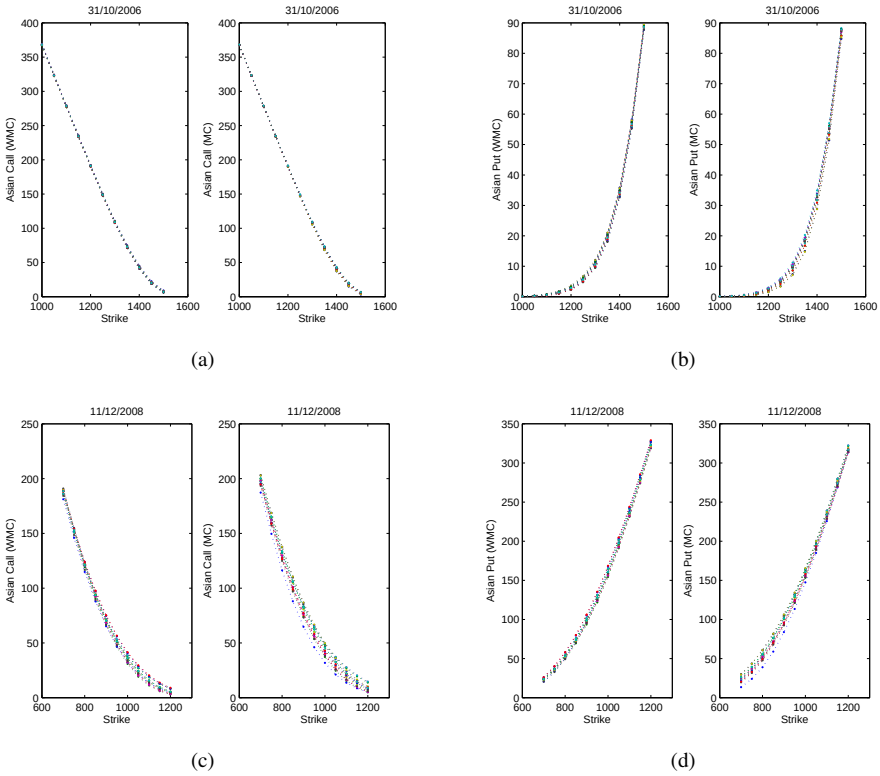
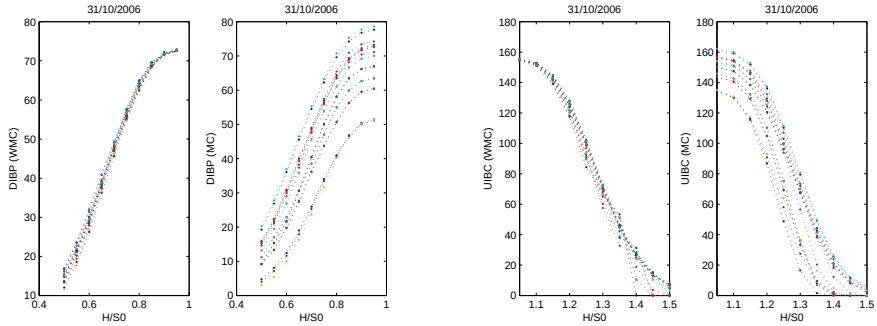
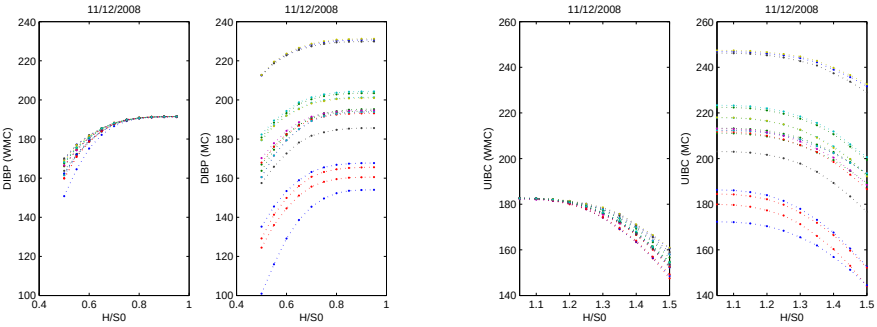


Figure 6.3: Asian option price



(a) (b)



(c) (d)

Figure 6.4: One-touch barrier option

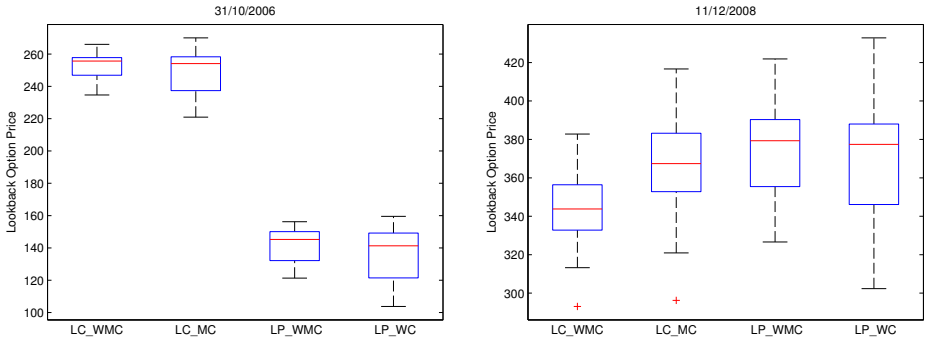


Figure 6.5: Lookback option price

Appendix

1. Model calibration

   Model calibration is a way to select those model parameters $\boldsymbol{\theta}^* = (\kappa^*, \eta^*, \lambda^*, \rho^*, v_0^*)$ that minimize the discrepancy $f\left(\{P_i\}, \{\hat{P}_i\}; \boldsymbol{\theta}\right)$ between the model prices $\hat{P}_i$ and the markets prices P_i of benchmark instruments. In this chapter, the root mean square error (RMSE) is used to measure the discrepancy between the model prices and the market prices of the benchmark options, i.e.,

   $$f\left(\{P_i\}, \{\hat{P}_i\}; \boldsymbol{\theta}\right) = \sqrt{\sum_{i=1}^{M} \frac{\left(P_i - \hat{P}_i\right)^2}{M}}, \tag{6.8}$$

   where M stands for the number of benchmark instruments. Alternatively, this discrepancy can also be measured by the average relative price error (ARPE), i.e.,

   $$f\left(\{P_i\}, \{\hat{P}_i\}; \boldsymbol{\theta}\right) = \frac{1}{M} \sum_{i=1}^{M} \frac{|P_i - \hat{P}_i|}{\hat{P}_i}. \tag{6.9}$$

   Guillaume and Schoutens [91] proposed other methods to measure the discrepancy between model and market prices of SPX options and to estimate model parameters with VIX data. Since different calibration methods lead to different estimations for the model parameters, calibration methods have a significant influence on the exotic option price. Calibration risk can be defined as the different optimal parameter sets arising from the different specifications of the discrepancy measures and the calibration methodology [91].

   In Section 7.4.1.2, we use the calibrated models of [91] with a fully free parameter set and with the RMSE (6.8) respectively ARPE (6.9) as the measure of discrepancy between the model prices and the market prices.

2. Simulation methods for the Heston model

   The variance process is always positive and zero is unattainable if $2\kappa\eta > \lambda^2$. However, even if this condition is satisfied, it is still possible to generate negative values for the variance process when it is discretized with the Euler-Maruyama scheme [185]. The balanced methods [113] can work well if $4\kappa\eta \geq \lambda^2$. If this condition is still not satisfied, many practical methods have been proposed in the literature (see [161] and references therein). In this chapter, the model parameters

satisfy $4\kappa\eta \geq \lambda^2$. Hence, we choose the balanced Milstein scheme to simulate the variance process and the Euler-Maruyama scheme to simulate the price process of the underlying asset. The Monte Carlo simulation is performed with ten million paths and a control variate.

7

Model-based and model-free upper bounds for discrete arithmetic Asian options

Confronted with model uncertainty, one may use the weighted Monte Carlo method to reduce its impact on the derivative value. Another remedy is to avoid using any parametric model. This model-free approach is immune to model uncertainty (see e.g. [19, 40, 53, 104, 165]). In this chapter, we propose an efficient method to calculate the comonotonicity-based value bounds for discrete arithmetic Asian options when the asset price model admits an analytical characteristic function.

The comonotonicity-based value bounds can be calculated either in the model-based framework or in the model-free framework. Quantities in the model-based framework can be compared with their model-free counterparts to detect model mis-specification. We address some practical issues.

7.1 Introduction

Let $S(t)$ be the price of an underlying asset at time t and $r > 0$ the risk-free interest rate, which will be assumed to be constant. The payoff of an arithmetic Asian option matured at T with strike K is then given by

$$\left(\xi \sum_{i=1}^{N} w_i S(t_i) - \xi K\right)^+,$$

where $(x)^+ = \max(x, 0)$ $\sum_{i=1}^{N} w_i = 1$ and $t_1, t_2, \ldots, t_N = T$ are the discrete monitoring times. When $\xi = 1$, it is a call option; when $\xi = -1$ it is a put option. Such options are difficult to be priced in the closed form even in the Black–Scholes market model. Although the Monte Carlo method or PDE-based method can be used to numerically calculate the Asian option price (see, e.g. [29, 172]), both approaches are rather time-consuming. It is a competitive alternative to efficiently calculate the sharp bounds of the option value. In [1, 37], the authors confirmed that the comonotonicity-based upper bound is an accurate approximation for the Asian option value under different parametric models. This value bound actually corresponds to the cheapest static hedging strategy with European options (see, e.g. [1, 39, 40]).

Simon et al. [160] pioneered the comonotonicity-based approach for calculating the value bound of this type of Asian options, and applied this approach in the Black–Scholes setting. Then, Albrecher et al. [1] and Chen and Ewald [37] applied this approach in the setting of exponential Lévy models and stochastic volatility models, respectively. Essentially, a comonotonicity-based value bound is a weighted sum of European option prices matured on each monitoring time with optimal strikes. To calculate the optimal strikes of the hedging instruments, Albrecher et al. [1] numerically built up the distribution function from the density function of the underlying asset price, and its inverse is then found by a bisection method. This algorithm can be computationally intensive, especially when the density function is complex and not given in a closed form. By simulating the stochastic differential system for the underlying asset price process, Chen and Ewald [37] approximated the distribution function of the underlying asset with its empirical distribution. In this chapter, we propose to recover the distribution function of the underlying asset price from its characteristic function by using Fourier-consine series.

Instead of the optimal strikes, model-independent strikes for European options in the hedging portfolio can also be used, such as the strikes proposed by Tchuindjo [167].

However, these strikes will be shown to be non-optimal, although its computation cost is very low in comparison with the scheme in [1]. Only for deeply in-the-money options, the approximation proposed by Tchuindjo [167] can be accurate.

Generalizing the work of Simon et al. [160], Chen et al. [40] proposed a static superhedging strategy consisting of a linear combination of the market-traded vanilla options, whose strikes are the solutions of an optimization problem to search for the cheapest superhedging strategy. This approach is *model-free* in the sense that it does not involve any parametric model, but finite market data. The non-uniqueness of the optimal static super-replicating strategy is discussed in [39]. Note that if the strikes do not correspond to the cheapest static hedging strategy, the resulting upper bound can be an arbitrage upper bound. As an illustrating example, the model-independent strikes proposed by Tchuindjo [167] will be used to calculate the value bound of Asian options in the model-free framework. We will show that these strikes lead to an arbitrage upper bound for an arithmetic Asian option.

Both the model-based approach and the model-free approach have their own advantages and shortcomings. The model-based approach can provide an exact value of a derivative, but it cannot avoid model uncertainty. The model-free approach is free from model uncertainty, but it may suffer from inaccuracy by approximating a derivative value with its bounds. However, the comparison between model-based and model-free (optimal) upper bounds can be used to detect model mis-specification. We will discuss some practical issues when calculating the model-free optimal upper bound, and when comparing it with the model-based optimal upper bound under the Heston model.

7.2 Comonotonic upper bounds and static super-replicating strategies

Consider a finite time horizon $T > 0$. The financial market is described via a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{0\leq t\leq T}, \mathbb{P})$, which satisfies the usual technical conditions of completeness and right-continuity, and where $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets of Ω. Price processes of traded financial instruments are modelled as stochastic processes on that probability space which are adapted to the filtration $(\mathcal{F}(t))_{0\leq t\leq T}$.
Market participants are assumed to have access to a number of European options with maturities t_i, $i = 1, \dots N$, with $0 = t_0 < t_1 < \cdots < t_N = T$. More precisely, they can trade in European calls and puts on the individual stocks. We in particular consider an asset with a non-negative stochastic price process denoted by $\{S(t), 0 \leq t \leq T\}$

that pays dividends continuously over time at a constant rate q per unit time. Further we consider a discretely monitored arithmetic Asian option of European-style maturing at T with strike price K. The pay-off at T depends on the underlying process S at the times t_i, $i = 1, \dots N$, weighted by corresponding positive weights w_i, $i = 1, \dots, N$, which sum up to one:

$$\left(\xi \sum_{i=1}^{N} w_i S(t_i) - \xi K \right)^+,$$

where $(x)^+ = \max(x, 0)$. When $\xi = 1$, it is a call option, when $\xi = -1$ it is a put option. Very often an equally weighted sum is used, i.e., all weights are chosen to be equal to $1/N$.

It is assumed that the financial market is arbitrage-free. In the model-based approach, we further assume that there exists a pricing measure $\mathbb{Q}$, equivalent to the physical probability measure $\mathbb{P}$, such that the current price of any pay-off at time t_i, $i \in \{1, \dots, N\}$, can be represented as the expectation of the discounted pay-off. In this price-recipe, the discounting factor is e^{-rt_i}, where r is the continuously compounded time-0 risk-free interest rate to expiration t_i, whereas expectations are taken with respect to $\mathbb{Q}$. For simplicity in notation and terminology, we assume deterministic interest rates. The case of a stochastic interest rate is covered in [39].

The price of an Asian call or put at time $t \in [0, T]$ with maturity T and strike K is given by:

$$A(t, T, K; \xi) = \mathrm{e}^{-r(T-t)} \mathbb{E} \left[\left(\xi \sum_{i=1}^{N} w_i S(t_i) - \xi K \right)^+ \Big| \mathcal{F}(t) \right]. \tag{7.1}$$

Note that without loss of generality we will assume that $t < t_1$ in what follows. When $t_1 < t < T$, we absorb the known asset prices in the strike.

Let us further introduce the following notation for the weighted sum given the information $\mathcal{F}(t)$, or in a Markovian setting given $S(t)$:

$$\mathcal{S} = \sum_{i=1}^{N} w_i S(t_i) \mid \mathcal{F}(t), \tag{7.2}$$

and for its corresponding comonotonic counterpart

$$\mathcal{S}^c = \sum_{i=1}^{N} w_i (F^t_{S(t_i)})^{-1}(U), \tag{7.3}$$

with U a uniform $(0,1)$-random variable and $F^t_{S(t_i)}$ the conditional cumulative distribution function (cdf) of $S(t_i)$ given the information $\mathcal{F}(t)$, or in a Markovian setting given $S(t)$, under the martingale measure $\mathbb{Q}$:

$$F^t_{S(t_i)}(x) = \mathbb{Q}(S(t_i) \leq x \mid \mathcal{F}(t)).$$

The conditional cdf of $\mathcal{S}^c$ given the information $\mathcal{F}(t)$ is analogously denoted by $F^t_{\mathcal{S}^c}$. When $t = 0$ we do not write the superscript.

$\mathcal{S}^c$ precedes $\mathcal{S}$ in the convex order sense, i.e.,

$$\mathbb{E}[\mathcal{S}] = \mathbb{E}[\mathcal{S}^c], \quad \mathbb{E}[(\mathcal{S}-d)^+] \leq \mathbb{E}[(\mathcal{S}^c-d)^+], \quad \forall d \in \mathbb{R}. \tag{7.4}$$

Let $f(d) = \mathbb{E}[(\mathcal{S}^c-d)^+] - \mathbb{E}[(\mathcal{S}-d)^+]$. $f(d)$ first increases as function of d, and then decreases (from some c on) but remains non-negative [61]. Therefore, not for out-of-the-money options or at-the-money options but for in-the-money options, it is much more accurate to approximate the option value by the corresponding comonotonic upper bound. Thus in the numerical experiments, we will focus on in-the-money options.

Before stating the main theorem about the optimal upper bound, we first recall here some definitions of inverses of cumulative distribution functions. The usual inverse F_X^{-1} of the cdf F_X of a random variable X is denoted by

$$F_X^{-1}(p) = \inf\{x \in \mathbb{R} \mid F_X(x) \geq p\}, \quad p \in [0,1],$$

with $\inf \emptyset = +\infty$, by convention. For any $x \in \mathbb{R}$ and $p \in [0,1]$, when F_X is strictly increasing, the following equivalence relation holds:

$$F_X^{-1}(p) = x \quad \Leftrightarrow \quad p = F_X(x).$$

However, when the cdf F_X is not strictly increasing the inverse of the cdf is not uniquely determined. We define the inverse F_X^{-1+} as follows:

$$F_X^{-1+}(p) = \sup\{x \in \mathbb{R} \mid F_X(x) \leq p\}, \quad p \in [0,1],$$

with $\sup \emptyset = -\infty$, by convention. Any convex combination of F_X^{-1} and F_X^{-1+} is also an inverse distribution function. As in [112] we define for any $\alpha \in [0,1]$ the α-inverse of the cdf F_X as

$$F_X^{-1(\alpha)}(p) = \alpha F_X^{-1}(p) + (1-\alpha)F_X^{-1+}(p), \quad p \in (0,1).$$

From the theory of comonotonic risks (see, e.g. [61, 112]), applied to Asian options see, e.g. [1, 40, 122, 160], we state

Theorem 1. *The Asian option price at $t \in [0,T]$ defined by (7.1) with a strike $K \in \big((F^t_{\mathcal{S}^c})^{-1+}(0),\ (F^t_{\mathcal{S}^c})^{-1}(1)\big)$ is bounded above as follows*

$$\begin{aligned} A(t,T,K;\xi) &\le \mathrm{e}^{-r(T-t)}\mathbb{E}\left[(\xi\mathcal{S}^c - \xi K)^+ \,\middle|\, \mathcal{F}(t)\right] \\ &= \mathrm{e}^{-rT}\sum_{i=1}^N w_i \mathrm{e}^{rt_i}\mathrm{e}^{-r(t_i-t)}\mathbb{E}\left[(\xi S(t_i) - \xi K_i)^+ \,\middle|\, \mathcal{F}(t)\right] \\ &=: \mathrm{e}^{-rT}\sum_{i=1}^N w_i \mathrm{e}^{rt_i} E(t,t_i,K_i;\xi), \end{aligned} \tag{7.5}$$

where the strikes $K_i \ge 0$ are given by

$$K_i = (F^t_{S(t_i)})^{-1(\alpha_i)}(F^t_{\mathcal{S}^c}(K)), \quad i = 1,\dots,N \tag{7.6}$$

with the α_i, $i = 1,\dots,N$, chosen in $[0,1]$ such that

$$\sum_{i=1}^N w_i K_i = K.$$

Moreover, the comonotonic upper bound (7.5) is optimal in the sense that for any set of strikes k_i, $i = 1,\dots,N$, such that $\sum_{i=1}^N w_i k_i = K$, it holds that

$$A(t,T,K;\xi) \le \sum_{i=1}^N w_i \mathrm{e}^{-r(T-t_i)} E(t,t_i,K_i;\xi) \le \sum_{i=1}^N w_i \mathrm{e}^{-r(T-t_i)} E(t,t_i,k_i;\xi). \tag{7.7}$$

This upper bound corresponds to the *infinite market case*, where it is assumed that all European option prices $E(t,t_i,K;\xi)$ for any maturity t_i, $i = 1,\dots,N$, and any strike $K \ge 0$ are known. Knowledge of the prices $E(t,t_i,K;\xi)$ for all $K \ge 0$ is equivalent to the knowledge of the conditional cdf $F^t_{S(t_i)}(x)$ for all x. This approach is a *model-based* approach as it is based on a particular asset price model such as the Black–Scholes model.

On the other hand when we assume that all these option prices are known because the price of any put and call for any strike $K \ge 0$ is observed in the market, we call this approach *model-free*. No assumption concerning the pricing measure $\mathbb{Q}$ that is actually used by the market is made in this case. The cdfs are as follows extracted from the market:

$$F^t_{S(t_i)}(x) = \delta_{1\xi} + e^{rt_i}\frac{\partial E(t,t_i,x+;\xi)}{\partial x},$$

where $\delta_{1\xi}$ stands for the Kronecker delta.
Following a reasoning as in Theorem 3 of [122] it holds that in the model-free approach an Asian call option can be super-replicated by trading in vanilla call options on the underlying asset of the Asian option, while an Asian put option can be superhedged by trading in vanilla put options on the underlying asset of the Asian option. Therefore, the price inequality

$$A(t,T,K;\xi) \leq \sum_{i=1}^{N} w_i \mathrm{e}^{-r(T-t_i)} E(t,t_i,K_i;\xi)$$

holds without the explicit assumption that the involved option prices are expectations of discounted pay-offs under some $\mathbb{Q}$-measure. We only have to assume that all option prices involved are traded prices in an arbitrage-free market, implying that a super-replicating strategy for the Asian option is more expensive than the Asian option itself. Notice however that in order to prove the equality

$$\mathrm{e}^{-rT} \sum_{i=1}^{N} w_i \mathrm{e}^{rt_i} E(t,t_i,K_i;\xi) = \mathrm{e}^{-r(T-t)} \mathbb{E}\left[(\xi \mathcal{S}^c - \xi K)^+ \,\middle|\, \mathcal{F}(t) \right]$$

we have to assume that option prices can be expressed as expectations of their discounted pay-offs.

Assuming that vanilla call and put options for any strike $K \geq 0$ are traded in the market is not realistic. Therefore, we will briefly discuss the *finite market case*. For more details and explanation of notations we refer to [40] and [122]. Assume that at time t (in practice it will be at time $t = 0$) only the prices $E(t,t_i,K_{i,j};\xi)$, $j = 0,1,\ldots,m_i$, $i = 1,\ldots,N$, are observed for a set of strikes satisfying

$$0 = K_{i,0} < K_{i,1} < \cdots < K_{i,m_i} < K_{i,m_i+1},$$

and where the strikes K_{i,m_i+1} are defined by

$$K_{i,m_i+1} = \sup\left\{ K \geq 0 \mid E(t,t_i,K;1) > 0 \right\}.$$

Linders et al. [122] proposed a method to determine these K_{i,m_i+1}, $i = 1,\ldots,N$.
Notice that we assume that the sets of traded strikes for the call and put options are identical. This assumption can be relaxed, see [122].

Theorem 2. *In the finite market case, the Asian option price at $t \in [0,T]$ defined by (7.1) with a strike $K \in ((F^t_{\mathcal{S}^c})^{-1+}(0), (F^t_{\mathcal{S}^c})^{-1}(1))$ is bounded above as follows*

$$A(t,T,K;\xi) \leq \sum_{i \in N_K} w_i \mathrm{e}^{-r(T-t_i)} E(t,t_i,K_{i,j_i};\xi)$$

$$+\sum_{i\in\overline{N}_K} w_i \mathrm{e}^{-r(T-t_i)} \left(\alpha_K E(t,t_i,K_{i,j_i};\xi) + (1-\alpha_K)E(t,t_i,K_{i,j_i+1};\xi)\right) \tag{7.8}$$

$$=: \mathrm{UB}(t,T,K;\xi)$$

where

$$N_K = \left\{ i\in\{1,\dots,N\} \,\middle|\, \begin{array}{l} \exists j_i \in \{0,\dots,m_i+1\} \quad \textit{such that} \\ \overline{F}^t_{S(t_i)}(K_{i,j_i-1}) < F^t_{\overline{\mathcal{S}}^c}(K) < \overline{F}^t_{S(t_i)}(K_{i,j_i}) \end{array} \right\},$$

$$\overline{N}_K = \left\{ i\in\{1,\dots,N\} \,\middle|\, \begin{array}{l} \exists j_i \in \{0,\dots,m_i\} \quad \textit{such that} \\ F^t_{\overline{\mathcal{S}}^c}(K) = \overline{F}^t_{S(t_i)}(K_{i,j_i}) \end{array} \right\},$$

with α_K *chosen in* $[0,1]$ *such that*

$$(F^t_{\overline{\mathcal{S}}^c})^{-1(\alpha_K)}(F^t_{\overline{\mathcal{S}}^c}(K)) = K.$$

Moreover, the upper bound (7.8) *is optimal in the sense that for any set of strikes* k_i, $i=1,\dots,N$, *such that* $\sum_{i=1}^N w_i k_i = K$, *it holds that*

$$\begin{aligned} &A(t,T,K;\xi) \\ &\quad\le \mathrm{UB}(t,T,K;\xi) \\ &\quad\le \sum_{i\in M_K} w_i \mathrm{e}^{-r(T-t_i)} E(t,t_i,K_{i,j_i};\xi) \\ &\quad+ \sum_{i\in\overline{M}_K} w_i \mathrm{e}^{-r(T-t_i)} \left(\beta_{i,K} E(t,t_i,K_{i,j_i};\xi) + (1-\beta_{i,K})E(t,t_i,K_{i,j_i+1};\xi)\right) \end{aligned} \tag{7.9}$$

where

$$\begin{aligned} M_K &= \{i\in\{1,\dots,N\} \mid \exists j_i \in \{0,\dots,m_i\} : k_i = K_{i,j_i}\} \\ \overline{M}_K &= \{i\in\{0,\dots,N\} \mid \exists j_i \in \{1,\dots,m_i\} : K_{i,j_i} < k_i < K_{i,j_i+1}\}, \end{aligned}$$

and with $\beta_{i,K}$ *chosen in* $(0,1)$ *such that*

$$\beta_{i,K} K_{i,j_i} + (1-\beta_{i,K}) K_{i,j_i+1} = k_i.$$

As an illustrating example of strikes k_i, $i=1,\dots,N$, different from the optimal strikes in Theorem 1 and Theorem 2, which may provide a non-optimal upper bound, we consider the model-independent strikes proposed by Tchuindjo [167] given by

$$k_i = \frac{K\mathrm{e}^{(r-q)(t_i-t)}}{\sum_{j=1}^N w_j \mathrm{e}^{(r-q)(t_j-t)}}, \quad i=1,\dots,N, \tag{7.10}$$

where q is the dividend rate. According to (7.7) and (7.9), the value bound calculated with the model-independent strikes (7.10) should not be smaller than the optimal value bound calculated with the optimal strikes in both frameworks, the model-based one and the model-free one. Since the optimal upper bound corresponds to an optimal static super-replicating strategy for Asian options, any non-optimal upper bound may lead to arbitrage. The non-uniqueness of the optimal static super-replicating strategy is discussed in [39].

From the theory of increasing convex ordering, it immediately follows that the upper bounds (7.8), respectively (7.9), in the finite market case will not be smaller than the upper bound (7.5), respectively (7.7), in the infinite market case. This can also be seen as follows. The option price curve is a convex function of the strike. In the finite market case an approximate option price curve is defined as a linear interpolation of the curve in the infinite market and hence the former lies above the latter.

As in the infinite market case, the approach is model-based when a particular price model for the underlying asset is assumed. The upper bounds (7.8) and (7.9) are the prices of static super-replicating strategies which are model-free when we do not assume that the involved option prices are expected discounted pay-offs under some $\mathbb{Q}$-measure. The only assumption that we have to make is that the market is free of arbitrage.

It is known from the literature that the comonotonic upper bound is a rather rough estimate for the option price, except for in-the-money option. Therefore, better model-based bounds are searched for. Comonotonic lower bounds and improved upper bounds based on conditioning are derived and discussed for Asian options and Asian basket options (see, e.g. [56, 120, 124, 136, 171, 182]). These bounds are, however, out of the scope of this thesis.

To conclude this section we derive a model-independent lower bound in terms of the t_i-forward prices at time t, with $t_i > t$, given by

$$\mathbb{E}[S(t_i) \mid \mathcal{F}(t)] = S(t)\mathrm{e}^{(r-q)(t_i-t)}. \tag{7.11}$$

Applying Jensen's inequality to the convex function $(\cdot)^+ = \max(\cdot, 0)$ in (7.1) and using (7.11), we easily arrive at the lower bound $\underline{A}(t, T, K; \xi)$:

$$\underline{A}(t, T, K; \xi) := \mathrm{e}^{-r(T-t)} \left(\xi \sum_{i=1}^{N} w_i S(t) \mathrm{e}^{(r-q)(t_i-t)} - \xi K \right)^+ \leq A(t, T, K; \xi). \tag{7.12}$$

7.3 Acceleration of the computation of the model-based optimal upper bound

The distribution function of the underlying asset at time t_i and its inverse play an important role in calculating the optimal strikes (7.6). Albrecher et al. [1] numerically built up the distribution function from the density function of the underlying asset price, and its inverse is then found by a bisection method. It is computationally intensive to build up the distribution function from the density function, especially when the density function is very complex or singular at some points, which is for example the case for the density function, see [65], in the Heston model [103]. To accelerate the procedure proposed by [1], we recover the distribution function from the characteristic function of the log-asset price, which is available in a large class of stochastic models, such as Lévy models [1] and Heston's model.

We approximate the distribution function $F_{\log(S_t)}$ by $\hat{F}_{\log(S_t)}$,

$$\hat{F}_{\log(S_t)}(x) = \begin{cases} 1 & \text{if } x \geq b, \\ \dfrac{x-a}{b-a} + \sum_{k=1}^{n} F_{k,t} \cdot \dfrac{b-a}{k\pi} \cdot \sin(k\pi \dfrac{x-a}{b-a}) & \text{if } x \in (a,b), \\ 0 & \text{if } x \leq a, \end{cases} \tag{7.13}$$

where $n \in \mathbb{N}$ is the number of the terms of Fourier-cosine series expansion for the density function f, and a, b $(-\infty < a < b < +\infty)$ are chosen such that

$$\int_a^b e^{i\omega x} f(x)\,\mathrm{d}x \approx \int_{\mathbb{R}} e^{i\omega x} f(x)\,\mathrm{d}x,$$

and

$$F_{k,t} = \frac{2}{b-a}\mathrm{Re}\left\{\phi\left(\frac{k\pi}{b-a}, t\right) \cdot \exp\left(-i\frac{ka\pi}{b-a}\right)\right\}. \tag{7.14}$$

In practice, we can set a to be sufficiently small while b is set to be sufficiently large. The derivation of (7.13)-(7.14) is given in Section 5.2.3. One may refer to [76] for an error analysis on the Fourier-cosine series expansion.

Correspondingly, the distribution function F_{S_t} can be approximated by $\hat{F}_{S_t}$, given by

$$\hat{F}_{S_t}(x) = \hat{F}_{\log(S_t)}\left(\log(x)\right).$$

Since $\hat{F}_{\log(S_t)}$ is given in closed form in terms of sine functions, its inverse can be easily calculated with the bisection method. Hence, the computational cost of the optimal strikes can be significantly reduced in the model-based framework.

7.4 Numerical examples

In this section, we consider the Asian option value (7.1) at time zero and with equal weights w_i, i.e.

$$A(0,T,K;\xi) = e^{-rT}\mathbb{E}\left[\left(\frac{1}{N}\xi\sum_{i=1}^{N}S(t_i) - \xi K\right)^{+}\Big| S(0) = S_0\right]. \tag{7.15}$$

The numerical experiments are first carried out to compare the quality of the upper bounds calculated with the strikes (7.6) and (7.10) in the model-based framework. The quality is assessed by their deviation from the model-independent lower bound (7.12) and the model price calculated with the Monte Carlo method.

Then, real market data are employed to calculate the model-free optimal value bound (7.8) and the model-free upper bound (7.9) with the strikes (7.10).

At last, we compare the model-free value bound with the optimal value bound based on the well-calibrated Heston models. The comparison can be used to detect model mis-specification, and some practical issues are discussed.

7.4.1 Model-based value bounds

According to Theorem 1, the comonotonic upper bound (7.5) should be smaller than or equal to the upper bound (7.7) with the model-independent strikes (7.10). In addition, the upper bound should not be less than the model price (MC-price) calculated with the Monte Carlo method. The MC-price is used to assess the quality of the upper bounds with the strikes (7.6) and (7.10). On the other hand, different pricing models may provide different MC-prices and different value bounds for each target option. In addition, the lower bound (7.12) can also work as a reference to assess the upper bounds for deeply in-the-money options. In this subsection, we assess the upper bounds under different models.

7.4.1.1 The Black–Scholes model and the Variance Gamma model

Here, we focus on assessing the quality of the model-based upper bounds in the Black–Scholes model and the Variance Gamma model with strikes given by (7.6) and (7.10). For an Asian put option we calculate the MC-prices and the value bounds with the same model parameters as those used in [167]. The results are plotted in Figure 7.1 and Figure 7.2, and the data are summarized in Table 7.1 and Table 7.2, respectively. Note that the

MC prices taken from Table VII and Table VIII in [167] are sometimes larger than the corresponding upper bounds[1]. The remedy provided here is to reduce the variance with a control variate[2].

Figure 7.1 and Figure 7.2 depict that, in each model, the comonotonic upper bound (7.5) is smaller than the upper bound (7.7) with the model-independent strikes (7.10). Comparing the results of Table 7.1 and Table 7.2, we conclude that a different pricing model for the underlying asset leads to different prices and model-based upper bounds for the Asian option when they are not too deep-in-the-money. Thus, if no information is available to differentiate the candidate models, one is confronted with model uncertainty.

On the other hand, as shown by [91], different calibration methods could yield different estimations for each model parameter. Calibration risk can be defined as the different optimal parameter sets arising from the different specifications of the discrepancy measures and the calibration methodology in the procedure of model calibration. We will investigate this problem with the well-calibrated Heston models from [91].

7.4.1.2 Heston's model

In the Heston model [103], the asset price process (S) and the variance process (v) under a risk-neutral probability measure $\mathbb{Q}$ evolve according to the following system of stochastic differential equations,

$$\begin{cases} \mathrm{d}S_t = (r-q)S_t\,\mathrm{d}t + \sqrt{v_t}S_t\,\mathrm{d}W_t, & S_0 > 0, \\ \mathrm{d}v_t = \kappa(\eta - v_t)\,\mathrm{d}t + \lambda\sqrt{v_t}\,\mathrm{d}\widetilde{W}_t, & v_0 = \sigma_0^2 > 0, \end{cases} \tag{7.16}$$

where $(W_t)_{0\le t\le T}$ and $(\widetilde{W}_t)_{0\le t\le T}$ are correlated Brownian motions satisfying $dW_t d\widetilde{W}_t = \rho dt$. The parameter η is the long-term average variance, while κ is the speed of the mean-reversion of the variance. The parameter λ is referred to as the volatility of variance since it scales the diffusion term of the variance process.

Different from the Lévy models used in [1] and [167], the density function of the underlying is not immediately available in closed form in the Heston model [65]. It is time-consuming to build up the distribution from the density function. However, the characteristic function of the log-asset price is available and reads

$$\phi(\omega, t) := \mathbb{E}\left[\exp\left(i\omega\log(S_t)\right) \mid S_0, v_0\right] = \exp(A + B + C), \tag{7.17}$$

[1] We refer to [163] for comments on those results.

[2] One may refer to the Appendix at the end of this chapter or [88] for more details on the technique of a control variate.

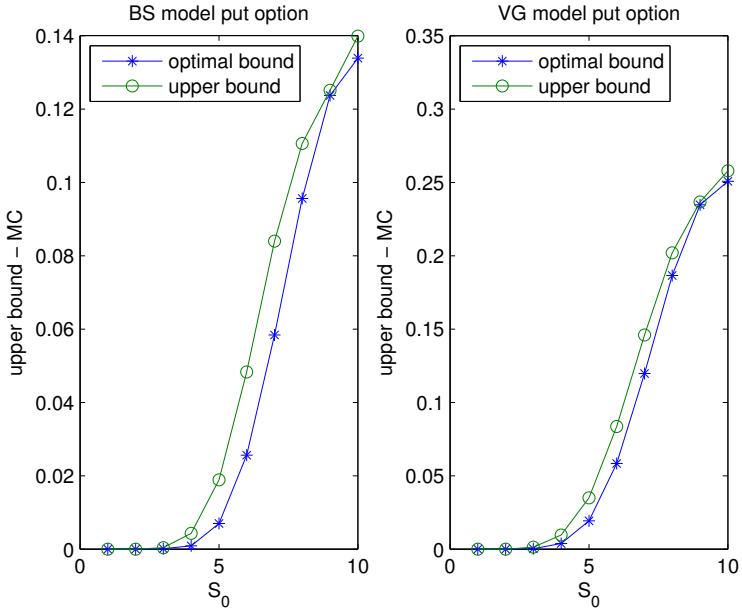


Figure 7.1: Spread between the (optimal) upper bound and the MC-price in the Black–Scholes model and the VG model

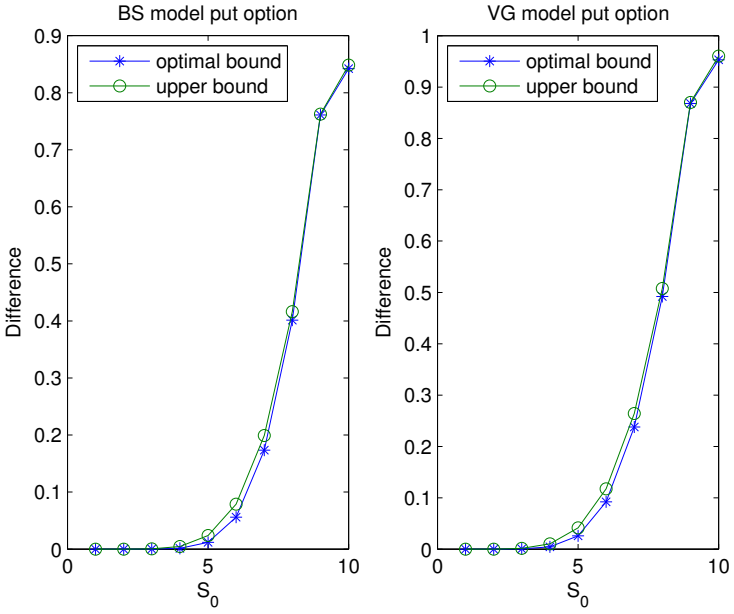


Figure 7.2: Spread between (optimal) upper bound and lower bound in the Black–Scholes model and the VG model

where

$$
\begin{aligned}
A &= i\omega\left(\log(S_0) + (r-q)t\right),\\
B &= \eta\kappa\lambda^{-2}\left((\kappa-\rho\lambda\omega i+d)\,t - 2\log\left(\left(1-ge^{dt}\right)/(1-g)\right)\right),\\
C &= v_0\lambda^{-2}(\kappa-\rho\lambda\omega i+d)(1-e^{dt})/(1-ge^{dt}),\\
d &= \sqrt{(\rho\lambda\omega i-\kappa)^2+\lambda^2(\omega i+\omega^2)},\\
g &= (\kappa-\rho\lambda\omega i+d)/(\kappa-\rho\lambda\omega i-d).
\end{aligned}
$$

Hence, we can use the method proposed in Section 7.3 to approximate the distribution function and calculate its inverse with a bisection method. The parameters in (7.13) are set to be $n = 2^{14}$, $a = -10$, $b = 2500$. The implementation is done in MATLAB (2014b) (Processor: Intel Core(TM) i7-3770 CPU @ 3.4GHz, RAM: 8GB). The computational cost to calculate the approximation of the distribution function (7.13) is less than $0.01s$. Following the lines of [1], we also approximated the distribution function by evaluating the integral of the density function with the *integral* function of MATLAB. Its computational cost is about $9.5s$. Hence, our method can efficiently accelerate the computation of the model-based optimal upper bound.

With the well-calibrated Heston models from [91], we calculate the MC-prices and the value bounds of the in-the-money Asian options monitored on the fourteen expiration dates of the benchmark SPX options (see Table 7.5 in the next section). We refer to the Appendix at the end of Chapter 6 for the computational aspects on model calibration and the Monte Carlo methods for the Heston model. In the present case, we use the Asian option value (7.15) with $T = 737$ days, $N = 14$ and

$$t_i \in \{1, 9, 20, 37, 72, 100, 110, 191, 201, 282, 293, 373, 555, 737\}.$$

The continuously compounded interest rate (r) is set to be a constant 0.0153. The strikes of Asian call options range from 500 to 860, while they range from 900 to 1200 for Asian put options. The difference between two consecutive strikes is 20. The model-free upper bounds of these options will be calculated in the next section.

The calculation of the optimal upper bound (7.5) and of the upper bound (7.7) involves the calculation of European option prices with the optimal strikes K_i by (7.6) and the strikes k_i by (7.10), where the optimal strikes are computed using the approximation of Section 7.3. The European option prices are calculated with the COS method [76].

The results are reported in columns two to five, seven and eight of Table 7.3 and Table 7.4. Besides the non-optimality of the upper bounds with the strikes (7.10), these

results confirm that although the well-calibrated models fit to the same data, they could also lead to different prices and model-based upper bounds.

In a nutshell, these numerical results under different models confirm that compared with the upper bound (7.7) with the model-independent strikes (7.10), the comonotonic upper bound (7.5) is a better approximation for the Asian option value, although both of them are good approximations for very deep-in-the-money options. The lower bounds are model-independent, however, different pricing models provide different upper bounds for the same option. The model-free upper bounds may work as a robust alternative.

$S(0)$	MC Tchuindjo	MC price	Standard error	Lower bound	Pricing error(%)	optimal upper bound	Pricing error(%)	upper bound	Pricing error(%)
10	0.0454	0.7085	$1.7011 \cdot 10^{-5}$	0.0000	-100.0000	0.8424	18.8990	0.8484	19.7459
9	0.5213	1.1287	$1.7600 \cdot 10^{-5}$	0.4913	-56.3095	1.2524	10.9595	1.2538	11.0835
8	1.4032	1.6987	$1.8847 \cdot 10^{-5}$	1.3930	-17.9865	1.7943	5.6278	1.8093	6.5109
7	2.3031	2.4098	$2.0456 \cdot 10^{-5}$	2.2948	-4.8985	2.4682	2.4234	2.4938	3.4858
6	3.2028	3.2271	$2.1228 \cdot 10^{-5}$	3.1966	-1.0463	3.2527	0.7933	3.2754	1.4967
5	4.1043	4.1031	$2.0207 \cdot 10^{-5}$	4.0983	-0.1608	4.1101	0.1706	4.1220	0.4606
4	5.0050	5.0004	$7.1867 \cdot 10^{-6}$	5.0001	-0.0120	5.0013	0.0180	5.0047	0.0860
3	5.9052	5.9018	$6.6788 \cdot 10^{-7}$	5.9018	-0.0017	5.9019	0.0017	5.9022	0.0068
2	6.8058	6.8036	$5.2797 \cdot 10^{-15}$	6.8036	0.0000	6.8036	0.0000	6.8036	0.0000
1	7.7064	7.7053	$1.5484 \cdot 10^{-14}$	7.7053	0.0000	7.7053	0.0000	7.7053	0.0000

$*$ Pricing error: (value bound $-$ MC-price)/MC-price

Table 7.1: Valuation of 3-year put option with monthly average for a geometric Brownian motion with drift: $t = 0$, $T = 3$, $N = 36$, $K = 10$, $r = 5\%$, $q = 2\%$ *and* $\sigma = 25\%$. *Monte Carlo simulations with control variate and* 10^7 *paths.*

$S(0)$	MC Tchuindjo	MC price	Standard error	Lower bound	Pricing error(%)	optimal upper bound	Pricing error(%)	upper bound	Pricing error(%)
10	0.0006	0.7023	$1.0493 \cdot 10^{-4}$	0.0000	-100.0000	0.9530	35.6970	0.9602	36.7222
9	0.7500	1.1245	$1.0506 \cdot 10^{-4}$	0.4913	-56.3095	1.3592	20.8715	1.3612	21.0494
8	1.6227	1.6985	$9.1384 \cdot 10^{-5}$	1.3930	-17.9865	1.8850	10.9803	1.9005	11.8928
7	2.4951	2.4130	$6.6868 \cdot 10^{-5}$	2.2948	-4.8985	2.5327	4.9606	2.5590	6.0506
6	3.3678	3.2304	$3.9222 \cdot 10^{-5}$	3.1966	-1.0463	3.2887	1.8047	3.3140	2.5879
5	4.2405	4.1049	$1.7545 \cdot 10^{-5}$	4.0983	-0.1608	4.1242	0.4702	4.1398	0.8502
4	5.1120	5.0007	$5.4677 \cdot 10^{-6}$	5.0001	-0.0120	5.0045	0.0760	5.0103	0.1920
3	5.9860	5.9019	$1.0194 \cdot 10^{-6}$	5.9018	-0.0017	5.9022	0.0051	5.9032	0.0220
2	6.8584	6.8036	$3.3962 \cdot 10^{-8}$	6.8036	0.0000	6.8036	0.0000	6.8036	0.0000
1	7.7307	7.7053	$1.0080 \cdot 10^{-8}$	7.7053	0.0000	7.7053	0.0000	7.7053	0.0000

$*$ Pricing error: (value bound $-$ MC-price)/MC-price

Table 7.2: Valuation of 3-year put option with monthly average for an exponential of a Variance Gamma process: $t = 0$, $T = 3$, $N = 36$, $K = 10$, $r = 5\%$, $q = 2\%$, $\sigma = 25\%$, $\nu = 10\%$ *and* $\theta = -1\%$. *Monte Carlo simulations with control variate and* $4 \cdot 10^7$ *paths.*

	lower	Heston MC		optimal upper bound			upper bound			
K	bound	RMSE full	ARPE full	RMSE full	ARPE full	finite market	RMSE full	ARPE full	finite market	toy model
500	365.4489	367.4518	366.7792	369.0387	366.8906	368.7650	377.4261	371.2725	376.9004	378.6416
520	346.0573	348.5431	347.6028	350.8762	348.1456	349.7404	359.7963	353.0933	359.6151	361.0733
540	326.6657	329.8178	328.5379	332.9937	329.6096	331.4062	342.3578	335.1025	342.3543	343.7281
560	307.2742	311.3189	309.6240	315.4228	311.3229	313.5263	325.1252	317.3199	325.2742	326.5898
580	287.8826	293.0906	290.9071	298.1945	293.3283	297.4075	308.1144	299.7667	308.3754	309.6731
600	268.4910	275.1790	272.4383	281.3387	275.6693	280.5639	291.3423	282.4654	291.5447	292.9944
620	249.0994	257.6328	254.2744	264.8838	258.3901	265.4279	274.8270	265.4396	275.4190	276.5712
640	229.7079	240.4985	236.4761	248.8565	241.5343	247.5432	258.5879	248.7145	259.3650	260.4224
660	210.3163	223.8214	219.1070	233.2817	225.1443	232.9213	242.6463	232.3171	243.5608	244.5685
680	190.9247	207.6446	202.2300	218.1826	209.2607	218.0064	227.0254	216.2762	228.0322	229.0322
700	171.5331	192.0092	185.9056	203.5799	193.9215	203.6710	211.7510	200.6232	212.6356	213.8387
720	152.1415	176.9539	170.1928	189.4926	179.1618	189.9480	196.8523	185.3925	197.6861	199.0165
740	132.7500	162.5132	155.1474	175.9371	165.0132	175.9123	182.3627	170.6226	182.9466	184.5985
760	113.3584	148.7144	140.8173	162.9279	151.5034	163.8125	168.3210	156.3567	168.6664	170.6228
780	93.9668	135.5832	127.2431	150.4768	138.6554	149.2638	154.7717	142.6429	154.7720	157.1334
800	74.5752	123.1359	114.4563	138.5931	126.4878	137.0476	141.7643	129.5336	141.3114	144.1790
820	55.1836	111.3872	102.4799	127.2833	115.0126	125.3100	129.3525	117.0842	128.4388	131.8130
840	35.7921	100.3455	91.3270	116.5512	104.2373	114.1645	117.6089	105.3690	116.1144	120.1073
860	16.4005	90.0111	81.0021	106.3977	94.1619	103.9151	106.6861	94.5404	104.5333	109.2141

$*$ RMSE: root mean square error
$*$ ARPE: average relative price error

Table 7.3: Valuation of 737-day Asian call options with discrete monitored dates $t_i \in \{1, 9, 20, 37, 72, 100, 110, 191, 201, 282, 293, 373, 555, 737\}$ *in the settings of the finite market and the calibrated Heston models:* $\kappa = 0.5527$, $\eta = 0.1271$, $\lambda = 0.3748$, $v_0 = 0.2403$, $r = 0.0153$, $q = 0.0088$, $S_0 = 873.59$ *(RMSE full) and* $\kappa = 3.1022$, $\eta = 0.0923$, $\lambda = 0.3285$, $v_0 = 0.2513$, $r = 0.0153$, $q = 0.0088$, $S_0 = 873.59$ *(ARPE full). The parameters in the toy model are* $\kappa = 0.3527$, $\eta = 0.1671$, $\lambda = 0.3748$, $v_0 = 0.2403$, $r = 0.0153$, $q = 0.0088$, $S_0 = 873.59$. *Monte Carlo (MC) simulations with control variate and* $1 \cdot 10^7$ *paths.*

	lower	Heston MC		optimal upper bound			upper bound			
K	bound	RMSE full	ARPE full	*RMSE full*	ARPE full	finite market	RMSE full	ARPE full	finite market	toy model
900	22.3827	93.9196	85.2710	110.1975	98.4652	112.0413	110.5076	98.7108	112.2632	113.0693
920	41.7742	105.0721	96.7637	121.1472	109.8240	123.2413	122.1717	110.6232	123.5515	124.7371
940	61.1658	116.8996	109.0194	132.6505	121.8267	133.1272	134.5757	123.3271	135.4547	137.1365
960	80.5574	129.3802	122.0031	144.6951	134.4479	144.6585	147.5944	136.6911	148.1108	150.1423
980	99.4897	142.4921	135.6745	157.2668	147.6597	157.5850	161.1607	150.6422	161.4074	163.6878
1000	119.3405	156.2090	149.9936	170.3496	161.4318	170.4008	175.2200	165.1211	175.0193	177.7189
1020	138.7321	170.5015	164.9148	183.9262	175.7329	183.7759	189.7233	180.0738	189.4947	192.1869
1040	158.1237	185.3406	180.3935	197.9783	190.5310	198.4330	204.6272	195.4519	204.3106	207.0489
1060	177.5153	200.6961	196.3825	212.4864	205.7936	212.6058	219.8945	211.2127	219.5252	222.2681
1080	196.9069	216.5342	212.8325	227.4307	221.4886	227.7004	235.4929	227.3188	235.1219	237.8128
1100	216.2984	232.8195	229.6980	242.7904	237.5840	240.7294	251.3948	243.7374	250.9844	253.6556
1120	235.6900	249.5156	246.9353	258.5444	254.0486	256.8384	267.5756	260.4394	267.3432	269.7727
1140	255.0816	266.5878	264.5003	274.6713	270.8519	273.9496	284.0136	277.3986	283.9480	286.1427
1160	274.4732	284.0014	282.3541	291.1490	287.9645	291.8913	300.6895	294.5914	300.8338	302.7467
1180	293.8648	301.7213	300.4598	307.9555	305.3581	309.1717	317.5853	311.9964	317.9052	319.5674
1200	313.2563	319.7158	318.7826	325.0685	323.0059	326.2583	334.6850	329.5941	335.1898	336.5890

∗ RMSE: root mean square error
∗ ARPE: average relative price error

Table 7.4: Valuation of 737-day Asian put options with discrete monitored dates $t_i \in \{1, 9, 20, 37, 72, 100, 110, 191, 201, 282, 293, 373, 555, 737\}$ *in the settings of the finite market and the calibrated Heston models:* $\kappa = 0.5527$, $\eta = 0.1271$, $\lambda = 0.3748$, $v_0 = 0.2403$, $r = 0.0153$, $q = 0.0088$, $S_0 = 873.59$ *(RMSE full) and* $\kappa = 3.1022$, $\eta = 0.0923$, $\lambda = 0.3285$, $v_0 = 0.2513$, $r = 0.0153$, $q = 0.0088$, $S_0 = 873.59$ *(ARPE full). The parameters in the toy model are* $\kappa = 0.3527$, $\eta = 0.1671$, $\lambda = 0.3748$, $v_0 = 0.2403$, $r = 0.0153$, $q = 0.0088$, $S_0 = 873.59$. *Monte Carlo (MC) simulations with control variate and* $1 \cdot 10^7$ *paths.*

7.4.2 Model-free value bounds

This subsection focuses on the model-free value bounds in the finite market case. Market data on SPX options is employed to calculate the optimal upper value bound (7.8), the upper bound (7.9) with strikes given by (7.10) and the lower bound (7.12) of arithmetic Asian options written on the SPX. We refer to [40] and [122] for some computational aspects for the comonotonic optimal upper bound. Here, we will point out some practical issues when dealing with real market data.

7.4.2.1 Market data and target options

We use the market data on SPX options from [91]. These European vanilla options were quoted on 2008/12/11 in terms of bid-ask prices. There are fourteen different maturities t_i, $i = 1, \ldots, 14$ ranging from 1 day to 737 days. Options with strikes $K_{i,j}$, $j = 0, 1, \ldots, m_i$ are traded on each expiration date. The call and put options with the same strike are traded in the market. The closing price (S_0) of the SPX was 873.59 on 2008/12/11, and the one-year continuous dividend rate q was 0.0088. The summary statistics are displayed in Table 7.5. The mid-price data of the benchmark options is used to calculate the optimal upper value bound (7.8) and the upper bound (7.9) with strikes given by (7.10). We filtered out the options whose mid-price violated the convexity of the option value function on each date.

The target options are the same as (7.15).

7.4.2.2 Practical issues and results

The calculation of the optimal upper value bound (7.8) involves the marginal distribution of the underlying asset price on each monitor date. Due to the call-put parity, the marginal distribution recovered with the benchmark data from either call or put options should be the same. However, the call-put parity is derived under the no friction assumption paired with the law of one price. The existence of a bid-ask spread is not consistent with these assumptions. Hence, the call-put parity does not strictly hold when the mid-price is used. In order to deal with this, we recover two marginal distributions on each monitor date: the first one is from the mid-prices of call options and the second one is from the mid-prices of put options. That is, each Asian option has two comonotonic optimal bounds. The mid-value of these two optimal bounds is plotted in Figure 7.3 and the data is summarized in the column *optimal upper bound/finite market* in Table 7.3 and Table 7.4.

Maturity	Strike		Number of
(days)	$K_{\min}$	$K_{\max}$	Options
1	825	1000	12
9	765	2000	54
20	525	1650	38
37	300	1500	74
72	200	1500	84
100	200	2000	76
110	500	1300	34
191	200	1900	84
201	600	1500	48
282	200	1500	80
293	500	1500	54
373	200	2500	74
555	200	1800	52
737	200	2000	62

Table 7.5: Summary statistics of SPX options quoted on 2008/12/11: $S_0 = 873.59$

Figure 7.3 highlights that the optimal upper bound (7.8) is lower than the upper bound (7.9) with strikes (7.10) in the model-free framework. In addition, these upper bounds are larger than the lower bound (7.12). The corresponding values for theses bounds can be found in the columns two, seven and ten of the Table 7.3 and Table 7.4.

In the finite market case, the marginal distribution function on each monitoring date is constructed by means of a finite set of traded European options, and it is not a strictly increasing function (thus a non-decreasing function). According to [39], the set of optimal strikes is not unique in this setting, but the value of the optimal bound is unique. However, the strikes defined in (7.10) are not optimal as confirmed by the numerical results. This is in line with our expectations, since this latter set of strikes is not found as a solution to an optimization problem.

As depicted in Figure 7.4, the spread between an upper bound and the lower bound decreases as the moneyness goes to be far deep in-the-money for both Asian call and put options. For far deep in-the-money options, although the upper bound is close to the lower bound, the optimal upper bound is a better approximation. It is consistent with the results in the model-based framework.

Compared with the computational cost of the model-free upper bound with the

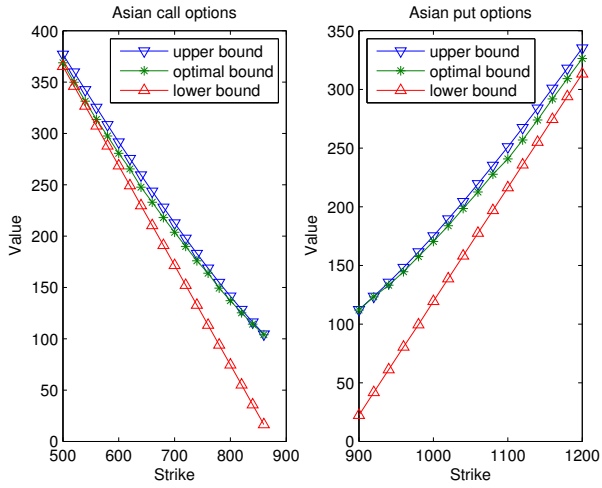


Figure 7.3: Upper bound, optimal upper bound and lower bound in the model-free framework

strikes given in (7.10), the computation of the model-free optimal upper bound requires an additional cost to search for the optimal strikes. However, the search algorithm proposed by [40] is quite efficient. We recommend to use the model-free optimal bound to approximate the Asian option value.

7.4.3 Discussion on model-based and model-free (optimal) upper bounds

Theoretically, if a pricing model is taken as the true dynamics of the underlying price in the infinite market, the model-based optimal bound should not exceed the optimal model-free upper bound in the finite market. Otherwise, the parametric model is mis-specified. As noted in [107], the model-free optimal upper bound can be employed to detect the model mis-specification. In practice, it is not necessary that the model-free optimal bound should strictly exceed the model-based optimal bound. The estimation method for the model parameters should be taken into account in this setting.

Model calibration is a popular method to estimate the model parameters which can minimize the difference between the model-based price and the benchmark data. Since it is an optimization problem for the difference, the calibrated model may not exactly reproduce the benchmark data, but it can work as an approximation of the dynamics of the underlying asset price process in the infinite market. It is reasonable that a well-

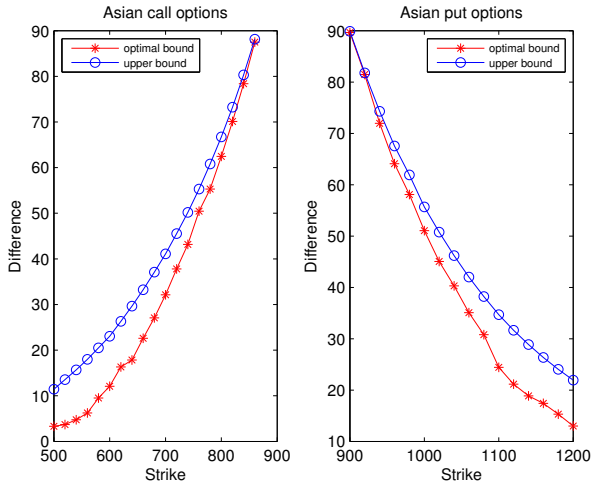


Figure 7.4: Spread between (optimal) upper bound and lower bound in the model-free framework

calibrated model leads to an optimal upper bound which exceeds the model-free bound with negligible difference. This rationale also holds if the upper bound with strikes (7.10) is used to detect model mis-specification.

In this subsection, we compare the numerical results of the model-free upper bounds in Section 7.4.2.2 with these reported in Section 7.4.1.2 under the calibrated Heston models [91]. We illustrate how to detect model mis-specification under the Heston model.

The numerical results summarized in Table 7.3 and Table 7.4 are plotted in Figure 7.5 and Figure 7.6. A simple criterion to detect model mis-specification is to compare the MC price with the model-free optimal bound. As depicted in Figure 7.5 and Figure 7.6, the Asian option prices under the calibrated Heston models are lower than their model-free optimal upper bound. Since the comonotonic upper bound is a rough estimate for the Asian option price, especially for around at-the-money options, it is not a strong criterion to assess a candidate model by comparing the model-based MC price with the model-free optimal upper bounds.

The model-based (optimal) upper bound is compared with the model-free (optimal) upper bound. Figure 7.5 and Figure 7.6 show that the (optimal) upper bound given by the first calibrated model (RMSE full) is close to, but not always lower than the corresponding model-free (optimal) upper bound. However, the difference is negligible

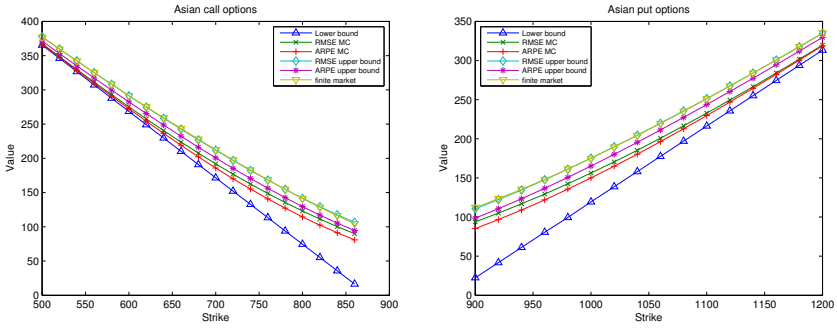


Figure 7.5: Comparison among model-based upper bounds, model-free upper bounds and MC prices of Asian options

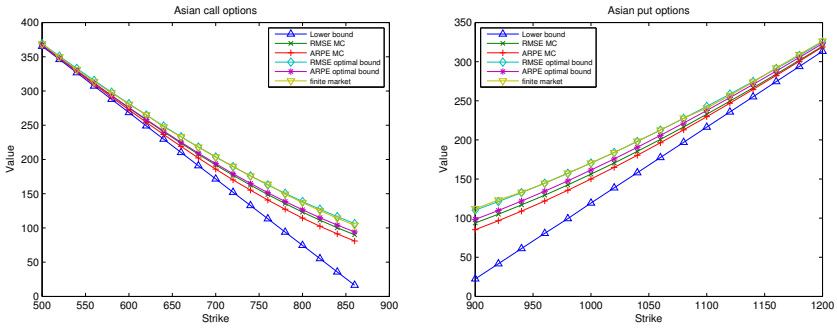


Figure 7.6: Comparison among model-based optimal upper bounds, model-free optimal upper bounds and MC prices of Asian options

between the model-based (optimal) upper bound based on the calibrated model (RMSE full) and the corresponding one in the finite market. Although the (optimal) upper bound given by the second calibrated model (ARPE full) is always lower than the model-free (optimal) upper bound, it is not convincing to reject the first calibrated model (RMSE full). Since the conclusion is the same from both perspectives, namely that of a comonotonic optimal upper bound and that of an upper bound with model-independent strikes, the upper bound is recommended to assess a pricing model. The model-based optimal upper bound, after all, involves the calculation of the optimal strikes.

If the model has been calibrated well to the market data in time, it seems redundant to detect model mis-specification with the approach based on these upper bounds. However, if the model parameters are not updated in time according to the market information, the model-free (optimal) upper bound can be employed to detect the model mis-specification. For example, we assume the Heston model was calibrated to the market data on some day before our benchmark data is available. However, the parameters are not updated when our benchmark data is available. As shown in Table 7.3 and Table 7.4, the upper bound under the toy model is strictly higher than the model-free upper bound.

Above all, the model-free optimal upper bound and the non-optimal upper bound can be used to detect model mis-specification, and the non-optimal upper bound is recommended. However, the estimation method for the model parameters should be taken into account to assess a model.

Appendix: control variate

The control variate technique, see e.g. [88], is applied to reduce the variance of the Monte Carlo estimator. In the present case of an Asian option, a possible control variate is the sum of the asset prices at the dates t_i, $i = 1, \ldots, N$:

$$X = \sum_{i=1}^{N} S(t_i),$$

with expected value, see (7.11) with $t = 0$,

$$\mathbb{E}[X] = \sum_{i=1}^{N} \mathbb{E}[S(t_i)] = S(0) \sum_{i=1}^{N} \mathrm{e}^{(r-q)t_i}.$$

For the experiments where the underlying is modelled by a geometric Brownian motion with volatility σ, we also consider the geometric Asian option with strike K and maturity

date T as control variate:

$$G(t,T,K;\xi) = e^{-r(T-t)}\mathbb{E}\left[\left(\xi\left(\prod_{i=1}^{N} S(t_i)\right)^{1/N} - \xi K\right)^{+} \middle| \mathcal{F}(t)\right]$$
$$= e^{-r(T-t)+(r-\delta)(\widetilde{T}-t)}E(t,\widetilde{T},K;\xi),$$

where the European option has maturity $\widetilde{T}$, strike K, interest rate $r-\delta$, and dividend rate q and is written on a price process modelled as a geometric Brownian motion with volatility $\widetilde{\sigma}$:

$$\widetilde{T} = \frac{1}{N}\sum_{i=1}^{N} t_i, \quad \widetilde{\sigma}^2 = \frac{\sigma^2}{N^2\widetilde{T}}\sum_{i=1}^{N}(2i-1)\,t_{N+1-i}, \quad \delta = \frac{1}{2}\sigma^2 - \frac{1}{2}\widetilde{\sigma}^2.$$

8 Conclusion

In financial markets, model uncertainty is a philosophical problem as well as a technical problem. Philosophically speaking, model uncertainty is ubiquitous whenever a probabilistic model is employed in financial economics. Although a model-free approach can be employed to avoid model uncertainty, probabilistic models are still extensively used in their own right. Hence, technically speaking, it is of great importance to quantify model uncertainty, reduce the impact of model uncertainty on quantities of interest and detect model mis-specification. These three topics constitute the theme of my thesis. The main contributions of this thesis can be summarized as follows:

1. An implementable numerical method was proposed to calculate the locally-risk minimising strategies for Asian options, spread options, basket options in a class of semimartingale models. These strategies are shown to be robust with respect to model uncertainty.

2. An efficient method was proposed to calculate the ensemble of the value of a derivative associated with a large ensemble of the model parameters. The ensemble of the derivative value allows us to quantify the uncertainty embedded in the derivative value with respect to parameter uncertainty. The proposed method can also be employed to accelerate the robust calibration procedure.

3. The T-COS method was proposed to calculate distorted expectations. This method can replace the Monte Carlo method to calculate distorted expectations if we know the density function or the characteristic function of a risk factor. The T-COS method can find extensive applications in finance, insurance and so on.

4. The weighted Monte Carlo method was empirically confirmed to be capable of reducing the impact of model uncertainty on the derivative value.

5. An efficient method is proposed to accelerate the calculation of the comonotonicity-based upper value bound for a discrete arithmetic Asian option in the models with the closed-form characteristic functions.

6. Some practical issues were addressed in detecting model mis-specification with the comparison between model-based quantities and their model-free counterparts of discrete arithmetic Asian options.

Based on the probabilistic models, financial economics has witnessed tremendous advance in theory and applications in financial markets. However, reflecting my research on model uncertainty, I find myself to be more ignorant of the essential mechanism of complex financial markets. It is so amazing and critical to quantify and manage uncertainty embedded in the whole markets. At the moment, the following topics are of interest for future research:

1. Categorising an investor's preference in the setting of model uncertainty.

2. Quantifying uncertainty embedded in the whole financial market.

3. Quantifying model uncertainty in the dynamic setting.

When these three topics are appropriately solved, theory on portfolio and derivative pricing can be investigated in more realistic settings.

Summary

Probabilistic models for financial markets play a crucial role in financial economics. The interplay between the theory of financial economics and its applications in financial markets promotes the tremendous advance in financial modeling. However, the complexity of financial markets and the limited capability of any single probabilistic model lead to a philosophical and technical problem — Model Uncertainty.

This thesis covers three topics involving model uncertainty in financial markets: numerical methods to quantify model uncertainty embedded in the pricing and hedging of derivative instruments, numerical methods to reduce the impact of model uncertainty, approaches to detect model mis-specification. Chapter 1 introduces the motivations for selecting model uncertainty as the theme of this thesis. After introducing some preliminaries of stochastic analysis and derivative pricing in Chapter 2, we present the main contributions in Chapter 3 to Chapter 7. Chapter 8 concludes this thesis and looks into the future of some promising topics on uncertainty in financial markets. The main contributions and ideas of this thesis can be summarized in the following way.

Chapter 3 investigates the discretisation of locally risk-minimising (LRM) strategies for vanilla options, Asian options, spread options and basket options in a class of semimartingale models, as well as the robustness of the LRM strategies with respect to model uncertainty.

We assume that the discounted stock price process S satisfies the stochastic differential equation (SDE)

$$\mathrm{d}S(t) = S(t-)\left[a(t)\,\mathrm{d}t + b(t)\,\mathrm{d}W(t) + \int_{\mathbb{R}} \gamma(t,z)\,\tilde{N}(\mathrm{d}t,\mathrm{d}z)\right], \qquad (3.9)$$

where a is Lipschitz, W and $\tilde{N}$ are a Brownian motion and a centered Poisson

random measure, respectively. Under some assumptions on the model parameters, the locally risk-minimising hedging strategy for a vanilla option or an Asian option can be formulated as the solution to a backward stochastic differential equation (BSDE),

$$\begin{cases} \mathrm{d}\tilde{V}(t) = \tilde{\Upsilon}(t)a(t)\,\mathrm{d}t + \tilde{\Upsilon}(t)b(t)\,\mathrm{d}W(t) + \displaystyle\int_{\mathbb{R}} \tilde{\Upsilon}(t)\gamma(t,z)\,\tilde{N}(\mathrm{d}t,\mathrm{d}z) + \mathrm{d}L^{FS}(t), \\ \tilde{V}(T) = \tilde{\xi}\,, \end{cases} \tag{3.12}$$

where $\tilde{\xi}$ is the discounted payoff, $\tilde{V}$ is the discounted value of the portfolio and the predictable process $\tilde{\Upsilon}$ defined as $\tilde{\Upsilon}(t) = \chi^{FS}(t)S(t-)$ is the amount of wealth to invest in an LRM strategy.

For spread options or basket options, the risky assets price processes S_i, $i = 1, 2$, satisfy the SDEs of the form (3.9) driven by correlated Brownian motions. Correspondingly, the counterpart of the BSDE (3.12) is

$$\begin{cases} \tilde{V}(t) = \tilde{V}(T) - \displaystyle\int_t^T \left\{\tilde{\Upsilon}_1(s)a_1(s) + \tilde{\Upsilon}_2(s)a_2(s)\right\}\mathrm{d}s - \int_t^T \tilde{\Upsilon}_1(s)\,\mathrm{d}\zeta_1(s) \\ \qquad\qquad - \displaystyle\int_t^T \tilde{\Upsilon}_2(s)\,\mathrm{d}\zeta_2(s) - L^{FS}(T) + L^{FS}(t)\,, \\ \tilde{V}(T) = \tilde{\xi}\,, \end{cases} \tag{3.59}$$

where $\tilde{\xi}$ is the discounted payoff, $\tilde{\Upsilon}_i(t) = \chi_i^{FS}(t)S_i(t-)$ is the amount of wealth invested in the risky asset S_i, $i = 1, 2$, in the setting of two underlying assets.

We simulate the LRM strategies by numerically solving the corresponding BSDEs. Using the discretisation scheme and the convergence results on backward stochastic differential equations as studied in [115], we show that the LRM strategies are robust towards the choice of the model and we derive an estimation of model uncertainty (see Theorem 3.12, Theorem 3.17, Theorem 3.18, Theorem 3.19 and Remark 3.20, Theorem 3.24, Theorem 3.25, Theorem 3.26).

Chapter 4 proposes an efficient numerical method to quantify uncertainty embedded in exotic derivatives.

Model uncertainty is characterized by a finite set of plausible models of different types, while parameter uncertainty is characterized by specifying each parameter in

an interval. Each parameter can take any value from its interval. Cont [49] proposed to quantify the uncertainty embedded in a derivative value in terms of its worst-case values. From this point of view, we only have to consider a special type of these models. Otherwise, we can calculate the worst-case values under the models of each type and find the bounds of these worst-case values. Hence, without loss of generality, we focus on parameter uncertainty in this chapter. In this setting, the key issue is how to efficiently calculate a large ensemble of the derivative value with a given large ensemble of the model parameters.

Combining the Monte Carlo method and the Smolyak interpolation algorithm (4.16), we propose an efficient numerical method to address the aforementioned key issue. With the resulting ensemble of derivative values, we can quantify the uncertainty embedded in these derivative values by at least two uncertainty measures.
The first one is the coherent uncertainty measure [49]

$$\mu_{\mathcal{Q}}(f) = \overline{P}(f) - \underline{P}(f)\,. \tag{4.4}$$

where $\overline{P}$ and $\underline{P}$ are the upper and lower bounds of the derivative value, which are approximated by

$$\overline{P}(f) \approx \max_{\mathbb{Q}\in\bar{\mathcal{Q}}} P^{\mathbb{Q}}(f) \quad \text{and} \quad \underline{P}(f) \approx \min_{\mathbb{Q}\in\bar{\mathcal{Q}}} P^{\mathbb{Q}}(f)\,, \tag{4.3}$$

where $\bar{\mathcal{Q}}$ is a collection of pricing models.
The second measure is the entropy of the derivative value

$$H(P) = \sum\nolimits_i p(x_i) \log \frac{1}{p(x_i)}\,, \tag{4.12}$$

where p is the empirical distribution of the derivative value.

Numerical examples are carried out to quantify parameter uncertainty embedded in the Bermudan put options and the down-and-out put options in the Heston model. The results show that for these two kinds of options, the at-the-money option ($K = 100, S_0 = 100$) embeds more uncertainty than the same kind of option with other moneyness. Comparatively speaking, more uncertainty is embedded in the in-the-money Bermudan put option ($K = 110, S_0 = 100$) than that in the out-of-the-money Bermudan put option ($K = 80, S_0 = 100$); it is the reverse order for the down-and-out put options. In addition, model uncertainty impacts more on the out-of-the-money and at-the-money down-and-out put options ($K = 80, 100; S_0 = 100$) than the corresponding Bermudan put options ($K = 80, 100; S_0 = 100$), respectively. However, the impact of parameter

uncertainty on the in-the-money down-and-out put option ($K = 110, S_0 = 100$) is stronger than that on the in-the-money Bermudan put option ($K = 110, S_0 = 100$). These results provide an investor with some necessary information to quote a derivative in the OTC market.

Chapter 5 proposes an efficient numerical method to calculate distorted expectations.

The distorted expectation of a random variable X is defined by

$$\mathbb{E}_\Psi[X] = \int_{-\infty}^{+\infty} x \, \mathrm{d}\Psi(F_X(x)),$$

where F_X is the distribution function and Ψ is a distortion function. Distorted expectations find extensive applications in finance and insurance.

Generally speaking, the distorted expectation does not admit an analytical formula, because the distorted density function is usually not given in the closed form. The Monte Carlo method is the standard approach to calculate distorted expectations. However, the Monte Carlo method has low efficiency. We propose an analytical approximation method for distorted expectations, the so-called T-COS method (Algorithm 5.1 and (5.20)).

In the Heston model, the results show that the T-COS method is about 2500 times more efficient than the standard Monte Carlo method with the saved and sorted paths. Note that simulating the price process costs much more time than loading the saved pathes. Hence, the numerical results highlight the advantage of the T-COS method over the standard Monte Carlo method.

Chapter 6 empirically investigates whether the weighted Monte Carlo method can reduce the uncertainty embedded in a derivative value.

The previous chapters mainly focus on how to quantify the impact of model uncertainty. In this chapter, we use the weighted Monte Carlo method to reduce the impact of model uncertainty on a derivative value.

Model uncertainty is characterized by the eighteen calibrated Heston models in [91]. Under each calibrated Heston model, we use the standard Monte Carlo method and the weighted Monte Carlo method to calculate the implied volatilities of vanilla options. The results are compared with the market data, respectively. The comparison shows that compared with the standard Monte Carlo method, the weighted Monte Carlo method can reduce the impact of model uncertainty in terms of implied volatility.

In addition, under each calibrated Heston model, we use the standard Monte Carlo method and the weighted Monte Carlo method to calculate the value of Asian options, barrier options and lookback options. The results also confirm that the weighted Monte Carlo method can reduce the impact of model uncertainty.

Chapter 7 proposes an efficient numerical method to calculate the comonotonicity-based upper value bounds of arithmetic Asian options, and empirically investigate how to detect model mis-specification.

Comonotonicity risk theory provides a super-replicating strategy and an upper value bound for a discrete arithmetic Asian option. The super-replicating strategy consists of trading European options with selected strikes, and the cost of this strategy corresponds to an upper value bound. This upper value bound is optimal in some sense. Calculation of the optimal strikes of the European options in the hedging portfolio is the key procedure in setting up the super-replicating strategy and calculating the optimal value bound. Recall that the super-replicating strategy can be set up either in the model-based framework or in the model-free framework. Hence, these optimal strikes have to be calculated in the model-based framework and the model-free framework with special numerical methods.

In the model-based framework, Albrecher et al. [1] proposed to recover the distribution function of the underlying asset price from its density function and use the bi-section method to search for the optimal strikes. However, it may be time-consuming to calculate the distribution function if the density function is not given in an analytical form, such as in the Heston model [103]. In this chapter, we propose an efficient method to accelerate the procedure of calculating the distribution function of the underlying asset in the model-based framework (7.13). In the model-free framework, Chen et al. [40] already proposed an efficient searching algorithm to calculate the optimal strikes. We follow this method in the model-free framework.

We show how to detect model mis-specification by comparing the model-based value bounds of Asian options with their model-free conterparts. We address some practical issues, such as the non-uniqueness of the marginal distribution for the underlying asset in the model-free setting.

Samenvatting

Probabilistische modellen voor financiële markten spelen een cruciale rol in de financiële economie. De wisselwerking tussen de theorie van de financiële economie en de toepassingen in financiële markten bevordert de enorme vooruitgang in de financiële modellering. Echter, de complexiteit van de financiële markten en de beperkte capaciteit van elk probabilistisch model afzonderlijk stimuleren de studie van een filosofisch en technisch probleem – Modelonzekerheid. Dit proefschrift behandelt drie thema's met betrekking tot modelonzekerheid in financiële markten: numerieke methoden om modelonzekerheid te kwantificeren die vervat zit in de prijsbepaling en het hedgen van afgeleide producten, numerieke methoden om de impact van modelonzekerheid te verminderen, technieken om verkeerde modelspecificatie te detecteren. Hoofdstuk 1 bevat de motivering waarom modelonzekerheid als onderwerp van deze scriptie gekozen werd. Na het invoeren in hoofdstuk 2 van een aantal basisbegrippen uit de stochastische analyse en de prijsbepaling van financieel afgeleide producten ook kortweg derivaten genoemd, presenteren we de belangrijkste bijdragen in de hoofdstukken 3 tot en met 7. Hoofdstuk 8 sluit dit proefschrift af en werpt een blik in de toekomst voor een aantal veelbelovende thema's betreffende onzekerheid in financiële markten. De belangrijkste bijdragen en ideeën van dit proefschrift kunnen als volgt worden samengevat.

Hoofdstuk 3 onderzoekt de discretisatie van lokaal-risicominimiserende (LRM) strategieën voor vanilla-opties, Aziatische opties, spread-opties en basket-opties voor een klasse van semimartingaalmodellen, alsook de robuustheid van de LRM strategieën met betrekking tot het modelleren van onzekerheid. We onderstellen dat het verdisconteerde aandelenkoersproces S voldoet aan de stochastische differenti-

aalvergelijking (SDE)

$$\mathrm{d}S(t) = S(t-)\left[a(t)\,\mathrm{d}t + b(t)\,\mathrm{d}W(t) + \int_{\mathbb{R}} \gamma(t,z)\tilde{N}(\mathrm{d}t,\mathrm{d}z)\right], \tag{3.9}$$

waarin a Lipschitz, W en $\tilde{N}$ respectievelijk een Brownse beweging en een gecentreerde Poisson toevalsmaat zijn. Onder bepaalde onderstellingen voor de modelparameters, kan de lokaal-risicominimiserende hedgingstrategie voor een vanilla-optie of een Aziatische optie geformuleerd worden als oplossing van een achterwaartse stochastische differentiaalvergelijking (BSDE),

$$\begin{cases} \mathrm{d}\tilde{V}(t) = \tilde{\Upsilon}(t)a(t)\,\mathrm{d}t + \tilde{\Upsilon}(t)b(t)\,\mathrm{d}W(t) + \displaystyle\int_{\mathbb{R}} \tilde{\Upsilon}(t)\gamma(t,z)\,\tilde{N}(\mathrm{d}t,\mathrm{d}z) + \mathrm{d}L^{FS}(t), \\ \tilde{V}(T) = \tilde{\xi}\,, \end{cases} \tag{3.12}$$

waarbij $\tilde{\xi}$ de verdisconteerde payoff voorstelt, $\tilde{V}$ de verdisconteerde waarde is van de hedgingportefeuille en het voorspelbare proces $\tilde{\Upsilon}$, gedefinieerd als $\tilde{\Upsilon}(t) = \chi^{FS}(t)S(t-)$, staat voor grootte van de investering in het onderliggende activum S.

De prijzen S_i, $i = 1,\ldots,d$, van de risicovolle activa voor spread-opties of basket-opties voldoen aan de SDEs van de vorm (3.9) aangedreven door gecorreleerde Brownse bewegingen. De tegenhanger van de BSDE (3.12) in het geval van twee onderliggende activa luidt als volgt

$$\begin{cases} \tilde{V}(t) = \tilde{V}(T) - \displaystyle\int_t^T \left\{\tilde{\Upsilon}_1(s)a_1(s) + \tilde{\Upsilon}_2(s)a_2(s)\right\}\mathrm{d}s - \int_t^T \tilde{\Upsilon}_1(s)\,\mathrm{d}\zeta_1(s) \\ \qquad\qquad - \displaystyle\int_t^T \tilde{\Upsilon}_2(s)\,\mathrm{d}\zeta_2(s) - L^{FS}(T) + L^{FS}(t)\,, \\ \tilde{V}(T) = h(S^{(\cdot)}(T))\,, \end{cases} \tag{3.59}$$

waarbij $\tilde{\xi}$ de verdisconteerde payoff is, $\tilde{\Upsilon}_i(t) = \chi_i^{FS}(t)S_i(t-)$ de grootte van investering in de risicovolle activa S_i, $i = 1,2$.

We simuleren de LRM strategieën door de overeenkomstige BSDEs numeriek op te lossen. Gebruik makend van een discretisatieschema en de convergentieresultaten voor achterwaartse stochastische differentiaalvergelijkingen zoals bestudeerd in [115] wordt aangetoond dat de LRM strategieën robuust zijn met betrekking tot de keuze

van het model en we leiden een schatting af voor modelonzekerheid (zie Stelling 3.12, Stelling 3.17, Stelling 3.18, Stelling 3.19 en Opmerking 3.20, Stelling 3.24, Stelling 3.25, Stelling 3.26).

Hoofdstuk 4 stelt een efficiënte numerieke methode voor om de onzekerheid ingebed in exotische derivaten te kwantificeren.

Modelonzekerheid wordt gekenmerkt door een eindige verzameling van mogelijke modellen van verschillende types, terwijl parameteronzekerheid wordt gekenmerkt door het specificeren van elke parameter in een interval. Elke parameter kan elke waarde in zo'n interval aannemen. Cont [49] suggereerde om de onzekerheid vervat in de waarde van een derivaat te kwantificeren via worst-case waarden. Vanuit dit oogpunt hoeven we enkel een speciaal type van deze modellen in overweging te nemen. Anderzijds kunnen we de waarde voor de worst-case berekenen voor de modellen van elk type en de grenzen van deze worst-case waarden zoeken. Vandaar dat we ons in dit hoofdstuk kunnen focussen op parameteronzekerheid zonder verlies van algemeenheid. In deze context is de belangrijkste vraag hoe we een grote verzameling van waarden voor derivaten efficiënt kunnen berekenen voor een grote verzameling van de modelparameters.

We stellen een efficiënte numerieke methode voor om deze belangrijke vraag aan te pakken en dit gebaseerd op een combinatie van de Monte-Carlomethode en het Smolyak-interpolatie-algoritme (4.16). Met de resulterende verzameling van waarden voor de derivaten kunnen we de onzekerheid ingebed in deze waarden kwantificeren aan de hand van twee onzekerheidsmaten. De eerste is de coherente onzekerheidsmaatstaf [49]

$$\mu_{\mathcal{Q}}(f) = \overline{P}(f) - \underline{P}(f)\,. \tag{4.4}$$

waarbij $\overline{P}$ en $\underline{P}$ de boven- en ondergrenzen zijn van de waarde van het derivaat, die benaderd worden door

$$\overline{P}(f) \approx \max_{\mathbb{Q}\in\bar{\mathcal{Q}}} P^{\mathbb{Q}}(f) \quad \text{and} \quad \underline{P}(f) \approx \min_{\mathbb{Q}\in\bar{\mathcal{Q}}} P^{\mathbb{Q}}(f)\,, \tag{4.3}$$

met $\bar{\mathcal{Q}}$ een verzameling van prijsmodellen.
De tweede maat is de entropie van de waarde van het derivaat

$$H(P) = \sum\nolimits_i p(x_i) \log \frac{1}{p(x_i)}\,, \tag{4.12}$$

waarbij p staat voor de empirische verdeling van de waarde van het derivaat.

Numerieke voorbeelden worden uitgewerkt om de parameteronzekerheid in Bermudan putopties en in down-and-out putopties te kwantificeren voor het Hestonmodel. De

resultaten tonen aan dat voor deze twee soorten opties, de in-the-money optie ($K = 100$, $S_0 = 100$) meer onzekerheid bevat dan dezelfde soort optie met een andere moneyness. Relatief gezien, bevat de in-the-money Bermudan putoptie ($K = 110$, $S_0 = 100$) meer onzekerheid dan de out-of-the-money Bermudan putoptie ($K = 80$, $S_0 = 100$); het omgekeerde geldt voor de down-and-out putopties. Bovendien heeft modelonzekerheid meer invloed op de out-of-the-money en at-the-money down-and-out putopties ($K = 80$, 100, $S_0 = 100$) dan op de overeenkomstige Bermudan putopties ($K = 80$, 100, $S_0 = 100$). Echter, de impact van parameteronzekerheid op een in-the-money down-and-out putoptie ($K = 110$, $S_0 = 100$) is groter dan die op de in-the-money Bermudan putoptie ($K = 110$, $S_0 = 100$). Deze resultaten geven een investeerder de nodige informatie om een afgeleid product te prijzen in de over-the-counter-markt.

Hoofdstuk 5 stelt een efficiënte numerieke methode voor om vervormde verwachtingswaarden te berekenen.

De vervormde verwachtingswaarde van een stochastische variabele X wordt gedefinieerd als

$$\mathbb{E}_\Psi[X] = \int_{-\infty}^{+\infty} x \, \mathrm{d}\Psi(F_X(x)),$$

waarbij F_X staat voor de distributiefunctie en Ψ voor een vervormingsfunctie. Vervormde verwachtingswaarden vinden uitgebreide toepassingen in financiën en verzekeringen. Over het algemeen bestaat er voor de vervormde verwachtingswaarde geen analytische formule, omdat de vervormde dichtheidsfunctie meestal niet in gesloten vorm bestaat. De Monte-Carlomethode is de standaard aanpak om dergelijke vervormde verwachtingswaarden te berekenen. Echter, de Monte-Carlomethode is niet heel efficiënt. Wij stellen een analytische benaderingsmethode voor voor de vervormde verwachtingswaarden, namelijk de zogenaamde T-COSmethode (algoritme 5.1 en (5.20)). In het Hestonmodel geven de numerieke resultaten aan dat de T-COSmethode ongeveer 2500 keer efficiënter is dan de standaard Monte-Carlomethode met opgeslagen en gesorteerde paden. Merk op dat het simuleren van het prijsproces veel meer tijd kost dan het laden van de opgeslagen paden. Vandaar dat de numerieke resultaten de voordelen van de T-COSmethode ten opzichte van de standaard Monte-Carlomethode extra in de verf zetten.

Hoofdstuk 6 onderzoekt op empirische wijze of de gewogen Monte-Carlomethode de onzekerheid in de waarde van een derivaat kan verminderen.

De focus van de voorgaande hoofdstukken was vooral hoe de impact van mod-

elonzekerheid te kwantificeren. In dit hoofdstuk gebruiken we de gewogen Monte-Carlomethode om de impact van modelonzekerheid op de waarde van een afgeleid product te verminderen. Modelonzekerheid wordt hier gekenmerkt door de achttien gecalibreerde Hestonmodellen uit [91]. Voor elke gecalibreerd Hestonmodel, gebruiken we de standaard Monte-Carlomethode en de gewogen Monte-Carlomethode om de impliciete volatiliteit bij vanilla-opties te berekenen. De resultaten worden vergeleken met de overeenkomstige marktgegevens. De vergelijking toont aan dat in tegenstelling tot de standaard Monte-Carlomethode, de gewogen Monte-Carlomethode de impact van modelonzekerheid bij de impliciete volatiliteit kan verminderen. Daarnaast berekenen we, onder elk gecalibreerd Hestonmodel, met behulp van zowel de standaard Monte-Carlomethode als de gewogen Monte-Carlomethode de waarde van Aziatische opties, barrier-opties en lookback-opties. De resultaten bevestigen ook hier dat de gewogen Monte-Carlomethode de impact van modelonzekerheid kan verminderen.

Hoofdstuk 7 stelt een efficiënte numerieke methode voor om de op comonotoniciteit gebaseerde bovengrenzen van rekenkundige Aziatische opties te berekenen en onderzoekt empirisch hoe foutieve modellen te detecteren.

Risicotheorie gebaseerd op comonotoniciteit zorgt voor een superreplicerende strategie en een bovengrens voor discrete rekenkundige Aziatische optieprijzen. De superreplicerende strategie bestaat uit het verhandelen van Europese opties met geselecteerde uitoefenprijzen en uitoefendata, en de kosten van deze strategie komen overeen met de waarden van een bovengrens. Deze bovengrens is optimaal in een welbepaalde zin. Berekening van de optimale uitoefenprijzen van de Europese opties in de hedgingportefeuille is de sleutelprocedure bij het opzetten van de superreplicerende strategie en bij het berekenen van de corresponderende optimale waarde. Deze superreplicerende strategie kan worden opgezet hetzij in een modelgebaseerde context hetzij in een modelvrije context. De optimale uitoefenprijzen moeten worden berekend met speciale numerieke methoden in de modelgebaseerde context en in een modelvrije context.

Albrecher et al. [1] stelden voor om in de modelgebaseerde context de distributiefunctie van de prijs van het onderliggend activum te halen uit de dichtheidsfunctie ervan en om via de bisectiemethode de optimale uitoefenprijzen te berekenen. De berekening van de verdelingsfunctie kan echter heel tijdrovend zijn als de dichtheidsfunctie niet in een analytische vorm beschikbaar is, zoals in het geval van het Hestonmodel [103]. In dit hoofdstuk stellen we een efficiënte manier voor om de

berekeningsprocedure voor de distributiefunctie (7.13) van de onderliggende waarde te versnellen in de modelgebaseerde context. In de modelvrije context hebben Chen et al. [40] reeds een efficiënt zoekalgoritme beschreven om de optimale uitoefenprijzen te berekenen. We volgen deze methode in het modelvrije kader.

We laten verder zien hoe voor Aziatische opties de specificatie van een verkeerd model kan gedetecteerd worden door het vergelijken van de modelgebaseerde grenzen met de overeenkomstige modelvrije grenzen. We bespreken een aantal praktische zaken, zoals het niet-unieke karakter van de marginale distributie voor de onderliggende waarde in het modelvrije kader.

Bibliography

[1] H. ALBRECHER, J. DHAENE, M. GOOVAERTS, AND W. SCHOUTENS, *Static hedging of Asian options under Lévy models*, The Journal of Derivatives, 12 (2005), pp. 63–72.

[2] H. ALBRECHER, F. GUILLAUME, AND W. SCHOUTENS, *Implied liquidity: model sensitivity*, Journal of Empirical Finance, 23 (2013), pp. 48–67.

[3] J. ANSEL AND C. STRICKER, *Lois de martingale, densités et décomposition de Föllmer-Schweizer*, Annales de l'Institut Henri-Poincaré Probabilités et Statistiques, 28 (1992), pp. 375–392.

[4] D. APPLEBAUM, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, Cambridge, 2nd ed., 2009.

[5] T. ARAI AND R. SUZUKI, *Local risk-minimization for Lévy markets*, International Journal of Financial Engineering, 02 (2015), p. 1550015.

[6] P. ARTZNER, F. DELBAEN, J. EBER, AND D. HEATH, *Coherent measures of risk*, Mathematical Finance, 9 (1999), pp. 203–228.

[7] S. ASMUSSEN AND J. ROSIŃSKI, *Approximations of small jumps of lévy processes with a view towards simulation*, Journal of Applied Probability, (2001), pp. 482–493.

[8] S. ASSING, S. JACKA, AND A. OCEJO, *Monotonicity of the value function for a two-dimensional optimal stopping problem*, The Annals of Applied Probability, 24 (2014), pp. 1554–1584.

[9] M. AVELLANEDA, R. BUFF, C. FRIEDMAN, N. GRANDECHAMP, L. KRUK, AND J. NEWMAN, *Weighted Monte Carlo: a new technique for calibrating asset-pricing models*, International Journal of Theoretical and Applied Finance, 04 (2001), pp. 91–119.

[10] M. AVELLANEDA, C. FRIEDMAN, R. HOLMES, AND D. SAMPERI, *Calibrating volatility surfaces via relative-entropy minimization*, Applied Mathematical Finance, 4 (1997), pp. 37–64.

[11] M. AVELLANEDA, A. LEVY, AND A. PARÁS, *Pricing and hedging derivative securities in markets with uncertain volatilities*, Applied Mathematical Finance, 2 (1995), pp. 73–88.

[12] V. BALLY, L. CARAMELLINO, AND A. ZANETTE, *Pricing and hedging American options by Monte Carlo methods using a Malliavin calculus approach*, Monte Carlo Methods and Applications, 11 (2005), pp. 97–133.

[13] K. BANNÖR AND M. SCHERER, *Capturing parameter risk with convex risk measures*, European Actuarial Journal, 3 (2013), pp. 97–132.

[14] N. BARBERIS, *American economic association thirty years of prospect theory in economics: a review and assessment*, Journal of Economic Perspectives, 27 (2013), pp. 173–195.

[15] N. BARBERIS AND M. HUANG, *Stocks as lotteries: the implications of probability weighting for security prices*, American Economic Review, 98 (2008), pp. 2066–2100.

[16] V. BARTHELMANN, E. NOVAK, AND K. RITTER, *High dimensional polynomial interpolation on sparse grids*, Advances in Computational Mathematics, 12 (2000), pp. 273–288.

[17] E. BAYRAKTAR AND S. YAO, *On the robust optimal stopping problem*, arXiv preprint arXiv: 1301.0091v8, (2014), pp. 1–50.

[18] E. BAYRAKTAR AND Y. ZHANG, *Minimizing the probability of lifetime ruin under ambiguity aversion*, SIAM Journal on Control and Optimization, 53 (2015), pp. 58–90.

[19] M. BEIGLBÖCK, P. HENRY-LABORDÈRE, AND F. PENKNER, *Model-independent bounds for option prices – a mass transport approach*, Finance and Stochastics, 17 (2013), pp. 477–501.

[20] F. BENTH AND N. DETERING, *Pricing and hedging Asian-style options in energy*, Available at SSRN 2330273, (2014), pp. 1–36.

[21] F. BENTH, G. DI NUNNO, AND A. KHEDHER, *Robustness of option prices and their Deltas in markets modelled by jump-diffusions*, Communications on Stochastic Analysis, 5 (2011), pp. 285–307.

[22] S. BIAGINI, B. BOUCHARD, C. KARDARAS, AND M. NUTZ, *Robust fundamental theorem for continuous processes*, working paper, (2014), pp. 1–31.

[23] BIS, *Revisions to the Basel II market risk framework – updated as of 31 December 2010*, no. December 2010, Bank for International Settlements, available online at: http://www.bis.org/publ/bcbs193.htm, 2011.

[24] F. BLACK AND M. SCHOLES, *The pricing of options and corporate liabilities*, The journal of political economy, 81 (1973), pp. 637–654.

[25] C. BOSE AND R. MURRAY, *Maximum entropy estimates for risk-neutral probability measures with non-strictly-convex data*, Journal of Optimization Theory and Applications, (2013), pp. 1–23.

[26] B. BOUCHARD AND R. ELIE, *Discrete time approximation of decoupled forward backward SDE with jumps*, Stochastic Processes and Their Applications, 118 (2008), pp. 53–75.

[27] J. BOYD, *Chebyshev and Fourier Spectral Methods*, Dover Publications, Inc., New York, 2nd ed., 2000.

[28] J. BRISSAUD, *The meanings of entropy*, Entropy, 7 (2005), pp. 68–96.

[29] M. BROADIE AND P. GLASSERMAN, *Estimating security price derivatives using simulation*, Management science, 42 (1996), pp. 269–285.

[30] P. BUCHEN AND M. KELLY, *The maximum entropy distribution of an asset inferred from option prices*, The Journal of Financial and Quantitative Analysis, 31 (1996), pp. 143–159.

[31] M. CAMBOU AND D. FILIPOVIĆ, *Model uncertainty and scenario aggregation*, Mathematical Finance(accepted), (2015).

[32] R. CARBONE, B. FERRARIO, AND M. SANTACROCE, *Backward stochastic differential equations driven by càdlàg martingales*, Theory of Probability and Its Applications, 52 (2008), pp. 304–314.

[33] R. CARMONA AND V. DURRLEMAN, *Pricing and hedging spread options*, SIAM Review, 45 (2003), pp. 627–685.

[34] P. CARR, H. GEMAN, AND D. MADAN, *Pricing and hedging in incomplete markets*, Journal of Financial Economics, 62 (2001), pp. 131–167.

[35] P. CARR AND D. MADAN, *Option valuation using the fast Fourier transform*, Journal of Computational Finance, 2 (1999), pp. 1–18.

[36] A. CHEN, F. HENTSCHEL, AND J. K. KLEIN, *A utility- and CPT-based comparison of life insurance contracts with guarantees*, Journal of Banking & Finance, 61 (2015), pp. 327 – 339.

[37] J. CHEN AND C. EWALD, *On the performance of the comonotonicity approach for pricing Asian option in some benchmark models from equities and commodities*, Working paper, (2014).

[38] N. CHEN AND Y. LIU, *American option sensitivities estimation via a generalized infinitesimal perturbation analysis approach*, Operations Research, 62 (2014), pp. 616–632.

[39] X. CHEN, G. DEELSTRA, J. DHAENE, D. LINDERS, AND M. VANMAELE, *On an optimization problem related to static super-replicating strategies*, Journal of Computational and Applied Mathematics, 278 (2015), pp. 213–230.

[40] X. CHEN, G. DEELSTRA, J. DHAENE, AND M. VANMAELE, *Static super-replicating strategies for a class of exotic options*, Insurance: Mathematics and Economics, 42 (2008), pp. 1067–1085.

[41] Z. CHEN AND L. EPSTEIN, *Ambiguity, risk, and asset returns in continuous time*, Econometrica, 70 (2002), pp. 1403–1443.

[42] X. CHENG AND F. RIEDEL, *Optimal stopping under ambiguity in continuous time*, Mathematics and Financial Economics, 7 (2013), pp. 29–68.

[43] A. CHERNY, *Weighted V@R and its properties*, Finance and Stochastics, 10 (2006), pp. 367–393.

[44] A. CHERNY AND D. MADAN, *New measures for performance evaluation*, Review of Financial Studies, 22 (2009), pp. 2571–2606.

[45] C. CHIARELLA, B. KANG, AND G. H. MEYER, *The evaluation of barrier option prices under stochastic volatility*, Computers & Mathematics with Applications, 64 (2012), pp. 2034–2048.

[46] T. CHOULLI, L. KRAWCZYK, , AND C. STRICKER, *$\mathcal{E}$-martingales and their applications in mathematical finance*, Annals of Probability, 26 (1998), pp. 853–876.

[47] T. CHOULLI, N. VANDAELE, AND M. VANMAELE, *The Föllmer–Schweizer decomposition: comparison and description*, Stochastic Processes and their Applications, 120 (2010), pp. 853 – 872.

[48] S. CHRISTENSEN, *Optimal decision under ambiguity for diffusion processes*, Mathematical Methods of Operations Research, 77 (2013), pp. 207–226.

[49] R. CONT, *Model uncertainty and its impact on the pricing of derivative instruments*, Mathematical Finance, 16 (2006), pp. 519–547.

[50] R. CONT AND R. DEGUEST, *Equity correlations implied by index options: estimation and model uncertainty analysis*, Mathematical Finance, 23 (2013), pp. 496–530.

[51] R. CONT AND P. TANKOV, *Financial Modelling with Jump Processes*, Chapman Hall, 2004.

[52] A. COX AND C. HOEGGERL, *Model-independent no-arbitrage conditions on American put options*, Mathematical Finance (accepted), (2013).

[53] A. COX AND J. OBŁÓJ, *Robust pricing and hedging of double no-touch options*, Finance and Stochastics, 15 (2011), pp. 573–605.

[54] J. DANIELSSON AND K. JAMES, *Model risk of risk models*, Available at SSRN: 2425689, (2015).

[55] C. DAVELOOSE, A. KHEDHER, AND M. VANMAELE, *Representations for conditional expectations and applications to pricing and hedging of financial products in Lévy and jump-diffusion setting*, UGent working paper, (2014).

[56] G. DEELSTRA, I. DIALLO, AND M. VANMAELE, *Bounds for Asian basket options*, Journal of Computational and Applied Mathematics, 218 (2008), pp. 215–228.

[57] F. DELBAEN AND W. SCHACHERMAYER, *The Mathematics of Arbitrage*, Springer-Verlag, Berlin, 2006.

[58] L. DENIS, M. HU, AND S. PENG, *Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths*, Potential Analysis, 34 (2010), pp. 139–161.

[59] L. DENIS AND M. KERVAREC, *Optimal investment under model uncertainty in nondominated models*, SIAM Journal on Control and Optimization, 51 (2013), pp. 1803–1822.

[60] L. DENIS AND C. MARTINI, *A theoretical framework for the pricing of contingent claims in the presence of model uncertainty*, The Annals of Applied Probability, 16 (2006), pp. 827–852.

[61] J. DHAENE, M. DENUIT, M. J. GOOVAERTS, R. KAAS, AND D. VYNCKE, *The concept of comonotonicity in actuarial science and finance: theory*, Insurance: Mathematics and Economics, 31 (2002), pp. 3–33.

[62] J. DHAENE, A. KUKUSH, D. LINDERS, AND Q. TANG, *Remarks on quantiles and distortion risk measures*, European Actuarial Journal, 2 (2012), pp. 319–328.

[63] G. DI NUNNO, A. KHEDHER, AND M. VANMAELE, *Robustness of quadratic hedging strategies in finance via backward stochastic differential equations with jumps*, Applied Mathematics and Optimization, 73 (2015), pp. 353–389.

[64] G. DI NUNNO, B. K. ØKSENDAL, AND F. PROSKE, *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer-Verlag, Berlin, 2009.

[65] A. DRĂGULESCU AND V. M. YAKOVENKO, *Probability distribution of returns in the Heston model with stochastic volatility*, Quantitative finance, 2 (2002), pp. 443–453.

[66] M. DUEMBGEN AND L. C. G. ROGERS, *Estimate nothing*, Quantitative Finance, 14 (2014), pp. 2065–2072.

[67] D. DUFFIE, N. GARLEANU, AND L. H. PEDERSEN, *Valuation in over-the-counter markets*, Review of Financial Studies, 20 (2007), pp. 1865–1900.

[68] E. EBERLEIN, D. MADAN, M. PISTORIUS, W. SCHOUTENS, AND M. YOR, *Two price economies in continuous time*, Annals of Finance, 10 (2013), pp. 71–100.

[69] E. EBERLEIN, D. MADAN, M. PISTORIUS, AND M. YOR, *Bid and ask prices as non-linear continuous time g-expectations based on distortions*, Mathematics and Financial Economics, 8 (2014), pp. 265–289.

[70] A. ELICES AND E. GIMÉNEZ, *Weighted Monte Carlo: calibrating the smile and preserving martingale condition*, arXiv:1102.3541, (2011).

[71] P. EMBRECHTS, G. PUCCETTI, AND L. RÜSCHENDORF, *Model uncertainty and VaR aggregation*, Journal of Banking & Finance, 37 (2013), pp. 2750–2764.

[72] L. EPSTEIN AND S. JI, *Ambiguous volatility and asset pricing in continuous time*, Review of Financial Studies, 26 (2013), pp. 1740–1786.

[73] L. EPSTEIN AND S. JI, *Ambiguous volatility, possibility and utility in continuous time*, Journal of Mathematical Economics, 50 (2014), pp. 269–282.

[74] B. ERAKER, *Do stock prices and volatility jump? Reconciling evidence from spot and option prices*, The Journal of Finance, 59 (2004), pp. 1367–1404.

[75] J. ETNER, M. JELEVA, AND J.-M. TALLON, *Decision theory under ambiguity*, Journal of Economic Surveys, 26 (2012), pp. 234–270.

[76] F. FANG AND C. OOSTERLEE, *A novel pricing method for European options based on Fourier-cosine series expansions*, SIAM Journal on Scientific Computing, 31 (2008), pp. 826–848.

[77] F. FANG AND C. OOSTERLEE, *A Fourier-based valuation method for bermudan and barrier options under Heston's model*, SIAM Journal on Financial Mathematics, 2 (2011), pp. 439–463.

[78] FEDERAL RESERVE, *Supervisory guidance on model risk management*, vol. 2011, Board of Governors of the Federal Reserve System Office, Office of the Comptroller of the Currency, available online at: http://www.occ.treas.gov/news-issuances/bulletins/2011/bulletin-2011-12.html, 2011.

[79] D. FILIPOVIĆ, M. KUPPER, AND N. VOGELPOTH, *Approaches to conditional risk*, SIAM Journal on Financial Mathematics, 3 (2012), pp. 402–432.

[80] H. FÖLLMER AND A. SCHIED, *Stochastic Finance*, De Gruyter, Berlin, Newyork, 3rd ed., 2011.

[81] H. FÖLLMER AND M. SCHWEIZER, *Hedging of contingent claims under incomplete information*, in Applied Stochastic Analysis, M. Davis and R. Elliott, eds., vol. 5 of Stochastic Monographs, Gordon and Breach Science Publishers, New York, 1991, pp. 389–414.

[82] H. FÖLLMER AND D. SONDERMANN, *Hedging of non-redundant contingent claims*, in Contributions to Mathematical Economics, W. Hildenbrand and A. Mas-Colell, eds., Elsevier, North-Holland, 1986, pp. 205–223.

[83] C. FRIEDMAN, W. CAO, Y. HUANG, AND Y. ZHANG, *Engineering more effective weighted Monte Carlo option pricing models*, Available at SSRN 2193807, (2013).

[84] A. GALICHON, P. HENRY-LABORDÈRE, AND N. TOUZI, *A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options*, The Annals of Applied Probability, 24 (2014), pp. 312–336.

[85] F. GAMBOA AND E. GASSIAT, *Bayesian methods and maximum entropy for ill-posed inverse problems*, The Annals of Statistics, 25 (1997), pp. 328–350.

[86] C. GASQUET AND P. WITOMSKI, *Fouier Analysis and Applications*, Springer-Verlag, New York, 1999.

[87] M. GASS, K. GLAU, M. MAHLSTEDT, AND M. MAIR, *Chebyshev interpolation for parametric option pricing*, arXiv preprint arXiv:1505.04648, (2015).

[88] P. GLASSERMAN, *Monte Carlo Methods in Financial Engineering*, Springer, New York, 2004.

[89] P. GLASSERMAN AND X. XU, *Robust risk measurement and model risk*, Quantitative Finance, 14 (2014), pp. 29–58.

[90] M. GOOVAERTS, R. KAAS, AND R. J. LAEVEN, *Decision principles derived from risk measures*, Insurance: Mathematics and Economics, 47 (2010), pp. 294–302.

[91] F. GUILLAUME AND W. SCHOUTENS, *Calibration risk: illustrating the impact of calibration risk under the Heston model*, Review of Derivatives Research, 15 (2012), pp. 57–79.

[92] L. GULKO, *The entropy theory of stock option pricing*, International Journal of Theoretical and Applied Finance, 02 (1999), pp. 331–355.

[93] T. GUO, S. ZHAO, AND X. ZENG, *The relations among the three kinds of conditional risk measures*, Science China Mathematics, 57 (2014), pp. 1753–1764.

[94] A. GUPTA AND C. REISINGER, *Robust calibration of financial models using Bayesian estimators*, Journal of Computational Finance, 17 (2014), pp. 3–36.

[95] T. HAENTJENS AND K. J. IN'T HOUT, *ADI schemes for pricing American options under the Heston model*, Applied Mathematical Finance, 22 (2015), pp. 207–237.

[96] L. HANSEN, *Nobel lecture: Uncertainty outside and inside economic models*, Journal of Political Economy, 122 (2014), pp. 945–987.

[97] L. HANSEN AND T. SARGENT, *Robustness*, Princeton University Press, 2008.

[98] ——, *Sets of models and prices of uncertainty*, working paper, (2015).

[99] X. HE AND X. Y. ZHOU, *Portfolio choice under cumulative prospect theory: an analytical treatment*, Management Science, 57 (2011), pp. 315–331.

[100] Z. HE, S. LI, B. WEI, AND J. YU, *Uncertainty, risk, and incentives: theory and evidence*, Management Science, 60 (2014), pp. 206–226.

[101] T. HELLMANN AND J. J. J. THIJSSEN, *Ambiguity in a real option game*, working paper, (2015), pp. 1–26.

[102] P. HÉNAFF AND C. MARTINI, *Model validation: theory, practice and perspectives*, The Journal of Risk Model Validation, 5 (2011), pp. 3–15.

[103] S. HESTON, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Review of Financial Studies, 6 (1993), pp. 327–343.

[104] D. HOBSON, *Robust hedging of the lookback option*, Finance and Stochastics, 2 (1998), pp. 329–347.

[105] ——, *The Skorokhod embedding problem and model-independent bounds for option prices*, in Paris-Princeton Lectures on Mathematical Finance 2010, A. Cousin, S. Crépey, O. Guéant, D. Hobson, M. Jeanblanc, J.-M. Lasry, J.-P. Laurent, P.-L. Lions, and P. Tankov, eds., vol. 2003 of Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, 2011, pp. 267–318.

[106] D. HOBSON AND M. KLIMMEK, *Model-independent hedging strategies for variance swaps*, Finance and Stochastics, 16 (2012), pp. 611–649.

[107] D. HOBSON, P. LAURENCE, AND T.-H. WANG, *Static-arbitrage upper bounds for the prices of basket options*, Quantitative Finance, 5 (2005), pp. 329–342.

[108] F. HUBALEK, J. KALLSEN, AND L. KRAWCZYK, *Variance-optimal hedging for processes with stationary independent increments*, Annals of Applied Probability, 16 (2006), pp. 853–885.

[109] K. IN'T HOUT AND S. FOULON, *ADI finite difference schemes for option pricing in the Heston model with correlation*, International Journal of Numerical Analysis & Modeling, 7 (2010), pp. 303–320.

[110] J. JACOD AND A. SHIRYAEV, *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin, 2nd ed., 2002.

[111] M. JEANBLANC, M. MANIA, M. SANTACROCE, AND M. SCHWEIZER, *Mean-variance hedging via stochastic control and BSDEs for general semimartingales*, The Annals of Applied Probability, 22 (2012), pp. 2388–2428.

[112] R. KAAS, J. DHAENE, AND M. J. GOOVAERTS, *Upper and lower bounds for sums of random variables*, Insurance: Mathematics and Economics, 27 (2000), pp. 151–168.

[113] C. KAHL AND H. SCHURZ, *Balanced Milstein methods for ordinary SDEs*, Monte Carlo Methods and Applications, 12 (2006), pp. 143–170.

[114] B. KANG, T. KIM, AND H. LEE, *Option-implied preference with model uncertainty*, Journal of Futures Markets, 34 (2014), pp. 498–515.

[115] A. KHEDHER AND M. VANMAELE, *Discretisation of FBSDEs driven by càdlàg martingales*, Journal of Mathematics Analysis and Application, 435 (2016), pp. 508–531.

[116] A. KLIMKE AND B. WOHLMUTH, *Algorithm 847: Spinterp: piecewise multilinear hierarchical sparse grid interpolation in MATLAB*, ACM Transactions on Mathematical Software, 31 (2005), pp. 561–579.

[117] F. KNIGHT, *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston, 1921.

[118] M. KOHLER, *A review on regression-based Monte Carlo methods for pricing American options*, Recent Developments in Applied Probability and Statistics, (2010), pp. 37–58.

[119] H. KUNITA AND S. WATANABE, *On square integrable martingales*, Nagoya Mathematical Journal, 30 (1967), pp. 209–245.

[120] D. LEMMENS, L. LIANG, J. TEMPERE, AND A. DE SCHEPPER, *Pricing bounds for discrete arithmetic Asian options under Lévy models*, Physica A: Statistical Mechanics and its Applications, 389 (2010), pp. 5193–5207.

[121] Q. LIN AND F. RIEDEL, *Optimal consumption and portfolio choice with ambiguity*, working paper, (2014).

[122] D. LINDERS, J. DHAENE, H. HOUNNON, AND M. VANMAELE, *Index options: a model-free approach*, KU Leuven: Working paper, (2012).

[123] F. LONGSTAFF AND E. SCHWARTZ, *Valuing American options by simulation: a simple least-squares approach*, Review of Financial Studies, 14 (2001), pp. 113–147.

[124] R. LORD, *Partially exact and bounded approximations for arithmetic Asian options*, Available at SSRN 678041, (2005).

[125] W. LV, X. PAN, AND T. HU, *Asymptotics of the risk concentration based on the tail distortion risk measure*, Statistics & Probability Letters, 83 (2013), pp. 2703 – 2710.

[126] T. LYONS, *Uncertain volatility and the risk-free synthesis of derivatives*, Applied Mathematical Finance, 2 (1995), pp. 117–133.

[127] D. MADAN, *Conic portfolio theory*, working paper, (2015).

[128] D. MADAN AND A. CHERNY, *Markets as a counterparty: an introduction to conic finance*, International Journal of Theoretical and Applied Finance, 13 (2010), pp. 1149–1177.

[129] D. MADAN, M. PISTORIUS, AND M. STADJE, *On consistent valuations based on distorted expectations: from multinomial random walks to Lévy processes*, working paper, (2013), pp. 1–42.

[130] ——, *On dynamic spectral risk measures and a limit theorem*, working paper, (2015), pp. 1–53.

[131] D. MADAN AND W. SCHOUTENS, *Tenor specific pricing*, International Journal of Theoretical and Applied Finance, 12 (2012).

[132] R. MERTON, *Theory of rational option pricing*, The Bell Journal of Economics and Management Science, 4 (1973), pp. 141–183.

[133] R. MERTON, *Option pricing when underlying stock returns are discontinuous*, Journal of Financial Economics, 3 (1976), pp. 125–144.

[134] A. NARAYAN AND D. XIU, *Constructing nested nodal sets for multivariate polynomial interpolation*, SIAM Journal on Scientific Computing, 35 (2013), pp. A2293–A2315.

[135] C. NERI AND L. SCHNEIDER, *Maximum entropy distributions inferred from option portfolios on an asset*, Finance and Stochastics, 16 (2012), pp. 293–318.

[136] J. NIELSEN AND K. SANDMANN, *Pricing bounds on Asian options*, Journal of Financial and Quantitative Analysis, 38 (2003), pp. 449–473.

[137] D. NUALART, *The Malliavin Calculus and Related Topics*, Springer-Verlag, Berlin, 2006.

[138] M. NUTZ, *Utility maximization under model uncertainty in discrete time*, Mathematical Finance, (2014), pp. 1–17.

[139] M. NUTZ AND H. M. SONER, *Superhedging and dynamic risk measures under volatility uncertainty*, SIAM Journal on Control and Optimization, 50 (2012), pp. 2065–2089.

[140] L. ORTIZ-GRACIA AND C. W. OOSTERLEE, *Efficient VaR and expected shortfall computations for nonlinear portfolios within the delta-gamma approach*, Applied Mathematics and Computation, 244 (2014), pp. 16–31.

[141] S. OULD ALY, *Monotonicity of prices in Heston model*, International Journal of Theoretical and Applied Finance, 16 (2013), p. 1350016.

[142] S. PARK AND A. BERA, *Maximum entropy autoregressive conditional heteroskedasticity model*, Journal of Econometrics, 150 (2009), pp. 219–230.

[143] S. PENG, *Nonlinear expectations and stochastic calculus under uncertainty*, arXiv preprint arXiv:1002.4546, (2010), pp. 1–149.

[144] E. PETROU, *Malliavin calculus in Lévy spaces and applications to finance*, Electronic Journal of Probability, 13 (2008), pp. 852–879.

[145] E. PLATEN, *An introduction to numerical methods for stochastic differential equations*, Acta Numerica, 8 (1999), pp. 197–172.

[146] P. PROTTER, *Stochastic Integration and Differential Equations*, Springer, Berlin, 2nd ed., 2005.

[147] J. RODRIGUEZ AND F. SANTOSA, *Estimation of asset distributions from option prices: analysis and regularization*, SIAM Journal on Financial Mathematics, 3 (2012), pp. 374–401.

[148] B. ROUTLEDGE AND S. ZIN, *Model uncertainty and liquidity*, Review of Economic Dynamics, 12 (2009), pp. 543–566.

[149] D. SAMPERI, *Calibrating a diffusion pricing model with uncertain volatility: regularization and stability*, Mathematical Finance, 12 (2002), pp. 71–87.

[150] P. SAMUELSON, *Rational theory of warrant pricing*, Industrial Management Review, 6 (1965), pp. 13–39.

[151] W. SCHOUTENS, E. SIMONS, AND J. TISTAERT, *A perfect calibration! Now what?*, Wilmott Magazine, March (2004), pp. 66–78.

[152] M. SCHWEIZER, *Risk–minimality and orthogonality of martingales*, Stochastics and Stochastics Reports, 3 (1990), pp. 123–131.

[153] ——, *Option hedging for semimartingales*, Stochastic Processes and their Applications, 37 (1991), pp. 339 – 363.

[154] ——, *Approximating random variables by stochastic integrals*, Annals of Probability, 22 (1993), pp. 1536–1575.

[155] ——, *Variance-optimal hedging in discrete time*, Mathematics of Operations Research, 20 (1995), pp. 1–32.

[156] ——, *A guided tour through quadratic hedging approaches*, in Option Pricing, Interest Rates and Risk Management, E. Jouini, J. Cvitanic, and M. Musiela, eds., Cambridge University Press, 2001, pp. 538–574.

[157] ——, *Local risk-minimization for multidimensional assets and payment streams*, Banach Center Publications, 83 (2008), pp. 213–229.

[158] C. SHANNON, *A mathematical theory of communication*, The Bell System Technical Journal, 27 (1948), pp. 379–423.

[159] S. SHREVE, *Stochastic Calculus for Finance II*, Springer, New York, 2004.

[160] S. SIMON, M. GOOVAERTS, AND J. DHAENE, *An easy computable upper bound for the price of an arithmetic Asian option*, Insurance: Mathematics and Economics, 26 (2000), pp. 175–183.

[161] X. SUN AND S. GAN, *An efficient semi-analytical simulation for the Heston model*, Computational Economics, 43 (2014), pp. 433–445.

[162] X. SUN, S. GAN, AND M. VANMAELE, *Analytical approximation for distorted expectations*, Statistics and Probability Letters, 107 (2015), pp. 246–252.

[163] X. SUN, D. HAESEN, AND M. VANMAELE, *Comment: "On approximating deep in-the-money Asian options under exponential Lévy processes"*, Journal of Futures Markets, 35 (2015), pp. 1220–1221.

[164] X. SUN AND M. VANMAELE, *Model-based and model-free upper bounds for discrete arithmetic Asian options*, working paper, (2015), pp. 1–27.

[165] X. TAN AND N. TOUZI, *Optimal transportation under controlled stochastic dynamics*, The Annals of Probability, 41 (2013), pp. 3201–3240.

[166] S. TANG AND X. LI, *Necessary conditions for optimal control of stochastic systems with random jumps*, SIAM Joural on Control and Optimization, 32 (1994), pp. 1447–1475.

[167] L. TCHUINDJO, *On approximating deep in-the-money Asian options under exponential Lévy processes*, Journal of Futures Markets, 32 (2012), pp. 75–91.

[168] H. TSUKAHARA, *Estimation of distortion risk measures*, Journal of Financial Econometrics, 12 (2013), pp. 213–235.

[169] A. TVERSKY AND D. KAHNEMAN, *Advances in prospect theory – cumulative representation of uncertainty*, Journal of Risk and Uncertainty, 5 (1992), pp. 297–323.

[170] N. VANDAELE, *Quadratic Hedging in Finance and Insurance*, PhD thesis, Ghent University, Belgium, 2010.

[171] M. VANMAELE, G. DEELSTRA, J. LIINEV, J. DHAENE, AND M. GOOVAERTS, *Bounds for the price of discretely sampled arithmetic Asian options*, Journal of Computational and Applied Mathematics, 185 (2006), pp. 51–90.

[172] J. VEČEŘ, *A new PDE approach for pricing arithmetic average Asian options*, Journal of Computational Finance, 4 (2001), pp. 105–113.

[173] J. VORBRINK, *Financial markets with volatility uncertainty*, Journal of Mathematical Economics, 53 (2014), pp. 64–78.

[174] R. WANG AND J. ZIEGEL, *Elicitable distortion risk measures: a concise proof*, Statistics & Probability Letters, 100 (2015), pp. 172–175.

[175] S. WANG, *A universal framework for pricing financial and insurance risks*, ASTIN Bulletin, 32 (2002), pp. 213–234.

[176] X. WANG AND Y. WANG, *Hedging strategies for discretely monitored Asian options under Lévy processes*, Journal of Industrial and Management Optimization, 10 (2014), pp. 1209–1224.

[177] G. WASILKOWSKI AND H. WOZNIAKOWSKI, *Explicit cost bounds of algorithms for multivariate tensor product problems*, Journal of Complexity, 11 (1995), pp. 1–56.

[178] D. XIU AND J. HESTHAVEN, *High-order collocation methods for differential equations with random inputs*, SIAM Journal on Scientific Computing, 27 (2005), pp. 1118–1139.

[179] G. XU AND H. ZHENG, *Basket options valuation for a local volatility jump-diffusion model with asymptotic expansion method*, Insurance : Mathematics and Economics, 47 (2010), pp. 415–422.

[180] S. YANG AND J. LEE, *Do affine jump-diffusion models require global calibration? Empirical studies from option markets*, Quantitative Finance, 14 (2014), pp. 111–123.

[181] C. YIN, *New class of distortion risk measures and their tail asymptotics with emphasis on VaR*, arXiv:1503.08586, (2015), pp. 1–35.

[182] P. ZENG AND Y. KWOK, *Pricing bounds and approximations for discrete arithmetic Asian options under time-changed Lévy processes*, Available at SSRN 2311667, (2013).

[183] ——, *Pricing Barrier and Bermudan style options under time-changed Lévy processes: fast Hilbert transform approach*, SIAM Journal on Scientific Computing, 36 (2014), pp. B450–B485.

[184] J. ZHANG, *A numerical scheme for BSDEs*, Annals of Applied Probability, 14 (2004), pp. 459–488.

[185] J. ZHU, *A simple and accurate simulation approach to the Heston model*, The Journal of Derivatives, 18 (2011), pp. 26–36.

List of Publications

Peer-Reviewed Journal Articles

1. Analytical Approximation for Distorted Expectations, *Statistics and Probability Letters*, 107 (2015), pp. 246–252. (with S. Gan and M. Vanmaele)

2. Comment: "On Approximating Deep in-the-money Asian Options under Exponential Lévy Processes", *Journal of Futures Markets*, 35 (2015), pp. 1220–1221. (with D. Haesen, and M. Vanmaele)

3. An Efficient Semi-Analytical Simulation for the Heston Model, *Computational Economics*, 43 (2014), pp. 433–445. (with S. Gan)

Submitted Papers

1. Uncertainty Quantification of Derivative Instruments, (2015). (with M. Vanmale)

2. Model Risk and Discretisation of Locally Risk-minimising Strategies, (2015). (with T. Schulz, A. Khedher, and M. Vanmaele)

3. Model-based and Model-free Upper Bounds for Discrete Arithmetic Asian Options, (2015). (with M. Vanmaele)

Working Papers

1. Weighted Monte Carlo Method: a Way to Reduce Calibration Risk? (2014). (with M. Vanmaele)