UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede amministrativa: Università degli Studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

CORSO DI DOTTORATO DI RICERCA IN SCIENZE MATEMATICHE

CURRICOLO MATEMATICA COMPUTAZIONALE

CICLO XXX

Essays on Quantization in Financial Mathematics

Coordinatore: Ch.mo Prof. Martino Bardi

Supervisore: Ch.mo Prof. Martino Grasselli

Co-Supervisore: Dott.ssa Giorgia Callegaro

Dottorando: Lucio Fiorin

Abstract

This thesis is devoted to the study of some applications of quantization to Financial Mathematics, especially to option pricing and calibration of financial data. Quantization is a technique that comes originally from numerical probability, and consists in approximating random variables and stochastic processes taking infinitely many values, with a discrete version of them, in order to simplify the quadrature algorithms for the computation of expected values.
The purpose of this thesis is to show the great flexibility that quantization can have in the area of numerical probability and option pricing. In the literature, often there are *ad hoc* methods for a particular type of model or derivative, but no general framework seems to exist. Finite difference methods are heavily affected by the curse of dimensionality, while Monte Carlo methods need intense computational effort in order to have good precision, and are not designed for calibration purposes. Quantization can give an alternative methodology for a broad class of models and derivatives.
The aim of the thesis is twofold: first, the extension of the literature about quantization to a broad class of models, namely local and stochastic volatility models, affine, pure jumps and polynomial processes, is an interesting theoretical exercise in itself. In fact, every time we deal with a different model we have to take in consideration the properties of the process and therefore the quantization algorithm must be adapted. Second, it is important to consider the computational results of the new types of quantization introduced. Indeed, the algorithms that we have developed turn out to be fast and numerically stable, and these aspects are very relevant, as we can overcome some of the issues present in literature for other types of approach.
The first line of research deals with a technique called *Recursive Marginal Quantization*. Introduced in Pagès and Sagna (2015), this methodology exploits the conditional distribution of the Euler scheme of a one dimensional stochastic differential equation in order to construct a step-by-step approximation of the process. In this thesis we deal with the generalization of this technique to systems of stochastic differential equations, in particular to the case of stochastic volatility models. The *Recursive Marginal Quantization* of multidimensional stochastic process allows us to price European and path dependent options, in particular American options, and to perform calibration on financial data, giving then an alternative, and sometimes overcoming, to the usual Monte Carlo techniques.
The second line of research takes a different perspective on quantization. Instead of using discretization schemes in order to compute the distribution of a stochastic process, we exploit the properties of the characteristic function and of the moment generating function for a broad class of processes. We consider the price process at maturity as a random variable, and we focus on the quantization of the stochastic variable, instead of focusing on the quantization of the whole stochastic process. This gives a faster and more precise technology for the pricing of options, and allows the quantization of a huge set of models for which the *Recursive Marginal Quantization* cannot be applied or is not numerically competitive.

Riassunto

Questa tesi si occupa dello studio delle applicazioni della quantizzazione alla Finanza Matematica, in particolare al prezzaggio di opzioni e alla calibrazione su dati finanziari. La quantizzazione è una tecnica che ha le sue origini dalla probabilità numerica, e consiste nell'approssimare variabili aleatorie e processi stocastici continui nello spazio delle realizzazioni con una versione discreta, allo scopo di semplificare gli algoritmi di quadratura per il calcolo di valori attesi.
L'obiettivo di questa tesi è di mostrare la grande flessibilità che può avere la quantizzazione nell'ambiente della probabilità numerica e del prezzaggio di opzioni. Nella letteratura spesso esistono metodi *ad hoc* per ogni tipo di modello e di derivato, ma non sembra esserci una metodologia unica. I metodi alle differenze finite soffrono fortemente della *curse of dimensionality*, mentre i metodi Monte Carlo necessitano di un grande sforzo computazionale, e non sono pensati per esercizi di calibrazione. La quantizzazione può a volte risolvere problemi specifici di queste tecnologie, e presenta una metodologia alternativa per una grande classe di modelli e di derivati. Lo scopo della tesi è duplice: in primo luogo, l'estensione della letteratura sulla quantizzazione ad un'ampia gamma di processi, cioè processi a volatilità locale e stocastica, affini, di puro salto e polinomiali, è di per se un interessante esercizio teorico. Infatti, per ogni tipo di processo dobbiamo considerare le sue proprietà specifiche, adattando quindi l'algoritmo di quantizzazione. Inoltre, è importante considerare i risultati computazioni dei nuovi tipi di quantizzazione introdotti, in quanto è fondamentale sviluppare algoritmi che siano veloci e stabili numericamente, allo scopo di superare le problematiche presenti nella letteratura per altri tipi di approcci.
Il primo filone di ricerca si occupa di una tecnica chiamata Quantizzazione Marignale Ricorsiva. Introdotta in Pagès and Sagna (2015), questa metodologia sfrutta la distribuzione condizionale dello schema di Eulero di un'equazione differenziale stocastica unidimensionale per costruire un'approssimazione passo passo del processo. In questa tesi generalizziamo questa tecnica ai sistemi di equazioni differenziali stocastiche, in particolare al caso dei modelli a volatilità stocastica. La Quantizzazione Marginale Ricorsiva di processi stocastici multidimensionali permette il prezzaggio di opzioni Europee e di opzioni path dependent, in particolare le opzioni Americane, e di effettuare calibrazione su dati finanziari, dando quindi un'alternativa, e spesso superandole, alle tipiche tecniche di Monte Carlo.
La seconda linea di ricerca tratta la quantizzazione da una prospettiva differente. Invece di usare schemi di discretizzazione per il calcolo della distribuzione di un processo stocastico, viene sfruttata le proprietà della funzione caratteristica e della funzione generatrice dei momenti di una vasta classe di processi. Consideriamo infatti il processo del prezzo a maturità come una variabile aleatoria, e ci focalizziamo sulla quantizzazione della variabile casuale, invece di considerare tutto il processo stocastico. Questo approccio porta a una tecnologia più veloce e precisa per il prezzaggio di opzioni, e permette la quantizzazione di un vasto insieme di modelli, che non potevano essere affrontati dalla Quantizzazione Marginale Ricorsiva.

Acknowledgements

First of all, my special thanks go to my advisors, Martino Grasselli and Giorgia Callegaro, who started working with me four years ago. Their mentorship has been incredibly important for me, and it is impossible to describe, or even measure, the amount of time, effort and resources that they have devoted to my M.Sc. and Ph.D. thesis. I am glad to have experienced with them the joy of mathematical research, and to have learned from them the value of hard work, the power of discussion, and the beauty of a well written mathematical result.
I would also like to thank Gilles Pagès, who is the Director of the Master *Probabilités et Finance* at University Paris 6 / École Polytechnique, for the opportunity of attending the "El Karoui" master, and for the experience of living the most intense yet stimulating six months of my academic life. The incredible amount of financial and mathematical knowledge acquired has been and will be fundamental throughout all my career.
Furthermore, I express my thanks to my co-authors Andrea Pallavicini, Abass Sagna and Wim Schoutens. It is never easy to interact with different approaches to the subject of option pricing, and with scientists from a different background, but I really appreciated the chance to work with smart, talented and experienced researchers as they are. I also thank Daniele Marazzina, for inviting me to present my research in a Ph.D. course on quantization, and Wolfgang Runggaldier, for having introduced me to the wonderful area of Financial Mathematics.
In addition, I would like to thank the Cross Asset - Quantitative Strategy team in Société Générale in London, where I spent six months doing an internship. Apart from the technical tools that I have learned, it has been an exceptional lesson on the importance of hard work and motivation to succeed in the professional environment.
Moreover, my thanks go to all the friends and colleagues in the Department of Mathematics in Padova and in the "El Karoui" master. I would like to thank especially the offices 339 and 712, Daniele Tovazzi, Veronica Dal Sasso, Elena Morotti, Francesco Ferraresso, Laura Cossu, Luisa Andreis, Anna Tovo in Padova, and Federico Borghese, Yaroslav Kopotilov, Michele Antonello, Kamila Domanska, Pawel Ambor, Nabil Manchih in Paris. The time spent together has been one of the best parts of the Ph.D. life. I thank also all the Ph.D. students that I have met during the conferences, winter and summer schools that I have attended.
Finally, my biggest thanks go to my family, especially my parents Silvano e Stefania and my sister Laura, for the support on every choice that I made during this period, for the encouragement in the hard periods and for the enormous amount of patience. Having the opportunity to study the things that I like most is a privilege, and this thesis is a thank you to all the people I have met during these three years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Basic tools in quantization

Quantization is a widely used tool in information theory, cluster analysis, pattern and speech recognition, numerical integration, data mining and, as in our case, numerical probability. The birth of optimal quantization dates back to the 50's, when the necessity to optimize signal transmission, by appropriate discretization procedures, arose, see e.g. Bennett (1948); Panter and Dite (1951); Max (1960).
Quantization consists in approximating a signal admitting a continuum of possible values, by a signal that takes values in a discrete set. *Vector quantization* deals with signals that are finite dimensional, such as random variables, while *functional quantization* extends the concepts to the infinite-dimensional setting, as it is the case of stochastic processes. Quantization of random vectors can be considered as a discretization of the probability space, providing in some sense the best approximation to the original distribution. It is therefore crucial for a given distribution to optimize the geometric location of these points and to evaluate the resulting error. Some numerical procedures have been developed in order to get optimal quadratic quantization of the Gaussian (and even non Gaussian) distribution in high dimension, mostly based on stochastic optimization algorithms, see e.g. Gersho and Gray (1992); Pagès and Printems (2003); Pagès and Yu (2013). Over the years many other application fields have been discovered, such as, in the 90's, numerical integration. This opened the door, especially in France and in Germany, to new research perspectives in Numerical Probability and applications to Mathematical Finance.
For a comprehensive introduction to optimal vector quantization and its applications, we refer to the recent review of Pagès (2015) and references therein.

A quantization grid, or simply a *quantizer* of level $N, N \geq 1$, is a subset of $\mathbb{R}^d, d \geq 1$, $\Gamma^N = \{x_1, \dots, x_N\}$, of size at most N having pairwise distinct components. A quantization function, or *quantizer*, is a Γ-valued[1] Borel function $q : \mathbb{R}^d \to \Gamma$, so that N-quantizing X means approximating X by a random vector $q(X)$ taking N values in $\Gamma \subset \mathbb{R}^d$. This procedure results in the error (with respect to the norm in use) $|X - q(X)|$.
The optimal choice of q and Γ represents the core of quantization. For what concerns q, one easily checks that, given $\xi \in \mathbb{R}^d$, $|\xi - q(\xi)|$ is minimized, i.e., it corresponds to the distance $\text{dist}(\xi, \Gamma) := \inf_{i \in \{1,\dots,N\}} |\xi - x_i|$ if and only if q is the nearest neighbour projection on Γ, i.e., if

$$q(\xi) = \text{Proj}_\Gamma(\xi) := \sum_{i=1}^{N} x_i \mathbb{1}_{C_i(\Gamma)}(\xi), \tag{1.1}$$

[1] For ease of notation, hereafter we drop N in the notation Γ^N as we consider a fixed number of points in the grid.

where $(C_i(\Gamma))_{1\le i\le N}$ is a Borel partition of $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, also known as Voronoi partition, satisfying

$$C_i(\Gamma) \subset \left\{\xi \in \mathbb{R}^d : |\xi - x_i| = \min_{i\le j\le N} |\xi - x_j|\right\}, \quad i = 1\ldots, N. \tag{1.2}$$

The dependence of the Voronoi partition on the norm is not explicit, even if the shape of the Voronoi cells in the partition depends on the norm in use (see e.g. (Graf and Luschgy, 2000, Ch. 1)). In particular, when one deals with Euclidean norms, the Voronoi cells are convex sets and we have

$$\left\{\xi \in \mathbb{R}^d : |\xi - x_i| < \min_{i\le j\le N, j\ne i} |\xi - x_j|\right\} = \overset{\circ}{C}_i(\Gamma) \subset \overline{C}_i(\Gamma) = \left\{\xi \in \mathbb{R}^d : |\xi - x_i| = \min_{i\le j\le N} |\xi - x_j|\right\}.$$

In what follows we will use the notation $\widehat{X}^\Gamma$ or $\widehat{X}$ to denote the Voronoi Γ-quantization of X:

$$\widehat{X}^\Gamma = \mathrm{Proj}_\Gamma(X).$$

As a consequence, in view of our application to quantitative finance, integrals of the form $\mathbb{E}[h(X)]$ (for a given Borel function $h : \mathbb{R}^d \to \mathbb{R}$) can be approximated by the finite sum below

$$\mathbb{E}[h(X)] \simeq \mathbb{E}\left[h(\widehat{X})\right] = \sum_{i=1}^{N} h(x_i)\, \mathbb{P}(\widehat{X} = x_i) \tag{1.3}$$

Clearly it still remains to clarify how to get the optimal or at least a "good" grid Γ and the associated weights $\mathbb{P}(\widehat{X} = x_i), i = 1, \ldots, N$ and to estimate the error. An example of optimal 50-dimensional quantization grid for the bivariate Gaussian distribution is given in Figure 1.1.

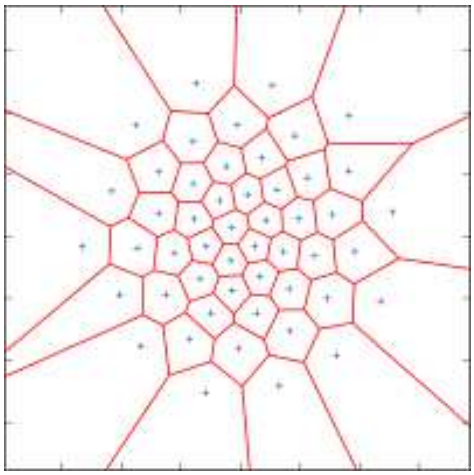


FIGURE 1.1: Example of optimal grid for a bivariate (standard) Gaussian distribution with $N = 50$.

The first issue is about the optimal choice for the grid Γ. We have that, for any given Borel function $\pi : \mathbb{R}^d \to \Gamma = \{x_1, \cdots, x_N\}$,

$$|X - \pi(X)| \ge \min_{i=1,\cdots,N} d(X, x_i) = d(X, \Gamma) = |X - \widehat{X}^\Gamma| \quad \mathbb{P}\text{ a.s}$$

so we can introduce the L^p-mean quantization error, for $p \in [1, +\infty)$, induced by a grid Γ:

$$\begin{aligned} e_{p,N}(X, \Gamma) &:= \left|\left|X - \widehat{X}^{\Gamma}\right|\right|_{L^p(\mathbb{P})} = \left|\left| \min_{1 \le i \le N} |X - x_i| \right|\right|_{L^p(\mathbb{P})} \\ &= \left(\int_{\mathbb{R}^d} \min_{1 \le i \le N} |\xi - x_i|^p \mu(d\xi) \right)^{1/p}. \end{aligned} \tag{1.4}$$

In order to optimally choose Γ, we look for a grid with size at most N which optimally approximates the distribution μ of X in an L^p-sense, when $X \in L^p(\mathbb{P})$. This means that Γ has to minimize the *distortion function* defined as follows (see (Graf and Luschgy, 2000, Equation (3.4))):

Definition 1.1.1. Let $p \in [1, +\infty)$ and let X be a random vector taking values in $\mathbb{R}^d$ belonging to $L^p(\mathbb{P})$. The L^p-distortion function D_p is a positive valued function defined on $(\mathbb{R}^d)^N$ by

$$D_p : (x_1, x_2, \ldots, x_N) \longmapsto \mathbb{E}\left[\min_{1 \le i \le N} |X - x_i|^p \right] = \left(e_{p,N}(X, \Gamma)\right)^p \tag{1.5}$$

when $\Gamma = (x_1, x_2, \ldots, x_N)$.

Optimal quantization boils down to finding a grid Γ which minimizes the L^p-mean quantization error $e_{p,N}(X, \Gamma)$, namely to computing

$$e_{p,N}(X) := \inf_{\Gamma \subset \mathbb{R}^d, \mathrm{card}(\Gamma) \le N} e_{p,N}(X, \Gamma) = \inf_{(x_1, \ldots, x_N) \in (\mathbb{R}^d)^N} D_p(x_1, x_2, \ldots, x_N)^{1/p}, \tag{1.6}$$

where the last equality follows from the fact that a grid with cardinality strictly smaller than N can always be written as a set of cardinality N with duplicate elements.

Of course, the natural question now is about the existence of an optimal grid, and eventually a procedure to find it. For the first issue, one can show (see e.g. (Pagès, 2015, Prop. 1.1)) that for $p \in [1, +\infty)$, if the random vector X belongs to $L^p(\mathbb{P})$, then the distortion function D_p attains (at least) one minimum $\Gamma^\star$. In the case when $\mathrm{card}(\mathrm{supp}(\mu)) \ge N$, then $\Gamma^\star$ has pairwise distinct components. Moreover $\lim_{N \to +\infty} e_{p,N}(X) = 0$. The grid $\Gamma^\star$ and $\mathrm{Proj}_{\Gamma^\star}$ are called *L^p-optimal quantizers*, respectively.

Remark 1.1.1. As soon as $d \ge 2$ uniqueness of the optimal quantization grid typically fails, due to invariance of μ with respect to various transformations. Nevertheless, in the one dimensional case, as soon as μ is absolutely continuous with respect to a log-concave density, then there exists exactly one optimal quantization grid at level N, see (Graf and Luschgy, 2000, Ch. 5) or (Pagès, 2015, Sect. 2).

As already pointed out, the error $e_{p,N}(X)$ is infinitesimal as the grid size N increases, and the rate of convergence of the quantization error is studied in the so-called Zador Theorem recalled below. There is also a non-asymptotic upper bound for optimal quantizers, called Pierce Lemma .

Theorem 1.1.1. *(a) Sharp asymptotic rate (***Zador** *Theorem, see Graf and Luschgy (2000); Zador (1982)). Let X be an random vector in $\mathbb{R}^d$ with finite $p + \eta$ moment, for some $\eta > 0$ and let $\mathbb{P}_X = g \cdot \lambda_d + P_s$ be the Lebesgue decomposition of $\mathbb{P}_X$ with respect to the Lebesgue measure λ_d , and we denote with P_s the singular part. Then*

$$\lim_{N \to +\infty} N^{\frac{1}{d}} e_{p,N}(X) = \tilde{Q}_p(\mathbb{P}_X), \tag{1.7}$$

where

$$\tilde{Q}_p(\mathbb{P}_X) = \tilde{J}_{p,d}\left(\int_{\mathbb{R}^d} g^{\frac{d}{d+p}} d\lambda_d\right)^{\frac{1}{p}+\frac{1}{d}} = \tilde{J}_{p,d}\,\|g\|_{\frac{d}{d+p}}^{1/p} \in [0,+\infty),$$

$$\tilde{J}_{p,d} = \inf_{N\geq 1} N^{\frac{1}{d}} e_{p,N}(U([0,1]^d)) \in (0,+\infty),$$

with $U([0,1]^d)$ stands for the uniform distribution over $[0,1]^d$.
*(b) Non-asymptotic bound (***Pierce** *Lemma, see Graf and Luschgy (2000); Luschgy and Pagès (2002)). Let $\eta > 0$. For every random vector X in $\mathbb{R}^d$ it is possible to find a constant $K_{p,d,\eta}$ such that*

$$\inf_{|\Gamma|\leq N} \|X - \widehat{X}^{\Gamma}\|_p \leq K_{p,d,\eta}\,\sigma_{p,\eta}(X) N^{-\frac{1}{d}}, \tag{1.8}$$

where

$$\sigma_{p,\eta}(X) = \inf_{\zeta\in\mathbb{R}^d} \|X-\zeta\|_{p+\eta} \leq +\infty.$$

Finally, given the existence, the question is how to obtain an optimal quantizer. The crucial point for the whole procedure to be efficient concerns the differentiability of the distortion function. We recall below a powerful result from (Graf and Luschgy, 2000, Lemma 4.10) (see also (Pagès, 2015, Prop. 1.1)) on the differentiability of the distortion function that holds in the case when the norm is smooth. We state it in the case of an Euclidean norm.

Proposition 1.1.2. Consider an Euclidean norm $|.|$ and $p \in (1,+\infty)$, then the L^p-distortion function is differentiable at any N-tuple having pairwise distinct components $\Gamma = \{x_1,\dots,x_N\}$ and the gradient is given by:

$$\begin{aligned}\nabla D_p(x_1,\dots,x_N) &= p\left(\int_{C_i(\Gamma)} \frac{x_i-\xi}{|x_i-\xi|}|x_i-\xi|^{p-1}\mu(d\xi)\right)_{1\leq i\leq N}\\ &= p\left(\mathbb{E}\left[\mathbb{1}_{X\in C_i(\Gamma)}\frac{x_i-X}{|x_i-X|}|x_i-X|^{p-1}\right]\right)_{1\leq i\leq N}.\end{aligned} \tag{1.9}$$

We conclude this subsection with the definition of $p-$stationary quantizers.

Definition 1.1.3. Let X be a random vector in $L^{p-1}(\mathbb{P})$ with distribution μ. A quantization grid $\Gamma = \{x_1,x_2,\dots,x_N\} \subset \mathbb{R}^d$ is *$p-$stationary* if the following two conditions are satisfied:

i) $\mu\left(\bigcup_{1\leq i\leq N}\partial C_i(\Gamma)\right) = 0$ (μ-negligibility of the boundary of the Voronoi diagram)

ii) $\forall i \in \{1,\dots,N\}, \mathbb{E}\left[\mathbb{1}_{X\in C_i(\Gamma)}\frac{x_i-X}{|x_i-X|}|x_i-X|^{p-1}\right] = 0$ (self-consistency of the centroids).

Based on the existence of the gradient of the distortion function, many stochastic algorithms like the gradient descent and fixed point procedures have been introduced, we refer to (Pagès, 2015, Section 3) for a detailed analysis of these methods. When the gradient itself is differentiable, it is even possible to apply the classical Newton-Raphson procedure, as we are going to see.
This is the essential on vector quantization we need in order to proceed. For a thorough treatment of this topic we refer to Graf and Luschgy (2000).

1.2 Main contributions of the thesis

1.2.1 Chapter 2

In this chapter we extend the results in Pagès and Sagna (2015) to a new type of local volatility model, namely the *quadratic normal volatility* (QNV) model. The QNV model assumes that the volatility component of the SDE describing the stock price is modeled using a polynomial of degree two. The breakthrough article of Pagès and Sagna (2015) introduced the *recursive marginal quantization* technique to one dimensional stochastic processes, approximating a process using an Euler scheme, and exploiting the conditional Gaussianity to recover the distribution of the process. We adapt this technique, which turns out to be very fast and accurate even with a small number N of quantization points, to the novel class of the QNV model. Surprisingly, the computational cost when pricing European style options via recursive marginal quantization is lower than using the closed form formulas derived in Carr et al. (2013). The low computational cost and the good accuracy in the pricing exercise lead to the possibility of calibrating the model to financial market data, so that we are able to provide the first known calibration exercise based on quantization methods and with the QNV model. As a further example, we price barrier options in order to prove the flexibility and performance of this new technique, especially when compared with a Monte Carlo benchmark.

1.2.2 Chapter 3

In this chapter we extend the results of Chapter 2 to the case of stochastic volatility models. In fact, in a stochastic volatility model, the Euler scheme of the variance process can be written independently of the scheme for the underlying, even when the correlation component between the volatility and the stock price processes is non zero. The Euler scheme of the price process still depends on the Euler scheme of the variance. This recursive dependence, that exploits a smart use of the Cholesky decomposition of two correlated Brownian motions, allows to approximate the distribution of the stock price process, the variance process, together with their joint distribution. We get the first known example of marginal recursive quantization applied to stochastic volatility models, in particular the Stein and Stein (1991) model, the SABR Model of Hagan et al. (2002) and the α-Hypergeometric Model of Da Fonseca and Martini (2016). The choice of three different models shows that this methodology is flexible enough to deal with different type of models, for which an Euler scheme is available. In order to show the potentiality of this method, in particular the possibility to compute the joint distribution at a given time of the stock price and variance process, we price also a particular type of path dependent option, namely the Volatility Corridor Swap option. This exotic option depends on the realized variance and the price process, and, once the quantization grids have been computed for both the price and the variance process, its price can be computed immediately, so that the resulting method over performs usual Monte Carlo based techniques.

1.2.3 Chapter 4

In this chapter we extend the results of Chapter 3 by introducing a smarter way to write the Euler scheme of the stock price and the variance process, in the presence of stochastic volatility. In fact, instead of using the Cholesky decomposition as in Chapter 3, the main idea here is to consider the two dimensional process as a whole, and

write the corresponding Euler scheme in a vector form. Since, as usual, the conditional distribution of an Euler scheme is Gaussian, it is possible to use the flexibility and potentiality of Gaussian vectors to write the joint distribution of the stock price and variance process. As seen in Chapter 3, it turns out that the Euler scheme associated to the variance process is somehow independent of the Euler scheme of the stock price process, so it can be easily quantized with the techniques developed and implemented in Chapter 1, while the stock price process can be quantized in a faster way. This allows to price in a fast and efficient way also path dependent options, in our case Bermudan / American options. In fact, recursive marginal quantization is a natural framework for pricing using dynamic programming, because it gives a multinomial lattice that approximates the process (in this case processes) in the best possible way, as studied in Bally et al. (2005). Due to the accuracy and low computational cost, this new technique allows us to calibrate on American options under the Heston (1993) model, giving the first known example of calibration using quantization of American options under a stochastic volatility model.

1.2.4 Chapter 5

In this chapter we tackle the application of the recursive marginal quantization when the dimension of the process is greater than one with a novel and different approach with respect to Chapter 3 and Chapter 4. This technique is more general and can actually be applied to diffusive processes of every dimension d, and exploits the fact that the optimal quantization grids for Gaussian vectors are already computed for different dimensions. It is also possible to derive closed form formulas for the computation of the transition probabilities of the Euler scheme at every time step. Given the flexibility and generality of this approach, we give numerical examples of the quantization of the Heston (1993) model as a case of stochastic volatility models, of a two correlated Black Scholes model with non zero correlation and the pricing of a basket option, and finally of a Markovian Backward Stochastic Differential Equation. The theory behind the quantization of a BSDE can be found in Pagès and Sagna (2016). Apart for the case of stochastic volatility, which has already been studied in Chapter 2 and Chapter 3, the other numerical examples provide the first example of the application of quantization to the pricing of basket options under non trivial models and of numerical examples of the quantization of a BSDE.

1.2.5 Chapter 6

In this chapter we introduce a new methodology to optimally discretize a stochastic process via quantization. The idea consists in using the Fourier methodology in order to efficiently compute the probability that the process at a given maturity belongs to a certain Voronoi region. In particular, we obtain a new discretization procedure for a random variable in an L^p-optimal sense, for any $p \geq 1$. With respect to recursive marginal quantization, our new Fourier methodology avoids the propagation of the discretization error through the time line in so far it preserves the law of the process at a given time. Also, we introduce a novel pricing methodology for a broad class of financial models, including stochastic volatility and pure jump processes, for which the characteristic function of the log-asset price can be efficiently computed. The new method reveals to be fast and accurate, to the point that we can calibrate the models on real data. Our approach is very flexible and allows for the possibility to price Vanilla but also American style options. Indeed we are dealing with an optimal multinomial lattice, where it is possible to compute the

transition probabilities of the underlying at each node, so that e.g. a backward dynamic programming algorithm can be used in this case. As an example, we consider the case of an American Put option in a Tempered Stable model, with constitutes the first application of quantization to a pure jump process.

1.2.6 Chapter 7

In this chapter we adapt the technology of Chapter 6 to a new class of models, based on Polynomial processes. They have been introduced by Cuchiero et al. (2009) and Filipovic and Larsson (2016), and the main property of these processes is that the expected value of any polynomial of the process is again a polynomial as a function of the initial value of the process. In this way, moments of all orders can be easily computed in closed form (up to a matrix exponential), even if the characteristic function of the process may be not known. We implement two different techniques, the first one is similar to the one in Chapter 6, where we consider the underlying process at maturity as a random variable, and we exploit the polynomial property to write the quantization algorithm. The second one is similar to the one in Chapter 4, but we present the technical details for the generalization to any dimension. We present, as an example, the application of the two methodologies to the pricing of European and Bermudan options under the Stochastic Volatility Jacobi (SVJ) model of Ackerer et al. (2016), which is a generalization of the Heston model. Nevertheless, the technology that we introduce can be applied to every process with explicit discrete moments, so also to all the affine processes of Chapter 6. This is the first example in literature of quantization of polynomial processes, and of pricing of Bermudan options under the SVJ model.

1.2.7 Chapter 8

In this chapter we overcome the limits of Chapter 6 in terms of computational cost of the quantization. We follow the same idea that the distribution of a process at a given maturity can be obtained via its characteristic function, and we exploit a property of the optimal quantizers for uniform random variables in order to obtain an incredibly fast quantization technology. We connect this technique with the Conic Finance theory as described in Madan and Schoutens (2016). Conic Finance is a theory that describes the possibility of having two prices in the market, i.e. the bid and the ask price of a derivative. Introducing a distortion of the distribution of the underlying price, it is possible to construct a rigorous theory that describes the behavior of the bid-ask spread, which is related to a somehow liquidity measure in the market. The quantization technique is able then to price European derivatives under stochastic volatility taking into account the presence of liquidity in the market. This motivates us to calibrate a stochastic volatility model, in our numerical example the Heston model, to a set of market data containing bid and ask prices for European options. It is then natural to define the concept of market implied liquidity, related to the bid-ask spread over strikes and maturities, thus giving the first example in literature of the implied liquidity surface under stochastic volatility.

Chapter 2

Local volatility models

This is a joint work with Giorgia Callegaro and Martino Grasselli. The content of this chapter has been published as a paper in *Risk Magazine*, see Callegaro et al. (2015) in the bibliography. This work has been awarded in the XXXIX AMASES meeting (Padova, September 10-12, 2015) for the "Best paper presented by a young researcher".

In this chapter we propose the first calibration exercise based on quantization methods. Pricing and calibration are typically difficult tasks to accomplish: pricing should be fast and accurate, otherwise calibration cannot be efficiently performed. We apply in a local volatility context the recursive marginal quantization methodology to the pricing of vanilla and barrier options. A successful calibration of the Quadratic Normal Volatility model is performed in order to show the potentiality of the method in a concrete example, while a numerical exercise on pricing of barrier options shows that quantization over performs Monte Carlo methods.

2.1 Introduction

Following the lines of Pagès and Sagna (2015), in this chapter we apply recursive marginal quantization to a special local volatility model, namely the Quadratic Normal Volatility model, that has been investigated by Blacher (2001), Ingersoll (1997), Lipton (2002), Zühlsdorff (2002) and lately revisited by Andersen (2011) and Carr et al. (2013). We find stationary quantizers via a Newton-Raphson method, in order to efficiently price vanilla and exotic derivatives. Indeed, the Newton-Raphson procedure, being deterministic, is very fast and it allows us to provide the first example of calibration based on quantization. The recursive marginal quantization is competitive even when closed form formulas for vanillas are available (as in the case of call and put prices for the Quadratic Normal Volatility model). Finally, we show the flexibility and the efficiency of the recursive marginal quantization in the pricing of non-vanilla contracts, when compared to the classic Monte Carlo simulation. Our numerical algorithms have performed quite well (wrt to Monte Carlo), so that in this chapter no speed-up procedure has been tested. As a consequence, this chapter does not provide the fastest possible numerical method, but a procedure that is competitive enough if compared to Monte Carlo.

The chapter is organized as follows: Section 2.2 extends the vector quantization method to the class of Markov diffusion processes, leading to the recursive marginal quantization. Section 2.3 introduces the Quadratic Normal Volatility model together with the well known results about closed form formulas for vanilla option prices.

Moreover, we apply the recursive marginal quantization approach to the pricing of barrier options. Section 2.4 illustrates our numerical results, with particular emphasis to the calibration exercise on real data. Section 2.5 concludes. Some technical details are given in the Appendix 2.6.

2.2 Recursive marginal quantization

In this section we consider the quantization of a continuous-time diffusive Markov process X, whose evolution is specified by the following SDE:

$$dX_t = b(t, X_t)dt + a(t, X_t)dW_t, \quad X_0 = x_0 > 0, \tag{2.1}$$

where W is a standard Brownian motion and the functions a and b satisfy the usual conditions ensuring the existence of a strong solution to the SDE. Following the approach presented in Pagès and Sagna (2015), we work on the Euler scheme of X and we discretize the process by exploiting its Markov property via vector quantization.

Having fixed a time horizon $T > 0$ and a time discretization grid $\{0 = t_0, t_1, \dots, t_M = T\}$, with constant step size $\Delta_k = t_k - t_{k-1}$, $k \geq 1$, such that $t_k = \frac{kT}{M}$, the Euler scheme for the process X is given by

$$\begin{aligned}
\widetilde{X}_{t_k} &= \widetilde{X}_{t_{k-1}} + b(t_{k-1}, \widetilde{X}_{t_{k-1}})\Delta_k + a(t_{k-1}, \widetilde{X}_{t_{k-1}})\Delta W_k, \\
\widetilde{X}_{t_0} &= \widetilde{X}_0 = x_0,
\end{aligned}$$

where $\Delta W_k := W_{t_k} - W_{t_{k-1}}$ is a centered Normal random variable with variance Δ_k, so that we have the following equality in distribution

$$\left(\widetilde{X}_{t_k} | \widetilde{X}_{t_{k-1}} = x\right) \overset{\text{Law}}{=} \mathcal{N}\left(m_{k-1}(x), \sigma^2_{k-1}(x)\right), \tag{2.2}$$

where

$$\begin{aligned}
m_{k-1}(x) &= x + b(t_{k-1}, x)\Delta_k, \\
\sigma^2_{k-1}(x) &= [a(t_{k-1}, x)]^2\Delta_k.
\end{aligned}$$

Our intention now is to use the vector quantization applied to every (one dimensional) random variable $\widetilde{X}_{t_k}$, $k \geq 1$, since we know its marginal distribution conditional to $\widetilde{X}_{t_{k-1}}$. This explains the term *marginal* of this quantization method. It can be seen in Pagès et al. (2003) that the error made by quantizing the Euler scheme can be easily controlled, under some mild regularity assumptions on the process. The distortion function relative to $\widetilde{X}_{t_{k+1}}$, denoted D_{k+1} reads

$$D_{k+1}(\mathbf{x}^{k+1}) = \sum_{i=1}^{N} \int_{C_i(\mathbf{x}^{k+1})} (y_{k+1} - x_i^{k+1})^2 \mathbb{P}(\widetilde{X}_{t_{k+1}} \in dy_{k+1}), \tag{2.3}$$

where $\mathbf{x}^{k+1} = \{x_1^{k+1}, x_2^{k+1}, \dots, x_N^{k+1}\}$ is the quantizer at time t_{k+1} and N is the (fixed) size of the quantizer at every time step. The delicate point here is that, in order to quantize $\widetilde{X}_{t_{k+1}}$ we have to apply the Newton-Raphson method without knowing its distribution. However, by using the conditional distribution in (2.2) we can rewrite the distortion function (2.3) in terms of $\widetilde{X}_{t_k}$, obtaining then a recursive formula to compute the stationary quantizer. In fact, the distribution function of $\widetilde{X}_{t_{k+1}}$ can be

written as follows

$$\mathbb{P}(\widetilde{X}_{t_{k+1}} \in dy_{k+1}) = dy_{k+1} \int_{\mathbb{R}} \phi_{m_k(y_k),\sigma_k(y_k)}(y_{k+1}) \mathbb{P}(\widetilde{X}_{t_k} \in dy_k) = dy_{k+1} \mathbb{E}\left[\phi_{m_k(\widetilde{X}_k),\sigma_k(\widetilde{X}_k)}(y_{k+1})\right],$$

where $\phi_{m,\sigma}$ denotes the density function associated with a Normal distribution $\mathcal{N}(m, \sigma^2)$. With this result it is possible to compute the Hessian matrix of the distortion function. Note that we are interested in the quantization of the Euler scheme $\widetilde{X}$ that we denote by $\widehat{X}_{t_k}$, $k \geq 0$, so that we substitute $\widetilde{X}_{t_k}$ with $\widehat{X}_{t_k}$ in (2.3). Due to the discrete nature of the quantizer, the integral in (2.3) becomes a finite sum, thus leading to extremely fast computations. In the sequel we will apply the recursive marginal quantization to a special local volatility model, namely the Quadratic Normal Volatility model. We refer the interested reader to Pagès and Sagna (2015) for a complete background including the analysis of the errors generated by the recursive quantization method.

2.3 The Quadratic Normal Volatility model

The class of Quadratic Normal Volatility (QNV) models has drawn much attention in the financial industry due to its analytic tractability and flexibility. We will refer to the works of Blacher (2001), Ingersoll (1997), Lipton (2002), Andersen (2011).
A QNV model is associated to an asset X evolving as follows

$$dX_t = (e_1 X_t^2 + e_2 X_t + e_3) dW_t, \; X_0 = x_0 > 0, \tag{2.4}$$

for some $e_1, e_2, e_3 \in \mathbb{R}$, where the Brownian motion W is taken under the risk neutral measure. This corresponds to the SDE (2.1) where $b(t,x) = 0$ (that is we consider the forward-price process) and $a(t,x) = e_1 x^2 + e_2 x + e_3$. Note that (2.4) includes, as special cases, the Brownian motion (for $e_1 = e_2 = 0$), the geometric Brownian motion (for $e_1 = e_3 = 0$) and the inverse of a three-dimensional Bessel process (for $e_2 = e_3 = 0$) which leads to a strict local martingale (we refer to Andersen (2011) and Carr et al. (2013) for other technical properties of the model). Apart from technicalities, the intuition underlying (2.4) is that mimicking a quadratic spot volatility gives some chances to get an implied volatility curve that is able to reproduce the smile and skew effects using a parsimonious number of parameters. This is more evident in the following parameterization (taken by Andersen (2011))

$$dX(t) = \sigma\left(qX(t) + (1-q)x_0 + \frac{1}{2}s\frac{(X(t)-x_0)^2}{x_0}\right)dW(t), \; X_0 = x_0 > 0. \tag{2.5}$$

Here $\sigma > 0$ is a proxy for the ATM volatility level, q is related to the implied volatility slope (that is q is the skew parameter) and s is a measure of the convexity of the quadratic volatility function (the vol-of-vol parameter).

2.3.1 Vanilla Options Pricing

The QNV model allows for closed form solutions for the prices of vanilla options (see also the technical Appendix 2.6.1 taken from Andersen (2011)). The corresponding formulas depend on the roots of the polynomial in (2.4). Note that even if closed form formulas are available for vanillas, their implementation is time consuming and it requires some care, especially in the truncation of the trigonometric series. Moreover, a calibration procedure based on these formulas should allow for the

possibility to switch from the first (real roots) to the second case (complex roots) without constraints. We will see in the calibration exercise that this is a real issue. On the contrary, in the recursive marginal quantization approach one never deals with this problem. Following the steps illustrated in the previous Section 2.2 one easily computes the critical points of the distortion function together with its Hessian. In Appendix 2.6.2 we present the formulas for the gradient, the Hessian matrix and the weights of the quantized random variable $\widehat{X}_T$.

2.3.2 Barrier Options Pricing

We focus now on barrier options. More precisely, on discrete time barrier options with daily monitoring. Indeed, although most models in the literature assume continuous monitoring of the barrier (which can lead to analytic solutions as in the Black and Scholes model), in practice most barrier options are discretely monitored. Unfortunately, this realistic setting in general does not allow for closed form solutions. We refer the reader interested to the pricing of discrete barrier options eg to Kou (2003) (for an introduction to the so-called continuity correction) and to Lipton and McGhee (2002) (for a PDE approach in a universal volatility model, that leads, in some benchmark cases, to analytic solutions). We also refer to Lipton et al. (2014) for a survey on pricing of barrier options in local-stochastic volatility models.
In order to apply recursive marginal quantization to this setting, we follow the approach in Sagna (2012), where the author presents an algorithm based on optimal marginal quantization, to approximate the price of knock out barrier options. We consider up-and-out put options. Pricing formulas in the other cases are just slight modifications of the ones that we are going to present here.
Given the Euler scheme $\widetilde{X}$ for the process X, the price of an up-and-out put option expiring at time T, with strike K and up-and-out barrier L can be approximated by

$$P^{\mathrm{LO}} := e^{-rT}\mathbb{E}\left((K-\widetilde{X}_T)^+\mathbb{1}_{\{\sup_{k=0,\dots,M}\widetilde{X}_{t_k}\leq L\}}\right) = e^{-rT}\mathbb{E}\left((K-\widetilde{X}_T)^+\prod_{k=1}^{M}G_{\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k}}(L)\right), \tag{2.6}$$

where

$$G_{x,y}(u) = \left(1-e^{-2M\frac{(x-u)(y-u)}{T\sigma^2(x)}}\right)\mathbb{1}_{\{u\geq\max(x,y)\}},$$

and where $\sigma(\cdot)$ is the volatility function of X. The last equality in the above equation can be obtained via an application of the so-called "regular Brownian bridge method", that is connected to the knowledge of the distribution of the minimum (or the maximum) of the continuous Euler scheme $\widetilde{X}$ relative to a process X over a time interval $[0,T]$, given its values at the discrete time observation points $0=t_0<t_1<\cdots<t_M=T$ (see eg Glasserman (2003)).
The expectation in (2.6) can be computed recursively, as soon as we have an approximation of the transition probability of $\widetilde{X}_{t_k}$ given $\widetilde{X}_{t_{k-1}}$. The idea now is to approximate this expectation using $\widehat{X}_{t_k}$ intead of $\widetilde{X}_{t_k}, k\geq 1$ and the transition matrix of $\widehat{X}_{t_k}$ given $\widehat{X}_{t_{k-1}}$. For all the detailed formulas we refer to the Appendix 2.6.3.

2.4 Numerical results

In this section we provide the first example of competitive and efficient calibration of a quantization-based method to real data and we then apply our result to the pricing of vanilla and non-vanilla derivatives. Note that, once we know the stationary grid

for each time step, the pricing of a generic option becomes immediate. For example, the price at $t = 0$ of a European Vanilla Put option on X with maturity T and strike K, that we have N-quantized at $t = t_M = T$ with an optimal grid $\mathbf{x}^M = (x_1^M, \dots, x_N^M)$ and associated optimal quantizer $\widehat{X}_T$, is given by

$$\mathbb{E}[(K - X_T)^+] \simeq \sum_{i=1}^{N} (K - x_i^M)^+ \, \mathbb{P}(\widehat{X}_T = x_i^M),$$

that can be immediately computed. Note also that $C_i(\mathbf{x}^M) = \left[\frac{x_{i-1}^M + x_i^M}{2}, \frac{x_i^M + x_{i+1}^M}{2}\right]$ since we work in a one dimensional setting.
The dimension of the quantization grids is taken to be constant over time, which is obviously not the optimal choice. Nevertheless, it represents a good trade-off between price precision and implementation cost. For more details on this aspect, as well as for an analysis on optimal dispatching, we refer to Pagès and Sagna (2015).

2.4.1 Calibration on vanillas

We first test the goodness of the pricing via recursive marginal quantization. Here we use 9 different strikes, equally spaced from 80% to 120% of the initial value of the underlying, and 6 different maturities, from 2 months to 2 years. As an error measure for this test we consider the Res. Norm, that is the sum of the squared differences between the model implied volatilities and the ones generated by the closed form formulas of the previous Section. We use 30-dimensional quantizers and 10 time steps for every maturity. Figure 2.1 shows the corresponding quantization grids. Computations are performed using *Matlab* on a CPU 2.4 GHz and 8 Gb memory computer. The inverse of the Hessian matrix, which is is tridiagonal and symmetric, is calculated using the LU-decomposition.

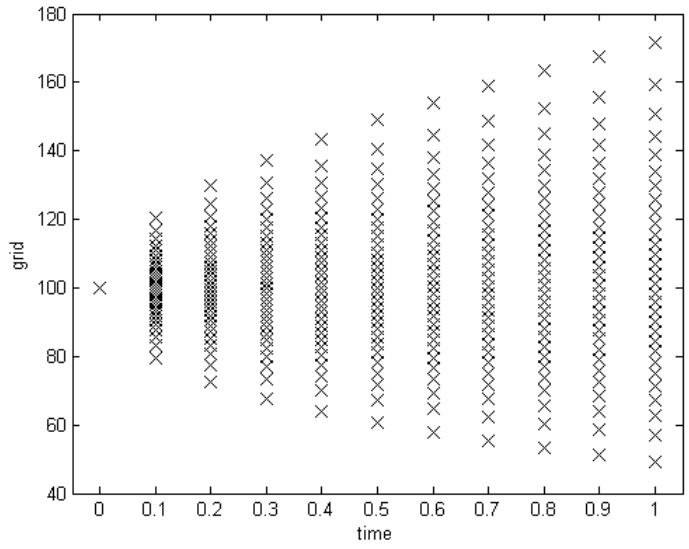


FIGURE 2.1: Quantization grids for the QNV model in the case: $\sigma = 0.2$; $q = 0.5$; $s = 0.1$ (in the (2.5) specification).

The results in Table 2.1 confirm the precision of prices generated by the quantization. Note that in the case of complex roots the quantization algorithm is faster than the

computation based on closed form formulas. This fact is relevant since market data calibration typically requires complex roots as we are going to show.

Real roots		
Analytic CT	Quantization CT	Res. Norm
0.03550 sec	1.07239 sec	$1.34891 * 10^{-4}$

Complex roots		
Analytic CT	Quantization CT	Res. Norm
10.25185 sec	1.14839 sec	$1.79906 * 10^{-4}$

TABLE 2.1: Comparison between pricing via closed form formulas and quantization with $N = 30$. Here CT stands for computational time. In the case of two real roots the parameters are taken from Andersen (2011): $\sigma = 0.2$; $q = 0.5$; $s = 0.1$; $x_0 = 100$; (in the (2.5) specification). We then perturb the s parameter in order to get two complex roots: $\sigma = 0.2$; $q = 0.5$; $s = 5$.

Let us now turn to real market data. Calibration is done via a standard non-linear least-squares optimizer that minimizes the total calibration error in terms of the difference between model and market implied volatilities $\sum_n (\sigma^{\text{imp}}_{n,\text{market}} - \sigma^{\text{imp}}_{n,\text{model}})^2$. Using a major provider, we take prices of European Vanilla Call-Put option on the Dax Index, as of 19 June 2014. Using the closed form formulas, it turns out that the implied volatility smile produced by the market is fitted better when the two roots are complex. As a consequence, quantization will be faster than closed form formulas. What is more, closed form formulas do not perform well for short maturities, to the point that we are not able to present results of the calibration based on closed form formulas in this case, while we note that the flexibility of the quantization approach permits to overcome these difficulties. We therefore show the joint results of calibration via closed form formulas and via quantization only with long maturities (from 1.5 years up to 3 years), while with short maturities (from 2 months up to 5 months) we only display the calibration results for the recursive marginal quantization. The calibrated parameters are displayed in Table 2.2.

	σ	q	s
Exact formulas / long maturities	0.16019	-0.04380	26.69999
Quantization / long maturities	0.17451	0.00005	7.62015
Quantization / short maturities	0.14536	-4.67521	16.74793

TABLE 2.2: Calibrated parameters of the Quadratic Normal Volatility model. Here $x_0 = 9837.63$.

With long maturities (resp. short maturities) the Res. Norm is given in Table 2.3, containing 4 maturities and 7 strikes (resp. 4 maturities and 5 strikes). Overall, the quality of the fit is not excellent (see Figure 2.2 and 2.3), but this is due to the particular model which is very parsimonious (only 3 parameters!). Nevertheless, we emphasize that despite the simplicity and the limits of the model, this represents

the first successful calibration example based on quantization. Moreover, it is important to notice that the procedure here illustrated is very robust since it can be easily applied to any local volatility (diffusive) model for which the Euler scheme is available.

	Closed form formulas		Quantization	
	Comp. time	Res. Norm	Comp. time	Res. Norm
Long maturities	339.15013 sec	$5.62922 * 10^{-4}$	221.15028 sec	$4.26904 * 10^{-4}$
Short maturities	-	-	159.02147 sec	$4.00141 * 10^{-4}$

TABLE 2.3: Computational times and calibration errors obtained via closed form formulas and quantization. For short maturities (from 2 to 5 months) we were not able to obtain meaningful results using the closed form formulas.

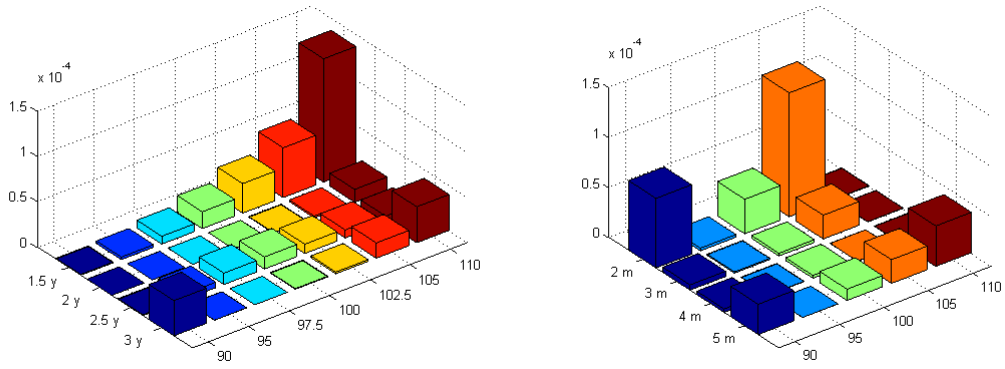


FIGURE 2.2: Implied volatility squared errors for the calibration via quantization. Long maturities on the left, short maturities on the right.

2.4.2 Pricing of barrier options

In order to test the goodness of this pure quantization method, we use the same data as in the previous subsection, focusing on short maturities. We fix the maturity $T = \frac{1}{3}$ and the strike $K = 100\%$ (ATM).
We compare the prices of an up-and-out put option obtained via quantization and with Monte Carlo simulation. The aim is to show that the quantization approach over performs the Monte Carlo method in terms of computational cost.
We first compute a Monte Carlo price with 10^7 simulations. The corresponding confidence interval is very sharp (about 0.3%), and we consider this price as our benchmark.
On the quantization side, we use 48-dimensional quantizers, which reveals to be a good tradeoff between precision and computational cost. Then we look for the number of paths required by the Monte Carlo which shares the same computation time required by quantization (about 2 seconds). It turns out that we need 10^4 simulations, as illustrated in Table 2.4. Notice that the quantization price falls within the confidence interval of the Monte Carlo which is quite large. What is more, Monte Carlo is much less precise than quantization when we fix the same computational

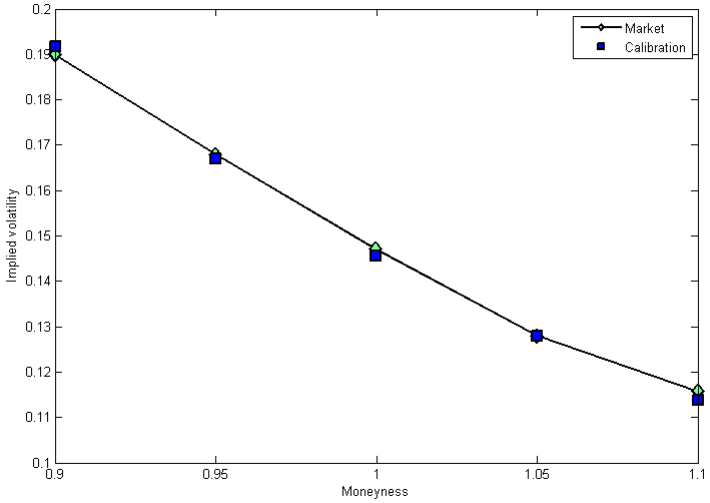


FIGURE 2.3: Example of fit of the implied volatility smile. For short maturities (here the maturity is 4 months) we consider 5 strikes.

time. Here, by precision we mean the distance between the benchmark and the quantization price, while for the Monte Carlo we mean the maximal distance between the benchmark price and the endpoints of the confidence interval.

We also perform another exercise, namely we look for the number of simulations required by Monte Carlo in order to match the precision of the quantization method. Table 2.5 shows that Monte Carlo requires $8 * 10^4$ simulations which take about 10 times the computational cost of the quantization.
In conclusion, the quantization method is a very good alternative to Monte Carlo.

2.5 Conclusion

We have applied recursive marginal quantization to the local volatility model QNV in order to provide an alternative way to compute prices, without the numerical problems due to the real/complex nature of the roots. The procedure gives a fast way to price vanilla as well as barrier options, compared to Monte Carlo simulation. A successful calibration of the QNV model on real data shows the flexibility and the robustness of the quantization method, which can be considered a model independent approach. Extensions of this work could include less parsimonious local volatility models, since the speed of the algorithm does not depend on the number of parameters, and pricing of structured contracts, in the spirit of Bardou et al. (2009) who investigated the energy market.

	Benchmark price	Q price	Q error	MC confidence interval
$L = 103.75$	305.17096	296.70996	2.77%	[295.43121, 321.15975]
$L = 105$	320.88575	313.47255	2.29%	[311.23967, 337.03611]
$L = 106.25$	327.77641	320.99300	2.00%	[317.02684, 342.86616]
$L = 107.5$	330.20446	323.76159	1.89%	[319.36426, 345.19677]
$L = 108.75$	330.87364	324.63297	1.83%	[319.90923, 345.73776]
Comp. time		2.11746 sec		2.71984 sec

TABLE 2.4: Results on the pricing of up-and-out put options with strike $K = 100\%$ via quantization and Monte Carlo simulation. The barrier L is in percentage of the initial price. Q stands for quantization, MC for Monte Carlo (10^4 simulations). We consider a similar computational cost and in the last column we display the confidence interval for the corresponding MC.

	Benchmark price	Q price	Q error	MC confidence interval
$L = 103.75$	305.17096	296.70996	2.77%	[299.38595, 308.44100]
$L = 105$	320.88575	313.47255	2.29%	[314.20251, 323.31032]
$L = 106.25$	327.77641	320.99300	2.00%	[321.12872, 330.24552]
$L = 107.5$	330.20446	323.76159	1.89%	[323.91769, 333.03919]
$L = 108.75$	330.87364	324.63297	1.83%	[324.54593, 333.66737]
Comp. time		2.11746 sec		19.12999 sec

TABLE 2.5: Results on the pricing of up-and-out put options with strike $K = 100\%$ via quantization and Monte Carlo simulation. The barrier L is in percentage of the initial price. Q stands for quantization, MC for the Monte Carlo ($8 * 10^4$ simulations) that shares the same precision as the Q method. In the last row we display the associated computation times.

2.6 Appendix

2.6.1 Closed form solution for the Call price

Formulas are taken from Andersen (2011). In the case of two distinct real roots $u, l \in \mathbb{R}, u > l$, the price at time 0 of a call option is found to be given by

$$p(0) = K_1\Phi_{0,1}\big(-d_-^{(1)}\big) - X_2\Phi_{0,1}\big(d_+^{(2)}\big) - X_1\Phi_{0,1}\big(d_+^{(1)}\big) + K_2\Phi_{0,1}\big(d_-^{(2)}\big),$$

where K_i, X_i, $d_\pm^{(i)}$ are given by

$$K_1 = \frac{(K-u)(x_0-l)}{u-l}, \quad X_1 = \frac{(x_0-u)(K-l)}{u-l},$$
$$K_2 = \frac{(K-l)(x_0-l)}{u-l}, \quad X_2 = \frac{(x_0-u)(K-u)}{u-l},$$

and

$$d_\pm^{(i)} = \frac{ln(\frac{X_i}{K_i}) \pm \frac{T}{2}}{\sqrt{T}}, \quad i = 1, 2,$$

and where $\Phi_{0,1}$ denotes the standard Normal cumulative distribution. In the case of two complex roots, denoted by $a \pm ib$, we have

$$p(0) = \sqrt{bA(x_0)}e^{\frac{1}{2}T}\left(\frac{2}{Z_U - Z_L}\sum_{n=1}^{+\infty} e^{-\alpha_n T}\sin(-a_n)\left(\widetilde{K}I_n^{(c)} - I_n^{(s)}\right)\right)$$
$$+\sqrt{\frac{A(x_0)}{A(L)}}(K-L)\left(\frac{\sin(Z_U - Z_0)}{\sin(Z_U - Z_L)} - \frac{1}{(Z_U - Z_L)^2}\sum_{n=1}^{+\infty} n\pi\sin(-a_n)\frac{e^{-(\alpha_n - \frac{1}{2})T}}{\alpha_n - \frac{1}{2}}\right).$$

Moreover,

$$Z_0 = \arctan\left(\tfrac{x_0 - a}{b}\right), \quad Z_U = \arctan\left(\tfrac{U-a}{b}\right), \quad Z_L = \arctan\left(\tfrac{L-a}{b}\right),$$

$$\widetilde{K} = \tfrac{K-a}{b}, \qquad A(x) = b\left(1 + \left(\tfrac{x-a}{b}\right)^2\right),$$
$$\alpha_n = \tfrac{n^2\pi^2}{2(Z_U - Z_L)^2} \qquad a_n = \tfrac{n\pi(Z_L - Z_0)}{Z_U - Z_L},$$

where U and L are the upper and lower bounds introduced to guarantee that the process is a martingale (a natural configuration is $U = +\infty$, $L = 0$) and

$$I_n^{(c)} = \tfrac{Z_L - Z_0}{2}\left(\tfrac{\cos(c_n^-)}{ZL - Z_0 - a_n} - \tfrac{\cos(c_n^+)}{ZL - Z_0 + a_n} + \tfrac{2\cos(Z_L)a_n}{a_n^2 - (ZL - Z_0)^2}\right),$$
$$I_n^{(s)} = \tfrac{Z_L - Z_0}{2}\left(\tfrac{\sin(c_n^-)}{ZL - Z_0 - a_n} - \tfrac{\sin(c_n^+)}{ZL - Z_0 + a_n} + \tfrac{2\sin(Z_L)a_n}{a_n^2 - (ZL - Z_0)^2}\right),$$
$$c_n^{\pm} = \left(1 \pm \tfrac{a_n}{Z_L - Z_0}\right)\left(\arctan(\widetilde{K} - Z_0)\right) \mp a_n + Z_0.$$

2.6.2 Formulas for the marginal quantization

Given the optimal quantization grid $\mathbf{x}^k = (x_1^k, \ldots, x_N^k)$ for $\widehat{X}_{t_k}$, the derivative of the distortion function relative to the grid $\mathbf{x}^{k+1} = (x_1^{k+1}, \ldots, x_N^{k+1})$ is

$$\frac{\partial D_{k+1}}{\partial x_j^{k+1}}(\mathbf{x}^{k+1}) = 2\sum_{i=1}^{N}\left\{\left(x_j^{k+1} - m_k(x_i^k)\right)\left(\Phi_{0,1}\left(\widehat{x}_{k+1,j^+}(x_i^k)\right) - \Phi_{0,1}\left(\widehat{x}_{k+1,j^-}(x_i^k)\right)\right)\right.$$
$$\left.+\sigma_k(x_i^k)\left(\phi_{0,1}\left(\widehat{x}_{k+1,j^+}(x_i^k)\right) - \phi_{0,1}\left(\widehat{x}_{k+1,j^-}(x_i^k)\right)\right)\right\}\mathbb{P}(\widehat{X}_{t_k} = x_i^k),$$

where

$$\widehat{x}_{k+1,j^+}(x_i^k) = \frac{x_j^{k+1} + x_{j+1}^{k+1} - 2m_k(x_i^k)}{2\sigma_k(x_i^k)},$$
$$\widehat{x}_{k+1,j^-}(x_i^k) = \frac{x_{j-1}^{k+1} + x_j^{k+1} - 2m_k(x_i^k)}{2\sigma_k(x_i^k)}.$$

For the Hessian matrix we get

$$\begin{aligned}\frac{\partial^2 D_{k+1}}{\partial^2 x_j^{k+1}}(\mathbf{x}^{k+1}) =& 2\sum_{i=1}^{N}\Bigg\{\Big(\Phi_{0,1}\big(\widehat{x}_{k+1,j^+}(x_i^k)\big)-\Phi_{0,1}\big(\widehat{x}_{k+1,j^-}(x_i^k)\big)\Big)\\ &-\frac{1}{4\sigma_k(x_i^k)}\phi_{0,1}\big(\widehat{x}_{k+1,j^+}(x_i^k)\big)\big(x_{j+1}^{k+1}-x_j^{k+1}\big)\\ &-\frac{1}{4\sigma_k(x_i^k)}\phi_{0,1}\big(\widehat{x}_{k+1,j^-}(x_i^k)\big)\big(x_j^{k+1}-x_{j-1}^{k+1}\big)\Bigg\}\mathbb{P}(\widehat{X}_{t_k}=x_i^k),\end{aligned}$$

$$\frac{\partial^2 D_{k+1}}{\partial x_j^{k+1}\partial x_{j-1}^{k+1}}(\mathbf{x}^{k+1}) = -\tfrac{1}{2}\sum_{i=1}^{N}\left\{\phi_{0,1}\big(\widehat{x}_{k+1,j^-}(x_i^k)\big)\left(\frac{x_j^{k+1}-x_{j-1}^{k+1}}{\sigma_k(x_i^k)}\right)\right\}\mathbb{P}(\widehat{X}_{t_k}=x_i^k),$$

$$\frac{\partial^2 D_{k+1}}{\partial x_j^{k+1}\partial x_{j+1}^{k+1}}(\mathbf{x}^{k+1}) = -\tfrac{1}{2}\sum_{i=1}^{N}\left\{\phi_{0,1}\big(\widehat{x}_{k+1,j^+}(x_i^k)\big)\left(\frac{x_{j+1}^{k+1}-x_j^{k+1}}{\sigma_k(x_i^k)}\right)\right\}\mathbb{P}(\widehat{X}_{t_k}=x_i^k).$$

Finally, the weights of the quantized process are given by the following approximation

$$\mathbb{P}(\widehat{X}_{t_{k+1}}\in C_i(\mathbf{x}^{k+1}))=\sum_{j=1}^{N}\left[\Phi_{0,1}\left(\widehat{x}_{k+1,i^+}(x_j^k)\right)-\Phi_{0,1}\left(\widehat{x}_{k+1,i^-}(x_j^k)\right)\right]\mathbb{P}(\widehat{X}_{t_k}=x_j^k).$$

For practical implementation, $x_0^k=-\infty, x_{N+1}^k=+\infty$.

2.6.3 Formulas for the barrier pricing

A key Lemma is given below, without proof, for reader's convenience.

Lemma 2.6.1.

$$\mathcal{L}\Big(\min_{t\in[0,T]}\widetilde{X}_t|\widetilde{X}_{t_k}=x_k,k=0,\cdots,M\Big)=\mathcal{L}\Big(\min_{k=0,\cdots,M-1}F^{-1}_{x_k,x_{k+1}}(U_k)\Big),\qquad(2.7)$$

where $(U_k)_{k=0,\cdots,M-1}$ are *iid* random variables uniformly distributed over the unit interval and $F^{-1}_{x_k,x_{k+1}}$ is the inverse function of the conditional cumulative function $F_{x_k,x_{k+1}}$, defined by

$$F_{x,y}(u):=\begin{cases}\exp\Big(-\frac{2M}{T\sigma^2(x)}(u-x)(u-y)\Big) & \text{if } u\le\min(x,y)\\ 1 & \text{otherwise.}\end{cases}$$

Introducing the functions $f(x):=(K-x)^+$ and $g_k(\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k}):=G_{\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k}}(L)$, $k=1,\dots,M$, the expectation in (2.6) reads

$$\mathbb{E}\left(f(\widetilde{X}_{t_M})\prod_{k=1}^{M}g_k(\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k})\right),,$$

so that the price can be obtained recursively (for all the details we refer to Sagna (2012)), as soon as we have an approximation of the transition probability of $\widetilde{X}_{t_k}$ given $\widetilde{X}_{t_{k-1}}$. The idea now is to use the transition matrix of $\widehat{X}_{t_k}$ given $\widehat{X}_{t_{k-1}}$.

Now, by applying recursively the tower property of conditional expectation we get

$$\begin{aligned}\pi_M f &:= \mathbb{E}\left(f(\widetilde{X}_{t_M})\prod_{k=1}^{M} g_k(\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k})\right) = \mathbb{E}\left(\mathbb{E}\left(f(\widetilde{X}_{t_M})\prod_{k=1}^{M} g_k(\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k})|\mathcal{F}_{t_{M-1}}\right)\right)\\ &= \mathbb{E}\left(\mathbb{E}\left(f(\widetilde{X}_{t_M})g_M(\widetilde{X}_{t_{M-1}},\widetilde{X}_{t_M})|\mathcal{F}_{t_{M-1}}\right)\prod_{k=1}^{M-1} g_k((\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k})\right)\\ &= \mathbb{E}\left(H_M(f(X_{t_{M-1}}))\prod_{k=1}^{M-1} g_k(\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k})\right), \qquad (2.8)\end{aligned}$$

where $H_k, k = 1,\dots,M$, is a family of bounded transition kernels related to the conditional densities g_k defined by:

$$H_k f(x) := \mathbb{E}\left[f(\widetilde{X}_{t_k})g_k(\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k})|\widetilde{X}_{t_{k-1}} = x\right]. \qquad (2.9)$$

Furthermore, we have $H_0 f(x) := \pi_0 f = \mathbb{E}\left[f(\widetilde{X}_0)\right]$. It follows from (2.8) that

$$\pi_k f = \pi_{k-1} H_k f, \quad k = 1,\dots,M,$$

so that we finally obtain the recursive expression

$$\pi_M = H_0 \circ H_1 \circ \cdots \circ H_M.$$

The potentiality of the algorithm developed in Section 2.2 gives us an immediate way to compute the transition matrix. In fact, given the quantization grids $\mathbf{x}^k = (x_1^k,\dots,x_N^k)$ of $\widehat{X}_{t_k}$ and $\mathbf{x}^{k-1} = (x_1^{k-1},\dots,x_N^{k-1})$ of $\widehat{X}_{t_{k-1}}$, the computation of the approximated transition probabilities is straightforward:

$$\widehat{H}_k = \sum_{j=1}^{N} g_k(x_i^{k-1},x_j^k)\widehat{p}_k^{ij}\delta_{x_i^{k-1}}, \quad k = 1,\dots,M,$$

and

$$\widehat{H}_0 = \sum_{i=1}^{N} \mathbb{P}(\widehat{X}_0 = x_0^i)\delta_{x_i^0}.$$

The approximated price follows then immediately:

$$P^{\mathrm{LO}} = e^{-rT}\mathbb{E}\left((K-\widetilde{X}_T)^+\prod_{k=1}^{M} G_{\widetilde{X}_{t_{k-1}},\widetilde{X}_{t_k}}(L)\right) \cong e^{-rT}(\widehat{H}_0 \circ \widehat{H}_1 \circ \cdots \widehat{H}_M)f,$$

where

$$\widehat{H}_k = \sum_{j=1}^{N} g_k(x_i^{k-1},x_j^k)\mathbb{P}(\widehat{X}_{t_k} \in C_j(\mathbf{x}^k)|\widehat{X}_{t_{k-1}} \in C_i(\mathbf{x}^{k-1})\delta_{x_i^{k-1}}, \quad k = 1,\dots,M.$$

The approximated transition probabilities are given by:

$$\widehat{p}_k^{ij} := \mathbb{P}\left(\widehat{X}_{t_k} \in C_j(\mathbf{x}^k)|\widehat{X}_{t_{k-1}} \in C_i(\mathbf{x}^{k-1})\right) = \Phi_{0,1}\left(\widehat{x}_{k,j^+}(x_i^{k-1})\right) - \Phi_{0,1}\left(\widehat{x}_{k,j^-}(x_i^{k-1})\right).$$

Chapter 3

Stochastic volatility models: a first approach

This is a joint work with Giorgia Callegaro and Martino Grasselli. The content of this chapter has been published as a paper in *Quantitative Finance*, see Callegaro et al. (2016) in the bibliography.

We provide the first recursive quantization based approach for pricing options in the presence of stochastic volatility. This method can be applied to any model for which an Euler scheme is available for the underlying price process and it allows to price vanillas, as well as exotics, thanks to the knowledge of the transition probabilities for the discretized stock process. We apply the methodology to some celebrated stochastic volatility models, including the Stein and Stein (1991) model and the SABR model introduced in Hagan et al. (2002). A numerical exercise shows that the pricing of vanillas turns out to be accurate; in addition, when applied to some exotics like equity-volatility options, the quantization-based method overperforms by far the Monte Carlo simulation.

3.1 Introduction

The problem of pricing derivative contracts in a stochastic volatility framework has been deeply investigated in the literature. Starting from the pioneering works of Hull and White (1987), Chesney and Scott (1989), Stein and Stein (1991) and Heston (1993), researchers have introduced different stochastic models for the spot volatility process, both with continuous trajectories and including jumps, in order to catch the stylized facts of the implied volatility surface, namely the smile and skew effects. The positivity of the volatility process is another crucial feature: in order to ensure it, Hull and White (1987) assume that the volatility itself follows a Geometric Brownian motion. Chesney and Scott (1989) instead propose to model the volatility as the exponential of a (stationary) Ornstein-Uhlenbeck process. Stein and Stein (1991) assume that the volatility is itself a stationary Ornstein-Uhlenbeck process. This assumption allows to recover the information about the risk neutral density of the underlying through a Fourier based method and to price vanillas in a very fast and efficient way. On the other hand, the Gaussian distribution of the volatility process opens delicate issues on the positivity. Heston (1993) assumes that the volatility is a CIR process, which remains positive under some parameter restrictions (the so called Feller condition). Both the models of Stein and Stein (1991) and Heston (1993)

belong to the class of affine models, for which Duffie et al. (2000) developed a systematic approach in order to perform the Fourier methodology, including the presence of jumps in both the underlying and the volatility process. Examples of models including jumps that have been successfully calibrated in both a single factor and a multi factor setting for the volatility process are e.g. Bates (1996), Jacobs and Li (2008) and Christoffersen et al. (2009).
Besides the undoubtful quality of the Fast Fourier Transform (henceforth FFT) methodology for pricing vanillas in the affine class of models, it is noteworthy that in the case of exotic derivatives the pricing problem still represents a challenge and the Monte Carlo simulation seems to be the only possibility (despite its potentially high computational cost), apart from approximations in the spirit of Sesana et al. (2014) and Date and Islyaev (2015). In addition, the pricing of equity volatility products, like Corridor Variance Swaps, may reveal some delicate numerical issues while applying the FFT approach, as it requires a rigorous check of the integrability of the corresponding characteristic function (see e.g. Da Fonseca et al. (2015)). What is more, the FFT approach clearly cannot be applied when dealing with models whose characteristic function is not known explicitly. Nevertheless, in some cases approximations are available, as in the case of the celebrated SABR model of Hagan et al. (2002), representing a standard in the banking industry, for which we know the shape of the implied volatility e.g. for small time to maturity or small vol of vol. The same holds true for the α-hypergeometric volatility model of Da Fonseca and Martini (2016), for which a Mellin transform and some asymptotics are available. Remarkably, in the last model the volatility remains positive as it is assumed to be the exponential of a mean reverting process.

In this chapter we introduce an alternative pricing method based on quantization. The method is model free, in that it only requires the knowledge of an Euler scheme for the stochastic equation modeling the evolution of the underlying process. It does not require any affine property of the model and we will show in the numerical section that our results over-perform by far the Monte Carlo simulation.

Starting from the results in Callegaro et al. (2015) (Chapter 2), we push forward the methodology to the case where the volatility is stochastic. Of course the passage from local to stochastic volatility models requires additional care, from both the theoretical and the numerical point of view. For example, a direct application of the argument presented in Pagès and Sagna (2015) is not possible in the multi dimensional case.
We will focus on the case where the volatility process is driven by a one dimensional stochastic factor. Nevertheless, most of our results can be easily extended to a multi dimensional setting, with the only drawback of a relevant increase of notational burden. We emphasize that the approach presented here only requires that an Euler scheme is available for the stochastic model, which is the case for most popular stochastic volatility models.

In the numerical illustration, we apply the new methodology to some stochastic volatility models:

- the Stein and Stein (1991) model, which was successively extended by Schöbel and Zhu (1999). This model is affine, so that a Fourier based approach is available. Prices obtained via Monte Carlo and via Fourier techniques will be

our benchmark values when testing the precision of the proposed quantization method.

- the celebrated SABR model, introduced in Hagan et al. (2002) and subsequently extended and analyzed by dozens of researchers both in financial institutions and in academia. Indeed, this model represents the standard reference in the banking industry for most financial models, ranging from equity, interest rates, FX and commodities. Although closed form formulas are not available for vanillas, excellent approximations have been provided in this setting for small maturities and/or small vol of vol. Benchmark prices will be obtained here via standard Monte Carlo.

- the α-hypergeometric stochastic volatility model recently introduced by Da Fonseca and Martini (2016). In this case no explicit formulas for vanillas are available, so that also here benchmark prices will be obtained via standard Monte Carlo.

We will apply the recursive marginal quantization method to all these models and we will compare prices of vanillas with the corresponding benchmark prices. Moreover, we will see that the quantization based approach can also be applied to the pricing of non-vanilla options. In order to give an idea of the flexibility of our approach, we consider a particular class of exotics, namely equity volatility options, that are receiving a growing attention in the financial community. Of course, our methodology can be easily applied as well to other stochastic volatility models, like e.g. the celebrated Heston (1993) one, for which a different technique, still based on Pagès and Sagna (2015), has been developed in Fiorin et al. (2015) in a general multi-dimensional setting. However, the approach of Fiorin et al. (2015) is not suitable for pricing equity volatility products, as the volatility has to be quantized in a very parsimonious number of points in order to be numerically tractable. Pricing of equity volatility options represents a challenging topic, especially in the case of non affine models, for which Monte Carlo represents the only alternative. We are going to show that our quantization based algorithm is much more performing than the standard Monte Carlo approach.

The chapter is organized as follows: in Section 3.2 we develop the quantization based approach in a general stochastic volatility setting. In Section 3.3 we apply the methodology to some stochastic volatility models: we show the performance of the recursive marginal quantization in the pricing of vanillas as well as in the case of equity volatility options, with particular focus on corridor variance swaps. Section 3.4 concludes, while we gather in the Appendix some technical proofs and additional material on quantization which may be useful for the reader in view of the implementation.

3.2 Quantization of a Stochastic Volatility Model

We now extend the one dimensional approach developed in the previous chapter to the case where there is a one dimensional stochastic volatility factor. We shall see that this extension is not trivial.

3.2.1 Model Dynamics and Euler Discretization

Let us consider a general stochastic volatility model described by the following dynamics:

$$\begin{cases} dS_t &= rS_t dt + \alpha(S_t, V_t)\left(\rho dB_t + \sqrt{1-\rho^2}dB_t^{\perp}\right) \\ dV_t &= b(t, V_t)dt + a(t, V_t)dB_t. \end{cases} \tag{3.1}$$

with $S_0 = s_0 > 0, V_0 = v_0 > 0$. Here S is the price process, V is the variance process (more precisely, the term $\alpha(S_t, V_t)/S_t$ is usually referred as to the volatility process) and B and $B^{\perp}$ are independent one-dimensional Brownian motions under the risk neutral probability measure $\mathbb{Q}$ and $\rho \in (-1, 1)$. The assumption of a constant short interest rate r can be easily relaxed.
The functions $\alpha : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $b : [0, T] \times \mathbb{R} \to \mathbb{R}$ and $a : [0, T] \times \mathbb{R} \to \mathbb{R}$ are measurable and they satisfy the boundedness and Lipschitz conditions which ensure the existence of a strong solution of the above SDE for every $s_0, v_0 \in \mathbb{R}$ (see e.g.(Rogers and Williams, 2000, Ch. V, Theorem 11.2)):

Assumption 3.2.1. We suppose that the following conditions hold:

i) there exists C_T such that, for every $t \in [0, T]$

$$|b(t, 0)| + |a(t, 0)| \leq C_T;$$

ii) there exists a real positive constant $[b]_{Lip}$ such that for every $t \in [0, T]$ and every $x, y \in \mathbb{R}$, $|b(t, x) - b(t, y)| \leq [b]_{Lip}|x - y|$;

iii) there exists a real positive constant $[\Sigma]_{Lip}$ such that for every $t \in [0, T]$ and every $u, w \in \mathbb{R}^2$, with $u = (u_1, u_2), v = (w_1, w_2)$, $||\Sigma(t, u) - \Sigma(t, w)|| \leq [\Sigma]_{Lip}|u - w|$, where

$$\Sigma(t, u) := \begin{pmatrix} \rho\alpha(u_1, u_2) & \sqrt{1-\rho^2}\alpha(u_1, u_2) \\ a(t, u_2) & 0 \end{pmatrix}. \tag{3.2}$$

Having fixed a time horizon $T > 0$ and a time discretization grid $\{0 = t_0, t_1, \ldots, t_M = T\}$, with constant time step $\Delta := T/M$, we consider the Euler schemes for S and V: $\widetilde{S}_0 = s_0, \widetilde{V}_0 = v_0$ and for $k \geq 1$

$$\widetilde{S}_{t_{k+1}} = \widetilde{S}_{t_k} + r\widetilde{S}_{t_k}\Delta + \alpha(\widetilde{S}_{t_k}, \widetilde{V}_{t_k})\left(\rho\sqrt{\Delta}Z_k + \sqrt{1-\rho^2}\sqrt{\Delta}Z_k^{\perp}\right) \tag{3.3}$$

$$\widetilde{V}_{t_{k+1}} = \widetilde{V}_{t_k} + b(t_k, \widetilde{V}_{t_k})\Delta + a(t_k, \widetilde{V}_{t_k})\sqrt{\Delta}Z_k, \tag{3.4}$$

where $t_k = k\Delta$ and Z_k and $Z_k^{\perp}, k = 0, \ldots, M$ are i.i.d. standard Gaussian random variables. From the Euler scheme of the process V in Equation (3.4) we obtain[1]

$$\sqrt{\Delta}Z_k = \frac{\widetilde{V}_{t_{k+1}} - \widetilde{V}_{t_k} - b(t_k, \widetilde{V}_{t_k})\Delta}{a(t_k, \widetilde{V}_{t_k})}. \tag{3.5}$$

[1] We implicitly assume that the model is non degenerate, i.e. $a(t_k, \widetilde{V}_{t_k}) \neq 0$ with probability 1, for every k.

Hence, the stochastic component in Equation (3.3), given (3.5), becomes

$$\rho\sqrt{\Delta}Z_k + \sqrt{1-\rho^2}\sqrt{\Delta}Z_k^{\perp} = \rho\left(\frac{\widetilde{V}_{t_{k+1}} - \widetilde{V}_{t_k} - b(t_k, \widetilde{V}_{t_k})\Delta}{a(t_k, \widetilde{V}_{t_k})}\right) + \sqrt{1-\rho^2}\sqrt{\Delta}Z_k^{\perp}, \quad (3.6)$$

which gives an equivalent formulation of the Euler scheme for the price process S in Equation (3.3). For the reader's ease we rewrite the Euler scheme for the pair (S, V):

$$\widetilde{S}_{t_{k+1}} = \widetilde{S}_{t_k} + r\widetilde{S}_{t_k}\Delta + \rho\alpha(\widetilde{S}_{t_k}, \widetilde{V}_{t_k})\left(\frac{\widetilde{V}_{t_{k+1}} - \widetilde{V}_{t_k} - b(t_k, \widetilde{V}_{t_k})\Delta}{a(t_k, \widetilde{V}_{t_k})}\right) + \sqrt{1-\rho^2}\alpha(\widetilde{S}_{t_k}, \widetilde{V}_{t_k})\sqrt{\Delta}Z_k^{\perp} \quad (3.7)$$

$$\widetilde{V}_{t_{k+1}} = \widetilde{V}_{t_k} + b(t_k, \widetilde{V}_{t_k})\Delta + a(t_k, \widetilde{V}_{t_k})\sqrt{\Delta}Z_k. \quad (3.8)$$

The following Lemma is an immediate consequence of equations (3.7) and (3.4) and it represents the natural extension of the one dimensional case to our bi-dimensional setting.

Lemma 3.2.2. For every $0 \leq k \leq M-1$, conditionally on the event $\{\widetilde{S}_{t_k} = s_k, \widetilde{V}_{t_{k+1}} = v_{k+1}, \widetilde{V}_{t_k} = v_k\}$, the random variable $\widetilde{S}_{t_{k+1}}$ is Gaussian:

$$\mathcal{L}\left(\widetilde{S}_{t_{k+1}} \middle| \{\widetilde{S}_{t_k} = s_k, \widetilde{V}_{t_{k+1}} = v_{k+1}, \widetilde{V}_{t_k} = v_k\}\right) = \mathcal{N}\left(\bar{m}_k(s_k, v_{k+1}, v_k), \bar{\sigma}_k^2(s_k, v_k)\right), \quad (3.9)$$

with

$$\bar{m}_k(s_k, v_{k+1}, v_k) := s_k + rs_k\Delta + \rho\alpha(s_k, v_k)\left(\frac{v_{k+1} - v_k - b(t_k, v_k)\Delta}{a(t_k, v_k)}\right)$$

$$\bar{\sigma}_k^2(s_k, v_k) := (1-\rho^2)(\alpha(s_k, v_k))^2\Delta.$$

In the following subsections we are going to explain how the bi-dimensional stochastic process (S, V) can be quantized via recursive marginal quantization. This will consist in quantizing the family of random variables $(\widetilde{S}_{t_k}, \widetilde{V}_{t_k})_{k\in\{0,M\}}$ in equations (3.7) and (3.4) in a recursive way, thanks to Lemma 3.2.2. The idea is the following:

- the (one dimensional) process V can be (independently) quantized as explained in Chapter 2. This will be done in Subsection 3.2.2;
- for what concerns S, it will be necessary to develop an *ad hoc* procedure, since for every $k \geq 0$ the stationary quantizer for $\widetilde{S}_{t_{k+1}}$ will depend on those of $\widetilde{S}_{t_k}$, $\widetilde{V}_{t_k}$ and $\widetilde{V}_{t_{k+1}}$ as the above Lemma 3.2.2 suggests. The quantization of S will be the object of Subsection 3.2.3.

Notation: Since we will deal with quantization grids for both S and V, from now on, for $k = 0, \ldots, M-1$ we will denote by $\widehat{\mathbf{v}}^{N_{k+1}^V} = (\widehat{v}_1^{N_{k+1}^V}, \ldots, \widehat{v}_{N_{k+1}^V}^{N_{k+1}^V})$ the quantizer for $\widetilde{V}_{t_{k+1}}$ and by $\widehat{\mathbf{s}}^{N_{k+1}^S} = (\widehat{s}_1^{N_{k+1}^S}, \ldots, \widehat{s}_{N_{k+1}^S}^{N_{k+1}^S})$ the quantizer for $\widetilde{S}_{t_{k+1}}$.

3.2.2 Quantization of the Process V

Let us consider Equation (3.4): the quantization of the random variables $(\widetilde{V}_{t_k})_{k\in\{0,\ldots,M\}}$ follows the same lines as in (Pagès and Sagna, 2015), as we explained in Chapter 2. We briefly recall here, for reader's convenience, the main steps in the quantization procedure, focusing on the procedure at time t_{k+1}, hence supposing that the first k steps of the recursive quantization algorithm have already been performed.

- **Obtaining a Stationary Grid for $\widetilde{V}_{t_{k+1}}$**
  Suppose we are at time t_{k+1} and we aim at quantizing $\widetilde{V}_{t_{k+1}}$. In the recursive marginal quantization setting, this means that we have already quantized all the random variables $\widetilde{V}_{t_j}$ for $j = 0, \dots, k$.

  Our final target is finding a stationary quantizer for $\widetilde{V}_{t_{k+1}}$, i.e., a critical point of the distortion function D_{k+1} in Equation (2.3), namely:

$$D_{k+1}(\widehat{\mathbf{v}}^{N_{k+1}^V}) = \sum_{i=1}^{N_{k+1}^V} \int_{C_i(\widehat{\mathbf{v}}^{N_{k+1}^V})} \left(v_{k+1} - \widehat{v}_i^{N_{k+1}^V}\right)^2 \mathbb{E}\left[\phi_{m_k(\widetilde{V}_{t_k}),\sigma_k(\widetilde{V}_{t_k})}(v_{k+1})\right] dv_{k+1}.$$

  As explained in Chapter 2, if we replace $\widetilde{V}$ by $\widehat{V}$, then the discrete distortion reads:

$$\widehat{D}_{k+1}(\widehat{\mathbf{v}}^{N_{k+1}^V}) = \sum_{i=1}^{N_{k+1}^V} \int_{C_i(\widehat{\mathbf{v}}^{N_{k+1}^V})} \left(v_{k+1} - \widehat{v}_i^{N_{k+1}^V}\right)^2 \sum_{j=1}^{N_k^V} \phi_{m_k(\widehat{v}_j^{N_k^V}),\sigma_k(\widehat{v}_j^{N_k^V})}(v_{k+1})\, \mathbb{P}(\widehat{V}_{t_k} \in C_j(\widehat{\mathbf{v}}^{N_k^V}))\, dv_{k+1}, \tag{3.10}$$

  where $\widehat{\mathbf{v}}^{N_k^V} = \{\widehat{v}_1^{N_k^V}, \widehat{v}_2^{N_k^V}, \dots, \widehat{v}_{N_k^V}^{N_k^V}\}$ is the N_k^V-quantizer at time t_k relative to $\widetilde{V}_{t_k}$, which is already known, together with the weights $\mathbb{P}(\widehat{V}_{t_k} \in C_j(\widehat{\mathbf{v}}^{N_k^V}))$, $j = 1, \dots, N_k^V$, associated to every point of the grid.

  It is, then, possible to compute the gradient and the Hessian matrix of $\widehat{D}_{k+1}$, hence to obtain a stationary quantizer at time t_{k+1} via, e.g., the Newton-Raphson procedure.

  For reader's convenience, we provide in Appendix 3.5 the gradient and the Hessian matrix for $\widehat{D}_{k+1}$.

- **Approximating the Distribution of $\widetilde{V}_{t_{k+1}}$**
  Remember that

$$\mathbb{P}\left(\widehat{V}_{t_{k+1}} = \widehat{v}_i^{N_{k+1}^V}\right) = \mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in C_i(\widehat{\mathbf{v}}^{N_{k+1}^V})\right).$$

  Once the quantizer $\widehat{\mathbf{v}}^{N_{k+1}^V}$ has been obtained, the probabilities (weights) associated to every point of the grid can be approximated as follows:

$$\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in C_i(\widehat{\mathbf{v}}^{N_{k+1}^V})\right) \approx \sum_{j=1}^{N_k^V} \left[\Phi_{0,1}\left(\widehat{v}_{k+1,i^+}(\widehat{v}_j^{N_k})\right) - \Phi_{0,1}\left(\widehat{v}_{k+1,i^-}(\widehat{v}_j^{N_k})\right)\right] \mathbb{P}(\widetilde{V}_{t_k} \in C_j(\widehat{\mathbf{v}}^{N_k^V})), \tag{3.11}$$

  where $\Phi_{0,1}$ denotes the cumulative distribution function of a standard Gaussian random variable and where

$$\begin{aligned}
\widehat{v}_{k+1,i^+}(\widehat{v}_j^{N_k^V}) &:= \frac{\widehat{v}_i^{N_{k+1}^V} + \widehat{v}_{i+1}^{N_{k+1}^V} - 2m_k(\widehat{v}_j^{N_k^V})}{2\sigma_k(\widehat{v}_j^{N_k^V})}, \\
\widehat{v}_{k+1,i^-}(\widehat{v}_j^{N_k^V}) &:= \frac{\widehat{v}_{i-1}^{N_{k+1}^V} + \widehat{v}_i^{N_{k+1}^V} - 2m_k(\widehat{v}_j^{N_k^V})}{2\sigma_k(\widehat{v}_j^{N_k^V})}.
\end{aligned} \tag{3.12}$$

- **Transition Probabilities from t_k to t_{k+1} for the Process $\widetilde{V}$**
  In the quantization phase we also obtain, for $i \in \{1, \dots, N_{k+1}^V\}$ and $j \in \{1, \dots, N_k^V\}$,

the approximated transition probabilities:

$$\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in C_i(\widehat{\mathbf{v}}^{N_{k+1}^V}) \middle| \widetilde{V}_{t_k} \in C_j(\widehat{\mathbf{v}}^{N_k^V})\right) = \Phi_{0,1}\left(\widehat{v}_{k+1,i^+}(\widehat{v}_j^{N_k^V})\right) - \Phi_{0,1}\left(\widehat{v}_{k+1,i^-}(\widehat{v}_j^{N_k^V})\right).$$

3.2.3 Quantization of the Price Process S

We now focus on the quantization of S at time t_{k+1}, i.e., on the recursive marginal quantization of $\widetilde{S}_{t_{k+1}}$, supposing the first k steps of the recursive quantization algorithm have been performed. As done in previous section, we will highlight the main steps of the procedure in the following subsections, which constitute the main contribution of the chapter.

Obtaining a Stationary Grid for $\widetilde{S}_{t_{k+1}}$

Recall that we denote by $\widehat{\mathbf{s}}^{N_{k+1}} = (\widehat{s}_1^{N_{k+1}}, \dots, \widehat{s}_{N_{k+1}}^{N_{k+1}}))$ an N_{k+1} quantizer of $\widetilde{S}_{t_{k+1}}$ at time t_{k+1}. The distortion function D_{k+1} takes here the following form (recall Equation (2.3)):

$$D_{k+1}(\widehat{\mathbf{s}}^{N_{k+1}}) = \sum_{i=1}^{N_{k+1}} \int_{C_i(\widehat{\mathbf{s}}^{N_{k+1}})} \left(s_{k+1} - \widehat{s}_i^{N_{k+1}}\right)^2 \mathbb{P}(\widetilde{S}_{t_{k+1}} \in ds_{k+1}). \tag{3.13}$$

Thanks to Lemma 3.2.2, we know that in our stochastic volatility setting the distribution of $\widetilde{S}_{t_{k+1}}$ depends not only on $\widetilde{S}_{t_k}$, but also on $\widetilde{V}_{t_k}$ and $\widetilde{V}_{t_{k+1}}$. So, in order to obtain a stationary quantizer for $\widetilde{S}_{t_{k+1}}$ by looking for the critical points of D_{k+1}, first of all it is necessary to make explicit $\mathbb{P}(\widetilde{S}_{t_{k+1}} \in ds_{k+1})$: this is done in the following Lemma whose proof is given in Appendix 3.6.

Lemma 3.2.3. For every $k \in \{0, \dots, M-1\}$ we have:

$$\mathbb{P}(\widetilde{S}_{t_{k+1}} \in ds_{k+1}) = \int_{\mathbb{R}^3} \mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k) \phi_{\bar{m}_k(s_k, v_{k+1}, v_k), \bar{\sigma}_k(s_k, v_k)}(s_{k+1}) ds_{k+1}, \tag{3.14}$$

where $\bar{m}_k(s_k, v_{k+1}, v_k)$ and $\bar{\sigma}_k(s_k, v_k)$ were defined in Lemma 3.2.2.

With this result, it is now possible to compute the gradient and the Hessian matrix of the distortion function in Equation (3.13) in a very efficient way. The interested reader can find in the Appendix 3.7 the gradient and Hessian matrix of the distortion function at time t_{k+1}. Moreover, in the Appendix 3.8 we compute the gradient and the Hessian matrix for the discrete version of the distortion function, to be used in the Newton-Raphson procedure to find a stationary quantizer for $\widetilde{S}_{t_{k+1}}$.

Approximating the Distribution of $\widetilde{S}_{t_{k+1}}$

We now have all the ingredients to prove the main result of the chapter, namely the formula giving the joint probability $\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}))$. This is a non trivial step in the recursive procedure, due to the presence of $\widetilde{V}$ in the dynamics of $\widetilde{S}$. As a corollary, we will obtain the weights associated to every point in the quantizer $\widehat{\mathbf{s}}^{N_{k+1}^S}$.

Notation: We will denote by N_k^S the size of a quantizer for $\widetilde{S}_{t_k}$, while N_k^V will be the size of a quantizer for $\widetilde{V}_{t_k}$.

Proposition 3.2.4. Let $\widehat{\mathbf{s}}^{N_{k-1}^S}$, $\widehat{\mathbf{v}}^{N_k^V}$ and $\widehat{\mathbf{v}}^{N_{k-1}^V}$ be stationary quantizers. The approximated joint probability of $(\widetilde{S}_{t_k}, \widetilde{V}_{t_{k+1}}, \widetilde{V}_{t_k})$ is given by the following recursive formula, for $k = 1, \dots, M$:

$$\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V})) \approx$$

$$\left(\Phi_{0,1}\left(\widehat{v}_{k+1,b^+}(v_c^{N_k^V})\right) - \Phi_{0,1}\left(\widehat{v}_{k+1,b^-}(v_c^{N_k^V})\right)\right)\left(\sum_{d=1}^{N_{k-1}^S}\sum_{e=1}^{N_{k-1}^V}\left[\Phi_{0,1}\left(\widehat{s}_{k,a^+}(s_d^{N_{k-1}^S}, v_c^{N_k^V}, v_e^{N_{k-1}^V})\right)\right.\right. \tag{3.15}$$

$$\left.\left.-\Phi_{0,1}\left(\widehat{s}_{k,a^-}(s_d^{N_{k-1}^S}, v_c^{N_k^V}, v_e^{N_{k-1}^V})\right)\right] \cdot \mathbb{P}(\widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V}))\right),$$

where $\widehat{v}_{k+1,b^-}(v_c^{N_k^V})$ are defined in (3.12) and

$$\begin{aligned}
\widehat{s}_{k+1,j^+}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V}) &= \frac{s_j^{N_{k+1}^S} + s_{j+1}^{N_{k+1}^S} - 2\bar{m}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})}{2\bar{\sigma}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_c^{N_k^V})} \\
\widehat{s}_{k+1,j^-}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V}) &= \frac{s_{j-1}^{N_{k+1}^S} + s_j^{N_{k+1}^S} - 2\bar{m}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})}{2\bar{\sigma}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_c^{N_k^V})}.
\end{aligned} \tag{3.16}$$

Moreover, for $k = 0$, i.e., at the first step of the quantization phase, we have that

$$\begin{aligned}
\mathbb{P}(\widetilde{S}_{t_0} \in C_a(\widehat{\mathbf{s}}^{N_0^S}), \widetilde{V}_{t_1} \in C_b(\widehat{\mathbf{v}}^{N_1^V}), \widetilde{V}_{t_0} \in C_c(\widehat{\mathbf{v}}^{N_0^V})) = &\left[\Phi_{0,1}\left(\widehat{v}_{1,b^+}(\widehat{v}_c^{N_0^V})\right) - \Phi_{0,1}\left(\widehat{v}_{1,b^-}(\widehat{v}_c^{N_0^V})\right)\right] \\
&\cdot \mathbb{P}(\widetilde{S}_{t_0} \in C_a(\widehat{\mathbf{s}}^{N_0^S}))\mathbb{P}(\widetilde{V}_{t_0} \in C_c(\widehat{\mathbf{v}}^{N_0^V})).
\end{aligned} \tag{3.17}$$

Proof. See Appendix 3.9. □

The following corollary, which is a direct consequence of the previous proposition, explains how to compute the weights of the stationary quantization grid (recall that $\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S})) = \mathbb{P}(\widehat{S}_{t_k} = \widehat{s}_a^{N_k^S})$).

Corollary 3.2.5. The weights associated to the points in the stationary quantizer $\widehat{\mathbf{s}}^{N_k^S}$ of $\widetilde{S}_{t_k}$ can be computed using the following formula:

$$\mathbb{P}\left(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S})\right) = \sum_{b=1}^{N_{k+1}^V}\sum_{c=1}^{N_k^V}\mathbb{P}\left(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V})\right). \tag{3.18}$$

Transition Probabilities for $\tilde{S}$

From Proposition 3.2.4, we can easily compute the transition probabilities also for the price process $\tilde{S}$.

Corollary 3.2.6. The transition probabilities from time t_{k-1} to time t_k can be computed using the following formula:

$$\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}) | \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}))$$
$$\approx \frac{1}{\mathbb{P}(\widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}))} \sum_{c=1}^{N_k^V} \sum_{e=1}^{N_{k-1}^V} \Big[\Phi_{0,1}\big(\widehat{s}_{k,a^+}(s_d^{N_{k-1}^S}, v_c^{N_k^V}, v_e^{N_{k-1}^V})\big)$$
$$- \Phi_{0,1}\big(\widehat{s}_{k,a^-}(s_d^{N_{k-1}^S}, v_c^{N_k^V}, v_e^{N_{k-1}^V})\big) \Big] \mathbb{P}(\widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})).$$

In conclusion, the quantization of $\widetilde{S}_{t_{k+1}}$ requires the knowledge of the quantizers of $\widetilde{S}_{t_k}$, of $\widetilde{V}_{t_{k+1}}$ and of $\widetilde{V}_{t_k}$. The quantizers of $\widetilde{V}$ can be computed independently using the formulas in Subsection 3.2.2 and they can be kept off-line. Then, stationary quantizers for $\widetilde{S}$ can be obtained recursively in a very fast way.

3.2.4 Error Analysis

Using quantization to approximate a random variable X allows to estimate an expected value of a function of X (i.e. $\mathbb{E}[f(X)]$) by the finite sum $\mathbb{E}[f(\widehat{X})]$, where $\widehat{X}$ is a quantization of X. In the case when we focus on prices of vanilla options, that is when $X = S_T$, with S in Equation (3.7), and f is the payoff function, the following result (see e.g. Pagès and Sagna (2015) and Pagès and Printems (2005)) is crucial: if f is Lipschitz continuous, then

$$\big|\mathbb{E}[f(X)] - \mathbb{E}[f(\widehat{X})]\big| \leq [f]_{\text{Lip}} ||X - \widehat{X}||_2,$$

where $[f]_{\text{Lip}}$ is the Lipschitz constant of f and $||X - \widehat{X}||_2$ is the quadratic quantization error. The above inequality suggests that quantization is a good approximation method, as long as the quantization error $||X - \widehat{X}||_2$ can be bounded. This motivates the error analysis developed in this subsection.
We denote by X the pair of price and variance processes introduced in Section 3.2, namely $X := (S, V)$.

Proposition 3.2.7. Consider the model in Equation (3.1) and recall the definition of Σ in Equation (3.2). Moreover, suppose that Assumption 3.2.1 holds true. Given a time horizon $T > 0$ and a time discretization grid $\{0 = t_0, \dots, t_M = T\}$, for every $k \in \{1, \dots, M\}$ and for any $\eta \in (0, 1]$ we have

$$||X_{t_k} - \widehat{X}_{t_k}||_2 \leq \sqrt{\frac{C_T}{M}} + K_{2,2,\eta} \sum_{\ell=1}^{k} a_\ell(b, \Sigma, t_k, \Delta, y_0, L, 2+\eta) \frac{1}{\sqrt{N_k^V}},$$

where C_T is a positive constant depending only on T, $K_{2,2,\eta}$ is a universal constant, $L := \max\Big\{[b]_{Lip}, \max_{s\in[0,T]} b(s,0), [\Sigma]_{Lip}, ||\Sigma(\cdot,0)||_{sup}\Big\}$ and a_ℓ are real constants depending on $[b]_{Lip}$ and $[\Sigma]_{Lip}$, that do not explode when M goes to infinity.

Proof. See Appendix 3.10. □

There are still some questions that deserve further investigation. For example, beside the Stein and Stein model, the stochastic volatility models considered in our numerical illustration, present some potential issues as the coefficients are not globally Lipschitz, so that the global error results presented in this Section have to be tailored to those settings. This issue is beyond the scope of this chapter and we leave it for future research.

3.3 Numerical Illustration

In this numerical section we apply the methodology to some stochastic volatility models. We first consider the Stein and Stein (1991) affine model, for which the Fourier based method is available for pricing. Of course our methodology can be easily applied as well to other affine stochastic volatility models like e.g. the Heston (1993) one, for which a slightly different approach has been developed in Fiorin et al. (2015). Then we consider some non affine models that do not admit closed form formulas for the price of vanillas, like the celebrated SABR model introduced by Hagan et al. (2002) and the α-hypergeometric volatility model recently introduced by Da Fonseca and Martini (2016). In these cases only approximated formulas based on asymptotic results are available, besides of course the Monte Carlo approach.
For the pricing of vanillas, we recall that having computed the grids $\mathbf{s}_T = (s_T^i)_{i=1,\dots,N_s}$ and the weights $\mathbb{P}(S_T \in C_i(\mathbf{s}_T))$, the price at time 0 of a Call option with maturity T and strike price K can be straightforwardly approximated as follows:

$$C_{0,T} = e^{-rT}\mathbb{E}[(S_T - K)^+] \approx e^{-rT}\sum_{i=1}^{N_s} \max(s_T^i - K, 0)\mathbb{P}(S_T \in C_i(\mathbf{s}_T)).$$

All the computations have been performed using *Matlab2015* on a CPU 2.4 GHz and 8 Gb memory computer. The *Matlab* code is available upon request.

3.3.1 Vanillas in the Stein and Stein (1991) Model

Stein and Stein (1991) assume that the instantaneous variance follows an Ornstein-Uhlenbeck process. The model is described by the following dynamics:

$$\begin{aligned} dS_t &= rS_t dt + V_t S_t(\rho dB_t + \sqrt{1-\rho^2}dB_t^\perp) \\ dV_t &= \kappa(\theta - V_t)dt + \xi dB_t. \end{aligned}$$

Let $t_k = k\Delta, k = 0, \dots, M$ and denote S_{t_k} by S_k and V_{t_k} by V_k. The Euler scheme reads

$$\begin{aligned} \widetilde{S}_{k+1} &= \widetilde{S}_k + r\widetilde{S}_k\Delta + \rho\widetilde{V}_k\widetilde{S}_k\sqrt{\Delta}Z_k + \sqrt{1-\rho^2}\widetilde{V}_k\widetilde{S}_k\sqrt{\Delta}Z_k^\perp, \\ \widetilde{V}_{k+1} &= \widetilde{V}_k + \kappa(\theta - \widetilde{V}_k)\Delta + \xi\sqrt{\Delta}Z_k, \end{aligned}$$

where $Z_k = B_{k+1} - B_k$ and $Z_k^\perp = B_{k+1}^\perp - B_k^\perp$ are i.i.d. Gaussian random variables. Exploiting the Cholesky decomposition we get

$$\begin{aligned} \widetilde{S}_{k+1} &= \widetilde{S}_k + \frac{\rho}{\xi}\left[\widetilde{V}_{k+1} - \widetilde{V}_k - \kappa(\theta - \widetilde{V}_k)\Delta\right]\widetilde{V}_k\widetilde{S}_k + \sqrt{1-\rho^2}\sqrt{\Delta}\widetilde{V}_k\widetilde{S}_k Z_k^\perp \\ \widetilde{V}_{k+1} &= \widetilde{V}_k + \kappa(\theta - \widetilde{V}_k)\Delta + \xi\sqrt{\Delta}Z_k. \end{aligned}$$

We consider the following set of parameters for the Stein and Stein (1991) model:

$$\begin{aligned} \xi = 0.05, \quad \kappa = 3, \quad \theta = 0.3, \quad \rho = -0.5, \\ r = 0.04, \quad T = 1, \quad V_0 = 0.25, \quad S_0 = 100, \end{aligned}$$

while for the time discretization and for the quantization we take

$$\Delta = \frac{1}{20}, \qquad N_k^V = 30, \qquad N_k^S = 30, \quad k = 1, \dots, M,$$

where N_k^V and N_k^S are the sizes, respectively, of the grids for the volatility and for the price.
We compare the price given by the quantization with a benchmark price given by the closed formulas as in the paper of Schöbel and Zhu (1999). The computation of the grids is made using a Newton-Raphson procedure both for the volatility and the price processes.
From Table 3.1 it turns out that pricing of vanillas via quantization is accurate. The quantization of the volatility process is almost immediate, while for the underlying it takes about 10 seconds, which is much less than what we can obtain with a classic quantization method based on stochastic algorithms. In Table 3.2 we provide relative errors in percentage for longer maturities up to 5 years. We emphasize that, although pricing via quantization cannot be comparable with the (faster) Fourier approach available for the affine models, quantization can be applied also to non vanilla products as we are going to show in the next section. What is more, quantization can be applied also to non affine models, as we will show in the next subsection. In Figure 3.1 we report the implied volatility smile induced by the Stein and Stein (1991) model of our example. Figure 3.2 displays the quantization grids used in the procedure.

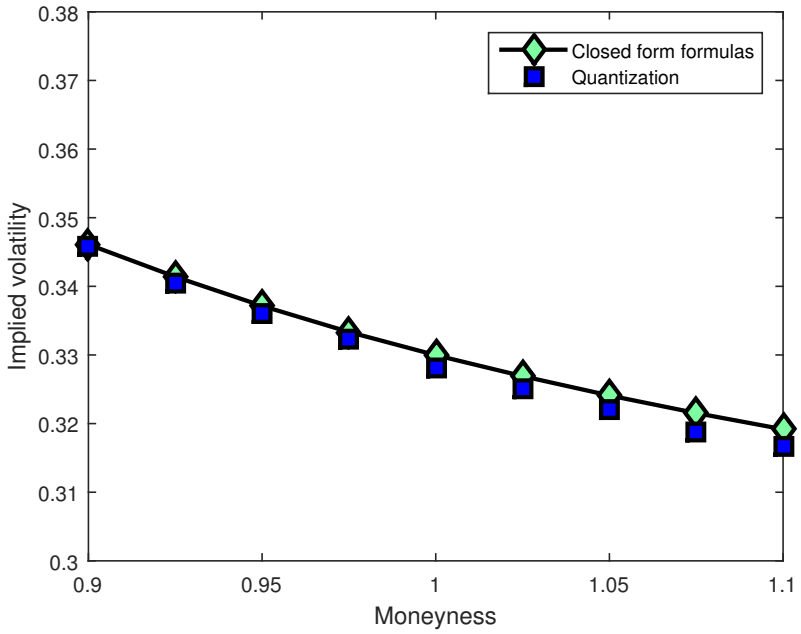


FIGURE 3.1: Example of fit of the implied volatility smile in the Stein and Stein (1991) model. We consider 9 strikes with maturity 1 year.

3.3.2 Vanillas in the SABR Model of Hagan et al. (2002)

The Hagan et al. (2002) (non affine) model is characterized by the following dynamics (here S denotes the forward price under the forward measure):

$$\begin{aligned} dS_t &= V_t S_t^{\beta}(\rho dB_t + \sqrt{1-\rho^2} dB_t^{\perp}) \\ dV_t &= \alpha V_t dB_t. \end{aligned}$$

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 60$	40.1947	40.2294	0.0864
$K = 65$	35.4500	35.4804	0.0858
$K = 70$	30.8727	30.9078	0.1138
$K = 75$	26.5234	26.5582	0.1311
$K = 80$	22.4604	22.4744	0.0623
$K = 85$	18.7361	18.7488	0.0677
$K = 90$	15.3903	15.3825	0.0506
$K = 95$	12.4451	12.4277	0.1403
$K = 100$	9.9070	9.8681	0.3925
$K = 105$	12.8284	12.6995	1.0053
$K = 110$	16.0608	15.9153	0.9058
$K = 115$	19.6277	19.4646	0.8308
$K = 120$	23.4893	23.3295	0.6801
$K = 125$	27.6002	27.4349	0.5990
$K = 130$	31.9187	31.7473	0.5370
$K = 135$	36.4042	36.2324	0.4719
$K = 140$	41.0213	40.8566	0.4016

TABLE 3.1: Comparison between pricing Call and Put options via closed form formulas and quantization in the Stein and Stein (1991) model. When the strike is lower or equal than 100 we are pricing a Call option, when it is greater than 100 we are pricing a Put option. The inverse of the Hessian matrix, which is tridiagonal and symmetric, is calculated using the LU-decomposition. Computational time for the quantization of the volatility process: 0.079 seconds. Computational time for the quantization of the price process: 10.9689 seconds.

The Euler scheme is given by

$$\begin{aligned} \widetilde{S}_{k+1} &= \widetilde{S}_k + \rho \widetilde{S}_k^{\beta} \widetilde{V}_k \sqrt{\Delta} Z_k + \sqrt{1-\rho^2} \widetilde{S}_k^{\beta} \widetilde{V}_k \sqrt{\Delta} Z_k^{\perp} \\ \widetilde{V}_{k+1} &= \widetilde{V}_k + \alpha \widetilde{V}_k \sqrt{\Delta} Z_k \end{aligned}$$

Strike	6 months	8 months	1 year	2 years	3 years	5 years
$K = 80$	0.0163	0.4872	0.0623	0.2878	0.7504	0.1215
$K = 85$	0.0919	0.6252	0.0677	0.2429	0.7759	0.0288
$K = 90$	0.2047	0.7637	0.0506	0.1525	0.7373	0.0196
$K = 95$	0.3347	1.0095	0.1403	0.0033	0.6099	0.2327
$K = 100$	0.5438	1.3913	0.3925	0.2134	0.6213	0.1305
$K = 105$	0.6965	0.6958	1.0053	1.0965	0.5548	0.4547
$K = 110$	0.5546	0.4912	0.9058	1.0417	0.5707	0.3362
$K = 115$	0.4218	0.3275	0.8308	0.9646	0.6164	0.2715
$K = 120$	0.3410	0.2650	0.6801	0.8993	0.6939	0.0645

TABLE 3.2: Relative error in percentage between pricing Call and Put options via closed form formulas and quantization in the Stein and Stein (1991) model for $T = 1/2, 2/3, 1, 2, 3, 5$ years. When the strike is smaller than 100 we consider Calls, while when it is greater than 100 we take Puts.

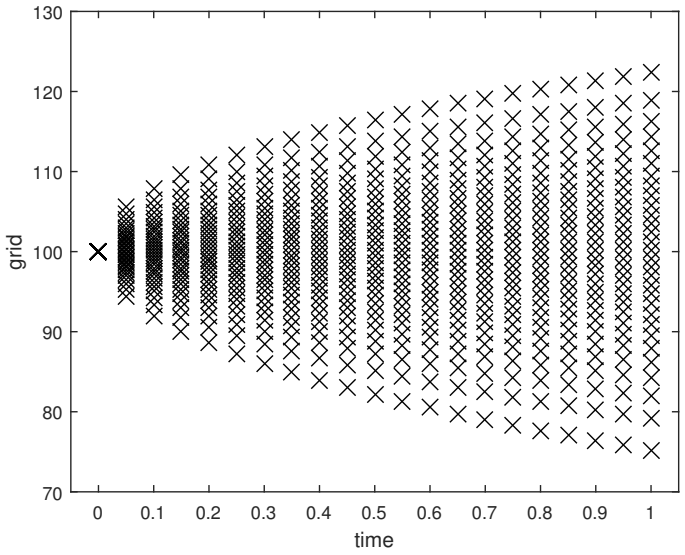


FIGURE 3.2: Quantization grids for the Stein and Stein (1991) model. Time step is taken as $\Delta = \frac{1}{20}$, while $N_k^V = N_k^S = 30, k = 1, \dots, M$ is the size of the grids for the volatility and the underlying.

and exploiting the Cholesky decomposition we get

$$\widetilde{S}_{k+1} = \widetilde{S}_k + \frac{\rho}{\alpha}(\widetilde{V}_{k+1} - \widetilde{V}_k)\widetilde{S}_k^\beta + \sqrt{1-\rho^2}\sqrt{\Delta}\widetilde{S}_k^\beta \widetilde{V}_k Z_k^\perp$$
$$\widetilde{V}_{k+1} = \widetilde{V}_k + \alpha \widetilde{V}_k \sqrt{\Delta} Z_k.$$

The parameters chosen for the pricing are

$$\alpha = 0.224 \qquad \beta = 0.75 \qquad \rho = -0.824$$
$$T = 1 \qquad V_0 = 0.0719 \qquad S_0 = 100,$$

and for the time discretization and the quantization we use as before:

$$\Delta = \frac{1}{20}, \qquad N_k^V = 30, \qquad N_k^S = 30, \quad k = 1, \dots, M.$$

The benchmark price is given by a Monte Carlo procedure with 10^6 paths. The quantization grids are obtained with a Newton-Raphson procedure both for the volatility and the price processes.
Table 3.3 displays the prices given by quantization and by the approximation given by Hagan et al. (2002). Also in this case we see that pricing is accurate even if not immediate when compared with the approximation formulas. Note however that pricing should be compared with the true benchmark given by the Monte Carlo simulation that takes around 30 seconds with 5×10^5 paths. In Table 3.4 we provide relative errors in percentage for longer maturities up to 5 years.

3.3.3 Vanillas in the α-Hypergeometric Model of Da Fonseca and Martini (2016)

The α-hypergeometric model of Da Fonseca and Martini (2016) is described by the following dynamics:

$$\begin{aligned} dS_t &= S_t e^{V_t}(\rho dB_t + \sqrt{1-\rho^2} dB_t^{\perp}), \\ dV_t &= (a - be^{\alpha V_t})dt + \sigma dB_t. \end{aligned}$$

Note that the volatility remains strictly positive in this model. The Euler scheme of the model is given by

$$\begin{aligned} \widetilde{S}_{k+1} &= \widetilde{S}_k + \rho \widetilde{S}_k e^{\widetilde{V}_k}\sqrt{\Delta} Z_k + \sqrt{1-\rho^2}\widetilde{S}_k e^{\widetilde{V}_k}\sqrt{\Delta} Z_k^{\perp} \\ \widetilde{V}_{k+1} &= \widetilde{V}_k + (a - be^{\alpha \widetilde{V}_k})\Delta + \sigma\sqrt{\Delta} Z_k \end{aligned}$$

from which we get

$$\begin{aligned} \widetilde{S}_{k+1} &= \widetilde{S}_k + \frac{\rho}{\sigma}\big(\widetilde{V}_{k+1} - \widetilde{V}_k - (a - be^{\alpha \widetilde{V}_k})\Delta\big)\widetilde{S}_k e^{\widetilde{V}_k} + \sqrt{1-\rho^2}\sqrt{\Delta}\widetilde{S}_k e^{\widetilde{V}_k} Z_k^{\perp} \\ \widetilde{V}_{k+1} &= \widetilde{V}_k + (a - be^{\alpha \widetilde{V}_k})\Delta + \sigma\sqrt{\Delta} Z_k. \end{aligned}$$

The parameters selected for pricing are

$$\begin{aligned} a = 1, \qquad b = 1, \qquad \alpha = 2, \qquad \sigma = 1, \\ \rho = -0.2, \qquad T = 1, \qquad V_0 = \log(0.25), \qquad S_0 = 100, \end{aligned}$$

and for the time discretization and the quantization we use as before:

$$\Delta = \frac{1}{20}, \qquad N_k^V = 30, \qquad N_k^S = 30, \quad k = 1, \dots, M.$$

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 60$	40.0475	40.0195	0.0699
$K = 65$	35.1042	35.0694	0.0992
$K = 70$	30.2179	30.1716	0.1531
$K = 75$	25.4346	25.3742	0.2375
$K = 80$	20.8257	20.7607	0.3120
$K = 85$	16.4909	16.4208	0.4252
$K = 90$	12.5524	12.4800	0.5766
$K = 95$	9.1346	9.0628	0.7859
$K = 100$	6.3321	6.2564	1.1946
$K = 105$	9.1780	9.1159	0.6768
$K = 110$	12.6315	12.5798	0.4098
$K = 115$	16.5922	16.5446	0.2869
$K = 120$	20.9336	20.8837	0.2386
$K = 125$	25.5358	25.5008	0.1371
$K = 130$	30.3040	30.2700	0.1121
$K = 135$	35.1719	35.1424	0.0838
$K = 140$	40.0975	40.0770	0.0511

TABLE 3.3: Comparison between pricing a Call via closed form (approximation) formulas and quantization in the SABR model of Hagan et al. (2002). When the strike is lower or equal than 100 we are pricing a Call option, when it is greater than 100 we are pricing a Put option. The inverse of the Hessian matrix, which is tridiagonal and symmetric, is calculated using the LU-decomposition. Computational time for the quantization of the volatility process: 0.0796 seconds. Computational time for the price process: 12.1071 seconds.

The benchmark price is given by a Monte Carlo procedure with 10^6 paths, due to the fact that, up to our knowledge, no closed form formulas are available for pricing vanillas in this setting. The quantization grids are obtained with a Newton-Raphson procedure both for the volatility and the price processes.

Strike	6 months	8 months	1 year	2 years	3 years	5 years
$K = 80$	0.1442	0.2010	0.3120	0.6003	0.9684	1.4529
$K = 85$	0.2365	0.3124	0.4252	0.7643	1.0255	1.5444
$K = 90$	0.4489	0.4702	0.5766	0.8719	1.0788	1.6065
$K = 95$	0.7356	0.6999	0.7859	1.0474	1.2036	1.6463
$K = 100$	1.0909	1.1682	1.1946	1.1475	1.5285	1.6887
$K = 105$	0.5121	0.6236	0.6768	0.8248	0.8341	1.0884
$K = 110$	0.2909	0.3296	0.4098	0.4961	0.5877	0.8479
$K = 115$	0.2128	0.2330	0.2869	0.4475	0.5004	0.7162
$K = 120$	0.1134	0.1472	0.2386	0.3428	0.4650	0.6570

TABLE 3.4: Relative error in percentage between pricing a Call via closed form (approximation) formulas and quantization in the SABR model of Hagan et al. (2002) for $T = 1/2, 2/3, 1, 2, 3, 5$ years. When the strike is smaller than 100 we consider Calls, while when it is greater than 100 we take Puts.

Table 3.5 shows the comparison between the quantization prices and the Monte Carlo ones for the α-hypergeometric model of Da Fonseca and Martini (2016). Here the computational time required for the quantization of the underlying process is lower than in the previous models, but it still remains quite high (about 7 seconds). On the other hand, a Monte Carlo simulation with 10^5 paths takes about 10 seconds and gives prices which are much less accurate (see column "MC price" in Table 3.5). In Table 3.6 we provide relative errors in percentage for longer maturities up to 5 years.

3.3.4 Pricing of an Exotic: Volatility Corridor Swap

Here we consider the pricing of a Volatility Corridor Swap, that is a non vanilla contract belonging to the class of equity-volatility derivatives. A Corridor Variance Swap is a generalisation of a standard variance swap in that the volatility is accumulated only when the underlying stock is within a pre-specified interval $[L, H]$, see Carr and Lewis (2004). For the affine models a Fourier approach has been developed for these products (see e.g. Da Fonseca et al. (2015) and references therein). In this subsection we focus on the SABR and the α- hypergeometric models for which such Fourier approach cannot be applied and we show that the quantization method we propose is very efficient when compared with the Monte Carlo benchmark.

Consider the formula giving the price at time 0 of a Volatility Corridor Swap (wlog

Strike	Bench. price	Quant. price	MC price	Rel. error (%)
$K = 60$	44.1617	44.4653	[42.7957, 45.5276]	0.6876
$K = 65$	40.2496	40.5186	[38.9443, 41.5548]	0.6684
$K = 70$	36.5165	36.7643	[35.2690, 37.7640]	0.6788
$K = 75$	32.9795	33.2305	[31.7867, 34.1723]	0.7611
$K = 80$	29.6571	29.8535	[28.5158, 30.7984]	0.6623
$K = 85$	26.5673	26.7371	[25.4743, 27.6603]	0.6391
$K = 90$	23.7232	23.8925	[22.6753, 24.7711]	0.7134
$K = 95$	21.1375	21.2545	[20.1315, 22.1435]	0.5537
$K = 100$	18.8159	18.8704	[17.8488, 19.7831]	0.2896
$K = 105$	21.8264	21.7331	[21.1575, 22.4954]	0.4276
$K = 110$	25.0123	24.8636	[24.2792, 25.7453]	0.5944
$K = 115$	28.4277	28.2293	[27.6277, 29.2277]	0.6980
$K = 120$	32.0491	31.8511	[31.1796, 32.9187]	0.6180
$K = 125$	35.8509	35.5602	[34.9094, 36.7924]	0.8109
$K = 130$	39.8070	39.5573	[38.7913, 40.8227]	0.6272
$K = 135$	43.8971	43.5695	[42.8051, 44.9891]	0.7462
$K = 140$	48.1015	47.8111	[46.9314, 49.2716]	0.6037

TABLE 3.5: Comparison between pricing a Call via closed form formulas and quantization in the α-hypergeometric model of Da Fonseca and Martini (2016). When the strike is lower or equal than 100 we are pricing a Call option, when it is greater than 100 we are pricing a Put option. The inverse of the Hessian matrix, which is tridiagonal and symmetric, is calculated using the LU-decomposition. Computational time for the quantization of the volatility process: 0.1109 seconds. Computational time for the price process: 6.8141 seconds.

we do not consider the presence of the strike price due to the linearity of the contract):

$$S(L,H,T) := \mathbb{E}\left[\frac{1}{T}\int_0^T V_u \mathbb{1}_{\{L<S_u<H\}} du\right],$$

Strike	6 months	8 months	1 year	2 years	3 years	5 years
$K = 80$	0.1727	0.3272	0.6128	0.6376	1.3259	0.8820
$K = 85$	0.1991	0.2761	0.5919	3.3031	0.9801	1.1970
$K = 90$	0.2367	0.2067	0.6676	0.6973	0.6082	0.9134
$K = 95$	0.4040	0.1172	0.5123	0.5973	1.2570	0.7741
$K = 100$	0.8696	0.0282	0.2568	0.6090	0.6490	0.7868
$K = 105$	0.9646	1.1712	0.2504	0.7204	1.0005	0.7519
$K = 110$	0.7996	1.1538	0.4478	1.7931	0.9324	0.7088
$K = 115$	0.7227	1.1119	0.5725	1.5081	0.8718	0.8042
$K = 120$	0.6479	0.9285	0.5040	1.0315	0.4906	0.8892

TABLE 3.6: Relative error in percentage between pricing a Call via closed form formulas and quantization in the α-hypergeometric model of Da Fonseca and Martini (2016) for $T = 1/2, 2/3, 1, 2, 3, 5$ years. When the strike is smaller than 100 we consider Calls, while when it is greater than 100 we take Puts.

that can be rewritten as

$$\frac{1}{T}\sum_{i=0}^{M-1}\mathbb{E}\int_{t_i}^{t_{i+1}} V_u \mathbb{1}_{\{L<S_u<H\}} du,$$

so that the expected value can be approximated using the quantization method, i.e.:

$$\begin{aligned}\frac{1}{T}\sum_{i=0}^{M-1}\mathbb{E}\int_{t_i}^{t_{i+1}} V_u \mathbb{1}_{\{L<S_u<H\}} du &\simeq \frac{1}{T}\sum_{i=0}^{M-1}\mathbb{E}\int_{t_i}^{t_{i+1}} \widehat{V}_u \mathbb{1}_{\{L<\widehat{S}_u<H\}} du \\ &= \frac{1}{T}\sum_{i=0}^{M-1}\int_{t_i}^{t_{i+1}} \mathbb{E}\left(\widehat{V}_u \mathbb{1}_{\{L<\widehat{S}_u<H\}}\right) du.\end{aligned}$$

Hence we can approximate $S(L,H,T)$ by means of $\widehat{S}(L,H,T)$:

$$\begin{aligned}\widehat{S}(L,H,T) &= \frac{1}{T}\sum_{i=0}^{M-1}\int_{t_i}^{t_{i+1}}\sum_{j,\ell}\widehat{v}_j \mathbb{1}_{\{L<\widehat{s}_\ell<H\}}\mathbb{P}(\widehat{S}_u=\widehat{s}_\ell,\widehat{V}_u=\widehat{v}_j)\,du \\ &= \frac{\Delta}{T}\sum_{i=0}^{M-1}\sum_{j=1}^{N_i^V}\sum_{\ell=1}^{N_i^S}\widehat{v}_j^{N_i^V}\mathbb{1}_{\{L<\widehat{s}_\ell^{N_i^S}<H\}}\mathbb{P}(\widehat{S}_{t_i}=\widehat{s}_\ell^{N_i^S},\widehat{V}_{t_i}=\widehat{v}_j^{N_i^V}) \\ &= \frac{1}{M}\sum_{i=0}^{M-1}\sum_{j=1}^{N_i^V}\sum_{\ell=1}^{N_i^S}\widehat{v}_j^{N_i^V}\mathbb{1}_{\{L<\widehat{s}_\ell^{N_i^S}<H\}}\mathbb{P}(\widehat{S}_{t_i}=\widehat{s}_\ell^{N_i^S},\widehat{V}_{t_i}=\widehat{v}_j^{N_i^V}).\end{aligned}$$

We use the same grids used in the pricing of vanillas, so that the pricing is immediate.

Corridor	Bench price	Q price	MC price	Rel error (%)
$L = 80, H = 120$	0.2487	0.2484	$[0.2397, 0.2576]$	0.0973
$L = 85, H = 115$	0.2447	0.2451	$[0.2344, 0.2549]$	0.1673
$L = 90, H = 110$	0.2278	0.2297	$[0.2131, 0.2430]$	0.8551

TABLE 3.7: Pricing of a Volatility Corridor Swap in the SABR model for different values of the corridor. We fixed $\alpha = 0.25, \beta = 0.75, \rho = -0.75$.

Parameter	Bench price	Q price	MC price	Rel error (%)
$\beta = 0.5$	0.2498	0.2492	$[0.2413, 0.2588]$	0.2296
$\beta = 0.7$	0.2498	0.2491	$[0.2408, 0.2583]$	0.2738
$\beta = 0.9$	0.2282	0.2301	$[0.2139, 0.2437]$	0.8554

TABLE 3.8: Pricing of a Volatility Corridor Swap in the SABR model for different values of the parameter β. We fixed $\alpha = 0.25, \rho = -0.75, L = 80, H = 120$.

Parameter	Bench price	Q price	MC price	Rel error (%)
$\alpha = 0.1$	0.2494	0.2496	$[0.2453, 0.2535]$	0.0751
$\alpha = 0.3$	0.2482	0.2476	$[0.2377, 0.2590]$	0.2138
$\alpha = 0.5$	0.2462	0.2440	$[0.2290, 0.2635]$	0.8969

TABLE 3.9: Pricing of a Volatility Corridor Swap in the SABR model for different values of the parameter α. We fixed $\beta = 0.75, \rho = -0.75, L = 80, H = 120$.

In Tables 3.7-3.10 we display the price of the Volatility Corridor Swap for different values of the parameters within the SABR model. We notice that pricing is fast and accurate. What is more, once the SABR model has been quantized, pricing of vanillas or Volatility Corridor Swaps is almost immediate, while a Monte Carlo simulation still needs a computational cost (about 25 seconds with 1.5×10^5 paths for a Volatility Corridor Swap in order to reach a similar precision of the recursive marginal

Parameter	Bench price	Q price	MC price	Rel error (%)
$\rho = -0.4$	0.2490	0.2486	$[0.2401, 0.2581]$	0.1522
$\rho = -0.6$	0.2490	0.2486	$[0.2399, 0.2579]$	0.1552
$\rho = -0.8$	0.2487	0.2484	$[0.2396, 0.2574]$	0.0861

TABLE 3.10: Pricing of a Volatility Corridor Swap in the SABR model for different values of the parameter ρ. We fixed $\alpha = 0.25, \beta = 0.75, L = 80, H = 120$.

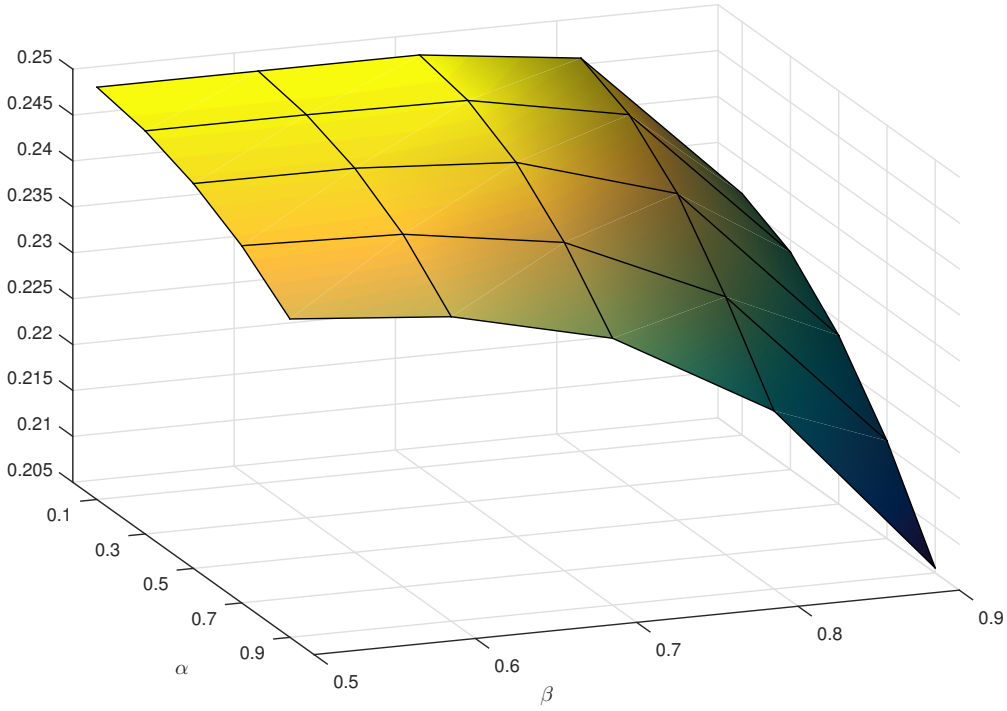


FIGURE 3.3: Price of a Volatility Corridor Swap in the SABR model as a function of parameters α, β.

quantization approach). Here the benchmark price is computed via a Monte Carlo simulation with 10^6 paths.
In Tables 3.11-3.12 we repeat the same experiments for the α-hypergeometric model. The benchmark price is still computed via a Monte Carlo simulation with 10^6 paths. Also in this case the pricing is fast and accurate. Moreover, once the model has been quantized, pricing of vanillas or Volatility Corridor Swaps is almost immediate, while a Monte Carlo simulation requires a higher computational cost (about 13 seconds with $\times 10^5$ paths for a Volatility Corridor Swap).

Parameter	Bench price	Q price	MC price	Rel error (%)
$\alpha = 1$	0.2344	0.2342	$[0.2131, 0.2557]$	0.0773
$\alpha = 1.5$	0.2396	0.2396	$[0.2176, 0.2617]$	0.0012
$\alpha = 2$	0.2427	0.2442	$[0.2203, 0.2651]$	0.6362

TABLE 3.11: Pricing of a Volatility Corridor Swap in the α-Hypergeometric model for different values of the parameter α. We fixed $a = 1, b = 1, \sigma = 1, \rho = -0.2, L = 80, H = 120$.

Parameter	Bench price	Q price	MC price	Rel error (%)
$\rho = -0.4$	0.2343	0.2348	$[0.2135, 0.2550]$	0.2415
$\rho = -0.6$	0.2349	0.2372	$[0.2150, 0.2548]$	0.9588
$\rho = -0.8$	0.2361	0.2384	$[0.2173, 0.2549]$	0.9623

TABLE 3.12: Pricing of a Volatility Corridor Swap in the α-Hypergeometric model for different values of the parameter ρ. We fixed $a = 1, b = 1, \sigma = 1, \alpha = 1, L = 80, H = 120$.

3.4 Conclusion

In this chapter we have presented the first recursive quantization based approach for pricing in the presence of stochastic volatility. Our framework is flexible enough to include most popular volatility models and it applies to affine as well as to non affine models. Pricing of vanilla options is accurate and, thanks to the knowledge of the transition density for the underlying, it is possible to price efficiently also non vanilla options. We tested our procedure on three different stochastic volatility settings with an increasing difficulty in terms of analytical tractability. Moreover, as an application to non vanilla options, we considered the pricing of a Corridor Volatility Swap under the SABR and the α-hypergeometric (non affine) models, where we showed that pricing is fast and accurate when compared with the Monte Carlo one. Recursive marginal quantization can be applied also to the pricing of other path dependent derivatives, for which some attempts have already been considered in the literature from an optimal quantization perspective (see e.g. Pagès and Wilbertz (2012) and Bally et al. (2005) for American options, and Sagna (2012) for barrier options). Although existing results on marginal quantization are still preliminary, we believe they are promising enough to motivate further research in the field. In particular, we are confident to be able to improve the numerical performance achieved with classic techniques like in Longstaff and Schwartz (2001) for American options and Lipton and McGhee (2002) for Barrier options. Another room for improvement relies in the possibility to allow for more sophisticated discretisation schemes, like e.g. Milstein, which are currently under investigation.

Let us also mention that hybrid models, such as stochastic local volatility (SLV) models, are becoming popular in practice, since they allow to improve pricing accuracy for exotic options, including volatility options, while matching the whole implied volatility surface. For a comprehensive study on the application of hybrid SLV models to option pricing we refer e.g. to the PhD thesis Tian (2013). In general a hybrid SLV model is specified in the form

$$\begin{cases} dS_t &= \mu_1(S_t,t)dt + L(S_t,t)\sigma_1(S_t,V_t,t)dW_t^1 \\ dV_t &= \mu_2(V_t,t)dt + \sigma_2(V_t,V_t,t)dW_t^2, \end{cases}$$

with $dW_t^1 dW_t^2 = \rho dt$ and where L is known as "leverage function", which has to be carefully determined by market information. We immediately see that our Lemma 3.2 and all the subsequent technical results in Section 3, needed to perform the quantization algorithm, still hold true, obviously as soon as the leverage function has been specified. Hence, our numerical procedure can be safely applied also to hybrid SLV models.
Finally, from a calibration perspective, we expect to gain in robustness, as in the case of local volatility models investigated in Callegaro et al. (2015) (Chapter 2), where they showed that marginal quantization performs very well and does not suffer the typical problems related to the choice of the starting point for the calibration.

3.5 Gradient and Hessian for the Quantization of $\widetilde{V}_{t_{k+1}}$.

The gradient and Hessian of the distortion function $\widehat{D}_{k+1}$ in Equation (3.10) are:

$$\begin{aligned} \frac{\partial \widehat{D}_{k+1}}{\partial \widehat{v}_j^{N_{k+1}^V}}(\widehat{\mathbf{v}}^{N_{k+1}^V}) =& 2\sum_{i=1}^{N_k^V} \Big\{ (\widehat{v}_j^{N_{k+1}^V} - m_k(\widehat{v}_i^{N_k^V})) \Big(\Phi_{0,1}(\widehat{v}_{k+1,j^+}(\widehat{v}_i^{N_k^V})) - \Phi_{0,1}(\widehat{v}_{k+1,j^-}(\widehat{v}_i^{N_k^V})) \Big) + \\ &+ \sigma_k(\widehat{v}_i^{N_k^V}) \Big(\phi_{0,1}(\widehat{v}_{k+1,j^+}(\widehat{v}_i^{N_k^V})) - \phi_{0,1}(\widehat{v}_{k+1,j^-}(\widehat{v}_i^{N_k^V})) \Big) \Big\} \mathbb{P}(\widetilde{V}_{t_k} \in C_i(\widehat{\mathbf{v}}^{N_k^V})). \end{aligned} \tag{3.19}$$

where

$$\begin{aligned} \widehat{v}_{k+1,j^+}(\widehat{v}_i^{N_k^V}) &:= \frac{\widehat{v}_j^{N_{k+1}^V} + \widehat{v}_{j+1}^{N_{k+1}^V} - 2m_k(\widehat{v}_i^{N_k^V})}{2\sigma_k(\widehat{v}_i^{N_k^V})}, \\ \widehat{v}_{k+1,j^-}(\widehat{v}_i^{N_k^V}) &:= \frac{\widehat{v}_{j-1}^{N_{k+1}^V} + \widehat{v}_j^{N_{k+1}^V} - 2m_k(\widehat{v}_i^{N_k^V})}{2\sigma_k(\widehat{v}_i^{N_k^V})}, \end{aligned} \tag{3.20}$$

and with $\Phi_{0,1}$ denoting the cumulative distribution function of a standard Gaussian random variable.

Focusing on the Hessian tri-diagonal matrix, the diagonal terms are:

$$\begin{aligned}\frac{\partial^2 \widehat{D}_{k+1}}{\partial^2 \widehat{v}_j^{N_{k+1}^V}}(\widehat{\mathbf{v}}^{N_{k+1}^V}) =& 2\sum_{i=1}^{N_k^V}\Bigg\{\Bigg(\Phi_{0,1}\big(\widehat{v}_{k+1,j^+}(\widehat{v}_i^{N_k^V})\big)-\Phi_{0,1}\big(\widehat{v}_{k+1,j^-}(\widehat{v}_i^{N_k^V})\big)\Bigg)\\ &-\frac{1}{4\sigma_k(\widehat{v}_i^{N_k^V})}\phi_{0,1}\big(\widehat{v}_{k+1,j^+}(\widehat{v}_i^{N_k^V})\big)\big(\widehat{v}_{j+1}^{N_{k+1}^V}-\widehat{v}_j^{N_{k+1}^V}\big)\\ &-\frac{1}{4\sigma_k(\widehat{v}_i^{N_k^V})}\phi_{0,1}\big(\widehat{v}_{k+1,j^-}(\widehat{v}_i^{N_k^V})\big)\big(\widehat{v}_j^{N_{k+1}^V}-\widehat{v}_{j-1}^{N_{k+1}^V}\big)\Bigg\}\mathbb{P}(\widetilde{V}_{t_k}\in C_i(\widehat{\mathbf{v}}^{N_k^V})).\end{aligned} \tag{3.21}$$

The sub-diagonal elements are:

$$\frac{\partial^2 \widehat{D}_{k+1}}{\partial \widehat{v}_j^{N_{k+1}^V}\partial \widehat{v}_{j-1}^{N_{k+1}^V}}(\widehat{\mathbf{v}}^{N_{k+1}^V}) = -\frac{1}{2}\sum_{i=1}^{N_k^V}\Bigg\{\phi_{0,1}\big(\widehat{v}_{k+1,j^-}(\widehat{v}_i^{N_k^V})\big)\Bigg(\frac{\widehat{v}_j^{N_{k+1}^V}-\widehat{v}_{j-1}^{N_{k+1}^V}}{\sigma_k(\widehat{v}_i^{N_k^V})}\Bigg)\Bigg\}\mathbb{P}(\widetilde{V}_{t_k}\in C_i(\widehat{\mathbf{v}}^{N_k^V})), \tag{3.22}$$

while the super-diagonal terms are given by

$$\frac{\partial^2 \widehat{D}_{k+1}}{\partial \widehat{v}_j^{N_{k+1}^V}\partial \widehat{v}_{j+1}^{N_{k+1}^V}}(\widehat{\mathbf{v}}^{N_{k+1}^V}) = -\frac{1}{2}\sum_{i=1}^{N_k^V}\Bigg\{\phi_{0,1}\big(\widehat{v}_{k+1,j^+}(\widehat{v}_i^{N_k^V})\big)\Bigg(\frac{\widehat{v}_{j+1}^{N_{k+1}^V}-\widehat{v}_j^{N_{k+1}^V}}{\sigma_k(\widehat{v}_i^{N_k^V})}\Bigg)\Bigg\}\mathbb{P}(\widetilde{V}_{t_k}\in C_i(\widehat{\mathbf{v}}^{N_k^V})). \tag{3.23}$$

3.6 Proof of Lemma 3.2.3

Proof. First of all, by using Bayes formula, Fubini Theorem and Lemma 3.2.2 we have:

$$\begin{aligned}\mathbb{P}(\widetilde{S}_{t_{k+1}}\le x) &= \int_{-\infty}^{x}\mathbb{P}(\widetilde{S}_{t_{k+1}}\in ds_{k+1})\\ &= \int_{-\infty}^{x}\int_{\mathbb{R}^3}\mathbb{P}(\widetilde{S}_{t_{k+1}}\in ds_{k+1}|(\widetilde{S}_{t_k}=s_k,\widetilde{V}_{t_{k+1}}=v_{k+1},\widetilde{V}_{t_k}=v_k))\mathbb{P}(\widetilde{S}_{t_k}\in ds_k,\widetilde{V}_{t_{k+1}}\in dv_{k+1},\widetilde{V}_{t_k}\in dv_k)\\ &= \int_{\mathbb{R}^3}\mathbb{P}(\widetilde{S}_{t_k}\in ds_k,\widetilde{V}_{t_{k+1}}\in dv_{k+1},\widetilde{V}_{t_k}\in dv_k)\int_{-\infty}^{x}\phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(s_{k+1})ds_{k+1}\\ &= \int_{\mathbb{R}^3}\mathbb{P}(\widetilde{S}_{t_k}\in ds_k,\widetilde{V}_{t_{k+1}}\in dv_{k+1},\widetilde{V}_{t_k}\in dv_k)\,\Phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(x),\end{aligned} \tag{3.24}$$

where as usual the Gaussian density ϕ and cumulative Gaussian distribution function Φ are defined resp. as follows:

$$\phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(x) = \frac{1}{\bar{\sigma}_k(s_k,v_k)\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\bar{m}_k(s_k,v_{k+1},v_k)}{\bar{\sigma}_k(s_k,v_k)}\right)^2\right],$$

$$\Phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(x) = \int_{-\infty}^{x}\phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(s_{k+1})ds_{k+1}.$$

We now focus on $\mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k)$ that can be rewritten as follows:

$$\mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k) = \int_{\mathbb{R}^2} \underbrace{\mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k, \widetilde{V}_{t_{k-1}} = v_{k-1})}_{(a)}$$
$$\cdot \underbrace{\mathbb{P}(\widetilde{S}_{t_{k-1}} \in ds_{k-1}, \widetilde{V}_{t_k} \in dv_k, \widetilde{V}_{t_{k-1}} \in dv_{k-1})}_{(b)}, \tag{3.25}$$

The term (b) in (3.25) can be easily computed via the recursive algorithm, so let us focus on the term (a).

As it can be seen from equations (3.7) and (3.4), conditionally on $\{\widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k, \widetilde{V}_{t_{k-1}} = v_{k-1}\}$, S_{t_k} and $V_{t_{k+1}}$ are independent (non centered) Gaussian, since B and $B^\perp$ are. Then we have that

$$\begin{aligned}(a) &= \mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k, \widetilde{V}_{t_{k-1}} = v_{k-1}) \\ &= \mathbb{P}(\widetilde{S}_{t_k} \in ds_k | \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k, \widetilde{V}_{t_{k-1}} = v_{k-1}) \mathbb{P}(\widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k, \widetilde{V}_{t_{k-1}} = v_{k-1}) \\ &= \mathbb{P}(\widetilde{S}_{t_k} \in ds_k | \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k, \widetilde{V}_{t_{k-1}} = v_{k-1}) \underbrace{\mathbb{P}(\widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k)}_{(c)},\end{aligned}$$

where the last equality follows from the Markov property of $\widetilde{V}$. We now focus on (c). Notice that $\widetilde{V}_{t_{k+1}}$ given $\widetilde{V}_{t_k}$ is an affine transformation of the standard Gaussian random variable Z_k (recall equation ((3.4))), hence it is independent of all the random variables Z_ℓ and $Z_\ell^\perp$, for $\ell = 0, \dots, k-1$. In particular, this means that, given $\widetilde{V}_{t_k}$, $\widetilde{V}_{t_{k+1}}$ is independent of $\widetilde{S}_{t_{k-1}}$. Using this independence and the Bayes rule we have:

$$\begin{aligned}\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k\right) &= \frac{\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k\right)}{\mathbb{P}\left(\widetilde{S}_{k-1} = s_{k-1}, \widetilde{V}_{t_k} = v_k\right)} \\ &= \frac{\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{S}_{t_{k-1}} = s_{k-1} | \widetilde{V}_{t_k} = v_k\right) \mathbb{P}\left(\widetilde{V}_{t_k} = v_k\right)}{\mathbb{P}\left(\widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k\right)} \\ &= \frac{\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{V}_{t_k} = v_k\right) \mathbb{P}\left(\widetilde{S}_{t_{k-1}} = s_{k-1} | \widetilde{V}_{t_k} = v_k\right) \mathbb{P}\left(\widetilde{V}_{t_k} = v_k\right)}{\mathbb{P}\left(\widetilde{S}_{t_{k-1}} = s_{k-1}, \widetilde{V}_{t_k} = v_k\right)} \\ &= \mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{V}_{t_k} = v_k\right)\end{aligned}$$

So, from Equation (2.2) and Lemma 3.2.2 it turns out that

$$(a) = \phi_{\bar{m}_k(s_{k-1}, v_k, v_{k-1}), \bar{\sigma}_k(s_{k-1}, v_{k-1})}(s_k) ds_k \cdot \phi_{m_k(v_k), \sigma_k(v_k)}(v_{k+1}) dv_{k+1}. \tag{3.26}$$

Therefore we have an iterative method to compute the distribution of the triplet:

$$\mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k) =$$
$$\int_{\mathbb{R}^2} \left(\phi_{\bar{m}_k(s_{k-1}, v_k, v_{k-1}), \bar{\sigma}_k(s_{k-1}, v_{k-1})}(s_k) ds_k\right) \left(\phi_{m_k(v_k), \sigma_k(v_k)}(v_{k+1}) dv_{k+1}\right) \mathbb{P}(\widetilde{S}_{t_{k-1}} \in ds_{k-1}, \widetilde{V}_{t_k} \in dv_k, \widetilde{V}_{t_{k-1}} \in dv_{k-1}). \tag{3.27}$$

Now, from (3.24) we have

$$\mathbb{P}(\widetilde{S}_{t_{k+1}} \leq x) = \int_{\mathbb{R}^3} \Phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(x)\mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k),$$

so that, using (3.27) and differentiating, we get

$$\mathbb{P}(\widetilde{S}_{t_{k+1}} \in ds_{k+1}) = \int_{\mathbb{R}^3} \mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k)\phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(s_{k+1})ds_{k+1} \tag{3.28}$$

and the proof is completed. □

3.7 Gradient and Hessian for the Quantization of $\widetilde{S}_{t_{k+1}}$

The gradient and the Hessian of the distortion function $\widehat{D}_{k+1}$ in Equation (3.13) are

$$\frac{\partial D_{k+1}}{\partial x_j}(\mathbf{x}) = \frac{\partial}{\partial x_j}\left(\sum_{i=1}^{N_{k+1}} \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} (x_j - y)^2 \mathbb{P}(\widetilde{S}_{t_{k+1}} \in ds_{k+1})\right)$$

$$= \frac{\partial}{\partial x_j}\left(\sum_{i=1}^{N_{k+1}} \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} (x_j - s_{k+1})^2 \int_{\mathbb{R}^3} \phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(s_{k+1})ds_{k+1}\mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k)\right)$$

$$= \int_{\mathbb{R}^3} \Bigg[\underbrace{\frac{\partial}{\partial x_j}\left(\sum_{i=1}^{N_{k+1}} \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} (x_j - s_{k+1})^2 \phi_{\bar{m}_k(s_k,v_{k+1},v_k),\bar{\sigma}_k(s_k,v_k)}(s_{k+1})ds_{k+1}\right)}_{(\star\star)} \mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k)\Bigg].$$

It can be easily deduced that the term $(\star\star)$ is given by

$$\begin{aligned}(\star\star) = \; & 2\Big\{(x_j - \bar{m}_k(s_k, v_{k+1}, v_k))\Big[\Phi_{0,1}\big(x_{k+1,j^+}(s_k, v_{k+1}, v_k)\big) - \Phi_{0,1}\big(x_{k+1,j^-}(s_k, v_{k+1}, v_k)\big)\Big] \\ & + \bar{\sigma}_k(s_k, v_{k+1}, v_k)\Big[\phi_{0,1}\big(x_{k+1,j^+}(s_k, v_{k+1}, v_k)\big) - \phi_{0,1}\big(x_{k+1,j^-}(s_k, v_{k+1}, v_k)\big)\Big]\Big\},\end{aligned}$$

where

$$\begin{aligned}\bar{x}_{k+1,j^+}(s_k, v_{k+1}, v_k) &= \frac{x_j + x_{j+1} - 2\bar{m}_k(s_k, v_{k+1}, v_k)}{2\bar{\sigma}_k(s_k, v_k)}, \\ \bar{x}_{k+1,j^-}(s_k, v_{k+1}, v_k) &= \frac{x_{j-1} + x_j - 2\bar{m}_k(s_k, v_{k+1}, v_k)}{2\bar{\sigma}_k(s_k, v_k)},\end{aligned} \tag{3.29}$$

so we get that

$$\begin{aligned}\frac{\partial D_{k+1}}{\partial x_j}(\mathbf{x}) = 2\int_{\mathbb{R}^3} & \Big\{(x_j - \bar{m}_k(s_k, v_{k+1}, v_k))\Big[\Phi_{0,1}\big(\bar{x}_{k+1,j^+}(s_k, v_{k+1}, v_k)\big) - \Phi_{0,1}\big(\bar{x}_{k+1,j^-}(s_k, v_{k+1}, v_k)\big)\Big] \\ & + \bar{\sigma}_k(s_k, v_k)\Big[\phi_{0,1}\big(\bar{x}_{k+1,j^+}(s_k, v_{k+1}, v_k)\big) - \phi_{0,1}\big(\bar{x}_{k+1,j^-}(s_k, v_{k+1}, v_k)\big)\Big]\Big\} \\ & \cdot \mathbb{P}(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_{k+1}} \in dv_{k+1}, \widetilde{V}_{t_k} \in dv_k).\end{aligned} \tag{3.30}$$

Similarly we can write the components of the (tri-diagonal and symmetric) Hessian matrix:

$$\begin{aligned}\frac{\partial^2 D_{k+1}}{\partial^2 x_j}(\mathbf{x}) = 2\int_{\mathbb{R}^3}\Bigg\{\Bigg(\Phi_{0,1}\big(\bar{x}_{k+1,j^+}(s_k,v_{k+1},v_k)\big)-\Phi_{0,1}\big(\bar{x}_{k+1,j^-}(s_k,v_{k+1},v_k)\big)\Bigg)\\ -\frac{1}{4\bar{\sigma}_k(s_k,v_k)}\phi_{0,1}\big(\bar{x}_{k+1,j^+}(s_k,v_{k+1},v_k)\big)\big(x_{j+1}-x_j\big)\\ -\frac{1}{4\bar{\sigma}_k(s_k,v_k)}\phi_{0,1}\big(\bar{x}_{k+1,j^-}(s_k,v_{k+1},v_k)\big)\big(x_j-x_{j-1}\big)\Bigg\}\\ \cdot\,\mathbb{P}(\widetilde{S}_{t_k}\in ds_k,\widetilde{V}_{t_{k+1}}\in dv_{k+1},\widetilde{V}_{t_k}\in dv_k);\end{aligned}$$

$$\frac{\partial^2 D_{k+1}}{\partial x_j\partial x_{j-1}}(\mathbf{x}) = -\frac{1}{2}\int_{\mathbb{R}^3}\left\{\phi_{0,1}\big(\bar{x}_{k+1,j^-}(s_k,v_{k+1},v_k)\big)\left(\frac{x_j-x_{j-1}}{\bar{\sigma}_k(s_k,v_k)}\right)\right\}\mathbb{P}(\widetilde{S}_{t_k}\in ds_k,\widetilde{V}_{t_{k+1}}\in dv_{k+1},\widetilde{V}_{t_k}\in dv_k);$$

$$\frac{\partial^2 D_{k+1}}{\partial x_j\partial x_{j+1}}(\mathbf{x}) = -\frac{1}{2}\int_{\mathbb{R}^3}\left\{\phi_{0,1}\big(\bar{x}_{k+1,j^+}(s_k,v_{k+1},v_k)\big)\left(\frac{x_{j+1}-x_j}{\bar{\sigma}_k(s_k,v_k)}\right)\right\}\mathbb{P}(\widetilde{S}_{t_k}\in ds_k,\widetilde{V}_{t_{k+1}}\in dv_{k+1},\widetilde{V}_{t_k}\in dv_k).$$

3.8 Towards the Newton Algorithm to Quantize $\widetilde{S}_{t_{k+1}}$

In this Appendix, we compute the gradient and the Hessian matrix for the discrete version of the distortion function, to be used in the Newton-Raphson procedure to find a stationary quantizer for $\widetilde{S}_{t_{k+1}}$. We denote by $\widehat{\mathbf{s}}^{N_{k+1}^S}=(\widehat{s}_1^{N_{k+1}^S},\ldots,\widehat{s}_{N_{k+1}^S}^{N_{k+1}^S})$ and by $\widehat{\mathbf{s}}^{N_k^S}=(\widehat{s}_1^{N_{k+1}^S},\ldots,\widehat{s}_{N_k^S}^{N_k^S})$ resp. the N_{k+1}^S and N_k^S-dimensional quantizers for $\widetilde{S}_{t_{k+1}}$ and $\widetilde{S}_{t_k}$ and by $\widehat{\mathbf{v}}^{N_{k+1}^V}=(\widehat{v}_1^{N_{k+1}^V},\ldots,\widehat{v}_{N_{k+1}^V}^{N_{k+1}^V})$ and $\widehat{\mathbf{v}}^{N_k}=(\widehat{v}_1^{N_{k+1}^V},\ldots,\widehat{v}_{N_k}^{N_k})$ the N_{k+1}^V and N_k^V-dimensional quantizers for for $\widetilde{V}_{t_{k+1}}$ and $\widetilde{V}_{t_k}$, respectively. Taking the gradient of the discrete distortion function $\widehat{D}_{k+1}(\widehat{\mathbf{s}}^{N_{k+1}^S})$ gives

$$\begin{aligned}\frac{\partial\widehat{D}_{k+1}}{\partial\widehat{s}_j^{N_{k+1}^S}}(\widehat{\mathbf{s}}^{N_{k+1}^S}) = 2\sum_{a=1}^{N_k^S}\sum_{b=1}^{N_{k+1}^V}\sum_{c=1}^{N_k^V}\Bigg\{\big(\widehat{s}_j^{N_{k+1}^V}-\bar{m}_k(\widehat{s}_a^{N_k^S},\widehat{v}_b^{N_{k+1}^V},\widehat{v}_c^{N_k^V})\big)\Big[\Phi_{0,1}\big(\widehat{s}_{k+1,j^+}(\widehat{s}_a^{N_k^S},\widehat{v}_b^{N_{k+1}^V},\widehat{v}_c^{N_k^V})\big)\\ -\Phi_{0,1}\big(\widehat{s}_{k+1,j^-}(\widehat{s}_a^{N_k^S},\widehat{v}_b^{N_{k+1}^V},\widehat{v}_c^{N_k^V})\big)\Big]+\bar{\sigma}_k(\widehat{s}_a^{N_k^S},\widehat{v}_c^{N_k^V})\\ \cdot\Big[\phi_{0,1}\big(\widehat{s}_{k+1,j^+}(\widehat{s}_a^{N_k^S},\widehat{v}_b^{N_{k+1}^V},\widehat{v}_c^{N_k^V})\big)-\phi_{0,1}\big(\widehat{s}_{k+1,j^-}(\widehat{s}_a^{N_k^S},\widehat{v}_b^{N_{k+1}^V},\widehat{v}_c^{N_k^V})\big)\Big]\Bigg\}\\ \cdot\,\mathbb{P}(\widetilde{S}_{t_k}\in C_a(\widehat{\mathbf{s}}^{N_k^S}),\widetilde{V}_{t_{k+1}}\in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}),\widetilde{V}_{t_k}\in C_c(\widehat{\mathbf{v}}^{N_k^V})),\end{aligned}$$

where $\widehat{s}_{k+1,j^+}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})$ and $\widehat{s}_{k+1,j^-}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})$ have been defined in (3.16). The components of the tridiagonal discrete distortion Hessian are

$$\frac{\partial^2 \widehat{D}_{k+1}}{\partial^2 \widehat{s}_j^{N_{k+1}^S}}(\widehat{\mathbf{s}}^{N_{k+1}^S}) = 2\sum_{a=1}^{N_k^S}\sum_{b=1}^{N_{k+1}^V}\sum_{c=1}^{N_k^V}\Bigg\{\Big(\Phi_{0,1}\big(\widehat{s}_{k+1,j^+}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})\big) - \Phi_{0,1}\big(\widehat{s}_{k+1,j^-}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})\big)\Big)$$
$$-\frac{1}{4\bar{\sigma}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_c^{N_k^V})}\phi_{0,1}\big(\widehat{s}_{k+1,j^+}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})\big)\big(s_{j+1}^{N_{k+1}^S} - s_j^{N_{k+1}^S}\big)$$
$$-\frac{1}{4\bar{\sigma}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_c^{N_k^V})}\phi_{0,1}\big(\widehat{s}_{k+1,j^-}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})\big)\big(s_j^{N_{k+1}^S} - s_{j-1}^{N_{k+1}^S}\big)\Bigg\}$$
$$\cdot\,\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}));$$

$$\frac{\partial^2 \widehat{D}_{k+1}}{\partial \widehat{s}_j^{N_{k+1}^S}\partial \widehat{s}_{j-1}^{N_{k+1}^S}}(\widehat{\mathbf{s}}^{N_{k+1}^S}) = -\frac{1}{2}\sum_{a=1}^{N_k^S}\sum_{b=1}^{N_{k+1}^V}\sum_{c=1}^{N_k^V}\Bigg\{\phi_{0,1}\big(\widehat{s}_{k+1,j^-}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})\big)\left(\frac{\widehat{s}_j^{N_{k+1}^V} - \widehat{s}_{j-1}^{N_{k+1}^V}}{\bar{\sigma}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_c^{N_k^V})}\right)\Bigg\}$$
$$\cdot\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}));$$

$$\frac{\partial^2 \widehat{D}_{k+1}}{\partial \widehat{s}_j^{N_{k+1}^S}\partial \widehat{s}_{j+1}^{N_{k+1}^S}}(\widehat{\mathbf{s}}^{N_{k+1}^S}) = -\frac{1}{2}\sum_{a=1}^{N_k^S}\sum_{b=1}^{N_{k+1}^V}\sum_{c=1}^{N_k^V}\Bigg\{\phi_{0,1}\big(\widehat{s}_{k+1,j^+}(\widehat{s}_a^{N_k^S}, \widehat{v}_b^{N_{k+1}^V}, \widehat{v}_c^{N_k^V})\big)\left(\frac{\widehat{s}_{j+1}^{N_{k+1}^V} - \widehat{s}_j^{N_{k+1}^V}}{\bar{\sigma}_k(\widehat{s}_a^{N_k^S}, \widehat{v}_c^{N_k^V})}\right)\Bigg\}$$
$$\cdot\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V})).$$

3.9 Proof of Proposition 3.2.4

Proof. We can decompose the initial joint distribution in this way:

$$\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V})) =$$
$$\sum_{d=1}^{N_{k-1}^S}\sum_{e=1}^{N_{k-1}^V}\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})),$$

where we can rewrite the generic element of the summation as a function of the extremes of the Voronoi regions $C_a(\widehat{\mathbf{s}}^{N_k^S})$ and $C_b(\widehat{\mathbf{v}}^{N_{k+1}^V})$ as follows:

$$\begin{aligned}
&\mathbb{P}(\widetilde{S}_{t_k} \in C_a(\widehat{\mathbf{s}}^{N_k^S}), \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})) \\
&= \mathbb{P}(\widetilde{S}_{t_k} \leq s_{a+1/2}^{N_k^S}, \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})) \\
&-\mathbb{P}(\widetilde{S}_{t_k} \leq s_{a-1/2}^{N_k^S}, \widetilde{V}_{t_{k+1}} \in C_b(\widehat{\mathbf{v}}^{N_{k+1}^V}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})) \\
&= \mathbb{P}(\widetilde{S}_{t_k} \leq s_{a+1/2}^{N_k^S}, \widetilde{V}_{t_{k+1}} \leq v_{b+1/2}^{N_{k+1}^V}, \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})) \\
&-\mathbb{P}(\widetilde{S}_{t_k} \leq s_{a+1/2}^{N_k^S}, \widetilde{V}_{t_{k+1}} \leq v_{b-1/2}^{N_{k+1}^V}, \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})) \\
&-\mathbb{P}(\widetilde{S}_{t_k} \leq s_{a-1/2}^{N_k^S}, \widetilde{V}_{t_{k+1}} \leq v_{b+1/2}^{N_{k+1}^V}, \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})) \\
&+\mathbb{P}(\widetilde{S}_{t_k} \leq s_{a-1/2}^{N_k^S}, \widetilde{V}_{t_{k+1}} \leq v_{b-1/2}^{N_{k+1}^V}, \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})),
\end{aligned}$$

with

$$\begin{aligned}
s_{a+1/2}^{N_k^S} &= \frac{s_a^{N_k^S} + s_{a+1}^{N_k^S}}{2}; \qquad s_{a-1/2}^{N_k^S} = \frac{s_{a-1}^{N_k^S} + s_a^{N_k^S}}{2}; \\
v_{b+1/2}^{N_{k+1}^V} &= \frac{v_b^{N_{k+1}} + v_{b+1}^{N_{k+1}}}{2}; \qquad v_{b-1/2}^{N_{k+1}^V} = \frac{v_{b-1}^{N_{k+1}} + v_b^{N_{k+1}}}{2}.
\end{aligned}$$

We now approximate the following term

$$\mathbb{P}(\widetilde{S}_{t_k} \leq s_k, \widetilde{V}_{t_{k+1}} \leq v_{k+1}, \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})),$$

for general $s_k, v_{k+1} \in \mathbb{R}$.
Remember that, from (3.27),

$$\begin{aligned}
&\mathbb{P}(\widetilde{S}_{t_k} \leq s_k, \widetilde{V}_{t_{k+1}} \leq v_{k+1}, \widetilde{V}_{t_k} \leq v_k) = \\
&\int_{-\infty}^{s_k} \int_{-\infty}^{v_{k+1}} \int_{-\infty}^{v_k} \int_{\mathbb{R}^2} \phi_{\bar{m}_k(s_{k-1},v_k,v_{k-1}),\bar{\sigma}_k(s_{k-1},v_{k-1})}(s_k) ds_k \phi_{m_k(v_k),\sigma_k(v_k)}(v_{k+1}) dv_{k+1} \\
&\cdot \mathbb{P}(\widetilde{S}_{t_{k-1}} \in ds_{k-1}, \widetilde{V}_{t_k} \in dv_k, \widetilde{V}_{t_{k-1}} \in dv_{k-1}),
\end{aligned}$$

so that we have

$$\begin{aligned}
&\mathbb{P}(\widetilde{S}_{t_k} \leq s_k, \widetilde{V}_{t_{k+1}} \leq v_{k+1}, \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})) \\
&= \int_{-\infty}^{s_k} \int_{-\infty}^{v_{k+1}} \int_{C_c(\widehat{\mathbf{v}}^{N_k})} \int_{C_d(\widehat{\mathbf{s}}^{N_{k-1}^S})} \int_{C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})} \phi_{\bar{m}_k(s_{k-1},v_k,v_{k-1}),\bar{\sigma}_k(s_{k-1},v_{k-1})}(s_k) ds_k \phi_{m_k(v_k),\sigma_k(v_k)}(v_{k+1}) dv_{k+1} \\
&\qquad \cdot \mathbb{P}(\widetilde{S}_{t_{k-1}} \in ds_{k-1}, \widetilde{V}_{t_k} \in dv_k, \widetilde{V}_{t_{k-1}} \in dv_{k-1})
\end{aligned}$$

$$= \int_{C_c(\widehat{\mathbf{v}}^{N_k})} \int_{C_d(\widehat{\mathbf{s}}^{N_{k-1}^S})} \int_{C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})} \left(\int_{-\infty}^{s_k} \phi_{\bar{m}_k(s_{k-1},v_k,v_{k-1}),\bar{\sigma}_k(s_{k-1},v_{k-1})}(s_k)ds_k \right) \left(\int_{-\infty}^{v_{k+1}} \phi_{m_k(v_k),\sigma_k(v_k)}(v_{k+1})dv_{k+1} \right)$$
$$\cdot \mathbb{P}(\widetilde{S}_{t_{k-1}} \in ds_{k-1}, \widetilde{V}_{t_k} \in dv_k, \widetilde{V}_{t_{k-1}} \in dv_{k-1})$$

$$= \int_{C_c(\widehat{\mathbf{v}}^{N_k})} \int_{C_d(\widehat{\mathbf{s}}^{N_{k-1}^S})} \int_{C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})} \Phi_{\bar{m}_k(s_{k-1},v_k,v_{k-1}),\bar{\sigma}_k(s_{k-1},v_{k-1})}(s_k)\Phi_{m_k(v_k),\sigma_k(v_k)}(v_{k+1})$$
$$\cdot \mathbb{P}(\widetilde{S}_{t_{k-1}} \in ds_{k-1}, \widetilde{V}_{t_k} \in dv_k, \widetilde{V}_{t_{k-1}} \in dv_{k-1}).$$

Finally,

$$\mathbb{P}(\widetilde{S}_{t_k} \leq s_k, \widetilde{V}_{t_{k+1}} \leq v_{k+1}, \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V}))$$

$$\approx \Phi_{\bar{m}_k(s_d^{N_{k-1}^S}, v_c^{N_k^V}, v_e^{N_{k-1}^V}), \bar{\sigma}_k(s_d^{N_{k-1}^S}, v_e^{N_{k-1}^V})}(s_k) \Phi_{m_k(v_c^{N_k^V}), \sigma_k(v_c^{N_k^V})}(v_{k+1})$$
$$\cdot \mathbb{P}(\widetilde{S}_{t_{k-1}} \in C_d(\widehat{\mathbf{s}}^{N_{k-1}^S}), \widetilde{V}_{t_k} \in C_c(\widehat{\mathbf{v}}^{N_k^V}), \widetilde{V}_{t_{k-1}} \in C_e(\widehat{\mathbf{v}}^{N_{k-1}^V})).$$

Now we normalize the c.d.f.s of the Gaussian distributions, we sum up all the terms and we get the result.
Let us now spend a couple of words about the first step of the algorithm, for $k = 0$, concerning the computation of the probability $\mathbb{P}(\widetilde{S}_{t_0} \in C_a(\widehat{\mathbf{s}}^{N_0^S}), \widetilde{V}_{t_1} \in C_b(\widehat{\mathbf{v}}^{N_1}), \widetilde{V}_{t_0} \in C_c(\widehat{\mathbf{v}}^{N_0}))$.
Given that $\widetilde{S}_{t_0} = S_{t_0} = s_0$ and that $\widetilde{V}_{t_0} = V_{t_0} = v_0$, it follows that

$$\mathbb{P}(\widetilde{S}_{t_0} \in C_a(\widehat{\mathbf{s}}^{N_0^S}), \widetilde{V}_{t_1} \in C_b(\widehat{\mathbf{v}}^{N_1}), \widetilde{V}_{t_0} \in C_c(\widehat{\mathbf{v}}^{N_0})) = \mathbb{P}(\widetilde{S}_{t_0} \in C_a(\widehat{\mathbf{s}}^{N_0^S}))\mathbb{P}(\widetilde{V}_{t_1} \in C_b(\widehat{\mathbf{v}}^{N_1}), \widetilde{V}_{t_0} \in C_c(\widehat{\mathbf{v}}^{N_0})).$$

Using the conditional Gaussian distribution of the variance process in Equation (2.2), we have

$$\mathbb{P}(\widetilde{V}_{t_1} \in C_b(\widehat{\mathbf{v}}^{N_1}), \widetilde{V}_{t_0} \in C_c(\widehat{\mathbf{v}}^{N_0})) = \left[\Phi_{0,1}\left(\widehat{v}_{1,b^+}(\widehat{v}_c^{N_0}) \right) - \Phi_{0,1}\left(\widehat{v}_{1,b^-}(\widehat{v}_c^{N_0}) \right) \right] \mathbb{P}(\widetilde{V}_{t_0} \in C_c(\widehat{\mathbf{v}}^{N_0})),$$

where $\widehat{v}_{1,b^\pm}(\widehat{v}_c^{N_0})$ are defined in (3.12).

□

3.10 Proof of Proposition 3.2.7

Proof. We have

$$||X_{t_k} - \widehat{X}_{t_k}||_2 \leq ||X_{t_k} - \widetilde{X}_{t_k}||_2 + ||\widetilde{X}_{t_k} - \widehat{X}_{t_k}||_2$$

where we recall that $\widetilde{X}_{t_k}$ denotes the Euler scheme at time t_k relative to the pair (S, V) in equations (3.7) and (3.4). Thus, the recursive marginal quantization error is two-sided, being the sum of the error due to the Euler scheme and the one due to the vector quantization of $\widetilde{X}_{t_k}$ via $\widehat{X}_{t_k}$. For what concerns the former, under our hypotheses we have

$$\mathbb{E}\left(\sup_{k \in \{0,\dots,M\}} |X_{t_k} - \widetilde{X}_{t_k}|^2 \right) \leq \frac{C_T}{M},$$

where C_T only depends on T. Hence, for every $k \in \{1, \dots, M\}$:

$$||X_{t_k} - \widetilde{X}_{t_k}||_2 \leq \sqrt{\frac{C_T}{M}}.$$

On the other hand, the error coming from the quantization of $\widetilde{X}_{t_k}$, can be controlled thanks to Theorem 3.1 in Pagès and Sagna (2015). Indeed, applying their result in our two-dimensional setting we find:

$$||\widetilde{X}_{t_k} - \widehat{X}_{t_k}||_2 \leq K_{2,2,\eta} \sum_{\ell=1}^{k} a_\ell(b, \Sigma, t_k, \Delta, y_0, L, 2+\eta) \frac{1}{\sqrt{N_k^V}}$$

and the proof is complete. □

Chapter 4

Stochastic volatility models: an alternative approach

This is a joint work with Giorgia Callegaro and Martino Grasselli. The content of this chapter has been submitted as a paper, see Callegaro et al. (2017a) in the bibliography. This work has been awarded in the 8th General AMaMeF conference (Amsterdam, June 19-23, 2017) for "Outstanding poster presentation".

We apply the recursive marginal quantization methodology to the pricing of vanillas and American style options, extending the results of Callegaro et al. (2015) (Chapter 2) to stochastic volatility models. The methodology is fast and accurate to the point that calibration can be efficiently performed. As a motivating example, we calibrate the Heston model on a book on Google stock that includes American style options.

4.1 Introduction

This chapter constitutes the sequel of Callegaro et al. (2015) (Chapter 2). In that chapter we calibrated a local volatility model using a quantization approach that had been introduced by Pagès and Sagna (2015). This technique has been further developed to be applied to pricing in financial markets in Callegaro et al. (2016) (Chapter 3), Fiorin et al. (2015) (Chapter 5) and McWalter et al. (2017). Here we extend the methodology to stochastic volatility models, in order to price vanillas and American style options. The challenge in our framework is represented by the presence of the volatility process, which increases the dimension of the problem. Starting from the Euler scheme of the stochastic volatility model, we propose a fast and accurate discretization of the asset price that improves the ones proposed in Callegaro et al. (2016) (Chapter 3) and Fiorin et al. (2015) (Chapter 5), see Remark 4.2.1 below. We first focus on the pricing of vanillas and we show that pricing can be efficiently performed. Then, we exploit the idea in Bally et al. (2005), where it is shown that pricing of American options can also be performed through a backward procedure, like in a multinomial lattice. As a result, we provide the first calibration example of the Heston (1993) model using a book of real data that includes American style options. Of course, for this affine model one could also calibrate on vanillas using the Fourier methodology of Carr and Madan (1999). What is more, we emphasize that our methodology is very flexible insofar it applies to any stochastic volatility model (well beyond the family of affine models), including the SABR model and many others.
The chapter is organized as follows: Section 4.2 extends the marginal quantization method to the class of stochastic volatility models and illustrates the idea of the algorithm. Further details can be found in Callegaro et al. (2016) (Chapter 3) with a

slightly different method. Section 4.3 illustrates our numerical results on the Heston (1993) model, with particular emphasis on the calibration exercise on real data including American style options. Section 4.4 concludes.

4.2 Recursive quantization of stochastic volatility models

Let us consider a continuous time Markov process $Y = (S, V)$ corresponding to the pair (price, volatility), whose evolution is specified by

$$\begin{pmatrix} dS_t \\ dV_t \end{pmatrix} = \begin{pmatrix} \mu_S(t, S_t, V_t) \\ \mu_V(t, V_t) \end{pmatrix} dt + \begin{pmatrix} \Sigma_{SS}(t, S_t, V_t) & \Sigma_{SV}(t, S_t, V_t) \\ \Sigma_{VS}(t, V_t) & \Sigma_{VV}(t, V_t) \end{pmatrix} dW_t, \quad \begin{pmatrix} S_0 = s_0 \\ V_0 = v_0 \end{pmatrix}, \tag{4.1}$$

where W is a 2 dimensional Brownian motion and $\mu_S, \mu_V, \Sigma_{SS}, \Sigma_{SV}, \Sigma_{VS}, \Sigma_{VV}$ are such that the SDE (4.1) is well defined. In this section we only introduce our novel methodology, so that we do not focus on the technical conditions which would ensure existence of a strong solution to the very general SDE (4.1). Nevertheless, in the following Section 6.3 we will provide more details on existence of a strong solution in the case of the model in use, ie, in Heston model.
Fix a time horizon T and a time grid with constant step size Δ and length M, ie $t_k = \frac{k}{M}T$ for $k = 0, \dots, M$, so that the Euler scheme for the pair (S, V) reads

$$\begin{pmatrix} \widetilde{S}_{k+1} \\ \widetilde{V}_{k+1} \end{pmatrix} = \begin{pmatrix} \widetilde{S}_k \\ \widetilde{V}_k \end{pmatrix} + \begin{pmatrix} \mu_S(t_k, \widetilde{S}_k, \widetilde{V}_k) \\ \mu_V(t_k, \widetilde{V}_k) \end{pmatrix} \Delta + \begin{pmatrix} \Sigma_{SS}(t_k, \widetilde{S}_k, \widetilde{V}_k) & \Sigma_{SV}(t_k, \widetilde{S}_k, \widetilde{V}_k) \\ \Sigma_{VS}(t_k, \widetilde{V}_k) & \Sigma_{VV}(t_k, \widetilde{V}_k) \end{pmatrix} \sqrt{\Delta}\, \widetilde{W}_k, \tag{4.2}$$

with initial condition

$$\begin{pmatrix} \widetilde{S}_0 = s_0 \\ \widetilde{V}_0 = v_0 \end{pmatrix},$$

where $\widetilde{W}_k := \frac{W_{t_{k+1}} - W_{t_k}}{\sqrt{\Delta}}$ is a bivariate standard Gaussian random variable having mean $(0, 0)$ and variance the identity matrix 2×2. Similarly to the one dimensional case (see Callegaro et al. (2015) (Chapter 2)), we have that

$$\left(\begin{pmatrix} \widetilde{S}_{k+1} \\ \widetilde{V}_{k+1} \end{pmatrix} \middle| \begin{pmatrix} \widetilde{S}_k \\ \widetilde{V}_k \end{pmatrix} = \begin{pmatrix} s \\ v \end{pmatrix} \right) \overset{\text{Law}}{=} \mathcal{N}\left(\mu(t_k, s, v), \Sigma(t_k, s, v)\right), \tag{4.3}$$

where

$$\mu(t_k, s, v) = \begin{pmatrix} s + \mu_S(t_k, s, v)\Delta \\ v + \mu_V(t_k, v)\Delta \end{pmatrix},$$

and

$$\Sigma(t_k,s,v)=\Delta\begin{pmatrix}\left(\Sigma_{SS}^2+\Sigma_{SV}^2\right)(t_k,s,v) & \left(\Sigma_{SS}\Sigma_{VS}+\Sigma_{SV}\Sigma_{VV}\right)(t_k,s,v)\\ \left(\Sigma_{SS}\Sigma_{VS}+\Sigma_{SV}\Sigma_{VV}\right)(t_k,s,v) & \left(\Sigma_{VV}^2+\Sigma_{VS}^2\right)(t_k,v)\end{pmatrix}.$$

As the dynamics of the volatility factor can be written independently of the price process, one can use Callegaro et al. (2015) (Chapter 2) in order to quantize the volatility. The difficult point is the construction of the distortion function relative to $\widetilde{S}_{k+1}$, that we denote by D_{k+1}, which depends also on the volatility process V. Let us fix a grid $\mathbf{x}^{k+1}=\{x_1^{k+1},\dots,x_N^{k+1}\}$. The distortion function for $\widetilde{S}_{k+1}$ reads

$$D_{k+1}\left(\mathbf{x}^{k+1}\right)=\sum_{i=1}^{N}\int_{C_i\left(\mathbf{x}^{k+1}\right)}\left(s_{k+1}-x_i^{k+1}\right)^2\mathbb{P}\left(\widetilde{S}_{k+1}\in ds_{k+1}\right). \tag{4.4}$$

Having quantized the pair $\left(\widetilde{S}_\ell,\widetilde{V}_\ell\right),\ell=0,\dots,k$, we can write the distribution of $\widetilde{S}_{k+1}$, using (4.3). This gives the recursive formula to compute the quantizers

$$\mathbb{P}\left(\widetilde{S}_{k+1}\in ds_{k+1}\right)=ds_{k+1}\int_{\mathbb{R}}\int_{\mathbb{R}}\phi_{\mu_S(t_k,s_k,v_k),\left(\Sigma_{SS}^2+\Sigma_{SV}^2\right)^{\frac{1}{2}}(t_k,s_k,v_k)}(s_{k+1})\,\mathbb{P}(\widetilde{S}_k\in ds_k,\widetilde{V}_k\in dv_k), \tag{4.5}$$

where ϕ is the density function of a scalar Normal distribution with mean $\mu_S(t_k,s_k,v_k)$ and variance $\left(\Sigma_{SS}^2+\Sigma_{SV}^2\right)(t_k,s_k,v_k)$. It is then possible to compute the gradient and the Hessian matrix of the distortion D_{k+1}, in order to find (numerically) its minima. Note that, once we have quantized $\left(\widetilde{S}_\ell,\widetilde{V}_\ell\right),\ell=0,\dots,k$, the integrals in (4.5) become a finite sum, thus leading to extremely fast computations. In the sequel, we will apply this methodology to the celebrated Heston (1993) stochastic volatility model.

Remark 4.2.1. Note that this result is different from the one in Callegaro et al. (2016) (Chapter 3), where we considered the Euler scheme separately for S and V. Indeed, using here the Euler scheme for the pair (Price, Volatility), it is possible to write the density of the price process more accurately, by taking into account properly the correlation between the price and the volatility. This allows us to price efficiently vanillas and American style options and then to calibrate the model.

Remark 4.2.2. Our technique can be applied to any stochastic volatility model for which an Euler discretization scheme is available, as also shown in Callegaro et al. (2016) (Chapter 3). Moreover, since recursive marginal quantization has been applied on top of a Milstein discretization scheme in McWalter et al. (2017) to discretize a geometric Brownian motion and a constant elasticity of variance process, we deem that an interesting research direction could be the extension of our methodology using a higher order discretization scheme.

4.3 Numerical results for the Heston model

4.3.1 Pricing of vanillas

The Heston (1993) model assumes the following risk-neutral dynamics for the pair (S, V)

$$\frac{dS_t}{S_t} = rdt + \sqrt{V_t}(\rho\, dW_t^1 + \sqrt{1-\rho^2} dW_t^2)$$
$$dV_t = \kappa(\theta - V_t)dt + \xi\sqrt{V_t} dW_t^1,$$

where W^1 and W^2 are two independent standard Brownian motions and where r is the risk free interest rate, θ is the long run average price variance, κ is the rate at which the variance V reverts to θ, ρ is the correlation between the asset and the instantaneous variance and ξ is the vol of vol parameter, which determines the volatility of the variance process. We assume that r, κ, θ and ξ are strictly positive. In this case, the above system of SDEs admits a strong solution (see eg (Andersen and Piterbarg, 2007a, Section 2)). The components of the Euler scheme here read

$$\mu(t_k, s, v) = \begin{pmatrix} s + rs\Delta \\ v + \kappa(\theta - v)\Delta \end{pmatrix}, \qquad \Sigma(t_k, s, v) = \Delta \begin{pmatrix} s^2 v & \rho\xi sv \\ \rho\xi sv & \xi^2 v \end{pmatrix}.$$

We first compare the pricing of European calls obtained with our quantization methods with the ones provided by the Fourier based methodology as in Carr and Madan (1999), that we take as our benchmark. In Table 4.1 we display the prices together with the errors (relative error and absolute difference of implied volatilities) for European call options with maturity $T = 1$ year. The strike K is in percentage of the initial price $S_0 = 100$. "Q" stands for quantization, where we considered $N = 20$ points for the underlying, $N = 10$ points for the volatility process and $M = 12$ time steps. This choice for N, M represents a good compromise between precision and computational time.
Results were obtained with *MATLAB*, on a laptop with a 2.7 GHz CPU and 8 Gb of memory. The computational time to get all the prices is less than two seconds. The results in Table 4.1 show the efficiency of our methodology, which is fast and accurate.

4.3.2 Pricing of American options

Many numerical methods computing prices of such options in the Heston model have been developed. They typically approximate the solution to the partial differential equation under early exercise constraints by ad hoc finite difference discretization grids (see eg Ikonen and Toivanen (2008)) or by exploiting recombining two dimensional lattices (see eg Beliaeva and Nawalkha (2010) or Vellekoop and Nieuwenhuis (2009)) or by means of Fourier based methods, such as the Wiener-Hopf factorization (see eg Boyarchenko and Levendorskiy (2013)) or the Cosine method (see eg Fang and Oosterlee (2011)).
An interesting feature of our quantization approach is that from the Euler-Maruyama discretization scheme one can directly deduce some information on the transition probability of the asset price, so that the pricing of American options can be performed through a backward procedure on a multinomial lattice, as illustrated in

	Bench. price	Q price	Rel error (%)	IV difference ($\times 10^{-3}$)
$K = 80$	24.91	24.99	0.32	3.87
$K = 85$	20.75	20.81	0.29	2.38
$K = 90$	16.75	16.79	0.24	1.33
$K = 95$	12.96	13.00	0.31	1.16
$K = 100$	9.43	9.45	0.21	0.92
$K = 105$	6.24	6.25	0.16	0.75
$K = 110$	3.50	3.52	0.57	0.72
$K = 115$	1.45	1.46	0.69	0.84
$K = 120$	0.40	0.40	0.75	0.89

TABLE 4.1: Results on the pricing of European call options via quantization and via the FFT approach of Carr and Madan (1999). The strike K is in percentage of the initial price $S_0 = 100$. The maturity is $T = 1$ year and $r = 0.04$. The parameters are those of the calibration exercise in Table 4.3. "Q" stands for quantization, where we considered grids of 20 points for the underlying, 10 points for the volatility process and 12 time steps. For the quantization, the computational times is less than 2 seconds for the whole book.

Bally et al. (2005). Let us assume that the set of possible exercise times of an American put option of maturity T is finite, meaning that we are approximating the American option value with the price of a Bermudan option, where the exercise times are $t_k = \frac{kT}{M}$, with $k = 1, \dots, M$. The parameters for the Heston model are the same as in the case of European options of the previous subsection. We show in Table 4.2 the comparison of the quantization price with a benchmark for the Heston model provided by the method of Vellekoop and Nieuwenhuis (2009) (for an implementation see for example PREMIA at `https://www.rocq.inria.fr/mathfi/Premia/`). For the benchmark price we consider 20 points for the discretization of V, 200 points for the discretization of S and 12 time steps.
The quantization approach leads to accurate prices and takes 1.9676 seconds to compute all the put prices. This opens the door to the possibility of calibrating a book including both European and American options, as we are going to do in the next subsection.

4.3.3 Calibration of American options to real data

In order to show the effectiveness of our pricing technique, we calibrate the Heston model to a book of American option prices on the GOOG Google stock as of date April, 26$^{\text{th}}$ 2017. The book includes 4 maturities (from 3 months to 14 months) with 12 strikes, ranging from 85% to 115% of the spot price, for a total of 96 options. Table

	Bench. price	Q price	Rel error (%)
$K = 80$	1.81	1.78	1.75
$K = 85$	2.48	2.50	0.69
$K = 90$	3.33	3.27	1.65
$K = 95$	4.42	4.47	1.13
$K = 100$	5.80	5.75	0.84
$K = 105$	7.61	7.62	0.10
$K = 110$	10.13	10.17	0.39
$K = 115$	14.79	14.87	0.53
$K = 120$	19.76	19.85	0.46

TABLE 4.2: Results on the pricing of American Put options via quantization and the method of Vellekoop and Nieuwenhuis (2009).The strike K is in percentage of the initial price $S_0 = 100$. The maturity is $T = 1$ year, $r = 0.04$. Q stands for quantization. The parameters of the Heston model are those of the calibration exercise in Table 4.3. The quantization grids are those used for the pricing in Table 4.1.

4.3 shows that the parameters calibrated with the quantization are in line with the ones found by the FFT methodology of Carr and Madan (1999). Res Norm indicates the average square error on implied volatilities (IV), defined as

$$\text{Res Norm} := \frac{1}{\#_{\text{strikes}}\,\#_{\text{maturities}}} \sum_{\ell=1}^{\#_{\text{strikes}}} \sum_{k=1}^{\#_{\text{maturities}}} \left(IV_{\ell,k}^{\text{market}} - IV_{\ell,k}^{\text{model}} \right)^2 .$$

Here the implied volatility for American options is defined in analogy with the European case. The numerical results presented in this subsection show that our procedure is also robust from the calibration point of view.
Figure 4.1 displays the implied volatility squared errors for the calibration on American Call options (resp. left hand side) and Put options (resp. right hand side). Overall, the error is in line with the performance of the Heston (1993) model, namely with a resnorm around 10^{-5}.
Finally, Figure 4.2 shows the quantization grids for the asset price in the Heston (1993) model for the parameters' set found in the calibration.

4.4 Conclusion

In this chapter we introduced a new efficient methodology to price options in a stochastic volatility environment through a quantization approach, which only assumes that a (Euler-Maruyama or more sophisticated) discretization scheme for the

	Quantization
ρ	-0.9250
κ	0.1269
ξ	0.4058
θ	0.1922
V_0	0.0319
Res Norm	5.2255×10^{-5}

TABLE 4.3: Heston (1993) model calibrated on a book of American Call and Put options on the GOOG Google stock as of date April, 26th 2017. The book includes 4 maturities (from 3 months to 14 months) with 12 strikes, ranging from 85% to 115% of the spot price $S_0 = 871.71$, for both Call and Put options, for a total of 96 options. Res Norm indicates the average square error on implied volatilities. In analogy with the calibration on vanilla options for the Heston model, see eg Da Fonseca and Grasselli (2011), note that also here the calibrated parameters do not satisfy the Feller's condition $2\kappa\theta \geq \xi^2$.

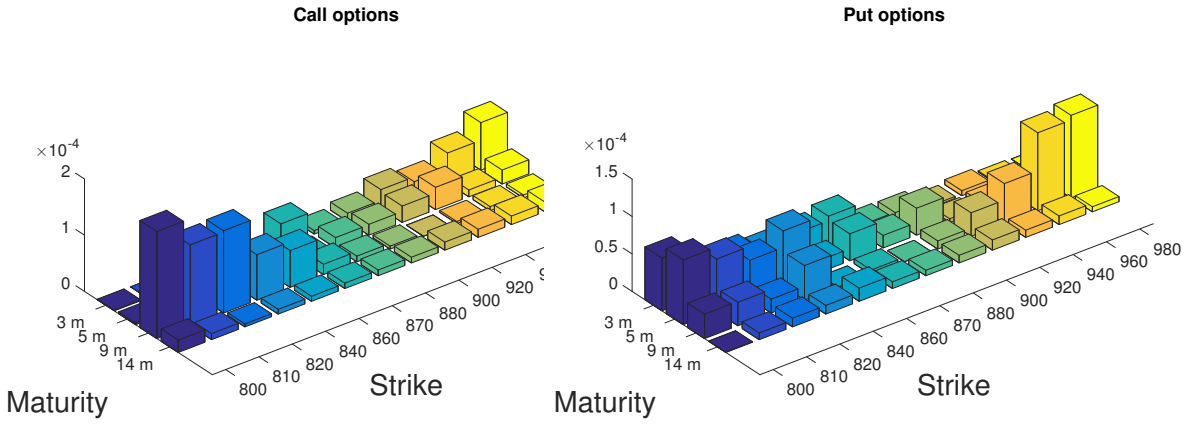


FIGURE 4.1: Implied volatility squared errors for the calibration on a book of American Call and Put options on the GOOG Google stock as of date April, 26th 2017 using the quantization method. American Call options on the left, American Put options on the right.

pair (price, volatility) is available. We have applied the methodology to the Heston (1993) model, for which an efficient benchmark is represented by the Fourier approach. However, we emphasize that our approach is flexible enough to consider also the pricing of American-style options, for which the Fourier technology is almost useless. As a motivating example, we have produced a calibration exercise on real data of the Heston model using a book of options that includes American style

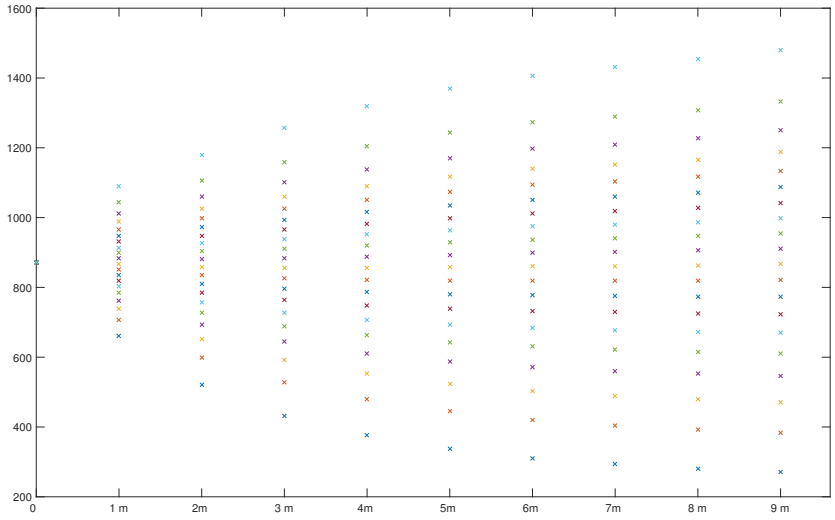


FIGURE 4.2: Quantization grids for the asset price in the Heston (1993) model for the parameter set found in the calibration exercise on the GOOG Google stock as of date April, 26$^{\text{th}}$ 2017, with $N = 20$ points at every time step and maturity 9 months.

contracts. This can be useful, for example, in view of reconstructing the volatility of the index starting from the implied volatility of its constituents, for which only American style options are typically available.

Chapter 5

Stochastic volatility models: a multidimensional approach

This is a joint work with Gilles Pagès and Abass Sagna. The content of this chapter has been submitted as a paper, see Fiorin et al. (2015) in the bibliography.

We introduce a new methodology for the quantization of the Euler scheme for a d-dimensional diffusion process. This method is based on a Markovian and componentwise product quantization and allows us, from a numerical point of view, to speak of *fast online quantization* in a dimension greater than one since the product quantization of the Euler scheme of the diffusion process and its companion weights and transition probabilities may be computed quite instantaneously. We show that the resulting quantization process is a Markov chain, then we compute the associated weights and transition probabilities from (semi-) closed formulas. From the analytical point of view, we show that the induced quantization errors at the k-th discretization step is a cumulative of the marginal quantization error up to that time. Numerical experiments are performed for the pricing of a Basket call option in a correlated Black Scholes framework, for the pricing of a European call option in a Heston model and for the approximation of the solution of backward stochastic differential equations in order to show the performances of the method.

5.1 Introduction

In Pagès and Sagna (2015) a Markovian (fast) quantization of an $\mathbb{R}^d$-valued Euler scheme of a diffusion process is proposed and analyzed. However, in practice, this approach allows to speak of fast quantization only in dimension one since, as soon as $d \geq 2$, one has to use recursive stochastic zero search algorithm (known to be very time consuming, compared to deterministic procedures like the Newton-Raphson algorithm, see Pagès and Printems (2003)) to compute optimal quantizers, their associated weights and transition probabilities. In order to overcome this limitation, we propose in this work another approach to quantize an $\mathbb{R}^d$-valued Euler scheme of a diffusion process. This method is based on a Markovian and componentwise product quantization. It allows again to speak of fast quantization in high dimension since the product quantization of the Euler scheme of the diffusion process and its transition probabilities can be computed almost instantaneously still using deterministic recursive zero search algorithms.
The chapter is organized as follows: Section 5.2 is the main part of this chapter: we present the algorithm and show the Markov property of the product quantization of the Euler scheme of a diffusion process. Then, we show how to compute the

weights and transition probabilities associated to the product quantizers and to its components. We also show how to compute the optimal quantizers associated to each component of the Euler scheme (keep in mind that this is the foundation of our method). Finally, we provide in Theorem 5.2.4 some a priori error bounds for the quantization error associated to the Markovian product quantization and show that, at every step discretization step t_k, this error is a cumulated (weighted) sum of the regular quantization errors, up to time t_k. In Section 5.3, we present some numerical results for the pricing of a European call Basket option and a European call option in the Heston model, as well for the approximation of BSDEs.

NOTATIONS. $\mathcal{M}(d,q,\mathbb{R})$, will denote the $d \times q$ matrices with real values. Given a matrix $A = [a_{ij}] \in \mathcal{M}(d,q,\mathbb{R})$, we denote by $A^\star$ the transpose , the norm of the matrix is defined as $\|A\| := \sqrt{\mathrm{Tr}(AA^\star)} = (\sum_{i,j} a_{ij}^2)^{1/2}$, where $\mathrm{Tr}(M)$ is the trace of the matrix M, for a squared matrix M of size $d \times d$. Given a generic function $f : \mathbb{R}^d \to \mathcal{M}(d,q,\mathbb{R})$, we will denote the Lipschitz constant $[f]_{\mathrm{Lip}} = \sup_{x \neq z} \frac{\|f(x)-f(z)\|}{|x-z|}$. Then we also denote $x \vee z = \max(x,z)$ for two real numbers x and z. If $x \in \mathbb{R}^d$, $|x|_2$ will stand for the Euclidean norm on $\mathbb{R}^d$. For every vectors x, z, the notation $(x|z)$ denotes the dot product of the vectors x and z. For a given vector $z \in \mathbb{R}^q$ and a matrix $M \in \mathcal{M}(d,q,\mathbb{R})$, z^i denotes the component i of z, $z^{(j:k)}$ the vector made up from the component j to the component k of the vector z and $M^{(i,j:k)}$ is the vector made up from the component j to the component k of the i-th row of the matrix M and M^{ij} for the component (i,j) of the matrix M. The notation $M^{i\bullet}$ stands for the i-th row of M. The function Φ_0 will denote the cumulative distribution function of the standard real valued Normal distribution and its derivative Φ_0' will stand for its density function.

5.2 Markovian product quantization of an $\mathbb{R}^d$-valued Euler process

Assume that $(X_t)_{t\geq 0}$ is an $\mathbb{R}^d$ valued stochastic process and a solution to the following SDE:

$$X_t = x_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \quad x_0 \in \mathbb{R}^d, \tag{5.1}$$

where W is a Brownian motion of dimension q , and the functions $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \to \mathcal{M}(d,q,\mathbb{R})$ satisfy the usual global Lipschitz continuity and linear growth assumptions, i.e. for every $t \in [0,T]$,

$$|b(t,x) - b(t,z)| \leq [b]_{\mathrm{Lip}}|x-z| \tag{5.2}$$
$$\|\sigma(t,x) - \sigma(t,z)\| \leq [\sigma]_{\mathrm{Lip}}|x-z| \tag{5.3}$$
$$|b(t,x)| \leq L(1+|x|) \text{ and } \|\sigma(t,x)\| \leq L(1+|x|) \tag{5.4}$$

for some $L > 0$. This guarantees the existence and pathwise uniqueness of a strong solution of (5.1), adapted to the (augmented) filtration of W. We also suppose that the matrix σ is positive definite. Throughout the chapter we will suppose that $\mathbb{R}^d$ is equipped with the canonical Euclidean norm.

5.2.1 The algorithm and the Markov property of the quantized process

Recall that the Euler scheme associated to the stochastic process $(X_t)_{t\geq 0}$ is defined recursively in the following procedure:

$$\overline{X}_{t_{k+1}} = \overline{X}_{t_k} + \Delta b(t_k, \overline{X}_{t_k}) + \sigma(t_k, \overline{X}_{t_k})(W_{t_{k+1}} - W_{t_k}), \quad \overline{X}_0 \in \mathbb{R}^d,$$

where $\Delta = \Delta_n = \frac{T}{n}$ and $t_k = \frac{kT}{n}$, for every $k \in \{0, \cdots, n\}$. To simplify the notation, we will often set $X_k := X_{t_k}$ to denote the process X evaluated at time t_k. We also set $b_k(x) := b(t_k, x)$ and $\sigma_k(x) = \sigma(t_k, x)$ for $x \in \mathbb{R}^d$. Recall also that the Euler operator associated to the conditional distribution of $\overline{X}_{k+1}$ given $\overline{X}_k = x$ is defined by

$$\mathcal{E}_k(x, z) := x + \Delta b(t_k, x) + \sqrt{\Delta}\sigma(t_k, x)z, \quad x \in \mathbb{R}^d,\ z \in \mathbb{R}^q$$

and that if Γ_{k+1} is an N_{k+1}-quantizer for $\overline{X}_{k+1}$, the distortion function $\overline{D}_{k+1}$ associated to $\overline{X}_{k+1}$ can be written for $k = 0, \cdots, n-1$, as

$$\begin{aligned} \overline{D}_{k+1}(\Gamma_{k+1}) &= \mathbb{E}\big((\mathrm{dist}(\overline{X}_{k+1}, \Gamma_{k+1})^2\big) \\ &= \mathbb{E}\big[\mathrm{dist}(\mathcal{E}_k(\overline{X}_k, Z_{k+1}), \Gamma_{k+1})^2\big] \end{aligned}$$

where $Z_{k+1} \sim \mathcal{N}(0; I_q)$ is independent from $\overline{X}_k$. The previous way to write the distortion function has been used in Pagès and Sagna (2015) to propose a fast recursive (and Markovian) quantization of the Euler process (using the Newton-Raphson algorithm for the numerical computation of the optimal grids) when $d = 1$.
Keep in mind that the conditional distribution of the discrete Euler process $\overline{X}$ is Gaussian and that one of the properties of a Gaussian vector is that any sub-component of the vector remains a Gaussian random vector. So, a natural alternative way to quantize the vector $\overline{X}_k \in \mathbb{R}^d$ is to quantize each component $\overline{X}_k^\ell$ by a grid Γ_k^ℓ of size N_k^ℓ, for $\ell = 1, \ldots, d$, and then to define its product quantization $\widehat{X}_k$ associated with the product quantizer $\Gamma_k = \bigotimes_{\ell=1}^d \Gamma_k^\ell$ of size $N_k = N_k^1 \times \ldots \times N_k^d$, as $\widehat{X}_k = (\widehat{X}_k^1, \ldots, \widehat{X}_k^d)$. The question now is how to quantize the $\overline{X}_k^i$'s. On the other hand, since the components of the vector $\overline{X}_k$ are not independent, it is also a challenging question to know how to compute (from closed formula) the companions weights and transition probabilities associated with the quantizations of the $\overline{X}_k^i$'s and the vector $\overline{X}_k$. We describe below the componentwise recursive Markovian quantization of the process $\{\overline{X}_k,\ k = 0, \ldots, n\}$.

It is clear that for every $\ell = 1, \ldots, d$, and for every $k = 0, \ldots, n-1$, the transition operator $\mathcal{E}_k^\ell(x, z)$ associated with the distribution of $\overline{X}_{k+1}^\ell$ given $\overline{X}_k = x$ reads as

$$\mathcal{E}_k^\ell(x, z) := m_k^\ell(x) + \sqrt{\Delta}(\sigma_k^{\ell\bullet}(x)|z), \quad x \in \mathbb{R}^d,\ z \in \mathbb{R}^q,$$

where

$$m_k^\ell(x) := x^\ell + \Delta b_k(x).$$

For every $k = 0, \ldots, n$, for every given $\ell \in \{1, \ldots, d\}$, we denote by $\widehat{X}_k^\ell$ the quantization of $\overline{X}_k^\ell$ on the grid $\Gamma_k^\ell = \{x_k^{\ell, i_\ell}, i_\ell = 1, \ldots, N_k^\ell\}$. We propose in what follows a recursive and componentwise product quantization of the process $\{\overline{X}_k,\ k = 0, \ldots, n\}$. In fact, for every $\ell = 1, \ldots, d$, we denote by Γ_k^ℓ an N_k^ℓ-quantizer (we suppose that we have access to it) of the ℓ-th component $\overline{X}_k^\ell$ of the vector $\overline{X}_k$ and by $\widehat{X}_k^\ell$, the resulting quantization of $\overline{X}_k^\ell$. Then, we define a componentwise recursive product quantizer

$\Gamma_k = \bigotimes_{\ell=1}^d \Gamma_k^\ell$ of size $N_k = N_k^1 \times \ldots \times N_k^d$ of the vector $\overline{X}_k = (\overline{X}_k^\ell)_{\ell=1,\ldots,d}$ by

$$\Gamma_k \quad = \quad \{(x_k^{1,i_1}, \ldots, x_k^{d,i_d}), \quad i_\ell \in \{1, \ldots, N_k^\ell\},\ \ell \in \{1, \ldots, d\}\}.$$

To define the Markovian product quantization, suppose that $\overline{X}_k$ has already been quantized and that we have access to the companion weights $\mathbb{P}(\widehat{X}_k = x_k^i)$, $i \in I_k$, where I_k are defined for every $k \in \{0, \ldots, n\}$ as

$$I_k = \{(i_1, \ldots, i_d),\ i_\ell \in \{1, \ldots, N_k^\ell\}\} \tag{5.5}$$

and for $i := (i_1, \ldots, i_d) \in I_k$, we set

$$x_k^i := (x_k^{1,i_1}, \ldots, x_k^{d,i_d}). \tag{5.6}$$

Setting $\tilde{X}_k^\ell = \mathcal{E}_k^\ell(\widehat{X}_k, Z_{k+1})$, it is possible to approximate the distortion function $\overline{D}_{k+1}^\ell$ associated to the ℓ-th component of the vector $\overline{X}_{k+1}^\ell$ by

$$\begin{aligned}
\tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell) \quad &:= \quad \mathbb{E}\big[\mathrm{dist}(\tilde{X}_{k+1}^\ell, \Gamma_{k+1}^\ell)^2\big] \\
&= \quad \mathbb{E}\big[\mathrm{dist}(\mathcal{E}_k^\ell(\widehat{X}_k, Z_{k+1}), \Gamma_{k+1}^\ell)^2\big] \\
&= \quad \sum_{i \in I_k} \mathbb{E}\big[\mathrm{dist}(\mathcal{E}_k^\ell(x_k^i, Z_{k+1}), \Gamma_{k+1})^2\big] \mathbb{P}\big(\widehat{X}_k = x_k^i\big).
\end{aligned}$$

This allows us to consider the sequence of product recursive quantizations of $(\widehat{X}_k)_{k=0,\cdots,n}$, defined for every $k = 0, \ldots, n-1$, by the following recursion:

$$\begin{cases}
\tilde{X}_0 = \widehat{X}_0, \quad \widehat{X}_k^\ell = \mathrm{Proj}_{\Gamma_k^\ell}(\tilde{X}_k^\ell),\ \ell = 1, \ldots, d, \\
\widehat{X}_k = (\widehat{X}_k^1, \ldots, \widehat{X}_k^d) \quad \text{and} \quad \tilde{X}_{k+1}^\ell = \mathcal{E}_k^\ell(\widehat{X}_k, Z_{k+1}),\ \ell = 1, \ldots, d, \\
\mathcal{E}_k^\ell(x, z) = m_k^\ell(x) + \sqrt{\Delta}(\sigma^{\ell\bullet}(t_k, x)|z),\ z = (z^1, \ldots, z^q) \in \mathbb{R}^q, \\
x = (x^1, \ldots, x^d),\ b = (b^1, \ldots, b^d) \text{ and } (\sigma^{\ell\bullet}(t_k, x)|z) = \sum_{m=1}^q \sigma^{\ell m}(t_k, x) z^m.
\end{cases} \tag{5.7}$$

where $(Z_k)_{k=1,\cdots,n}$ are *i.i.d.*, $\mathcal{N}(0; I_q)$-distributed, independent of $\overline{X}_0$.
In the following result, we show that the sequence $(\widehat{X}_k)_{k\geq 0}$ of Markovian and product quantizations is in fact a Markov chain. Its transition probabilities will be computed further on.

Remark 5.2.1. We may remark that the process $(\widehat{X}_k)_{k\geq 0}$ is a Markov chain on $\mathbb{R}^d$.
In fact, setting $\mathcal{F}_k^{\widehat{X}} = \sigma(\widehat{X}_0, \ldots, \widehat{X}_k)$, we have for any bounded function $f : \mathbb{R}^d \to \mathbb{R}$

$$\begin{aligned}
\mathbb{E}(f(\widehat{X}_{k+1})|\mathcal{F}_k^{\widehat{X}}) \quad &= \quad \sum_{j \in I_{k+1}} \mathbb{E}\left(f(x_{k+1}^j) \mathbb{1}_{\{\widehat{X}_{k+1} = x_{k+1}^j\}} | \mathcal{F}_k^{\widehat{X}}\right) \\
&= \quad \sum_{j \in I_{k+1}} f(x_{k+1}^j) \mathbb{E}\left(\mathbb{1}_{\{\mathcal{E}_k(\widehat{X}_k, Z_{k+1}) \in \prod_{\ell=1}^d C_{j_\ell}(\Gamma_{k+1}^\ell)\}} | \mathcal{F}_k^{\widehat{X}}\right),
\end{aligned}$$

where $\mathcal{E}_k(\widehat{X}_k, Z_{k+1}) = (\mathcal{E}_k^1(\widehat{X}_k, Z_{k+1}), \ldots, \mathcal{E}_k^d(\widehat{X}_k, Z_{k+1}))$. It follows that

$$\mathbb{E}(f(\widehat{X}_{k+1})|\mathcal{F}_k^{\widehat{X}}) = \sum_{j \in I_{k+1}} f(x_{k+1}^j) h_j(\widehat{X}_k),$$

where for every $x \in \mathbb{R}^d$,

$$h_j(x) = \mathbb{P}\big(\mathcal{E}_k(x, Z_{k+1}) \in \prod_{\ell=1}^{d} C_{j_\ell}(\Gamma_{k+1}^{\ell})\big).$$

As a consequence, $\mathbb{E}(f(\widehat{X}_{k+1})|\mathcal{F}_k^{\widehat{X}}) = \varphi(\widehat{X}_k)$, so that $(\widehat{X}_k)_{k\geq 0}$ is a Markov chain.

Now, for a given componentwise (quadratic) optimal quantizers $\Gamma_k = \bigotimes_{\ell=1}^{d} \Gamma_k^{\ell}$, let us explain how to compute the companion transition probability weights associated with the quantizations of the $\overline{X}_k^{\ell}$'s and the whole vector $\overline{X}_k$. We write all the quantities of interest as an expectation of a function of a standard $\mathbb{R}^{q-1}$-valued Normal distribution. These transformations are the key step of this work. In fact, since the optimal quantization grids associated to standard Normal random vectors (up to dimension 10) and their companion weights are available on `www.quantize.maths-fi.com`, these quantities of interest may be computed instantaneously using a cubature formula.

5.2.2 Computing the companion weights and transition probabilities of the marginal quantizations

First of all we define the following quantities which will be needed in the sequel. For every $k \in \{0, \ldots, n-1\}$ and for every $j \in I_{k+1}$ we set

$$x_{k+1}^{i,j_i-1/2} = \frac{x_{k+1}^{i,j_i} + x_{k+1}^{i,j_i-1}}{2},\ x_{k+1}^{i,j_i+1/2} = \frac{x_{k+1}^{i,j_i} + x_{k+1}^{i,j_i+1}}{2},\ \text{with } x_{k+1}^{i,1/2} = -\infty, x_{k+1}^{i,N_{k+1}^i+1/2} = +\infty,$$

and if $Z_k^{(2:q)} = z \in \mathbb{R}^{q-1}$ and $x \in \mathbb{R}^d$, we set (if $\sigma_k^{i1}(x) > 0$)

$$x_{k+1}^{i,j_i-}(x,z) := \frac{x_{k+1}^{i,j_i-1/2} - m_k^i(x) - \sqrt{\Delta}\big(\sigma_k^{(i,2:q)}(x)|z\big)}{\sqrt{\Delta}\sigma_k^{i1}(x)}$$

$$\text{and} \quad x_{k+1}^{i,j_i+}(x,z) := \frac{x_{k+1}^{i,j_i+1/2} - m_k^i(x) - \sqrt{\Delta}\big(\sigma_k^{(i,2:q)}(x)|z\big)}{\sqrt{\Delta}\sigma_k^{i1}(x)}.$$

We also set

$$\mathbb{J}_{k,j_i}^0(x) = \Big\{z \in \mathbb{R}^{q-1},\quad \sqrt{\Delta}\big(\sigma_k^{(i,2:q)}(x)|z\big) \in \big(x_{k+1}^{i,j_i-1/2} - m_k^i(x), x_{k+1}^{i,j_i+1/2} - m_k^i(x)\big)\Big\}$$

and

$$\begin{aligned}
\mathbb{J}_k^0(x) &= \big\{i \in \{1,\ldots,d\},\quad \sigma_k^{i1}(x) = 0\big\}\\
\mathbb{J}_k^-(x) &= \big\{i \in \{1,\ldots,d\},\quad \sigma_k^{i1}(x) < 0\big\}\\
\mathbb{J}_k^+(x) &= \big\{i \in \{1,\ldots,d\},\quad \sigma_k^{i1}(x) > 0\big\}.
\end{aligned}$$

The following result allows us to compute the weights and the transition probabilities associated to the quantizations $\widehat{X}_k$, $k = 0, \ldots, n$.

Proposition 5.2.1. *Let $\{\widehat{X}_k, k = 0, \ldots, n\}$ be the sequence defined from the algorithm (5.7). The transition probability $\mathbb{P}(\widehat{X}_{k+1} = x_{k+1}^j|\widehat{X}_k = x_k^\ell)$, $\ell \in I_k$, $j \in I_{k+1}$, is given by*

$$\mathbb{P}\big(\widehat{X}_{k+1} = x_{k+1}^j|\widehat{X}_k = x_k^i\big) = \mathbb{E}\prod_{\ell \in \mathbb{J}_k^0(x_k^i)} \mathbf{1}_{\{\zeta \in \mathbb{J}_{k,j_\ell}^0(x_k^i)\}} \max\big(\Phi_0(\beta_j(x_k^i,\zeta)) - \Phi_0(\alpha_j(x_k^i,\zeta)), 0\big) \quad (5.8)$$

where $\zeta \sim \mathcal{N}(0; I_{q-1})$ *and, for every* $x \in \mathbb{R}^d$ *and* $z \in \mathbb{R}^{q-1}$,

$$\alpha_j(x,z) = \max\Big(\sup_{\ell \in \mathbb{J}_k^+(x)} x_{k+1}^{\ell, j_\ell -}(x,z), \sup_{\ell \in \mathbb{J}_k^-(x)} x_{k+1}^{\ell, j_\ell +}(x,z)\Big)$$
$$\text{and } \beta_j(x,z) = \min\Big(\inf_{\ell \in \mathbb{J}_k^+(x)} x_{k+1}^{\ell, j_\ell +}(x,z), \inf_{\ell \in \mathbb{J}_k^-(x)} x_{k+1}^{\ell, j_\ell -}(x,z)\Big).$$

Before proving this result, remark that we may deduce the probability weights associated to the quantizations $(\widehat{X}_{k+1})$ by

$$\mathbb{P}\big(\widehat{X}_{k+1} = x_{k+1}^j\big) = \sum_{i \in I_k} \mathbb{P}\big(\widehat{X}_{k+1} = x_{k+1}^j | \widehat{X}_k = x_k^i\big)\mathbb{P}\big(\widehat{X}_k = x_k^i\big) \tag{5.9}$$

where the conditional probabilities are computed using the formula (5.8).

Proof. See Appendix 5.4.1. □

Now, we focus on in the particular case where the matrix $\sigma(t,x)$, for $(t,x) \in [0,T]\times\mathbb{R}^d$, is diagonal with positive diagonal entries $\sigma^{\ell\ell}(t,x)$, $\ell = 1,\dots,d$. The following result says how to compute the transition probability weights of the $\widehat{X}_k$'s. Let us set for every $x \in \mathbb{R}^d$, every $\ell \in \{1,\dots,d\}$ and $j_\ell \in \{1,\dots,N_{k+1}^\ell\}$,

$$x_{k+1}^{\ell, j_\ell -}(x,0) := \frac{x_{k+1}^{\ell, j_\ell - 1/2} - m_k^\ell(x)}{\sqrt{\Delta}\sigma_k^{\ell\ell}(x)} \quad \text{and} \quad x_{k+1}^{\ell, j_\ell +}(x,0) := \frac{x_{k+1}^{\ell, j_\ell + 1/2} - m_k^\ell(x)}{\sqrt{\Delta}\sigma_k^{\ell\ell}(x)}.$$

Proposition 5.2.2. *Let* $\{\widehat{X}_k, k = 0,\dots,n\}$ *be the sequence of quantizers defined by the algorithm* (5.7) *and associated with the solution* (X_t) *of* (5.1). *Suppose that the volatility matrix* $\sigma(t,x)$ *of* $(X_t)_{t\geq 0}$ *is diagonal with positive diagonal entries* $\sigma^{\ell\ell}(t,x)$, $\ell = 1,\dots,d$. *Then, the transition probability weights* $\mathbb{P}(\widehat{X}_{k+1} = x_{k+1}^j | \widehat{X}_k = x_k^i)$, $i \in I_k, j \in I_{k+1}$, *are given by*

$$\mathbb{P}\big(\widehat{X}_{k+1} = x_{k+1}^j | \widehat{X}_k = x_k^i\big) = \prod_{\ell=1}^d \mathbb{P}\big(\widehat{X}_{k+1}^\ell = x_{k+1}^{j_\ell} | \widehat{X}_k = x_k^i\big) \tag{5.10}$$
$$= \prod_{\ell=1}^d \big[\Phi_0\big(x_{k+1}^{\ell, j_\ell +}(x_k^i, 0)\big) - \Phi_0\big(x_{k+1}^{\ell, j_\ell -}(x_k^i, 0)\big)\big], \tag{5.11}$$

and the companion probability weights $\mathbb{P}\big(\widehat{X}_{k+1} = x_{k+1}^j\big)$ *are given for every* $k = 0,\dots,n-1$ *and every* $j \in I_{k+1}$ *by*

$$\mathbb{P}\big(\widehat{X}_{k+1} = x_{k+1}^j\big) = \sum_{i \in I_k} \prod_{\ell=1}^d \big[\Phi_0\big(x_{k+1}^{\ell, j_\ell +}(x_k^i, 0)\big) - \Phi_0\big(x_{k+1}^{\ell, j_\ell -}(x_k^i, 0)\big)\big]\mathbb{P}\big(\widehat{X}_k = x_k^i\big) \tag{5.12}$$

Proof. See Appendix 5.4.2. □

The following result is useful in the situation where we need to approximate the expectation of a function of one component of the vector $\overline{X}_k$ as for example in the pricing of European options in the Heston model.

Proposition 5.2.3. *Let* Γ_{k+1}^ℓ *be an optimal quantizer for the random variable* $\tilde{X}_{k+1}^\ell$. *Suppose that the optimal product quantizer* Γ_k *for* $\tilde{X}_k$ *and its companion weights* $\mathbb{P}(\widehat{X}_k = x_k^i)$, $i \in I_k$, *have already been computed.*

For any $\ell \in \{1,\dots,d\}$ and any $j_\ell \in \{1,\dots,N_{k+1}^\ell\}$, the transition probability weights $\mathbb{P}(\tilde{X}_{k+1}^\ell \in C_{j_\ell}(\Gamma_{k+1}^\ell)|\widehat{X}_k = x_k^i)$ are given by

$$\mathbb{P}\big(\tilde{X}_{k+1}^\ell \in C_{j_\ell}(\Gamma_{k+1}^\ell)|\widehat{X}_k = x_k^i\big) = \Phi_0\big(x_{k+1}^{\ell,j_\ell+}(x_k^i)\big) - \Phi_0\big(x_{k+1}^{\ell,j_\ell-}(x_k^i)\big). \tag{5.13}$$

Proof. See Appendix 5.4.3. □

Notice that the companion probability $\mathbb{P}(\tilde{X}_{k+1}^\ell \in C_{j_\ell}(\Gamma_{k+1}))$ is given, for every $\ell \in \{1,\dots,d\}$ and for every $j_\ell \in \{1,\cdots,N_{k+1}^\ell\}$, by

$$\mathbb{P}\big(\tilde{X}_{k+1}^\ell \in C_{j_\ell}(\Gamma_{k+1}^\ell)\big) = \sum_{i\in I_k}\left[\Phi_0\left(\frac{x_{k+1}^{\ell,j_\ell+1/2} - m_k^\ell(x_k^i)}{\sqrt{\Delta}|\sigma_k^{\ell\bullet}(x_k^i)|_2}\right) - \Phi_0\left(\frac{x_{k+1}^{\ell,j_\ell-1/2} - m_k^\ell(x_k^i)}{\sqrt{\Delta}|\sigma_k^{\ell\bullet}(x_k^i)|_2}\right)\right]\mathbb{P}\big(\widehat{X}_k = x_k^i\big). \tag{5.14}$$

We may note that the ℓ-th component process $(\widehat{X}_k^\ell)_{k\geq 0}$ is not a Markov chain. We may however compute the transition probabilities

$$\mathbb{P}(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}|\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}}), \quad \ell,\ell' \in \{1,\dots,d\},\ j_\ell \in \{1,\dots,N_{k+1}^\ell\},\ j_{\ell'} \in \{1,\dots,N_k^{\ell'}\}.$$

This is the aim of the following remark which follows from *Bayes* formula.

Remark 5.2.2. For $\ell,\ell' \in \{1,\dots,d\}$, $j_\ell \in \{1,\dots,N_{k+1}^\ell\}$ and $j_{\ell'} \in \{1,\dots,N_k^{\ell'}\}$, we have

$$\mathbb{P}\big(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}|\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}}\big) = \sum_{i\in I_k}\mathbb{1}_{\{i_{\ell'}=j_{\ell'}\}}\frac{\mathbb{P}(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}|\widehat{X}_k = x_k^i)}{\mathbb{P}(\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}})}\mathbb{P}\big(\widehat{X}_k = x_k^i\big) \tag{5.15}$$

where the terms $\mathbb{P}(\widehat{X}_k = x_k^i)$, $\mathbb{P}(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}|\widehat{X}_k = x_k^i)$ and $\mathbb{P}(\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}})$ are computed from (5.9), (5.13) and (5.14), respectively.
As a matter of fact, applying *Bayes* formula and summing over $i \in I_k$ yields:

$$\begin{aligned}\mathbb{P}\big(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}|\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}}\big) &= \sum_{i\in I_k}\frac{\mathbb{P}(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}, \widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}}, \widehat{X}_k = x_k^i)}{\mathbb{P}(\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}})}\\ &= \sum_{i\in I_k}\mathbb{1}_{\{i_{\ell'}=j_{\ell'}\}}\frac{\mathbb{P}(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}, \widehat{X}_k = x_k^i)}{\mathbb{P}(\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}})}\\ &= \sum_{i\in I_k}\mathbb{1}_{\{i_{\ell'}=j_{\ell'}\}}\frac{\mathbb{P}(\widehat{X}_{k+1}^\ell = x_{k+1}^{\ell,j_\ell}|\widehat{X}_k = x_k^i)}{\mathbb{P}(\widehat{X}_k^{\ell'} = x_k^{\ell',j_{\ell'}})}\mathbb{P}\big(\widehat{X}_k = x_k^i\big).\end{aligned}$$

In the foregoing, we assume that we have access to the N_k^ℓ-quantizers Γ_k^ℓ of the ℓ-th component $\overline{X}_k^\ell$ of the vector $\overline{X}_k$, for every $\ell = 1,\dots,d$. We show how to compute the distortion functions associated with every component of the vector $\tilde{X}_{k+1}$, $k = 0,\dots,n-1$. From the numerical point of view, this will allow us to use the Newton-Raphson algorithm to compute the optimal quantizers associated to each component $\tilde{X}_{k+1}^\ell$, $\ell = 1,\dots,d$, of the vector $\tilde{X}_{k+1}$, for $k = 0,\dots,n-1$. Then, the quantization $\widehat{X}_{k+1}$ of $\tilde{X}_{k+1}$ is defined as the product quantization $\widehat{X}_k = (\widehat{X}_k^1,\dots,\widehat{X}_k^d)$, where $\widehat{X}_k^\ell = \mathrm{Proj}_{\Gamma_{k+1}^\ell}(\tilde{X}_{k+1}^\ell)$.

5.2.3 Computing the distortion, the gradient and the Hessian matrix associated to a componentwise quantizer

Our aim, for the computation of the componentwise optimal quantizations, is to use the Newton-Raphson algorithm in $\mathbb{R}^{N_k}$ which involves the gradient and the Hessian matrix of the distortion functions $\tilde{D}_k^\ell$, $k = 0, \dots, n$; $\ell = 1, \dots, d$. In the following, we give useful expressions for the distortion functions $\tilde{D}_k^\ell$, their gradient vectors $\nabla \tilde{D}_k^\ell$ and their Hessian matrices $\nabla^2 \tilde{D}_k^\ell$. We state these results in the next proposition. Above all, recall that for every $\ell = 1, \dots, d$, for every $k = 0, \dots, n-1$,

$$\tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell) = \sum_{i \in I_k} \mathbb{E}\big[d(\mathcal{E}_k^\ell(x_k^i, Z_{k+1}), \Gamma_{k+1}^\ell)^2\big] \mathbb{P}\big(\widehat{X}_k = x_k^i\big)$$

and notice the distortion function $\tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell)$ is continuously differentiable as a function of the N_{k+1}-quantizer $\Gamma_{k+1}^\ell = \{x_{k+1}^{\ell,j_\ell},\ j_\ell = 1, \dots N_{k+1}^\ell\}$ (having pairwise distinct components so that it can be viewed as an N_{k+1}^ℓ-tuple) and its gradient vector reads

$$\nabla \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell) = 2\left[\sum_{i \in I_k} \mathbb{E}\Big(\mathbb{1}_{\{\mathcal{E}_k^\ell(x_k^i, Z_{k+1}) \in C_{j_\ell}(\Gamma_{k+1}^\ell)\}}\big(x_{k+1}^{\ell,j_\ell} - \mathcal{E}_k^\ell(x_k^i, Z_{k+1})\big)\Big)\mathbb{P}(\widehat{X}_k = x_k^\ell)\right]_{j_\ell = 1, \cdots, N_{k+1}^\ell}.$$

We recall that key point of our method is to deal with the product quantization of the components of the process $(\overline{X}_k)_{0 \le k \le n}$. From a numerical point of view, each component will be quantized using the Newton-Raphson algorithm. To this end, we have to compute (explicitly) the distortion function $\tilde{D}_{k+1}^\ell(\cdot)$, the components of its gradient vector and the components its Hessian matrix. This is the purpose of the following remark. Its proof relies on tedious though elementary computation. Therefore, we have deliberately omitted the proof.

Remark 5.2.3. Recall that for every $\ell \in \{1, \dots, d\}$, $\vartheta_k^\ell(x)^2 = \sum_{p=1}^q \Delta\big(\sigma_k^{\ell p}(x)\big)^2$.
a) *Distortion*. We have for every $\ell = 1, \dots, d$ and every $k = 0, \dots, n-1$,

$$\tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell) = \sum_{j_\ell = 1}^{N_{k+1}^\ell} \sum_{i \in I_k} \Psi_{\ell, j_\ell}(x_k^i)\, p_k^i = \sum_{j_\ell = 1}^{N_{k+1}^\ell} \mathbb{E}\Psi_{\ell, j_\ell}(\widehat{X}_k), \tag{5.16}$$

where for every $x \in \mathbb{R}^d$,

$$\begin{aligned}\Psi_{\ell, j_\ell}(x) = &\Big(\big(m_k^\ell(x) - x_{k+1}^{\ell,j_\ell}\big)^2 + \vartheta_k^\ell(x)^2\Big)\Big(\Phi_0\big(x_{k+1}^{\ell,j_\ell +}(x)\big) - \Phi_0\big(x_{k+1}^{\ell,j_\ell -}(x)\big)\Big)\\ &+ 2\vartheta_k^\ell(x)\Big(x_{k+1}^{\ell,j_\ell} - m_k^\ell(x)\Big)\Big(\Phi_0'\big(x_{k+1}^{\ell,j_\ell +}(x)\big) - \Phi_0'\big(x_{k+1}^{\ell,j_\ell -}(x)\big)\Big)\\ &- \vartheta_k^\ell(x)^2\Big(x_{k+1}^{\ell,j_\ell +}(x)\Phi_0'\big(x_{k+1}^{\ell,j_\ell +}(x)\big) - x_{k+1}^{\ell,j_\ell -}(x)\Phi_0'\big(x_{k+1}^{\ell,j_\ell -}(x)\big)\Big).\end{aligned}$$

b) *Gradient*. The components of the gradient $\nabla \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell)$ are given for every $j_\ell = 1, \dots, N_{k+1}^\ell$ by

$$\frac{\partial \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell)}{\partial x_{k+1}^{\ell,j_\ell}} = \sum_{i \in I_k} \Psi_{j_\ell}'(x_k^i)\, p_k^i = \mathbb{E}\Psi_{j_\ell}'(\widehat{X}_k) \tag{5.17}$$

where for every $x \in \mathbb{R}^d$,

$$\Psi'_{j_\ell}(x) = \big(x_{k+1}^{\ell,j_\ell} - m_k^\ell(x)\big)\Big(\Phi_0\big(x_{k+1}^{\ell,j_\ell+}(x)\big) - \Phi_0\big(x_{k+1}^{\ell,j_\ell-}(x)\big)\Big) + \vartheta_k^\ell(x)\Big(\Phi'_0\big(x_{k+1}^{\ell,j_\ell+}(x)\big) - \Phi'_0\big(x_{k+1}^{\ell,j_\ell-}(x)\big)\Big).$$

c) *Hessian*. The sub-diagonal, the super-diagonals and the diagonal terms of the Hessian matrix are given respectively by

$$\frac{\partial^2 \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell)}{\partial x_{k+1}^{\ell,j_\ell}\partial x_{k+1}^{\ell,j_\ell-1}} = \sum_{i\in I_k} \Psi''_{j_\ell,j_\ell-1}(x_k^i)\, p_k^i = \mathbb{E}\Psi''_{j_\ell,j_\ell-1}(\widehat{X}_k),$$

$$\frac{\partial^2 \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell)}{\partial x_{k+1}^{\ell,j_\ell}\partial x_{k+1}^{\ell,j_\ell+1}} = \sum_{i\in I_k} \Psi''_{j_\ell,j_\ell+1}(x_k^i)\, p_k^i = \mathbb{E}\Psi''_{j_\ell,j_\ell+1}(\widehat{X}_k)$$

$$\text{and} \quad \frac{\partial^2 \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^\ell)}{\partial^2 x_{k+1}^{\ell,j_\ell}} = \sum_{i\in I_k} \Psi''_{j_\ell,j_\ell}(x_k^i)\, p_k^i = \mathbb{E}\Psi''_{j_\ell,j_\ell}(\widehat{X}_k)$$

where, for every $x \in \mathbb{R}^d$,

$$\Psi''_{j_\ell,j_\ell-1}(x) = -\frac{1}{4}\frac{1}{\vartheta_k^\ell(x)}(x_{k+1}^{\ell,j_\ell} - x_{k+1}^{\ell,j_\ell-1})\Phi'_0\big(x_{k+1}^{\ell,j_\ell-}(x)\big),$$

$$\Psi''_{j_\ell,j_\ell+1}(x) = -\frac{1}{4}\frac{1}{\vartheta_k^\ell(x)}(x_{k+1}^{\ell,j_\ell+1} - x_{k+1}^{\ell,j_\ell})\Phi'_0\big(x_{k+1}^{\ell,j_\ell+}(x)\big),$$

$$\Psi''_{j_\ell,j_\ell}(x) = \Phi_0\big(x_{k+1}^{\ell,j_\ell+}(x)\big) - \Phi_0\big(x_{k+1}^{\ell,j_\ell-}(x)\big) + \Psi''_{j_\ell,j_\ell-1}(x) + \Psi''_{j_\ell,j_\ell+1}(x).$$

Once we have access to the gradient vector and the Hessian matrix associated with $\tilde{X}_{k+1}^\ell$ and to the optimal grids and companions weights associated with the $\widehat{X}_{k'}$'s, $k' = 0,\ldots,k$, it is possible to write down (at least formally) a Newton-Raphson zero search procedure to compute the optimal quantizer Γ_{k+1}^ℓ. The Newton-Raphson algorithm is in fact indexed by $p \geq 0$, where a current grid $\Gamma_{k+1}^{\ell,p}$ is updated as follows:

$$\Gamma_{k+1}^{\ell,p+1} = \Gamma_{k+1}^{\ell,p} - \big(\nabla^2 \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^{\ell,p})\big)^{-1}\nabla \tilde{D}_{k+1}^\ell(\Gamma_{k+1}^{\ell,p}), \quad p \geq 1, \tag{5.18}$$

starting from a $\Gamma_{k+1}^{\ell,0} \in \mathbb{R}^{N_{k+1}^\ell}$ (with increasing components).

Remark 5.2.4. (*Stationarity property*) If Γ_{k+1}^ℓ is an optimal Markovian product quantizer for $\tilde{X}_{k+1}^\ell$ and if $\widehat{X}_{k+1}^\ell$ denotes the quantization of $\tilde{X}_{k+1}^\ell$ by the grid Γ_{k+1}^ℓ, then Γ_{k+1}^ℓ is a stationary quantizer for $\tilde{X}_{k+1}^\ell$, means, $\mathbb{E}\Big(\tilde{X}_{k+1}^\ell \big| \widehat{X}_{k+1}^\ell\Big) = \widehat{X}_{k+1}^\ell$. Equivalently, this means that if $\Gamma_{k+1}^\ell = \big\{x_{k+1}^{\ell,j_\ell},\, j_\ell = 1,\ldots,N_{k+1}^\ell\big\}$ with

$$x_{k+1}^{\ell,j_\ell} = \frac{\sum_{i\in I_k} \mathbb{E}\big(\mathcal{E}_k^\ell(x_k^i, Z_{k+1})\mathbb{1}_{\{\mathcal{E}_k^\ell(x_k^i,Z_{k+1})\in C_{j_\ell}(\Gamma_{k+1}^\ell)\}}\big)\mathbb{P}(\widehat{X}_k = x_k^i)}{p_{k+1}^{j_\ell}} \tag{5.19}$$

$$\text{and} \quad p_{k+1}^{j_\ell} = \sum_{i\in I_k} \mathbb{P}\big(\mathcal{E}_k^\ell(x_k^i, Z_{k+1}) \in C_{j_\ell}(\Gamma_{k+1}^\ell)\big)\mathbb{P}(\widehat{X}_k = x_k^i),\ j_\ell = 1,\ldots,N_{k+1}^\ell \tag{5.20}$$

A straightforward computation leads to the following result: if $\Gamma_{k+1}^\ell = \{x_{k+1}^{\ell,j_\ell},\, j_\ell = 1,\ldots,N_{k+1}^\ell\}$ is a stationary quantizer for $\tilde{X}_{k+1}^\ell$ then, for every $\ell \in \{1,\ldots,d\}$ and for

every $j_\ell \in \{1, \dots, N_{k+1}^\ell\}$,

$$x_{k+1}^{\ell,j_\ell} = \frac{\sum_{i\in I_k}\left[m_k^\ell(x_k^i)\gamma_{\ell,k}(x_k^i) - \vartheta_k^\ell(x_k^i)\gamma'_{\ell,k}(x_k^i)\right] p_k^i}{\sum_{i\in I_k}\gamma_{\ell,k}(x_k^i)\, p_k^i} \tag{5.21}$$

where for every $x \in \mathbb{R}^d$,

$$\gamma_{\ell,k}(x) = \Phi_0\big(x_{k+1}^{\ell,j_\ell+}(x)\big) - \Phi_0\big(x_{k+1}^{\ell,j_\ell-}(x)\big) \text{ and } \gamma'_{\ell,k}(x) = \Phi'_0\big(x_{k+1}^{\ell,j_\ell+}(x)\big) - \Phi'_0\big(x_{k+1}^{\ell,j_\ell-}(x)\big). \tag{5.22}$$

5.2.4 The error analysis

Our aim is now to compute the quadratic quantization error bound $\|\overline{X}_T - \widehat{X}_T\|_2 := \|\overline{X}_n - \widehat{X}_n^{\Gamma_n}\|_2$. The analysis of this error bound is the subject of the following theorem. We suppose that $x_0 = X_0 = \tilde{X}_0$. We consider here a regular time discretization $(t_k)_{0\le k\le n}$ with step $\Delta = T/n$: $t_k = \frac{kT}{n}$, $k = 0, \dots, n$.

Theorem 5.2.4. *Assume the coefficients b, σ satisfy the classical Lipschitz assumptions (5.2), (5.3) and (5.4). Let, for every $k = 0, \dots, n$, Γ_k be a Markovian product quantizer for $\tilde{X}_k$ at level N_k. Then, for every $k = 0, \cdots, n$, for any $\eta \in]0, 1]$,*

$$\|\overline{X}_k - \widehat{X}_k^{\Gamma_k}\|_2 \le K_{2,\eta} \sum_{k'=1}^{k} e^{(k-k')\Delta C_{b,\sigma}} a_{k'}\big(b, \sigma, d, k, \Delta, x_0, L, 2+\eta\big) N_{k'}^{-1/d} \tag{5.23}$$

where for every $p \in (2, 3]$,

$$a_{k'}(b, \sigma, d, k, d, \Delta, x_0, L, p) := e^{\Delta C_{b,\sigma}\frac{(k-k')}{p}}\left[e^{(\kappa_p+K_p)t_{k'}}|x_0|^p + \frac{d^{k(\frac{p}{2}-1)}e^{\kappa_p\Delta}L + K_p}{d^{\frac{p}{2}-1}(\kappa_p + K_p)}\left(e^{(\kappa_p+K_p)t_{k'}} - 1\right)\right]^{\frac{1}{p}},$$

with $C_{b,\sigma} = [b]_{\mathrm{Lip}} + \frac{1}{2}[\sigma]^2_{\mathrm{Lip}}$, $K_{2,\eta} := K_{2,1,\eta}$ is a universal constant defined in Equation (1.8);

$$\kappa_p := \left(\frac{(p+1)(p-2)}{2} + 2pL\right) \text{ and } K_\theta := 2^{p-1}L^p\left(1 + p + \frac{p(p-1)}{2}\Delta^{\frac{p}{2}-1}\right)\mathbb{E}|Z|^p,\ Z \sim \mathcal{N}(0; I_d).$$

Proof. See Appendix 5.4.4. □

5.3 Numerical examples

It is noteworthy that the pricing of vanilla and exotic options still represents a challenge from a numerical point of view when the dimension of the process considered is greater than one. Monte Carlo techniques seem to be the only possibility to tackle these issues. This quantization technique is an alternative to simulation based procedures, as it gives the possibility to price options in both (correlated) local volatility models and stochastic volatility models.
The computations for all numerical examples have been implemented in *MATLAB*, and have been performed on a CPU 2.7 GHz and 4 Gb memory computer.

5.3.1 Pricing of a Basket European option

We consider a European Basket option with maturity T and strike K, based on two stocks with prices S^1 and S^2 with associated weights w_1 and w_2. We suppose that S^1

and S^2 evolve following the dynamics

$$\begin{cases} dS_t^1 = rS_t^1 dt + \rho\,\sigma_1 S_t^1 dW_t^1 + \sqrt{1-\rho^2}\,\sigma_1 S_t^1 dW_t^2 \\ dS_t^2 = rS_t^2 dt + \sigma_2 S_t^2 dW_t^1 \end{cases} \tag{5.24}$$

where W^1 and W^2 are two independent Brownian motion, r is the interest rate and $\rho \in [-1,1]$, is the correlation term.
We know that in this case, the price at time $t=0$ of the call option reads

$$e^{-rT}\mathbb{E}\big[\max(w_1 S_T^1 + w_2 S_T^2 - K, 0)\big] = e^{-rT}\mathbb{E}F(X_T), \quad X = (S^1, S^2), \tag{5.25}$$

where the function F is defined, for every $x = (s^1, s^2) \in \mathbb{R}^2$, by $F(x) = \max(w_1 s^1 + w_2 s^2 - K, 0)$. Using the Markovian product quantization, the price of the Basket European option is approximated by

$$e^{-rT} \sum_{j \in I_n} F(x_n^j)\mathbb{P}(\widehat{X}_n = x_n^j). \tag{5.26}$$

For the numerical exercises we will use the following set of parameters:

$$r = 0.04, \quad \sigma_1 = 0.3, \quad \sigma_2 = 0.4, \quad \rho = 0.5$$
$$w_1 = w_2 = 0.5, \quad S_0^1 = 100, \quad S_0^2 = 100, \quad T = 1.$$

The benchmark price is given by the algorithm developed in Ju (2002). The results are given in Table 5.1 and Table 5.2. For both tables, we consider Call prices for the strikes $K \in \{80, 85, 90, 95, 100\}$ and Put prices for strikes K lying to $\{100, 105, 110, 115, 120\}$.
We also depict in Table 5.1 the associated relative error $\left(\text{Err.} = \frac{|\text{Price} - MQ_N^n|}{\text{Price}}\right)$ between the benchmark prices (Price) and the prices obtained from the Markovian and product quantization method of size $N = N_1 = N_2$ with n discretization steps (denoted by MQ_N^n) and the computation time (in seconds) for the Markovian and product quantization method.
In Table 5.1, we set the number of time steps $n = 10$ and make the sizes N_1 and N_2 of the marginal quantizers varying whereas, for the results of Table 5.2, we set $N_1 = N_2 = 30$ and make varying the number n of time steps.
We verify, as expected, that increasing the size of the marginal quantizers lead to more precise results (see Theorem 5.2.4). However, this increases the computation time. On the other hand, it is also clear from Theorem 5.2.4 that fixing the marginal quantization size, the number n of the time steps increases the global quantization error. From the numerical results, the choice $N_1 = N_2 = 20$ and $n = 10$ seems to be a good compromise.

5.3.2 Pricing of a European option in the Heston model

Stochastic volatility models have been introduced in order to catch the smile and skew property of the implied volatility surface. Most of the models use Fourier techniques to compute the price of vanilla options, as in Hull and White (1987); Heston (1993); Christoffersen et al. (2009). In this example, we consider a European option with maturity T and strike K, in the Heston model, introduced in Heston (1993), where the stock price S and its stochastic variance V evolve following the

K	Price	MQ_{10}^{10}	Err. (%)	MQ_{20}^{10}	Err. (%)	MQ_{30}^{10}	Err. (%)
80	25.9491	25.4427	1.9516	25.8721	0.2966	25.9656	0.0636
85	22.4481	21.9007	2.4384	22.3543	0.4177	22.4532	0.0229
90	19.2736	18.6934	3.0101	19.1596	0.5915	19.2612	0.0645
95	16.4323	15.8139	3.7633	16.2935	0.8450	16.3964	0.2183
100	13.9197	13.2858	4.5541	13.7537	1.1929	13.8566	0.4535
100	9.9987	9.3727	6.2602	9.8406	1.5810	9.9435	0.5515
105	12.6050	11.9509	5.1895	12.4218	1.4536	12.5218	0.6603
110	15.5060	14.8264	4.3828	15.2981	1.3408	15.3965	0.7062
115	18.6768	18.0111	3.5646	18.4441	1.2461	18.5422	0.7204
120	22.0904	21.4207	3.0317	21.8432	1.1189	21.9345	0.7055
Time (s)		0.49		8.41		41.82	

TABLE 5.1: Prices of a Basket option. The prices correspond to Call prices for strikes $K \in \{80, 85, 90, 95, 100\}$ and to Put prices for strikes $K \in \{100, 105, 110, 115, 120\}$. The number of time steps $n = 10$. Size of the grids $N_1 = N_2 = N$. The error Err. $= \frac{|\text{Price} - \text{MQ}_N^n|}{\text{Price}}$ corresponds to the relative error between the benchmark price (Price) and Product Markovian quantization price MQ_N^n of size N with n discretization steps.

dynamics

$$\begin{cases} dS_t = rS_t dt + \sqrt{V_t} S_t dW_t^1 \\ dV_t = \kappa(\theta - V_t)dt + \rho\,\sigma\sqrt{V_t} dW_t^1 + \sqrt{1-\rho^2}\,\sigma\sqrt{V_t} dW_t^2. \end{cases} \tag{5.27}$$

In the previous equation, the parameter r is the interest rate; $\kappa > 0$, is the rate at which V reverts to the long running average variance $\theta > 0$; the parameter $\sigma > 0$, is the volatility of the variance and $\rho \in [-1, 1]$, is the correlation term. In this case, the price of the call at time $t = 0$ reads

$$e^{-rT}\mathbb{E}\big[\max(S_T - K, 0)\big] = e^{-rT}\mathbb{E}H(X_T), \quad x = (S, V), \tag{5.28}$$

where $H(x) = \max(x^1 - K, 0)$, for $x = (x^1, x^2) \in \mathbb{R}^2$.

K	Price	MQ_{30}^{20}	Err. (%)	MQ_{30}^{30}	Err. (%)	MQ_{30}^{40}	Err. (%)
80	25.9491	25.8461	0.3969	25.7646	0.7112	25.6937	0.9844
85	22.4481	22.3335	0.5107	22.2473	0.8943	22.1731	1.2252
90	19.2736	19.1460	0.6618	19.0599	1.1086	18.9837	1.5041
95	16.4323	16.2916	0.8561	16.2074	1.3686	16.1310	1.8336
100	13.9197	13.7660	1.1041	13.6861	1.6781	13.6117	2.2130
100	9.9987	9.8490	1.4971	9.7677	2.3095	9.6926	3.0608
105	12.6050	12.4428	1.2869	12.3675	1.8842	12.2954	2.4566
110	15.5060	15.3328	1.1173	15.2642	1.5593	15.1966	1.9956
115	18.6768	18.4954	0.9714	18.4342	1.2990	18.3717	1.6337
120	22.0904	21.9045	0.8413	21.8506	1.0854	21.7931	1.3459
Time (s)		84.13		151.90		194.05	

TABLE 5.2: Prices of a Basket option. The prices correspond to Call prices for strikes $K \in \{80, 85, 90, 95, 100\}$ and to Put prices for strikes $K \in \{100, 105, 110, 115, 120\}$. Size of the grids $N_1 = N_2 = 30$ and the number of time steps $n \in 20, 30, 40$. The error Err. $= \frac{|\text{Price} - \text{MQ}_N^n|}{\text{Price}}$ corresponds to the relative error between the benchmark price (Price) and Product Markovian quantization price MQ_N^n of size N with n discretization steps.

Using the Markovian and product quantization method, the price of a call in the Heston model is approximated as

$$e^{-rT} \sum_{j_1=1}^{N_n^1} \max(s_n^{1j_1} - K, 0)\, \mathbb{P}(\widehat{S}_n^1 = s_n^{1j_1}), \tag{5.29}$$

where $\mathbb{P}(\widehat{S}_n^1 = s_n^{1j_1})$ is computed according to (5.9).

For the numerical experiments we will use the following parameters, obtained from a calibration on market prices:

$$\kappa = 2.3924, \quad \theta = 0.0929, \quad \sigma = 0.6903, \quad \rho = -0.82,$$
$$r = 0.04, \quad S_0 = 100, \quad V_0 = 0.0719, \quad T = 1.$$

The pricing of European options under local and stochastic volatility models using recursive quantization techniques has already been studied, see e.g. Pagès and Sagna (2015) and Callegaro et al. (2015) (Chapter 2) for the local volatility case, and Callegaro et al. (2016) for the stochastic volatility case. However, the method we

present here is more general and is model free compared to Callegaro et al. (2016) (Chapter 3) where the method depends on the structure of the model.
The benchmark price is obtained using a Fourier based approach like in Carr and Madan (1999) , since the Heston model is affine and the characteristic function is known in closed form.
Due to the fact that the derivative only depends on the price and not on the variance, it seems reasonable to choose the marginal grid size N_1 greater than N_2. To guarantee a good balance between precision and computational time, setting $N_1 = 2N_2$ seems to be a good trade-off. As for the previous example, we consider Call prices for the strikes $K \in \{80, 85, 90, 95, 100\}$ and Put prices for strikes K belonging to $\{100, 105, 110, 115, 120\}$ and depict in Table 5.3 and Table 5.4 the corresponding prices (by making varying even n or the sizes N_1 and N_2 of the quantizers) and the associated relative errors $\left(\text{Err.} = \frac{|\text{Price} - MQ^n_{N_1,N_2}|}{\text{Price}}\right)$ between the Fourier prices (Price) and the prices obtained from the Markovian and product quantization method of size $N = N_1 \times N_2$ with n discretization steps (denoted by $\text{MQ}^n_{N_1,N_2}$) and the computation time (in seconds) for the Markovian and product quantization method.
The best choice from the point of view of accuracy and computational effort, is obtained by taking the size N_1 of the quantization of the price process S equal to 20, by taking $N_2 = 10$ for the variance process V, and by setting $n = 20$. The flexibility of this approach would easily allow the pricing of exotic options, as seen in Callegaro et al. (2015) (Chapter 2) and Callegaro et al. (2016) (Chapter 3).

5.3.3 Approximation of BSDE

In this section, we consider a Markovian BSDE

$$Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \tag{5.30}$$

where W is a q-dimensional Brownian motion, $(Z_t)_{t\in[0,T]}$ is a square integrable progressively measurable process taking values in $\mathbb{R}^q$, $f : [0, T]\times\mathbb{R}^d \times\mathbb{R}\times\mathbb{R}^q \to \mathbb{R}$. The terminal condition is of the form $\xi = h(X_T)$, for a given Borel function $h : \mathbb{R}^d \to \mathbb{R}$, where X_T is the value at time T of a Brownian diffusion process $(X_t)_{t\geq 0}$, strong solution to the stochastic differential equation:

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \qquad x \in \mathbb{R}^d. \tag{5.31}$$

As pointed out in the introduction, many (time) discretization schemes and several (spacial) numerical approximation method of the solution of such as BSDE are proposed in the literature (we refer for example to Bally et al. (2001); Bouchard and Touzi (2004); Crisan et al. (2010); Gobet and Turkedjiev (2016); Hu et al. (2011); Gobet et al. (2016); Bender and Denk (2007); Pagès and Sagna (2016)). Our aim in this section is to test the performances of our method to the numerical scheme proposed in Pagès and Sagna (2016). To this end, we first show that the Markovian product quantization method allows us to compute the term appearing in the numerical schemes proposed in Pagès and Sagna (2016) (as well as for several numerical schemes) using (semi)-closed formula. We then test the performance of our method to a BSDE associated to the price of the Call option in the Black-Scholes model and to a multidimensional BSDE.

K	Price	$\mathrm{MQ}^{20}_{10,6}$	Err. (%)	$\mathrm{MQ}^{20}_{20,10}$	Err. (%)	$\mathrm{MQ}^{20}_{30,16}$	Err. (%)
80	26.3910	25.8790	1.9401	26.2684	0.4645	26.3705	0.0777
85	22.6069	22.0686	2.3813	22.4879	0.5264	22.5804	0.1174
90	19.0506	18.5191	2.7897	18.9360	0.6012	19.0478	0.0148
95	15.7524	15.2279	3.3295	15.6471	0.6681	15.7634	0.0700
100	12.7422	12.2750	3.6668	12.6532	0.6987	12.7560	0.1083
100	8.8212	8.3579	5.2515	8.7361	0.9639	8.8390	0.2017
105	10.9308	10.4676	4.2371	10.8690	0.5655	10.9861	0.5056
110	13.3794	13.0481	2.4763	13.3460	0.2497	13.4429	0.4741
115	16.1828	15.8037	2.3425	16.1796	0.0201	16.2862	0.6389
120	19.3456	19.1549	0.9855	19.3697	0.1250	19.4603	0.5929
Time (s)		0.52		4.24		24.82	

TABLE 5.3: Prices in the Heston model. The prices correspond to Call prices for strikes $K \in \{80, 85, 90, 95, 100\}$ and to Put prices for strikes $K \in \{100, 105, 110, 115, 120\}$. The number of time steps $n = 20$. The error Err. $= \frac{|\text{Price} - \text{MQ}^n_{N_1,N_2}|}{\text{Price}}$ corresponds to the relative error between the benchmark price (Price) and Product Markovian quantization price $\text{MQ}^n_{N_1,N_2}$ of sizes $N = N_1 \times N_2$ with n discretization steps.

Explicit numerical scheme for the BSDE

Let us set for $i \in I_k$, $j \in I_{k+1}$,

$$p_k^i = \mathbb{P}(\widehat{X}_k = x_k^i),\ k = 0, \cdots, n$$
$$\text{and} \quad p_k^{ij} = \mathbb{P}(\widehat{X}_{k+1} = x_{k+1}^j | \widehat{X}_k = x_k^i),\ k = 0, \cdots, n-1.$$

Setting $\widehat{Y}_k = \widehat{y}_k(\widehat{X}_k)$, for every $k \in \{0, \cdots, n\}$, the quantized BSDE scheme reads as

$$\begin{cases} \widehat{y}_n(x_n^i) = h(x_n^i) & x_n^i \in \Gamma_n \\ \widehat{y}_k(x_k^i) = \widehat{\alpha}_k(x_k^i) + \Delta_n f\big(t_k, x_k^i, \widehat{\alpha}_k(x_k^i), \widehat{\beta}_k(x_k^i)\big) & x_k^i \in \Gamma_k \end{cases}$$

where for $k = 0, \ldots, n-1$,

$$\widehat{\alpha}_k(x_k^i) = \sum_{j \in I_{k+1}} \widehat{y}_{k+1}(x_{k+1}^j)\, p_k^{ij} \quad \text{and} \quad \widehat{\beta}_k(x_k^i) = \frac{1}{\sqrt{\Delta_n}} \sum_{j \in I_{k+1}} \widehat{y}_{k+1}(x_{k+1}^j)\, \Lambda_k^{ij}, \qquad (5.32)$$

Strike	Price	$MQ^{10}_{20,10}$	Err. (%)	$MQ^{30}_{20,10}$	Err. (%)	$MQ^{40}_{20,10}$	Err. (%)
80	26.3910	26.4632	0.2737	26.1709	0.8339	26.0900	1.1406
85	22.6069	22.7001	0.4121	22.3799	1.0041	22.2960	1.3752
90	19.0506	19.1619	0.5841	18.8229	1.1954	18.7418	1.6207
95	15.7524	15.8851	0.8428	15.5342	1.3851	15.4616	1.8460
100	12.7422	12.9032	1.2635	12.5458	1.5415	12.4863	2.0083
100	8.8212	8.9902	1.9155	8.6274	2.1964	8.5673	2.8784
105	10.9308	11.1520	2.0235	10.7704	1.4670	10.7266	1.8678
110	13.3794	13.6555	2.0633	13.2609	0.8860	13.2347	1.0820
115	16.1828	16.5099	2.0212	16.1090	0.4562	16.0992	0.5165
120	19.3456	19.7143	1.9063	19.3123	0.1717	19.3157	0.1545
Time (s)		2.18		6.69		8.98	

TABLE 5.4: Prices in the Heston model. The prices correspond to Call prices for strikes $K \in \{80, 85, 90, 95, 100\}$ and to Put prices for strikes $K \in \{100, 105, 110, 115, 120\}$. The error Err. $= \frac{|\text{Price} - \text{MQ}^n_{N_1,N_2}|}{\text{Price}}$ corresponds to the relative error between the benchmark price (Price) and Product Markovian quantization price $\text{MQ}^n_{N_1,N_2}$ of sizes $N_1 = 20, N_2 = 20$ with n discretization steps.

with

$$\Lambda_k^{ij} = \mathbb{E}\big(Z_{k+1}\mathbb{1}_{\{\widehat{X}_{k+1}=x_{k+1}^j\}}\big|\widehat{X}_k = x_k^i\big).$$

In the following, we give closed formula for the Λ_k^{ij}'s. We will first suppose that the components of $\overline{X}_k$ are independent, for every $k = 0, \dots, n$.

Proposition 5.3.1. *Suppose that $q = d$ and $\mathcal{E}_k^\ell(x, Z_{k+1}) = \mathcal{E}_k^\ell(x, Z_{k+1}^\ell)$, for every $\ell \in \{1, \dots, d\}$ and $x \in \mathbb{R}^d$. Then*

$$\Lambda_k^{ij,\ell} = \big(\Phi_0'(x_{k+1}^{\ell,j_\ell-}(x_k^i)) - \Phi_0'(x_{k+1}^{\ell,j_\ell+}(x_k^i))\big)\prod_{\ell'\neq\ell}^{d}\big[\Phi_0\big(x_{k+1}^{\ell',j_{\ell'}+}(x_k^i)\big) - \Phi_0\big(x_{k+1}^{\ell',j_{\ell'}-}(x_k^i)\big)\big] \tag{5.33}$$

Proof. See Appendix 5.4.5. □

In the following, we compute the p-th component $\Lambda_k^{ij,p}$ of Λ_k^{ij} in a general setting. Let us set

$$\mathbb{J}_{k,j_\ell}^{0,p}(x) = \Big\{z \in \mathbb{R}, \quad \sqrt{\Delta}\sigma_k^{\ell p}(x)z \in \big(x_{k+1}^{\ell,j_\ell-1/2} - m_k^\ell(x), x_{k+1}^{\ell,j_\ell+1/2} - m_k^\ell(x)\big)\Big\}$$

and

$$\mathbb{L}_k^{0,p}(x) = \Big\{ \ell \in \{1,\dots,d\}, \quad \sum_{p' \neq p} \big(\sigma_k^{\ell p'}(x)\big)^2 = 0 \Big\}.$$

We also set

$$x_{k+1}^{\ell,p,j_\ell -}(x,z) = \frac{x_{k+1}^{\ell,j_\ell - 1/2} - m_k^\ell(x) - \sqrt{\Delta}\sigma_k^{\ell p}(x) z}{\sqrt{\Delta}\Big(\sum_{p' \neq p} \big(\sigma_k^{\ell p'}(x)\big)^2\Big)^{1/2}};\; x_{k+1}^{\ell,p,j_\ell +}(x,z) = \frac{x_{k+1}^{\ell,j_\ell + 1/2} - m_k^\ell(x) - \sqrt{\Delta}\sigma_k^{\ell p}(x) z}{\sqrt{\Delta}\Big(\sum_{p' \neq p} \big(\sigma_k^{\ell p'}(x)\big)^2\Big)^{1/2}}.$$

Proposition 5.3.2. *For every $p \in \{1,\dots,q\}$, the p-th component $\Lambda_k^{ij,p}$ of Λ_k^{ij} reads*

$$\Lambda_k^{ij,p} = \mathbb{E}\,\zeta \prod_{\ell \in \mathbb{L}_k^{0,p}(x_k^i)} \mathbb{1}_{\{\zeta \in \mathbb{J}_{k,j_\ell}^{0,p}(x_k^i)\}} \big(\Phi_0(\alpha_j^p(x_k^i,\zeta)) - \Phi_0(\beta_j^p(x_k^i,\zeta))\big)^+, \quad \zeta \sim \mathcal{N}(0;1) \tag{5.34}$$

(convention: $\prod_{\ell \in \emptyset}(\cdot) = 1$) where for every $x \in \mathbb{R}^d$ and $z \in \mathbb{R}$,

$$\alpha_j^p(x,z) = \sup_{\ell \in \left(\mathbb{L}_k^{0,p}(x)\right)^c} x_{k+1}^{\ell,p,j_\ell -}(x,z) \quad \text{and} \quad \beta_j^p(x,z) = \inf_{\ell \in \left(\mathbb{L}_k^{0,p}(x)\right)^c} x_{k+1}^{\ell,p,j_\ell +}(x,z). \tag{5.35}$$

In particular, if $p \in \{1,\dots,q\}$ and if for every $\ell \in \{1,\dots,d\}$ there exists $p' \neq p$ such that $\sigma_k^{\ell p'}(x) \neq 0$, then,

$$\Lambda_k^{ij,p} = \mathbb{E}\,\zeta \big(\Phi_0(\alpha_j^p(x_k^i,\zeta)) - \Phi_0(\beta_j^p(x_k^i,\zeta))\big)^+. \tag{5.36}$$

Proof. See Appendix 5.4.6 □

Pricing a risk neutral Black-Scholes Call under the historical probability

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. We consider a call option with maturity T and strike K on a stock price $(X_t)_{t\in[0,T]}$ with dynamics

$$dX_t = \mu X_t dt + \sigma X_t dW_t.$$

Considering a self financing portfolio Y_t with φ_t assets and bonds with risk free return r. We know that (see El Karoui et al. (1997)) the portfolio evolves according to the following dynamics:

$$Y_t = Y_T + \int_t^T f(Y_s, Z_s) ds - \int_t^T Z_s dW_s \tag{5.37}$$

where the payoff $Y_T = (X_T - K)^+$, the hedging strategy $Z_t = \sigma\varphi_t X_t$ and $f(y,z) = -ry - \frac{\mu - r}{\sigma} z$. It is clear that the function f is linear with respect to y and z and, it is Lipschitz continuous with $[f]_{\text{Lip}} = \max(r, \frac{\mu - r}{\sigma})$. We perform the numerical tests from the algorithm we propose with the following parameters

$$X_0 = 100, \quad r = 0.1, \quad \mu = 0.2, \quad K = 100, \quad T = 0.5$$

and make varying the volatility σ.

σ	$\widehat{Y}_0$ $(n=20)$	$\widehat{Y}_0$ $(n=40)$	Y_0	$\widehat{Z}_0$ $(n=20)$	$\widehat{Z}_0$ $(n=40)$	Z_0
0.05	04.97	05.01	05.00	04.67	04.58	04.62
0.07	05.23	05.26	05.27	06.04	05.95	05.95
0.10	05.81	05.84	05.85	07.83	07.72	07.71
0.30	10.88	10.89	10.91	19.00	18.91	19.01
0.40	13.56	13.56	13.58	24.91	24.82	24.99
0.50	16.26	16.25	16.26	31.07	30.98	31.24

TABLE 5.5: Call price in the BS model: $N_k = 100$, $\forall k = 1, \ldots, n; n \in \{20, 40\}$. Computational time: < 1 second for $n = 20$ and around 1 second for $n = 40$.

Multidimensional example

We consider the following example due to J.-F. Chassagneux: let $t \in [0, T]$. Set

$$e_t = \exp(W_t^1 + \ldots + W_t^d + t)$$

where W is a d-dimensional Brownian motion. Consider the following BSDE:

$$dX_t = dW_t, \qquad -dY_t = f(t, Y_t, Z_t)dt - Z_t \cdot dW_t, \qquad Y_T = \frac{e_T}{1 + e_T},$$

where $f(t, y, z) = (z_1 + \ldots + z_d)\left(y - \frac{2+d}{2d}\right)$. The solution of this BSDE is given by

$$Y_t = \frac{e_t}{1 + e_t}, \qquad Z_t = \frac{e_t}{(1 + e_t)^2}. \tag{5.38}$$

For the numerical experiments, we put the (regular) time discretization mesh to $n = 20$, with discretization step Δ. We use the uniform dispatching grid allocation and define the quantization $(\widehat{W}_{t_k})_{0 \le k \le n}$ of the Brownian trajectories $(W_{t_k})_{0 \le k \le n}$ from the following recursive procedure

$$\widehat{W}_{t_{k+1}} = \widehat{W}_{t_k} + \sqrt{\Delta}\,\widehat{\varepsilon}, \tag{5.39}$$

$\widehat{W}_0 = 0$ and where $\widehat{\varepsilon}$ is the optimal quantization of the d-dimensional standard Gaussian random variable. We choose $t = 0.5$, $d = 2, 3$, so that $Y_0 = 0.5$ and $Z_0^i = 0.25$, for every $i = 1, \ldots, d$.
Using the Markovian product quantization method we get

1. for $d = 2$, with $N_1 = N_2 = 30$: $\widehat{Y}_0 = 0.504$, $\widehat{Z}_0^1 = \widehat{Z}_0^2 = 0.24$. The computation time is around 4 seconds.
2. for $d = 3$, with $N_1 = N_2 = N_3 = 15$: $\widehat{Y}_0 = 0.547$, $\widehat{Z}_0^1 = \widehat{Z}_0^2 = \widehat{Z}_0^1 = 0.22$. The computation time is around 1 minute.

Remark that the only reason motivating the choice of this example is the fact that the considered backward has an explicit solution. Nevertheless our method works for a general local volatility diffusion process X.

5.4 Appendix

5.4.1 Proof of Proposition 5.2.1

Proof. Let us set $v^{i,j_i+} := x_{k+1}^{i,j_i+1/2}$ and $v^{i,j_i-} = x_{k+1}^{i,j_i-1/2}$. We have

$$\begin{aligned}
\mathbb{P}(\widehat{X}_{k+1} = x_{k+1}^j | \widehat{X}_k = x_k^\ell) &= \mathbb{P}\Big(\bigcap_{i=1}^d \big\{ \widetilde{X}_{k+1}^i \in \big(v^{i,j_i-}, v^{i,j_i+}\big) \big\} \big| \widehat{X}_k = x_k^\ell \Big) \\
&= \mathbb{P}\Big(\bigcap_{i=1}^d \big\{ \mathcal{E}_k^i(x_k^\ell, Z_{k+1}) \in \big(v^{i,j_i-}, v^{i,j_i+}\big) \big\} \Big) \\
&= \mathbb{E}\Big(\mathbb{E}\Big(\mathbf{1}_{\bigcap_{i=1}^d \left\{ \mathcal{E}_k^i(x_k^\ell, Z_{k+1}) \in \left(v^{i,j_i-}, v^{i,j_i+}\right) \right\}} \Big) \big| Z_{k+1}^{(2:d)} \Big) \\
&= \mathbb{E}\big(\Psi(x_k^\ell, Z_{k+1}^{(2:d)}) \big)
\end{aligned}$$

where, for every $u \in \mathbb{R}^{q-1}$,

$$\Psi(x,u) = \mathbb{P}\Big(\bigcap_{i=1}^d \big\{ m_k^i(x_k^\ell) + \sqrt{\Delta}\sigma_k^{i1}(x) Z_{k+1}^1 + \sqrt{\Delta}\big(\sigma_k^{(i,2:q)}(x)|u\big) \in \big(v^{i,j_i-}, v^{i,j_i+}\big) \big\} \Big).$$

Let us set

$$A_{i,k} = \Big\{ m_k^i(x_k^\ell) + \sqrt{\Delta}\sigma_k^{i1}(x) Z_{k+1}^1 + \sqrt{\Delta}\big(\sigma_k^{(i,2:q)}(x)|u\big) \in \big(v^{i,j_i-}, v^{i,j_i+}\big) \Big\}.$$

We know that if $i \in \mathbb{J}_k^0(x)$ then $A_{i,k} = \{u \in \mathbb{J}_{k,j_i}^0(x)\}$ and we deduce that

$$\Psi(x,u) = \prod_{i_0 \in \mathbb{J}_k^0(x)} \mathbf{1}_{\{u \in \mathbb{J}_{k,j_i}^0(x)\}} \mathbb{P}\Big(\big(\bigcap_{i_- \in \mathbb{J}_k^-(x)} A_{i_-,k} \big) \cap \big(\bigcap_{i_+ \in \mathbb{J}_k^+(x)} A_{i_+,k} \big) \Big).$$

Furthermore, notice that if $i_+ \in \mathbb{J}_k^+(x)$ then

$$A_{i_+,k} = \big\{ Z_{k+1}^1 \in (x_{k+1}^{i,j_i-}(x,u), x_{k+1}^{i,j_i+}(x,u)) \big\}$$

and $i_- \in \mathbb{J}_k^-(x)$ then

$$A_{i_-,k} = \big\{ Z_{k+1}^1 \in (x_{k+1}^{i,j_i+}(x,u), x_{k+1}^{i,j_i-}(x,u)) \big\}.$$

It follows that (remark that the sets $\mathbb{J}_k^-(x)$ or $\mathbb{J}_k^+(x)$ may be empty)

$$\begin{aligned}
\mathbb{P}\Big(\big(\bigcap_{i_- \in \mathbb{J}_k^-(x)} A_{i_-,k} \big) \cap \big(\bigcap_{i_+ \in \mathbb{J}_k^+(x)} A_{i_+,k} \big) \Big) = \mathbb{P}\Big(Z_{k+1}^1 \in \big(\sup_{i \in \mathbb{J}_k^+(x)} x_{k+1}^{i,j_i-}(x,u), \inf_{i \in \mathbb{J}_k^+(x)} x_{k+1}^{i,j_i+}(x,u) \big) \\
\cap \big(\sup_{i \in \mathbb{J}_k^-(x)} x_{k+1}^{i,j_i+}(x,u), \inf_{i \in \mathbb{J}_k^-(x)} x_{k+1}^{i,j_i-}(x,u) \big) \Big).
\end{aligned}$$

This completes the proof since $Z_{k+1}^{(2:d)} \sim \mathcal{N}(0; I_{q-1})$. □

5.4.2 Proof of Proposition 5.2.2

Proof. 1. Set $v^{\ell,j_\ell+} := x_{k+1}^{\ell,j_\ell+1/2}$ and $v^{\ell,j_\ell-} = x_{k+1}^{\ell,j_\ell-1/2}$, for $j \in I_{k+1}$ and $\ell = 1,\dots,d$. We have

$$\begin{aligned}\mathbb{P}(\widehat{X}_{k+1} = x_{k+1}^j | \widehat{X}_k = x_k^i) &= \mathbb{P}\Big(\bigcap_{\ell=1}^d \big\{ \tilde{X}_{k+1}^\ell \in \big(v^{\ell,j_\ell-}, v^{\ell,j_\ell+}\big) \big\} \big| \widehat{X}_k = x_k^i \Big) \\ &= \mathbb{P}\Big(\bigcap_{\ell=1}^d \big\{ \mathcal{E}_k^\ell(x_k^i, Z_{k+1}) \in \big(v^{\ell,j_\ell-}, v^{\ell,j_\ell+}\big) \big\} \Big).\end{aligned}$$

Since for every $k = 0,\dots,n-1$, $\sigma(t_k, x)$ is a diagonal matrix, it follows that the operators $\mathcal{E}_k^\ell(x_k^i, Z_{k+1}) = \mathcal{E}_k^\ell(x_k^i, Z_{k+1}^\ell)$, for $\ell = 1,\dots,d$, are independent, so that

$$\begin{aligned}\mathbb{P}(\widehat{X}_{k+1} = x_{k+1}^j | \widehat{X}_k = x_k^i) &= \prod_{\ell=1}^d \mathbb{P}\Big(\mathcal{E}_k^\ell(x_k^i, Z_{k+1}^\ell) \in \big(v^{\ell,j_\ell-}, v^{\ell,j_\ell+}\big) \Big) \\ &= \prod_{\ell=1}^d \big[\Phi_0\big(x_{k+1}^{\ell,j_\ell+}(x_k^i)\big) - \Phi_0\big(x_{k+1}^{\ell,j_\ell-}(x_k^i)\big) \big].\end{aligned}$$

The second assertion immediately follows. □

5.4.3 Proof of Theorem 5.2.3

Proof. 1. For every $k \in \{1,\dots,n-1\}$, for every $\ell = 1,\dots,d$ and for every $j_\ell = 1,\dots,N_{k+1}^\ell$, we have

$$\begin{aligned}\mathbb{P}\big(\tilde{X}_{k+1}^\ell \in C_{j_\ell}(\Gamma_{k+1}^\ell) | \widehat{X}_k = x_k^i\big) &= \mathbb{P}\big(\tilde{X}_{k+1}^\ell \le x_{k+1}^{\ell,j_\ell+1/2} | \widehat{X}_k = x_k^i\big) - \mathbb{P}\big(\tilde{X}_{k+1}^\ell \le x_{k+1}^{\ell,j_\ell-1/2} | \widehat{X}_k = x_k^i\big) \\ &= \mathbb{P}\big(\mathcal{E}_k^\ell(x_k^i, Z_{k+1}) \le x_{k+1}^{\ell,j_\ell+1/2}\big) - \mathbb{P}\big(\mathcal{E}_k^\ell(x_k^i, Z_{k+1}) \le x_{k+1}^{\ell,j_\ell-1/2}\big).\end{aligned}$$

To complete the proof we just have to remark that $\mathcal{E}_k^\ell(x_k^i, Z_{k+1}) \sim \mathcal{N}\big(m_k^\ell(x_k^i); \Delta |\sigma_k^{\ell\bullet}(x_k^i)|_2^2\big)$. □

5.4.4 Proof of Theorem 5.2.4

Proof. Recall that for every $k \ge 0$, $\widehat{X}_k = (\widehat{X}_k^1, \dots, \widehat{X}_k^d)$, where $\widehat{X}_k^\ell$ is the quantization of the ℓ-th component $\overline{X}_k^\ell$ of the vector $\overline{X}_k$. Therefore, following step by step the proof of Lemma 3.2. in Pagès and Sagna (2015), we obtain for every $k \ge 1$:

$$\|\overline{X}_k - \widehat{X}_k\|_2 \le \sum_{k'=1}^k e^{(t_k - t_{k'})C_{b,\sigma}} \|\tilde{X}_{k'} - \widehat{X}_{k'}^{\Gamma_{k'}}\|_2,$$

where $C_{b,\sigma} = [b]_{\mathrm{Lip}} + \frac{1}{2}[\sigma]_{\mathrm{Lip}}^2$. Using the definition of $\widehat{X}_k$ combined with Pierce's Lemma (see Theorem 1.1.1(b)) yields for every $k = 1,\dots,n$, for any $\eta \in (0,1]$,

$$\|\overline{X}_k - \widehat{X}_k\|_2 \le K_{2,\eta} \sum_{k'=1}^k e^{(k-k')\Delta C_{b,\sigma}} \|\tilde{X}_{k'}^\ell\|_{2+\eta} (N_{k'}^\ell)^{-1/d}.$$

Recall that each component $\widehat{X}_k^\ell$ of the vector $\widehat{X}_k$ is defined as

$$\widehat{X}_k^\ell = \widehat{\tilde{X}_k^\ell},\ \ell = 1,\dots,d,$$

where $\widehat{\tilde{X}_k^\ell}$ is an optimal quadratic quantization of $\tilde{X}_k^\ell$. Hence each component of $\widehat{X}_k$ is stationary with respect to $\tilde{X}_k^\ell$, that is $\widehat{X}_k = \mathbb{E}(\tilde{X}_k^\ell \,|\, \widehat{\tilde{X}_k^\ell}) = \widehat{\tilde{X}_k^\ell}$. We deduce that for every $p \in (2,3]$,

$$\mathbb{E}|\widehat{X}_k|^p \le d^{\frac{p}{2}-1}\mathbb{E}|\tilde{X}_k|^p.$$

Following the lines of the proof of Lemma 3.2 in Pagès and Sagna (2015), we easily show that for every $\ell \in \{1,\dots,d\}$, $\|\tilde{X}_{k'}^\ell\|_{2+\eta} \le a_{k'}(b,\sigma,k,d,\Delta,x_0,L,2+\eta)$, for $\eta \in (0,1]$. This completes the proof. □

5.4.5 Proof of Proposition 5.3.1

Proof. Let us set $v^{\ell,j_\ell+} := x_{k+1}^{\ell,j_\ell+1/2}$ and $v^{\ell,j_\ell-} = x_{k+1}^{\ell,j_\ell-1/2}$. We have

$$\begin{aligned}
\Lambda_k^{ij} &= \mathbb{E}\big(Z_{k+1}\mathbb{1}_{\{\mathcal{E}_k(\widehat{X}_k,Z_{k+1})\in C_j(\Gamma_{k+1})\}}\big|\widehat{X}_k = x_k^i\big)\\
&= \mathbb{E}\big(Z_{k+1}\mathbb{1}_{\{\mathcal{E}_k(x_k^i,Z_{k+1})\in C_j(\Gamma_{k+1})\}}\big)\\
&= \mathbb{E}\Big(Z_{k+1}\prod_{\ell'=1}^{d}\mathbb{1}_{\left\{\mathcal{E}_k^{\ell'}(x_k^i,Z_{k+1}^{\ell'})\in\left(v^{\ell',j_{\ell'}-},v^{\ell,j_{\ell'}+}\right)\right\}}\Big).
\end{aligned}$$

Since the components of Z_{k+1} are independent, it follows that for $\ell = 1,\dots,d$, the component (i_ℓ, j) of Λ_k^{ij} reads

$$\begin{aligned}
\Lambda_k^{ij,\ell} &= \mathbb{E}\Big(Z_{k+1}^\ell\prod_{\ell'=1}^{d}\mathbb{1}_{\left\{\mathcal{E}_k^{\ell'}(x_k^i,Z_{k+1}^{\ell'})\in\left(v^{\ell',j_{\ell'}-},v^{\ell,j_{\ell'}+}\right)\right\}}\Big)\\
&= \mathbb{E}\Big(Z_{k+1}^\ell\mathbb{1}_{\left\{\mathcal{E}_k^{\ell}(x_k^i,Z_{k+1}^{\ell})\in\left(v^{\ell',j_{\ell}-},v^{\ell,j_{\ell}+}\right)\right\}}\Big)\times\mathbb{E}\Big(\prod_{\ell'\neq\ell}^{d}\mathbb{1}_{\left\{\mathcal{E}_k^{\ell'}(x_k^i,Z_{k+1}^{\ell'})\in\left(v^{\ell',j_{\ell'}-},v^{\ell',j_{\ell'}+}\right)\right\}}\Big).
\end{aligned}$$

It is clear that

$$\mathbb{E}\Big(\prod_{\ell'\neq\ell}^{d}\mathbb{1}_{\left\{\mathcal{E}_k^{\ell'}(x_k^i,Z_{k+1}^{\ell'})\in\left(v^{\ell',j_{\ell'}-},v^{\ell',j_{\ell'}+}\right)\right\}}\Big) = \prod_{\ell'\neq\ell}^{d}\left[\Phi_0\big(x_{k+1}^{\ell',j_{\ell'}+}(x_k^i,0)\big) - \Phi_0\big(x_{k+1}^{\ell',j_{\ell'}-}(x_k^i,0)\big)\right].$$

On this other hand,

$$\mathbb{E}\Big(Z_{k+1}^\ell\mathbb{1}_{\left\{\mathcal{E}_k^{\ell}(x_k^i,Z_{k+1}^{\ell})\in\left(v^{\ell,j_{\ell}-},v^{\ell,j_{\ell}+}\right)\right\}}\Big) = \Phi_0'(x_{k+1}^{\ell,j_\ell-}(x_k^i,0)) - \Phi_0'(x_{k+1}^{\ell,j_\ell+}(x_k^i,0)).$$

Combining both previous equalities gives the announced result. □

5.4.6 Proof of Proposition 5.3.2

Proof. Let us set $v^{\ell,j_\ell+} := x_{k+1}^{\ell,j_\ell+1/2}$ and $v^{\ell,j_\ell-} = x_{k+1}^{\ell,j_\ell-1/2}$. We have

$$\begin{aligned}
\Lambda_k^{ij} &= \mathbb{E}\big(Z_{k+1}\mathbb{1}_{\{\mathcal{E}_k(\widehat{X}_k,Z_{k+1})\in C_j(\Gamma_{k+1})\}}\big|\widehat{X}_k = x_k^i\big)\\
&= \mathbb{E}\big(Z_{k+1}\mathbb{1}_{\{\mathcal{E}_k(x_k^i,Z_{k+1})\in C_j(\Gamma_{k+1})\}}\big)\\
&= \mathbb{E}\big(Z_{k+1}\mathbb{1}_{A_j^i}\big)
\end{aligned}$$

where

$$A_j^i = \bigcap_{\ell=1}^{d} \{\mathcal{E}_k^\ell(x_k^i, Z_{k+1}) \in (v^{\ell,j_\ell-}, v^{\ell,j_\ell+})\}$$
$$= \bigcap_{\ell=1}^{d} \{m_k^\ell(x_k^i) + \sqrt{\Delta}\sigma_k^{\ell p}(x_k^i)Z_{k+1}^p + \sum_{p'\neq p} \sqrt{\Delta}\sigma_k^{\ell p'}(x_k^i)Z_{k+1}^{p'} \in (v^{\ell,j_\ell-}, v^{\ell,j_\ell+})\}.$$

Then, conditioning by Z_{k+1}^p shows that the component $\Lambda_k^{ij,p}$ of Λ_k^{ij} reads

$$\Lambda_k^{ij,p} = \mathbb{E}\big(\mathbb{E}\big(Z_{k+1}^p \mathbb{1}_{A_{j_\ell}^i}\big)\big|Z_{k+1}^p\big) = \mathbb{E}(\Psi(\zeta)), \qquad \zeta \sim \mathcal{N}(0;1),$$

where for every u,

$$\Psi(u) = u\,\mathbb{P}\Big(\bigcap_{\ell=1}^{d} A_{\ell,k}^p(u)\Big)$$

with

$$A_{\ell,k}^p(u) = \{m_k^\ell(x_k^i) + \sqrt{\Delta}\sigma_k^{\ell p}(x_k^i)u + \sum_{p'\neq p} \sqrt{\Delta}\sigma_k^{\ell p'}(x_k^i)Z_{k+1}^{p'} \in (v^{\ell,j_\ell-}, v^{\ell,j_\ell+})\}.$$

Keep in mind that

$$\sum_{p'\neq p} \sqrt{\Delta}\sigma_k^{\ell p'}(x_k^i)Z_{k+1}^{p'} \overset{\mathcal{L}}{=} \Big(\Delta \sum_{p'\neq p} \big(\sigma_k^{\ell p'}(x_k^i)\big)^2\Big)^{1/2} Z, \quad Z \sim \mathcal{N}(0;1).$$

Then, we may write

$$\Psi(u) = u\,\mathbb{1}_{\big\{\bigcap_{\ell_0\in\mathbb{L}_k^{0,p}(x_k^i)} A_{\ell_0,k}^p(u)\big\}}\mathbb{P}\Big(\bigcap_{\ell_+\in\big(\mathbb{L}_k^{0,p}(x_k^i)\big)^c} A_{\ell_+,k}^p(u)\Big)$$

with

$$A_{\ell_+,k}^p(u) = \big\{Z \in (x_{k+1}^{\ell,p,j_\ell-}(x_k^i,u), x_{k+1}^{\ell,p,j_\ell+}(x_k^i,u))\big\}, \quad Z \sim \mathcal{N}(0;1).$$

It follows that

$$\Psi(u) = u\,\mathbb{1}_{\big\{\bigcap_{\ell_0\in\mathbb{L}_k^{0,p}(x_k^i)} A_{\ell_0,k}^p(u)\big\}}\mathbb{P}\Big(\bigcap_{\ell_+\in\big(\mathbb{L}_k^{0,p}(x_k^i)\big)^c} A_{\ell_+,k}^p(u)\Big)$$
$$= u \prod_{\ell\in\mathbb{L}_k^{0,p}(x_k^i)} \mathbb{1}_{\big\{u\in\mathbb{J}_{k,j_\ell}^{0,p}(x_k^i)\big\}}\mathbb{P}\big(Z \in \alpha_j^p(x_k^i,u), \beta_j^p(x_k^i,u)\big).$$

The result follows immediately. □

Chapter 6

Affine models and beyond

This is a joint work with Giorgia Callegaro and Martino Grasselli. The content of this chapter has been submitted as a paper, see Callegaro et al. (2017b) in the bibliography.

In this chapter we introduce a novel pricing methodology for a broad class of models for which the characteristic function of the log-asset price can be efficiently computed. The new method avoids the numerical integration required by the Fourier-based approaches and reveals to be fast and accurate, to the point that we can calibrate the models on real data. Our approach allows to price also American-style options, as it is possible to compute the transition probabilities for the underlying. This is accomplished through an efficient multinomial lattice discretization of the asset price based on a new quantization procedure which exploits the knowledge of the Fourier transform of the process at a given time. As a motivating example, we price an American Put option in a Tempered Stable model, with constitutes the first application of quantization to a pure jump process.

6.1 Introduction

A necessary feature in order to be able to reproduce the smile and skew typically observed in the option market consists in allowing the volatility to vary stochastically. The problem of pricing derivatives in a stochastic volatility framework has been deeply investigated in the literature. Starting from the pioneering works of Hull and White (1987), Chesney and Scott (1989), Stein and Stein (1991) and Heston (1993), researchers introduced dozens of models, with continuous trajectories or including jumps, with different levels of analytical tractability. Both models of Stein and Stein (1991) and Heston (1993) belong to the affine class, for which the log-asset price admits an explicit characteristic function in terms of the solution of ordinary differential equations of Riccati type, so that options prices can be obtained efficiently through the computation of a numerical integral, for which many ad-hoc algorithms have been introduced, like e.g. the celebrated Fast Fourier Transform (FFT), a method well known since long time ago by engineers and introduced in quantitative finance by Carr and Madan (1999). Duffie et al. (2000) developed a systematic approach in order to perform the Fourier methodology for affine models, including the presence of jumps in both the underlying and the volatility. Many affine models, including jumps, have been indeed calibrated in both a single and a multi factor setting for the volatility process, like e.g. Bates (1996), Jacobs and Li (2008), Christoffersen et al. (2009) for the Heston based models and more recently also Da Fonseca and Grasselli (2011) under the multi-dimensional matrix-valued

(Wishart) specification of the volatility process. Also, some non affine models admitting an explicit characteristic function of the log-asset have been introduced in literature. For example, the 3/2 model of Heston (1997) and the 4/2 model of Grasselli (2016) consider specifications for the volatility process involving non linear transformations of the square root process, thus breaking the affine property. These models are useful as they are able to capture interesting features, like the strict positivity of the volatility and a mean reversion depending on the level of the volatility. Interestingly, for these models one can still apply the Fourier technique. Another important example of non affine model admitting explicit characteristic function of the log-asset is the Variance Gamma model introduced by Madan et al. (1998), which belongs to the more general class of Tempered Stable models introduced by Cont et al. (1997), see also Eberlein et al. (2010).

While the Fast Fourier Transform methodology for pricing vanillas is extremely efficient, it is noteworthy that in the case of exotic derivatives the pricing problem still represents a challenge, and Monte Carlo simulation is often the only possibility, despite its potentially high computational cost. What is more, for American-style options the Fourier approach is almost useless.

Starting from the results in Pagès and Sagna (2015) and Callegaro et al. (2015) (Chapter 2), Callegaro et al. (2016) (Chapter 3), Callegaro et al. (2017a) (Chapter 4) and Fiorin et al. (2015) (Chapter 5) extended the methodology to the case where the volatility is stochastic. The passage from local to stochastic volatility models requires additional care, from both the theoretical and the numerical point of view. For example, a direct application of the argument presented in Pagès and Sagna (2015) is not possible in the multi dimensional case. The approach in Callegaro et al. (2016) (Chapter 3), Callegaro et al. (2017a) (Chapter 4) and Fiorin et al. (2015) (Chapter 5) only requires that an Euler scheme is available for the stochastic model, which is the case for most popular stochastic volatility models. A further extension of the method to higher order discretization schemes has been recently investigated by McWalter et al. (2017).

The approaches considered above can be applied in principle to a broad class of stochastic volatility models, even non affine, but they do not lead to an efficient pricing methodology in terms of computational time, so that the calibration becomes a difficult task. What is more, the quantization of an Euler scheme (or even a more sophisticated higher order scheme like in McWalter et al. (2017)) leads to a bias in the approximation of the true probability density of the underlying at any time step of the time discretization.

The chapter is organized as follows: in Section 6.2 we recall some basic notions on quantization, with a particular focus on the recent recursive marginal quantization, and we illustrate our new methodology. We also analyze the error generated by the quantization and we provide some ex-ante estimates, thus providing the mathematical foundation to the whole procedure. In Section 6.3 we apply the methodology to six stochastic volatility and pure jump models. We show the performance of the Fourier-quantization methodology in the pricing of vanillas and we calibrate one of them (a Tempered Stable model) on a book of options on the DAX index. We also go beyond vanillas by pricing an American Put option in a Tempered Stable model. Section 6.4 concludes, while we gather in Appendix 6.5.1, 6.5.2 and 6.5.3 the technical proofs.

6.2 The Fourier Based Quantization

Consider a random vector X. The aim of this section is to find an optimal discretization of X, in a sense that will be made precise in what follows, by means of a vector taking a finite number of values.
Let us introduce a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ where the filtration satisfies the usual hypotheses and consider the (positive) value of an asset $S = (S_t)_{t\in[0,T]}$ at time $T > 0$, and take $X = S_T$, namely $d = 1$. Suppose now that the conditional characteristic function $\phi_{t,T}(u)$ of $\log(S_T)$, for $t \in [0,T]$,

$$\phi_{t,T}(u) := \mathbb{E}\left[e^{iu\log(S_T)}\middle|\mathcal{F}_t\right], \quad u \in \mathbb{R}$$

is explicitly known, or at least it can be computed efficiently. We will denote by $\phi_T(u) := \phi_{0,T}(u)$.
Standard arguments, see e.g. Shephard (1991), show that the density and the distribution of S_T can be recovered by the knowledge of ϕ_T through Fourier inversion:

$$\mathbb{P}(S_T \in dz) = \left(\frac{1}{\pi}\frac{1}{z}\int_0^\infty Re\left(e^{-iu\log(z)}\phi_T(u)\right)du\right)dz =: f(z)dz, \tag{6.1}$$

and

$$\mathbb{P}(S_T \le z) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty Re\left(\frac{e^{-iu\log(z)}\phi_T(u)}{iu}\right)du, \quad z \in (0,+\infty). \tag{6.2}$$

Here is where quantization meets Fourier analysis: we are going to express the density of the random variable S_T appearing in the distortion function in terms of the characteristic function through the Fourier inversion formula above. This will allow us to quantize efficiently all random variables admitting a characteristic function in closed form. This is a new contribution in the theory of quantization of stochastic processes at a fixed time.
In order to do that, we proceed as follows:

i) we express the distortion function in terms of the Fourier representation of the price density;

ii) we differentiate the distortion function and we notice the tridiagonal structure of its Hessian;

iii) we provide the *Master Equation* that has to be solved in order to find the stationary quantizers.

6.2.1 The Distortion Function and the Master Equation

First of all, we recall that in the case $d = 1$ the Voronoi partition consists of simple intervals. That is, given a quantization grid $\Gamma = \{x_1, \dots, x_N\}$ of size N, a single Voronoi cell $C_j(\Gamma)$ takes the form $C_j(\Gamma) = \left[x_j^-, x_j^+\right]$, where

$$x_j^- = \frac{1}{2}\left(x_{j-1} + x_j\right) \quad \forall j = 2, \dots, N, \qquad x_j^+ = \frac{1}{2}\left(x_j + x_{j+1}\right) \quad \forall j = 1, \dots, N-1.$$

Note that, since we consider positive processes (hence random variables), we have $x_1^- = 0$ and $x_N^+ = +\infty$.
Let us denote by $\mathbb{P}_{S_T}$ the distribution of S_T. Exploiting the discrete nature of the N-quantizer $q(S_T)$, and using the fact that it is constant over the Voronoi cells $C_j(\Gamma)$,

we can equivalently write the distortion function D_p in Definition 1.1.1 as

$$D_p(\Gamma) = D_p(x_1, \dots, x_N) = \sum_{j=1}^{N} \int_{C_j(\Gamma)} |z - x_j|^p d\mathbb{P}_{S_T}(z). \tag{6.3}$$

As the density of the random variable S_T is assumed to be continuous and concentrated on $(0, +\infty)$, we shall differentiate the distortion function inside the integral in order to find the Master Equation satisfied by the stationary quantizers.

Theorem 6.2.1 (The Master Equation). Let $\Gamma = \{x_1, \dots, x_N\}$ be a quantization grid of size N and $p \in (1, +\infty)$. Then Γ is p-stationary for the distortion function (6.3) if for all $j = 1, ..., N$:

$$\int_0^{+\infty} Re\left[\phi_T(u) e^{-iu\log(x_j)} \left(\bar{\beta}\left(\frac{x_j^-}{x_j}, -iu, p\right) - \bar{\beta}\left(\frac{x_j}{x_j^+}, 1 - p + iu, p\right)\right)\right] du = 0, \tag{6.4}$$

where for $a \in \mathbb{C}, Re(b) > 0$ and $x \in (0, 1)$, the function $\bar{\beta}$ is defined as

$$\bar{\beta}(x, a, b) := \int_x^1 t^{a-1}(1-t)^{b-1} dt^{1}. \tag{6.5}$$

Proof. See Appendix 6.5.1 □

6.2.2 The Newton−Rapshom Algorithm

The problem of finding the $p-$stationary grid Γ for a generic random variable S_T boils down to solving the nonlinear system of equations (6.4). As far as the derivatives of the Master Equation can be computed analytically, like in our case, the most efficient way consists in implementing a Newton-Raphson algorithm, as suggested in Pagès and Sagna (2015), Callegaro et al. (2015) (Chapter 2), Callegaro et al. (2016) (Chapter 3), Callegaro et al. (2017a) (Chapter 4), Fiorin et al. (2015) (Chapter 5) and McWalter et al. (2017).
Let us denote with L_j the derivative with respect to x_j of the distortion function (6.3), that is:

$$L_j(\Gamma) := \frac{p}{\pi} x_j^{p-1} \int_0^{+\infty} Re\left[\phi_T(u) e^{-iu\log(x_j)} \left(\bar{\beta}\left(\frac{x_j^-}{x_j}, -iu, p\right) - \bar{\beta}\left(\frac{x_j}{x_j^+}, 1 - p + iu, p\right)\right)\right] du. \tag{6.6}$$

Finding the stationary quantizers consists in solving the N-dimensional system $L(\Gamma) = 0$, which is equivalent to (6.4). Now, the Newton Raphson method needs the computation of the $N \times N$ Jacobian matrix ∇L of the function L. Notice that L_j depends

[1] The function $\bar{\beta}(x, a, b)$ can be written as

$$\bar{\beta}(x, a, b) = \beta(a, b) - \beta(x, a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt - \int_0^x t^{a-1}(1-t)^{b-1} dt,$$

where $\beta(a, b)$ (resp. $\beta(x, a, b)$) is the complete (resp. incomplete) Euler beta function, see e.g. Abramowitz and Stegun (1970). Therefore, $\bar{\beta}(x, a, b)$ can be denoted as complementary incomplete Euler beta function. The complete Euler beta function is defined only when $Re(a) > 0, Re(b) > 0$, because otherwise the integral would diverge. Nevertheless, its definition can be extended, by regularization, to negative values of $Re(a)$ and $Re(b)$, see Ozcag et al. (2008) for a comprehensive study of these special functions. From an implementation perspective, softwares like e.g. Wolfram System already include a regularized version of the Beta and incomplete Beta functions defined for most arguments, by taking into account singular cases.

only on x_{j-1}, x_j and x_{j+1} (when $j = 1$ the function L_j depends just on x_1 and x_2, while for $j = N$ it depends on x_{N-1} and x_N), so that the Jacobian of L turns out to be tridiagonal.

Theorem 6.2.2 (The Fourier Quantization Algorithm). Let $\Gamma^{(0)} = \{x_1, \dots, x_N\}$ be a quantization grid of size N and $p \in (1, +\infty)$. The Newton−Raphson algorithm for the determination of the p-stationary quantizers takes the following recursive form:

$$\Gamma^{(n+1)} = \Gamma^{(n)} - \left(\nabla L(\Gamma^{(n)})\right)^{-1} \cdot L(\Gamma^{(n)}), \quad n = 0, 1, ...$$

where the components of the $N-$dimensional vector $L(\cdot)$ are given by (6.6) and the tridiagonal $N \times N$ matrix $\nabla L(\cdot)$ is as follows:

$$\nabla L_{j,j+1}(\Gamma) = -\frac{p}{2\pi}\frac{1}{x_j^+}\left(\frac{x_{j+1} - x_j}{2}\right)^{p-1} \int_0^{+\infty} Re\left[\phi_T(u)e^{-iu\log(x_j^+)}\right] du, \qquad j = 1, ..., N-1; \tag{6.7}$$

$$\nabla L_{j,j-1}(\Gamma) = -\frac{p}{2\pi}\frac{1}{x_j^-}\left(\frac{x_j - x_{j-1}}{2}\right)^{p-1} \int_0^{+\infty} Re\left[\phi_T(u)e^{-iu\log(x_j^-)}\right] du, \qquad j = 2, ..., N; \tag{6.8}$$

$$\begin{aligned}\nabla L_{j,j}(\Gamma) = &\frac{p(p-1)}{\pi} x_j^{p-2} \int_0^{+\infty} Re\left[\phi_T(u)e^{-iu\log(x_j)}\left(\bar{\beta}\left(\frac{x_j^-}{x_j}, -iu, p-1\right)\right.\right. \\ &\left.\left. + \bar{\beta}\left(\frac{x_j}{x_j^+}, 2 - p + iu, p-1\right)\right)\right] du + \nabla L_{j,j+1}(\Gamma) + \nabla L_{j,j-1}(\Gamma), \qquad j = 1, ..., N.\end{aligned} \tag{6.9}$$

Proof. See Appendix 6.5.2 □

Remark 6.2.1. At the generic step k, the algorithm needs the computations of the $N-$dimensional vectors $L(\Gamma^{(n)})$ and $\nabla L_{j,j}(\Gamma^{(n)})$ and the $(N-1)-$dimensional vectors associated to the upper and lower diagonal of the Jacobian. In addition, the algorithm requires the inversion of a tridiagonal matrix, that can be performed efficiently using for example an LU decomposition, at a very small computational cost.

6.2.3 Transition Probabilities

In view of the pricing of American (and possibly more general path dependent) options, the study of the transition probability of a process $(X_t)_{t\in[0,T]}$ between two time steps is fundamental. For a fixed maturity $T > 0$, let us define a discretization mesh $t_k = k\Delta, k = 0, \dots, M$, where $\Delta = \frac{T}{M}$. Let us assume that we have computed the optimal quantization grids of size N for X_{t_k}, namely $\Gamma_k = \{x_1^k, \dots, x_N^k\}$, and for $X_{t_{k+1}}$, namely $\Gamma_{k+1} = \{x_1^{k+1}, \dots, x_N^{k+1}\}$[2]. We denote with $\widehat{X}_k$ (resp. $\widehat{X}_{k+1}$) the optimal quantizer of X_{t_k} (resp. $X_{t_{k+1}}$). We are interested in computing for $k = 0, ..., M$, the following transition probability between the quantization grid Γ_k of the process X at time t_k and the quantization grid Γ_{k+1} at time t_{k+1}:

$$\widehat{\pi}_{j\ell}^k := \mathbb{P}\left(\widehat{X}_{k+1} = x_\ell^{k+1} \middle| \widehat{X}_k = x_j^k\right) = \mathbb{P}\left(X_{t_{k+1}} \in C_\ell\left(\Gamma_{k+1}\right) \middle| X_{t_k} \in C_j\left(\Gamma_k\right)\right).$$

That is, we are reducing the dynamics of the process X to a multinomial lattice with a constant number of nodes N, and we are interested in the transition probability

[2]Note that the size N of the grids could depend on time, that is $N = N^k$, but here we will consider the case of a constant N for the sake of simplicity.

between two subsequent nodes. An immediate application of the Bayes' rule gives the following result.

Lemma 6.2.3. The transition matrix $(\widehat{\pi}^k_{j\ell})_{j\ell}$ takes the following form:

$$\widehat{\pi}^k_{j\ell} = \frac{1}{\mathbb{P}\left(\widehat{X}_k = x^k_j\right)} \int_{C_j(\Gamma_k)} \mathbb{P}\left(X_{t_{k+1}} \in C_\ell\left(\Gamma_{k+1}\right) \middle| X_{t_k} = z\right) \mathbb{P}\left(X_{t_k} \in dz\right). \tag{6.10}$$

Note that, when $(X_t)_{t\in[0,T]}$ is a process of dimension one with characteristic function $\phi_{t,T}(u)$, it is possible to compute the elements in (6.10) using the Fourier inversion formulas is (6.1) and (6.2).
In addition to this result, one can also prove that the multinomial lattice obtained by quantizing at each time step the price process is actually a Markov chain. In fact, using the arguments of (Pagès et al., 2003, Section 3), we get directly the following

Proposition 6.2.4. Let X be a Markov process on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. For $k = 0, ..., M$, let $\widehat{X}_k$ be the optimal quantization of the process X at time t_k, and $\Gamma_k = \{x^k_1, \ldots, x^k_N\}$ the corresponding optimal quantization grid. Then the process $\left(\widehat{X}_k\right)_{k\in\{0,\ldots,M\}}$, equipped with the transition probabilities in Equation (6.10), is a discrete-time Markov chain.

6.2.4 A Priori Estimates of the Quantization Error

Optimal $p-$quantizers minimize the quantization error, while *stationary* quantizers are just stationary points of the distortion function. As a consequence, optimal quantizers are stationary, but the converse is not true in general[3]. It is therefore important to provide some *a priori* estimates for the quantization error. In this subsection we present two different results for the asymptotic behavior of the L^p-mean quantization error $e_{p,N}(S_T, \Gamma)$. The first result is a celebrated lemma due to Pierce (1970), who gives an *a priori* estimation of the error depending on the $(p+\delta) - th$ moment of the random variable S_T.

Lemma 6.2.5. (Pierce Lemma, 1970) Suppose $\mathbb{E}\left(|S_T|^{p+\delta}\right) < +\infty$ for some $\delta > 0$. Then the L^p-mean quantization error satisfies:

$$\left(Ne_{p,N}(S_T, \Gamma)\right)^p \leq \left(C_1 \mathbb{E}\left(|S_T|^{p+\delta}\right) + C_2\right), \quad \delta > 0, N \geq C_3, \tag{6.11}$$

for some constants $C_1, C_2, C_3 > 0$ depending on δ and p, but not on the distribution of S_T and on the size N of the quantization grid Γ.

Notice that the quality of the error estimate in the Pierce Lemma depends on the level of regularity (integrability) of the underlying. In a stochastic volatility framework, some typical phenomena like moment explosion may jeopardize the usefulness of the result as the constants C_1, C_2, C_3 may depend on δ. We now present a result in the analysis of the convergence of the quantization error where the constants do not depend on the moment explosion.

Theorem 6.2.6. Let us assume that S_T has a finite $p-$th moment, and let assume that its density f (as in Equation (6.1)) has a polynomial behavior at 0 and $+\infty$. Then the

[3] An important exception is the case when the density is log-concave, for which the Fourier quantization algorithm allows to determine quantizers that are also optimal, see Remark 1.1.1.

L^p-mean quantization error $e_{p,N}(S_T,\Gamma)$ has the following asymptotic behavior:

$$\lim_{N\to+\infty} Ne_{p,N}(S_T,\Gamma) \leq \frac{1}{2}\left(\frac{||f||_{\frac{1}{p+1}}}{(p+1)}\right)^{\frac{1}{p}}.$$

Proof. See Appendix 6.5.3. □

Remark 6.2.2. The assumption of Theorem 6.2.6 is slightly stronger then the hypothesis of the Pierce Lemma 6.2.5. In fact, the finiteness of the $p-th$ moment and the polynomial behavior at infinity together imply the existence of the $(p+\delta)-th$ moment, for δ small enough. Nevertheless, it is important to notice that this result gives an estimation which depends on the distribution of the stochastic process at time T, and not on the value of δ as in the Pierce Lemma. For the asymptotic behavior of some density functions in a stochastic volatility framework, see for example Gulisashvili and Stein (2010).

6.2.5 The Role of the Parameter p

In this subsection we briefly discuss the behavior of the stationary quantizers according to the choice of the parameter p in the L^p-distortion function. The usual choice in the literature is $p=2$, namely when one looks for *quadratic* optimal quantizer. We present a simple example based on a random variable with known density, namely the standard Gaussian $\mathcal{N}(0,1)$, in order to get an insight about the behavior of the quantization grids.

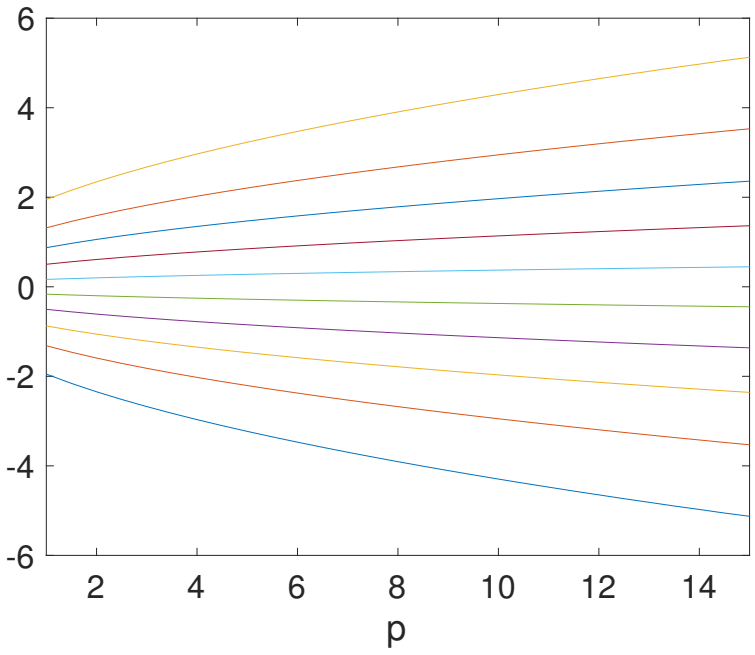


FIGURE 6.1: Plot of the p-optimal quantizers for a $\mathcal{N}(0,1)$ random variable with a fixed number of points for the quantization grids $N=10$ for some value of the parameter p ranging between 1 and 15.

From Figure 6.1 we see that as p increases, for N fixed and equal to $N=10$, the points in the quantization grid are more and more sparse around the origin. In other

words, the choice of p depends on whether we are interested in the tails of the distribution. In this sense, higher values for p are related to more conservative distortion functions.

The choice of p affects also the behavior of the distortion function when considering the number of points N in the grid. From Figure 6.2(a) it follows that the (log-)distortion function decreases more and more quickly as N increases. Figure 6.2(b) illustrates the rate of convergence α_p of the distortion function when expressed as $D_p(\Gamma) = \frac{C}{N^{\alpha_p}}$ in terms of the number of points of the grid N. We see that as $N \to +\infty$ the rate α_p tends to the corresponding value of the parameter p, in line with the asymptotic result in Theorem 6.2.6 and Lemma 6.2.5.

6.3 Numerical Results

6.3.1 Structure of the Numerical Section

This section is devoted to applications of our new quantization technique to option pricing. In Subsection 6.3.2, we consider six models for which the characteristic function of the asset price can be computed efficiently. These models include affine and non affine jump-diffusion processes, where jumps are driven by Poisson and more general Lévy processes, like in the tempered stable model. We will also consider a bi-dimensional specification of the Heston (1993) model in order to emphasize that our methodology can deal with a multi factor stochastic volatility.

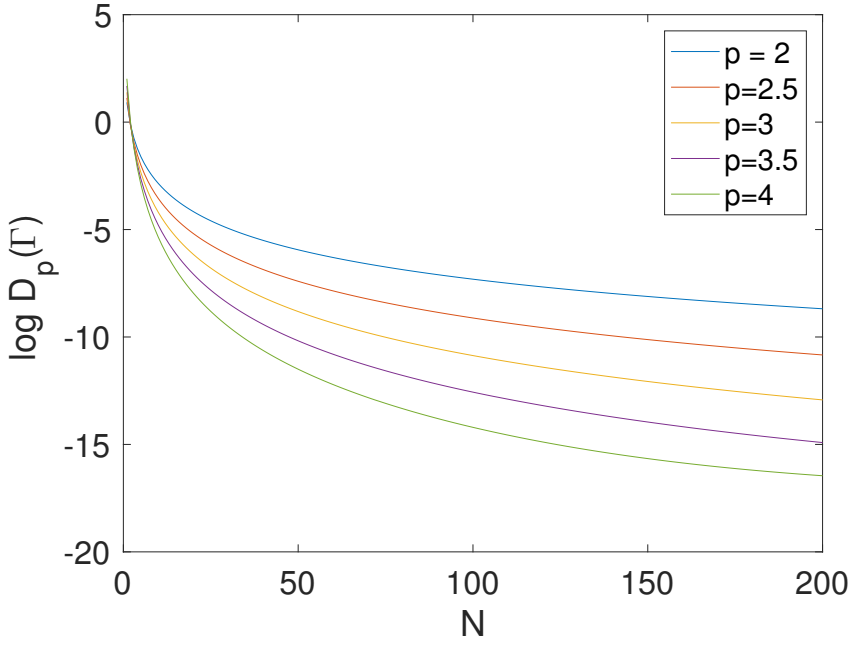
As it is classical in the literature, first of all we will show in Subsection 6.3.3 the numerical performance of our methodology in pricing vanilla options. Pricing vanillas will be revealed to be very efficient, to the point that it will be possible to calibrate the models using real data. In Subsection 6.3.4 we will provide an example of calibration for the Tempered Stable model[4] on real data.

Thanks to the results in Subsection 6.2.3 on the transition probabilities, in Subsection 6.3.5 we will consider the pricing of American options. In fact, exploiting the discrete nature, with respect to time and space, of our discretization procedure, American option pricing reduces to a classic backward induction algorithm applied to a somehow *optimal multinomial lattice*.

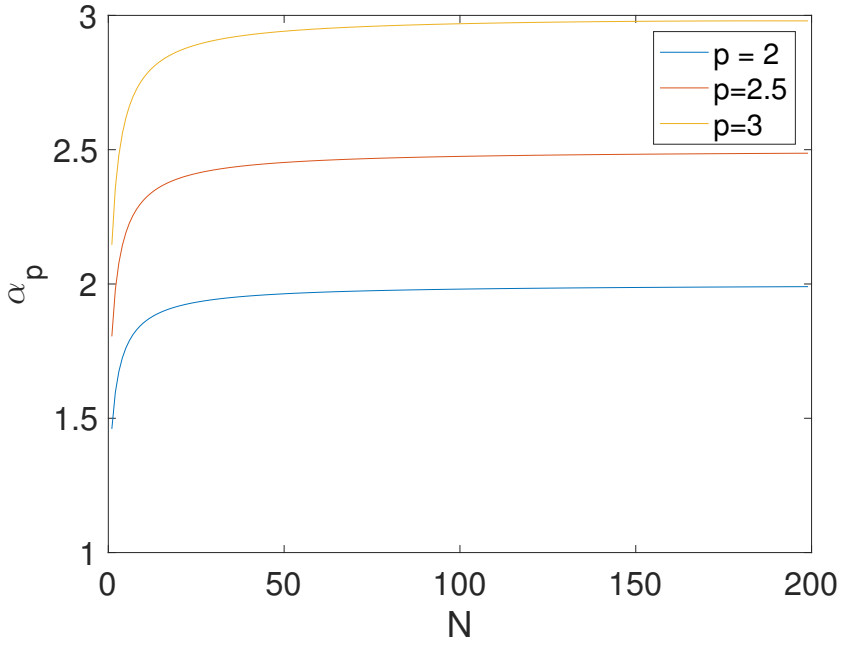
6.3.2 The Models

We resume in Table 6.1 the main features of the six models we are going to implement in our numerical tests. Three of them are not affine, but they also admit a characteristic function for the log-asset that can be expressed in closed form. Jumps are driven by compensated Poisson as well as by Lévy processes. We have also increased the dimension of the problem, by including a two-factor specification for the volatility in the Heston (1993) model, thus leading to the so-called double-Heston model investigated by Christoffersen et al. (2009).
We now briefly review the risk neutral dynamics of the models in Table 6.1. In Table 6.2, Table 6.3 and Table 6.4 we recall the closed form expression of the characteristic function of the log-price for all models, together with the numerical values of the parameters we used in our pricing example. For each model one should check whether

[4]For sake of brevity we omit the calibration for the other models, for which results are available upon request.



(a) Plot of $\log(D_p(\Gamma))$ as a function of the quantization grid size N.



(b) Plot of α_p as a function of the quantization grid size N.

FIGURE 6.2: Plot of the log-distortion $D_p(\Gamma)$ and of the parameter α_p in the function $D_p(\Gamma) = \frac{C}{N^{\alpha_p}}$ for a $\mathcal{N}(0,1)$ random variable as a function of the size of the grid $N = 1, ..., 200$ for some values of the parameter p.

Model	Affine	Non Affine	Jumps	$2d-$Vol
Heston	✓			
Double-Heston	✓			✓
Bates	✓		✓(Poisson)	
4/2		✓		
Tempered stable		✓	✓(Lévy)	
Variance Gamma		✓	✓(Lévy)	

TABLE 6.1: Features of the six models implemented in the numerical section.

the hypotheses required in either Lemma 6.2.5 or Theorem 6.2.6 are satisfied. In particular, since a few results are known relatively to the (polynomial) behavior of the density of S_T, we will mainly focus on assumptions of Lemma 6.2.5, namely we will provide some reference on finiteness of moments of S_T.

The Heston Model

The Heston (1993) model considers the following dynamics for the pair (S, V):

$$\frac{dS_t}{S_t} = rdt + \sqrt{V_t}\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right),$$
$$dV_t = \kappa(\theta - V_t) + \xi\sqrt{V_t}dW_t^1,$$

where V denotes the instantaneous variance of the price process and it is modeled as a CIR process. Here W^1 and W^2 are two independent standard Brownian motions and r is the interest rate. The parameters of the models are θ (the long-run average variance), κ (the speed of the mean reversion of the variance), ξ (the vol of vol parameter) and ρ[5]. The characteristic function in Table 6.2 is the one computed in Albrecher et al. (2007). The conditions on the finiteness of moments of S_T, which depend on model parameters, are provided in Proposition 3.1 in Andersen and Piterbarg (2007b). More generally, we refer to Gulisashvili and Stein (2010) for the asymptotic behavior of the density function of S_T.

[5]Of course, also the initial variance V_0 should be considered as a parameter to be estimated/calibrated, as it is unobservable.

The Double-Heston Model

In order to improve the fit of the implied volatility surface, Christoffersen et al. (2009) introduced an extension of the Heston (1993) model where the instantaneous variance is a bi-dimensional CIR process (V^1, V^2):

$$\begin{aligned}
\frac{dS_t}{S_t} =& rdt + \sqrt{V_t^1}\left(\rho_1 dW_t^1 + \sqrt{1-\rho_1^2}dW_t^2\right) + \sqrt{V_t^2}\left(\rho_2 dB_t^1 + \sqrt{1-\rho_2^2}dB_t^2\right),\\
dV_t^1 =& \kappa_1\left(\theta_1 - V_t^1\right) + \xi_1\sqrt{V_t^1}dW_t^1,\\
dV_t^2 =& \kappa_2\left(\theta_2 - V_t^2\right) + \xi_2\sqrt{V_t^2}dB_t^1.
\end{aligned}$$

Hereafter all Brownian motions are assumed to be mutually independent. The conditions on the finiteness of moments of S_T can be easily deduced again from Andersen and Piterbarg (2007b). In fact, as the volatility factors are independent, the characteristic function splits into one-dimensional problems, see also Section 3.1.1 in the extended online version of Da Fonseca et al. (2015).

The Bates Model

The Bates (1996) model extends the Heston (1993) one by including the presence of (Poisson) jumps in the price process:

$$\begin{aligned}
\frac{dS_t}{S_t} &= rdt + \sqrt{V_t}(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2) + dZ_t,\\
dV_t &= \kappa(\theta - V_t) + \xi\sqrt{V_t}dW_t^1,
\end{aligned}$$

where Z is a (independent) compound Poisson process with intensity λ and log-normal distribution of the jump size J, namely $\log(1+J) \sim \mathcal{N}(a - \frac{1}{2}b^2, b^2)$. We refer e.g. to Sections 6.1 and 6.2 in Keller-Ressel (2011) for a study of the finiteness of moments of S_T.

The $4/2$ Stochastic Volatility Model

Grasselli (2016) introduced a model where the instantaneous variance is given by a superposition of a CIR process and its inverse. This allows to recover as special cases the square root Heston (1993) model and the (non affine) $3/2$ model of Heston (1997) and Platen (1997):

$$\begin{aligned}
\frac{dS_t}{S_t} &= rdt + \left(a\sqrt{V_t} + \frac{b}{\sqrt{V_t}}\right)(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2),\\
dV_t &= \kappa(\theta - V_t) + \xi\sqrt{V_t}dW_t^1
\end{aligned}$$

where the parameters have the same financial meaning as in the Heston (1993) model[6]. We refer to Grasselli (2016) for a study of the finiteness of the moments of S, which can become a strict local martingale under some parameter conditions.

[6]Note that taking $a = 0$ and $b \neq 0$ corresponds to the $3/2$ model, for which the mean reversion speed is proportional to the level of the instantaneous variance, contrarily to the Heston (1993) model that has a constant speed. See Grasselli (2016) for further details.

The Tempered Stable and Variance Gamma Models

The Tempered Stable model, introduced in Cont et al. (1997), assumes that the underlying is an exponential Lévy process:

$$S_t = S_0 \exp[(r + \gamma_c)t + X_t], \quad t \in [0, T], \tag{6.12}$$

where γ_c is a convexity correction parameter and X denotes a Lévy process associated with a tempered stable distribution[7]: $X \sim TS(c^+, \alpha^+, \lambda^+, c^-, \alpha^-, \lambda^-)$. The parameters λ^+ and λ^- are tail decay rates, α^+ and α^- describe the measure of positive and negative jumps, while c^+ and c^- determine the arrival rate of jumps of given size. For further details on tempered stable processes and applications to finance we refer to Cont and Tankov (2004). In particular, the moments of $X_t, t \in [0, T]$ up to the fourth order are known explicitly.

The Variance Gamma model, introduced by Madan et al. (1998), is an important specification of the tempered stable family. It is constructed by taking a Brownian motion with drift and changing its time scale:

$$X_t = \theta\gamma(t; 1, \nu) + \sigma W_{\gamma(t;1,\nu)}, \quad t \in [0, T],$$

where $\gamma(t; \mu, \nu)$ denotes a Gamma process (with mean rate μ and variance rate ν), that is an increasing Lévy process (also called subordinator) with independent and stationary Gamma increments over non-overlapping intervals of time $(t, t + h)$[8]. The parameter σ describes the variance of the log-price process, while θ is related to the skewness.

6.3.3 Pricing of Vanilla Options

As soon as an optimal quantization grid $\Gamma = \{x_1, \dots, x_N\}$ of size N for the random variable S_T is given, pricing of vanilla options is immediate, since for any Lipschitz continuous function $F : \mathbb{R}^d \to \mathbb{R}$ we can exploit the following cubature formula

$$\mathbb{E}[F(S_T)] \cong \mathbb{E}[F(\widehat{S_T})] = \sum_{\ell=1}^{N} F(x_\ell)\mathbb{P}(\widehat{S_T} = x_\ell).$$

We consider a book of vanilla options with maturity $T = 1$ year written on an underlying with initial price $S_0 = 100$, interest rate $r = 2\%$ and strike K ranging between 80 and 120[9]. We fix the grid size to $N = 120$ and we set $p = 2$. For the parameter of the different models we used the numerical values declared in Table 6.2, Table 6.3 and Table 6.4. We implemented the pricing using *MATLAB*, on a laptop with a 2.7 GHz CPU and 8 Gb of memory.

[7] Tempered stable distributions form a six parameter family of infinitely divisible distributions. They include several well-known subclasses like Variance Gamma distributions of Madan et al. (1998), bilateral Gamma distributions of Kuchler and Tappe (2008) and the CGMY distributions of Carr et al. (2002).

[8] The density $f_h(g)$ of the Gamma increment $g = \gamma(t + h; \mu, nu) - \gamma(t; \mu, \nu)$ is given by the Gamma density function with mean μh and variance νh. See e.g. Madan et al. (1998) and Cont et al. (1997) for background on Gamma processes.

[9] As usual, Call (resp. Put) prices are provided when the strike is less (resp. greater) than 100.

Heston

$$\phi_T(u) = e^{C(T,u)+D(T,u)V0+riuT+iu\log(S0)}$$

$$C(T,u) = \frac{\kappa\theta}{\xi^2}\left((\kappa - iu\rho\xi - d(u))\,T - 2\log\left(\frac{c(u)-e^{-d(u)T}}{c(u)-1}\right)\right)$$

$$D(T,u) = \frac{\kappa - iu\rho\xi + d(u)}{\xi^2}\left(\frac{1-e^{-d(u)T}}{c(u)-e^{-d(u)T}}\right)$$

$$c(u) = \frac{\kappa - iu\rho\xi + d(u)}{\kappa - iu\rho\xi - d(u)}$$

$$d(u) = \sqrt{(iu\rho\xi - \kappa)^2 + iu\xi^2 + u^2\xi^2}$$

$$\rho = -0.6677 \qquad \kappa = 1.1646 \qquad \theta = 0.0682$$

$$\xi = 0.5360 \qquad r = 0.02 \qquad S_0 = 100 \qquad V_0 = 0.0451.$$

Double Heston

$$\phi_T(u) = e^{C_1(T,u)+C_2(T,u)+D_1(T,u)V_0^1+D_2(T,u)V_0^2+riuT+iu\log(S0)}$$

$$C_j(T,u) = \frac{\kappa_j\theta_j}{\xi_j^2}\left((\kappa_j - iu\rho_j\xi_j - d_j(u))\,T - 2\log\left(\frac{c_j(u)-e^{-d_j(u)T}}{c_j(u)-1}\right)\right), \quad j = 1,2$$

$$D_j(T,u) = \frac{\kappa_j - iu\rho_j\xi_j + d_j(u)}{\xi_j^2}\left(\frac{1-e^{-d_j(u)T}}{c_j(u)-e^{-d_j(u)T}}\right), \quad j = 1,2$$

$$c_j(u) = \frac{\kappa_j - iu\rho_j\xi_j + d_j(u)}{\kappa_j - iu\rho_j\xi_j - d_j(u)}, \quad j = 1,2.$$

$$d_j(u) = \sqrt{(iu\rho_j\xi_j - \kappa_j)^2 + iu\xi_j^2 + u^2\xi_j^2}, \quad j = 1,2.$$

$$\rho_1 = -0.834 \qquad \rho_2 = -0.957 \qquad \kappa_1 = 2.597 \qquad \theta_1 = 0.053$$

$$\xi_1 = 0.280 \qquad \kappa_2 = 2.597 \qquad \theta_2 = 0.153 \qquad \xi_2 = 3.667$$

$$r = 0.02 \qquad S_0 = 100 \qquad V_0^1 = 0.03 \qquad V_0^2 = 0.03.$$

TABLE 6.2: Characteristic functions and parameters of the Heston and Double Heston models.

Table 6.5, Table 6.6 and Table 6.7 illustrate the results. For each model we display four columns. The first one (*Benchmark*) shows the prices obtained via the methodology of Carr and Madan (1999) that we take as a benchmark. In the second column (*Quantization*) we display the prices obtained with our new methodology. In the third column (*Rel error*) there is the relative error in basis points, obtained as the absolute value of the relative error between the benchmark and quantization prices. Finally, in the last column (denoted by IVSD) we show the square of the difference between implied volatilities obtained with the benchmark and the quantization prices.

Bates
$\phi_T(u) = e^{C(T,u)+D(T,u)V0+E(T,u)+riuT+iu\log(S0)}$
$C(T,u) = \frac{\kappa\theta}{\xi^2}\left(\left(\kappa - iu\rho\xi - d(u)\right)T - 2\log\left(\frac{c(u)-e^{-d(u)T}}{c(u)-1}\right)\right)$
$D(T,u) = \frac{\kappa - iu\rho\xi + d(u)}{\xi^2}\left(\frac{1-e^{-d(u)T}}{c(u)-e^{-d(u)T}}\right)$
$E(T,u) = \lambda T\left(-iua + e^{iu\log(1+a)+\frac{b^2}{2}\left(-u^2-iu\right)} - 1\right)$
$d(u) = \sqrt{(iu\rho\xi - \kappa)^2 + iu\xi^2 + u^2\xi^2}$
$c(u) = \frac{\kappa - iu\rho\xi + d(u)}{\kappa - iu\rho\xi - d(u)}$
$\rho = -0.8210 \qquad \kappa = 2.3924 \qquad \theta = 0.0929 \qquad \xi = 0.6903$
$\lambda = 0.1 \qquad a = 0.1 \qquad b = 0.1$
$r = 0.02 \qquad S_0 = 100 \qquad V_0 = 0.0719$

4/2
$\phi_T(u) = e^{B(T,u)+C(T,u)V0+riuT+iu\log(S0)}V_0^{d^-(u)}e(u)^{m(u)+1}n(u)^{-d^+(u)}\cdot$
$\cdot\frac{\Gamma(d^+(u))}{\Gamma(m(u)+1)}{}_1F_1\left(d^+(u), m(u)+1, \frac{e^2(u)V_0}{n(u)}\right)$
$A(u) = \kappa^2 - 2\sigma^2\left(iu\left(\frac{a\rho\kappa}{\xi} - \frac{1}{2}a^2\right) - \frac{1}{2}u^2\left(1-\rho^2\right)a^2\right)$
$B(T,u) = \frac{\kappa^2\theta}{\xi^2}T + iu\left(-ab - \frac{a\rho\kappa\theta + b\rho\kappa}{\xi}\right)T - u^2\left(1-\rho^2\right)abT$
$C(T,u) = \frac{1}{\xi^2}\left(-\sqrt{A(u)}\coth\left(\frac{\sqrt{A(u)}}{2}T\right) + \kappa - iua\rho\xi\right)$
$m(u) = \frac{2}{\xi^2}\sqrt{\left(\kappa\theta - \frac{\xi^2}{2}\right)^2 - 2\xi^2\left(iu\left(\frac{b\rho}{\xi}\left(\frac{\xi^2}{2} - \kappa\theta\right) - \frac{b^2}{2}\right) - \frac{1}{2}u^2\left(1-\rho^2\right)b^2\right)}$
$K(T,u) = \frac{1}{\xi^2}\left(\sqrt{A(u)}\coth\left(\frac{\sqrt{A(u)}}{2}T\right) + \kappa\right) \qquad e(u) = \frac{\sqrt{A(u)}}{\xi^2\sinh\left(\frac{\sqrt{A(u)}}{2}T\right)}$
$d^{\pm}(u) = \frac{1}{2} + \frac{m(u)}{2} \pm \left(\frac{iub\rho}{\xi} + \frac{\kappa\theta}{\xi^2}\right) \qquad n(u) = K(T,u) - \frac{iua\rho}{\xi}$
$a = 0.3, \qquad b = 0.025, \qquad \rho = -0.7$
$\kappa = 1.8 \qquad \theta = 0.04 \qquad \xi = 0.2$
$r = 0.02 \qquad S_0 = 100 \qquad V_0 = 0.04$

TABLE 6.3: Characteristic functions and parameters of the Bates and 4/2 models

Tempered Stable
$\phi_T(u) = e^{A^+(T,u)+A^-(T,u)+(r+\omega)iuT+iu\log(S0)}$
$A^{\pm}(T,u) = Tc^{\pm}\Gamma(-\alpha^{\pm})\left((\lambda^{\pm}-iu)^{\alpha^{\pm}} - (\lambda^{\pm})^{\alpha^{\pm}}\right)$
$\omega = -\alpha^+\Gamma(-\beta^+)\left((\lambda^+-1)^{\beta^+} - \lambda^{+\beta^+}\right) - \alpha^-\Gamma(-\beta^-)\left((\lambda^-+1)^{\beta^-} - \lambda^{-\beta^-}\right)$
$c^+ = 1 \quad \alpha^+ = 0.5 \quad \lambda^+ = 6 \quad c^- = 1$
$\alpha^- = 0.5 \quad \lambda^- = 4 \quad r = 0.02 \quad S_0 = 100$

Variance Gamma
$\phi_T(u) = \left(1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu u^2\right)^{-\frac{T}{\nu}} e^{(r+\omega)iuT+iu\log(S0)}$
$\omega = \frac{1}{\kappa}\log\left(1 - \kappa\theta - \frac{1}{2}\sigma^2\kappa\right)$
$\sigma = 0.1213 \quad \nu = 0.1686 \quad \theta = 0.0436$
$r = 0.02 \quad S_0 = 100$

TABLE 6.4: Characteristic functions and parameters of the Tempered Stable and Variance Gamma models.

Table 6.5, Table 6.6 and Table 6.7 definitely show that the Fourier-quantization methodology is very accurate. What is more, our methodology is very efficient: the computational time to get all the prices is less than 0.5 seconds for each model[10].

6.3.4 Calibration to Real Data

In this subsection, we perform the calibration of the Tempered Stable model to a book of 80 option prices on the DAX index as of date August, 28$^{\text{th}}$ 2008[11]. The book includes 8 maturities (from 6 weeks to 3 years) with 20 strikes, ranging from 80% to 120% of the forward spot price. The calibration exercise consists in minimizing the average square error on implied volatilities, defined as

$$\text{Res Norm} := \frac{1}{\#_{\text{strikes}}\,\#_{\text{maturities}}} \sum_{\ell=1}^{\#_{\text{strikes}}} \sum_{k=1}^{\#_{\text{maturities}}} \left(IV_{\ell,k}^{\text{market}} - IV_{\ell,k}^{\text{model}}\right)^2.$$

[10] We did not perform any optimization in the code. On top of that, we guess that the computational time can be even reduced by migrating to other languages like e.g. *C++*.

[11] The calibration can be easily performed for each of the six models presented, we skip the results (available upon request) for the other models for the sake of brevity. The choice of the dataset is also arbitrary: we tested several dates by getting similar results.

Heston				
Strike	Benchmark	Quantization	Rel error (bp)	IVSD (10^{-9})
$K = 80$	23.8545	23.8548	0.1259	0.1707
$K = 85$	19.7998	19.7996	0.0801	0.0326
$K = 90$	15.9937	15.9939	0.1158	0.0327
$K = 95$	12.4947	12.4946	0.0950	0.0107
$K = 100$	9.3709	9.3713	0.4384	0.1104
$K = 100$	7.3908	7.3897	1.4249	0.7254
$K = 105$	9.6153	9.6135	1.8997	2.1005
$K = 110$	12.3500	12.3483	1.4093	2.0872
$K = 115$	15.6221	15.6210	0.7253	1.1215
$K = 120$	19.3972	19.3957	0.7741	2.9009
Double Heston				
Strike	Benchmark	Quantization	Rel error (bp)	IVSD (10^{-9})
$K = 80$	26.1012	26.1006	0.2592	0.6079
$K = 85$	22.3013	22.3013	0.0098	0.0005
$K = 90$	18.7483	18.7474	0.4638	0.6442
$K = 95$	15.4781	15.4769	0.7789	1.0599
$K = 100$	12.5224	12.5215	0.7214	0.5369
$K = 100$	10.5423	10.5414	0.8569	0.5369
$K = 105$	12.8268	12.8257	0.8638	0.7724
$K = 110$	15.4652	15.4636	1.0334	1.6390
$K = 115$	18.4605	18.4590	0.7857	1.4771
$K = 120$	21.8027	21.8011	0.6945	1.9000

TABLE 6.5: Pricing of Call and Put options, comparative performance of the Heston and Double Heston models.

Bates				
Strike	Benchmark	Quantization	Rel error (bp)	IVSD (10^{-9})
$K = 80$	25.1892	25.1891	0.0569	0.0301
$K = 85$	21.4161	21.4145	0.7462	2.6017
$K = 90$	17.8933	17.8930	0.1970	0.1084
$K = 95$	14.6531	14.6529	0.1057	0.0176
$K = 100$	11.7262	11.7260	0.1408	0.0179
$K = 100$	9.7460	9.7450	1.0836	0.7322
$K = 105$	12.0606	12.0592	1.1650	1.2409
$K = 110$	14.7359	14.7335	1.6546	3.8517
$K = 115$	17.7822	17.7803	1.0617	2.5818
$K = 120$	21.1953	21.1940	0.6524	1.6919
4/2				
Strike	Benchmark	Quantization	Rel error (bp)	IVSD (10^{-10})
$K = 80$	22.2976	22.2975	0.0438	0.4389
$K = 85$	18.1556	18.1554	0.0701	0.3200
$K = 90$	14.4630	14.4627	0.1752	0.7178
$K = 95$	11.2904	11.2894	0.8825	7.7704
$K = 100$	8.6592	8.6591	0.1165	0.0666
$K = 100$	6.6791	6.6791	0.0436	0.0055
$K = 105$	9.4665	9.4658	0.7249	2.9660
$K = 110$	12.7152	12.7147	0.4218	1.9542
$K = 115$	16.3530	16.3530	0.0359	0.0276
$K = 120$	20.3049	20.3049	0.0126	0.0067

TABLE 6.6: Pricing of Call and Put options, comparative performance of the Bates and 4/2 models.

Tempered Stable				
Strike	Benchmark	Quantization	Rel error (bp)	IVSD (10^{-9})
$K = 80$	27.4464	27.4433	1.1301	1.1594
$K = 85$	24.2328	24.2291	1.5040	1.2875
$K = 90$	21.2982	21.2932	2.3803	2.1063
$K = 95$	18.6452	18.6394	3.1076	2.4338
$K = 100$	16.2692	16.2649	2.6318	1.2236
$K = 100$	14.2890	14.2847	2.9965	1.2236
$K = 105$	17.0806	17.0770	2.1142	0.8311
$K = 110$	20.1232	20.1198	1.7291	0.7607
$K = 115$	23.3979	23.3940	1.6465	0.9456
$K = 120$	26.8831	26.8779	1.9504	1.8195

Variance Gamma				
Strike	Benchmark	Quantization	Rel error (bp)	IVSD (10^{-10})
$K = 80$	21.6791	21.6780	0.5022	4.9870
$K = 85$	16.9923	16.9917	0.3270	0.2344
$K = 90$	12.6260	12.6240	1.5411	0.8184
$K = 95$	8.8247	8.8239	0.9782	0.0722
$K = 100$	5.8004	5.7998	0.9101	0.0184
$K = 100$	3.8202	3.8197	1.3818	0.0184
$K = 105$	6.5342	6.5331	1.7495	0.0847
$K = 110$	9.9808	9.9799	0.9047	0.0688
$K = 115$	13.9749	13.9744	0.3720	0.0367
$K = 120$	18.3358	18.3351	0.3705	0.1181

TABLE 6.7: Pricing of Call and Put options, comparative performance of the Tempered Stable and Variance Gamma models.

Parameters	Quantization & Fourier	FFT
c_+	0.2899	0.3380
α_+	0.2800	0.0100
λ_+	18.9716	15.9982
c_-	0.0145	0.0185
α_-	1.7029	1.6574
λ_-	0.0100	0.2934
Res Norm	9.1902×10^{-5}	1.2551×10^{-4}

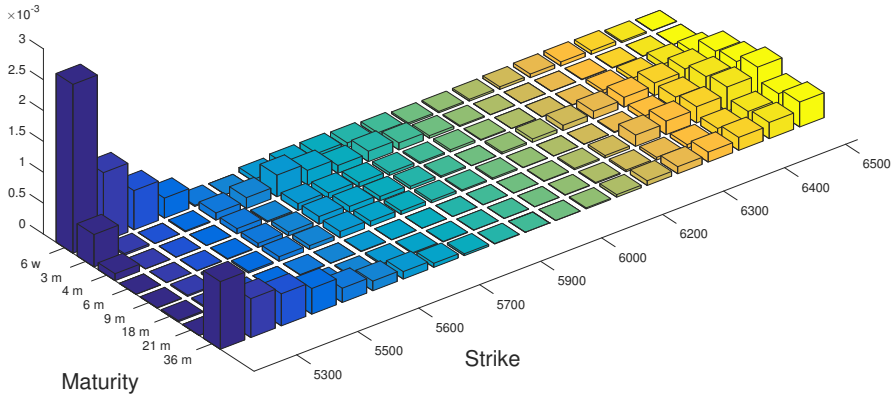
TABLE 6.8: Tempered Stable model calibrated on a book of vanillas on the DAX index as of date August, 28th 2008. The book includes 8 maturities (from 6 weeks to 3 years) with 10 strikes, ranging from 80% to 120% of the spot price, for a total of 80 options. The left (resp. right) column corresponds to the parameters calibrated using the quantization (resp. FFT) procedure. Res Norm indicates the average square error on implied volatilities.

The optimization is performed using a nonlinear least squares minimizer (the function **lsqnonlin** implemented in *MATLAB*, which uses the Levenberg-Marquardt algorithm). In Table 6.8 we compare the calibrated Tempered Stable parameters, obtained with our methodology, with the ones obtained with the Fast Fourier Transform (FFT) approach of Carr and Madan (1999). The average square error on implied volatilities (Res Norm) is similar for the two methods. Figure 6.3 provides the squared errors in the two cases. In conclusion, our new methodology can be efficiently adopted also in a calibration perspective.

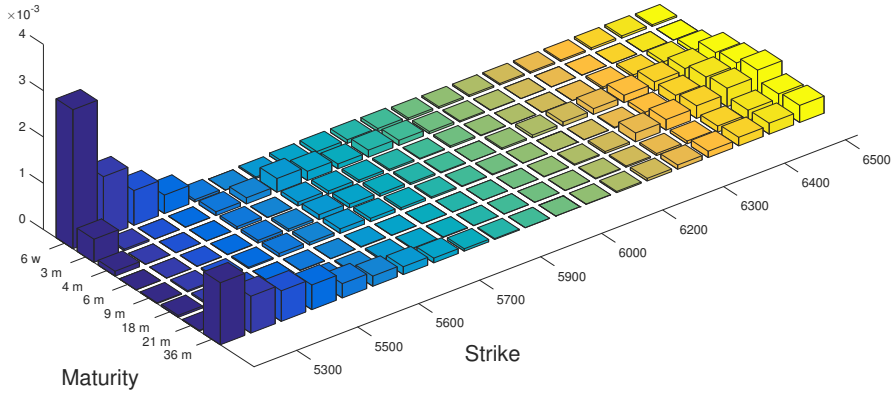
6.3.5 Pricing of American Options

Motivated by the results in Subsection 6.2.3 on transition probabilities, we now consider the pricing of American options. Let us assume that the set of possible exercise times of the American option of maturity T is finite, i.e. we are approximating the American option value with the price of a Bermudan option, where the exercise time are $t_k = \frac{kT}{M}$, with $k = 1, \dots, M$. Following the argument developed in Bally et al. (2005), it is possible to write a quantized version of the backward dynamic programming formula for the optimal value process giving the price of the American option with associated payoff $h(t, X)$:

$$\begin{cases} \widehat{V}_M & := h(t_N, \widehat{X}_M) \\ \widehat{V}_k & := \max\left(h(t_k, \widehat{X}_k), \mathbb{E}\left[\widehat{V}_{k+1} | \widehat{X}_k\right]\right), \quad k = 0, \dots, M-1. \end{cases}$$



(a) Squared errors using the quantization based procedure.



(b) Squared errors using the FFT algorithm.

FIGURE 6.3: Plot of the square of the difference between the calibrated model implied volatilities and the market implied volatilities, for all the 10 strikes and the 8 maturities of the book of vanillas on the DAX index as of date August, 28th 2008. The model used is a Tempered Stable.

Assume that we have computed the transition probabilities $\pi^k_{j\ell} = \mathbb{P}\left(\widehat{X}_{k+1} = x^{k+1}_\ell | \widehat{X}_k = x^k_j\right)$ as defined in Section 6.2.3. By backward induction, one can define the function $\widehat{v}_k$ by

$$\begin{aligned} \widehat{v}_M(x^M_j) &:= h(t_M, x^M_j), \quad j = 1, \dots, N, \\ \widehat{v}_k(x^k_j) &:= \max\left(h(t_k, x^k_j), \sum_{\ell=1}^N \pi^k_{j\ell}\, \widehat{v}_{k+1}(x^{k+1}_\ell)\right), \quad k = 0, \dots, M-1. \end{aligned} \tag{6.13}$$

Then $\widehat{V}_k = \widehat{v}_k\left(\widehat{X}_k\right)$ satisfies the dynamic programming equation.
We compute now the price of an American Put option in the Tempered Stable model, where we consider the parameters as obtained in the calibration exercise of the previous subsection, see Table 6.8. The relative error is computed as the absolute value of the relative error between the benchmark and quantization prices, and is expressed in basis points. Here the benchmark price is obtained using a finite difference method as in Cont and Voltchkova (2005).
From Table 6.9 we see that with $N = 400$ the relative error is around 1 basis point on average. In conclusion, contrary to the FFT approach, here we can price also non vanilla options using the same methodology. In fact, exploiting the discrete

Strike	Bench.	$N = 200$	RelErr	$N = 300$	RelErr	$N = 400$	RelErr
$K = 80$	2.2589	2.2805	95.6218	2.2640	22.5774	2.2598	3.9842
$K = 85$	3.1986	3.2187	62.8400	3.2030	13.7560	3.1994	2.5011
$K = 90$	4.4940	4.5136	43.6137	4.4980	8.9008	4.4946	1.3351
$K = 95$	6.2169	6.2357	30.2402	6.2206	5.9515	6.2174	0.8043
$K = 100$	8.4182	8.4364	21.6198	8.4219	4.3952	8.4187	0.5940
$K = 105$	11.1182	11.1367	16.6394	11.1217	3.1480	11.1186	0.3598
$K = 110$	14.3026	14.3216	13.2843	14.3063	2.5869	14.3030	0.2797
$K = 115$	17.9269	17.9461	10.7102	17.9310	2.2871	17.9275	0.3347
$K = 120$	21.9262	21.9463	9.1671	21.9305	1.9611	21.9268	0.2736
Mean Err	–	–	33.7485	–	7.2849	–	1.1629

TABLE 6.9: Pricing of American Put options under the Tempered Stable model according to the size of the quantization grid N and the strike price. The maturity is $T = 1$ year and the parameters of the Tempered Stable model are as in the first column of Table 6.8. The relative error is computed as the absolute value of the relative error between the benchmark and quantization prices, and is expressed in basis points. The benchmark price is obtained using a finite difference method as in Cont and Voltchkova (2005).

nature, with respect to time and space, of our discretization of a stochastic process based on quantization, we can apply all the pricing techniques that have been so far developed for trees and lattices.

6.4 Conclusion

In any advanced stochastic model, pricing of a derivative should be fast and accurate, otherwise calibration cannot be performed. In this chapter we introduced a new efficient methodology to price options through an algorithm that mixes quantization with Fourier analysis. The new technique is fast, accurate and flexible enough to consider the pricing of American-style options. In fact, our Fourier-quantization based procedure approximates the continuous time dynamics of the model with a somehow optimal multinomial lattice, for which many backward algorithms are available once the transition probabilities are known, which is the case in our approach.

6.5 Appendix

6.5.1 Proof of Theorem 6.2.1

Proof. Let us fix the index j and rewrite the distortion function in Equation (6.3) by emphasizing in the summation the four components depending on the variable x_j:

$$\begin{aligned} D_p(\Gamma) =& (\textit{terms not depending on } x_j) + \int_{x_{j-1}}^{x_j^-} (z - x_{j-1})^p d\mathbb{P}_{S_T}(z) + \int_{x_j^-}^{x_j} (x_j - z)^p d\mathbb{P}_{S_T}(z) \\ &+ \int_{x_j}^{x_j^+} (z - x_j)^p d\mathbb{P}_{S_T}(z) + \int_{x_j^+}^{x_{j+1}} (x_{j+1} - z)^p d\mathbb{P}_{S_T}(z). \end{aligned}$$

We then derive with respect to x_j, recalling the definition of the density f in (6.1):

$$\begin{aligned} \frac{\partial D_p(\Gamma)}{\partial x_j} =& \frac{1}{2}(x_j^- - x_{j-1})^p f(x_j^-) - \frac{1}{2}(x_j - x_j^-)^p f(x_j^-) + \int_{x_j^-}^{x_j} p(x_j - z)^{p-1} f(z) dz \\ &+ \frac{1}{2}(x_j^+ - x_j)^p f(x_j^+) - \int_{x_j}^{x_j^+} p(z - x_j)^{p-1} f(z) dz - \frac{1}{2}(x_{j+1} - x_j^+)^p f(x_j^+) \\ =& \frac{1}{2}\left(\frac{x_j - x_{j-1}}{2}\right)^p f(x_j^-) - \frac{1}{2}\left(\frac{x_j - x_{j-1}}{2}\right)^p f + \int_{x_j^-}^{x_j} p(x_j - z)^{p-1} f(z) dz \\ &+ \frac{1}{2}\left(\frac{x_{j+1} - x_j}{2}\right)^p f(x_j^+) - \int_{x_j}^{x_j^+} p(z - x_j)^{p-1} f(z) dz - \frac{1}{2}\left(\frac{x_{j+1} - x_j}{2}\right)^p f(x_j^+) \\ =& \int_{x_j^-}^{x_j} p(x_j - z)^{p-1} \left(\frac{1}{\pi}\frac{1}{z}\left(\int_0^{+\infty} Re\left[\phi_T(u) e^{-iu\log(z)}\right] du\right)\right) dz \\ &- \int_{x_j}^{x_j^+} p(z - x_j)^{p-1} \left(\frac{1}{\pi}\frac{1}{z}\left(\int_0^{+\infty} Re\left[\phi_T(u) e^{-iu\log(z)}\right] du\right)\right) dz \\ =& \frac{p}{\pi}\int_0^{+\infty} Re\left[\phi_T(u)\left(\int_{x_j^-}^{x_j} (x_j - z)^{p-1} z^{-1-iu} dz - \int_{x_j}^{x_j^+} (z - x_j)^{p-1} z^{-1-iu} dz\right)\right] du, \end{aligned}$$

where we used the Equation (6.1). Now we study separately the two last integrals. For the first one we use the change of variable $t = z/x_j$ to get

$$\begin{aligned} \int_{x_j^-}^{x_j} (x_j - z)^{p-1} z^{-1-iu} dz =& x_j^{p-1-iu} \int_{\frac{x_j^-}{x_j}}^1 (1 - t)^{p-1} t^{-1-iu} dt \\ =& x_j^{p-1-iu} \bar{\beta}\left(\frac{x_j^-}{x_j}, -iu, p\right), \end{aligned}$$

where the function $\bar{\beta}$ is defined in (6.5), while for the second integral we use $1/t = z/x_j$ and we find

$$\begin{aligned} \int_{x_j}^{x_j^+} (z - x_j)^{p-1} z^{-1-iu} dz =& x_j^{p-1-iu} \int_1^{\frac{x_j}{x_j^+}} \left(\frac{1}{t} - 1\right)^{p-1} \left(\frac{1}{t}\right)^{-1-iu} \left(-\frac{1}{t^2}\right) dt \\ =& x_j^{p-1-iu} \bar{\beta}\left(\frac{x_j}{x_j^+}, 1 - p + iu, p\right). \end{aligned}$$

Summing up the two expressions we get

$$\frac{\partial D_p(\Gamma)}{\partial x_j} = \frac{p}{\pi} x_j^{p-1} \int_0^{+\infty} Re\left[\phi_T(u) e^{-iu\log(x_j)} \left(\bar{\beta}\left(\frac{x_j^-}{x_j}, -iu, p\right) - \bar{\beta}\left(\frac{x_j}{x_j^+}, 1-p+iu, p\right)\right)\right] du,$$

which gives the result. □

6.5.2 Proof of Theorem 6.2.2

Proof. First of all, notice that we can rewrite $\frac{\partial D_p(\Gamma)}{\partial x_i}$ as seen in the proof of Theorem 6.2.1:

$$\frac{\partial D_p(\Gamma)}{\partial x_j} = \int_{x_j^-}^{x_j} p(x_j - z)^{p-1} f(z) dz - \int_{x_j}^{x_j^+} p(z - x_j)^{p-1} f(z) dz.$$

We now compute the upper and lower diagonal components. Noting that x_{j+1} and x_{j-1} appear only in the endpoints of the interval of integration, we easily have that

$$\begin{aligned}\frac{\partial^2 D_p(\Gamma)}{\partial x_j \partial x_{j+1}} &= -\frac{1}{2} p(x_j^+ - x_j)^{p-1} f(x_j^+) \\ &= -\frac{p}{2\pi} \frac{1}{x_j^+} \left(\frac{x_{j+1} - x_j}{2}\right)^{p-1} \int_0^{+\infty} Re\left[\phi_T(u) e^{-iu\log(x_j^+)}\right] du,\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 D_p(\Gamma)}{\partial x_j \partial x_{j-1}} &= -\frac{1}{2} p(x_j - x_j^-)^{p-1} f(x_j^-) \\ &= -\frac{p}{2\pi} \frac{1}{x_j^-} \left(\frac{x_j - x_{j-1}}{2}\right)^{p-1} \int_0^{+\infty} Re\left[\phi_T(u) e^{-iu\log(x_j^-)}\right] du.\end{aligned}$$

Let us now consider the main diagonal:

$$\begin{aligned}\frac{\partial^2 D_p(\Gamma)}{\partial x_j \partial x_j} =& -\frac{1}{2} p(x_j^+ - x_j)^{p-1} f(x_j^+) + \int_{x_j^-}^{x_j} p(p-1)(x_j - z)^{p-2} f(z) dz \\ &+ \int_{x_j}^{x_j^+} p(p-1)(z - x_j)^{p-2} f(z) dz - \frac{1}{2} p(x_j - x_j^-)^{p-1} f(x_j^-) \\ =& p(p-1)\left(\int_{x_j^-}^{x_j} (x_i - z)^{p-2} f(z) dz + \int_{x_j}^{x_j^+} (z - x_j)^{p-2} f(z) dz\right) \\ &+ \frac{\partial^2 D_p(\Gamma)}{\partial x_j \partial x_{j+1}} + \frac{\partial^2 D_p(\Gamma)}{\partial x_j \partial x_{j-1}}.\end{aligned}$$

Note that the integrals in the last expression are similar to the ones encountered in the proof of Theorem 6.2.1 when replacing the exponent $p-2$ with $p-1$, so the same computations lead to the result. □

6.5.3 Proof of Theorem 6.2.6

Proof. The proof adapts to our context the arguments of Graf and Luschgy (2000) and Cambanis and Gerr (1983) and will be developed in several steps.

Step 0. As a preliminary step, we observe that from

$$f(z) \sim z^{-\alpha} \quad \text{when } z \to +\infty$$

and the fact that S_T has finite $p-th$ moment, it follows that $\alpha > p+1$. Similarly, the polynomial behavior at 0

$$f(z) \sim z^{\gamma} \quad \text{when } z \to 0$$

and $\mathbb{E}\left[S_T^p\right] < +\infty$ implies that $\gamma > -(1+p)$. These conditions together guarantee that $||f||_{\frac{1}{p+1}} < +\infty$.

Step 1. Now we develop the expression of the distortion function in the Voronoi partition as follows:

$$\begin{aligned}
D_N(x_1,\dots,x_N) &= \sum_{i=1}^{N} \int_{C_i} |z-x_i|^p f(z)dz \\
&= \int_0^{x_1} (x_1-z)^p f(z)dz + \int_{x_1}^{\frac{x_1+x_2}{2}} (z-x_1)^p f(z)dz \\
&+ \int_{\frac{x_1+x_2}{2}}^{x_2} (x_2-z)^p f(z)dz + \int_{x_2}^{\frac{x_2+x_3}{2}} (z-x_2)^p f(z)dz + \dots \\
&+ \int_{\frac{x_{N-2}+x_{N-1}}{2}}^{x_{N-1}} (x_{N-1}-z)^p f(z)dz + \int_{x_{N-1}}^{\frac{x_{N-1}+x_N}{2}} (z-x_{N-1})^p f(z)dz \\
&+ \int_{\frac{x_{N-1}+x_N}{2}}^{x_N} (x_N-z)^p f(z)dz + \int_{x_N}^{+\infty} (z-x_N)^p f(z)dz.
\end{aligned}$$

Let $\xi_1 = \underset{z\in\left[x_1,\frac{x_1+x_2}{2}\right]}{\operatorname{argmax}} f(z), \xi_i = \underset{z\in C_i}{\operatorname{argmax}} f(z)$, for $i = 2,\dots,N-1$, and $\xi_N = \underset{z\in\left[\frac{x_{N-1}+x_N}{2},x_N\right]}{\operatorname{argmax}} f(z)$, then

$$\begin{aligned}
D_N(x_1,\dots,x_N) &\le \int_0^{x_1} (x_1-z)^p f(z)dz + f(\xi_1)\int_{x_1}^{\frac{x_1+x_2}{2}} (z-x_1)^p dz \\
&+ f(\xi_2)\int_{\frac{x_1+x_2}{2}}^{x_2} (x_2-z)^p dz + f(\xi_2)\int_{x_2}^{\frac{x_2+x_3}{2}} (z-x_2)^p dz + \dots \\
&+ f(\xi_{N-1})\int_{\frac{x_{N-2}+x_{N-1}}{2}}^{x_{N-1}} (x_{N-1}-z)^p dz + f(\xi_{N-1})\int_{x_{N-1}}^{\frac{x_{N-1}+x_N}{2}} (z-x_{N-1})^p dz \\
&+ f(\xi_N)\int_{\frac{x_{N-1}+x_N}{2}}^{x_N} (z-x_N)^p dz + \int_{x_N}^{+\infty} (z-x_N)^p f(z)dz \\
&= \int_0^{x_1} (x_1-z)^p f(z)dz + \sum_{i=1}^{N-1} \frac{f(\xi_i)+f(\xi_{i+1})}{p+1}\left(\frac{x_{i+1}-x_i}{2}\right)^{p+1} \\
&+ \int_{x_N}^{+\infty} (z-x_N)^p f(z)dz.
\end{aligned}$$

Step 2. Let us define a quantization grid $(\bar{x}_1,\dots,\bar{x}_N)$ such that

$$\frac{\int_0^{\bar{x}_i} f^{\frac{1}{p+1}}(z)dz}{||f||_{\frac{1}{p+1}}^{\frac{1}{p+1}}} = \frac{\int_0^{\bar{x}_i} f^{\frac{1}{p+1}}(z)dz}{\int_0^{+\infty} f^{\frac{1}{p+1}}(z)dz} = \frac{2i-1}{2N}, \quad \text{for } i = 1,\dots,N.$$

This definition is well posed from Step 0. We have that, for $i = 1, \dots, N-1$

$$\int_{\bar{x}_i}^{\bar{x}_{i+1}} f^{\frac{1}{p+1}}(z)dz = \frac{||f||_{\frac{1}{p+1}}^{\frac{1}{p+1}}}{N},$$

and, using the mean value theorem (notice that f is continuous), we obtain that for $i = 1, \dots, N-1$ there exists $\zeta_i \in [\bar{x}_i, \bar{x}_{i+1}]$ such that

$$(\bar{x}_{i+1} - \bar{x}_i)^p = \frac{||f||_{\frac{1}{p+1}}^{\frac{p}{p+1}}}{f^{\frac{p}{p+1}}(\zeta_i)N^p},$$

where the denominator is always bounded away from zero.
Step 3. In this step we provide a bound for the quantization error. We define

$$E_N(x_1, \dots, x_N) := \sum_{i=1}^{N-1} \frac{f(\xi_i) + f(\xi_{i+1})}{p+1} \left(\frac{x_{i+1} - x_i}{2}\right)^{p+1},$$

so that

$$D_N(x_1, \dots, x_N) \leq E_N(x_1, \dots, x_N) + \int_0^{x_1} (x_1 - z)^p f(z)dz + \int_{x_N}^{+\infty} (z - x_N)^p f(z)dz.$$

We now use the quantization grid $(\bar{x}_1, \dots, \bar{x}_N)$ defined in Step 2. Let us define, in a similar way as in Step 1, $\bar{\xi}_1 = \underset{z \in \left[\bar{x}_1, \frac{\bar{x}_1 + \bar{x}_2}{2}\right]}{\operatorname{argmax}} f(z)$, $\bar{\xi}_i = \underset{z \in \left[\frac{\bar{x}_{i-1} + \bar{x}_i}{2}, \frac{\bar{x}_i + \bar{x}_{i+1}}{2}\right]}{\operatorname{argmax}} f(z)$, for $i = 2, \dots, N-1$, and $\bar{\xi}_N = \underset{z \in \left[\frac{\bar{x}_{N-1} + \bar{x}_N}{2}, x_N\right]}{\operatorname{argmax}} f(z)$, then:

$$\begin{aligned} E_N(\bar{x}_1, \dots, \bar{x}_N) &= \sum_{i=1}^{N-1} \frac{f(\bar{\xi}_i) + f(\bar{\xi}_{i+1})}{2^{p+1}(p+1)} (\bar{x}_{i+1} - \bar{x}_i)^p (\bar{x}_{i+1} - \bar{x}_i) + \\ &= \frac{||f||_{\frac{1}{p+1}}^{\frac{p}{p+1}}}{2^{p+1}(p+1)N^p} \sum_{i=1}^{N-1} \frac{f(\bar{\xi}_i) + f(\bar{\xi}_{i+1})}{f^{\frac{p}{p+1}}(\zeta_i)} (\bar{x}_{i+1} - \bar{x}_i). \end{aligned}$$

By definition of the quantization error $e_{p,N}(S_T, \Gamma)$, we have that

$$\begin{aligned} \left(e_{p,N}(S_T, \Gamma)\right)^p \leq D_N(\bar{x}_1, \dots, \bar{x}_N) \leq E_N(\bar{x}_1, \dots, \bar{x}_N) &+ \int_0^{\bar{x}_1} (\bar{x}_1 - z)^p f(z)dz \\ &+ \int_{\bar{x}_N}^{+\infty} (z - \bar{x}_N)^p f(z)dz. \end{aligned}$$

In the next step we study the asymptotic behavior of $e_{p,N}$ when $N \to +\infty$.

Step 4 Note that when $N \to +\infty$, $\bar{x}_1 \to 0$ and $\bar{x}_N \to +\infty$, so that

$$\sum_{i=1}^{N-1} \frac{f(\bar{\xi}_i) + f(\bar{\xi}_{i+1})}{f^{\frac{p}{p+1}}(\zeta_i)} (\bar{x}_{i+1} - \bar{x}_i) \to 2 \int_0^{+\infty} f^{\frac{1}{p+1}}(z)dz = 2||f||_{\frac{1}{p+1}}^{\frac{1}{p+1}}.$$

Now we show that the integrals $\int_{\bar{x}_N}^{+\infty}(z-\bar{x}_N)^p f(z)dz$ and $\int_{\bar{x}_N}^{+\infty}(z-\bar{x}_N)^p f(z)dz$ are of order $o\left(\frac{1}{N^p}\right)$. For the first one we recall that, from Step 2,

$$\int_{\bar{x}_N}^{+\infty} f^{\frac{1}{p+1}}(z)dz = \frac{||f||_{\frac{1}{p+1}}^{\frac{1}{p+1}}}{2N},$$

so that

$$\lim_{N\to+\infty} N^p \int_{\bar{x}_N}^{+\infty}(z-\bar{x}_N)^p f(z)dz = \frac{2^p}{||f||_{\frac{1}{p+1}}^{\frac{p}{p+1}}} \lim_{N\to+\infty} \frac{\int_{\bar{x}_N}^{+\infty}(z-\bar{x}_N)^p f(z)dz}{\left(\int_{\bar{x}_N}^{+\infty} f^{\frac{1}{p+1}}(z)dz\right)^p},$$

therefore we are reduced to the study of

$$\lim_{y\to+\infty} \frac{\int_y^{+\infty}(z-y)^p f(z)dz}{\left(\int_y^{+\infty} f^{\frac{1}{p+1}}(z)dz\right)^p}.$$

From Step 0 we have that $f(z)\sim z^{-\alpha}$ when $z\to+\infty$ for $\alpha > p+1$, then

$$\lim_{y\to+\infty} \frac{\int_y^{+\infty}(z-y)^p f(z)dz}{\left(\int_y^{+\infty} f^{\frac{1}{p+1}}(z)dz\right)^p} = \lim_{y\to+\infty} \frac{\int_y^{+\infty}(z-y)^p z^{-\alpha}dz}{\left(\frac{y^{-\frac{\alpha}{p+1}+1}}{\frac{\alpha}{p+1}-1}\right)^p}.$$

Let us focus on the numerator:

$$\begin{aligned}\int_y^{+\infty}(z-y)^p z^{-\alpha}dz &\overset{z=\frac{y}{t}}{=} y^{p+1-\alpha}\int_0^1 (1-t)^p\, t^{\alpha-2-p}dt \\ &= y^{p+1-\alpha}\beta(\alpha-1-p,p+1),\end{aligned}$$

where the Beta function is well defined here as both $\alpha-1-p$ and $p+1$ are positive. In conclusion we get

$$\lim_{N\to+\infty} N^p \int_{\bar{x}_N}^{+\infty}(z-\bar{x}_N)^p f(z)dz = 0,$$

that is $\int_{\bar{x}_N}^{+\infty}(z-\bar{x}_N)^p f(z)dz$ is $o\left(\frac{1}{N^p}\right)$. The second integral $\int_{\bar{x}_N}^{+\infty}(z-\bar{x}_N)^p f(z)dz$ can be handled similarly using the fact that $f(z)$ behaves as z^γ at 0 with $\gamma > -1$.
Step 5 Using the results in Step 4, we have that, when $N\to+\infty$,

$$\begin{aligned}\left(e_{p,N}(S_T,\Gamma)\right)^p \le D_N(\bar{x}_1,\dots,\bar{x}_N) &\le \frac{||f||_{\frac{1}{p+1}}^{\frac{p}{p+1}}}{2^{p+1}(p+1)}\frac{2||f||_{\frac{1}{p+1}}^{\frac{1}{p+1}}}{N^p} + o\left(\frac{1}{N^p}\right) + o\left(\frac{1}{N^p}\right) \\ &\sim \frac{||f||_{\frac{1}{p+1}}}{2^p(p+1)}\frac{1}{N^p},\end{aligned}$$

and the theorem is proved. □

Chapter 7

Polynomial processes

This is a joint work with Giorgia Callegaro and Andrea Pallavicini.

Quantization algorithms are recently successfully adopted in option pricing problems to speed up Monte Carlo simulations thanks to the high convergence rate of the numerical approximation. In particular, recursive marginal quantization has been proven a flexible and versatile tool when applied to stochastic volatility processes. In this chapter we apply for the first time these techniques to the family of polynomial processes, by exploiting whenever possible their peculiar properties. We derive theoretical results to assess the approximation errors, and we describe in numerical examples practical tools for fast exotic option pricing.

7.1 Introduction

Recently a new class of Markov processes, termed polynomial processes, has been introduced to model stock prices in view of financial applications. We refer to Cuchiero et al. (2012) and Filipovic and Larsson (2016) for an introduction and a review of the main properties of this family of processes, which include, e.g., the Brownian motion, the geometric Brownian motion, Ornstein-Uhlenbeck processes, Jacobi processes (on which we will focus here), Lévy processes and, more generally, affine processes. The main property of polynomial processes is that conditional expectations of polynomial functions of the process are again of polynomial type. In particular, expected values of any polynomial of the process is again a polynomial in the initial value of the process, so that moments of all orders can be easily computed in closed form (up to a matrix exponential), even if the characteristic function of the process may be not known.
This fundamental property allows to implement fast calibration algorithms for plain-vanilla options quoted by the market, which usually depend only on the marginal probability distribution of the underlying asset price at option expiry date. On the other hand, pricing exotic options also requires transition probabilities because of the path-dependent feature of such derivative products, such as time averages, continuous barriers and early redemptions. Joint probability distributions of the process at different times can be in principle derived by exploiting the polynomial property of the model, but the computational time rapidly explodes as the number of time observations increases. The work of Filipovic et al. (2016) represents an attempt to deal with this issue by introducing an approximating Markov chain, which reduces the dimensionality of the problem. In the present chapter we follow a different approach and we apply, for the first time, quantization techniques to polynomial processes.
In this chapter we focus on a particular polynomial process, the Stochastic Volatility Jacobi process (hereafter SVJ) first introduced in Ackerer et al. (2016), but our results

can be extended to any polynomial model. The SVJ model is a diffusion model for stock prices where the log-price squared volatility follows a Jacobi process with values in some compact interval. It includes as limiting cases the Black-Scholes model and the Heston model, so that it can be viewed as a possible alternative to these models in practical applications of option pricing. Our analysis on SVJ quantization provides both practical tools to develop fast exotic option pricing algorithms and theoretical results to assess the approximation errors.
In particular, we describe two alternative approaches to apply quantization techniques to polynomial processes. First, we directly quantize the price probability distributions obtained by exploiting the polynomial property. In this way we provide an alternative approach which could improve the efficiency of the pricing algorithm proposed in Ackerer et al. (2016), even if we are bound to the dimensionality problems arising in pricing path-dependent derivatives. So, to overcome these limitations, we extend to multidimensional models the framework of Callegaro et al. (2017a) (Chapter 4), which does not rely on the polynomial property of the model, and we use it to price Bermudan options.

The chapter is organized as follows. In Section 7.2 the SVJ model is presented. Two different quantization approaches are then described: firstly, in Section 7.3 quantization techniques are adapted to polynomial models, leading to new pricing formulas for plain-vanilla options, whose approximation error is discussed. Then, in Section 7.4 recursive marginal quantization (which does not take into account the polynomial nature of our stochastic process) is introduced in a multidimensional setting and it is applied to price path-dependent exotic options. Numerical results for all the introduces algorithms, along with a discussion, are presented in Section 7.5.

7.2 The Stochastic Volatility Jacobi Model

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$, where $\mathbb{Q}$ is a risk neutral probability measure and where the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual hypotheses and models all the randomness in our model. We assume that the stock price process S follows a SVJ model as in Ackerer et al. (2016), namely we fix $0 \leq v_{min} < v_{max}$ and we define

$$S_t = e^{X_t} \tag{7.1}$$

where the dynamics of (V, X) follows the stochastic volatility model

$$\begin{cases} dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{Q(V_t)}dW_t \\ dX_t = (r - \delta - V_t/2)dt + \rho\sqrt{Q(V_t)}dW_t + \sqrt{V_t - \rho^2 Q(V_t)}dW_t^{\perp} \end{cases} \tag{7.2}$$

with $X_0 = x_0 \in \mathbb{R}$, $V_0 = v_0 \in [v_{min}, v_{max}]$ and where the interest rate $r > 0$, $\rho \in [-1, 1]$, the mean reversion speed is $\kappa \geq 0$, the reversion level θ belongs to $[v_{min}, v_{max}]$,

$$Q(v) := \frac{(v - v_{min})(v - v_{max})}{\left(\sqrt{v_{max}} - \sqrt{v_{min}}\right)^2}$$

and where W and $W^{\perp}$ are independent standard Brownian motions. Clearly, we have that $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^{W^{\perp}}$, $t \in [0, T]$. It is known that as special limiting cases of a SVJ model we obtain the Black-Scholes (take $v_0 = \theta = v_{max}$) and the Heston model (take $v_{min} = 0$ and $v_{max} \to \infty$).

Remark 7.2.1. **(Existence and Uniqueness of SVJ SDE Solution)**
The name SVJ is motivated by the model being clearly a stochastic volatility one, with the instantaneous squared volatility V having a dynamics of Jacobi type, bounded on the interval $[v_{min}, v_{max}]$. Indeed, the following result holds (see (Ackerer et al., 2016, Theorem 2.1)): for any deterministic initial state $(v_0, x_0) \in [v_{min}, v_{max}] \times \mathbb{R}$, there exists a unique solution (V, X) to the system (7.2), taking values in $[v_{min}, v_{max}] \times \mathbb{R}$. Furthermore, it is possible to show that if $(v_0, x_0) \in (v_{min}, v_{max}) \times \mathbb{R}$, then (V_t, X_t) takes values in $(v_{min}, v_{max}) \times \mathbb{R}$ if and only if

$$\frac{\sigma^2(v_{max} - v_{min})}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \leq 2\kappa \min\{v_{max} - \theta, \theta - v_{min}\}.$$

Moments in the SVJ model are known in closed form up to a matrix exponential. Indeed, if we write the generator $\mathcal{G}$ of the SVJ process, namely

$$\mathcal{G}f(v,x) = b^\top(v)\nabla f(v,x) + \frac{1}{2}\mathrm{Tr}\left(a(v)\nabla^2 f(v,x)\right)$$

with drift vector $b(v)$ and the diffusion matrix $a(v)$ given by

$$b(v) = \begin{bmatrix} \kappa(\theta - v) \\ r - \delta - v/2 \end{bmatrix}, \quad a(v) = \begin{bmatrix} \sigma^2 Q(v) & \rho\sigma Q(v) \\ \rho\sigma Q(v) & v \end{bmatrix}$$

we have that $\mathcal{G}$ maps any polynomial of degree n onto a polynomial of degree n or less as shown in Filipovic and Larsson (2016). As a consequence is possible to evaluate the conditional moments of (V_T, X_T) as follows. Let Pol_n be the vector space of polynomials in (v, x) of degree less than or equal to n. For any positive integer N, we term $M = (N+2)(N+1)/2$ the dimension of Pol_N, we introduce a basis $h_1(v,x), \ldots, h_M(v,x)$ of polynomials of Pol_N, and we denote by G the matrix representation of the linear map $\mathcal{G}$ restricted to Pol_N with respect to this basis. Thus, from Theorem 3.1 in Filipovic and Larsson (2016) we get that for any polynomial $p \in \mathrm{Pol}_N$ we have

$$\mathbb{E}\left[p(V_T, X_T)|\mathcal{F}_t\right] = \left[h_1(V_t, X_t) \ldots h_M(V_t, X_t)\right] e^{(T-t)G}\vec{p} \quad \forall t < T$$

where $\vec{p} \in \mathbb{R}^M$ is the coordinate representation of the polynomial $p(v,x)$ with respect to the basis h. In this chapter we term this relationship as *polynomial property*. We recall here the technical results on the closed form pricing of European options, that we will need from now on. Let us define the weighted Lebesgue space

$$L^2_w = \left\{ f(x) : ||f||^2_w = \int_{\mathbb{R}} f^2(x) w(x) dx < \infty \right\},$$

equipped with the scalar product

$$\langle f, g \rangle_w = \int_{\mathbb{R}} f(x) g(x) w(x) dx,$$

where w is the Gaussian weight function, i.e., the Gaussian density with mean μ_w and variance σ^2_w. The space L^2_w, which is an Hilbert space, admits an orthonormal

basis, called the generalized Hermite polynomials H_n, given by

$$H_n(x) = \frac{1}{\sqrt{n!}} \mathcal{H}_n\left(\frac{x-\mu_w}{\sigma_w}\right), \tag{7.3}$$

where $\mathcal{H}_n$ are the probabilist Hermite polynomials defined as

$$\mathcal{H}_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}. \tag{7.4}$$

If we assume that g_T is the density of the log price X_T, then we can define $\ell(x) = \frac{g_T(x)}{w(x)}$. If we want to price a European option with payoff $f \in L^2_w$ then we get

$$\mathbb{E}[f(X_T)] = \int_{\mathbb{R}} f(x) g_T(x) dx = \int_{\mathbb{R}} f(x)\ell(x)w(x)dx = \langle f, \ell \rangle_w.$$

Since L^2_w is an Hilbert space with orthonormal basis H_n, we can rewrite the previous formula as

$$\mathbb{E}[f(X_T)] = \sum_{n\geq 0} f_n \ell_n, \tag{7.5}$$

for the *Fourier coefficients*

$$f_n = \langle f, H_n \rangle_w, \tag{7.6}$$

and the *Hermite moments*

$$\ell_n = \langle \ell, H_n \rangle_w = \int_{\mathbb{R}} H_n(x)\ell(x)w(x)dx = \int_{\mathbb{R}} H_n(x) g_T(x) dx. \tag{7.7}$$

Notice that the last equality shows that ℓ_n is a linear combination of moments of X_T, since H_n is a polynomial. It is then possible to compute ℓ_n in closed form, because of the polynomial nature of the process.

7.3 Quantization of a Polynomial Process

We will now discuss how to deal with polynomial processes. In particular, we consider the SVJ model, but our approach is general and flexible enough to be applied to any polynomial process. As a first approach, in this section, we exploit the polynomial property and we focus on the quantization of the log price process X at a fixed date. Then, in Section 7.4, in order to deal with path-dependent options, we forget about the polynomial nature of (V, X) and we extend the general framework in Callegaro et al. (2017a) (Chapter 4) to discretize the bidimensional process (V, X) at a whole set of dates via RMQ.

7.3.1 Exploiting the Polynomial Property

In this section we will consider the problem of finding a (sub-)optimal quantizer of the log price process X at a given time T.
The main result of this section is the possibility of writing the Master equation in closed form, thanks to the polynomial nature of our processes.

Theorem 7.3.1. **(Polynomial Process Quantization)**
Consider the Master equation

$$\mathbb{E}\left[(X_T - x_i)\, \mathbb{1}_{X_T \in C_i(\Gamma)}\right] = 0, \quad i = 1, \dots, N \tag{7.8}$$

and its i-th component

$$E_i(x_1,\dots,x_N) := \mathbb{E}\left[(X_T - x_i)\,\mathbb{1}_{X_T\in C_i(\Gamma)}\right] = 0 \tag{7.9}$$

where $C_i(\Gamma) = \left[\frac{x_{i-1}+x_i}{2}, \frac{x_i+x_{i+1}}{2}\right]$, $C_1(\Gamma) = \left[-\infty, \frac{x_1+x_2}{2}\right]$ and $C_N(\Gamma) = \left[\frac{x_{N-1}+x_N}{2}, +\infty\right]$. If X is a polynomial process with Hermite moments ℓ_n, then equation (7.9) reads

$$\sum_{n\geq 0} f_n^i \ell_n = 0, \tag{7.10}$$

where the (Fourier) coefficients f_n^i are given by

$$f_n^i = h_n\left(\frac{x_{i-1}+x_i}{2}\right) - h_n\left(\frac{x_i+x_{i+1}}{2}\right) - x_i\left(l_n\left(\frac{x_{i-1}+x_i}{2}\right) - l_n\left(\frac{x_i+x_{i+1}}{2}\right)\right)$$

with

$$\begin{aligned}
h_0(K) &= \sigma_w\phi\left(\frac{K-\mu_w}{\sigma_w}\right) + \mu_w\Phi\left(\frac{\mu_w-K}{\sigma_w}\right)\\
h_1(K) &= \sigma_w\left[\frac{K-\mu_w}{\sigma_w}\phi\left(\frac{K-\mu_w}{\sigma_w}\right) + \Phi\left(\frac{K-\mu_w}{\sigma_w}\right)\right] + \mu_w\phi\left(\frac{K-\mu_w}{\sigma_w}\right)\\
h_n(K) &= \frac{1}{\sqrt{n!}}\phi\left(\frac{K-\mu_w}{\sigma_w}\right)\left[\sigma_w\mathcal{H}_n\left(\frac{K-\mu_w}{\sigma_w}\right) + n\sigma_w\mathcal{H}_{n-2}\left(\frac{K-\mu_w}{\sigma_w}\right)\right.\\
&\quad \left. + \mu_w\mathcal{H}_{n-1}\left(\frac{K-\mu_w}{\sigma_w}\right)\right], n\geq 2,
\end{aligned} \tag{7.11}$$

and

$$\begin{aligned}
l_0(K) &= \Phi\left(\frac{K-\mu_w}{\sigma_w}\right)\\
l_n(K) &= \frac{1}{\sqrt{n!}}\mathcal{H}_{n-1}\left(\frac{K-\mu_w}{\sigma_w}\right)\phi\left(\frac{K-\mu_w}{\sigma_w}\right), n\geq 1,
\end{aligned} \tag{7.12}$$

and where ϕ and Φ are, respectively, the density and the cumulative distribution functions of a standard univariate Gaussian random variable.

Proof. See Appendix 7.7.1. □

7.3.2 Calculation of the Sub-Optimal Quantizer

Even if Equation (7.9) can be written in closed form for every $i = 1,\dots,N$, it is impossible to find an analytical expression for the solution to the nonlinear system, which corresponds to the (sub)-optimal quantizer. Hence, we need to solve this system numerically. As already noted in the previous chapters, the literature suggests the Newton-Raphson method as the best first choice to tackle the system of equations. The proposition below provides the Jacobian matrix to be used in the Newton-Raphson procedure.

Proposition 7.3.2. Consider the system of equations (7.8)

$$\begin{cases} E_1(x_1,\dots,x_N) := \mathbb{E}\left[(X_T - x_1)\,\mathbb{1}_{X_T\in C_1(\Gamma)}\right] = 0 \\ E_2(x_1,\dots,x_N) := \mathbb{E}\left[(X_T - x_2)\,\mathbb{1}_{X_T\in C_2(\Gamma)}\right] = 0 \\ \qquad\qquad \vdots \\ E_N(x_1,\dots,x_N) := \mathbb{E}\left[(X_T - x_N)\,\mathbb{1}_{X_T\in C_N(\Gamma)}\right] = 0 \end{cases} \tag{7.13}$$

When X is a polynomial process with Hermite moments ℓ_n, the Jacobian matrix J of the vector function $E = (E_1,\dots,E_N)$ is tridiagonal and symmetric, and its components have the following form:

$$\begin{aligned} J_{i,i-1} &= \frac{1}{2}\left(\frac{x_i - x_{i-1}}{2}\right) g_T\left(\frac{x_{i-1}+x_i}{2}\right) \quad i = 2,\dots,N \\ J_{i,i} &= J_{i,i-1} + J_{i,i+1} - \mathbb{P}\left(S_T \in C_i(\Gamma)\right) \quad i = 1,\dots,N, \end{aligned}$$

with $J_{1,0} = J_{N,N+1} = 0$ and where

$$g_T(x) = \sum_{n\geq 0} \ell_n H_n(x) w(x), \tag{7.14}$$

$$\mathbb{P}\left(X_T \in C_i(\Gamma)\right) = \sum_{n\geq 0} \ell_n \left(l_n\left(\frac{x_{i-1}+x_i}{2}\right) - l_n\left(\frac{x_i + x_{i+1}}{2}\right)\right), \tag{7.15}$$

and the coefficients l_n are computed in (7.12).

Proof. See Appendix 7.7.2. □

The Newton-Raphson algorithm has then the following structure: starting from an initial guess $\Gamma^{(0)}$, the k-th iteration is

$$\Gamma^{(k)} = \Gamma^{(k-1)} - J^{-1}\left(\Gamma^{(k-1)}\right) E\left(\Gamma^{(k-1)}\right) \quad k = 1,2,\dots \tag{7.16}$$

where $E = (E_1,\dots,E_N)$ is defined in Proposition 7.3.2 and it is computed thanks to Theorem 7.3.1. Given a stopping criterion, the final iteration is the (sub-)optimal quantization grid Γ^*.
Then, let us assume that we have found, numerically, the solution $\Gamma^* = \{x_1^*,\dots,x_N^*\}$. This grid Γ^* is the (sub-)optimal quantization grid associated to X_T. In order to compute an expected value as the pricing of a Vanilla option, we need to know the weights associated to every Voronoi cell $C_i(\Gamma^*)$, for $i = 1,\dots,N$. The weights are straightforwardly given by (7.15).

7.3.3 Analysis of the Approximation Error

We focus on pricing of at time 0 of a European option with payoff f. We consider, without loss of generality, f as the payoff of a Call option written on S having expiry $T > 0$ and strike price K, i.e.,

$$f = f(S_T) = (S_T - K)^+$$

In what follows we will need the following three versions of the price:

- π_f is the exact price at time 0, i.e.,

$$\pi_f := \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}(S_T - K)^+\right] = e^{-rT}\int_{\mathbb{R}}(e^x - K)^+ g_T(x)dx,$$

where g_T is the density of the log price X at time T, given by (7.14). This formula contains an infinite sum, so the function g_T function is not computable in closed form.

- $\pi_f^{(M)}$ is the price computed using the polynomial approximation at level M, i.e., approximating the density $g_T(x)$ with

$$g_T^{(M)}(x) = \sum_{n=0}^{M} \ell_n H_n(x) w(x), \tag{7.17}$$

namely

$$\pi_f^{(M)} = e^{-rT}\int_{\mathbb{R}}(e^x - K)^+ g_T^{(M)}(x)dx = \sum_{n=0}^{M}\ell_n f_n,$$

- $\widehat{\pi}_f^{(M,N)}$ is the price computed by approximating the log-spot price at maturity by means of quantization on a grid with N points:

$$\begin{aligned}\widehat{\pi}_f^{(M,N)} &= e^{-rT}\sum_{i=1}^{N}\int_{C_i(\Gamma^X)}(e^{x_i} - K)^+ g_T^{(M)}(x)dx \\ &= e^{-rT}\sum_{i=1}^{N}(e^{x_i} - K)^+ \mathbb{P}\left(X_T^{(M)} \in C_i(\Gamma^X)\right)\end{aligned} \tag{7.18}$$

where $\Gamma^X = \{x_1, \ldots, x_N\}$ is the optimal quantizer relative to $X_T^{(M)}$, the log price with (approximate) density $g_T^{(M)}$. We denote by $\widehat{X}_T^{(M,N)}$ the quantization of $X_T^{(M)}$. Notice that we also have

$$\widehat{\pi}_f^{(M,N)} = e^{-rT}\sum_{i=1}^{N}(e^{x_i} - K)^+ \mathbb{P}\left(\widehat{X}_T^{(M,N)} = x_i\right).$$

The accuracy of our methodology can be studied by analyzing the (asymptotic) behavior of the price approximation, namely

$$\mathrm{err}_{M,N} := \left|\pi_f - \widehat{\pi}_f^{(M,N)}\right|. \tag{7.19}$$

We split the error in two parts that we study separately:

$$\mathrm{err}_{M,N} \le \left|\pi_f - \pi_f^{(M)}\right| + \left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right|,$$

where $\left|\pi_f - \pi_f^{(M)}\right|$ is the truncation error $|\varepsilon^{(M)}|$, depending on M, as defined and studied in Ackerer et al. (2016, Section 4), to which we refer for a detailed analysis, while $\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right|$ is the quantization error, depending on N, on which we focus in the remaining part of this section.
We focus now the quantization error $\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right|$, via an intermediate lemma and a theorem providing a precise asymptotic behavior of the quantization error.

Lemma 7.3.3. The quantization error $\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right|$ satisfies

$$\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right| \leq \left|\left|S_T^{(M)} - \widehat{S}_T^{(M,N)}\right|\right|_2,$$

where $S_T^{(M)} := e^{X_T^{(M)}}$ and $\widehat{S}_T^{(M,N)}$ is the N-quantizer relative to $S_T^{(M)}$.

Proof. See Appendix 7.7.3. □

Now we are ready to size the quantization error. As known from Zador theorem, the distance $||S_T^{(M)} - \widehat{S}_T^{(M,N)}||_2$ has an asymptotic linear decay when N goes to infinity. The precise expression of this limit can be derived by using the recent error estimates obtained in a different setting by Callegaro et al. (2017b, Theorem 2.11) (Theorem 6.2.6), where the second order moment of $S_T^{(M)}$ is required to be finite and the density of $S_T^{(M)}$ at 0 and at $+\infty$ must have polynomial behavior. Here, we can show that we can relax these assumptions thanks to the explicit form of the density $h_T^{(M)}$ of $S_T^{(M)}$.

Theorem 7.3.4. **(Quantization Error Estimate)**
In our setting, for any given $M > 0$, we have the following result:

$$\lim_{N\to+\infty} N \left|\left|S_T^{(M)} - \widehat{S}_T^{(M,N)}\right|\right|_2 \leq \frac{\left|\left|h_T^{(M)}\right|\right|_{\frac{1}{2}}^{\frac{1}{2}}}{2\sqrt{3}}. \tag{7.20}$$

where $S_T^{(M)} = e^{X_T^{(M)}}$ has density $h_T^{(M)}(s) = \frac{g_T^{(M)}(\ln(s))}{s}$ for $s \in (0, +\infty)$ and $\widehat{S}_T^{(M,N)}$ is the N-quantizer relative to $S_T^{(M)}$.

Proof. See Appendix 7.7.4. □

7.4 An Alternative Approach: Multidimensional Recursive Quantization

Here, we consider the quantization of a system of SDEs. We will present a general framework, that we will then apply to the case of the SVJ model. Let us consider the following d-dimensional SDE:

$$d\mathbf{X}_t = \mu(t, \mathbf{X}_t)\, dt + \Sigma(t, \mathbf{X}_t)\, d\mathbf{W}_t, \quad \mathbf{X}_0 = \mathbf{x}_0. \tag{7.21}$$

where $\mu : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$, $\Sigma : \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}^d \times \mathbb{R}^q$ and $\mathbf{W}$ is a q-dimensional Brownian motion. We suppose that μ and Σ are sufficiently regular so that to ensure existence and uniqueness of a solution to the SDE (7.21).
Let us now fix a time discretization grid $t_k = k\Delta, k = 0, \ldots, L$, with $\Delta = \frac{T}{L}$, where T is a given maturity, Δ is the time step size and L is the number of discretization points of the time grid. A general discretization scheme can be written in the following iterative form:

$$\widetilde{\mathbf{X}}_{t_{k+1}} = A\left(t_k, \Delta, \widetilde{\mathbf{X}}_{t_k}\right) + B\left(t_k, \Delta, \widetilde{\mathbf{X}}_{t_k}, \Delta\mathbf{W}_{t_k}\right), \quad \widetilde{\mathbf{X}}_{t_0} = x_0, \tag{7.22}$$

where $A : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ and $B : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^q \longrightarrow \mathbb{R}^d$ depend on the discretization scheme and $\Delta\mathbf{W}_{t_k} = \frac{1}{\sqrt{\Delta}}\left(\mathbf{W}_{t_{k+1}} - \mathbf{W}_{t_k}\right)$ is a q-dimensional Gaussian vector with mean $\mathbf{0}$ and variance-covariance matrix I_q.

Depending on the time discretization scheme in use, it is possible to know the law of $(\widetilde{\mathbf{X}}_{t_{k+1}}|\widetilde{\mathbf{X}}_{t_k})$, that clearly depends on A and B. In particular, in the case of the Euler-Maruyama (or simply Euler) scheme, that we will choose, $(\widetilde{\mathbf{X}}_{t_{k+1}}|\{\widetilde{\mathbf{X}}_{t_k} = \mathbf{x}\}), \mathbf{x} \in \mathbb{R}^d$, has a multivariate Gaussian distribution, while in the case of the Milstein scheme it has a generalized Chi-squared distribution. For higher order schemes, the conditional distribution has to be determined on a case by case basis.

7.4.1 Mathematical Foundation of the Algorithm

Henceforth we consider the Euler scheme, so that, conditioning on $\{\widetilde{\mathbf{X}}_{t_k} = \mathbf{x}\}$, we have (recall Equation (7.22))

$$A(t_k, \Delta, \mathbf{x}) = \mathbf{x} + \mu(t_k, \mathbf{x})\Delta, \qquad B(t_k, \Delta, \mathbf{x}, \Delta\mathbf{W}_{t_k}) = \sqrt{\Delta}\Sigma(t_k, \mathbf{x})\Delta\mathbf{W}_{t_k}, \tag{7.23}$$

and the following lemma holds:

Lemma 7.4.1. For every $0 \leq k \leq L$, conditionally on the event $\{\widetilde{\mathbf{X}}_{t_k} = \mathbf{x}\}$, the random vector $\widetilde{\mathbf{X}}_{t_{k+1}}$ is Gaussian:

$$\mathcal{L}(\widetilde{\mathbf{X}}_{t_{k+1}}|\{\widetilde{\mathbf{X}}_{t_k} = \mathbf{x}\}) \sim \mathcal{N}\left(\mathbf{x} + \mu(t_k, \mathbf{x})\Delta, \Delta(\Sigma\Sigma^T)(t_k, \mathbf{x})\right). \tag{7.24}$$

In particular, if $\mathbf{X}_t = \left(X_t^1, \ldots, X_t^d\right)$ and $\widetilde{\mathbf{X}}_{t_k} = \left(\widetilde{X}_{t_k}^1, \ldots, \widetilde{X}_{t_k}^d\right)$ for $k = 0, \ldots, L$, and $\mathbf{x} = \left(x^1, \ldots, x^d\right)$, we have that for every $i = 1, \ldots, d$

$$\mathcal{L}(\widetilde{X}_{t_{k+1}}^i|\{\widetilde{\mathbf{X}}_{t_k} = \mathbf{x}\}) \sim \mathcal{N}\left(m_i(t_k, \mathbf{x}), s_i(t_k, \mathbf{x})\right), \tag{7.25}$$

where

$$m_i(t_k, \mathbf{x}) := x^i + \mu_i(t_k, \mathbf{x})\Delta$$

is the i-th component of the vector $\mathbf{x} + \mu(t_k, \mathbf{x})$ and $s_i(t_k, \mathbf{x})$ is the i-th diagonal element of the (symmetric) matrix $\Sigma\Sigma^T$.

It is then possible to write the distribution of $\widetilde{X}_{t_{k+1}}^i$ in a closed form:

$$\mathbb{P}\left(\widetilde{X}_{t_{k+1}}^i \in dx_{k+1}^i\right) = \int_{\mathbb{R}^d} \phi_{m_i(t_k,\mathbf{x}_k), s_i(t_k,\mathbf{x}_k)}\left(x_{k+1}^i\right) \mathbb{P}\left(\widetilde{\mathbf{X}}_{t_k} \in d\mathbf{x}_k\right), \tag{7.26}$$

where $\phi_{m,s}$ is the probability density function of a one dimensional Gaussian variable with mean m and variance s.

Let us fix a quantization grid $\mathbf{\Gamma}_{i,k+1} = \left(\gamma_{i,k+1}^1, \ldots, \gamma_{i,k+1}^N\right)$ of size N relative to $\widetilde{X}_{t_{k+1}}^i$. The distortion function associated with $\mathbf{\Gamma}_{i,k+1}$ reads

$$D_{i,k+1}(\mathbf{\Gamma}_{i,k+1}) = \sum_{j=1}^{N} \int_{C_j(\mathbf{\Gamma}_{i,k+1})} \left(x_{k+1}^i - \gamma_{i,k+1}^j\right)^2 \mathbb{P}\left(\widetilde{X}_{t_{k+1}}^i \in dx_{k+1}^i\right) \tag{7.27}$$

where $\left(C_j(\mathbf{\Gamma}_{i,k+1})\right)_{j=1,\ldots,N}$ is the Voronoi tessellation associated with the grid $\mathbf{\Gamma}_{i,k+1}$.

It is now possible to write the recursive quantization algorithm. Having quantized every i-th component of the vector $\widetilde{\mathbf{X}}_{t_k}$, via an N^i-dimensional grid, it is possible to approximate the distribution in (7.26) as

$$\mathbb{P}\left(\widetilde{X}_{t_{k+1}}^i \in dx_{k+1}^i\right) \approx \sum_{j_1=1}^{N^1} \cdots \sum_{j_d=1}^{N^d} \phi_{m_i(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}), s_i(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d})}\left(x_{k+1}^i\right) \mathbb{P}\left(\widetilde{\mathbf{X}}_{t_k} = \left(x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}\right)\right), \tag{7.28}$$

where $x_{\ell,k}^{j_\ell}$ corresponds to the j_ℓ-th element of the optimal quantization grid of the ℓ-th component of the vector $\widetilde{\mathbf{X}}_{t_k}$. It is immediate to see that it is possible to compute in closed form also the distribution of the vector $\widetilde{\mathbf{X}}_{t_{k+1}}$: indeed, we have that

$$\mathbb{P}\left(\widetilde{\mathbf{X}}_{t_{k+1}} \in dx_{k+1}\right) \approx \sum_{j_1=1}^{N^1} \cdots \sum_{j_d=1}^{N^d} \bar{\phi}_{m(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}), s(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d})}(x_{k+1}) \, \mathbb{P}\left(\widetilde{\mathbf{X}}_{t_k} = \left(x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}\right)\right), \tag{7.29}$$

where $\bar{\phi}$ is the density function of a d-dimensional Gaussian random variable with mean $m(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}) = \left(x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}\right) + \mu\left(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}\right)$ and variance-covariance matrix $s(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}) = \Delta(\Sigma\Sigma^T)\left(t_k, x_{1,k}^{j_1}, \ldots, x_{d,k}^{j_d}\right)$.
Having computed all these elements, it is possible to compute the (approximate) distortion function (7.27), its gradient and its Hessian function and to implement the Newton-Raphson algorithm as in Callegaro et al. (2015) (Chapter 2), Callegaro et al. (2016) (Chapter 3) and Callegaro et al. (2017a) (Chapter 4).

7.4.2 Recursive Quantization of the SVJ Model

We focus now on the application of the arguments in Section 7.4.1 to the specific model considered. We consider the Euler scheme of the price S, instead of the log price X, since quantizing S instead of X is crucial if we want to be in the setting of Section 7.3.3 devoted to the study of the numerical error of our procedure. Using the notation of the previous section, $\mathbf{X}_t = (S_t, V_t)$ and $\widetilde{X}_{t_k} = \left(\widetilde{S}_{t_k}, \widetilde{V}_{t_k}\right)$, and the Euler scheme reads

$$\begin{pmatrix} \widetilde{S}_{t_{k+1}} \\ \widetilde{V}_{t_{k+1}} \end{pmatrix} = \begin{pmatrix} \widetilde{S}_{t_k} \\ \widetilde{V}_{t_k} \end{pmatrix} + \begin{pmatrix} (r-\delta)\Delta \\ \kappa\left(\theta - \widetilde{V}_{t_k}\right)\Delta \end{pmatrix} + \sqrt{\Delta} \begin{pmatrix} \rho \widetilde{S}_{t_k} \sqrt{Q\left(\widetilde{V}_{t_k}\right)} & \widetilde{S}_{t_k} \sqrt{\widetilde{V}_{t_k} - \rho^2 Q\left(\widetilde{V}_{t_k}\right)} \\ \sigma \sqrt{Q\left(\widetilde{V}_{t_k}\right)} & 0 \end{pmatrix} \begin{pmatrix} \Delta W_k^1 \\ \Delta W_k^2 \end{pmatrix}, \tag{7.30}$$

with the initial condition for the scheme

$$\begin{pmatrix} \widetilde{S}_{t_0} \\ \widetilde{V}_{t_0} \end{pmatrix} = \begin{pmatrix} S_0 \\ V_0 \end{pmatrix}.$$

We have then that

$$\mathbb{P}\left(\widetilde{S}_{t_{k+1}} \in ds_{k+1}\right) = \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_{m_1(t_k, s_k, v_k), \varsigma_1(t_k, s_k, v_k)}(s_{k+1}) \, \mathbb{P}\left(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_k} \in dv_k\right), \tag{7.31}$$

where $m_1(t_k, s_k, v_k) = s_k + (r-\delta)\Delta$ and $\varsigma_1(t_k, s_k, v_k) = (s_k)^2 v_k$. In the case of the variance process

$$\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in dv_{k+1}\right) = \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_{m_2(t_k, s_k, v_k), \varsigma_2(t_k, s_k, v_k)}(v_{k+1}) \, \mathbb{P}\left(\widetilde{S}_{t_k} \in ds_k, \widetilde{V}_{t_k} \in dv_k\right), \tag{7.32}$$

where $m_2(t_k, s_k, v_k) = v_k + \kappa(\theta - v_k)\Delta = m_2(t_k, v_k)$ and $\varsigma_2(t_k, s_k, v_k) = \sigma^2 Q(v_k) = \varsigma_2(t_k, v_k)$. Moreover, we notice that, since m_2 and ς_2 do not depend on s_k, we can simplify (7.32):

$$\mathbb{P}\left(\widetilde{V}_{t_{k+1}} \in dv_{k+1}\right) = \int_{\mathbb{R}} \phi_{m_2(t_k, v_k), \varsigma_2(t_k, v_k)}(v_{k+1}) \, \mathbb{P}\left(\widetilde{V}_{t_k} \in dv_k\right). \tag{7.33}$$

This allows to use the technique developed in Pagès and Sagna (2015) and Callegaro et al. (2015) (Chapter 2) for the quantization of the variance process, since it is one dimensional and it can be discretized independently of S. On the other hand, of course, the quantization grids for S will depend on the ones for V.

Remark 7.4.2. **(Discretization of the Jacobi Process)**
We recall here that the Euler scheme is not well-defined for the Jacobi process (which corresponds to the volatility process), since it would require a priori to consider the square-root of a negative real number. Many solutions to this problem have been proposed and we refer to e.g. Alfonsi (2015, Chapter 6). Notice that in practice our methodology does not require the simulation of S and V via the Euler scheme. Nevertheless, the effect of using such a scheme are visible in the quantization phase (see equations below).

We now give an idea on how it is possible to recursively obtain the quantization grids $\mathbf{\Gamma}_{1,k} =: \mathbf{\Gamma}_{S,k}$ and $\mathbf{\Gamma}_{2,k} =: \mathbf{\Gamma}_{V,k}, k = 1, \dots, L$. We suppose that the cardinality of the grids is fixed: $|\mathbf{\Gamma}_{1,k}| = N^S$ and $|\mathbf{\Gamma}_{2,k}| = N^V$, for every k. Moreover, we recall that the quantization grids at time $t_0 = 0$, $\mathbf{\Gamma}_{S,0}$ and $\mathbf{\Gamma}_{V,0}$, are vectors whose components correspond, respectively, with s_0 and v_0.
Let us assume now that we have computed the optimal grids for the price and the variance process, namely $\mathbf{\Gamma}_{S,k} = \left(s_k^1, \dots, s_k^{N_S}\right)$ for the price process and $\mathbf{\Gamma}_{V,k} = \left(v_k^1, \dots, v_k^{N_V}\right)$, up to time t_k and that we want to obtain $\mathbf{\Gamma}_{S,k+1}$ and $\mathbf{\Gamma}_{V,k+1}$. To do this, we look for the zeros of the gradient of the distortion function (7.27) when the probability (7.29) takes the form

$$\mathbb{P}\left(\widetilde{S}_{t_{k+1}} \in ds_{k+1}, \widetilde{V}_{t_{k+1}} \in dv_{k+1}\right) \approx \sum_{i=1}^{N_S}\sum_{j=1}^{N_V} \bar{\phi}_{m(t_k,s_k^i,v_k^j),\varsigma(t_k,s_k^i,v_k^j)}(s_{k+1}, v_{k+1})\mathbb{P}\left(\widetilde{S}_{t_k} = s_k^i, \widetilde{V}_{t_k} = v_k^j\right), \tag{7.34}$$

where $\bar{\phi}$ is the density of a bivariate Gaussian with mean

$$m(t_k, s_k^i, v_k^j) = \begin{pmatrix} s_k^i + (r-\delta)\,\Delta \\ v_k^j + \kappa\left(\theta - v_k^j\right)\Delta \end{pmatrix}$$

and variance - covariance matrix

$$\varsigma(t_k, s_k^i, v_k^j) = \Delta \begin{pmatrix} \left(s_k^i\right)^2 v_k^j & \rho\sigma s_k^i Q(v_k^j) \\ \rho\sigma s_k^i Q(v_k^j) & \sigma^2 Q(v_k^j) \end{pmatrix}.$$

Remark 7.4.3. **(Calculation of Transition Probabilities)**
Notice that, as a byproduct of recursive quantization, we instantaneously get for free also the transition probabilities. Indeed, from (7.34) we immediately have the transition densities

$$\mathbb{P}\left(\widetilde{S}_{t_{k+1}} \in ds_{k+1}, \widetilde{V}_{t_{k+1}} \in dv_{k+1} | \widetilde{S}_{t_k} = s_k^i, \widetilde{V}_{t_k} = v_k^j\right) \approx \sum_{i=1}^{N_S}\sum_{j=1}^{N_V} \bar{\phi}_{m(t_k,s_k^i,v_k^j),\varsigma(t_k,s_k^i,v_k^j)}(s_{k+1}, v_{k+1}),$$

for $i = 1, \dots, N_S$ and $j = 1, \dots, N_V$.

7.5 Numerical Results

In this section we present numerical results on pricing of European and Bermudan options. Polynomial quantization is only used to price vanilla options, while recursive marginal quantization, allowing for an immediate approximation of the transition probabilities, is exploited both to price European and Bermudan derivatives. Before showing our results, we show in Figure 7.1 the behavior of the density function $g_T^{(M)}$ introduced in (7.17) for different values of M.

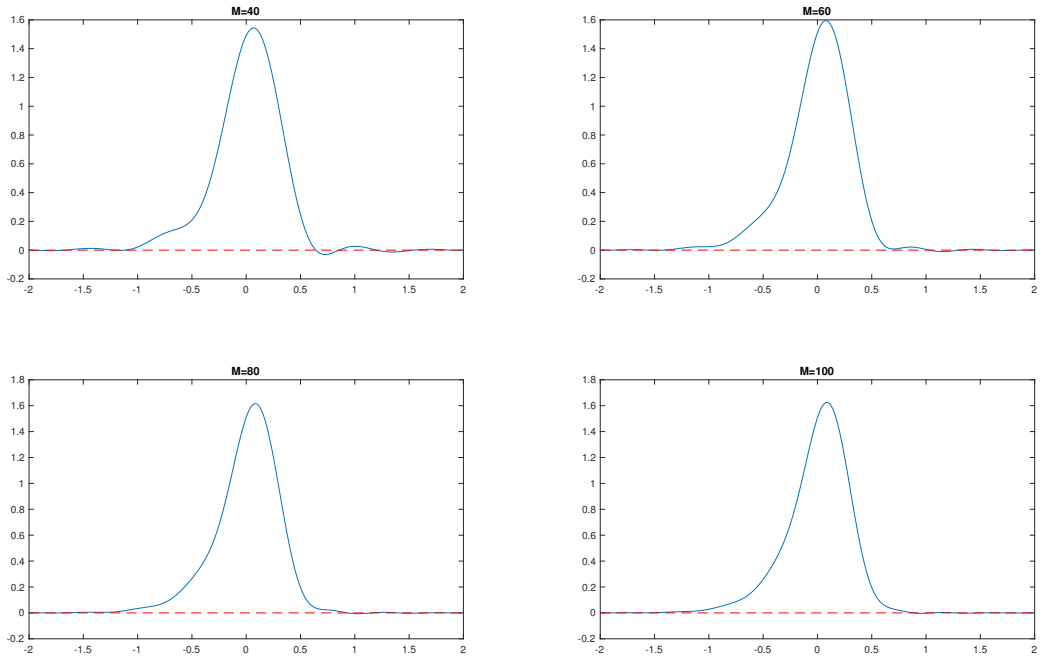


FIGURE 7.1: Plot of the density function $z \to g_T^M(z)$ for different values of M. This analysis shows that for values of M lower than 100 the density may be negative, so we have to carefully choose M. The parameters of the SJV model are the same of the pricing exercise.

This analysis shows that for values of M lower than 100 the density might become negative, so we have to carefully choose M. As we can see in the numerical section of Ackerer et al. (2016), the pricing of European options is accurate also for small values of M, but the fact that the density function $g_T^{(M)}(x)$ can be negative for x in sets which are not negligible, could be an issue, at least from a financial point of view. Nevertheless, the choice of M does not affect the accuracy of the quantization algorithm. In all the numerical examples we will consider the parameters in Table 7.1:

7.5.1 Polynomial Quantization

We use the technique developed in Section 7.3.1, and we compute the quantization grid associated to the log price process at time T, i.e. we approximate X_T using an optimal grid $\Gamma^*_{X_T} = \{x_1, \dots, x_N\}$. The price of a Call option with maturity T and

$\kappa = 1.7$	$\theta = 0.06$	$\sigma = 0.5$	$\rho = -0.5$
$V_0 = 0.1$	$v_{min} = 10^{-2}$	$v_{max} = 1$	$r = 0.04$
$\delta = 0$	$S_0 = 100$	$T = 1$	$M = 80$

TABLE 7.1: Parameters of the SVJ model.

strike K is then approximated as

$$e^{-rT}\mathbb{E}\left[\left(e^{X_T} - K\right)^+\right] \approx e^{-rT}\sum_{i=1}^{N}\left(e^{x_i} - K\right)^+ \mathbb{P}\left(X_T \in C_i(\Gamma^*_{X_T})\right),$$

where $C_i(\Gamma^*) = \left[\frac{x_{i-1}+x_i}{2}, \frac{x_i+x_{i+1}}{2}\right]$ and the weights are given by (7.15). The results in Table 7.2 show that the quantization technique is accurate. Moreover, the computational cost is comparable to the execution time declared in Ackerer et al. (2016), that we used as a benchmark.

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 80$	25.8992	25.8774	0.0841
$K = 85$	22.1171	22.0920	0.1137
$K = 90$	18.6064	18.5621	0.2379
$K = 95$	15.4078	15.3277	0.5198
$K = 100$	12.5529	12.5351	0.1418
$K = 105$	10.0595	10.0666	0.0709
$K = 110$	7.9301	7.9392	0.1148
$K = 115$	6.1520	6.1214	0.4981
$K = 120$	4.7001	4.6816	0.3933

TABLE 7.2: Pricing comparison between the benchmark price and the price obtained via polynomial quantization of a European Call option for the SVJ model with parameters as in Table 7.1. The quantization grids have size $N = 20$.

7.5.2 Recursive Quantization

We use the methodology implemented in Section 7.4.2. Note that we do not exploit the fact that S is the exponential of a polynomial process, but we construct the optimal quantizers starting from the Euler scheme (7.30). We then compute, at every time step t_k, for $k = 1, \dots, L$, such that $t_L = T$, the quantization of the price process S at time t_k, that we call $\widehat{S}_{t_k}$. In order to price a European call option with strike K and maturity T we need only $\Gamma^*_{S,L} = \{s_1^L, \dots, s_N^L\}$, the optimal quantization grid associated to $\widehat{S}_T$, and we have the following approximation:

$$e^{-rT}\mathbb{E}\left[\left(e^{X_T} - K\right)^+\right] \approx \sum_{i=1}^{N}\left(s_i^L - K\right)^+ \mathbb{P}\left(\widehat{S}_T = s_i^L\right),$$

where the weights are computed using (7.31). The results in Table 7.3 show that, since we have simplified the formulas in Callegaro et al. (2016) (Chapter 3), the computational efficiency of the Recursive Marginal quantization for stochastic volatility models is improved.

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 80$	25.8992	25.9082	0.0348
$K = 85$	22.1171	22.1462	0.1315
$K = 90$	18.6064	18.6430	0.1969
$K = 95$	15.4078	15.4395	0.2060
$K = 100$	12.5529	12.5677	0.1175
$K = 105$	10.0595	10.0789	0.1930
$K = 110$	7.9301	7.9508	0.2606
$K = 115$	6.1520	6.1692	0.2792
$K = 120$	4.7001	4.7106	0.2234

TABLE 7.3: Pricing comparison between the benchmark price and the price obtained via recursive quantization of a European Call option for the SVJ model with parameters as in Table 7.1. The quantization grids have size $N_S = 20$, $N_V = 10$ for every time step, and $L = 12$.

7.5.3 Bermudan Options

The advantage of the Recursive Marginal quantization algorithm developed in Section 7.4.1 is the possibility to price path dependent options, since we approximate the process at every time step of the Euler scheme, and the transition densities are given directly by the algorithm, as shown in (7.34). This motivates us to show an application of this methodology to the pricing of Bermudan options. Pricing such

options can be done via a backward procedure on the multinomial tree obtained via quantization, as presented e.g. in Bally et al. (2005, Proposition 2.1). The benchmark used for comparison is a Longstaff Schwarz algorithm of Longstaff and Schwartz (2001), since, up to our knowledge, there is no specific algorithm designed for the pricing of Bermudan option under the SVJ model. The results in Table 7.4 show the accuracy of the methodology.

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 80$	3.0410	2.9984	1.3997
$K = 85$	4.1040	4.1077	0.0899
$K = 90$	5.4579	5.5012	0.7931
$K = 95$	7.1493	7.2222	1.0199
$K = 100$	9.2192	9.3151	1.0404
$K = 105$	11.6984	11.8285	1.1120
$K = 110$	14.6035	14.7564	1.0470
$K = 115$	17.9352	18.0969	0.9015
$K = 120$	21.6788	21.8295	0.6951

TABLE 7.4: Pricing comparison between the benchmark price and the price obtained via recursive quantization of a Bermudan Put option for the SVJ model with parameters as in Table 7.1. The quantization grids have size $N_S = 20$, $N_V = 10$ for every time step, and $L = 12$.

7.6 Conclusion

In this chapter we presented how to apply quantization techniques to polynomial processes. In particular, we focused on the SVJ model, but our results can be extended to any polynomial model. Our analysis on SVJ quantization provided numerical tools to develop fast exotic option pricing algorithms. We followed two approaches. Firstly, we exploited the polynomial property, and we provided theoretical results to assess the approximation errors. As a result we obtained alternative pricing tools for polynomial models, although limited to the dimensionality problems arising in pricing path-dependent derivatives. Numerical examples were provided. Then, we overcame these limitations by applying RMQ to polynomial processes, by viewing them as a particular case of stochastic volatility processes. Numerical examples for Bermudan options were described. We leave for a future work the comparison of these two quantization solutions with Markov chain approximations based on cubature techniques.

7.7 Appendix

7.7.1 Proof of Proposition 7.3.1

Proof. We want to compute the expected value in (7.9). Please note that the function $f^i(y) := (y - x_i)\,\mathbb{1}_{y\in\left[\frac{x_{i-1}+x_i}{2},\frac{x_i+x_{i+1}}{2}\right]} \in L^2_w \quad \forall i = 1,\dots,N$. In fact

$$\begin{aligned}\int_{\mathbb{R}} \left(f^i(y)\right)^2 w(y)dy &\leq \int_{\mathbb{R}} (y-x_i)^2 w(y)dy \\ &= \int_{\mathbb{R}} y^2 w(y)dy - 2x_i \int_{\mathbb{R}} yw(y)dy + x_i^2 \int_{\mathbb{R}} w(y)dy \\ &= \sigma_w^2 + \mu_w^2 - 2x_i\mu_w + x_i^2 \leq \infty \quad \forall i = 1,\dots,N.\end{aligned}$$

Using the polynomial property in (7.5), we can rewrite it in the form of (7.10), where

$$\begin{aligned}f_n^i &= \int_{\mathbb{R}} f^i(y)H_n(y)w(y)dy \\ &= \underbrace{\int_{\mathbb{R}} y\mathbb{1}_{\left[\frac{x_{i-1}+x_i}{2},\frac{x_i+x_{i+1}}{2}\right]}(y)H_n(y)w(y)dy}_{a_n^i} - x_i \underbrace{\int_{\mathbb{R}} \mathbb{1}_{\left[\frac{x_{i-1}+x_i}{2},\frac{x_i+x_{i+1}}{2}\right]}(y)H_n(y)w(y)dy}_{b_n^i}\end{aligned}$$

We focus first on the computation of a_n^i. Let us define

$$h_n(K) = \int_{\mathbb{R}} y\mathbb{1}_{[K,\infty]}(y)H_n(y)w(y)dy,$$

then $a_n^i = h_n\left(\frac{x_{i-1}+x_i}{2}\right) - h_n\left(\frac{x_i+x_{i+1}}{2}\right)$. When $n = 0$ we have, integrating by parts, that

$$\begin{aligned}h_0(K) &= \int_{\mathbb{R}} y\mathbb{1}_{[K,\infty]}(y)w(y)dy \\ &= \sigma_w\phi\left(\frac{K-\mu_w}{\sigma_w}\right) + \mu_w\Phi\left(\frac{\mu_w-K}{\sigma_w}\right).\end{aligned}$$

When $n \geq 1$ we have that

$$\begin{aligned}h_n(K) &= \int_K^\infty yH_n(y)w(y)dy \\ &= \frac{1}{\sqrt{n!}}\int_K^\infty y\mathcal{H}_n\left(\frac{y-\mu_w}{\sigma_w}\right)\frac{1}{\sigma_w}\phi\left(\frac{y-\mu_w}{\sigma_w}\right)dy \\ &= \frac{1}{\sqrt{n!}}\int_{\frac{K-\mu_w}{\sigma_w}}^\infty (\sigma_w z + \mu_w)\,\mathcal{H}_n(z)\,\phi(z)\,dz \\ &= \frac{\sigma_w}{\sqrt{n!}}\int_{\frac{K-\mu_w}{\sigma_w}}^\infty z\mathcal{H}_n(z)\,\phi(z)\,dz + \frac{\mu_w}{\sqrt{n!}}\int_{\frac{K-\mu_w}{\sigma_w}}^\infty \mathcal{H}_n(z)\,\phi(z)\,dz.\end{aligned}$$

We exploit the recursive relation between the Hermite polynomials for $n \geq 1$, : $z\mathcal{H}_n(z) = \mathcal{H}_{n+1}(z) + n\mathcal{H}_{n-1}(z)$, in order to get

$$h_n(K) = \frac{\sigma_w}{\sqrt{n!}}\int_{\frac{K-\mu_w}{\sigma_w}}^\infty \mathcal{H}_{n+1}(z)\,\phi(z)\,dz + \frac{n\sigma_w}{\sqrt{n!}}\int_{\frac{K-\mu_w}{\sigma_w}}^\infty \mathcal{H}_{n-1}(z)\,\phi(z)\,dz + \frac{\mu_w}{\sqrt{n!}}\int_{\frac{K-\mu_w}{\sigma_w}}^\infty \mathcal{H}_n(z)\,\phi(z)\,dz.$$

The case $n = 1$ can be obtained directly using integration by parts, while Ackerer et al. (2016, Theorem 3.7) proves that, for $n \geq 1$,

$$\int_x^\infty \mathcal{H}_n(z)\,\phi(z)\,dz = \mathcal{H}_{n-1}(x)\,\phi(x),$$

so we have the result for $h_n(K)$. The b_n^i coefficients can be computed similarly. In fact, if we define

$$l_n(K) = \int_{\mathbb{R}} \mathbb{1}_{[K,\infty]}(y) H_n(y) w(y) dy,$$

then $b_n^i = l_n\left(\frac{x_{i-1}+x_i}{2}\right) - l_n\left(\frac{x_i+x_{i+1}}{2}\right)$. The case when $n = 0$ is trivial, instead when $n \geq 1$ we have that

$$\begin{aligned}
l_n(K) &= \int_K^\infty H_n(y) w(y) dy \\
&= \frac{1}{\sqrt{n!}} \int_K^\infty \mathcal{H}_n\left(\frac{y-\mu_w}{\sigma_w}\right) \frac{1}{\sigma_w} \phi\left(\frac{y-\mu_w}{\sigma_w}\right) dy \\
&= \frac{1}{\sqrt{n!}} \int_{\frac{K-\mu_w}{\sigma_w}}^\infty \mathcal{H}_n(z)\,\phi(z)\,dz \\
&= \frac{1}{\sqrt{n!}} \mathcal{H}_{n-1}\left(\frac{K-\mu_w}{\sigma_w}\right) \phi\left(\frac{K-\mu_w}{\sigma_w}\right).
\end{aligned}$$

□

7.7.2 Proof of Proposition 7.3.2

Proof. Remember that g_T is the density of X_T. We can then rewrite $E_i(x_1, \ldots, x_N)$ as

$$E_i(x_1, \ldots, x_N) = \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} y g_T(y) dy - x_i \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} g_T(y) dy.$$

This shows that E_i depends only on x_{i-1}, x_i and x_{i+1}, so that the Jacobian matrix J is tridiagonal. Moreover the lower diagonal has components:

$$\begin{aligned}
J_{i,i-1} = \frac{\partial E_i}{\partial x_{i-1}}(x_{i-1}, x_i, x_{i+1}) &= -\frac{1}{2}\frac{x_{i-1}+x_i}{2} g_T\left(\frac{x_{i-1}+x_i}{2}\right) + \frac{1}{2} x_i g_T\left(\frac{x_{i-1}+x_i}{2}\right) \\
&= \frac{1}{2}\left(\frac{x_i - x_{i-1}}{2}\right) g_T\left(\frac{x_{i-1}+x_i}{2}\right).
\end{aligned}$$

The upper diagonal:

$$\begin{aligned}
J_{i,i+1} = \frac{\partial E_i}{\partial x_{i+1}}(x_{i-1}, x_i, x_{i+1}) &= \frac{1}{2}\frac{x_i+x_{i+1}}{2} g_T\left(\frac{x_i+x_{i+1}}{2}\right) - \frac{1}{2} x_i g_T\left(\frac{x_i+x_{i+1}}{2}\right) \\
&= \frac{1}{2}\left(\frac{x_{i+1} - x_i}{2}\right) g_T\left(\frac{x_i+x_{i+1}}{2}\right).
\end{aligned}$$

We can deduce immediately that $J_{i,i-1} = J_{i-1,i}$, so that J is also symmetric. Finally the diagonal has components:

$$\begin{aligned} J_{i,i} = \frac{\partial E_i}{\partial x_i}(x_{i-1}, x_i, x_{i+1}) &= \frac{1}{2}\frac{x_i + x_{i+1}}{2} g_T\left(\frac{x_i + x_{i+1}}{2}\right) - \frac{1}{2}\frac{x_i + x_{i+1}}{2} g_T\left(\frac{x_i + x_{i+1}}{2}\right) \\ &\quad - \frac{1}{2}x_i\left(g_T\left(\frac{x_i + x_{i+1}}{2}\right) - g_T\left(\frac{x_{i-1} + x_i}{2}\right)\right) - \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} g_T(y)dy \\ &= J_{i,i-1} + J_{i,i+1} - \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} g_T(y)dy, \end{aligned}$$

and the integral in the last equality is exactly the weight of the i-th Voronoi cell. The expression for the density in (7.14) comes from the following fact: the pricing of a derivative with payoff f is, see (7.5),

$$\begin{aligned} \mathbb{E}\left[f(X_T)\right] &= \sum_{n\geq 0} f_n \ell_n \\ &= \sum_{n\geq 0} \int_{\mathbb{R}} f(y) H_n(y) w(y) dy \ell_n \\ &= \int_{\mathbb{R}} f(y) \sum_{n\geq 0} H_n(y) \ell_n w(y) dy, \end{aligned}$$

where the fact that we can change the order of the infinite sum and the integral is proved in Ackerer et al. (2016). Since the price of the derivative can be seen also as

$$\int_{\mathbb{R}} f(y) g_T(y) dy,$$

(7.14) follows. Finally, the expression for $\mathbb{P}\left(X_T \in C_i(\Gamma)\right)$ comes immediately from the proof of Theorem 7.3.1. □

7.7.3 Proof of Lemma 7.3.3

Proof. First of all remember that

$$\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right| = \left|e^{-rT}\int_{\mathbb{R}} (e^x - K)^+ g_T^{(M)}(x)dx - e^{-rT}\sum_{i=1}^{N}\int_{C_i(\Gamma^X)} (e^{x_i} - K)^+ g_T^{(M)}(x)dx\right|.$$

By introducing $s := e^x$ (notice that the payoff $(s-K)^+$ is Lipschitz with respect to s and this will be crucial), we have

$$\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right| = e^{-rT}\left|\int_0^{\infty} (s-K)^+ \frac{g_T^{(M)}(\ln s)}{s} ds - \sum_{i=1}^{N}\int_{C_i(\Gamma^S)} (s_i - K)^+ \frac{g_T^{(M)}(\ln s)}{s} ds\right|$$

where $S_T^{(M)}$ is a random variable with density $h_T^{(M)}(s) := \frac{g_T^{(M)}(\ln(s))}{s}$ for $s \in (0, +\infty)$ and where $\Gamma^S = \{s_1, \ldots, s_N\}$ is an N-quantizer for $S_T^{(M)}$. We denote by $\widehat{S}_T^{(M,N)}$ the quantization of $S_T^{(M)}$ on Γ^S. We have

$$\mathbb{P}\left(\widehat{S}_T^{(M,N)} = s_i\right) = \int_{C_i(\Gamma^S)} h_T^{(M)}(s) ds.$$

Thus, working on the error $\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right|$ corresponds to estimating the error coming from pricing a European Call option on $S_T^{(M)}$ via quantization. Now, for every Lipschitz function f with Lipschitz constant $[f]_{\text{Lip}}$, using twice the tower property, we have the following (classic) result:

$$\begin{aligned}\left|\pi_f^{(M)} - \widehat{\pi}_f^{(M,N)}\right| &= \left|\mathbb{E}\left[f\left(S_T^{(M)}\right) - f\left(\widehat{S}_T^{(M,N)}\right)\right]\right| \\ &\leq \mathbb{E}\left[\left|\mathbb{E}\left[f\left(S_T^{(M)}\right)|\widehat{S}_T^{(M,N)}\right] - f\left(\widehat{S}_T^{(M,N)}\right)\right|\right] \\ &\leq [f]_{\text{Lip}}\left|\left|S_T^{(M)} - \widehat{S}_T^{(M,N)}\right|\right|_1 \\ &\leq [f]_{\text{Lip}}\left|\left|S_T^{(M)} - \widehat{S}_T^{(M,N)}\right|\right|_2,\end{aligned}$$

where

$$||S_T^{(M)} - \widehat{S}_T^{(M,N)}||_r = \left(\sum_{i=1}^N \int_{C_i(\Gamma^S)} |s - s_i|^r \, h_T^{(M)}(s)ds\right)^{\frac{1}{r}}$$

is the L^r distance between the random variable with density $h_T^{(M)}$ and its quantization with N points. The Lipschitz constant for the payoff of a Call option is 1, and we have the result. □

7.7.4 Proof of Theorem 7.3.4

Proof. It is worth noticing that in this polynomial setting, by definition (recall Equation (7.17)), the density $g_T^{(M)}(s), s \in (0, +\infty)$, behaves like $s^M e^{-\frac{s^2}{2}}$, so that $h_T^{(M)}(s)$ behaves like $(\ln s)^M e^{-\frac{(\ln s)^2}{2}} \frac{1}{s} =: \widetilde{h}_T^{(M)}(s)$ at 0 and at infinity.
The proof of Callegaro et al. (2017b, Theorem 2.11) (Theorem 6.2.6) consists of five steps, from zero to four. We now adapt it to our setting and in the case of quadratic quantization, namely in the case when p in the cited paper is equal to 2. The first three steps remain the same, so we sketch them.
Step 0
We have to prove that $||h_T^{(M)}||_{\frac{1}{p+1}} = ||h_T^{(M)}||_{\frac{1}{3}} < +\infty$. We hence study the convergence at 0 and at $+\infty$ of the integral of $\left(\widetilde{h}_T^{(M)}\right)^{\frac{1}{3}}$. In the rest of the proof, without loss of generality, we will assume that M is a multiple of 3, so that computations will be explicit. If we denote by $\overline{M} := \frac{M}{3}$, then, a primitive function is

$$\int \left(\widetilde{h}_T^{(M)}(s)\right)^{\frac{1}{3}} ds = \beta_{\overline{M}} \, Erf\left(\frac{-2 + \ln s}{\sqrt{6}}\right) + \sum_{n=1}^{\overline{M}} \alpha_n \, s^{\frac{2}{3}} \, e^{-\frac{1}{6}\ln^2 s} \, (\ln s)^{\overline{M}-n} =: \widetilde{H}_T^M(s),$$

where Erf is the error function, defined as $Erf(s) = \frac{2}{\pi}\int_0^s e^{-t^2}dt$, and the coefficients $\alpha_n, n = 1, \ldots, \overline{M}$ and $\beta_{\overline{M}}$ can be explicitly computed, e.g. using a symbolic programming language as *Mathematica*. Given that

$$\lim_{s\to+\infty} \widetilde{H}_T^M(s) = \beta_{\overline{M}},$$

and

$$\lim_{s\to 0} \widetilde{H}_T^M(s) = -\beta_{\overline{M}},$$

we obtain the finiteness of $||h_T^{(M)}||_{\frac{1}{3}}$.

Step 1

Here it can be shown the following estimation for the distortion function D associated to $S_T^{(M)}$ and calculated in a generic grid $\Gamma = \{s_1, \dots, s_N\}$:

$$D(s_1,\dots,s_N) \leq \int_0^{s_1} (s_1 - y)^2 h_T^{(M)}(s)ds + \sum_{i=1}^{N-1} \frac{h_T^{(M)}(\xi_i) + h_T^{(M)}(\xi_{i+1})}{3} \left(\frac{s_{i+1} - s_i}{2}\right)^3 + \int_{s_N}^{+\infty} (y - s_N)^2 h_T^{(M)}(s)ds,$$

for some $\xi_1, \dots, \xi_N \in \mathbb{R}$.

Step 2

There exists a grid $\overline{\Gamma} = \{\bar{s}_1, \dots, \bar{s}_N\}$, and $\zeta_1, \dots, \zeta_{N-1}$, with $\zeta_i \in [\bar{s}_i, \bar{s}_{i+1}]$, such that

$$\int_0^{\bar{s}_i} \left(h_T^{(M)}(s)\right)^{\frac{1}{3}} ds = \int_{\bar{s}_N}^{+\infty} \left(h_T^{(M)}(s)\right)^{\frac{1}{3}} ds = \frac{1}{2N} ||h_T^{(M)}||_{\frac{1}{3}}^{\frac{1}{3}},$$

and

$$(\bar{s}_{i+1} - \bar{s}_i)^2 = \frac{||h_T^{(M)}||_{\frac{1}{3}}^{\frac{2}{3}}}{\left(h_T^{(M)}(\zeta_i)\right)^{\frac{2}{3}} N^2}.$$

Step 3

We provide the following bound for the quantization error:

$$||S_T^{(M)} - \widehat{S}_T^{(M,N)}||_2^2 \leq \frac{||h_T^{(M)}||_{\frac{1}{3}}^{\frac{2}{3}}}{24N^2} \sum_{i=1}^{N-1} \frac{h_T^{(M)}(\bar{\xi}_i) + h_T^{(M)}(\bar{\xi}_{i+1})}{\left(h_T^{(M)}(\zeta_i)\right)^{\frac{2}{3}}} (\bar{s}_{i+1} - \bar{s}_i) + \int_0^{\bar{s}_1} (\bar{s}_1 - s)^2 h_T^{(M)}(s)ds + \int_{\bar{s}_N}^{+\infty} (s - \bar{s}_N)^2 h_T^{(M)}(s)ds.$$

Step 4

In Step 2 we have proved that

$$\frac{1}{N^2} = \frac{4}{||h_T^{(M)}||_{\frac{1}{3}}^{\frac{2}{3}}} \left(\int_0^{\bar{s}_i} \left(h_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2 = \frac{4}{||h_T^{(M)}||_{\frac{1}{3}}^{\frac{2}{3}}} \left(\int_{\bar{s}_N}^{+\infty} \left(h_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2,$$

So in order to prove that, when $N \to \infty$, $\int_0^{\bar{s}_1} (\bar{s}_1 - s)^2 h_T^{(M)}(s)ds$ and $\int_{\bar{s}_N}^{+\infty} (s - \bar{s}_N)^2 h_T^{(M)}(s)ds$ are $o\left(\frac{1}{N^2}\right)$, we just need to prove that

$$\lim_{y \to +\infty} \frac{\int_y^{+\infty} (s-y)^2 h_T^{(M)}(s)ds}{\left(\int_y^{+\infty} \left(h_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2} = 0,$$

and that

$$\lim_{y\to 0}\frac{\int_0^y (s-y)^2\, h_T^{(M)}(s)ds}{\left(\int_0^y \left(h_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2} = 0.$$

Since, both at 0 and at infinity, $h_T^{(M)} \sim \widetilde{h}_T^{(M)}$, we can equivalently prove that

$$\lim_{y\to+\infty}\frac{\int_y^{+\infty} s^2\, \widetilde{h}_T^{(M)}(s)ds - 2y\int_y^{+\infty} s\, \widetilde{h}_T^{(M)}(s)ds + y^2\int_y^{+\infty} \widetilde{h}_T^{(M)}(s)ds}{\left(\int_y^{+\infty} \left(\widetilde{h}_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2} = 0,$$

and that

$$\lim_{y\to 0}\frac{\int_0^y s^2\, \widetilde{h}_T^{(M)}(s)ds - 2y\int_0^y s\, \widetilde{h}_T^{(M)}(s)dz + y^2\int_0^y \widetilde{h}_T^{(M)}(s)ds}{\left(\int_0^y \left(\widetilde{h}_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2} = 0.$$

Up to a constant, we have that, for $\ell = 0, 1, 2$,

$$\int s^\ell \widetilde{h}_T^{(M)}(s)ds = \beta_{\ell,M} Erf\left(\frac{-\ell + \ln s}{\sqrt{2}}\right) + \sum_{n=0}^{M-1} \alpha_{\ell,n} e^{-\frac{1}{2}(\ln s)^2} s^\ell\, (\ln s)^n =: \widetilde{H}_{\ell,T}^{(M)}(s),$$

where, as before, Erf is the error function and $\alpha_{\ell,n}$ and $\beta_{\ell,M}$ can be computed, for $\ell = 0, 1, 2$, with *Mathematica* in closed form. Please note that $\lim_{y\to+\infty} \widetilde{H}_{\ell,T}^{(M)}(y) = \beta_{\ell,M}$ and $\lim_{y\to 0} \widetilde{H}_{\ell,T}^{(M)}(y) = -\beta_{\ell,M}$, for $\ell = 0, 1, 2$. We have than that

$$\lim_{y\to+\infty}\frac{\int_y^{+\infty} s^2\widetilde{h}_T^{(M)}(s)dz - 2y\int_y^{+\infty} s\widetilde{h}_T^{(M)}(s)ds + y^2\int_y^{+\infty} \widetilde{h}_T^{(M)}(s)ds}{\left(\int_y^{+\infty} \left(\widetilde{h}_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2} =$$

$$\lim_{y\to+\infty}\frac{\beta_{2,M} - \widetilde{H}_{2,T}^{(M)}(y) - 2y\left(\beta_{1,M} - \widetilde{H}_{1,T}^{(M)}(y)\right) + y^2\left(\beta_{0,M} - \widetilde{H}_{0,T}^{(M)}(y)\right)}{\left(\beta_M - \widetilde{H}_T^{(M)}(y)\right)^2} = 0,$$

and, in a similar way,

$$\lim_{y\to 0}\frac{\int_0^y s^2\widetilde{h}_T^{(M)}(s)ds - 2y\int_0^y s\widetilde{h}_T^{(M)}(s)ds + y^2\int_0^y \widetilde{h}_T^{(M)}(s)ds}{\left(\int_0^y \left(\widetilde{h}_T^{(M)}(s)\right)^{\frac{1}{3}} ds\right)^2} =$$

$$\lim_{y\to 0}\frac{\widetilde{H}_{2,T}^{(M)}(y) + \beta_{2,M} - 2y\left(\widetilde{H}_{1,T}^{(M)}(y) + \beta_{1,M}\right) + y^2\left(\widetilde{H}_{0,T}^{(M)}(y) + \beta_{0,M}\right)}{\left(\widetilde{H}_T^{(M)}(y) + \beta_M\right)^2} = 0.$$

Since

$$\sum_{i=1}^{N-1} \frac{h_T^{(M)}(\bar{\xi}_i) + h_T^{(M)}(\bar{\xi}_{i+1})}{\left(h_T^{(M)}(\zeta_i)\right)^{\frac{2}{3}}} \left(\bar{s}_{i+1} - \bar{s}_i\right) \to 2||h_T^{(M)}||_{\frac{1}{3}}^{\frac{1}{3}},$$

when $N \to +\infty$, we conclude that

$$\lim_{N\to+\infty} N^2 ||S_T^{(M)} - \widehat{S}_T^{(M,N)}||_2^2 \leq \frac{1}{12} ||h_T^{(M)}||_{\frac{1}{3}}.$$

□

Chapter 8

Conic quantization

This is a joint work with Wim Schoutens.

In this chapter we introduce a new technology for the pricing of European options under a wide class of price processes. The method is based on an evolution of the quantization technique introduced in Callegaro et al. (2017b) (Chapter 6), and is quick and accurate enough to calibrate on financial data. Moreover, it allows to introduce the concept of market implied liquidity, linking the pricing under stochastic volatility with the Conic Finance theory of two prices, studied in Madan and Schoutens (2016). As a motivating example, we construct for the first time the market implied liquidity surface under stochastic volatility.

8.1 Introduction

We introduce an alternative technique for the quantization of a stochastic process which has a characteristic function in closed form. Using some ideas from Fourier analysis, as already seen in Chapter 6, it is possible to compute the (optimal) quantization of a stochastic process at a given maturity in a fast and accurate way. This quantization technique can be used, together with the theory of Conic Finance, to compute bid and ask prices for European put and call options in every model for which the characteristic function can be computed.
As studied in Callegaro et al. (2017b) (Chapter 6), it is possible to derive the density of the process directly from the characteristic function, and this result, together with the knowledge of the optimal quantizers for uniformly distributed random variables, leads to the computation of the quantizers for the process at a given time (usually the maturity of the European option). This quantizers are optimal in the sense that they are the best approximation of the time process at maturity using a finite number of points, where every point has the same weight of the others.
The Conic Finance theory deals with the existence of two prices in the market for a derivative, i.e. the best ask and best bid prices. These prices are computed by using a transformation of the distribution of the underlying, via the so-called distortion functions. The pricing of bid-ask prices for European Calls and Puts under non trivial dynamics for the underlying has been designed in a Monte Carlo framework, see Madan and Schoutens (2016, Chapter 3.3). In this chapter we introduce the application of the distortion functions to the optimal quantizers of a given model at a given maturity. Up to our knowledge, this is the first technique which considers the presence of a liquidity factor in the pricing of a derivative under stochastic volatility.
The chapter is organized as follows: in Section 8.2 we recall some notion on Conic Finance, especially in the distortion of the distribution density of the price. In Section

8.3 we introduce the new quantization algorithm, and we connect Quantization and Conic Finance. In Section 8.4 we show the performance of the method on pricing options and bid-ask spreads. Finally, in Section 8.5 we present a calibration exercise on market data, and we introduce the concept of implied liquidity under stochastic volatility. Section 8.6 concludes.

8.2 Conic Finance and the bid - ask spread pricing

Let us consider a filterered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$, equipped with a risk-neutral measure $\mathbb{Q}$ and such that the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual assumptions and models the randomness in the dynamics of the models considered. We also work under the assumptions of zero cost cash flows and constant interest rate factor r. When pricing a derivative, that in our case will be a European Call option on the underlying S, with strike K and maturity T, we typically look at zero cost cash flows of the following form:

$$(S_T - K)^+ - e^{rT}b,$$

in the case we agree to pay at time T a cash amount $e^{rT}b$, i.e., equivalently, a cash amount b at time 0, and receive the payoff, or

$$e^{rT}a - (S_T - K)^+,$$

in the case we agree to receive at time T a cash amount $e^{rT}a$, i.e., equivalently, a cash amount a at time 0, and payout the payoff. We will call b the bid price, that is the price the market wants to pay for the risk $(S_T - K)^+$, and a the ask price, the price that the market wants to receive for the risk $(S_T - K)^+$. Any price higher than b and any price lower than a will make these prices unacceptable. Under this construction, the risk neutral price will be lower than the ask price and greater than the bid price. In fact, let us consider the set of zero cost cash flows Z defined as

$$\mathcal{A}^* = \{Z | e^{-rT}\mathbb{E}^{\mathbb{Q}}[Z] \geq 0\},$$

The set $\mathcal{A}^*$ is an half space containing all the non-negative random variables. This means that we are considering all the possible arbitrages, but we need to take into consideration also zero cost cash flows that are not necessary arbitrages, in our set of acceptable cash flows. Note that the non-negative random variables constitute a convex cone, and also $\mathcal{A}^*$ is a convex cone. The half space $\mathcal{A}^*$ is then too big for our purposes, and in our two price economy we will consider a set $\mathcal{A}$ for the acceptable zero cost cash flows such that $\mathcal{A}$ contains all the arbitrages and $\mathcal{A} \subset \mathcal{A}^*$. It is shown in Artzner et al. (1999) that the elements of $\mathcal{A}$ can be defined by a convex set of probability measures $\mathcal{M}$ where

$$Z \in \mathcal{A} \iff e^{-rT}\mathbb{E}^{Q}[Z] \geq 0 \quad \text{for all } Q \in \mathcal{M}.$$

As we have seen, the market accepts the risk to buy at the best bid price b and the risk to sell at the best ask price a, this means that the zero cost cash flows $(S_T - K)^+ - e^{rT}b$ and $e^{rT}a - (S_T - K)^+$ are acceptable. Equivalently this means that for all $Q \in \mathcal{M}$

$$(S_T - K)^+ - e^{rT}b \in \mathcal{A} \iff b \leq e^{-rT}\mathbb{E}^{Q}[(S_T - K)^+],$$

and

$$e^{rT}a - (S_T - K)^+ \in \mathcal{A} \iff a \geq e^{-rT}\mathbb{E}^Q[(S_T - K)^+].$$

This construction shows the intuition that the risk neutral price of the European option will be between the bid and the ask price. Finally, the best bid and the best ask price at time zero for the European call are given by

$$bid((S_T - K)^+) = e^{-rT} \inf_{Q \in \mathcal{M}} \mathbb{E}^Q[(S_T - K)^+], \tag{8.1}$$

and

$$ask((S_T - K)^+) = e^{-rT} \sup_{q \in \mathcal{M}} \mathbb{E}^Q[(S_T - K)^+]. \tag{8.2}$$

We refer to this construction as Conic Finance, since the prices are defined by convex cones of acceptable zero cost cash flows.
A market model can then be constructed giving a specification of a set of supporting measures, and the bid and ask prices can be computed as in (8.1) and (8.2). In Cherny and Madan (2009) the authors define operational cones, and make use of concave distortion function in order to construct bid and ask prices for European options. A concave distortion function is a concave function which maps the unit interval to itself. It is shown in Kusuoka (2001) that the bid and ask price must be an expectation under a concave distortion. More specifically, it is possible to prove that there exists a concave distortion $\Psi(u)$ from the unit interval to itself such that the price of a Call option on the underlying S with strike K and maturity T is given by

$$bid((S_T - K)^+) = e^{-rT} \int_{\mathbb{R}} (x - K)^+ d\Psi\left(F_{S_T}(x)\right),$$

where $F_{S_T}(x)$ is the risk neutral distribution of S at time T.
In this work we will use only the MINMAXVAR distortion function, studied in Cherny and Madan (2010), defined as

$$\Psi_\lambda(u) = 1 - \left(1 - u^{\frac{1}{\lambda+1}}\right)^{1+\lambda},$$

where λ is a positive real value.

This is all we need from the Conic Finance theory. The interested reader can found additional details in Madan and Schoutens (2016).

8.3 The uniform Fourier quantization algorithm

The link between Fourier analysis and quantization techniques has been already studied and implemented in Callegaro et al. (2017b) (Chapter 6). Here we use a different approach, in order to improve the efficiency of the pricing algorithm, as the numerical results in Section 8.4 will show.

8.3.1 A simulation result

We consider an $\mathbb{R}$-valued random variable X. We call F the distribution function of X, i.e.

$$F(x) = \mathbb{P}(X \leq x), \quad x \in \mathbb{R}.$$

F is always non-decreasing, cadlag and such that $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to+\infty} F(x) = 1$, and it is possible to associate to F the canonical left inverse F_l^{-1}:

$$F_l^{-1}(u) = \inf\{x|F(x) \geq u\}.$$

Then $F_l^{-1}(u)$ is non decreasing, left continuous and satisfies

$$F_l^{-1}(u) \leq x \iff u \leq F(x)$$

It is then immediate to see that, if U is a uniform random variable in $[0,1]$, then $F_l^{-1}(U) \stackrel{d}{=} X$. This means that, for every measurable function $h \in L^1$, the following holds:

$$\mathbb{E}\left[h(X)\right] = \mathbb{E}\left[h(F_l^{-1}(U))\right]. \tag{8.3}$$

8.3.2 Optimal quantization of uniform random variables

The link between the distribution of a generic random variable and (a modification of) the distribution of a uniform random variable has important applications also in quantization. It is easy to prove, see Graf and Luschgy (2000, Example 4.17), that the optimal quantization grid of size N, $\Gamma = \{u_1, \ldots, u_N\}$, for a uniform random variable $U \sim \mathcal{U}([0,1])$ is given by

$$u_j = \frac{2j-1}{2N}, \quad j = 1, \ldots, N.$$

Let us define the optimal quantizer $\widehat{U}^\Gamma$ of U in the canonical way, i.e.,

$$\widehat{U}^\Gamma = \sum_{j=1}^{N} u_j \mathbb{1}_{C_j(\Gamma)}(U),$$

where $C_j(\Gamma)$ is the Voronoi cell associated to u_j, namely

$$C_j(\Gamma) = \left[\frac{j-1}{N}, \frac{j}{N}\right] \quad j = 1, \ldots, N.$$

Note that, because of the geometric construction of the Voronoi cells, it follows immediately that

$$\mathbb{P}\left(\widehat{U}^\Gamma = u_j\right) = \mathbb{P}\left(U \in C_j(\Gamma)\right) = \frac{1}{N} \quad \forall j = 1, \ldots, N.$$

Motivated by these properties, we can approximate the expectation of a generic measurable function $g : [0,1] \to \mathbb{R}$ in the following way:

$$\mathbb{E}\left[g(U)\right] \approx \sum_{j=1}^{N} g(u_j)\mathbb{P}\left(U \in C_j(\Gamma)\right) = \frac{1}{N}\sum_{j=1}^{N} g(u_j),$$

so the expectation in (8.3) can be approximated, setting $g = h \circ F_l^{-1}$, as

$$\mathbb{E}\left[h(X)\right] = \mathbb{E}\left[h(F_l^{-1}(U))\right] \approx \frac{1}{N}\sum_{j=1}^{N} h\left(F_l^{-1}(u_j)\right). \tag{8.4}$$

The study of the error boils down to the study of the convergence of the optimal quantizer to the original random variable. As seen in Theorem 1.1.1, it is possible to have an *a priori* estimation of the convergence, which is linear. This is a reason why quantization should be preferred to Monte Carlo simulations, where the rate of convergence is of order $o\left(\frac{1}{\sqrt{N}}\right)$, where N is the number of simulations.

8.3.3 A Fourier quantization approach

The issue now is how to derive F_l^{-1} in a closed form or, at least, efficiently. From now on we will consider X as the value of a log price process S at maturity T, and h as the payoff of a derivative at maturity (for example, in the case of a call option with strike K, $h(x) = (e^x - K)^+$).
In the case where S follows a Black Scholes dynamics, it is possible to compute the distribution function F of the log process at time T using the log normal property of S, but this is not always possible when we consider more complicated models, such as affine or Lévy processes. We will focus on processes for which the characteristic function can be easily computed.
Suppose that the characteristic function of the log process

$$\phi_T(u) := \mathbb{E}\left[e^{iu\log(S_t)}\right], \quad u \in \mathbb{R} \tag{8.5}$$

is known in closed form, or it can be easily computed. We omit the dependence from the initial value of the price process at time 0.
The characteristic function can be seen as the Fourier - Laplace transform of the density of the price process, so it can be proved (see e.g. Shephard (1991)) that

$$F(x) = \mathbb{P}\left(\log(S_T) \leq x\right) = \frac{1}{2} - \frac{1}{\pi}\int_0^{+\infty} Re\left(\frac{e^{-iux}\phi_T(u)}{iu}\right) du. \tag{8.6}$$

If we assume enough regularity of the characteristic function $\phi_T(u)$, the distribution function of the log price $F(x)$ is continuous, so the canonical left inverse is exactly the inverse. Up to our knowledge it is not possible to write F^{-1} in a closed form using the characteristic function ϕ_T, but we can find a numerical way to solve the problem. Motivated by (8.4), we need to compute

$$y_j := F^{-1}\left(\frac{2j-1}{2N}\right) \quad \forall j = 1, \dots, N$$

The determination of the values of the vector $\mathbf{y} = (y_1, \dots, y_N)$ boils down to solve the following system:

$$G_j(\mathbf{y}) := F(y_j) - \frac{2j-1}{2N} = 0, \quad j = 1, \dots, N. \tag{8.7}$$

This system can be easily solved computing the function F in a fine grid of points, and then the solution $\mathbf{y} = (y_1, \dots, y_N)$ can be obtained via interpolation. This approach is much faster then the Newton - Raphson algorithm implemented in Callegaro et al. (2017b) (Chapter 6). In fact we are able to compute prices for European options in milliseconds, improving by a factor 15 the performance.

8.3.4 Conic pricing via quantization

The idea is similar to the one in Madan and Schoutens (2016, Chapter 5.3) but instead of simulating the stock price paths using Monte Carlo, we use the quantization approach. Note that, since the function F is non-decreasing, and since $u_1 \leq u_2 \leq \cdots \leq u_N$, then also $y_1^* \leq \cdots \leq y_N^*$. If we consider the pricing of a call option with strike K, maturity T, and considering a risk free interest rate r, then we get that

$$\text{call}_{K,T} = \mathbb{E}e^{-rT}\left[(e^{\log(S_T)} - K)^+\right] \approx \frac{1}{N}e^{-rT}\sum_{j=1}^{N}(e^{y_j^*} - K)^+. \tag{8.8}$$

In the case of Call options, if we define $\text{payoff}_j = (e^{y_j^*} - K)^+$, then $\text{payoff}_1 \leq \cdots \leq \text{payoff}_N$. This observation is important because in conic pricing we need to attach higher weights to lower payoffs and lower weights to higher payoffs in order to calculate the bid price.
In order to compure the ask and the bid prices, instead of assigning a weight $\frac{1}{N}$ to payoff_j, we assign a distorted weight to it. Let us call Ψ a generic distortion function. Then, according to the conic pricing theory, payoff_j receives the weight

$$p_j^* = \Psi\left(\frac{j}{N}\right) - \Psi\left(\frac{j-1}{N}\right), \quad j = 1\ldots, N, \tag{8.9}$$

for the bid price and

$$\tilde{p}_j = \Psi\left(\frac{N-j+1}{N}\right) - \Psi\left(\frac{N-j}{N}\right), \quad j = 1\ldots, N, \tag{8.10}$$

for the ask price. So we have that

$$\text{call}_{K,T}^{\text{bid}} \approx e^{-rT}\sum_{j=1}^{N}(e^{y_j^*} - K)^+ p_j^*, \qquad \text{call}_{K,T}^{\text{ask}} \approx e^{-rT}\sum_{j=1}^{N}(e^{y_j^*} - K)^+ \tilde{p}_j.$$

8.4 Numerical examples

In this section we will show that this pricing technique, when computing the mid price of a Vanilla option, is comparable to a standard FFT approach, and then we will show how we can create bid-ask smiles. The quantization method previously developed is very general, in the sense that the pricing algorithm is not tailor-made for a particular model. Instead, as long as the characteristic function is known, then it is possible to price European options on every underlying. This allows us to price a huge set of models, and we present here a (non exhaustive) collection:

- Stochastic volatility:
  - Heston model of Heston (1993).
  - Bates model of Bates (1996).
  - Bi Heston model of Christoffersen et al. (2009).
  - Hull - White model of Hull and White (1987).
  - Stein - Stein model of Stein and Stein (1991).
  - Wishart model of Da Fonseca et al. (2008).

- Exponential Levy:
  - Variance Gamma model of Madan et al. (1998).
  - Normal Inverse Gaussian model of Barndorff-Nielsen (1997).
  - CGMY model of Carr et al. (2002).
  - Tempered Stable model of Koponen (1995).
- Non affine:
  - $\frac{3}{2}$ model of Platen (1997).
  - $\frac{4}{2}$ model of Grasselli (2016).

8.4.1 Pricing exercise

We present here a numerical example for the pricing of a European option. We recall that the price of a Call option on the underlying S with maturity T and strike K can be approximated as

$$\text{call}_{K,T} \approx \frac{1}{N} e^{-rT} \sum_{j=1}^{N} (e^{y_j^*} - K)^+,$$

where y_j^* are the solution of the system (8.7). We present here two examples for the Heston model, which is the most famous and most used stochastic volatility model, and the Tempered Stable model, which generalizes exponential Lévy models. The application of our methodology to other models is straightforward.

The Heston model

The Heston model considers the following dynamics for the pair of price and volatility (S, V):

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{V_t}(\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2), \\ dV_t &= \kappa(\theta - V_t) + \xi \sqrt{V_t} dW_t^1, \end{aligned}$$

where V denotes the instantaneous variance of the price process and it is modeled as a CIR process. Here W^1 and W^2 are two independent standard Brownian motions and r is the interest rate. The parameters of the models are θ (the long-run average variance), κ (the speed of the mean reversion of the variance), ξ (the vol of vol parameter) and the correlation ρ.

The Tempered Stable model

The Tempered Stable model has been deeply studied in Cont et al. (1997), assumes that the stock has the dynamic of an exponential Lévy process:

$$S_t = S_0 \exp[(r + \gamma_c)t + X_t], \quad t \in [0, T], \tag{8.11}$$

where γ_c is needed to ensure under the risk neutral measure and X denotes a Lévy process associated with a tempered stable distribution. Tempered stable distributions form a six parameter family of infinitely divisible distributions. They include

Strike	Bench.price	Q price	Rel. error (bp)	ImpVol sqdiff (10^{-8})
$K = 80.0$	25.9677	25.9587	3.4800	14.4584
$K = 85.0$	22.2394	22.2316	3.4921	7.5735
$K = 90.0$	18.7952	18.7877	4.0037	5.4109
$K = 95.0$	15.6696	15.6632	4.0685	3.1682
$K = 100.0$	12.8884	12.8821	4.8749	2.6873
$K = 105.0$	10.4650	10.4585	6.1758	2.6531
$K = 110.0$	8.3985	8.3910	8.8745	3.5080
$K = 115.0$	6.6730	6.6667	9.4736	2.6618
$K = 120.0$	5.2601	5.2517	15.9185	5.1973

TABLE 8.1: Pricing comparison between the benchmark price and the price obtained via quantization of a Call option for the Heston model with parameters $\kappa = 2.3924$, $\xi = 0.6903$, $V_0 = 0.0719$, $\theta = 0.0929$, $\rho = -0.8210$, $r = 0.04$, $T = 1$. The quantization grids have size $\mathbf{N = 1000}$, and the computational time is around 30 **milliseconds**. The benchmark price is computed using an FFT pricer.

the Variance Gamma distributions of Madan et al. (1998), the bilateral Gamma distributions of Kuchler and Tappe (2008) and the CGMY distributions of Carr et al. (2002). Here we assume that $X \sim TS(c^+, \alpha^+, \lambda^+, c^-, \alpha^-, \lambda^-)$. The parameters λ^+ and λ^- describe the rates of the tail decay, α^+ and α^- involve positive and negative jumps, and c^+ and $c^-$are the arrival rate of jumps of given size.

The Variance Gamma model, introduced by Madan et al. (1998), is an important specification of the tempered stable family. In this framework, the Conic quantization approach is a generalization of Madan and Schoutens (2016, Example 5.7), where the authors compute bid and ask prices for European Put and Calls under the VG model in a semi-closed form.

Strike	Bench. price	Q price	Rel. error (bp)	ImpVol sqdiff (10^{-7})
$K = 80.0$	28.5214	28.4966	8.6903	8.0272
$K = 85.0$	25.2837	25.2592	9.7120	6.2545
$K = 90.0$	22.3119	22.2878	10.7988	5.0239
$K = 95.0$	19.6104	19.5870	11.9197	4.1266
$K = 100.0$	17.1775	17.1541	13.6605	3.7794
$K = 105.0$	15.0054	14.9824	15.3371	3.4267
$K = 110.0$	13.0812	13.0569	18.6174	3.7360
$K = 115.0$	11.3883	11.3643	21.0614	3.6304
$K = 120.0$	9.9072	9.8813	26.1757	4.3594

TABLE 8.2: Pricing comparison between the benchmark price and the price obtained via quantization of a Call option for the Tempered Stable model with parameters $\alpha^+ = 1, \lambda^+ = 0.5, c^+ = 6$, $\alpha^- = 1, \lambda^- = 0.5, c^- = 4$, $r = 0.04$, $T = 1$. The quantization grids have size $\mathbf{N = 1000}$, and the computational time is around 35 **milliseconds**. The benchmark price is computed using a FFT pricer.

The computational time in our examples prove that this method is precise and computationally comparable to a FFT approach, that we used as benchmark. We used two measures for the error. The first one is the relative error $\frac{|\text{Benchmark} - \text{Quantization price}|}{\text{Benchmark}}$ that we present in basis points ($1bp = 0.01\%$), while ImpVol sqdiff is a measure that involves the implied volatility of the prices, i.e. $(\text{impvol}_{\text{Benchmark}} - \text{impvol}_{\text{Quantization}})^2$, and is mostly used in the industry practice.
The two examples here study the behavior of the pricing of European call options for a given choice of the size N of the quantization grids. Note that the computational time is composed by:

1. The solution of (8.7), which is the most computationally intense.

2. The finite sum in (8.8), which is just a finite sum and is immediate to compute.

The benchmark price is computed using a Fourier approach as in Carr and Madan (1999). This pricing exercise shows that the uniform Fourier quantization technique can price different types of models in a fast and precise way. The interesting part of this technique is that it can be easily implemented in a Conic pricing framework, as we will see later. Up to our knowledge, it is not possible to do the same using the technology of Carr and Madan (1999) for a generic model with a characteristic function in closed form.

8.4.2 Conic quantization pricing

We show here the possibility of recreating a implied volatility smile using our methodology with the conic pricing theory. We choose the same model as the pricing exercise, the Heston model, and we use the generalized distortion function MINMAXVAR defined as

$$\Psi_{\lambda,\gamma}(u) = 1 - \left(1 - u^{\frac{1}{\lambda+1}}\right)^{1+\gamma}, \tag{8.12}$$

where both λ and γ are two real positive numbers. We present here a plot of the implied volatilities presenting the mid, the ask and the bid price. The implied volatilities of call option are on the left of the smile, while put options are on the right. The parameters λ and γ are chosen to be $\gamma = \lambda = 0.01$. Figure 8.1 shows that we are able to price with a smile, using in this particular example the Heston model.

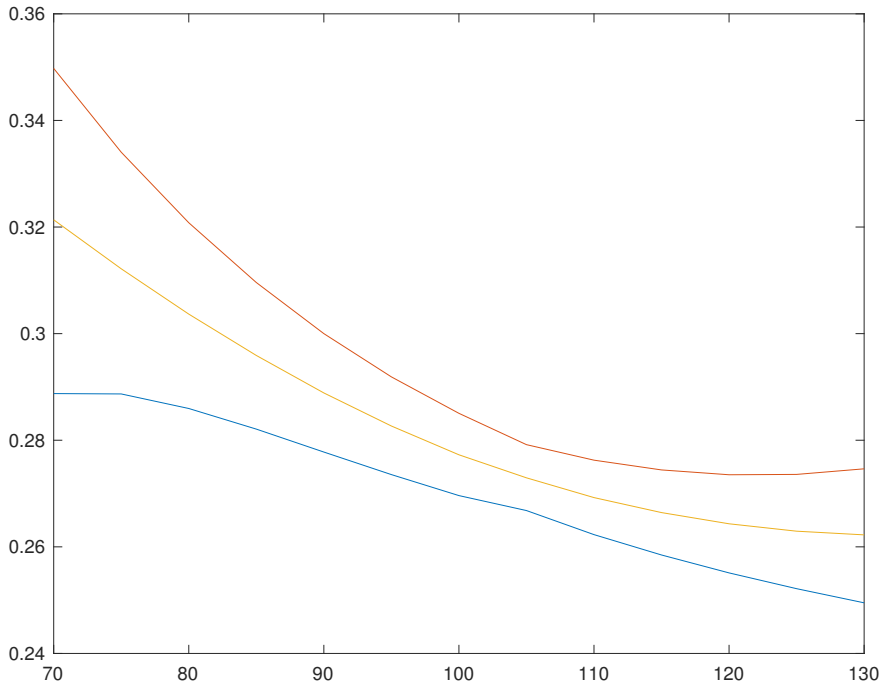


FIGURE 8.1: Bid - Ask smile using the MINMAXVAR distortion function with parameter $\lambda = 0.01$ in the Heston model as in Table 8.1 and $N = 1000$ points for the quantization grid.

8.5 Calibration on market data

In this section, we perform an exercise of the calibration on bid and ask prices under stochastic volatility, taking into consideration a factor linked to the liquidity of the market. We use, as an example, the Heston model, but we could in fact use every model that has a characteristic function in closed form. The market data are taken from the SP500 index as of date July, 11$^{\text{th}}$ 2017.

8.5.1 Global fit

A first approach to calibration of the bid-ask smile is considering a generalized MINMAXVAR distortion function where the parameters λ and γ are fixed for every strike and maturity. This approach gives a global value of this liquidity parameters in the market.
The book of options considered includes 8 maturities with 18 different strikes, for a total of 288 options, since we are considering both ask and bid prices in the calibration routine. The calibration algorithm that we implemented consists in finding the minimum of the average square implied volatility error, i.e. we minimize

$$\frac{1}{\#_{\text{strikes}} \times \#_{\text{maturities}}} \sum_{i=1}^{\#_{\text{strikes}}} \sum_{j=1}^{\#_{\text{maturities}}} \left[\left(IV_{i,j}^{\text{bid,mkt}} - IV_{i,j}^{\text{bid,mod}}(\text{param}) \right)^2 + \left(IV_{i,j}^{\text{ask,mkt}} - IV_{i,j}^{\text{ask,mod}}(\text{param}) \right)^2 \right], \tag{8.13}$$

where $IV_{i,j}^{\text{bid,mkt}}$ and $IV_{i,j}^{\text{ask,mkt}}$ are the bid and ask implied volatility taken from market data, while $IV_{i,j}^{\text{bid,mod}}(\text{param})$ and $IV_{i,j}^{\text{ask,mod}}(\text{param})$ are the bid and ask implied volatility computed using the conic quantization technique, where param contains both the model parameters and the distortion function parameters. In Table 8.3 we present the results of the calibration exercise, while Figure 8.2 and 8.3 provide the squared errors for both bid and ask implied volatilities.

8.5.2 Implied liquidity

It is also possible to consider the calibration of the bid-ask spread under a different perspective. Using the usual MINMAXVAR distortion function, we assume that every derivative has a different parameter, so there is a different λ for every strike and every maturity. We define the market implied liquidity as the parameter λ which, for every strike and every maturity, minimizes the difference between the model ask/bid prices and the market prices.

In this case the calibration is done on the mid price, and then the parameters $\lambda_{K,T}$ are obtained minimizing the distance between the model bid ask prices and the market bid ask prices. This calibration approach is more flexible, and it leads to a better fit of the volatility smiles. In analogy with the implied volatility surface, we call the surface composed by $\lambda_{K,T}$ the implied liquidity surface of the market. Note that this surface depends on the choice of the model, in this case the Heston model, and on the choice of the distortion function, in this case the MINMAXVAR function. With different models and distortion function we could expect a different shape of different levels of implied liquidity. Figure 8.4 shows the implied volatility surface of the market. It shows that the implied liquidity is decreasing for increasing values of the maturity and of the moneyness.

Parameters	Conic Quantization
κ	2.2299
ξ	0.4574
ρ	-0.6452
θ	0.0308
V_0	0.0163
λ	0.0001
γ	0.0108
Res Norm	1.8460×10^{-4}

TABLE 8.3: Heston model calibrated on a book of bid and ask vanillas on the SP500 index as of date July, 11$^{\text{th}}$ 2017. The book includes 8 maturities (from 6 weeks to 3 years) with 18 strikes, for a total of 288 options. Res Norm indicates the average square error on implied volatilities.

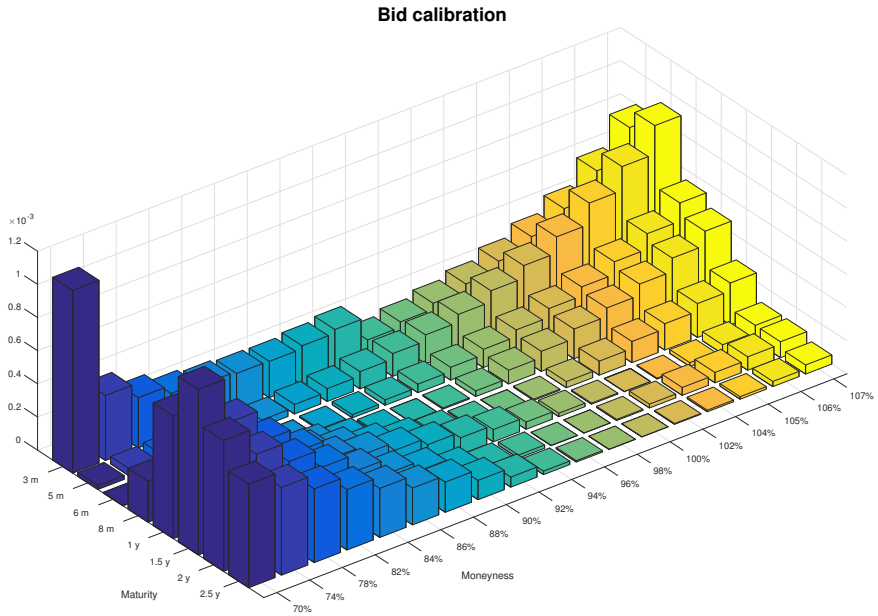


FIGURE 8.2: Plot of the square of the difference between the calibrated model implied volatilities and the market implied volatilities on *bid prices*, for all the 18 strikes and the 8 maturities of the book of vanillas on the SP500 index.

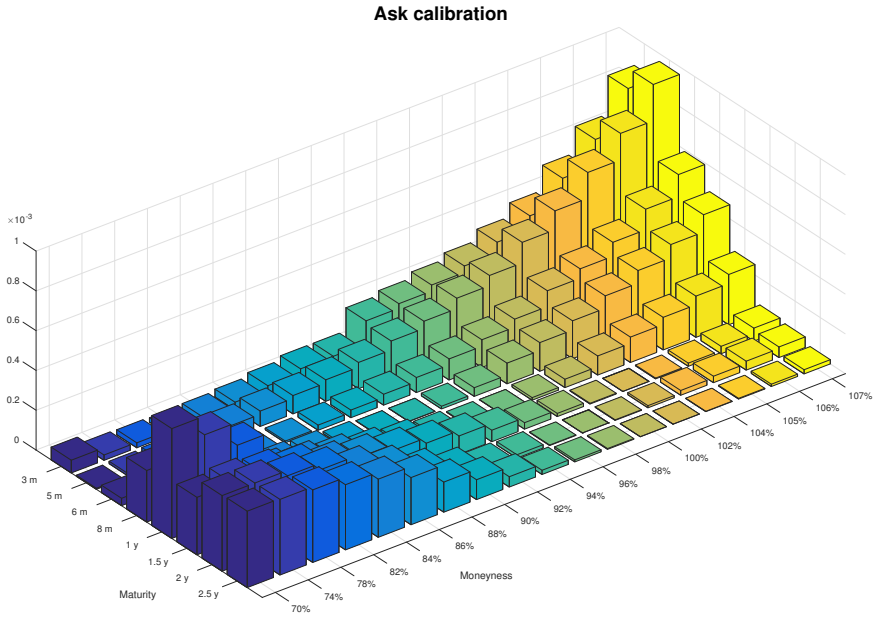


FIGURE 8.3: Plot of the square of the difference between the calibrated model implied volatilities and the market implied volatilities on *ask prices*, for all the 18 strikes and the 8 maturities of the book of vanillas on the SP500 index.

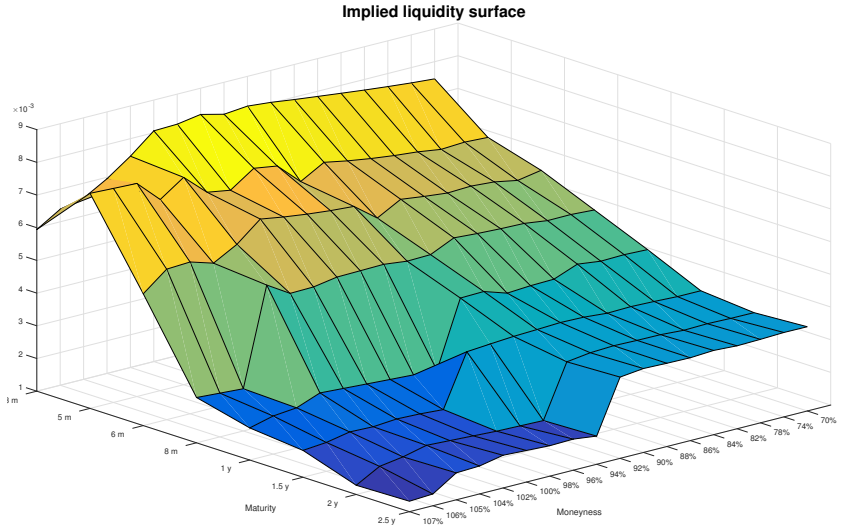


FIGURE 8.4: Plot of the implied liquidity surface for the SP500 index as of date July, 11th 2017 under the Heston model and the MINMAX-VAR distortion function.

8.6 Conclusion

The presence of liquidity in the market of derivatives is an important feature that has to be taken into account when calibrating a model to financial data. Indeed, the mid price is a price that does not exist in reality, so calibration exercises can be biased. The theory of two prices, studied in Madan and Schoutens (2016) tackles this issue and allows us to consider the bid and ask prices for a European derivative. It is then possible, using the new quantization algorithm, to consider the liquidity of the market and the presence of stochastic volatility in the same pricing mechanism, and we can consider for the first time in literature the implied liquidity surface under stochastic volatility.

Bibliography

Abramowitz, M. and Stegun, I. A. (1970). *Handbook of Mathematical Functions*. Dover Publications, New York, first edition.

Ackerer, D., Filipovic, D., and Pulido, S. (2016). The Jacobi Stochastic Volatility Model. *Preprint available at:* https://ssrn.com/abstract=2782486.

Albrecher, H., Mayer, P., Schoutens, W., and Tistaert, J. (2007). The little Heston trap. *Wilmott Magazine*, pages 83–92.

Alfonsi, A. (2015). *Affine Diffusions and Related Processes: Simulation, Theory and Applications*. Springer.

Andersen, L. (2011). Option Pricing with Quadratic Volatility: a Revisit. *Finance and Stochastics*, 15(2):191–219.

Andersen, L. and Piterbarg, V. (2007a). Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50.

Andersen, L. B. G. and Piterbarg, V. V. (2007b). Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Bally, V., Pagès, G., and Printems, J. (2001). A stochastic quantization method for non-linear problems. *Monte Carlo Methods and Appl.*, 1:21–34.

Bally, V., Pagès, G., and Printems, J. (2005). A quantization tree method for pricing and hedging multidimensional American options. *Mathematical Finance*, 15(1):119 – 168.

Bardou, O., Bouthemy, S., and Pagès, G. (2009). Optimal quantization for the pricing of swing options. *Applied Mathematical Finance*, 16:183–217.

Barndorff-Nielsen, O. E. (1997). Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scand. J. Statist.*, 24:1–13.

Bates, D. S. (1996). Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *The Review of Financial Studies*, 9:69–107.

Beliaeva, N. and Nawalkha, S. (2010). A simple approach to pricing American options under the Heston stochastic volatility model. *The Journal of Derivatives*, 17(4):25–43.

Bender, C. and Denk, R. (2007). A forward scheme for backward SDEs. *Stochastic Processes and their Applications*, 117(12):1793–1812.

Bennett, W. R. (1948). Spectra of quantized signals. *Bell System Technical Journal*, 27(3):446–472.

Blacher, G. (2001). A new approach for designing and calibrating stochastic volatility models for optimal delta-vega hedging of exotics. *In: ICBI Global Derivatives, Conference Presentation. ICBI Global Derivatives, Juan-Les-Pins*.

Bouchard, B. and Touzi, N. (2004). Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, 111(2):175–206.

Boyarchenko, S. and Levendorskiy, S. (2013). American options in the Heston model with stochastic interest rate and its generalizations. *Applied Mathematical Finance*, 20(1):26–49.

Callegaro, G., Fiorin, L., and Grasselli, M. (2015). Quantized calibration in local volatility. *Risk Magazine*, 28(4):62 – 67.

Callegaro, G., Fiorin, L., and Grasselli, M. (2016). Pricing via recursive quantization in stochastic volatility models. *Quantitative Finance*, Issue 6:855–872.

Callegaro, G., Fiorin, L., and Grasselli, M. (2017a). American Quantized Calibration in Stochastic Volatility. *Preprint*.

Callegaro, G., Fiorin, L., and Grasselli, M. (2017b). Quantization meets Fourier: A new tecnology for pricing options. *Preprint available at: https://ssrn.com/abstract=2951755*.

Cambanis, S. and Gerr, N. (1983). A simple class of asymptotically optimal quantizers. *IEEE Transactions on Information Theory*, 29(5):664–676.

Carr, P., Fischer, T., and Ruf, J. (2013). Why are quadratic normal volatility models analytically tractable? *SIAM J. Financial Mathematics*, 4:185–202.

Carr, P., Geman, H., Madan, D., and Yor, M. (2002). The fine structure of asset returns: an empirical investigation. *The Journal of Business*, 75(2):305–332.

Carr, P. and Lewis, K. (2004). Corridor Variance Swaps. *Risk Magazine*, February:67–72.

Carr, P. and Madan, D. B. (1999). Option Valuation Using the Fast Fourier Transform. *Journal of Computational Finance*, 2:61–73.

Cherny, A. and Madan, D. (2009). New measures for performance evaluation. *The Review of Financial Studies*, 22(7):2571–2606.

Cherny, A. and Madan, D. (2010). Markets as a counterparty: An introduction to conic finance. *International Journal of Theoretical and Applied Finance*, 13:1149–1177.

Chesney, M. and Scott, L. (1989). Pricing European Currency Options: A Comparison of the Modified Black-Scholes Model and a Random Variance Model. *Journal of Financial and Quantitative Analysis*, 24:267–284.

Christoffersen, P., Heston, S., and Jacobs, K. (2009). The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well. *Management Science*, 55(12):1914–1932.

Cont, R., Bouchaud, J.-P., and Potters, M. (1997). Scaling in financial data: Stable laws and beyond. *Scale Invariance and Beyond, Dubrulle, B., Graner, F., and Sornette, D., eds., Springer. Berlin, 1997*.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC.

Cont, R. and Voltchkova, E. (2005). A Finite Difference Scheme for Option Pricing in Jump Diffusion and Exponential Lévy Models. *SIAM Journal on Numerical Analysis*, 43(4):1596 – 1626.

Crisan, D., Manolarakis, K., and Touzi, N. (2010). On the Monte Carlo simulation of BSDE's: an improvement on the Malliavin weights. *Stochastic Processes and their Applications*, 120:1133–1158.

Cuchiero, C., Filipovic, D., Mayerhofer, E., and Teichmann, J. (2009). Affine Processes on Positive Semidefinite Matrices. *Annals of Applied Probability, to appear*.

Cuchiero, C., Keller-Ressel, M., and Teichmann, J. (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 67(4):711–740.

Da Fonseca, J., Gnoatto, A., and Grasselli, M. (2015). Analytic pricing of volatility-equity options within Wishart-based stochastic volatility models. *Operations Research Letters*, 43(6):601–607.

Da Fonseca, J. and Grasselli, M. (2011). Riding on the Smiles. *Quantitative Finance*, 11:1609–1632.

Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2008). A multifactor volatility heston model. *Quantitative Finance*, 8(6):591–604.

Da Fonseca, J. and Martini, C. (2016). The α-Hypergeometric Stochastic Volatility Model. *Stochastic Processes and their Applications*, 5:1472–1502.

Date, P. and Islyaev, S. (2015). A fast calibrating volatility model for option pricing. *European Journal of Operational Research*, 243(2):599–606.

Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusion. *Econometrica*, 68:1343–1376.

Eberlein, E., Glau, K., and Papapantoleon, A. (2010). Analysis of Fourier transform valuation formulas and applications. *Applied Mathematical Finance*, 17:211–240.

El Karoui, N., Peng, S., and Quenez, M. C. (1997). Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71.

Fang, F. and Oosterlee, C. W. (2011). A Fourier-based valuation method for Bermudan and Barrier options under Heston's model. *SIAM Journal on Financial Mathematics*, 2(1):439–463.

Filipovic, D. and Larsson, M. (2016). Polynomial diffusions and applications in finance. *Finance and Stochastics*, 20(4):931–972.

Filipovic, D., Larsson, M., and Pulido, S. (2016). Markov cubature rules for polynomial processes. *Swiss Finance Institute Research Paper*, 16–79.

Fiorin, L., Sagna, A., and Pagès, G. (2015). Product Markovian quantization of a diffusion process with applications to finance. *Working Paper, available at: http://arxiv.org/abs/1511.01758*.

Gersho, A. and Gray, R. (1992). *Vector Quantization and Signal Compression*. Kluwer Academic Press, Boston.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Springer Series: Applications of Mathematics.

Gobet, E., Lopez-Salas, J., Turkedjiev, P., and Vasquez, C. (2016). Stratified regression Monte-Carlo scheme for semilinear PDEs and BSDEs with large scale parallelization on GPUs. *SIAM Journal on Scientific Computing*, 38(6):C652–C677.

Gobet, E. and Turkedjiev, P. (2016). Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Mathematics of Computation*, 85:1359–1391.

Graf, S. and Luschgy, H. (2000). *Foundations of quantization for probability distributions*. Springer, New York.

Grasselli, M. (2016). The 4/2 stochastic volatility model: A unified approach for the heston and the 3/2 model. *Mathematical Finance*, 27(4):1013–1034.

Gulisashvili, A. and Stein, E. M. (2010). Asymptotic behavior of the stock price distribution density and implied volatility in stochastic volatility models. *Applied Mathematics and Optimization*, 61(3):287–315.

Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. (September 2002). Managing smile risk. *Wilmott magazine*, pages 84–108.

Heston, S. (1997). A simple new formula for options with stochastic volatility. *Technical report, Washington University of St. Louis*.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343.

Hu, Y., Nualart, D., and Song, X. (2011). Malliavin calculus for backward stochastic differential equations and applications to numerical solutions. *The Annals of Applied Probability*, 21(6):2379–2423.

Hull, J. and White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, 42:281–300.

Ikonen, S. and Toivanen, J. (2008). Efficient numerical methods for pricing American options under stochastic volatility. *Numerical Methods for Partial Differential Equations*, 24(1):104–126.

Ingersoll, J. (1997). Valuing foreign exchange rate derivatives with a bounded exchange process. *Review of Derivatives Research*, 1:159–181.

Jacobs, K. and Li, X. (2008). Modeling the Dynamics of Credit Spreads with Stochastic Volatility. *Management Science*, 54(6):1176–1188.

Ju, N. (2002). Pricing Asian and Basket options via Taylor expansion. *Journal of Computational Finance*, 5 (3):79–103.

Keller-Ressel, M. (2011). Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance*, 21:73–98.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. *Physics Review E*, 52:1197 – 1199.

Kou, S. (2003). On pricing of discrete barrier options. *Statistica Sinica*, 13(4):955–964.

Kuchler, U. and Tappe, S. (2008). Bilateral gamma distributions and processes in financial mathematics. *Stochastic Processes and their Applications*, 118(2):261 – 283.

Kusuoka, S. (2001). On law invariant coherent risk measures. *Kusuoka S., Maruyama T. (eds) Advances in Mathematical Economics*, 3.

Lipton, A. (2002). The vol smile problem. *Risk Magazine*, 15(2):61–65.

Lipton, A., Gal, A., and Lasis, A. (2014). Pricing of vanilla and first-generation exotic options in the local stochastic volatility framework: survey and new results. *Quantitative Finance*, 14(11):1899–1922.

Lipton, A. and McGhee, W. (2002). Universal barriers. *Risk Magazine*, 15:81–85.

Longstaff, F. A. and Schwartz, E. S. (2001). Valuing American Options by Simulation: A Simple Least-Squares Approach. *Review of Financial Studies*, 14(1):113 – 147.

Luschgy, H. and Pagès, G. (2002). Functional quantization of Gaussian processes. *Journal of Functional Analysis*, 196:486–531.

Madan, D. and Schoutens, W. (2016). *Applied Conic Finance*. Cambridge.

Madan, D. B., Carr, P. P., and Chang, E. C. (1998). The Variance Gamma process and option pricing. *European finance review*, 2(1):79–105.

Max, J. (1960). Quantizing for minimum distortion. *IRE Transactions on Information Theory*, 6(1):7–12.

McWalter, T., Rudd, R., Kienitz, J., and Platen, E. (2017). Recursive Marginal Quantization of Higher-Order Schemes. *Working Paper, available on SSRN*.

Ozcag, E., Ege, I., and Gurcay, H. (2008). An extension of the incomplete beta function for negative integers. *Journal of Mathematical Analysis and Applications*, 338(2):984 – 992.

Pagès, G. (2015). Introduction to vector quantization and its applications for numerics. *ESAIM: proceedings and surveys*, 48:29–79.

Pagès, G., Pham, H., and Printems, J. (2003). *Optimal quantization methods and applications to numerical problem in finance*. In Handbook of Computational and Numerical Methods in Finance , S.T. Rachev ed., Birkhauser, Boston.

Pagès, G. and Printems, J. (2003). Optimal quadratic quantization for numerics: the Gaussian case. *Monte Carlo Methods and Applications*, 9(2):135–165.

Pagès, G. and Printems, J. (2005). Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods and Applications*, 11(4):407–446.

Pagès, G. and Sagna, A. (2015). Recursive marginal quantization of the Euler scheme of a diffusion process. *Applied Mathematical Finance*, 22(5):463–498.

Pagès, G. and Sagna, A. (2016). Improved error bounds for quantization based numerical schemes for bsde and nonlinear filtering. *Stochastic Processes and Their Applications. To appear.*

Pagès, G. and Wilbertz, B. (2012). *Optimal Delaunay and Voronoi quantization schemes for pricing American style options*, pages 171 – 217. Springer.

Pagès, G. and Yu, J. (2013). Pointwise convergence of the Lloyd algorithm in higher dimension. *SIAM Journal on Control and Optimization. To appear.*

Panter, P. F. and Dite, W. (1951). Quantization distortion in pulse-count modulation with nonuniform spacing of levels. *Proceedings of the IRE*, 39(1):44–48.

Pierce, J. (1970). Asymptotic quantizing error for unbounded random variables. *IEEE Trans. Inf. Theor.*, 16(1):81–83.

Platen, E. (1997). A non-linear stochastic volatility model. *Financial Mathematics Research Report No. FMRR 005-97, Center for Financial Mathematics, Australian National University, Canberra.*

Rogers, L. and Williams, D. (2000). *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, volume 2. Cambridge university press.

Sagna, A. (2012). Pricing of barrier options by marginal functional quantization. *Monte Carlo Methods and Applications*, 17:371 – 398.

Schöbel, R. and Zhu, J. (1999). Stochastic volatility with an Ornstein Uhlenbeck process: An extension. *European Finance Review*, 3:23 – 46.

Sesana, D., Marazzina, D., and Fusai, G. (2014). Pricing exotic derivatives exploiting structure. *European Journal of Operational Research*, 236(1):369–381.

Shephard, N. G. (1991). From characteristic function to distribution function: A simple framework for the theory. *Econometric Theory*, 7(4):519–529.

Stein, E. M. and Stein, J. C. (1991). Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4:727 – 752.

Tian, Y. (2013). The Hybrid Stochastic-Local Volatility Model with Applications in Pricing FX Options. *PhD thesis, School of Mathematical Sciences Monash University, Australia. Available at SSRN 2399935 (2013).*

Vellekoop, M. and Nieuwenhuis, H. (2009). A tree-based method to price American options in the Heston model. *Journal of Computational Finance*, 13(1):1 – 21.

Zador, P. (1982). Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Trans. Inform. Theory*, 28:139–149.

Zühlsdorff, C. (2002). Extended LIBOR market models with affine and quadratic volatility. *Working Paper, University of Bonn (2002).* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=228250.